2. argitalpena

Estatistikaren oinarriak

Ariketak

Elena Agirre Basurko

ueu

ESTATISTIKAREN OINARRIAK. ARIKETAK

2. argitalpena

Elena Agirre Basurko

Udako **Euskal Unibertsitatea**
Bilbo, 2010

ISBN: 978-84-8438-288-1
Lege-gordailua: BI-1780-2010

Inprimategia: CUATROAS, Bilbo
Azalaren diseinua: Iñigo Ordozgoiti
Hizkuntza-zuzenketen arduraduna: Ander Altuna Gabiola

Banatzaileak: UEU. Erribera 14, 1. D BILBO telf. 946790546 Faxa. 944793039
Helbide elektronikoa: argitalpenak@ueu.org
www.ueu.org

Elkar Banaketa: Igerabide, 88 DONOSTIA

Olatz, Naroa eta nire gurasoei

Aurkibidea

Hitzaurrea

*Estatistikaren Oinarriak. Ariketak i*zeneko liburu honetan, Euskal Herriko Unibertsitateko Bilboko Industria Ingeniaritza Teknikoko Unibertsitate Eskolan, *Ingeniaritzaren Metodo Estatistikoak* irakasgaian aztertzen diren gaiak garatu dira. Liburu hau, azken urteetan *Ingeniaritzaren Metodo Estatistikoak* irakasgaiaren prestakuntzan egin dudan lanaren ondorioa dela esan daiteke. Liburuko gai bakoitzaren hasieran gaiari buruz jakin beharreko kontzeptu eta emaitza teorikoak azaldu dira, eta ondoren, gaiari buruzko ariketa ebatziak daude. Liburuko teoria, era laburrean, frogapenik gabe azaldu da, ikaslearen helburu praktikoetarako erabilgarria izan dadin. Halaber, gai bakoitzaren bukaeran ikasleak egiteko zenbait ariketa proposatu dira, liburuaren bukaeran ariketa proposatu horien emaitzak daudelarik. Oro har, ariketen enuntziatuak Ingeniaritza Teknikoa ikasten duten ikasleentzako egokiak direlakoan nago. Hala ere, liburu hau, Estatistikaren hastapenak aztertu, ikasi edo landu behar dituen edozein erabiltzailerentzat baliagarria izatea nahiko nuke.

Bukatzeko, nire eskerrik beroenak Albaro Antari, zeinaren kolaboraziorik gabe ez zen burutuko liburua; Carlosi, lan hau egiteko eskaini didan denboragatik eta liburuaren argitalpen-prozesuan aritu izan diren UEUko kide guztiei, nigan jarritako konfiantzagatik.

Elena Agirre Basurko

Bigarren argitalpenaren hitzaurrea

Iaz argitaratu zen *Kalkulu Diferentziala eta Integrala* (Angulo eta beste, 2009) izeneko Piskunoven liburuaren euskal bertsioaren bigarren argitalpenaren aitzinsolasa honela hasten da: "Ez da ohiko kontua euskarazko liburuen bigarren argitalpena ateratzea; are gutxiago, unibertsitatean erabiltzen den eskuliburu batena". Beraz, esku artean duzun liburu honen egile naizen neurrian, oso pozik nago bigarren argitalpen honen aurrean. Dirudienez estatistikaren oinarriak ezagutzeko eta ulertzeko liburu hau zenbait pertsonarentzat lagungarria izan da, eta horrek benetan pozten nau. Aurrerantzean ere horrela izatea nahi nuke. 2010-2011 ikasturtean UPV/EHUn Ingeniaritzako Graduak abiatuko dira eta *Ingeniaritzaren Metodo Estatistikoak* izeneko irakasgaian estatistikaren oinarrizko kontzeptu, emaitza eta teknikak garatuko dira. Garai berri bat hasiko da unibertsitatean; irakasteko metodo ezberdinak konbinatuz, ezagutza transmititzeko era aldatuko da, baina ez da aldatuko transmititu beharreko ezagutza. Horregatik liburu hau aurrera begira ere baliozkoa izatea espero dut.

Bigarren argitalpen honek liburua hobetzeko zenbait zuzenketa eta aldaketa egiteko aukera eman dit. Zentzu honetan eskerrak eman nahi nizkieke azken urteotan Bilboko Industria Ingeniaritza Teknikoko Unibertsitate Eskolan izan ditudan ikasleei, eta bereziki nire lankide diren Naiara Barrado eta Patxi de la Hozi. Eskerrak ere Martxel Ensunza, Jose Ramon Etxebarria eta Juan Carlos Odriozolari, nire lan-jarduera euskaraz burutzeko irakatsitakoagatik. UEUko Nekane Intxaurtza eta Ander Altuna, mila esker zuen lanagatik. Eta bukatzeko, Carlos, eskerrik beroenak zuretzat, oraingo honetan ere nire ondoan izan zarelako.

Elena Agirre

Gasteizen, 2010eko maiatzaren 28an

1. Estatistika deskribatzailea. Aldagai bakuna

1.1. SARRERA

Estatistika multzo bati dagozkion zenbakizko datuak biltzen, sailkatzen eta aztertzen dituen zientzia da. Estatistikaren barruan bi arlo nagusi bereiz daitezke: **Estatistika Deskribatzailea** eta **Estatistika Induktiboa.** Estatistika deskribatzailearen bidez, aztergai den multzoari dagozkion datuak bildu, antolatu eta egoera deskribatzeko ezaugarriak lortuko dira. Estatistika induktiboaren bidez, aldiz, emaitzak orokortu, ondorioak atera edo aurresanak egin daitezke.

Azterketa estatistikoa multzo batean gauzatuko da. Multzo oso horri **populazioa** deritzo. Hala ere, kasu gehienetan azterketa populazioaren azpimultzoetan egiten da. Populazioaren edozein azpimultzori **lagina** deritzo. Populazioaren elementu bakoitza **unitate estatistiko** edo **ale estatistikoa** da. Populazioaren (edo laginaren) **tamaina** populazioaren (edo laginaren) elementu-kopurua da.

Azterketa estatistikoan aztergaiak har ditzakeen balioek **aldagai estatistikoa** determinatzen dute. Aldagaiak hartzen dituen balioak zenbakiak badira, **aldagai kuantitatiboa** dugu. Aldagaiak hartzen dituen balioak zenbakizkoak ez badira, aldiz, **aldagai kualitatiboa** dugu.

Aldagai aleatorio kuantitatiboak diskretuak edo jarraituak izan daitezke. **Aldagai kuantitatibo diskretua**k balio isolatuak hartzen ditu eta **aldagai kuantitatibo jarraitua**k tarte bateko edozein balio har dezake.

1.2. MAIZTASUN-TAULAK

Datuak bildu ondoren, komeni da datuak ordenatu eta tauletan adieraztea. Horretarako, honako kontzeptuak hartuko dira kontuan:

1. **Modalitateak:** x_i gaiak dira aldagaiaren balioak. Balio hauek zenbakiak direnean, oro har, txikienetik handienera ordenatzen dira.

2. **Maiztasun absolutuak:** f_i gaiak balio bakoitza zenbat aldiz jaso den adierazten du.

3. **Maiztasun metatuak:** $F_i = f_1 + f_2 + \dots + f_i$ maiztasun absolutuen batura da.

Batzuetan interesgarria izan daiteke **maiztasun erlatiboa** $h_i = \dfrac{f_i}{n}$ eta **maiztasun erlatibo metatua** $H_i = \dfrac{F_i}{n}$ kalkulatzea, non laginaren tamaina n eta $i = 1, 2, \ldots, k$ diren.

Askotan datu-kopurua handia denean, komenigarria da **maiztasun elkartuen metodoa** erabiltzea. Horretarako, ondoko pausoak eman daitezke:

i. Aldagai estatistikoaren *heina* kalkulatuko da, heina balio handienaren eta txikienaren arteko kendura delarik.

ii. Tartea, azpitartetan edo klasetan banatuko da. Klase-kopurua zehazteko orduan klase gutxi hartuz gero, informazioa gal daiteke, eta, aldiz, klase gehiegik azterketa zaildu dezakete. Beraz, klase-kopurua $k = \sqrt{n}$ aukera daiteke, non n gaia laginaren (populazioaren) tamaina den.

iii. Klaseak zehaztuko dira. Praktikan erosoa da $[l_i\,,\ l_{i+1})$ erako klaseak aukeratzea. Klase bakoitzeko balioen ordezkari moduan klase-marka dugu, hots, tarte bakoitzeko erdiko puntua: $x_i = \dfrac{l_i + l_{i+1}}{2}$.

iv. Maiztasunak zenbatu eta taula eratuko da.

$[l_i\,, l_{i+1})$	x_i	f_i	F_i	h_i	H_i
$[l_1\,, l_2)$	x_1	f_1	$F_1 = f_1$	$h_1 = f_1/n$	$H_1 = F_1/n$
$[l_2\,, l_3)$	x_2	f_2	$F_2 = f_1 + f_2$	$h_2 = f_2/n$	$H_2 = F_2/n$
…	…	…	…	…	…
$[l_k\,, l_{k+1})$	x_k	f_k	$F_k = f_1 + f_2 + \ldots + f_k = n$	$h_k = f_k/n$	$H_k = F_k/n = 1$
		$\sum_{i=1}^{k} f_i = n$		$\sum_{i=1}^{k} h_i = 1$	

1.3. ADIERAZPEN GRAFIKOAK

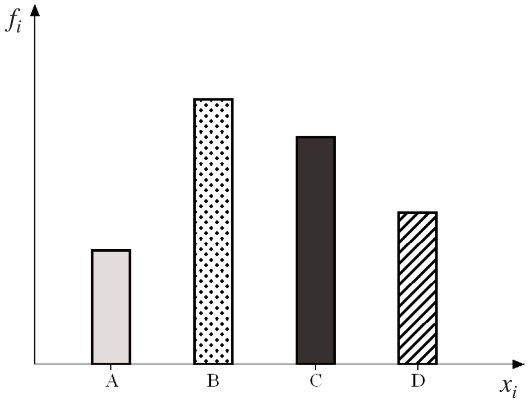
Adierazpen grafikoak oso lagungarriak izaten dira. Zenbaitetan begirada soil batez banaketaren ideia ematen dute. Aldagaiaren izaeraren arabera, hurrengo adierazpen grafikoak azter daitezke:

$$\text{Aldagaia}\begin{cases}\text{Kualitatiboa}\begin{cases}\text{Barra-diagrama}\\ \text{Sektore-diagrama}\end{cases}\\[2ex] \text{Kuantitatiboa}\begin{cases}\text{Diskretua}\begin{cases}\text{Barra-grafikoa}\\ \text{Maiztasun metatuen grafikoa}\end{cases}\\[2ex] \text{Jarraitua}\begin{cases}\text{Histograma}\\ \text{Maiztasun absolutuen poligonoa}\\ \text{Maiztasun metatuen histograma}\\ \text{Maiztasun metatuen poligonoa}\end{cases}\end{cases}\end{cases}$$

1.3.1. Aldagai kualitatiboaren adierazpen grafikoak

1.3.1.1. Barra-diagrama

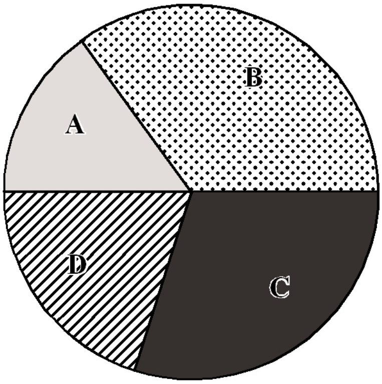
Barrez osaturiko diagrama bat da, non abzisa-ardatzean aldagaiaren balioak eta ordenatu-ardatzean maiztasun absolutuak adieraziko diren. Barrak ez dira elkarturik egoten.



1.1. irudia.

1.3.1.2. Sektore-diagrama

Zirkulu baten barruan aldagai kualitatiboaren balio bakoitzari dagokion sektorea adierazten da. Sektore bakoitzaren kalkulurako $\alpha_i = 360° \cdot \frac{f_i}{n}$ adierazpena erabiliko da, non f_i maiztasun absolutua eta n laginaren tamaina diren.

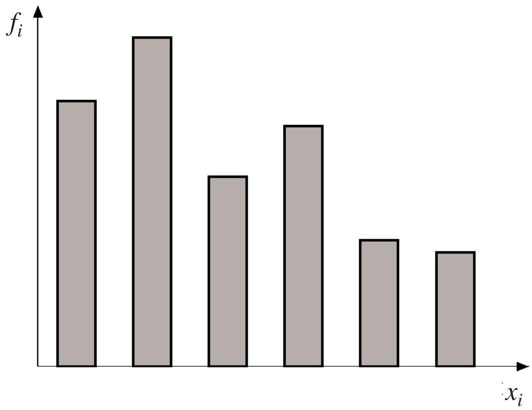


1.2. irudia.

1.3.2. Aldagai kuantitatibo diskretuaren adierazpen grafikoak

1.3.2.1. Barra-grafikoa

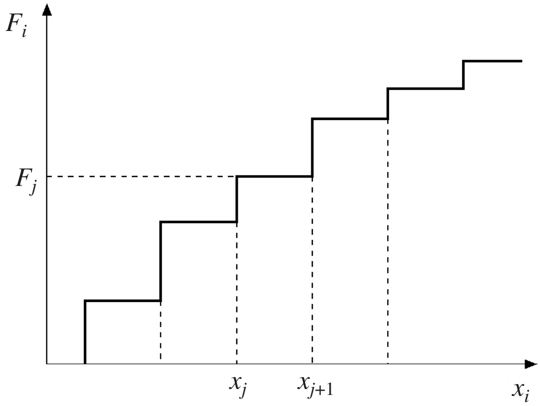
Abzisa-ardatzean aldagaiaren balioak eta ordenatu-ardatzean maiztasun absolutuak adierazten dituzten barrez osaturiko grafikoa da.



1.3. irudia.

1.3.2.2. Maiztasun metatuen grafikoa

Grafiko honetan abzisa-ardatzean aldagaiaren balioak eta ordenatu-ardatzean maiztasun metatuak adieraziko dira. Grafiko honek eskailera-itxura du. x_j balioari F_j altuera eta x_{j+1} balioari F_{j+1} altuera emango zaizkie, hurrenez hurren, (x_j, x_{j+1}) tarteko balioei F_j altuerako maila dagokielarik.

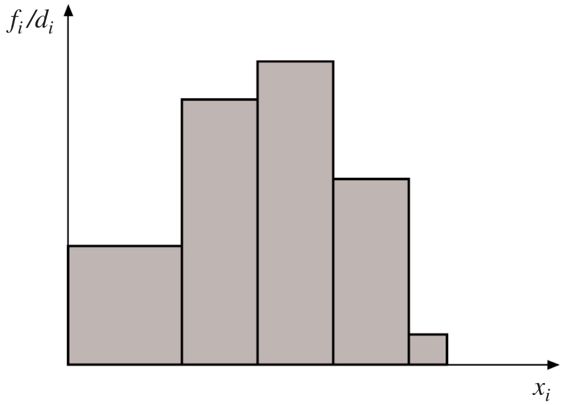


1.4. irudia.

1.3.3. Aldagai kuantitatibo jarraituaren adierazpen grafikoak

1.3.3.1. Histograma

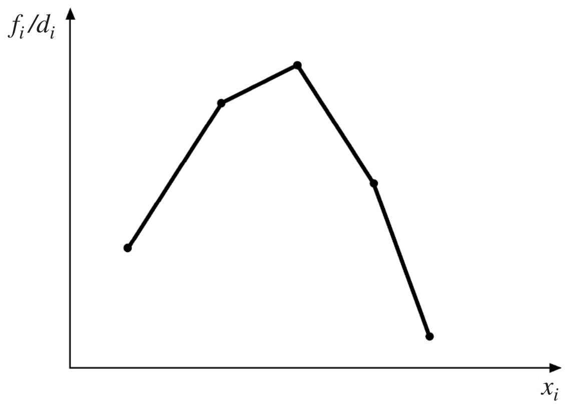
Oinarrian klaseak dituen errektangelu-multzoa da. Klase bakoitza errektangelu baten oinarria izango da. Klase guztiak anplitude berekoak badira, errektangelu bakoitzaren altuera klaseari dagokion maiztasun absolutua izango da. Klaseak anplitude berekoak ez direnean, altuerak f_i / d_i balioak izango dira, non kasu bakoitzean f_i maiztasun absolutua eta d_i klasearen anplitudea izango diren.



1.5. irudia.

1.3.3.2. Maiztasun absolutuen poligonoa

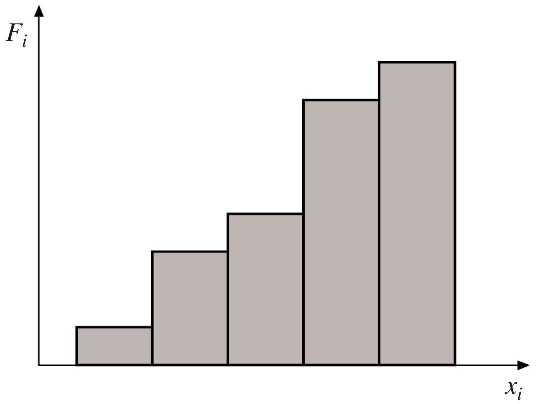
Aurreko histograman errektangeluetako goiko aldeetako erdiko puntuak lotuz, maiztasun absolutuen poligonoa eraikiko da.



1.6. irudia.

1.3.3.3 Maiztasun metatuen histograma

Errektangelu-multzo bat da, zeinaren oinarriak klaseak eta altuerak maiztasun metatuak diren.



1.7. irudia.

1.3.3.4 Maiztasun metatuen poligonoa

Aurreko histograman (l_{i+1}, F_i) puntuak lotuz lortuko den poligonoa da.

1.8. irudia.

1.4. ESTATISTIKO DESKRIBATZAILEAK

Estatistiko deskribatzaileak laginaren menpeko funtzioak dira. Aztergai den fenomenoaren ezaugarriak era laburtuan deskribatzeko balio dute. Hurrengo taulan zenbait estatistiko adieraziko dira era laburtuan. Taula azaldu baino lehen ohar batzuk kontuan hartzea komeni da:

1. X aldagai estatistiko kuantitatiboak $x_1, x_2, ..., x_k$ modalitateak $f_1, f_2, ..., f_k$ maiztasun absolutuz hartzen ditu eta laginaren tamaina n da.

2. Estatistikoek hurrengo sailkapena onartzen dute: i) joera zentralekoak (batez bestekoa, mediana, moda), ii) sakabanatzekoak (heina, kuartilarteko heina, bariantza, desbideratze tipikoa, aldakuntza-koefizientea), iii) posiziokoak (pertzentilak, dezilak eta kuartilak) eta iv) formakoak (alborapena, kurtosia).

3. Kasu jarraituan erabilitako notazioa:

   l_i = estatistikoa daukan klasearen behe-muturra

   f_i = estatistikoa daukan klasearen maiztasun absolutua

   d_i = estatistikoa daukan klasearen luzera

   f_{i-1} = estatistikoa daukan aurreko klasearen maiztasun absolutua

   F_{i-1} = estatistikoa daukan aurreko klasearen maiztasun metatua

   n = laginaren tamaina

4. Bariantza adierazteko $V(x)$, s^2 edo $s^2\ (x)$ notazioak erabil daitezke.

   Kurtosiaren adierazpenean, bigarren ordenako momentua bariantza da, $m_2 = s^2$ alegia.

Estatistikoa	*Adierazpena*	*Esanahia*
Batez besteko aritmetikoa	$\bar{x} = \dfrac{\sum_{i=1}^{k} x_i f_i}{n}$	Aldagaiak hartzen dituen balio guztien baturaren eta laginaren tamainaren arteko proportzioa da.
Mediana	Kasu diskretua: • n bakoitia denean, mediana $(n+1)/2$ posizioan dagoen balioa da. • n bikoitia denean, mediana $n/2$ eta $(n/2)+1$ posizioetako balioen erdiko puntua da. Kasu jarraitua: $M_e = l_i + \dfrac{\frac{n}{2} - F_{i-1}}{f_i} \cdot d_i$	Bere ezkerrean eta bere eskuinean ale-kopuru berdina uzten duen balioa da, eta aldez aurretik aldagaiaren balioak ordenatuta izango dira.
Moda	Kasu diskretua: f_i handiena duen x_i Kasu jarraitua: $M_o = l_i + \dfrac{\Delta_1}{\Delta_1 + \Delta_2} \cdot d_i$ $\Delta_1 = \dfrac{f_i}{d_i} - \dfrac{f_{i-1}}{d_{i-1}}, \ \Delta_2 = \dfrac{f_i}{d_i} - \dfrac{f_{i+1}}{d_{i+1}}$	Maiztasun handieneko balioa da.

Heina edo anplitudea	$R = max(x_i) - min(x_i)$	Laginaren balio handienaren eta txikienaren arteko kendura da.
Bariantza	$V(x) = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 f_i}{n}$ $= \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2$	Batez bestekoarekiko errore karratuen batez bestekoa da.
Desbideratze tipikoa	$s(x) = \sqrt{V(x)}$	Bariantzaren erro karratu positiboa da.
c parametroarekiko r. ordenako momentua	$M_r(c) = \frac{\sum_{i=1}^{k}(x_i - c)^r f_i}{n}$	Parametroarekiko erroreen potentzien baturaren eta laginaren tamainaren arteko proportzioa da.
Momentu zentrala	$m_r = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^r f_i}{n}$	Aurreko momentuan $c = \bar{x}$ eginez lortzen da.
Jatorriarekiko momentua	$a_r = \frac{\sum_{i=1}^{k} x_i^r f_i}{n}$	c-rekiko momentuan $c = 0$ eginez lortzen da.
Aldakuntza-koefizientea	$CV_x = \frac{s_x}{\bar{x}}$	Desbideratze tipikoaren eta batez bestekoaren arteko zatidura da. Aldagai batek bi multzotan edo bi aldagai ezberdinen aldakuntza konparatzeko balio du.
k. ordenako pertzentila	Kasu diskretua: Esanahia aplikatu. Kasu jarraitua: $P_k = l_i + \frac{\frac{kn}{100} - F_{i-1}}{f_i} d_i$	Banaketaren % k balio bere ezkerrean uzten du, non $k = 1, 2, \ldots, 99$ den.

k. ordenako dezila	$D_1 \equiv P_{10}$ $D_2 \equiv P_{20}$ .... $D_9 \equiv P_{90}$	Banaketaren % 10*k* balio bere ezkerrean uzten du, non $k = 1, 2, \ldots, 9$ den.
k. ordenako kuartila	$Q_1 \equiv P_{25}$ $Q_2 \equiv P_{50}$ $Q_3 \equiv P_{75}$	Banaketaren % 25*k* balio bere ezkerrean uzten du, $k = 1, 2, 3$ delarik.
Kuartilarteko heina	$R_I = Q_3 - Q_1$	3. ordenako eta 1. ordenako kuartilen arteko diferentzia da.
Alborapena	$v = \dfrac{\bar{x} - M_o}{s_X}$	Banaketaren simetria neurtzen du. $v > 0$ →eskuinerantz alboratua $v < 0$ →ezkerrerantz alboratua $v = 0$ →banaketa, alboragabea
Kurtosia	$g_2 = \dfrac{m_4}{m_2^2}$	Banaketaren zorroztasuna neurtzen du. $g_2 > 0$→banaketa leptokurtikoa $g_2 = 0$→banaketa mesokurtikoa $g_2 < 0$→banaketa platikurtikoa

1.5. ARIKETA EBATZIAK

1.5.1. ariketa

Ondoko balioak 30 torlojuren lodierak (mm) dira:

1	2	3	3	2	1	2	5	2	4
4	4	5	3	2	5	3	4	1	4
2	3	1	1	2	5	3	4	1	3

a. Eraiki bedi maiztasun-taula.

b. Irudika bedi barra-grafikoa.

c. Adieraz bedi “oso mehea” = 1, “mehea” = 2, “ertaina” = 3, “lodia” = 4 eta “oso lodia” = 5 lodierei dagokien sektore-diagrama.

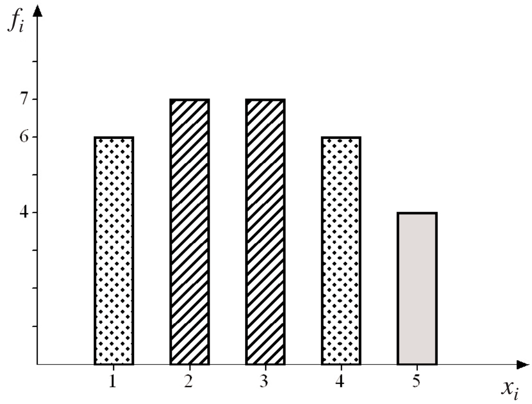
d. Esanahia azalduz, kalkula bitez batez bestekoa, moda, mediana eta desbideratze tipikoa.

Ebazpena

a. Hasteko, datuak ordenatu eta aldagai kuantitatibo diskretu honi dagokion maiztasun-taula eratuko da.

x_i	f_i	F_i	h_i	H_i
1	6	6	$6/30 = 0{,}20$	$6/30 = 0{,}20$
2	7	13	$7/30 \approx 0{,}23$	$13/30 \approx 0{,}43$
3	7	20	$7/30 \approx 0{,}23$	$20/30 \approx 0{,}67$
4	6	26	$6/30 = 0{,}20$	$26/30 \approx 0{,}87$
5	4	30	$2/15 \approx 0{,}13$	$30/30 = 1$

b. Barra-grafikoa:

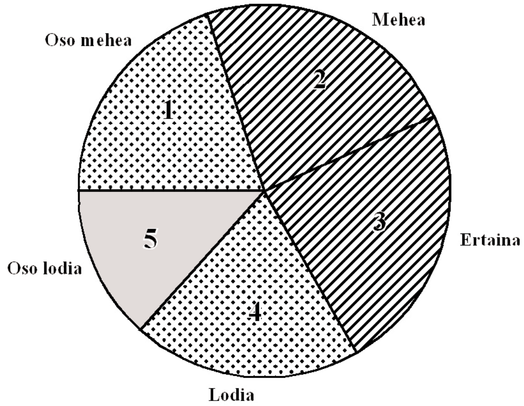


1.9. irudia.

c. Sektore-diagrama:

Hasteko, gai bakoitzari dagokion sektorea kalkulatuko da:

$$\alpha_1 = f_1 \cdot \frac{360}{n} = 6 \cdot \frac{360}{30} = 72° = \alpha_4, \alpha_2 = f_2 \cdot \frac{360}{n} = 7 \cdot \frac{360}{30} = 84° = \alpha_3, \alpha_5 = 48° .$$



1.10. irudia.

d. Batez bestekoaren kalkulua:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} =$$

$$= \frac{1 \cdot 6 + 2 \cdot 7 + 3 \cdot 7 + 4 \cdot 6 + 5 \cdot 4}{30} = \frac{85}{30} = 2,83 \text{ mm.}$$

Beraz, torloju-multzo horren batez besteko lodiera 2,83 mm-koa da. Batez bestekoak, beste estatistikorik kalkulatu gabe, multzoaren adierazle den lodiera azaltzen du.

Banaketak bi moda ditu, 2 eta 3 balioak hain zuzen ere, maiztasun handieneko balioak baitira. Kasu honetan, 2 eta 3 mm-ko lodierako torlojuak dira gehien azaltzen direnak laginean.

Medianaren balioa $M_e = 3$ da, neurtutako lodieren erdiak gehienez 3 mm-koak eta beste erdiak gutxienez 3 mm-koak direlako. Beraz, torlojuen erdien lodiera [1, 3] mm-koa eta beste erdien lodiera [3, 5] mm-koa da.

Bariantzaren balioa hurrengoa da:

$$s_x^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{1^2 \cdot 6 + 2^2 \cdot 7 + 3^2 \cdot 7 + 4^2 \cdot 6 + 5^2 \cdot 4}{30} - (2,83)^2 = 1,76 \text{ mm}^2.$$

Bariantzaren erro karratu positiboa kalkulatuz,

$$s = \sqrt{1,76} = 1,33 \text{ mm}$$

lor daiteke. Desbideratze tipikoak torlojuen batez besteko lodierarekiko erroreen batez bestekoaren erro karratua adierazten du eta bere balioa 1,33 mm-koa da.

1.5.2. ariketa

Hurrengo taulan 110 enpresaren fakturazioa azaltzen da:

Fakturazioa (milaka euro)	[0,10)	[10,20)	[20,30)	[30,40)	[40,50]
Enpresa-kopurua	15	25	32	23	15

a. Kalkula bedi gehien fakturatu den kantitatea zein den.

b. Zenbat enpresak fakturatzen dituzte gutxienez 32,17 mila euro? Adieraz ezazu emaitza maiztasun metatuen poligonoan.

Ebazpena

a. Adierazitako enpresen fakturazioa aldagai kuantitatibo jarraitua da. Hona hemen dagokion maiztasun absolutuen eta maiztasun metatuen taula:

$[l_i, l_{i+1})$	[0,10)	[10,20)	[20,30)	[30,40)	[40,50]
f_i	15	25	32	23	15
F_i	15	40	72	95	110

Bestalde, modaren kalkuluak gehien fakturatu den kantitatea adieraziko du.

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot d_i = 20 + \frac{\frac{32-25}{10}}{\left(\frac{32-25}{10}\right) + \left(\frac{32-23}{10}\right)} \cdot 10 = 20 + 4,375 = 24,375.$$

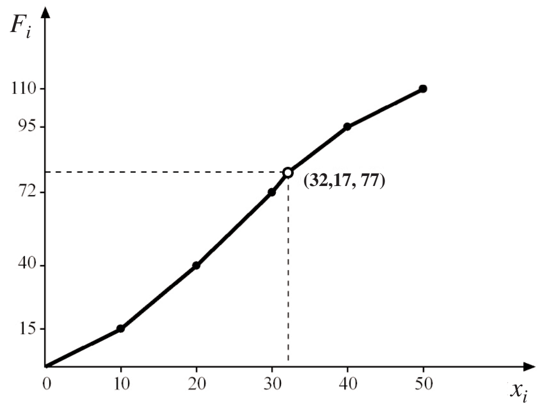
Beraz, maiztasun handieneko fakturazioa 24.375 eurokoa da.

b. Bigarren atal honetan eskatzen dena posizio-estatistiko bat zehaztea da. Horrela, $P_j = 32,17$ balioaren bidez % j enpresak 32,17 mila euro baino gutxiago fakturatzen dituztela adierazten da. Datuak ordezkatuz, pertzentilaren ordena zehaztuko da.

$$P_j = l_i + \frac{\frac{j \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 30 + \frac{\frac{j \cdot 110}{100} - 72}{23} \cdot 10 = 32,17 \quad \rightarrow \quad j = 69,99.$$

Beraz, $j \approx 70$ da, hots, enpresen % 70ek (77 enpresak, alegia) 32.170 euro baino gutxiago fakturatzen ditu. Ondorioz, enpresen beste % 30ak gutxienez 32.170 euro fakturatuko ditu, kasu horretan 33 enpresa izango direlarik.

Grafikoki ikus daitekeenez, (32,17, 77) puntuak 77 enpresak 32,17 mila euro baino gutxiago fakturatzen dituztela adierazten du.



1.11. irudia.

1.5.3. ariketa

Ingurumenari kalteak eragiteagatik, 1.000 egunetan izandako isun-kopurua neurtu da eta egun bakoitzean 0 eta 5 isun bitartean jaso dira, hurrengo taulan ikus daitekeen moduan:

Isun-kopurua	0	1	2	3	4	5
Egun-kopurua	?	260	150	190	100	90

a. Kalkula bitez batez bestekoa, mediana, moda eta desbideratze tipikoa.

b. Lor itzazu 1. eta 3. ordenako kuartilak.

c. Adieraz bitez barra-grafikoa eta maiztasun metatuen grafikoa.

Ebazpena

a. Guztira 1.000 egunetako isun-kopurua neurtu denez, taulan falta den datua $f_1 = 210$ da.

$$1.000 = f_1 + 260 + 150 + 190 + 100 + 90 \rightarrow f_1 = 210.$$

Hori kontuan hartuz,

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} =$$

$$= \frac{0 \cdot 210 + 1 \cdot 260 + 2 \cdot 150 + 3 \cdot 190 + 4 \cdot 100 + 5 \cdot 90}{1.000} = \frac{1.980}{1.000} = 1{,}98 \text{ isun}$$

jaso dira, batez beste.

Mediana2 da. Hau da egunen erdietan gehinez 2 isun eta beste egunen erdietan 2 eta 5 isun bitartean jaso dira.

Modaren balioa 1 da. Hau da, gehien jaso den isun-kopurua 1 da.

Bariantzaren balioa hurrengoa da:

$$s_x^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{1^2 \cdot 260 + 2^2 \cdot 150 + 3^2 \cdot 190 + 4^2 \cdot 100 + 5^2 \cdot 90}{1.000} - (1,98)^2 = 2,5 \text{ (isun)}^2.$$

Bariantzaren erro karratu positiboa kalkulatuz, desbideratze tipikoa 1,58 isunekoa dela ondorioztatzen da.

$$s = \sqrt{2,5} = 1,58 \text{ isun}.$$

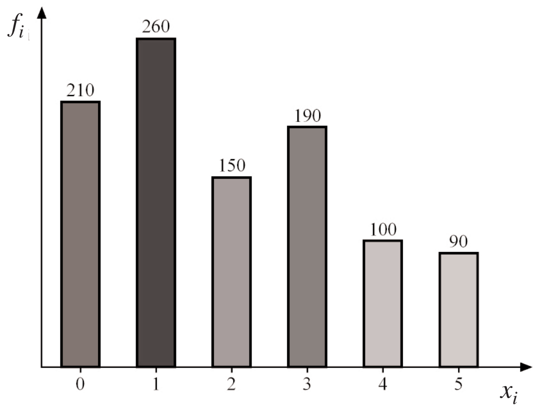
b. Kalkula ditzagun Q_1 eta Q_3 kuartilak. Horretarako osa ditzagun taulako datuak, maiztasun metatuen zutabea erantsiz.

x_i	0	1	2	3	4	5
f_i	210	260	150	190	100	90
F_i	210	470	620	810	910	1.000

Orduan, $\frac{25 \cdot n}{100} = \frac{25 \cdot 1.000}{100} = 250 < F_2$ denez, 1. ordenako kuartila $Q_1 = 1$ balioa da.

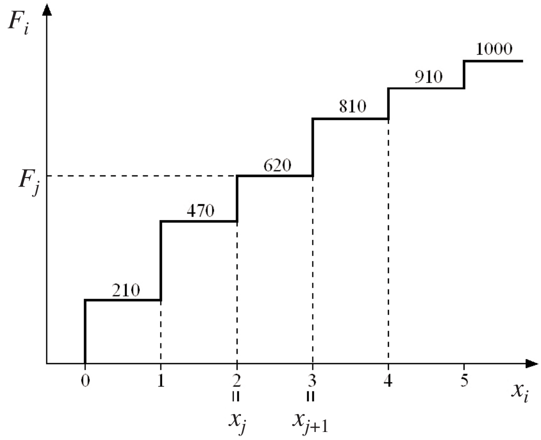
Era berean, $\frac{75 \cdot n}{100} = \frac{75 \cdot 1.000}{100} = 750 < F_4$ denez, 3. ordenako kuartila $Q_3 = 3$ balioa da.

c. Barra-grafikoa:



1.12. irudia.

Maiztasun metatuen grafikoa:



1.13. irudia.

1.5.4. ariketa

Ondoko taulan 80 langileren hileko soldatak azaltzen dira:

Soldatak (euro)	Ingeniariak (%)
1.500-1.560	15
1.560-1.600	15
1.600-1.680	25
1.680-1.740	20
1.740-1.800	15
1.800-1.860	10

a. Kalkula bedi langile horien zein proportziok irabazten duen hilero $\bar{x} + s$ euro baino gutxiago.

b. Lagin horretako 4 langileren soldata gutxienez *A* eurokoa da. Estatistiko egoki bat erabiliz, zehatz bedi *A* kantitatea.

c. Azal bedi aurreko ataleko *A* kantitatearen interpretazio grafikoa aldagaiaren histograman.

Ebazpena

a. Enuntziatuko datuei honako maiztasun-taula dagokie:

$[l_i, l_{i+1})$	h_i	x_i	f_i	F_i
[1.500,1.560)	0,15	1.530	12	12
[1.560, 1.600)	0,15	1.580	12	24
[1.600, 1.680)	0,25	1.640	20	44
[1.680, 1.740)	0,20	1.710	16	60
[1.740, 1.800)	0,15	1.770	12	72
[1.800, 1.860]	0,10	1.830	8	80

Batez bestekoaren kalkulua:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{1.530 \cdot 12 + 1.580 \cdot 12 + \ldots + 1.830 \cdot 8}{80} = \frac{133.360}{80} = 1.667 \text{ euro.}$$

Bariantzaren kalkulua:

$$s^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{(1.530)^2 \cdot 12 + (1.580)^2 \cdot 12 + \ldots + (1.830)^2 \cdot 8}{80} - (1.667)^2 = 8.751 \text{ (euro)}^2.$$

Ondorioz, desbideratze tipikoaren balioa hurrengoa da:

$$s = \sqrt{8.751} = 93,55 \text{ euro.}$$

$\bar{x}$ + s = 1.667+ 93,55 = 1.760,55 euro da. Orduan, hilero 1.760,55 euro baino gutxiago irabazten dituzten langileen proportzioa kalkulatu behar da. Horretarako P_k pertzentilaren ordena zehaztu behar da.

$$P_k = l_i + \frac{\frac{k \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i.$$

Kasu honetan, hurrengo ekuazioa ebatziko da:

$$1.760,55 = 1.740 + \frac{\frac{k \cdot 80}{100} - 60}{12} \cdot 60 \quad \rightarrow$$

$$k = 80,14.$$

Beraz, 80 langileen % 80,14k hilero 1.760,55 euro baino gutxiago irabazten ditu.

b. Erabiliko den estatistikoa berriro ere pertzentila da. 4 langilek laginaren % 5 osatzen dutenez, erantzuna 95. ordenako pertzentilaren bidez kalkula daiteke.

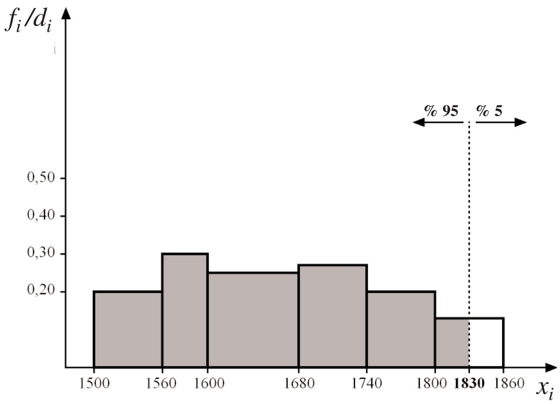
$$P_{95} = l_i + \frac{\frac{95 \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 1.800 + \frac{\frac{95 \cdot 80}{100} - 72}{8} \cdot 60 = 1.830 \text{ euro.}$$

Hau da, 4 langileren hileko soldatak gutxienez A = 1.830 eurokoak dira.

c. Aldagai estatiskoari dagokion histograman abzisa-ardatzean klaseak eta ordenatu-ardatzean f_i/d_i balioak adieraziko dira, non f_i gaiak maiztasun absolutuak eta d_i balioak klaseen anplitudeak diren.

$[l_i, l_{i+1})$	f_i	d_i	f_i/d_i
[1.500,1.560)	12	60	0,20
[1.560, 1.600)	12	40	0,30
[1.600, 1.680)	20	80	0,25
[1.680, 1.740)	16	60	0,27
[1.740, 1.800)	12	60	0,20
[1.800, 1.860]	8	60	0,13

Histogramaren adierazpen grafikoa honako hau da:



1.14. irudia.

Bertan ikus daitekeenez, A = 1.830 balioa posizio estatistiko bat da, 95. ordenako pertzentila hain zuzen ere. Beraz, kasu honetan banaketaren % 95 geratzen da 1.830 balioaren ezkerrean eta beste % 5a bere eskuinean dago.

1.5.5. ariketa

Hurrengo datuek pertsona baten iazko elektrizitate-gastuak (euro) adierazten dituzte:

Urt.	Ots.	Mar.	Api.	Mai.	Eka.	Uzt.	Abu.	Ira.	Urr.	Aza.	Abe.
35,6	36	36,2	35,6	27,6	25,8	27,6	32,4	32,4	32,4	32,4	33,6

a. Nolakoa da aldagai hau, kuantitatiboa ala kualitatiboa?

b. Zein da urteko batez besteko elektrizitate-gastua?

c. Lor bedi desbideratze tipikoaren balioa.

d. Kalkula bedi medianaren balioa eta adierazi bere esanahia.

Ebazpena

a. Pertsona baten iazko elektrizitate-gastuak zenbakizko balioak direnez, aldagai kuantitatiboa da. Gainera, balio puntualak daudenez, aldagai kuantitatibo diskretua dela erantsi daiteke.

Halaber, aldagaiari hurrengo maiztasun-taula dagokio:

x_i	25,8	27,6	32,4	33,6	35,6	36	36,2
f_i	1	2	4	1	2	1	1
F_i	1	3	7	8	10	11	12

b. Urteko batez besteko elektrizitate-gastua:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{25,8 \cdot 1 + 27,6 \cdot 2 + \ldots + 36,2 \cdot 1 +}{12} = \frac{387,6}{12} = 32,3 \text{ euro.}$$

Bariantzaren kalkulua:

$$s_x^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{(25,8)^2 \cdot 1 + (27,6)^2 \cdot 2 + \ldots + (36,2)^2 \cdot 1}{12} - (32,3)^2 = 11,57 \text{ (euro)}^2.$$

Desbideratze tipikoa, bariantzaren erro karratu positiboaren emaitza da:

$$s = \sqrt{11,57} = 3,4 \text{ euro.}$$

d. Mediana kalkulatzeko, kasu honetan n = 12 bikoitia denez, mediana seigarren eta zazpigarren balioen erdiko puntua da. Ondorioz,

$$M_e = 32{,}4 \text{ euro}$$

da. Horrek esan nahi du, sei hilabetetako gastuak gehienez 32,4 eurokoak eta beste sei hilabeteetako gastuak gutxienez 32,4 eurokoak izan direla.

1.5.6. ariketa

Lan-talde batek produkzio-plantako produkzio-unitate bakoitza egiteko behar izan den denbora neurtzen du. Produkzio-planta batean 30 egunetako datuak bildu dira eta datu horien bidez egun bakoitzean produkzio-unitateak egiteko guztira behar izan dituzten orduak neurtu dira:

128	119	96	97	124	128	142	98	108	120
114	109	124	132	97	138	133	136	120	112
144	128	103	135	114	109	100	111	131	113

Datuok anplitude bereko bost klasetan bilduz:

a. Kalkula bitez produkzio-unitate horien ekoizpenerako batez besteko denbora eta maiztasun handieneko ordu-kopurua.

b. Alborapen-koefizientearen balioaren arabera, zer esan daiteke?

Ebazpena

a. Emandako banaketa, datuak bost klasetan bilduz, honela adieraz daiteke:

$[l_i, l_{i+1})$	x_i	f_i
[95, 105)	100	6
[105, 115)	110	8
[115, 125)	120	5
[125, 135)	130	6
[135, 145)	140	5

Produkzio-unitateen ekoizpenerako, batez beste, 118,67 ordu behar dira:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{3.560}{30} = 118{,}67 \text{ ordu.}$$

Bestalde, modaren balioa kalkulatuko da.

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot d_i = 105 + \frac{\frac{8-6}{10}}{\left(\frac{8-6}{10}\right) + \left(\frac{8-5}{10}\right)} \cdot 10 = 105 + 4 = 109 \text{ ordu.}$$

Orduan, maiztasun handienez 109 ordu behar dira pieza horiek ekoizteko.

b. Alborapen-koefizientearen balioa zehazteko, batez bestekoa eta modaren balioak aurreko atalean lortu direnez, orain desbideratze tipikoaren balioa kalkulatu behar da.

Lehenago, bariantzaren kalkulua egin behar da:

$$s^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 = \frac{428.200}{30} - (118,67)^2 = 190,76 \text{ (ordu)}^2.$$

Orduan, desbideratze tipikoa $s = \sqrt{190,76} = 13,81$ ordu da.

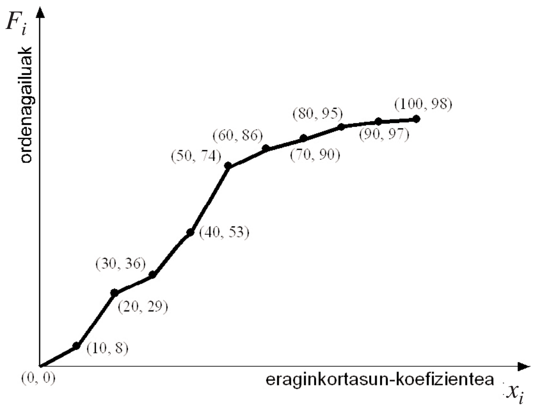
Bukatzeko, alborapen-koefizientearen balioa kalkulatuko da.

$$v = \frac{\bar{x} - M_o}{s} = \frac{118,67 - 109}{13,81} = 0,7.$$

Alborapen-koefizientea positiboa denez, banaketa eskuinerantz alboratua dela esan daiteke.

1.5.7. ariketa

Ordenagailuen eraginkortasuna neurtzeko 0 eta 100 bitarteko koefizienteak ezarri dira. Horrela, ondoko grafikoan 98 ordenagailuren eraginkortasuna neurtzen duen koefizientea maiztasun metatuen poligonoaren bidez adierazten da:



1.15. irudia.

a. Eraiki bedi aldagai kuantitatibo jarraitu honi dagokion maiztasun absolutu eta metatuen taula.

b. Kalkula itzazu bariantza, desbideratze tipikoa eta aldakuntza-koefizientearen balioak.

c. Jakinda 49 ordenagailuren eraginkortasunaren koefizientea x balioa baino txikiagoa dela, zenbatekoa da x balioa?

Ebazpena

a. Maiztasun metatuen grafikoan (l_{i+1}, F_i) puntuak adierazten dira, l_{i+1} balioa klasearen goi-muturra eta F_i gaia x_i balio bakoitzari dagokion maiztasun metatua izanik. Orduan, maiztasun absolutu eta metatuen taula osatzeko, aldagai kuantitatibo jarraitua hamar anplitudeko klasetan bana daiteke.

$[l_i, l_{i+1})$	x_i	f_i	F_i
[0, 10)	5	8	8
[10, 20)	15	21	29
[20, 30)	25	7	36
[30, 40)	35	17	53
[40, 50)	45	21	74
[50, 60)	55	12	86
[60, 70)	65	4	90
[70, 80)	75	5	95
[80, 90)	85	2	97
[90, 100)	95	1	98

b. Lehenengo batez bestekoaren balioa, eta gero, eskatutako estatistikoen balioak kalkulatuko dira.

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{3.630}{98} = 37,04.$$

Bariantza:

$$s^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 = \frac{177.450}{98} - (37,04)^2 = 438,69.$$

Desbideratze tipikoa:

$$s = \sqrt{438,69} = 20,94.$$

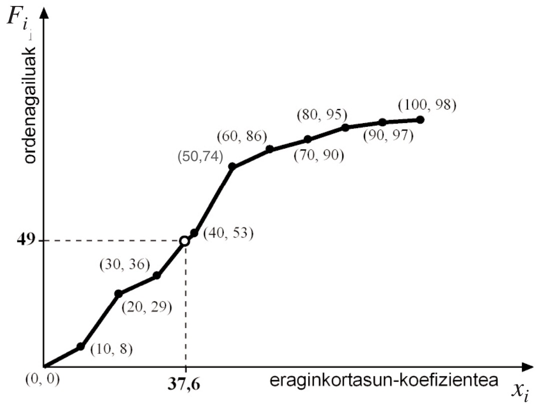
Aldakuntza-koefizientea:

$$CV_x = \frac{s_x}{\bar{x}} = \frac{20,94}{37,04} = 0,565.$$

c. P_k pertzentilaren bidez, laginaren % k ordenagailuk duten eraginkortasuna P_k balioa baino txikiagoa dela adierazten da. Kasu honetan, 49 ordenagailu laginaren erdia direla ikus daiteke. Beraz, problema mediana kalkulatuz ebatz dezakegu:

$$M_e = l_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot d_i = 30 + \frac{49 - 36}{17} \cdot 10 = 37,6.$$

Ondorioz, ordenagailu erdien eraginkortasunaren koefizientea 37,6 baino txikiagoa da.



1.16. irudia.

1.5.8. ariketa

A termometroarekin neurtutako tenperaturen erroreek 1 ºC-eko batez bestekoa eta 0,01 ºC-ko desbideratze tipikoa erakutsi dituzte. Bestalde, B termometroarekin neurtutako tenperaturen erroreek 0,2 ºC-ko batez bestekoa eta 0,08 ºC-ko desbideratze tipikoa adierazi dituzte.

a. Zein termometro da erlatiboki zehatzagoa, A ala B?

b. Gaur goizeko zortzietan 8 ºC-ko tenperatura zegoenean, A termometroak 9,5 ºC eta B termometroak 9,6 ºC neurtu dituzte. Zein termometrok erakutsi du erlatiboki errore txikiagoa goizeko zortzietan?

Oharra: errorea = $|$benetako tenperatura – neurtutako tenperatura$|$

Ebazpena

a. Erlatiboki zehatzagoa den termometroa zehazteko, aldakuntza-koefizienteak kalkulatuko dira:

$$CV_A = \frac{s_A}{\bar{x}_A} = \frac{0{,}01}{1} = 0{,}01,$$

$$CV_B = \frac{s_B}{\bar{x}_B} = \frac{0{,}08}{0{,}2} = 0{,}4.$$

B termometroarekin egindako neurketei dagokien aldakuntza-koefizienteak % 40ko aldakuntza adierazten du. A termometroari dagokionez, aldiz, % 1eko aldakuntza du. Ondorioz, $CV_A < CV_B$ denez, A termometroa zehatzagoa da.

b. Errorea benetako tenperaturaren eta neurtutako tenperaturaren arteko diferentziaren balio absolutua denez, A eta B termometroek goizeko zortzietan jasotako neurrien erroreak

$$x_A = |8 - 9{,}5| = 1{,}5,$$

$$x_B = |8 - 9{,}6| = 1{,}6$$

dira, hurrenez hurren. Tipifikazioa aplikatuz, balio bakoitzaren posizio erlatiboa kalkula daiteke:

$$z_A = \frac{x_A - \bar{x}_A}{s_A} = \frac{1{,}5 - 1}{0{,}01} = 50,$$

$$z_B = \frac{x_B - \bar{x}_B}{s_B} = \frac{1{,}6 - 0{,}2}{0{,}08} = 17{,}50$$

$z_B < z_A$ denez, B termometroaren goizeko zortzietako neurriaren errorea erlatiboki txikiagoa da.

1.5.9. ariketa

Lan-talde batean kualifikazioa eta produktibitatea kalifikatu dira: kualifikazioaren batez besteko puntuazioa 10 eta desbideratze tipikoa 2,2 eta produktibitatearen batez besteko puntuazioa 30 eta desbideratze tipikoa 13 izan dira. Marta eta Mikel lan-talde horretako bi partaide dira. Martak kualifikazioan 12 eta produktibitatean 35 puntuazioak lortu ditu. Mikelen puntuazioak kualifikazioan eta produktibitatean 11 eta 38 dira, hurrunez hurren. Kualifikazioa eta produktibitatea batera hartuz, globalki, nork du puntuazio hobea?

- -

Ebazpena

Enuntziatuko datuen arabera, bi alderdi neurtzen dira:

Kualifikazioa	Produktibitatea
$\bar{x}_1 = 10$	$\bar{x}_2 = 30$
$s_1 = 2,2$	$s_2 = 13$

Martaren eta Mikelen puntuazioak kualifikazioan eta produktibitatean:

	Kualifikazioa	Produktibitatea
Mikel	$x_1^1 = 11$	$x_2^1 = 38$
Marta	$x_1^2 = 12$	$x_2^2 = 35$

Bi alderdiak batera tratatzeko puntuazioak tipifikatu behar dira.

$$\text{Mikel: } z_1^1 = \frac{11-10}{2,2} = 0,45, \qquad z_2^1 = \frac{38-30}{13} = 0,62,$$

$$\text{Marta: } z_1^2 = \frac{12-10}{2,2} = 0,91, \qquad z_2^2 = \frac{35-30}{13} = 0,38.$$

Puntuazio globala lortzeko, bi puntuazioen batez bestekoa kalkulatuko da.

$$\text{Mikelen puntuazio globala: } \frac{z_1^1 + z_2^1}{2} = \frac{0,45+0,62}{2} = 0,535,$$

$$\text{Martaren puntuazio globala: } \frac{z_1^2 + z_2^2}{2} = \frac{0,91+0,38}{2} = 0,645.$$

Ondorioz, Martaren puntuazio globala Mikelena baino hobea da.

1.5.10. ariketa

Hogei hiritako hileko euri-kantitateak (l/m^2) neurtu dira eta emaitzak anplitude ezberdinetako sei klasetan bildu dira. Anplitudeak honakoak dira: $d_1 = 8$, $d_2 = 8$, $d_3 = 4$, $d_4 = 4$, $d_5 = 8$, $d_6 = 10$, non d_i balioa *i*. klasearen anplitudea den. Bestalde, dagozkien maiztasun erlatibo metatuak hauek dira, hurrenez hurren: H_1= 0,10, H_2 = 0,10, H_3 = 0,55, H_4 = 0,75, H_5 = 0,95, H_6 = 1.

a. Jaso den euri-kantitate txikiena 65 l/m^2-koa dela jakinda, kalkula bedi zein den hogei hirietako batez besteko hileko euri-kantitatea.

b. Egia al da hamar hiritako euri-kantitatea 73 l/m^2 eta 86 l/m^2 bitartekoa izan zela?

Ebazpena

a. Hasteko, euri-kantitate txikiena x_i balio txikiena 65 denez, eta klaseen anplitudeak ezagunak direnez, klaseak eta klase-markak zehatz daitezke. Bestealde, H_i maiztasun erlatibo metatuak ezagunak direnez eta $n = 20$ denez, maiztasun erlatiboak, maiztasun absolutuak eta maiztasun absolutu metatuak kalkula daitezke. Horretarako kontuan hartuko diren adierazpenak hurrengoak dira: $h_i = H_i - H_{i-1}$, $F_i = 20H_i$ eta $f_i = F_i - F_{i-1}$.

$[l_i, l_{i+1})$	x_i	H_i	h_i	F_i	f_i
[65, 73)	69	0,10	0,10	2	2
[73, 81)	77	0,10	0	2	0
[81, 85)	83	0,55	0,45	11	9
[85, 89)	87	0,75	0,20	15	4
[89, 97)	93	0,95	0,20	19	4
[97, 107)	102	1	0,05	20	1

Batez besteko hileko euri-kantitatea hauxe da:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{69 \cdot 2 + 77 \cdot 0 + 83 \cdot 9 + 87 \cdot 4 + 93 \cdot 4 + 102 \cdot 1}{20} = \frac{1.707}{20} = 85{,}35 \text{ l/m}^2.$$

b. Pertzentilak erabiliz, 10. ordenako pertzentila kalkulatuz, bi hiritako euri-kantitateak 65-73 l/m^2-koak izan zirela ikus daiteke.

$$73 = P_k = l_i + \frac{\frac{k \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 65 + \frac{\frac{k \cdot 20}{100} - 0}{2} \cdot 8 \quad \rightarrow \quad k = 10.$$

Era berean, 60. ordenako pertzentila kalkulatuz, 12 hiritako euri-kantitateak 86 l/m^2 baino txikiagoak izan zirela nabari da.

$$86 = P_m = l_j + \frac{\frac{m \cdot n}{100} - F_{j-1}}{f_j} \cdot d_j = 85 + \frac{\frac{m \cdot 20}{100} - 11}{4} \cdot 4 \quad \rightarrow \quad m = 60.$$

Ondorioz, hamar hiritako euri-kantitateak 73-86 l/m^2 bitartekoak izan ziren.

1.5.11. ariketa

Berrogei pertsonari test psikometriko bat aplikatu zaie eta hurrengo puntuazioak lortu dira:

Puntuazioak	0-4	4-8	8-12	12-16	16-19
Pertsona-kopurua	6	15	10	5	4

a. Kalkula bitez mediana eta moda. Azaldu beraien esanahia.

b. Erabil ezazu estatistiko egoki bat 13 puntu baino gutxiago lortu duen pertsona-kopurua zehazteko.

Ebazpena

a. Biz X = "test psikometrikoaren puntuazioak" aldagai estatistikoa. Taulako puntuazioak klasetan bil daitezke. Ondoren maiztasun metatuak kalkulatuko dira.

$[l_i, l_{i+1})$	[0, 4)	[4, 8)	[8, 12)	[12,16)	[16,19)
f_i	6	15	10	5	4
F_i	6	21	31	36	40

Mediana da bere eskuinaldean banaketaren % 50 uzten duen balioa.

$$M_e = l_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot d_i = 4 + \frac{\frac{40}{2} - 6}{15} \cdot 4 = 7,73.$$

Beraz, lagineko pertsonen erdiek ez dute gainditu 7,73 puntuazioa.

Moda, maiztasun absolutu handieneko balioa da. Kasu honetan, hurrengo balioa du:

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot d_i = 4 + \frac{\frac{15-6}{4}}{\frac{15-6}{4} + \frac{15-10}{4}} \cdot 4 = 6,57.$$

Beraz, gehien lortu den puntuazioa 6,57 izan da.

b. P_k pertzentilaren bidez, banaketaren % k pertsonak lortu duten puntuazioa P_k baino txikiagoa dela adierazten da.

$$P_k = l_i + \frac{\frac{k \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 12 + \frac{\frac{k \cdot 40}{100} - 31}{5} \cdot 4 = 13,$$

$$k = 80,625.$$

Ondorioz, lagineko pertsonen % 80,625ek lortu duen puntuazioa 13 baino txikiagoa da, hots, 32 pertsonak 13 baino puntuazio txikiagoa lortu dute.

1.5.12. ariketa

Gasteizen eta Bilbon kokatutako bi enpresa ekoizpena aztertzen ari dira. Ondoko taulan egunean ekoitzitako pieza-kopurua neurtu da:

Pieza-kopurua	Enplegatu-kopurua Gasteiz	Enplegatu-kopurua Bilbo
0-30	94	85
30-50	140	160
50-70	160	160
70-90	98	89
90-100	8	6

Enpresen plangintzetan, ekoizpenaren arabera (ekoizpen txikienetik handienera), lau langile-kategoria ezarriko dira: A kategorian enplegatuen % 65, B kategorian enplegatuen % 20, C kategorian enplegatuen % 10 eta D kategorian enplegatuen % 5.

a. D kategoriako langile izateko ekoizpen minimoa, non da handiagoa, Gasteizen ala Bilbon?

b. Irudika bitez kasu bakoitzeko histograma eta maiztasun metatuen poligonoa, aurreko atalean kalkulatutako estatistikoen balioak adieraziz.

Ebazpena

a. Hasteko, hona hemen kasu bakoitzari dagokion taula:

Gasteiz				Bilbo		
f_i^G	f_i^G / d_i	F_i^G	$[l_i, l_{i+1})$	f_i^B	f_i^B / d_i	F_i^B
94	3,13	94	[0, 30)	85	2,83	85
140	7	234	[30, 50)	160	8	245
160	8	394	[50, 70)	160	8	405
98	4,9	492	[70, 90)	89	4,45	494
8	0,8	500	[90, 100)	6	0,6	500

Lehenengo, kalkula dezagun Gasteizko kasurako 95. ordenako pertzentila.

$\frac{500 \cdot 95}{100} = 475$ denez, orduan $P_{95}^{G} \in [70, 90)$ dago.

$$P_{95}^{G} = 70 + \frac{475 - 394}{98} \cdot 20 = 86,53.$$

Ondorioz, Gasteizko langileen artean 87 eta 100 pieza bitartean ekoizten dituzten langileak D kategorian daude.

Orain, kalkula dezagun Bilboko langileen kasurako 95. ordenako pertzentilaren balioa.

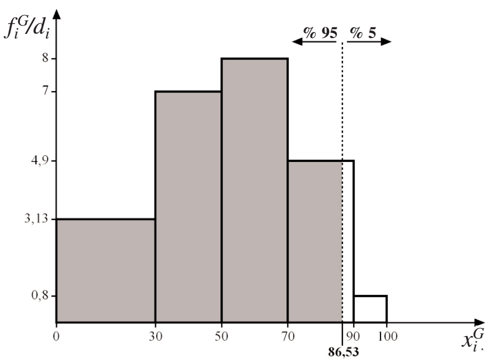
$\frac{500 \cdot 95}{100} = 475$ denez, orduan $P_{95}^{B} \in [70, 90)$ dago.

$$P_{95}^{B} = 70 + \frac{475 - 405}{89} \cdot 20 = 85,73.$$

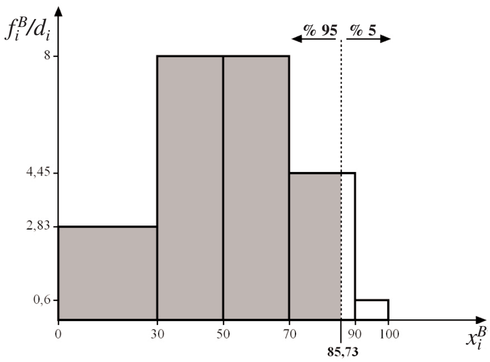
Hau da, Bilboko langileen artean 86 eta 100 pieza bitartean ekoizten dituztenak daude D kategorian.

Gasteizko kasuan 95. ordenako pertzentilaren balioa handiagoa denez, D kategoriako langile izateko ekoizpen minimoaren balioa Gasteizen handiagoa da.

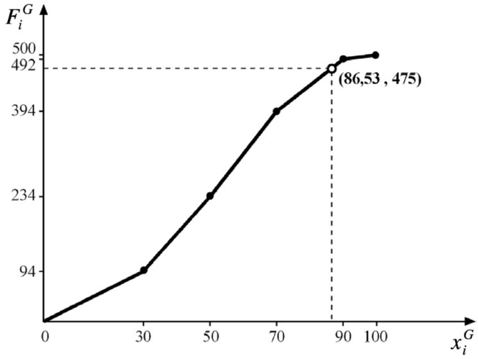
b. 1.17. eta 1.18. irudietan kasu bakoitzeko histogramak adierazi dira. Era berean, 1.19. eta 1.20. irudietan Gasteizko eta Bilboko kasuetarako maiztasun metatuen poligonoak nabari daitezke.



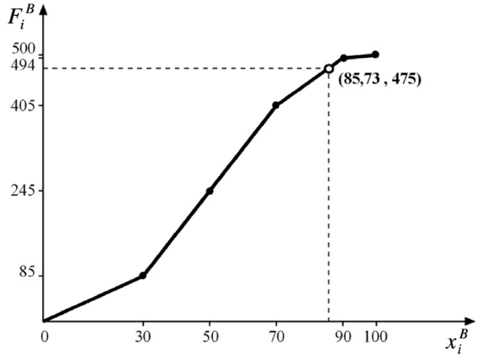
1.17. irudia



1.18. irudia



1.19. irudia



1.20. irudia

1.5.13. ariketa

Ondoko taulan 40 lanpararen iraupenak azaltzen dira:

Iraupena (urte)	2	6	10	14	18
Lanpara-kopurua	6	15	10	5	4

a. Kalkula bedi lanparen batez besteko iraupena, mediana eta desbideratze tipikoa.

b. Beste mota bateko lanparen batez besteko iraupena 11 urte eta desbideratze tipikoa 5,5 urte izan dira. Zein datu daude erlatiboki kontzentratuago?

c. Enuntziatuko 40 lanparetako batek 10 urte eta b) ataleko beste lanpara batek 12 urte iraun dute. Zein lanparak izan du erlatiboki iraupen handiagoa?

Ebazpena

a. Biz X = "lanparen iraupena" aldagai aleatorio diskretua. Zehatz dezagun modalitateaz, maiztasun absolutuaz eta maiztasun metatuaz osaturiko taula:

x_i	2	6	10	14	18
f_i	6	15	10	5	4
F_i	6	21	31	36	40

Lanparen batez besteko iraupena honela kalkulatuko da:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{344}{40} = 8{,}6 \text{ urte.}$$

Medianaren kalkulua:

n = 40 bikoitia denez, mediana 20. eta 21. posizioetako balioen erdiko puntua da. Beraz, M_e = 6 urte da. Ondorioz, lanparen erdiek 2 edo 6 urteko iraupena eta beste lanparen erdiek 6, 10, 14 edo 18 urteko iraupena dute.

Bariantza:

$$s_x^2 = \frac{\sum_{i=1}^{5} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{2^2 \cdot 6 + 6^2 \cdot 15 + 10^2 \cdot 10 + 14^2 \cdot 5 + 18^2 \cdot 4}{40} - (8,6)^2 = 22,04 \text{ (urte)}^2.$$

Bariantzaren erro karratu positiboa kalkulatuz, desbideratze tipikoaren balioa lortuko da:

$$s = \sqrt{22,04} = 4,69 \text{ urte.}$$

b. Kalkula ditzagun aldakuntza-koefizienteen balioak bi lanpara-multzoetarako.

Lehenengo kasuan, aldakuntza-koefizientearen balioa hauxe da:

$$CV_x = \frac{s_x}{\bar{x}} = \frac{4,69}{8,6} = 0,545.$$

Bigarren atal honetako lanparei dagokien aldakuntza-koefizientea honakoa da:

$$CV_y = \frac{s_y}{\bar{y}} = \frac{5,5}{11} = 0,5.$$

Ondorioz, bigarren multzoko lanparen iraupenari dagokion aldakuntza-koefizientea txikiagoa denez, datu hauek erlatiboki kontzentratuago daude.

c. Galdera honi erantzuna emateko tipifikazioa aplikatuko da.

I. multzoko lanparak	II. multzoko lanparak
$\bar{x}_1 = 8,6$	$\bar{x}_2 = 11$
$s_1 = 4,69$	$s_2 = 5,5$
$x_1 = 10$	$x_2 = 12$

I. multzoko lanpararen balio tipifikatua: $z_1 = \dfrac{x_1 - \bar{x}_1}{s_1} = \dfrac{10 - 8,6}{4,69} = 0,299.$

II. multzoko lanpararen balio tipifikatua: $z_2 = \dfrac{x_2 - \bar{x}_2}{s_2} = \dfrac{12 - 11}{5,5} = 0,182.$

$z_1 > z_2$ denez, lehenengo multzotik hartutako lanparak erlatiboki gehiago irauten du.

1.5.14. ariketa

Hona hemen marka ezagun bateko 27 autoren gasolina-kontsumoa (litro) 100 km-ko ibilbidean:

2,1	3,3	4,4	3,0	4,0	5,0	2,7	2,6	4,8
4,7	2,8	4,8	3,9	2,3	3,8	2,8	3,0	3,7
3,3	4,4	3,1	4,0	3,7	2,5	2,7	5,1	4,7

Aldagai kuantitatibo jarraitu hau [2,1, 5,1] tartean definituta dagoela jakinda, kalkula bitez batez bestekoa, mediana, moda, desbideratze tipikoa, aldakuntza-koefizientea, Q_1 eta Q_3 kuartilak, heina, kuartilarteko heina, alborapena eta kurtosia.

Ebazpena

Egin dezagun azterketa aldagai kuantitatiboa era jarraituan tratatuz. Heina R = 5,1 − 2,1 = 3 denez eta klase-kopurua $k = \sqrt{n} = \sqrt{27} \approx 5$ har daitekeenez, datuak $\frac{R}{k} = \frac{3}{5} = 0,6$ anplitudeko 5 klasetan banatuko dira. Argudio hauen arabera, hona hemen aldagai kuantitatibo jarraituari dagokion maiztasun-taula:

$[l_i, l_{i+1})$	x_i	f_i	F_i	h_i	H_i
[2,1, 2,7)	2,4	4	4	0,15	0,15
[2,7, 3,3)	3,0	7	11	0,26	0,41
[3,3, 3,9)	3,6	5	16	0,18	0,59
[3,9, 4,5)	4,2	5	21	0,18	0,78
[4,5, 5,1]	4,8	6	27	0,22	1

Batez bestekoa:

$$\overline{x} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{2,4 \cdot 4 + 3,0 \cdot 7 + 3,6 \cdot 5 + 4,2 \cdot 5 + 4,8 \cdot 6}{27} = \frac{98,4}{27} = 3,64 \text{ litro.}$$

Mediana: $F_3 = 16 > \frac{27}{2} = 13,5$ denez, orduan $M_3 \in [3,3, 3,9)$ dago.

$$M_e = l_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot d_i = 3,3 + \frac{13,5 - 11}{5} \cdot 0,6 = 3,6 \text{ litro.}$$

Moda:

Maiztasun handiena $f_2 = 7$ da. Orduan, moda [2,7, 3,3) tartean dago.

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot d_i = 2,7 + \frac{\dfrac{7-4}{0,6}}{\left(\dfrac{7-4}{0,6}\right) + \left(\dfrac{7-5}{0,6}\right)} \cdot 0,6 = 3,06 \text{ l.}$$

Desbideratze tipikoa:

Desbideratze tipikoa bariantzaren erro karratu positiboa denez, lehenengoz bariantzaren balioa lortuko da.

$$s_x^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i}{n} - (\bar{x})^2 =$$

$$= \frac{(2,4)^2 \cdot 4 + 3^2 \cdot 7 + (3,6)^2 \cdot 5 + (4,2)^2 \cdot 5 + (4,8)^2 \cdot 6}{27} - (3,64)^2 = 0,72 \text{ (litro)}^2.$$

Beraz, desbideratze tipikoaren balioa hurrengoa da:

$$s = \sqrt{0,72} = 0,85 \text{ litro.}$$

Aldakuntza-koefizientea:

$$CV = \frac{s_x}{\bar{x}} = \frac{0,85}{3,64} = 0,23.$$

Q_1 eta Q_3 kuartilen kalkulua:

$$F_2 = 11 > \frac{25 \cdot n}{100} = \frac{25 \cdot 27}{100} = 6,75 \text{ denez, orduan } Q_1 \in [2,7, 3,3) \text{ dago.}$$

$$Q_1 = l_i + \frac{\dfrac{25 \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 2,7 + \frac{6,75 - 4}{7} \cdot 0,6 = 2,94 \text{ litro.}$$

$$F_4 = 21 > \frac{75 \cdot n}{100} = \frac{75 \cdot 27}{100} = 20,25 \text{ denez, orduan } Q_3 \in [3,9, 4,5) \text{ dago.}$$

$$Q_3 = l_i + \frac{\dfrac{75 \cdot n}{100} - F_{i-1}}{f_i} \cdot d_i = 3,9 + \frac{20,25 - 16}{5} \cdot 0,6 = 4,41 \text{ litro.}$$

Heina:

$$R = max(x_i) - min(x_i) = 5,1 - 2,1 = 3.$$

Kuartilarteko heina:

$$Q_3 - Q_1 = 4{,}41 - 2{,}94 = 1{,}47.$$

Alborapen-koefizientea:

$$v = \frac{\bar{x} - M_o}{s_X} = \frac{3,64 - 3,06}{0,85} = 0,68.$$

Koefiziente horren balioa positiboa denez, banaketa eskuinerantz alboratua da.

Kurtosia:

$$g_2 = \frac{m_4}{m_2^2}.$$

Lehenengo, laugarren ordenako momentuaren balioa lortuko da.

$$m_4 = \frac{\sum_{i=1}^{k} \left(x_i - \bar{x}\right)^4 f_i}{n} = \frac{\left(2,4 - 3,64\right)^4 \cdot 4 + \left(3 - 3,64\right)^4 \cdot 7 + \ldots + \left(4,8 - 3,64\right)^4 \cdot 6}{27} =$$

$$= \frac{21,97}{27} = 0,81 \text{ (litro)}^4.$$

Bestalde, bariantzaren balioa $m_2 = 0{,}72$ (litro)^2 denez, orduan $m_2^2 = 0,52$ (litro)^4 da.

Ondorioz, kurtosiaren balioa lortuko da:

$$g_2 = \frac{0,81}{0,52} = 1,56.$$

Kurtosi-koefizientearen balioa 0 baino handiagoa denez, banaketa leptokurtikoa da.

1.6. ARIKETA PROPOSATUAK

1.6.1. ariketa

Airearen kutsatzaile batzuen kontzentrazioen arabera, 1, 2, 3, 4 eta 5 indizeak adierazten dira, indize horiek airearen kalitatea hurrengo eran definitzen dutelarik:

Indizea	Airearen kalitatea
1	Oso ona
2	Ona
3	Egokia
4	Txarra
5	Oso txarra

Hurrengo zenbakiek hemezortzi egunetako airearen kalitatea adierazten dute:

1	2	2	3	3	2	3	3	4
3	1	3	1	5	4	3	2	3

a. Nolakoa da aldagaia?

b. Adieraz bitez barra-grafikoa eta sektore-diagrama.

c. Kalkula bitez batez besteko indizea eta mediana. Azaldu esanahia.

1.6.2. ariketa

X aldagai estatistikoari hurrengo taula dagokio:

x_i	f_i	F_i	h_i
10		2	
12		5	0,12
15	3		
20		13	
35	6		0,24
38			0,24

a. Osa bedi taula.

b. Eraiki bitez barra-grafikoa eta maiztasun metatuen diagrama.

b. Kalkula bitez mediana eta moda.

1.6.3. ariketa

Aleazio forjatuko 40 barra mozteko hurrengo bihurdurak behar izan dira:

33	21	32	44	35	22	40	36	22	37
20	37	42	31	23	44	32	30	44	44
42	35	40	36	32	31	37	43	24	40
25	30	26	35	33	41	25	44	36	27

a. Bil itzazu balio hauek anplitude bereko sei klasetan eta osa ezazu maiztasun-taula.

b. Adieraz bitez histograma eta maiztasun metatuen poligonoa.

c. Konpara bitez heina eta kuartilarteko heina.

1.6.4. ariketa

Berrogei mekanikariren gaitasuna neurtzeko, proba berezi bat burutu da. Proba horretako puntuazioak hurrengoak izan dira:

Aptitudea	37	37,20	37,50	38	38,10	38,50
Mekanikari-kopurua	1	5	13	6	10	5

a. Kalkula bitez batez bestekoa, desbideratze tipikoa, mediana eta moda.

b. Lor bitez 65. eta 90. ordenako pertzentilak. Zer adierazten dute kasu honetan?

1.6.5. ariketa

Hona hemen makina-multzo batek egun batean kontsumitzen duen erregai-kantitatea:

Erregai-kantitatea (litro)	Makina-kopurua
[10, 20)	8
[20, 30)	12
[30, 40)	10
[40,50)	22
[50,60)	13

a. Irudika bitez histograma eta maiztasun metatuen poligonoa.

b. Kalkula bitez heina, kuartilarteko heina, desbideratze tipikoa, eta aldakuntza-koefizientea.

c. Zein bi pertzentil hurbilen artean dago 42,5 litroko erregai-kantitatea?

1.6.6. ariketa

Berrogei hiritako urteko batez besteko tenperaturak (ºC) hauexek dira:

Tenperaturak	10	14	17	20	25
Hiri-kopurua	5	10	15	7	3

a. Kalkula bedi H_4 maiztasun erlatibo metatua eta adierazi beraren esanahia kasu honetan.

b. Kalkula bitez batez besteko tenperatura, mediana, moda eta desbideratze tipikoa.

c. Zein da 65. ordenako pertzentilaren balioa? Zer esan nahi du?

d. Beste hogeita bost hiri hartu dira, batez besteko tenperatura 18 ºC eta desbideratze tipikoa 4,2 ºC izanik. Zein hirik dituzte tenperatura erlatiboki kontzentratuagoak, lehenengo berrogeiek edo oraingo hogeita bostek?

1.6.7. ariketa

Ondoko taulan hogei enpresaren urteko irabaziak eta dagozkien maiztasun erlatiboak azaltzen dira:

Irabaziak (milaka euro)	Enpresak (%)
[65, 81)	10
[81, 85)	45
[85, 89)	20
[89, 97)	20
[97, 107)	5

a. Irudika bitez histograma eta maiztasun metatuen poligonoa.

b. 65 mila euroko irabazia minimoa izanda, zein da enpresen erdiek irabazitako kantitate maximoa urte horretan? Lortutako emaitza adieraz bedi histograman.

c. Enpresak errentagarria izateko urtean gutxienez 95 mila euro irabazi behar baditu, hauetariko zenbat dira errentagarriak? Adieraz bedi lortutako emaitza maiztasun metatuen poligonoan.

1.6.8. ariketa

Herrialde batean etxebizitzaren batez besteko alokairua 600 euro eta desbideratze tipikoa 20 euro dira. Beste herrialde batean, etxebizitzaren batez besteko alokairua 400 euro eta desbideratze tipikoa 40 euro dira. Lehenengo herrialdeko etxe bat 520 eurotan alokatu da eta bigarren herrialdeko beste etxe bat 335 eurotan alokatu da. Zein etxek du erlatiboki prezio baxuagoa?

1.6.9. ariketa

Hurrengo taulan industrien sufre dioxidoaren emisioak daude:

Sufre dioxidoa (tona)	Industria (%)
[6, 14)	15
[14, 22)	25
[22, 30)	40
[30, 42)	20

a. Industrien zein portzentajek emititzen ditu gutxienez 18 tona sufre dioxido?

b. Jakinda emisio minimoa 6 tonakoa dela, zein da industrien % 75ek emititzen duen sufre dioxidorako goi-muga?

1.6.10. ariketa

Neurgailu batek hezetasun erlatiboko hurrengo datuak jaso ditu:

29	12	34	17	26	28	23	28	29	34
17	23	12	26	34	29	12	34	28	17

a. Eraiki bedi aldagai estatistiko kuantitatibo diskretu honi dagokion maiztasun-taula.

b. Kalkula bitez batez bestekoa, mediana, moda, desbideratze tipikoa, aldakuntza-koefizientea, Q_1 eta Q_3 kuartilak, heina, kuartilarteko heina, alborapena eta kurtosia.

2. Probabilitatearen oinarriak

2.1. OINARRIZKO KONTZEPTUAK

Edozein fenomeno edo esperimentutan behaketak lortzen dira. Bi fenomeno-mota daude: **fenomeno deterministak** eta **fenomeno aleatorioak.** Fenomeno deterministetan emaitza zehatza aurresan daiteke. Adibidez, zenbait baldintza ezagunetan jazotzen diren lege fisikoak. Esperimentuaren emaitza zehatza ezin denean aurresan, aldiz, fenomeno aleatorioa dugu (herri batean bihar jasoko den euri-kantitatea, hurrengo hauteskundeen emaitza zehatzak…).

Esperimentu baten emaitza posible guztien multzoari **lagin-espazioa** deritzo. Lagin-espazioko edozein azpimultzo **gertaera** da. Lagin-espazioa adierazteko Ω letra eta gertaeretarako *A, B*... letra larriak erabiliko ditugu.

Ω lagin-espazioko elementu sinpleenari **oinarrizko gertaera** deritzo.

Ω lagin-espazioko *A* eta *B* gertaeren artean **A parte B** dela diogu $A \subset B$ betetzen denean.

$A \subset B$ eta $B \subset A$ gertatzen denean, *A* eta *B* **gertaera berdinak** dira, $A = B$ alegia.

A gertaeraren **kontrako gertaera**k, $\overline{A}$ adieraziko denak, *A* ez jazotzea esan nahi du.

A edo *B* gertaera jazotzea, $A \cup B$ **bildura** da. $A \cup \overline{A} = \Omega$ denez, $A \cup \overline{A}$ gertaerari **gertaera segurua** deritzo.

$A \cap B$ gertaerari **ebakidura** deritzo. $A \cap B = \varnothing$ bada, A eta B **gertaera bateraezinak** dira.

Bira $A_1, A_2, \ldots, A_n$ gertaerak. $\bigcup_{i=1}^{n} A_i = \Omega$ eta $A_i \cap A_j = \varnothing, \ \forall i \neq j$ betetzen badira, orduan $\{A_i\}_{i=1}^{n}$ multzoa **gertaera-sistema osoa** da.

Gertaeren arteko erlazioak multzo-teoriako erlazioen parekoak dira. Horrela, gertaeren arteko erlazioak ere Venn-en diagramen bidez azal daitezke.

2.2. PROBABILITATEAREN DEFINIZIOAK ETA PROPIETATEAK

Probabilitatearen definizioen artean, azpimarratzekoak dira definizio klasikoa eta definizio axiomatikoa.

2.2.1. Probabilitatearen definizioak

2.2.1.1. Definizio klasikoa

A gertaeraren probabilitatea, gertaera jazotzearen aldeko kasuen eta kasu posibleen arteko zatidura da. Laburbilduz,

$$P(A) = \frac{A \text{ gertatzearen aldeko kasu-kopurua}}{\text{Kasu posible guztien kopurua}}.$$

2.2.1.2. Definizio matematiko edo axiomatikoa

Bira esperimentu aleatorio bati lotutako edozein *A* gertaera eta Ω lagin-espazioa. $P(A)$ gaiari *A* gertaeraren probabilitatea deritzo, ondoko axiomak betetzen baditu:

i. $0 \leq P(A) \leq 1$

ii. $P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$, $A_i \cap A_j = \emptyset$, $\forall\, i \neq j$

iii. $P(\Omega) = 1$

Hiru axioma horiei **Kolgomorov-en axiomak** deritze.

2.2.2. Probabilitatearen propietateak

Aurreko axiometatik abiatuz, hurrengo propietateak egiazta daitezke:

1. $P(\overline{A}) = 1 - P(A)$
2. $A \subseteq B \Rightarrow P(A) \leq P(B)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. $P(A \cup B) \leq P(A) + P(B)$.

2.3. BALDINTZAZKO PROBABILITATEA

Bira Ω lagin-espazioa eta *A* eta *B* bi gertaera. ***A* gertaerarekiko baldintzazko *B* gertaeraren probabilitatea,** honela definitzen da:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}, \qquad P(A) > 0.$$

Era berean, ***B* gertaerarekiko baldintzazko *A* gertaeraren probabilitatea,** ondokoa da:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) > 0.$$

Bira A eta B bi gertaera. Aurreko definizioetatik honako hau ondorioztа daiteke:

$$P(A \cap B) = P(A)\, P(B \mid A), \qquad P(A) \neq 0,$$

$$P(A \cap B) = P(B)\, P(A \mid B), \qquad P(B) \neq 0.$$

Oro har, $A_1, A_2, ..., A_n$ gertaeren ebakiduraren probabilitatea honela definitzen da:

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)P(A_2 \mid A_1) \ldots P(A_n \mid A_1 \cap \ldots \cap A_{n-1})$$

baldin $P(A_1 \cap A_2 \cap ... \cap A_n) > 0$ bada.

A eta B **gertaera independenteak** dira, baldin eta soilik baldin,

$$P(A \cap B) = P(A)P(B).$$

Bira $A_1, A_2, ..., A_n$ gertaerak. $A_1, A_2, ..., A_n$ **gertaerak elkarrekiko independenteak** dira, baldin eta hurrengo berdintzak betetzen badira:

$$P(A_i \cap A_j) = P(A_i)\, P(A_j), \ \forall i \neq j$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)\, P(A_j)\, P(A_k), \ \forall i \neq j \neq k$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1)\, P(A_2)\, ...\, P(A_n).$$

2.4. BAYES-EN TEOREMA

2.4.1. Probabilitate osoaren teorema

Bira $\{A_1, A_2, \ldots, A_n\}$ gertaera-sistema osoa eta B beste gertaera bat. Orduan,

$$P(B) = P(B \mid A_1)\, P(A_1) + P(B \mid A_2)\, P(A_2) + ... + P(B \mid A_n)\, P(A_n)$$

non $P(A_i) > 0, \ i = 1, ..., n$ diren.

2.4.2. Bayes-en teorema

Bira $\{A_1, A_2, \ldots, A_n\}$ gertaera-sistema osoa eta B beste gertaera bat, zeintzuentzat $P(A_i) > 0$, $i = 1, ..., n$ eta $P(B) > 0$ diren. Orduan, sistema osoko edozein gertaeraren B gertaerarekiko baldintzazko probabilitatea, hurrengo eran kalkula daiteke:

$$P\left(A_r \middle| B\right) = \frac{P\left(B \middle| A_r\right) P\left(A_r\right)}{P\left(B \middle| A_1\right) P\left(A_1\right) + \ldots + P\left(B \middle| A_n\right) P\left(A_n\right)}, \qquad r = 1, 2, \ldots n.$$

$P(A_1), P(A_2), \ldots, P(A_n)$ zenbakiak “a priori” probabilitateak dira, eta $P(A_1 \mid B)$, $P(A_2 \mid B)$, …, $P(A_n \mid B)$ gaiei “a posteriori” probabilitateak deritze.

2.5. ARIKETA EBATZIAK

2.5.1. ariketa

Adieraz bitez lagin-espazioa eta gertaera, eta ondoren, aurki bedi gertaera bakoitzaren probabilitatea hurrengo kasu hauetan:

a. Dado orekatu baten jaurtiketaren emaitza 1 edo 3 izatea.

b. Txanpon baten bi jaurtiketatan, aurpegia ez ateratzea.

c. Bi dado jaurti ondoren, lortutako puntuazioen batura 6 baino txikiagoa izatea.

- -

Ebazpena

Erabiliko den notazioa hurrengoa da: Ω lagin-espazioa, A gertaera eta $P(A)$ gertaeraren probabilitatea.

a. $\Omega = \{1,2,3,4,5,6\}$, $A = \{1,3\}$, $P(A) = \frac{1}{3}$.

b. $\Omega = \{(a_1,a_2) \,/\, a_1 =$ "1. jaurtiketaren emaitza", $a_2 =$ "2. jaurtiketaren emaitza", a_1, $a_2 \in \{U =$ "aurpegia", $G =$ "gurutzea"$\}\} = \{(U,U), (U,G), (G,U), (G,G)\}$

$A = \{$aurpegirik, ez$\} = \{(G,G)\}$

$$P(A) = \frac{1}{4} = 0,25.$$

c. $\Omega = \{(\text{i, j}) \,/\, \text{i, j} \in \{1,2,3,4,5,6 \quad = \{(1,1), (1,2),\ldots, (1,6), (2,1), (2,2),\ldots, (2,6), \ldots, (6,1),\ldots(6,6)\}$

$A = \{(\text{i, j}) \in \Omega \,/\, \text{i} + \text{j} < 6\} = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (3,1), (3,2), (4,1)\}$

$$P(A) = \frac{10}{36} = 0,28.$$

___

2.5.2. ariketa

Kutxa batean hamar pila daude, eta hamarretatik bat akastuna da. Aleatorioki kutxatik hiru pila hartzen dira. Zein da akastuna ez hartzeko probabilitatea?

- -

Ebazpena

Bira A = "akastuna hartu" eta $\overline{A}$ = "akastuna ez hartu" gertaerak. Emaitza posible guztiak $\binom{10}{3} = 120$ eta A gertaeraren aldeko kasuak $\binom{9}{2}\binom{1}{1} = 36$ direla kontuan hartuz, A gertaeraren probabilitatea zehatz daiteke.

$$P(A) = \frac{36}{120} = 0{,}3.$$

Kontrako gertaeraren probabilitatearen adierazpena erabiliz, akastuna ez hartzeko probabilitatea kalkulatuko da:

$$P(\overline{A}) = 1 - P(A) = 1 - 0{,}3 = 0{,}7.$$

2.5.3. ariketa

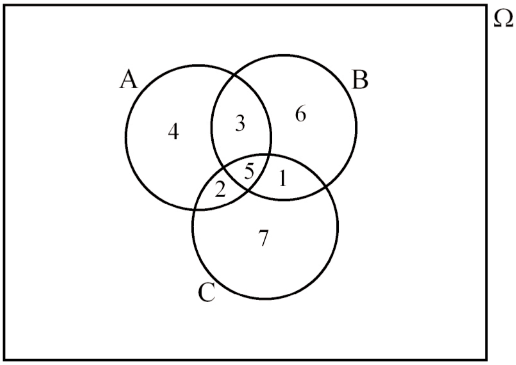
Lantegi batean A, B, C atalak daude. Lantegiko 28 langileetatik, A atalean 14, B atalean 15 eta C atalean 15 langile ditugu. Gainera, 8 langile A eta B ataletan, 6 langile B eta C ataletan eta 7 langile A eta C ataletan ari dira aldi berean. Lantegiko edozein langile era aleatorioan hartuz, zein da aldi berean hiru ataletan aritzeko probabilitatea? Eta A eta B ataletan bai, baina ez C atalean? Eta soilik A atalean? Eta B edo C ataletan bai, baina A atalean ez? Eta gutxienez hiru ataletakoren batean?

Ebazpena

Bira A = "A atalean lan egin", B = "B atalean lan egin" eta C = " C atalean lan egin" gertaerak.

Lantegian 28 langile daude, hots, $|\Omega| = 28$. Gainera, $|A| = 14$, $|B| = 15$, $|C| = 15$, $|A \cap B| = 8$, $|A \cap C| = 7$ eta $|B \cap C| = 6$.

Venn-en diagrama baten bidez atal bakoitzeko langile-kopurua zehatz daiteke:



2.1. irudia.

Ondoren hurrengo azpimultzoetako langile-kopurua kalkulatuko da:

"Hiru ataletan batera":

$$|A \cap B \cap C| = |\Omega| - |A| - |B| - |C| + |A \cap B| + |A \cap C| + |B \cap C| = 5.$$

"*A* eta *B* ataletan bai, baina ez *C* atalean": $\left|A \cap B \cap \overline{C}\right| = 3$.

"Soilik *A* atalean": $\left|A \cap \overline{B} \cap \overline{C}\right| = 4$.

"*B* edo *C* ataletan bai, baina *A* atalean ez": $\left|(B \cup C) \cap \overline{A}\right| = 14$.

"Gutxienez hiru ataletakoren batean": $|\Omega| = 28$.

Ondorioz, eskatutako gertaeren probabilitateak hauexek dira:

$P(\text{"hiru ataletan batera"}) = \dfrac{5}{28} = 0{,}18.$

$P(\text{"}A\text{ eta }B\text{ ataletan bai, baina ez }C\text{ atalean"}) = \dfrac{3}{28} = 0{,}11.$

$P(\text{"soilik }A\text{ atalean"}) = \dfrac{4}{28} = 0{,}14.$

$P(\text{"}B\text{ edo }C\text{ ataletan bai, baina }A\text{ atalean ez"}) = \dfrac{14}{28} = 0{,}5.$

$P(\text{"gutxienez hiru ataletakoren batean"}) = \dfrac{28}{28} = 1.$

2.5.4. ariketa

Elaborazio-planta batean aleatorioki makina bat hartu da. Makina ekoizpenerako erabiltzeko probabilitatea 0,31 da; diseinurako erabiltzeko probabilitatea 0,82; ekoizpenerako eta diseinurako erabiltzeko probabilitatea 0,18 eta beste eragiketetarako erabiltzeko probabilitatea 0,12 da. Onargarriak al dira probabilitate hauek guztiak batera?

Ebazpena

Makinaren erabilerak hurrengo gertaeren bidez adieraz daitezke: E = "ekoizpenerako", D = "diseinurako", $E \cap D$ = "ekoizpenerako eta diseinurako".

$P(E) = 0{,}31$, $P(D) = 0{,}82$, $P(E \cap D) = 0{,}18$ dira. Orduan, aleatorioki aukeratutako makina ekoizpenerako edo diseinurako erabiltzeko probabilitatea kalkula daiteke:

$$P(E \cup D) = P(E) + P(D) - P(E \cap D) =$$
$$= 0{,}31 + 0{,}82 - 0{,}18 = 0{,}95.$$

Biz $\overline{E \cup D} =$ "beste erabileretarako" gertaera. Enuntziatuaren arabera, $P(\overline{E \cup D}) = 0{,}12$ da. Horrela bada,

$$P(\Omega) = P(E \cup D) + P(\overline{E \cup D}) = 0{,}95 + 0{,}12 \neq 1$$

dugu, eta ondorioz, probabilitate hauek guztiak batera ez dira onargarriak.

2.5.5. ariketa

Esperimentu batean gerta daitezkeen hiru emaitza posible daude, *A, B* eta *C,* hirurak elkarrekiko bateraezinak direnak. Froga bedi hurrengo *P* eragilea probabilitatea denetz:

a. $P(A) = 0{,}34$, $P(B) = 0{,}43$, $P(C) = 0{,}21$

b. $P(A) = 0{,}35$, $P(B) = 0{,}40$, $P(C) = 0{,}25$

c. $P(A) = 0{,}27$, $P(B) = 0{,}75$, $P(C) = -0{,}02$

Ebazpena

Kasu bakoitzean, Kolmogorov-en hiru axiomak betetzen direnetz ikusi behar da.

a. *P* eragilea ez da probabilitatea,

$$P(A) + P(B) + P(C) = 0{,}34 + 0{,}43 + 0{,}21 \neq 1$$

baita.

b. *P* eragilea probabilitatea da, Kolgomorov-en hiru axiomak betetzen direlako:

i. $P(A), P(B), P(C) \in [0,1]$

ii. $P(A \cup B \cup C) = P(A) + P(B) + P(C) = 0{,}35 + 0{,}40 + 0{,}25 = 1$,

$$A \cap B = A \cap C = B \cap C = \emptyset$$

iii. $P(\Omega) = P(A \cup B \cup C) = 1$

c. *P* eragilea ez da probabilitatea, $P(C) = -0{,}02 < 0$ baita.

2.5.6. ariketa

Froga bitez hurrengo baieztapenak, betetzen badira eta ez badira betetzen, idatz itzazu hori erakusten duten adibideak:

a. Bi gertaera bateraezinak badira, orduan independenteak dira.

b. Bi gertaera independenteak badira, orduan bateraezinak dira.

Ebazpena

Aurreko bi ataletako baieztapenak, oro har, ez dira betetzen.

a. Demagun lantegi baterako 40 makina erosi direla. Makina horietatik 30 Txinatik datoz. Bira *A* = “makina, Txinakoa” eta *B* = “makina, ez Txinakoa” gertaerak. *A* eta *B* bateraezinak dira, $P(A \cap B) = 0$ baita. Bestalde, $P(A) = 30/40 = 0{,}75$ eta $P(\mathrm{B}) = 10/40 = 0{,}25$ dira.

A eta B gertaerak independenteak dira, baldin eta soilik baldin $P(A \cap B) = P(A)P(B)$ bada, eta kasu horretan $P(A \cap B) = 0$ eta $P(A)P(B) = 0{,}75 \cdot 0{,}25 \neq 0$. Beraz, A eta B gertaerak ez dira independenteak.

b. Demagun dado orekatu baten bi jaurtiketatan bi seiko ateratzea aztergai dela. Bira E_1 = "sei, lehenengo jaurtiketan", E_2 = "sei, bigarren jaurtiketan" eta $E_1 \cap E_2$ = "sei, bi jaurtiketetan" gertaerak. Ezagunak dira $P(E_1) = 1/6 = P(E_2)$ balioak. Gainera, lehenengo jaurtiketako emaitzak bigarren jaurtiketako emaitza ez duenez baldintzatuko, $P(E_1 \cap E_2) = P(E_1)\ P(E_2) = 1/36$ da. Beraz, bi gertaerak independenteak dira, baina ez dira bateraezinak, $P(E_1 \cap E_2) \neq 0$ baita.

2.5.7. ariketa

Bira A eta B gertaerak, $P(A) = 0{,}5$, $P(B) = 0{,}3$ eta $P(A \cap B) = 0{,}15$ izanik. Kalkula bitez $P(B \mid \overline{A})$ eta $P(A \mid \overline{B})$ probabilitateen balioak.

Ebazpena

a. Probabilitateen balioak kalkulatzeko baldintzazko probabilitatearen definizioa erabiliko da:

$$P(B \mid \overline{A}) = \frac{P\left(B \cap \overline{A}\right)}{P\left(\overline{A}\right)} \tag{1}$$

$B = (B \cap A) \cup (B \cap \overline{A})$ eta $P(B) = P(B \cap A) + P(B \cap \overline{A})$ adierazpenak kontuan hartuz, $P(B \cap \overline{A})$ probabilitatearen balioa kalkula daiteke:

$$\begin{aligned} P(B \cap \overline{A}) &= P(B) - P(B \cap A) \\ &= 0{,}30 - 0{,}15 =,015. \end{aligned} \tag{2}$$

Bestalde,

$$P(\overline{A}) = 1 - P(A) = 0{,}5. \tag{3}$$

Azkenik, (2) eta (3) balioak (1) adierazpenean ordezkatuz, $P(B \mid \overline{A})$ balioa lortuko da:

$$P(B \mid \overline{A}) = \frac{0{,}15}{0{,}50} = 0{,}3.$$

Era berean, aurreko urratsak elkartuz, $P(A \mid \overline{B})$ balioa kalkulatuko da:

$$P\left(A\middle|\overline{B}\right)=\frac{P\left(A\cap\overline{B}\right)}{P\left(\overline{B}\right)}$$
$$=\frac{P(A)-P(A\cap B)}{1-P(B)}$$
$$=\frac{0,5-0,15}{1-0,3}=0,5.$$

2.5.8. ariketa

Hogeita hamar ingeniari teknikoren artean, sei mekanikakoak dira. Aleatorioki, ingeniari teknikoetatik bi aukeratzen dira.

Kalkula ezazu bi ingeniarietakoren bat mekanikakoa izateko probabilitatea. Zein da bigarrena mekanikakoa izateko probabilitatea?

Ebazpena

P(mekanikakoa) = p = 6/30 = 0,2, P(ez-mekanikakoa) = q = 0,8.

Bira M_1 = “aukeraturiko lehenengo ingeniari teknikoa, mekanikakoa” eta M_2 = “aukeraturiko bigarren ingeniari teknikoa, mekanikakoa” gertaerak.

Bi ingeniarietakoren bat mekanikakoa izateko probabilitatea kalkulatzeko, bilduraren probabilitatea aplika daiteke:

$$P(M_1 \cup M_2) = P(M_1) + P(M_2) - P(M_1 \cap M_2) \tag{1}$$

Alde batetik, (1) adierazpenean ebakiduraren probabilitatea kalkulatu behar da.

$$P(M_1 \cap M_2) = P(M_1)\, P(M_2 \mid M_1) = \frac{6}{30}\cdot\frac{5}{29}. \tag{2}$$

Bestalde aukeratutako bi ingeniarietatik bigarrena mekanikakoa izateko $P(M_2)$ probabilitatearen balioa zehazteko dago.

$$P(M_2)=P(M_1)P\left(M_2\middle|M_1\right)+P\left(\overline{M_1}\right)P\left(M_2\middle|\overline{M}_1\right)=\frac{6}{30}\cdot\frac{5}{29}+\frac{24}{30}\cdot\frac{6}{29}. \tag{3}$$

Bukatzeko, (2) eta (3) balioak (1) adierazpenean ordezkatuko dira:

$$P(M_1\cup M_2)=\frac{6}{30}+\frac{6}{30}\cdot\frac{5}{29}+\frac{24}{30}\cdot\frac{6}{29}-\frac{6}{30}\cdot\frac{5}{29}=0,366.$$

2.5.9. ariketa

Unibertsitate bateko talde eragilean 80 irakasle daude, eta horietatik 16 Medikuntza Fakultatekoak dira. Talde eragileko irakasleen artean bi kargu berri ezartzeko asmoa dago, karguak aleatorioki banatuko direlarik. Kalkula bedi bi kargu berriak Medikuntza Fakultateko irakasleei egokitzeko probabilitatea:

a. Irakasle bakoitzak kargu berri bat onar dezakeelarik.

b. Irakasle batek bi kargu berriak onar ditzakeelarik.

Ebazpena

Bira $M_i =$ " *i.* kargua, Medikuntza Fakultateko irakasleari" gertaerak, $i = 1, 2$ izanik.

a. Lehenengo kasuan bi gertaeraren ebakiduraren probabilitatearen adierazpena planteatuko da, kontuan hartuta lehenengo kargua irakasle batentzat izango dela eta bigarren kargua beste irakasle batentzat izango dela

$$P\left(M_1 \cap M_2\right) = P\left(M_1\right)P\left(M_2 | M_1\right) = \frac{16}{80} \cdot \frac{15}{79} = 0,038.$$

b. Bigarren kasuan, bi karguak irakasle berdinarentzat izan daitezke. Beraz, ebakiduraren probabilitatean gertaeren arteko independentzia dago:

$$P\left(M_1 \cap M_2\right) = P\left(M_1\right)P\left(M_2 | M_1\right) = P\left(M_1\right)P\left(M_2\right) = \frac{16}{80} \cdot \frac{16}{80} = 0,04.$$

2.5.10. ariketa

Dado orekatu bat bost aldiz jaurtitzen da. Kalkula bedi:

a. Zehazki lehenengo hiru jaurtiketetan hiru seiko ateratzeko probabilitatea.

b. Hiru seiko ateratzeko probabilitatea.

Ebazpena

Bira S_i = "sei, *i.* jaurtiketan" eta $\overline{S_i}$ = "ez-sei, *i.* jaurtiketan" gertaerak, non $i = 1, 2, \ldots, 5$ den.

$P(S_i) = \dfrac{1}{6}$ eta $P(\overline{S_i}) = \dfrac{5}{6}$ dira, $i = 1,2, \ldots, 5$ izanik. Kontuan hartu behar da jaurtiketetako emaitzak elkarrekiko independenteak direla.

a. $P\left(S_1 \cap S_2 \cap S_3 \cap \overline{S}_4 \cap \overline{S}_5\right) = P\left(S_1\right)P\left(S_2\right)P\left(S_3\right)P\left(\overline{S}_4\right)P\left(\overline{S}_5\right) =$

$$= \frac{1}{6}\frac{1}{6}\frac{1}{6}\frac{5}{6}\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 0,0032.$$

b. Bigarren ataleko galderan hiru seikoak edozein posiziotan atera daitezke. Hau da, emaitza hurrengo gertaeren probabilitateen batura izango da:

$$P\left(S_1 \cap S_2 \cap S_3 \cap \overline{S}_4 \cap \overline{S}_5\right) = P\left(S_1\right)P\left(S_2\right)P\left(S_3\right)P\left(\overline{S}_4\right)P\left(\overline{S}_5\right) =$$

$$= \frac{1}{6}\frac{1}{6}\frac{1}{6}\frac{5}{6}\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 0,0032$$

$$P\left(\overline{S}_1 \cap S_2 \cap S_3 \cap S_4 \cap \overline{S}_5\right) = P\left(\overline{S}_1\right)P\left(S_2\right)P\left(S_3\right)P\left(S_4\right)P\left(\overline{S}_5\right) =$$

$$= \frac{5}{6}\frac{1}{6}\frac{1}{6}\frac{1}{6}\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 0,0032$$

....

$$P\left(\overline{S}_1 \cap \overline{S}_2 \cap S_3 \cap S_4 \cap S_5\right) = P\left(\overline{S}_1\right)P\left(\overline{S}_2\right)P\left(S_3\right)P\left(S_4\right)P\left(S_5\right) =$$

$$= \frac{5}{6}\frac{5}{6}\frac{1}{6}\frac{1}{6}\frac{1}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 0,0032.$$

Honelako $\binom{5}{3}$ batugai daude. Ondorioz, hurrengo emaitza lortuko da:

P(hiru seiko, bost jaurtiketetan) $= \binom{5}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = 10 \cdot 0,0032 =$ 0,032.

2.5.11. ariketa

Exekutibo batek hiru agentziatatik alokatzen ditu kotxeak: % 20 Alfa agentziatik; % 20 Beta agentziatik eta % 60 Omega agentziatik. Gainera, Alfa agentziako kotxeetatik % 10ek, Beta agentziako kotxeetatik % 12k eta Omega agentziako kotxeetatik % 4k ez dute airbag sistemarik. Zenbatekoa da exekutiboak airbag sistemadun kotxe bat alokatzeko probabilitatea?

Ebazpena

Biz S = “airbag sistema ez eduki” gertaera. Kasu honetan, kotxeak hiru agentzietatik lor daitezke. Bira A = “kotxea, Alfa agentziakoa”, B = “kotxea, Beta agentziakoa” eta C = “kotxea, Omega agentziakoa” gertaerak. $\{A, B, C\}$ gertaera-sistema osoa da, hots, $A \cup B \cup C = \Omega$ eta binaka bateraezinak dira. Orduan, probabilitate osoaren legea aplikatuz, exekutiboak alokatzen duen kotxe batek airbag sistema ez edukitzeko probabilitatea lortuko da.

$$P(S) = P(S \mid A)\,P(A) + P(S \mid B)\,P(B) + P(S \mid C)\,P(C)$$

$$= 0{,}20 \cdot 0{,}10 + 0{,}20 \cdot 0{,}12 + 0{,}60 \cdot 0{,}04 = 0{,}068.$$

Ondorioz, exekutiboak airbag sistemadun kotxe bat alokatzeko probabilitatea hurrengoa da:

$$P\left(\overline{S}\right) = 1 - P(S) = 1 - 0{,}068 = 0{,}932.$$

2.5.12. ariketa

Ordenagailu-denda batean bi teknikari daude. C teknikariak konpoketen % 40 eta D teknikariak konponketen % 60 egiten dituzte. C teknikariak egindako 10 konponketatik bat berrikusi egin behar da, eta D teknikariak egindako 20 konponketatik bat berrikusteko dago. Aleatorioki, berrikuspena behar izan duen konponketa bat hartu da. Zein da konponketa hori C teknikariak egina izateko probabilitatea?

Ebazpena

Bira C = "C teknikariak, konpondu", D = "D teknikariak, konpondu", B = "konponketa berrikusi". $\{C, D\}$ gertaera-sistema osoa da, $P(C) = 0{,}4$ eta $P(D) = 0{,}6$ probabilitateen balioak ezagunak izanik. Halaber, $P(B \mid C) = 0{,}10$ eta $P(B \mid D) = 0{,}05$ dira. Orduan, Bayes-en teorema aplika daiteke:

$$P(C|B) = \frac{P(B|C)P(C)}{P(B|C)P(C) + P(B|D)P(D)} = \frac{0{,}10 \cdot 0{,}40}{0{,}10 \cdot 0{,}40 + 0{,}05 \cdot 0{,}60} = 0{,}57.$$

2.5.13. ariketa

Enpresa batek hiru lantegi ditu eta lantegi bakoitzean osagai elektronikoen ekoizpenerako merkatuan dauden I, II eta III teknikak erabiltzen dira, hurrenez hurren. Enpresa honek ekoitzitako edozein osagai elektronikok kalitate-kontrola gainditzeko 0,74ko probabilitatea du. Horrez gain, jakina da osagai elektroniko bat I teknikaz egina izateko probabilitatea 0,40koa dela eta II teknikaz egina izateko probabilitatea 0,30ekoa dela.

Azken urteko emaitzak aztertuz, I teknikaz egindako osagai elektronikoek kontrol-kalitateko azterketa gainditzeko probabilitatea 0,80koa eta II teknikaz 0,60koa dela ikusi da.

a. Zein da III teknikaz egindako osagai elektronikoek kontrol-kalitateko azterketa gainditzeko probabilitatea?

b. Urtearen bukaeran kontrol-kalitatea gainditu duten osagai elektronikoen artean, zein proportzio izango da II teknikaz egin ez dena?

Ebazpena

Datuen arabera honako gertaerak eta probabilitateak ezagunak dira: G = "kontrol-kalitatea gainditu", A = "I teknika erabili", B = "II teknika erabili" eta C = "III teknika erabili". Bestalde, $\{A, B, C\}$ gertaera-sistema osoa da. Halaber, hurrengo probabilitateen balioak ezagunak dira: $P(A) = 0{,}40$, $P(B) = 0{,}30$, $P(C) = 0{,}30$, $P(G) = 0{,}74$, $P(G \mid A) = 0{,}80$, $P(G \mid B) = 0{,}60$.

a. III teknikaz egindako osagai elektronikoek kontrol-kalitateko azterketa gainditzeko probabilitatea kalkulatzeko, probabilitate osoaren teorema aplika daiteke:

$$P(G) = P(G \mid A)\, P(A) + P(G \mid B)\, P(B) + P(G \mid C)\, P(C)$$

$$0{,}74 = 0{,}8 \cdot 0{,}4 + 0{,}6 \cdot 0{,}3 + P(G \mid C) \cdot 0{,}3$$

$$P(G \mid C) = \frac{0{,}74 - 0{,}32 - 0{,}18}{0{,}3} = 0{,}8.$$

b. Kontrol-kalitatea gainditu duen osagai elektroniko bat aleatorioki aukeratuz, II teknikaz egina izateko probabilitatea Bayes-en teorema aplikatuz kalkula daiteke:

$$P\left(B \middle| G\right) = \frac{P\left(G \middle| B\right) P(B)}{P\left(G \middle| A\right) P(A) + P\left(G \middle| B\right) P(B) + P\left(G \middle| C\right) P(C)} =$$

$$= \frac{0{,}6 \cdot 0{,}3}{0{,}74} = 0{,}243.$$

Kontrako gertaeraren probabilitatea erabiliz, hurrengoa lortzen da:

$$P\left(\overline{B} \middle| G\right) = 1 - P\left(B \middle| G\right) = 1 - 0{,}243 = 0{,}757.$$

Ondorioz, urtearen bukaeran kontrol-kalitatea gainditu duten osagai elektronikoen artean % 75,7 ez da II teknikaz egina.

2.5.14. ariketa

Hiri jakin bateko inprimategietatik % 30ek euskaraz, % 52k gaztelaniaz eta % 40k ingelesez inprimatzen dute. Halaber, inprimategietako % 10ek euskaraz eta gaztelaniaz, % 8k euskaraz eta ingelesez, % 12k gaztelaniaz eta ingelesez eta % 5ek hiru hizkuntzetan inprimatzen dute. Aleatorioki, hiri horretako inprimategi bat hartuz, zein da euskara, ingelesa edo gaztelaniaz aparteko beste hizkuntza batean inprimitzeko probabilitatea?

- -

Ebazpena

Bira E = "euskaraz inprimatu", G = "gaztelaniaz inprimatu", I = "ingelesez inprimatu" gertaerak.

Ezagunak dira $P(E) = 0{,}30$, $P(G) = 0{,}52$, $P(I) = 0{,}40$, $P(E \cap G) = 0{,}10$, $P(E \cap I) = 0{,}08$, $P(G \cap I) = 0{,}12$ eta $P(E \cap I \cap G) = 0{,}05$ balioak.

Aleatorioki, hiri horretako inprimategi bat hartuz, hiru hizkuntzakoren batean inprimatzeko probabilitatea honela kalkula daiteke:

$$P(G \cup E \cup I) = P(E) + P(G) + P(I) - P(E \cap G) - P(E \cap I) - P(G \cap I) + P(E \cap I \cap G)$$
$$= 0{,}30 + 0{,}52 + 0{,}40 - 0{,}10 - 0{,}08 - 0{,}12 + 0{,}05 = 0{,}97.$$

Ondorioz, kontrako gertaeraren probabilitatea erabiliz, beste hizkuntza batean inprimitzeko probabilitatea hurrengoa da:

$$P\left(\overline{G \cup E \cup I}\right) = 1 - P\left(G \cup E \cup I\right) = 1 - 0{,}97 = 0{,}03.$$

2.5.15. ariketa

Hurrengo taulan, zenbait lantegitako langileek prestakuntza-ikastaroa jaso ala ez, produkzio-kalitatearen sarien banaketa azaltzen da:

	Z lantegiko langileak		Beste lantegietako langileak	
	Prestakuntza bai	Prestakuntza ez	Prestakuntza bai	Prestakuntza ez
Saria bai	1.000	250	75	4.000
Saria ez	7.000	750	25	4.000

a. Prestakuntza jaso ez duten langileen artean, zein da produkzio-kalitatearen saria lortzeko probabilitatea?

b. Prestakuntza jaso ez duten eta *Z* lantegikoak ez diren langileen artean, zein da produkzio-kalitatearen saria lortzeko probabilitatea?

c. Oro har, zenbatekoa da produkzio-kalitatearen saria lortzeko probabilitatea?

Ebazpena

Bira *S* = “saria lortu”, *F* = “prestakuntza jaso” eta *Z* = “*Z* lantegikoa izan” gertaerak.

a. Aleatorioki prestakuntzarik jaso ez duen langile bat aukeratuz, produkzio-kalitatearen saria lortzeko probabilitatea hauxe da:

$$P\left(S|\overline{F}\right) = \frac{P\left(S \cap \overline{F}\right)}{P\left(\overline{F}\right)} = \frac{250 + 4.000}{250 + 750 + 4.000 + 4.000} = \frac{4.250}{9.000} = 0{,}47.$$

b. Aleatorioki prestakuntzarik jaso ez duen Z lantegitik kanpoko langile bat aukeratuz, produkzio-kalitatearen saria lortzeko probabilitatea hurrengo eran kalkula daiteke.

$$P\left(S\middle|\overline{F}\cap\overline{Z}\right)=\frac{4.000}{8.000}=0,5.$$

c. Aleatorioki langile bat aukeratuz, produkzio-kalitatearen saria lortzeko probabilitatea honako hau da:

$$P(S)=\frac{5.325}{17.100}=0,31.$$

2.5.16. ariketa

Bira esperimentu bati dagozkion A eta B gertaerak. Azal bedi $P(A)$ eta $P(A \mid B)$ probabilitateen arteko diferentzia.

Ebazpena

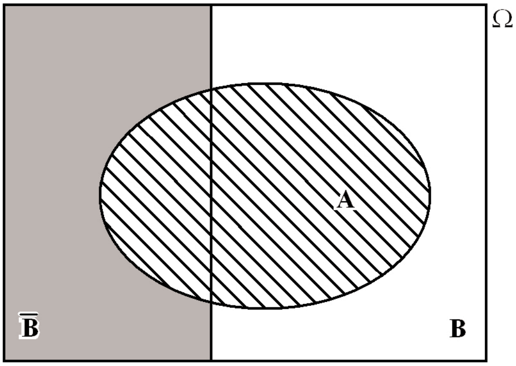
Esperimentuari dagokion emaitza posible guztien multzoa Ω lagin-espazioa da. Beraz, A gertaeraren probabilitatea kalkulatzean lagin-espazioa Ω da.

$$P(A)=\frac{|A|}{|\Omega|}.$$

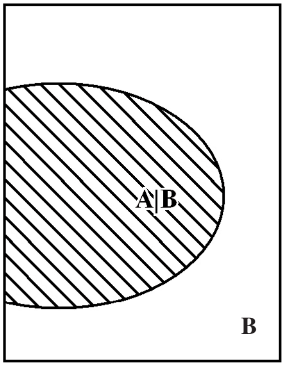
$P(A \mid B)$ probabilitatearen kalkuluan, aldiz, lagin-espazioa B azpimultzoa da.

$$P\left(A\middle|B\right)=\frac{P(A\cap B)}{P(B)}=\frac{|A\cap B|}{|B|}.$$

Grafikoki honela adieraz daiteke egoera:



2.2. irudia.



2.3. irudia.

2.5.17. ariketa

2003. urteko selektibitateko lehenengo deialdira Hego Euskal Herrian aurkeztutako ikasleak hauek izan ziren: % 13 Araban, % 28 Gipuzkoan, % 42 Bizkaian eta % 17 Nafarroan. Arabako ikasleetatik % 35ek, Gipuzkoakoetatik % 64k, Bizkaikoetatik % 42k eta Nafarroakoetatik % 19k selektibitatea euskaraz burutu zuten.

a. Hego Euskal Herriko ikasleen artean, zein da selektibitatea gaztelaniaz burutu zutenen proportzioa?

b. Selektibitatea euskaraz egin zutenen artean, zein da Arabako ikasleen proportzioa?

Ebazpena

Bira A = "selektibitatea, Araban", G = "selektibitatea, Gipuzkoan", B = "selektibitatea, Bizkaian", N = "selektibitatea, Nafarroan" eta E = "selektibitatea, euskaraz", gertaerak. $\{A, G, B, N\}$ gertaera-sistema osoa da. Gainera, $P(A) = 0{,}13$, $P(G) = 0{,}28$, $P(B) = 0{,}42$, $P(N) = 0{,}17$, $P(E|A) = 0{,}35$, $P(E|G) = 0{,}64$, $P(E|B) = 0{,}42$ eta $P(E|N) = 0{,}19$ probabilitateen balioak ezagunak dira.

a. Hego Euskal Herriko ikasleen artean, selektibitatea euskaraz burutu zutenen proportzioa kalkulatzeko probabilitate osoaren legea aplika daiteke:

$$P(E) = P(E|A)P(A) + P(E|G)P(G) + P(E|B)P(B) + P(E|N)P(N) =$$

$$= 0{,}35 \cdot 0{,}13 + 0{,}64 \cdot 0{,}28 + 0{,}42 \cdot 0{,}42 + 0{,}19 \cdot 0{,}17 = 0{,}433.$$

Ondorioz, kontrako gertaeraren probabilitatearen adierazpena erabiliz, Hego Euskal Herriko ikasleen artean, selektibitatea gaztelaniaz egiteko probabilitatea kalkulatuko da:

$$P\left(\overline{E}\right) = 1 - P(E) = 1 - 0{,}433 = 0{,}567.$$

Hau da, Hego Euskal Herriko ikasleen artean % 56,7k burutu zuen selektibitatea gaztelaniaz.

b. Bayes-en teorema aplikatuz, eskatutako probabilitatea kalkulatuko da.

$$P(A|E) = \frac{P(E|A)P(A)}{P(E|A)P(A) + P(E|G)P(G) + P(E|B)P(B) + P(E|N)P(N)} =$$

$$= \frac{0,35 \cdot 0,13}{0,35 \cdot 0,13 + 0,64 \cdot 0,28 + 0,42 \cdot 0,42 + 0,19 \cdot 0,17} = 0,105.$$

Beraz, selektibitatea euskaraz egin zutenen artean, Arabako ikasleen proportzioa % 10,5 da.

2.5.18. ariketa

Enpresa batean eskaeraren gorakada badago, produktuen prezioek gora egiteko probabilitatea 0,80koa da. Eskaeraren gorakadarik ez badago, produktuen prezioek gora ez egiteko probabilitatea 0,65ekoa da. Enpresako eskaeraren gorakada % 70ekoa da.

a. Zein izango da enpresako produktuen prezioek gora egiteko probabilitatea?

b. Produktuen prezioek gora egin badute, zein da enpresaren eskaerak gorakadarik ez izateko probabilitatea?

Ebazpena

Bira honako gertaerak: E = "eskaeraren gorakada", G = "prezioek, gora". Enuntziatuko datuen arabera: $P(G \mid E) = 0{,}80$, $P(\overline{G}|\overline{E}) = 0{,}65$, P($E$) = 0,70. Gainera, E eta $\overline{E}$ gertaerek gertaera-sistema osoa osatzen dute, $P(E) = 0{,}70$ eta $P(\overline{E}) = 0{,}30$ izanik.

a. Probabilitate osoaren teorema aplikatuz, enpresako produktuen prezioek gora egiteko probabilitatea kalkula daiteke:

$$P(G) = P(G|E)P(E) + P(G|\overline{E})P(\overline{E}) =$$

$$= 0,80 \cdot 0,70 + (1 - 0,65)(1 - 0,70) = 0,665.$$

b. $P(\overline{E}|G)$ probabilitatearen balioa kalkulatu behar da. Horretarako, Bayes-en teorema erabil daiteke:

$$P(\overline{E}|G) = \frac{P(G|\overline{E})P(\overline{E})}{P(G|\overline{E})P(\overline{E}) + P(G|E)P(E)} =$$

$$= \frac{0,35 \cdot 0,30}{0,35 \quad 0,30 + 0,80 \quad 0,70} = 0,158.$$

2.5.19. ariketa

Alkoholemia-egoeretan, alkoholimetroak alkoholemia-kasuetako % 99 detektatzen ditu, eta alkoholemia-kasurik ez dagoenean, alkoholimetroak kasuetako % 3an detektatzen du alkoholemia. Demagun kasuetako % 80an alkoholemiako egoerak daudela.

a. Zein da alkoholimetroak era egokian funtzionatzeko probabilitatea?

b. Zein da alkoholemia detektatu den kasu batean, benetan alkoholemia izateko probabilitatea?

Ebazpena

Bira A = "alkoholemia izan", D = "alkoholemia detektatu" gertaerak. Enuntziatuko datuen arabera, $P(A)=0{,}80$, $P(D|A)=0{,}99$ eta $P(D|\overline{A})=0{,}03$ dira. $\{A,\overline{A}\}$ gertaera-sistema osoa da.

Biz E = "alkoholimetroak era egokian funtzionatu" gertaera.

a. Alkoholimetroak era egokian funtzionatzeko probabilitatea hurrengo eran kalkula daiteke:

$$P(E)=P(A)P(D|A)+P(\overline{A})P(\overline{D}|\overline{A})=$$
$$=0{,}80\cdot 0{,}99+(1-0{,}80)(1-0{,}03)=0{,}986.$$

b. Bayes-en teorema aplikatzeko baldintzak betetzen direnez, hots, $\{A,\overline{A}\}$ gertaera-sistema osoa eta $P(A)>0$ eta $P(\overline{A})>0$ direnez, eskatutako probabilitatearen balioa hauxe da:

$$P(A|D)=\frac{P(D|A)P(A)}{P(D|A)P(A)+P(D|\overline{A})P(\overline{A})}=$$
$$=\frac{0{,}99\cdot 0{,}80}{0{,}99\cdot 0{,}80+0{,}03\cdot 0{,}20}=0{,}992.$$

2.5.20. ariketa

Eskola Politekniko batean iaz ikasketak bukatu zituzten ikasleetatik % 27k Industri Ingeniaritza, % 35ek Informatika eta % 38k Telekomunikazio Ingeniaritza ikasketak bukatu zituzten, ikasle bakoitzak ikasketa bakarra burutu zuelarik. Industri Ingeniaritza ikasi zuten ikasleetatik % 90, Informatika ikasi zutenetatik % 80 eta Telekomunikazio Ingeniaritza ikasi zutenetatik % 95 lanean daude.

Eskola politeknikoan iaz ikasketak bukatu zituen ikasleetako bat aleatorioki aukeratuz gero, zein da Industri Ingeniaria edo lanean izateko probabilitatea?

Ebazpena

Iaz ikasketak bukatu zituen Eskola Politeknikoko ikasleetako bat aleatorioki aukeratuko da. Bira I = “Industri Ingeniaritza”, F = “Informatika”, T = “Telekomunikazio Ingeniaritza” eta L = “lanean” gertaerak.

Enuntziatuko datuetatik, hurrengo balioak ezagunak dira: $P(I) = 0{,}27$, $P(F) = 0{,}35$, $P(T) = 0{,}38$, $P(L|I) = 0{,}90$, $P(L|F) = 0{,}80$, $P(L|T) = 0{,}95$

Kalkulatu behar den probabilitatea $P(I \cup L)$ da eta horretarako bi gertaeren bilduraren probabilitatearen adierazpena kontuan hartuko da.

$$P(I \cup L) = P(I) + P(L) - P(I \cap L). \tag{1}$$

Bestalde, $\{I, F, T\}$ gertaera-sistema osoa denez, probabilitate osoaren teorema aplika daiteke:

$$\begin{aligned} P(L) &= P(I)\,P(L|I) + P(F)\,P(L|F) + P(T)\,P(L|T) = \\ &= 0{,}27 \cdot 0{,}90 + 0{,}35 \cdot 0{,}80 + 0{,}38 \cdot 0{,}95 = 0{,}884. \end{aligned} \tag{2}$$

Halaber, I eta L gertaeren ebakiduraren probabilitatea hurrengoa da:

$$\begin{aligned} P(I \cap L) &= P(I)\,P(L|I) = \\ &= 0{,}27 \cdot 0{,}90 = 0{,}243. \end{aligned} \tag{3}$$

Bukatzeko, lortutako (2) eta (3) balioak (1) adierazpenean ordezkatuko dira:

$$\begin{aligned} P(I \cup L) &= P(I) + P(L) - P(I \cap L) = \\ &= 0{,}27 + 0{,}884 - 0{,}243 = 0{,}911. \end{aligned}$$

2.6. ARIKETA PROPOSATUAK

2.6.1. ariketa

Bira A, B eta C gertaerak. $P(A) = 0{,}4$, $P(B) = 0{,}41$, $P(C) = 0{,}53$, $P(A \cap B) = 0{,}11 = P(A \cap C)$, $P(B \cap C) = 0{,}23$, $P(A \cap B \cap C) = 0{,}01$.

a. Bateraezinak al dira A, B eta C gertaerak?

b. Kalkula bitez $P(A|B)$, $P(B|A)$, $P(A|C)$, $P(\overline{B}|A)$, $P(A \cap B|C)$, $P(A \cup B|C)$ probabilitateak.

2.6.2. ariketa

Esan zergatik diren ezinezkoak hurrengo enuntziatuak:

a. Ikasle batek Estatistika irakasgaian A kalifikazioa lortzeko probabilitatea 0,32 da, eta A edo B kalifikazioa lortzeko probabilitatea 0,27 da.

b. Bi merkataritza-gune eraikitzen ari dira. Lehenengoa garaiz bukatua izateko probabilitatea 0,35 da, eta biak garaiz bukatuak izateko probabilitatea 0,42 da.

2.6.3. ariketa

Enpresa bateko langileek fabrikazio, salmenta edo marketineko sailetan lan egiten dute. Enpresan 44 langile daude: 20 fabrikazioan, 18 salmentan, 16 marketinean, 5 fabrikazio eta salmentan, 4 fabrikazio eta marketinean, 3 salmenta eta marketinean. Enpresako langile bat era aleatorioan hartuz, zein da fabrikazio, marketin eta salmentan, hiruretan lan egiteko probabilitatea? Eta soilik salmentan? Eta soilik fabrikazio eta marketinean? Eta fabrikazioan edo salmentan? Eta salmenta ez den beste sail batean?

2.6.4. ariketa

Eskola bateko gela batean hogei ume daude eta lau etorkinak dira. Aleatorioki gelako hiru ume aukeratu dira. Kalkula bitez hurrengo balioak:

a. Ume etorkinik ez izateko probabilitatea.

b. Gehienez ume bat etorkina izateko probabilitatea.

2.6.5. ariketa

Sistema mekaniko batek era egokian funtzionatzen du, bere osagai guztiek era egokian funtzionatzen badute. Sistema mekanikoak 65 osagai ditu eta osagaien % 99k era egokian funtzionatzen du. Demagun osagai mekanikoen funtzionamendua elkarrekiko independentea dela. Zein da sistema mekanikoak era egokian ez funtzionatzeko probabilitatea?

2.6.6. ariketa

Denda batean, ehun musika-kateren artean 81 goi-fidelitatekoak dira. Aleatorioki dendako lau musika-kate aukeratzen dira. Kalkula ezazu lau musika-kateak goi-fidelitatekoak izateko probabilitatea hurrengo kasuetan:

a. Aukeratzen den musika-kate bakoitza behin bakarrik aukera daiteke (itzulerarik gabe).

b. Aukeratzen den musika-katea, hartu ondoren berriro ere aukeratua izan daiteke (itzuleraz).

2.6.7. ariketa

Herri konkretu batean unibertsitatera joan diren emakumeak eta gizonak zenbatu dira, eta ondoko taula eratu:

	Unibertsitatera joan	Unibertsitatera ez joan
Emakume	1.215	400
Gizon	1.500	355

Datu horietan oinarrituz, herriko emakume bat era aleatorioan hartuz gero, zenbatekoa da unibertsitatera joateko probabilitatea? Eta unibertsitatera joan den pertsona bat aleatorioki aukeratuz gero, zein da emakumea izateko probabilitatea? Zer ondoriozta daiteke bi emaitza horietatik?

2.6.8. ariketa

Ordenagailuetarako biruskontrako programen artean erabiltzaileen % 57k McAfee, % 38k Panda eta % 5ek beste biruskontrako programa batzuk erabiltzen dituzte. Demagun erabiltzaileek ordenagailu bakoitzean biruskontrako programa bakarra instalatzen dutela. McAfee programa % 99an da eraginkorra, Panda % 95ean eta besteak % 93an. Biruskontrako programa eraginkorren artean, zein da McAfee erabiltzen dutenen proportzioa?

2.6.9. ariketa

EJK enpresak hiru lantegi ditu eta lantegi bakoitzean osagai elektronikoen produkziorako merkatuan dauden A, B eta C teknikak erabiltzen dira, hurrenez hurren. Enpresa honek produzitutako edozein osagai elektronikok kalitate-kontrola gainditzeko 0,74ko probabilitatea du. Jakina da gainera osagai elektroniko bat A teknikaz egina izateko probabilitatea 0,40koa dela eta B teknikaz egina izateko probabilitatea 0,30ekoa dela.

Azken urteko emaitzak aztertuz, A teknikaz egindako osagai elektronikoek kontrol-kalitateko azterketa gainditzeko probabilitatea 0,80koa eta B teknikaz 0,60koa dela ikusi da.

a. Zein da C teknikaz egindako osagai elektronikoek kontrol-kalitateko azterketa gainditzeko probabilitatea?

b. Urtearen bukaeran kontrol-kalitatea gainditu duen osagai elektroniko bat aukeratzen bada, zein proportziotan izango da B teknikaz egin ez dena?

2.6.10. ariketa

Lantegi batean ekoizten dituzten produktuei buruz jaso diren kexetatik % 35 arazo mekanikoengatik eta % 65 arazo elektrikoengatik izan dira. Arazo mekanikoengatik jasotako kexen artean % 16 bermeko lehen urtean izan dira, eta arazo elektrikoengatik jasotako kexen artean % 24 bermeko lehen urtean izan dira. Bermeko lehen urtean kexa izan duen produktu bat aleatorioki hartuz, zein da kexa arazo mekanikoengatik izateko probabilitatea?

2.6.11. ariketa

Eguraldiaren iragarleak egun eguzkitsuen % 99 eguzkitsu izango zela iragarri zuen. Aldiz, eguzkitsuak izan ez ziren egunetatik % 3 eguzkitsu izango zela iragarri zuen. Demagun biharko eguna ez-eguzkitsua izateko probabilitatea 0,80 dela.

a. Zein da iragarleak biharko egingo duen iragarpena egokia izateko probabilitatea?

b. Egun eguzkitsuen iragarpenen artean, zenbatekoa da egun eguzkitsuen proportzioa?

3. Aldagai aleatorio diskretua

3.1. ALDAGAI ALEATORIOA

Esperimentu bati dagozkion emaitzak (balioak) eta emaitza bakoitza lortzeko probabilitateak adierazten dituen aldagaiari **aldagai aleatorio** deritzo. Esperimentuaren emaitzak aldagaiak har ditzakeen balioak dira, eta balio horiek aleatorioki har daitezkeenez, aldagaia aleatorioa dela esaten da.

Oro har, aldagai aleatorioa, esperimentu aleatorio baten emaitza posible guztien multzoan definituriko funtzio erreala da. Aldagaiak hartzen dituen balioak diskretuak direnean, **aldagai aleatorio diskretua** dugu. Aldagaiak tarte bateko edozein balio har badezake, aldiz, **aldagai aleatorio jarraitua** da.

3.2. ALDAGAI ALEATORIO DISKRETUA

3.2.1. Probabilitate-funtzioa eta banaketa-funtzioa

Biz X aldagai aleatorio diskretua, zeinak $x_1, x_2, ..., x_n$ balioak hartzen dituen.

$P(X = x_1) = f(x_1) = p_1$, $P(X = x_2) = f(x_2) = p_2$, ..., $P(X = x_n) = f(x_n) = p_n$ balioek X aldagai aleatorioaren **probabilitate-funtzioa** definituko dute, baldin eta ondoko propietateak betetzen badira:

i. $f(x_i) \geq 0, \quad i = 1, ..., n$

ii. $\sum_{i=1}^{n} f(x_i) = 1.$

Bira X aldagai aleatorio diskretuaren $x_1, x_2, ..., x_n$ balioak eta $f(x)$ probabilitate-funtzioa. X aldagai aleatorio diskretuaren $F(x)$ **banaketa-funtzioa** honela definitzen da:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

Beraz, probabilitate-funtzioa ezagutuz banaketa-funtzioa zehatz daiteke. Alderantziz, banaketa-funtzioa ezaguna bada, probabilitate-funtzioa lortzeko hurrengo propietatea aplika daiteke:

$$f(x_i) = F(x_i) - F(x_{i-1}), \qquad i = 1, 2, \ldots, n.$$

3.2.2. Aldagai aleatorio diskretuaren batez bestekoa edo itxaropen matematikoa

Biz X aldagai aleatorio diskretua, zeinak $x_1, x_2, ..., x_n$ balioak $p_1, p_2, ..., p_n$ probabilitatez hartzen dituen, hurrenez hurren.

X aldagai aleatorio diskretuaren **batez bestekoa** edo **itxaropen matematikoa** honela definitzen da:

$$\mu = E(x) = \sum_{i=1}^{n} x_i p_i = \sum_{i=1}^{n} x_i f(x_i).$$

Aldagai aleatorio diskretuaren itxaropen matematikoaren propietate garrantzitsuenak hauek dira:

1. $E(k) = k$, k konstantea izanik.

2. Bira $X_1, X_2, ..., X_n$ aldagai aleatorio diskretuak.

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n).$$

3. $E(kX) = kE(X)$, k konstantea izanik.

4. Bira $X_1, X_2, ..., X_n$ aldagai aleatorio diskretuak eta k_i konstanteak, $i = 1, 2, ..., n$.

$$E(k_1X_1 + k_2X_2 + ... + k_nX_n) = k_1E(X_1) + k_2E(X_2) + ... + k_nE(X_n).$$

5. Bira $X_1, X_2, ..., X_n$ aldagai aleatorio diskretu independenteak.

$$E(X_1 \cdot X_2 \cdot ... \cdot X_n) = E(X_1) \cdot E(X_2) \cdot ... \cdot E(X_n).$$

3.2.3. Aldagai aleatorio diskretuaren bariantza eta desbideratze tipikoa

Biz X aldagai aleatorio diskretua, zeinak $x_1, x_2, ..., x_n$ balioak $p_1, p_2, ..., p_n$ probabilitatez hartzen dituen, hurrenez hurren.

X aldagai aleatorio diskretuaren **bariantza** honela definitzen da:

$$\sigma_x^2 = Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p_i.$$

Nabari daitekeenez, $\sigma_x^2 = Var(X) = E[(X - \mu)^2]$ *da.*

Askotan, kalkuluak errazteko asmoz, beste adierazpen hau erabiltzen da:

$$\sigma_x^2 = Var(X) = \sum_{i=1}^{n} x_i^2 p_i - \mu^2.$$

Bariantzaren erro karratu positiboari **desbideratze tipikoa** deritzo.

$$\sigma_x = \sqrt{Var(x)}.$$

X aldagai aleatorio diskretuaren bariantzaren propietateen artean hurrengoak aipa daitezke:

1. $Var(X) \geq 0$, edozein X aldagai aleatorio diskretutarako.

2. $Var(k) = 0$, k konstantea izanik.

3. $Var(X) = E[(X - \mu)^2] = E(X^2) - \mu^2$.

4. Bira $X_1, X_2, ..., X_n$ aldagai aleatorio diskretu independenteak.

$$Var(X_1 + X_2 + ... + X_n) = Var(X_1) + Var(X_2) + ... + Var(X_n).$$

5. $Var(kX) = k^2 Var(X)$, k konstantea izanik.

6. $Var(X + k) = Var(X)$, k konstantea izanik.

3.2.4. Tchebychev-en teorema

Biz X aldagai aleatorioa, itxaropen eta bariantza finitukoa. Orduan, ondokoa betetzen da:

$$P(|X - E(X)| > k\sigma) \leq \frac{1}{k^2}, \qquad \forall k > 0.$$

Hemendik honakoa ondorioztatuko da:

$$P(|X - E(X)| \leq k\sigma) \geq 1 - \frac{1}{k^2}, \qquad \forall k > 0.$$

Teorema honen bidez X aldagai aleatorioaren balioak $[\mu - k\sigma, \mu + k\sigma]$ tartean izateko probabilitatearen behe-bornea kalkula daiteke.

3.3. BANAKETA DISKRETU GARRANTZITSUENETARIKO BATZUK

Atal honen helburua, Ingeniaritzako aplikazioetan maiztasun handienez gertatzen diren fenomenoei dagozkien banaketak azaltzea da. Arrazoi horregatik, beste banaketei garrantzirik kendu gabe, banaketa diskretuen artean binomiala, hipergeometrikoa eta Poisson-en banaketa azpimarratuko dira.

3.3.1. Banaketa binomiala

Banaketa binomiala duen aldagai aleatorioaren zenbait adierazgarri:

i. Esperimentu batean n proba daude eta proba bakoitzak bi emaitza posible ditu: A = "arrakasta" ala $\overline{A}$ = "porrota".

ii. Arrakastaren probabilitatearen balioa proba guztietan berdina da, $P(A) = p$ eta $P(\overline{A}) = 1 - p = q$ izanik.

iii. n probak elkarrekiko independenteak dira.

Biz X = "arrakasta-kopurua, n probatan" aldagai aleatorio diskretua. n probatan x arrakasta izateko probabilitatea hurrengoa da:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \qquad x = 0, 1, \ldots, n.$$

Banaketa honi, n eta p parametroetako **banaketa binomiala** deritzo. Laburbilduz, X: $Bin(n,p)$ dela adieraz daiteke.

X aldagai aleatorioari dagokion banaketa-funtzioa honako hau da:

$$F(x) = P(X \le x) = \sum_{k \le x} \binom{n}{k} p^k (1 - p)^{n-k}.$$

n eta p parametroetako banaketa binomiala duen X aldagai aleatorioaren batez bestekoa eta bariantza hurrengoak dira:

$$\mu = E(x) = np, \qquad \sigma_x^2 = Var(X) = npq.$$

3.3.2. Banaketa hipergeometrikoa

Banaketa binomialaren modukoa da, baina itzulerarik gabe. Demagun N elementuen artean r arrakasta daudela. N elementuetatik n elementu aldi berean edo elkarren segidan itzulerarik gabe aukeratzen dira. Bira X = "arrakasta-kopurua, n elementuen artean" aldagai aleatorio diskretua, $p = r/N = P$(arrakasta) eta $q = (N-r)/N = P$(porrota) probabilitateak. n elementuetatik x arrakasta izateko probabilitatea hurrengo eran kalkulatuko da:

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \qquad 0 \le x \le min(r,n).$$

Aldagai honek N, n eta p parametroetako **banaketa hipergeometrikoa** du. Laburbilduz, X: $H(N,n,p)$.

X aldagai aleatorioari dagokion banaketa-funtzioa hurrengoa da:

$$F(x) = P(X \le x) = \sum_{k \le x} \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}.$$

N, n eta p parametroetako banaketa hipergeometrikodun X aldagai aleatorioaren batez bestekoa eta bariantza honela kalkulatzen dira:

$$\mu = E(X) = np, \qquad \sigma_x^2 = Var(X) = npq\frac{(N-n)}{(N-1)}.$$

3.3.3. Poisson-en banaketa

Poisson-en banaketari jarraitzen dioten aldagai aleatorioek denbora-tarte, azalera edo beste neurri-unitate batean behatutako gertaera independenteen kopurua aztertzen dute. Gertaera bat neurri-unitate jakin batean jasotzeko probabilitatea konstante mantentzen da unitate guztietarako. Horrela, egunean zehar orduro neurtutako ozono-mailak, enpresa baten hilabeteko fakturazioak edota asteko lan-istripuen kopurua Poisson-en banaketen adibide modura har daitezke.

X aldagai aleatorioari dagokion probabilitate-funtzioa

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

bada, orduan X aldagaiak **λ parametroko Poisson-en banaketa** du. X: Poisson (λ) adieraziko dugu, $\lambda > 0$ delarik.

λ parametroko Poisson-en banaketa duen X aldagai aleatorioari dagokion banaketa-funtzioa honela definitzen da:

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} \frac{e^{-\lambda}\lambda^k}{k!}, \qquad x = 0, 1, 2, \ldots$$

Gainera, X: Poisson(λ) banaketaren batez besteko eta bariantzaren balioak λ parametroa dira:

$$\mu = E(X) = \lambda \qquad \sigma_x^2 = Var(X) = \lambda.$$

3.4. BANAKETEN ARTEKO KONBERGENTZIA

3.4.1. Banaketa binomialaren eta Poisson-en banaketaren arteko erlazioa

Demagun X aldagai aleatorio diskretuak n, p parametroetako banaketa binomiala duela. n "handia" eta p "txikia" direnean, orduan banaketa binomiala np parametroko Poisson-en banaketaren bidez hurbil daiteke. Hau da, $n \to \infty$ eta $p \to 0$ konbergitzen dutenean, hurrengo hurbilketa erabil daiteke:

$$Bin(n,p) \cong \text{Poisson}\,(np).$$

Praktikan, $p \leq 0{,}1$ eta $np < 5$ direnean, ontzat hartzen da hurbilketa.

3.4.2. Banaketa hipergeometrikoaren eta banaketa binomialaren arteko erlazioa

Demagun X aldagai aleatorio diskretuak N, n eta p parametroetako banaketa hipergeometrikoa duela. N zenbakia n delakoarekiko "handia" denean, orduan banaketa hipergeometrikoa n eta $p = r/N$ parametroetako banaketa binomialaren bidez hurbil daiteke. Hau da, $N \to \infty$ doanean, hurbilketa hau erabil daiteke:

$$H(N,n,p) \cong Bin(n,p).$$

Praktikan, hurbilketa egokia da $\frac{n}{N} < 0{,}1$ denean.

3.5. ARIKETA EBATZIAK

3.5.1. ariketa

Biz X aldagai aleatorio diskretua, $x = 0, 1, 2$ balioak hartzen dituena.

Hurrengo funtzioak, probabilitate-funtzioak al dira? Arrazoitu zergatia.

a. $f(0) = 0{,}36,\ f(1) = 0{,}36,\ f(2) = 0{,}36.$

b. $f(0) = 0{,}1,\ f(1) = 0{,}6,\ f(2) = 0{,}3.$

c. $f(0) = 0{,}3,\ f(1) = 0{,}8,\ f(2) = -0{,}1.$

Ebazpena

Hiru kasuetan, $f(x)$ funtzioak probabilitate-funtzioa izateko baldintzak betetzen diren ikusi behar da, baldintza horiek hurrengo biak direlarik:

i. $f(x_i) \geq 0, \quad i = 1, ..., n.$

ii. $\sum_{i=1}^{n} f(x_i) = 1.$

a. Kasu honetan $f(x)$ ez da probabilitate-funtzioa,

$$f(0) + f(1) + f(2) = 0{,}36 + 0{,}36 + 0{,}36 \neq 1$$

delako.

b. $f(x)$ funtzioa probabilitate-funtzioa da,

$$f(x) \geq 0, \quad x = 0{,}1{,}2$$

$$f(0) + f(1) + f(2) = 1$$

betetzen direlako.

c. Oraingoan $f(x)$ ez da probabilitate-funtzioa, $f(2) < 0$ delako.

3.5.2. ariketa

X aldagai aleatorio diskretuaren probabilitate-funtzioa honakoa da:

$$f(x) = \begin{cases} k, & x = 1{,}3 \\ 2k, & x = 2 \\ 0, & \text{beste} \end{cases}$$

a. Aurki bedi k konstante positiboaren balioa.

b. Kalkula bitez $P(X \leq 2)$ eta $P(1 < X \leq 3)$.

c. Lor bitez X aldagai aleatorio diskretuaren banaketa-funtzioa, batez bestekoa eta desbideratze tipikoa.

Ebazpena

a. $f(x)$ probabilitate-funtzioa bada,

i. $f(x_i) \geq 0, \quad i = 1, ..., n$

ii. $\sum_{i=1}^{n} f(x_i) = 1$

axiomak bete behar ditu.

Kasu honetan, lehenengo axiomatik $k > 0$ ondoriozta daiteke, eta bigarrenetik k konstantearen balioa lortuko da.

$$f(1) + f(2) + f(3) = k + 2k + k = 1 \rightarrow$$
$$k = 1/4$$

b. $P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$.

$P(1 < X \leq 3) = P(X = 2) + P(X = 3) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$.

c. X aldagai aleatorio diskretuari dagokion banaketa-funtzioa hurrengoa da:

$x < 1$ bada: $P(X \leq x) = 0$.

$1 \leq x < 2$ bada: $P(X \leq x) = P(X = 1) = 1/4$.

$2 \leq x < 3$ bada: $P(X \leq x) = P(X = 1) + P(X = 2) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$.

$x \geq 3$ bada: $P(X \leq x) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$.

Ondorioz, $F(x)$ banaketa-funtzioaren adierazpen analitikoa hauxe da:

$$F(x) = \begin{cases} 0, & x < 1 \\ 1/4, & 1 \leq x < 2 \\ 3/4, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

Bestalde, X aldagai aleatorio diskretuaren batez bestekoa kalkulatuko dugu.

$$\mu = E(x) = \sum_{i=1}^{n} x_i f_i = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} = 2.$$

Era berean, bariantzaren balioa kalkula daiteke.

$$\sigma_x^2 = Var(X) = \sum_{i=1}^{n} x_i^2 f(x_i) - \mu^2 =$$

$$= \left(1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{2} + 3^2 \cdot \frac{1}{4}\right) - 2^2 = 0,5.$$

Erro karratua aplikatuz, desbideratze tipikoaren balioa ondorioztatzen da.

$$\sigma_x = \sqrt{Var(x)} = \sqrt{0,5} = 0,707.$$

3.5.3. ariketa

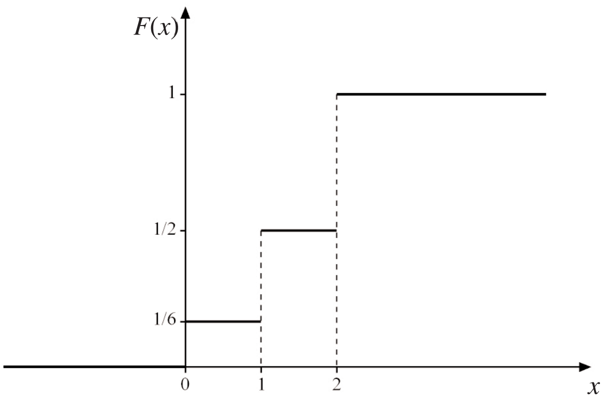
Biz X aldagai aleatorio diskretuaren hurrengo banaketa-funtzioa:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/6, & 0 \le x < 1 \\ 1/2, & 1 \le x < 2 \\ 1, & x \ge 2 \end{cases}$$

Kalkula bitez $f(x)$ probabilitate-funtzioa, $E(X)$ eta $Var(X)$.

Ebazpena

$F(x)$ banaketa-funtzioaren adierazpen grafikoan hiru etenune dituela nabari daiteke:



3.1. irudia.

$x = 0, 1, 2$ etenuneak dira probabilitate-funtzioak hartzen dituen balioak. Horrela,

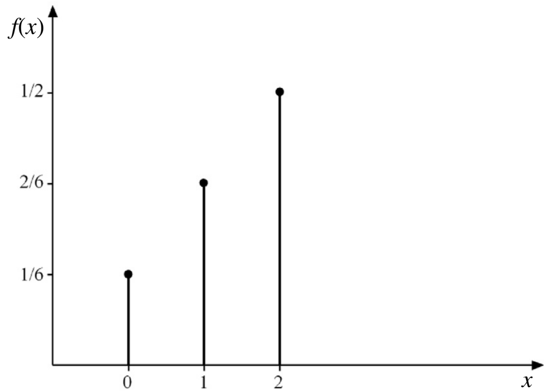
$$f(x_i) = F(x_i) - F(x_{i-1}), \qquad i = 1, 2, \ldots, n$$

propietatea aplikatuz, X aldagai aleatorio diskretuari dagokion banaketa-funtzioa lortuko da:

$$f(0) = F(0) = 1/6$$
$$f(1) = F(1) - F(0) = 1/2 - 1/6 = 1/3$$
$$f(2) = F(2) - F(1) = 1 - 1/2 = 1/2$$

Gainerako balioetan $f(x)$ funtzioak balio nulua hartzen du. Beraz, $f(x)$ probabilitate-funtzioaren adierazpen analitikoa hurrengoa da:

$$f(x) = \begin{cases} 1/6, & x = 0 \\ 1/3, & x = 1 \\ 1/2, & x = 2 \\ 0, & \text{beste} \end{cases}$$



3.2. irudia.

Bestalde, X aldagai aleatorio diskretuaren batez bestekoa kalkulatuko dugu.

$$\mu = E(x) = \sum_{i=1}^{n} x_i f_i = 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{2} = \frac{4}{3}.$$

Era berean, bariantzaren balioa kalkula daiteke, eta azken horren erro karratua kalkulatuz, desbideratze tipikoa lortuko da.

$$\sigma_x^2 = Var(x) = \sum_{i=1}^{n} x_i^2 f(x_i) - \mu^2 =$$
$$= \left(0^2 \cdot \frac{1}{6} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{2}\right) - \left(\frac{4}{3}\right)^2 = \frac{5}{9}.$$

$$\sigma_x = \sqrt{Var(x)} = \sqrt{5/9} = 0{,}745.$$

3.5.4. ariketa

Bira X eta Y bi aldagai aleatorio diskretu independente. X aldagai aleatorio diskretuaren itxaropen matematikoa –5 eta bariantza 3 dira. Y aldagai aleatorio diskretuaren itxaropen matematikoa 1 eta bariantza 4 dira. Kalkula bitez:

a. $E(3X + 5Y + 2)$

b. $\sigma(3X + 5Y + 2)$

Ebazpena

a. $E(3X + 5Y + 2) = E(3X) + E(5Y) + E(2) = 3 \cdot E(X) + 5 \cdot E(Y) + 2 =$

$$= 3 \cdot (-5) + 5 \cdot 1 + 2 = -8$$

b. $Var\,(3X + 5Y + 2) = Var(3X) + Var(5Y) + Var(2) =$

$$= 3^2 \cdot Var(X) + 5^2 \cdot Var(Y) + 0 = 3^2 \cdot 3 + 5^2 \cdot 4 = 127.$$

Ondorioz, desbideratze tipikoaren balioa honako hau da:

$$\sigma(3X + 5Y + 2) = \sqrt{127} = 11{,}27.$$

3.5.5. ariketa

Sistema elektroniko bat balioztatzean –2, 0, 2 puntuazioak eman dira, dagokion probabilitate-legea $f(x) = k\,(|x| + 2)^2$ izanik.

a. k konstantearen balioa zehaztuz, kalkula bedi sistema elektronikorako itxaron daitekeen puntuazioa eta bariantza.

b. Zer esango zenuke hurrengo baieztapenari buruz: “0,75 balioa, $P(|X| \le 3{,}772)$ probabilitatearen behe-borne bat da”? Arrazoi ezazu erantzuna.

Ebazpena

a. Biz X = “sistema elektronikoari emandako puntuazioak” aldagai aleatorio diskretua.

$f(x) = k\,(|x| + 2)^2$ probabilitate-legea izateko bete behar diren axiomak:

i. $f(x) \ge 0, \quad x = -2, 0, 2 \quad \rightarrow \quad k > 0.$

ii. $\displaystyle\sum_x f(x) = \sum_x P(X = x) = f(-2) + f(0) + f(2) = 1.$

$$f(-2) + f(0) + f(2) = k\,(2 + 2)^2 + k\,(0 + 2)^2 + k\,(2 + 2)^2 = 1 \quad \rightarrow \quad k = \frac{1}{36}.$$

Sistema elektronikorako itxaron daitekeen puntuazioa:

$$E(X) = \mu = \sum_x xf(x) = -2f(-2) + 0f(0) + 2f(2) = -2\cdot\frac{4^2}{36} + 2\cdot\frac{4^2}{36} = 0.$$

Sistema elektronikoaren bariantza:

$$Var(X) = \sum_x x^2 f(x) - \mu^2 = 2^2 f(-2) + 0^2 f(0) + 2^2 f(2) = 32/9.$$

b. Tchebychev-en teorema aplikatuz hurrengoa ondoriozta daiteke:

$$P(|X - E(X)| \leq c\sigma) \geq 1 - \frac{1}{c^2}.$$

Kasu honetan desbideratze tipikoaren balioa σ = 1,886 da. Orain Chebychev-en teoreman c = 2 ordezkatuz:

$$P(|X - 0| \leq 3,772) \geq 1 - \frac{1}{2^2} = 0,75.$$

Beraz, 0,75 balioa $P(|X - 0| \leq 3{,}772)$ probabilitatearen behe-borne bat da.

3.5.6. ariketa

Lehiaketa batera aurkezten diren arkitektoek saria lortzeko probabilitatea 0,70 da. Lehiaketa horretara sei arkitekto aurkeztu dira. Kalkula ezazu:

a. Sei arkitektoak sarituak izateko probabilitatea.

b. Gutxienez bi eta gehienez lau arkitekto sarituak izateko probabilitatea.

c. Zenbat arkitekto itxaron daiteke sarituak izatea?

Ebazpena

Biz X="arkitekto sarituen kopurua lehiaketan" aldagai aleatorio diskretua. Lehiaketara aurkezten den edozein arkitekto saritua izateko probabilitatea p = 0,70 konstante mantentzen denez eta n = 6 arkitekto aurkeztu direnez, X aldagai aleatorio diskretuak banaketa binomiala du, non p = P(saritua) = 0,70 eta q = P(ez-saritua) = 0,30 diren. Orduan,

$$P(X = x) = \binom{6}{x}(0,70)^x(0,30)^{6-x}, \qquad x = 0,1,\ldots,6.$$

a. $P(6 \text{ arkitekto saritu}) = P(X = 6) = \binom{6}{6}(0,70)^6(0,30)^0 = (0,70)^6 = 0,12.$

b. P(gutxienez bi eta gehienez lau arkitekto saritu) = $P(2 \leq X \leq 4)$ =

$$= \binom{6}{2}(0,70)^2(0,30)^4 + \binom{6}{3}(0,70)^3(0,30)^3 + \binom{6}{4}(0,70)^4(0,30)^2 =$$

$$= \frac{6!}{2!4!} \cdot 0,49 \cdot 0,0081 + \frac{6!}{3!3!} \cdot 0,343 \cdot 0,027 + \frac{6!}{2!4!} \cdot 0,2401 \cdot 0,09 =$$

= 0,059535 + 0,18522 + 0,324135 = 0,56889.

c. $E(X) = np = 6 \cdot 0,7 = 4,2$ denez, 4,2 arkitekto sarituak izatea itxaron daiteke.

3.5.7. ariketa

AXX kableak ekoizten dituen lantegi batean, edozein kable akastuna izateko probabilitatea 0,05 da. Produktuen kalitatea aztertzeko asmoz, egunero aleatorioki 15 kable hartzen dira, hurrengo araua aplikatzen delarik: “15 kabletik bat baino gehiago akastuna bada, produkzioa gelditzea erabakiko da”.

a. Azal bedi aldagai aleatorioari dagokion banaketa.

b. Kalkula bedi edozein egunetan produkzioa gelditzeko probabilitatea.

c. Produkzioaren kalitatea aztertzeko egunero 120 kable hartu eta arau berbera aplikatuko balitz, zenbatekoa izango litzateke produkzioa gelditzeko probabilitatea?

Ebazpena

a. Edozein *AXX* kable akastuna izateko probabilitatea $p = 0,05$ da. $n = 15$ kable hartzen dira. Demagun 15 kableen kalitatea elkarren artean independentea dela eta edozein kable akastuna izateko probabilitatea 0,05 dela. Orduan X = “*AXX* kable akastunen kopurua” aldagai aleatorio diskretuak $n = 15$ eta $p = 0,05$ parametroetako banaketa binomiala du, $X : Bin(15, 0,05)$ alegia.

Banaketa honi dagokion probabilitate-legea ondokoa da:

$$P(X = x) = \binom{15}{x}(0,05)^x(0,95)^{15-x}, \qquad x = 0, \ldots, 15.$$

b. P(produkzioa gelditu) = $P(X > 1) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$ =

$$= 1 - \binom{15}{0}(0,05)^0(0,95)^{15} - \binom{15}{1}(0,05)^1(0,95)^{14} = 0,17.$$

c. $n = 120$ denean, banaketa binomialaren hurbilketa moduan banaketa normala erabil daiteke. Kasu honetarako, $X \approx N(np, \sqrt{npq})$, $np = 120 \cdot 0,05 = 6$ eta

$\sqrt{npq} = \sqrt{120 \cdot 0,05 \cdot 0,95} = \sqrt{5,7} = 2,39$ izanik. 0,5eko jarraitutasunagatiko zuzenketa-faktorea aplikatuz, eskatutako probabilitatea kalkulatuko da (ikus 102 or.).

$$P(\text{produkzioa gelditu}) = P(X > 1) = P(X > 1,5) = P\left(\frac{X-6}{2,39} > \frac{1,5-6}{2,39}\right) =$$
$$= P(Z > -1,88) = P(Z \leq 1,88) = \phi(1,88) = 0,9699.$$

3.5.8. ariketa

Produktu jakin bati buruzko inkestan "bai" ala "ez" erantzuneko 21 galdera daude. Erantzunen herena gehienez ezezkoa bada, emaitza positibotzat hartuko da, eta kontrako kasuan emaitza negatiboa dela esango da. Demagun "bai" ala "ez" probabilitate beraz erantzuteko pertsona bati inkesta egiten zaiola eta galdera guztiei erantzuten diela.

a. Zein da emaitza negatiboa izateko probabilitatea?

b. Eta 21 galdera ordez 48 galdera badaude, zein da emaitza negatiboa izateko probabilitatea?

Ebazpena

Biz X = "ezezko emaitzen kopurua" aldagai aleatorio diskretua. P("ez" erantzun) = = 1/2 = p eta P("bai" erantzun) = 1/2 = q probabilitateek edozein galderatarako balio bera hartzen dute. Kasu honetan, $X: Bin(n = 21,\ p = 1/2)$ da. Beraz, dagokion banaketa-funtzioa hurrengoa da:

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} \binom{21}{k} \left(\frac{1}{2}\right)^{k} \left(\frac{1}{2}\right)^{21-k} = \sum_{k=0}^{x} \binom{21}{k} \left(\frac{1}{2}\right)^{21}.$$

a. $P(\text{emaitza positiboa}) = P\left(X \leq \frac{21}{3}\right) = \sum_{x=0}^{7} P(X = x) = \sum_{x=0}^{7} \binom{21}{x} \left(\frac{1}{2}\right)^{21} = 0,095.$

P(emaitza negatiboa) = 1 − P(emaitza positiboa) = 1 − 0,095 = 0,905

b. Hasteko, P(emaitza positiboa) = $P(X \leq 16)$ probabilitatea kalkulatuko da. Kasu honetan, $np = 48 \cdot \frac{1}{2} = 24 > 5$ eta $nq = 24 > 5$ direnez, $np = 24$ eta $\sqrt{npq} = 3,46$ parametroetako banaketa normalaren bidezko hurbilketa erabiliko dugu, hots, $X \approx N(24,\ 3,46)$. Halaber, 0,5eko jarraitutasunagatiko zuzenketa aplikatuko da:

$$P(X \leq 16) = P(X < 16,5) = P\left(\frac{X-24}{3,46} < \frac{16,5-24}{3,46}\right) = P(Z < -2,17) =$$
$$= 1 - P(Z < 2,17) =$$
$$= 1 - \phi(2,17) = 1 - 0,9850 = 0,0150.$$

Ondorioz, kontrako gertaeraren probabilitatearen adierazpena erabiliko da:

$$P(\text{emaitza negatiboa}) = 1-P(\text{emaitza positiboa}) = 0{,}985.$$

3.5.9. ariketa

Administrazio-enpresa batek egunero, batez beste, bi kexa jasotzen ditu.

a. Kalkula bedi administrazio-enpresaren kexarik gabeko egunen portzentajea.

b. Zein da egun batean gutxienez kexa bat eta gehienez lau kexa jasotzeko probabilitatea?

c. Zein da bi egunetan gehienez bost kexa jasotzeko probabilitatea?

d. Zenbatekoa da hiru egunetan sei kexa baino gehiago jasotzeko probabilitatea?

Ebazpena

Biz X = “administrazio-enpresak egunero jasotzen duen kexa-kopurua”. X aldagai aleatorioak $\lambda = 2$ parametroko Poisson-en banaketa duela suposa daiteke, aldagaiak denbora-tarte batean jasotzen den arrakasta-kopurua adierazten baitu. Orduan, aldagaiari dagokion probabilitate-legea hurrengoa da:

$$P(X = x) = \frac{e^{-2}2^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

a. Horren arabera, egun batean kexarik ez jasotzeko probabilitatea kalkulatuko da:

$$P(X = 0) = \frac{e^{-2}2^0}{0!} = e^{-2} = 0{,}135.$$

Beraz, egunetako % 13,5ean administrazio-enpresak ez du kexarik jasotzen.

b. $P(1 \le X \le 4)$ probabilitatea kalkulatzeko, probabilitate-legearen adierazpena erabil daiteke:

$$P(1 \le X \le 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) =$$

$$= e^{-2}\left(\frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}\right) = 0{,}812.$$

Beste era batean, banaketa-funtzioaren bidez ere kalkula daiteke aurreko probabilitatea:

$$P(1 \le X \le 4) = F_X(4) - F_X(0) = 0{,}947 - 0{,}135 = 0{,}812.$$

c. Biz Y = “administrazio-enpresak bi egunetan jasotzen duen kexa-kopurua”. Y aldagai aleatorioak $\lambda_y = \lambda \cdot 2 = 4$ parametroko Poisson-en banaketa duela suposa daiteke. Horren arabera, bi egunetan gehienez bost kexa jasotzeko probabilitatea kalkulatuko da:

$$P(Y \le 5) = F_y(5) = e^{-4}\left(1 + \frac{4^1}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} + \frac{4^5}{5!}\right) = 0{,}785.$$

d. Biz Z = “administrazio-enpresak hiru egunetan jasotzen duen kexa-kopurua”. Z aldagai aleatorioak $\lambda_z = \lambda \cdot 3 = 6$ parametroko Poisson-en banaketa badu, orduan:

$$P(Z > 6) = 1 - P(Z \le 6) = 1 - F_z(6) =$$
$$= 1 - e^{-6}\left(1 + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!} + \frac{6^5}{5!} + \frac{6^6}{6!}\right) = 1 - 0{,}606 = 0{,}394.$$

3.5.10. ariketa

Eskola batean urtero, batez beste, 1.050 ikaslek egiten dute sartzeko eskaera.

a. Zenbatekoa da datorren urtean gehienez 950 ikaslek sartzeko eskaera egiteko probabilitatea?

b. Demagun edozein ikasle onartua izateko probabilitatea 1/6 dela. 100 ikasletik zenbat itxaron daiteke onartuak izatea?

Ebazpena

Biz X = “urtero sartzeko eskaera egiten duen ikasle-kopurua” aldagai aleatorioa. X aldagaiak $\lambda = 1.050$ parametroko Poisson-en banaketa du. Parametroaren balioa handia denez, 102 orriko banaketa normalaren bidezko $X \approx N(\lambda, \sqrt{\lambda})$ hurbilketa erabil daiteke.

a. Horren arabera, 0,5eko jarraitutasunagatiko zuzenketa-faktorea aplikatuz, kalkula dezagun eskatutako probabilitatearen balioa:

$$P(X \le 950) = P(X < 950{,}5) = P\left(\frac{X - 1.050}{\sqrt{1.050}} < \frac{950{,}5 - 1.050}{\sqrt{1.050}}\right) =$$
$$= P(Z < -3{,}07) = 1 - P(Z < 3{,}07) =$$
$$= 1 - 0{,}9989 = 0{,}0011.$$

b. Biz Y = “100 ikasletik onartua izan daitekeen ikasle-kopurua” aldagai aleatorioa. Demagun edozein ikasle onartua izateko probabilitatea 1/6 dela eta ehun ikasleren artean bat onartua izateak besteen onarpena ez duela baldintzatuko. Orduan Y aldagai aleatorioak $n = 100$ eta $p = 1/6$ parametroetako banaketa binomiala du. $Y : Bin(100, 1/6)$. Kasu horretan,

$$E(Y) = np = 100 \cdot \frac{1}{6} = 16,67 \text{ ikasle}$$

onartuak izatea itxaron daiteke.

3.5.11 ariketa

Aholkularitza-etxe batek urtero, batez beste, 1.200 pertsonari ematen die aholkua.

a. Zenbatekoa da hurrengo urtean 1.085 pertsona baino gehiagori aholkua emateko probabilitatea?

b. Kalkula bedi 1.085-1.200 pertsonari aholkua emateko probabilitatea.

- -

Ebazpena

Demagun X = "urtero aholkua jasotzen duen pertsona-kopurua" aldagai aleatorio diskretuak $\lambda = 1.200$ parametroko Poisson-en banaketa duela.

a. $P(X > 1.085)$ kalkulatzeko banaketa normalaren bidezko hurbilketa erabiliko da, $X \approx N(1.200,\ \sqrt{1.200})$ hain zuzen ere.

0,5eko jarraitutasunagatiko zuzenketa-faktorea aplikatuz:

$$P(X > 1.085) = P(X > 1.085,5) = P\left(Z > \frac{1.085,5 - 1.200}{\sqrt{1.200}}\right) =$$

$$= P(Z > -3,31) = P(Z < 3,31\ ) = 0,9995.$$

b. Aurreko atalean egin den moduan eragingo da.

$$P(1.085 \leq X \leq 1.200) = P(1.084,5 < X < 1.200,5) =$$

$$= P\left(\frac{1.084,5 - 1.200}{\sqrt{1.200}} < Z < \frac{1.200,5 - 1.200}{\sqrt{1.200}}\right) =$$

$$= P(-3,33 < Z < 0,01) = \phi(0,01) - \phi(-3,33) =$$

$$= \phi(0,01) - 1 + \phi(3,33) = 0,5040 - 1 + 0,9996 = 0,5036.$$

3.5.12. ariketa

Elektronikaren arloko enpresa ezagun batek teknikariak kontratatzeko azterketa egin du. Azterketara aurkeztu ziren hautagaietatik, % 15ek gai zela adierazi zuen. Demagun hautagai guztiek formakuntza bera dutela.

a. Azterketara aurkeztutako 5 pertsona aleatorioki aukeratuz gero, zenbatekoa da gutxienez 4 gai izateko probabilitatea?

b. Eta azterketara aurkeztutako 50 pertsonetatik, zenbatekoa da gutxienez 16 gai izateko probabilitatea?

Ebazpena

a. Biz X = "azterketara aurkeztutako 5 pertsonatik, gai den pertsona-kopurua" aldagai aleatorio diskretua. X: $Bin(5, 0{,}15)$, non $p = P(\text{gai}) = 0{,}15$, $q = P(\text{ez-gai}) = 0{,}85$ eta $n = 5$ diren.

Banaketa honi dagokion probabilitate-legea ondokoa da:

$$P(X = x) = \binom{5}{x}(0{,}15)^x(0{,}85)^{5-x}, \qquad x = 0, \ldots, 5.$$

$$P(X \geq 4) = P(X = 4) + P(X = 5) =$$

$$= \binom{5}{4}(0{,}15)^4(0{,}85)^1 + \binom{5}{5}(0{,}15)^5(0{,}85)^0 = 0{,}0022.$$

b. Demagun $n = 50$. Orduan, $N(np = 7{,}5,\ \sqrt{npq} = 2{,}52)$ banaketa normalaren bidezko hurbilketa erabiliko da:

$$P(X \geq 16) = P(X > 15{,}5) = P\left(\frac{X - 7{,}5}{2{,}52} > \frac{15{,}5 - 7{,}5}{2{,}52}\right) =$$

$$= P(Z \geq 3{,}17) = 1 - 0{,}9992 = 0{,}0008.$$

3.5.13. ariketa

Gas naturaleko berogailuen instalazioen % 60 era egokian egiten da.

a. Kalkula bedi 6 instalaziotik 3 egokiak izateko probabilitatea.

b. Lor bedi 6 instalaziotik gutxienez 5 egokiak izateko probabilitatea.

c. Zein da itxaron daitekeen instalazio egokien kopurua? Eta bariantza?

Ebazpena

Bira X = "instalazio egokien kopurua, n instalaziotik" aldagai aleatorio binomiala, A = "instalazio egokia" gertaera eta $p = P(A) = 0{,}60$ eta $q = 1-p = 0{,}40$ probabilitateak. Orduan, X aldagaiari dagokion probabilitate-legea hauxe da:

$$P(X = x) = \binom{n}{x}(0{,}60)^x(0{,}40)^{n-x}, \qquad x = 0, 1, 2, \ldots, n.$$

a. Probabilitate-legean $x = 6$ eta $n = 3$ ordezkatuko dira.

$$P(X = 3) = \binom{6}{3}(0,60)^3(0,40)^3 = 0,2765.$$

b. Oraingoan $P(X \geq 5)$ kalkulatu behar da.

$$P(X \geq 5) = P(X = 5) + P(X = 6) = \binom{6}{5}(0,60)^5(0,40)^1 + \binom{6}{6}(0,60)^6 =$$

$$= 0,1866 + 0,0467 = 0,2333.$$

c. $E(X) = np = 6 \cdot 0,6 = 3,6$ berogailu zuzen instalatzea itxaron daiteke.

$Var(X) = npq = 6 \cdot 0,60 \cdot 0,40 = 1,44$ (berogailu)^2

3.5.14. ariketa

Fabrika batean 600 langile berri hartu dira. Edozein langilek lehenengo urtean lana galtzeko probabilitatea 0,01ekoa bada, zein da lehenengo urtean 10 langilek lana galtzeko probabilitatea? Zenbat langilek lana galtzea itxaron daiteke lehen urtean?

Ebazpena

Biz X = "lehenengo urtean lana galduko duen langile-kopurua" aldagai aleatorioa. Demagun aldagaiak λ parametroko Poisson-en banaketa duela, $\lambda = np = 600 \cdot 0,01 = 6$ izanik. Orduan aldagai horri dagokion probabilitate-funtzioa honako hau da:

$$P(X = x) = \frac{e^{-6}6^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Ondorioz, lehenengo galderaren erantzuna hauxe da:

$$P(X = 10) = \frac{e^{-6}6^{10}}{10!} = 0,0413\,.$$

Bigarren galderari buruz, itxaron daiteke $E(X) = \lambda = 6$ langilek lana galtzea lehenengo urtean.

3.5.15. ariketa

Industriako hondakinetako TCDD substantziaren kantitateari buruzko azken ikerkuntzen arabera, 100 hondakinetik 6 hondakinetan TCDD substantziaren kantitatea baimendutakoa da. 100 hondakinetatik aleatorioki hiru hondakin hartu dira.

a. Zein da TCDD substantziaren kantitatea hiru hondakinetan baimendutakoa ez izateko probabilitatea?

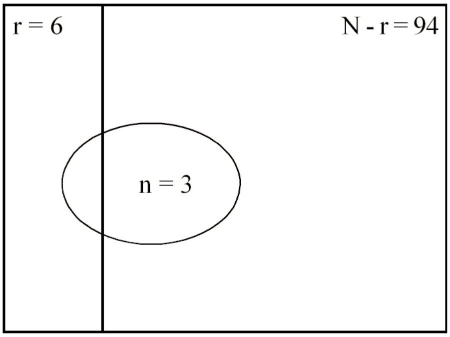
b. Kalkula bedi gehienez hondakin batek baimendutako TCDD substantziaren kantitatea izateko probabilitatea.

c. Lor bitez batez bestekoaren eta desbideratze tipikoaren balioak.

Ebazpena

Bira X = "baimendutako TCDD kantitatea duten hondakinen kopurua, 3 hondakinetik" aldagai aleatorioa A = "TCDD kantitate baimendua edukitzea" gertaera eta $p = P(A) = \dfrac{6}{100} = 0,06$ probabilitatea. X aldagai aleatorio diskretuak N = 100, n = 3 eta p = 0,06 parametroetako banaketa hipergeometrikoa du.

$$X\text{: } H(100,\ 3,\ 0{,}06)$$



3.3. irudia.

Orduan, aldagai honi dagokion probabilitate-legea hurrengo adierazpenaren bidez kalkula daiteke:

$$P(X = x) = \frac{\binom{6}{x}\binom{94}{3-x}}{\binom{100}{3}}, \qquad x = 0, 1, 2, 3.$$

a. Probabilitate-legean $x = 0$ balioa ordezkatuko da:

$$P(X = 0) = \frac{\binom{6}{0}\binom{94}{3}}{\binom{100}{3}} = \frac{1 \cdot \dfrac{94 \cdot 93 \cdot 92}{3!}}{\dfrac{100 \cdot 99 \cdot 98}{3!}} = 0,829.$$

b. Oraingoan $P(X \leq 1)$ kalkulatu behar da.

$$P(X \leq 1) = P(X = 0) + P(X = 1) = 0,829 + \frac{\binom{6}{1}\binom{94}{2}}{\binom{100}{3}} =$$

$$= 0,829 + \frac{\frac{6}{1!}\frac{94 \cdot 93}{2!}}{\frac{100 \cdot 99 \cdot 98}{3!}} = 0,829 + 0,162 = 0,991.$$

c. $E(X) = np = 3 \cdot 0,06 = 0,18$ hondakinek baimendutako TCDD kantitatea izatea itxaron daiteke.

Bariantzaren kalkulua burutu ondoren, erro karratua aplikatuz desbideratze tipikoaren balioa lortuko da.

$$Var(X) = npq\frac{N-n}{N-1} = 3 \cdot 0,06 \cdot 0,94 \cdot \frac{97}{99} = 1,658$$

$$\sigma_x = \sqrt{1,658} = 1,288.$$

3.5.16. ariketa

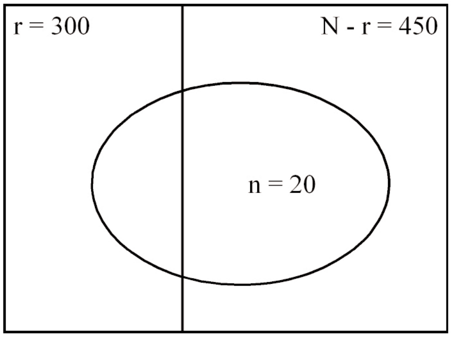
Ordenagailu-areto batean 750 ordenagailu daude eta 450 ordenagailuk A sistema eragilea eta 300 ordenagailuk B sistema eragilea erabiltzen dituzte. Ordenagailu bakoitzak sistema eragile bakarra erabiltzen du. Aleatorioki 20 ordenagailu hartu dira.

a. Zein da 6 ordenagailuk B sistema eragilea erabiltzeko probabilitatea?

b. Kalkula bedi gutxienez 12 ordenagailuk B sistema eragilea erabiltzeko probabilitatea?

Ebazpena

Biz X = "hogei ordenagailutik, B sistema eragilea erabiltzen duen ordenagailu-kopurua". X aldagai aleatorioak $N = 750$, $n = 20$ eta $p = \frac{300}{750} = 0,4$ parametroetako banaketa hipergeometrikoa du.



3.4. irudia.

Aldagai honi dagokion probabilitate-legea honela kalkula daiteke:

$$P(X=x)=\frac{\binom{300}{x}\binom{450}{20-x}}{\binom{750}{20}}, \qquad x=0,\ 1,\ \ldots,\ 20.$$

a. P(sei ordenagailuk B sistema) $=P(X=6)=\dfrac{\binom{300}{6}\binom{450}{14}}{\binom{750}{20}}$ adierazpenaren balioa kalkulatuz ebatzi daiteke. Hala ere, $\dfrac{n}{N}=\dfrac{20}{750}=0{,}027<0{,}1$ betetzen denez, banaketa binomialaren bidezko $X\approx Bin(n=20,\ p=0{,}4)$ hurbilketa erabiltzea komeni da. Ondorioz, emaitza honela kalkulatuko da:

$$P(X=6)=\binom{20}{6}(0{,}4)^6(0{,}6)^{14}=0{,}1244.$$

Beste era batean, azken emaitza hori banaketa binomialaren banaketa-funtzioaren taulako emaitzen bidez kalkula daiteke:

$$P(X=6)=F(6)-F(5)=0{,}2499-0{,}1255=0{,}1244.$$

b. Aurreko atalean bezala, X aldagaiari banaketa binomialaren hurbilketa egokituko zaio, eta banaketa horren banaketa-funtzioaren balioen taula erabiliz, gutxienez 12 ordenagailuk B sistema eragilea edukitzeko probabilitatea honela kalkulatuko da:

$$P(X\geq 12)=1-P(X\leq 11)=1-F(11)=1-0{,}9435=0{,}0565.$$

3.5.17. ariketa

a. Ikerketa estatistiko baten ondorioek erakusten dutenaren arabera, gaur egun enpresen % 45ek Interneten bidezko salmentak egiten ditu.

Aleatorioki 12 enpresa aukeratzen badira, zein da horietatik gutxienez 5 enpresak Interneten bidezko salmentak egiteko probabilitatea?

b. Herri bateko 10 enpresatik 3 enpresak Interneten bidezko salmentak egiten dituzte. Hamar enpresa horietatik aleatorioki 5 enpresa hartzen badira, zein da gutxienez enpresa batek Interneten bidezko salmentak egiteko probabilitatea?

Ebazpena

a. Biz X = “Interneten bidez salmentak egiten dituen enpresa-kopurua, 12 enpresen artean” aldagai aleatorioa. X aldagai aleatorioak n = 12 eta p = 0,45 parametroetako banaketa binomiala du, hots, X: Bin(12, 0,45).

Ondorioz,

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - \sum_{k=0}^{4} \binom{12}{k} (0,45)^k (0,55)^{12-k}$$

ebatziz kalkula daiteke. Beste era batean, banaketa binomialaren eranskineko banaketa-funtzioaren balioen taula erabiliz emaitza zuzenean lortuko da:

$P(X \geq 5) = 1 - P(X < 5) = 1 - P(X \leq 4) = 1 - F(4) = 1 - 0,3045 = 0,6955.$

b. Biz Y = “Interneten bidez salmentak egiten dituen enpresa-kopurua, 5 enpresen artean” aldagai aleatorioa. Y aldagai aleatorioak N = 10, n = 5 eta p = 0,3 parametroetako banaketa hipergeometrikoa du, Y: H(10, 5, 0,3) alegia.

Y aldagaiari dagokion probabilitate-legea hurrengo hau da:

$$P(Y = y) = \frac{\binom{3}{y}\binom{7}{5-y}}{\binom{10}{5}}, \quad y = 0, 1, 2, 3.$$

Beraz, probabilitate-legearen balioak erabiliko dira ondoko gertaeraren probabilitatea kalkulatzeko:

$P(Y \geq 1) = 1 - P(Y < 1) = 1 - P(Y = 0) =$

$$= 1 - \frac{\binom{3}{0}\binom{7}{5}}{\binom{10}{5}} = 1 - \frac{\frac{3!}{0!\ 3!} \cdot \frac{7!}{5!\ 2!}}{\frac{10!}{5!\ 5!}} = 1 - \frac{21}{252} = 1 - 0,0833 = 0,9167.$$

3.6. ARIKETA PROPOSATUAK

3.6.1. ariketa

Biz X aldagai aleatorio diskretua, $x = 10, 20, 30$ balioak hartzen dituena. Zehatz bedi hurrengo funtzioetatik zeintzuk diren probabilitate-funtzioak:

a. $f(10) = 0{,}3$, $f(20) = 0{,}6$, $f(30) = 0{,}1$.

b. $f(10) = 1/3$, $f(20) = 1/2$, $f(30) = 1/3$.

c. $f(10) = 0{,}4$, $f(20) = 0{,}8$, $f(30) = -0{,}2$.

3.6.2. ariketa

Biz $f(x)$ funtzioa:

$$f(x) = \begin{cases} 1/4, & x = 0 \\ 1/3, & x = 1 \\ 5/12, & x = 2 \\ 0, & \text{beste} \end{cases}$$

a. Froga bedi $f(x)$ funtzioa X aldagai aleatorioaren probabilitate-funtzioa dela.

b. Banaketa-funtzioa zehaztuz, lor bitez $P(X>1)$ eta $P(0 < X \leq 2)$ balioak.

c. Kalkula bitez X aldagai aleatorioaren batez bestekoa eta bariantza.

3.6.3. ariketa

Hurrengo funtzioa X aldagai aleatorio diskretu bati dagokion banaketa-funtzioa da:

$$F(x) = \begin{cases} 0, & x < 1 \\ 1/3, & 1 \leq x < 2 \\ 1/2, & 2 \leq x < 3 \\ 2/3, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

Kalkula bitez $f(x)$ probabilitate-funtzioa, $E(X)$ eta $Var(X)$.

3.6.4. ariketa

Bira X eta Y bi aldagai aleatorio diskretu independente, $E(X) = 1$, $E(Y) = 0$, $\sigma_x = 2$ eta $\sigma_y = 2$ direlarik. Froga ezazu ondoko baieztapenak egiak diren ala ez:

a. $X-Y+2$ aldagai aleatorioaren batez bestekoa 3 da.

b. $X-Y+2$ aldagai aleatorioaren desbideratze tipikoa 0 da.

3.6.5. ariketa

Lantegi batean istripuen % 70 laneko segurtasun-faltagatik gertatzen da.

a. Kalkula bedi aleatorioki hartutako lau istriputik bi laneko segurtasun-faltagatik izateko probabilitatea.

b. Zein da lau istriputik erdia baino gehiago laneko segurtasun-faltagatik izateko probabilitatea?

c. Aleatorioki zortzi istripu hartuz gero, zein da laneko segurtasun-faltagatik itxaron daitekeen istripu-kopurua?

3.6.6. ariketa

Ingeniaritza-enpresa batek goizero, batez beste, 1,5 eskari jasotzen ditu.

a. Kalkula bedi goiz batean ingeniaritza-enpresak eskaririk ez jasotzeko probabilitatea.

b. Zein da goiz batean ingeniaritza-enpresak hiru eskari baino gutxiago jasotzeko probabilitatea?

c. Lor bedi aste bateko laneko bost goizetan bost lan-egunetan ingeniaritza-enpresak bost eskari baino gutxiago jasotzeko probabilitatea

3.6.7. ariketa

Demagun inprimategi zehatz batean inprimatutako hamabost liburutik hiru liburuk inprimatze-erroreak dituztela. Hamabost liburutik aleatorioki bi aukeratu dira.

a. Zein da gutxienez liburu bat inprimatze-errorerik gabea izateko probabilitatea?

b. Kalkula bitez inprimatze-erroreen batez bestekoa eta desbideratze tipikoa.

3.6.8. ariketa

Trokelatzeko makina batek 50 pieza egin ditu, zeintzuetatik 45 pieza onargarriak diren. Aleatorioki 50 piezatik 20 pieza hartu dira.

a. Zein da pieza horietatik guztiak onargarriak izateko probabilitatea?

b. Zein da pieza horietatik hamabost edo hamasei pieza onargarriak izateko probabilitatea?

c. Zenbat pieza onargarri itxaron daitezke?

3.6.9. ariketa

Gripearen aurkako txertoak pertsona bati erreakzioa egiteko probabilitatea 0,01ekoa da. Aleatorioki 300 pertsona hartu dira.

a. Zein da gehienez 3 pertsonari txertoak erreakzioa egiteko probabilitatea.

b. Zein da gutxienez 4 eta gehienez 7 pertsonari gripearen aurkako txertoak erreakzioa egiteko probabilitatea?

3.6.10. ariketa

Lanpara fluoreszente baten iraupena gutxienez 5.000 ordukoa izateko probabilitatea 0,85ekoa da. Aleatorioki mota horretako 20 lanpara hartu dira.

a. Zein da 20 lanparatik 18 lanparak gutxienez 5.000 orduko iraupena izateko probabilitatea

b. Kalkula bedi 20 lanparatik gutxienez 15 lanparak gutxienez 5.000 orduko iraupena izateko probabilitatea.

c. Hogei lanparatik 17 lanpara gutxienez 5.000 orduko iraupenekoak dira. Hogei lanparatik aleatorioki lau lanpara hartu dira. Zein da bi edo hiru lanpara gutxienez 5.000 orduko iraupenekoak izateko probabilitatea?

3.6.11. ariketa

Tornu automatiko batean egunero, batez beste, bi aldaketa egiten dira. Tornu automatikoa ez dabil ongi egunean lau aldaketa baino gehiago egin behar direnean. Demagun tornuan egun batean izandako aldaketa-kopuruak ez duela eragiten hurrengo egunean tornuan izango den aldaketa-kopuruan. Zein da bi egun jarraian tornua ongi aritzeko probabilitatea?

4. Aldagai aleatorio jarraitua

4.1. BANAKETA-FUNTZIOA ETA DENTSITATE-FUNTZIOA

X aldagaiak tarte bateko edozein balio har badezake, X **aldagai aleatorio jarraitua** da. Orduan, X aldagai aleratorioaren heina $(-\infty,+\infty)$, (a, b), $(-\infty, a]$, $(b, +\infty)$, ... erakoa izan daiteke.

X aldagai aleatorio jarraituaren $F(x)$ **banaketa-funtzioa**k $F(x) = P(X \leq x)$ adierazten du, $x \in \mathbb{R}$ delarik.

$F(x)$ banaketa-funtzioaren propietate adierazgarrienak ondokoak dira:

i. $F(-\infty) = 0$, $F(+\infty) = 1$.

ii. $F(x)$ funtzioa [0,1] tartean bornatua da.

iii. $F(x)$ funtzioa ez-beherakorra da.

iv. $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$.

$F(x)$ banaketa-funtzioa deribagarria bada, $F'(x) \equiv f(x)$ funtzioari **dentsitate-funtzioa** deritzo.

$f(x)$ funtzioa dentsitate-funtzioa da baldin eta ondoko bi propietateak betetzen baditu:

i. $f(x) \geq 0$, $x \in \mathbb{R}$.

ii. $\int_{-\infty}^{+\infty} f(x)dx = 1.$

Gainera, $f(x)$ dentsitate-funtzioa ezagututa, $P(x_1 \leq X \leq x_2)$ kalkula daiteke:

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx.$$

Ondorioz, $P(x_1 \leq X \leq x_2)$, $P(x_1 < X < x_2)$, $P(x_1 \leq X < x_2)$ eta $P(x_1 < X \leq x_2)$ probabilitateen balioak berdinak dira. Halaber, $P(X = x) = 0$ da.

Bestalde, $f(x)$ dentsitate-funtzioa ezagutuz gero, $F(x)$ banaketa-funtzioa lor daiteke:

$$F(x) = \int_{-\infty}^{x} f(t)dt, \qquad \forall x \in \mathbb{R}.$$

4.2. ALDAGAI JARRAITUAREN BATEZ BESTEKOA ETA BARIANTZA

4.2.1. Batez bestekoa edo itxaropen matematikoa

X aldagai aleatorio jarraituaren **batez bestekoa** edo **itxaropen matematikoa**, honela definitzen da:

$$E(X) = \mu = \int_{-\infty}^{+\infty} xf(x)dx.$$

Bere propietate garrantzitsuenak hauek dira:

1. $E(k) = k$, k konstantea izanik.

2. Bira X_1 , X_2 , ..., X_n aldagai aleatorio jarraituak.

$$E(X_1 + X_2 + ...+ X_n) = E(X_1) + E(X_2)+ ... + E(X_n).$$

3. $E(kX) = kE(X)$, k konstantea izanik.

4. Bira X_1 , X_2 , ..., X_n aldagai aleatorio jarraituak eta k_i konstanteak, $i = 1, 2, ..., n$.

$$E(k_1X_1 + k_2X_2 + ...+ k_nX_n) = k_1E(X_1) + k_2E(X_2)+ ... + k_nE(X_n).$$

5. Bira X_1, X_2, ..., X_n aldagai aleatorio jarraitu independenteak.

$$E(X_1 \cdot X_2 \cdot ... \cdot X_n) = E(X_1)E(X_2) \dots E(X_n).$$

4.2.2. Aldagai aleatorio jarraituaren bariantza eta desbideratze tipikoa

X aldagai aleatorio jarraituaren **bariantza** honela definitzen da:

$$\sigma_x^2 = Var(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx.$$

X aldagai aleatorio jarraituaren bariantzaren propietate garrantzitsuenak ondokoak dira:

1. $Var(X) \geq 0$, edozein X aldagai aleatorio jarraitutarako.

2. $Var(k) = 0$, k konstantea izanik.

3. $Var(X) = E[(X - \mu)^2] = E(X^2) - \mu^2$.

4. Bira X_1, X_2, ..., X_n aldagai aleatorio jarraitu independenteak.

$$Var(X_1 + X_2 + ...+ X_n) = Var(X_1) + Var(X_2)+ ... + Var(X_n).$$

5. $Var(kX) = k^2Var(X)$, k konstantea izanik.

6. $Var(X + k) = Var(X)$, k konstantea izanik.

Hirugarren propietatea kontuan hartuz, praktikan, kalkuluak errazteko bariantzaren balioa hurrengo adierazpenaren bidez lor daiteke:

$$\sigma_x^2 = Var(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2.$$

Bariantzaren erro karratu positiboari **desbideratze tipikoa** deritzo:

$$\sigma_x = \sqrt{Var(x)}.$$

Tipifikazioa. Biz X aldagai aleatorio jarraitua, μ batez besteko eta σ desbideratze tipikoduna. Orduan, $Z = \dfrac{X-\mu}{\sigma}$ aldagai aleatorio jarraituak 0 batez bestekoa eta 1 desbideratze tipikoa ditu. Z aldagaiari X aldagaiaren **tipifikatua** deritzo.

4.2.3. Tchebychev-en teorema

Biz $f(x)$ dentsitate-funtzioko X aldagai aleatorio jarraitua, zeinak μ batez bestekoa eta σ desbideratze tipiko finitua dituen. Orduan, ondokoa beteko da:

$$P\left(|X-\mu| \le k\sigma\right) \ge 1 - \frac{1}{k^2}, \qquad \forall k > 0.$$

Teorema horren bidez X aldagai aleatorioaren balioak $[\mu - k\sigma, \mu + k\sigma]$ tartean izateko probabilitatearen behe-bornea kalkula daiteke.

4.3. ZENBAIT BANAKETA JARRAITU

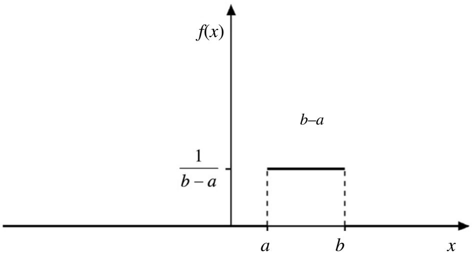
4.3.1. Banaketa uniformea

X aldagai aleatorio jarraituak **banaketa uniformea** du **[a,b] tartean,** hau da, X*:* $U(a,b)$, baldin eta ondoko dentsitate-funtzioa badu:

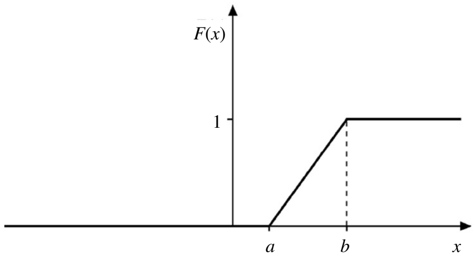
$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{beste} \end{cases}$$

Aldagai aleatorio honi dagokion banaketa-funtzioa hurrengoa da:

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$



4.1. irudia.



4.2. irudia.

[a,b] tartean banaketa uniformea duen *X* aldagai aleatorioaren batez bestekoa eta bariantzaren balioak erraz kalkula daitezke:

$$E(X)=\frac{a+b}{2}, \qquad Var(X)=\frac{(b-a)^2}{12}.$$

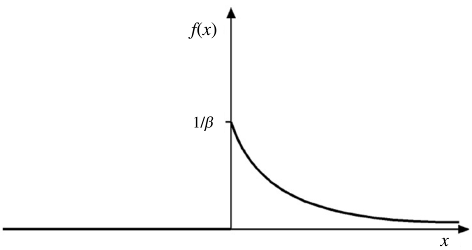
4.3.2. Banaketa esponentziala

X aldagai aleatorio jarraituak β parametroko **banaketa esponentziala** du baldin eta bere dentsitate-funtzioa hurrengoa bada:

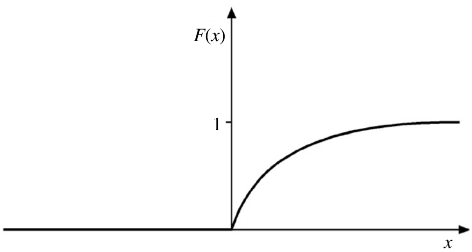
$$f(x)=\begin{cases}\frac{1}{\beta}e^{\frac{-x}{\beta}}, & x>0\\ 0, & x\le 0\end{cases} \qquad (\beta>0).$$

Banaketa-funtzioa lortzeko, $F(x)=\int_{-\infty}^{x} f(t)dt$ integrala ebatziko da. Horrela, banaketa esponentzialaren banaketa-funtzioa hauxe dugu:

$$F(x)=\begin{cases}1-e^{\frac{-x}{\beta}}, & x>0.\\ 0, & x\le 0.\end{cases}$$



4.3. irudia.



4.4. irudia.

Erraz kalkula daitezke banaketa esponentzialerako batez bestekoa eta bariantzaren balioak.

$$E(X)=\beta, \qquad Var(X)=\beta^2.$$

4.3.3. Banaketa normala

X aldagai aleatorio jarraituak μ eta σ parametroetako **banaketa normala** du, baldin eta ondoko dentsitate-funtzioa badagokio:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

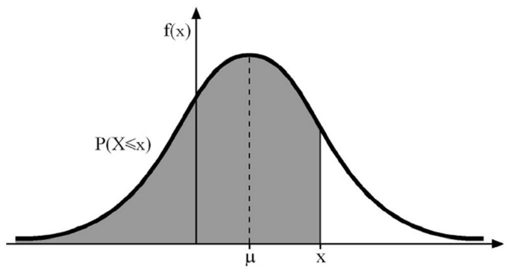
non μ eta σ zenbaki errealak diren. Laburbilduz, $X: N(\mu,\sigma)$.

Banaketa normalaren batez bestekoa μ eta desbideratze tipikoa σ dira.

$$E(X) = \mu, \qquad Var(X) = \sigma^2.$$

$N(\mu,\sigma)$ banaketadun X aldagai aleatorioaren banaketa-funtzioa honela definitzen da:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt.$$



4.5. irudia.

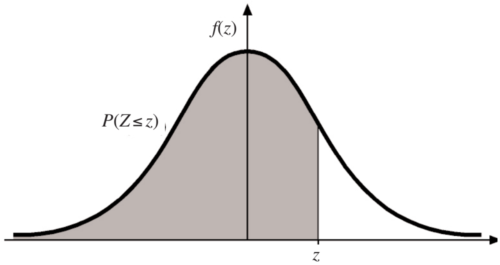
Integral hau zenbakizko metodoen bidez ebatzi behar da, eta kasu bakoitzean μ eta σ parametro ezberdinekiko egin beharko litzateke. Hala ere, X aldagaiak $N(\mu,\sigma)$ banaketa badu, $Z = \frac{X-\mu}{\sigma}$ aldagai tipifikatuaren banaketa N(0,1) izango da, eta $Z: N(0,1)$ aldagai aleatoriorako banaketa-funtzioaren balioak taularatuta daudenez, X aldagairako banaketa-funtzioaren balioak kalkulatu ahal izango dira:

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P(Z \le z).$$

non $Z = \frac{X-\mu}{\sigma}$ aldagaia eta $z = \frac{x-\mu}{\sigma}$ balioa diren.

$\mu = 0$ eta $\sigma = 1$ parametroetako **banaketa normal tipiko** edo **estandarra** duen Z aldagai aleatorio jarraituari ondoko dentsitate-funtzioa dagokio:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}, \ \forall z \in \mathbb{R}.$$



4.6. irudia.

4.4. BANAKETEN ARTEKO KONBERGENTZIA

Atal honetan banaketa binomial, Poisson-en banaketa eta banaketa normalaren arteko erlazioak adieraziko dira. Batzuetan hurbilketa aplikatzean, aldagai aleatorio diskretua aldagai aleatorio jarraitu batez hurbilduko da. Kasu horietan 0,5eko faktorearen bidezko **jarraitutasun-zuzenketa** aplikatuko da.

4.4.1. Banaketa binomialaren eta banaketa normalaren arteko erlazioa

Biz X aldagai aleatorioa n eta p parametroetako banaketa binomialekoa. n "handia" eta p "txikia" direnean, orduan banaketa binomiala $N(np, \sqrt{npq})$ banaketaz hurbil daiteke. Hau da, $n \to \infty$ eta $p \to 0$ konbergitzen dutenean, hurrengo hurbilketa erabil daiteke:

$$Bin(n, p) \cong N(np, \sqrt{npq}).$$

Praktikan, hurbilketa egokia da $np > 5$ eta $nq > 5$ direnean.

4.4.2. Poisson-en banaketaren eta banaketa normalaren arteko erlazioa

Demagun X aldagai aleatorio diskretuak λ parametroko Poisson-en banaketa duela. Orduan, $\lambda \to \infty$ jotzen duenean, Poisson-en banaketa $N(\lambda, \sqrt{\lambda})$ banaketaz hurbil daiteke. Hau da, $\lambda \to \infty$ konbergitzen duenean, ondoko hurbilketa erabiliko da:

$$\text{Poisson}\ (\lambda) \cong N(\lambda, \sqrt{\lambda}).$$

Praktikan, hurbilketa egokia da $\lambda > 18$ denean.

4.5. ARIKETA EBATZIAK

4.5.1. ariketa

Biz X aldagai aleatorioari dagokion $f(x)$ dentsitate-funtzioa hurrengoa:

$$f(x) = \begin{cases} kx^{-1/2}, & 0 < x \le 20 \\ 0, & \text{beste} \end{cases}$$

a. Zein da k konstantearen balioa?

b. Lor bedi X aldagai aleatorio jarraituaren banaketa-funtzioa.

c. Kalkula bitez P($10 < X < 30$) eta $E(X)$ balioak.

Ebazpena

a. $f(x)$ dentsitate-funtzioa denez,

$$1 = \int_{-\infty}^{+\infty} f(x)dx = k\int_0^{20} \frac{1}{\sqrt{x}}\,dx = 2kx^{1/2}\Big|_0^{20} = 2k\sqrt{20} = 4k\sqrt{5} \qquad \rightarrow$$

$$k = \frac{\sqrt{5}}{20}.$$

b. X aldagai aleatorio jarraituaren $F(x)$ banaketa-funtzioa:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

$$x \le 0: \qquad F(x) = \int_0^x 0dt = 0$$

$$0 < x \le 20: \qquad F(x) = \int_{-\infty}^{0} 0dt + \frac{\sqrt{5}}{20}\int_0^x \frac{dt}{\sqrt{t}} = \frac{\sqrt{5t}}{10}\Bigg|_0^x = \frac{\sqrt{5x}}{10}$$

$$x > 20: \qquad F(x) = \int_{-\infty}^{0} 0dt + \frac{\sqrt{5}}{20}\int_0^{20} \frac{dt}{\sqrt{t}} + \int_{20}^{x} 0dt = \frac{\sqrt{5t}}{10}\Bigg|_0^{20} = 1.$$

Ondorioz, banaketa-funtzioa hauxe da:

$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{\sqrt{5x}}{10}, & 0 < x \le 20 \\ 1, & x > 20 \end{cases}$$

c. $P(10 < X < 30) = F(30) - F(10) = 1 - \dfrac{\sqrt{50}}{10} = 0{,}293$

$$E(X) = \int_0^{20} xf(x)dx = \frac{\sqrt{5}}{20}\int_0^{20} x \cdot x^{-1/2}dx = \frac{x\sqrt{5x}}{30}\Bigg|_0^{20} = 6{,}67.$$

4.5.2. ariketa

X aldagai aleatorio jarraituari ondoko dentsitate-funtzioa dagokio:

$$f(x) = \begin{cases} kx^3 e^{-x^2}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

a. Zein da k konstantearen balioa?

b. Lor bedi X aldagai aleatorio jarraituaren banaketa-funtzioa.

c. Kalkula bitez $P(X > 2{,}5)$ eta $P(0 < X < 1{,}5)$ balioak.

Ebazpena

a. $f(x)$ funtzioa dentsitate-funtzioa denez, hurrengo axiomak beteko dira:

i. $f(x) \geq 0, \quad \forall x \in \mathbb{R}$.

ii. $\int f(x)\,dx = 1$.

Beraz, lehenengo axiomatik $k > 0$ ondorioztatzen da. Era berean, bigarren axiomatik, k konstantearen balioa lor daiteke, non $u = x^2$ den.

$$1 = \int f(x)\,dx = \int_0^\infty kx^3 e^{-x^2}\,dx = \frac{k}{2}\int_0^\infty ue^{-u}\,du = \frac{k}{2}\cdot I.$$

Zatikako integrazioa erabiliz I integralaren balioa kalkula daiteke:

$$1 = \int_0^\infty ue^{-u}\,du = -ue \quad \Big|_0^\infty + \int_0^\infty e^{-u}\,du = 0 - e \quad \Big|_0^\infty = 0 - (0-1) = 1.$$

Ondorioz, k konstantearen balioa zehaztuko da.

$$1 = \int f(x)\,dx = \int_0^\infty kx^3 e^{-x^2}\,dx = \frac{k}{2}\int_0^\infty ue^{-u}\,du = \frac{k}{2} \quad \to$$

$$k = 2.$$

b. Definizioz $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$ da. Kasu honetarako hurrengoa lor daiteke:

$$x \leq 0: \qquad F(x) = \int_{-\infty}^{x} 0\,dt = 0$$

$$x > 0: \qquad F(x) = \int_0^x 2t^3 e^{-t^2}\,dt = \int_0^{x^2} ue^{-u}\,du = -ue^{-u} - e^{-u}\Big|_0^{x^2} = 1 - e^{-x^2}(x^2+1).$$

X aldagaiari dagokion banaketa-funtzioaren adierazpen analitikoa:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-x^2}(\mathrm{x}^2+1), & x > 0 \end{cases}$$

c. $P(X > 2{,}5) = 1 - F(2{,}5) = e^{-(2{,}5)^2}((2{,}5)^2 + 1) = 0{,}0140$

$P(0 < X < 1{,}5) = F(1{,}5) - F(0) = 1 - e^{-(1{,}5)^2}((1{,}5)^2 + 1) = 0{,}6575$

4.5.3. ariketa

Demagun lantegi bateko jardueren iraupenak banaketa uniformea duela, 12 orduko batez bestekoa eta 3 ordu karratuko bariantza edukiz.

a. Kalkula bedi edozein jarduerak 12 ordu baino gutxiago irauteko probabilitatea.

b. Lor bedi *c* konstantearen balioa, edozein jarduerak *c* ordu baino gehiago irauteko probabilitatea 0,75ekoa dela jakinik.

Ebazpena

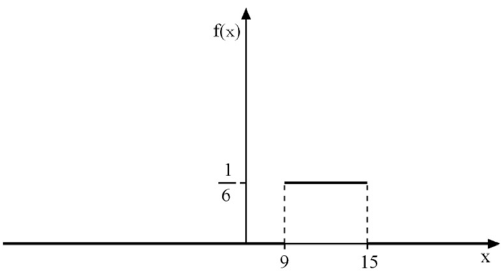
Biz X = "lantegiko jardueren iraupena" aldagai aleatorio jarraitua banaketa uniformekoa $[a, b]$ tartean.

$X: U(a,b)$ aldagaiaren batez bestekoa eta bariantzaren balioak ezagunak direnez,

$$E(X) = \frac{a+b}{2} = 12, \qquad Var(X) = \frac{(b-a)^2}{12} = 3$$

ekuazio-sistema ebatziz banaketaren a eta b parametroen balioak lortuko dira.

$$a = 9, \qquad b = 15$$



4.7. irudia.

a. $P(X < 12) = F(12)$ kalkulatu behar da. Kasu honetan, $X{:}U(9,15)$ aldagaiari dagokion banaketa-funtzioa honako hau da:

$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x-9}{6}, & 9 \le x \le 15 \\ 1, & x > 15 \end{cases}$$

Ondorioz, hona hemen edozein jarduerak 12 ordu baino gutxiago irauteko probabilitatea:

$$P(X < 12) = F(12) = \frac{12-9}{6} = 0,5.$$

b. Oraingoan, 0,75 = $P(X > c)$ dela jakinik, c konstantearen balioa zehaztu behar da.

$$0,75 = P(X > c) \quad \rightarrow$$

$$1 - 0,75 = P(X \le c) = \frac{c-9}{6} \quad \rightarrow$$

$$c = 6 \cdot (1 - 0,75) + 9 = 10,5.$$

4.5.4. ariketa

Bira X = “LH motorraren kontsumoa” aldagai aleatorioa eta $f(x)$ dagokion dentsitate-funtzioa:

$$f(x) = \begin{cases} ke^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

a. Zein da k konstantearen balioa?

b. Lor bedi X aldagai aleatorio jarraituaren banaketa-funtzioa.

c. Kalkula bedi LH motorraren kontsumoa 10-100 litro bitartekoa izateko probabilitatea eta adierazi bere esanahia.

Ebazpena

a. $f(x)$ dentsitate-funtzioa denez,

$$1 = \int_{-\infty}^{+\infty} f(x)dx = k\int_{0}^{+\infty} e^{-x}dx = -ke^{-x}\Big|_0^{\infty} = k \quad \rightarrow$$

$$k = 1.$$

Nabaria denez, X aldagai aleatorioak $\beta = 1$ parametroko banaketa esponentziala du.

b. X aldagai aleatorio jarraituaren $F(x)$ banaketa-funtzioa:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

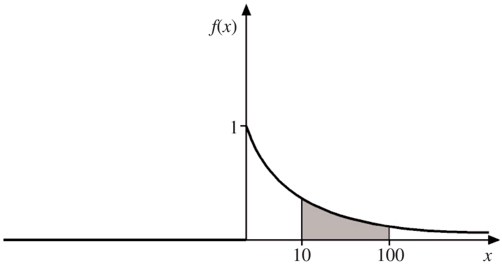
$$x \le 0: \quad F(x) = \int_{-\infty}^{x} 0dt = 0$$

$$x > 0: \quad F(x) = \int_{0}^{x} e^{-t}dt = -e^{-t}\Big|_0^{x} = 1 - e^{-x}.$$

Ondorioz, X aldagai aleatorioari dagokion banaketa-funtzioa hauxe da:

$$F(x) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

c. $P(10 < X < 100) = F(100) - F(10) = (1-e^{-100})-(1-e^{-10}) = e^{-10}- e^{-100} = 0{,}000045$



4.8. irudia.

$P(10 < X < 100) = 0{,}000045$ balioak, dentsitate-funtzioak, abzisa-ardatzak eta $x = 10$ eta $x = 100$ zuzenek mugatzen duten eremuaren azalera adierazten du.

4.5.5. ariketa

Aldagai aleatorio batek $\mu = 65{,}6$ batez bestekodun banaketa normala du.

a. Aldagaiak 60 baino txikiagoak diren balioak hartzeko probabilitatea 0,352 dela jakinda, kalkula bedi aldagai horren desbideratze tipikoa.

b. Zein x baliok uzten du banaketaren % 87,9 bere eskuinean?

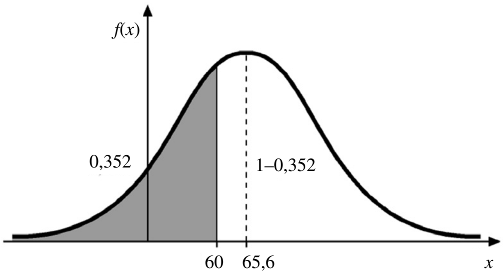
Ebazpena

Biz X: $N(65{,}6, \sigma)$.

a. Enuntziatuko datutik $0{,}352 = P(X < 60)$ da. Aldagaiaren tipifikazioa erabiliz,

$$0{,}352 = P(X < 60) = P\left(\frac{X-\mu}{\sigma} < \frac{60-65{,}6}{\sigma}\right)$$

ondorioztatzen da, non $Z = \dfrac{X-\mu}{\sigma}$ aldagaia banaketa normal estandarrekoa den.



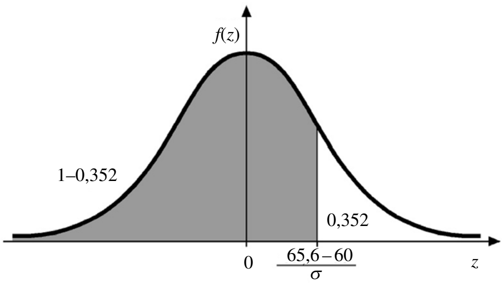
4.9. irudia.

Bestalde, kontrako gertaeraren probabilitatea erabiliz,

$$1-0{,}352 = P(X \geq 60) = P\left(Z \geq \frac{60-65{,}6}{\sigma}\right)$$

da, non $\dfrac{60-65{,}6}{\sigma} < 0$ den. Orduan, baliokidea den hurrengo probabilitatea plantea daiteke:

$$1-0{,}352 = P\left(Z \leq \frac{65{,}6-60}{\sigma}\right) = \phi\left(\frac{65{,}6-60}{\sigma}\right).$$



4.10. irudia.

Ondorioz, banaketa normal estandarraren banaketa-funtzioaren taula kontsultatuz,

$$\frac{65{,}6-60}{\sigma} = 0{,}38$$

da, eta hemendik desbideratze tipikoa banan daiteke:

$$\sigma = \frac{65{,}6-60}{0{,}38} = 14{,}74.$$

b. Hasteko, $P(X \geq x) = 0{,}879$ bada, x balioa 65,6 baino txikiagoa da eta $\dfrac{x-65{,}6}{14{,}74} < 0$ da. Orduan,

$$0{,}879 = P\left(\frac{X-65{,}6}{14{,}74} \geq \frac{x-65{,}6}{14{,}74}\right) = \phi\left(\frac{65{,}6-x}{14{,}74}\right).$$

lor daiteke, non $Z = \dfrac{X - 65,6}{14,74}$ aldagaia banaketa normal estandarrekoa den.

Banaketa normal tipikoaren banaketa-funtzioaren taula kontsultatuz,

$$\frac{-x + 65,6}{14,74} = 1,17$$

lortuko da, eta hemendik *x* gaiaren balioa ondoriozta daiteke.

$$x = 48,35.$$

4.5.6. ariketa

Proba berezi batean aztertutako populazioaren % 25,14k 6,5 puntuazioa gainditu zuen, % 61,51ren puntuazioa 4,5 eta 6,5 bitartekoa izan zen eta % 13,35ek ez zuen gainditu 4,5 puntuazioa. Demagun puntuazioen banaketa normala dela. Kalkula bitez batez besteko puntuazioa eta desbideratze tipikoaren balioak.

Ebazpena

Enuntziatuko datuetatik:

$P(X > 6,5) = 0,2514$, $P(4,5 \leq X \leq 6,5) = 0,6151$, $P(X < 4,5) = 0,1335$

X = "proban lortutako puntuazioa" aldagai aleatorio jarraituak banaketa $N(\mu, \sigma)$ du.

Kalkula ditzagun banaketaren bi parametroen balioak:

$$0,2514 = P(X > 6,5) = P\left(\frac{X - \mu}{\sigma} > \frac{6,5 - \mu}{\sigma}\right) \quad \rightarrow$$

$$0,7486 = P\left(\frac{X - \mu}{\sigma} \leq \frac{6,5 - \mu}{\sigma}\right)$$

$Z = \dfrac{X - \mu}{\sigma}$ aldagaiak banaketa normal tipikoa duenez, dagokion banaketa-funtzioaren balioen taulan kontsultatuz:

$$\frac{6,5 - \mu}{\sigma} = 0,67. \tag{1}$$

Era berean, jakinda:

$$0,1335 = P(X < 4,5) = P\left(\frac{X - \mu}{\sigma} < \frac{4,5 - \mu}{\sigma}\right),$$

$\frac{4,5-\mu}{\sigma} < 0$ dela ondoriozta daiteke. Orduan, simetria erabiliz:

$$0,1335 = P\left(\frac{X-\mu}{\sigma} > -\frac{4,5-\mu}{\sigma}\right)$$

berdintza beteko da, eta kontrako gertaeraren probabilitatea erabiliz:

$$1-0,1335 = P\left(\frac{X-\mu}{\sigma} \leq -\frac{4,5-\mu}{\sigma}\right)$$

probabilitatea dugu.

Banaketa normal tipikoaren banaketa-funtzioaren balioen taula kontsultatuko da:

$$-\frac{4,5-\mu}{\sigma} = 1,11 \qquad (2)$$

(1)-(2) sistema ebatziz, batez bestekoa eta desbideratze tipikoaren balioak lor daitezke.

$\mu = 5,74 \qquad \sigma = 1,12.$

__4.5.7. ariketa__

Substantzia erradioaktibo baten X bizitza-aldia, 5 minutuko batez bestekodun banaketa esponentzialekoa da.

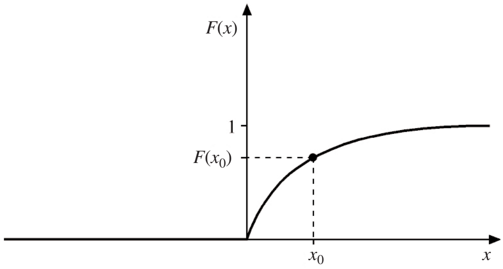
a. Zein da substantzia horrek 4 minutu eta 6 minutu arteko bizitza-aldia izateko probabilitatea?

b. Jakinda substantziaren bizitza-aldia gutxienez 4 minutukoa dela, zein da 5 minutu baino gutxiagoko bizitza-aldia izateko probabilitatea?

__Ebazpena__

Biz X = “substantzia erradioaktiboaren bizitza-aldia” aldagai aleatorioa. X aldagai aleatorioak banaketa esponentziala duenez, eta $E(X) = 5 = \beta$ denez, orduan X aldagai aleatorioari dagokion banaketa-funtzioa hurrengoa da:

$$F(x) = \begin{cases} 1-e^{-\frac{1}{5}x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



4.11. irudia.

a. Substantziak 4 minutu eta 6 minutu arteko bizitza-aldia izateko probabilitatea honela kalkula daiteke:

$$P(4 \le X \le 6) = F(6) - F(4) = \left(1 - e^{\frac{-6}{5}}\right) - \left(1 - e^{\frac{-4}{5}}\right) = e^{\frac{-4}{5}} - e^{\frac{-6}{5}} = 0{,}148.$$

b. Oraingoan baldintzazko probabilitatea aplikatu behar da:

$$P(X < 5 | X \ge 4) = \frac{P\left[(X < 5) \cap (X \ge 4)\right]}{P(X \ge 4)} = \frac{P(4 \le X < 5)}{P(X \ge 4)} = \frac{F(5) - F(4)}{1 - F(4)} = \frac{e^{-4/5} - e^{-1}}{e^{-4/5}} = 0{,}1812.$$

4.5.8. ariketa

Konpainia batean asteko telefono-gastuak, uniformeki, 100 euro eta 150 euro bitartekoak dira.

a. Kalkula bitez konpainiaren asteko batez besteko telefono-gastua eta desbideratze tipikoa.

b. Zein da aste batean 120 euro baino gutxiagoko telefono-gastua izateko probabilitatea?

Ebazpena

Biz X = “konpainiaren asteko telefono-gastua” aldagai aleatorioa. X aldagai aleatorioak [100, 150] tartean banaketa uniformea du.

a. Oro har, banaketa uniforme bateko batez bestekoa eta bariantza honela kalkulatzen dira:

$$E(X) = \frac{a+b}{2}, \qquad Var(X) = \frac{(b-a)^2}{12}.$$

Kasu honetan, $a = 100$ eta $b = 150$ balioak ordezkatuko dira.

$$E(X) = \frac{100 + 150}{2} = 125 \text{ euro}$$

$$Var(X) = \frac{(150 - 100)^2}{12} = 208,33 \text{ (euro)}^2 .$$

Desbideratze tipikoa, bariantzaren erro karratu positiboa kalkulatuz lortuko da.

$$\sigma_x = \sqrt{Var(X)} = 14,43 \text{ euro.}$$

b. $P(X < 120) = F(120)$ probabilitatearen balioa kalkulatu behar da. Horretarako, banaketa uniformearen banaketa-funtzioaren adierazpena kontuan har daiteke.

$$F(x) = \begin{cases} 0, & x < 100 \\ \dfrac{x - 100}{50}, & 100 \le x \le 150 \\ 1, & x > 150 \end{cases}$$

Ondorioz, aste batean 120 euro baino gutxiagoko telefono-gastua izateko probabilitatea hauxe da:

$$P(X < 120) = F(120) = \frac{120 - 100}{50} = 0,4 .$$

4.5.9. ariketa

Saiakera baterako hodien diametroen luzerek 5 cm-ko batez besteko eta 0,04 cm-ko desbideratze tipikodun banaketa normalari jarraitzen diote. Saiakera aurrera eramateko, hodien diametroek gutxienez 4,90 cm-ko neurria izan behar dute eta gehienez 5,04 cm-koa. Zein da saiakerarako baliagarria den hodien proportzioa?

Ebazpena

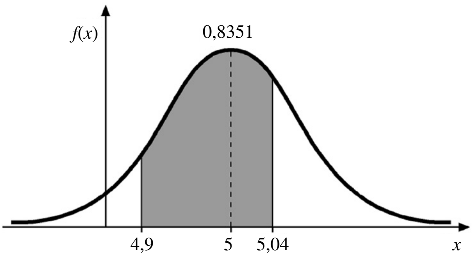
Biz X = “hodien diametroen luzerak” aldagai aleatorioa. X aldagai aleatorioak $N(5,\ 0{,}04)$ banaketa du.

$$P(4,90 \le X \le 5,04) = P\left(\frac{4,90 - 5}{0,04} \le \frac{X - 5}{0,04} \le \frac{5,04 - 5}{0,04}\right) = P(-2,5 \le Z \le 1) =$$

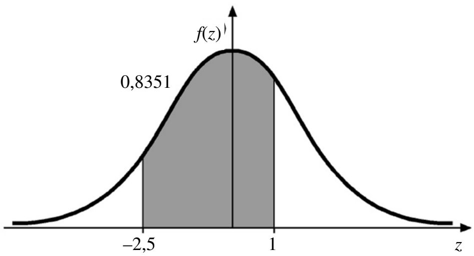
$$= \phi(1) - \phi(-2{,}5) = \phi(1) - [1 - \phi(2{,}5)] =$$

$$= 0{,}8413 - 1 + 0{,}9938 = 0{,}8351.$$

Adierazpen grafikoa:



4.12. irudia.



4.13. irudia.

Beraz, saiakerarako baliagarria den hodien proportzioa % 83,51 dela esan daiteke.

4.5.10. ariketa

Produktu jakin bati buruzko inkestan, galdetutako pertsonen artean hamarretik batek produktua gustuko du. Aleatorioki 120 pertsonako lagina hartu da eta guztiek erantzun diote produktuari buruzko inkestari.

a. Zein da gutxienez 18 pertsonak produktu hori gustuko izateko probabilitatea?

b. Zein da gutxienez 12 pertsonak eta gehienez 19 pertsonak produktu hori gustuko izateko probabilitatea?

Ebazpena

Biz X=“produktua gustuko duen pertsona-kopurua”, $p = P$(produktua, gustuko) = 0,1 edozein pertsonarentzat eta n = 120 laginaren tamaina. X aldagai aleatorioa banaketa binomialekoa dela onar daiteke.

$$X\text{: } Bin(120,\ 0{,}1).$$

a. Banaketa binomialaren probabilitate-legea erabiliz, hurrengo kalkulua egin beharko litzateke:

$$P(X \geq 18) = 1 - P(X < 18) = 1 - \sum_{k=0}^{17} \binom{120}{k} (0{,}1)^k (0{,}9)^{120-k}.$$

Hala ere, np >5 eta nq > 5 direnez, $N(np, \sqrt{npq})$ banaketa normalaren bidezko hurbilketa erabil daiteke, non np = 12 eta $\sqrt{npq} = \sqrt{120 \cdot 0{,}1 \cdot 0{,}9} = 3{,}29$ diren. Orduan, banaketa diskretu bat banaketa jarraitu baten bidez hurbilduko denez, 0,5eko jarraitutasunagatiko zuzenketa-faktorea aplikatuko da.

$$P(X \geq 18) = P(X > 17{,}5) = P\left(\frac{X-12}{3{,}29} > \frac{17{,}5-12}{3{,}29}\right) = P(Z > 1{,}67) =$$

$$= 1 - \phi(1{,}67) = 1 - 0{,}9525 = 0{,}0475.$$

b. Era berean, hurbilketa berdina erabiliz $P(12 \leq X \leq 19)$ probabilitatearen balioa kalkula daiteke.

$$P(12 \leq X \leq 19) = P(11{,}5 < X < 19{,}5) = P\left(\frac{11{,}5-12}{3{,}29} < \frac{X-12}{3{,}29} < \frac{19{,}5-12}{3{,}29}\right) =$$

$$= P(-0{,}15 < Z < 2{,}28) = \phi(2{,}28) - \phi(-0{,}15) =$$

$$= \phi(2{,}28) - 1 + \phi(0{,}15) = 0{,}9887 - 1 + 0{,}5596 = 0{,}5483.$$

4.5.11. ariketa

Hozte-sistemak konpontzen dituen enpresa batek hilero, batez beste, 20 hozte-sistema konpontzen ditu. Kalkula bedi hil batean enpresak duen:

a. Zortzi hozte-sistema konpontzeko probabilitatea.

b. Gutxienez bost hozte-sistema konpontzeko probabilitatea.

c. Gehienez sei hozte-sistema konpontzeko probabilitatea.

Ebazpena

Biz X = "enpresak hilean konpontzen dituen hozte-sistemen kopurua". Demagun X aldagai aleatorioak Poisson (λ) banaketa duela. Poisson-en banaketaren batez bestekoa $E(X) = \lambda$ denez, eta kasu honetan $E(X) = 20$ denez, orduan X: Poisson (20) da.

$\lambda = 20 > 18$ denez, Poisson-en banaketa hau $N(20, \sqrt{20})$ banaketaren bidez hurbil daiteke, eta ondoko probabilitateak hurbilketa erabiliz kalkulatuko dira:

a. $P(X = 8) = P(7{,}5 < X < 8{,}5) = P\left(\frac{7{,}5-20}{\sqrt{20}} < \frac{X-20}{\sqrt{20}} < \frac{8{,}5-20}{\sqrt{20}}\right) =$

$$= P(-2{,}80 < Z < -2{,}57) = P(2{,}57 < Z < 2{,}80) =$$

$$= \phi(2{,}80) - \phi(2{,}57) = 0{,}9974 - 0{,}9949 = 0{,}0025.$$

b. $P(X \geq 5) = P(X > 4{,}5) = P\left(\frac{X-20}{\sqrt{20}} > \frac{4{,}5-20}{\sqrt{20}}\right) = P(Z > -3{,}47) =$

$$= P(Z < 3{,}47) = \phi(3{,}47) = 0{,}9997.$$

c. $P(X \leq 6) = P(X < 6{,}5) = P\left(\frac{X-20}{\sqrt{20}} < \frac{6{,}5-20}{\sqrt{20}}\right) = P(Z < -3{,}02) =$

$$= \phi\,(-3{,}02) = 1 - \phi(3{,}02) = 1 - 0{,}9987 = 0{,}0013.$$

4.5.12. ariketa

Mila ordu baino gehiago etengabe lan egitean, mikrotxipen % 15ek porrot egiten du. Aleatorioki, 200 mikrotxip hartu dira.

a. Mila ordu baino gehiago etengabe lan egitean, zenbat mikrotxipek itxaron daiteke porrot egitea?

b. Kalkula bedi 45 mikrotxip baino gutxiagok mila ordu baino gehiago etengabe lan egitean porrot egiteko probabilitatea.

Ebazpena

Biz X = "mila ordu baino gehiago etengabe lan egitean, porrot egiten duen mikrotxip-kopurua". P(porrot, mila ordu baino gehiago etengabe lan egitean) = p = 0,15 konstantea bada mikrotxip guztietarako eta n = 200 mikrotxip independente badira, orduan X aldagai aleatorioak banaketa binomiala duela esan daiteke.

$$X\text{: } Bin\ (200,\ 0{,}15).$$

a. Ondorioz, mila ordu baino gehiago etengabe lan egitean, $E(X) = np = 200 \cdot 0{,}15 = 30$ mikrotxipek porrot egitea itxaron daiteke.

b. $P(X < 45)$ balioa kalkulatu behar da. Kasu honetan banaketa binomialaren probabilitate-legea erabili ordez, hobe da banaketa normalaren bidezko hurbilketa aplikatzea. $np = 30 > 5$ eta $nq = 170 > 5$ direnez:

$$Bin(200,\ 0{,}15) \cong N(30,\ 5{,}05)$$

hurbilketa aplikatuko da, non $np = 30$ eta $\sqrt{npq} = \sqrt{200 \cdot 0{,}15 \cdot 0{,}85} = 5{,}05$ diren.

Orduan, 0,5eko jarraitutasunagatiko zuzenketa-faktorea aplikatuz:

$$P(X < 45) = P(X < 44{,}5) = P\left(\frac{X-30}{5{,}05} < \frac{44{,}5-30}{5{,}05}\right) = P(Z < 2{,}87) = 0{,}9979$$

da, non Z aldagaiak banaketa normal tipikoa duen.

4.5.13. ariketa

Demagun trafikogune batera orduro iristen den ibilgailu-kopuruak 60 batez beste-kodun Poisson-en banaketari jarraitzen diola.

a. Kalkula bedi ordu batean trafikogunera 50 eta 70 ibilgailu bitartean iristeko probabilitatea.

b. Lor bedi minutu batean trafikogunera gutxienez bi ibilgailu iristeko probabilitatearen balioa.

Ebazpena

Enuntziatuko datuen arabera, X = "trafikogunera orduro iristen den ibilgailu-kopurua" aldagai aleatorioak $E(X) = 60$ batez bestekodun Poisson-en banaketa du. Laburbilduz, X: Poisson (60) da.

$\lambda = 60 > 18$ denez, Poisson-en banaketa hau $N(60, \sqrt{60})$ banaketaren bidez hurbil daiteke.

a. $$P(50 \le X \le 70) = P(49,5 < X < 70,5) = P\left(\frac{49,5-60}{\sqrt{60}} < \frac{X-60}{\sqrt{60}} < \frac{70,5-60}{\sqrt{60}}\right) =$$

$$= P(-1,36 < Z < 1,36) = 2\phi(1,36) - 1 = 2 \cdot 0,9131 - 1 = 0,8262.$$

b. Biz Y = "trafikogunera minuturo iristen den ibilgailu-kopurua" aldagai aleatorioa. Y aldagai aleatorioak $\lambda = 1$ parametroko Poisson-en banaketa duela esan daiteke. Orduan, banaketa horren probabilitate-legearen adierazpena kontuan hartuz:

$$P(Y = y) = \frac{e^{-1}1^y}{y!} = \frac{e^{-1}}{y!}, \qquad y = 0,\ 1,\ 2,\ \ldots$$

eskatutako probabilitatea honela kalkula daiteke:

$$P(Y \ge 2) = 1 - P(Y < 2) = 1 - P(Y = 0) - P(Y = 1) =$$

$$= 1 - \frac{e^{-1}}{0!} - \frac{e^{-1}}{1!} = 1 - 2e^{-1} = 0,264.$$

4.6. ARIKETA PROPOSATUAK

4.6.1. ariketa

Biz $f(x)$ funtzioa:

$$f(x) = \begin{cases} kx^3, & 2 \le x \le 4 \\ 0, & \text{beste} \end{cases}$$

a. Zein da k konstantearen balioa $f(x)$ funtzioa X aldagai aleatorio jarraituaren dentsitate-funtzioa izateko?

b. Lor bedi X aldagai aleatorio jarraituari dagokion banaketa-funtzioa.

c. Kalkula bitez $P(0 < X < 2,5)$, $P(X > 3)$ eta $P(X \le 3,5)$ probabilitateen balioak.

4.6.2. ariketa

X aldagai aleatorio jarraituaren banaketa-funtzioa honakoa da:

$$F(x) = \begin{cases} 0, & x < 0 \\ 3x, & 0 \le x \le 1/3 \\ 1, & x > 1/3 \end{cases}$$

a. Lor bedi X aldagai aleatorioari dagokion dentsitate-funtzioa.

b. Kalkula bitez $P(0{,}2 \le X \le 0{,}7)$, $P(X < 0{,}32)$ eta $P(X \ge 0{,}27)$ probabilitateen balioak.

4.6.3 ariketa

Demagun osagai-mota berezi baten iraupena urtetan aztertu nahi dela. Jakinda aldagaiari dagokion dentsitate-funtzioa

$$f(x) = \begin{cases} \dfrac{e^{-x/2}}{2}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

dela, kalkula bitez hurrengo gertaeren probabilitateen balioak:

a. Osagaiek urte bat eta hiru urte bitarteko iraupena izatekoa.

b. Osagaiek gutxienez urte-erdiko iraupena izatekoa.

4.6.4. ariketa

Biz X = “bi konposatu kimikoren artean erreakzioa gertatzeko behar den denbora”. Demagun X aldagai aleatorioak 10 minutu eta 20 minutu bitarteko banaketa uniformea duela. Bi konposatu kimikoen artean erreakzioa gertatzeko behar den denbora gutxienez 12 minutukoa bada, kalkula bedi erreakzioa gertatzeko 15 eta 17 minutu bitarteko denbora behar izateko probabilitatea.

4.6.5. ariketa

Aldagai aleatorio batek $\mu = 60$ batez besteko banaketa normala du. Aldagaiak 40,8 baino txikiagoak diren balioak hartzeko probabilitatea 0,209 dela jakinda, kalkula bedi aldagaiaren desbideratze tipikoa.

4.6.6. ariketa

Makina batean pieza bat muntatzeko behar den denbora $\mu = 12{,}8$ minutu eta $\sigma^2 = 2{,}25$ (minutu)^2 parametroetako banaketa normala duen aldagai aleatorioa da.

a. Zein da pieza muntatzen bukatzeko gutxienez 11,5 minutu behar izateko probabilitatea?

b. Kalkula bedi piezaren muntaketa bukatzeko 11 minutu baino gehiago eta 14 minutu baino gutxiago behar izateko probabilitatea.

4.6.7. ariketa

Kalkula bedi β desbideratze tipikodun banaketa esponentzialeko aldagai batek, gutxienez bere batez bestekoa eta gehienez batez bestekoaren bikoitza balio izateko probabilitatea.

4.6.8. ariketa

Biltegi batean, sei motorretik bat injekziogabea da. Biltegiko 200 motor hartu dira aleatorioki.

a. Zein da motor injekziogabeen kopurua 25 eta 35 bitartekoa izateko probabilitatea?

b. Zenbat motor injekziogabe itxaron daiteke egotea?

4.6.9. ariketa

Aireportu batera 20 minuturik behin iristen den hegazkin-kopuruak $\lambda = 100$ parametroko Poisson-en banaketa du. Kalkula bedi aleatorioki aukeratutako 20 minutuko denboratartean, 80 eta 120 hegazkin bitartean iristeko probabilitatea.

4.6.10. ariketa

Biz X = “enpresa batek urtero kontratatzen duen langile-kopurua” aldagai aleatorioaren dentsitate-funtzioa:

$$f(x) = \begin{cases} \dfrac{1}{k^2} x e^{-x/4}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

a. k^2 konstantearen balioa zehaztuz, kalkula bedi enpresak hurrengo urtean gehienez 12 langile kontratatzeko probabilitatea.

b. Zenbat langile itxaron daiteke enpresak kontratatzea datorren urtean?

4.6.11. ariketa

Orduz orduko ozono-mailak jasotzen dituen neurgailu batean, datuetako % 9 baliogabea da. Aleatorioki neurgailuak jasotako 120 datu hartu dira.

a. Zein da datu horien % 85 baino gehiago baliokoak izateko probabilitatea?

b. Eta 120 datu hartu ordez 20 datu hartzen badira, zein da datu horien % 85 baino gehiago baliokoak izateko probabilitatea?

5. Laginketarako sarrera

5.1. OROKORTASUNAK

Laginketak ahalbidetzen du populazioaren eta populaziotik atera daitezkeen laginen arteko erlazioa aztertzea. Inferentziaren bidez laginetatik ondorioztatutako emaitzak populaziora orokor daitezke, konfiantza-maila zehatz batez. Laginketa egokia izateko, laginetatik lortutako emaitzek populazio osoaren adierazgarriak izan behar dute. Laginketaren oinarriak erakusteko asmoz, liburu honetan lagin aleatorio bakunetan oinarritutako laginketa aztertuko da.

Bira N elementuko populazioa eta bertatik atera daitezkeen n elementuetako laginak. Populazioko edozein elementuk laginerako aukeratua izateko probabilitate bera duenean, **lagin aleatorio bakunak** eratzen dira. Adibidez, demagun biltegi batean 200.000 lanpara daudela, eta beraien funtzionamenduaren iraupena aztertzeko 50 lanparako laginak hartuko direla. Lagin bat aukeratu eta beraren batez besteko iraupena kalkulatuko da. Bigarren lagin bat aukeratu eta dagokion batez bestekoa kalkulatuko da... 50 elementuko $C_{200.000,50}$ lagin aleatorio aukera daitezke, eta kasu bakoitzean, batez bestekoak kalkulatuz, $\overline{X}$ aldagai aleatorioak batez bestekoaren banaketa deskribatuko du.

Laginaren edozein funtziori **estatistikoa** deritzo. Adibidez, $\overline{X}$ aldagaia estatistiko bat da. Estatistikoak lagin ezberdinetan hartzen dituen balioak aztertuz, estatistikoaren **lagin-banaketa** adieraz daiteke.

Aztergai diren estatistiko aipagarrienak hurrengoak dira:

1. Laginaren batez bestekoa: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

2. Laginaren bariantza: $s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.

3. Laginaren kuasibariantza: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.

Hurrengo atalean, laginketako zenbait banaketa azalduko dira. Ondorengo ataletan laginketa-teoriako emaitza garrantzitsuenak aurki daitezke. Enuntziatuko diren hipotesietan lagin aleatorio bakunak hartuko dira eta hautaketa guztiak itzulerarik gabe egingo dira, hau da, populazioko elementu bat hautatzen denean ez da berriro populaziora itzuliko.

5.2. LAGINKETA-TEORIAKO ZENBAIT BANAKETA

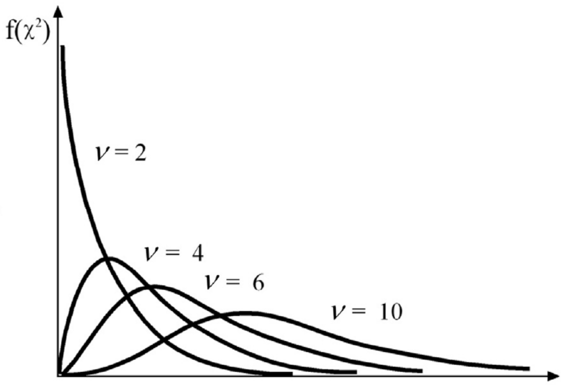
5.2.1. Pearson-en χ^2 banaketa

Bira X_1, X_2, …, X_n aldagai aleatorioak, $N(0, 1)$ banaketadunak eta independenteak. $Y = X_1 + X_2 + \ldots + X_n$ aldagai aleatorioak **$v = n$ askatasun-graduko χ^2 (khi karratu) banaketa** du, laburbilduz χ_n^2 .

Banaketa horri dagokion dentsitate-funtzioa hurrengoa da:

$$f(y) = \begin{cases} 0, & y \le 0 \\ \dfrac{y^{\frac{n}{2}-1}}{n^{\frac{n}{2}}\Gamma(n/2)} exp(-y/2), & y > 0 \end{cases}$$

Dentsitate-funtzioaren adierazpen grafikoa askatasun-graduen araberakoa da.

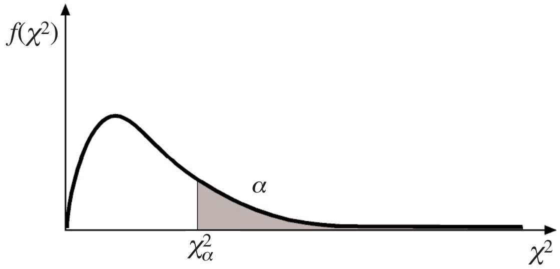


5.1. irudia.

χ_v^2 banaketak honako bereizgarritasunak ditu:

i. [0, ∞) tartean definiturik dago.
ii. Definizio-eremuan jarraitua da.
iii. Ez da simetrikoa.
iv. Askatasun-gradua $n \geq 30$ denean, banaketa normalaren bidez hurbil daiteke.
v. Batez bestekoa $\mu = n$ eta bariantza $\sigma^2 = 2n$ dira.

$P(\chi_v^2 > \chi_\alpha) = \alpha$ probabilitateen balioak kalkulatzeko, liburu honetako taulen eranskina kontsulta daiteke.



5.2. irudia.

5.2.2. Student-en t banaketa

Bira X, X_1, X_2, …, X_n aldagai aleatorioak, $N(0,\sigma)$ banaketadunak eta independenteak.

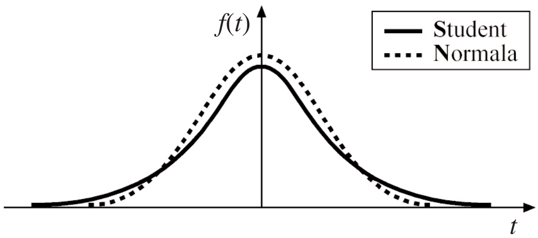
$t = \dfrac{X}{\sqrt{\dfrac{1}{n}\sum_{i=1}^{n} X_i}}$ aldagai aleatorioak $\boldsymbol{v=n}$ **askatasun-graduko Student-en** ***t*** **banaketa** du.

Zenbakitzailea eta zatitzailea desbideratze tipikoaz zatituz, hurrengo adierazpenak ere n askatasun-graduko Student-en t banaketa du:

$$t = \frac{X/\sigma}{\sqrt{\dfrac{1}{n}\sum_{i=1}^{n} X_i \quad \sigma}} = \frac{Z}{\sqrt{\chi_n^2/n}}$$

non Z aldagai aleatorioak banaketa normal tipikoa duen.

Student-en t banaketari dagokion dentsitate-funtzioaren adierazpen grafikoa askatasun-graduen araberakoa da. Gainera, banaketa normal tipikoari dagokion dentsitate-funtzioaren antzeko itxura du.



5.3. irudia.

Hona hemen Student-en t banaketaren propietate batzuk:

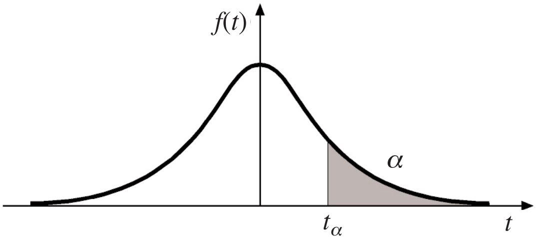
a. $(-\infty, \infty)$ tartean definiturik dago.

b. Batez bestekoarekiko simetrikoa da.

c. Kanpai-forma du eta askatasun-gradua $n \geq 30$ denean, banaketa normal tipikoaren bidez hurbil daiteke.

d. Batez bestekoa $\mu = 0$ eta bariantza $\sigma^2 = \dfrac{n}{n-2} (n > 2)$ dira.

Liburu honetako taulen eranskinean $P\,(t_v > t_\alpha) = \alpha$ balioak adierazi dira.



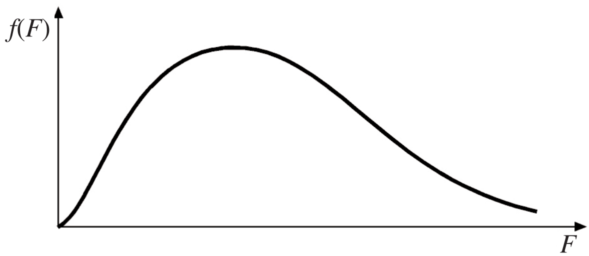
5.4. irudia.

5.2.3. Snedecor-en F banaketa

Bira X eta Y aldagai aleatorio independenteak, $\chi^2_{n_1}$ eta $\chi^2_{n_2}$ banaketadunak, hurrenez hurren. $F = \dfrac{X/n_1}{Y/n_2}$ aldagai aleatorioak $\boldsymbol{v_1 = n_{11}}$**eta** $\boldsymbol{v_2 = n_2}$ **askatasun-gradutako Snedecor-en *F* banaketa** du, non v_1 gaiak zenbakitzailearen askatasun-graduak eta v_2 gaiak zatitzailearen askatasun-graduak adierazten dituzten.

$$F = \frac{X/n_1}{Y/n_2} : F_{n_1, n_2} .$$

F_{v_1, v_2} banaketaren dentsitate-funtzioaren adierazpen grafikoa askatasun-graduen araberakoa da.



5.5. irudia.

Snedecor-en F banaketaren adierazgarri garrantzitsuenak honako hauek dira:

i. $[0, \infty)$ tartean definiturik dago.

ii. Definizio-eremuan jarraitua da.

iii. Ez da simetrikoa.

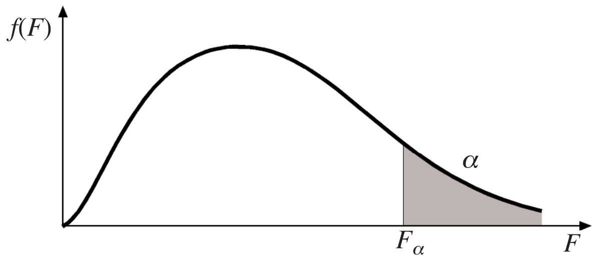
iv. Batez bestekoa $\mu = \dfrac{n_2}{n_2 - 2}(n_2 > 2)$ eta bariantza $\sigma^2 = \dfrac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}$ $(n_2 > 4)$ dira.

$P(F_{v_1,v_2} > F_\alpha) = \alpha$ balioak kalkulatzeko liburuaren bukaeran dauden taulak kontsulta daitezke. Halaber, F banaketaren hurrengo propietatea oso erabilgarria da:

$$F_{1-\alpha;v_1,v_2} = \frac{1}{F_{\alpha;v_2,v_1}}.$$

Propietate horren aplikazio modura, adibidez, $F_{0,95;9,15} = \dfrac{1}{F_{0,05;15,9}} = \dfrac{1}{2,59} = 0,386$ kalkula daiteke.

Liburu honetako eranskineko tauletan $P(F_{v_1,v_2} > F_\alpha) = \alpha$ balioak adierazi dira, $\alpha = 0,05$ eta $\alpha = 0,10$ izanik.



5.6. irudia.

5.3. BATEZ BESTEKOAREN LAGIN-BANAKETA

5.3.1. σ^2 ezaguna denean

1 teorema. Biz μ batez bestekoa eta σ^2 bariantza ezaguneko populazio normala eta populaziotik ateratako n tamainako edozein lagin aleatorio bakun. Orduan, $\overline{X}$ aldagai aleatorioak $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$ banaketa du.

2 teorema. Demagun σ^2 bariantza ezaguneko populaziotik $n \geq 30$ tamainako lagin aleatorio bakuna hartu dela. Orduan, $\overline{X}$ aldagai aleatorioaren banaketa $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$ banaketaz hurbil daiteke. Teorema honi **limite zentralaren teorema** deritzo.

5.3.2. σ^2 ezezaguna denean

3 teorema. Demagun desbideratze tipiko ezezaguneko populazio normal batetik $n \geq 30$ tamainako lagin aleatorio bakuna hartzen dela. Orduan, $\overline{X}$ aldagai aleatorioak $N\left(\mu, \frac{S}{\sqrt{n}}\right)$ banaketa du, non $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ laginaren kuasibariantza den.

4 teorema. Demagun desbideratze tipiko ezezaguneko populazio normal batetik $n < 30$ tamainako lagin aleatorio bakuna hartzen dela. Orduan,

$$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

aldagai aleatorio jarraituak $\nu = n-1$ askatasun-graduko Student-en t banaketa du, non S^2 laginaren kuasibariantza den.

5.4. BI LAGINEN BATEZ BESTEKOEN ARTEKO KENDURAREN BANAKETA

5.4.1. Populazioen bariantzak ezagunak direnean

5.4.1.1. Populazioak banaketa normalekoak eta independenteak

5 teorema. Bira X_1: N(μ_1,σ_1) eta X_2: N(μ_2,σ_2) aldagai aleatorio independenteak, zeintzuen σ_1^2 eta σ_2^2 bariantzak ezagunak diren. Populazio bakoitzetik n_1 eta n_2 tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren. Orduan, $\overline{X}_1 - \overline{X}_2$ aldagaiak $N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$ banaketa du.

5.4.1.2. Populazioak banaketa ez normalekoak eta independenteak ($n_1 \geq 30$, $n_2 \geq 30$)

6 teorema. Bira X_1 eta X_2 aldagai aleatorio independenteak, zeintzuen σ_1^2 eta σ_2^2 bariantzak ezagunak diren. Populazio bakoitzetik $n_1 \geq 30$ eta $n_2 \geq 30$ tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren. Orduan, $\overline{X}_1 - \overline{X}_2$ aldagaia $N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$ banaketaren bidez hurbil daiteke.

5.4.2. Populazioen bariantzak ezezagunak direnean

5.4.2.1 Populazioak banaketa normalekoak eta independenteak ($n_1 \geq 30$, $n_2 \geq 30$)

7 teorema. Bira X_1: N(μ_1,σ_1) eta X_2: N(μ_2,σ_2) aldagai aleatorio independenteak, bariantza ezezagunekoak. Populazio bakoitzetik $n_1 \geq 30$ eta $n_2 \geq 30$ tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren. Orduan, $\overline{X}_1 - \overline{X}_2$ aldagaia

$N\left(\mu_1-\mu_2,\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)$ banaketaren bidez hurbil daiteke, non S_1^2 eta S_2^2 bi populazioetatik hartutako lagin aleatorio bakunen kuasibariantzak diren.

5.4.2.2. Populazioak banaketa normalekoak, independenteak eta bariantza berdinekoak ($n_1 < 30$, $n_2 < 30$)

8 teorema. Bira X_1: N(μ_1,σ_1) eta X_2: N(μ_2,σ_2) aldagai aleatorio independenteak, bariantzak ezezagunak eta berdinak $(\sigma_1^2=\sigma_2^2)$ izanik. Populazio bakoitzetik $n_1 < 30$ eta $n_2 < 30$ tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren. Orduan,

$$t_{n_1+n_2-2}=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$ aldagai aleatorioak

$\nu = n_1 + n_2 - 2$ askatasun-graduko Student-en banaketa du.

5.4.2.3. Populazioak banaketa normalekoak, independenteak eta bariantza ezberdinekoak ($n_1 < 30$, $n_2 < 30$)

9 teorema. Bira X_1: N(μ_1,σ_1) eta X_2: N(μ_2,σ_2) aldagai aleatorio independenteak, bariantza ezezagunak ezberdinak izanik. Populazio bakoitzetik $n_1 < 30$ eta $n_2 < 30$ tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren. Orduan,

$$t_v=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$ aldagai aleatorioak

$$v=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1+1}+\frac{\left(S_2^2/n_2\right)^2}{n_2+1}}-2$$ askatasun-graduko Student-en banaketa du.

5.5. LAGINAREN PROPORTZIOAREN BANAKETA

10 teorema. Bira populazio binomialeko p proportzioa eta populaziotik ateratako n tamainako lagin aleatorio bakuna, $np > 5$ eta $nq > 5$ izanik. Laginaren proportzioaren $\hat{p}$ aldagai aleatorioa $N\left(p,\sqrt{\frac{pq}{n}}\right)$ banaketaren bidez hurbil daiteke.

5.6. BI LAGINEN PROPORTZIOEN ARTEKO KENDURAREN BANAKETA

11 teorema. Bira bi populazio binomialetako p_1 eta p_2 proportzioak. Demagun populazio bakoitzetik n_1 eta n_2 tamainetako lagin aleatorio bakunak aukeratzen direla, hurrenez hurren, non $n_i p_i > 5$ eta $n_i q_i > 5$ $(i = 1,2)$ diren. $\hat{p}_1 - \hat{p}_2$ laginen proportzioen arteko kendura $N\left(p_1 - p_2, \sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}\right)$ banaketaren bidez hurbil daiteke.

5.7. BARIANTZAREN LAGIN-BANAKETA

12 teorema. Bira σ^2 bariantzako populazio normala eta n tamainako lagin aleatorio bakuna. Orduan,

$$\chi^2_{n-1} = \frac{ns^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \text{ aldagai aleatorio jarraituak}$$

$\nu = n - 1$ askatasun-graduko Pearson-en χ^2 banaketa du.

μ ezaguna bada, $\dfrac{\sum_{i=1}^{n}(x_i - \mu)^2}{\sigma^2}$ aldagai aleatorioak χ^2_n banaketa du.

5.8. KUASIBARIANTZEN ARTEKO ZATIDURAREN LAGIN-BANAKETA

13 teorema. Bira σ_1^2 eta σ_2^2 bariantzetako bi populazio normal independente. Lehenengo populaziotik n_1 tamainako lagin aleatorio bakuna eta bigarren populaziotik n_2 tamainako lagin aleatorio bakuna ateratzen badira, orduan

$$F_{n_1-1,n_2-1} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \text{ aldagai aleatorio jarraituak}$$

$\nu_1 = n_1 - 1$ eta $\nu_2 = n_2 - 1$ askatasun-graduko Snedecor-en F banaketa du.

5.9. ARIKETA EBATZIAK

5.9.1. ariketa

X aldagai aleatorioaren batez bestekoa 4 eta desbideratze tipikoa 12 dira. Populaziotik 41 elementuko lagin aleatorio bakuna hartu da.

a. Zein da laginaren batez bestekoak (0,8) tartean balioak hartzeko probabilitatea?

b. Kalkula bedi laginaren batez bestekoa gutxienez sei izateko probabilitatea.

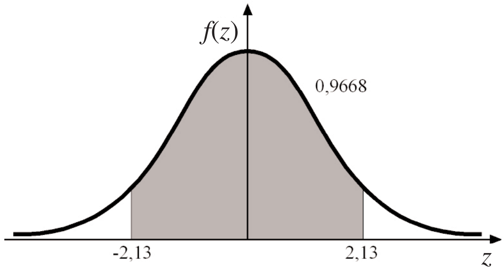
Ebazpena

Bira X aldagai aleatorioa, $\mu = 4$, $\sigma = 12$ eta $n = 41$ tamainako lagin aleatorio bakuna.

a. Limite zentralaren teoremaren ondorioz, $\overline{X}$ aldagai aleatorioa $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ banaketaren bidez hurbil daiteke. Orduan, eskatutako probabilitatearen balioa honela lor daiteke:

$$P(0 < \overline{X} < 8) = P\left(\frac{0-4}{12/\sqrt{41}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{8-4}{12/\sqrt{41}}\right) = P(-2{,}13 < Z < 2{,}13) =$$

$$= 2\phi(2{,}13) - 1 = 2 \cdot 0{,}9834 - 1 = 0{,}9668.$$



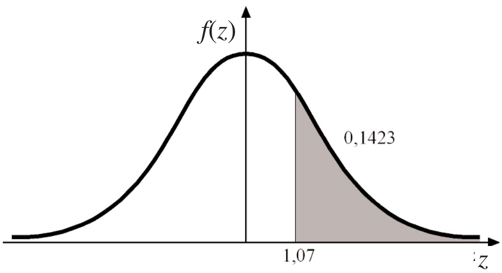
5.7. irudia.

b. Era berean,

$$P(\overline{X} \geq 6) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{6-4}{12/\sqrt{41}}\right) = P(Z \geq 1{,}07) = 1 - \phi(1{,}07) =$$

$$= 1 - 0{,}8577 = 0{,}1423$$

balioa ondorioztatuko da.



5.8. irudia.

5.9.2. ariketa

Autoak konpontzen dituzten tailer batean, konponketen iraupenak egun bateko batez bestekodun banaketa normalari jarraitzen dio. Bertan konpondu diren zazpi auto aleatorioki aukeratu dira, eta konponduak izateko denboraren desbideratze tipikoa 0,4 egunekoa izan da.

a. Zenbatekoa da auto horiek konponduak izan daitezen batez besteko denbora gehienez egun erdikoa izateko probabilitatea?

b. Kalkula bedi aurreko ataleko probabilitatea, jakinda, oro har, autoak konponduak izateko denboraren desbideratze tipikoa egun batekoa dela.

Ebazpena

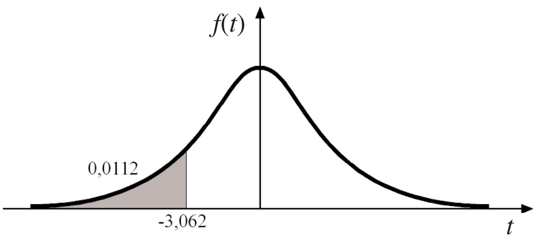
Biz X = "autoak konponduak izateko behar den denbora" aldagai aleatorioa, $N(\mu = 1, \sigma)$ banaketakoa.

$n = 7$ tamainako eta $s = 0{,}4$ desbideratze tipikodun lagin aleatorio bakun bat hartu da.

a. Aldagai aleatorioa banaketa normalekoa, σ ezezaguna eta $n < 30$ direnez, 4 teoremaren arabera $\dfrac{\overline{X} - \mu}{s/\sqrt{n-1}}$ aldagaiak $n - 1$ askatasun-graduko Student-en banaketa du. Ondorioz, hurrengo emaitza lor daiteke:

$$P(\overline{X} \le 0{,}5) = P\left(\frac{\overline{X} - \mu}{s/\sqrt{n-1}} \le \frac{0{,}5 - 1}{0{,}4/\sqrt{6}}\right) =$$

$$= P(t_6 \le -3{,}062) = P(t_6 \ge 3{,}062) = 0{,}012.$$



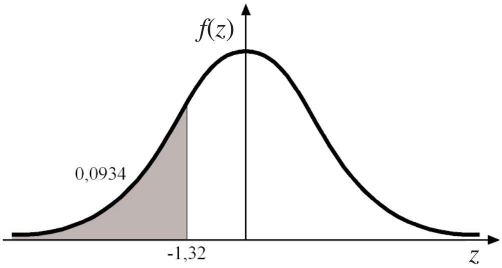
5.9. irudia.

Azken probabilitatearen balioa lortzeko, Student-en taulan interpolazioa aplikatu da.

b. Oraingoan datu berri bat dago, $\sigma = 1$ alegia. Orduan, 1 teorema aplikatu behar da.

$$P(\overline{X} \le 0{,}5) = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \le \frac{0{,}5 - 1}{1/\sqrt{7}}\right) =$$

$$= P(Z \le -1{,}32) = 1 - \phi(1{,}32) = 1 - 0{,}9066 = 0{,}0934.$$



5.10. irudia.

5.9.3. ariketa

Lan zehatz bat burutzeko emakumeek behar duten batez besteko denbora 460 ordukoa da. Gizonek, aldiz, lan berdina burutzeko batez beste 480 ordu behar dituzte. Emakumeek lana burutzeko erabiltzen duten denborak $N(\mu_e,\sigma_e)$ banaketa du eta gizonek lan berdina burutzeko erabiltzen duten denbora $N(\mu_g,\sigma_g)$ banaketakoa da. Aleatorioki 100 emakume eta 125 gizon hartu dira, lana burutzeko emakumeek 52 orduko desbideratze tipikoa izan dute eta gizonek 60 ordukoa.

Laginei dagokienez, kalkula bedi gizonek lana burutzeko behar duten batez besteko denbora emakumeek lana burutzeko behar duten batez besteko denbora baino hamar ordu gehiagokoa izateko probabilitatea.

Ebazpena

Bira X_e = "emakumeek behar duten ordu-kopurua lana burutzeko" eta X_g = "gizonezkoek behar duten ordu-kopurua lana burutzeko" aldagai aleatorioak banaketa normalekoak, X_e: $N(\mu_e = 460,\sigma_e)$ eta X_g: $N(\mu_g = 480,\sigma_g)$ hain zuzen ere. X_e eta X_g independenteak direla onar daiteke. Bestalde, lagineko datuak emakumeen kasuan $n_e = 100$ emakume eta $s_e = 52$ h dira. Gizonezkoen kasuan, lagineko datuak $n_g = 125$ gizon eta $s_g = 60$ h dira.

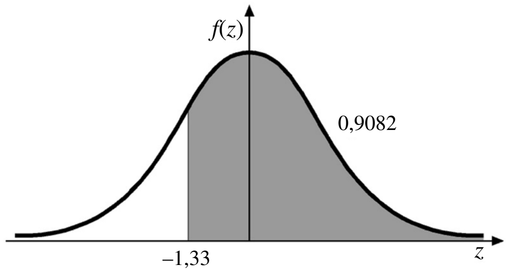
7 teoremaren emaitza kontuan hartuz, eskatutako probabilitatearen balioa kalkulatuko da.

$$P(\bar{x}_g > \bar{x}_e + 10) = P(\bar{x}_g - \bar{x}_e > 10) = P\left(Z > \frac{10 - (\mu_g - \mu_e)}{\sqrt{\frac{s_g^2}{n_g - 1} + \frac{s_e^2}{n_e - 1}}} \right) =$$

$$= P\left(Z > \frac{10 - (480 - 460)}{\sqrt{\frac{60^2}{124} + \frac{52^2}{99}}} \right) = P(Z > -1,33) =$$

$$= P(Z < 1,33) = \phi(1,33) = 0,9082.$$

non $Z = \dfrac{(\bar{x}_g - \bar{x}_e) - (\mu_g - \mu_e)}{\sqrt{\dfrac{s_g^2}{n_g - 1} + \dfrac{s_e^2}{n_e - 1}}}$ banaketa normal tipikodun aldagaia den.



5.11. irudia.

5.9.4 ariketa

Hotel batek hilero jasotzen duen bezero-kopuruak banaketa normalari jarraitzen dio, hileko bezero-kopuruaren desbideratze tipikoa 30 bezerokoa delarik. Era aleatorioan, bost hilabetetan jasotako bezeroak hartu dira.

a. Zein da bost hilabetetako edozein denboralditan jasotako bezero-kopuruaren kuasibariantza gutxienez 64 (bezero)^2-koa izateko probabilitatea?

b. Zein da laginaren kuasibariantza populazioaren bariantzaren gutxienez bikoitza izateko probabilitatea?

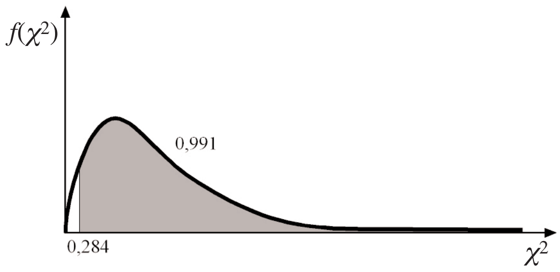
Ebazpena

Biz X = “hileko bezero-kopurua”, X: $N(\mu, \sigma = 30)$.

a. $P(S^2 \geq 64)$ balioa lortzeko 12 teorema aplikatuko da, zeinaren arabera $\dfrac{(n-1)S^2}{\sigma^2}$ aldagai aleatorio jarraituak $\nu = n - 1$ askatasun-graduko Pearson-en khi karratuaren banaketa duen.

$$P(S^2 \geq 64) = P\left(\chi_4^2 \geq \frac{(n-1)}{\sigma^2} 64\right) =$$

$$= P\left(\chi_4^2 \geq \frac{(5-1)}{30^2} 64\right) = P\left(\chi_4^2 \geq 0{,}284\right) = 0{,}991.$$

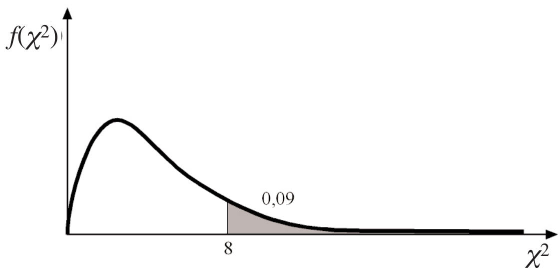


5.12. irudia.

Azken probabilitatearen balioa kalkulatzeko χ^2 banaketaren taulan interpolazioa aplikatu da.

b. Era berean, $P(S^2 \geq 2\sigma^2)$ balioa kalkulatzeko, 12 teorema erabiliko da.

$$P(S^2 \geq 2\sigma^2) = P(\frac{S^2}{\sigma^2} \geq 2) = P\left(\chi_4^2 \geq 2(n-1)\right) = P\left(\chi_4^2 \geq 8\right) = 0,09.$$



5.13. irudia.

5.9.5. ariketa

Azken azterketen arabera, gaur egun populazioaren % 90ek sakelako telefonoa erabiltzen du. Aleatorioki 200 pertsonako lagin aleatorio bakuna hartu da.

a. Kalkula bedi 200 pertsona horietatik sakelako telefonoa erabiltzen dutenen proportzioa gutxienez 0,85 izateko probabilitatea.

b. Zein da sakelako telefonoen erabiltzaileen benetako proportzioaren eta lagineko erabiltzaileen proportzioaren arteko diferentzia gehienez % 2 izateko probabilitatea?

Ebazpena

Bira sakelako telefonoen erabiltzaileen proportzioa p = 0,90 eta n = 200 tamainako lagin aleatorio bakuna.

a. Lagineko sakelako telefonoen erabiltzaileen proportzioa $\hat{p}$ gutxienez 0,85 izateko probabilitatea kalkulatzeko 10 teorema aplikatuko da, non $np = 200 \cdot 0{,}90 = 180 > 5$ eta $nq = 200 \cdot 0{,}10 = 20 > 5$ ezberdintzak betetzen diren.

$$P(\hat{p} \geq 0{,}85) = P\left(\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \geq \frac{0{,}85-0{,}90}{\sqrt{\frac{0{,}90\cdot 0{,}10}{200}}}\right) = P(Z \geq -2{,}38) =$$

$$= P(Z < 2{,}38) = \phi(2{,}38) = 0{,}9913.$$

b. Era berean $P\left(|p-\hat{p}| \leq 0{,}02\right)$ probabilitatearen balioa kalkulatzeko 10 teorema aplikatuko da, non $Z = \frac{\hat{p}-p}{\sqrt{pq/n}}$ banaketa normal tipikodun aldagaia den.

$$P\left(|p-\hat{p}| \leq 0{,}02\right) = P\left(|Z| \leq \frac{0{,}02}{\sqrt{\frac{0{,}90\cdot 0{,}10}{200}}}\right) = P\left(|Z| \leq 0{,}95\right) =$$

$$= 2 \cdot \phi(0{,}95) - 1 = 2 \cdot 0{,}8289 - 1 = 0{,}6578.$$

5.9.6. ariketa

Bideoak ekoizten dituen enpresa batek bi mikrozirkuitu-mota diseinatzen ditu. Bi mikrozirkuitu-mota horien fluxu-kantitateak banaketa normalekoak dira, desbideratze tipikoak 3 Wb eta 6 Wb izanik, hurrenez hurren. Bi mikrozirkuituen fluxu-kantitateak elkarrekiko independenteak direla suposatuz, I. motako mikrozirkuituaren hamasei unitate eta II. motako mikrozirkuituaren beste hamasei unitate hartu dira.

Aipatutako azken mikrozirkuituetarako, kalkula bedi I. motako mikrozirkuituen bariantzaren eta II. motako mikrozirkuituen bariantzaren arteko proportzioa gehienez 0,6 izateko probabilitatea.

- -

Ebazpena

Bira X_1 = "I mikrozirkuituaren fluxu-kantitatea" eta X_2 = "II mikrozirkuituaren fluxu-kantitatea", X_1: $N(\mu_1, \sigma_1 = 3)$ eta X_2: $N(\mu_2, \sigma_2 = 6)$ aldagai aleatorio independenteak. Populazio bakoitzetik $n_1 = 16$ eta $n_2 = 16$ tamainetako lagin aleatorio bakunak hartu dira, hurrenez hurren.

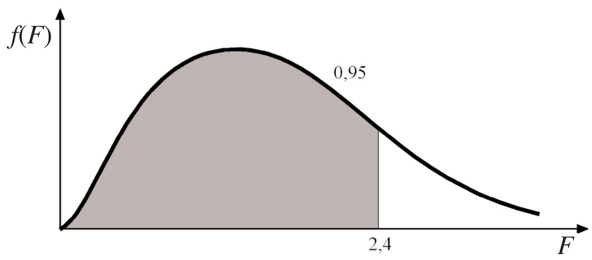
$P\left(\frac{s_1^2}{s_2^2} \leq 0{,}6\right)$ kalkulatu behar da.

$$\text{Jakinda } \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = F_{n_1-1,n_2-1} \text{ eta } \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} = \frac{\frac{s_1^2 n_1}{(n_1-1)}\sigma_2^2}{\frac{s_2^2 n_2}{(n_2-1)}\sigma_1^2} = \frac{s_1^2 n_1(n_2-1)\sigma_2^2}{s_2^2 n_2(n_1-1)\sigma_1^2}$$

orduan, eskatutako probabilitatearen balioa kalkulatzeko 13 teorema aplikatuko da.

$$P\left(\frac{s_1^2}{s_2^2} \le 0,6\right) = P\left(\frac{s_1^2 n_1(n_2-1)\sigma_2^2}{s_2^2 n_2(n_1-1)\sigma_1^2} \le 0,6\cdot\frac{n_1(n_2-1)\sigma_2^2}{n_2(n_1-1)\sigma_1^2}\right) = P\left(F_{15,15} \le 0,6\cdot\frac{16\cdot15\cdot6^2}{16\cdot15\cdot3^2}\right) =$$

$$= P(F_{15,15} \le 2,4) = 0,95.$$



5.14. irudia.

5.9.7 ariketa

Lasterketa bateko pertsonen puntuazioek banaketa normala dute, desbideratze tipikoa 1,12 delarik. Lasterketako lagin aleatorio bakun bat hartu eta bere batez bestekoa kalkulatu da. Laginaren batez bestekoaren eta populazioaren batez bestekoaren arteko diferentzia gehienez puntu-erdikoa izateko probabilitatea 0,98 izateko, zein izan behar da laginaren tamaina?

Ebazpena

Enuntziatuko datuetatik, hurrengoa ondorioztatzen da:

$$P\left(\left|\overline{X} - \mu\right| \le 0,5\right) = 0,98.$$

Populazioa banaketa normalekoa eta desbideratze tipikoa ezaguna denez, orduan $\overline{X}$ aldagaiak $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ banaketa du, $\sigma = 1,12$ delarik. Beraz, hurrengoa ondoriozta daiteke:

$$0,98 = P\left(\left|\overline{X} - \mu\right| \le 0,5\right) = P\left(-0,5 \le \overline{X} - \mu < 0,5\right) = P\left(\frac{-0,5}{\sigma/\sqrt{n}} \le \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \le \frac{0,5}{\sigma/\sqrt{n}}\right)$$

$Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ aldagaiaren banaketa $N(0, 1)$ dela kontuan hartuko da.

$$0,98 = 2\phi\left(\frac{0,5\sqrt{n}}{\sigma}\right) - 1$$

$$\phi\left(\frac{0,5\sqrt{n}}{\sigma}\right) = 0,99.$$

Banaketa normal estandarreko banaketa-funtzioaren taulan kontsultatuz, honako balio hau lortzen da:

$$\frac{0,5\sqrt{n}}{\sigma} = 2,33.$$

Eta hortik, hartu behar den laginaren tamaina zehatz daiteke.

$$n = \left(\frac{2,33}{0,5} \cdot 1,12\right)^2 \approx 27.$$

5.9.8. ariketa

Hiri bateko airearen kalitatea neurtzeko sareko bi estaziotan sufre dioxidoaren hileko batez bestekoak aztergai dira. Era aleatorioan A estaziotik 10 hilabetetako batez besteko sufre dioxido mailak eta B estaziotik 12 hilabetetako batez besteko sufre dioxido mailak hartu dira. A estazioko 10 datuen desbideratze tipikoa 2 $\mu g/m^3$ eta B estazioko 12 datuen desbideratze tipikoa 1 $\mu g/m^3$ dira. Sufre dioxidoaren mailak A estazioan $N(15, \sigma_1)$ banaketa du eta B estazioan $N(13, \sigma_2)$ banaketa du, bariantza ezezagunak eta ezberdinak direla onartuz. Demagun bi aldagaiak elkarrekiko independenteak direla.

Kalkula bedi A estazioko 10 datuz eta B estazioko 12 datuz osaturiko laginak hartzerakoan, A estazioko batez besteko sufre dioxido mailak B estazioko batez besteko sufre dioxido maila 1 $\mu g/m^3$ baino gehiagotan gainditzeko probabilitatea.

Ebazpena

Bira X_1= “A estazioko hileko batez besteko sufre dioxido maila” eta X_2 = “B estazioko hileko batez besteko sufre dioxido maila” aldagai aleatorio independenteak.

X_1: $N(\mu_1 = 15, \sigma_1)$ eta X_2: $N(\mu_2 = 13, \sigma_2)$ dira.

A estazioan: $s_1 = 2\ \mu g/m^3$, $n_1 = 10$.

B estazioan: $s_2 = 1\ \mu g/m^3$, $n_2 = 12$.

$P(\overline{X}_1 > \overline{X}_2 + 1)$ probabilitatearen balioa kalkulatu behar da. 9 teoremako emaitza kontuan hartuko da, zeinaren arabera enuntziatuko baldintzetan

$$t_v = \frac{\left(\overline{X}_1 > \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}} \text{ aldagai aleatorioak}$$

$$v = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(S_1^2 / n_1\right)^2}{n_1 + 1} + \dfrac{\left(S_2^2 / n_2\right)^2}{n_2 + 1}} - 2 \text{ askatasun-graduko Student-en banaketa du.}$$

Kasu honetan, askatasun-graduak zehazteko hurrengo adierazpen baliokidea erabiliko da, non s_1^2 eta s_2^2 lagin-bariantzak diren:

$$v = \frac{\left(\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1 - 1}\right)^2}{n_1 + 1} + \dfrac{\left(\dfrac{s_2^2}{n_2 - 1}\right)^2}{n_2 + 1}} - 2 = \frac{\left(\dfrac{2^2}{10 - 1} + \dfrac{1^2}{12 - 1}\right)^2}{\dfrac{\left(\dfrac{2^2}{10 - 1}\right)^2}{10 + 1} + \dfrac{\left(\dfrac{1^2}{12 - 1}\right)^2}{12 + 1}} - 2 \approx 13.$$

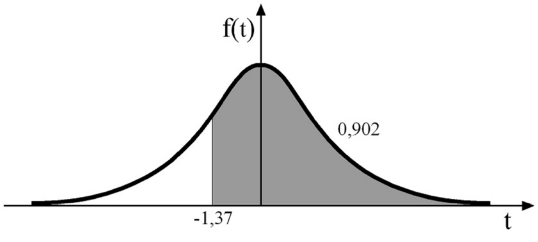
Horrela, eskatutako probabilitatearen balioa hurrengo hau da:

$$P(\overline{X}_1 > \overline{X}_2 + 1) = P(\overline{X}_1 - \overline{X}_2 > 1) =$$

$$= P\left(\frac{\left(\overline{X}_1 > \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}} > \frac{1 - (15 - 13)}{\sqrt{\dfrac{2^2}{10 - 1} + \dfrac{1^2}{12 - 1}}}\right) =$$

$$= P(t_{13} > -1{,}37) = P(t_{13} < 1{,}37) = 1 - 0{,}098 = 0{,}902.$$

Probabilitate hau kalkulatzeko Student-en taulako balioak hartu eta interpolazioa aplikatu da.



5.15. irudia.

5.9.9. ariketa

Duela bi urtetik gaur egunera produktu zehatz baten kontsumitzaileen kopurua % 60tik % 85era igo dela onartu da. Onarpen hori ontzat hartzeko, duela bi urte aleatorioki 250 pertsonari inkesta bat egin zitzaien eta gaur egun inkesta bera 400 pertsonari egin zaie.

a. Zein da gaur egungo kontsumitzaileen laginaren proportzioa duela bi urteko kontsumitzaileen laginaren proportzioa baino % 20 baino handiagoa izateko probabilitatea?

b. Kalkula bedi duela bi urteko laginaren proportzioaren eta gaur egungo laginaren proportzioaren arteko diferentzia gutxienez % 20koa eta gehienez % 30ekoa izateko probabilitatea.

Ebazpena

Bira X_1 = "produktuaren kontsumitzaile-kopurua duela bi urte" eta X_2 = "produktuaren kontsumitzaile-kopurua gaur egun" aldagai aleatorioak.

Problemako datuetatik abiatuz, kontsumitzaileen proportzioa duela bi urte $p_1 = 0{,}60$ zen eta gaur egun kontsumitzaileen proportzioa $p_2 = 0{,}85$ da.

a. Gaur egungo kontsumitzaileen laginaren proportzioa duela bi urteko laginarena baino % 20 handiagoa izateko probabilitatearen balioa kalkulatzeko 11 teorema aplikatuko da, non $n_1p_1 = 250 \cdot 0{,}60 = 150 > 5$, $n_1q_1 = 250 \cdot 0{,}40 = 100 > 5$, $n_2p_2 = 400 \cdot 0{,}85 = 340 > 5$ eta $n_2q_2 = 400 \cdot 0{,}15 = 60 > 5$ ezberdintzak betetzen diren eta $Z = \dfrac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)}{\sqrt{\dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}}}$ den.

$$P(\hat{p}_2 > \hat{p}_1 + 0{,}20) = P(\hat{p}_2 - \hat{p}_1 > 0{,}20) =$$

$$= P\left(\frac{\hat{p}_2 - \hat{p}_1 - (p_2 - p_1)}{\sqrt{\dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}}} > \frac{0{,}20 - (0{,}85 - 0{,}60)}{\sqrt{\dfrac{0{,}60 \cdot 0{,}40}{250} + \dfrac{0{,}85 \cdot 0{,}15}{400}}} \right) =$$

$$= P(Z > -1{,}39) = \phi(1{,}39) = 0{,}9177.$$

b. Era berean, bigarren ataleko probabilitatearen balioa kalkulatzeko 11 teorema aplikatuko da.

$$P(0{,}20 \le |\hat{p}_2 - \hat{p}_1| \le 0{,}30) =$$

$$= P\left(\frac{0,20 - (0,85 - 0,60)}{\sqrt{\frac{0,60 \cdot 0,4}{250} + \frac{0,85 \cdot 0,15}{400}}} \le Z \le \frac{0,30 - (0,85 - 0,60)}{\sqrt{\frac{0,60 \cdot 0,4}{250} + \frac{0,85 \cdot 0,15}{400}}} \right) =$$

$$= P(-1,39 \le Z \le 1,39) = 2\phi(1,39) - 1 = 2 \cdot 0,9177 - 1 = 0,8354.$$

5.9.10. ariketa

Lantegi batean piezak ekoizteko A eta B prozedurak erabiltzen dira. Demagun prozedura horiek aztertzeko eguneko ekoitzitako pieza-kopurua zenbatu dela. Bi prozeduren bidez ekoitzitako piezak neurtzen dituzten aldagai aleatorioak, banaketa normalekoak eta elkarrekiko independenteak dira. A prozeduraz batez beste 480 pieza eta B prozeduraz batez beste 460 pieza ekoizten dira, bariantzak ezezagunak baina berdinak direlarik.

a. Era aleatorioan, A prozeduraz ekoitzitako 21 egunetako piezak eta B prozeduraz ekoiztutako 21 egunetako piezak hartu dira, laginen kuasibariantzak 30 eta 25 pieza karratu izanik, hurrenez hurren. Datu hauetan oinarrituz, kalkula bedi $P\left(\overline{X}_1 \le \overline{X}_2 + 16\right)$ probabilitatea.

b. A prozeduraz ekoitzitako beste 21 egunetako piezak eta B prozeduraz ekoitzitako beste 16 egunetako piezak hartu dira. Kalkula bedi $P\left(\frac{s_1^2}{s_2^2} \le k\right) = 0,99$ berdintza betetzeko k konstantearen balioa.

Ebazpena

Bira X_1 = "A prozeduraz ekoitzitako eguneko pieza-kopurua" eta X_2 = "B prozeduraz ekoitzitako eguneko pieza-kopurua" aldagai aleatorioak, banaketa normalekoak eta independenteak, X_1: $N(480,\sigma)$ eta X_2: $N(460,\sigma)$ izanik. Kasu honetan bi populazioek desbideratze tipiko bera dute, σ alegia:

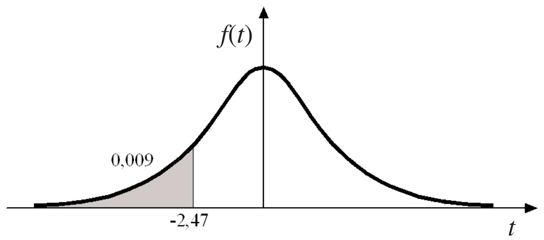
a. $P(\overline{X}_1 \le \overline{X}_2 + 16)$ probabilitatearen balioa kalkulatzeko 8 teoremaren emaitza kontuan hartuko da, zeinaren arabera bi populazio normal independenteren kasuan, bariantzak ezezagunak eta berdinak $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$ eta n_1<30 eta n_2 < 30 tamainetako lagin aleatorio bakunak hartzen direnean, orduan

$$t_{n_1+n_2-2} = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ aldagai aleatorioak}$$

$\nu = n_1 + n_2 - 2$ askatasun-graduko Student-en banaketa du, non $n_1 = 21 = n_2$, $S_1^2 = 30$ eta $S_2^2 = 25$ diren.

$$P\left(\overline{X}_1 \le \overline{X}_2 + 16\right) = P\left(\overline{X}_1 - \overline{X}_2 \le 16\right) = P\left(t_{40} \le \frac{16 - (480 - 460)}{\sqrt{\dfrac{20 \cdot 30 + 20 \cdot 25}{21 + 21 - 2}}\sqrt{\dfrac{1}{21} + \dfrac{1}{21}}} \right) =$$

$$= P(t_{40} \le -2{,}47) = P(t_{40} \ge 2{,}47) = 0{,}009.$$



5.16. irudia.

b. Konstantearen balioa zehazteko, 13 teorema aplikatuko da, non $\dfrac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} = F_{n_1 - 1, n_2 - 1}$ den. Kasu honetan bi populazioen bariantzak berdinak direnez, $F_{n_1 - 1, n_2 - 1} = \dfrac{S_1^2}{S_2^2}$ da. Gainera, bariantzaren eta kuasibariantzaren arteko erlazioa kontuan hartuko da,

$$F_{n_1 - 1, n_2 - 1} = \frac{S_1^2}{S_2^2} = \frac{\dfrac{s_1^2 n_1}{(n_1 - 1)}}{\dfrac{s_2^2 n_2}{(n_2 - 1)}} = \frac{s_1^2 n_1 (n_2 - 1)}{s_2^2 n_2 (n_1 - 1)} \text{ delarik.}$$

$$0{,}99 = P\left(\frac{s_1^2}{s_2^2} \le k \right) = P\left(F_{20,15} \le k \cdot \frac{n_1 (n_2 - 1)}{n_2 (n_1 - 1)} \right) = P\left(F_{20,15} \le k \cdot \frac{21 \cdot (16 - 1)}{16 \cdot (21 - 1)} \right) =$$

$$= P(F_{20,15} \le k \cdot 0{,}984).$$

Ondorioz,

$$0{,}01 = P(F_{20,15} > k \cdot 0{,}984)$$

beteko da. Snedecor-en taulan kontsultatuz, k konstantearen balioa zehatz daiteke:

$$k \cdot 0{,}984 = 3{,}37 \quad \to$$

$$k \approx 3{,}42.$$

5.9.11. ariketa

Laborategi bateko materialaren % 40 birziklatzera eramaten da.

a. Aleatorioki, laborategiko materialaren 88 unitate hartu dira. Zein da horien gutxienez laurdena eta gehienez erdia birziklatzera eramateko probabilitatea?

b. Zehatz bedi lagin aleatorio bakunaren m tamaina, birziklatzera eramateko materialaren lagineko proportzioa gutxienez % 20 izateko probabilitatea 0,95 dela jakinik.

Ebazpena

Birziklatzera eramaten den laborategiko materialaren proportzioa p = 0,40 da. Biz $\hat{p}$ gaia birziklatzera eramaten den laborategiko materialaren lagineko proportzioa.

a. Eskatutako probabilitatea kalkulatzeko 10 teorema aplikatuko da, non n = 88 lagin aleatorio bakunaren tamaina, $np = 88 \cdot 0{,}40 = 35{,}2 > 5$ eta $nq = 88 \cdot 0{,}60 = 52{,}8 > 5$ diren.

$$P(0,25 \leq \hat{p} \leq 0,50) = P\left(\frac{0,25 - 0,40}{\sqrt{\dfrac{0,40 \cdot 0,60}{88}}} \leq \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} \leq \frac{0,50 - 0,40}{\sqrt{\dfrac{0,40 \cdot 0,60}{88}}} \right) =$$

$$= P(-2{,}87 \leq Z \leq 1{,}91) = \phi(1{,}91) - \phi(-2{,}87) =$$

$$= \phi(1{,}91) - [1 - \phi(2{,}87)] = 0{,}9719 - 1 + 0{,}9979 = 0{,}9698.$$

b. Era berean, 10 teorema aplikatuz,

$$0,95 = P(\hat{p} \geq 0,20) = P\left(\frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}} \geq \frac{0,20 - 0,40}{\sqrt{\dfrac{0,40 \cdot 0,60}{m}}} \right) =$$

$$= P(Z \geq -0,41\sqrt{m}) = P(Z \leq 0,41\sqrt{m}).$$

Liburu honetako taulen eranskinean banaketa normalari dagokion taulan kontsultatuz,

$$0,95 = \phi(0,41\sqrt{m}) \quad \rightarrow$$

$$0,41\sqrt{m} = 1,645$$

$$m = \left(\frac{1,645}{0,41} \right)^2 \approx 16$$

lagin aleatorio bakunaren tamaina zehaztuko da.

5.9.12. ariketa

Altzairuzko xaflen gogortasuna neurtzen duen koefizienteak $N(1{,}6,\ 0{,}3)$ banaketa du. Aleatorioki, altzairuzko bost xafla hartu dira.

a. Kalkula bedi aleatorioki hartutako altzairuzko bost xaflen batez besteko gogortasuna gehienez 1,5 izateko probabilitatea.

b. Demagun σ^2 ezezaguna dela. Aleatorioki aukeratutako altzairuzko bost xaflen gogortasuna neurtzen duten koefizienteen bariantza 0,2 dela jakinik, lor bedi laginaren batez bestekoa gehienez 2 izateko probabilitatea.

Ebazpena

X = "altzairuzko xaflen gogortasun-koefizientea" aldagai aleatorioak banaketa normala du, $\mu = 1{,}6$ eta $\sigma = 0{,}3$ direlarik.

Enuntziatuko laginean $n = 5$ xaflaren gogortasun-koefizienteen balioak adierazi dira.

a. Aldagai aleatorioak banaketa normala duenez eta σ ezaguna denez, 1 teoremaren arabera $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ aldagaiak banaketa normal tipikoa du.

$$P(\overline{X} \le 1{,}5) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{1{,}5-1{,}6}{0{,}3/\sqrt{5}}\right) = P(Z \le -0{,}75) =$$

$$= 1 - \phi(0{,}75) = 1 - 0{,}7734 = 0{,}2266.$$

b. Oraingoan σ ezezaguna eta laginaren tamaina $n = 5 < 30$ direnez, 4 teorema aplikatuko da.

$t = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} = \dfrac{\overline{X}-\mu}{s/\sqrt{n-1}}$ aldagai aleatorio jarraituak $\nu = n-1$ askatasun-graduko Student-en banaketa du, non S^2 laginaren kuasibariantza eta $s^2 = 0{,}2$ laginaren bariantza diren.

$$P(\overline{X} \le 2) = P\left(\frac{\overline{X}-\mu}{s/\sqrt{n-1}} \le \frac{2-1{,}6}{0{,}45/\sqrt{4}}\right) = P(t_4 \le 1{,}78) = 1 - P(t_4 \ge 1{,}78) = 1-\alpha.$$

Student-en taulako balioen arteko interpolazioa aplikatuz,

$$\alpha = P(t_4 \ge 1{,}78) = 0{,}079$$

balioa lor daiteke. Balio hori ordezkatuz, eskatutako probabilitatea lortuko da.

$$P(\overline{X} \le 2) = 1 - 0{,}079 = 0{,}921.$$

5.10. ARIKETA PROPOSATUAK

5.10.1. ariketa

Informatikan matrikulatutako ikasleen batez besteko kalifikazioa 5,7 eta bariantza 5,97 izan dira. Aleatorioki, 50 ikasle aukeratu dira.

a. Zein da 50 ikasle hauen batez besteko kalifikazioa 6 baino handiagoa izateko probabilitatea?

b. Kalkula bedi horien batez besteko kalifikazioa gutxienez 5 izateko probabilitatea.

5.10.2. ariketa

Burdinazko bolen diametroak neurtzen dituen aldagai aleatorioak banaketa normala du, batez besteko diametroa 72 cm-koa delarik. Aleatorioki aukeratutako burdinazko zazpi bolaren diametroen neurriak hauek izan dira:

78,9 – 65,1 – 55,2 – 80,9 – 57,4 – 55,4 – 62,3.

Aleatorioki, beste zazpi bola hartu dira.

a. Kalkula bedi diametroen batez besteko neurria aurreko beste zazpi bolena baino handiagoa izateko probabilitatea.

b. Lor bedi diametroen batez besteko neurria gehienez 64 cm-koa izateko probabilitatea.

5.10.3. ariketa

Biruskontrako programa batek, kasuetako % 75ean birus konkretu bat desagerraraztea lortu du. Biruskontrako programa hori aleatorioki 150 ordenagailuri aplikatu zaie.

a. Zein da ordenagailu errekuperatuen proportzioa gutxienez % 65 eta gehienez % 85 izateko probabilitatea?

b. Kalkula bedi ordenagailu errekuperatuen proportzioa % 85 baino handiagoa izateko probabilitatea.

5.10.4. ariketa

Banaketa normaleko populazio batetik hartzen den 14 tamainako lagin aleatorio bakunaren bariantza 40,5 baino handiagoa bada, orduan populazioaren bariantza 20 delako baieztapena errefusatu egingo da.

a. Populazioaren bariantza 20 bada ere, zein probabilitatez errefusatuko da baieztapena?

b. Lor bedi laginaren bariantza gutxienez 40,5 eta gehienez 42 izateko probabilitatea.

5.10.5. ariketa

Bi aldagai aleatorio independenteren bidez I eta II estazioetatik egunean zehar igarotzen diren ibilgailuak zenbatzen dira. Demagun aldagaiak banaketa normalekoak direla, bariantzak 20 eta 36 ibilgailu karratu izanik, hurrenez hurren. Aleatorioki, I estazioan 8 egunetako ibilgailu-kopurua eta II estazioan 10 egunetako ibilgailu-kopurua hartu dira.

a. Kalkula bedi lagin hauen bariantzen arteko proportzioa gehienez 1,78 izateko probabilitatea.

b. Lor bedi $\dfrac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2}$ zatidura gehienez 2,44 izateko probabilitatea.

5.10.6. ariketa

Muntaketa-kate batean egiten diren piezetatik % 7 errepikatu behar dira. Berarekiko independentea den muntaketa-katean, aldiz, egiten diren piezetatik % 2 errepikatu behar dira. Aleatorioki lehenengo muntaketa-kateko 180 pieza eta bigarren muntaketa-kateko 300 pieza hartu dira.

a. Kalkula bedi errepikatu behar diren lehenengo lagineko piezen proportzioa, errepikatu behar diren bigarren lagineko piezen proportzioa baino handiagoa izateko probabilitatea.

b. Zein da $\hat{p}_1 - \hat{p}_2$ bi laginen proportzioen arteko kendura gutxienez 0,10 izateko probabilitatea?

5.10.7. ariketa

Kirol-firma baten salmentek bi herrialdetan zein bilakaera izan duten aztertu nahi da. Demagun salmentak banaketa normalekoak eta elkarrekiko independenteak direla, kirol-firmaren hileko batez besteko salmentak I herrialdean 136 produktu eta hileko batez besteko salmentak II herrialdean 83 produktu izanik. I herrialdean eta II herrialdean, aleatorioki 31 egunetako salmentak zenbatu dira eta salmenten kuasibariantzak 25 eta 16 produktu karratukoak dira, hurrenez hurren.

a. Kalkula bedi 31 egunetako I herrialdeko batez besteko salmentak II herrialdeko batez besteko salmenta gutxienez 50 produktutan gainditzeko probabilitatea.

b. Zein da 31 egunetako I herrialdeko batez besteko salmentak II herrialdeko batez besteko salmenta gehienez 57 produktutan gainditzeko probabilitatea?

5.10.8. ariketa

Zigarro-fabrikatzaile baten arabera XX markako zigarroek duten nikotina-kantitatearen desbideratze tipikoa 0,05 mg-koa da. Aleatorioki marka horretako 15 zigarro aukeratu dira. Demagun XX markako zigarroen nikotina-kantitateak banaketa normala duela.

a. Zein da hamabost zigarroen nikotina-kantitatearen desbideratze tipikoa gehienez 0,065 mg-koa izateko probabilitatea?

b. Lor bedi hamabost zigarroen nikotina-kantitatearen kuasibariantza gutxienez 0,0015 $(mg)^2$ izateko probabilitatea.

5.10.9. ariketa

A eta B motako piezen ekoizpenek banaketa normala dute, desbideratze tipikoak 3 eta 4,3 pieza izanik, hurrenez hurren. Ekoizpen hauek neurtzen dituzten aldagai aleatorioak elkarrekiko independenteak dira. Aleatorioki aukeratutako aste bateko 5 egunetako ekoizpenak hartu dira.

a. Kalkula bedi lagineko A motako piezen ekoizpenaren kuasibariantza lagineko B motako piezen ekoizpenaren kuasibariantzaren gutxienez bikoitza izateko probabilitatea.

b. Zein da a zenbakiaren balioa, $P\left(\dfrac{S_1^2}{S_2^2} \leq a\right) = 0{,}95$ bada?

5.10.10. ariketa

Eremu zehatz bateko enpresa defizitarioen kopurua urte hasieran eta urte bukaeran aztertu nahi da. X_1 = “enpresa defizitarioen kopurua urte hasieran” eta X_2 = “enpresa defizitarioen kopurua urte bukaeran” aldagaiak banaketa normalekoak dira, eta urte hasieran batez beste 150 enpresa defizitario eta urtearen bukaeran batez beste 100 enpresa defizitario daude. Bi aldagaiak independenteak eta bariantza ezezagunekoak direla suposatuko da. Aleatorioki, urtearen hasieran 16 enpresa eta urtearen bukaeran 12 enpresa hartu dira, urtearen hasierako enpresa defizitarioen kuasibariantza 64 eta urtearen bukaerako enpresa defizitarioen kuasibariantza 36 izanik. Lor bedi $\overline{X}_1 - \overline{X}_2$ diferentzia gehienez 45 enpresakoa izateko probabilitatea hurrengo baldintzetan:

a. Populazioen bariantzak berdinak direnean

b. Populazioen bariantzak ezberdinak direnean.

5.10.11. ariketa

Autoetako fokoen iraupenak $N(10, 3)$ banaketa du. Aleatorioki n foko aukeratu dira. Foko hauen batez besteko iraupenaren eta orokorreko batez besteko iraupenaren arteko diferentzia gehienez unitatea izateko probabilitatea 0,99 bada, zenbat autoz osaturiko lagina hartu behar da?

6. Estimazioa

6.1. ESTIMAZIOAREN KONTZEPTUA

Lagin aleatorio bakunetatik ateratako informaziotik populaziorako orokortasunak, ondorioak edota aurresanak lortzea ahalbidetzen duen alorrari **Inferentzia Estatistikoa** edo **Estatistika Induktiboa** deritzo.

Errealitateko problema ugaritan populazioaren parametro ezezagun (μ, σ, p…) baten balioa lortu nahi da. Horretarako, n tamainako lagin aleatorio bakunak hartuz eta laginketaren emaitzez baliatuz, populazioaren parametroen balio hurbilak kalkula daitezke. Prozesu horri **estimazioa** deritzo. Populazioaren parametrorako balio zehatz bat lortzen denean, **puntu-estimazioa** burutu da. Populazioaren parametroa zein tartetan dagoen zehazten denean, aldiz, **tarte-estimazioa** aplikatu da. Tarte horiei **konfiantza-tarteak** deritze. Bi kasuetan, estimazioz lortzen den populazioaren parametroaren balioa ez da ziurra; estimazioa burutzean probabilitateak parte hartuko du.

6.2. PUNTU-ESTIMAZIOA

Biz populazioaren ezaugarri bat aztertzen duen X aldagai aleatorioa. Populazioaren θ parametroaren balio hurbila edo puntu-estimazioa lortzeko, n tamainako lagin aleatorio bakuna x_1, x_2, …, x_n hartuko da. Parametroaren estimazioa zehazteko behar den lagin aleatorio bakunaren funtzioari $\hat{\theta} = \hat{\theta}(x_1, x_2, \ldots, x_n)$ **estimatzaile** deritzo eta lortuko den balio hurbilari **estimazio** izena emango zaio. Ahal den kasu guztietan **estimatzaile hoberena** erabiliko da, hots, bariantza txikieneko estimatzaile zentratua (bere itxarotako balioa parametroaren berdina duena, $E(\hat{\theta}) = \theta$) aukeratuko da.

Estimatzaile hoberenak zehazteko erabiltzen diren prozeduren artean **egiantza handieneko metodoa** azpimarra daiteke. Metodo horren bidez, estimatzailerik hoberena lortzeko, lagin aleatorio bakunaren eta parametroaren menpeko $F(x_1, x_2, \ldots, x_n; \theta)$ probabilitate-funtzioaren maximoa zehaztuko da.

Hona hemen estimatzaile hoberenen zenbait adibide:

i. n eta p parametroetako banaketa binomialeko populazioaren p arrakastaren proportziorako estimatzaile hoberena lagineko arrakasta-kopuruaren proportzioa da.

ii. Poisson-en banaketako λ parametroa estimatzeko estimatzailerik hoberena laginaren batez bestekoa da.

iii. μ eta σ parametroetako banaketa normalaren batez bestekorako estimatzailerik hoberena laginaren batez bestekoa da.

iv. μ eta σ parametroetako banaketa normalaren bariantzarako estimatzailerik hoberena laginaren kuasibariantza da.

6.3. TARTE-ESTIMAZIOA

Bira populazioaren θ parametro ezezaguna, $x_1, x_2, \ldots, x_n$ lagin aleatorio bakuna eta θ_1 eta θ_2 estimatzaileak. Bi estimatzaile horietan lagineko balioak ordezkatzean, L_1 eta L_2 balioak finkatuko dira. $[L_1, L_2]$ tartea θ parametroa estimatzeko **konfiantza-tartea** da, baldin eta

$$P(L_1 \le \theta \le L_2) = 1-\alpha$$

bada, $1-\alpha$ zenbakia $(0 < \alpha < 1)$ konfiantza-maila delarik. Laburbilduz, $I_\theta^{1-\alpha} = [L_1, L_2]$ delakoa θ parametroa estimatzeko $1-\alpha$ konfiantza-mailako konfiantza-tartea da. Emaitza fidagarria izateko, oro har, erabiltzen diren konfiantza-mailak 0,90, 0,95, 0,99… izaten dira.

6.3.1. Zenbait konfiantza-tarte

Azalduko diren konfiantza-tarteetan, $A_{\alpha/2}$ gaiak bere eskuinean dagokion banaketaren kasuan $\alpha/2$ azalerako eremua mugatuko du, non $A = z, t, \chi^2, F$ izan daitekeen. z hizkia banaketa normal tipikorako, t hizkia Student-en banaketarako, χ^2 delakoa khi karratuaren banaketarako eta F hizkia Snedecor-en banaketarako erabiliko dira. Bestalde, n gaiak laginaren tamaina, $1-\alpha$ zenbakiak konfiantza-maila eta S^2 balioak laginaren kuasibariantza adieraziko dituzte.

6.3.1.1. Populazio normalaren batez bestekorako konfiantza-tartea

Parametroa	*Populazioa*	*Lagina*	*Konfiantza-tartea*
μ	Normala σ ezaguna		$\left[\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$
μ	Normala σ ezezaguna	$n \ge 30$	$\left[\bar{x} - z_{\alpha/2}\dfrac{S}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\dfrac{S}{\sqrt{n}}\right]$
μ	Normala σ ezezaguna	$n < 30$	$\left[\bar{x} - t_{\alpha/2;n-1}\dfrac{S}{\sqrt{n}}, \bar{x} + t_{\alpha/2;n-1}\dfrac{S}{\sqrt{n}}\right]$

6.3.1.2. Bi banaketa normal independenteren batez bestekoen arteko kendurarako konfiantza-tartea

Param.	*Populazioa*	*Lagina*	*Konfiantza-tartea*
$\mu_1 - \mu_2$	Normalak, independenteak σ_1, σ_2 ezagunak		$\left[\left(\bar{x}_1 - \bar{x}_2\right) \mp z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$
$\mu_1 - \mu_2$	Normalak, independenteak σ_1, σ_2 ezezagunak	$n_1 \geq 30$ $n_2 \geq 30$	$\left[\left(\bar{x}_1 - \bar{x}_2\right) \mp z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right]$
$\mu_1 - \mu_2$	Normalak, independenteak σ_1, σ_2 ezezagunak $\sigma_1 = \sigma_2$	$n_1 < 30$ $n_2 < 30$	$\left[\left(\bar{x}_1 - \bar{x}_2\right) \mp t_{\alpha/2;v}\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$ $v = n_1 + n_2 - 2$
$\mu_1 - \mu_2$	Normalak, independenteak σ_1, σ_2 ezezagunak $\sigma_1 \neq \sigma_2$	$n_1 < 30$ $n_2 < 30$	$\left[\left(\bar{x}_1 - \bar{x}_2\right) \mp t_{\alpha/2;v}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right]$

Oharra: Azken kasuan $v = \dfrac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1+1} + \frac{\left(S_2^2/n_2\right)^2}{n_2+1}} - 2$ da.

6.3.1.3. Banaketa normalaren bariantzarako konfiantza-tartea

Parametroa	*Populazioa*	*Konfiantza-tartea*
σ^2	Normala, μ ezezaguna	$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right]$
σ^2	Normala, μ ezaguna	$\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\alpha/2;n}}, \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\alpha/2;n}}\right]$

6.3.1.4. Bi banaketa normal independenteren bariantzen arteko zatidurarako konfiantza-tartea

Param.	*Populazioa*	*Konfiantza-tartea*
$\dfrac{\sigma_1^2}{\sigma_2^2}$	Normalak, μ_1, μ_2 ezezagunak	$\left[\dfrac{S_1^2/S_2^2}{F_{\alpha/2;n_1-1,n_2-1}}, \dfrac{S_1^2/S_2^2}{F_{1-\alpha/2;n_1-1,n_2-1}}\right]$
$\dfrac{\sigma_1^2}{\sigma_2^2}$	Normalak, μ_1, μ_2 ezagunak	$\left[\dfrac{\dfrac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2}{n_1}}{\dfrac{\sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{n_2}}\dfrac{1}{F_{\alpha/2;n_1,n_2}}, \dfrac{\dfrac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2}{n_1}}{\dfrac{\sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{n_2}}\dfrac{1}{F_{1-\alpha/2;n_1,n_2}}\right]$

6.3.1.5. Banaketa binomialaren parametrorako konfiantza-tartea ($n \geq 30$)

$$I_p^{1-\alpha} = \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right].$$

6.3.1.6. Bi banaketa binomial independenteren proportzioen arteko kendurarako konfiantza-tartea ($n_1, n_2 \geq 30$)

$$I_{p_1-p_2}^{1-\alpha} = \left[\left(\hat{p}_1 - \hat{p}_2\right) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}, \left(\hat{p}_1 - \hat{p}_2\right) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}\right].$$

6.3.1.7. Bi banaketa normal ez independenteren batez bestekoen diferentziarako konfiantza-tartea

Parametroa	*Populazioa*	*Lagina*	*Konfiantza-tartea*
$\mu_1 - \mu_2$	Normalak ez independenteak	$n \geq 30$	$\left[\bar{d} - z_{\alpha/2}\dfrac{S}{\sqrt{n}}, \bar{d} + z_{\alpha/2}\dfrac{S}{\sqrt{n}}\right]$
$\mu_1 - \mu_2$	Normalak ez independenteak	$n < 30$	$\left[\bar{d} - t_{\alpha/2;n-1}\dfrac{S}{\sqrt{n}}, \bar{d} + t_{\alpha/2;n-1}\dfrac{S}{\sqrt{n}}\right]$

X eta Y aldagai aleatorioak ez dira independenteak. $D = X - Y$ aldagai aleatorioak banaketa normala du; batez bestekoa $\bar{d} = \sum_{i=1}^{n}\dfrac{d_i}{n} = \sum_{i=1}^{n}\dfrac{(x_i - y_i)}{n}$ da eta kuasibariantza $S^2 = \sum_{i=1}^{n}\dfrac{(d_i - \bar{d})^2}{n-1}$.

6.4. LAGINAREN TAMAINAREN DETERMINAZIOA

Estimazioa era egokian egiteko laginaren tamaina zehaztu behar da. Aurreko atalean adierazi diren konfiantza-tarteetan laginaren tamainak jokatzen duen papera oso garrantzitsua da; izan ere, laginaren tamaina zenbat eta handiagoa izan, konfiantza-tartearen luzera hainbat eta txikiagoa baita, eta ondorioz estimazioa zehatzagoa izango da. Ez da komenigarria tamaina handiegiko laginak hartzea, azterketa estatistikoa egiteko diru, denbora, edo beste mota bateko arazoak ekar baititzake. Laginaren tamaina txikiegia izateak, aldiz, ateratako emaitza oso fidagarria ez izatea ekar dezake. Ondorioz, kasu bakoitzean laginaren tamaina zehaztea komeni da. Adibide modura, gara dezagun σ^2 ezaguneko populazio normalaren batez bestekoa estimatzeko laginaren tamainaren determinazioa.

Kasu honetan, $1-\alpha$ konfiantza-mailaz, μ estimatzeko konfiantza-tartea hurrengo hau da:

$$I_{\mu}^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

Beraz, $1-\alpha$ probabilitatez, $\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ ezberdintza beteko da, edo era baliokidean:

$$1-\alpha = P\left(|\bar{x}-\mu| \le z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right). \tag{6.1}$$

Bestalde, demagun populazioaren batez bestekoaren eta laginaren batez bestekoaren arteko diferenzia gehienez ε errorea izatea nahi dela.

$$1-\alpha = P\left(|\bar{x}-\mu| \le \varepsilon\right). \tag{6.2}$$

$1-\alpha$ probabilitatez, (6.1) eta (6.2) adierazpenetatik $z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \varepsilon$ ondoriozta daiteke, eta hemendik laginaren tamaina zehaztuko da.

$$n = \left(z_{\alpha/2}\frac{\sigma}{\varepsilon}\right)^2.$$

6.5. ARIKETA EBATZIAK

6.5.1. ariketa

Demagun pila alkalinoen iraupena banaketa normalekoa dela. Aleatorioki 18 pila hartu dira, batez besteko iraupena 10.500 ordukoa eta kuasibariantza 2.304 (ordu)^2 direlarik.

a. % 98ko konfiantza-mailaz, zehatz bedi pila alkalinoen batez besteko iraupena.

b. % 98ko konfiantza-mailaz, estima bedi pila alkalinoen iraupenaren bariantza.

Ebazpena

Biz X = “pila alkalinoen iraupena” aldagai aleatorioa $N(\mu,\sigma)$ banaketakoa. Lagin aleatorio bakunaren datuak n = 18, $\bar{x}$ = 10.500 eta S = 48 dira.

a. Populazioaren bariantza ezezaguna eta laginaren tamaina n < 30 direnez, populazioaren batez bestekoa estimatzeko konfiantza-tartearen adierazpen orokorra hurrengoa da:

$$I_{\mu}^{1-\alpha} = \left[\bar{x} - t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}, \bar{x} + t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}\right].$$

Aurreko datuak eta $1-\alpha = 0{,}98$ konfiantza-maila ordezkatuz, konfiantza-tartea zehatz daiteke.

$$I_{\mu}^{0,98} = \left[10.500 - t_{0,01;17}\frac{48}{\sqrt{18}}, 10.500 + t_{0,01;17}\frac{48}{\sqrt{18}}\right].$$

Eranskineko Student-en taulatik $t_{0,01;17}$ = 2,567 balioa lortu eta konfiantza-tartean ordezkatuko da.

$$I_{\mu}^{0,98} = \left[10.500 - 2{,}567 \cdot 11{,}314, 10.500 + 2{,}567 \cdot 11{,}314\right] =$$

$$= [10.500 - 29,\ 10.500 + 29] =$$

$$= [10.417,\ 10.529].$$

Ondorioz, % 98ko konfiantza-mailaz, pila alkalinoen batez besteko iraupena gutxienez 10.471 ordukoa eta gehienez 10.529 ordukoa dela onar daiteke.

b. Kalkula dezagun σ^2 estimatzeko % 98ko konfiantza-mailako konfiantza-tartea hurrengo adierazpenean datuak ordezkatuz:

$$I_{\sigma^2}^{1-\alpha} = \left[\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right].$$

Kasu honetan,

$$I_{\sigma^2}^{0,98} = \left[\frac{17 \cdot (48)^2}{\chi^2_{0,01;17}}, \frac{17 \cdot (48)^2}{\chi^2_{0,99;17}}\right].$$

da, non $1-\alpha = 0{,}98$, n = 18, S = 48 diren. Bukatzeko, khi karratuaren banaketaren taulan $\chi^2_{0,01;17} = 33{,}4$ eta $\chi^2_{0,99;17} = 6{,}41$ balioak kalkulatu eta konfiantza-tartean ordezkatuko dira.

$$I_{\sigma^2}^{1-\alpha} = \left[\frac{17 \cdot (48)^2}{33{,}4}, \frac{17 \cdot (48)^2}{6{,}41}\right] =$$

$$= [1.172{,}69,\ 6.110{,}45].$$

6.5.2. ariketa

Multinazionaletako Elektronika Sailetako emakumeen proportzioa aztertu nahi da. Aleatorioki multinazionaletako Elektronika Sailetako 120 langileko talde bat hartu da, langile horietako 24 emakumeak izanik.

a. Kalkula bedi errorearen balioa, konfiantza-maila % 99 denean.

b. % 90eko konfiantza-mailaz, zein tartetan koka daiteke multinazionaletako Elektronika Sailetako emakumeen proportzioa?

Ebazpena

Bira X = “emakume-kopurua multinazionaletako Elektronika Sailetan” aldagai aleatorioa eta p = emakumeen proportzioa multinazionaletako Elektronika Sailetan. Kasu honetan, laginaren tamaina n = 120 eta lagineko emakumeen proportzioa $\hat{p} = \dfrac{24}{120} = 0,2$ dira.

a. Hurrengo konfiantza-tarteak proportzio bat estimatzea ahalbidetzen du, $1-\alpha$ konfiantza-mailaz:

$$I_p^{1-\alpha} = \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right].$$

Era baliokidean,

$$P\left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right] = 1-\alpha$$

edo

$$P\left[|p-\hat{p}| \le z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right] = 1-\alpha$$

da, non errorea $\varepsilon = z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$ den.

Errorearen kalkulua:

$$\varepsilon = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\varepsilon = z_{0,005}\sqrt{\frac{0,2\cdot 0,8}{120}}$$

$$\varepsilon = 2,575 \cdot 0,037 = 0,095$$

z = 2,575 balioa banaketa normal tipikoaren taulatik atera da.

b. Kalkula dezagun p estimatzeko $1-\alpha = 0{,}90$ konfiantza-mailako konfiantza-tartea.

$$I_p^{0,90} = \left[0,2 - z_{0,05}\sqrt{\frac{0,2 \cdot 0,8}{120}}, 0,2 + z_{0,05}\sqrt{\frac{0,2 \cdot 0,8}{120}}\right].$$

Banaketa normal tipikoaren taulatik $z_{0,05} = 1{,}645$ balioa aterako da, eta balio hori konfiantza-tartean estimatuz honako emaitza lortuko da:

$$\begin{aligned} I_p^{0,90} &= [0{,}2 - 1{,}645 \cdot 0{,}037,\ 0{,}2 + 1{,}645 \cdot 0{,}037] = \\ &= [0{,}2 - 0{,}061,\ 0{,}2 + 0{,}061] = \\ &= [0{,}139,\ 0{,}261]. \end{aligned}$$

Ondorioz, multinazionaletako Elektronika Sailetako emakumeen proportzioa % 13,9-% 26,1 bitartekoa da, % 90eko konfiantza-mailaz.

6.5.3. ariketa

Makina paketatzaile batek azukre paketeak betetzen ditu. Demagun paketeko azukre-kantitatea neurtzen duen aldagai aleatorioak banaketa normala duela. Aleatorioki 29 azukre pakete hartu dira, zeintzuen azukre-kantitatearen bariantza 0,0025 (gramo)^2 den. Zer esango zenuke hurrengo baieztapenari buruz: "% 95eko konfiantza-mailaz, azukre-kantitatearen desbideratze tipikoa gramo-erdi baino handiagoa da".

- -

Ebazpena

Biz X = "azukre-kantitatea paketeko" banaketa normalekoa. Bestalde, $n = 29$ tamainako lagin aleatorio bakunaren bariantza $s^2 = 0{,}0025$ (gr)^2 da.

Kalkula dezagun σ^2 estimatzeko % 95eko konfiantza-mailako konfiantza-tartea. Kasu honetan laginaren bariantza ezaguna denez, konfiantza-tartearen adierazpen orokorra hauxe da:

$$I_{\sigma^2}^{1-\alpha} = \left[\frac{ns^2}{\chi^2_{\alpha/2;n-1}}, \frac{ns^2}{\chi^2_{1-\alpha/2;n-1}}\right].$$

Kasu honetan,

$$I_{\sigma^2}^{0,95} = \left[\frac{29 \cdot 0,0025}{\chi^2_{0,025;28}}, \frac{29 \cdot 0,0025}{\chi^2_{0,975;28}}\right]$$

da, non $1-\alpha = 0{,}95$, $n = 29$, $s^2 = 0{,}0025$ diren. Bukatzeko, khi karratuaren banaketaren taulan $\chi^2_{0,025;28}$ eta $\chi^2_{0,975;28}$ balioak kalkulatu eta konfiantza-tartean ordezkatuko dira.

$$I_{\sigma^2}^{0,95} = \left[\frac{29 \cdot 0,0025}{44,5}, \frac{29 \cdot 0,0025}{15,31}\right] =$$

$$= [0,0016,\ 0,0047].$$

Erro karratu positiboa aplikatuz desbideratze tipikoa estimatzeko konfiantza-tartea lor daiteke.

$$I_{\sigma}^{0,95} = [0,040,\ \ 0,069].$$

Ondorioz, hurrengoa baiezta daiteke: "% 95eko konfiantza-mailaz, azukre-kantitatearen desbideratze tipikoa 0,040-0,069 gramo bitartekoa da". Beraz, enuntziatuan emandako baieztapena ez genuke ontzat hartuko.

6.5.4. ariketa

Bi herrialdetako energia elektrikoaren kontsumoa aztergai da. Bi herrialdeetan urteko energia elektrikoaren kontsumoak banaketa normala du, A herrialdeko energia elektrikoaren kontsumoak 100 milioi kWh-ko desbideratze tipikoa eta B herrialdean 90 milioi kWh-ko desbideratze tipikoa izanik. Demagun energia elektrikoaren kontsumoak elkarrekiko independenteak direla. Azken 12 urteetan kontsumitutako energia elektrikoa neurtu da eta A herrialdean batez beste 2.346 milioi kWh eta B herrialdean batez beste 2.247 milioi kWh kontsumitu dira. % 99ko konfiantza-mailaz, onar al daiteke A herrialdeko energia elektrikoaren kontsumoa B herrialdekoa baino handiagoa izatea?

Ebazpena

Bira X_1 = "A herrialdeko energia elektrikoaren kontsumoa" eta X_2 = "B herrialdeko energia elektrikoaren kontsumoa" aldagai aleatorioak, banaketa normalekoak.

$$X_1\!: N(\mu_1,100), \qquad X_2\!: N(\mu_2,90).$$

Bi lagin aleatorio bakunen datuak hauek dira:

$$n_1 = 12 \text{ hilabete}, \qquad \bar{x}_1 = 2.346 \text{ milioi Kwh}.$$

$$n_2 = 12 \text{ hilabete}, \qquad \bar{x}_2 = 2.247 \text{ milioi Kwh}.$$

Eraiki dezagun $\mu_1 - \mu_2$ estimatzeko $1-\alpha = 0,99$ konfiantza-mailako konfiantza-tartea.

$$I_{\mu_1-\mu_2}^{1-\alpha} = \left[(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

$$I_{\mu_1-\mu_2}^{0,99} = \left[(2.346 - 2.247) - z_{0,005}\sqrt{\frac{100^2}{12} + \frac{90^2}{12}}, (2.346 - 2.247) + z_{0,005}\sqrt{\frac{100^2}{12} + \frac{90^2}{12}}\right]$$

Banaketa normal estandarraren taulatik $z_{0,005} = 2{,}575$ balioa lortu da.

$$I^{0,99}_{\mu_1-\mu_2} = \left[(2.346 - 2.247) - z_{0,005}\sqrt{\frac{100^2}{12}+\frac{90^2}{12}}, (2.346 - 2.247) + z_{0,005}\sqrt{\frac{100^2}{12}+\frac{90^2}{12}}\right] =$$

$$= \left[99 - 2{,}575\sqrt{\frac{18.100}{12}}, 99 + 2{,}575\sqrt{\frac{18.100}{12}}\right] =$$

$$= [99 - 100, 99 + 100] =$$

$$= [-1, 199].$$

Tarte horretan balio negatiboak eta balio nulua daude. Ondorioz, % 99ko konfiantza-mailaz ezin da onartu A herrialdeko energia elektrikoaren kontsumoa B herrialdekoa baino handiagoa delako baieztapena.

6.5.5. ariketa

A eta B motako pieza mekanikoen ekoizpenek banaketa normalari jarraitzen diote eta elkarrekiko independenteak dira. Hurrengo taulan egunero ekoitzitako pieza-kopurua adierazi da:

	Astelehena	Asteartea	Asteazkena	Osteguna	Ostirala
A mota	91	88	96	95	87
B mota	92	86	98	90	88

a. % 90eko konfiantza-mailaz, onar al daiteke hurrengo baieztapena: “A motako piezen ekoizpenaren bariantza 5 $(\text{pieza})^2$ da”, alegia?

b. % 90eko konfiantza-mailaz, zein da A motako piezen ekoizpenaren bariantzaren eta B motako piezen ekoizpenaren bariantzaren arteko zatidurarako konfiantza-tartea?

Ebazpena

Bira X_1 = “A motako eguneko pieza-kopurua” eta X_2 = “B motako eguneko pieza-kopurua” aldagai aleatorioak, banaketa normalekoak eta elkarrekiko independenteak.

$$n_1 = 5 \text{ egun}, \ \bar{x}_1 = 91{,}4 \text{ pieza}, \ s_1^2 = 13{,}04 \ (\text{pieza})^2$$

$$n_2 = 5 \text{ egun}, \ \bar{x}_2 = 90{,}8 \text{ pieza}, \ s_2^2 = 16{,}96 \ (\text{pieza})^2$$

a. Lehenengo galderari erantzuteko, populazioaren σ_1^2 bariantza estimatzeko % 90eko konfiantza-mailako konfiantza-tartea eraikiko da.

$$I_{\sigma^2}^{1-\alpha} = \left[\frac{ns^2}{\chi^2_{\alpha/2;n-1}}, \frac{ns^2}{\chi^2_{1-\alpha/2;n-1}}\right]$$

$$I_{\sigma^2}^{0,90} = \left[\frac{5\cdot 13,04}{\chi^2_{0,05;4}}, \frac{5\cdot 13,04}{\chi^2_{0,95;4}}\right].$$

Khi karratuaren banaketaren taulatik $\chi^2_{0,05;4} = 9,49$ eta $\chi^2_{0,95;4} = 0,711$ balioak lortu eta konfiantza-tartean ordezkatuko dira.

$$I_{\sigma^2}^{0,90} = \left[\frac{65,2}{9,49}, \frac{65,2}{0,711}\right]$$

$$I_{\sigma^2}^{0,90} = [6,87,\ 91,70].$$

Ondorioz, $5 \notin I_{\sigma^2}^{0,90}$, orduan enuntziatuko baieztapena ez da onargarria, % 90eko konfiantza-mailaz.

b. Oraingoan, $1-\alpha = 0{,}90$ hartuz, bi bariantzen arteko zatidura estimatzeko konfiantza-tartea eraikiko da, aldagaiak banaketa normalekoak, independenteak eta populazioen batez bestekoak ezezagunak direlarik.

$$I_{\sigma_1^2/\sigma_2^2}^{1-\sigma} = \left[\frac{S_1^2/S_2^2}{F_{\alpha/2;n_1-1,n_2-1}}, \frac{S_1^2/S_2^2}{F_{1-\alpha/2;n_1-1,n_2-1}}\right].$$

Laginen bariantzak kalkulatuta daudenez, bariantzen eta kuasibariantzen arteko erlazioa aplikatuko da.

$$\frac{S_1^2}{S_2^2} = \frac{\frac{n_1}{n_1-1}s_1^2}{\frac{n_2}{n_2-1}s_2^2} = \frac{\frac{5}{4}\cdot 13,04}{\frac{5}{4}\cdot 16,96} = 0,769.$$

Orduan, eskatutako konfiantza-tartea hurrengo hau da:

$$I_{\sigma_1^2/\sigma_2^2}^{0,90} = \left[\frac{0,769}{F_{0,05;4,4}}, \frac{0,769}{F_{0,95;4,4}}\right].$$

Snedecor-en taulan bilatuz, $F_{0,05;4,4} = 6{,}39$ balioa lortzen da. Bestalde, Snedecor-en banaketaren propietatea aplikatuz, $F_{0,95;4,4} = \frac{1}{6,39} = 0,156$ balioa kalkula daiteke. Bi balioak konfiantza-tartean ordezkatuz, eskatutakoa lortuko da.

$$I_{\sigma_1^2/\sigma_2^2}^{0,90} = \left[\frac{0,769}{6,39}, \frac{0,769}{1/6,39}\right] =$$

$$= [0{,}12,\ 4{,}91].$$

6.5.6. ariketa

Bi patata-motaren ekoizpena konparatzeko asmoz, lursailak bi zatitan banatu dira eta lursail bakoitzean patata-mota bat landatu da. Aleatorioki hartutako sei lursailetan izandako patata-ekoizpena (kilo) hurrengoa da:

A mota	21	24	18	19	20	18
B mota	19	20	19	20	22	17

Demagun patataren ekoizpenak banaketa normala duela.

a. % 95eko konfiantza-mailaz, lor bedi patata-ekoizpenaren batez bestekoen arteko diferentziarako konfiantza-tartea.

b. Eta % 99ko konfiantza-mailaz?

Ebazpena

Bira X_1 = "A motako pataten ekoizpena" eta X_2 = "B motako pataten ekoizpena" aldagai aleatorioak, banaketa normalekoak. X_1 eta X_2 ez dira independenteak, bi patata-moten landaketak lursail berdinean egiten baitira. Beraz, kasu honetan batez bestekoen diferentziarako konfiantza-tartea hauxe da:

$$I_{\mu_1-\mu_2}^{1-\alpha} = \left[\bar{d} - t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}, \bar{d} + t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}\right]$$

$\bar{d} = \sum_{i=1}^{n} \frac{d_i}{n} = \sum_{i=1}^{n} \frac{(x_{i,1} - x_{i,2})}{n}$ eta $S^2 = \sum_{i=1}^{n} \frac{(d_i - \bar{d})^2}{n-1}$ dira.

d_i diferentzien balioak:

d_i	2	4	–1	–1	–2	1

Binakako datuen batez bestekoa eta kuasibariantza:

$$\bar{d} = \sum_{i=1}^{n} \frac{d_i}{n} = \frac{2+4-1-1-2+1}{6} = 0,5$$

$$S^2 = \sum_{i=1}^{n} \frac{(d_i - \bar{d})^2}{n-1} = \frac{(1,5)^2 + (3,5)^2 + (-1,5)^2 + (-1,5)^2 + (-2,5)^2 + (0,5)^2 +}{5} = 5,1.$$

Kuasibariantzaren erro karratu positiboa kalkulatuz, $S = 2{,}26$ lortuko da.

a. Balio hauek eta Student-en taulako $t_{0,025;5} = 2{,}571$ balioa % 95eko konfiantza-mailako konfiantza-tartearen adierazpenean ordezkatuko dira.

$$I^{0,95}_{\mu_1-\mu_2} = \left[0{,}5 - t_{0,025;5}\frac{2{,}26}{\sqrt{6}}, 0{,}5 + t_{0,025;5}\frac{2{,}26}{\sqrt{6}}\right] =$$

$$= [0{,}5 - 2{,}571 \cdot 0{,}92,\ 0{,}5 + 2{,}571 \cdot 0{,}92] =$$

$$= [0{,}5 - 2{,}37,\ 0{,}5 + 2{,}37] =$$

$$= [-1{,}87,\ 2{,}87].$$

b. Oraingoan egoera berdina, baina konfiantza-maila ezberdina da. Beraz, $1-\alpha = 0{,}99$ dela kontuan hartuz eta Student-en taulako $t_{0,005;5} = 4{,}032$ balioa ordezkatuz, hurrengo konfiantza-tartea lor daiteke:

$$I^{0,99}_{\mu_1-\mu_2} = \left[0{,}5 - t_{0,005;5}\frac{2{,}26}{\sqrt{6}}, 0{,}5 + t_{0,005;5}\frac{2{,}26}{\sqrt{6}}\right] =$$

$$= [0{,}5 - 4{,}032 \cdot 0{,}92,\ 0{,}5 + 4{,}032 \cdot 0{,}92] =$$

$$= [0{,}5 - 3{,}71,\ 0{,}5 + 3{,}71] =$$

$$= [-3{,}21,\ 4{,}21].$$

6.5.7. ariketa

Makina batek goizeko txandan eta arratsaldeko txandan ekoizten dituen pieza akastunen proportzioak aztertu nahi dira. Aleatorioki goizeko txandan ekoitzitako 200 pieza hartu dira, horietatik 18 pieza akastunak direlarik. Era berean, aleatorioki arratsaldeko txandan ekoitzitako 180 pieza aukeratu dira, 18 pieza akastunak direlarik. Demagun goizeko txandan ekoizten den pieza akastunen kopurua arratsaldeko txandan ekoizten denarekiko independentea dela.

a. % 95eko konfiantza-mailaz, lor bedi proportzioen diferentziarako konfiantza-tartea.

b. Demagun goizeko txandako eta arratsaldeko txandako piezetatik tamaina bereko bi lagin hartzen direla, hurrenez hurren. % 95eko konfiantza-mailaz, zenbat pieza hartu behar dira proportzioen diferentzia estimatzeko, errorea gehienez % 1 denean?

Ebazpena

Bira p_1 = "goizeko txandan ekoizten diren pieza akastunen proportzioa", p_2 = "arratsaldeko txandan ekoizten diren pieza akastunen proportzioa", $\hat{p}_1 = \frac{18}{200} = 0{,}09$ eta

$\hat{p}_2 = \frac{18}{180} = 0,10$ hartutako lagin aleatorio bakunetako pieza akastunen proportzioak.

a. Bi proportzioen arteko diferentzia estimatzeko $1-\alpha$ konfiantza-mailako konfiantza-tartea honako hau da:

$$I_{p_1-p_2}^{1-\alpha} = \left[\left(\hat{p}_1 - \hat{p}_2\right) \mp z_{\alpha/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \right].$$

Kasu honetan,

$$I_{p_1-p_2}^{0,95} = \left[\left(0,09 - 0,10\right) \mp z_{0,025} \sqrt{\frac{0,09 \cdot 0,91}{200} + \frac{0,10 \cdot 0,90}{180}} \right]$$

da, non $z_{0,025} = 1,96$ den.

Orduan, $p_1 - p_2$ estimatzeko % 95eko konfiantza-mailako konfiantza-tartea hauxe da:

$$I_{p_1-p_2}^{0,95} = [-0,01 - 1,96 \cdot 0,03, -0,01 + 1,96 \cdot 0,03] =$$
$$= [-0,069,\ 0,049\ ]$$

b. Bi proportzioen diferentzia estimatzeko konfiantza-tartetik

$$P\left[\left(\hat{p}_1 - \hat{p}_2\right) - z_{a/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \le p_1 - p_2 \le \left(\hat{p}_1 - \hat{p}_2\right) + z_{a/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \right] = 1 - \alpha$$

edo era baliokidean

$$P\left[\left| (p_1 - p_2) - (\hat{p}_1 - \hat{p}_2) \right| \le z_{a/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \right] = 1 - \alpha$$

ondoriozta daiteke. Gainera, kasu honetan $n_1 = n_2 = n$ da. Beraz, errorearen adierazpena hauxe da:

$$\varepsilon = z_{a/2} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n} + \frac{\hat{p}_2\hat{q}_2}{n}}$$

Kasu honetan errorea gehienez 0,01 denez, balioak ordezkatuz

$$0,01 = 1,96 \sqrt{\frac{0,09 \cdot 0,91 + 0,10 \cdot 0,90}{n}}$$

lor daiteke, eta n bakanduz txanda bakoitzean zenbat pieza hartu behar diren ebatziko da.

$$n = \left(\frac{1,96}{0,01} \right)^2 (0,09 \cdot 0,91 + 0,10 \cdot 0,90)$$

$n \approx 6.604$ pieza txanda bakoitzean.

6.5.8. ariketa

Erreakzio kimiko mota baten erreakzio-denborak banaketa normala du. Aleatorioki 67 erreakzio kimiko burutu dira eta erreakzio-denbora batez beste 2,3 ordu eta desbideratze tipikoa 0,71 ordu izan dira.

a. Estima bedi batez besteko erreakzio-denbora % 97ko konfiantza-mailako konfiantza-tartea zehaztuz.

b. Zein da aurreko konfiantza-mailaz eta 85 tamainako lagin aleatorio bakuna hartuta atera daitekeen ondorioa?

Ebazpena

Biz X = "erreakzio kimikoaren erreakzio-denbora" aldagai aleatorioa $N(\mu,\sigma)$ banaketakoa. Lagin aleatorio bakunaren datuak $n = 67$, $\bar{x} = 2{,}3$ eta $s = 0{,}71$ dira.

a. Kasu honetan, populazioaren batez bestekoa estimatzeko konfiantza-tartearen adierazpen orokorra hurrengoa da:

$$I_\mu^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{S}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right].$$

Laginaren kuasibariantza ezezaguna, baina laginaren desbideratze tipikoaren balioa ezaguna da. Orduan populazioaren batez bestekoa estimatzeko hurrengo konfiantza-tarte baliokidea erabil daiteke:

$$I_\mu^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n-1}}, \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n-1}}\right]$$

$$I_\mu^{0,97} = \left[2{,}3 - z_{0{,}015}\frac{0{,}71}{\sqrt{66}}, 2{,}3 + z_{0{,}015}\frac{0{,}71}{\sqrt{66}}\right].$$

Banaketa normal tipikoaren taulatik $z_{0{,}015} = 2{,}17$ balioa lortu eta aurreko adierazpenean ordezkatuz, batez besteko erreakzio-denborarako konfiantza-tartea zehaztuko da.

$$I_\mu^{0,97} = [2{,}3 - 0{,}19,\ 2{,}3 + 0{,}19] =$$

$$= [2{,}11,\ 2{,}49].$$

b. Aurreko ataleko azterketa errepikatu behar da, baina $n = 85$ hartuz.

$$I_\mu^{0,97} = \left[2{,}3 - 2{,}17\cdot\frac{0{,}71}{\sqrt{84}}, 2{,}3 + 2{,}17\cdot\frac{0{,}71}{\sqrt{84}}\right] =$$

$$= [2{,}3 - 0{,}17,\ 2{,}3 + 0{,}17] =$$

$$= [2{,}13,\ 2{,}47].$$

Laginaren tamaina handituz, errorea txikiagoa egiten da, eta ondorioz, konfiantza-tartearen luzera txikiagoa da, estimazioa zehatzagoa delarik.

6.5.9. ariketa

Demagun mezularitza-enpresa bateko langileek eguneko egiten duten distantzia neurtzen duen aldagai aleatorioa banaketa normalekoa dela. Aleatorioki 100 egunetan egindako distantzia zenbatu da, batez beste 550 km eta kuasibariantza 10.000 km^2 izanik. Zein konfiantza-mailaz esan daiteke mezularitza-enpresako langileek eguneko egiten duten batez besteko distantzia gutxienez 520 km-koa eta gehienez 580 km-koa dela?

Ebazpena

X = “mezularitza-enpresako langileek eguneko egiten duten distantzia” aldagai aleatorioak banaketa normala duela suposatu da, non desbideratze tipikoa ezezaguna den.

Laginaren tamaina n =100 (> 30) egunetako distantzia, $\bar{x}$ = 550 km eta S = 100 km dira.

Baldintza hauetan, mezularitza-enpresako langileek eguneko egiten duten batez besteko distantzia estimatzeko 1−α konfiantza-mailako konfiantza-tartea honako hau da:

$$I_{\mu}^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{S}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right].$$

Era baliokidean,

$$1 - \alpha = P\left[\bar{x} - z_{\alpha/2}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

adierazpena dugu. Kasu konkretu honetan:

$$1-\alpha = P(520 \leq \mu \leq 580)$$

Beraz,

$$\begin{cases} 520 = \bar{x} - z_{\alpha/2}\dfrac{S}{\sqrt{n}} \\ 580 = \bar{x} + z_{\alpha/2}\dfrac{S}{\sqrt{n}} \end{cases}$$

ekuazio-sisteman n = 100, $\bar{x}$ = 550 eta S = 100 datuak ordezkatuz,

$$\begin{cases} 520 = 550 - z_{\alpha/2}\dfrac{100}{\sqrt{100}} \\ 580 = 550 + z_{\alpha/2}\dfrac{100}{\sqrt{100}} \end{cases}$$

ekuazio-sistema ebatz daiteke eta honako balio hau lortu:

$$z_{\alpha/2} = 3$$

Banaketa normal tipikoaren taula kontsultatuz, balio hori

$$\frac{\alpha}{2} = (1 - 0,9987) = 0,0013$$

gaiari dagokio. Ondorioz, eskatutako konfiantza-maila hauxe da:

$$1-\alpha = 1 - 2 \cdot 0,0013 = 0,9974.$$

6.5.10. ariketa

Azterketa baterako bi motatako bateriak behar dira. I. motako eta II. motako baterien tentsioek banaketa normala dute eta elkarrekiko independenteak dira. Demagun I. motako eta II. motako baterien tentsioen bariantzak berdinak direla. Aleatorioki I. motako 22 bateria hartu dira, batez besteko tentsioa 19,7 V eta desbideratze tipikoa 0,6 V izanik. Era berean, aleatorioki II. motako 20 bateria hartu dira, eta batez beste 18,2 V-eko tentsioa eta 0,4 V-eko desbideratze tipikoa dituzte. % 98ko konfiantza-mailaz, onargarria al da bi bateria-motek batez besteko tentsio berdina izatea?

Ebazpena

Bira X_1= "I. motako baterien tentsioa" eta X_2 = "II. motako baterien tentsioa" aldagai aleatorioak, banaketa normalekoak eta bariantza berekoak.

$$X_1: N(\mu_1,\sigma), \qquad X_2: N(\mu_2,\sigma).$$

Bi lagin aleatorio bakunen datuak hauek dira:

$n_1 = 22$ bateria, $\bar{x}_1 = 19,7$ V, $s_1 = 0,6$ V.

$n_2 = 20$ bateria, $\bar{x}_2 = 18,2$ V, $s_2 = 0,4$ V.

Egindako galderari erantzuteko, $\mu_1 - \mu_2$ estimatzeko $1 - \alpha = 0,98$ konfiantza-mailako konfiantza-tartea eraikiko da, non populazioaren bariantzak ezezagunak eta berdinak eta hartutako lagin aleatorio bakunak txikiak diren.

$$I^{1-\alpha}_{\mu_1-\mu_2} = \left[\left(\bar{x}_1 - \bar{x}_2\right) \mp t_{\alpha/2;n_1+n_2-2} \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right]$$

$$I^{0,98}_{\mu_1-\mu_2} = \left[(19,7 - 18,2) \mp t_{0,01;40} \sqrt{\frac{22 \cdot (0,6)^2 + 20 \cdot (0,4)^2}{22 + 20 - 2}} \sqrt{\frac{1}{22} + \frac{1}{20}} \right].$$

Student-en taulan kontsultatuz, $t_{0,01;40} = 2{,}423$ balioa lor daiteke. Azkenik, konfiantza-tartea zehazteko, balio hau konfiantza-tartearen adierazpenean ordezkatuko da.

$$I_{\mu_1-\mu_2}^{0,98} = \left[(19{,}7-18{,}2) \pm 2{,}423\sqrt{\frac{22\cdot(0{,}6)^2+20\cdot(0{,}4)^2}{22+20-2}}\sqrt{\frac{1}{22}+\frac{1}{20}} \right] =$$

$$= [1{,}5 - 2{,}423 \cdot 0{,}163,\ 1{,}5 + 2{,}423 \cdot 0{,}163] =$$

$$= [1{,}5 - 0{,}396,\ 1{,}5 + 0{,}396] =$$

$$= [1{,}104,\ 1{,}896].$$

0 balioa ez dago tarte horretan. Beraz, % 98ko konfiantza-mailaz ezin da onartu bi bateria-moten batez besteko tentsioak berdinak direnik.

6.5.11. ariketa

Nitrogeno dioxidoaren mailak neurtzen dituen makina batek ez ditu datu guztiak jasotzen, hau da, eguneko zenbait ordutan izandako nitrogeno dioxidoaren mailak balio galdutzat hartzen dira. Aleatorioki hartutako udako bi egunetan orduro jasotako nitrogeno dioxidoaren datuetatik 6 balio galdu dira. Era berean, aleatorioki hartutako neguko bi egunetan neurtutako datuetatik 8 balio galdu dira.

a. % 90eko konfiantza-mailaz, estima bedi udako nitrogeno dioxidoaren mailen artean balio galduen proportzioa.

b. % 90eko konfiantza-mailaz, estima bedi $p_1 - p_2$ udako balio galduen proportzioaren eta neguko balio galduen proportzioaren arteko diferentzia.

Ebazpena

Bira p_1 = “nitrogeno dioxidoaren mailen artean, udako balio galduen proportzioa” eta p_2 = “nitrogeno dioxidoaren mailen artean, neguko balio galduen proportzioa”. Bestalde, hartu diren lagin aleatorio bakunen tamainak $n_1 = n_2 = 48 (n > 30)$ eta laginetako proportzioak $\hat{p}_1 = \frac{6}{48} = 0{,}125$ eta $\hat{p}_2 = \frac{8}{48} = 0{,}167$ dira.

a. Hurrengo konfiantza-tartearen bidez udako balio galduen proportzioa estima daiteke, $1-\alpha$ konfiantza-mailaz:

$$I_{p_1}^{1-\alpha} = \left[\hat{p}_1 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}},\ \hat{p}_1 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}} \right].$$

$1 - \alpha = 0{,}90$ konfiantza-mailari taula normal estandarreko $z_{\alpha/2} = z_{0,05} = 1{,}645$ balioa dagokio. Balio guztiak konfiantza-tartean ordezkatuz, eskatutako tartea zehatz daiteke:

$$I_{p_1}^{0,90} = \left[0,125 - 1,645\sqrt{\frac{0,125\cdot 0,875}{48}}, 0,125 + 1,645\sqrt{\frac{0,125\cdot 0,875}{48}}\right] =$$

$$= [0,125 – 0,079, 0,125 + 0,079] =$$

$$= [0,046, 0,204].$$

Ondorioz, % 90eko konfiantza-mailaz, udako nitrogeno dioxidoaren mailei dagokien balio galduen proportzioa gutxienez % 4,6 eta gehienez % 20,4 izango da.

b. Oraingoan, $p_1 - p_2$ estimatzeko $1 - \alpha = 0,90$ konfiantza-mailako konfiantza-tartea eraiki behar da:

$$I_{p_1-p_2}^{1-\alpha} = \left[(\hat{p}_1 - \hat{p}_2) \mp z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}\right].$$

Datu guztiak eta banaketa normal tipikoaren taulatik ateratako $z_{0,05} = 1,645$ balioa ordezkatuz, hurrengo estimazio-tartea lor daiteke:

$$I_{p_1-p_2}^{0,90} = \left[(0,125 - 0,167) \mp 1,645\sqrt{\frac{0,125\cdot 0,875}{48} + \frac{0,167\cdot 0,833}{48}}\right] =$$

$$= [–0,042 – 0,118, \ –0,042 + 0,118] =$$

$$= [–0,16, 0,076].$$

Ondorioz, $p_1 - p_2$ estimatzeko % 90eko konfiantza-mailako konfiantza-tarteak balio ez-positiboak dituenez, ezin izango da ondorioztatu nitrogeno dioxidoaren mailei dagokienez, neguko balio galduen proportzioa udako balio galduen proportzioa baino handiagoa denik.

6.5.12. ariketa

Lantegi batean burdinazko habeak ekoizten dira. Burdinazko habeen luzera neurtzen duen aldagai aleatorioak banaketa normala du, desbideratze tipikoa 1,7 cm-koa izanik.

a. Zenbat habez osaturiko lagin aleatorio bakuna hartu behar da, burdinazko habeen batez besteko luzeraren eta lagineko burdinazko habeen batez besteko luzeraren arteko diferentzia gehienez 1 cm-ekoa izateko probabilitatea 0,992 bada?

b. Aurreko atalean lortutako tamainako lagin aleatorio bakuna hartu da eta horren batez besteko luzera 14 cm-koa izan da. Oro har, zein da burdinazko habeen batez besteko luzera 13-15 cm-ko tartetik kanpo izateko probabilitatea?

Ebazpena

Biz X = "burdinazko habeen luzera" aldagai aleatorioa, X: $N(\mu,\ 1{,}7)$.

a. Alde batetik, populazioaren banaketa normala eta bariantza ezagunak direnean, populazioaren batez bestekoa estimatzeko konfiantza-tartea hau da:

$$I_\mu^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

Era baliokidean:

$$1-\alpha = P\left[|\mu - \bar{x}| \le z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

$\sigma = 1{,}7$, $1 - \alpha = 0{,}992$, $\alpha/2 = 0{,}004$, eta $z_{\alpha/2} \equiv z_{0{,}004} = 2{,}65$ balioak kontuan hartuz, honakoa betetzen da:

$$0{,}992 = P\left[|\mu - \bar{x}| \le 2{,}65 \cdot \frac{1{,}7}{\sqrt{n}}\right]. \qquad (1)$$

Bestalde, enuntziatuak dioena honela adieraz daiteke:

$$0{,}992 = P\left(|\mu - \bar{x}| \le 1\right). \qquad (2)$$

(1)-(2) ekuazioen ondorioz, 0,992ko probabilitateaz, laginaren tamaina zehazteko hurrengo ekuazioa ebatziko da:

$$2{,}65 \cdot \frac{1{,}7}{\sqrt{n}} = 1$$

$$n = \left(2{,}65 \cdot \frac{1{,}7}{1}\right)^2 \approx 20.$$

b. Lagin aleatorio bakunaren datuak n = 20 burdinazko habe eta $\bar{x}$ = 14 cm dira. Bestalde, burdinazko habeen batez besteko luzera estimatzeko $1 - \alpha$ konfiantza-mailako konfiantza-tartea honako hau da:

$$I_\mu^{1-\alpha} = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

Era baliokidean,

$$1-\alpha = P\left(\mu \in \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]\right)$$

betetzen da, eta

$$\alpha = P\left(\mu \notin \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]\right)$$

ondoriozta daiteke.

Enuntziatuko datuen arabera,

$$\alpha = P(\mu \notin [13,15])$$

berdintzatik zehaztea eskatzen da. Ondorioz;

$$\begin{cases} 13 = \bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} \\ 15 = \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} \end{cases}$$

ekuazio-sisteman $n = 20$, $\bar{x} = 14$ eta $\sigma = 1{,}7$ balioak ordezkatu eta gero, ezezagun bakarra $z_{\alpha/2}$ balioa lortuko da

$$\begin{cases} 13 = 14 - z_{\alpha/2}\dfrac{1{,}7}{\sqrt{20}} \\ 15 = 14 + z_{\alpha/2}\dfrac{1{,}7}{\sqrt{20}} \end{cases}$$

Orduan,

$$z_{\alpha/2} \approx 2{,}63$$

balioa lortuko da, eta banaketa normal tipikoaren taula kontsultatuz, balio hau

$$\frac{\alpha}{2} = (1 - 0{,}9957) = 0{,}0043$$

probabilitateari dagokio. Ondorioz, eskatutako probabilitatearen balioa hurrengoa da:

$$\alpha = 2 \cdot 0{,}0043 = 0{,}0086.$$

6.5.13. ariketa

Produktu bat egiteko A eta B prozedurak erabiltzen dira. Prozedura horietatik garestiena zein den ikusi nahi da. Aleatorioki A prozeduraz eginiko 10 produktu hartu dira, horien kostua batez beste 2.600 euro eta kuasibariantza 101.124 (euro)^2 izanik. Era berean, B prozeduraz eginiko 16 produkturen batez besteko kostua 2.200 euro eta kuasibariantza 40.000 (euro)^2 dira. Demagun prozedura ezberdinez egindako produktuen kostuek banaketa normala eta elkarrekiko independentea dutela.

a. Lor bedi batez besteko kostuen diferentzia estimatzeko % 90eko konfiantza-mailako konfiantza-tartea.

b. % 90eko konfiantza-mailaz, estima bedi A prozedurari dagokion kostuaren bariantzaren eta B prozedurari dagokion kostuaren bariantzaren arteko zatidura.

Ebazpena

Bira X_1 = "A prozeduraz eginiko produktuen kostua" eta X_2 = "B prozeduraz eginiko produktuen kostua" aldagai aleatorioak, banaketa normalekoak eta elkarrekiko independenteak, bariantza ezezagunekoak.

$$X_1: N(\mu_1,\sigma_1), \qquad X_2: N(\mu_2,\sigma_2).$$

Bi lagin aleatorio bakunen datuak hauek dira:

$$n_1 = 10 \text{ produktu}, \quad \bar{x}_1 = 2.600 \text{ euro}, \quad S_1 = 318 \text{ euro}.$$

$$n_2 = 16 \text{ produktu}, \quad \bar{x}_2 = 2.200 \text{ euro}, \quad S_2 = 200 \text{ euro}.$$

a. % 90eko konfiantza-mailaz ($1 - \alpha = 0{,}90$), $\mu_1 - \mu_2$ batez besteko kostuen diferentzia estimatzeko konfiantza-tartea eraikiko da. Populazioaren bariantzak ezberdinak direla suposa daitekeenez eta laginen tamainak txikiak direnez, konfiantza-tartea zehazteko hurrengo adierazpena erabiliko da:

$$I^{1-\alpha}_{\mu_1-\mu_2} = \left[\left(\bar{x}_1 - \bar{x}_2 \right) \mp t_{\alpha/2;v} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \right]$$

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1+1} + \frac{\left(S_2^2/n_2\right)^2}{n_2+1}} - 2 = \frac{\left(\frac{318^2}{10} + \frac{200^2}{16} \right)^2}{\frac{\left(318^2/10\right)^2}{11} + \frac{\left(200^2/16\right)^2}{17}} - 2 \approx 14 \text{ izanik.}$$

Kasu honetan,

$$I^{0,90}_{\mu_1-\mu_2} = \left[(2.600 - 2.200) \mp t_{0,05;14} \sqrt{\frac{318^2}{10} + \frac{200^2}{16}} \right]$$

konfiantza-tartea zehazteko, azkenik Student-en taulako $t_{0,05;14} = 1{,}761$ balioa ordezkatuko da.

$$I^{0,90}_{\mu_1-\mu_2} = [400 - 1{,}761 \cdot 112{,}3,\ \ 400 + 1{,}761 \cdot 112{,}3] =$$

$$= [400 - 197{,}76,\ \ 400 + 197{,}76] =$$

$$= [202{,}24,\ \ 597{,}76].$$

Ondorioz, $\mu_1 - \mu_2$ estimatzeko tarteko balioak positiboak direnez, A prozeduraren bidezko batez besteko kostuak B prozeduraren bidezko kostuak baino handiagoak direla onar daiteke, % 90eko konfiantza-mailaz.

b. Eraiki dezagun bi bariantzen arteko zatidura estimatzeko konfiantza-tartea, $1 - \alpha = 0{,}90$ konfiantza-mailaz.

$$I^{1-\alpha}_{\sigma_1^2/\sigma_2^2} = \left[\frac{S_1^2 / S_2^2}{F_{\alpha/2;n_1-1,n_2-1}}, \frac{S_1^2 / S_2^2}{F_{1-\alpha/2;n_1-1,n_2-1}} \right] =$$

$$= \left[\frac{(318)^2 /(200)^2}{F_{0,05;9,15}}, \frac{(318)^2 /(200)^2}{F_{0,95;9,15}} \right].$$

Azkenik, tarteko muturrak zehazteko Snedecor-en taulako $F_{0,05;9,15} = 2{,}59$ lortu behar da. Ondoren, Snedecor-en banaketaren propietatea aplikatuz, $F_{0,95;9,15} = \dfrac{1}{F_{0,05;15,9}} = \dfrac{1}{3{,}01}$ balioa kalkula daiteke. Bi balioak konfiantza-tartean ordezkatuz, konfiantza-tartea hurrengoa da:

$$I^{0,90}_{\sigma_1^2/\sigma_2^2} = \left[\frac{2{,}53}{2{,}59}, \frac{2{,}53}{1/3{,}01} \right] =$$

$$= [0{,}977,\ 7{,}615].$$

6.6. ARIKETA PROPOSATUAK

6.6.1. ariketa

Automozioaren industriako langileen batez besteko adina estimatu nahi da. Horretarako, aleatorioki automozioko 61 langile hartu dira eta beraien batez besteko adina 45 urte eta desbideratze tipikoa 15 urte dira. Demagun automozioaren industriako langileen adinak banaketa normala duela.

a. % 95eko konfiantza-mailaz, eraiki bedi automozioko langileen adinaren batez bestekorako konfiantza-tartea.

b. Zein da automozioko langileen adinaren desbideratze tipikorako % 95eko konfiantza-mailako konfiantza-tartea?

6.6.2. ariketa

Airearen kutsadurak arnasketa-arazoak sorrarazten ditu. Horrela, aleatorioki aukeratutako 100 pertsonatik 18 pertsonak arnasa hartzeko arazoak dituztela ikusi da.

a. % 99ko konfiantza-mailaz, estima bedi airearen kutsaduragatik arnasketa-arazoak dituzten pertsonen proportzioa.

b. Airearen kutsaduragatik arnasketa-arazoak dituzten pertsonen proportzioei buruz, zein konfiantza-mailaz esan daiteke laginaren proportzioaren eta populazioaren proportzioaren arteko diferentzia gehienez % 9 dela?

6.6.3. ariketa

Aireportu batean hilero galtzen den ekipajetik zein ehuneko berreskuratzen den jakin nahi da. Horretarako, aleatorioki galdutako n ekipajez osaturiko lagina hartu da, non ekipajearen % 92 berreskuratu den.

a. Kalkula bedi laginaren tamaina, jakinda laginean berreskuratuko den ekipajearen proportzioaren eta, oro har, berreskuratuko den ekipajearen proportzioaren arteko diferentzia 0,05 baino txikiagoa izateko probabilitatea 0,995ekoa dela.

b. Beste aireportu batean galdutako 200 ekipaje hartu dira aleatorioki eta haien % 97 berreskuratu da. Bi aireportuetan galdutako 200 ekipaje aukeratzen direla kontuan hartuz, eraiki bedi $p_1 - p_2$ berreskuratutako ekipajearen proportzioen arteko diferentziarako % 99,5eko konfiantza-mailako konfiantza-tartea.

6.6.4. ariketa

Lantegi batean burdinazko eta gomazko bolak egiten dira. Demagun burdinazko bolak egiteko behar den denbora eta gomazko bolak egiteko behar den denbora banaketa normalekoak eta elkarrekiko independenteak direla. Aleatorioki burdinazko 41 bola eta gomazko 31 bola hartu dira; horiek egiteko batez besteko denborak 12,2 ordu eta 9,8 ordu eta desbideratze tipikoak 1,3 ordu eta 0,9 ordu dira, hurrenez hurren.

a. % 90eko konfiantza-mailaz, estima bedi burdinazko bolak egiteko behar den batez besteko denboraren eta gomazko bolak egiteko behar den batez besteko denboraren arteko diferentzia.

b. % 90eko konfiantza-mailaz, estima bedi burdinazko bolak egiteko denboraren bariantzaren eta gomazko bolak egiteko denboraren bariantzaren arteko zatidura.

6.6.5. ariketa

Enpresen irabaziak handitzeko asmoz azken urteetan planifikazio-estrategiak aplikatu dira. Orain planifikazio-estrategia berriak daude. Aleatorioki planifikazio-estrategia zaharrak erabiltzen dituzten 8 enpresa hartu dira, beraien irabazien batez besteko hazkundea (aurreko urteko irabaziekiko) 10 puntukoa eta desbideratze tipikoa 2,8 unitatekoa izanik. Era berean, estrategia berriak aplikatu dituzten 10 enpresa hartu dira, beraien irabazien batez besteko hazkundea (aurreko urteko irabaziekiko) 7 puntukoa eta desbideratze tipikoa 3,2 unitatekoa izanik. Demagun enpresen irabazien hazkundeek banaketa normala dutela eta elkarrekiko independenteak direla. % 99,9ko konfiantza-mailaz, eraiki bedi planifikazio zaharren eta berrien bidez lortutako batez besteko hazkundeen diferentzia estimatzeko konfiantza-tartea:

a. Populazioen bariantzak berdinak direla suposatuz.

b. Populazioen bariantzak ezberdinak direla suposatuz.

6.6.6. ariketa

Aleatorioki aukeratutako hemeretzi altzairuzko xaflaren batez besteko lodiera 8 mm eta desbideratze tipikoa 3 mm dira. Demagun altzairuzko xaflen lodierek banaketa normala dutela.

a. Aurki bedi μ batez besteko lodierarako % 99,8ko konfiantza-mailako konfiantza-tartea.

b. Demagun altzairuzko xaflen lodieraren bariantza 2 mm^2 dela. Zein da kasu honetan altzairuzko xaflen batez besteko lodierarako % 99,8ko konfiantza-mailako konfiantza-tartea?

6.6.7. ariketa

Ura tratatzeko instalazio batean ura tratatua izan aurretik eta tratatua izan ondoren duen kloro-kantitatea neurtu da azken bost hilabeteetan. Hurrengo datuek bost lagin aleatoriotan neurtutako kloro-kantitatea erakusten dute:

Tratatu aurreko kloroa	2,0	2,5	2,2	2,8	2,3
Tratatu ondorengo kloroa	1,5	2,1	2,0	2,4	2,0

Demagun kloro-kantitateak banaketa normala duela.

a. Lor bedi batez besteko kloro-kantitateen arteko diferentzia estimatzeko % 99ko konfiantza-mailako konfiantza-tartea.

b. Ba al dago ezberdintasunik kloroaren batez bestekoen arteko diferentzia estimatzeko konfiantza-tartearen kalkuluan bost hilabetetako datuak hartu ordez hogeita hamabost hilabetetako datuak hartzen badira?

6.6.8. ariketa

Hiri batean aleatorioki hartutako bost egunetan 16, 17, 18, 32 eta 40 l/m^2 euri jaso dira. Demagun euri-kantitateak banaketa normala duela.

a. Kalkula bedi euri-kantitatearen bariantza estimatzeko % 95eko konfiantza-mailako konfiantza-tartea.

b. Demagun batez besteko euri-kantitatea 21 l/m^2 dela. Zein da orduan euri-kantitatearen bariantza estimatzeko % 95eko konfiantza-mailako konfiantza-tartea?

c. Aurreko bi tarteak konparatuz, zer ondoriozta daiteke?

6.6.9. ariketa

Demagun hozkailuen zehaztasuna neurtzeko 0 ºC-tan adierazten duten tenperaturarekiko errorea neurtzen dela. Bi hozkailu-marka aztergai dira. A markako eta B markako hozkailuek 0 ºC-rekiko dituzten erroreak banaketa normalekoak eta elkarrekiko independenteak dira, 0,01 ºC-ko batez bestekoa dutelarik. Aleatorioki A markako 5 hozkailu eta B markako 4 hozkailu hartu dira eta beraien tenperaturak 0ºC-tan hurrengoak izan dira:

A marka	0,01	0,02	–0,01	0,03	–0,02
B marka	0,011	0,009	–0,09	0,01	

a. Kalkula bedi bariantzen arteko zatidurararako % 90eko konfiantza-mailako konfiantza-tartea.

b. Nola ebatziko litzateke aurreko atala, populazioen batez bestekoak ezezagunak balira?

6.6.10. ariketa

Ikerkuntza-proiektuen artean helburua lortzen dutenen proportzioa estimatu nahi da. Horretarako, aleatorioki 126 ikerkuntza-proiektu hartu dira eta horietatik 107 proiektuk helburua lortu dute.

a. % 98,4ko konfiantza-mailaz, onargarria al da helburua lortzen duten proiektuen ehunekoa gutxi gorabehera % 77,3 - % 92,7 delako baieztapena?

b. Zenbat ikerkuntza-proiektu hartu behar dira aleatorioki, % 98,4ko konfiantza-mailaz helburua lortzen dutenen laginaren proportzioaren eta populazioaren proportzioaren arteko diferentzia gehienez % 5ekoa izateko?

6.6.11. ariketa

Bi marka berritako autoen artean, lehenengo bi urteetan konponketak behar dituztenen kopurua jakin nahi da. Aleatorioki hartutako I markako 200 autotik 9 autok eta II markako 300 autotik 15 autok konponketak behar izan dituzte lehenengo bi urteetan.

a. % 95eko konfiantza-mailaz, onartuko al zenuke lehenengo bi urteetan konponketak behar dituzten autoetatik II markakoen proportzioa handiagoa delako baieztapena?

b. Eta % 99ko konfiantza-mailaz?

7. Hipotesi-kontrastea

7.1. OINARRIZKO KONTZEPTUAK

Hipotesi estatistikoa populazioaren ezaugarri bati buruz egiten den baieztapena da. Baieztapen hori egia denetz frogatzeko populazioko elementu bakoitzerako betetzen den ikusi beharko litzateke. Prozesu hau askotan ezinezkoa, oso garestia edo denboran oso luzea da. Horregatik, hipotesia onargarria denetz ikusteko, lagin aleatorio bakun bat eta probarako estatistikoa aukeratuko dira. Estatistikoak laginean balio bat hartuko du, eta laginketa eta estimazioaren emaitzez baliatuz, baieztapena onartu ala errefusatzeko erabaki estatistikoa hartuko da. Erabaki hori zuzena ala okerra izan daiteke, hau da, baieztapena onar daiteke onargarria ez denean edo errefusa daiteke errefusagarria ez denean.

Horrela, **hipotesi-kontrastea** populazioaren ezaugarri bati buruz egindako hipotesia onargarria ala errefusagarria den erabakitzeko prozedura estatistikoa da. Baieztapena populazioaren parametro bati buruzkoa denean, **hipotesi-kontraste parametrikoa** dugu. Kontrastatu nahi den hipotesiari H_0 **hipotesi nulua** deritzo. Hipotesi nulua H_1 **hipotesi alternatiboa**z kontrastatuko da; hipotesi nulua onartzen bada, hipotesi alternatiboa errefusatzen da, eta alderantziz, hipotesi nulua errefusatzen bada, hipotesi alternatiboa onartzen da.

Hipotesi-kontrastea egiteko $x_1, x_2, \ldots, x_n$ lagin aleatorio bakunean oinarritutako funtzio bat erabiliko da. Laginaren menpeko funtzio horri $T(x_1, x_2, \ldots, x_n)$ **kontrasteko estatistikoa** deritzo. Kontrasteko estatistikoak hartzen duen balioa S_0 **onarpen-eremuan** badago, hipotesi nulua onar daiteke. Kontrasteko estatistikoaren balioa, aldiz, S_1 **eremu kritikoan** badago, orduan hipotesi nulua errefusatuko da.

Bukaerako erabakia hartzean erroreak egin daitezke. Hipotesi nulua egia denean hipotesi nulua errefusatzen bada, orduan **I. motako errorea** egin da. Bestalde, hipotesi nulua gezurra denean hipotesi nulua onartzen bada, **II. motako errorea** egin da. I. motako errorea egiteko probabilitateari α **adierazgarritasun-maila** deritzo eta II. motako errorea betetzeko probabilitateari β izena ematen zaio. **Kontrastearen ahalmena** edo **potentzia** adierazteko $1 - \beta$ notazioa erabiltzen da, eta hipotesi nulua gezurra denean hipotesi nulua errefusatzeko probabilitatea adierazten du.

$$\alpha = P(E_I) = P(H_0 \text{ errefusatu} \mid H_0 \text{ egia})$$

$$\beta = P(E_{II}) = P(H_0 \text{ onartu} \mid H_0 \text{ gezurra})$$

$$1 - \beta = P(H_0 \text{ errefusatu} \mid H_0 \text{ gezurra})$$

Hipotesi-kontrastean komeni da α adierazgarritasun-maila aldez aurretik finkatzea eta $1-\beta$ potentzia maximoa (edo bigarren motako errorea egiteko β probabilitate minimoa) duen kontrastea egitea.

Bestalde, ***p*-balioa** edo α_c **adierazgarritasun-maila kritikoa** estatistikoaren balioa onarpen-eremuan dagoeneko adierazgarritasun-maila maximoa da.

$$\alpha_c = \max\{\alpha / T(x_1,x_2,\ldots,x_n) \in S_0\}$$

Beraz, hipotesi-kontrastea egiteko α_0 adierazgarritasun-maila hartzen bada, $\alpha_0 > \alpha_c$ izanik, orduan $T(x_1, x_2, \ldots, x_n) \notin S_0$, eta ondorioz, H_0 errefusatuko da α_0 adierazgarritasun-mailaz.

Bukatzeko, eremu kritikoaren arabera, bi motatako hipotesi-kontrasteak daude: **alde bakarreko hipotesi-kontrastea**n eremu kritikoa zuzen errealeko multzo bakarra da eta **bi aldeko hipotesi-kontrastea**n eremu kritikoa zuzen errealeko bi multzo disjuntuz osaturik dago.

7.2. HIPOTESI-KONTRASTEAREN URRATSAK

Hipotesi-kontrastearen urratsak honela labur daitezke:

1. *Hipotesi nulua eta hipotesi alternatiboa zehaztu.*
   Lehenengo urrats honetan kontrastatu nahi diren hipotesiak zehaztuko dira eta aztergai den populazioaren parametroari buruzko baieztapenak ezarriko dira.

2. *Adierazgarritasun-maila finkatu.*
   I. motako errorea betetzeko α probabilitatea edo adierazgarritasun-maila aldez aurretik finkatuko da. Hipotesi-kontrastea adierazgarria izateko hartzen diren adierazgarritasun-mailak, oro har, 0,005, 0,01 edo 0,05 dira.

3. *Probarako estatistiko egokia hartu.*
   Lagin aleatorio bakuna hartu ondoren, hipotesi nulua egia denetz erabakitzeko beharrezkoa den probarako estatistikoa aukeratu behar da. Probarako estatistikoa izango da hipotesietan adierazitako populazioko parametroaren estimatzailearen menpeko funtzioa. Lagineko datuak probarako estatistikoan ordezkatuko dira eta era horretan estatistikoak hartzen duen balioa kalkulatuko da.

4. *Eremu kritikoa edota onarpen-eremua zehaztu.*
   Hipotesi nulua egia denean estatistikoaren banaketa ezaguna bada, orduan eremu kritikoa edota onarpen-eremua zehatz daiteke.

5. *Erabaki estatistikoa hartu.*
   Lagineko datuak ordezkatuz lortutako probarako estatistikoaren balioa eremu kritikoan badago, hipotesi nulua errefusatuko da, α adierazgarritasun-mailaz. Kontrako kasuan, probarako estatistikoaren balioa onarpen-eremuan badago, hipotesi nulua onartuko da, α adierazgarritasun-mailaz.

7.3. ZENBAIT HIPOTESI-KONTRASTE

Atal honetan zenbait hipotesi-kontrastetarako hipotesien araberako estatistikoak eta onarpen-eremuak laburtu dira, non hipotesi-kontraste guztiak α adierazgarritasun-mailaz gauzatuko diren.

7.3.1. Banaketa normalaren batez bestekorako hipotesi-kontrastea

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
H_o: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ H_o: $\mu \geq \mu_0$, H_1: $\mu < \mu_0$ H_o: $\mu \leq \mu_0$, H_1: $\mu > \mu_0$ (σ ezaguna)	$z = \dfrac{\overline{x} - \mu_o}{\sigma / \sqrt{n}}$	$[-z_{\alpha/2}, z_{\alpha/2}]$ $[-z_{\alpha}, \infty)$ $(-\infty, z_{\alpha}]$
H_o: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ H_o: $\mu \geq \mu_0$, H_1: $\mu < \mu_0$ H_o: $\mu \leq \mu_0$, H_1: $\mu > \mu_0$ (σ ezezaguna, $n \geq 30$)	$z = \dfrac{\overline{x} - \mu_o}{S / \sqrt{n}}$	$[-z_{\alpha/2}, z_{\alpha/2}]$ $[-z_{\alpha}, \infty)$ $(-\infty, z_{\alpha}]$
H_o: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$ H_o: $\mu \geq \mu_0$, H_1: $\mu < \mu_0$ H_o: $\mu \leq \mu_0$, H_1: $\mu > \mu_0$ (σ ezezaguna, $n < 30$)	$t = \dfrac{\overline{x} - \mu_o}{S / \sqrt{n}}$	$[-t_{\alpha/2;n-1}, t_{\alpha/2;n-1}]$ $[-t_{\alpha;n-1}, \infty)$ $(-\infty, t_{\alpha;n-1}]$

7.3.2. Bi banaketa normal independenteren batez bestekoen arteko diferentziarako hipotesi-kontrastea

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 \neq \mu_2$ H_0: $\mu_1 \geq \mu_2$, H_1: $\mu_1 < \mu_2$ H_0: $\mu_1 \leq \mu_2$, H_1: $\mu_1 > \mu_2$ (σ_1, σ_2 ezagunak)	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$[-z_{\alpha/2}, z_{\alpha/2}]$ $[-z_\alpha, \infty)$ $(-\infty, z_\alpha]$
H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 \neq \mu_2$ H_0: $\mu_1 \geq \mu_2$, H_1: $\mu_1 < \mu_2$ H_0: $\mu_1 \leq \mu_2$, H_1: $\mu_1 > \mu_2$ ($\sigma_1 \neq \sigma_2$ ezezagunak) n_1, $n_2 < 30$	$t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	$[-t_{\alpha/2;v}, t_{\alpha/2;v}]$ $[-t_{\alpha;v}, \infty)$ $(-\infty, t_{\alpha;v}]$ $v = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1+1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2+1}} - 2$
H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 \neq \mu_2$ H_0: $\mu_1 \geq \mu_2$, H_1: $\mu_1 < \mu_2$ H_0: $\mu_1 \leq \mu_2$, H_1: $\mu_1 > \mu_2$ ($\sigma_1 = \sigma_2$ ezezagunak) n_1, $n_2 < 30$	$t = \dfrac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$	$[-t_{\alpha/2;n_1+n_2-2}, t_{\alpha/2;n_1+n_2-2}]$ $[-t_{\alpha;n_1+n_2-2}, \infty)$ $(-\infty, t_{\alpha;n_1+n_2-2}]$

7.3.3. Populazio normalaren bariantzarako hipotesi-kontrastea

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
$H_0:\sigma^2=\sigma_0^2,\ H_1:\sigma^2\neq\sigma_0^2$ $H_0:\sigma^2\geq\sigma_0^2,\ H_1:\sigma^2<\sigma_0^2$ $H_0:\sigma^2\leq\sigma_0^2,\ H_1:\sigma^2>\sigma_0^2$ (μ ezezaguna)	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$	$\left[\chi^2_{1-\alpha/2;n-1},\chi^2_{\alpha/2;n-1}\right]$ $\left[\chi^2_{1-\alpha;n-1},\infty\right)$ $\left[0,\chi^2_{\alpha;n-1}\right]$
$H_0:\sigma^2=\sigma_0^2,\ H_1:\sigma^2\neq\sigma_0^2$ $H_0:\sigma^2\geq\sigma_0^2,\ H_1:\sigma^2<\sigma_0^2$ $H_0:\sigma^2\leq\sigma_0^2,\ H_1:\sigma^2>\sigma_0^2$ (μ ezaguna)	$\chi^2=\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma_0^2}$	$\left[\chi^2_{1-\alpha/2;n},\chi^2_{\alpha/2;n}\right]$ $\left[\chi^2_{1-\alpha;n},\infty\right)$ $\left[0,\chi^2_{\alpha;n}\right]$

7.3.4. Banaketa normaleko bi populazio independenteren bariantzen arteko zatidurarako hipotesi-kontrastea

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
$H_0:\sigma_1^2=\sigma_2^2,\ H_1:\sigma_1^2\neq\sigma_2^2$ $H_0:\sigma_1^2\geq\sigma_2^2,\ H_1:\sigma_1^2<\sigma_2^2$ $H_0:\sigma_1^2\leq\sigma_2^2,\ H_1:\sigma_1^2>\sigma_2^2$ (μ_1, μ_2 ezezaguna)	$F=\dfrac{S_1^2}{S_2^2}$	$\left[F_{1-\alpha/2;n_1-1,n_2-1},F_{\alpha/2;n_1-1,n_2-1}\right]$ $\left[F_{1-\alpha;n_1-1,n_2-1},\infty\right)$ $\left[0,F_{\alpha;n_1-1,n_2-1}\right]$
$H_0:\sigma_1^2=\sigma_2^2,\ H_1:\sigma_1^2\neq\sigma_2^2$ $H_0:\sigma_1^2\geq\sigma_2^2,\ H_1:\sigma_1^2<\sigma_2^2$ $H_0:\sigma_1^2\leq\sigma_2^2,\ H_1:\sigma_1^2>\sigma_2^2$ (μ_1, μ_2 ezezaguna)	$F=\dfrac{\dfrac{\sum_{i=1}^{n_1}(x_{1i}-\mu_1)^2}{n_1}}{\dfrac{\sum_{i=1}^{n_2}(x_{2i}-\mu_2)^2}{n_2}}$	$\left[F_{1-\alpha/2;n_1,n_2},F_{\alpha/2;n_1,n_2}\right]$ $\left[F_{1-\alpha;n_1,n_2},\infty\right)$ $\left[0,F_{\alpha;n_1,n_2}\right]$

7.3.5. *Banaketa binomialaren parametrorako hipotesi-kontrastea* ($n \geq 30$)

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
H_0: $p = p_0$, H_1: $p \neq p_0$	$z = \dfrac{\hat{p} - p_o}{\sqrt{\dfrac{p_o q_o}{n}}}$	$[-z_{\alpha/2}\,,\, z_{\alpha/2}]$
H_0: $p \geq p_0$, H_1: $p < p_0$		$[-z_{\alpha}\,,\, \infty)$
H_0: $p \leq p_0$, H_1: $p > p_0$		$(-\infty\,,\, z_{\alpha}]$

7.3.6. *Bi banaketa binomial independenteren parametroen arteko diferentziarako hipotesi-kontrastea* ($n_1 \geq 30$, $n_2 \geq 30$)

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
H_0: $p_1 = p_2$, H_1: $p_1 \neq p_2$	$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}}$	$[-z_{\alpha/2}\,,\, z_{\alpha/2}]$
H_0: $p_1 \geq p_2$, H_1: $p_1 < p_2$		$[-z_{\alpha}\,,\, \infty)$
H_0: $p_1 \leq p_2$, H_1: $p_1 > p_2$		$(-\infty\,,\, z_{\alpha}]$

7.3.7. *Bi banaketa normal ez independenteren batez bestekoen arteko diferentziarako hipotesi-kontrastea*

Hipotesi-kontrasteak	*Estatistikoa*	*Onarpen-eremua*
H_0: $\mu_D = \mu_0$, H_1: $\mu_D \neq \mu_0$	$t = \dfrac{\overline{d} - \mu_o}{S / \sqrt{n}}$	$[-t_{\alpha/2;n-1}, t_{\alpha/2;n-1}]$
H_0: $\mu_D \geq \mu_0$, H_1: $\mu_D < \mu_0$		$[-t_{\alpha;n-1}\,,\, \infty)$
H_0: $\mu_D \leq \mu_0$, H_1: $\mu_D > \mu_0$		$(-\infty\,,\, t_{\alpha;n-1}]$
$D = X - Y$: $N(\mu_D, \sigma_D)$ σ_D ezezaguna		

Oharra: $\overline{d} = \sum_{i=1}^{n} \dfrac{d_i}{n} = \sum_{i=1}^{n} \dfrac{(x_i - y_i)}{n}, \quad S^2 = \sum_{i=1}^{n} \dfrac{(d_i - \overline{d})^2}{n-1}.$

7.4. ARIKETA EBATZIAK

7.4.1. ariketa

Bizikletetarako osagaiak egiten dituen enpresa batek urtero batez beste 60 eskaera jasotzen dituela dio. Baieztapen hori onargarria denetz ikusteko, aleatorioki hartutako lau urtetako eskaerak zenbatu dira eta batez beste 50 eskaera izan dira. Demagun bizikletetarako osagaiak egiten dituen enpresak jasotzen dituen eskaera-kopuruak banaketa normala duela, desbideratze tipikoa 6 eskaerakoa izanik. % 5eko adierazgarritasun-mailaz, zer esango zenuke enpresaren baieztapenari buruz?

Ebazpena

Biz X = “bizikletetarako osagaiak egiten dituen enpresak jasotzen duen eskaera-kopurua” aldagai aleatorioa, $\sigma = 6$ desbideratze tipikodun banaketa normalekoa. Kasu honetan bariantza ezaguneko banaketa normala duen populazioaren batez bestekorako alde biko hipotesi-kontrastea egin behar da.

Hauek dira hipotesi-nulua eta hipotesi alternatiboa, hurrenez hurren:

$H_0 : \mu = 60$ hipotesti nulua
$H_1 : \mu \neq 60$ hipotesi alternatiboa.

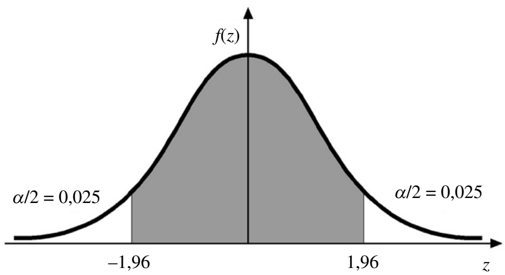
Adierazgarritasun-maila $\alpha = 0{,}05$ da.

Probarako estatistikoa

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

da, eta kasu honetan $z = \dfrac{50 - 60}{6/\sqrt{4}} = -3{,}33$ balioa hartzen du.

Onarpen-eremua $S_0 = [-z_{\alpha/2}, z_{\alpha/2}] = [-z_{0{,}025}, z_{0{,}025}] = [-1{,}96\ ,\ 1{,}96]$ da.



7.1. irudia.

Estatistikoaren balioa onarpen-eremutik kanpo dagoenez, orduan hipotesi nulua errefusatuko da, % 5eko adierazgarritasun-mailaz.

7.4.2. ariketa

Torno metaliko batek 0,5 cm-ko batez besteko lodierako piezak egiten ditu. Kalitate-kontroleko arduradunak piezen lodiera orokorrean 0,5 cm baino txikiagoa bada, prozesua geldiarazten du. Demagun piezen lodiera banaketa normalekoa dela.

a. Aleatorioki hartutako 200 piezaren lodierak neurtu dira, batez besteko lodiera 0,48 cm eta desbideratze tipikoa 0,2 cm direlarik. % 1eko adierazgarritasun-mailaz, erabaki ezazu prozesua gelditu ala ez.

b. Piezen batez besteko lodiera gutxienez 0,5 cm-koa bada, zein da erabaki ez-zuzena hartzeko probabilitatea?

Ebazpena

Enuntziatuaren arabera, X = "torno metalikoak egiten dituen piezen lodiera" aldagai aleatorioak banaketa normala duela suposatuko da.

a. Kalitate-kontroleko arduradunak $\mu < 0{,}5$ bada, prozesua gelditzea erabakitzen du. Beraz, hipotesi hau frogatzeko hurrengo hipotesi-kontrastea egin daiteke:

$H_0 : \mu \geq 0{,}5$ hipotesi nulua
$H_1 : \mu < 0{,}5$ hipotesi alternatiboa.

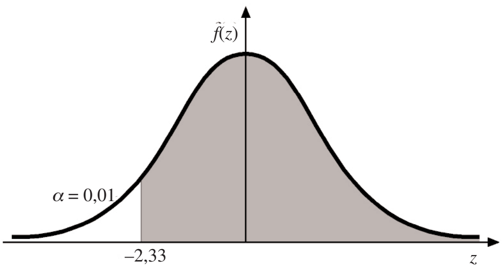
Adierazgarritasun-maila: $\alpha = 0{,}01$.

Probarako estatistikoa: $z = \dfrac{\bar{x} - \mu_0}{S / \sqrt{n}}$.

Kasu honetan probarako estatistikoaren balioa hauxe da:

$$z = \frac{\bar{x} - \mu_0}{S / \sqrt{n}} = \frac{\bar{x} - \mu_0}{s / \sqrt{n-1}} = \frac{0{,}48 - 0{,}5}{0{,}2 / \sqrt{199}} = -1{,}41 \ .$$

Onarpen-eremua: $S_0 = [-z_\alpha, \infty) = [-z_{0,01}, \infty) = [-2{,}33 \ , \infty)$.



7.2. irudia.

Probarako estatistikoaren balioa $-1{,}41 \in S_0$ dagoenez, orduan % 1eko adierazgarritasun-mailaz hipotesi nulua onartuko genuke. Hau da, ez genuke geldiaraziko prozesua.

b. Definizioa aplikatuz, atal honetan eskatzen dena adierazgarritasun-mailaren balioa da.

$P(\text{erabaki ez-zuzena} \mid \mu \geq 0{,}5) = P(H_0 \text{ ez onartu} \mid H_0 \text{ egia}) = P(E_I) = \alpha = 0{,}01.$

7.4.3. ariketa

X aldagai aleatorioaren desbideratze tipikoa 9 da. H_0: $\mu = 6$ hipotesi nulua H_1: $\mu = 3{,}8$ hipotesi alternatiboaren aurka kontrastatzeko, 81 tamainako lagin aleatorio bakuna hartu da. Kalkula bedi hipotesi-kontrastearen ahalmena, jakinda laginaren batez bestekoaren eremu kritikoa (1,4) tartea dela.

Ebazpena

Definizioz, hipotesi-kontrastearen $1 - \beta$ ahalmena, hipotesi nulua gezurra denean, hipotesi nulua errefusatzeko probabilitatea da.

$$1 - \beta = P(H_0 \text{ errefusatu} \mid H_0 \text{ gezurra}).$$

Kasu honetan,

$$1 - \beta = P(\bar{x} \in S_1 = (1,4) \mid H_1 : \mu = 3{,}8 \text{ egia})$$

probabilitatearen balioa kalkulatu behar da. Laginketaren gaian aztertutako emaitzen arabera, limite zentralaren teoremak dioenez, σ^2 bariantza ezaguneko populaziotik $n \geq 30$ tamainako lagin aleatorio bakuna hartzen bada, $\overline{X}$ aldagai aleatorioaren banaketa $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ banaketaz hurbil daiteke.

Orduan, ahalmenaren balioa zehazteko, $\mu = 3{,}8$, $\sigma = 9$ eta $n = 81$ balioak ordezkatuko dira eta $\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N(3{,}8,\ 1)$ dela kontuan hartuko da.

$$\begin{aligned}
1 - \beta &= P(1 < \bar{x} < 4 \mid H_1 : \mu = 3{,}8 \text{ egia}) = \\
&= P\left(\frac{1 - 3{,}8}{1} < \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < \frac{4 - 3{,}8}{1} \middle| \mu = 3{,}8\right) = \\
&= P\left(-2{,}8 < Z < 0{,}2 \mid \mu = 3{,}8\right) = \\
&= \phi(0{,}2) - \phi(-2{,}8) = \phi(0{,}2) - [1 - \phi(2{,}8)] = \\
&= 0{,}5793 - 1 + 0{,}9974 = 0{,}5767.
\end{aligned}$$

7.4.4. ariketa

Bideoak ekoizten dituen enpresa batek bi mikrozirkuitu-mota diseinatzen ditu. Bi mikrozirkuitu-mota horien fluxu-kantitateak banaketa normalekoak dira, desbideratze tipikoak 3 Wb eta 6 Wb izanik, hurrenez hurren. Bi mikrozirkuituetako fluxu-kantitateak elkarrekiko independenteak direla suposatuko da.

Ekoizlearen arabera, I. motako mikrozirkuituen fluxu-kantitatearen batez bestekoa gutxienez II. motako mikrozirkuituen fluxu-kantitatearen batez bestekoaren berdina da. Aleatorioki hamaseina mikrozirkuitu hartu dira eta honako emaitzak lortu:

I mota: $n_1 = 16$ mikrozirkuitu, $\bar{x}_1 = 12$ Wb, $s_1 = 3{,}5$ Wb.

II mota: $n_2 = 16$ mikrozirkuitu, $\bar{x}_2 = 15$ Wb, $s_2 = 5{,}0$ Wb.

Zer esan daiteke ekoizleak dioenari buruz, % 1,5eko adierazgarritasun-mailaz?

Ebazpena

Bira X_1 = "I mikrozirkuituaren fluxu-kantitatea" eta X_2 = "II mikrozirkuituaren fluxu-kantitatea", X_1: $N(\mu_1,\ \sigma_1 = 3)$ eta X_2: $N(\mu_2,\ \sigma_2 = 6)$ aldagai aleatorio independenteak direlarik. Populazio bakoitzetik $n_1 = 16 = n_2$ tamainako lagin aleatorio bakunak hartu dira, hurrenez hurren.

Baieztapenari buruzko erabakia hartzeko, honako hipotesi-kontrastea egin daiteke:

H_0: $\mu_1 \geq \mu_2$ hipotesi nulua
H_1: $\mu_1 < \mu_2$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}015$.

Probarako estatistikoa: $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$.

Probarako estatistikoaren balioa: $\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{12 - 15}{\sqrt{\dfrac{9}{16} + \dfrac{36}{16}}} = -1,79.$

Onarpen-eremua: $S_0 = [-z_\alpha, \infty) = [-z_{0,015}, \infty) = [-2{,}17, \infty)$.

Estatistikoaren balioa onarpen-eremuan dagoenez, H_0: $\mu_1 \geq \mu_2$ hipotesi nulua onartuko genuke % 1,5eko adierazgarritasun-mailaz.

7.4.5. ariketa

Alkoholimetroen funtzionamendua neurtzeko, aleatorioki 400 alkoholimetro hartu dira eta 328 alkoholimetrok era egokian funtzionatu dute. Alkoholimetroen fabrikatzaileak dioenez, "alkoholimetroen % 90ek funtzionamendu egokia du".

a. % 2,5eko adierazgarritasun-mailaz, zer esango zenuke fabrikatzailearen baieztapenari buruz?

b. % 2,5eko adierazgarritasun-mailaz, zer esango zenuke fabrikatzaileak "alkoholimetroen % 90ek baino gehiagok du funtzionamendu egokia" esango balu?

c. Aurreko bi atalen emaitzetatik, zer ondoriozta daiteke?

Ebazpena

Banaketa binomialeko p parametroari buruzko baieztapena egia denetz ikusi nahi da, p = "era egokian funtzionatzen duen alkoholimetroen proportzioa" delarik. Hartu den lagin aleatorio bakunean era egokian funtzionatzen duten alkoholimetroen proportzioa $\hat{p} = \dfrac{328}{400} = 0,82$ da, non $n = 400$ den.

a. Enuntziatuko baieztapenari buruzko erabakia hartzeko, hurrengo hipotesi-kontrastea egin daiteke:

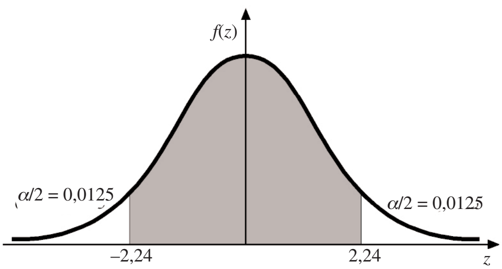
H_0: $p = 0,90$ hipotesi nulua
H_1: $p \neq 0,90$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0,025$.

Probarako estatistikoa: $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$.

Probarako estatistikoaren balioa: $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} = \dfrac{0,82 - 0,90}{\sqrt{\dfrac{0,90 \cdot 0,10}{400}}} = -5,3$.

Onarpen-eremua: $S_0 = [-z_{\alpha/2}, z_{\alpha/2}] = [-z_{0,0125}, z_{0,0125}] = [-2,24, 2,24]$.



7.3. irudia.

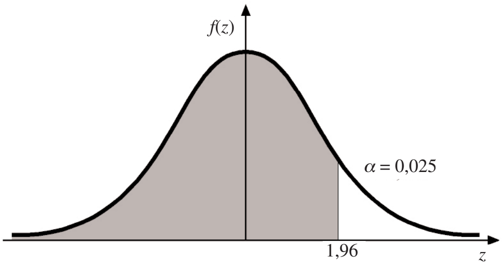
Estatistikoaren balioa onarpen-eremutik kanpo dagoenez, $-5{,}3 \notin S_0$ alegia, ez genuke onartuko H_0: $p = 0{,}90$ hipotesi nulua % 2,5eko adierazgarritasun-mailaz. Ondorioz, ez genuke onartuko fabrikatzailearen baieztapena.

b. Fabrikatzailearen beste hipotesia $p > 0{,}90$ da. Oraingoan eskuineko hurrengo hipotesi-kontrastea egingo da:

H_0: $p \leq 0{,}90$ hipotesi nulua
H_1: $p > 0{,}90$ hipotesi alternatiboa.

Adierazgarritasun-maila eta probarako estatistikoaren balioa aurreko ataleko berberak dira, baina onarpen-eremua beste bat da.

Onarpen-eremua: $S_0 = (-\infty, z_\alpha] = (-\infty, z_{0,025}] = (-\infty, 1{,}96]$.



7.4. irudia.

Estatistikoaren balioa onarpen-eremuan dagoenez, $-5{,}3 \in S_0$, ez genuke onartuko H_1: $p > 0{,}90$ hipotesi alternatiboa % 2,5eko adierazgarritasun-mailaz.

c. Aurreko bi ataletatik, % 2,5eko adierazgarritasun-mailaz onar daitekeen hipotesia hauxe da: "funtzionamendu zuzena duen alkoholimetroen portzentajea % 90 baino txikiagoa da".

7.4.6. ariketa

Bi hiritan aleatorioki aukeratutako bost egunetako euri-kantitateak (l/m^2) hauek izan dira:

A hiria	17	18	19	22	23
B hiria	17,5	19	18,7	21	22,7

Normalitatearen hipotesia suposatuz eta % 5eko adierazgarritasun-mailaz, egiatzat hartuko al zenuke A hiriko euri-kantitatearen bariantza B hiriko euri-kantitatearen bariantza baino handiagoa izatea?

Ebazpena

Bira X_1 = "A hiriko euri-kantitatea" eta X_2 = "B hiriko euri-kantitatea" banaketa normaleko bi aldagai aleatorio. Demagun aldagai aleatorio hauek independenteak direla.

Aldez aurretik, aleatorioki hartutako bost egunetako euri-kantitateetarako batez bestekoa eta kuasibariantza kalkulatuko dira.

$$\bar{x}_1 = \frac{99}{5} = 19{,}80 \text{ l/m}^2, \quad S_1^2 = 6{,}70 \text{ (l/m}^2)^2$$

$$\bar{x}_2 = \frac{98{,}9}{5} = 19{,}78 \text{ l/m}^2, \quad S_2^2 = 4{,}25 \text{ (l/m}^2)^2$$

Egingo den hipotesi-kontrastea hurrengoa da:

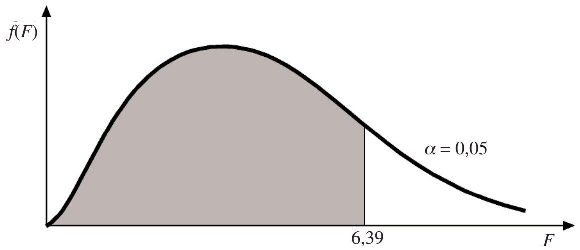
H_0: $\sigma_1^2 \leq \sigma_2^2$ hipotesi nulua

H_1: $\sigma_1^2 > \sigma_2^2$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}05$.

Probarako estatistikoa: $\dfrac{S_1^2}{S_2^2}$.

Onarpen-eremua: $S_0 = [0, F_{\alpha;n_1-1,n_2-1}] = [0, F_{0{,}05;4,4}] = [0,\ 6{,}39]$.



7.5. irudia.

Kasu honetan estatistikoaren balioa $\dfrac{S_1^2}{S_2^2} = \dfrac{6{,}70}{4{,}25} = 1{,}58$ onarpen-eremuan dagoenez, $H_0 : \sigma_1^2 \leq \sigma_2^2$ hipotesi nulua onartuko da % 5eko adierazgarritasun-mailaz.

7.4.7. ariketa

Zenbait ikerketaren arabera, airearen kalitatearen baldintza arruntetan karbono monoxidoaren mailen desbideratze tipikoak 0,1 μg/m³ baino txikiagoa izan behar du. Karbono monoxidoaren orduko mailek estazio ezberdinetan banaketa normala dutela suposatu da. Ikertzaileek egunean zehar Albiako estazioan behatutako karbono monoxidoaren mailak neurtu dituzte eta hamar ordutako ondoko mailak jaso dira:

7,96 7,90 7,98 8,01 7,97 7,96 8,03 8,02 8,04 8,02

a. % 5eko adierazgarritasun-mailaz, esan al daiteke karbono monoxidoaren mailen desbideratze tipikoa 0,1 μg/m³ baino txikiagoa dela Albiako estazioan?

b. Demagun beste kasu batean Mazarredoko estazioko karbono monoxidoaren 25 ordutako mailen balioak jaso direla, horien kuasibariantza $(0{,}039)^2$ $(\mu g/m^3)^2$ izanik. % 10eko adierazgarritasun-mailaz, zer esango zenuke hurrengo baieztapenari buruz: "Albiako eta Mazarredoko karbono monoxidoaren mailen desbideratze tipikoak ez dira berdinak".

Ebazpena

Hasteko, kalkula ditzagun Albiako datuei buruzko batez bestekoa eta desbideratze tipikoa:

$$\bar{x}_1 = 7{,}989 \ \mu g/m^3, \quad s_1^2 = 0{,}00167 \ (\mu g/m^3)^2, \quad S_1^2 = 0{,}00186 \ (\mu g/m^3)^2.$$

Bira X_1 = "karbono monoxidoaren orduko maila, Albiako estazioan" eta X_2 = "karbono monoxidoaren orduko maila, Mazarredoko estazioan" banaketa normaleko bi aldagai aleatorio independente.

a. Plantea dezagun hurrengo hipotesi-kontrastea:

H_0: $\sigma_1 \geq 0{,}1$ hipotesi nulua
H_1: $\sigma_1 < 0{,}1$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}05$.

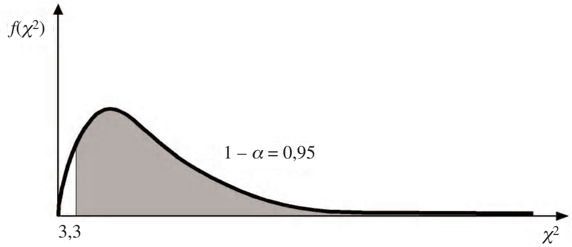
Probarako estatistikoa:

$$\chi^2 = \frac{(n_1 - 1)S_1^2}{\sigma_0^2}.$$

Kasu honetan hartzen duen balioa hauxe da:

$$\chi^2 = \frac{(10-1)0{,}00186}{(0{,}1)^2} = 1{,}674.$$

Onarpen-eremua: $S_0 = [\chi^2_{1-\alpha;n-1}, \ \infty] = [\chi^2_{0{,}95;9}, \ \infty] = [3{,}33, \ \infty)$.



7.6. irudia.

Orduan, probarako estatistikoaren balioa, 1,674 alegia, onarpen-eremutik kanpo dagoenez, ondorioz H_1: $\sigma_1 < 0{,}1$ hipotesi alternatiboa onar daiteke % 5eko adierazgarritasun-mailaz.

b. Oraingoan, hurrengo hipotesi-kontrastea egingo da:

$$H_0: \ \sigma_1^2 = \sigma_2^2 \ \text{hipotesi nulua}$$

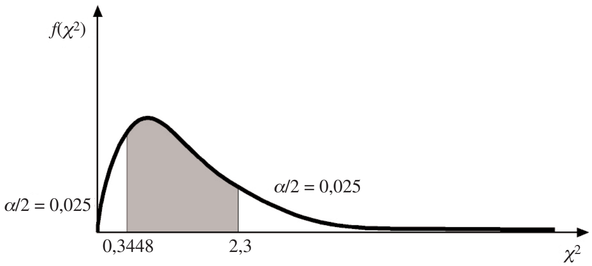
$$H_1: \ \sigma_1^2 \neq \sigma_2^2 \ \text{hipotesi alternatiboa.}$$

Adierazgarritasun-maila $\alpha = 0{,}10$ da.

Probarako estatistikoak hartzen duen balioa $\dfrac{S_1^2}{S_2^2} = \dfrac{0{,}00186}{(0{,}039)^2} = 1{,}22$ da.

Onarpen-eremua:

$$S_0 = [F_{1-\alpha/2;n_1-1,n_2-1},\ F_{\alpha/2;n_1-1,n_2-1}] = [F_{0,95;9,24},\ F_{0,05;9,24}] = [0{,}3448,\ 2{,}3].$$



7.7. irudia.

Estatistikoaren balioa onarpen-eremuan dagoenez, hots, 1,22 betetzen denez, Mazarredoko eta Albiako monoxido karbonoaren mailen bariantzak berdinak direla onar daiteke, % 10eko adierazgarritasun-mailaz. Beste hitz batzuetan, Mazarredoko eta Albiako monoxido karbonoaren mailen desbideratze tipikoak ezberdinak direlako hipotesia errefusatuko genuke, % 10eko adierazgarritasun-mailaz.

7.4.8. ariketa

Enpresa batek lan egiteko bi metodo berriren ekoizpena neurtu nahi du. Horretarako era aleatorioan 12 eta 8 langile aukeratu dira: lehenengoek I metodoa eta bigarrenek II metodoa erabiliko dute. Aste baten buruan ondoko emaitzak jaso dira:

I metodoa: $\bar{x}_1 = 4.300$ pieza, $S_1 = 1.100$ pieza, $n_1 = 12$ langile.

II metodoa: $\bar{x}_2 = 3.600$ pieza, $S_2 = 1.900$ pieza, $n_2 = 8$ langile.

Normaltasunaren hipotesia suposatuz eta % 10eko adierazgarritasun-mailaz, onar al daiteke I metodoaren bidez II metodoaren bidez baino gehiago ekoizten dela dioen hipotesia?

Ebazpena

Bira X_1 = "I metodoa erabiliz ekoitzitako pieza-kopurua" eta X_2 = " II metodoa erabiliz ekoitzitako pieza-kopurua" aldagai aleatorio independenteak, banaketa normalekoak. Pieza gehiago zein metodok ekoizten dituen ebatzi baino lehen, bariantzak berdinak direnetz zehaztu behar da. Horretarako hurrengo kontrastea egingo da.

i. Alde biko kontrastea:

H_0: $\sigma_1^2 = \sigma_2^2$ hipotesi nulua

H_1: $\sigma_1^2 \neq \sigma_2^2$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}1$.

Probarako estatistikoa: $F = \dfrac{S_1^2}{S_2^2}$.

Kasu honetan probarako estatistikoaren balioa $\dfrac{S_1^2}{S_2^2} = \dfrac{(1.100)^2}{(1.900)^2} = 0{,}335$ da.

Onarpen-eremua: $S_0 = [F_{1-\alpha/2;n_1-1,n_2-1},\ F_{\alpha/2;n_1-1,n_2-1}] = [F_{0,95;11,7},\ F_{0,05;11,7}]$.

Snedecor-en taulan kontsultatuz, $F_{0,05;11,7} = 3{,}6$ lor daiteke. Bestalde, Snedecor-en banaketaren propietatea aplikatuz $F_{0,95;11,7} = \dfrac{1}{F_{0,05;7,11}} = \dfrac{1}{3{,}01} = 0{,}332$ balioa zehaztuko da. Bi balio horiek onarpen-eremuan ordezkatuko dira. Horrela, onarpen-eremua $S_0 = [0{,}332,\ 3{,}6]$ da.

Estatistikoaren balioa onarpen-eremuan dagoenez, $0{,}335 \in [0{,}332,\ 3{,}6]$ alegia, orduan % 10eko adierazgarritasun-mailaz hipotesi nulua onar daiteke. Hau da, $\sigma_1^2 = \sigma_2^2$ bariantzak berdinak direla suposatuko da.

Baldintza hauetan, bi populazio normal, independente eta bariantza ezezagun berdinetakoak ditugunean, % 10eko adierazgarritasun-mailaz, hurrengo hipotesi-kontrastea egingo dugu:

ii. Ezkerretiko hipotesi-kontrastea:

H_0: $\mu_1 \leq \mu_2$ hipotesi nulua

H_1: $\mu_1 > \mu_2$ hipotesi alternatiboa.

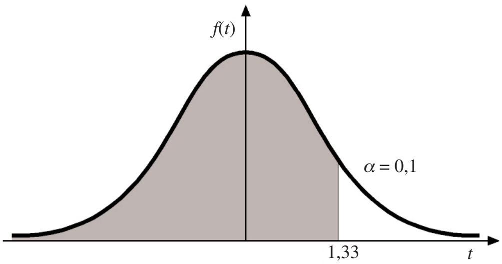
Adierazgarritasun-maila: $\alpha = 0{,}1$.

Probarako estatistikoa: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$.

Estistikoaren balioa hauxe da:

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} = \frac{4.300 - 3.600}{\sqrt{\dfrac{11\cdot(1.100)^2 + 7\cdot(1.900)^2}{12+8-2}}\sqrt{\dfrac{1}{12}+\dfrac{1}{8}}} = 1{,}05.$$

Onarpen-eremua: $S_0 = (-\infty, t_{\alpha;n_1+n_2-2}] = (-\infty, t_{0{,}1;18}] = (-\infty, 1{,}33]$.



7.8. irudia.

Estatistikoaren balioa 1,05 ∈ (–∞, 1,33] dagoenez, H_0: $\mu_1 \leq \mu_2$ hipotesi nulua onar daiteke, % 1eko adierazgarritasun-mailaz. Beraz, ez genuke onartuko enuntziatuko hipotesia.

7.4.9. ariketa

Esperimentu kimiko bat egiteko nahasketan, esperimentuaren hasieran eta bukaeran kloro-kantitatea (litro) aztergai da. Aleatorioki sei nahasketa hartu dira eta dagozkien kloro-kantitateak neurtu dira.

Kloroa, hasieran	7,0	9,1	7,8	8,1	7,2	9,0
Kloroa, bukaeran	7,5	8,7	7,6	8,4	7,5	9,1

Normaltasunaren hipotesia suposatuz eta % 2ko adierazgarritasun-mailaz, onartuko al zenuke esperimentu kimikoaren hasieran eta bukaeran kloro-kantitatea berdina delako hipotesia?

Ebazpena

Bira X_1= “kloro-kantitatea esperimentu kimikoaren hasieran” eta X_2 = “kloro-kantitatea esperimentu kimikoaren bukaeran” aldagai aleatorio normalak. Aldagai aleatorio hauek ez dira independenteak, esperimentu hasierako kloroak bukaerako kloro-kantitatean eragingo baitu.

Emandako baieztapena onargarria denetz ikusteko, hurrengo hipotesi-kontrastea egingo da:

H_0: $\mu_D = 0$ hipotesi nulua
H_1: $\mu_D \neq 0$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}02$.

Probarako estatistikoa: $t = \dfrac{\bar{d} - \mu_0}{S/\sqrt{n}}$.

Probarako estatistikoaren balioa kalkulatu aurretik, balioen diferentzien taula kalkulatuko da.

d_i	-0,5	0,4	0,2	-0,3	-0,3	-0,1

Datu horien batez bestekoa eta kuasibariantzaren balioak hauek dira:

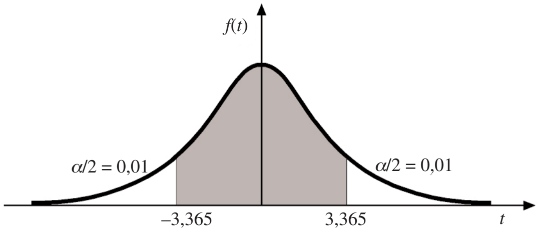
$$\bar{d} = \sum_{i=1}^{n} \frac{d_i}{n} = -0{,}1$$

$$S^2 = \sum_{i=1}^{n} \frac{(d_i - \bar{d})^2}{n-1} = 0{,}116$$

Balio horiek probarako estatistikoan ordezkatuko dira:

$$t = \frac{\bar{d} - \mu_0}{S/\sqrt{n}} = \frac{-0{,}1 - 0}{\sqrt{0{,}116}/\sqrt{6}} = -0{,}72$$

Onarpen-eremua: $S_0 = [-t_{\alpha/2;n-1}, t_{\alpha/2;n-1},] = [-t_{0,01;5}, t_{0,01;5}] = [-3{,}365, 3{,}365]$.



7.9. irudia.

Probarako estatistikoaren balioa onarpen-eremuan dagoenez, $-0{,}72 \in S_0$ alegia, orduan % 2ko adierazgarritasun-mailaz esperimentu kimikoaren hasieran eta bukaeran kloro-kantitatea berdina dela onar daiteke.

7.4.10. ariketa

Enpresa batek bere produktuen salmentak aztertzeko, duela bi urte aleatorioki aukeratutako 250 pertsonari inkesta bat egin zien eta haietatik 75 pertsonak enpresako produktuak kontsumitzen zituztela adierazi zuten. Gaur egun enpresak aleatorioki hartutako 400 pertsonari luzatu die inkesta berdina eta horietatik 128 bere produktuen kontsumitzaileak dira. Pertsona guztiek erantzun dute inkesta.

% 3ko adierazgarritasun-mailaz, esan al daiteke kontsumitzaileen kopurua gaur egunera arte hazi egin dela?

Ebazpena

Bira X_1 = "kontsumitzaile-kopurua duela bi urte" eta X_2 = "kontsumitzaile-kopurua gaur egun" aldagai aleatorioak. Laginetako datuetatik abiatuz, laginetako kontsumitzaileen proportzioak duela bi urte eta gaur egun $\hat{p}_1 = \frac{75}{250} = 0{,}30$ eta $\hat{p}_2 = \frac{128}{400} = 0{,}32$ dira, hurrenez hurren. % 3ko adierazgarritasun-mailaz, kontsumitzaileen kopurua gaur egunera arte hazi egin denentz ikusteko, honako hipotesi-kontrastea egingo da:

H_0: $p_1 \geq p_2$ hipotesi nulua
H_1: $p_1 < p_2$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}03$.

Probarako estatistikoa: $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}}$.

Kasu honetan, estatistikoaren balioa hurrengoa da:

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}} = \frac{0,30 - 0,32}{\sqrt{\frac{0,30 \cdot 0,70}{250} + \frac{0,32 \cdot 0,68}{400}}} = -0,54.$$

Onarpen-eremua: $S_0 = [-z_\alpha, \infty) = [-z_{0,03}, \infty) = [-1,88, \infty)$

Estatistikoaren balioa $-0,54 \in S_0 = [-1,88, \infty)$ dagoenez, orduan H_0: $p_1 \geq p_2$ hipotesi nulua onartuko genuke, $\alpha = 0,03$ adierazgarritasun-mailaz. Hau da, % 3ko adierazgarritasun-mailaz, kontsumitzaileen kopurua bi urtetik hona ez dela hazi esan daiteke.

7.4.11. ariketa

Makina batek ondoz ondoko 12 txandatan ekoitzitako pieza akastunen kopurua hurrengoa izan da:

15 11 16 14 13 12 16 10 9 11 14 15

Demagun pieza akastunen ekoizpena banaketa normalari darraiola.

a. % 5eko adierazgarritasun-mailaz, onargarria al da makinak ekoizten duen pieza akastunen desbideratze tipikoa gutxienez lau unitatekoa delako baieztapena?

b. Zein da desbideratze tipikoa gutxienez lau unitatekoa izateko hipotesia ez onartzeko adierazgarritasun-maila minimoa?

Ebazpena

Biz X = “makinak ekoizten dituen pieza akastunen kopurua” aldagai aleatorioa. X: $N(\mu,\sigma)$.

Laginaren batez bestekoa:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{156}{12} = 13 \text{ pieza akastun ondoz ondoko hamabi txandatan.}$$

Laginaren kuasibariantza:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 =$$

$$= \frac{(9-13)^2 + (10-13)^2 + 2(11-13)^2 + (12-13)^2 + (13-13)^2 + 2(14-13)^2 + 2(15-13)^2 + 2(16-13)^2}{11} =$$

$$= \frac{62}{11} = 5,64 \text{ (pieza)}^2 \text{ ondoz ondoko hamabi txandatan.}$$

a. Plantea dezagun hurrengo hipotesi-kontrastea:

H_0: $\sigma \geq 4$ hipotesi nulua
H_1: $\sigma < 4$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0{,}05$.

Probarako estatistikoa: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$.

Kasu honetan hartzen duen balioa $\chi^2 = \dfrac{(12-1)(62/11)}{4^2} = 3{,}875$ da.

Hipotesi-kontrastearen onarpen-eremua:

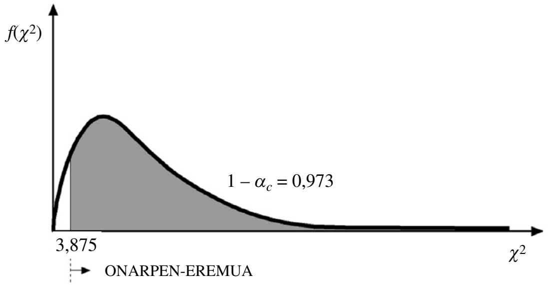
$$S_0 = \left[\chi^2_{1-\alpha;n-1},\ \infty\right) = \left[\chi^2_{0{,}95;11},\ \infty\right) = [4{,}57,\ \infty).$$

Probarako estatistikoaren balioa, 3,875 alegia, onarpen-eremutik kanpo dagoenez, desbideratze tipikoa lau unitatekoa baino txikiagoa dela onar daiteke % 5eko adierazgarritasun-mailaz.

b. Desbideratze tipikoa gutxienez lau unitatekoa delako hipotesia ez onartzeko adierazgarritasun-maila minimoa edo desbideratze tipikoa gutxienez lau unitatekoa delako hipotesia onartzeko adierazgarritasun-maila maximoa p-balioa da.

$$p\text{-}balioa = \alpha_c = \max\{\alpha / T(x_1, x_2, \ldots, x_n) \in S_0\}$$

$T(x_1, x_2, \ldots, x_n)$ gaia estatistikoak hartzen duen balioa da. Onarpen-eremua $S_0 = [3{,}875,\ \infty)$ denean lortuko da p-balioa. Khi karratuaren banaketaren taula kontsultatu eta interpolazioa aplikatuz, $\alpha_c = 0{,}027$ p-balioa lortuko da.



7.10. irudia.

7.4.12. ariketa

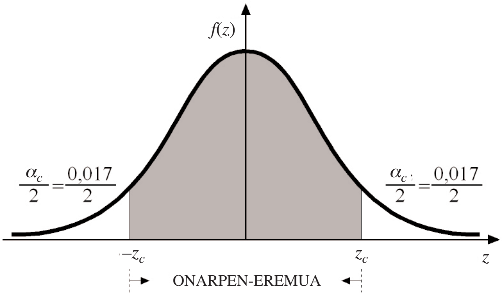
Fabrikatzaile baten arabera, bere produktuetako % $100p_0$-k Europar Batasunak agindutako espezifikazioak betetzen ditu. H_0: $p = p_0$ hipotesi nulua onargarria denetz jakiteko, % 2ko adierazgarritasun-mailako hipotesi-kontrastea egin da eta lortutako p-balioa 0,017 izan da.

a. Onartuko al zenuke fabrikatzailearen baieztapena? Arrazoi ezazu zergatia.

b. Zein adierazgarritasun-mailaz onartuko zenuke fabrikatzailearen baieztapena?

Ebazpena

a. Definizioz, p-balioa hipotesi nulua onartzeko adierazgarritasun-maila maximoa da. Beraz, H_0: $p = p_0$ hipotesi nulua, $\alpha = 0{,}02$ adierazgarritasun-maila eta $\alpha_c = 0{,}017$ p-balioa diren kasuan, $\alpha > \alpha_c$ denez, ez genuke onartuko fabrikatzailearen baieztapena % 2ko adierazgarritasun-mailaz.



7.11. irudia.

b. Fabrikatzailearen baieztapena onartzeko hartu behar den adierazgarritasun-maila $\alpha \leq 0{,}017$ p-balioa baino txikiagoa edo berdina izango da.

7.4.13. ariketa

Ikertzaile batek behin eta berriro jarri du martxan esperimentu zehatz bat eta esperimentua osorik burutzeko batez beste 72 ordu erabili ditu. Demagun esperimentuaren iraupena banaketa normalari darraiola. Zazpi esperimentu burutzeko ondoko denborak (orduak) erabili ditu:

78,9 65,1 55,2 80,9 57,4 55,4 62,3

Esperimentu-mota horri buruz beste ikertzaile batek dioenez, esperimentuak burutzeko batez besteko denbora ez da 72 ordura iristen.

a. Zer esango zenuke baieztapen horri buruz % 0,5eko adierazgarritasun-mailaz?

b. Kalkula bedi hipotesi nulua onartzeko adierazgarritasun-maila maximoaren balioa.

Ebazpena

Hasteko zazpi esperimentu hauen batez besteko iraupena eta bariantza kalkulatuko dira.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{455,2}{7} = 65 \text{ h}$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - (\bar{x})^2 = \frac{30.300,28}{7} - (65)^2 = 103,61 \text{ h}^2$$

$$s = \sqrt{103,61} = 10,18 \text{ h}$$

X = "esperimentuen iraupena" aldagai aleatorioak 72 batez bestekodun banaketa normala du, desbideratze tipikoa ezezaguna izanik.

a. Egin dezagun hurrengo hipotesi-kontrastea:

H_0: $\mu \geq 72$ hipotesi nulua
H_1: $\mu < 72$ hipotesi alternatiboa.

Adierazgarritasun-maila: $\alpha = 0,005$.

Probarako estatistikoa: $t = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n-1}}$.

Kasu honetan, probarako estatistikoaren balioa: $t = \dfrac{65 - 72}{10,18 / \sqrt{6}} = -1,68$.

Onarpen-eremua: $S_0 = [-t_{\alpha;\, n-1}, \infty) = [-t_{0,005;\, 6}, \infty) = [-3,707, \infty)$.

Probarako estatistikoaren balioa onarpen-eremuan dagoenez, orduan H_0: $\mu \geq 72$ hipotesia onartuko genuke, 0,005eko adierazgarritasun-mailaz. Hau da, bigarren ikertzaileak dioena errefusatu egingo genuke, % 0,5eko adierazgarritasun-mailaz.

b. Atal honetan α_c p-balioaren kalkulua eskatzen da.

p-balioari dagokion onarpen-eremua $S_0 = [-t_{\alpha_c;6}, \infty) = [-1,68, \infty)$ da. Student-en taulan kontsultatu eta interpolazioa aplikatuz, $\alpha_c = 0,076$ p-balioa lortuko da.

7.4.14. ariketa

Demagun prozesu mekaniko bat garatzeko kostuek banaketa normala dutela. Prozesu horiek garatzeko kostuen desbideratze tipikoa gehienez 1,8 mila eurokoa dela suposatu da.

a. Hipotesi honen kontrasteko eremu kritikoa $S_1 = (118,1, +\infty)$ da. Zein probabilitatez onartuko da enuntziatuko hipotesia, hipotesia egia bada?

b. Hipotesia kontrastatzeko 81 prozesu mekanikotako kostuak neurtu dira, lortutako p-balioa 0,05ekoa izanik. Zenbatekoa izango da 81 prozesu mekaniko horiek garatzeko kostuen desbideratze tipikoa?

- -

Ebazpena

Biz X = “prozesu mekanikoa garatzeko kostua”. X: $N(\mu,\sigma)$. Enuntziatuko hipotesiaren arabera, $\sigma \leq 1{,}8$ mila euro suposatu da.

a. Hipotesi-kontrasteen lehen motako errorearen probabilitatea α dela kontuan hartuz:

$$\alpha = P(E_I) = P(H_0 \text{ errefusatu} \mid H_0 \text{ egia})$$

lehenengo galderari erantzuna emateko

$$P(H_0 \text{ onartu} \mid H_0 \text{ egia}) = 1 - \alpha$$

probabilitatearen balioa kalkulatu behar da.

H_0: $\sigma \leq 1{,}8$ hipotesi nulua H_1: $\sigma > 1{,}8$ hipotesi alternatiboaren aurka kontrastatzean eremu kritikoa $S_1 = (118{,}1, +\infty)$ bada, khi karratuaren banaketaren taula kontsultatuz, $118{,}1 = \chi^2_{0{,}025;90}$ lortzen da. Beraz, kasu honetan $\alpha = 0{,}025$ da, eta

$$P(H_0\text{: } \sigma \leq 1{,}8 \text{ onartu} \mid H_0\text{: } \sigma \leq 1{,}8 \text{ egia}) = 1 - \alpha = 0{,}975$$

dela ondoriozta daiteke.

b. p-balioa = 0,05 eta n = 81, orduan $\chi^2_{0{,}05;80} = 101{,}9$. Adierazgarritasun-maila kritikoa edo p-balioaren definizioa kontuan hartuz,

$$\frac{ns^2}{\sigma_o^2} = \chi^2_{\text{p-balioa};n-1}$$

berdintza beteko da. Adierazpen honetan ezagunak diren balioak ordezkatuz eta ekuazioa ebatziz, desbideratze tipikoaren balioa lortuko da.

$$\frac{81 \cdot s^2}{(1{,}8)^2} = 101{,}9$$

$$s = 1{,}8\sqrt{\frac{101{,}9}{81}} \approx 2 \text{ mila euro.}$$

7.5. ARIKETA PROPOSATUAK

7.5.1. ariketa

Kotxe-marka ezagun baten kontsumoa banaketa normalari darraio. Batez besteko kontsumoa (litro/100 km) estimatzeko, aleatorioki horietariko 10 kotxe hartu dira eta ondoko kontsumoak izan dituzte:

5,1	5	5,3	5,5	5,2	5,4	5,7	6	6,1	5

a. % 1eko adierazgarritasun-mailaz, onargarria al da kotxe-marka horren batez besteko kontsumoa gutxienez 5,5 litro/100 km delako hipotesia?

b. Kotxe-marka hori ekoizten duen enpresak kontsumoaren desbideratze tipikoa 1,3 litro/100 km-koa dela baieztatzen du. Aurreko lagineko datuetan oinarrituz, kontrasta bedi baieztapen hori $\sigma \neq 1{,}3$ hipotesiaren aurka, $\alpha = 0{,}01$ hartuz.

7.5.2. ariketa

Aleatorioki aukeratutako 64 lanordutan, trokelagailu automatiko batek batez beste 1.038 pieza onargarri egin ditu, kuasidesbideratze tipikoa 146 piezakoa delarik. Demagun trokelagailu automatikoak egindako pieza onargarrien kopurua neurtzen duen aldagai aleatorioak banaketa normala duela.

a. Batez besteko pieza onargarriak 1.000 baino gehiago direlako susmoa dago. Zer esan daiteke horri buruz % 5eko adierazgarritasun-mailaz?

b. Batez besteko pieza onargarriak gehienez 1.000 direlako hipotesia egia bada, zein da hipotesi hori onartzeko probabilitatea?

7.5.3. ariketa

Enpresa batean kalitate-kontroleko arduradunek diote makinak produzitzen duen pieza-kopuruaren bariantza 25 baino txikiagoa denean makinaren produkzioa egokia dela. Demagun produkzioak banaketa normala duela. Aleatorioki hartutako hamabi egunetako produkzioaren bariantza 25 piezakoa izan da.

a. % 5eko adierazgarritasun-mailaz, onargarria al da makinaren produkzioaren bariantza 10 unitatekoa delako baieztapena?

b. Eta % 10eko adierazgarritasun-mailaz?

7.5.4. ariketa

Industria-instalazio batek sufre dioxidoa ($\mu g/m^3$) emititzen du. Emisioak murrizteko asmoz, bi iragazki berri diseinatu dira. Bi iragazkien eraginkortasuna aztertzeko, instalazioaren sei egunetako sufre dioxidoaren emisioak neurtu dira.

Sufre dioxidoa (I iragazkia)	Sufre dioxidoa (II iragazkia)
22	20
18	19
15	16
20	18
25	23
19	22

Demagun sufre dioxidoaren emisioek banaketa normala dutela. II iragazkia I iragazkia baino eraginkorragoa delako susmoa dago. % 1eko adierazgarritasun-mailaz, zer esango zenuke susmoari buruz?

7.5.5. ariketa

Enpresa baten hileko irabaziek banaketa normalari jarraitzen diote. Enpresaren hileko irabazien desbideratze tipikoa gehienez bi mila eurokoa bada, enpresaren erritmoa egokia dela esaten da. Azken sei hilabeteetan irabazien batez bestekoa 5.900 euro eta desbideratze tipikoa 2.717 euro izan dira.

a. Zein da enpresaren erritmoa egokia dela onartzeko adierazgarritasun-maila maximoa?

b. Baliorik kalkulatu gabe, defini itzazu hurrengo probabilitateak:

$$P\left(\frac{ns^2}{\sigma^2} > \chi^2_{\alpha;n-1} \middle| \sigma \leq 2.000\right) \qquad P\left(\frac{ns^2}{\sigma^2} > \chi^2_{\alpha;n-1} \middle| \sigma > 2.000\right).$$

7.5.6. ariketa

Dado baten 200 jaurtiketa aleatoriotan 122 seiko atera dira.

a. % 2,5eko adierazgarritasun-mailaz, onar al daiteke dadoa trukatuta egotearen hipotesia?

b. Kalkula bedi p-balioa eta azaldu esanahia kasu honetan.

7.5.7. ariketa

A eta B hirietako eguneko batez besteko tenperaturek banaketa normalari jarraitzen diote eta elkarrekiko independenteak dira. Azken astean neurtu diren eguneko batez besteko tenperaturak (ºC) hauek izan dira:

	Astelehena	Asteartea	Asteazkena	Osteguna	Ostirala
A hiria	9,1	8,8	9,6	9,5	8,7
B hiria	9,2	8,6	9,8	9,0	8,8

A hiriko tenperaturak oro har B hiriko tenperaturak baino altuagoak direlako hipotesia, onargarria al da % 10eko adierazgarritasun-mailaz?

7.5.8. ariketa

Burdinazko habeak egiten dituen lantegi batean aleatorioki 250 habe hartu dira eta horietatik 75 habek erdoiltzeko joera erakutsi dute. Erdoiltzearen aurkako tratamendu bat aplikatu ondoren, lantegian aleatorioki beste 250 burdinazko habe hartu dira eta horietatik 25 habek erdoiltzeko joera erakutsi dute.

a. % 1eko adierazgarritasun-mailaz, onargarria al da “erdoiltzeko joera erakusten duten habeen proportzioa txikiagoa da tratamendua jaso duten habeen kasuan” dioen hipotesia?

b. Zein adierazgarritasun-maila kritikok eramango gintuzke kontrako erabakia hartzera?

7.5.9. ariketa

Trafikoa aztertzen lan egiten duen teknikari batek ezaugarri ezberdinetako bi egun-mota bereizten ditu. Estazio konkretu batean, egun euritsuetan eta egun ez-euritsuetan sailkatutako 25 egunetako batez besteko trafikoa eta desbideratze tipikoa aztertu dira, ondorio hauekin:

Egun euritsuetan: $\bar{x}_1 = 353$ ibilgailu, $s_1 = 25$ ibilgailu

Egun ez-euritsuetan: $\bar{x}_2 = 279$ ibilgailu, $s_2 = 37$ ibilgailu

Demagun trafiko-fluxua banaketa normalari darraiola eta egun euritsu eta ez-euritsuetako trafiko-datuak elkarrekiko independenteak direla.

a. % 5eko adierazgarritasun-mailaz, zer esango zenuke ondoko baieztapenari buruz: “Egun euritsuetako trafikoaren bariantza egun ez-euritsuetako trafikoaren bariantza baino txikiagoa da”.

b. Kalkula bedi aurreko baieztapena onartzeko probabilitatea, baieztapena gezurra denean.

7.5.10. ariketa

Unibertsitate zehatz bateko A sailak batez beste B sailak baino ikerkuntza-proiektu gehiago zuzentzen dituelako susmoa dago. Horretarako azken sei urteetan sail horiek zuzendu dituzten ikerkuntza-proiektuak zenbatu dira, A sailari dagokion batez bestekoa 17 proiektu eta desbideratze tipikoa 7 proiektu, eta B sailari dagokion batez bestekoa 12 proiektu eta desbideratze tipikoa 5 proiektu izanik, hurrenez hurren. Demagun A sailak eta B sailak zuzendutako ikerkuntza-proiektuen kopuruek banaketa normalak dituztela eta elkarrekiko independenteak eta bariantza ezberdinekoak direla. % 2,5eko adierazgarritasun-mailaz, zer esango zenuke enuntziatuko susmoari buruz?

7.5.11. ariketa

Biz $X{:}N(\mu, 5)$ aldagai aleatorioa.

a. Populazioaren batez bestekoak μ_0 balioa hartzen duela baieztatu da. Baieztapen hori egia denetz ikusteko, % 1eko adierazgarritasun-mailako alde biko hipotesi-kontrastea burutu da, non p-balioa 0,009 den. Onartuko al zenuke emandako baieztapena? Zergatik?

b. H_0: $\mu = 3{,}2$ hipotesi nulua H_1: $\mu = 4{,}7$ hipotesi alternatiboaren aurka kontrastatzeko, 60 tamainako lagin aleatorio bakuna hartu da. Kalkula bedi hipotesi-kontrastearen ahalmena, jakinda laginaren batez bestekoaren eremu kritikoa (3,5) tartea dela.

8. χ^2 banaketaren aplikazioak

8.1. SARRERA

Estatistikako zenbait problematan, batzuetan, populazioaren parametro ezezagun bati buruzko hipotesiak onargarriak direnetz ikusi behar da eta horretarako aurreko gaian aztertutako hipotesi-kontraste parametrikoak egingo dira. Beste batzuetan, populazioaren adierazgarri orokorrago bati buruzko erabaki estatistikoak hartu behar dira. Orduan, hipotesi-kontraste ez-parametrikoak ebatziko dira. Kasu horietako zenbaitetan, bi faktoreren arteko independentzia, homogeneotasun-probak edo lagin bateko datuei dagokien banaketa aztergai direnean, hain zuzen ere, χ^2 banaketa aplikatuko da.

8.2. DOIKUNTZAREN EGOKITASUNA

Demagun populazio batean X izaera aztergai dela. Bira n elementu baztertzaile askez osaturiko $x_1, x_2, \ldots, x_k$ lagina eta $o_1, o_2, \ldots, o_k$ behatutako maiztasunak, $\sum_{i=1}^{k} o_i = n$ izanik.

Bira $p_1, p_2, \ldots, p_k$ probabilitate teorikoak, non $p_i = P(X = x_i)$ diren. Itxarotako maiztasunak

$$e_i = n \, . \, p_i \qquad (i = 1, 2, \ldots, k)$$

eran definitzen dira, $\sum_{i=1}^{k} p_i = 1$ eta $\sum_{i=1}^{k} e_i = n$ izanik.

Aukeratutako banaketa teorikoak behatutakoarekin bat egiten dueneko hipotesi nulua onargarria denetz ikusteko, probarako hurrengo estatistikoa erabiliko da:

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

zeinek $v = k - 1$ askatasun-graduko khi karratu banaketa duen.

Horrela, $\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \leq \chi^2_{\alpha;k-1}$ bada, hipotesi nulua onartuko da, α adierazgarritasun-mailaz. Aldiz, $\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} > \chi^2_{\alpha;k-1}$ bada, hipotesi nulua errefusatuko da, α adierazgarritasun-mailaz.

Kalkuluak errazteko asmoz, probarako testaren berdintza aplika daiteke:

$$\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = \sum_{i=1}^{k} \frac{o_i^2}{e_i} - n$$

Praktikan, χ^2-ren testa aplikatzeko hurrengo baldintzak hartu behar dira kontuan:

i. Itxarotako maiztasun guztiek $e_i = n \,.\, p_i > 5 (i = 1, 2, \ldots, k)$ bete behar dute.

ii. Itxarotako maiztasunen batek baldintza hori betetzen ez badu, elkarren segidako modalitateak elkartuko dira, itxarotako maiztasun berriak bost zenbakia gainditu arte.

iii. Itxarotako maiztasunak lortzeko r parametro estimatu behar badira, orduan χ^2 banaketaren askatasun-graduak $k - r - 1$ dira.

8.3. BI FAKTOREREN ARTEKO INDEPENDENTZIA-PROBA

Bira X eta Y faktoreak. X eta Y independenteak direnetz ikusteko, χ^2 proba erabiliko duen hurrengo hipotesi-kontrastea gauza daiteke:

H_0: X eta Y independenteak dira.

H_1: X eta Y ez dira independenteak.

Hasteko, aztergai den populaziotik n tamainako lagina hartuko da. Ondoren, X eta Y faktoreei dagokien kontingentzia-taula eraikiko da, non gelaxka bakoitzean behatutako maiztasunak adieraziko diren.

Y \ X	y_1	…	y_j	…	y_m	X-ren maiztasunak
x_1	o_{11}	…	o_{1j}	…	o_{1m}	o_{x1}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_i	o_{i1}	…	o_{ij}	…	o_{im}	o_{xi}
⋮	⋮	⋮	⋮	⋮	⋮	
x_k	o_{k1}	…	o_{kj}	…	o_{km}	o_{xk}
Y-ren maiztasunak	o_{y1}	…	o_{yj}	…	o_{ym}	n

X eta Y independenteak direla suposatuz, e_{ij} itxarotako maiztasuna hurrengo eran kalkulatuko da:

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{n}, \qquad i = 1, 2, \ldots, k, \qquad j = 1, 2, \ldots, m$$

non o_{xi} gaia x_i balioari dagokion total marginala eta o_{yj} gaia y_j balioari dagokion total marginala diren.

Gogoratzekoa da itxarotako maiztasun guztiek $e_{ij} > 5$ ($i = 1, 2, \ldots, k, j = 1, 2, \ldots, m$) bete behar dutela. Kontrako kasuan, ondoz ondoko errenkada edo zutabeak bildu behar dira, itxarotako maiztasun guztiek bost zenbakia gainditu arte.

Ondoren, kontingentzia-taulako datuei dagokien probarako estatistikoaren balioa kalkulatuko da:

$$\chi^2_{(k-1)(m-1)} = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

zeinek $v = (k-1)(m-1)$ askatasun-graduko khi karratu banaketa duen.

Horrela, $\displaystyle\sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \leq \chi^2_{\alpha;(k-1)(m-1)}$ bada, hipotesi nulua onartuko da, α adierazgarritasun-mailaz. Bestela, $\displaystyle\sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} > \chi^2_{\alpha;(k-1)(m-1)}$ bada, hipotesi nulua errefusatuko da, α adierazgarritasun-mailaz.

8.4. HOMOGENEOTASUN-PROBAK

Bi faktoreren arteko independentzia aztertzeko orduan χ^2 proba aplikatu da, populazioko lagin bakarra hartuta. Homogeneotasun-probetan, aldiz, populazioko zenbait lagin aske desberdin aukeratuko dira, izaera konkretu batekiko azterketa egitean, laginak populazio beretik ateratakoak direnetz ikusi nahi baita.

Kalkulu matematikoen aldetik, independentziaren proba eta homogeneotasun-proba berdinak dira. Bi kasuetan Yates-en zuzenketa aplika daiteke.

8.4.1. Yates-en zuzenketa

Bi bider bi motako kontingentzia-taula honelakoa da:

Y \ X	y_1	y_2	X-ren maiztasunak
x_1	o_{11}	o_{12}	o_{x1}
x_2	o_{21}	o_{22}	o_{x2}
Y-ren maiztasunak	o_{y1}	o_{y2}	n

Kasu honetan, probarako estatistikoa hau da:

$$\chi_1^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$$

Hala ere, errore txikiagoa egiten da hurrengo estatistikoa erabiliz:

$$\chi_Y^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(\left|o_{ij}-e_{ij}\right|-\frac{1}{2}\right)^2}{e_{ij}}.$$

Estatistiko hau orain arte erabilitako estatistikoaren **Yates-en zuzenketa** da. Bere aplikazioa egokia da $\left|o_{11}\cdot o_{22}-o_{12}\cdot o_{21}\right|>\frac{n}{2}$ denean.

8.5. ARIKETA EBATZIAK

8.5.1. ariketa

Hurrengo taulan telefonia-enpresa bateko zentralitak bost minutuko 530 tartetan jasotako mezu-kopurua adierazten da:

Mezu-kopurua	0	1	2	3	4	5	6	7	8
Maiztasuna	4	90	140	175	60	31	20	6	4

% 5eko adierazgarritasun-mailaz, onargarria al da datu hauek Poisson-en banaketatik ateratakoak direlako baieztapena?

Ebazpena

Kasu honetan datuak λ parametroko Poisson-en banaketara doitzen direnetz ikusi nahi da. Orduan honako hipotesi-kontrastea egin behar da:

H_0: X aldagaiak Poisson-en banaketa du.

H_1: X aldagaiak ez du Poisson-en banaketa.

Poisson-en banaketako λ parametroa estimatzeko, laginaren batez bestekoa kalkulatuko da.

$$\hat{\lambda} = \bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i o_i = \frac{1.484}{530} = 2{,}8.$$

χ^2 testa aplikatuta, doikuntzaren egokitasuna azter daiteke. Probarako erabiliko den

$$\chi^2_{k-r-1} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

estatistikoak $v = k - r - 1$ askatasun-graduko khi karratu banaketa du, non k modalitate-kopurua eta r estimatutako parametro-kopurua diren. Kasu honetan, datuak ez dira bildu modalitate gehiagotan, itxarotako maiztasun guztiak bost zenbakia baino handiagoak direlako.

x_i	o_i	$p_i = P(X = x_i)$	$e_i = np_i$	$(o_i - e_i)^2 / e_i$
0	4	0,061	32,33	24,82
1	90	0,170	90,10	0,0001
2	140	0,238	126,14	1,52
3	175	0,223	118,19	27,31
4	60	0,156	82,68	6,22
5	31	0,087	461,10	4,95
6	20	0,041	217,30	0,14
7, 8	6, 4 → 10	0,016, 0,006 → 0,022	8,48, 3,18 → 11,66	0,24
	$n = 530$			65,20

Taulako probabilitate-legea Poisson-en banaketari dagozkion

$$P(X = x) = \frac{e^{-2{,}8}(2{,}8)^k}{k!}, \qquad k = 0,\ 1,\ \ldots,\ 8$$

adierazpenen bidez kalkulatu da.

Bestalde, probarako estatistikoak $v = k - r - 1 = 8-1-1 = 6$ askatasun-graduko banaketa du, non $k = 8$ modalitate eta $r = 1$ estimatutako parametro-kopurua diren.

Horrela, $\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = 65,20 > \chi^2_{0,05;6} = 12,592$ denez, hipotesi nulua errefusatuko da, $\alpha = 0,05$ adierazgarritasun-mailaz.

8.5.2. ariketa

Makina batek eguneko ekoizten dituen piezak kalitate hobezinekoak direnetz jakiteko, proba berezi bat aplikatzen da. Egunero zazpi piezaren kalitatea aztertzen da. Aleatorioki hartutako 65 kalitate-kontrol burutu dira eta kalitate hobezineko piezen banaketa honakoa izan da:

Kalitate hobezineko piezak	0	1	2	3	4	5	6	7
Egun	10	15	10	15	5	3	4	3

$\alpha = 0,005$ hartuz, onar al daiteke 65 egunetako ekoizpenaren kalitate hobezineko piezen kopuruak banaketa binomiala izatea?

Ebazpena

Biz X = "kalitate hobezineko piezen kopurua, eguneko" aldagai aleatorioa. X aldagaiak banaketa binomiala duenetz jakiteko, X: *Bin*(m,p) alegia, χ^2 testa aplikatuko da. Hasteko, $m = 7$ dela onartuko da, egunero zazpi piezaren kalitatea aztertzen delako. Bestalde, $p = P$(pieza, kalitate hobezinekoa) parametroa estimatzeko, laginaren batez bestekoa kalkulatuko da.

$$\hat{p} = \frac{\overline{X}}{m} = \frac{\overline{X}}{7}.$$

Kasu honetan:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{k} x_i o_i = \frac{160}{65} = 2,46.$$

Beraz, p parametroaren hurrengo estimazioa lor daiteke:

$$\hat{p} = \frac{\overline{X}}{m} = \frac{2,46}{7} = 0,35.$$

Ondorioz, enuntziatuan emandako datuek *Bin*$(m = 7,\ p = 0,35)$ banaketa dutenetz ikusi nahi da. Doikuntzaren egokitasuna aztertzeko, probarako $\chi^2_{k-r-1} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$ estatistikoa erabil daiteke.

Halaber, probabilitate-legearen balioak zehazteko, banaketa binomialari dagozkion

$$P(X=x)=\binom{m}{x}p^x(1-p)^{m-x}, \qquad x=0,\,1,\,2,\,\ldots,\,m$$

balioak kalkulatu dira.

x_i	o_i		$p_i=P(X=x_i)$	$e_i=np_i$		$(o_i-e_i)^2/e_i$
0	10	25	0,0490	3,185	15,197	6,324
1	15		0,1848	12,012		
2	10		0,2985	19,403		4,557
3	15		0,2679	17,414		0,335
4	5	15	0,1442	9,373	12,987	0,312
5	3		0,0466	3,029		
6	4		0,0084	0,546		
7	3		0,0006	0,039		
	65					11,528

$\nu = k - r - 1$ askatasun-graduko khi karratu banaketa du, non k modalitate-kopurua eta r estimatutako parametro-kopurua diren. Kasu honetan, aurreko taulan ikus daitekeenez, datuak modalitatetan batu dira itxarotako maiztasun guztiak bost zenbakia baino handiagoak izan daitezen. Bestalde, probarako estatistikoak $\nu = k - r - 1 = 4 - 1 - 1 = 2$ askatasun-graduko khi karratu banaketa du.

Bukatzeko, $\displaystyle\sum_{i=1}^{k}\frac{(o_i-e_i)^2}{e_i}=11{,}528>\chi^2_{0{,}005;2}=10{,}60$ denez, 65 egunetako ekoizpenaren kalitate hobezineko piezen kopuruak banaketa binomiala ez duela onartuko da, $\alpha = 0{,}005$ adierazgarritasun-mailaz.

8.5.3. ariketa

Baratze bat fumigatzen den 101 alditan aireko hondar tionikoen mailak (μg/m^3) neurtu dira:

Hondar tionikoa	[40,50)	[50,60)	[60,70)	[70,80)	[80,90)	[90,100)
Fumigazio-aldiak	28	25	33	8	4	3

% 1eko adierazgarritasun-mailaz, puntuazioek banaketa normala al dute?

Ebazpena

Hasteko, datuak era egokian klasetan banatuko dira, eta ondoren banaketa normalaren μ batez bestekoa eta σ desbideratze tipikoa estimatuko dira. Horretarako honako kalkuluak egin daitezke:

$[l_i, l_{i+1})$	$x_i = (l_i + l_{i+1})/2$	o_i	$x_i o_i$	$(x_i - \bar{x})^2 o_i$
[40,50)	45	28	1.260	5.854,56
[50,60)	55	25	1.375	497,29
[60,70)	65	33	2.145	1.012,82
[70,80)	75	8	600	1.931,93
[80,90)	85	4	340	2.609,17
[90,100)	95	3	285	3.789,27
		101	6.005	15.695,04

Batez bestekoaren eta desbideratze tipikoaren estimazioak:

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i o_i = \frac{6.005}{101} = 59,46$$

$$\hat{\sigma} = s = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 o_i} = \sqrt{\frac{15.695,04}{101}} = 12,47.$$

Doikuntzaren egokitasunerako χ^2 testa aplikatuko da. Probarako estatistikoak

$$\chi^2_{k-r-1} = \sum_{i=1}^{k}\frac{(o_i - e_i)^2}{e_i}$$

$v = k - r - 1$ askatasun-graduko khi karratu banaketa du, k klase-kopurua eta r estimatutako parametro-kopurua izanik.

Banaketa normaleko banaketa-funtzioaren balioen bidez klase bakoitzari dagokion probabilitatea kalkulatuko da, Z aldagaiak banaketa normal tipikoa duelarik.

$$P(40 \le X < 50) = P\left(\frac{40 - 59,46}{12,47} \le \frac{X - 59,46}{12,47} < \frac{50 - 59,46}{12,47}\right) =$$

$$= P(-1,56 \le Z < -0,76) = P(0,76 \le Z < 1,56) =$$

$$= \phi(1,56) - \phi(0,76) = 0,9406 - 0,7764 = 0,1642.$$

Era honetan hurrengo taulako probabilitate guztiak zehaztuko dira.

x_i	o_i		p_i		$e_i = np_i$		$(o_i - e_i)^2 / e_i$
45	28		0,1642		16,584		7,86
55	25		0,2924		29,532		0,70
65	33		0,2863		28,916		0,58
75	8		0,1482		14,968		
85	4	15	0,0424	0,1971	4,282	19,907	1,21
95	3		0,0065		0,657		
	101						10,35

Pobarako estatistikoak $v = k - r - 1 = 4 - 2 - 1 = 1$ askatasun-graduko banaketa du, $k = 4$ eta $r = 2$ izanik.

Horrela, $\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = 10,35 > \chi^2_{0,01;1} = 6,63$ denez, hipotesi nulua errefusatuko da, $\alpha = 0,05$ adierazgarritasun-mailaz. Hau da emandako datuak ez dira banaketa normaletik ateratakoak.

8.5.4. ariketa

Hurrengo taulan hiru enpresatako azken lau hilabeteetako eskaerak (produktu-kopurua) adierazten dira:

	1 hilabetea	2 hilabetea	3 hilabetea	4 hilabetea	
A enpresa	2.200	2.500	2.700	1.900	9.300
B enpresa	1.900	2.200	2.000	2.000	8.100
C enpresa	2.000	2.600	1.950	2.100	8.650
	6.100	7.300	6.650	6.000	26.050

% 5eko adierazgarritasun-mailaz, froga ezazu enpresaren eta hilabetearen arteko independentzia dagoenetz.

Ebazpena

Kasu honetan egingo den hipotesi-kontrastea hurrengoa da:

H_0: enpresa eta hilabetea independenteak dira.

H_1: enpresa eta hilabetea ez dira independenteak.

Bi faktoreren independentziaren probarako testa aplikatu da. Enuntziatuko taulako gelaxketan behatutako o_{ij} maiztasunak adierazten dira. Hurrengo taulan itxarotako maiztasunen balioak adieraziko dira, non itxarotako maiztasun bakoitza kalkulatzeko

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{n}, \qquad i = 1, 2, 3 \qquad j = 1, 2, 3, 4$$

adierazpena aplikatuko den, n = 26.050 izanik.

e_{ij}	1 hilabetea	2 hilabetea	3 hilabetea	4 hilabetea
A enpresa	2.178	2.606	2.374	2.142
B enpresa	1.897	2.270	2.068	1.866
C enpresa	2.026	2.424	2.208	1.992

Ondoren, probarako estatistikoaren balioa kalkulatuko da.

$$\sum_{i=1}^{3}\sum_{j=1}^{4}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} = \frac{(2.200-2.178)^2}{2.178} + \frac{(2.500-2.606)^2}{2.606} +$$

$$+\frac{(2.700-2.374)^2}{2.374} + \ldots + \frac{(1.950-2.208)^2}{2.208} + \frac{(2.100-1.992)^2}{1.992} = 139,78.$$

Estatistiko honek $v = (3 - 1)(4 - 1) = 6$ askatasun-graduko khi karratu banaketa du.

Kasu honetan $139,78 = \sum_{i=1}^{k}\sum_{j=1}^{m}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} > \chi^2_{0,05;6} = 12,59$ denez, independentziaren hipotesia errefusatu egingo da, α = 0,05 izanik.

8.5.5. ariketa

Ondoko taulan A eta B lantegietako goizeko, arratsaldeko eta gaueko txandetan ekoitzitako pieza-kopurua adierazten da:

	A lantegia	B lantegia	
Goizekoa	750	1.000	1.750
Arratsaldekoa	850	1.200	2.050
Gauekoa	725	1.100	1.825
	2.325	3.300	5.625

% 1eko adierazgarritasun-mailaz, froga ezazu txandak ez duela eraginik lantegien ekoizpenean.

Ebazpena

Bi faktoreren independentzia aztertu behar da. Hasteko, itxarotako maiztasunen taula kalkula daiteke.

e_{ij}	A lantegia	B lantegia
Goizekoa	723	1.027
Arratsaldekoa	847	1.203
Gauekoa	754	1.071

Itxarotako maiztasunen kalkulurako adierazpena:

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{5.625}, \qquad i = 1, 2, 3 \qquad j = 1, 2.$$

Ondoren, probarako estatistikoaren balioa kalkulatuko da, zeinek $v = (3 - 1)(2 - 1) = 2$ askatasun-graduko khi karratu banaketa duen.

$$\sum_{i=1}^{3} \sum_{j=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = \frac{(750 - 723)^2}{723} + \frac{(1.000 - 1.027)^2}{1.027} +$$

$$+ \frac{(850 - 847)^2}{847} + \ldots + \frac{(1.100 - 1.071)^2}{1.071} = 3,64.$$

Oraingoan $\sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = 3,64 \le \chi^2_{0,01;2} = 9,21$ denez, independentziaren hipotesia onartu egingo da, % 1eko adierazgarritasun-mailaz.

8.5.6. ariketa

Ordenagailuen biruskontrako A, B, C eta D programen eraginkortasuna aztertu nahi da. Birus baten aurka lau programak aplikatu dira; kasu batzuetan, ordenagailua garbi dago (birusik gabe) eta beste batzuetan oraindik ez dago garbi:

	Garbi	Ez garbi	
A programa	70	30	100
B programa	68	10	78
C programa	95	35	130
D programa	74	18	92
	307	93	400

% 2,5eko adierazgarritasun-mailaz, froga ezazu ez dagoela diferentzia adierazgarririk lau programen eraginkortasunean.

Ebazpena

Honetan, bi faktoreren homogeneotasun-proba egingo da.

Hurrengo taulan itxarotako maiztasunen balioak lortuko dira:

e_{ij}	Garbi	Ez garbi
A programa	76,75	23,25
B programa	59,87	18,14
C programa	99,78	30,23
D programa	70,61	21,39

Itxarotako maiztasunen kalkulurako adierazpena:

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{400}, \qquad i = 1, 2, 3, 4 \qquad j = 1, 2.$$

Probarako estatistikoaren balioa:

$$\sum_{i=1}^{4} \sum_{j=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = \frac{(70 - 76,75)^2}{76,75} + \frac{(30 - 23,25)^2}{23,25} + $$
$$+ \frac{(68 - 59,87)^2}{59,87} + \ldots + \frac{(18 - 21,39)^2}{21,39} = 8,99$$

zeinek $v = (4 - 1)(2 - 1) = 3$ askatasun-graduko khi karratu banaketa duen.

Probarako estatistikoaren balioa $\sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = 8,99 < \chi^2_{0,025;3} = 9,35$ denez, lau programen artean eraginkortasunean diferentzia adierazgarririk ez dagoela onar daiteke, % 2,5eko adierazgarritasun-mailaz.

8.5.7. ariketa

Oposaketa zehatz batean, 90 mutiletik 70ek gainditu dute eta 15 neskatik 12k gainditu dute. % 2,5eko adierazgarritasun-mailaz, froga bedi oposaketa gainditu duten mutilen eta nesken kopuruan ez dagoela diferentzia adierazgarririk.

Ebazpena

Enuntziatuko datuak hurrengo kontingentzia-taulan adieraz daitezke:

	Gainditu	Ez gainditu	
Mutilak	70	20	90
Neskak	12	3	15
	82	23	105

Homogeneotasun-proba honetan, faktore bakoitzak bi balio hartzen ditu. Balio horiei dagozkien itxarotako maiztasunak:

$$e_{11} = \frac{90 \cdot 82}{105} = 70,29, \quad e_{12} = \frac{90 \cdot 23}{105} = 19,71.$$
$$e_{21} = \frac{15 \cdot 82}{105} = 11,71, \quad e_{22} = \frac{15 \cdot 23}{105} = 3,29\,.$$

Probarako estatistikoaren balioa:

$$\chi_1^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} =$$
$$= \frac{(70 - 70,29)^2}{70,29} + \frac{(20 - 19,71)^2}{19,71} + \frac{(12 - 11,71)^2}{11,71} + \frac{(3 - 3,29)^2}{3,29} = 0,038.$$

Kasu honetan ez da Yates-en zuzenketa aplikatu, $\left|o_{11} \cdot o_{22} - o_{12} \cdot o_{21}\right| > \frac{n}{2}$ betetzen ez delako.

$\sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = 0,038 < \chi_{0,025;1}^2 = 5,02$ denez, oposaketa gainditu duten mutilen eta nesken kopuruan diferentzia adierazgarririk ez dagoela onartuko da, % 2,5eko adierazgarritasun-mailaz.

8.5.8. ariketa

Azken urteetan zenbait kutsatzaileren emisioak murrizteko, EAEko industriek iragazkiak ezarri dituzte beren instalazioetan. Zenbait industriak isunak ordaindu dituzte emisio altuegiak izateagatik. Bestalde, iragazkiak goi-teknologiakoak edo arruntak izan daitezke. Hurrengo taulan aleatorioki hartutako 240 industriaren egoera azaltzen dugu:

	Goi-teknologiakoa	Arrunta	
Isuna, bai	30	90	120
Isuna, ez	20	100	120
	50	190	240

% 5eko adierazgarritasun-mailaz, froga bedi ez dagoela dependentziarik isuna jasotzearen eta iragazki-motaren artean.

Ebazpena

Bira X = "isuna jaso" eta Y = "iragazki-mota" faktoreak. Kasu honetan hurrengo hipotesi-kontrastea egin behar da:

H_0: X eta Y independenteak dira.

H_1: X eta Y ez dira independenteak.

Behatutako maiztasunak bi bider bi motako kontingentzia-taulan adierazi dira. Bestalde, datu hauei dagozkien itxarotako maiztasunak kalkulatzeko hurrengo adierazpena aplikatuko da:

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{n}, \qquad i = 1, 2, \qquad j = 1, 2, \quad n = 240$$

$$e_{11} = \frac{120 \cdot 50}{240} = 25, \quad e_{12} = \frac{120 \cdot 190}{240} = 95$$
$$e_{21} = \frac{120 \cdot 50}{105} = 25, \quad e_{22} = \frac{120 \cdot 190}{240} = 95.$$

Oraingoan,

$$|o_{11} \cdot o_{22} - o_{12} \cdot o_{21}| = |30 \cdot 100 - 90 \cdot 20| = 1200 > \frac{n}{2} = \frac{240}{2} = 120$$

betetzen denez, Yates-en zuzenketa aplikatuko da. Beraz, probarako erabiliko den estatistikoa hau da:

$$\chi_Y^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\left(|o_{ij} - e_{ij}| - \frac{1}{2}\right)^2}{e_{ij}}.$$

Kalkula dezagun estatistikoaren balioa:

$$\chi_Y^2 = \sum_{i=1}^{2} \sum_{i=1}^{2} \frac{\left(|o_{ij} - e_{ij}| - \frac{1}{2}\right)^2}{e_{ij}} =$$
$$= \frac{\left(|30 - 25| - \frac{1}{2}\right)^2}{25} + \frac{\left(|90 - 95| - \frac{1}{2}\right)^2}{95} + \frac{\left(|20 - 25| - \frac{1}{2}\right)^2}{25} + \frac{\left(|100 - 95| - \frac{1}{2}\right)^2}{95} = 2,05.$$

Orduan, $\chi_Y^2 = 2,05 < \chi^2_{0,05;1} = 3,84$ denez, isuna jasotzearen eta iragazki-motaren artean ez dagoela dependentziarik onartuko da, % 5eko adierazgarritasun-mailaz.

8.5.9. ariketa

Papergintzaren arabera, gaur egungo eskarien % 40 dina4 tamainakoa, % 45 orri-tamainakoa eta % 15 beste tamaina batzuetakoak dira. Aleatorioki hartutako inprimategi batek dina4 tamainako 1.900 lote, orri-tamainako 1.400 lote eta beste tamainetako 700 lote eskatzen ditu.

% 1eko adierazgarritasun-mailaz, onar al dezake inprimategiak papergintzaren baieztapena?

- -

Ebazpena

Biz X = "inprimategiak erabiltzen duen paper-mota (tamainaren arabera)". Papergintzaren arabera, hurrengo taula eraiki daiteke:

x_i	p_i
Dina4	0,40
Orria	0,45
Beste	0,15

Doikuntzaren egokitasuna aztertzeko, hurrengo hipotesi-kontrastea egin daiteke:

H_0: $p_1 = 0{,}40, \quad p_2 = 0{,}45, \quad p_3 = 0{,}15.$

H_1: $p_1 \neq 0{,}40, \quad p_2 \neq 0{,}45, \quad p_3 \neq 0{,}15.$

Kontrasterako $\chi^2_{k-r-1} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$ estatistikoa erabil daiteke, zeinen askatasun-graduak $v = k - r - 1 = 3 - 0 - 1 = 2$ diren, $k = 3$ modalitate-kopurua eta $r = 0$ estimatutako parametro-kopurua izanik.

x_i	o_i	$p_i = P(X = x_i)$	$e_i = np_i$	$(o_i - e_i)^2 / e_i$
Dina4	1.900	0,40	1.600	56,25
Orria	1.400	0,45	1.800	88,89
Beste	700	0,15	600	16,67
	4.000			161,81

Beraz, $\sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = 161,81 > \chi^2_{0,01;2} = 9,21$ denez, hipotesi-nulua errefusa daiteke % 1eko adierazgarritasun-mailaz. Beste hitz batzuetan, industriaren baieztapena ez da onargarria izango, $\alpha = 0{,}01$ izanik.

8.5.10 ariketa

Enpresa batek A, B eta C marketako motorrak erosten ditu. Motorrek I., II. edo III. motako akatsak badituzte, motorrak etxera itzultzen dira. Aleatorioki aukeratutako motorren laginean hurrengo akatsak behatu dira:

	I akatsa	II akatsa	III akatsa
A motorrak	27	20	23
B motorrak	22	25	27
C motorrak	19	24	26

% 2,5eko adierazgarritasun-mailaz, froga bedi akatsak motorraren markarekiko independenteak direla.

Ebazpena

Bi faktoreren independentzia aztertzeko, hasteko, balio bakoitzari dagozkion total marjinalak kalkulatu behar dira.

	I akatsa	II akatsa	III akatsa	
A motorrak	27	20	23	70
B motorrak	22	25	27	74
C motorrak	19	24	26	69
	68	69	76	213

Ondoren, itxarotako maiztasunen balioak kalkulatuko dira.

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{n}, \qquad i = 1, 2, 3 \qquad j = 1, 2, 3.$$

$$e_{11} = \frac{70 \cdot 68}{213} = 22{,}35, \quad e_{12} = \frac{70 \cdot 69}{213} = 22{,}68, \quad e_{13} = \frac{70 \cdot 76}{213} = 24{,}98$$

$$e_{21} = \frac{74 \cdot 68}{213} = 23{,}62, \quad e_{22} = \frac{74 \cdot 69}{213} = 23{,}97, \quad e_{23} = \frac{74 \cdot 76}{213} = 26{,}40$$

$$e_{31} = \frac{69 \cdot 68}{213} = 22{,}03, \quad e_{32} = \frac{69 \cdot 69}{213} = 22{,}35, \quad e_{33} = \frac{69 \cdot 76}{213} = 24{,}62.$$

Ondoren, probarako estatistikoaren balioa kalkulatuko da.

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}=\frac{(27-22,35)^2}{22,35}+\frac{(20-22,68)^2}{22,68}+$$

$$+\frac{(23-24,98)^2}{24,98}+\ldots+\frac{(26-24,62)^2}{24,62}=2{,}23.$$

Probarako estatistikoak $\nu = (3 - 1)(3 - 1) = 4$ askatasun-graduko khi karratu banaketa du.

Bukatzeko, $\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} = 2{,}23 < \chi^2_{0,05;4} = 9{,}49$ denez, independentziaren hipotesia onartuko da, % 5eko adierazgarritasun-mailaz. Beraz, motorraren markak motorraren akatsetan eragingo duelako hipotesia errefusa daiteke, $\alpha = 0{,}05$ izanik.

8.5.11. ariketa

Hiri batean aleatorioki hartu den laginak, hiriko eremuen arabera eta biztanleen adinaren arabera, hurrengo banaketa erakutsi du:

	18 urtetik beherakoak	18-32	32-46	46-60	60 urte edo gehiagokoak
A eremua	124	272	288	212	75
B eremua	188	212	195	127	80
C eremua	169	195	199	188	89

% 10eko adierazgarritasun-mailaz, egiazta ezazu badagoela biztanleen adinaren banaketaren aldetik hiriko eremuaren araberako diferentzia adierazgarririk.

Ebazpena

Hiriko biztanleriak eremuaren arabera adin-diferentzia adierazgarririk ez badu, laginak homogeneoak direla esan daiteke. Beraz, kasu honetan egingo den hipotesi-kontrastea honakoa da:

H_0: Laginak homogeneoak dira.

H_1: Laginak ez dira homogeneoak.

Homogeneotasun-probarekin hasi aurretik, enuntziatuko 3x5 motako kontingentzia-taulari dagozkion total marjinalak kalkulatuko dira.

$o_{x1} = 971 \qquad o_{x2} = 802 \qquad o_{x3} = 840$

$o_{y1} = 481 \qquad o_{y2} = 679 \qquad o_{y3} = 682 \qquad o_{y4} = 527 \qquad o_{y5} = 244.$

X= “hiriko eremua” faktorearen total marjinalak batuz edo Y = “biztanleriaren adina” faktorearen total marjinalak batuz, laginaren tamaina lor daiteke:

$$n = \sum_{i=1}^{3} o_{xi} = \sum_{j=1}^{5} o_{yj} = 2.613.$$

Homogeneotasun-proba ebatzi aurretik, itxarotako maiztasunen balioak lortu behar dira.

$$e_{ij} = \frac{o_{xi} \cdot o_{yj}}{n}, \qquad i = 1, 2, 3 \qquad j = 1, 2, 3, 4, 5.$$

$$\begin{array}{lllll} e_{11} = 178{,}74 & e_{12} = 252{,}32 & e_{13} = 253{,}43 & e_{14} = 195{,}84 & e_{15} = 90{,}67 \\ e_{21} = 147{,}63 & e_{22} = 208{,}40 & e_{23} = 209{,}32 & e_{24} = 161{,}75 & e_{25} = 74{,}89 \\ e_{31} = 154{,}63 & e_{32} = 218{,}28 & e_{33} = 219{,}24 & e_{34} = 169{,}41 & e_{35} = 78{,}44 \end{array}$$

Orduan, probarako estatistikoaren balioa hauxe da:

$$\sum_{i=1}^{3}\sum_{j=1}^{5} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = \frac{(124 - 178{,}74)^2}{178{,}74} + \frac{(272 - 252{,}32)^2}{252{,}32} +$$

$$+ \frac{(288 - 253{,}43)^2}{253{,}43} + \frac{(212 - 195{,}84)^2}{195{,}84} + \ldots + \frac{(89 - 78{,}44)^2}{78{,}44} = 56{,}10.$$

Probarako estatistikoak $v = (3 - 1)(5 - 1) = 8$ askatasun-graduko khi karratu banaketa du.

Bukatzeko, $\sum_{i=1}^{3}\sum_{j=1}^{5} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = 56{,}10 > \chi^2_{0,10;8} = 13{,}36$ denez, hiriko eremuak biztanleen adinean diferentzia adierazgarriak ezartzen dituela ondoriozta daiteke, % 10eko adierazgarritasun-mailaz.

8.6. ARIKETA PROPOSATUAK

8.6.1. ariketa

Biz X = “ingeniaritza-enpresetan ardura-postuetan dagoen emakume-kopurua”. Aleatorioki 1.000 enpresa hartu dira eta bertako ardura-postuetako emakume-kopurua zenbatu da:

Emakume-kopurua	0	1	2	3	4
Enpresa-kopurua	200	155	300	270	75

% 1eko adierazgarritasun-mailaz, onar al daiteke datu hauek banaketa binomiala betetzen dutelako hipotesia?

8.6.2. ariketa

Hurrengo taulan 80 harriren dentsitateak (gr/cm^3) adierazten dira:

Dentsitateak	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)
Harri-kopurua	2	6	8	34	24	6

% 5eko adierazgarritasun-mailaz, onargarria al da datu hauek banaketa normaletik datozelako hipotesia?

8.6.3. ariketa

Aleatorioki aukeratutako 463 telebista-hargailutan egindako fidelitate- eta selektibitate-proben emaitzak hauek izan dira:

Selektibitatea

Fidelitatea	Baxua	Ertaina	Altua
Baxua	22	46	85
Ertaina	58	33	67
Altua	32	76	44

% 1eko adierazgarritasun-mailaz, froga bedi selektibitatearen eta fidelitatearen artean badagoela dependentzia-erlazioa.

8.6.4. ariketa

Soldadura-teknika egokienaren bila dabilen lantegi batean, A, B eta C soldadura-motak erabiltzen dira. Bestalde, soldatzaileak I, II eta III mailakoak dira.

Soldadura-motak

Langileen mailak	A	B	C
I	20	8	4
II	8	15	26
III	12	48	14

Datu hauen arabera, % 5eko adierazgarritasun-mailaz, froga ezazu langileen mailak eta soldadura-motak ez direla independenteak.

8.6.5. ariketa

Hurrengo taulan biltegi batean orduro prestatzen den pakete-kopurua adierazi da:

Paketeak	0	1	2	3	4	5
Ordu-kopurua	1	2	10	10	10	2

% 5eko adierazgarritasun-mailaz, froga bedi datu hauek Poisson-en banaketari jarraitzen dion populaziotik hartu direla.

8.6.6. ariketa

Hiru material ezberdinek, tenperatura altuak jaso ondoren, honako emaitza erakutsi dute:

	I materiala	II materiala	III materiala
Desintegratuak	62	36	28
Hasierako egoeran	105	72	90

% 2,5eko adierazgarritasun-mailaz, froga bedi hiru materialek desintegratzeko probabilitate bera dutela.

8.6.7. ariketa

Bulego batean kafe-makina jarri nahi dute. Bi kafe-makina mota daude. Mota bakoitzeko makina 200 aldiz probatu da. Lehenengo motako makina 190 alditan ongi dabil eta bigarren motako makina 185 alditan ongi dabil. % 5eko adierazgarritasun-mailaz, ba al dago makinen funtzionamenduan diferentzia adierazgarririk?

8.6.8. ariketa

Azken merkatu-azterketaren arabera, zulatzeko makinen salmentak honakoak izan dira: % 40 A motakoak, % 35 B motakoak, % 20 C motakoak eta % 5 beste motatakoak. Aleatorioki hartutako burdindegi baten salmentak aztertu dira eta A motako 32, B motako 29, C motako 27 eta beste mota batzuetako 12 zulatzeko makina saldu dira.

% 2,5eko adierazgarritasun-mailaz, burdindegiko nagusiarentzat onargarriak al dira merkatu-azterketaren emaitzak?

8.6.9. ariketa

Aseguru-etxe bateko bezero batzuk jasotako tratamenduarekin pozik daude eta beste batzuk ez. Bezeroen artean, batzuek ezbeharrak izan dituzte eta besteek ez.

Ezbeharrak \ Bezeroa	Pozik	Ez-pozik
Bai	35	25
Ez	55	65

% 5eko adierazgarritasun-mailaz, onargarria al da ezbeharrak izateak bezeroaren pozik ala ez pozik egotean eragina duelako hipotesia?

8.6.10. ariketa

Laneko estatutu berriak onartzeko galdeketan, bi sailen erantzunak honako hauek izan dira:

	A saila	B saila
Alde	60	85
Kontra	30	65
Ez alde/ ez kontra	10	15

% 1eko adierazgarritasun-mailaz, onargarria al da estatutuen aldeko langileen proportzioa homogeneoa delako hipotesia?

9. Erregresio lineal bakuna

9.1. SARRERA

Erregresio-analisia zenbait aldagairen arteko erlazio funtzionala aztertzeko aplika daitekeen teknika estatistikoa da. Gai honetan bi aldagairen kasua aztertuko da, non X aldagai askea (x balio finkoak hartzen dituena) eta Y menpeko aldagaia (aleatorioa) diren. Erregresioaren helburua menpeko aldagaiak aldagai askearen funtzioan adierazten duen eredua zehaztea da, era horretan aldagai askearen x baliorako menpeko aldagaiaren aurresana lortu ahal izango baita.

Erregresioa aztertzen hasteko kasu sinpleena **erregresio lineal bakuna** da, non X eta Y aldagaien arteko erlazio funtzionala lineala den. Erlazio lineal hori

$$y = \alpha x + \beta + \varepsilon$$

ereduaren bidez emanik dator, non ε errorea eta α eta β parametroak diren. Oro har, ezin da zuzen honen adierazpen matematikoa zehaztu, baina X eta Y aldagaien balioez osaturiko $\left\{\left(x_i, y_i\right)\right\}_{i=1}^{n}$ puntu-hodeiari doitutako

$$y = ax + b$$

***X* gaineko *Y* aldagaiaren erregresio-zuzeneko** a eta b koefizienteetatik abiatuz, α eta β parametroen estimazioak lor daitezke.

Puntu-hodeiari doituriko erregresio-zuzena finkatu ondoren, aldagai askearen x balio jakin bati dagokion y aldagai aleatorioaren balioa aurresan ahal izango da.

Hurrengo urrats batean, doikuntzaren egokitasuna aztertzea komeni da, kalkulatutako erregresio-zuzeneko balioen eta puntu-hodeiko balioen arteko errorea aztertzeko. Horretarako **estimazio-errore estandarra** kalkulatuko da.

Halaber, dependentzia linealaren maila aztertzeko **Pearson-en korrelazio-koefizientea** lor daiteke.

Puntu-hodeiari doitzen zaion lerroaren arabera, erregresio parabolikoa, koadratikoa edo esponentziala izan daitezke. Gai honetako aztergaia erregresio lineal bakuna da.

9.2. ERREGRESIO LINEAL BAKUNA

Demagun X eta Y aldagaien artean dependentzia lineala dagoela. Orduan,

$$y = \alpha x + \beta + \varepsilon$$

erregresio lineal bakuneko eredua dugu, non ε errorea eta α eta β parametroak diren.

Erregresioaren hurrengo hipotesiak beteko dira:

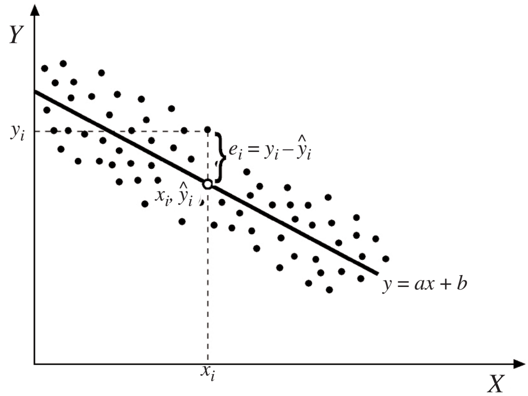
i. ε erroreek 0 batez bestekodun eta bariantza konstanteko banaketa normala dute.

ii. Erroreak elkarrekiko independenteak dira.

Eredu linealaren α eta β parametroak estimatzeko $\left\{(x_i, y_i)\right\}_{i=1}^{n}$ puntu-hodeira doitutako

$$y = ax + b$$

erregresio-zuzena lortuko da. Erregresio-zuzeneko a eta b koefizienteen kalkulurako **minimo karratuen metodoa** aplika daiteke. Metodo hau errore karratuen batura minimoa lortzean datza.



9.1. irudia.

Minimo karratuen metodoaren oinarria hurrengo funtzioaren minimoa kalkulatzea da:

$$F(a,b) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (ax_i + b))^2.$$

Orduan, a eta b parametroekiko deribatuak kalkulatu eta zerora berdinduko dira.

$$\left.\begin{aligned} \frac{\partial F(a,b)}{\partial a} &= 2\sum_{i=1}^{n} (y_i - (ax_i + b))(-x_i) = 0 \\ \frac{\partial F(a,b)}{\partial b} &= 2\sum_{i=1}^{n} (y_i - (ax_i + b))(-1) = 0 \end{aligned}\right\}.$$

Era baliokidean, beste bi ekuazio hauek ondoriozta daitezke:

$$\left.\begin{aligned}\sum_{i=1}^{n} y_i x_i &= a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i &= a\sum_{i=1}^{n} x_i + b\cdot n\end{aligned}\right\}.$$

Orduan, **laginaren erregresio-koefizienteak** lortuko dira:

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b = \overline{y} - a\overline{x}$$

koefiziente hauek α eta β parametroen estimazioak dira, hurrenez hurren.

Bukatzeko, X gaineko Y aldagaiaren erregresio-zuzena, era normalean, honela idatz daiteke:

$$y - \overline{y} = a(x - \overline{x}).$$

9.3. PEARSON-EN KORRELAZIO LINEALAREN KOEFIZIENTEA

X eta Y aldagai aleatorioen artean dependentzia linealaren maila aztertzeko, **Pearson-en *r* korrelazio linealaren koefizientea** kalkulatuko da:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}}$$

non $x_1, x_2, \ldots, x_n$ eta $y_1, y_2, \ldots, y_n$ balioak X eta Y aldagaien laginak osatzen dituzten balioak diren. Pearson-en korrelazio-koefizientearen balioa [-1, 1] tartean dago.

$|r| = 1$ denean, dependentzia lineal absolutua dago. $|r| \geq 0{,}8$ bada, dependentzia lineal handia dagoela esaten da. Halaber, $0{,}5 \leq |r| < 0{,}8$ bada, dependentzia lineal onargarria dagoela esan daiteke. Bestalde, $|r| < 0{,}5$ denean, dependentzia linealik ez dagoela onartzen da.

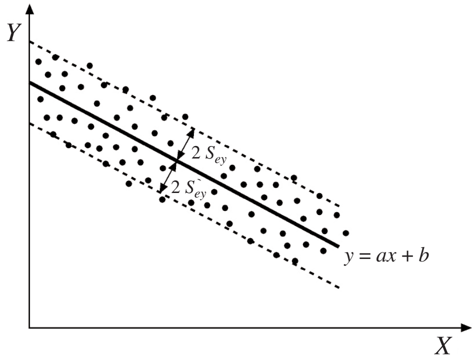
Gainera, $r > 0$ denean dependentzia gorakorra da eta $r < 0$ denean dependentzia beherakorra da.

9.4. ESTIMAZIO-ERRORE ESTANDARRA

Erregresio linealaren azterketan, **estimazio-errore estandarra** honela definitzen da:

$$S_{ey} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - ax_i - b)^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} y_i}{n-2}}.$$

Estimazio-errore estandarraren bidez, oro har, erregresio-zuzeneko $\hat{y}$ balioen eta puntu-hodeiko y balioen arteko errorea azter daiteke. $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ laginaren tamaina $n \geq 30$ denean, puntu-hodeiko % 95 $(-2S_{ey}, 2S_{ey})$ mendelean dago.



9.2. irudia.

9.5. MINIMO KARRATUEN METODOKO ESTIMATZAILEEI BURUZKO INFERENTZIAK

Azpimarratzekoa da $y = ax + b$ erregresio-zuzena $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ lagineko datuak erabiliz lortu dela. Helburua zuzen hau populazio osoari dagokion eredu linealaren adierazgarri izatea da. Hona hemen lan honetan lagungarri diren zenbait konfiantza-tarte:

9.5.1. Populazioaren α erregresio-koefizienterako konfiantza-tartea

α koefizientea estimatzeko $1-\gamma$ konfiantza-mailako konfiantza-tartea hauxe da:

$$I_\alpha^{1-\gamma} = \left[a - t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \bar{x})^2}},\ a + t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \bar{x})^2}} \right].$$

9.5.2. β koefizienterako konfiantza-tartea

β koefizientea estimatzeko $1-\gamma$ konfiantza-mailako konfiantza-tartea ondokoa da:

$$I_\beta^{1-\gamma} = \left[b - t_{\gamma/2;n-2}S_{ey}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}},\ b + t_{\gamma/2;n-2}S_{ey}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \right].$$

9.5.3. $y = \alpha x_0 + \beta$ ordenaturako konfiantza-tartea

$x = x_0$ abzisari dagokion ordenatua estimatzeko $1 - \gamma$ konfiantza-mailako konfiantza-tartea hauxe da:

$$I^{1-\gamma}_{\alpha x_0+\beta} = \left[a x_0 + b - t_{\gamma/2;n-2} S_{ey} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}},\ a x_0 + b + t_{\gamma/2;n-2} S_{ey} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \right].$$

9.5.4. Populazioaren ρ korrelazio-koefizienterako konfiantza-tartea

Populazioaren ρ korrelazio-koefizientea estimatzeko $1 - \gamma$ konfiantza-mailako konfiantza-tartea hau da:

$$I^{1-\gamma}_{\rho} = \left[\frac{1}{2} \ln \frac{1+r}{1-r} - z_{\gamma/2} \sqrt{\frac{1}{n-3}},\ \frac{1}{2} \ln \frac{1+r}{1-r} + z_{\gamma/2} \sqrt{\frac{1}{n-3}} \right]$$

non r gaia laginaren korrelazio-koefizientea, n laginaren tamaina eta $z_{\gamma/2}$ gaia banaketa normal estandarraren taulako balioa diren.

9.6. ARIKETA EBATZIAK

9.6.1. ariketa

Ondoko taulan zenbait pieza bukatzeko behar izan den denbora adierazi da:

x (egun)	70	65	77	70	67	62
y (pieza)	127	112	146	131	116	97

a. Lor bedi korrelazio-koefizientearen balioa.

b. Doi ezazu puntu-hodeia X gaineko Y-ren erregresio-zuzenera.

c. Zenbat pieza bukatuko dira 80 egunetan?

Ebazpena

a. Hasteko, taula batean Pearson-en korrelazio-koefizientearen balioa kalkulatzeko lagungarri izango diren gaiak adieraziko dira:

x	y	xy	x^2	y^2
70	127	8.890	4.900	16.129
65	112	7.280	4.225	12.544
77	146	11.242	5.929	21.316
70	131	9.170	4.900	17.161
67	116	7.772	4.489	13.456
62	97	6.014	3.844	9.409
411	729	50.368	28.287	90.015

Pearson-en korrelazio-koefizientearen balioa hurrengoa da:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{50.368 - \frac{1}{6}\cdot 411 \cdot 729}{\sqrt{28.287 - \frac{1}{6}(411)^2}\sqrt{90.015 - \frac{1}{6}(729)^2}} =$$

$$= \frac{431,5}{11,55 \cdot 37,97} = 0,984.$$

Beraz, korrelazio-koefizientearen balioa 1etik hurbil dagoenez, x eta y datuen artean erlazio lineal handia dago.

b. Zehatz bedi $y = ax + b$ erregresio-zuzena.

Hasteko a koefizientearen balioa kalkulatuko da.

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{50.368 - \frac{1}{6}\cdot 411 \cdot 729}{28.287 - \frac{1}{6}\cdot(411)^2} = \frac{431,5}{133,5} = 3,23.$$

Bestalde, b koefizientearen balioa kalkulatzeko

$$b = \overline{y} - a\overline{x}$$

adierazpena erabiliko da eta hurrengo emaitza lortuko da:

$$b = \overline{y} - a\overline{x} = \frac{729}{6} - 3,23 \cdot \frac{411}{6} = -99,76.$$

Ondorioz, erlazio lineala adierazten duen erregresio-zuzena hauxe da:

$$y = 3,23\, x - 99,76.$$

c. Erregresio-zuzenaren adierazpenean $x = 80$ egun ordezkatuta,

$$y = 3,23 \cdot 80 - 99,76 = 158,64$$

gutxi gorabehera 159 pieza bukatuko dira.

9.6.2. ariketa

X aldagaiak A enpresaren azken zazpi hilabeteetako irabaziak (milaka euro) eta *Y* aldagaiak B enpresaren azken zazpi hilabeteetako irabaziak (milaka euro) adierazten dituzte.

x	11,1	10,3	12,0	15,1	13,7	17,3	18,5
y	10,9	12,2	13,8	16,0	13,2	21,1	16,4

a. Datu horien artean, ba al dago erlazio linealik?

b. Lor bedi *X* gaineko *Y*-ren erregresio-zuzena.

c. B enpresak hilabete batean 15 mila euro irabazten baditu, zenbatekoak dira A enpresaren irabaziak hilabete horretan?

Ebazpena

a. Hurrengo taulan Pearson-en korrelazio-koefizientea kalkulatzeko behar diren balioak daude:

x	y	xy	x^2	y^2
11,1	10,9	120,99	123,21	118,81
10,3	12,2	125,66	106,09	148,84
12,0	13,8	165,60	144,00	190,44
15,1	16,0	241,60	228,01	256,00
13,7	13,2	180,84	187,69	174,24
17,3	21,1	365,03	299,29	445,21
18,5	16,4	303,40	342,25	268,96
98,0	103,6	1.503,12	1.430,54	1.602,50

Pearson-en korrelazio-koefizientearen balioa:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{51.503{,}02 - \frac{1}{7}\cdot 98 \cdot 103{,}6}{\sqrt{1.430{,}54 - \frac{1}{7}(98)^2}\sqrt{1.602{,}5 - \frac{1}{7}(103{,}6)^2}} =$$

$$= \frac{52{,}72}{7{,}65 \cdot 8{,}32} = 0{,}828.$$

Beraz, korrelazio-koefizientearen balioa 1etik hurbil dagoenez, X eta Y aldagaien artean erlazio lineal handia dago.

b. Lor dezagun X gaineko Y-ren erregresio-zuzenaren adierazpena.

$$y = ax + b.$$

a koefizientearen kalkulua:

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{1.503,12 - \frac{1}{7}\cdot 98 \cdot 103,6}{1.430,54 - \frac{1}{7}\cdot (98)^2} = \frac{52,72}{58,54} = 0,90.$$

b koefizientearen kalkulurako

$$b = \bar{y} - a\bar{x}$$

adierazpena erabiliko da. Oraingoan b koefizientearen balioa 2,2 da.

$$b = \bar{y} - a\bar{x} = \frac{103,6}{7} - 0,90 \cdot \frac{98}{7} = 2,2.$$

Ondorioz, erlazio lineala adierazten duen erregresio-zuzena hauxe da:

$$y = 0,9\,x + 2,2.$$

c. Aurreko erregresio-zuzenean $y = 15$ ordezkatuz,

$$15 = 0,9\,x + 2,2$$

$$x = \frac{15 - 2,2}{0,9} = 14,22.$$

A enpresaren irabaziak hilabete horretan gutxi gorabehera 14.220 eurokoak izatea itxaron daiteke.

9.6.3. ariketa

Hurrengo taulan 10 nahasteren karbonoaren edukia eta iragazkortasun-indizeak azaltzen dira:

Karbono-edukia (%)	4,5	4,3	4,2	4,4	4,1	4,8	5	4,9	4,7	4,6
Iragazkortasun-indizea	15	18	12	16	17	11	8	4	6	7

a. Kalkula bedi r koefizientearen balioa.

b. Lor bedi populaziorako korrelazio-koefizientearen % 95eko konfiantza-mailako konfiantza-tartea.

Ebazpena

a. Hurrengo taulan Pearson-en korrelazio linealaren koefizientearen kalkulurako balioak daude:

x	y	xy	x^2	y^2
4,5	15	67,5	20,25	225
4,3	18	77,4	18,49	324
4,2	12	50,4	17,64	144
4,4	16	70,4	19,36	256
4,1	17	69,7	16,81	289
4,8	11	52,8	23,04	121
5	8	40,0	25,00	64
4,9	4	19,6	24,01	16
4,7	6	28,2	22,09	36
4,6	7	32,2	21,16	49
45,5	114	508,2	207,85	1.524

Ondorioz, kasu honetan Pearson-en korrelazio-koefizientearen balioa hauxe da:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{508{,}2 - \frac{1}{10}\cdot 45{,}5\cdot 114}{\sqrt{207{,}85 - \frac{1}{10}\cdot(45{,}5)^2}\sqrt{1.524 - \frac{1}{10}\cdot(114)^2}} =$$

$$= \frac{-10{,}5}{0{,}91\cdot 14{,}98} = -0{,}77.$$

b. Populazioaren ρ korrelazio-koefizientea estimatzeko $1-\gamma$ konfiantza-mailako konfiantza-tartea hau da:

$$I_\rho^{1-\gamma} = \left[\frac{1}{2}\ln\frac{1+r}{1-r} - z_{\gamma/2}\sqrt{\frac{1}{n-3}}, \frac{1}{2}\ln\frac{1+r}{1-r} + z_{\gamma/2}\sqrt{\frac{1}{n-3}}\right].$$

Kasu honetan, $r = -0{,}77$, $n = 10$, $1-\gamma = 0{,}95$ dira eta banaketa normal tipikoaren taulatik $z_{\gamma/2} = z_{0{,}025} = 1{,}96$ balioa aterako da. Ondorioz, balio horiek ordezkatuz, ρ estimatzeko % 95eko konfiantza-mailako konfiantza-tartea lor daiteke.

$$I_\rho^{0{,}95} = \left[\frac{1}{2}\ln\frac{1-0{,}77}{1+0{,}77} - 1{,}96\sqrt{\frac{1}{7}}, \frac{1}{2}\ln\frac{1-0{,}77}{1+0{,}77} + 1{,}96\sqrt{\frac{1}{7}}\right] =$$

$$= [-1{,}02 - 0{,}74, -1{,}02 + 0{,}74] =$$

$$= [-1{,}76, -0{,}28].$$

9.6.4. ariketa

A eta B garbiketa-produktuen salmenten artean erlazio linealik dagoen jakiteko, aleatorioki produktu horien bost hilabetetako salmentak hartu dira. Hurrengo taulan X aldagaiaren bidez A produktuaren salmentak eta Y aldagaiaren bidez B produktuaren salmentak adierazi dira:

x	102	110	120	108	112
y	200	198	250	205	212

a. Kalkula bitez r korrelazio-koefizientea eta X gaineko Y-ren erregresio-zuzena.

b. Lor bedi errore estandarraren balioa.

Ebazpena

a. Kalkula dezagun Pearson-en korrelazio linealaren r koefizientea.

x	y	xy	x^2	y^2
102	200	20.400	10.404	40.000
110	198	21.780	12.100	39.204
120	250	30.000	14.400	62.500
108	205	22.140	11.664	42.025
112	212	23.744	12.544	44.944
552	1.065	118.064	61.112	228.673

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{118.064 - \frac{1}{5}\cdot 552 \cdot 1.065}{\sqrt{61.112 - \frac{1}{5}\cdot(552)^2}\sqrt{228.673 - \frac{1}{5}\cdot(1.065)^2}} =$$

$$= \frac{488}{13,08 \cdot 42,76} = 0,87.$$

Korrelazio-koefizientearen balioak puntu-hodeirako erlazio lineal handia erakusten du. Beraz, zentzuzkoa da $y = ax + b$ erregresio-zuzena proposatu eta zehaztea.

Erregresio-zuzeneko koefizienteen kalkulua:

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{118.064 - \frac{1}{5}\cdot 552\cdot 1.065}{61.112 - \frac{1}{5}\cdot (552)^2} = \frac{488}{171,2} = 2,85$$

$$b = \overline{y} - a\overline{x} = \frac{1.065}{5} - 2,85\cdot\frac{552}{5} = -101,64.$$

Ondorioz, X gaineko Y-ren erregresio-zuzena hauxe da:

$$y = 2,85\,x - 101,64.$$

b. Estimazio-errore estandarra:

$$S_{ey} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} y_i}{n-2}} =$$

$$= \sqrt{\frac{228.673 - 2,85\cdot 118.064 + 101,64\cdot 1.065}{3}} =$$

$$= \sqrt{\frac{473,2}{3}} = 12,07\,.$$

9.6.5. ariketa

Erreakzio kimiko jakin baten zortzi saiakuntzetan hurrengo tenperatura eta presioak neurtu dira:

x, tenperatura (ºC)	33	30	20	25	29	24	34	22
y, presioa (Pa)	100	99	75	80	90	85	111	70

a. Lor bedi Y gaineko X aldagaiaren erregresio-zuzena.

b. 105 Pa eta 110 Pa bitarteko presioetarako zein tenperatura itxaron daiteke?

Ebazpena

a. Ondoko taulan geroko eragiketetarako beharrezko datuak kalkulatu dira:

x	y	xy	x^2	y^2
33	100	3.300	1.089	10.000
30	99	2.970	900	9.801
20	75	1.500	400	5.625
25	80	2.000	625	6.400
29	90	2.610	841	8.100
24	85	2.040	576	7.225
34	111	3.774	1.156	12.321
22	70	1.540	484	4.900
217	710	19.734	6.071	64.372

$x = cy + d$ erregresio-zuzenaren determinazioa, c eta d koefizienteak zehaztuz lortuko da.

$$c = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2} = \frac{19.734 - \frac{1}{8}\cdot 217\cdot 710}{64.372 - \frac{1}{8}\cdot(710)^2} = \frac{475,25}{1.359,5} = 0,35$$

$$d = \bar{x} - c\bar{y} = \frac{217}{8} - 0,35\cdot\frac{710}{8} = -3,94.$$

Ondorioz, Y gaineko X-ren erregresio-zuzena hau da:

$$x = 0,35\, y - 3,94.$$

b. Erregresio-zuzenaren adierazpenean $y_1 = 105$ Pa ordezkatuta,

$$x_1 = 0,35 \cdot 105 - 3,94 = 32,81 \text{ ºC}$$

lortzen da. Era berean, $y_2 = 110$ Pa baliorako

$$x_2 = 0,35 \cdot 110 - 3,94 = 34,56 \text{ ºC}$$

lortuko da. Beraz, 105 Pa eta 110 Pa bitarteko presioetarako 32,81 ºC eta 34,56 ºC bitarteko tenperaturak itxaron daitezke.

9.6.6. ariketa

Aleatorioki hartutako 5 zilindroren diametroen luzerak eta zilindroak apurtzeko behar diren indarrak honakoak dira:

x (cm)	10	20	18	22	26
y (kg)	92	96	94	100	110

non X aldagaiaren bidez zilindroen diametroen luzerak eta Y aldagaiaren bidez zilindroak apurtzeko indarra adierazi diren.

a. Datu hauetan oinarrituz, baiezta al daiteke korrelazio positiboa dagoela? Arrazoi ezazu erantzuna.

b. Zenbateko indarra aplikatu behar litzateke 35 cm-ko diametrodun zilindroa apurtzeko?

Ebazpena

a. Taulan emaniko datuak ikusita honakoa esan daiteke: zilindroak zenbat eta diametro handiagoa izan, hainbat eta indar handiago aplikatu behar da bera apurtzeko. Hala ere, datuek erakusten dutena r korrelazio-koefizientearen bidez zehatz daiteke.

x	y	xy	x^2	y^2
10	92	920	100	8.464
20	96	1.920	400	9.216
18	94	1.692	324	8.836
22	100	2.200	484	10.000
26	110	2.860	676	12.100
96	492	9.592	1.984	48.616

Pearson-en korrelazio-koefizientearen balioa hurrengoa da:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{9.592 - \frac{1}{5}\cdot 96\cdot 492}{\sqrt{1.984 - \frac{1}{5}\cdot(96)^2}\sqrt{48.616 - \frac{1}{5}\cdot(492)^2}} =$$

$$= \frac{145,6}{11,87\cdot 14,25} = 0,86.$$

Korrelazio-koefizientearen balioa positiboa denez eta unitatetik hurbil dagoenez, datu hauetan oinarrituz, X eta Y aldagaien artean erlazio lineal positiboa dagoela esan daiteke.

b. 35 cm-ko diametrodun zilindroa apurtzeko zenbateko indarra aplikatu behar den jakiteko, lehenengo, X gaineko Y-ren erregresio-zuzena

$$y = ax + b$$

lortu behar da.

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{9.592 - \frac{1}{5}\cdot 96\cdot 492}{1.984 - \frac{1}{5}\cdot(96)^2} = \frac{145,6}{140,8} = 1,03$$

$$b = \bar{y} - a\bar{x} = \frac{492}{5} - 1,03\cdot\frac{96}{5} = 78,62.$$

Ondorioz, X gaineko Y-ren erregresio-zuzena hauxe da:

$$y = 1,03\, x + 78,62.$$

Erregresio-zuzenean $x = 35$ cm ordezkatuz, zilindroa apurtzeko zenbateko indarra aplikatu behar den lortuko da:

$$y = 1,03 \cdot 35 + 78,62 = 114,67 \text{ kg}$$

9.6.7. ariketa

Ibilgailu-mota zehatz baten abiaduraren arabera CO_2-aren emisioak aztertu dira:

x, abiadura(km/h)	50	70	80	90	100	120
y, CO_2 emisioak (g/km)	147	149	171	180	193	200

Abiadura aldagai asketzat harturik, estima bedi erregresio-zuzena % 99ko konfiantza-mailaz.

Ebazpena

Ariketa honetan $Y = \alpha X + \beta$ erlazio funtzionala ezarriko da, non α eta β koefizienteak estimatuko diren. Horretarako, enuntziatuan emaniko datuetara doituko den $y = ax + b$ erregresio-zuzena zehaztu behar da. Lagungarria izango den hurrengo taula erabil daiteke:

x	y	xy	x^2	y^2	$(x-\bar{x})^2$
50	147	7.350	2.500	21.609	1.225
70	149	10.430	4.900	22.201	225
80	171	13.680	6.400	29.241	25
90	180	16.200	8.100	32.400	25
100	193	19.300	10.000	37.249	225
120	200	24.000	14.400	40.000	1.225
510	1.040	90.960	46.300	182.700	2.950

Pearson-en korrelazio-koefizientearen balioa:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{90.960 - \frac{1}{6}\cdot 510 \cdot 1.040}{\sqrt{46.300 - \frac{1}{6}\cdot(510)^2}\sqrt{182.700 - \frac{1}{6}\cdot(1.040)^2}} =$$

$$= \frac{2.560}{54,31 \cdot 49,33} = 0,96.$$

$y = ax + b$ erregresio-zuzena zehazteko, a eta b koefizienteen balioak kalkulatuko dira.

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{90.960 - \frac{1}{6}\cdot 510\cdot 1.040}{46.300 - \frac{1}{6}\cdot(510)^2} = \frac{2.560}{2.950} = 0,87$$

$$b = \bar{y} - a\bar{x} = \frac{1.040}{6} - 0,87\cdot\frac{510}{6} = 99,38.$$

Ondorioz, X gaineko Y-ren erregresio-zuzena hauxe da:

$$y = 0,87\, x + 99,38.$$

a eta b koefizienteen balioak ezagututa, kalkula ditzagun α eta β estimatzeko % 99ko konfiantza-mailako konfiantza-tarteak.

$$I_\alpha^{1-\gamma} = \left[a - t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \bar{x})^2}},\ a + t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \bar{x})^2}} \right]$$

$$I_{\beta}^{1-\gamma} = \left[b - t_{\gamma/2;n-2} S_{ey} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}},\ \ b + t_{\gamma/2;n-2} S_{ey} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \right].$$

Bi adierazpenetan estimazio-errore estandarra dago. Beraz, lehenengo S_{ey} errore estandarraren balioa kalkulatuko da.

$$S_{ey} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} y_i}{n-2}} =$$

$$= \sqrt{\frac{182.700 - 0{,}87 \cdot 90.960 - 99{,}38 \cdot 1.040}{3}} =$$

$$= \sqrt{\frac{209{,}6}{4}} = 7{,}24.$$

Bestalde, bi konfiantza-tarteetan Student-en balio bat dago, $t_{0,005;4}$ hain zuzen ere. Student-en taulan kontsultatuz, $t_{0,005;4} = 4{,}604$ dela froga daiteke.

Orduan, α koefizientea estimatzeko % 99ko konfiantza-mailako konfiantza-tartea hauxe da:

$$I_{\alpha}^{0,99} = \left[0{,}87 - t_{0,005;4} \frac{7{,}24}{(2.950)^{1/2}},\ 0{,}87 + t_{0,005;4} \frac{7{,}24}{(2.950)^{1/2}} \right] =$$

$$= [0{,}87 - 4{,}604 \cdot 0{,}13,\ \ 0{,}87 + 4{,}604 \cdot 0{,}13] =$$

$$= [0{,}87 - 0{,}60,\ \ 0{,}87 + 0{,}60] =$$

$$= [0{,}27,\ \ 1{,}47].$$

Era berean, β koefizientea estimatzeko % 99ko konfiantza-mailako konfiantza-tartea ondokoa da:

$$I_{\beta}^{0,99} = \left[99{,}38 - 4{,}604 \cdot 7{,}24 \cdot \sqrt{\frac{1}{6} + \frac{(85)^2}{2.950}},\ \ 99{,}38 + 4{,}604 \cdot 7{,}24 \cdot \sqrt{\frac{1}{6} + \frac{(85)^2}{2.950}} \right] =$$

$$= [99{,}38 - 4{,}604 \cdot 7{,}24 \cdot 1{,}62,\ \ 99{,}38 + 4{,}604 \cdot 7{,}24 \cdot 1{,}62] =$$

$$= [99{,}38 - 54,\ \ 99{,}38 + 54] =$$

$$= [45{,}38,\ 153{,}38].$$

9.6.8. ariketa

Bost haberi x tentsioa aplikatu zaie eta presio horren ondorioz habeek y luzapena jaso dute.

x (milaka libra)	1	2	3	4	5
y (hazbete)	12	29	38	60	75

a. Doi bedi erregresio-zuzena.

b. % 99ko konfiantza-mailaz, estima bedi populaziorako korrelazio-koefizientea.

Ebazpena

a. X gaineko Y-ren erregresio-zuzena lortzeko, $y = ax + b$ zuzeneko koefizienteak zehaztu behar dira. Horretarako,

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b = \overline{y} - a\overline{x}$$

adierazpenak erabiliko dira. Bi koefiziente hauen balioak lortzeko, aldez aurretik zenbait eragiketa egingo dira.

x	y	xy	x^2	y^2
1	12	12	1	144
2	29	58	4	841
3	38	114	9	1.444
4	60	240	16	3.600
5	75	375	25	5.625
15	214	799	55	11.654

a koefizientearen kalkulua:

$$a = \frac{799 - \frac{1}{5}\cdot 15 \cdot 214}{55 - \frac{1}{5}\cdot (15)^2} = \frac{157}{10} = 15,7.$$

b koefizientearen kalkulua:

$$b = \frac{214}{5} - 15,7 \cdot \frac{15}{5} = -4,3.$$

Ondorioz, erregresio-zuzena hau da:

$$y = 15{,}7\,x - 4{,}3.$$

b. Populaziorako ρ korrelazio-koefizientearen estimazioarekin hasi aurretik, laginerako Pearson-en r korrelazio-koefizientearen balioa:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{799 - \frac{1}{5}\cdot 15\cdot 214}{\sqrt{55 - \frac{1}{5}(15)^2}\sqrt{11.654 - \frac{1}{5}\cdot(214)^2}} =$$

$$= \frac{157}{3{,}16\cdot 49{,}95} = 0{,}99.$$

Populazioaren ρ korrelazio-koefizientea estimatzeko $1-\gamma$ konfiantza-mailako konfiantza-tartea hauxe da:

$$I_{\alpha}^{1-\gamma} = \left[\frac{1}{2}\ln\frac{1+r}{1-r} - z_{\gamma/2}\sqrt{\frac{1}{n-3}},\ \frac{1}{2}\ln\frac{1+r}{1-r} + z_{\gamma/2}\sqrt{\frac{1}{n-3}}\right].$$

$r = 0{,}99$, $n = 5$, $1-\gamma = 0{,}99$ eta taula normal tipikotik ateratako $z_{\gamma/2} = z_{0{,}005} = 2{,}575$ balioa ordezkatuz, ρ estimatzeko % 99ko konfiantza-mailako konfiantza-tartea honakoa da:

$$I_{\rho}^{0{,}99} = \left[\frac{1}{2}\ln\frac{1+0{,}99}{1-0{,}99} - 2{,}575\cdot\sqrt{\frac{1}{2}},\ \frac{1}{2}\ln\frac{1+0{,}99}{1-0{,}99} + 2{,}575\cdot\sqrt{\frac{1}{2}}\right] =$$

$$= [2{,}65 - 1{,}82,\ 2{,}65 + 1{,}82] =$$

$$= [0{,}83,\ 4{,}47].$$

9.6.9. ariketa

Hurrengo taulan azaltzen diren datuak produktu baten eskaera eta sei merkatu ezberdinetan dituen prezioak dira:

x, eskaera (produktu)	20	15	9	12	14	13
y, prezioa (milaka euro)	25	44	130	90	50	70

a. Doi bitez datu hauek eskaeraren gaineko prezioaren erregresio-zuzenera.

b. % 95eko konfiantza-mailaz, lor bedi eskaera 25 produktukoa denerako prezioaren aurresanaren estimazioa.

Ebazpena

a. Hasteko, erregresio-zuzena lortu aurretik, Pearson-en korrelazio linealaren koefizientearen bidez, ikus dezagun erregresio lineala kasu honetan egokia dela.

Hurrengo taulako datuak ondoko kalkuluetarako baliagarriak izango dira:

x	y	xy	x^2	y^2	$(x-\bar{x})^2$
20	25	500	400	625	1.308,03
15	44	660	225	1.936	3.154,69
9	130	1.170	81	16.900	4.378,03
12	90	1.080	144	8.100	5.801,36
14	50	700	196	2.500	7.424,69
13	70	910	169	4.900	11.271,36
83	409	5.020	1.215	34.961	33.338,17

Pearson-en korrelazio-koefizientearen balioa:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{5.020 - \frac{1}{6}\cdot 83 \cdot 409}{\sqrt{1.215 - \frac{1}{6}\cdot(83)^2}\sqrt{34.961 - \frac{1}{6}\cdot(409)^2}} =$$

$$= \frac{-637,83}{8,18\cdot 84,15} = -0,93.$$

Korrelazio-koefizientearen balioak produktuaren prezioaren eta eskaeraren artean erlazio lineal handia eta negatiboa dagoela adierazten du. Beraz, zentzuzkoa da X gaineko Y aldagaiaren erregresio-zuzena planteatzea. Orduan, $y = ax + b$ zuzeneko a eta b koefizienteak kalkulatu behar dira.

a koefizientearen balioa:

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{5.020 - \frac{1}{6}\cdot 83\cdot 409}{1.215 - \frac{1}{6}\cdot(83)^2} = \frac{-637,83}{66,83} = -9,54.$$

b koefizientearen balioa:

$$b = \bar{y} - a\bar{x} = \frac{409}{6} + 9,54 \cdot \frac{83}{6} = 200,14.$$

Ondorioz, eskaeraren gaineko prezioaren erregresio-zuzena hauxe da:

$$y = -9,54x + 200,14.$$

b. Kalkula dezagun $x_0 = 25$ abzisari dagokion ordenatua estimatzeko % 95eko konfiantza-mailako konfiantza-tartea. Horri dagokion adierazpen orokorra honako hau da:

$$I^{1-\gamma}_{\alpha x_0+\beta} = \left[ax_0 + b \mp t_{\gamma/2;n-2} S_{ey} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sqrt{(x_i - \bar{x})^2}}} \right].$$

Adierazpen horretan estimazio-errore estandarra dago. Beraz, lehenengo S_{ey} errore estandarraren balioa kalkulatuko da.

$$S_{ey} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} y_i}{n-2}} =$$

$$= \sqrt{\frac{34.961 + 9,54 \cdot 5.020 - 200,14 \cdot 409}{4}} =$$

$$= \sqrt{\frac{994,54}{4}} = 15,77.$$

Gainera, kalkulatu behar den konfiantza-tartean Student-en balio bat dago, $t_{0,025;4}$ hain zuzen ere. Student-en taulan kontsultatuz, $t_{0,025;4} = 2,776$ dela nabari daiteke.

Orduan, $x_0 = 25$ abzisari dagokion ordenatua estimatzeko % 95eko konfiantza-mailako konfiantza-tartea hauxe da:

$$I^{1-\gamma}_{\alpha 25+\beta} = \left[-9,54 \cdot 25 + 200,14 \mp 2,776 \cdot 15,77 \sqrt{\frac{1}{6} + \frac{(25 - 13,83)^2}{33.338,17}} \right] =$$

$$= [-9,54 \cdot 25 + 200,14 \mp 2,776 \cdot 15,77 \cdot 0,41] =$$

$$= [-38,36 - 17,95, -38,36 + 17,95] =$$

$$= [-56,31, -20,41].$$

9.6.10. ariketa

Hiri batean sei hilabetetan jaso den euri-kantitatea (l/m²) eta lurrunketa (l/m²) honakoak dira:

x, euri-kantitatea	89,3	78,4	40,3	221,4	95,7	70,3
y, lurrunketa	89	82,2	87,0	178,9	119,8	82,5

a. Lor bedi X gaineko Y-ren erregresio-zuzena.

b. % 90eko konfiantza-mailaz, estima bitez α eta β koefizienteak.

Ebazpena

a. Eskatutako erregresio-zuzena lortu aurretik, Pearson-en korrelazio linealaren koefizientearen balioa lortuko da.

Ondoko taulako balioak ondoren egingo diren kalkuluetarako baliagarriak izango dira:

x	y	xy	x^2	y^2	$(x-\bar{x})^2$
89,3	89,0	7.947,70	7.974,49	7.921,00	98,67
78,4	82,2	6.444,48	6.146,56	6.756,84	434,03
40,3	87,0	3.506,10	1.624,09	7.569,00	3.473,14
221,4	178,9	39.608,46	49.017,96	32.005,21	14.924,69
95,7	119,8	11.464,86	9.158,49	14.352,04	12,48
70,3	82,5	5.799,75	4.942,09	6.806,25	837,14
595,4	639,4	74.771,35	78.863,68	75.410,34	19.780,15

Pearson-en korrelazio-koefizientearen balioa:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2}} =$$

$$= \frac{74.771,35 - \frac{1}{6}\cdot 595,4 \cdot 639,4}{\sqrt{78.863,68 - \frac{1}{6}\cdot(595,4)^2}\sqrt{75.410,34 - \frac{1}{6}\cdot(639,4)^2}} =$$

$$= \frac{11.321,56}{140,64 \cdot 85,27} = 0,9.$$

Korrelazio-koefizientearen balioak erlazio lineal handia dagoela adierazten du. Ondorioz, X gaineko Y aldagaiaren $y = ax + b$ erregresio-zuzena planteatuko da, non a eta b koefizienteak zehaztu behar diren.

a koefizientearen kalkulua:

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{74.771,35 - \frac{1}{6}\cdot 595,4\cdot 639,4}{78.863,68 - \frac{1}{6}\cdot (595,4)^2} = \frac{11.321,56}{19.780,15} = 0,57.$$

b koefizientearen kalkulua:

$$b = \overline{y} - a\overline{x} = \frac{639,4}{6} - 0,57\cdot\frac{595,4}{6} = 50.$$

Ondorioz, X gaineko Y-ren erregresio-zuzena honakoa da:

$$y = 0,57x + 50.$$

b. α eta β estimatzeko % 90eko konfiantza-mailako konfiantza-tarteen adierazpen orokorrak hauek dira:

$$I_{\alpha}^{1-\gamma} = \left[a - t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \overline{x})^2}},\ a + t_{\gamma/2;n-2}\frac{S_{ey}}{\sqrt{\sum (x_i - \overline{x})^2}} \right]$$

$$I_{\beta}^{1-\gamma} = \left[b - t_{\gamma/2;n-2}S_{ey}\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{\sqrt{\sum (x_i - \overline{x})^2}}},\ b + t_{\gamma/2;n-2}S_{ey}\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{\sqrt{\sum (x_i - \overline{x})^2}}} \right]$$

$t_{0,05;4} = 2,132$ balioa Student-en taulatik lor daiteke.

Bi konfiantza-tarteetan S_{ey} estimazio-errore estandarra azaltzen denez, estimazio-errore estandarraren balioa behar izango da.

$$S_{ey} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - a\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} y_i}{n-2}} =$$

$$= \sqrt{\frac{75.410,34 - 0,57\cdot 74.771,35 - 50\cdot 639,4}{4}} =$$

$$= \sqrt{\frac{820,67}{4}} = 14,32.$$

Orduan, α koefizientea estimatzeko % 90eko konfiantza-mailako konfiantza-tartea hauxe da:

$$I_{\alpha}^{0,90} = \left[0,57 - 2,132 \cdot \frac{14,32}{\sqrt{19.780,15}},\quad 0,57 + 2,132 \cdot \frac{14,32}{\sqrt{19.780,15}}\right] =$$

$$= [0,57 - 2,132 \cdot 0,10,\ 0,57 + 2,132 \cdot 0,10] =$$

$$= [0,57 - 0,2132,\ 0,57 + 0,2132] =$$

$$= [0,3568,\ 0,7832].$$

Era berean, β koefizienterako % 90eko konfiantza-mailako konfiantza-tartea ondokoa da:

$$I_{\beta}^{0,90} = \left[50 - 2,132 \cdot 14,32 \cdot \sqrt{\frac{1}{6} + \frac{(99,23)^2}{19.780,15}},\quad 50 + 2,132 \cdot 14,32 \cdot \sqrt{\frac{1}{6} + \frac{(99,23)^2}{19.780,15}}\right] =$$

$$= [50 - 2,132 \cdot 14,32 \cdot 0,82,\ 50 + 2,132 \cdot 14,32 \cdot 0,82] =$$

$$= [50 - 25,03,\ 50 + 25,03] =$$

$$= [24,97,\ 75,03].$$

9.7. ARIKETA PROPOSATUAK

9.7.1. ariketa

Ondoko taulan 1996-2000 denboraldiko etxe bateko elektrizitate-kontsumoa adierazten da:

x (urtea)	1996	1997	1998	1999	2000
y (kW)	360	370	390	380	400

a. Kalkula bedi r korrelazio-koefizientea.

b. Zein elektrizitate-kontsumo itxaron daiteke 2001. urterako?

9.7.2. ariketa

Hurrengo taulan X eta Y aldagaiek I. eta II. motako makinek informazioa prozesatzeko behar dituzten denborak adierazten dituzte, hurrenez hurren:

x (minutu)	11,1	10,3	12,0	15,1	13,7	14,0
y (minutu)	10,9	12,0	13,8	21,5	13,2	16,0

a. Datu horien artean, ba al dago erlazio linealik?

b. Lor bedi Y gaineko X-ren erregresio-zuzena.

9.7.3. ariketa

Aleatorioki aukeratu diren 7 langileren bidez, formazio-ikastaro baten eraginkortasuna kontrastatu nahi da. Horretarako langile horien formazio-ikastaroko kalifikazioak eta produktibitate-indizea hartu dira.

x (ikastaroko kalifikazioak)	50	40	37	46	25	22	39
y (produktibitate-indizea)	0,96	0,78	0,80	0,90	0,77	0,76	0,85

a. Zehatz bedi X gaineko Y-ren erregresio-zuzena.

b. Kalkula bitez α eta β koefizienteetarako % 95eko konfiantza-mailako konfiantza-tarteak.

9.7.4. ariketa

Trafiko-fluxu handia duen errepide batean, auto-istripuei eta euri-egunei buruzko ondoko datuak bildu dira:

Auto-istripuak	5	8	12	20	15
Euri-egunak	1	3	2	5	4

a. Aurki bedi auto-istripuen eta euri-egunen arteko korrelazioa.

b. Kalkula bedi populazioaren korrelazio-koefizientea estimatzeko % 95eko konfiantza-mailako konfiantza-tartea.

9.7.5. ariketa

Hurrengo taulan X eta Y aldagaiek enpresa baten irabaziak eta kaleratutako langileen kopurua adierazten dituzte, hurrenez hurren:

x (milaka euro)	250	180	120	100	50
y (langile)	0	5	8	10	20

a. Kalkula bitez r korrelazio-koefizientea eta X gaineko Y-ren erregresio-zuzena.

b. Lor bedi errore estandarraren balioa.

9.7.6. ariketa

Produktu jakin baten lotea ekoizteko kostua, lotearen tamainaren araberakoa da. Ondoko taulan bost loteren tamainak (produktu-kopurua) eta kostuak (euro) adierazten dira:

x (produktu-kopurua)	1	10	45	90	150
y (euro)	30	150	500	1.000	2.000

a. Doi bitez datu horiek lotearen taminaren gaineko kostuaren erregresio-zuzenera.

b. % 99ko konfiantza-mailaz, lor bedi lotea 200 produktukoa denerako kostuaren aurresanaren estimazioa.

9.7.7. ariketa

Ondoko datuek makina batean ekoitzitako pieza akastunen kopurua eta horiek egiteko eguneko ordua erakusten dituzte:

x (pieza akastunak)	5	7	4	9	12	16
y (ordua)	8:00	14:00	18:00	21:00	23:00	24:00

a. Lor bedi errore estandarraren balioa.

b. Kalkula bitez α eta β koefizienteetarako % 95eko konfiantza-mailako konfiantza-tarteak.

9.7.8. ariketa

Prozesu kimiko jakin batean tenperaturak eta pH-ak erlazioa dutelako susmoa dago. Hurrengo taulan bost prozesu kimikotan jasotako datuak azaltzen dira:

x (pH)	2	0	1	4	3
y (ºC)	36	38	40	33	35

a. Taulako datuak erabiliz, doi bedi erregresio-zuzena.

b. % 95eko konfiantza-mailaz, estima bedi populaziorako korrelazio-koefizientea.

Ariketa proposatuen emaitzak

1. GAIA. ESTATISTIKA DESKRIBATZAILEA. ALDAGAI BAKUNA

1.6.1. a. Airearen kalitatearen definizioa aztertzen bada, aldagai kualitatiboa da. Aldiz, indizeak aztertzen badira, aldagai kuantitatibo diskretua dugu.

b. Barra-diagrama eta sektore-diagrama

c. $\bar{x} = 2,67$, $M_e = 3$. Batez bestekoaren arabera, airearen kalitatea *ona* eta *oso ona*ren artean dago. Medianak adierazten duenaren arabera, egunen erdiak *oso onak, onak* eta *egokiak* dira.

1.6.2. a. Taula:

x_i	f_i	F_i	h_i
10	2	2	0,08
12	3	5	0,12
15	3	8	0,12
20	5	13	0,20
35	6	19	0,24
38	6	25	0,24

b. Adierazpen grafikoak

c. $M_e = 20$.

Banaketa bimodala da, hots, bi moda ditu: 35 eta 38 balioak.

1.6.3. a. Taula:

$[l_i, l_{i+1})$	f_i	F_i	h_i	H_i
[20, 24)	5	5	0,125	0,125
[24, 28)	5	10	0,125	0,25
[28, 32)	4	14	0,10	0,35
[32, 36)	8	22	0,20	0,55
[36, 40)	6	28	0,15	0,70
[40, 44]	12	40	0,30	1

b. Adierazpen grafikoak.

c. Heina, $R = 44 - 20 = 24$.

Kuartilarteko heina $R_1 = Q_3 - Q_1 = 40{,}67 - 28 = 12{,}67$.

1.6.4. a. $\bar{x} = 37{,}8$, $s = 0{,}42$, $M_e = 38$, $M_o = 37{,}5$.

b. $P_{65} = 38{,}1$, $P_{90} = 38{,}5$. Esanahia: mekanikarien % 65en gaitasuna 38,1 puntuazioaren azpitik dago eta mekanikarien % 90en gaitasuna 38,5 puntuazioaren azpitik dago.

1.6.5. a. Adierazpen grafikoak.

b. $R = 50$, $Q_3 - Q_1 = 21{,}645$, $s = 13{,}12$, $CV = 0{,}34$.

c. 42,5 balioa P_{54} eta P_{55} pertzentilen artean dago.

1.6.6. a. $H_4 = 0{,}925$. Hau da, hirien % 92,5en tenperaturak 10, 14, 17 edo 20 ºC dira.

b. $\bar{x} = 16{,}5$ ºC, $M_e = M_o = 17$ ºC, $s = 3{,}81$ ºC.

c. $P_{65} = 17$ ºC. Kasu honetan, horrek esan nahi du hirien % 65en tenperaturak gehienez 17 ºC-koak direla.

d. Lehenengo berrogei hiriek.

1.6.7. a. Adierazpen grafikoak.

b. 84.556 euro.

c. Bi enpresa.

1.6.8. Tipifikazioa erabiliz, lehenengo herrialdeko etxearen alokairua bigarrenekoa baino merkeagoa dela froga daiteke.

1.6.9. a. % 72,5.

b. 29 tona.

1.6.10. a.

x_i	f_i	F_i	h_i	H_i
12	3	3	0,15	0,15
17	3	6	0,15	0,30
23	2	8	0,10	0,40
26	2	10	0,10	0,50
28	3	13	0,15	0,65
29	3	16	0,15	0,80
34	4	20	0,20	1

b. $\bar{x} = 24{,}6$, $M_e = 27$, $M_o = 34$, $s = 7{,}43$, $CV = 0{,}30$, $Q_1 = 17$, $Q_3 = 29$, $R = 22$, $R_1 = 12$, $g_2 = 38{,}8$.

2 GAIA. PROBABILITATEAREN OINARRIAK

2.6.1. a. *A, B* eta *C* gertaerak ez dira bateraezinak, $P(A \cap B)$, $P(A \cap C)$, $P(B \cap C)$ eta $P(A \cap B \cap C)$ probabilitateak ez-nuluak baitira.

b. $P(A|B) = \dfrac{0{,}11}{0{,}41} = 0{,}268$, $P(B|A) = \dfrac{0{,}11}{0{,}40} = 0{,}275$, $P(A|C) = \dfrac{0{,}11}{0{,}53} = 0{,}208$,

$P(\overline{B}|A) = \dfrac{0{,}29}{0{,}53} = 0{,}725$, $P(A \cap B|C) = 0{,}019$, $P(A \cup B|C) = 0{,}623$.

2.6.2. a. $P(A \cup B) \geq P(A)$ izan behar delako.

b. $P(A \cap B) \leq P(A)$ izan behar delako.

2.6.3. 2/44, 12/44, 12/44, 33/44, 26/44.

2.6.4. a. $\dfrac{560}{1.140}$, b. $\dfrac{1.040}{1.140}$.

2.6.5. S = “sistemak funtzionatu”, E_i = “osagaiak funtzionatu”.

$P(S) = P(E_1 \cap E_2 \cap \ldots \cap E_{65}) = P(E_1)P(E_2)\ldots P(E_{65}) = (0{,}99)^{65} = 0{,}52$; $P(\overline{S}) = 0{,}48$

2.6.6. F_i = “*i.* musika-katea, goi-fidelitatekoa”, $i = 1, 2, 3, 4$.

a. $P(F_1 \cap F_2 \cap F_3 \cap F_4) = P(F_1)P(F_2 | F_1)P(F_3 | F_1F_2)P(F_4 | F_1F_2F_3) = 0{,}424$.

b. $P(F_1 \cap F_2 \cap F_3 \cap F_4) = P(F_1)P(F_2)P(F_3)P(F_4) = 0{,}430$.

2.6.7. E = “emakume”, U = “unibertsitatera joan” eta $\overline{U}$ = “unibertsitatera ez joan”. $P(U|E) = 0{,}752$, $P(E|U) = 0{,}448$. Ondorioz, herri horretan jasotako datuen arabera, emakumeen artean unibertsitatera joan diren proportzioa % 75,2koa izan arren, unibertsitatera joan diren pertsonen artean gizonezkoak nagusitzen dira, emakumeen proportzioa % 44,8koa izanik.

2.6.8. M = “McAfee”, D = “Panda” eta B = “beste”. $\{M,D,B\}$ gertaera-sistema osoa da. Bayes-en teorema aplikatuz, $P(M|E) = 0{,}581$ da.

2.6.9. G = “kontrol-kalitatea gainditu”, A = “A teknika erabili”, B = “B teknika erabili”, C = “C teknika erabili” gertaerak. $P(A) = 0{,}40$, $P(B) = 0{,}30$, $P(C) = 0{,}30$, $P(G) = 0{,}74$, $P(G|A) = 0{,}80$, $P(G|B) = 0{,}60$. $\{A, B, C\}$ gertaera-sistema osoa da.

a. Probabilitate osoaren teorema aplikatuz: $P(G|C) = 0{,}8$.

b. Bayes-en teorematik, $P(B|G) = 0{,}243$ eta $P(\overline{B}|G) = 0{,}757$ ondorioztа daitezke.

2.6.10. M = “kexa jaso, arazo mekanikoengatik”, E = “kexa jaso, arazo elektrikoengatik”, B = “bermeko lehen urtean, kexa jaso” gertaerak. $\{E, M\}$ gertaera-sistema osoa. $P(M) = 0{,}35$, $P(E) = 0{,}65$, $P(B|M) = 0{,}16$ eta $P(B|E) = 0{,}24$. Bayes-en teorema aplikatuz $P(M|B) = 0{,}264$ lortuko da.

2.6.11. E = “eguzkitsua”, I = “eguzkitsu iragarri” eta Z = “iragarpen zuzena” $\{E,\overline{E}\}$ gertaera-sistema osoa. $P(\overline{E}) = 0{,}80$, $P(I|E) = 0{,}99$ eta $P(I|\overline{E}) = 0{,}03$.

a. $P(Z) = 0{,}974$.

b. $P(E|I) = 0{,}892$.

3. GAIA. ALDAGAI ALEATORIO DISKRETUA

3.6.1. a. bada probabilitate-funtzioa, b. eta c. ez dira probabilitate-funtzioak.

3.6.2. a. Probabilitate-funtzioa da, b. $P(X > 1) = 5/12$, $P(0 < X \leq 2) = 0{,}75$, c. $E(X) = 7/6$, $Var(X) = 23/36$.

3.6.3. Probabilitate-funtzioa: $f(1) = 1/3$, $f(2) = 1/6$, $f(3) = 1/6$, $f(4) = 1/3$, $f(x) = 0$ $x \neq 1, 2, 3, 4$.

$E(X) = 5/2$, $Var(X) = 19/12$.

3.6.4. a. egia eta b. gezurra.

3.6.5. a. 0,2646, b. 0,6517, c. $E(X) = 5{,}6$.

3.6.6. a. 0,223, b. 0,809, c. 0,1321.

3.6.7. a. 0,971, b. $E(X) = 0{,}4$, $\sigma_x = 0{,}545$.

3.6.8. a. 0,067, b. 0,076, c. $E(X) = 18$.

3.6.9. a. 0,647, b. 0,341.

3.6.10. a. 0,2293, b. 0,9327, c. 0,5053.

3.6.11. 0,8974.

4. GAIA. ALDAGAI ALEATORIO JARRAITUA

4.6.1. a. $k = 1/60$.

b. $F(x) = \begin{cases} 0, & x < 2 \\ \dfrac{x^4 - 16}{240}, & 2 \le x \le 4 \\ 1, & x > 4 \end{cases}$

c. $P(0 < X < 2{,}5) = 0{,}096$, $P(X > 3) = 0{,}729$, eta $P(X \le 3{,}5) = 0{,}559$.

4.6.2. a. $f(x) \equiv F'(x) = \begin{cases} 3, & 0 \le x \le 1/3 \\ 0, & \text{beste} \end{cases}$

b. $P(0{,}2 \le X \le 0{,}7) = 0{,}4$, $P(X < 0{,}32) = 0{,}96$, $P(X \ge 0{,}27) = 0{,}19$

4.6.3. a. $P(1 \le X \le 3) = 0{,}383$, b. $P(X \ge 0{,}5) = 0{,}779$

4.6.4. 0,25.

4.6.5. $\sigma = 23{,}7$.

4.6.6. a. 0,8078, b. 0,673.

4.6.7. 0,233.

4.6.8. a. 0,6126 , b. 33.

4.6.9. 0,9596.

4.6.10. a. $k^2 = 16$, $P(X \le 12) = 0{,}8$, b. 8.

4.6.11. a. 0,9838, b. 0,7334.

5. GAIA. LAGINKETARAKO SARRERA

5.10.1. a. 0,1922, b. 0,9788.

5.10.2. a. 0,924, b. 0,052.

5.10.3. a. 0,9958, b. 0,0021.

5.10.4. a. 0,008, b. 0,002.

5.10.5. a. 0,95, b. 0,90.

5.10.6. a. 0,9938, b. 0,0062.

5.10.7. a. 0,9955, b. 0,9997.

5.10.8. a. 0,954, b. 0,906.

5.10.9. a. 0,10, b. 3,11.

5.10.10. a. 0,042, b. 0,036.

5.10.11. $n \approx 60$.

6. GAIA. ESTIMAZIOA

6.6.1. a. [41,2, 48,8], b. [12,84, 18,41].

6.6.2. a. [0,081, 0,279], b. 0,9822.

6.6.3. a. 857, b. [–0,11, 0,01].

6.6.4. a. [1,97, 2,83], b. [1,16, 3,60].

6.6.5. a. [–3,04, 9,04], b. [–3,49, 9,49].

6.6.6. a. [5,82, 10,18], b. [7, 9].

6.6.7. a. [0,13, 0,59].

b. Bai. Konfiantza-tarteko muturretako adierazpenean $n > 30$ denean $z_{\alpha/2}$ balioa kalkulatu behar da.

6.6.8. a. [41,94, 965,29].

b. [41,47, 640,19].

c. Konfiantza-maila bera izanik, bigarren tartearen luzera txikiagoa denez, bigarren estimazioa zehatzagoa da.

6.6.9. a. [0,022, 0,727], b. [0,019, 1,133].

6.6.10. a. Bai, b. 296.

6.6.11. a. Ez, estimazio-tartean balio positiboak baitaude.

b. Ez, aurreko ataleko arrazoi beragatik.

7. GAIA. HIPOTESI-KONTRASTEA

7.5.1. a. Onartu, % 1eko adierazgarritasun-mailaz.

b. Ez onartu, % 1eko adierazgarritasun-mailaz.

7.5.2. a. Onartu, % 5eko adierazgarritasun-mailaz.

b. 0,95.

7.5.3. a. Errefusatu, % 5eko adierazgarritasun-mailaz.

b. Errefusatu, % 10eko adierazgarritasun-mailaz.

7.5.4. Ez genuke onartuko susmoa, % 1eko adierazgarritasun-mailaz.

7.5.5. a. 0,05.

b. Adierazgarritasun-maila eta kontrastearen ahalmena.

7.5.6. a. Bai, % 2,5eko adierazgarritasun-mailaz, onartu.

b. p-balioa = 0,0018. Orduan, hipotesi-kontrastearen adierazgarritasun-maila 0,0018 baino handiagoa denean, dadoa trukatuta dagoela onar daiteke.

7.5.7. Ez onartu, % 10eko adierazgarritasun-mailaz.

7.5.8. a. Onartu, % 1eko adierazgarritasun-mailaz

b. Adierazgarritasun-maila kritikoa edo p-balioa oso txikia da, ia nulua. Beraz, p-balioa baino handiagoak diren adierazgarritasun-mailek kontrako erabakia hartzera eramango gintuzketen, kasu honetan beti.

7.5.9. a. Onartu, % 5eko adierazgarritasun-mailaz.

b. 0,05.

7.5.10. Ez onartu, % 2,5eko adierazgarritasun-mailaz.

7.5.11. a. Ez, adierazgarritasun-maila p-balioa baino handiagoa delako.

b. 0,6767.

8. GAIA. χ^2 BANAKETAREN APLIKAZIOAK

8.7.1. Errefusatu, % 1eko adierazgarritasun-mailaz.

8.7.2. Errefusatu, % 5eko adierazgarritasun-mailaz.

8.7.3. Badago dependentzia, % 1eko adierazgarritasun-mailaz.

8.7.4. Dependentzia-erlazioa dago, % 5eko adierazgarritasun-mailaz.

8.7.5. Datuek Poisson-en banaketa dute, % 5eko adierazgarritasun-mailaz.

8.7.6. Desintegratzeko probabilitate bera dute, % 2,5eko adierazgarritasun-mailaz.

8.7.7. Diferentzia adierazgarririk ez, % 5eko adierazgarritasun-mailaz.

8.7.8. Ez onartu emaitzak, % 2,5eko adierazgarritasun-mailaz.

8.7.9. Independenteak dira, % 5eko adierazgarritasun-mailaz.

8.7.10. Homogeneotasunaren hipotesia onartu, % 1eko adierazgarritasun-mailaz.

9. GAIA. ERREGRESIO LINEAL BAKUNA

9.7.1. a. $r = 0{,}90$.

b. 407 kW.

9.7.2. a. Bai, $r = 0{,}84$.

b. $x = 0{,}41y + 6{,}73$.

9.7.3. a. $y = 0{,}006x + 0{,}60$.

b. $I_{\alpha}^{0,95} = [0{,}0055,\ 0{,}0065]$, $I_{\beta}^{0,95} = [0{,}155,\ 1{,}045]$.

9.7.4. a. $r = 0{,}89$.

b. [0,03, 2,81].

9.7.5. a. $r = -0{,}95$, $y = -0{,}09x + 21{,}31$.

b. $S_{ey} = 2{,}59$.

9.7.6. a. $y = 12{,}90x - 27{,}91$.

b. [1.774,73, 3.329,45].

9.7.7. a. $S_{ey} = 4{,}47$.

b. $I_{\alpha}^{0,95} = [1{,}0198,\ 1{,}0202]$, $I_{\beta}^{0,95} = [3{,}90,\ 14{,}06]$.

9.7.8. a. $y = -1{,}5x + 39{,}4$.

b. [–2,26, –0,5,].

Taulak

1. taula. Banaketa binomialaren banaketa-funtzioa

$$F(x) = P(X \le x) = \sum_{k \le x} \binom{n}{k} p^k (1-p)^{n-k}$$

		p									
n	*x*	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	0	0,95	0,9	0,85	0,8	0,75	0,7	0,65	0,6	0,55	0,5
2	0	0,9025	0,81	0,7225	0,64	0,5625	0,49	0,4225	0,36	0,3025	0,25
	1	0,9975	0,99	0,9775	0,96	0,9375	0,91	0,8775	0,84	0,7975	0,75
3	0	0,8574	0,729	0,6141	0,512	0,4219	0,343	0,2746	0,216	0,1664	0,125
	1	0,9928	0,972	0,9392	0,896	0,8438	0,784	0,7182	0,648	0,5748	0,5
	2	0,9999	0,999	0,9966	0,992	0,9844	0,973	0,9571	0,936	0,9089	0,875
4	0	0,8145	0,6561	0,522	0,4096	0,3164	0,2401	0,1785	0,1296	0,0915	0,0625
	1	0,986	0,9477	0,8905	0,8192	0,7383	0,6517	0,563	0,4752	0,391	0,3125
	2	0,9995	0,9963	0,988	0,9728	0,9492	0,9163	0,8735	0,8208	0,7585	0,6875
	3	1	0,9999	0,9995	0,9984	0,9961	0,9919	0,985	0,9744	0,959	0,9375
5	0	0,7738	0,5905	0,4437	0,3277	0,2373	0,1681	0,116	0,0778	0,0503	0,0313
	1	0,9774	0,9186	0,8352	0,7373	0,6328	0,5283	0,4284	0,337	0,2562	0,1876
	2	0,9988	0,9915	0,9734	0,9421	0,8965	0,837	0,7648	0,6826	0,5931	0,5001
	3	0,9999	0,9996	0,9978	0,9933	0,9844	0,9693	0,9459	0,913	0,8688	0,8126
	4	0,9999	1	1	0,9997	0,999	0,9977	0,9947	0,9898	0,9816	0,9689
6	0	0,7351	0,5314	0,3771	0,2621	0,178	0,1176	0,0754	0,0467	0,0277	0,0156
	1	0,9672	0,8857	0,7764	0,6553	0,534	0,4201	0,3191	0,2333	0,1636	0,1094
	2	0,9977	0,9841	0,9526	0,9011	0,8306	0,7442	0,6471	0,5443	0,4416	0,3438
	3	0,9998	0,9987	0,9941	0,983	0,9624	0,9294	0,8826	0,8208	0,7448	0,6563
	4	0,9999	0,9999	0,9996	0,9984	0,9954	0,9889	0,9777	0,959	0,9309	0,8907
	5	0,9999	1	1	0,9999	0,9998	0,9991	0,9982	0,9959	0,9918	0,9845
7	0	0,6983	0,4783	0,3206	0,2097	0,1335	0,0824	0,049	0,028	0,0152	0,0078
	1	0,9556	0,8503	0,7166	0,5767	0,445	0,3295	0,2338	0,1586	0,1024	0,0625
	2	0,9962	0,9743	0,9263	0,852	0,7565	0,6472	0,5323	0,4199	0,3164	0,2266
	3	0,9998	0,9973	0,988	0,9667	0,9295	0,8741	0,8002	0,7102	0,6082	0,5
	4	1	0,9999	0,9989	0,9954	0,9872	0,9713	0,9444	0,9037	0,847	0,7734
	5	1	1	1	0,9997	0,9987	0,9963	0,991	0,9811	0,9642	0,9375
	6	1	1	1	1	1	0,9999	0,9994	0,9983	0,9962	0,9922

8	0	0,6634	0,4305	0,2725	0,1678	0,1001	0,0576	0,0319	0,0168	0,0084	0,0039
	1	0,9427	0,8131	0,6572	0,5033	0,3671	0,2553	0,1692	0,1064	0,0632	0,0352
	2	0,9942	0,9619	0,8948	0,7969	0,6786	0,5518	0,4279	0,3154	0,2201	0,1446
	3	0,9996	0,995	0,9787	0,9437	0,8862	0,8059	0,7065	0,5941	0,4769	0,3634
	4	1	0,9996	0,9972	0,9896	0,9727	0,942	0,894	0,8263	0,7396	0,6368
	5	1	1	0,9998	0,9988	0,9958	0,9887	0,9748	0,9502	0,9115	0,8556
	6	1	1	1	0,9999	0,9996	0,9987	0,9965	0,9915	0,9818	0,965
	7	1	1	1	1	1	0,9999	0,9998	0,9994	0,9982	0,9963
9	0	0,6302	0,3874	0,2316	0,1342	0,0751	0,0404	0,0207	0,0101	0,0046	0,002
	1	0,9287	0,7748	0,5995	0,4362	0,3004	0,196	0,1211	0,0706	0,0385	0,0196
	2	0,9916	0,947	0,8592	0,7382	0,6007	0,4628	0,3373	0,2318	0,1495	0,0899
	3	0,9993	0,9916	0,9661	0,9144	0,8343	0,7296	0,6089	0,4826	0,3614	0,254
	4	0,9999	0,999	0,9944	0,9805	0,9511	0,9011	0,8283	0,7334	0,6214	0,5001
	5	0,9999	0,9998	0,9994	0,997	0,99	0,9746	0,9464	0,9006	0,8342	0,7462
	6	0,9999	0,9999	1	0,9998	0,9987	0,9956	0,9888	0,9749	0,9502	0,9103
	7	0,9999	0,9999	1	1	0,9999	0,9995	0,9986	0,9961	0,9909	0,9806
	8	0,9999	0,9999	1	1	1	0,9999	0,9999	0,9996	0,9992	0,9982
10	0	0,5987	0,3487	0,1969	0,1074	0,0563	0,0282	0,0135	0,006	0,0025	0,001
	1	0,9138	0,7361	0,5443	0,3758	0,244	0,1493	0,086	0,0463	0,0232	0,0108
	2	0,9884	0,9298	0,8202	0,6778	0,5256	0,3828	0,2617	0,1672	0,0995	0,0547
	3	0,9989	0,9872	0,95	0,8791	0,7759	0,6496	0,5139	0,3822	0,266	0,1719
	4	0,9999	0,9984	0,9901	0,9672	0,9219	0,8497	0,7516	0,633	0,5044	0,377
	5	1	0,9999	0,9986	0,9936	0,9803	0,9526	0,9052	0,8337	0,7384	0,6231
	6	1	1	0,9998	0,9991	0,9965	0,9894	0,9741	0,9452	0,898	0,8282
	7	1	1	0,9999	0,9999	0,9996	0,9984	0,9953	0,9877	0,9726	0,9454
	8	1	1	0,9999	1	1	0,9998	0,9996	0,9983	0,9955	0,9893
	9	1	1	0,9999	1	1	0,9999	1	0,9999	0,9997	0,9991
11	0	0,5688	0,3138	0,1673	0,0859	0,0422	0,0198	0,0088	0,0036	0,0014	0,0005
	1	0,8981	0,6973	0,4921	0,3221	0,1971	0,113	0,0606	0,0302	0,0139	0,0059
	2	0,9848	0,9104	0,7787	0,6174	0,4552	0,3128	0,2001	0,1189	0,0652	0,0328
	3	0,9985	0,9814	0,9304	0,8389	0,7133	0,5696	0,4255	0,2963	0,1911	0,1134
	4	0,9999	0,9972	0,984	0,9496	0,8854	0,7897	0,6683	0,5328	0,3971	0,2745
	5	1	0,9997	0,9972	0,9884	0,9657	0,9218	0,8513	0,7535	0,6331	0,5001
	6	1	1	0,9995	0,9981	0,9925	0,9784	0,9498	0,9006	0,8262	0,7257
	7	1	1	0,9998	0,9998	0,9989	0,9957	0,9877	0,9707	0,939	0,8868
	8	1	1	0,9998	1	1	0,9994	0,9979	0,9941	0,9852	0,9674
	9	1	1	0,9998	1	1	0,9999	0,9997	0,9993	0,9978	0,9943
	10	1	1	0,9998	1	1	0,9999	0,9999	1	0,9999	0,9997
12	0	0,5404	0,2824	0,1422	0,0687	0,0317	0,0138	0,0057	0,0022	0,0008	0,0002
	1	0,8817	0,659	0,4434	0,2749	0,1584	0,085	0,0425	0,0196	0,0083	0,0031
	2	0,9805	0,8891	0,7358	0,5584	0,3907	0,2528	0,1513	0,0835	0,0422	0,0192
	3	0,9978	0,9743	0,9078	0,7946	0,6488	0,4925	0,3467	0,2254	0,1345	0,0729
	4	0,9999	0,9956	0,9761	0,9275	0,8424	0,7236	0,5834	0,4382	0,3045	0,1937
	5	1	0,9994	0,9954	0,9807	0,9456	0,8821	0,7873	0,6652	0,527	0,3871
	6	1	0,9999	0,9994	0,9962	0,9857	0,9613	0,9154	0,8418	0,7394	0,6127
	7	1	0,9999	1	0,9995	0,9972	0,9904	0,9745	0,9427	0,8883	0,8061
	8	1	0,9999	1	1	0,9996	0,9982	0,9944	0,9847	0,9645	0,9269
	9	1	0,9999	1	1	1	0,9997	0,9992	0,9972	0,9922	0,9806
	10	1	0,9999	1	1	1	0,9999	1	0,9997	0,999	0,9967
	11	1	0,9999	1	1	1	0,9999	1	1	1	0,9996

13	0	0,5133	0,2542	0,1209	0,055	0,0238	0,0097	0,0037	0,0013	0,0004	0,0001
	1	0,8645	0,6214	0,3983	0,2337	0,1267	0,0637	0,0296	0,0126	0,0049	0,0017
	2	0,9754	0,8662	0,692	0,5017	0,3326	0,2025	0,1132	0,0579	0,0269	0,0112
	3	0,9968	0,9659	0,882	0,7474	0,5843	0,4206	0,2783	0,1686	0,0929	0,0461
	4	0,9996	0,9936	0,9658	0,9009	0,794	0,6543	0,5005	0,3531	0,2279	0,1334
	5	0,9999	0,9991	0,9924	0,97	0,9198	0,8346	0,7159	0,5745	0,4268	0,2905
	6	0,9999	0,9999	0,9987	0,993	0,9757	0,9376	0,8705	0,7713	0,6437	0,5
	7	0,9999	1	0,9998	0,9988	0,9943	0,9818	0,9538	0,9025	0,8212	0,7095
	8	0,9999	1	0,9999	0,9999	0,999	0,996	0,9874	0,9681	0,9301	0,8666
	9	0,9999	1	0,9999	1	0,9999	0,9994	0,9975	0,9924	0,9796	0,9539
	10	0,9999	1	0,9999	1	1	1	0,9997	0,9989	0,9958	0,9888
	11	0,9999	1	0,9999	1	1	1	1	1	0,9994	0,9983
	12	0,9999	1	0,9999	1	1	1	1	1	0,9999	0,9999
14	0	0,4877	0,2288	0,1028	0,044	0,0178	0,0068	0,0024	0,0008	0,0002	0,0001
	1	0,847	0,5847	0,3567	0,1979	0,101	0,0475	0,0205	0,0081	0,0029	0,001
	2	0,9699	0,8417	0,6479	0,448	0,2812	0,1609	0,0839	0,0398	0,017	0,0066
	3	0,9958	0,9559	0,8535	0,6981	0,5214	0,3552	0,2205	0,1243	0,0632	0,0288
	4	0,9995	0,9908	0,9533	0,8701	0,7416	0,5842	0,4227	0,2792	0,1672	0,0899
	5	0,9999	0,9986	0,9885	0,9561	0,8884	0,7805	0,6405	0,4858	0,3373	0,2121
	6	0,9999	0,9999	0,9978	0,9883	0,9618	0,9067	0,8164	0,6924	0,5461	0,3954
	7	0,9999	1	0,9997	0,9975	0,9898	0,9685	0,9246	0,8498	0,7413	0,6049
	8	0,9999	1	1	0,9995	0,998	0,9917	0,9756	0,9416	0,8811	0,7882
	9	0,9999	1	1	0,9998	0,9998	0,9983	0,9939	0,9824	0,9573	0,9104
	10	0,9999	1	1	0,9998	1	0,9997	0,9988	0,996	0,9885	0,9715
	11	0,9999	1	1	0,9998	1	0,9999	0,9998	0,9993	0,9978	0,9937
	12	0,9999	1	1	0,9998	1	0,9999	0,9999	0,9998	0,9997	0,9993
	13	0,9999	1	1	0,9998	1	0,9999	0,9999	0,9999	0,9999	1
15	0	0,4633	0,2059	0,0874	0,0352	0,0134	0,0047	0,0016	0,0005	0,0001	0
	1	0,8291	0,5491	0,3186	0,1671	0,0802	0,0352	0,0142	0,0052	0,0017	0,0005
	2	0,9639	0,816	0,6042	0,398	0,2361	0,1268	0,0618	0,0271	0,0107	0,0037
	3	0,9946	0,9445	0,8226	0,6481	0,4613	0,2968	0,1728	0,0905	0,0425	0,0176
	4	0,9995	0,9873	0,9382	0,8357	0,6865	0,5154	0,352	0,2173	0,1205	0,0593
	5	1	0,9978	0,9831	0,9389	0,8516	0,7215	0,5643	0,4032	0,2609	0,1509
	6	1	0,9997	0,9963	0,9819	0,9433	0,8687	0,7549	0,6098	0,4523	0,3036
	7	1	1	0,9993	0,9957	0,9826	0,9498	0,8868	0,7869	0,6536	0,5
	8	1	1	0,9998	0,9992	0,9957	0,9846	0,9578	0,905	0,8183	0,6964
	9	1	1	0,9999	0,9999	0,9991	0,9962	0,9876	0,9662	0,9231	0,8491
	10	1	1	0,9999	1	0,9998	0,9992	0,9972	0,9907	0,9746	0,9407
	11	1	1	0,9999	1	0,9999	0,9998	0,9996	0,9981	0,9937	0,9824
	12	1	1	0,9999	1	0,9999	0,9999	1	0,9997	0,9989	0,9963
	13	1	1	0,9999	1	0,9999	0,9999	1	1	0,9999	0,9995
	14	1	1	0,9999	1	0,9999	0,9999	1	1	1	1
16	0	0,4401	0,1853	0,0743	0,0281	0,01	0,0033	0,001	0,0003	0,0001	0
	1	0,8107	0,5147	0,284	0,1407	0,0635	0,0261	0,0097	0,0033	0,001	0,0002
	2	0,957	0,7892	0,5615	0,3518	0,1971	0,0993	0,045	0,0183	0,0066	0,002
	3	0,9929	0,9315	0,79	0,5981	0,405	0,2458	0,1338	0,0651	0,0281	0,0105
	4	0,999	0,9829	0,9211	0,7982	0,6302	0,4498	0,2891	0,1665	0,0853	0,0383
	5	0,9998	0,9966	0,9766	0,9183	0,8104	0,6597	0,4899	0,3288	0,1976	0,105
	6	0,9999	0,9994	0,9946	0,9733	0,9205	0,8246	0,6881	0,5271	0,366	0,2272
	7	0,9999	0,9998	0,9991	0,993	0,9729	0,9256	0,8405	0,716	0,5629	0,4018
	8	0,9999	0,9999	1	0,9985	0,9926	0,9743	0,9328	0,8577	0,7441	0,5982
	9	0,9999	0,9999	1	0,9997	0,9984	0,9928	0,977	0,9417	0,8759	0,7728
	10	0,9999	0,9999	1	0,9999	0,9998	0,9984	0,9937	0,9809	0,9514	0,895
	11	0,9999	0,9999	1	0,9999	1	0,9997	0,9986	0,9951	0,9851	0,9617
	12	0,9999	0,9999	1	0,9999	1	0,9999	0,9997	0,9991	0,9966	0,9895
	13	0,9999	0,9999	1	0,9999	1	0,9999	0,9999	0,9999	0,9995	0,998
	14	0,9999	0,9999	1	0,9999	1	0,9999	0,9999	1	1	0,9998
	15	0,9999	0,9999	1	0,9999	1	0,9999	0,9999	1	1	1

17	0	0,4181	0,1668	0,0631	0,0225	0,0075	0,0023	0,0007	0,0002	0	0
	1	0,7922	0,4818	0,2524	0,1182	0,0501	0,0192	0,0067	0,0021	0,0005	0,0001
	2	0,9497	0,7618	0,5197	0,3096	0,1637	0,0773	0,0327	0,0123	0,004	0,0011
	3	0,9912	0,9174	0,7556	0,5489	0,353	0,2018	0,1028	0,0464	0,0184	0,0063
	4	0,9988	0,9779	0,9013	0,7582	0,5739	0,3886	0,2348	0,126	0,0595	0,0245
	5	0,9998	0,9954	0,9681	0,8943	0,7653	0,5967	0,4197	0,2639	0,147	0,0717
	6	0,9999	0,9993	0,9917	0,9623	0,8929	0,7751	0,6188	0,4478	0,2902	0,1661
	7	0,9999	1	0,9982	0,989	0,9597	0,8952	0,7873	0,6405	0,4743	0,3145
	8	0,9999	1	0,9996	0,9974	0,9876	0,9596	0,9007	0,8011	0,6626	0,5
	9	0,9999	1	0,9999	0,9995	0,9969	0,9872	0,9618	0,9081	0,8166	0,6855
	10	0,9999	1	0,9999	0,9999	0,9994	0,9967	0,9881	0,9652	0,9174	0,8339
	11	0,9999	1	0,9999	1	0,9999	0,9993	0,9971	0,9894	0,9699	0,9283
	12	0,9999	1	0,9999	1	1	0,9999	0,9995	0,9975	0,9914	0,9755
	13	0,9999	1	0,9999	1	1	1	1	0,9996	0,9982	0,9937
	14	0,9999	1	0,9999	1	1	1	1	1	0,9998	0,9989
	15	0,9999	1	0,9999	1	1	1	1	1	1	0,9999
	16	0,9999	1	0,9999	1	1	1	1	1	1	1
18	0	0,3972	0,1501	0,0536	0,018	0,0056	0,0016	0,0004	0,0001	0	0
	1	0,7735	0,4503	0,224	0,0991	0,0394	0,0142	0,0046	0,0013	0,0003	0,0001
	2	0,9418	0,7338	0,4796	0,2714	0,1352	0,06	0,0236	0,0082	0,0025	0,0007
	3	0,9891	0,9018	0,7202	0,5011	0,3056	0,1646	0,0783	0,0328	0,012	0,0038
	4	0,9984	0,9718	0,8794	0,7164	0,5186	0,3327	0,1887	0,0942	0,0411	0,0155
	5	0,9998	0,9936	0,9581	0,8671	0,7174	0,5344	0,3551	0,2088	0,1077	0,0482
	6	1	0,9988	0,9882	0,9487	0,861	0,7217	0,5492	0,3743	0,2258	0,119
	7	1	0,9998	0,9973	0,9837	0,943	0,8593	0,7284	0,5635	0,3915	0,2404
	8	1	1	0,9995	0,9957	0,9806	0,9404	0,8611	0,7369	0,5779	0,4073
	9	1	1	0,9999	0,999	0,9945	0,979	0,9405	0,8653	0,7473	0,5928
	10	1	1	1	0,9998	0,9987	0,9939	0,979	0,9424	0,8721	0,7597
	11	1	1	1	0,9999	0,9997	0,9985	0,9941	0,9798	0,9463	0,8811
	12	1	1	1	0,9999	0,9999	0,9997	0,9988	0,9943	0,9817	0,9519
	13	1	1	1	0,9999	0,9999	0,9999	1	0,9988	0,9951	0,9846
	14	1	1	1	0,9999	0,9999	0,9999	1	0,9999	0,999	0,9963
	15	1	1	1	0,9999	0,9999	0,9999	1	1	0,9999	0,9994
	16	1	1	1	0,9999	0,9999	0,9999	1	1	1	1
	17	1	1	1	0,9999	0,9999	0,9999	1	1	1	1
19	0	0,3774	0,1351	0,0456	0,0144	0,0042	0,0011	0,0003	0,0001	0	0
	1	0,7548	0,4203	0,1985	0,0829	0,031	0,0104	0,0032	0,0009	0,0002	0
	2	0,9335	0,7055	0,4413	0,2369	0,1113	0,0462	0,017	0,0055	0,0015	0,0003
	3	0,9868	0,8851	0,6841	0,4551	0,263	0,1331	0,0592	0,023	0,0077	0,0021
	4	0,998	0,9649	0,8555	0,6733	0,4653	0,2822	0,1501	0,0697	0,028	0,0095
	5	0,9998	0,9915	0,9462	0,8369	0,6676	0,4738	0,2969	0,163	0,0777	0,0317
	6	1	0,9984	0,9836	0,9324	0,825	0,6654	0,4813	0,3081	0,1726	0,0835
	7	1	0,9998	0,9958	0,9767	0,9224	0,8179	0,6657	0,4878	0,3169	0,1796
	8	1	1	0,999	0,9933	0,9711	0,916	0,8146	0,6675	0,494	0,3238
	9	1	1	0,9997	0,9984	0,9909	0,9674	0,9126	0,8139	0,6711	0,5
	10	1	1	0,9998	0,9997	0,9975	0,9894	0,9654	0,9115	0,816	0,6762
	11	1	1	0,9998	1	0,9993	0,9971	0,9887	0,9647	0,913	0,8204
	12	1	1	0,9998	1	0,9997	0,9993	0,997	0,9884	0,9659	0,9165
	13	1	1	0,9998	1	0,9998	0,9998	0,9994	0,9969	0,9892	0,9683
	14	1	1	0,9998	1	0,9998	0,9999	1	0,9993	0,9974	0,9905
	15	1	1	0,9998	1	0,9998	0,9999	1	0,9998	0,9996	0,9979
	16	1	1	0,9998	1	0,9998	0,9999	1	0,9999	1	0,9997
	17	1	1	0,9998	1	0,9998	0,9999	1	0,9999	1	1
	18	1	1	0,9998	1	0,9998	0,9999	1	0,9999	1	1

20	0	0,3585	0,1216	0,0388	0,0115	0,0032	0,0008	0,0002	0	0	0
	1	0,7359	0,3918	0,1756	0,0691	0,0243	0,0076	0,0022	0,0005	0,0001	0
	2	0,9246	0,677	0,4049	0,206	0,0912	0,0354	0,0122	0,0036	0,0009	0,0002
	3	0,9842	0,8671	0,6477	0,4114	0,2251	0,107	0,0445	0,0159	0,0049	0,0013
	4	0,9975	0,9569	0,8298	0,6296	0,4148	0,2374	0,1183	0,0509	0,0188	0,0059
	5	0,9997	0,9888	0,9326	0,8042	0,6171	0,4163	0,2455	0,1255	0,0553	0,0207
	6	1	0,9977	0,978	0,9133	0,7857	0,6079	0,4167	0,2499	0,1299	0,0577
	7	1	0,9997	0,994	0,9678	0,8981	0,7722	0,6011	0,4158	0,252	0,1316
	8	1	1	0,9986	0,99	0,959	0,8866	0,7625	0,5955	0,4143	0,2517
	9	1	1	0,9997	0,9974	0,9861	0,952	0,8783	0,7552	0,5914	0,4119
	10	1	1	0,9999	0,9994	0,996	0,9828	0,9469	0,8723	0,7507	0,5881
	11	1	1	0,9999	0,9999	0,999	0,9948	0,9805	0,9433	0,8692	0,7483
	12	1	1	0,9999	1	0,9998	0,9987	0,9941	0,9788	0,9419	0,8684
	13	1	1	0,9999	1	1	0,9997	0,9986	0,9934	0,9785	0,9423
	14	1	1	0,9999	1	1	0,9999	0,9998	0,9983	0,9935	0,9793
	15	1	1	0,9999	1	1	0,9999	1	0,9996	0,9984	0,9941
	16	1	1	0,9999	1	1	0,9999	1	0,9999	0,9997	0,9987
	17	1	1	0,9999	1	1	0,9999	1	0,9999	0,9999	0,9998
	18	1	1	0,9999	1	1	0,9999	1	0,9999	0,9999	1
	19	1	1	0,9999	1	1	0,9999	1	0,9999	0,9999	1

2. taula. Poisson-en banaketaren banaketa-funtzioa

$$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{e^{-\lambda}\lambda^k}{k!}$$

$x \setminus \lambda$	0,005	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,995	0,99	0,9802	0,9704	0,9608	0,9512	0,9418	0,9324	0,9231	0,9139
1	1	0,9999	0,9998	0,9995	0,9992	0,9988	0,9983	0,9977	0,9969	0,9962
2	1	0,9999	1	0,9999	1	1	1	1	0,9999	0,9999
3	1	0,9999	1	0,9999	1	1	1	1	1	1

$x \setminus \lambda$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679
1	0,9953	0,9824	0,963	0,9384	0,9098	0,8781	0,8442	0,8088	0,7725	0,7358
2	0,9998	0,9988	0,9963	0,992	0,9856	0,9769	0,9659	0,9526	0,9372	0,9197
3	1	0,9999	0,9996	0,9992	0,9982	0,9967	0,9943	0,9909	0,9866	0,981
4	1	1	0,9999	0,9999	0,9998	0,9997	0,9993	0,9986	0,9977	0,9963
5	1	1	0,9999	1	1	1	1	0,9998	0,9997	0,9994

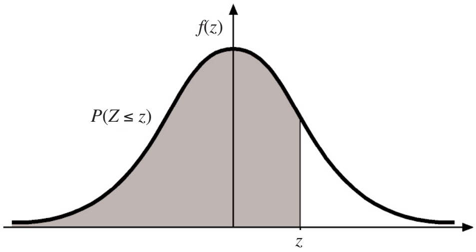
$x \setminus \lambda$	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2
0	0,3329	0,3012	0,2725	0,2466	0,2231	0,2019	0,1827	0,1653	0,1496	0,1353
1	0,6991	0,6626	0,6268	0,5918	0,5578	0,5249	0,4933	0,4628	0,4338	0,406
2	0,9005	0,8795	0,8571	0,8335	0,8088	0,7833	0,7573	0,7306	0,7038	0,6767
3	0,9743	0,9662	0,9569	0,9463	0,9343	0,9211	0,9069	0,8913	0,8748	0,8571
4	0,9946	0,9922	0,9893	0,9858	0,9814	0,9762	0,9705	0,9636	0,956	0,9473
5	0,9991	0,9984	0,9977	0,9969	0,9955	0,9938	0,9921	0,9896	0,9869	0,9834
6	0,9999	0,9996	0,9995	0,9995	0,999	0,9985	0,9982	0,9974	0,9967	0,9954
7	1	0,9998	0,9998	1	0,9998	0,9996	0,9997	0,9994	0,9994	0,9988

$x \backslash \lambda$	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3
0	0,1225	0,1108	0,1003	0,0907	0,0821	0,0743	0,0672	0,0608	0,055	0,0498
1	0,3797	0,3546	0,3309	0,3084	0,2873	0,2674	0,2487	0,2311	0,2146	0,1992
2	0,6497	0,6227	0,5961	0,5697	0,5438	0,5184	0,4937	0,4695	0,446	0,4232
3	0,8387	0,8193	0,7994	0,7787	0,7576	0,736	0,7142	0,692	0,6697	0,6472
4	0,9379	0,9275	0,9163	0,9041	0,8912	0,8774	0,863	0,8477	0,8319	0,8152
5	0,9796	0,9751	0,9701	0,9643	0,958	0,9509	0,9434	0,9349	0,9259	0,916
6	0,9942	0,9925	0,9907	0,9884	0,9858	0,9828	0,9796	0,9756	0,9714	0,9664
7	0,9986	0,998	0,9975	0,9967	0,9957	0,9946	0,9935	0,9919	0,9902	0,988
8	0,9997	0,9995	0,9994	0,9992	0,9988	0,9984	0,9982	0,9976	0,997	0,9961
9	1	0,9999	0,9999	0,9999	0,9997	0,9995	0,9996	0,9994	0,9992	0,9988
10	1	1	1	1	0,9999	0,9998	1	0,9999	0,9998	0,9996

$x \backslash \lambda$	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9	4
0	0,045	0,0408	0,0369	0,0334	0,0302	0,0273	0,0247	0,0224	0,0202	0,0183
1	0,1847	0,1712	0,1586	0,1469	0,1359	0,1257	0,1162	0,1074	0,0991	0,0916
2	0,4012	0,3799	0,3594	0,3398	0,3209	0,3028	0,2854	0,2689	0,253	0,2381
3	0,6249	0,6025	0,5803	0,5584	0,5367	0,5153	0,4941	0,4735	0,4531	0,4335
4	0,7982	0,7806	0,7626	0,7442	0,7255	0,7065	0,6872	0,6679	0,6482	0,6289
5	0,9057	0,8946	0,8829	0,8706	0,8577	0,8442	0,8301	0,8156	0,8004	0,7852
6	0,9612	0,9554	0,9491	0,9422	0,9348	0,9268	0,9182	0,9092	0,8993	0,8894
7	0,9858	0,9832	0,9803	0,977	0,9733	0,9693	0,9648	0,96	0,9544	0,9489
8	0,9953	0,9943	0,9932	0,9918	0,9902	0,9884	0,9863	0,9841	0,9813	0,9787
9	0,9986	0,9983	0,9979	0,9974	0,9968	0,996	0,9952	0,9943	0,9929	0,9919
10	0,9996	0,9996	0,9995	0,9993	0,9991	0,9988	0,9985	0,9982	0,9974	0,9972
11	0,9999	1	1	0,9999	0,9998	0,9997	0,9996	0,9995	0,999	0,9991
12	1	1	1	1	1	1	0,9999	0,9999	0,9995	0,9997
13	1	1	1	1	1	1	1	1	0,9997	0,9999

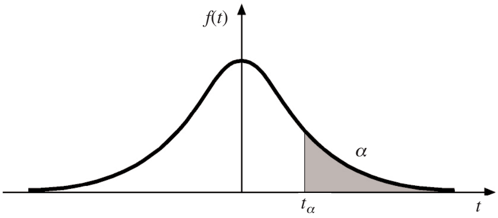
3. taula. Banaketa normal estandarra

$$P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



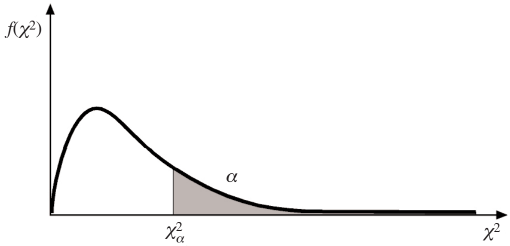
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

4. taula. Student-en t banaketako $P(t_\nu > t_\alpha) = \alpha$ balioak



ν\α	0,25	0,2	0,15	0,1	0,05	0,025	0,01	0,005	0,0005
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,656	636,578
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,600
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,768
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,689
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,660
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,679	0,848	1,045	1,296	1,671	2,000	2,390	2,660	3,460
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,290

5. taula. χ^2 banaketako $P(\chi_\nu^2 > \chi_\alpha) = \alpha$ balioak (I)



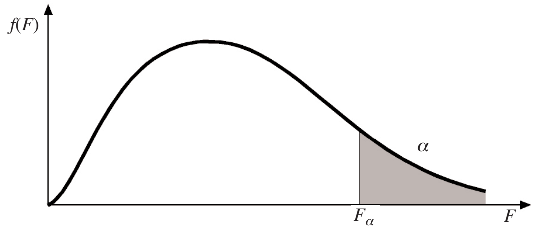
ν\α	0,995	0,99	0,975	0,95	0,9	0,75	0,5
1	0,000039	0,00016	0,00098	0,00393	0,0158	0,102	0,455
2	0,010025	0,02010	0,0506	0,103	0,211	0,575	1,386
3	0,07172	0,115	0,216	0,352	0,584	1,213	2,37
4	0,207	0,297	0,484	0,711	1,064	1,923	3,36
5	0,412	0,554	0,831	1,145	1,610	2,67	4,35
6	0,676	0,872	1,237	1,635	2,20	3,45	5,35
7	0,989	1,239	1,690	2,17	2,83	4,25	6,35
8	1,344	1,647	2,18	2,73	3,49	5,07	7,34
9	1,735	2,09	2,70	3,33	4,17	5,90	8,34
10	2,16	2,56	3,25	3,94	4,87	6,74	9,34
11	2,60	3,05	3,82	4,57	5,58	7,58	10,34
12	3,07	3,57	4,40	5,23	6,30	8,44	11,34
13	3,57	4,11	5,01	5,89	7,04	9,30	12,34
14	4,07	4,66	5,63	6,57	7,79	10,17	13,34
15	4,60	5,23	6,26	7,26	8,55	11,04	14,34
16	5,14	5,81	6,91	7,96	9,31	11,91	15,34
17	5,70	6,41	7,56	8,67	10,09	12,79	16,34
18	6,26	7,01	8,23	9,39	10,86	13,68	17,34
19	6,84	7,63	8,91	10,12	11,65	14,56	18,34
20	7,43	8,26	9,59	10,85	12,44	15,45	19,34
21	8,03	8,90	10,28	11,59	13,24	16,34	20,3
22	8,64	9,54	10,98	12,34	14,04	17,24	21,3
23	9,26	10,20	11,69	13,09	14,85	18,14	22,3
24	9,89	10,86	12,40	13,85	15,66	19,04	23,3
25	10,52	11,52	13,12	14,61	16,47	19,94	24,3
26	11,16	12,20	13,84	15,38	17,29	20,8	25,3
27	11,81	12,88	14,57	16,15	18,11	21,7	26,3
28	12,46	13,56	15,31	16,93	18,94	22,7	27,3
29	13,12	14,26	16,05	17,71	19,77	23,6	28,3
30	13,79	14,95	16,79	18,49	20,6	24,5	29,3
40	20,7	22,2	24,4	26,5	29,1	33,7	39,3
50	28,0	29,7	32,4	34,8	37,7	42,9	49,3
60	35,5	37,5	40,5	43,2	46,5	52,3	59,3
70	43,3	45,4	48,8	51,7	55,3	61,7	69,3
80	51,2	53,5	57,2	60,4	64,3	71,1	79,3
90	59,2	61,8	65,6	69,1	73,3	80,6	89,3
100	67,3	70,1	74,2	77,9	82,4	90,1	99,3

5. taula. χ^2 banaketako $P(\chi_\nu^2 > \chi_\alpha) = \alpha$ balioak (II)

ν\α	0,25	0,1	0,05	0,025	0,01	0,005
1	1,323	2,71	3,84	5,02	6,63	7,88
2	2,77	4,61	5,99	7,38	9,21	10,60
3	4,11	6,25	7,81	9,35	11,34	12,84
4	5,39	7,78	9,49	11,14	13,28	14,86
5	6,63	9,24	11,07	12,83	15,09	16,75
6	7,84	10,64	12,59	14,45	16,81	18,55
7	9,04	12,02	14,07	16,01	18,48	20,3
8	10,22	13,36	15,51	17,53	20,1	22,0
9	11,39	14,68	16,92	19,02	21,7	23,6
10	12,55	15,99	18,31	20,5	23,2	25,2
11	13,70	17,28	19,68	21,9	24,7	26,8
12	14,85	18,55	21,0	23,3	26,2	28,3
13	15,98	19,81	22,4	24,7	27,7	29,8
14	17,12	21,1	23,7	26,1	29,1	31,3
15	18,25	22,3	25,0	27,5	30,6	32,8
16	19,37	23,5	26,3	28,8	32,0	34,3
17	20,5	24,8	27,6	30,2	33,4	35,7
18	21,6	26,0	28,9	31,5	34,8	37,2
19	22,7	27,2	30,1	32,9	36,2	38,6
20	23,8	28,4	31,4	34,2	37,6	40,0
21	24,9	29,6	32,7	35,5	38,9	41,4
22	26,0	30,8	33,9	36,8	40,3	42,8
23	27,1	32,0	35,2	38,1	41,6	44,2
24	28,2	33,2	36,4	39,4	43,0	45,6
25	29,3	34,4	37,7	40,6	44,3	46,9
26	30,4	35,6	38,9	41,9	45,6	48,3
27	31,5	36,7	40,1	43,2	47,0	49,6
28	32,6	37,9	41,3	44,5	48,3	51,0
29	33,7	39,1	42,6	45,7	49,6	52,3
30	34,8	40,3	43,8	47,0	50,9	53,7
40	45,6	51,8	55,8	59,3	63,7	66,8
50	56,3	63,2	67,5	71,4	76,2	79,5
60	67,0	74,4	79,1	83,3	88,4	92,0
70	77,6	85,5	90,5	95,0	100,4	104,2
80	88,1	96,6	101,9	106,6	112,3	116,3
90	98,6	107,6	113,1	118,1	124,1	128,3
100	109,1	118,5	124,3	129,6	135,8	140,2

6.a. taula. Snedecor-en F banaketako $P(F_{\nu_1,\nu_2} > F_\alpha) = \alpha$ *balioak (I)*

$$\alpha = 0{,}05$$



ν_2 \ ν_1	Zenbakitzailearen askatasun-graduak									
Zatitzailearen askatasun-graduak	1	2	3	4	5	6	7	8	9	10
1	161,00	199,00	216,00	225,00	230,00	234,00	237,00	239,00	241,00	242,00
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

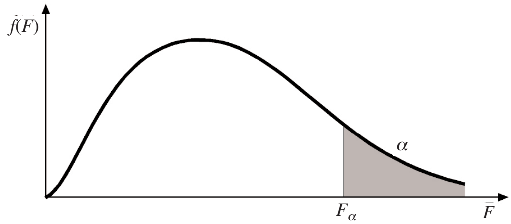
6.a. taula. Snedecor-en F banaketako $P(F_{\nu_1,\nu_2} > F_\alpha) = \alpha$ *balioak (II)*

$$\alpha = 0{,}05$$

ν_2 \ ν_1	Zenbakitzailearen askatasun-graduak								
Zatitzailearen askatasun-graduak	12	15	20	24	30	40	60	120	∞
1	244,00	245,90	248,00	249,00	250,00	251,00	252,20	253,30	254,00
2	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,37
6	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

6.b. taula. Snedecor-en F banaketako $P(F_{v_1,v_2} > F_\alpha) = \alpha$ *balioak (I)*

$$\alpha = 0{,}01$$



Zatitzailearen askatasun-graduak (rows, v_2); Zenbakitzailearen askatasun-graduak (columns, v_1)

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10
1	4.052	5.000	5.404	5.624	5.764	5.859	5.928	5.981	6.022	6.056
2	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47
	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32

6.b. taula. Snedecor-en F banaketako $P(F_{\nu_1,\nu_2} > F_\alpha) = \alpha$ *balioak (II)*

$$\alpha = 0{,}01$$

$\nu_2 \backslash \nu_1$	Zenbakitzailearen askatasun-graduak								
Zatitzailearen askatasun-graduak	12	15	20	24	30	40	60	120	∞
1	6.107	6.157	6.209	6.234	6.261	6.286	6.313	6339	6,37
2	99,42	99,43	99,45	99,46	99,47	99,48	99,48	99,49	99,50
3	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	14,37	14,20	14,02	13,93	13,84	13,75	13,65	15,56	13,46
5	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	4,16	4,01	3,86	3,78	3,70	3,62	3,54	5,45	3,36
13	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,85	3,17
14	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	2,96	2,70	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	2,93	1,76	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	2,90	1,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	2,87	1,74	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	2,84	1,73	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	2,66	2,52	2,37	2,29	2,20	2,11	2,01	1,92	1,81
60	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	2,34	2,19	2,03	1,95	1,86	1,77	1,06	1,55	1,38
∞	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Bibliografia

Ardanuy, R. eta Quintín M. (1993): *Estadística para Ingenieros*, Hespérides, Salamanca.

Canavos, G.C. (1988): *Probabilidad y Estadística. Aplicaciones y métodos,* McGraw-Hill, Mexiko.

Casas, J.M. (1996): *Inferencia Estadística para economía y administración de empresas,* Editorial Centro de estudios Ramón Areces, Madril.

Eguzkitza, J.M. (1998): *Apuntes de Métodos Estadísticos de la Ingeniería,* Geneve, Bilbo.

Etxeberria, J. (1998): *Estatistika. Teoria eta praktika SPSSWIN erabiliz,* Elhuyar, Usurbil.

Fernandez, K.; Orbe, J. eta Zubia, M. (1997): *Estatistikarako sarrera. Ariketak,* UEU, Bilbo.

Fernandez, K.; Orbe, J. eta Zubia, M. (1997): *Estatistika I eta Estatistika II. Ariketak,* UEU, Bilbo.

Johnson, R.A. (1997): *Probabilidad y Estadística para ingenieros de Miller y Freund*, Prentice Hall Hispanoamericana, Mexico.

Mendenhall, W. eta Sincich, T. (1997): *Probabilidad y Estadística para Ingeniería y Ciencias*, Prentice Hall Hispanoamericana, Mexiko.

Pérez, C. (2003): *Estadística. Problemas resueltos y aplicaciones,* Prentice Hall, Madrid.

Pérez, G. (1999): *Estadística Matemática,* UPV/EHU, Bilbo.

Sincich, T. eta MacClave, J.T. (2003): *Statistics,* Prentice Hall.

Spiegel, M.R. (1991): *Estadística,* McGraw-Hill, Madril.

Urkaregi, A. (1991): *Bioestatistika*, UEU, Bilbo.

Wilks, D. S. (1995): *Statistical methods in atmospheric sciences,* Academic Press, San diego, California.

Matematika Sailean argitaratu diren beste liburu batzuk

Bioestatistika

Arantza Urkaregi
1991n argitaratua
ISBN: 84-86967-37-6

Kalkulu diferentziala eta integrala I

Joserra Aizpurua eta Patxi Angulo (koord.)
1992an argitaratua
ISBN: 84-86967-47-3

Kalkulu diferentziala eta integrala II

Joserra Aizpurua eta Patxi Angulo (koord.)
1992an argitaratua
ISBN: 84-86967-48-1

Ekuazio diferentzialak. Aplikazioak eta ariketak

Elena Agirre (euskaratzailea)
1994an argitaratua
ISBN: 84-86967-63-5

Optimizazioa. Programazio lineala

Victoria Fernandez eta Ana Zelaia
1992an argitaratua
ISBN: 84-86967-65-1

Estatistikarako Sarrera

Karmele Fernandez
1996an argitaratua
ISBN: 84-86967-70-8

Estatistikarako Sarrera. Ariketak

Karmele Fernandez, Josu Orbe, Marian Zubia
1997an argitaratua
ISBN: 84-86967-80-5

Estatistika I eta Estatistika II. Ariketak

Karmele Fernandez, Josu Orbe, Marian Zubia
1996an argitaratua
ISBN: 84-86967-79-1

R Hasiberrientzat

Gorka Azkune eta Yosu Yurramendi (itzul.)
2005ean PDFn argitaratua
ISBN: 84-86967-077-7

Estatistikaren oinarriak. Ariketak

Elena Agirre Basurko
2006an argitaratua
ISBN: 84-8438-079-3

Kalkulu diferentziala eta integrala (2. argitalpena)

Patxi Angulo Martin (koord.)
2009an argitaratua
ISBN: 978-84-8438-236-2