INGENIARITZAKO METODO ESTATISTIKOAK
EBALUAZIO JARRAITURIK GABE
BIGARREN DEIALDIA 2015-2016

***Emaitzen argitalpena**: 2016ko ekainaren 27an, 19:00etan*
***Azterketen berrikuspena**: 2016ko ekainaren 30ean, 13:00etan (7 I 1 gela).*

1. ARIKETA

4 bola zuri (Z), 6 beltz (B) eta 8 gorri (G) dituen kutxa batetik hiru erauzketa egiten dira. Ondorengo atalak bi era desberdinetara ebatzi: itzulerarekin eta itzulerarik gabe.

(1.) Kalkula ezazu kolore bakoitzeko bola bat ateratzeko probabilitatea (**2 puntu**).

(A) Bolak zein ordenan agertzen diren kontuan izan gabe eta itzulerarik gabe:

$$P(E)=3!\cdot P(Z)\cdot P(G|Z)\cdot P(B|Z\cap G)=3!\cdot\frac{4}{18}\cdot\frac{8}{17}\cdot\frac{6}{16}=0{,}2352$$

Ordenak garrantzirik ez duenez hiru koloreak $P_3=3!=6$ era desberdinetara ager daitezke, ondorioz, aurreko probabilitatean 3! agertzen da. Ondorengo atalean gauza bera gertatzen da.

(B) Bolak zein ordenan agertzen diren kontuan izan gabe eta itzulerarekin:

$$P(E)=3!\cdot P(Z)\cdot P(G|Z)\cdot P(B|Z\cap G)=3!\cdot\frac{4}{18}\cdot\frac{8}{18}\cdot\frac{6}{18}=0{,}1975$$

Bolen ordena kontuan izan beharko balitz, emaitza desberdina izango litzateke. Adibidez, itzulerarekin lehenengo bola zuria, bigarrena gorria eta hirugarrena beltza izango balira, probabilitatea hurrengoa izango litzateke:

$$P(E)=P(Z)\cdot P(G|Z)\cdot P(B|Z\cap G)=\frac{4}{18}\cdot\frac{8}{18}\cdot\frac{6}{18}=0{,}0329$$

Azkenik, itzulerarik gabe:

$$P(E)=P(Z)\cdot P(G|Z)\cdot P(B|Z\cap G)=\frac{4}{18}\cdot\frac{8}{17}\cdot\frac{6}{16}=0{,}0392$$

(2.) Kalkula ezazu bi kolore desberdinetako bolak ateratzeko probabilitatea (**4 puntu**).

Itzulerarik gabe:

Gertaera posibleek ZB, GB eta ZG egitura dute (binaka hartutako hiru koloreen konbinazioak $C_2^3=\binom{3}{2}=3$)

Bestalde, hiru bola ateratzen ditugunez (dagokion) kasuan konkretuki ZZB, ZBZ, BZZ,BBZ,BZB,BBZ posibilitateak ditugu (permutazioak errepikapenekin, hiru elementu daude, bat bi aldiz errepikatzen da eta beste bat baten bakarrik: $PR_3^{21} = \frac{3!}{2!} = 3$, agertzen den kolore bakoitzeko).

Laburbilduz, S gertaerak 16 oinarrizko gertaera ditu:

GGB	**GBG**	**BGG**	**BBG**	**BGB**	**GBB**
ZZB	**ZBZ**	**BZZ**	**BBZ**	**BZB**	**ZBB**
ZZG	**ZGZ**	**GZZ**	**GGZ**	**GZG**	**ZGG**

Elkartutako probabilitateak ondorengoak izanik:

	$\frac{8}{18}\frac{7}{17}\frac{6}{16} = \frac{336}{4896} = 0.06863$			$\frac{6}{18}\frac{5}{17}\frac{8}{16} = \frac{240}{4896} = 0.04902$	
	$\frac{4}{18}\frac{3}{17}\frac{6}{16} = \frac{72}{4896} = 0.01471$			$\frac{6}{18}\frac{5}{17}\frac{4}{16} = \frac{120}{4896} = 0.024510$	
	$\frac{4}{18}\frac{3}{17}\frac{8}{16} = \frac{96}{4896} = 0.01961$			$\frac{8}{18}\frac{7}{17}\frac{4}{16} = \frac{224}{4896} = 0.04575$	

Aurreko taulan koloreek permutazio desberdinak adierazten dituzte.

Ondorioz, eskatutako probabilitatea ondokoa da:

$$\mathbb{P}(S) = 3\times\frac{(8\times7\times6)+(6\times5\times8)+(4\times3\times6)+(6\times5\times4)+(4\times3\times8)+(8\times7\times4)}{18\times17\times16} = 3\frac{1088}{4896} = 0.6667$$

Beste era batera, emaitza hurrengoa da:

$$\mathbb{P}(S) = \mathbb{P}(ZB)+\mathbb{P}(ZG)+\mathbb{P}(BG) = \frac{2}{17}+\frac{10}{51}+\frac{6}{17} = \frac{2}{3} = 0.6667$$

izan ere

$$\mathbb{P}(ZB) = P_3^{21}\left[\mathbb{P}(2Z\wedge 1B)+\mathbb{P}(1Z\wedge 2B)\right] = 3\left[\frac{4}{18}\frac{3}{17}\frac{6}{16}+\frac{4}{18}\frac{6}{17}\frac{5}{16}\right] = \frac{2}{17} = 0.11764$$

$$\mathbb{P}(ZG) = P_3^{21}\left[\mathbb{P}(2Z\wedge 1G)+\mathbb{P}(1Z\wedge 2G)\right] = 3\left[\frac{4}{18}\frac{3}{17}\frac{8}{16}+\frac{4}{18}\frac{8}{17}\frac{7}{16}\right] = \frac{10}{51} = 0.196078$$

$$\mathbb{P}(BG) = P_3^{21}\left[\mathbb{P}(2B\wedge 1G)+\mathbb{P}(1B\wedge 2G)\right] = 3\left[\frac{6}{18}\frac{5}{17}\frac{8}{16}+\frac{6}{18}\frac{8}{17}\frac{7}{16}\right] = \frac{6}{17} = 0.3529$$

Beste era batera eginda:

$$\mathbb{P}(ZB)=\mathbb{P}(2Z\wedge 1B)+\mathbb{P}(1Z\wedge 2B)=\frac{\binom{4}{2}\binom{6}{1}}{\binom{18}{3}}+\frac{\binom{4}{1}\binom{6}{2}}{\binom{18}{3}}=\frac{2}{17}=0.117647$$

$$\mathbb{P}(ZG)=\mathbb{P}(2Z\wedge 1G)+\mathbb{P}(1Z\wedge 2G)=\frac{\binom{4}{2}\binom{8}{1}}{\binom{18}{3}}+\frac{\binom{4}{1}\binom{8}{2}}{\binom{18}{3}}=\frac{10}{51}=0.196078$$

$$\mathbb{P}(BG)=\mathbb{P}(2B\wedge 1G)+\mathbb{P}(1B\wedge 2G)=\frac{\binom{6}{2}\binom{8}{1}}{\binom{18}{3}}+\frac{\binom{6}{1}\binom{8}{2}}{\binom{18}{3}}=\frac{6}{17}=0.352941$$

<u>Itzulerarekin:</u>

Kasu honetan, lagin espazioa berdina da, baina ondorengo irudian letra lodiz idatzita ez daudenei elkartutako probabilitateak desberdinak dira:

GGB	**GB**G	**BG**G	**BBG**	**BG**B	**GB**B
ZZB	**ZB**Z	**BZ**Z	**BBZ**	**BZ**B	**ZB**B
ZZG	**ZG**Z	**GZ**Z	**GGZ**	**GZ**G	**ZG**G

Oraingoan elkartutako probabilitateak hurrengoak dira:

	$\frac{8}{18}\frac{8}{18}\frac{6}{18}=\frac{384}{5832}=0.06584$			$\frac{6}{18}\frac{6}{18}\frac{8}{18}=\frac{288}{5832}=0.04938$	
	$\frac{4}{18}\frac{4}{18}\frac{6}{18}=\frac{96}{5832}=0.01646$			$\frac{6}{18}\frac{6}{18}\frac{4}{18}=\frac{144}{5832}=0.02469$	
	$\frac{4}{18}\frac{4}{18}\frac{8}{18}=\frac{128}{5832}=0.02195$			$\frac{8}{18}\frac{8}{18}\frac{4}{18}=\frac{256}{5832}=0.04390$	

Azkenik, probabilitatea kasu honetan ondokoa da:

$$\mathbb{P}(S)=3\times\frac{(8\times 8\times 6)+(6\times 6\times 8)+(4\times 4\times 6)+(6\times 6\times 4)+(4\times 4\times 8)+(8\times 8\times 4)}{18\times 18\times 18}=3\frac{1296}{5832}=0.6667$$

Era berean, emaitza aurreko kasuan lortutako era desberdinetara ere lor daiteke.

(3.) Kalkula ezazu ateratako bigarren bola gorria izateko probabilitatea, lehenengo bola beltza eta hirugarren bola zuria direla jakinik (**3 puntu**).

Itzulerarik gabe:

Ateratako lehenengo bola beltza dela dakigunez, kutxan 4 bola zuri, 8 bola gorri eta 5 bola beltz baino ez daudela suposa dezakegu:

Aintzat hartu behar diren kasuak ondokoak dira:

$$P(B \cap B \cap Z) = \frac{5}{17} * \frac{4}{16} = \frac{20}{272} = 0{,}0735$$

$$P(B \cap G \cap Z) = \frac{8}{17} * \frac{4}{16} = \frac{32}{272} = 0{,}1176$$

$$P(B \cap Z \cap Z) = \frac{4}{17} * \frac{3}{16} = \frac{12}{272} = 0{,}04412$$

Bayes-en teorema aplikatuz:

$$P(G/Z) = \frac{\frac{8}{17} * \frac{4}{16}}{\frac{5}{17} * \frac{4}{16} + \frac{8}{17} * \frac{4}{16} + \frac{4}{17} * \frac{3}{16}} = \frac{32}{64} = 0{,}5$$

Hurrengo eran ere emaitza bera lortzen da:

$$\mathbb{P}\left(G_2 \middle| B_1 \cap Z_3\right) = \frac{\mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right)}{\mathbb{P}\left(B_1 \cap Z_3\right)} = \frac{2/51}{4/51} = \frac{1}{2} = 0.50$$

izan ere

$$\mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right) = \frac{6}{18}\frac{8}{17}\frac{4}{16} = \frac{2}{51} = 0.0392157$$

eta

$$\mathbb{P}\left(B_1 \cap Z_3\right) = \mathbb{P}\left(B_1 \cap B_2 \cap Z_3\right) + \mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right) + \mathbb{P}\left(B_1 \cap Z_2 \cap Z_3\right) =$$
$$= \frac{6}{18}\frac{5}{17}\frac{4}{16} + \frac{6}{18}\frac{8}{17}\frac{4}{16} + \frac{6}{18}\frac{4}{17}\frac{3}{16} = \frac{4}{51} = 0.0784314$$

Itzulerarekin:

Kasu honetan, ateratako lehenengo bola beltza dela dakigun arren bola kutxara itzultzen denez, kutxan 4 bola zuri, 8 bola gorri eta 6 bola beltz daude. Aurreko kasuan egindako era berean eginez:

Oraingoan, aintzat hartu behar diren kasuak hurrengoak dira:

$$P(B \cap B \cap Z) = \frac{6}{18} * \frac{4}{18} = \frac{24}{324} = 0{,}0740741$$

$$P(B \cap G \cap Z) = \frac{8}{18} * \frac{4}{18} = \frac{32}{324} = 0{,}0987654$$

$$P(B \cap Z \cap Z) = \frac{4}{18} * \frac{4}{18} = \frac{16}{324} = 0{,}0493827$$

Bayes-en teorema aplikatuz, ondokoa lortzen da:

$$P(G/Z) = \frac{\frac{8}{18} * \frac{4}{18}}{\frac{6}{18} * \frac{4}{18} + \frac{8}{18} * \frac{4}{18} + \frac{4}{18} * \frac{4}{18}} = \frac{32}{72} = 0{,}444444$$

Era berean:

$$\mathbb{P}\left(G_2 \middle| B_1 \cap Z_3\right) = \frac{\mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right)}{\mathbb{P}\left(B_1 \cap Z_3\right)} = \frac{8/243}{2/27} = \frac{216}{486} = 0.444444$$

izan ere

$$\mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right) = \frac{6}{18}\frac{8}{18}\frac{4}{18} = \frac{192}{5832} = \frac{8}{243} = 0.0329218$$

eta

$$\mathbb{P}\left(B_1 \cap Z_3\right) = \mathbb{P}\left(B_1 \cap B_2 \cap Z_3\right) + \mathbb{P}\left(B_1 \cap G_2 \cap Z_3\right) + \mathbb{P}\left(B_1 \cap Z_2 \cap Z_3\right) =$$
$$= \frac{6}{18}\frac{6}{18}\frac{4}{18} + \frac{6}{18}\frac{8}{18}\frac{4}{18} + \frac{6}{18}\frac{4}{18}\frac{4}{18} = \frac{432}{5832} = \frac{2^4 3^3}{2^3 3^6} = \frac{2}{3^3} = 0.0740741$$

(4.) Lortutako emaitzetatik zer ondoriozta dezakezu? (**puntu 1**).

Ondorio nagusia emaitza desberdinak lortzen ditugula da:

a) Ordena kontuan izanez gero probabilitatea nabarmen jaisten da,

b) itzulera kontsideratzen bada probabilitatea ere jaitsi egiten da. Kasu honetan, orokorrean, ordena kontuan izaten denean baino gutxiago jaisten da.

2. ARIKETA

A ATALA: Metroak Bidezabal-etik daukan igaro-kadentzia 10 minutukoa dela dakigu. Geltokira heltzen den erabiltzaile batek itxaron behar duen denbora ondorengo dentsitate funtzioa duen zorizko aldagaia da:

$$f_1(x) = \begin{cases} \frac{1}{\alpha} & 0 < x < 10 \\ 0 & \text{beste kasuetan} \end{cases}$$

Metroa atzerapenez heltzen bada, itxaroteko denborak ondorengo dentsitate funtzioa jarraitzen du:

$$f_2(z) = \begin{cases} \frac{1}{\beta} e^{-z/7} & z > 0 \\ 0 & \text{beste kasuetan} \end{cases}$$

(1.) Kalkulatu α eta β balio errealak $f_1(x)$ eta $f_2(z)$ probabilitateko dentsitate funtzioak izateko (**2 puntu**).

X aldagai jarraitu orok ondorengoa bete behar du: $F(\infty) \equiv \mathbb{P}(X < \infty) @\int_{-\infty}^{\infty} f(z)\,dz = 1$. α-ren balioa kalkulatzeko, emandako banaketa uniformea eta jarraitua denez (0, 10) tartean, badakigu definizioz bere dentsitate funtzioa honakoa izango dela:

$$f_1(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & beste\ kasuetan \end{cases}$$

Banaketa biak berdinduz $\alpha = 10$ daukagu.

β-ren kalkulurako, dentsitate funtzioaren integrala aldagaiaren definizio-eremuan bat izan behar dela aplikatuko dugu:

$$1 = \int_0^\infty \frac{1}{\beta} e^{-z/7}\, dz$$

Integral inpropioa da, limite bat infinitu delako, beraz, limiteak hartuz, integralaren balioa kalkula dezakegu:

$$1 = \int_0^\infty \frac{1}{\beta} e^{-z/7} dz = \lim_{x\to\infty} \int_0^x \frac{1}{\beta} e^{-z/7} dz = \lim_{x\to\infty} \left(-\frac{7}{\beta} e^{-z/7} \right)_0^x = \lim_{x\to\infty} \left(-\frac{7}{\beta} e^{-x/7} + \frac{7}{\beta} \right) = \frac{7}{\beta}$$

Beraz:

$$1 = \frac{7}{\beta} \to \beta = 7$$

Beraz, banaketa era honetan geldituko da:

$$f_2(z) = \begin{cases} \frac{1}{7} e^{-z/7} & z \geq 0 \\ 0 & beste\ kasuetan \end{cases}$$

(2.) Metroa orduan heltzen bada, kalkulatu erabiltzaileak batezbeste itxaron behar duen denbora. Errepikatu prozesua metroa atzerapenez heltzen den kasurako (**puntu 1**).

X aldagaiaren batezbestekoaren eta bariantzaren kalkuluetarako, aldagaiak banaketa uniformea jarraitzen duela ohartu beharra daukagu:

$$\mu = \frac{b-a}{2} = \frac{10-0}{2} = 5\ minutu$$

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(10-0)^2}{12} = \frac{25}{3} = 8.3333\ minutu^2$$

a = 7 parametroa duen banaketa esponentzialarentzat, batezbestekoa eta bariantza kalkulatzeko, teorian ikusitakoaren arabera:

$$\mu = a = 7\ minutu$$
$$\sigma^2 = a^2 = 7^2 = 49\ minutu^2$$

(3.) Metroa astean egun baten atzerapenez heltzen dela jakinik, kalkulatu erabiltzaileak 5 minutu baino gehiago itxaron behar izateko probabilitatea (**3 puntu**).

$$P(T > 5) = P(R) * P(T > 5/R) + P(NR) * P(T > 5/NR)$$

$$P(T > 5/R) = \int_5^\infty \frac{1}{7} e^{-z/7} dz = \lim_{x\to\infty} \int_5^x \frac{1}{7} e^{-z/7} dz = e^{-5/7}$$

$$P(T > 5/NR) = \int_5^{10} \frac{1}{10} dx = \frac{1}{2}$$

Beraz, eskatutako probabilitatea:

$$P(T > 5) = P(R) * P(T > 5/R) + P(NR) * P(T > 5/NR) = \frac{1}{7} * e^{-5/7} + \frac{6}{7} * \frac{1}{2} = 0.4985$$

B ATALA: Egun zehatz batean, finantza-erakunde bateko akzioen kotizazioek -1€ eta 7€ artean oszilatzen dutela frogatu da, banaketa uniformea (errektangularra) jarraituz.

(1.) Aldakuntza horretarako *X* zorizko aldagaia eta dentsitate funtzioa definitu (**puntu 1**).

X zorizko aldagaiak banaketa uniformea jarraitzen du -1 eta 7 balioen artean, beraz, dentsitate funtzioa ondorengoa da:

$$f(x) = \begin{cases} 0 & x < -1 \\ \frac{1}{b-a} = \frac{1}{7+1} = \frac{1}{8} & -1 \le x \le 7 \\ 0 & x > 7 \end{cases}$$

(2.) Dentsitate funtzioa abiapuntutzat hartuz, kalkulatu banaketa funtzioa eta itxaropen matematikoa eta desbiderazio tipikoa (**3 puntu**). **Atal honetan, dagokien formulen aplikazioen ordez, emaitzak lortzeko erabilitako prozesua ebaluatuko da.**

X zorizko aldagai jarraituaren banaketa funtzioa ondorengoa da:

$$F(x)=\mathbb{P}(X\le x)=\int_{-\infty}^{x} f(z)\,dz=\begin{cases}\int_{-\infty}^{x} f(z)\,dz=0 & x<-1\\ \int_{-1}^{x}\frac{1}{8}dz=\frac{1}{8}(z|_{-1}^{x}=\frac{1}{8}(x+1) & -1\le x\le 7\\ 1 & x>7\end{cases}$$

Zorizko aldagai honen itxaropen matematikoa (itxarondako balioa):

$$\mu_X=E[X]=\int_{-\infty}^{x} z\ f(z)\,dz=\begin{cases}\int_{-\infty}^{x} f(z)\,dz=0 & x<-1\\ \int_{-1}^{x}\frac{z}{8}dz=\frac{1}{16}(z^2|_{-1}^{7}=\frac{1}{16}(49-1)=3 & -1\le x\le 7\Rightarrow \mu_X=E[X]=3\text{ €}\\ 0 & x>7\end{cases}$$

Analogoki, desbiderazio tipikoa honakoa izango da:

$$\sigma_X=\sigma[X]=Var[X]=+\sqrt{\int_{-\infty}^{x} z^2\ f(z)\,dz}=+\sqrt{E\left[X^2\right]-\left(E[X]\right)^2}$$

Baina $x<-1$ edo $x>7$ bada, f(z) = 0 denez, kalkulua murriztu egiten da:

$$\sigma_X^2=E\left[X^2\right]-\left(E[X]\right)^2=\int_{-1}^{7} z^2\ f(z)\,dz-3^2=\int_{-1}^{7}\frac{1}{8}z^2\ dz-3^2=\frac{1}{24}(z^3|_{-1}^{7}-9=\frac{344}{24}-9=\frac{128}{24}=\frac{16}{3}\Rightarrow$$

$$\sigma_X=\sqrt{\frac{16}{3}}=2.3094\text{ €}$$

3. ARIKETA

Igeriketako hautatzaile nazionalak Rio de Janeiro 2016an herrialdea ordezkatuko duen talde olinpikoa osatu behar du. Plaza baterako, 1500 metro libre probarako marka minimoa lortu duten bi igerilarien artean aukeratu behar du. Hautaketa egiteko, prestakuntza faseko bi urteetan bi igerilariek egindako denboren erregistro historikoa dauka (banaketa normala jarraitzen dute). Erabakia hartzeko, Igeriketa Federazio Nazionalak (IFN) azken sei hilabeteetan parte hartu duten probetan lortutako denborak erabiltzea erabaki du. Independentzia bermatzeko, bi igerilariak bat etorri ez diren probetan lortutako denborak aintzat hartuko dira soilik. Informazioa honako taulan agertzen da:

	PROBA							
	1	2	3	4	5	6	7	8
1. igerilaria	15:00:90	15:01:03	14:59:20	15:01:05	14:58:90	15:00:01	15:01:01	14:57:20
2. igerilaria	14:59:10	15:02:10	15:02:30	15:01:99	15:01:20	15:00:99	15:00:97	

IFN-k %90eko konfidantza maila ezarri eta ikerketa bat agindu du, aurkeztuko zaion informazio eta kalkulu guztien justifikazioa eskatzen duelarik.

OHARRA. Ariketaren ebazpenerako erabiliko den neurri-unitatea segundoa (s.) da. Ondorengo taulak probetan behar izandako denbora, segundotan, erakusten du:

	PROBA (denbora segundutan)							
	1	2	3	4	5	6	7	8
1. igerilaria	900,90	901,03	899,20	901,05	898,90	900,01	901,01	897,20
2. igerilaria	899,10	902,10	902,30	901,99	901,20	900,99	900,97	

Populazioko batezbestekoen eta bariantzen estimazio puntualak ondorengoak dira:

$$\hat{\mu}_1 = \bar{x}_1 = 899.9125\ s. \qquad \hat{\sigma}_1^2 = S_1^2 = 1.9461\ s.^2$$
$$\hat{\mu}_2 = \bar{x}_2 = 901.2357\ s. \qquad \hat{\sigma}_1^2 = S_2^2 = 1.1904\ s.^2$$

(1.) Hautatzaileak 1. igerilaria aukeratua izango dela uste du, prestakuntza osoan zehar egindako ibilbide hobea izan baita, baina IFN-k ezarritako irizpideak onartu behar ditu. Ikerketaren emaitzen arabera, zuzen al dago hautatzailea eta 1. igerilaria aukeratuko da? (**3 puntu**).

Bi batezbestekoen diferentziaren kontraste bat da. Hautatzaileak 1. igerilaria aukeratua izango dela uste duenez, egin beharreko hipotesian bere denboren batezbestekoa beste igerilariarena baino txikiagoa den kontrastatu behar da. Beraz, ondorengo **hipotesiak** planteatuko dira:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_a : \mu_1 < \mu_2 \end{cases} \Rightarrow \begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 < 0 \end{cases}$$

$n_1 = 8$ eta $n_2 = 7$ tamainak dituzten bi laginak independenteak dira, desbiderazio tipikoak ezezagunak dituzten populazio normaletatik datoz.

Beraz, lehenengo eta behin, laginak desbiderazio tipiko berdina duten bi populaziotik datozen kontrastatu beharko dugu. Horrela bada, bi bariantzen zatidurarako kontraste bat planteatuko dugu. Hipotesiak hauek dira:

$$\begin{cases} H_0 : \dfrac{\sigma_1^2}{\sigma_2^2} = 1 \\ \\ H_a : \dfrac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases} \Rightarrow \begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_a : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

$n_1 = 8$ eta $n_2 = 7$ tamainak dituzten bi laginak independenteak dira, eta μ_1 eta μ_2 batezbestekoak ezezagunak dituzten populazio normaletatik datoz.

Kasu honetan erabilitako **kontrastearen estatistikoa**: $\boxed{KE = \dfrac{S_1^2}{S_2^2}}$

Estatistiko honek Fisher−Snedecor-en banaketa jarraitzen du, zenbatzaile eta izendatzaileko askatasun graduak $\nu_1 = n_1 - 1 = 7$ eta $\nu_2 = n_2 - 1 = 6$ izanik, hurrenez hurren: $\boxed{KE : F_{7;6}}$.

Eskualde kritikoa, EK, ondorengo baldintza betetzen duten kontrasterako estatistikoaren balioekin osatuta dago:

$$\boxed{EK = \left\{ ke \notin \left[F_{\nu_1;\nu_2;1-\frac{\alpha}{2}}, F_{\nu_1;\nu_2;\frac{\alpha}{2}} \right] \mid EK : F_{7;6} \right\} = \left\{ ke > F_{\nu_1;\nu_2;\frac{\alpha}{2}} \cup ke < F_{\nu_1;\nu_2;1-\frac{\alpha}{2}} \mid KE : F_{7;6} \right\}}$$

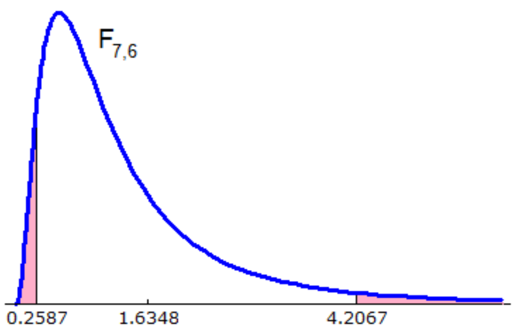
Emandako adierazgarritasunarentzat ($\alpha = 0.10$) eskualde kritikoaren limiteak:

- $L_1 = F_{7;6;0.95} = 0.258667$ (*R:* $qf(0.05,7,6)$; *Excel:* =DISTR.F.INV(0,95;7;6)
- $L_2 = F_{7;6;0.05} = \dfrac{1}{F_{6;7;0.95}} = 4.206658$ (*R:* $qf(0.05,6,7)$; *Excel:* =DISTR.F.INV(0,95;6;7)

Eta kontrasterako estatistikoaren balioa:

$$ke = \frac{S_1^2}{S_2^2} = \frac{1.9461}{1.1904} = 1.634839 \in (0.258687, 4.206658) \Rightarrow \boxed{ke \notin EK}$$

Kontrastearen estatistikoa ez dago eskualde kritikoaren barne. Beraz, $\alpha = 0.10$ adierazgarritasun mailarekin ez dago ebidentzia estatistikorik hipotesi nulua errefusatzeko. Hau dela eta, **laginak desbiderazio tipiko berdineko populazioei dagozkiela onartzen da**.



Orain, hasierako kontrastera bueltatuko gara. **Hipotesiak**:

$$\boxed{\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_a : \mu_1 < \mu_2 \end{cases} \Rightarrow \begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 < 0 \end{cases}}$$

$n_1 = 8$ eta $n_2 = 7$ tamainak dituzten bi laginak independenteak dira, desbiderazio tipikoak ezezagunak dituzten populazio normaletatik datoz; Hala ere, aurreko atalean desbiderazioak berdinak direla kontrastatu da.

Kasu honetan, erabilitako **estatistikoaren kontrastea** hauxe da: $\boxed{\dfrac{\bar{X}_1 - \bar{X}_2}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}}$

non: $S_P^2 = \dfrac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{n_1 + n_2 - 2}$

Estatistiko honek t-student banaketa jarraitzen du, askatasun graduak $\nu = n_1 + n_2 - 2 = 13$ izanik: $\boxed{KE : t_{13}}$

Eskualde kritikoa, EK, ondorengo baldintza betetzen duten kontrasterako estatistikoaren balioekin osatuta dago:

$$\boxed{EK = \left\{ ke \in \left(-\infty, -t_{\nu;\alpha}\right] \mid KE : t_{13} \right\} = \left\{ ke < -t_{\nu;\alpha} \mid KE : t_{13} \right\}}$$

$\alpha = 0.10$ adierazgarritasunarentzako eskualde kritikoaren limitea:

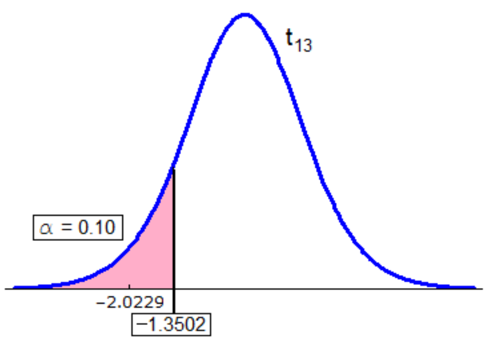
$L = -t_{13;\,0.10} = -1.350171$

(R: *qt(0.10,13,lower.tail=FALSE)=1.350171*; Excel: = DISTR.T.INV(0,2;13))

Eta kontrasterako estatistikoaren balioa:

$$\left.\begin{array}{l} \dfrac{\bar{X}_1 - \bar{X}_2}{S_P \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \\ S_P = 1.263832 \end{array}\right\} \Rightarrow ke = -2.022969 \in \left(-\infty, -1.350171\right] \Rightarrow \boxed{ec \in EK}$$

Kontrastearen estatistikoa eskualde kritikoan dago. Hau dela eta, $\alpha = 0.10$ adierazgarritasunarekin, ebidentzia estatistikoa dago hipotesi nulua errefusatzeko. Beraz, **1. igerilariaren denboren batezbestekoa beste igerilariarena baino txikiagoa dela onartzen da**, eta hautatzailearen ustea zuzena zen.



Behin erabakia hartuta eta igerilaria aukeratuta %5eko adierazgarritasunarekin:

(2.) Kalkulatu bere denboren batezbestekorako konfidantza tartea (**2 puntu**).

$n_1 = 8$ tamaina duen lagina desbiderazio tipiko ezezaguna duen populazio normaletatik datoz:

$$\left.\begin{array}{l} Normala \\ \sigma \ ezezaguna \end{array}\right\} \Rightarrow I_\mu^{1-\alpha} = \left[\bar{x} - t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}, \bar{x} + t_{n-1,\alpha/2} \frac{S}{\sqrt{n}} \right]$$

Ariketa ebazteko beharrezkoak diren lagineko estatistikoak hauek dira:

$$\bar{x} = 899.9125 s; \; S = \sqrt{1.9461} = 1.3950 s$$

Beraz, $t_{n-1,\alpha/2} = t_{7,0.025} = 2.3646$ (*R:* $qt(0.975,7)$; *Excel:* = DISTR.T.INV(0.05;7))

$$I_\mu^{0.95} = \left[899.9125 - 2.3646 \frac{1.395}{\sqrt{8}}, 899.9125 + 2.3646 \frac{1.395}{\sqrt{8}} \right] \Rightarrow I_\mu^{0.95} = [898.7463, 901.0787]$$

(3.) Lagin handia hartzea posiblea izango balitz, kalkula ezazu laginaren tamaina igerilariaren populazioaren eta laginaren batezbestekoen arteko diferentzia segundo laurdena baino txikiagoa izateko (**2 puntu**).

Tamaina handiko lagina izango balitz, hartu beharreko konfidantza tartea honakoa izango litzateke:

$$I_\mu^{1-\alpha} = \left[\bar{x} - z_{\alpha/2} \frac{S}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{S}{\sqrt{n}} \right]$$

$$P\left(\bar{x} - z_{\alpha/2} \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2} \frac{S}{\sqrt{n}} \right) = 1 - \alpha \Leftrightarrow P\left(|\mu - \bar{x}| \le z_{\alpha/2} \frac{S}{\sqrt{n}} \right) = 1 - \alpha$$

$$z_{\alpha/2} \frac{S}{\sqrt{n}} \le 0.25 \Rightarrow \sqrt{n} \ge z_{\alpha/2} \frac{S}{0.25}$$

$\alpha = 0.5 \Rightarrow \alpha/2 = 0.025 \Rightarrow z_{0.025} = 1.96$, orduan: $\sqrt{n} \ge 10.9368 \Rightarrow n \ge 119.6135 \Rightarrow n \ge 120$

(4.) Londresen izandako azken Joko Olinpikoen esperientziagatik, hautatzailearen iritziz marka 15 minutu baino baxuagoa izan behar da finalean sartzeko. Hautatzaileak bere igerilaria finalera heltzeko probabilitatea 0,25 baino altuagoa izatea kontrastatu nahi du. Horretarako, ondorengo erabaki-araua hartzen du: "erregistratutako zortzi marketatik gutxienez bost marketan 15 minututik jaitsi bada, hipotesi nulua errefusatzen da". Kontraste egokia planteatu eta bere potentzia kalkulatu (**3 puntu**).

Planteatu beharreko kontrastea:

$$\begin{cases} H_0 : p = 0.25 \\ H_a : p > 0.25 \end{cases}$$ non p= "igerilaria finalera heltzeko probabilitatea" den

Kontrastearen potentzia kalkulatzeko ondorengo probabilitate baldintzatua kalkulatu behar da:

$$Kontrastearen\ Potentzia = 1-\beta = P(\ H_0\ errefusatu \,|\, H_0\ \ gezurra)$$

Horretarako, erabaki-arauan definitutako zorizko aldagaia hartu behar da:

X = "15 minututik beherako marka kopurua (erregistratutako zortzi marketatik)"

non $X:\ Bin(8,p)$, p ="marka 15 minutu azpitik izateko probabilitatea" izanik

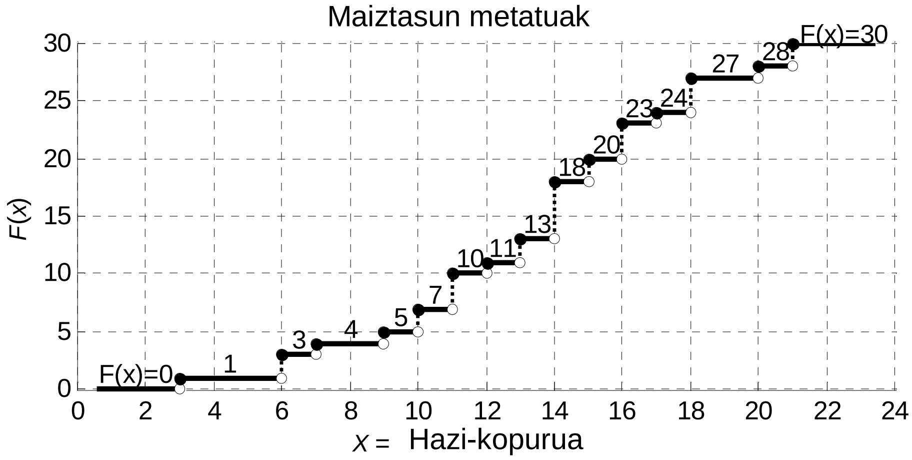
Beraz, erabaki-araua honakoa da: H_0 errefusatu $X \geq 5$ bada

$$Kontrastearen\ Potentzia = 1-\beta = P(\ H_0\ errefusatu \,|\, H_0\ \ gezurra) =$$
$$= P\left(X \geq 5 \,|\, \hat{p} = 3/8\right) = P\left(X \geq 5 \,|\, X:\ Bin\left(8, 3/8\right)\right) =$$
$$1 - P\left(X < 5 \,|\, X:\ Bin\left(8, 3/8\right)\right) = 1 - P\left(X \leq 4 \,|\, X:\ Bin\left(8, 3/8\right)\right) = 0.1374$$

(*R:* $pbinom\left(4,8,3/8\right)$; *Excel:* =DISTR.BINOM(4;8;3/8;1))

4. ARIKETA

Sagarrek fruitu bakoitzeko izaten duten hazi kopurua 0 eta 24 bitartekoa izaten da. Zoriz, sagardi batetik 30 sagar aukeratu dira eta bakoitzak dituen hazi kopurua aztertu da. Lortutako maiztasun metatuen diagrama honakoa da:



(1.) Arrazoitu *X* aldagaia kuantitatiboa edo kualitatiboa den, eta diskretua edo jarraitua den (**puntu 1**).

Emandako serie estatistikoa informazio kuantitatiboa (zenbakiak dira) eta diskretua (sagarrak ezin du hazi bat eta erdi eduki).

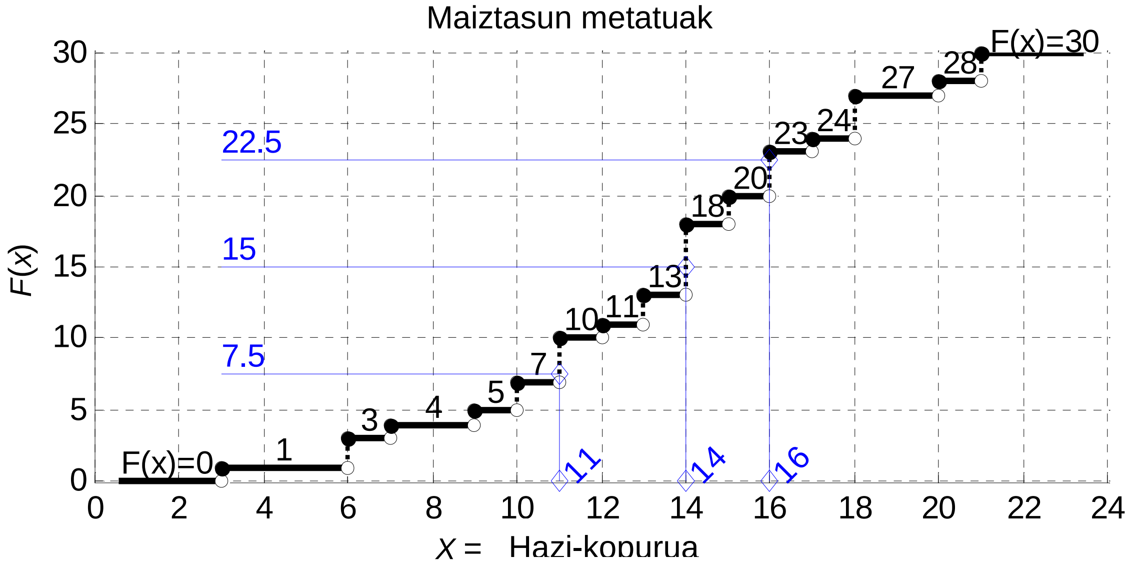
Aldagaia hala tratatuz:

(2.) Eman maiztasun taula (**1⁵ puntu**).

x_i	f_i	F_i	h_i	H_i
3	1	1	0.033333	0.033333
6	2	3	0.066667	0.1
7	1	4	0.033333	0.13333
9	1	5	0.033333	0.16667
10	2	7	0.066667	0.23333
11	3	10	0.1	0.33333
12	1	11	0.033333	0.36667
13	2	13	0.066667	0.43333
14	5	18	0.16667	0.6
15	2	20	0.066667	0.66667
16	3	23	0.1	0.76667
17	1	24	0.033333	0.8
18	3	27	0.1	0.9
20	1	28	0.033333	0.93333
21	2	30	0.066667	1
	30		1	

(3.) Lortu mediana eta Q_1 eta Q_3 kuartilak (**2 puntu**).

Mediana Q_2 kuartila da. Q_1 (25 % - 75 %), Q_2 (50 % - 50 %) eta Q_3 (75 % - 25 %) kuartilak oso erraz lor daitezke emandako maiztasun metatuen gainean:



Beraz, $Q_1 = 11$, $Me = Q_2 = 14$, $Q_3 = 16$; beraien unitateak haziak izanik

(4.) Kalkulatu moda (**puntu 1**).

Moda maiztasun maximoa duen X-ren balioa da, ariketa honetan maiztasun maximoa $f_i = 5$ (ikus lehenengo ataleko taula) da, beraz:

$\boxed{Moda = 14\ hazi}$.

Datuak 0 eta 24 arteko zabalera berdineko 6 klasetan taldekatuz:

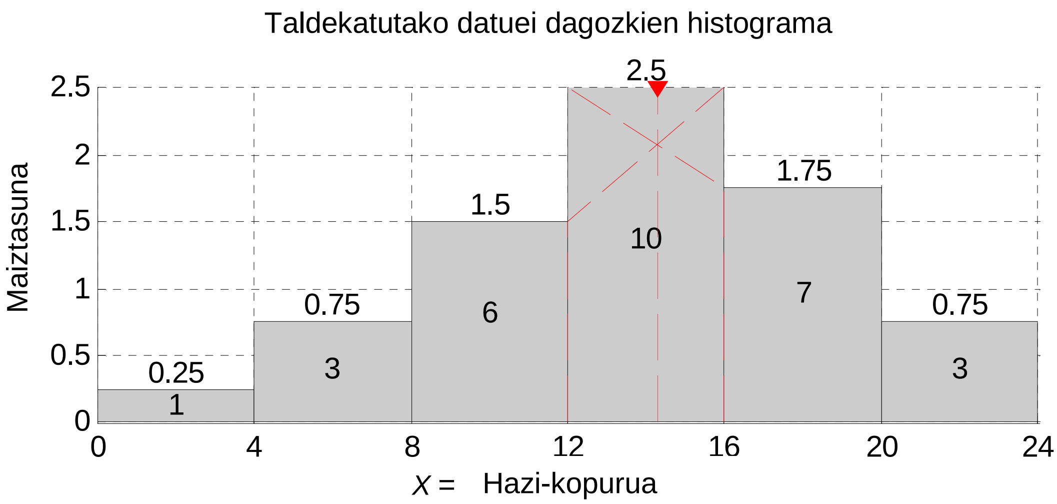
(5.) Eman arrazoituz maiztasun taula (**1^{5} puntu**).

Ariketak adierazten duenez, datuak tartekatzeko zabalera berdineko 6 klase hartu behar dira, beraz, klase bakoitzaren tarte-zabalera:

$$h\ @\ \frac{R}{6} = \frac{24-0}{6} = \frac{24}{6} = 4\ hazi$$

Class	l_i	u_i	d_i	m_i	f_i	F_i	h_i	H_i	fd_i
[0, 4)	0	4	4	2	1	1	0.033333	0.033333	0.25
[4, 8)	4	8	4	6	3	4	0.1	0.13333	0.75
[8, 12)	8	12	4	10	6	10	0.2	0.33333	1.5
[12, 16)	12	16	4	14	10	20	0.33333	0.66667	2.5
[16, 20)	16	20	4	18	7	27	0.23333	0.9	1.75
[20, 24]	20	24	4	22	3	30	0.1	1	0.75
			24		30		1		

(6.) Irudikatu histograma (**puntu 1**).



(7.) Estimatu batezbesteko aritmetikoa (**puntu 1**).

Taldekatutako datuen batezbestekoa lortzeko klase-markak eta klase bakoitzaren maiztasun behar dira:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{6} x_i \cdot f_i = \frac{2\cdot 1 + 6\cdot 3 + 10\cdot 6 + 14\cdot 10 + 18\cdot 7 + 22\cdot 3}{30} \approx 13.7\overline{3}\ hazi$$

(8.) Estimatu mediana. Hirugarren atalean lortutakoarekin bat al dator? Arrazoitu erantzuna. (**puntu 1**).

Jarraian agertzen den irudia ondoren datozen kalkuluetarako erabiliko da

[8, 12)	8	12	4	10	6	10	0.2	0.33333	1.5
[12, 16)	12	16	4	14	10	20	0.33333	0.66667	2.5
[16, 20)	16	20	4	18	7	27	0.23333	0.9	1.75

Medianak serie estatistikoaren datuen %50a (15 datu) ezkerrean uzten du, eta beste %50a eskuinean. Orduan, $Me \in [12, 16)$. Honakoa ondorioztatzen da:

$$Mediana = l_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} h = 12 + \frac{\frac{30}{2} - 10}{10} 4 = 12 + \frac{5}{10} 4 = 14 \; hazi$$

Modaren definizioa kontutan hartuz: $Mo \in [12, 16)$; kalkuluaren espresioa:

$$Moda = l_i + \frac{f_i - f_{i-1}}{2 f_i - f_{i-1} - f_{i+1}} h = 12 + \frac{10 - 6}{2 \times 10 - 6 - 7} 4 = 12 + \frac{4}{7} 4 = 14.2857 \; hazi$$

Lortutako emaitzetan, medianaren balioa serie estatistikoaren balioekin (taldekatu gabeko datuak) eta tarteetan banatutakoekin (taldekatutako datuak) berdina da (hain zuzen ere, $Mediana = 14 \; hazi$). Hala ere, hau ez da bat ere ohikoa. Honela, estatistikoen (batezbestekoa, desbiderazio tipikoa) kalkulurako jatorrizko serie estatistikoa erabiltzen bada, egin daitekeen errore bakarra biribiltze-errorea da. Aldiz, balioak klase-tarteetan taldekatzen baditugu, taldekatzearen ondorioz egindako mozketa-errorea (tarte bakoitzean dauden balioak klase-markaren balioagatik ordezkatzen dira) ere kontutan hartu beharra dago. Mozketa-errorea txikitu egiten da tarte kopurua handitzen den heinean.