eman ta zabal zazu
euskal herriko unibertsitatea — universidad del país vasco

Matematika Aplikatua Saila
Industria Ingeniaritza Teknikoko Unibertsitate Eskola
Euskal Herriko Unibertsitatea
Rafael Moreno, "Pitxitxi", 3 Pasealekua
48013 Bilbao

# INGENIARITZAKO METODO ESTATISTIKOAK

BIGARREN DEIALDIA 2012-2013

(a) *Azterketa egiteko denbora hiru ordukoa da, laborategiko proba barne, laborategiko proba ordu erdira jasoko da.*
(b) *Ezin da inolako bibliografiarik erabili, taula estatistikoak (taula hauetan ezin daiteke ezer idatzita egon) eta teoriako orria izan ezik.*
(c) *Ariketa bakoitza orri desberdin batean hasi behar da.*
(d) *Azterketa egiteko ikasle bakoitzak bere identifikaziorako agiria aurkeztu behar du.*
(e) ***Noten argitalpena:** 2013ko uztailaren 12an, 18:00etan.*
***Azterketen berrikusketa:** 2013ko uztailaren 16an, 10:00etan (MAL, 7I1 gela)*

## *1. ARIKETA* (ESTATISTIKA DESKRIBATZAILEA)

Kalitate zehatz bateko altzairu bereziz, manganesoarekin aleatutako altzairuz, osatutako 240 lagin aztertu dira. Zorizko aldagaiak banaketa normala duela suposatuz, azterketa honen emaitzak manganesoaren ehunekoetan (%) hurrengoak izan dira: $\overline{x} = 1.35\%,\ s = 0.21\%$. Beste analisi bat egiteko, enpresa bateko Kalitate-Sailak eskala kualitatibo bat ezarri nahi du, eskala honetan manganesoaren balio posibleak hurrengo irizpideen arabera sailkatzen dira:

| IRIZPIDEA | ESANAHIA | INTERPRETAZIOA |
|---|---|---|
| I | baxuegiak diren balioak | balio txikienen %5a |
| II | balio baxu onargarriak | hurrengo %20a |
| III | balio onargarriak | zentroko %50a |
| IV | balio altu onargarriak | hurrengo %20a |
| V | altuegiak diren balioak | balio altuenen %5a |

Zehaztu: (1. - 6 PUNTU) Aurreko sailkapenetik eratorritako Mn kontzentrazioa mugak.

Era ez zuzen batean, maiztasun-taula eman da. Demagun lagina era normalean banaturik dagoela. Orduan:

$$z = \frac{x - \overline{x}}{s} \Leftrightarrow x = \overline{x} + z \cdot s$$

Erreferentziazko probabilitatea erabiliz, klase bakoitzaren behe-muturra eta goi-muturra lortuko dira. Adibidez balio txikienen %5 (I. motakoak) hartuz, tipifikatutako puntuazioa honakoa izango da:

$$\mathbb{P}\left(Z \le z_1\right) = 0.05 \underset{simetria}{\Leftrightarrow} \mathbb{P}\left(Z \le |z_1|\right) = 0.95 \Rightarrow z_1 \underset{interpolatuz}{=} -1.644853627$$

$$\mathbb{P}\left(z_1 \le Z \le z_2\right) = 0.20 \Leftrightarrow \mathbb{P}\left(Z \le z_2\right) = 0.25 \underset{simetria}{\Leftrightarrow} \mathbb{P}\left(Z \le |z_2|\right) = 0.75 \Rightarrow z_2 \underset{interpolatuz}{=} -0.67448975$$

$$\mathbb{P}\left(z_2 \le Z \le z_3\right) = 0.50 \Leftrightarrow \mathbb{P}\left(Z \le z_3\right) = 0.75 \underset{simetria}{\Leftrightarrow} z_3 = -z_2 = 0.67448975$$

$$\mathbb{P}\left(z_3 \le Z \le z_4\right) = 0.20 \Leftrightarrow \mathbb{P}\left(Z \le z_4\right) = 0.95 \underset{simetria}{\Leftrightarrow} z_4 = -z_2 = 1.644853627$$

Honela manganesoaren kontzentrazio mailei buruzko maiztasun-taula lor daiteke:

| SAILKAPENA | DEFINIZIOA | $\ell_i(\%)$ | $\mathcal{L}_i(\%)$ |
|---|---|---|---|
| I | 5% | 0 | 1.004581 |
| II | 20% | 1.004581 | 1.208357 |
| III | 50% | 1.208357 | 1.491643 |
| IV | 20% | 1.491643 | 1.695419 |
| V | 5% | 1.695419 | ∞ |

(2.- 3 PUNTU) Egindako azterketaren balio zentralen %50a mugatzen duten Mn kontzentrazioak. Lagineko zein estatistiko zehaztu berri da? Kalkulu gehiago egin gabe, zein informazio gehigarri ondoriozta daiteke emaitza honetatik?

Emandako informaziotik hurrengoa ondoriozta daiteke:
Balio zentralen %50a barnean daukan erdiko tartea [1.208357, 1.491643] da, kuartilarteko heinari dagokion tartea izanik. Hortaz, hemendik lehen eta hirugarren ordenako kuartilak zehatz daitezke:

$$Q_1 = 1.208357\,\% \text{ y } Q_3 = 1.491643\,\%$$

(3.- 1 PUNTU) Zein da balio batek %1.63 gainditzeko probabilitatea?

$$\mathbb{P}(X \geq 1.63) = \mathbb{P}\left(Z \geq z_1 = \frac{1.63-1.35}{0.21}\right) = \mathbb{P}(Z \geq z_1 = 1.3333) = 1 - \mathbb{P}(Z \leq z_1 = 1.3333) =$$
$$= 1 - 0.908788726 = 9.1211274\,\%$$

## 2. ARIKETA (PROBABILITATEA)

Ezaguna da etil-zelulosan funtsatutako erretxinaren kolorea, disoluzioaren biskositatea eta etoxiloaren ehunekoa ezaugarri independenteak direla. Bestalde, aurreko ezaugarri bakoitzean banan-banan oinarritutako erretxinaren balorazioetan errore bat egoteko probabilitateak neurtu dira, probabilitate hauek 0.03, 0.05 eta 0.02 izanez, hurrenez hurren.

Izan bitez ondorengo gertaerak:

$E_1$ := "Kolorearen ebaluazio zuzena" $\quad p(\overline{E}_1) = 0.03$

$E_2$ := "Biskositatearen ebaluazio zuzena" $\quad p(\overline{E}_2) = 0.05$

$E_3$ := "Etoxiloaren ehunekoaren ebaluazio zuzena" $\quad p(\overline{E}_3) = 0.02$

2 PUNTU

Ondorioz, hiru ezaugarri hauek banan – banan era zuzenean neurtzeko, hau da errorerik gabe neurtzeko, probabilitateak hurrengoak dira:

$$p(E_1) = 0.97;\ p(E_2) = 0.95;\ p(E_3) = 0.98;$$

2 PUNTU

(1.-) Hiru ezaugarriak aldi berean baloratzen badira, zein da balorazioan errorerik ez egoteko probabilitatea?

$$p(E_1 \cap E_2 \cap E_3) = 0.97 \times 0.95 \times 0.98 = 0.903070$$

3 PUNTU

(2.-) Kalkula ezazu hiru ezaugarri horietatik ezaugarriren batek parte hartzen duen balorazio batean erroreren bat egoteko probabilitatea.

$$p(\overline{E}_1 \cup \overline{E}_2 \cup \overline{E}_3) =$$
$$p(\overline{E}_1) + p(\overline{E}_2) + p(\overline{E}_3) - p(\overline{E}_1 \cap \overline{E}_2) - p(\overline{E}_1 \cap \overline{E}_3) - p(\overline{E}_2 \cap \overline{E}_3) + p(\overline{E}_1 \cap \overline{E}_2 \cap \overline{E}_3) = 0.0969$$

Beste era batera:

$$p(\overline{E}_1 \cup \overline{E}_2 \cup \overline{E}_3) = 1 - p(E_1 \cap E_2 \cap E_3) = 1 - 0.903070 = 0.096930$$

3 PUNTU

## *<u>3. ARIKETA</u>* (ESTIMAZIOA ETA HIPOTESI-KONTRASTEA)

Beira desberdinez, konposizio ezberdina duten beirez, egindako bi ontzi-talde independenteen presio kritikoek hurrengo balioak zehaztu dituzte:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 TALDEA | 100 | 102 | 96 | 106 | 110 | 110 | 120 | 112 | 112 | 90 |
| 2 TALDEA | 104 | 88 | 100 | 98 | 102 | 92 | 96 | 100 | 96 | 96 |

Bi konposizioen populazioek banaketa normala dutela suposatzen da.

Oharra: 2 PUNTU ariketaren planteamendurako eta ebazpenaren justifikaziorako gordeko dira.

(1.- 2 PUNTU) Ezagunak izan ez arren, $\alpha = \%1$ adierazgarritasun mailaz, populazio bien bariantzak berdinak direla erabakitzeko ebidentzia estatistiko nahikoa al dago?

Populazioko bariantzei buruzko bi aldeko hipotesi-kontrastea da (populazioen banaketa normala dela suposatzen da):

$$\begin{cases} H_0: & \sigma_1^2 = \sigma_2^2 \\ H_a: & \sigma_1^2 \neq \sigma_2^2 \end{cases} \quad (\textbf{0.5 PUNTU})$$

Bi zorizko lagin bakunen estatistikoak hurrengoak dira:

| ESTATISTIKOA | 1 TALDEA | 2º TALDEA |
|---|---|---|
| $n$ | 10 | 10 |
| $\sum_{i=1}^{10} x_i$ | 1058 | 972 |
| $\bar{x}_i$ | 105.8 | 97.2 |
| $\sum_{i=1}^{10} x_i^2$ | 112644 | 94680 |
| $S$ | 8.8692 | 4.7329 |

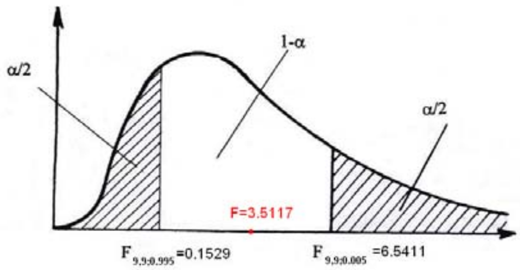
Eskualde kritikoa: $\left\{ S_1^2 / S_2^2 \notin \left[ F_{n-1,m-1;1-\alpha/2}, F_{n-1,m-1;\alpha/2} \right] \right\}$

Beraz, probabilitate banaketa Fisher-Snedecor-en banaketa da, onarpen eremuko muturrak hurrengoak izanik: (**0.5 PUNTU**):

$$\begin{cases} F_{n-1,m-1;\alpha/2} = F_{9,9;0.005} = IDF.F(0.995;9,9) = 6.5411 \\ F_{n-1,m-1;1-\alpha/2} = F_{9,9;0.995} = 1/F_{9,9;0.005} = 1/6.5411 = 0.1529 \end{cases}$$

Probarako estatistikoa (**0.5 PUNTU**):

$$F = \frac{S_1^2}{S_2^2} = \frac{8.8692^2}{4.7329^2} = \frac{78.6627}{22.4003} = 3.5117$$



Matematika Aplikatua Saila
Industria Ingeniaritza Teknikoko Unibertsitate Eskola
Euskal Herriko Unibertsitatea
Rafael Moreno, "Pitxitxi", 3 Pasealekua
48013 Bilbao

Beraz $F \in S_0$ dugunez, $\alpha$ =% 1 adierazgarritasun mailaz, aztertutako laginean oinarrituz ez da hipotesi nulua errefusatzeko ebidentzia estatistiko nahikoa existitzen (ondorioz, datu-talde biak bariantza berdina duten populazioetatik datoz) (**0.5 PUNTU**).

(2.- 3 PUNTU) $\alpha = \%1$ adierazgarritasun mailaz, 2. taldeak 1. taldeak baino presio kritiko altuagoa jasaten duela erabakitzeko ebidentzia estatistiko nahikoa al dago?

Oraingo honetan, batez bestekoei buruzko hipotesi-kontraste bat egin behar da. Hipotesiak ondorengoak dira (**0.5 PUNTU**):

$$\begin{cases} H_0: & \mu_1 \geq \mu_2 \\ H_a: & \mu_1 < \mu_2 \end{cases} \Leftrightarrow \begin{cases} H_0: & \mu_1 = \mu_2 \\ H_a: & \mu_1 < \mu_2 \end{cases}$$

Eskualde kritikoa (desbiderazio tipikoak ezezagunak dira baina berdinak, lehenengo atalean onartutako $H_0$ hipotesiaren arabera):

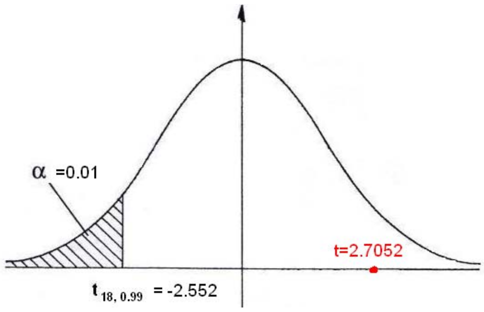
$$S_1 = \left\{ \bar{x}_1 - \bar{x}_2 < t_{n+m-2;1-\alpha} S_p \sqrt{\frac{1}{n} + \frac{1}{m}} \right\} = \left\{ \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}} < t_{n+m-2;1-\alpha} \right\}$$

Probarako estatistikoa (**0.5 PUNTU**)

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}} \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{8.6}{3.1790} = 2.7052$$

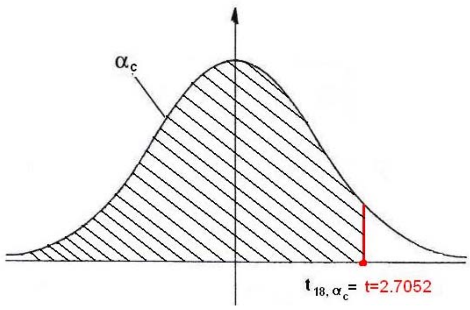
$\alpha$ = %1 adierazgarritasun maila bada, onarpen eremua eta eskualde kritikoa banatzen dituen balio kritikoa ondokoa da:

$$t_{18,0.99} = IDF.T(0.01,18) = -IDF.T(0.99,18) = -2.552379618$$




Hortaz $t_{18,0.99} \leq t$ denez (onarpen eremuan gaude), $\alpha$ =% 1 adierazgarritasun mailaz, aztertutako laginean oinarrituz ez da hipotesi nulua errefusatzeko ebidentzia estatistiko nahikoa existitzen (beraz, 2. taldeak ez du 1. taldeak baino presio kritiko altuagoa jasaten) (**1 PUNTU**).

(3.- 1 PUNTU) Kalkula ezazu kontrastearen p-balio hurbildua lortutako emaitza justifikatuz.




$$t_{18,\alpha_c} = 2.7052 \Rightarrow \alpha_c = p - balioa = CDF.T(2.7052,18) = \%99.27$$

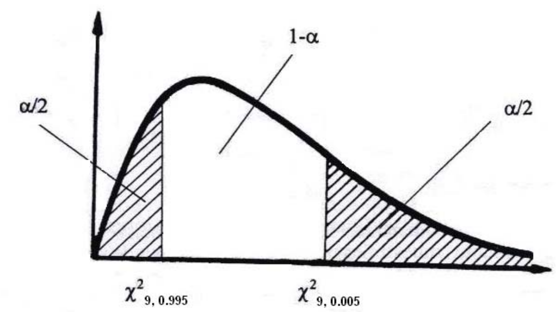
Beraz:

$\alpha > \alpha_c = p - balioa = \%99.27$ bada, $H_0$ errefusatu egingo da.

$\alpha \leq \alpha_c = p - balioa = \%99.27$ bada, $H_0$ onartu egingo da.

(4.- 2 PUNTU) %99eko konfiantza-mailaz, kalkula 1. taldeko bariantzaren konfiantza-tartea. Interpretatu aurreko estimazioa.

$\sigma_1^2$ -ren konfidantza tartea ($\mu$ ezezaguna): $I_{\sigma_1^2}^{1-\alpha} = \left[ \dfrac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}}, \dfrac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}} \right]$




$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.05,9} = 23.5589$$
$$\chi^2_{0.05,9} = IDF.CHISQ(0.995,9) = 23.5589$$
$$\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.995,9} = 1.7349$$
$$\chi^2_{0.995,9} = IDF.CHISQ(0.005,9) = 1.7349$$

Eskatutako konfiantza-tartea (**1 PUNTU**):

$$I_{\sigma_1^2}^{0.99} = \left[ \frac{9 \cdot 8.8692^2}{23.5589}, \frac{9 \cdot 8.8692^2}{1.7349} \right] = [30.0125, 408.0486]$$

**OHAR ARGIGARRIA**: Ariketa honetan, enuntziatuak adierazi bezala, erabilitako neurri unitatea presio kritikoa da.

euskal herriko unibertsitatea — universidad del país vasco

Matematika Aplikatua Saila
Industria Ingeniaritza Teknikoko Unibertsitate Eskola
Euskal Herriko Unibertsitatea
Rafael Moreno, “Pitxitxi”, 3 Pasealekua
48013 Bilbao

## *4. ARIKETA*

Izan bedi X “makina batera heltzen diren lanen helduera”, orduro 8 bezeroko tasa duen Poisson-en banaketari jarraitzen dion zorizko aldagaia. Izan bedi Y “makinak lan bat egiteko erabilitako denbora”, 5 minutuko batezbestekodun banaketa esponentziala duen beste zorizko aldagai bat.

Kalkulatu:

(1.- 2 PUNTU) Hurrengo orduan gutxienez lau bezero heltzeko probabilitatea.

$$\mathbb{P}(X \geq 4 \text{ bezero}) = 1 - \mathbb{P}(X < 4 \text{ bezero}) = 1 - \mathbb{P}(X \leq 3 \text{ bezero}) = 1 - F(3) = 1 - \sum_{i=0}^{3} \frac{e^{-8} 8^i}{i!} = 0.957619888$$

(2.- 2 PUNTU) X zorizko aldagaiaren banaketa-funtzioa lortu

$$F_X(a) = \sum_{k=0}^{a} \mathbb{P}(X = k) = \sum_{k=0}^{a} \frac{e^{-8}\lambda^i}{i!} = \sum_{k=0}^{a} \frac{e^{-8}8^i}{i!}$$

lortu heldutako bezero kopurua gutxienez 5 izateko baina zortzi baino gehiago ez izateko probabilitatea.

$$\mathbb{P}(5 \leq X \leq 8 \text{ bezero}) = F(8) - F(4) = 0.592547341\text{-}0.0996324\text{=}0.49291494$$

(3.- 2 PUNTU) Makina batek lan bat egiteko 3 minutu edo gutxiago erabiltzeko probabilitatea.

$$\mathbb{P}(Y \leq 3 \text{ minutu}) = \int_0^3 \tfrac{1}{5} e^{-\frac{1}{5}y} dy = -\left[ e^{-\frac{1}{5}t} \right|_0^3 = 1 - e^{-\frac{3}{5}} = 0.451188$$

(4.- 2 PUNTU) Makinak lan zehatz bat egiteko jadanik 3 minutu erabili baditu, zein da lana egiteko beste hiru minutu osagarri erabiltzeko probabilitatea?

$$\mathbb{P}(Y \geq 3 \text{ minutu + 3 minutu} | Y \geq 3 \text{ minutu}) =$$

$$\begin{cases} \text{(probabilitate baldintzatua)} & \dfrac{\mathbb{P}(Y \geq 6 \text{ minutu})}{\mathbb{P}(Y \geq 3 \text{ minutu})} = \dfrac{\int_6^\infty \frac{1}{5} e^{-\frac{1}{5}y} dy}{\int_3^\infty \frac{1}{5} e^{-\frac{1}{5}y} dy} = \dfrac{-\left[ e^{-\frac{1}{5}t} \right|_6^\infty}{-\left[ e^{-\frac{1}{5}t} \right|_3^\infty} = \dfrac{e^{-\frac{6}{5}}}{e^{-\frac{3}{5}}} = 0.5488 \\ \text{(zorizko aldagai esponentzialak ez dauka memoriarik)} & \int_3^\infty \frac{1}{5} e^{-\frac{1}{5}y} dy = -\left[ e^{-\frac{1}{5}t} \right|_3^\infty = e^{-\frac{3}{5}} = 0.5488 \end{cases}$$

(5.- 2 PUNTU) Makinak lanen %83.64a T minututan bukatzen baditu, zenbat balio du T-k?

$$\mathbb{P}(Y \leq T \text{ minutu}) = \int_0^T \tfrac{1}{5} e^{-\frac{1}{5}y} dy = -\left[ e^{-\frac{1}{5}t} \right|_0^T = 1 - e^{-\frac{T}{5}} = 0.8364 \Leftrightarrow T = 9.05165 \text{ minutu}$$

---

Ariketak ebazteko erabilgarriak izan daitezkeen SPSS-ko hainbat komando:

| | |
|---|---|
| CDF.NORMAL(1.3333,0,1) = 0.90878 | CDF.NORMAL(1.3333,1,1) = 0.63546 |
| IDF.F(0.99,9,9)=5.3511 | CDF.NORMAL(0.75,0,1)=0.7734 |
| IDF.NORMAL(0.95,0,1)=1.6448 | CDF.T(0.99,18)=0.8323 |
| IDF.CHISQ(0.005,9) = 1.7349 | IDF.T(0.99,18)=2.552 |
| CDF.T(2.7052,18)=0.9927 | CDF.CHISQ(2.7052,18)=0.00001 |
| IDF.NORMAL(0.75,0,1)=0.6745 | CDF.NORMAL(0.95,0,1)=0.8289 |
| IDF.F(0.995,9,9)=6.5411 | IDF.CHISQ(0.995,9) = 23.5589 |