

INGENIARITZAKO METODO ESTATISTIKOAK

BIGARREN DEIALDIA (UZTAILA 2019)

Ohar orokorrak:

Ariketaren iraupena: 2 ordu eta 30 minutu

1 ARIKETA

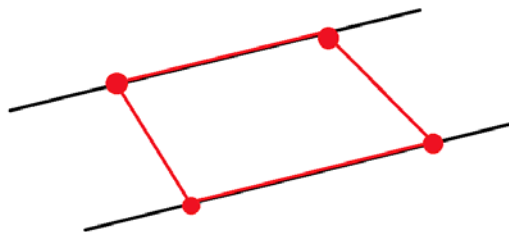
A ATALA

Izan bitez R_1 , R_2 eta R_3 izendatutako hiru zuzen paralelo. Zoriz, R_1 zuzenean zazpi puntu marrazten dira, R_2 zuzenean hamar puntu marrazten dira eta R_3 zuzenean lau puntu.

(A.1) Zenbat lauki ezberdin eraiki daitezke hauen erpinak R_1 eta R_2 zuzenetan dauden hamazazpi puntuetako lau puntutan egon daitezen? (puntu 1)

Lauki bat zehazteko R_1 eta R_2 zuzen bakoitzetik bi puntu hartu behar dira. Izan bedi C_1 puntuak binaka hartzeko modu ezberdin guztien kopurua R_1 zuzenean, eta C_2 puntuak binaka hartzeko modu ezberdin guztien kopurua R_2 zuzenean. Eraiki daitezkeen lauki guztien kopurua, G , ondorengoa da:

$$G = C_1 C_2 = \binom{7}{2} \binom{10}{2} = \frac{7!}{2!5!} \frac{10!}{2!8!} = 21 \times 45 = 945$$

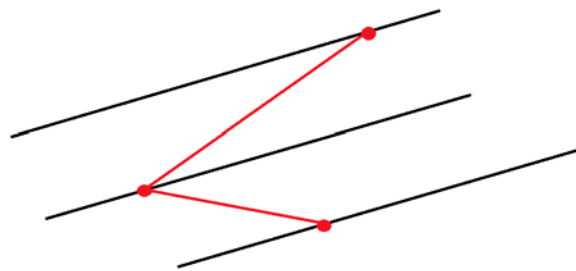


(A.2) Zenbat modu ezberdinetan elkar daitezke hiru zuzenak, hauetako puntu bakoitzetik igarotzen den lerro hautsi baten bitartez? (puntu 1)

Izan bitez: C_1 10 puntu banaka hartzeko modu ezberdin kopurua R_1 zuzenean; C_2 7 puntu banaka hartzeko modu ezberdin kopurua R_2 zuzenean; eta C_3 4 puntu banaka hartzeko modu ezberdin kopurua R_3 zuzenean. Modu berean ondorioztatzen da:



$$G_T = C_1 C_2 C_3 = \binom{10}{1} \binom{7}{1} \binom{4}{1} = 10 \times 7 \times 4 = 280$$



R_1 , R_2 eta R_3 zuzen bakoitzean gaizki marraztutako puntu bat dagoela suposatzen da. Zoriz zuzen bat aukeratzen da eta ondoren, zuzeneko puntu bat aukeratzen da.

(A.3) Zein da gaizki marraztutako puntu bat aukeratzeko probabilitatea?

(2 puntu)

Ondorengo gertaerak definitzen dira: $\bar{D} \triangleq$ "puntua gaizki marraztuta" eta $E_{R_i} \triangleq$ " R_i zuzena aukeratu, $i = 1, 2, 3$ ". Enuntziatuaren informazioaz baliatuz nabaria da $\mathbb{P}(E_{R_i}) = \frac{1}{3} \forall i$; $\mathbb{P}(\bar{D}|E_{R_1}) = \frac{1}{7}$; $\mathbb{P}(\bar{D}|E_{R_2}) = \frac{1}{10}$ eta $\mathbb{P}(\bar{D}|E_{R_3}) = \frac{1}{4}$. Probabilitate osoaren teorema erabiliz

$$\mathbb{P}(\bar{D}) = \sum_{i=1}^3 \mathbb{P}(E_{R_i}) \mathbb{P}(\bar{D}|E_{R_i}) = \frac{1}{3} \left(\frac{1}{7} + \frac{1}{10} + \frac{1}{4} \right) = \frac{1}{3} \frac{69}{140} = \frac{23}{140} = 0.164286$$

(A.4) Gaizki marraztutako puntu bat aukeratu dela jakinik, zein da puntu hori R_3 zuzeneko izateko probabilitatea?

(puntu 1)

Bayes-en teorema aplikatuz

$$\mathbb{P}(E_{R_3}|\bar{D}) = \frac{\mathbb{P}(E_{R_3}) \mathbb{P}(\bar{D}|E_{R_3})}{\mathbb{P}(\bar{D})} = \frac{\frac{1}{3} \frac{1}{4}}{\frac{23}{140}} = \frac{\frac{1}{12}}{\frac{23}{140}} = \frac{35}{69} = 0.507246$$

B ATALA

Izan bitez A eta B gertaerak, non $P(A) = 0.25$, $P(B|A) = 0.50$ eta $P(A|B) = 0.25$. Frogatu:

(B.1) \bar{A} eta \bar{B} gertaerak independenteak direla.

(3 puntu)

\bar{A} eta \bar{B} bi gertaera independenteak izan daitezen honakoa bete behar da: $\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\bar{A}) \mathbb{P}(\bar{B})$.

Morgan-en legeak aplikatuz, ondorioztatzen da:

$$\begin{aligned} \bar{A} \cap \bar{B} &= \overline{A \cup B} \Rightarrow \mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\overline{A \cup B}) = 1 - \mathbb{P}(A \cup B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) \\ &= 1 - \mathbb{P}(A) - \mathbb{P}(B) [1 - \mathbb{P}(A)] = [1 - \mathbb{P}(A)] [1 - \mathbb{P}(B)] = \mathbb{P}(\bar{A}) \mathbb{P}(\bar{B}) \end{aligned}$$

Hau da, A eta B gertaera independenteak badira hauen kontrako gertaerak ere independenteak izango dira, beste modu batean esanda:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Leftrightarrow \mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\bar{A})\mathbb{P}(\bar{B})$$

$$\mathbb{P}(B|A) = 0.50 = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A) = 0.50 \times 0.25 = 0.125$$

$$\mathbb{P}(A|B) = 0.25 = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Rightarrow \mathbb{P}(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A|B)} = \frac{0.125}{0.25} = 0.50$$

Beraz: $\mathbb{P}(A \cap B) = 0.125 = \mathbb{P}(A)\mathbb{P}(B) = 0.25 \times 0.50$, hau da, A eta B gertaera independenteak dira. Ondorioz, hauen kontrako gertaerak ere independenteak dira eta

$$\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\bar{A})\mathbb{P}(\bar{B}) = 0.75 \times 0.50 = 0.375.$$

(B.2) $(A|B) \cup (\bar{A}|\bar{B})$ gertaera segurua dela.

(2 puntu)

Kalkulatu dezagun:

$$\mathbb{P}[(A|B) \cup (\bar{A}|\bar{B})] = \mathbb{P}(A|B) + \mathbb{P}(\bar{A}|\bar{B}) - \mathbb{P}[(A|B) \cap (\bar{A}|\bar{B})] = 0.25 + 0.75 - 0 = 1$$

$$\Rightarrow (A|B) \cup (\bar{A}|\bar{B}) = \Omega$$

Zeren,

$$\mathbb{P}(\bar{A}|\bar{B}) = \frac{\mathbb{P}(\bar{A} \cap \bar{B})}{\mathbb{P}(\bar{B})} = \frac{0.375}{0.50} = 0.75$$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)} = \mathbb{P}(B); \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

$$\mathbb{P}(\bar{B}|\bar{A}) = \frac{\mathbb{P}(\bar{A} \cap \bar{B})}{\mathbb{P}(\bar{A})} = \mathbb{P}(\bar{B}); \mathbb{P}(\bar{A}|\bar{B}) = \frac{\mathbb{P}(\bar{A} \cap \bar{B})}{\mathbb{P}(\bar{B})} = \mathbb{P}(\bar{A})$$

$$(A|B) \cap (\bar{A}|\bar{B}) = A \cap \bar{A} = \emptyset \Rightarrow \mathbb{P}[(A|B) \cap (\bar{A}|\bar{B})] = \mathbb{P}(\emptyset) = 0$$

2 ARIKETA

A ATALA

Izan bedi $N(\mu, 40)$ banaketa normala jarraitzen duen populazio bat. Bertatik, n tamainako lagin bat hartzen da.

(A.1) Populazioko batezbestekoaren estimazio bat burutzen bada, zer egin beharko litzateke errorea laurdenera murrizteko? Justifikatu erantzuna. (2 puntu)

Populazio baten batezbestekoaren estimazioaren **errore estandarra** ondorengo espresioaren bidez zehazten da:

$$\sigma_{\hat{\mu}} = \frac{\sigma}{\sqrt{n}}$$

Populazio baten batezbestekoaren estimazioaren **errorea** ondorengo espresioaren bidez zehazten da:

$$\varepsilon = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Laurdenera murriztu nahi badugu, laginaren tamaina $16n$ izan beharko litzateke.

(A.2) %96-ko konfiantza-mailaz, zehaztu populazioko batezbestekoaren estimaziorako laginak eduki beharko lukeen tamaina, konfiantza tartearen zabalera zortzi unitatekoa izan dadin. (puntu 1)

Estimazioaren konfiantza tartearen zabalera, \mathcal{R} , honela definitzen da:

$$\mathcal{R}_{\alpha} = \mathcal{L}_{\alpha} - l_{\alpha} = \left(\hat{\mu} + z_{\alpha} \frac{\sigma}{\sqrt{n}} \right) - \left(\hat{\mu} - z_{\alpha} \frac{\sigma}{\sqrt{n}} \right) = 2z_{\alpha} \frac{\sigma}{\sqrt{n}} = 8 \text{ unitate}$$

Ordezkatuz,

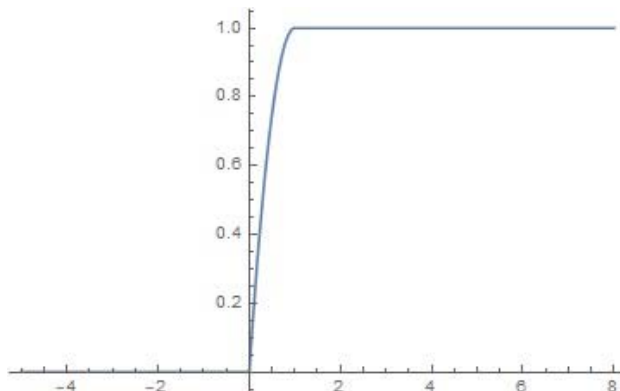
$$2z_{\alpha} \frac{\sigma}{\sqrt{n}} = 8 \Rightarrow n = \left(2z_{\alpha} \frac{\sigma}{8} \right)^2 = \left(2z_{96\%} \frac{40}{8} \right)^2 = (2 \times 2.053749 \times 5)^2 = 20.53749^2 = 421.7885 \Rightarrow n = 422$$

B ATALA

Izan bedi ondorengo banaketa-funtzioa duen X zorizko aldagaia:

$$F(x) = \begin{cases} 0 & x < 0 \\ x(2-x) & x \in [0,1] \\ 1 & x > 1 \end{cases}$$

Grafikoki,



(B.1) Lortu X zorizko aldagaiaren $f(x)$ dentsitate-funtzioa.

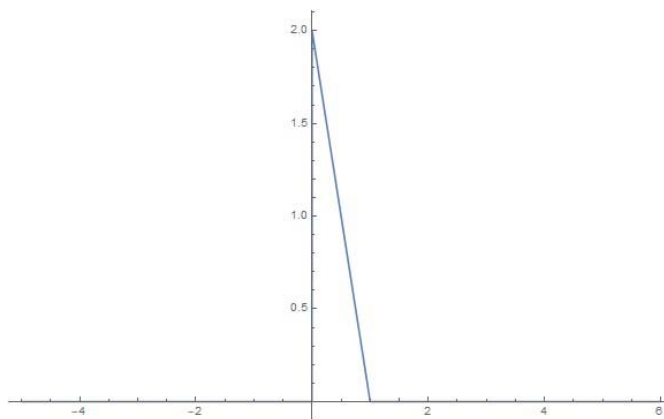
(2 puntu)

Jakina da $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(x)dx$. Beraz: $f(x) = \frac{dF(x)}{dx}$. Banaketa-funtzioa definituta dagoen zati

bakoitza deribatuz $f(x) = \frac{dF(x)}{dx} = \frac{d[x(2-x)]}{dx} = 2 - 2x \quad \forall x \in [0, 1]$. Laburbilduz, X zorizko aldagaiaren dentsitate-funtzioa:

$$f(x) = \begin{cases} 0 & x < 0 \\ 2 - 2x & x \in [0, 1] \\ 0 & x > 1 \end{cases}$$

Bere adierazpen grafikoa



(B.2) Kalkulatu X aldagaiaren ondorengo estatistikoak: batezbestekoa, moda, mediana eta bariantza.

(4 puntu)

Zorizko aldagaiaren batezbestekoa kalkulatzeko:

$$\mu_X = E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x(2-2x)dx = \left(x^2 - \frac{2}{3}x^3 \right) \Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3} = 0.3333 \text{ unitate}$$

Moda, Mo , dentsitate-funtzioak maximo erlatibo bat duen puntuan kokatzen da. Kasu honetan



densitate-funtzioak $x = 0$ puntuan du maximo erlatiboa, beraz, $M_0 = 0$.

Medianak, Me , banaketaren % 50, bere ezkerrean eta bere eskuinean uzten ditu, hau da:

$$F(x) = \mathbb{P}(X \leq Me) = \int_{-\infty}^{Me} f(x) dx = \int_0^{Me} (2-2x) dx = \left(2x - x^2\right) \Big|_0^{Me} = 2Me - Me^2 = 0.5$$

Bigarren mailako ekuazioa ebatziz, $Me = 0.292893$.

Azkenik, zorizko aldagaiaren bariantza, honela definitzen da:

$$\begin{aligned} \sigma_x^2 &= E\left[(X - \mu_x)^2\right] = E(X^2) - \mu_x^2 = \left(\int_{-\infty}^{\infty} x^2 f(x) dx\right) - \mu_x^2 = \int_0^1 x^2 (2-2x) dx - \frac{1}{3^2} = \\ &= \left(\frac{2}{3}x^3 - \frac{1}{2}x^4\right) \Big|_0^1 - \frac{1}{3} = \frac{2}{3} - \frac{1}{2} - \frac{1}{3^2} = \frac{1}{18} = 0.05556 \text{ unitate}^2 \end{aligned}$$

(B.3) Kalkulatu $\mathbb{P}(X \leq 0.40)$ eta $\mathbb{P}(X > 0.25 | X \leq 0.4)$

(2 puntu)

$$\mathbb{P}(X \leq 0.40) = F(a = 0.40) = \int_{-\infty}^{0.40} f(x) dx = \int_0^{0.40} (2-2x) dx = \left(2x - x^2\right) \Big|_0^{0.40} = 0.8 - 0.16 = 0.64$$

$$\mathbb{P}(X > 0.25 | X \leq 0.4) = \frac{\mathbb{P}[(X > 0.25) \cap (X \leq 0.4)]}{\mathbb{P}(X \leq 0.4)} = \frac{0.2025}{0.64} = 0.31640625$$

Zeren,

$$\mathbb{P}[(X > 0.25) \cap (X \leq 0.4)] = \int_{0.25}^{0.4} f(x) dx = \int_{0.25}^{0.4} (2-2x) dx = \left(2x - x^2\right) \Big|_{0.25}^{0.40} = 0.2025$$

Emitza probabilitate honen ezberdina izanik:

$$\mathbb{P}(X > 0.25) = 1 - \mathbb{P}(X \leq 0.25) = 1 - \int_{-\infty}^{0.25} f(x) dx = 1 - \int_0^{0.25} (2-2x) dx = 1 - \left(2x - x^2\right) \Big|_0^{0.25} = 0.4375$$

3 ARIKETA

Iparramerikako automobilgintzako ingeniari talde batek auto mota baten frenatze-sistamarako diseinu berri bat proposatu du. Oraingo sistemarako, zehaztutako baldintzapean, frenatze-tartearen batezbestekoa 40 milla/ordu-tan 120 oinetakoa da. Diseinu berria praktikan jartzeko lagineko datuek, frenatze-tartearen batezbestekorako murrizketa bat adierazi behar dute. **(A)** Kontrasterako aztertu nahi den parametroa definitu eta beharrezko hipotesiak zehaztu. (2 puntu)

Kontrasterako aztertu nahi den parametroa, zehaztutako baldintzapean, frenatze-tartearen batezbestekoa, μ , da. Egin beharrezko hipotesi kontrastea:

$$\begin{cases} H_0: \mu = 120 \text{ oin} \\ H_a: \mu < 120 \text{ oin} \end{cases}$$

(B) Suposatu frenatze-sistema berrirako frenatze-tarteak banaketa normal bat jarraitzen duela $\sigma = 10$ oin izanik. Izan bedi \bar{X} 36 behaketa dituen zorizko lagin bakun baterako frenatze-tartearen batezbestekoa, ondorengo zein eskualde da egokia hipotesi nulua errefusatzeko? (2 puntu)

$$\begin{aligned} R_1 &\triangleq \{ \bar{X} / \bar{X} \geq 124.80 \text{ oin} \} \\ R_2 &\triangleq \{ \bar{X} / \bar{X} \leq 115.20 \text{ oin} \} \\ R_3 &\triangleq \{ \bar{X} / \bar{X} \leq 114.87 \text{ oin} \cup \bar{X} \geq 125.13 \text{ oin} \} \end{aligned}$$

\bar{X} geroz eta 120 oin baina txikiagoa izan, orduan eta hipotesi nulua errefusatzeko probabilitate handiagoa egongo da. Beraz, R_2 da **(A)** atalean zehaztutako hipotesi nulua errefusatzeko eskualde egokia.

(C) Zein da **(B)** atalean aukeratutako eskualderako adierazgarritasun-maila? (2 puntu)

(B) atalaren arabera, μ -ren hipotesi kontraste bat burutu behar da populazioaren bariantza ezaguna izanik. Kontrastearen estatistikoa z_{KE} honakoa da:

$$z_{KE} = \frac{\bar{X} - \mu_0}{\sigma_{\hat{\mu}}} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \stackrel{H_0}{\sim} \mathcal{N}(\mu = 0, \sigma = 1)$$

Definizioz, adierazgarritasun-maila I. motako errorea egitearen probabilitatea da, hau da:

$$\begin{aligned} \alpha &= \mathbb{P}(H_0 \text{ errefusatu} | H_0 \text{ egia}) = \mathbb{P}(\bar{X} \leq 115.20 \text{ oin} | \mu = 120 \text{ oin}) = \mathbb{P}\left(Z \leq z_{KE} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{115.20 - 120}{10 / \sqrt{36}}\right) = \\ &= \mathbb{P}(Z \leq z_c = -2.88) = 1 - \mathbb{P}(Z \leq 2.88) = 1 - 0.9980116 = 0.001988376 \end{aligned}$$

(D) Nola aldatuko litzateke eskualde kritikoa 0.001-eko adierazgarritasun-maila duen kontraste bat edukitzeko? (2 puntu)

Eskualde kritiko berria definitzeko z_l balio bat aurkitu behar dugu non $\alpha = \mathbb{P}(Z \leq z_l) = 0.001$ izango den. Banaketa normalaren taula erabiliz edo R-ko komandoa erabiliz (`qnorm(0.001, 0, 1)`), $z_l = -3.090232$ izango da. Beraz:

$$\alpha = 0.001 = \mathbb{P}\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - 120}{10/\sqrt{36}} \leq z_l = -3.090232\right) = \mathbb{P}(\bar{X} \leq 114.85 \text{ oin})$$

Ondorioz, $\alpha = 0.001$ adierazgarritasun-maila duen kontrastearen eskualde kritikoa:

$$R \triangleq \{\bar{X} / \bar{X} \leq 114.85 \text{ oin}\}$$

(E) Zein da diseinu berria martxan ez jartzeko probabilitatea frenatze-tarte erreala 115 oinekoa izanik eta (B) ataleko eskualde kritikoa erabiltzen bada? (2 puntu)

Kasu honetan II. motako errorea egiteko probabilitatea kalkulatu behar dugu, hau da, hipotesi-nulua onartzeko probabilitatea $\mu = 115$ oin izanik. Beraz,

$$\beta = \mathbb{P}(H_0 \text{ onartu} | H_0 \text{ gezurra}) = \mathbb{P}(\bar{X} \geq 115.20 \text{ oin} | \mu = 115 \text{ oin}) = \mathbb{P}\left(Z \geq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{115.20 - 115}{10/\sqrt{36}}\right) = \mathbb{P}(Z \geq z_1 = 0.12) = 1 - \mathbb{P}(Z \leq z_1 = 0.12) = 1 - 0.5478 = 0.4522$$

Ariketak egiteko erabilgarria izan daitekeen informazioa:

R-ko funtzioak erabiliz lortutako balioak

| | | |
|---|---|---|
| <code>qt(0.95, 5) = 2.015048</code> | <code>pt(0.95, 5) = 0.8071313</code> | <code>pnorm(0.12, 0, 1) = 0.5478</code> |
| <code>pnorm(2.88, 0, 1) = 0.998</code> | <code>pbinom(4, 200, 0.0168) = 0.7525971</code> | <code>tinvs(0.025, 5) = -2.5705818</code> |
| <code>tcdf(0.975, 5) = 0.81283044</code> | <code>tcdf(2.824263, 5) = 0.981538519</code> | <code>pbinom(4, 200, 0.0007) = 0.9999996</code> |
| <code>pbinom(4, 200, 0.019608) = 0.6443085</code> | <code>distr.t(0.95; 5; 1) = 0.192869</code> | <code>qnorm(0.98, 0, 1) = 2.0537</code> |
| <code>pbinom(4, 200, 0.000817) = 0.9999992</code> | <code>tinvs(0.05, 5) = -2.0150484</code> | <code>pt(2.82426, 5) = 0.9815385</code> |
| <code>qt(0.975, 5) = 2.570582</code> | <code>qnorm(0.001, 0, 1) = -3.0902</code> | <code>binocdf(4, 200, 0.0007) = 0.9999996196</code> |