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

INGENIERITZAKO METODO ESTATISTIKOAK

LEHENENGO DEIALDIA 2013-2014

***Emaitzen argitalpena**: 2014ko maiatzaren 29an 17:00etan.*
***Azterketen berrikuspena**: 2014ko ekainaren 3an 10:00etan (7 I 1 gela).*

---

1. ARIKETA

A ATALA: Batetik seira zenbakitutako sei gutun-azaleko multzo batetik biko tamaina duten laginak ateratzen dira itzulerarik gabe, eta M=”**lortutako balio maximoa**” zorizko aldagaia kontsideratzen da. (1.) Esperimentuaren lagin-espazio egoki bat adierazi (2 puntu).

Ω lagin espazioak $\binom{6}{2} = 15$ oinarrizko elementu ditu (hau da, 15 lagin daude):

	1	2	3	4	5	6	7	8	9
Lagina	{1, 2}	{1, 3}	{1, 4}	{1, 5}	{1, 6}	{2, 3}	{2, 4}	{2, 5}	{2, 6}
M	2	3	4	5	6	3	4	5	6

	10	11	12	13	14	15
Lagina	{3, 4}	{3, 5}	{3, 6}	{4, 5}	{4, 6}	{5, 6}
m	4	5	6	5	6	6

(2.) M-ren probabilitate funtzioa eta banaketa funtzioa lortu (2 puntu).

Oinarrizko elementuak taldekatuz eta probabilitateak taula batean elkartuz, hurrengo probabilitate funtzioa daukagu

	1	2	3	4	5
M	2	3	4	5	6
p(m)	$1/15$	$2/15$	$3/15$	$4/15$	$5/15$

Banaketa funtzioa ondorengoa da

$$F(M) = \begin{cases} 0 & si \quad M < 2 \\ \frac{1}{15} & si \quad 2 \le M < 3 \\ \frac{3}{15} & si \quad 3 \le M < 4 \\ \frac{6}{15} & si \quad 4 \le M < 5 \\ \frac{10}{15} & si \quad 5 \le M < 6 \\ 1 & si \quad M \ge 6 \end{cases}$$

(3.) M bere batezbestekoa baino handiagoa izateko probabilitatea kalkulatu (2 puntu).

$\mu_M = E[M] = \sum_{\forall i} x_i \mathbb{P}(M = x_i) = \frac{70}{15} = 4.6667$ denez, ondorengoa ondorioztatzen da

$$\mathbb{P}(M > E[M] = 4.6667) = \frac{4}{15} + \frac{5}{15} = \frac{9}{15} = 0.6000$$

B ATALA: A pieza ekoizteko A1 eta A2 osagaiak erabiltzen dira. A akastuna ez izateko A1 eta A2 ezin daitezke akastunak izan. Osagai hauek akastunak izateko probabilitatea 0.2 eta 0.05 dira, hurrenez hurren. A piezekin batera B piezak ere ekoizten dira. B pieza hauek, ez akastunak izango dira kasuen %90ean.
Oharra: Pieza eta osagai bakoitza akastuna izatea edo ez, beste pieza eta osagaiekiko independentea da

Gertaera ondorengoa da: $X \triangleq$ "pieza edo osagaia egokia (ez akastuna) da". Enuntziatutik hurrengoa ondorioztatzen da

$$\mathbb{P}(A) = \mathbb{P}(A1 \cap A2) = \mathbb{P}(A1) \times \mathbb{P}(A2) = 0.8 \times 0.95 = 0.76$$

$$\mathbb{P}(B) = 0.9$$

(1º) A eta B piezen proportzio bera duen multzo batetik pieza bat zoriz hartuz gero, zein da pieza hori akats gabea izateko probabilitatea? (2 puntu).

Probabilitate osoaren teorema aplikatuz

$$P(\text{Ez akastuna}) = P(\text{Ez akastuna}|A)P(A) + P(\text{Ez akastuna}|B)P(B) =$$

$$0.76 \times 0.5 + 0.9 \times 0.5 = 0.83$$

(2º) Hartutako pieza akastuna bada, zein da A motakoa izateko probabilitatea? (2 puntu).

Bayes-en teorema aplikatuz

$$P(A|\text{Akastuna}) = \frac{P(\text{Akastuna}|A)P(A)}{P(\text{Akastuna}|A)P(A) + P(\text{Akastuna}|B)P(B)} =$$

$$\frac{0.24 \times 0.5}{0.24 \times 0.5 + 0.1 \times 0.5} = 0.706$$

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

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

2. ARIKETA

SOLETE herriko kaian neguan olatuek hartzen duten altuerari buruzko ikerketa bat egin da. Lortutako serie estatistikoa honakoa da:

OLATUEN ALTUERA (m)	BEHAKETA KOPURUA
0	6
1	4
2	9
3	10
4	8
5	5
6	6
7	2
8	2

Klaseetan taldekatu gabeko datuen zorizko lagina da.

(1.) Maiztasun erlatiboen eta maiztasun metatuen taula eraiki (puntu 1).

ALTURA DE LA OLA (m)	f_i	F_i	h_i	H_i
0	6	6	11.5384615	11.5384615
1	4	10	7.69230769	19.2307692
2	9	19	17.3076923	36.5384615
3	10	29	19.2307692	55.7692308
4	8	37	15.3846154	71.1538462
5	5	42	9.61538462	80.7692308
6	6	48	11.5384615	92.3076923
7	2	50	3.84615385	96.1538462
8	2	52	3.84615385	100
	52		1	

(2.) Batezbesteko aritmetikoa, desbiderazio tipikoa, bariantza eta moda kalkulatu (2 puntu).

Batezbesteko aritmetikoa: $\overline{x} = \frac{1}{n}\sum_{i=0}^{8} f_i x_i = \frac{175}{52} = 3.3654\ m$

Bariantza: $s^2 = \frac{1}{n}\sum_{i=0}^{8} f_i\left(x_i - \overline{x}\right)^2 = 4.5396\ \mathrm{m}$

Desbiderazio tipikoa: $s = 2.1306\ \mathrm{m}$

Moda (horiz) $Mo = x_i / \max\ f_i = 3\ m$

(3.) P_{12} kalkulatu, eta lortutako emaitza interpretatu (puntu 1).

P_{12} pertzentila serie estatistikoaren (zorizko lagin adierazgarriaren) datuen %12a bere ezkerrean uzten duen datua da, hau da:

$$\frac{12n}{100} = 6,24 \text{ denez, } \left.\begin{array}{l} F_1 < 6,24 \\ F_2 > 6,24 \end{array}\right\} \Rightarrow P_{12} = x_2 = 1\ m.$$

Beraz, olatuen %12ak 1m edo gutxiago neurtzen dute.

Beste era batera:

$P_{12} = x_{\frac{12n}{100}} = x_{0.12n} = x_{6.24} = 1.06\ m$ balio hau serieko lehenengo eta bigarren datuen arteko interpolazioa egin ondoren lortu da (maiztasun taulako zati laranja begiratu).

(4.) Zein olatu portzentajek du 5.87m baino altuera handiagoa? (2 puntu).

5.87 metro baino altuera gehiago duten olatuak 6, 7 eta 8 metroko olatuak dira, hortaz:

$$P = 11.5384615 + 3.84615385 + 3.84615385 = \%19.2307692$$

Beste era batera:

Kalkulatu beharreko ehunekoa:

$$P = p_1 + 11.5384615 + 3.84615385 + 3.84615385 = 20.4808\ \%$$

da, $p_1 = \dfrac{0.13 \times 5 \times 100}{52} = 1.2500$, $x_6 = 5\ \mathrm{m}$ datuari dagokion proportzioa izanik (maiztasun taulan berdez dauden tarteak begiratu)

(5.) Zein da zorizko lagin bakunaren olatu txikienen %30aren altuera maximoa? (2 puntu)

$$\frac{30n}{100} = 15.6 \text{ denez, } \left.\begin{array}{l} F_2 < 15,6 \\ F_3 > 15,6 \end{array}\right\} \Rightarrow P_{30} = x_3 = 2\ m.$$

Beste era batera:

Atal hau egiteko maiztasun erlatibo metatuetan oinarri gaitezke (maiztasun taulan urdinez dauden tarteak begiratu)

$$h_{X \le 30\,\%} = 1 + \frac{17.2076923}{30 - 19.2307692} = 2.6071\ \mathrm{m}$$

(6.) Emandako serie estatistikoan datu atipikoak al daude? Arrazoitu egindako justifikazioak (2 puntu).

Datu arraroak (existitzen badira) finkatu ahal izateko kuartil desberdinak kalkulatzen ditugu:

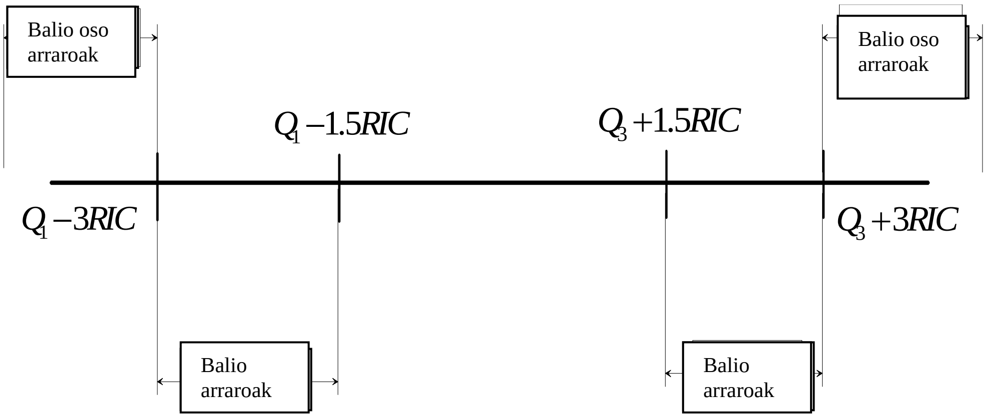
$$Q_1 = P_{25} = x_{\frac{25n}{100}} = x_{0.25n} = x_{13} = 2\ m$$

$$Q_2 = P_{50} = Me = x_{\frac{50n}{100}} = x_{0.5n} = x_{26} = 3\ m \qquad RIC = Q_3 - Q_1 = 5 - 2 = 3\ m$$

$$Q_3 = P_{75} = x_{\frac{75n}{100}} = x_{0.75n} = x_{39} = 5\ m$$

Kanpo eta barne heziak ondorengoak izanik:

$$M_1 = Q_1 - 3RIC = -7\ m$$
$$m_1 = Q_1 - 1.5RIC = -2.5\ m$$
$$m_3 = Q_3 + 1.5RIC = 9.5\ m$$
$$M_3 = Q_3 + 3RIC = 14\ m$$



Ondorioz, ez dago balio arrarorik serie estatistiko osoa $(m_1, m_3) = (-2.5\ m, 9.5\ m)$ tartean baitago.

Aurreko kalkulu guztiak taldekatutako datuak erabiliz egin daitezke, kasu honetan tarteen banaketa ondorengoa da:

ALTURA DE LA OLA	l_i	L_i	f_i
0	-0,5	0,5	6
1	0,5	1,5	4
2	1,5	2,5	9
3	2,5	3,5	10
4	3,5	4,5	8
5	4,5	5,5	5
6	5,5	6,5	6
7	6,5	7,5	2
8	7,5	8,5	2
			52

Datuak taldekatzean lortutako emaitzak eta ebazpen honetan lortutako emaitzak konparatuz, datuak taldekatzen hurbilketak lortzen direla esan beharra dago.

3. ARIKETA

Populazio bateko 2000 pertsonako lagin batetik, 3 pertsonek medikamendu baten kontrako erreakzioa izan dute.

Behatu daitekeen bezalaxe, zorizko aldagaiaren banaketa binomiala da:

$X \triangleq \mathcal{B}\left(n = 2000,\ p = \text{z.l.b. erabiliz kalkulatzear}\right).$

(1.) Estimatu populazio horretan kontrako erreakzioa duten pertsonen proportzioa, estimazioaren desbiderazio tipikoa adieraziz (2 puntu).

$$\hat{p} = \frac{3}{2000} = 0.0015 = 1.5\times10^{-3}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{1.5\times10^{-3}\times0.9985}{2000}} = 0.000865376$$

(2.) %90eko konfiantza-mailaz, populazioaren proportzioaren konfiantza-tartea kalkulatu (2 puntu).

$np = 3$ eta $nq = 1997 > 5$ (edo baliokidea dena: n altua eta p<1) betetzen direnez, hurrengo hurbilketa aplika daiteke $X \triangleq \mathcal{B}(n,\ p) \to \mathcal{N}(\mu,\sigma)$, hortaz $z_\alpha = \pm1.64485363$ izanik. Bukatzeko kalkulatu behar den konfiantza tartea ondorengoa da:

$$I_p^{1-\alpha} = \left[\hat{p} \pm z_\alpha\sqrt{\frac{pq}{n}}\right] = \left[0.0015 \pm 1.64485363\times0.000865376\right] =$$

$$= \left[0.000076583,\ 0.002923417\right]$$

(3.) Zein da 1000 pertsonako lagin batetik

(a) 3 pertsonek kontrako erreakzioa izateko probabilitatea? (2 puntu)

Probabilitatea era zehatzean kalkulatuz:

$$\mathbb{P}(X = 3) = F(3) - F(2) = \mathbb{P}(X \le 3) - \mathbb{P}(X \le 2) = \binom{1000}{3} 0.0015^3 (1-0.0015)^{997} =$$

$$(500\times333\times998)\,0.0015^3\,(0.9985)^{997} = (166167000)\times\left(3.3750\times10^{-9}\right)\times\left(2.238850\times10^{-1}\right) =$$

$$0.12555776$$

Probabilitatea era hurbilduan kalkulatuz:

$$\lambda = np = 0.0015\times1000 = 1.5 < 5$$
$$p < 0.1$$

betetzen direnez, Poisson-en hurbilketa aplika daiteke. Hortaz, Poisson-en banaketa erabiliz:

$$\mathbb{P}(X=3)=e^{-\lambda}\left.\frac{\lambda^k}{k!}\right|_{k=3}=e^{-1.5}\frac{1.5^3}{3!}=0.12551072$$

(b) 2 pertsona baino gutxiagok kontrako erreakzioa izateko probabilitatea? (2 puntu).

Hau da: $F(2)=\mathbb{P}(X<2)=\mathbb{P}(X\le 1)=\mathbb{P}\left(X=0\right)+\mathbb{P}\left(X=1\right)$

Banaketa Binomiala zuzenean aplikatuz:

$$\mathbb{P}(X<2)=\sum_{k=0}^{1}\binom{1000}{k}0.0015^k\left(1-0.0015\right)^{1000-k}=0.9985^{999}\left(0.9985+1000\times 0.0015\right)=0.5576998$$

Poisson-en banaketa aplikatuz, berriz:

$$\mathbb{P}(X<2)=\sum_{k=0}^{1}e^{-1.5}\frac{1.5^k}{k!}=e^{-1.5}\left(1+1.5\right)=2.5e^{-1.5}=0.5578254$$

(4.) 100000 pertsonako lagin batetik, zein da 140 pertsonek baino gutxiagok kontrako erreakzioa izateko probabilitatea? (2 puntu)

Lagina tamaina kontutan hartuz hurbilketak aplikatzea komeni da, kasu honetan

$$pn=0.0015\times 10^5=150>5$$
$$qn=0.9985\times 10^5=150>5$$

betetzen direnez, ezin daiteke Poisson-en hurbilketa aplikatu, Normalaren hurbilketa aplikatu behar da. Horretarako hurrengo estatistikoen balioak ezagutu behar dira:

$$\mu=np=150$$
$$\sigma_x=\sqrt{npq}=\sqrt{np(1-p)}=\sqrt{10^5\times 1.5\times 10^{-3}\times 0.9985}=12.23825968$$

Bestalde, jatorrizko banaketa diskretua denez jarraitutasun zuzenketa aplikatu behar da. Hortaz,

$$p(X<140)=p(X\le 139.5)$$

$$X=139.5\Rightarrow Z=\frac{X-\overline{x}}{s_x}=\frac{139.5-150}{12.23825968}=-0.857965125$$

$$p(X\le 139.5)=p(Z\le -0.857965125)=0.1955$$

4. ARIKETA

Journal of Human Movement Studies-en argitaratutako "*Practice and Fatigue Effects on the Programming of a Coincident Timing Response*" artikuluaren arabera, neke baldintzek erantzunaren portaera azaltzen duten mekanismoak distortsionatu egiten dituzte. 15 ikaslerekin esperimentu bat egin zen. Ikasle hauek zenbait mugimendu egiteko trebatu ziren, eta hauek egiteko beharrezkoa zuten denbora kontrolatu zen. Esperimentua jarduera egin aurretik (nekerik gabe) eta ondoren (nekearekin) egin zen eta hurrengo balioak erregistratu ziren:

NORBANAKOA	DENBORA SEGUNDUTAN	
	NEKERIK GABE	NEKEAREKIN
1	158	91
2	92	59
3	65	215
4	98	226
5	33	223
6	89	91
7	148	92
8	58	177
9	142	134
10	117	116
11	74	153
12	66	219
13	109	143
14	57	164
15	85	100

Ariketak egiteko denborek nekerik gabe eta nekearekin banaketa normala jarraitzen dutela suposatzen bada

Hurrengo taulan laburbilduta dauden balio guztiek enuntziatuarekin bat datorren denbora unitatea dute: segundo (s). Lagin hauetako estatistikoak ondorengoak dira:

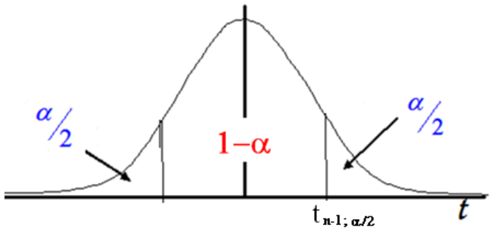
ESTATISTIKOAK	NEKERIK GABE	NEKEAREKIN	BIKOTEEN DIFERENTZIA
n	15		
$\overline{x}$	92.7333	146.8667	-54.1333
s	35.19747466	53.81804737	80.1880013
$S \equiv \hat{s}$	36.43284755	55.70697144	83.002467

(1.) %5eko adierazgarritasun mailaz, lortu ariketak nekerik gabe egiteko batezbesteko denboraren konfiantza-tartea (2 puntu).

Populazio bateko (lagin bateko) batezbesteko aritmetikoari buruzko ariketa bat da, laginaren tamaina txikia izanik (n = 15 < 30), hortaz, Student-en t banaketa erabili behar da. Hortaz, konfiantza tartea hurrengo eran adieraz daiteke:

$$I_{\mu_1}^{1-\alpha} = \left[\overline{x}_1 \pm t_{n-1;\alpha/2} \frac{S_1}{\sqrt{n}} \right]$$

Kalkuluak egin ondoren tartea ondorengoa izanik



$$t_{14;0.025} = IDF.T\left(0.975,14\right) = 2.14478668$$

$$I_{\mu_1}^{1-\alpha} = \left[\overline{x}_1 \pm t_1 \frac{S_1}{\sqrt{n}}\right] = \left[92.7333 \pm 2.144786681 \times 9.406920789\right]$$
$$I_{\mu_1}^{0.975} = \left[72.55749491, 112.9091718\right]$$

Hau da, tarte-estimazioaren interpretazioa

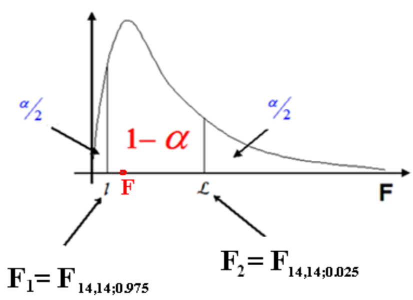
$$P\left(72.55749491 \leq \mu_1 \leq 112.9091718\right) = 0.95$$

(2.) Bi laginak independenteak direla suposatuz esperimentu bietan kontrol gradua bera dela kontrastatu %5eko adierazgarritasun mailaz (2 puntu).

Bi populazioko bariantzen kontraste bati buruzko ariketa bat da

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

Hau da, bi aldeko hipotesi kontrastea da: eskualde edo eremu kritikoa probabilitate banaketaren bi muturretan dago. Ariketa honetan erabili beharreko banaketa eredua Fisher-Snedecor da, balio kritikoak (eskualde kritikoaren mugak) ondorengoak izanik:



$$\begin{cases} F_1 = F_{14,14:0.975} = F_{14,14;0.025}^{-1} = 2.978587524^{-1} = 0.335729601 \\ F_2 = F_{14,14;0.025} = 2.978587524 \end{cases}$$

Probarako edo kontrasterako estatistikoa:

$$F = \frac{S_1^2}{S_2^2} = \frac{36.43284755^2}{55.70697144^2} = 0.427727464$$

$F_1 \leq F \leq F_2$ betetzen denez, α = %5 adierazgarritasun mailaz, aztertutako laginetan oinarrituz ez da hipotesi nulua errefusatzeko ebidentzia estatistiko nahikoa existitzen (hortaz, bietan kontrol gradua bera dela ondoriozta daiteke)

(3.) %95eko konfiantza-mailaz ariketak nekearekin eta nekerik gabe egiteko batezbesteko denboren arteko diferentziaren konfiantza-tartea zehaztu (2 puntu).

Batezbesteko aritmetikoei buruzko ariketa bat da (datuak parekatutako datuak dira, enuntziatuan datuak pertsona berarenak direla ziurtatzen baita). Tamaina txikiko bi populazio (bi lagin) daude, (n = 15 < 30), hortaz aplikatu beharreko banaketa eredua Student-en t banaketa da:

Kasu honetan diferentziak erabiliz lan egin behar da: $d_i = x_1 - x_2, \forall i \in [1, 15]$. Ondorioz, konfiantza tartea hurrengo eran adieraz daiteke:

$$I_{\bar{d}}^{1-\alpha} = \left[\bar{d} \pm t_{14;0.025} \frac{S_{\bar{d}}}{\sqrt{n}}\right] = [-54.13333333 \pm 2.144786681 \times 21.43114482] =$$

Kalkuluak egin ondoren tartea hurrengoa izanik

$$I_{\bar{d}}^{0.95} = [-100.0985673, -8.168099361]$$

Tarte-estimazioaren interpretazioa hurrengoa da

$$P(-100.0985673 \leq \bar{d} \leq -8.168099361) = 0.95$$

(4.) %95eko konfiantza mailaz, nekeak ariketak egiteko denbora aldatzen duen zehaztu (2 puntu).

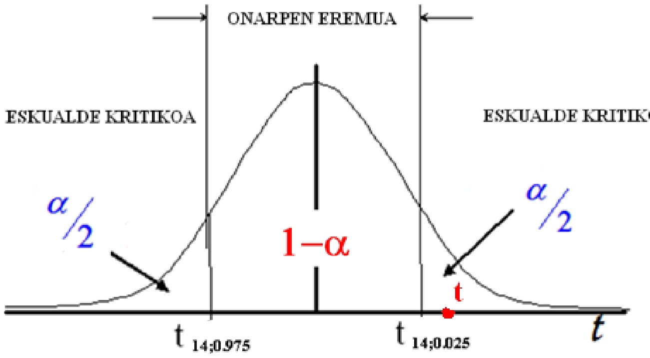
Hipotesi kontraste hau, bi populazio ez independenteen (parekatutako datuen) kontraste bat dena, bi aldeko hipotesi kontraste bat bezala bidera daiteke (ondorioz, eskualde edo eremu kritikoa banaketaren mutur bietan egonik).

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_a: & \mu_1 \neq \mu_2 \end{cases} \Leftrightarrow \begin{cases} H_0: & \mu_1 - \mu_2 = 0 \\ H_a: & \mu_1 - \mu_2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0: & \mu_d = 0 = \mu_0 \\ H_a: & \mu_d \neq 0 = \mu_0 \end{cases}$$

Laginaren tamaina txikia denez erabili beharreko probabilitate eredua Student-en t banaketa da, hortaz, balio kritikoak (eskualde kritikoaren mugak) ondorengoak dira:

$$\left\{|\bar{d} - \mu_0| > t_{n-1;\alpha/2} S / \sqrt{n}\right\} = \left\{\frac{|\bar{d} - \mu_0|}{S / \sqrt{n}} > t_{n-1;\alpha/2}\right\}$$

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



$$t_{14;0.025} = IDF.T(0.975,14) = +2.144786681$$

$$-t_{14;0.025} = t_{14;0.975} = IDF.T(0.025,14) = -2.144786681$$

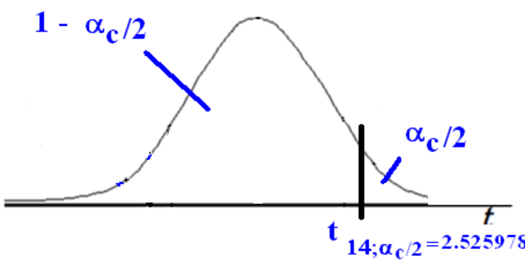
Kontrasterako estatistikoa:

$$t = \frac{\left|\bar{d} - 0\right|}{\frac{S_{\bar{d}}}{\sqrt{n}}} = \frac{54.133}{21.4305} = 2.525978$$

$t > t_{14;0.025}$ denez, α = %5 adierazgarritasun mailaz, aztertutako lagin bietan oinarrituz hipotesi nulua errefusatzeko ebidentzia estatistiko nahikoa existitzen dela ondoriozta daiteke; hau da, nekeak ariketak egiteko denbora aldatzen duela suposatzeko ebidentziak daude.

(5.) Aurreko kontrastearen p-balioa kalkulatu (2 puntu).

p-balioaren interpretazioa ondorengoa da



$$1-\alpha_c / 2 = CDF.T\left(2.525978,14\right) = 0.98789058 \Leftrightarrow \text{p-valor} = \alpha_c = 0.02421884 = \%2.421884$$