

# ESTADÍSTIKA ENPRESARA APLIKATUA

2023: Azterketa ebatziak

Irakaslea: Josemari Sarasola

Donostiako Ekonomia eta Enpresa Fakultatea. EHU

Egilea: Josemari Sarasola



Gizapedia

[gizapedia.org](http://gizapedia.org)

## ESTADÍSTIKA ENPRESARA APLIKATUA

Irakaslea: Josemari Sarasola (Bernoulli prozesuetatik inferentzia eta balidazio-probetara)

Data: 2023ko ekainaren 2an, 10:00

Iraupena: 1h 20min.

Oharra: Axotaz zein arkatzez egin daiteke, baino beti garbi eta txukun.

**I. ebazkizuna** (*Ebaluazio ez jarraitua: 1.5 puntu*)

Larrialdiko mediku pediateren ordutegiak planifikatzeko, haurrak larrialdietara egunean zehar noiz heltzen diren aztertzen duen txosten bat egin behar da. Susmoa dago goizeko 6etatik 12etara, haur larrialdiak behereneko maila batera iristen direla, zoriz banatuta ordu tarte horri legezkiokoenak baino haur larrialdi gutxiago izaten dela alegia, eta beraz ordu horretan pediatriako larrialdi-zerbitzuek lehentasuna galdu beharko luketela. Susmo hori baieztatzeko, azken 20 larrialdietako iriste-orduak jaso dira, eta haietatik 2 baino ez ziren izan aipatu ordu-tartean. %5eko esangura-mailaz, erabaki ezazu benetan ordu tarte horretan proportzionalki haur larrialdi gutxiago izaten den, hipotesi nulua zehatz planteatuz eta justifikatuz eta p-balioaren kalkulua eskuz nola egingo zenukeen (adierazpen aljebraikoa ematearekin aski da) nahiz Larson-en nomogramaren bitartez garatuz.

Aurretiko kalkulu batzuk egin ditzagun egoera argitzeko. Berez, eta egunean zehar larrialdiak zoriz gertatzen badira, 6etatik 12etara gertatzen diren larrialdiak %25 izan beharko liriateke, 0.25eko probabilitateaz alegia, tarte horretan 6 ordu daudela aintzat harturik,  $6/24$  eginez. Beraz, batez beste  $20 \times 6/24 = 5$  larrialdi suertatu beharko liriateke, gauzak berez edo zoriz izanez gero esan bezala. Ebidentzia da 2 larrialdi izan dela epe horretan, beraz, badirudi larrialdiak ordu-tarte horretan gutxiago direla.

Enuntziatuan ere tarte horretan larrialdi gutxiago sartzen delako susmoa aipatzen da. Beraz, zuhurtziaz, susmoaren nahiz ebidentziak erakusten duenaren *aurka* jokatzuz, hau izan behar da hipotesi nulua,  $p$  izanik larrialdi bat 6etatik 12etara izateko probabilitatea:

$$H_0 : p \geq 0.25$$

$X$  6etatik 12etara larrialdi kopuruaren erreferentziazko banaketa, hipotesi nulupean beti ere, hau da:

$$X \sim B(n = 20, p = 0.25)$$

p-balioa kalkulatzeko, probaren norabidea ezkeretik (azpitik) hartu behar da, izan ere, hipotesi nulua ( $p > 0.25$  baino handiagoa edo berdina) baztertu egingo da, larrialdi kopurua aipatu ordu-tartean *txikia* denean:

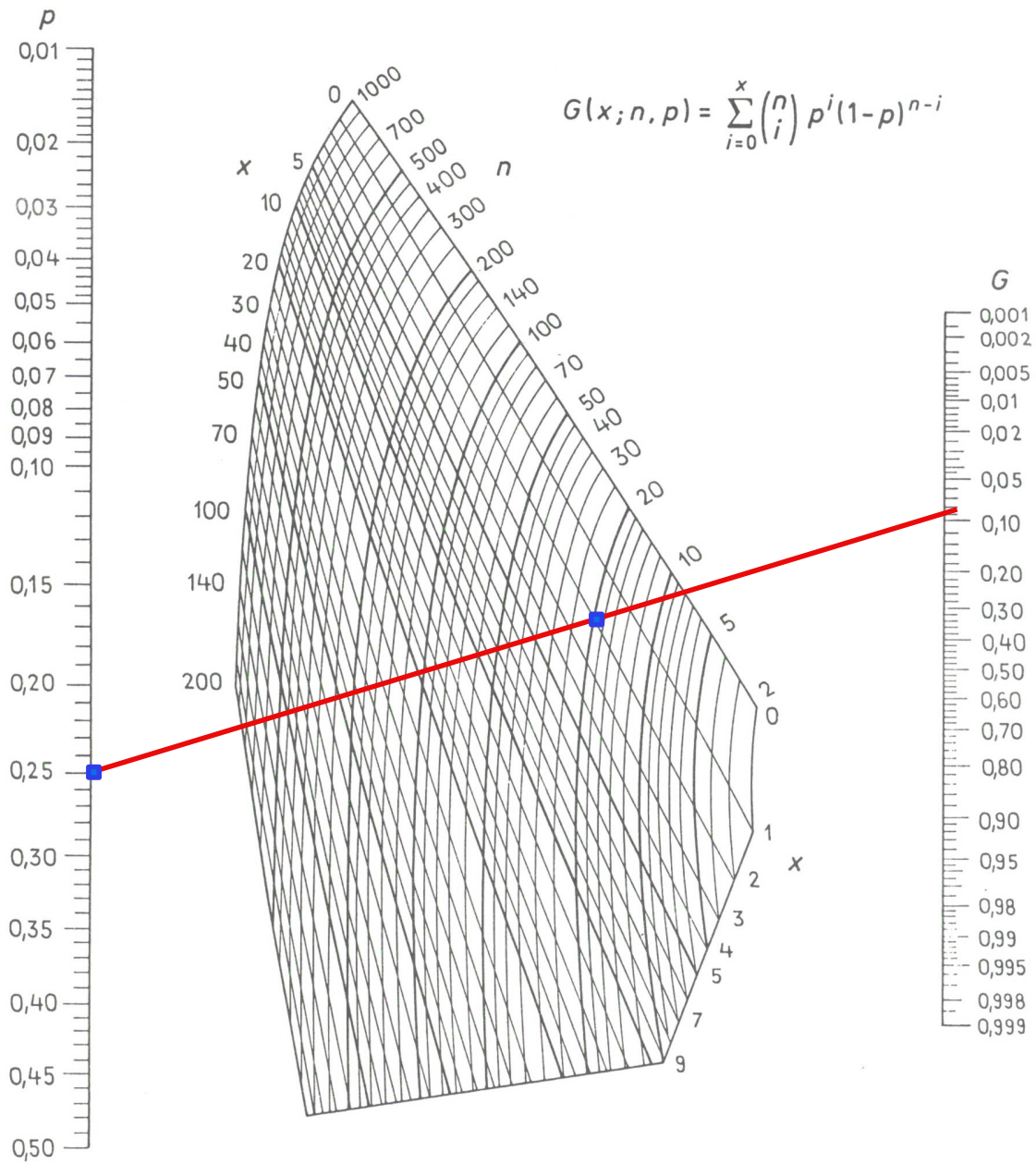
$$p = P[X \leq 2/B(n = 20, p = 0.25)] =$$

$$P[X = 0/B(n = 20, p = 0.25)] + P[X = 1/B(n = 20, p = 0.25)] + P[X = 2/B(n = 20, p = 0.25)] =$$

$$0.25^2 \times 0.75^{18} \times \frac{20!}{18!2!} + 0.25^1 \times 0.75^{19} \times \frac{20!}{19!1!} + 0.25^0 \times 0.75^{20} \times \frac{20!}{20!0!} =$$

$$0.25^2 \times 0.75^{18} \times \frac{20 \times 19}{2} + 0.25^1 \times 0.75^{19} \times 20 + 0.75^{20} = 0.09126043$$

p-balioaren balio zehatza kalkulatu da, beharrezkoa ez den arren: esangura-maila baino handiagoa denez, hipotesi nulua onartu eta beraz, goizeko txandan larrialdi gutxiago dagoela baieztatzeko ebidentzia aski sendorik ez dagoela erabaki behar da. Hurrengo orrialdean, p-balioaren hurbilketa egiten da, emaitza eta ondorio berdinekin.



## II. ebazkizuna (Ebaluazio ez jarraitua: 1.5 puntu)

Hileta-zerbitzuen inguruan ekonomia zirkularra sustatzeko helburuarekin HUMAN FATTY sortu berri den enpresa batek hildakoen gorpuak birziklatu nahi ditu, haietatik gantza erauziz eta haiekin xaboiak egiteko. Gorpu batetik ekoiz daitekeen xaboi kantitatea 2 kilo (kopuru finkoa) gehi gorpuaren pisuaren %10 dela zenbatetsi da.

(a) Hilotz batzuen pisuak jaso dira: 49-63-73-83-67-39-51-65-61. Paper normala erabiliz eta bertan puntuak zehaztasunez kokatuz, erabaki ezazu pisuari buruz ezarritako eredu normala egokia izan daitekeen, eta baiezkotan estimatu itzazu batezbestekoa eta desbideratze estandarra. Oharra: har ezazu ekarri duzun paper normaleko orri bat, bertan egin marrazketa eta beharrezko kalkuluak, eta azterketa entregatzearekin batera barruan beste orriekin sartu.

Datuak ordenatu egiten ditut: 39-49-51-61-63-65-67-73-83.

Haien maiztasun metatuak kalkulatzeko ditut, formula egokiarekin:

$x_i$	39	49	51	61	63	65	67	73	83
$i/(n+1)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

Puntu bakoitza (datua/maiztasun metatua) paper normalean irudikatzen dut (azterketaren bukaeran), eta hortik hilotz bakoitzaren  $P$  pisua honela banatzen dela estimatzen dut:

$$P \sim N(62, 16)$$

Desbideratze estandarra grafikoko erreferentziako 78 eta 46 puntuetatik  $(78-46)/2=16$  eginez zenbatetsi da. Ohartarazi behar da zuzena nola marrazten den, estimazioak ere desberdinak izango direla.

(b) Pisuari buruz estimatu dituzun batezbestekoa eta desbideratzea harturik, betiere pisua banaketa normalaren arabera dela suposatuz, nola banatzen da gorpu bakoitzetik atera daitekeen xaboi kantitatea, horretarako aldagai-aldaketa egokia eginda?

$X$  xaboi kantitatea honela erlazionatzen da  $P$  pisuarekin:  $X = 2 + 0.1P$ .

Aldaketa linealaren propietatea erabiliz honela banatzen da xaboi kantitatea hilotz bakoitzeko:

$$X \sim N(\mu = 2 + 0.1 \times 62 = 8.2, \sigma = 0.1 \times 16 = 1.6)$$

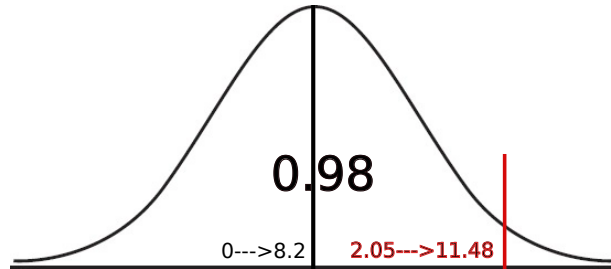


gizapedia.org

(c) Aurreko erantzunetik abiatuz, kalkula ezazu gorpu batetik %98ko probabilitateaz atera daitekeen xaboi kopuru maximoa, %95eko probabilitateaz atera daitekeen xaboi kopuru minimoa eta batezbestekoaren inguruko %90eko tarte simetriko bat, kasu bakoitzean probabilitatea eta aldagaiaren balioa azaltzen dituen banaketa normalaren diagrama lagungarri bat marraztuz.

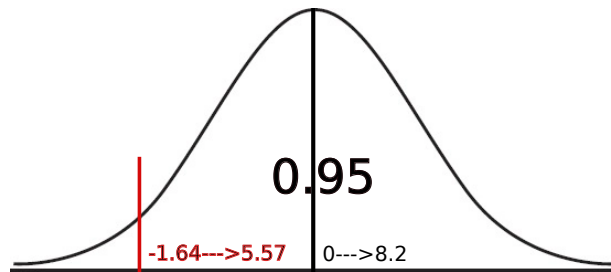
%98ko kopuru maximoa

$$P[X < x] = P\left[Z < \frac{x - 8.2}{1.6}\right] = 0.98 \rightarrow \frac{x - 8.2}{1.6} = 2.05 \rightarrow x = 2.05 \times 1.6 + 8.2 = 11.48kg$$



%95eko kopuru minimoa

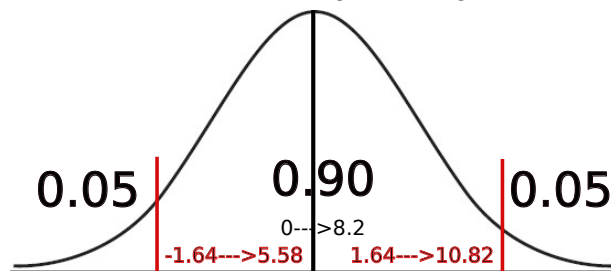
$$P[X > x] = P\left[Z > \frac{x - 8.2}{1.6}\right] = 0.95 \rightarrow \frac{x - 8.2}{1.6} = -1.64 \rightarrow x = -1.64 \times 1.6 + 8.2 = 5.57kg$$



%90eko tarte simetrikoa

$$P[\mu - k < X < \mu + k] = 0.9 \rightarrow P[X < \mu + k] = 0.95 \rightarrow P[X < \mu + k] = P\left[Z < \frac{\mu + k - 8.2}{1.6}\right] = P\left[Z < \frac{8.2 + k - 8.2}{1.6}\right] = P\left[Z > \frac{k}{1.6}\right] = 0.95 \rightarrow \frac{k}{1.6} = 1.64 \rightarrow k = 2.62$$

Beraz, hauxe izango %90eko tarte simetrikoa:  $8.2 \pm 2.62 : 5.58kg - 10.82kg$ .



(d) Zenbat gorpu izan behar dira eskuragarri, 100kg xaboi ekoitzi ahal izateko probabilitatea %975 izan dadin?

Gorpu kopurua  $n$  izendatuz,  $n$  gorpu horietatik ateratzen den xaboi-kopurua honela banatzen da, banaketaren normalaren ugalkortasunaren propietatea erabiliz:

$$\mathbf{X} \sim N(\mu = 8.2n, \sigma = \sqrt{1.6^2 n} = 1.6\sqrt{n})$$

Proposatutako galdera honela garatzen da:

$$P[\mathbf{X} > 100] = P\left[Z > \frac{100 - 8.2n}{1.6\sqrt{n}}\right] = 0.975 \rightarrow \frac{100 - 8.2n}{1.6\sqrt{n}} = 1.96$$

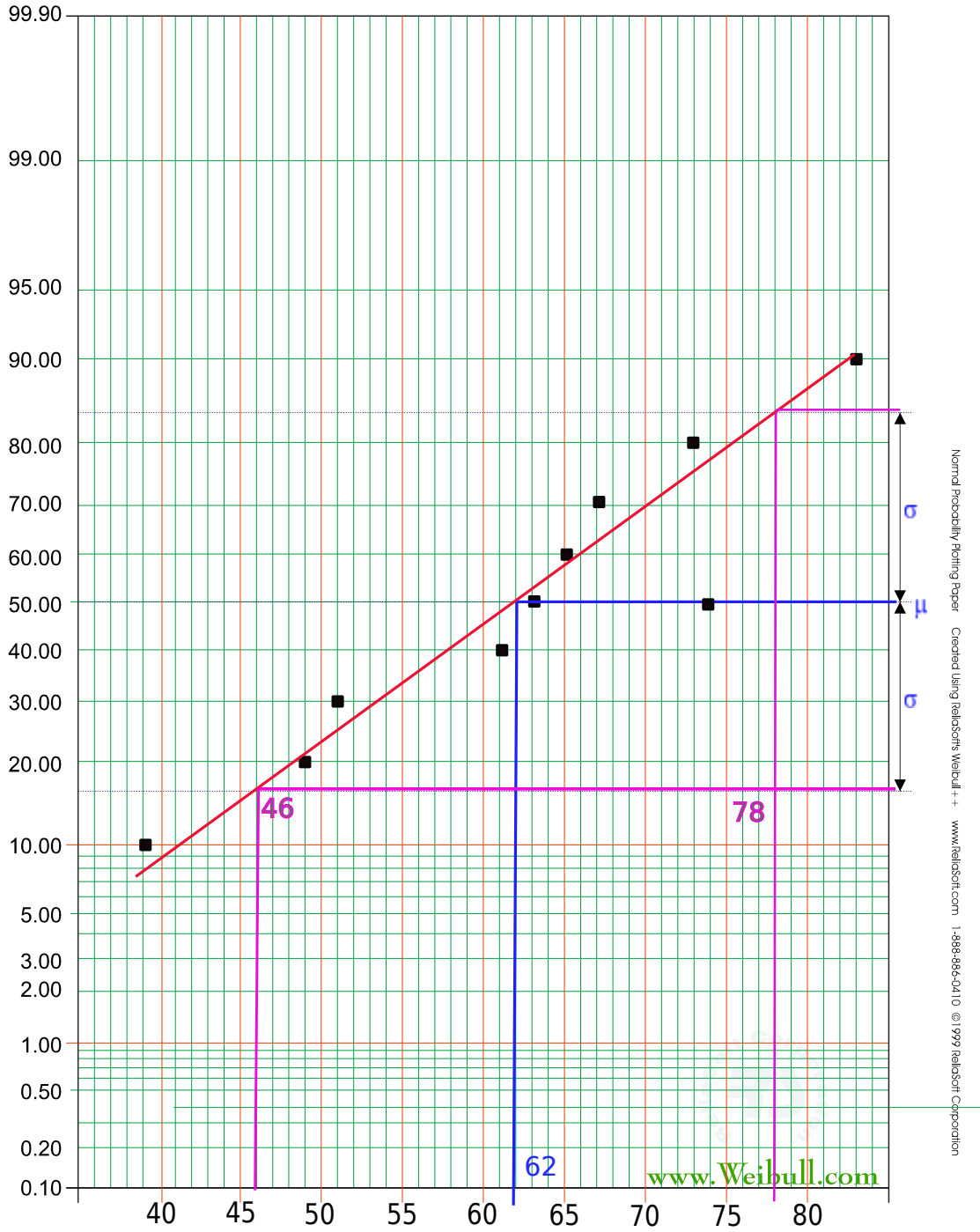
Hortik ekuazio koadratiko hau suertatzen da:

$$8.2n + 3.136\sqrt{n} - 100 = 0 \rightarrow 8.2x^2 + 3.136x - 100 = 0$$

$$\rightarrow x = 3.3062(\text{soluzio positiboa bakarrik hartuz}) \rightarrow n = 3.3062^2 = 10.93 \rightarrow 11 \text{ gorpu (ezkertasunez)}$$



[gizapedia.org](http://gizapedia.org)



### III. ebazkizuna (Ebaluazio ez jarraitua: 1.5 puntu)

DENAGEZURRA alderdi politiko berria Nafarroa eta EAeko zenbait herritan aurkeztu da. Honako hauek dira herri horietan izan dituen boto-portzentajeak (20 herri dira herri bakoitzeko, 40 datu guztira beraz, eta lurralde bakoitzean datuak ordenaturik ematen zaizkizu):

Nafarroa: 25 27 28 29 31 32 32 33 33 33 37 38 39 40 41 43 43 44 50 54

EAE: 9 13 15 16 16 17 17 17 18 21 22 22 22 23 24 24 25 25 26 27

Wilcoxonon hein-proba erabiliz, bi lurralde horietan alderdi horretako boto-emaeleek edo haien portzentajeek populazio homogeneo bat osatzen dutela esan al daiteke? Lagin-tamaina handietarako formula harturik, p balioaren nahiz balio kritikoaren metodoa erabil ezazu, %1eko esangura-mailarekin. Laguntza: lehenengo 16 zenbaki naturalen batura (1+2+...+16) 136 da.

Datu guztiak bateratu eta ordena ditzagun:

$x_i$	9	13	15	16	16	17	17	17	18	21
Non	EAE	EAE	EAE	EAE	EAE	EAE	EAE	EAE	EAE	EAE
Heina	1	2	3	4	5	6	7	8	9	10
$x_i$	22	22	22	23	24	24	25	25	25	26
Non	EAE	EAE	EAE	EAE	EAE	EAE	EAE	EAE	NA	EAE
Heina	11	12	13	14	15	16	18	18	18	20
$x_i$	27	27	—							
Non	EAE	NA	—							
Heina	21.5	21.5	—							

$$W_{EAE} = 136 + 18 + 18 + 20 + 21.5 = 213.5$$

$W_{EAE}$  honela banatzen da lagin-tamaina handietarako:

$$W_{EAE} \sim N\left(\mu = \frac{20 \times 41}{2} = 410, \sigma = \sqrt{\frac{20 \times 20 \times 41}{12}} = 36.96\right)$$

p-balioaren metodoaz

Suertatu den  $W$  estatistikoa 410-eko batezbestekotik behera dago; beraz, "arraroa" behetik dago:

$$p = P[W < 213.5] = P\left[Z < \frac{213.5 - 410}{36.96}\right] = P[Z < -5.31] \approx 0$$

p-balioa esangura-maila erdia baino txikaigoa denez (proba alde bikoia da, EAeko  $W$  estatistikoa handia nahiz txikia denean baztertzeko baitugu bi populazioen arteko berdintasuna), bi lurraldeen berdintasuna baztertu eta ondorioz Nafarroa lurralde desberdina dela (kasu honetan, DENAGEZURRA alderdiaren aldekoagoa) baieztatu behar da. Beraz, bi lurraldeek ez dute osatzen populazio homogeneo bat, eta hori baieztatu behar da gutxienez.



gizapedia.org



## balio kritikoak zehaztuz

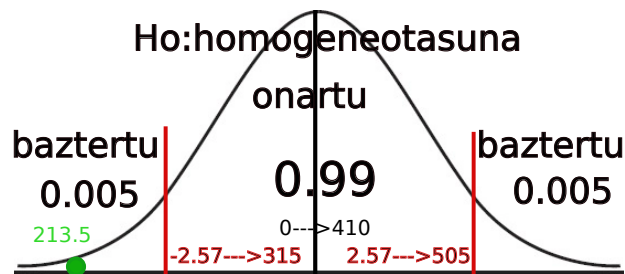
Balio kritikoak bi aldeetara daude, ezkerrera nahiz eskuinera, alde bakoitzean esangura-maila erdiko probabilitatearekin:

$$P[W < W_0] = P\left[Z < \frac{W_0 - 410}{36.96}\right] = 0.005 \rightarrow \frac{W_0 - 410}{36.96} = -2.57 \rightarrow W_0 = 315$$

Banaketa normala simetrikoa denez, eskuin aldetik balio kritikoa  $410 + (410 - 315) = 505$  izango da:

$$P[W < W_0] = P\left[Z < \frac{W_0 - 410}{36.96}\right] = 0.995 \rightarrow \frac{W_0 - 410}{36.96} = 2.57 \rightarrow W_0 = 505$$

Bi populazioen arteko berdintasuna baztertzen da, EAEko W estatistikoa 315 baino txikiagoa edo 505 baino handiagoa denean.  $W_{EAE} = 213.5$  izanik, lurraldeen arteko homogeneotasuna baztertu egiten dugu.



gizapedia.org

## ESTADISTIKA ENPRESARA APLIKATUA

Irakaslea: Josemari Sarasola (Bernoulli prozesuetatik inferentzia eta balidazio-probetara)

Data: 2023ko uztailaren 5ean, 10:00

Iraupena: 1h 20min (estimazioa)

Oharra: Axotaz zein arkatzez egin daiteke, baina beti argi eta txukun.

**I. ebazkizuna** (1.5 puntu)

Ukrainiako Armadak drone kamikazeak garatu ditu errusiar tankeak suntsitzeko. Egindako probetan, probabilitate-estimazioa da 10 dronetatik batek lortzen duela tankea suntsitzea.

**Egin beharreko atazak:**

- (a) Zenbat da tanke baten aurka 6 drone batera gidatu ondoren, tankea suntsitzeko probabilitatea? Egin kalkulua zehatz formula binomiala erabiliz, ahalik eta modu sinpleenean, nahiz Larson nomogramarekin lagunduta.
- (b) Zenbat drone gidatu behar dira batera tanke baten aurka tankea suntsitzeko probabilitatea 0.9 izan dadin? Kopuru desberdinekin proba eginez egin behar duzu, probabilitate zehatzak kalkulatu, ahalik eta modu sinpleenean, 12 droneko kopuru batetik abiatu.
- (c) Aurreko ataleko emaitza Larson nomograman erakutsi.

- (a) Tankearen suntsiketa eragiten duten droneak  $X$  izendatuz:  $X \sim B(n = 6, p = 1/10 = 0.1)$

$$P[\text{tankea suntsitu}] = P[X \geq 1] = 1 - P[X = 0] = 1 - 0.1^0 \times 0.9^6 \times \frac{6!}{0!6!} = 0.468559$$

Larson nomograman ezin dira "handiagoa edo berdin" probabilitateak adierazi eta beraz osagarriaren probabilitatea adierazten da,  $P[X = 0]$  probabilitatea alegia (marra gorria)

- (b)

18 dronerekin probatuz:

$$P[\text{tankea suntsitu}] = P[X \geq 1/B(n = 18, p = 0.1)] = 1 - P[X = 0/B(n = 18, p = 0.1)] = 1 - 0.1^0 \times 0.9^{18} \times \frac{18!}{0!18!} = 0.85$$

19 dronerekin probatuz:

$$P[\text{tankea suntsitu}] = P[X \geq 1/B(n = 19, p = 0.1)] = 1 - P[X = 0/B(n = 19, p = 0.1)] = 1 - 0.1^0 \times 0.9^{19} \times \frac{19!}{0!19!} = 0.8649$$

20 dronerekin probatuz:

$$P[\text{tankea suntsitu}] = P[X \geq 1/B(n = 20, p = 0.1)] = 1 - P[X = 0/B(n = 20, p = 0.1)] = 1 - 0.1^0 \times 0.9^{20} \times \frac{20!}{0!20!} = 0.8784$$

21 dronerekin probatuz:

$$P[\text{tankea suntsitu}] = P[X \geq 1/B(n = 21, p = 0.1)] = 1 - P[X = 0/B(n = 21, p = 0.1)] = 1 - 0.1^0 \times 0.9^{21} \times \frac{21!}{0!21!} = 0.8905$$

22 dronerekin probatuz:

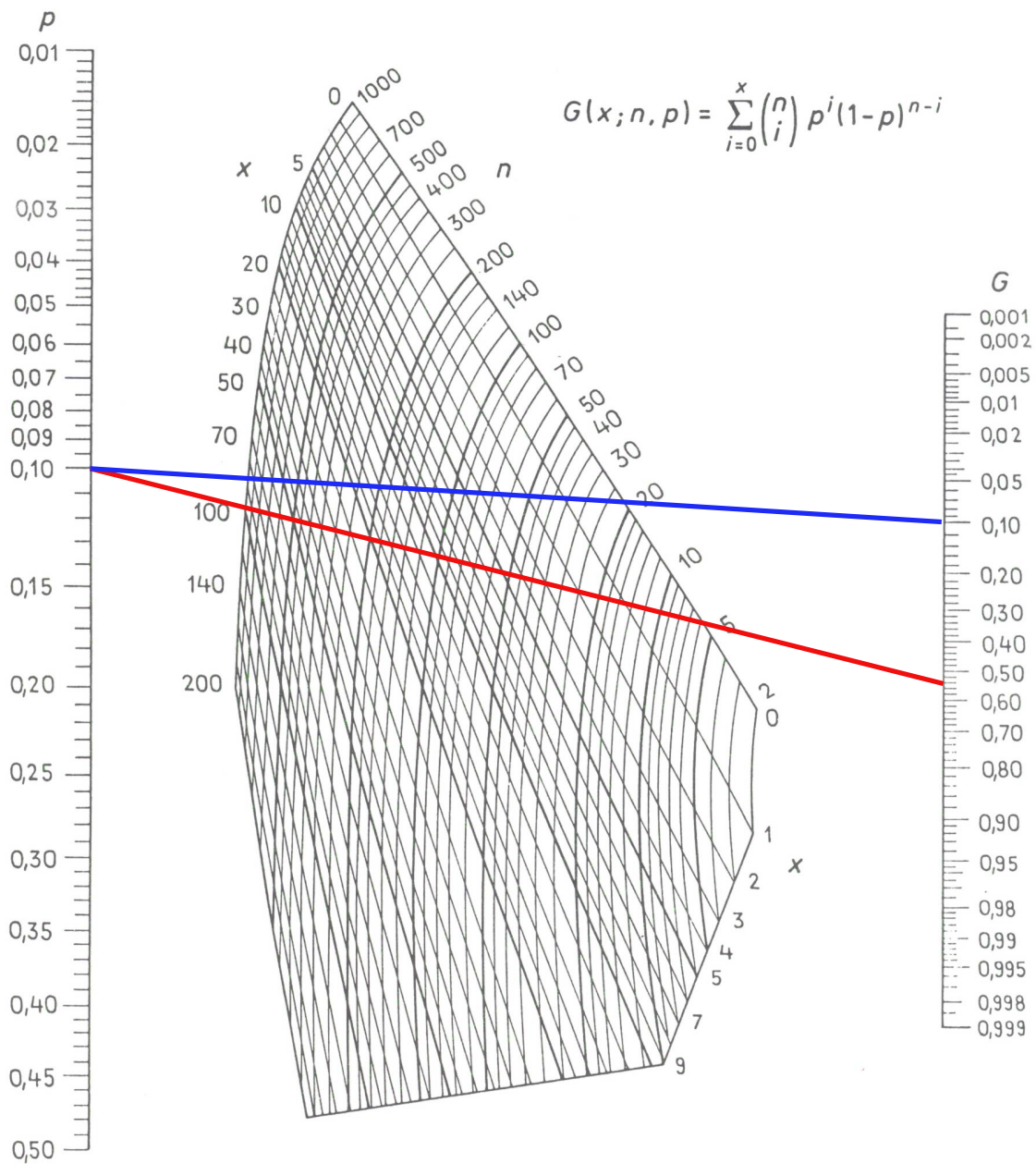
$$P[\text{tankea suntsitu}] = P[X \geq 1/B(n = 22, p = 0.1)] = 1 - P[X = 0/B(n = 22, p = 0.1)] = 1 - 0.1^0 \times 0.9^{22} \times \frac{22!}{0!22!} = 0.9015$$

Beraz, 22 drone jaurti behar dira tanke baten aurka, tankearen suntsiketa 0.9ko probabilitateaz ziurtatzeko.

Nomograman marra urdinaz adierazten da.



gizapedia.org



[gizapedia.org](http://gizapedia.org)

## II. ebazkizuna (3 puntu)

Ewing sarkoma hezurretan eta horien inguruetan garatzen den minbizi mota bat da, bereziki haurrek eta gazteek pairatzen dutena. 6-30 urte bitarteko pertsonen artean, intzidentzia-probabilitatea urteko 0.00002 da. Osakidetzan alta emanda dauden pertsonak adin horrekin 400.000 dira.

### Egin beharreko atazak:

- (a) Adierazi, kalkulatu gabe, banaketa binomialaren bitartez, 400.000 pertsona horien artean urtean 2 kasu izateko probabilitatea?
- (b) Adierazi eta kalkulatu 2 kasu edo gutxiago probabilitate berdina egoerari dagokion hurbilketa egokienaren bitartez.
- (c) Azken urtean 12 kasu jaso ziren 400.000 pertsona horien artean. Intzidentzia gehitu dela esan al daiteke? Hipotesi nulua argi eta garbi adierazi eta justifikatu. Erabil ezazu hurbilketa, aurreko atalean bezalaxe. Laguntza:  $P[X \geq 14] = 0.034$ . Esangura-maila: %10.
- (d) Gaixotasun horren aurkako txerto bat garatu da, eta 400.000 pertsona horiei proposatu zaie. Pertsona horietako batek edo haren guraso edo tutoreak txertoari uko egiteko probabilitatea 0.6 da. Zenbat da 160.500 dosi baino gehiago behar izateko probabilitatea?
- (e) Zenbat dosi prestatu behar dira eskatzen duten guztiei dosia eman ahal izateko probabilitatea 0.975 izan dadin?

- (a)  $X$  kasu kopurua banaketa binomial honen arabera da:  $B(n = 400.000, p = 0.00002)$ .

$$P[X = 2] = 0.00002^2 \times 0.9998^{399998} \frac{400000!}{399998!2!}$$

- (b) Banaketa binomialean  $n$  handia eta  $p$  txikia direnez, Poisson banaketa erabiltzen da hurbilketa gisa:

$$B(n = 400.000, p = 0.00002) \rightarrow P(\lambda = np = 8)$$

Hartara:

$$P[X \leq 2] = P[X = 0] + P[X = 1] + P[X = 2] = \frac{e^{-8}8^0}{0!} + \frac{e^{-8}8^1}{1!} + \frac{e^{-8}8^2}{2!} = 0.01375$$

- (c) Hipotesi nulua galderan planteatzen denaren aurkakoa da:  $H_0 : \lambda \leq 8$  (kasu kopurua ez da gehitu).

Hipotesi nulua baztertuko da kasu kopurua oso handia denean, beraz, probaren norabidea eskuin aldetikakoa da (bestela esanda,  $\lambda = 8$  izanik, 12 kasu asko dira):

$$p = P[X \geq 12] = P[X = 12] + P[X = 13] + P[X \geq 14] = \frac{e^{-8}8^{12}}{12!} + \frac{e^{-8}8^{13}}{13!} + 0.034 = 0.112 > \alpha$$

Beraz, hipotesi nulua onartu eta kasu kopurua gehitu dela baieztatzeko arrazoi askirik ez dagoela erabaki behar da.

- (d)  $X$  eskatzen den txerto kopurua izanik,  $n$  handia dela eta  $p$  txikia ez dela kontuan harturik, De Moivre-Laplace teorema erabiliz (kontuan hartu uko egiteko probabilitatea 0.6 dela, eta hartara onartzeko probabilitatea 0.4):

$$B(n = 400.000, p = 0.4) \rightarrow N(\mu = np = 160.000, \sigma = \sqrt{400.000 \times 0.4 \times 0.6} = 309.8)$$

Beraz:

$$P[X > 160.500] = P\left[Z > \frac{160.500 - 160.000}{309.8}\right] = P[Z > 1.61] = 1 - P[Z < 1.61] = 0.0537$$

- (e)  $t$  behar den txerto kopurua izanik,  $t$  kopuru nahikoa izango da, 0.975eko probabilitateaz, eskatzen den txerto kopurua  $t$  edo txikiagoa denean:

$$P[X \leq t] = P\left[Z < \frac{t - 160.000}{309.8}\right] = 0.975 \rightarrow \frac{t - 160.000}{309.8} = 1.96 \rightarrow t = 160.607, 2 \rightarrow 160.608 \text{ txerto}$$

### III. ebazkizuna (Ebaluazio ez jarraitua: 1.5 puntu)

Donostiako KITZIKA sexologia institutuak eskuratu duen nazioarteko txosten baten arabera, koito edo larru joaldiaren iraupena uniformeki banatzen da 2min-22min tartean. Txostenak baieztatutakoa euskaldunen artean ere betetzen den ikertzeko, 200 bikote heterosexual euskalduneri euren azken larru joaldiaren denbora galdetu zaie eta hauek dira jasotako datuak:

Tartea	Bikoteak
2min-5min	26
5min-10min	74
10min-15min	58
15min-22min	42
	200

Khi-karratu proba erabiliz, erabaki ezazu euskaldunen arteko larru joaldiak nazioartean bezala banatzen diren, aipatutako banaketa uniformearen arabera alegia. Esangura-maila: %1.

Adierazburuan maiztasun empirikoak ( $O$  : *observed*) ematen dira.  $U(2, 22)$  banaketa uniformeari dagokion maiztasun teorikoak aise kalkulatu dira banaketa uniformean probabilitateak tartearen zabalerarekiko proportzionalak direnez:

$$P[2 < X < 5] = \frac{3}{20} = 0.15 ; P[5 < X < 10] = \frac{5}{20} = 0.25 ; ; P[10 < X < 15] = \frac{5}{20} = 0.25 ; ; P[15 < X < 22] = \frac{7}{20} = 0.35$$

Maiztasun teorikoak probabilitate horiek lagin-tamainarekin bidertuz kalkulatu dira, bider eginez alegia.

Tartea	$O_i$	$E_i$	$(O_i - E_i)^2/E_i$
2min-5min	26	30	0.53
5min-10min	74	50	11.52
10min-15min	58	50	1.28
15min-22min	42	70	11.2
	200		24.53

Enpirikoen eta teorikoen arteko diskrepantzia estatistikoa da khi-karratu:

$$\mathbf{X}^2 = \sum (O_i - E_i)^2/E_i = 24.53$$

Eta esanguratsua den jakiteko,  $\chi_{4-1}^2$  banaketan  $X_{0.01}^2$  balio kritikoa bilatzen dut:

$$X_{0.01}^2 = 6.25$$

Estatistikoak gaitzen du balio kritikoa; beraz, hipotesi nulua, erdua egokia dela alegia, baztertu egiten da, eta ondorioz, euskaldunen artean koitoak desberdinak direla esan behar da, ez laburragoak edo luzeagoak, baizik eta besterik gabe desberdinak.



gizapedia.org