Universidad del País Vasco Euskal Herriko Unibertsitatea

EUP-UEP-D

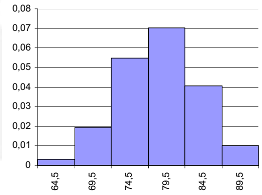
4.Gaia: Probabilitate-banaketa jarraituak

Cristina Alcalde - Arantxa Zatarain

Donostiako Unibertsitate Eskola Politeknikoa - UPV/EHU

Probabilitate-banaketa jarraituak

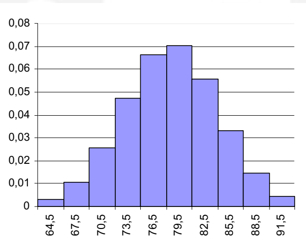
Probabilitate-banaketa, jarraitua dela esango dugu zorizko aldagaia jarraitua denean. Zorizko aldagai jarraituak $[\alpha, \beta]$ tartean balioak hartzen duenean, tartea azpitartetan zatitzen dugu maiztasunak lortu, eta maiztasun erlatiboen histograma egiten dugu



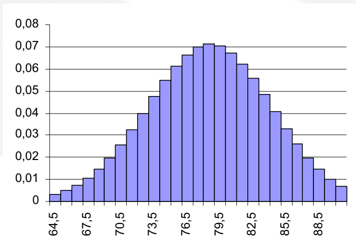
tartearen dentsitatea “maiztasun erlatiboa/luzera” eran definitzen dugu.

Probabilitate-banaketa jarraituak

Tarteen kopurua handitzen doanean, dentsitate histograma leuntzen doa

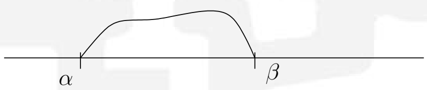


Probabilitate-banaketa jarraituak



Probabilitate-banaketa jarraituak

Probabilitatearen banaketa $[\alpha, \beta]$ tartean jarraiki, (uniformeki) banatzen bada masa unitarioa bezala kontsidera daiteke.



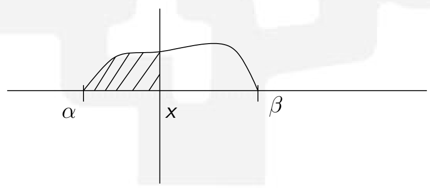
Probabilitate-banaketa jarraituak

Definizioa

Honako funtzioari banaketa-funtzioa deitzen zaio:

$$F(x) = P(X \leq x)$$

$[\alpha, x]$ *tartean dagoen probabilitatearen masa bezala uler daiteke.*

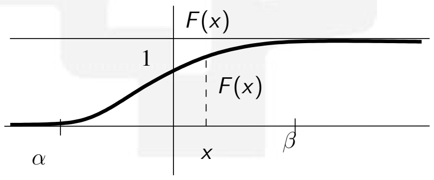


Hedapena

X aldagaia $[\alpha, \beta]$ tartean aldatzen bada, $F(x)$ funtzioa $\mathbb{R}$ zuzen erreal osora heda daiteke era honetan

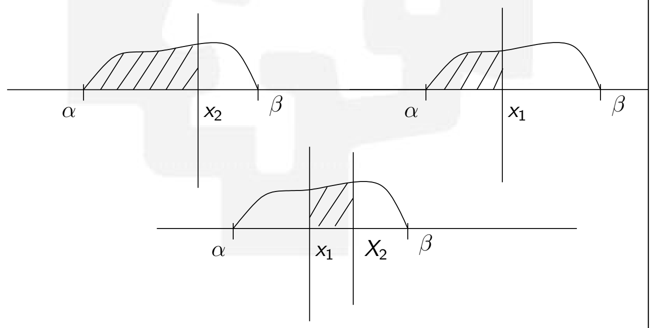
$$F(x) = \left\{ \begin{array}{ll} 0 & x \leq \alpha \text{ denean} \\ F(x) = P(X \leq x) & x \in (\alpha, \beta) \\ 1 & x \geq \beta \text{ denean} \end{array} \right\}$$

eta bere grafikoa irudian ikusten dena da.



Propietatea

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$$



Propietatea

$$P(X > x_2) = 1 - P(X \leq x_2) = 1 - F(x_2)$$

Dentsitate-funtsioa

Definizioa

$F(x)$ banaketa-funtzioa baldin bada, $f(x)$ era honetan definitutako funtzioari dentsitate-funtzioa deitzen diogu

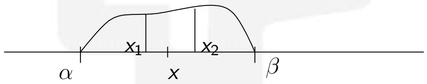
$$f(x) = F'(x) = \frac{dF(x)}{dx}$$

Definizioaren zergatia.

Aldagaia jarraitua denean $X = x$ izateko probabilitatea nulua da

$$P(X = x) = 0$$

beraz, logikoa dirudi x puntuan ingurune infinitesimalaren probabilitatearen masa neurtzea, hau da, banaketa-funtzioaren $dF(x)$ elementu infinitesimala. Ikus dezagun nola egiten dugun.



Definizioaren zergatia.

Izan bitez x_1 eta x_2 non $x_1 < x < x_2$ den. Orduan $F(x_2) - F(x_1)$ $[x_1, x_2]$ tartearen probabilitatearen masa da. Masa hau tartearen luzeraz zatituz tartearen batezbesteko dentsitatea lortzen da.

$$\text{batezbesteko dentsitatea} = \frac{F(x_2) - F(x_1)}{x_2 - x_1}$$

$x_2 \longrightarrow x_1$ doanean limitea hartuz $f(x)$ "aldiuneko dentsitatea" lortzen dugu.

$$f(x) = \lim_{x_2 \to x_1} \frac{F(x_2) - F(x_1)}{x_2 - x_1}$$

baina limite hori ,$f(x)$, $F(x)$ funtzioaren deribatua x puntuan da. Beraz

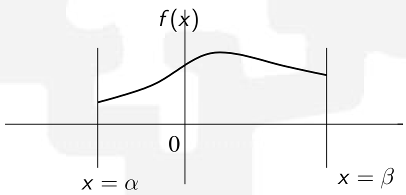
$$f(x) = F'(x) = \frac{dF}{dx}$$

Propietateak

1. $dF(x) = f(x)dx$
2. $F(x) = P(\alpha \le X \le x) = \int_\alpha^x f(t)dt$
3. $P(x_1 \le X \le x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$
4. $F(x) = \int_{-\infty}^x f(t)dt$
5. $\int_\alpha^\beta f(t)dt = \int_{-\infty}^\infty f(t)dt = 1$

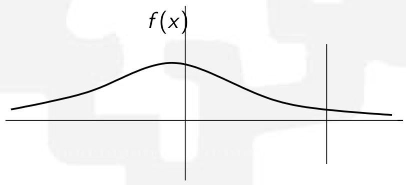
Adierazpen grafikoak

X aldagaiaren balioen tartea $[\alpha, \beta]$ bada, $f(x)$ dentsitate-funtzioaren adierazpen grafikoa irudian ikusten den adierazpenaren itxurakoa izango da,



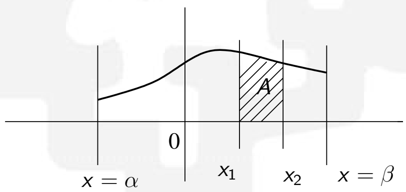
Adierazpen grafikoak

eta X aldagaiaren balioen eremua $(-\infty, \infty)$ bada irudian dugunarena.



Probabilitatea

$f(x)$ kurbak eta OX ardatzak tarte horretan ($[\alpha, \beta]$ edo $(-\infty.\infty)$ mugatzen duten azalera 1 da.
Beste aldetik $P(x_1 \le x \le x_2)$ probabilitatea irudian adierazten den eremuaren azalera da.



A azalera $= P(x_1 \le x \le x_2)$

Batezbestekoa

Definizioa
Izan bedi X, $[\alpha, \beta]$ tartean aldatzen den zorizko aldagai jarraitua. Orduan X aldagaiaren batezbestekoa edo itxaropen matematikoa honako hau da.

$$\mu_x = E[x] = \int_\alpha^\beta x \cdot f(x)dx$$

($\alpha = -\infty$ eta $\beta = \infty$ izan daitezke)

Bariantza

Definizioa
X, $[\alpha, \beta]$ tartean aldatzen den zorizko aldagai jarraitua bada, X aldagaiaren bariantza honako hau da.

$$\sigma_x^2 = Bar(x) = \int_\alpha^\beta (x - \mu_x)^2 \cdot f(x)dx$$

Teorema

$$\sigma_x^2 = Bar(x) = \int_\alpha^\beta x^2 \cdot f(x)dx - \mu_x^2$$

Gaussen banaketa normala

Definizioa

X, $[-\infty, \infty]$ tartean aldatzen den zorizko aldagai jarraitua Gaussen banaketa normalari jarraitzen duela esango dugu bere dentsitate funtzioa honako hau denean:

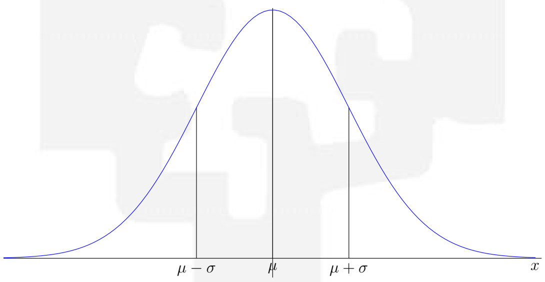
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in (-\infty, \infty)$$

Propietateak

1. Definitu ditugun funtzioak dentsitate-funtzioak betetzen duten propietatea betetzen du.
$$\int_{-\infty}^{\infty} f(x)dx = 1$$
2. Definizioaren agertzen diren μ eta σ parametroak X aldagaiaren batezbestekoa eta desbidazio tipikoa dira
3. $f(x)$ funtzioaren adierazpen grafikoa lor daiteke honako puntuak kontutan hartuz:
   - 3.1 $f(x)$ ez da negatiboa
   - 3.2 OX ardatza, $f(x)$ funtzioaren asintota horizontala da.
   - 3.3 $x = \mu$ funtzioaren simetria ardatza da.
   - 3.4 $(\mu, \frac{1}{\sigma\sqrt{2\pi}})$ puntua f-ren maximoa da.
   - 3.5 $x = \mu \pm \sigma$ puntutan f-ren inflexio-puntuak daude.

Gaussen dentsitate-funtzioa

Gaussen dentsitate-funtzioaren adierazpen grafikoa:



Gaussen dentsitate-funtzioa

Oharra

- *X aldagaia μ batezbestekoa eta σ^2 bariantza duen banaketa normalari jarraitzen bazaio X, $N(\mu, \sigma^2)$ dela esango dugu.*
- *$\mu = 0$ eta $\sigma = 1$ badira, banaketa normal tipikoa deituko diogu eta adierazteko Z hizkia erabiliko dugu. Z $N(0, 1)$ da.*

Banaketa-funtzioa

Definizioa

X zorizko aldagaia $N(\mu, \sigma^2)$ bada, bere banaketa-funtzioa honako hau da:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(X-\mu)^2}{2\sigma^2}} dX, \qquad -\infty < x < \infty$$

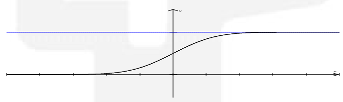
Banaketa normal tipikoa

Definizioa

Z $N(0, 1)$ aldagaiaren banaketa funtzioa

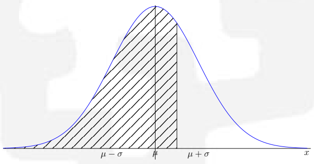
$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{z^2}{2}} dZ, -\infty < z < \infty$$

Eta adierazpen grafikoa hau da:



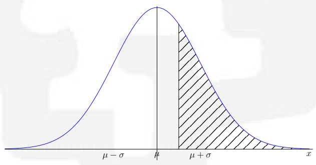
Propietatea

$F(z)$ banaketa funtzioak irudian adierazitako eremuaren azalera adierazten du.



Oharra

Probabilitate kalkuluan, askotan, erosoagoa da irudian adierazten den eremuaren azalera α kalkulatzea

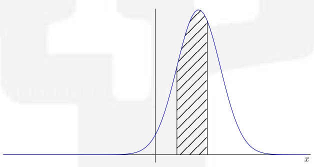


Helburua

X zorizko aldagai jarraitua $N(\mu, \sigma^2)$ bada, Honako probabilitate hau kalkulatzea nahi dugu

$$p(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$$

hau da, irudian azaltzen den R eremuaren azalera.



Tipifikazioa

Kalkulua errezteko aldagai-aldaketa hau egiten dugu:

$$z = \frac{x - \mu}{\sigma}$$

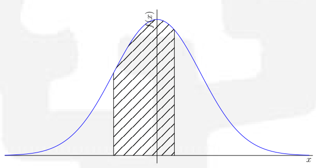
aldagai-aldaketa honi aldagaiaren tipifikazioa deitzen zaio. Aldaketa honen bidez X $N(\mu, \sigma^2)$ aldagaiaren $P(x_1 \leq X \leq x_2)$ probabilitatearen kalkulua, $P(z_1 \leq Z \leq z_2)$ probabilitatearen kalkulua bihurtzen da, baina Z aldagaia $N(0, 1)$ da eta probabilitateen balioak taulan daude.

Tipifikazioa

$$\begin{aligned}
P(x_1 \leq X \leq x_2) &= F(x_2) - F(x_1) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} \mathrm{e}^{-\frac{(X-\mu)^2}{2\sigma^2}} dX = \\
&= \left\{ \begin{array}{lcl} Z = \frac{X-\mu}{\sigma} & \rightarrow & dX = \sigma dZ \\ z_1 = \frac{x_1-\mu}{\sigma} & \text{eta} & z_2 = \frac{x_2-\mu}{\sigma} \end{array} \right. = \\
&= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} \mathrm{e}^{-\frac{z^2}{2}} dZ = \\
&= P(z_1 \leq Z \leq z_2) \qquad \text{non } ZN(0,1) \text{ den.}
\end{aligned}$$

Tipifikazioa

Grafikoki, irudian azaltzen den R' eremuaren azalera kalkulatzen dugu

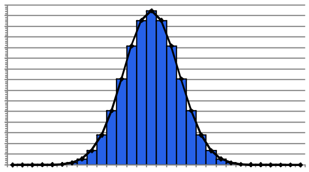


Banaketa binomialaren hurbilketa

Moivre-ren teorema

X zorizko aldagai diskretua binomiala bada eta $n \to \infty$, orduan Z aldagaia

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \qquad N(0,1) \text{ da.}$$



Hurbilketa hona da $n \geq 30$ eta $n \cdot p \geq 5$ eta $n \cdot q \geq 5$ direnean.