

APRAKSTOŠĀ STATISTIKA

Vidējais aritmētiskais:

$$AVERAGE(Arr) = \bar{x} = \frac{\sum i = 1nx_i}{n}$$

Mediānas pozīcija:

$$MEDIAN(Arr) = \frac{n+1}{2}$$

Ģeometriskais vidējais:

$$\bar{x} = \sqrt[n]{x_1 * x_2 * \dots * x_n}$$

$$\bar{x} = GEOMEAN(number1, [number2], ...)$$

Amplitūda (rangs):

$$range(X) = X_{max} - X_{min}$$

Dispersija (variance) izlasei:

$$s^2 = VAR.S(X)$$

$$s^2 = \frac{\sum_{i=0}^n (x_i - \bar{x})^2}{n-1}$$

Dispersija (variance) ģenerālkopai:

$$\sigma^2 = VAR.P(X)$$

$$\sigma^2 = \frac{\sum_{i=0}^n (x_i - \bar{x})^2}{N}$$

Standartnovirze (standartklūda) izlasei:

$$s = STDEV.S(X)$$

$$s = \sqrt{s^2}$$

Standartnovirze (standartklūda) ģenerālkopai:

$$s = STDEV.P(X)$$

$$\sigma = \sqrt{\sigma^2}$$

Standartizētā vērtība (Z-score):

$$Z = \frac{x - \bar{x}}{s}$$

$$Z = \frac{x - \bar{x}}{\sigma}$$

Asimetrijas koeficients:

| Nobīde pa labi | Simetrija | Nobīde pa kreisi |
|---|---|---|
| Mediāna - X_{min} > X_{max} - Mediāna | Mediāna - X_{min} ≈ X_{max} - Mediāna | Mediāna - X_{min} < X_{max} - Mediāna |
| Q_1 - X_{min} > X_{max} - Q_3 | Q_1 - X_{min} ≈ X_{max} - Q_3 | Q_1 - X_{min} < X_{max} - Q_3 |
| Mediāna - Q_1 > Q_3 - Mediāna | Mediāna - Q_1 ≈ Q_3 - Mediāna | Mediāna - Q_1 < Q_3 - Mediāna |

Kvartiles:

If we have an ordered dataset x_1, x_2, \dots, x_n , we can interpolate between data points to find the p th empirical quantile if x_i is in the $i/(n+1)$ quantile. If we denote the integer part of a number a by $\lfloor a \rfloor$, then the empirical quantile function is given by,

$$q(p/4) = x_k + \alpha(x_{k+1} - x_k) = QUARTILE.EXC(Arr, p)$$

$$k = \lfloor p(n+1)/4 \rfloor$$

$$\alpha = p(n+1)/4 - \lfloor p(n+1)/4 \rfloor$$

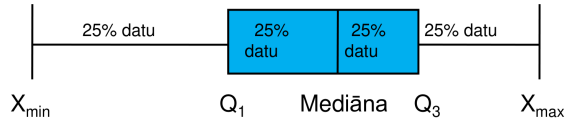
To find the first, second, and third quartiles of the dataset we would evaluate $q(0.25)$, $q(0.5)$, and $q(0.75)$ respectively.

Starpkvartīļu rangs (IQR):

$$IQR = Q_3 - Q_1$$

Kastes diagramma (Box plot)

Example:



Kovariācija

$$cov_s(X, Y) = \frac{\sum_{i=0}^n (X_i - \bar{X})(Y_i - \bar{Y})'}{n - 1}$$

$$cov_p(X, Y) = \frac{\sum_{i=0}^n (X_i - \bar{X})(Y_i - \bar{Y})'}{n}$$

$$cov_s(X, Y) = COVARIANCE.S(array1, array2)$$

$$cov_p(X, Y) = COVARIANCE.P(array1, array2)$$

Kovariācijas koeficients (r; Pearson's)

$$cov_s(X, Y) = \frac{\sum_{i=0}^n (X_i - \bar{X})(Y_i - \bar{Y})'}{n - 1}$$

$$r = \frac{COV(x, y)}{\sigma_x \sigma_y} = CORREL(array1, array2)$$

Notikumi

Kombinācijas (secība nav svarīga):

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} = COMBIN(n, k)$$

Kombinācijas ar atkārojumiem (secība nav svarīga):

$$\overline{C}_n^k = COMBINA(n, k) = TODO$$

Variācijas (secība ir svarīga):

$$A_n^k = \frac{n!}{(n-k)!} = PERMUT(n, k)$$

Pretējais notikums:

$$P(\neg A) = 1 - P(A)$$

Neatkarīgo notikumu šķelums:

$$P(A \wedge B) = P(A) * P(B)$$

Neatkarīgo notikumu disjunktija:

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

Nesavienojamo notikumu disjunktija:

$$P(A \vee B) = P(A) + P(B)$$

Nosacītā varbūtība:

$P(B|A)$ - B, ja zināms, ka A notika.

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}$$

Nosacītā varbūtība (neatkarīgiem notikumiem):

$P(B|A)$ - B, ja zināms, ka A notika.

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

Atkarīgo notikumu reizināšana:

$P(B|A)$ - B, ja zināms, ka A notika.

$$P(A) = P(A)P(B|A)$$

Pilnā varbūtība

$$P(A) = \sum_n P(A | B_n)P(B_n)$$

Beiesa teorēma

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

Kur A un B ir notikumi un $P(B) \neq 0$.

GADĪJUMLIELUMI

————— Sagaidāmā vērtība: —————

$$E[X] = \mu = \sum_{i=1}^N x_i P(X = x_i)$$

————— Dispersija: —————

$$\sigma^2 = \sum_{i=1}^N [x_i - E[X]]^2 P(X = x_i) = E[x^2] - (E[X])^2$$

————— Standartnovirze: —————

$$\sigma = \sqrt{\sum_{i=1}^N [x_i - E[X]]^2 P(X = x_i)} = \sqrt{E[x^2] - (E[X])^2}$$

————— Kovariācija: —————

$$= \sum_{i=1}^N [x_i - E[X]][y_i - E[Y]]P(X = x_i, Y = y_i)$$

————— Divu diskrēto gadījumu lielumu summa —————

$$E(X + Y) = E(X) + E(Y)$$

————— Divu diskrēto gadījumu lielumu dispersija —————

$$Var(X + Y) = \sigma_{x+y}^2 = \sigma x + \sigma y + 2\sigma_{xy}$$

— Divu diskrēto gadījumu lielumu standartnovirze —

$$\sigma_{x+y} = \sqrt{\sigma_{x+y}^2}$$

▪ Portfeļa ienesīgums (vidējais svērtais ienesīgums) ▪

$$E(P) = wE(X) + (1 - w)E(Y)$$

Kur w = aktīva X īpatsvars portfelī $(1 - w)$ = aktīva Y īpatsvars portfelī.

————— Portfeļa risks (svērtā izkļiede) —————

$$\sigma_P = \sqrt{w^2\sigma^2 + (1 - w)^2\sigma_Y^2 + 2w(1 - w)\sigma_{XY}}$$

Kur w = aktīva X īpatsvars portfelī $(1 - w)$ = aktīva Y īpatsvars portfelī.

————— Binomiālais sadalījums —————

$$P(X = x|n, p) = \frac{n!}{x!(n - x)!} (1 - p)^{1-x} p^x$$

Kur p - notikuma varbūtība, X - diskrēts gadījumu lielums, n - notikumu (izmēģinājumu) skaits.

————— Binomiālais sadalījuma raksturlielumi —————

$$\begin{aligned}\mu &= E(X) = np \\ \sigma^2 &= np(1 - p) \\ \sigma &= \sqrt{np(1 - p)}\end{aligned}$$

————— Puasona sadalījums (binom. aproksimācija) —————

$$\begin{aligned}P(X = x|\lambda) &= \frac{e^{-\lambda} \lambda^x}{x!} \\ P(X = x|\lambda) &= POISSON.DIST(X, \mu, CUMM.)\end{aligned}$$

Kur x = notikuma iestāšanās skaits, λ = sagaidāmais notikumu skaits (Puasona sadalījuma vidējā vērtība), $\lambda = np$, e = naturālā logaritma bāze (2.71828... vai $EXP(1)$)

————— Puasona sadalījuma raksturlielumi —————

$$\begin{aligned}\mu &= \lambda \\ \sigma^2 &= \lambda \\ \sigma &= \sqrt{\lambda}\end{aligned}$$

— Ģeometriskais sadalījums (Hiperģeometriskais) —

$$P(X = x|n, N, E) = \frac{\left(\frac{E}{x}\right) \left(\frac{N-E}{n-x}\right)}{\left(\frac{N}{n}\right)}$$

Kur N = ģenerālkopas lielums, E = interesējošo vienumu skaits ģenerālkopā, n = izlases lielums, x = interesējošo vienumu skaits izlasē.

$$E[x] = \frac{nE}{N}$$

$$\sigma = \sqrt{\frac{nE(N-E)}{N^2}} * \sqrt{\frac{N-n}{N-1}}$$

$$f(X) = EXPON.DIST(x, lambda, 0)$$

$$f(X) = \lambda e^{-\lambda X}, X > 0$$

$$F(X) = EXPON.DIST(x, lambda, 1)$$

$$F(X) = 1 - e^{-\lambda X}$$

Kur λ - vidējais notikuma iestāšanās skaits laika vienībā.

$$x = NORM.DIST(X, \mu, \sigma, CUMM)$$

$$Z = NORM.S.INV(x)$$

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

$$\mu \pm 1\sigma = 68,26\%$$

$$\mu \pm 2\sigma = 95,44\%$$

$$\mu \pm 3\sigma = 99,73\%$$