STATISTICĂ DESCRIPTIVĂ

Mediana:

n=impar	n=par
$Me = X_{\frac{n+1}{2}}$	$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$

Media aritmetică:

Populație	Eșantion
$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$	$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$

Asimetrie	Boltirea
$M_3 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{n}$	$\alpha_4 = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^4}{S^4} - 3$

Amplitudinea: $R = X_{max} - X_{min}$

Valoarea centrală (VC): VC = $(X_{min}+X_{max})/2$

Media deviației:

De la medie:	De la mediană:
$R_{\overline{X}} = \frac{\sum_{i=1}^{n} \lvert X_i - \overline{X} \rvert}{n}$	$R_{Me} = \frac{\sum_{i=1}^{n} \lvert X_i - Me \rvert}{n}$

Variația:

Populației	Eșantionului
$\sigma^2 = \frac{SS}{N} = \frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N}$	$s^2 = \frac{SS}{n-1} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$

Deviația standard:

$$s = \sqrt{s^2} = \sqrt{\frac{SS}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$$

Eroarea standard: $ES = \frac{s}{\sqrt{n}}$

Coeficientul de variație: $CV = \frac{s}{\overline{X}}$

PROBABILITĂȚI

$$Pr(A) = \frac{\text{nr. cazuri favorabile}}{\text{nr. cazuri posibile}}$$

SENSIBILITATEA: Se = Pr(B|A)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
SPECIFICITATEA: Sp = Pr(nonB|nonA)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
VALOAREA PREDICTIVĂ POZITIVĂ: VPP = Pr(A|B)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
VALOAREA PREDICTIVĂ NEGATIVĂ: VPN = Pr(nonA|nonB)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
Rata falșilor pozitivi: RFP = Pr(B|nonA)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
Rata falșilor negativi: RFN = Pr(nonA|B)
unde Pr(B) = probabilitatea unui test pozitiv și Pr(A) = probabilitatea existenței unei patologii de interes.
Adunare: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
Adunare în cazul evenimentelor mutual exclusive:
$Pr(A \cup B) = Pr(A) + Pr(B)$
Înmulțire: $Pr(A \cap B) = Pr(A) \cdot Pr(B|A)$
Înmulțire evenimente independente:
$Pr(A \cap B) = Pr(A) \cdot Pr(B)$
Formula lui BAYES:

$$Pr(A \mid B) = \frac{Pr(B \mid A) \cdot Pr(A)}{Pr(B \mid A) \cdot Pr(A) + Pr(B \mid \overline{A}) \cdot Pr(\overline{A})}$$

VARIABILE ALEATOARE

Media = valoare expectată = speranța matematică:

$$M(X) = \sum_{i=1}^{n} X_i \cdot Pr(X_i)$$

Variația: $V(X) = \sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)$

Abaterea standard: $\sigma(X) = \sqrt{V(X)} = \sqrt{\sum_{i=1}^{n} (X_i - M(X))^2 \cdot Pr(X_i)}$

VARIABILA ALEATOARE BINOMIALĂ

Distribuția binomială: $Pr(X = k) = C_n^k p^k q^{n-k}$

unde $C_n^k = \frac{n!}{k! \cdot (n-k)!}$

Media sau speranța matematică: $M(X) = n \cdot p$
Variația: $V(X) = n \cdot p \cdot q$
Abaterea standard: $\sigma(X) = \sqrt{V(X)} = \sqrt{n \cdot p \cdot q}$

VARIABILA ALEATOARE POISSON

$$X : \begin{pmatrix} k \\ e^{-\theta} \cdot \frac{\theta^k}{k!} \end{pmatrix} \quad Pr(X = k) = \frac{e^{-\theta} \cdot \theta^k}{k!}$$

e = 2.718281828; $\theta = n \cdot p$ (n = volumul eșantionului, p = probabilitatea de apariție a evenimentului de interes).

INTERVALE DE ÎNCREDERE

Medie: $\left[\overline{X} - Z_\alpha \frac{s}{\sqrt{n}}, \overline{X} + Z_\alpha \frac{s}{\sqrt{n}}\right]$

Frecvențe: $\left[f - Z_\alpha \sqrt{\frac{f(1-f)}{n}}; f + Z_\alpha \sqrt{\frac{f(1-f)}{n}}\right]$

COEFICIENTUL DE CORELAȚIE PEARSON

$$r = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sqrt{\sum (X - \overline{X})^2 \sum (Y - \overline{Y})^2}}$$

TESTUL HI-PĂTRAT (χ^2)

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{(f_i^0 - f_i^t)^2}{f_i^t}$$

unde χ^2 = parametrul testului χ^2; f_i^o = frecvența observată; f_i^t = frecvența teoretică. Regiunea critică pentru α = 0,05 este [3,84, ∞).

- Dacă $\chi^2 \in [3,84; \infty)$ se respinge H_0 cu un risc de eroare de tip I (α).
- Dacă $\chi^2 \notin [3,84; \infty)$ se acceptă H_0 cu un risc de eroare de tip II (β).

TESTUL Z PENTRU PROPORȚII

Compararea unei frecvențe observate cu o frecvență teoretică

$$z = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}}$$

p = frecvenţă teoretică (într-o populaţie), f = frecvenţă observată, n = volumul eşantionului.

Testarea egalităţii a două frecvenţe

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},\quad p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$

p_1 = frecvenţa în primul eşantion; n_1 = volumul primului eşantion; p_2 = frecvenţa în cel de-al doilea eşantion; n_2 = volumul celui de-al doilea eşantion.

RISCURI ŞI RATE – TABELUL DE CONTINGENŢĂ 2×2

	Patologie+	Patologie-	Total
Test +	AP	FP	= AP+FP
Test -	FN	AN	= FN+AN
Total	= AP+FN	=FP+AN	= AP+FP+FN+AN = n

unde AP = adevărat pozitivi, AN = adevărat negativi, FP = fals pozitivi, FN = fals negatici

Denumire	Formula
Rata falşilor pozitivi	=FP/(FP+AN)
Rata falşilor negativi	=FN/(FN+AP)
Sensibilitatea	=AP/(AP+FN)
Specificitatea	=AN/(AN+FP)
Acurateţea	=(AP+AN)/n
Valoarea predictivă pozitivă	=AP/(AP+FP)
Valoarea predictivă negativă	=AN/(AN+FN)
Riscul relativ	=AP(FP+AN)/FN(AP+FP)
Rata şansei	=(TP·TN)/(FN·FP)
Risul atribuabil	=AP/(AP+FP)-FN/(FN+AN)

TESTUL Z DE COMPARARE A MEDIEI UNUI EŞANTION CU MEDIA UNEI POPULAŢII

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$ unde μ_0 = media populaţiei; $\overline{X}$ = media eşantionului; σ = deviaţia standard a populaţiei şi n = volumul eşantionului

- **Regiunea critică** pentru $\alpha = 0{,}05$ (testul bilateral):

$(-\infty , -1{,}96] \cup [1{,}96 , \infty)$.

TESTUL STUDENT (T) DE COMPARARE A UNEI MEDII CU O MEDIE CUNOSCUTĂ (VARIAŢII NECUNOSCUTE)

$$t = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$ unde μ_0 = media populaţiei; $\overline{X}$ = media eşantionului; s = deviaţia standard a eşantionului şi n = volumul eşantionului.

- **Numărul de grade de libertate (df)*:*** df = n-1
- **Regiunea critică** pentru α = 0,05 (testul bilateral): $(-\infty, -t_{n-1,0.025}] \cup [t_{n-1,0.025}, +\infty)$

TESTUL STUDENT (T) DE COMPARARE A DOUĂ MEDII (VARIAŢII EGALE)

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ unde $$s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$$

$\overline{X}_1$ = media primului eşantionului; n_1 = volumul primului eşantion; s_1^2 = variaţia primului eşantion; $\overline{X}_2$ = media celui de-al doilea eşantion; n_2 = volumul celui de-al doilea eşantion; s_2^2 = variaţia celui de-al doilea eşantion.

TESTUL STUDENT (T) DE COMPARARE A MEDIILOR A DOUĂ EŞANTIOANE PERECHI

$$t = \frac{\overline{d}}{\frac{s}{\sqrt{n}}}$$, unde $$\overline{d} = \frac{(d_1 + d_2 + ... + d_n)}{n}$$; d_i = diferenţa dintre valoarea iniţială şi valoarea finală ($1 \le i \le n$) ; s = deviaţia standard a diferenţelor; n = volumul eşantionului

TESTUL Z DE COMPARARE A MEDIILOR A DOUĂ POPULAŢII (VARIAŢII CUNOSCUTE ŞI INEGALE)

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$ unde $\overline{X}_1$ = media primului eşantionului; n_1 = volumul primului eşantion; s_1^2 = variaţia primului eşantion; $\overline{X}_2$ = media celui de-al doilea eşantion; n_2 = volumul celui de-al doilea eşantion; s_2^2 = variaţia celui de-al doilea eşantion.

ESTIMAREA VOLUMULUI EŞANTIONULUI

Compararea mediilor (date normal distribuite)

$$n = \frac{(z_{1-\alpha} - z_{1-\beta})^2 \sigma^2}{(m - \mu)^2}$$

Valoarea critică pentru testul bilateral: $z_{1\text{-}5\%} = 1{,}960$, $z_{1\text{-}\beta(\beta=20\%)} = 0{,}842$

Compararea a două medii (date normal distribuite)

$$n_1 = n_2 = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 (\sigma_1^2 + \sigma_2^2)}{d^2}$$

Medii (datele nu respectă distribuţia normală)

$$n = \frac{\sigma^2}{\alpha k^2}$$

unde k = diferenţa pe care dorim să o identificăm (valoare clinică de interes aleasă arbitrar)

Medii: nu există nici un obiectiv a priori desfăşurării experimentului

$$n = \frac{(z_{1-\alpha} - z_{1-\beta})^2 \sigma^2}{(m - \mu)^2}$$

PROPORŢII

Proporţia centrală (valoarea nu este în apropierea valorilor extreme 0 sau 1):

$$n = \left[\frac{z_{1-\alpha/2}\sqrt{\pi(1-\pi)} + z_{1-\beta}\sqrt{p(1-p)}}{p - \pi}\right]^2$$

unde π = proporţia teoretică; p = proporţia dorită

Proporţia extremă (în apropierea valorii de 0 sau 1):

$$n = \left[\frac{z_{1-\alpha/2}\sqrt{\pi} + z_{1-\beta}\sqrt{p}}{p - \pi}\right]^2$$