

# ESTADÍSTICA

Mitjana	$M_e = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$	- Valor que ocupa el lloc central de totes les dades ordenades de menor a major - Si el total de dades és par = mitja dels dos valors centrals
Mitja / Valors agr.	$\bar{X} = \frac{\sum_{i=1}^N x_i}{N} \quad \bar{X} = \frac{\sum_{i=1}^K x_i \cdot f_i}{N}$	- sumar totes les dades i dividir entre el n total - k=classes - taula freq. → límits   xi   fi   xi·fi
Moda Valors agr.	$M_o = L_i + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot a_i$ Taula freq. → límits   fi	- 2 modes: bimodal // +2 modes: multimodal - Si tots els valors tenen la mateixa freq.: no hi ha moda - Si la freq. màxima són valors adjacents es fa la mitja
Moda Valors agr. Dif. amplituds	$h_i = \frac{f_i}{a_i} M_o = L_i + \frac{(h_i - h_{i-1})}{(h_i - h_{i-1}) + (h_i - h_{i+1})} \cdot a_i$	1. calculem les diferents altures - taula freq. → amplituds   fi   hi 2. La classe modal és la que té major altura
Quartils	2, 3, 4, 5, 6, 7, 9 ↓ ↓ ↓ Q <sub>1</sub> Q <sub>2</sub> Q <sub>3</sub> n pars 1, 2, 3, 4, 5, 6, 7, 9 2.5 4.5 6.5 ↓ ↓ ↓ Q <sub>1</sub> Q <sub>2</sub> Q <sub>3</sub> n impars	- 3 n per dividir grup de dades ordenades(4 parts) - Q1,Q2,Q3 = 25%,50%,75% // Q2 = mitjana - Dades ordenades de menor a major
Quartils Valors agr.	$Q_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i$	- taula freq. → límits   fi   Fi
Decils	$D_k = L_i + \frac{\frac{k \cdot N}{10} - F_{i-1}}{f_i} \cdot a_i$ k = 1,2,3... L <sub>i</sub> = límit inf. N = suma f <sub>i</sub> F <sub>i-1</sub> = Fr.Ac. Ant. a <sub>i</sub> = amplitud	- 9n per dividir grup de dades ordenades(10 parts) - 10%,20%...90% // D5=mitjana, Q2
Percentils	$P_k = L_i + \frac{\frac{k \cdot N}{100} - F_{i-1}}{f_i} \cdot a_i$	- 99n per dividir grup de dades ordenades(100 parts) - 1%,2%...99% // P50=mitjana, Q2, D5
Variància	$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}$	- mesura de dispersió de dades respecte la mitja - és igual que la desviació típica però al quadrat - calculem primer la mitja
Variància Valors agr.	$\sigma^2 = \sum_{i=1}^n \frac{x_i^2 \cdot f_i}{N} - \bar{x}^2$	- taula freq. → límits   xi   fi   xi·fi   xi^2·fi
Desviació típica	$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}}$	- dispersió de dades respecte la mitja - arrel quadrada de la variància
Desv. Típica Valors agr.	$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot f_i}{N}}$	- taula freq. → límits   xi   fi   xi·fi   xi^2·fi
Coefficient de variació	$C.V. = \frac{\sigma}{\bar{X}} \cdot 100$	- relació entre desv. típica d'una mostra i la mitja - s'expressa en percentatge - compara les dispersions de 2 distribucions (si les mitges son +)
Puntuacions diferencials	$x_i = X_i - \bar{X}$	- es resta a les puntuacions la mitja
Puntuacions típiques	$z = \frac{X_i - \bar{X}}{\sigma}$	- tipificació: comparar puntuacions de diferent distribucions - es divideix les puntuacions entre la desviació típica - són independents de les unitats utilitzades
Distribucions bidimensionals	- <b>Relació funcional</b> → coneixent X sabem el valor de Y    - <b>Relació estadística</b> → coneixent X estímem el valor de Y - <b>Var. Est. Bidimensionals</b> → cada un està definit per (X, Y), tenen rel. estadística, una es independent i l'altre dependent - <b>Distr. Bidimensionals</b> → cada un està definit per (Xi, Yi), si els representem amb coords.= núvol de punts/diagrama dispersió → recta regr.	
Rectes de regressió	$y - \bar{y} = \frac{s_{xy}}{\sigma_x^2} (x - \bar{x})$	- recta que s'ajusta el millor possible al núvol de punts - passa per el punt $(\bar{x}, \bar{y})$ = punt de gravetat
Covariància	$s_{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i - (\bar{x} \cdot \bar{y})$	- mitja aritmètica dels productes respecte les seves mitges $\alpha_{xy} > 0$ → Correlació directa $\alpha_{xy} < 0$ → Correlació inversa
Coefficient de Correlació	$r = \frac{s_{xy}}{\sigma_x \cdot \sigma_y}$ - r proper a -1 o 1 → correlació forta - r proper a 0 → correlació dèbil - r = -1 o 1 → els punts estan sobre la recta = dependència funcional	- Quocient entre la covariància i el producte de les derivacions típiques - És un nombre real entre -1 i 1

Taula de freqüències

$X_i$ valors	$f_i$ Freq. Abs.	$F_i$ Freq. Abs. Ac.	$n_i$ Freq. Rel.	$N_i$ Freq. Rel. Ac.
Valors ordenats <	Vegades que apareix un valor, total = N	Valor final = N	Freq. Abs. / N	Valor final = 1

Taula de freqüències agrupades

límits classe	$C_i$ marca classe
[ l.inferior , l.superior ) amplitud = l.sup - l.inf	punt mig de l'interval (l.sup - l.inf)/2

Amplitud →  
restem el menor i el major i elegim un nombre més gran que la dif. i divisible  
el nombre ha de ser entre 6 i 15 (l.inf pertany a L'interval, el l.sup no)