Formulari Estadística Descriptiva

- Percentil p de dades agrupades en intervals: $P_p = L_p + (L_{p+1} - L_p)\dfrac{N \cdot p - N_{p-1}}{n_p}$
- Coeficient de simetria: $g_1 = \frac{m_3}{s^3}$, s: desviació típica
  - Dades brutes: $m_3 = \dfrac{(x_1-\bar{x})^3 + (x_2-\bar{x})^3 + \cdots + (x_N-\bar{x})^3}{N}$
  - Dades en taula de freqüències: $m_3 = \dfrac{(x_1-\bar{x})^3 n_1 + (x_2-\bar{x})^3 n_2 + \cdots + (x_k-\bar{x})^3 n_k}{N}$
- Coeficient d'apuntament: $g_2 = \frac{m_4}{s^4} - 3$, s: desviació típica
  - Dades brutes: $m_4 = \dfrac{(x_1-\bar{x})^4 + (x_2-\bar{x})^4 + \cdots + (x_N-\bar{x})^4}{N}$
  - Dades en taula de freqüències: $m_4 = \dfrac{(x_1-\bar{x})^4 n_1 + (x_2-\bar{x})^4 n_2 + \cdots + (x_k-\bar{x})^4 n_k}{N}$
- Recta de regressió: $\hat{Y} = aX + b$, $\quad a = \frac{\text{Cov(X,Y)}}{\text{Var(X)}} \quad b = \bar{y} - a\bar{x}$
- Coeficient de contingència: $0 \le C \le \sqrt{1 - \dfrac{1}{\text{mín}(k,l)}}$, $\quad C = \sqrt{\dfrac{\chi^2}{\chi^2 + N}}$, $\quad \chi^2 = \sum_i \sum_j \dfrac{(n_{ij} - e_{ij})^2}{e_{ij}}$, $\quad e_{ij} = \dfrac{n_{i*} n_{*j}}{N}$

Formulari Estadística Inferencial

Variables aleatòries usuals

V.A. (X)	$f_X(x)$		$E(X)$	$Var(X)$	Altres propietats
Binomial $B(n,p)$ $\Omega_X = \{0, 1, \cdots, n\}$	$\binom{n}{x} p^x (1-p)^{n-x}$ 0	si $x \in \Omega_X$ si $x \notin \Omega_X$	np	$np(1-p)$	
Poisson $Po(\lambda)$ $\Omega_X = \{0, 1, \cdots\}$	$\frac{\lambda^x}{x!} e^{-\lambda}$ 0	si $x \in \Omega_X$ si $x \notin \Omega_X$	λ	λ	$B(n,p) \approx Po(np)$ (n gran, p petit)
Uniforme $\mathcal{U}(a,b)$ $\Omega_X = [a,b]$	$\frac{1}{b-a}$ 0	si $x \in [a,b]$ si $x \notin [a,b]$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$F_X(x) = \begin{cases} \frac{x-a}{b-a} & x \in [a,b] \\ 0 & x < a \\ 1 & x > b \end{cases}$
Gaussiana $X(\mu, \sigma^2)$ $\Omega_X = \mathbb{R}$			μ	σ^2	$Z \sim N(0,1)$ normal estándar $F_Z(-z) = 1 - F_Z(z)$ $F_X(x) = F_Z(\frac{x-\mu}{\sigma})$ $B(n,p) \approx N(np, np(1-p))$ (n gran) $Po(\lambda) \approx N(\lambda, \lambda)$ (λ gran)

Estadístics més usuals

Paràmetre mostral (estadístic)	Esperança	Variància	Distribució de probabilitat	
$\bar{X}$	$E(\bar{X}) = \mu$	$\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$	$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ $\frac{\bar{X}-\mu}{\hat{s}_X/\sqrt{n}} \sim t_{n-1}$ $\bar{X} \sim N(\mu, \frac{\hat{s}_X^2}{n})$	població normal, σ conegut població normal, σ desconegut, $n \le 30$ σ desconegut, $n > 30$
$\hat{s}_X^2$	$E(\hat{s}_X^2) = \sigma^2$	$\text{Var}(\hat{s}_X^2) = \frac{2\sigma^4}{n-1}$	$\frac{n-1}{\sigma^2} \hat{s}_X^2 \sim \chi^2_{n-1}$	població normal
$\hat{p}_X$	$E(\hat{p}_X) = p$	$\text{Var}(\hat{p}_X) = \frac{p(1-p)}{n}$	$\hat{p}_X \sim N(p, \frac{p(1-p)}{n})$ $\hat{p}_X \sim t_{n-1}$	$n > 30$ població normal, $n \le 30$

Intervals de confiança més usuals

Paràmetre mostral	Interval de confiança	
Mitjana	$\bar{X} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$	població normal, σ conegut
	$\bar{X} \pm t_{n-1,\alpha/2} \dfrac{\hat{s}_X}{\sqrt{n}}$	població normal, σ desconegut i $n \leq 30$
	$\bar{X} \pm z_{\alpha/2} \dfrac{\hat{s}_X}{\sqrt{n}}$	si $n > 30$
Variància	$\left[\dfrac{n-1}{\chi^2_{n-1,1-\alpha/2}} \hat{s}^2_X, \dfrac{n-1}{\chi^2_{n-1,\alpha/2}} \hat{s}^2_X \right]$	si la població segueix una llei normal
Proporció	$\hat{p}_X \pm z_{\alpha/2} \sqrt{\dfrac{\hat{p}_X(1-\hat{p}_X)}{n}}$	si $n > 30$