S T Q Q S S D ___/___/___

MA201 statistique vendredi 21/10/2022

Résumé théorique

n-échantillon : sous ensemble de taille n de la population

moyenne : $\bar{x} = 1/n \sum_{i=1}^{n} x_i$

variance : $v = 1/n \sum_{i=1}^{n} (x_i - \bar{x})^2$

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

$\mathbb{E}[aX] = a\,\mathbb{E}[X]$

$Var(\hat{\theta}_n) = \mathbb{E}[\hat{\theta}_n^2] - \mathbb{E}[\hat{\theta}_n]^2$

$Var(a\hat{\theta}_n + b) = a^2\,Var(\hat{\theta}_n)$

$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$: estimateur convergent à cause de la variance asymptotiquement nulle

biais : $B_\theta(\hat{\theta}_n, \theta) = \mathbb{E}[\hat{\theta}_n] - \theta$ | $\neq 0 \to$ biaisé ; $= 0 \to$ non biaisé

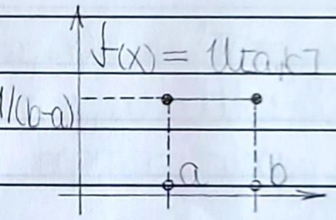
estimateur ; espérance de la variable θ

$\Rightarrow$ représente le décallage entre l'estimation et du vrai valeur

loi uniforme $\mathbb{1}_{[a,b]} = \begin{cases} 1/(b-a) \,;\; x \in [a,b] \\ 0 \,;\; x \notin [a,b] \end{cases}$

$\mathbb{E}[\mathbb{1}_{[a,b]}] = (a+b)/2$

$Var(\mathbb{1}_{[a,b]}) = (b-a)^2/12$



$F(x) = \begin{cases} 0 \,;\; x < a \\ (x-a)/(b-a) \\ 1 \,;\; x > b \end{cases}$

$$\mathbb{E}[X^n] = \frac{(b^{n+1} - a^{n+1})}{(n+1)(b-a)}$$

espérance

continue ↪ fonction densité

$\mathbb{E}[X_i] = \int x \cdot f_{X_i}(x)\,dx$

$\mathbb{E}[X_i^n] = \int x^n \cdot f_{X_i}(x)\,dx$

$\mathbb{E}[1/X] = \int 1/x \cdot f_{X_i}(x)\,dx$

discrète

$\mathbb{E}[X_i] = 1/n \sum_{i=1}^{n} X_i$

fonction de répartition

$\mathbb{1}_{[a,b]}(x) = \begin{cases} 1 & x \in [a,b] \\ 0 & x \notin [a,b] \end{cases}$; $\int_{-\infty}^{+\infty} f(x)\,\mathbb{1}_{[a,b]}\,dx = \int_a^b f(x)\,dx$

__/__/__ S T Q Q S S D

ma201 statistique — vendredi 21/10/2022

Résume théorique

47 fonction de densité de probabilité

$$f_X(x) = \frac{d}{dx} F_X(x)$$

← Fonction de Répartition

16 matrice d'information de fisher

$$\mathbb{E}\begin{bmatrix} \partial^2/\partial\theta_1^2 A & \partial^2/\partial\theta_1\partial\theta_2 A \\ \partial^2/\partial\theta_2\partial\theta_1 A & \partial^2/\partial\theta_2^2 A \end{bmatrix}$$

loi dérivée pour chaque variable

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial \overset{\ln}{\log} f(x_i;\theta)}{\partial\theta}\right)^2\right] = -\mathbb{E}\left[\frac{\partial^2 \log(f(x_i;\theta)|\theta)}{\partial\theta^2}\right]$$

$A = \ln(f(x_i;\theta))$ ln peut être utilisé ln

19 maximum de vraisemblance

$$L(\theta; x_1,\dots,x_n) = \prod_{i=1}^{n} f_i(x_i;\theta)$$

$$\ln(L) = \sum_{i=1}^{n} \ln(f_i(x_i;\theta))$$

$\frac{\partial \ln(L)}{\partial\theta} = 0$, point critique et $\frac{\partial^2 \ln(L)}{\partial\theta^2} < 0$ point de maximum; il faut montrer que c'est maximale

47 loi de Bayes

→ likelihood

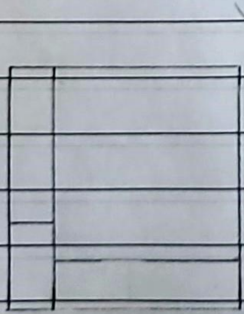
→ probabilité a posteriori

$$f(\theta|x) = \frac{f(x|\theta)\cdot f(\theta)}{f(x)}$$

$f(\theta)$ → probabilité a priori

$f(x)$ → probabilité marginale

peut être une fonction de theta différent comme θ^2, θ^3



modèle linéaire gaussien $Z_K = H_K\theta + B_K$

quand $B_K \sim \mathcal{N}(m_K, R_K)$ donc, par déduction,

$Z_K \sim \mathcal{N}(H_K\theta + m_K, R_K)$

→ peut être autre loi

S T Q Q S S D   __/__/__

MAP01 statistique   vendredi 21/10/2022

Résume théorique

24 Loi a priori de Jeffreys

$J(\theta)=\sqrt{I(\theta)}$

Loi normale (gaussienne)

$$\mathcal{N}(\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-(x-\mu)^2/2\sigma^2)$$

$\mathbb{E}[\mathcal{N}(\mu,\sigma^2)]=\mu$; $\mathbb{E}[\mathcal{N}(0,1)]=0$

$Var(\mathcal{N}(\mu,\sigma^2)]=\sigma^2$

$$f(z|\theta)=\frac{1}{\sqrt{2\pi}^N\sigma^N}\cdot\exp\left(\frac{-1}{2\sigma^2}(z-\theta U)^T(z-\theta U)\right)$$

$Z=[z_1,\dots,z_n]^T$

$U=[u_1,\dots,u_n]^T$

$z_n=\theta u_n+b_n$

$b_n\sim\mathcal{N}(0,\sigma^2)$

25 Borne de Cramer et Rao

$BCR(\theta)=I(\theta)^{-1}$

26 Gamma Distribution

$$X\sim\Gamma(\alpha,\beta)=\frac{x^{\alpha-1}e^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)}\xrightarrow{\beta=1/\theta}\Gamma(k,\theta)=\frac{x^{k-1}e^{-x/\theta}}{\theta^{k}\Gamma(k)}$$

$\Gamma(\alpha)=(\alpha-1)!$ $\mathbb{E}[\Gamma(\alpha,\beta)]=\alpha/\beta$ $\mathbb{E}[\Gamma(k,\theta)]=k\theta$

gamma fonction $Var(\Gamma(\alpha,\beta))=\alpha/\beta^2$ $Var(\Gamma(k,\theta))=k\theta^2$

__ / __ / __ S T Q Q S S D

MA201 STATISTIQUE vendredi 21/10/2022

Résumé Théorique

ESTIMATEUR des moments — définition espérance — moments

moment d'ordre 1 $\mathbb{E}_{\theta_1,\theta_2}[X_i] = 1/n \sum_{i=1}^{n} x_i$

moment d'ordre n $\mathbb{E}[X_i^n] = 1/n \sum_{i=1}^{n} x_i^n$

loi de Bernoulli $\mathcal{B}e(k,p) = p^k (1-p)^{1-k} \quad \forall k \in \{0,1\}$

$\mathbb{E}[\mathcal{B}e(k,p)] = p$

$Var(\mathcal{B}e(k,p)) = pq = p(1-p)$

loi Exponentiel $\mathcal{E}(x,\lambda) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x) = f(x;\lambda)$; $F(x;\lambda) = (1-e^{-\lambda x})\mathbb{1}_{\mathbb{R}^+}(x)$

$\mathbb{E}[\mathcal{E}(x,\lambda)] = 1/\lambda$; $\mathbb{E}[\mathcal{E}(x,\lambda)] = 2/\lambda^2$

$Var(\mathcal{E}(x,\lambda)) = 1/\lambda^2$

loi de Pareto $Pa(\alpha,\theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} \quad \forall x \geq \theta = f(x,\theta)$; $F(x,\theta) = \begin{cases} 0 ; & x < \theta \\ 1-(\theta/x)^\alpha ; & x \geq \theta \end{cases}$

$\mathbb{E}[Pa(\alpha,\theta)] = \alpha\theta/(\alpha-1) \quad \forall \alpha > 1$

$Var(Pa(\alpha,\theta)) = (\theta/(\alpha-1))^2 \, \alpha/(\alpha-2) \quad \forall \alpha > 2$

} sinon ∞

ESTIMATEUR du RISQUE moyen en valeur absolue

$r_{MVA} = \mathbb{E}_{\theta}[|\theta - \hat{\theta}|] = \mathbb{E}_{\theta|z}[|\theta-\hat{\theta}|] = \int_z^{+\infty} |\theta-\hat{\theta}| f(\theta|z)\, d\theta$

on doit TROUVER $\hat{\theta}$ tel que $\frac{\partial}{\partial\theta}\int_z^{+\infty} |\theta-\hat{\theta}| f(\theta|z)\, d\theta = 0$

le RISQUE EST minimiser

ATTENTION pour manipuler le module il faut considérer les limits

53 2

S T Q Q S S D ___/___/___

ma201 statistique dimanche 23/10/2022

estimateur du risque quadratique

$$rMVA = \underset{\theta}{E}[(\theta-\hat{\theta})^2] = \underset{\theta|z}{E}[(\theta-\hat{\theta})^2] = \int_z^{+\infty} (\theta-\hat{\theta})^2 f(\theta|z)\,d\theta$$

on doit minimiser le risque
donc $\hat{\theta}$ doit satisfaire $\frac{\partial}{\partial\hat{\theta}} \int_z^{+\infty} (\theta-\hat{\theta})^2 f(\theta|z)\,d\theta = 0$

estimateur du maximum de vraisemblance a posteriori

$$\hat{\theta}^{MAP} = \underset{\theta}{\arg\max}\{f_{\theta|z}(\theta|z)\}$$

il faut chequer si l'estimateur est dedans l'intervale correcte, si non on considere les limites de θ

$$\arg\max_S f(x) := \underset{x\in S}{\arg\max} f(x) := \{x \in S \mid f(s) \leq f(x)\ \forall s \in S\}$$

$$\arg\min_S f(x) := \arg\min_S f(x) := \{x \in S \mid f(s) \geq f(x)\ \forall s \in S\}$$

approche fréquentiste: paramètrers deterministes

approche bayésienne: variables aléatoires

spiral®