TD 2 feuille 1

1, 2, 3, 5

Statistiques - TD 2 -

Exercice 1 :

$X_1 \to X_n$ iid $\quad \mu = E(X_1)$

$\sigma^2 = V(X_1)$

$$S_{n-1}^2 = \frac{1}{n-1} \sum^n (X_i - \overline{X_n})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i\overline{X_n} + \overline{X_n}^2)$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - 2\overline{X_n} \sum_{i=1}^n X_i + \sum_{i=1}^n \overline{X_n}^2 \right]$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - \frac{2}{n} \left(\sum_{i=1}^n X_i\right)^2 + n\overline{X_n}^2 \right]$$

$\overline{X_n} = \frac{1}{n}\sum X_i$

$-2\overline{X_n}\sum X_i = -\frac{2}{n}(\sum X_i)^2$

$$(\overline{X_n})^2 = \left(\sum \frac{X_i}{n}\right)^2 = \frac{(\sum X_i)^2}{n^2}$$

$$\text{d'où } S_{n-1}^2 = \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - \frac{2}{n}\left(\sum_{i=1}^n X_i\right)^2 + \frac{1}{n}\left(\sum_{i=1}^n X_i\right)^2 \right]$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - \frac{1}{n}\left(\sum_{i=1}^n X_i\right)^2 \right]$$

$$\left(\sum_{i=1}^n X_i\right)^2 = \left(\sum_{i=1}^n X_i\right)\left(\sum_{j=1}^n X_j\right) = \sum_{i=1}^n X_i^2 + \sum_{i\neq j} X_iX_j$$

$$= \sum_{i=1}^n X_i^2 + 2\sum_{j=1}^n X_iX_j$$

$$\text{Donc } S_{n-1}^2 = \frac{1}{n-1}\left[\sum X_i^2 - \frac{1}{n}\sum X_i^2 - \frac{2}{n}\sum_{j>n} X_iX_j \right]$$

$= \dots$

$$= E(S_{n-1}^2) = \sigma^2$$

$$E(S_{n-1}^2) = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2 = \frac{1}{n-1}\sum_{i=1}^n [V(X_1 - \overline{X_n}) + E(X_1 - \overline{X_n})^2]$$

$$= \frac{1}{n-1}\sum_{i=1}^n [V(X_1) + V(\overline{X_n}) + E(X_1 - X_n)^2 + E(X_1) - E(\overline{X_n})$$

Car $E(X_1\overline{X_n}) = E(X_1\overline{X_n}) \qquad X_i$ à la même loi que X_1.

$$= E\Big(X_1 \cdot \underbrace{\frac{1}{n}\sum_{j=j}^n X_j}_{X_1 + \sum X_j}\Big)$$

$$= \frac{1}{n}E(X_1^2) + \frac{1}{n}E\left(X_1\sum_{j=2} X_j\right)$$

$$= \frac{1}{n}(\sigma^2 + \mu^2) + \frac{1}{n}\sum_{j=2}^n (E(X_iX_j))$$

$\sigma^2 + \mu^2$: $V(X_1)$, $(E(X_1))^2$

Car X_1 et X_j

$$= \frac{1}{n}(\sigma^2 + \mu^2) + \frac{1}{n}\sum_{j=2}^n E(X_1)E(X_j)$$

$$\text{d'où } E(X_i\overline{X_n}) = \frac{1}{n}(\sigma^2 + \mu^2) + \frac{1}{n}(n-1)\mu^2$$

$$= \mu^2 + \frac{\sigma^2}{n}$$

$$E(S_{n-1}^2) = \frac{1}{m-1}\sum_{i=1}^{m} [\sigma^2 + \frac{\sigma^2}{m} - 2(\mu^2 + \frac{\sigma^2}{m}) + 2\mu^2]$$

$$= \frac{1}{m-1}\sum_{i=1}^{m} [\frac{m-1}{m}\sigma^2] = \sigma^2$$

$\longrightarrow$ S_{n-1}^2 est un estimateur sans biais de σ^2.

$E(\hat{\theta}) = \theta$ alors $\hat{\theta}$ est sans biais.

Exercice 5 : X continue

- $(X_1, \dots, X_m)$ iid de m̂ loi que X.
- $L(\theta) = \ln(\mathcal{L}(\theta))$

$$= \ln \prod_{i=1}^{m} f(\theta; x_i)$$

$$= \sum_{i=1}^{m} f(\theta_i, x_i) = \sum_{i=1}^{m} \left(\frac{\theta}{2\sqrt{x_i}} e^{-\theta x_i}\right) = \sum_{i=1}^{m} \ln\left(\frac{\theta}{2\sqrt{x_i}}\right) + \sum_{i=1}^{m} \ln(e^{-\theta x_i})$$

$$= \sum_{i=1}^{m} \ln(\theta) - \sum \ln(2\sqrt{x_i}) - \theta \sum_{i=1}^{m} x_i$$

$$= m\ln(\theta) - \sum_{i=1}^{m} \ln(2\sqrt{x_i}) - \theta \sum_{i=1}^{m} x_i$$

- $\frac{\partial L(\theta)}{\partial \theta} = 0 \iff \frac{m}{\theta} - \sum_{i=1}^{m} x_i = 0$

$$\iff \hat{\theta} = \frac{m}{\sum_{i=1}^{m} x_i}$$

- $\frac{\partial^2 L(\theta)}{\partial \theta^2} = -\frac{m}{\sigma^2} < 0 \implies \hat{\theta}_{MV} = \frac{m}{\sum X_i} = \frac{1}{\bar{X}_m}$

parfois on n'arrive pas à prouver que $\frac{\partial^2 L(\theta)}{\partial \theta^2} < 0$
Il suffit alors de vérifier que $\left.\frac{\partial^2 L(\theta)}{\partial \theta^2}\right|_{\theta = \hat{\theta}} < 0$.

Exercice 2 : X continue, v.a.r de densité

$$f(\mu; x) = \frac{1}{x\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(\ln x - \mu)^2\right\} \mathbb{1}_{\mathbb{R}^+}(x)$$

- $(X_1, \dots, X_m)$ m échantillon i.i.d de m̂ loi que X.
- vraisemblance

$$\mathcal{L}(\mu) = \prod_{i=1}^{m} f(\mu, x_i)$$

- Log. vraisemblance

$$L(\mu) = \ln(\mathcal{L}(\mu)) = \sum_{i=1}^{m} \ln(f(\mu, x_i))$$

d'où $L(\mu) = \sum_{i=1}^{m} \ln\left(\frac{1}{x_i\sqrt{2\pi}} e^{-\frac{1}{2}((\ln x - \mu)^2)}\right)$

$$= \sum_{i=1}^{m} \ln\left(\frac{1}{x_i\sqrt{2\pi}}\right) + \sum_{i=1}^{m} \ln\left(e^{-\frac{1}{2}(\ln x_i - \mu)^2}\right)$$

$$= \sum \ln\left(\frac{1}{x_i\sqrt{2\pi}}\right) - \frac{1}{2}\sum(\ln x_i - \mu)^2$$

- L est dérivable par rapport à μ

$$\frac{\partial L(\mu)}{\partial \mu} = \sum(\ln x_i - \mu)$$

$\frac{\partial L(\mu)}{\partial \mu} = 0 \iff \sum \ln x_i - m\mu = 0 \iff \mu = \frac{1}{m}\sum \ln x_i$

Stat TD2 feuille 2

Pile : $P(X=0^k) = \frac{1}{2}$

Face : $P(X=1^q) = \frac{1}{2}$

$E(X) = 0^k \times \frac{1}{2} + 1^q \times \frac{1}{2}$.

$$\frac{\partial^2 L(\mu)}{\partial \mu} = -m < 0$$

Donc $\hat{\mu}_{MV} = \frac{1}{m} \Sigma \ln x_i$

e) Loi de $Y = \ln X$

→ Trop long, trop compliqué ... tant pis

Exercice 3 : 1) $\forall k \geq 3$, $P(X=k) = C_{k-1}^{2} p^3 (1-p)^{k-3}$

$$L(p) = \ln(\mathcal{L}(p)) = \Sigma \ln (P(X=x_i))$$

d'où $L(p) = \Sigma \ln \left(C_{k-1}^{2} p^3 (1-p)^{k-3} \right)$

$$= \Sigma \ln (C_{k-1}^{2}) + \Sigma 3 \ln p + \Sigma \ln(1-p)(x_i - 3)$$

$$= \Sigma \ln \left(C_{x_i - 1}^{2}\right) + 3m \ln p + \ln(1-p) \Sigma (x_i - 3)$$

$$\frac{\partial L(p)}{\partial p} = \frac{3m}{p} - \frac{1}{1-p} \Sigma (x_i - 3)$$

et $\frac{\partial L(p)}{\partial p} = 0 \iff \frac{3m(1-p) - p \Sigma(x_i - 3)}{p(1-p)} = 0$

$$\iff 3m(1-p) - p \Sigma(x_i - 3) = 0$$

$$\iff 3m - 3\cancel{m}p - p\Sigma x_i + 3\cancel{m}p = 0$$

$$\iff 3m - p \Sigma x_i = 0$$

$$\iff \boxed{p = \frac{3m}{\Sigma x_i}}$$

$$\frac{\partial^2 L(p)}{\partial^2 p} = -\frac{3m}{p^2} - \frac{\Sigma (x_i - 3)}{(1-p)^2} < 0$$

Donc $\hat{p}_{MV} = \frac{3m}{\sum_{i=1}^{m} x_i} = \frac{3}{\bar{x}_m}$

2) 15 fois "3" + 3 × 4 + 2 × 5 = $\frac{67}{20}$

$$\hat{p}_{MV} = \frac{60}{67} = 0,89.$$

Exercice 7 :

$$\begin{cases} X \text{ Tension de rupture} \hookrightarrow \mathcal{N}(\mu, \sigma^2) \\ 1-\alpha = 95\% \Rightarrow \alpha = 5\% \Rightarrow 1 - \frac{\alpha}{2} = 0{,}975 \end{cases}$$

On veut trouver IC pour la moyenne à 95%.

En fait, on cherche IC pour μ ($\mu \Rightarrow$ moyenne)

On a 2 cas :
- si $\sigma = 45$ connu
- si σ inconnu

1er cas : $IC_1 = \left[\overline{x_n} - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \,;\, \overline{x_n} + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$ || à savoir par cœur.

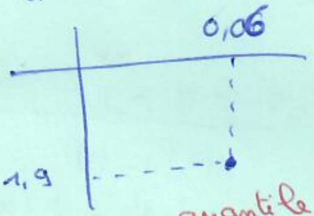
$\rightarrow$ quantile de la loi $\mathcal{N}(0,1)$ d'ordre $1-\alpha$.

Pour l'application numériques on cherche $\overline{x_n}$, $q_{1-\frac{\alpha}{2}}$, n.

$$\left.\begin{array}{l} \overline{x_n} = 307{,}8 \\ q_{1-\frac{\alpha}{2}} = q_{0{,}975} = 1{,}96 \\ n = 10 \text{ (car 10 cordes)} \end{array}\right\} IC_1 = [279{,}91 \,;\, 335{,}65]$$

Pour trouver $q_{0{,}975}$ $\Rightarrow$

$\mathcal{N}(0,1)$ (lecture de la table).



2ème cas : si σ inconnu

$\rightarrow$ quantile de $St(n-1)$ d'ordre $1-\frac{\alpha}{2}$

$$IC_2 = \left[\overline{x_n} - t_{1-\frac{\alpha}{2}}(n-1) \frac{s_{n-1}}{\sqrt{n}} \,;\, \overline{x_n} + t_{1-\frac{\alpha}{2}}(n-1) \frac{s_{n-1}}{\sqrt{n}} \right]$$ || à savoir par cœur

où s_{n-1}^2 est la réalisation de $S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X_n})^2$ (estimateur sans biais de σ^2)

⚠ $t_{1-\frac{\alpha}{2}}(n-1)$ $\rightarrow$ On ne multiplie pas par $(n-1)$ c'est juste une notation.

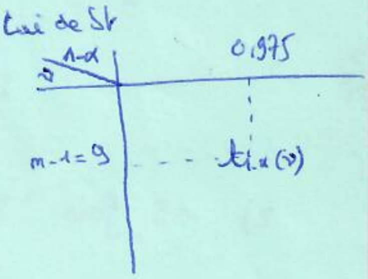
$t \Rightarrow$ quantile de la loi de student.

$\overline{x_n} = 307{,}8$

$t_{1-\frac{\alpha}{2}}(n-1) = t_{0{,}975}(9) = 2{,}262$

$n = 10$

$s_{n-1} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - x_n)^2} = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n \overline{x_n}^2 \right)} = 47{,}56$



donc $IC_2 = [273{,}8 \,;\, 311{,}8]$

Donc le fabricant ne peut pas affirmer que toute les cordes ont une tension supérieur à 300.

Exercice 9.

X durée de vie d'une ampoule $\hookrightarrow \mathcal{E}(\lambda)$

1) Loi de $2\lambda n \overline{X_n} \hookrightarrow \chi^2(2n)$ (ADMIS)

2) Intervalle de confiance de la moyenne du niveau $1-\alpha$

Moyenne : $E(X) = \frac{1}{\lambda}$

$$P\left(k_{1-\frac{\alpha}{2}}(2n) < 2\lambda \overline{X_n} < k_{\frac{\alpha}{2}}(2n)\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\frac{k_{1-\frac{\alpha}{2}}(2n)}{2n\overline{X_n}} < \lambda < \frac{k_{\frac{\alpha}{2}}(2n)}{2n\overline{X_n}}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\frac{2n\overline{X_n}}{k_{\frac{\alpha}{2}}(2n)} < \frac{1}{\lambda} < \frac{2n\overline{X_n}}{k_{1-\frac{\alpha}{2}}(2n)}\right) = 1-\alpha$$

I C de confiance au niveau $1-\alpha$:

$$IC = \left[\frac{2n\overline{x_n}}{k_{\frac{\alpha}{2}}(2n)} ; \frac{2n\overline{x_n}}{k_{1-\frac{\alpha}{2}}(2n)}\right]$$

Pour $1-\alpha = 95\%$

$n = 10$

$\overline{x_n} = 1,77$

$k_{\frac{\alpha}{2}}(2n) = k_{0,025}(20) = 34,2$

$k_{1-\frac{\alpha}{2}}(2n) = k_{0,975}(20)$

$= 9,59$

Donc il faut calculer

$IC =$

Exercice 10 : sondages de taille 1000 $\rightarrow$

1) IC au niveau $1-\alpha = 99\%$ pour p_1 proportion d'électeurs de A.

$$IC = \left[\hat{p}_1 - q_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} ; \hat{p}_1 + q_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}\right]$$

$\mathcal{N}(0,1)$

$\hat{p}_1 = \frac{384}{1000}$

$n_1 = 1000$

$q_{1-\frac{\alpha}{2}} = q_{0,995} = 2,576$ donc $IC = [0,34 ; 0,42]$.

2) $IC_2 = [0,38 ; 0,45]$