Statisque

1/2

TD n°4

Exercice 1

(1)

Modélisation

Soit X_i la va prend la valeur $x_i = 0$ si l'article est bon

$x_i = 1$ sinon

Donc l'estimateur optimal pour p est la fréquence observée d'articles défaillants.

$$\hat{\mu} = \hat{p} = \frac{1}{600} \sum_{i=1}^{600} x_i = \frac{78}{600} = 0,13.$$

(2) Loi de l'estimateur

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

chaque $x_i \rightsquigarrow$ Bernoulli (p) où $0,13$.

$$\Rightarrow \sum_{i=1}^{n} X_i \rightsquigarrow \mathcal{B}(n, p)$$

c'est-à-dire que $n\overline{X} \rightsquigarrow \mathcal{B}(n, p)$

Rappel.

· Si $n > 30$ et $np < 15$ alors la loi $\mathcal{B}(n, p)$ converge vers $P(n.p)$

· Si $np(1-p) > 15$ alors

$$\mathcal{L}[\mathcal{B}(n, p)] = \mathcal{L}[\mathcal{N}(np; np(1-p))]$$

loi

Ici $np(1-p) = 600 \times 0,13(1-0,13) = 67,4 > 15$
Donc on peut approximer $\mathcal{L}(n\overline{X})$ par
$\mathcal{L}(\mathcal{N}(np\,;\,np(1-p))$
et donc $\overline{X} \leadsto N\left(p\,;\,\frac{p(1-p)}{n}\right)$

Exercice 2.

$X_i \leadsto$ Bernoulli (p) avec $p = 0,4$.
$S_n = \sum_{i=1}^{n} X_i \leadsto \mathcal{B}(n,p)$ avec $n = 400$

On cherche $P(S_n \geq 150)$
Comme $n(1-p) = 400 \times 0,4 \times (1-0,4) = 96 > 15$
Donc $S_n \leadsto \mathcal{N}(np, np(1-p))$

$$\Rightarrow \frac{S_n - np}{\sqrt{np(1-p)}} \leadsto \mathcal{N}(0,1)$$

$$P(S_n \geq 150) = P\left(Z_n \geq \frac{150 - np}{\sqrt{np(1-p)}}\right)$$

$$= P\left(Z_n > \frac{150 - 160}{\sqrt{96}}\right)$$

$$= P(Z_n > -1,02) \qquad (F(-x) = 1 - F(x))$$

$$= 0,85 = 1 - F(-1,02)$$

Exercice 3

X_1, X_2, X_3 où chaque X_i à même loi que X : $E(X) = \mu_X$
$V(X) = \sigma_X^2$

μ_X et σ_X^2 sont inconnues.
On veut estimer μ_X.

$$x_1 = \frac{x_1 + 2x_2 + 3x_3}{6}$$

$$\overline{x}_2 = \frac{x_1 + x_2 + x_3}{3}$$

Rappel : si T_m est un estimateur de θ.
Le biais $B(T_m) = E(T_m) - \theta$

$E(\overline{X_1}) = \frac{1}{6}\left(E(X_1) + 2E(X_2) + 3E(X_3)\right)$

$= \frac{6\mu_X}{6} = \mu_X$

$\Rightarrow \overline{X_1}$ est un estimateur de μ_X sans biais :

$E(\overline{X_2}) = \frac{1}{3}\left(E(X_1) + E(X_2) + E(X_3)\right)$

$= \frac{1}{3} 3\mu_X = \mu_X$

$\Rightarrow \overline{X_2}$ est aussi un estimateur sans biais

$V(\overline{X_1}) = \frac{1}{36}\left[V(X_1) + 4V(X_2) + 9V(X_3)\right] = \frac{14}{36}\sigma_X^2 = \frac{7\sigma_X^2}{18}$

$V(\overline{X_2}) = \frac{1}{9}\left(V(X_1) + V(X_2) + V(X_3)\right) = \frac{3\sigma_X^2}{9} = \frac{\sigma_X^2}{3}$

On a $V(\overline{X_2}) < V(\overline{X_1})$
On choisit $\overline{X_2}$.

Exercice 4

$X \sim \mathcal{N}(\mu, \sigma^2)$ avec $\mu = \sigma = \theta$
On propose deux estimateurs pour θ
$T = X$ et $V = \frac{X}{2}$

$E(T) = E(X) = \theta \Rightarrow T$ est sans biais.

$E(U) = \frac{1}{2}E(X) = \frac{\theta}{2} \Rightarrow U$ est borné

- $V(T) = V(X) = \sigma^2$
- $V(U) = \frac{1}{4}V(X) = \frac{\sigma^2}{4}$

$V(U) < V(T)$

Conclusion : il faut rajouter un critère pour faire notre choix.

Exercise 5

X a pour densité $f(x,\theta) = \begin{cases} \frac{2}{\theta}\left(1-\frac{x}{\theta}\right) & \text{si } 0 \leq x \leq \theta \\ 0 & \text{sinon} \end{cases}$

1) $E(X) = \int_{-\infty}^{+\infty} x f(x,\theta)\,dx = \int_0^\theta x \frac{2}{\theta}\left(1-\frac{x}{\theta}\right) dx = \frac{\theta}{3}$

$V(X) = E(X^2) - (E(X))^2 = E(X^2) - \frac{\theta^2}{9}$

or $E(X^2) = \int_0^\theta x^2 \cdot \frac{2}{\theta}\left(1-\frac{x}{\theta}\right) dx$

$V(X) = \frac{\theta^2}{16}$

2) on estime θ par $T_m = \overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$

a) le biais de T_m :

$B(T_m) = E(T_m)$

or $E(T_m) = E(\overline{X}) = E(X) = \frac{\theta}{3}$

$E(\overline{X}) = \frac{1}{m}\sum_{i=1}^{m} E(X_i)$
$= \frac{1}{m}\sum_{i=1}^{m} E(X)$
$= m\,\frac{E(X)}{m} = E(X)$

On modifie T_m en posant :

$T_m = 3\overline{X}$

On a bien $E(T_m) = 3E(\overline{X}) = 3 \times \frac{\theta}{3} = \theta$

b) T_m est convergent si $E(T_m) = \theta$
(vers θ) $V(T_m) \xrightarrow[m\to\infty]{} 0$

or $V(T_m) = V(3\overline{X}) = 9V(\overline{X}) = 9 \times \frac{V(X)}{m} = \frac{9}{m} \times \frac{\theta^2}{18} = \frac{\theta^2}{2m}$

$V(T_m) \xrightarrow[m\to+\infty]{} 0$

Statistique
TD n° 4
2/2

c) $T_n \xrightarrow[n\to\infty]{mq} \theta \iff E(T_n-\theta)^2 \xrightarrow[n\to+\infty]{} 0$

On a $E(T_n) = \theta$, donc :

$E(T_n-\theta)^2 = E(T_n - E(T_n))^2$

// $V(X) = E(X-E(X)) = E(X^2) - [E(X)]^2$

$E(T_n-\theta)^2 = E(T_n - E(T_n))^2 = V(T_n) = \frac{\theta^2}{2n} \xrightarrow[n\to+\infty]{} 0$

d) TCL :

$\frac{T_n - E(T_n)}{\sigma_{T_n}} \sim \mathcal{N}(0,1)$

$\frac{T_n - \theta}{\sqrt{\frac{\theta^2}{2n}}} \sim \mathcal{N}(0,1)$

$\iff T_n \sim \mathcal{N}\left(\theta, \frac{\theta^2}{2n}\right)$

Exercice 6

$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ $\quad \sigma_X^2 = 4$

1) On estime μ_X par :

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

Mq $\bar{X}$ est l'estimateur le plus précis.

$E(\bar{X}) = \mu_X$

$V(\bar{X}) = \frac{\sigma_X^2}{n} = \frac{4}{n}$

$\bar{X}$ est le plus précis si son risque quadratique

$R(\bar{X}) = E[(\bar{X}-\mu)^2] = V(\bar{X}) + \text{Biais}^2(\bar{X})$

$= V(\bar{X}) + B^2(\bar{X})$ est minimal

Comme $B(\bar{X}) = 0$

$\Rightarrow$ Donc $R(\bar{X}) = V(\bar{X}) = \frac{4}{n}$

Il faut donc montrer que la variance de tout autre estimateur sans biais soit $\geqslant \frac{4}{n}$.

Tout autre estimateur est fonction des X_i

$T_n = f(X_1, \dots, X_n)$

Tout estimateur sans biais est une combinaison linéaire des X_i.

On peut proposer :

$$T_n = \frac{\sum_{k=1}^{n} a_k X_k}{\sum_{k=1}^{n} a_k} \quad / \ \sum a_k \neq 0$$

$$V(T_n) = \frac{1}{\left(\sum_{k=1}^{n} a_k\right)^2} \sum_{k=1}^{n} V(\alpha_k X_k)$$

$$= \frac{\sum a_k^2 V(X_k)}{(\sum a_k)^2} = \frac{\sigma_x^2 \sum a_k^2}{(\sum a_k)^2}$$

pour $n = 2$.

$$V(T_2) = \frac{\sigma_x^2 (a_1^2 + a_2^2)}{(a_1 + a_2)^2} \overset{?}{\geqslant} \frac{\sigma_x^2}{2}$$

$2(a_1^2 + a_2^2) > (a_1 + a_2)^2$?

$2a_1^2 + 4a_2^2 > a_1^2 + a_2^2 + 2a_1 a_2$?

$a_1^2 + a_2^2 - 2a_1 a_2 = (a_1 - a_2)^2 > 0$.