Université Blida 1
Faculdes Sciences
Département de Maths

Série d'exercices n°3 de Statistique Inférentielle

Exercice1: Une urne contient 3 boules numérotées 1,2 et 3.
On fait une suite de $n = 100$tirages non exhaustifs et indépendants et on note le $n°$ de la boule tirée lors du ième tirage.
Soit $S_{100} = X_1 + X_2 + ... + X_{100}$. Indiquer le plus petit nombre s que l'on peut trouver tel que $P(200 - s \leq S_{100} \leq 200 + s) \succeq \frac{9}{10}$ (*)
1) A l'aide de l'inégalité de Bienaymé-Tchebitchev
2) A l'aide du théorème Central Limite.

Exercice2: Soit $X_1; X_2$ un échantillon de taille 2 de la loi normale.
Soit $T_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2$; $T_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2$ et $T_3 = \frac{X_1 + X_2}{2}$
a) Les estimateurs T_i ;i=1,2,3 sont-ils sans biais?
b) Lequel des trois estimateurs est le plus efficace?
(Indication prendre la variance comme critère)

Exercice3:a) $X \rightsquigarrow N(25, 36)$. Quelle taille n d'échantillon faut il prendre pour que $\overline{X}$ ait une variance inférieure à 2?
b) $X \leftarrow N(0,1)$ et $Y \rightsquigarrow \chi^2_5$ avec X et Y indépendantes
calculer $P(X \succeq 2\sqrt{\frac{Y}{5}})$

Exercice4: Soit X_1 un estimateur sans biais de μ de variance σ^2
Soit X_2 un estimateur sans biais de μ de variance $4\sigma^2$
Quel est l'estimateur sans biais de μ de la forme $\lambda X_1 + (1 - \lambda)X_2$ le plus efficace ?(cad de variance minimale)

b) Utiliser la variance comme critère de qualité pour choisir le meilleur estimateur.

Exercice 5: Soit X une observation de densité
$f(x, s) = s(1 - x)^{s-1} sis \in [0, 1]$
1) Determiner la densité de $Z = -Ln(1 - X)$ et calculer $E(Z)$
2) Determiner un estimateur G du maximum de vraisemblance. pour s. G est il sans biais? asymptôtiquement sans biais?

Exercice 6: On considère une grandeur aléatoire uniformément distribuée entre 0 et $a(a \succ 0$, inconnu)
a) Si $Y = SupX_i$ est un estimateur de a, déduire de Y un estimateur á qui soit sans biais.
b) Montrer que $\frac{n+\alpha}{n}Y^\alpha$ est un estimateur sans biais de a^α

Exercice 7:Soit X_1,X_2,....,X_n un n-échantillon de vaiid selon la loi

exponentielle de paramètre λ.
a) Montrer que nY et $\frac{S}{n}$ sont des estimateurs sans biais de $\frac{1}{\lambda}$
ou $Y = MinX_i$ et $S = \sum_{i=1}^{n} X_i$.
b) Utiliser la variance comme critère de qualité pour choisir le meilleur estimateur.

Exercice 8: Soit T_1 et T_2 2 estimateurs différents du même paramètre θ. On suppose que $E(T_1) = \theta + b_1$ et $E(T_2) = \theta + b_2$ ou b_1 et b_2 sont des valeurs numériques connues.

a) Soit $T = \alpha T_1 + \beta T_2$. Calculer α et β pour que T soit un estimateur sans
biais de θ.

b) On suppose que $b_1 = b_2 = 0$ et $Cov(T_1, T_2) = 0$.Calculer α et β pour que T soit un estimateur sans biais ayant la plus petite variance possible. Quelle cette variance?

c) Si T_1 et T_2 sont les moyennes empiriques des 2 échantillons au hasard tirées d'une même population au cours d'expéreiences indépendantes, à quoi correspond l'estimateur T à variance minimale du b)?

Serie n° 3

Ex1 : { Inégalité de Bienaymé Tchebichev : $P(|X - E(X)| > \sigma t) < 1/t^2$ }

TCL CV

① $P(200 - \Delta < S < 200 + \Delta) \geq 9/10$

X_i peut prendre les valeurs 1, 2 et 3.

$$P(X_i = 1) = P(X_i = 2) = P(X_i = 3) = 1/3.$$

$$E(X_i) = \sum_{k=1}^{3} k P(X_i = k) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} = \frac{6}{3} = 2 \Rightarrow E(X_i) = 2.$$

$$Var(X_i) = E(X_i^2) - E^2(X_i) = 1 \cdot \frac{1}{3} + 4 \cdot \frac{1}{3} + 9 \cdot \frac{1}{3} - 4 = 2/3 \Rightarrow Var(X_i) = \frac{2}{3}$$

$$E(S) = \sum_{i=1}^{100} E(X_i) = 200.$$

$$\sigma_S^2 = Var(S) = Var \sum_{i=1}^{100} X_i = 100 \, Var X_i = \frac{200}{3} \Rightarrow \sigma_S = 10\sqrt{2/3}$$

$$P(200 - \Delta < S < 200 + \Delta) \geq 9/10 = P(-\Delta < S - 200 < \Delta) \geq 9/10$$

$$P(|S - 200| < \Delta) \geq 9/10 = 1 - P(|S - 200| \geq \Delta) \geq 9/10$$

$$P(|S - 200| \geq \Delta) \leq \frac{1}{10}$$ A l'aide de l'inégalité de Bienaymé Tchebichev

Posons $\Delta = \sigma t \Rightarrow t = \Delta/\sigma$; $P(|S - E(S)| \geq \sigma t) \leq \frac{1}{t^2} = \frac{\sigma^2}{\Delta^2} < \frac{1}{10}$

$$\Delta^2 > 10\sigma^2 \Rightarrow \Delta > \sigma\sqrt{10} = 10\sqrt{10} \cdot \sqrt{2/3} \Rightarrow \Delta > 25,79.$$

② TCL : $\frac{S - E(S)}{\sigma_S} \rightsquigarrow N(0,1)$

$$\frac{S - E(S)}{\sigma_S} = \frac{S - 200}{10\sqrt{2/3}} \rightsquigarrow N(0,1)$$

$$P(200 - \Delta \leq S \leq 200 + \Delta) \geq 9/10$$

$$P\left(\frac{-\Delta}{10\sqrt{2/3}} \leq \frac{S - 200}{10\sqrt{2/3}} \leq \frac{\Delta}{10\sqrt{2/3}}\right) \geq 9/10$$

$$\Phi\left(\frac{\Delta}{10\sqrt{2/3}}\right) - \Phi\left(\frac{-\Delta}{10\sqrt{2/3}}\right) \geq 9/10$$

$$\Phi\left(\frac{\Delta}{10\sqrt{3/2}}\right) \geq 19/20 = 0,95 \xrightarrow{\text{table}} \frac{\Delta}{10\sqrt{3/2}} = 1,645 \Rightarrow \Delta = 13,42$$

$0,9495 \longrightarrow 1,64$
$0,9505 \longrightarrow 1,65$
$\frac{0,9495 + 0,9505}{2} = 0,95$
$\frac{1,64 + 1,65}{2} = 1,645$

Serie n° 3 :

Ex2 : $X_i \sim N(m, \sigma)$; $T_1 = \frac{2}{3}X_1$ $E(\hat{\theta}) = \theta$

a) $T_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2$ $E(T_1) = \frac{2}{3}E(X_1) + \frac{1}{3}E(X_2) = \frac{2}{3}m + \frac{1}{3}m = m$

$T_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2$ $E(T_2) = \frac{1}{4}E(X_1) + \frac{3}{4}E(X_2) = \frac{1}{4}m + \frac{3}{4}m = m$

$T_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ $E(T_3) = \frac{1}{2}E(X_1) + \frac{1}{2}E(X_2) = \frac{1}{2}m + \frac{1}{2}m = m$

$\begin{cases} E(aX_1 + bX_2) = aE(X_1) + bE(X_2) \\ Var(aX_1 + bX_2) = a^2 Var(X_1) + b^2 Var(X_2) \end{cases}$

T_i est un e.s.b de m $\forall i = \overline{1,3}$

b) $Var(T_1) = \frac{4}{9}Var(X_1) + \frac{1}{9}Var(X_2) = \frac{4}{9}\sigma^2 + \frac{1}{9}\sigma^2 = \frac{5}{9}\sigma^2 = \frac{40}{72}\sigma^2$

$Var(T_2) = \frac{1}{16}Var(X_1) + \frac{9}{16}Var(X_2) = \frac{1}{16}\sigma^2 + \frac{9}{16}\sigma^2 = \frac{9}{16}\sigma^2 = \frac{45}{72}\sigma^2$

$Var(T_3) = \frac{36\sigma^2}{72}$

On voit que $Var\,T_3 < Var\,T_1 < Var\,T_2$, donc T_3 est le meilleur estimateur

Ex3 : $Z = -\ln(1-X)$

1) $F_Z(z) = P(Z < z) = P(-\ln(1-X) < z) = P(\ln(1-X) > -z) = P(1-X > e^{-z})$

$= P(X < 1 - e^{-z}) = F_X(1 - e^{-z})$

On dérive : $f_Z(z) = e^{-z} f_X(1-e^{-z}) = e^{-z}\lambda(1-1+e^{-z})^{\lambda-1}$ $Z \sim Exp(\lambda)$

$= \lambda e^{-\lambda z}$ pour $z > 0$ $= \Gamma(1,\lambda)$

$E(Z) = \int_0^{+\infty} z\lambda e^{-\lambda z}\,dz = \frac{1}{\lambda}\boxed{\int_0^{+\infty} t e^{-t}\,dt} = \frac{1}{\lambda}$, $t = \lambda z \Rightarrow z = \frac{t}{\lambda}$

$\Gamma(2) = 1$ $dt = \lambda dz \Rightarrow dz = \frac{dt}{\lambda}$

2) $\mathcal{L}(z_i^n; \lambda) = \prod_{i=1}^{n} f(z_i) = \lambda^n e^{-\lambda\sum_i z_i}$

On cherche la valeur de λ pour laquelle $\mathcal{L}(z^n, \lambda)$ est maximum.

$\max_\lambda(\mathcal{L}(z^n, \lambda)) = \max_\lambda(\lambda^n e^{-\lambda\sum z_i}) \Leftrightarrow (\lambda^n e^{-\lambda\sum z_i})' = 0$

$n\lambda^{n-1}e^{-\lambda\sum z_i} - \lambda^n \sum z_i e^{-\lambda\sum z_i} = 0 \Leftrightarrow n\lambda^{n-1} = \lambda^n \sum z_i$

$\Rightarrow \lambda = \frac{n}{\sum z_i} \longrightarrow \boxed{\hat{\lambda} = \frac{1}{\bar{Z}}} \Rightarrow \frac{1}{\hat{\lambda}} = \bar{Z}$

$E\left(\frac{1}{\hat{\lambda}}\right) = E(\bar{Z}) = E(Z) = \frac{1}{\lambda}$; $\bar{Z} = \frac{1}{\hat{\lambda}}$ est est. s.b pour $\frac{1}{\lambda}$.

Ex4 : $E(X_1) = \mu$ et $Var X_1 = \sigma^2$
$E(X_2) = \mu$ et $Var X_2 = 4\sigma^2$ $\quad T = \lambda X_1 + (1-\lambda) X_2$; $E(T) = \mu$

$Var T = \lambda^2 Var X_1 + (1-\lambda)^2 Var X_2$
$= \lambda^2 \sigma^2 + 4\sigma^2 (1-\lambda)^2 = f(\lambda)$

Var T minimum $\Longrightarrow f'(\lambda) = 0 \Rightarrow 2\sigma^2 \lambda - 8\sigma^2(1-\lambda) = 0 \Rightarrow \lambda = 4/5$
$f''(\lambda) < 0 \quad 2\sigma^2 + 8\sigma^2 = 10\sigma^2 > 0$

$T = \frac{4}{5} X_1 + \frac{1}{5} X_2$

$2\lambda - 8 + 8\lambda = 0$
$10\lambda = 8 \Rightarrow \lambda = 4/5$

Ex5 : $X \sim N(25, 36)$

a) $Var \bar{X} \leq 2 \iff \frac{36}{n} \leq 2 \Rightarrow \underline{n \geq 18}$

b) $X \sim N(0,1)$; $Y \sim \chi^2_5$ et $X \perp Y$

$X \sim N(0,1)$
$Y \sim \chi^2_n$ $\Big\} \Rightarrow \frac{X}{\sqrt{Y/n}} \sim t_n$

$P(X \geq 2\sqrt{Y/5})$

$P(X \geq 2\sqrt{Y/5}) = P\left(\frac{X}{\sqrt{Y/5}} \geq 2\right) = P(T \geq 2)$ avec $T \sim t_5$

|| table

$\frac{0,10}{2} = 0,05$

__Exo6__ : $X_i \sim \mathcal{U}_{[0,a]}$ $(a>0)$ $\forall i = \overline{1,n}$.

1) $Y = \sup_i X_i$. Il faut chercher la densité de Y.

$$F_Y(y) = P(Y<y) = P(\sup_i X_i < y) = P(X_i < y, \forall i = \overline{1,n}) = \left(P(X<y)\right)^n = \left(F_X(y)\right)^n$$

$$\Rightarrow f_Y(y) = n F_X^{n-1}(y) \cdot f_X(y). \qquad F_X(x) = \begin{cases} 0 & \text{si } x<0 \\ x/a & \text{si } 0<x<a \\ 1 & \text{si } x>a \end{cases} \Rightarrow f_X(x) = \{$$

$$\Rightarrow f_Y(y) = n\left(\frac{y}{a}\right)^{n-1} \cdot \frac{1}{a} \cdot 1_{[0,a]}(y) = n\frac{y^{n-1}}{a^n} \cdot 1_{[0,a]}(y).$$

$$E(Y) = \int_0^a y f(y)\,dy = \int_0^a \frac{n}{a^n} y^n\,dy = \frac{n}{a^n} \cdot y^n dy = \frac{n}{a^n} \cdot \frac{a^{n+1}}{n+1} = \frac{na}{n+1} \neq a.$$

Y est biaisé mais asymptotiquement s.b.

$\tilde{a} = \frac{n+1}{n} Y = \frac{n+1}{n} \sup X_i$ est un estimateur s.b pour a.

Car $E(\tilde{a}) = \frac{n+1}{n} E(Y) = \frac{n+1}{n} \cdot \frac{n}{n+1} a = a$.

b) $T = \frac{n+\alpha}{n} Y^\alpha$

$$E(T) = \frac{n+\alpha}{n} E(Y^\alpha) = \frac{n+\alpha}{n} \int_0^a y^\alpha \cdot \frac{n y^{n-1}}{a^n}\,dy = \frac{n+\alpha}{n} \cdot \frac{n}{a^n} \int_0^a y^{n+\alpha-1}\,dy$$

$$= \frac{n+\alpha}{n+\alpha} \cdot \frac{1}{a^n} \cdot a^{n+\alpha} = a^\alpha.$$

__Exo7__ : $Y = \min_i X_i$

$$F_Y(y) = P(Y<y) = P(\min_i X_i < y) = 1 - P(\min_i X_i \geq y) = 1 - P(X_i \geq y, \forall i = \overline{1,n})$$

$$= 1 - \left(P(X \geq y)\right)^n = 1 - \left(1 - P(X \leq y)\right)^n$$

$F_Y(y) = 1 - (1 - F_X(y))^n$ on dérive : $f_Y(y) = n(1 - F_X(y))^{n-1} \cdot f_X(y)$

$$= n e^{-(n-1)\lambda y} \cdot \lambda e^{-\lambda y} \text{ si } y>0$$

$f_Y(y) = n\lambda e^{-n\lambda y}$ si $y>0$

$$E(nY) = n E(Y) = n \int_0^{+\infty} y n\lambda e^{-n\lambda y}\,dy \qquad t = n\lambda y, \; y = \frac{t}{n\lambda}$$

$$= \frac{n}{n\lambda} \int_0^{+\infty} t e^{-t}\,dt = \frac{1}{\lambda} \qquad \Gamma(2) = 1$$

$\Rightarrow nY$ est un estimateur s.b pour $1/\lambda$

$E\left(\frac{\Sigma}{n}\right) = E(\bar{X}) = E(X) = 1/\lambda$

$E((nY)^2) = n^2 E(Y^2) = n^2 \int_0^{+\infty} y^2 n\lambda e^{-n\lambda y}\,dy$; $\begin{cases} t = n\lambda y \\ y = \frac{t}{n\lambda} \end{cases}$

$$= n^2\lambda \int_0^{+\infty} \frac{t^2}{n^2\lambda^2} e^{-t} \cdot \frac{dt}{n\lambda} = \frac{1}{\lambda^2}\Gamma(3) = \frac{\cdot}{\lambda}$$

$Var(nY) = E((nY)^2) - E(nY)^2 = 1/\lambda^2$

$Var(\bar{X}) = \frac{Var X}{n} = \frac{1}{n\lambda^2}$; $\bar{X}$ est meilleur est ... var $V(\bar{X}) < V(nY)$