III - EXERCICES SUR L'ESTIMATION DE PARAMÈTRES

Exercice 1

Début juin 2008, le gouvernement chinois donnait une consigne de boycottage du tourisme en France. De ce fait le nombre de demandes de visas de tourisme déposées à l'ambassade de France à Pekin relevé au cours des trois dernières semaines de juin et des deux premières de juillet a été de : 627 878 925 1105 1213
tandis qu'à la même période l'année précédente, on enregistrait en moyenne 2000 demandes par semaine avec un écart type de 250.
On notera X le nombre de demandes de visas de tourisme déposées à l'ambassade de France à Pekin chaque semaine au cours l'été 2008.
On fera l'hypothèse que X est normalement distribuée et que l'écart type est resté inchangé. On souhaite évaluer le paramètre m = E(X).

1) Proposer un estimateur pour le paramètre m. Cet estimateur est-il sans biais ? Est-il convergent ? Calculer sa variance. A quelle estimation de m, cet estimateur et les observations conduisent-ils ?
2) Si on prend pour m la valeur estimée à la question 1, peut-on dire, comme le titrait *Le Monde* dans son édition du 20 juillet 2008, que "les touristes chinois boudent la France" ? Peut-on dire comme l'affirmait l'article sous jacent de ce même journal que le nombre de demandes "a diminué des 2/3" ?

Exercice 2

Reprendre le contexte de l'exercice précédent. On fait encore l'hypothèse que X est normalement distribuée, mais on ne suppose plus que l'écart type σ est resté inchangé. Désormais on veut estimer σ.

1) Proposer un estimateur pour σ^2. Est-il sans biais ? Etudier sa convergence.
2) Proposer un estimateur sans biais pour σ^2.
3) Quelles estimations de σ^2 permettent les 2 estimateurs précédents et les observations rapportées plus haut ? (cf. énoncé de l'exercice 1)
4) En déduire la meilleure estimation possible pour σ, compte tenu des observations. Préciser quel estimateur a été utilisé pour estimer σ ainsi que ses propriétés.

Exercice 3

Une enquête nationale a montré que 15% des adolescents âgés de 14/15 ans et scolarisés en classe de 3e, présentent un surpoids (obésité incluse). Dans la région R., cette proportion semble plus élevée mais reste à évaluer. On notera ρ la prévalence de surpoids dans la région R. Une enquête menée auprès de 145 élèves scolarisés en classe de 3e et âgés de 14/15 ans a révélé que parmi eux 27 présentent un surpoids.

1) Décrire le modèle d'inférence adapté à la résolution de ce problème, en particulier préciser quelle est la variable du modèle et donner sa loi.
2) Proposer un estimateur pour ρ et étudier ses propriétés. Quelle estimation de ρ amènent cet estimateur et les observations rapportées ?

CamScanner

3) Peut-on affirmer que la proportion d'adolescents présentant un surpoids dans la région R est plus élevée que dans l'ensemble du territoire national ?

Exercice 4

Démontrer que pour qu'un estimateur sans biais T soit convergent en moyenne quadratique, il suffit que sa variance tende vers 0 quand n tend vers l'infini.
Cette propriété est-elle encore vraie si T est seulement asymptotiquement sans biais ?

Exercice 5

Soit X la variable aléatoire dont la loi est donnée dans le tableau ci-dessous :

x_i	-1	0	1
p_i	a	1-2 a	a

où a est un paramètre réel de valeur inconnue.

1) Pour quelles valeurs de a le tableau précédent définit-il bien une loi de probabilité ? Calculer E(X), V(X) et P(X < 0).

2) On pose : $M = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ où $(X_1,..,X_n)$ est un n-échantillon de la variable X.

Déterminer α de sorte que la variable $T = \alpha M$ soit un estimateur sans biais pour a. Montrer que T est convergent et calculer sa variance.

3) Pour i=1...n, on définit une variable Z_i par : $Z_i = 1$ ssi $X_i < 0$ et $Z_i = 0$ sinon ; alors on pose $Y = \sum_{i=1}^{n} Z_i$. Calculer E(Y) et V(Y) et donner une interprétation de Y.

Montrer que la variable $W = \frac{Y}{n}$ est un estimateur sans biais pour a. Etudier sa convergence.

4) Quelles estimations amènent les estimateurs T et W et les observations suivantes :

-1	0	1
37	68	45

Laquelle préférer et pourquoi ?

Exercice 6

Deux urnes identiques comportent la même proportion p de boules blanches. Dans l'urne i, (i = 1, 2) on tire avec remise un échantillon de taille n_i et on note F_i la propostion de boules blanches dans l'échantillon tiré. On supposera n_1 différent de n_2.

On pose $M = \frac{F_1 + F_2}{2}$ et $S = a F_1 + (1 - a) F_2$ où a est un réel, $0 < a < 1$.

1) Montrer que M et S sont deux estimateurs sans biais pour p ~~meilleurs que F_1 et F_2~~.
2) Déterminer a de sorte que S soit de variance minimale. Vérifier que S ainsi défini est meilleur que M et que de plus S et M sont identiques ssi $n_1 = n_2$.

CamScanner

Exercice 7

Le service des études sur la sécurité routière a montré que le nombre d'accidents survenus chaque semaine à un certain carrefour, peut être modélisé par une variable X de loi de poisson de paramètre λ.

Pour estimer λ on dispose des observations suivantes rapportant le nombre d'accidents par semaine tout au long d'une année :

Nombre d'accidents	0	1	2	3	4	Total
Effectifs observés	12	16	15	5	4	52

1) Déterminer un estimateur W pour λ par la méthode du maximum de vraisemblance.
2) Vérifier que cet estimateur W, est l'estimateur "naturel" pour λ, et préciser ses propriétés.
3) En utilisant la loi estimée pour X, calculer la probabilité qu'une semaine donnée on n'enregistre aucun accident. Comparer cette valeur avec la fréquence observée. Conclure.

Exercice 8

Avec le développement d'internet, la diffusion du livre est mise à mal et au cours de la dernière décennie, Paris a vu de nombreuses libairies fermer. un économiste s'est saisi de cette question et propose de modéliser la durée de vie d'une librairie par une variable T de loi exponentielle de densité :

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre de valeur inconnue, $\theta > 0$.

On peut remarquer que de ce fait $\theta = E(T)$ autrement dit θ représente la DDV moyenne d'une librairie.

Par ailleurs en consultant les Bottins du commerce, pour un échantillon de n = 15 librairies il a obtenu : $\sum_{i=1}^{15} t_i = 77$ et $\sum_{i=1}^{15} t_i^2 = 743$.

1) On note $(t_1, t_2, ..., t_n)$ les durées de vie observées sur les n librairies de l'échantillon, calculer la vraisemblance des observations $(t_1, t_2, ..., t_n)$ sachant θ.
2) Calculer alors un estimateur, noté W, pour le paramètre θ, par la méthode du maximum de vraisemblance. Etudier ses propriétés.
3) Quelle estimation de θ permettent les observations et cet estimateur ?

Exercice 9

On considère une variable continue de densité :

$$f(x) = \begin{cases} a & \text{si } x \in [-1, 0] \\ 1 - a & \text{si } x \in]0, 1] \\ 0 & \text{sinon} \end{cases}$$ où a est un paramètre de valeur inconnue, $0 < a < 1$.

1) Montrer que f définit bien une densité de probabilité et calculer E(X), V(X) et P(X<0).

CamScanner

2) Calculer un estimateur T pour a, par la méthode des moments.

3) On note Y l'effectif empirique de l'événement (X<0) dans un n-échantillon (X_1, X_2, ..., X_n) de X. Montrer que si on a observé la valeur (x_1, x_2, ..., x_n) pour (X_1, X_2, ..., X_n), la vraisemblance des observations peut s'écrire $L(x_1, x_2, ..., x_n ; a) = a^y (1 - a)^{n-y}$ où y est la valeur de Y associée aux observations.

4) Calculer alors un estimateur W pour a, par la méthode du maximum de vraisemblance. Montrer qu'il est sans biais et calculer sa variance.

5) Comparer les deux estimateurs T et W.

Exercice 10

On considère une variable discrète dont la loi n'est que partiellement connue. En effet, X peut prendre les valeurs 0, 1, 2 et P(X=1) = P(X=2) = a, où a est un paramètre de valeur inconnue.

1) Pour quelles valeurs de a le tableau ci-dessus définit-il bien une loi de probabilité ? Calculer alors E(X) et V(X).

2) On note Nj le nombre de variables prenant la valeur j dans le n-échantillon (X_1, X_2, ..., X_n). Donner en la justifiant la loi de Nj, j= 0, 1, 2 et montrer que ces variables sont liées par un relation simple que l'on explicitera.

3) Montrer que $\frac{N_1}{n}$ est un estimateur sans biais pour a et calculer sa variance.

4) Déterminer un estimateur T pour a, par la méthode des moments et étudier ses propriétés.

5) Si (x_1, x_2, ..., x_n) est une réalisation de (X_1, X_2, ..., X_n) pour laquelle la variable N_0 prend la valeur n_0, montrer que la vraisemblance des observations (x_1, x_2, ..., x_n) s'écrit : $L(x_1, x_2, ..., x_n ; a) = a^{n-n_0} (1 - 2a)^{n_0}$.

Calculer alors un estimateur W pour a, par la méthode du M.V.

6) On dispose des observations suivantes :

Valeurs de X	0	1	2
Effectifs observés	8	6	31

A quelles estimations les trois estimateurs et ces observations conduisent ils ? Laquelle aura votre préférence ? Justifier.

Exercice 11

Une association d'usagers constate que le temps séparant l'arrivée de deux rames successives sur une certaine ligne de métro, peut être modélisé par une variable aléatoire T de loi uniforme sur [1 ; a] où a > 0, et que les interarrivées sont des v.a. indépendantes. A la demande des usagers et disposant d'une série de n interarrivées (t_1, ..., t_n), on souhaite estimer le paramètre a, temps maximum d'attente entre deux rames.

1) Rappeler les valeurs de E(T) et V(T).

2) Calculer un estimateur A pour a par la méthode des moments. montrer qu'il est sans biais et convergent en un sens à préciser.

3) Calculer un estimateur W pour a par la méthode du MV. Calculer la loi de W. En déduire les propriétés de W.

4) Comparer les estimateurs A et W. Lequel préférer ?

CamScanner

Exercice 1:

1°)

2°) $\hat{m} = 950$

$X \sim N(\hat{m}, \sigma)$; c'à-dire $N(950, 250) \Rightarrow T = \frac{X - 950}{250} \sim N(0,1)$

on cherche t/ $P(-t < T < t) = 0,95 = 1-\alpha \Rightarrow t = 1,96$

$\Rightarrow -1,96 < \frac{X-950}{250} < 1,96 \Rightarrow -1,96 \times 250 + 950 < X < 1,96 \times 250 + 950$

$\Rightarrow 460 < X < 1440 \Rightarrow X \in [460, 1440]$

Comme $2000 > 1440$, donc les touristes chinois bouderont la France

* $\frac{2000 \times 1}{3} \simeq 667$ donc l'article du journal "le Monde" n'était pas précis

Exercice 2:

1°/ on peut proposer $S_n^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ comme estimateur pour σ^2

. $S_n^2 \xrightarrow{cv\,en\,P} \sigma^2$ (Thm - LGN - Slutsky page 26 du cours) loi de grand nombre

2°/ on peut proposer $S_n'^2 = \frac{n}{n-1} S_n^2$ comme estimateur sans biais pour σ^2

(. Asy sans biais $\lim_{n \to +\infty} E(S_n^2) = \sigma^2$
. sans biais : $E(S_n'^2) = \sigma^2$)

* 627 878 925 1105 1213

3) Le premier estimateur S_n^2 : $\frac{1}{5}[(627-950)^2 + (818-950)^2 + (925-950)^2 + (1105-950)^2 + (1213-950)^2]$

$\hat{\sigma}^2 = 40666 \Rightarrow \hat{\sigma} \simeq 201,66$

Le second estimateur $S_n'^2$: $\hat{\sigma}^2 = 50832 \Rightarrow \hat{\sigma} \simeq 225$

4) La plus proche de 250 c'est $\hat{\sigma} = 225$ donc le meilleur estimateur est

$S_n' = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$

Exercice 3:

-1- Modèle d'influence :

• La Population : Les adolescents agés 14/15 ans et scolarisé en classe de 3e dans la région R.

• caractère étudie : surpoids / non

on note p la prévalence de surpoids

p : est un paramètre de valeur inconnu, $0 < p \leq 1$.

• Échantillon prelevé : de taille 145 supposé tiré au hasard avec un taux de sondage inférieur à 10%.

• La variable de modèle : Z défini par $Z = \begin{cases} 1 & \text{si un surpoids} \\ 0 & \text{sinon} \end{cases}$

$$P(Z=1) = p$$

-2- On peut proposer comme estimateur p, la frequence empirique F

$$F = \frac{1}{n} \sum_{i=1}^{n} Z_i \; ; \; F \text{ est sans biais et convergent.}$$

Cet estimateur et les observations rapportés amènent l'estimation

$$\hat{p}_c = \frac{25}{145} = 18,6\%$$

-3- A voir dans l'exercice 19.

Exercice 4:

T : estimateur sans biais.

EX5 :

1). $0 < 1-2a < 1 \Rightarrow \boxed{0 < a < \frac{1}{2}}$

- $E(X) = \sum_{i=1}^{3} x_i p_i = -a + a = 0$
- $V(X) = \sum_{i=1}^{n} x_i^2 p_i = a + a = 2a$
- $P(X<0) = P(X=-1) = a$

2) on pose : $M = \frac{1}{n} \sum_{i=1}^{n} X_i^2$

$T = \alpha M$: $\alpha = ?$ T un estimateur sans biais

$E(T) = \alpha E(M) = \frac{\alpha}{n} \sum_{i=1}^{n} E(X_i^2) = \frac{\alpha}{n} . n2 . a = 2a$

$= a$

$\Rightarrow \boxed{\alpha = \frac{1}{2}}$ et par suite : $\boxed{T = \frac{1}{2} M}$

* T est convergent ? { sans biais asymp sans biais

T C.V en m.q (moyenne quadratique) { $V(T) \xrightarrow[n \to +\infty]{} 0$

$V(T) = \frac{1}{4} V(M) = \frac{1}{4n^2} \sum_{i=1}^{n} V(X_i^2)$

$V(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = 2a - 4a^2$

$\Rightarrow V(T) = \frac{1}{4n^2} n(2a - 4a^2) = \frac{a}{2n}(1-2a)$

$\Rightarrow$ T C.V en m.q vers a, donc T C.V en p $\xrightarrow{n \to +\infty}$ 0 vers a.

la Propriété n'est pas encore vraie si T est seulement asymptotiquement biais.

C-à-d, si T est asymptotiquement sans biais alors $\lim_{n \to \infty} E(T) = m$

---

* la convergence en moyenne quadratique de T par : $\lim_{n \to +\infty} E((T_n - m)^2) = 0$

si $\lim (T_n) \xrightarrow[n \to \infty]{} 0 \Rightarrow \lim_{n \to \infty} E((T_n - m)^2) =$

et T_n est p

- $E(Y) = n.a$, $V(Y) = n.a(1-a)$
- $Y \rightsquigarrow B(n, a)$
- $W = \frac{Y}{n}$ est un estimateur sans biais ?

$E(W) = \frac{1}{n} E(Y) = a$

- $\frac{Y}{n} \xrightarrow{en\ P} a$ (en utilisant la L.G.N. page 25/26 du cours)

- Pour l'estimateur T : les observations amènent l'estimation : $\hat{a} = \frac{57 + 45}{2 * 150} = 0,27$
- Pour l'estimateur W : les observations amènent l'estimation : $\hat{a} = \frac{37}{150} = 0,24$

$V(T) = \frac{a}{2n}(1-2a)$; $V(W) = \frac{a(1-a)}{n}$

$V(T) - V(W) = \frac{a}{2n}(1-2a) - \frac{2a(1-a)}{2n} = \frac{-a}{2n}$

$\Rightarrow V(T) < V(W)$, on préfère T comme estimateur pour le paramètre a.

EX 6 :

F_1 ; $M = \frac{F_1+F_2}{2}$; $S = aF_1 + (1-a)F_2$; $0<a<1$

1°) $E(M) = \frac{1}{2}(E(F_1) + E(F_2)) = p$

$E(S) = aE(F_1) + (1-a)E(F_2) = p$

$\Rightarrow$ M et S sont deux estimateurs sans biais pour p.

2°) $V(S) = a^2 V(F_1) + (1-a)^2 V(F_2)$

$= a^2 \frac{p(1-p)}{n_1} + (1-a)^2 \frac{p(1-p)}{n_2}$

$= p(1-p)\left(\frac{a^2}{n_1} + \frac{(1-a)^2}{n_2}\right)$

* on pose $f(a) = \frac{a^2}{n_1} + \frac{(1-a)^2}{n_2}$

f est dérivable sur $]0,1[$ et $\forall a \in]0,1[$ on a :

on a $f'(a) = \frac{2a}{n_1} - \frac{2(1-a)}{n_2}$

$f'(a) = 0 \Rightarrow \frac{2a}{n_1} = \frac{2(1-a)}{n_2} \Rightarrow a = \frac{n_1}{n_1+n_2}$

• Tableau de variation pour f :

a	0		$\frac{n_1}{n_1+n_2}$		1
$f'(a)$		−	0	+	
$f(a)$	$\frac{1}{n_2}$	↘	$\frac{1}{n_1+n_2}$	↗	$\frac{1}{n_1}$

$f\left(\frac{n_1}{n_1+n_2}\right) = \frac{n_1}{(n_1+n_2)^2} + \frac{n_2}{(n_1+n_2)^2} = \frac{1}{(n_1+n_2)}$

$\boxed{a = \frac{n_1}{n_1+n_2}}$

* vérifions que S est meilleur que M.

$V(M) = \frac{1}{4}\frac{p(1-p)}{n_1} + \frac{1}{4}\frac{p(1-p)}{n_2}$

$= \frac{p(1-p)}{4}\left[\frac{1}{n_1} + \frac{1}{n_2}\right] = p(1-p)\left[\frac{1}{4n_1} + \frac{1}{4n_2}\right]$

Ça revient à comparer : $\frac{1}{4n_1} + \frac{1}{4n_2}$ et $\frac{1}{n_1+n_2}$

$\frac{1}{4}\left(\frac{n_1+n_2}{n_1 n_2}\right) \geq \frac{1}{n_1+n_2}$ car $(n_1+n_2)^2 \geq 4n_1n_2$

$\Rightarrow V(S) \leq V(M)$; par suite S est le meilleur estimateur pour a.

* $S = M \Leftrightarrow n_1 = n_2$

"$\Rightarrow$" : $S = M \Rightarrow \frac{F_1}{2} + \frac{F_2}{2} = aF_1 + (1-a)F_2 \Rightarrow a = \frac{1}{2} \Rightarrow \frac{1}{2} = \frac{n_1}{n_1+n_2} \Rightarrow n_1 = n_2$

"$\Leftarrow$" : $n_1 = n_2 \Rightarrow a = \frac{1}{2} \Rightarrow S = M$

feridbabadiabaz@yahoo.fr

EX 7 :

1. loi de Poisson : $P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

$L(x_1, \dots, x_n) = \prod_{i=1}^{n} P_\lambda(X=x_i) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$

$= e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} = e^{-n\lambda} \frac{\lambda^{n\bar{x}}}{\prod_{i=1}^{n} x_i!}$

-3- ~~$P(X=0)$~~ = Notre estimateur et l'observation amènent l'estimation pour λ : $\hat{\lambda} = \frac{16\times1 + 12\times2 + 3\times5 + 16}{52}$

$\hat{\lambda} = 1,48$

$P(X=0) = e^{-1,48} = 0,227$

• La fréquence observée $= \frac{12}{52} = 0,23$

Notre estimateur est relativement précis

$$L(x_1,\dots,x_n,\lambda) = e^{-n\lambda}\frac{\lambda^{x_1}}{x_1!}\times\frac{\lambda^{x_2}}{x_2!}\times\dots\times\frac{\lambda^{x_n}}{x_n!} = e^{-n\lambda}\frac{\lambda^{n\bar{x}}}{\prod_{i=1}^{n} x_i!}$$

on cherche à maximiser la fct $\lambda \to L(x_1,\dots,x_n,\lambda)$

ça revient à maximiser la fct $\lambda \to Ln(L(x_1\dots x_n,\lambda$

cond° nécessaire :

$$\frac{\partial Ln(L)}{\partial\lambda} = \frac{\partial}{\partial\lambda}\left(Ln\left(e^{-n\lambda}\frac{\lambda^{n\bar{x}}}{\prod_{i=1}^{n} x_i!}\right)\right.$$

$$= \frac{\partial}{\partial\lambda}\left(-n\lambda + n\bar{x}\,Ln(\lambda) - Ln\left(\prod_{i=1}^{n} x_i\right)\right.$$

$$= -n + \frac{n\bar{x}}{\lambda}$$

$$\frac{\partial}{\partial\lambda}\left(Ln(L)\right) = 0 \iff \boxed{\lambda = \bar{x}}$$

cond° suffissante : $\frac{\partial^2}{\partial\lambda^2} Ln(L) = -\frac{n\bar{x}}{\lambda^2} < 0 \quad \forall\lambda$

$\Rightarrow$ notre estimateur pour λ est $\boxed{W = \bar{X}}$

- Les estimateurs obtenus par la MV sont toujours asymptotiquement sans biais

Dans notre cas $E(W) = E(\bar{X}) = E(X) = \lambda$

donc notre estimateur est sans biais

Daprès L.G.N, $\bar{X} \xrightarrow{P_n} E(X)$ donc $\bar{X}$ est un estimateur "naturel" pour λ.

9:

$$f(x)=\begin{cases} a & \text{si } x\in[-1,0] \\ 1-a & \text{si } x\in\,]0,1] \\ 0 & \text{sinon} \end{cases} \quad \text{où } 0<a<1$$

-1- Montrer que f définie bien une densité de Proba : * $f\geq 0$ * f est continue p.p * $\int_{-\infty}^{+\infty} f(x)\,dx = 1$?

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-1}^{0} a\,dx + \int_{0}^{1} 1-a\,dx = ax\Big]_{-1}^{0} + (1-a)x\Big]_{0}^{1}$$

$= a + (1-a) = 1$ Donc définie bien une densité de Probabilité

$$* E(x) = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_{-1}^{0} ax\,dx + \int_{0}^{1} (1-a)x\,dx = a\frac{x^2}{2}\Big]_{-1}^{0} + (1-a)\frac{x^2}{2}\Big]_{0}^{1}$$

$$= \frac{-a}{2} + \left(\frac{1}{2} - \frac{a}{2}\right) = \frac{1}{2} - a$$

$$* V(X) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - [E(x)]^2 = \int_{-1}^{0} ax^2\,dx + \int_{0}^{1} (1-a)x^2\,dx - \left[E(X)\right]^2$$

$$= a\frac{x^3}{3}\Big]_{-1}^{0} + (1-a)\frac{x^3}{3}\Big]_{0}^{1} - \left(\frac{1}{2}-a\right)^2$$

$$= \frac{a}{3} + (1-a) - \left(\frac{1}{2}-a\right)^2 = \frac{1}{3} - \left(\frac{1}{2}-a\right)^2$$

$* P(X<0) = ?$

ona $P(X<0) = F(0)$; pour $x\in[-1,0]$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-1}^{x} 0\,dt = at\Big]_{-1}^{x} = ax+a$

$= a(x+1)$

Donc, $P(X<0) = F(0) = a$

-2- Calculons un estimateur pour a par la méthode $E(X) = \frac{1}{2} - a = \varphi(a)$; avec $\varphi(x) = \frac{1}{2} - x$

• φ est bijective sur $]0,1[$, donc φ est inversible et $\varphi^{-1}(x) = \frac{1}{2} - x$; Donc notre estimateur est $T = \frac{1}{2} - \overline{X}$

• $E(T) = \frac{1}{2} - E(\overline{X}) = \frac{1}{2} - E(X) = \frac{1}{2} - \left(\frac{1}{2} - a\right) = a$; Par conséquent notre estimateur est sans biais.

$\overline{X} \xrightarrow[\text{(LGN)}]{\text{CV en Pr}} E(X) = \frac{1}{2} - a \Rightarrow T \xrightarrow{\text{CV Pr}} a$

Loi de grand nbre

-3- On note Y l'effectif empirique de l'évenement $(X<0)$ dans un n-échantillon de X. $Y \rightsquigarrow B(n,a)$, $p(X\leq 0) = a$

et si on a observé la valeur $(x_1, x_2, \ldots, x_n)$. Pour $(X_1, X_2, \ldots X_n)$, la vraisemblance des observations : $L(x_1, x_2, \ldots, x_n, a) = \prod_{i=1}^{n} f(x_i, a)$

avec $f(x_i, a) = a$ si $x_i \in [-1, 0]$

$f(x_i, a) = 1 - a$ si $x_i \in]0, 1]$ $\Rightarrow L(x_1, x_2, \ldots, x_n, a) = a^y \cdot (1-a)^{n-y}$

y l'effectif de l'évenement $(X < 0)$ dans l'échantillon.

y : Le nombre d'observations < 0 dans $x_1, x_2, \ldots, x_n$.

-4- on cherche à maximiser la fct : $a \to a^y (1-a)^{n-y}$ ou encore la fct : $a \to \ln\left(a^y (1-a)^{n-y}\right)$.

* C.N : $\frac{\partial \ln(L)}{\partial a} = 0 \Rightarrow \frac{\partial}{\partial a}\left(y \ln(a) + (n-y)\ln(1-a)\right)$

$= \frac{y}{a} - \frac{n-y}{1-a} = 0$

$\Rightarrow y(1-a) = a(n-y) \Rightarrow \boxed{a = \frac{y}{n}}$

* C.S : $\frac{\partial^2 \ln(L)}{\partial a^2} = \frac{-y}{a^2} - \frac{(n-y)}{(1-a)^2} < 0, \forall a \in]0, 1[$.

par suite notre estimateur $\boxed{W = \frac{Y}{n}}$

- $E(W) = \frac{1}{n} E(Y) = \frac{1}{n} na = a \Rightarrow$ W est un estimateur sans biais pour a
- $V(W) = \frac{1}{n^2} V(Y) = \frac{1}{n^2} n.a(1-a) = \frac{a(1-a)}{n}$

-5- on va comparer les 2 estimateurs T et W

$V(T) = V\left(\frac{1}{2} - \overline{X}\right) = V(\overline{X}) = \frac{V(X)}{n} = \frac{1}{n}\left(\frac{1}{3}\left(\frac{1}{2} - a\right)^2\right)$

$V(W) = a\frac{(1-a)}{n}$; ça revient à comparer : $a(1-a)$ et $\frac{1}{3} - \left(\frac{1}{2} - a\right)^2$

$= a - a^2$ $= \frac{1}{3} - \left(\frac{1}{4} + a^2 - a\right)$

$= \frac{1}{12} + a - a^2$

donc $V(T) > V(W)$

. par suite notre estimateur préféré pour le paramètre a est W.

Exercice 11 :

-1- T ↝ $U_{[1,a]}$ $\quad E(T) = \frac{1+a}{2}$; $V(T) = \frac{(a-1)^2}{12}$

-2- $E(T) = \frac{1+a}{2} = \varphi(a)$ avec $\varphi(x) = \frac{1+x}{2}$

on a φ bijectif sur $\mathbb{R}_+^*$ donc inversible et $\varphi^{-1}(x) = 2x - 1$

(car $\varphi(x) = \frac{1+x}{2} = y \Rightarrow x = 2y - 1$ donc $\varphi^{-1}(x) = 2x - 1$)

Par suite notre estimateur $A = 2\bar{X} - 1$

$E(A) = 2E(\bar{X}) - 1 = 2E(T) - 1 = a$

donc A est estimateur sans biais pour a.

- $A \xrightarrow{CV.Pr} a$ $\left(\bar{X} \xrightarrow{LGN} E(T)\right)$

3 - $L(t_1 t_2 \cdots t_n, a) = \prod_{i=1}^{n} f(t_i, a) = \prod_{i=1}^{n} \frac{1}{1-a} = \left(\frac{1}{a-1}\right)^n$ (les observations)

Elle est ↓ en a, donc $a = \min(t_1, \ldots, t_n)$

Par suite notre estimateur $W = \min(T_1, \ldots, T_n)$ ← (les v.a)

$$P(W \leqslant u) = P(\min(T_1, \ldots, T_n) \leqslant u)$$

$$= P(\{T_1 \leqslant u\} \cap \{T_2 \leqslant u\} \cdots \cap \{T_n \leqslant u\})$$

$$= \prod_{i=1}^{n} P(\{T_i \leqslant u\}) = (P(T_i \leqslant u))^n = (F(u))^n$$

avec $F(u) = \int_1^u \frac{1}{a-1} dt = \frac{u-1}{a-1}$

$$= \left(\frac{u-1}{a-1}\right)^n$$

$$G(u) = \begin{cases} 0 & \text{si } u < 1 \\ \left(\frac{u-1}{a-1}\right)^n & \text{si } u \in [1, a] \\ 1 & \text{si } u > a \end{cases} \quad ; \quad g(u) = G'(u) = n\left(\frac{u-1}{a-1}\right)^{n-1}\left(\frac{1}{a-1}\right) ; u \in [1, a]$$

$$\Rightarrow g(u) = \begin{cases} n\left(\frac{u-1}{a-1}\right)^{n-1}\left(\frac{1}{a-1}\right) & \text{si } u \in [1, a] \\ 0 & \text{sinon} \end{cases}$$

$$E(W) = \int_1^a u\, g(u)\, du = \int_1^a u\, n\left(\frac{u-1}{a-1}\right)^{n-1}\left(\frac{1}{a-1}\right) du .$$

$$= n\left(\frac{1}{a-1}\right)\int_1^a u\left(\frac{u-1}{a-1}\right)^{n-1} du \quad \text{IPP}$$

$w = u \longrightarrow w' = 1$

$v' = \left(\frac{u-1}{a-1}\right)^{n-1} \longrightarrow \frac{\left(\frac{u-1}{a-1}\right)^n (a-1)}{n}$

$$E(W) = \left[u\left(\frac{u-1}{a-1}\right)^n\right]_1^a = \int_1^a \left(\frac{u-1}{a-1}\right)^n du = a - \int_1^a \left(\frac{u-1}{a-1}\right)^n du$$

$= a - \left[\frac{1}{n+1}\left(\frac{u-1}{a-1}\right)^{n+1}(a-1)\right]_1^a = a - \left[\frac{1}{n+1}\frac{(u-1)^{n+1}}{(a-1)^n}\right]_1^a$

$= a - (a-1)\left[\frac{1}{n-1}\right]$ converge vers a lorsque $n \to +\infty$

$\Rightarrow$ W est un estimateur asymptotiquement sans biais.

-4- A est un estimateur sans biais

W est un estimateur asymptotiquement sans biais

Pour comparer A et W il suffit de comparer leur E.Q.M (erreur quadratique moyenne)

e.q.m $(A) = V(A)$ (car l'estimateur sans biais)

$= V(2\overline{X}-1) = 4V(\overline{X}) = 4\frac{V(X)}{n} = \frac{4}{n}\frac{(1-a)^2}{12}$

$= \frac{(1-a)^2}{3n}$

e.q.m $(W) = V(W) + b_n^2$

$V(W) = \int_1^a u^2 g(u)\,du - E(W)^2$: double IPP

$b_n = E(W) - a$

donc $b_n^2 = \left(-(a-1)\left(\frac{1}{n-1}\right)\right)^2 = \left(\frac{a-1}{n-1}\right)^2$