Correction des exercices

estimation

Ex n°1

1) $\overline{X}$ est un estimateur pour m

$E(\overline{X}) = m$ donc notre estimateur est sans biais.

$\overline{X}$ converge en probabilité vers m

$V(\overline{X}) = \frac{V}{n} = \frac{\sigma^2}{n} = \frac{(250)^2}{5} = 12500$

notre estimateur et les observations amènent à l'estimation $\hat{m} = \frac{\sum x_i}{n} = 950$

2) $\hat{m} = 950$

$X \rightsquigarrow N(\hat{m}, \sigma)$ c'est à dire $N(950, 250)$

$\Rightarrow T = \frac{X - 950}{250} \rightsquigarrow N(0,1)$

On cherche t / $P(-t \leq T \leq t) = 0,95$

$\Rightarrow t = 1,96$

$\Rightarrow -1,96 \leq \frac{X - 950}{250} \leq 1,96$

$\Rightarrow -1,96 \times 250 + 950 \leq X \leq 1,96 \times 250 + 950$

$\Rightarrow 460 \leq X \leq 1440$

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

Scanné avec CamScanner

Comme $\bar{x}_{200} > 1440$, donc les touristes chinois bombent la France.

$$\frac{200 \times 1}{3} \simeq 667$$

donc l'article du journal "le Monde" n'était pas assez précis

Ex N°2

1) On peut prendre $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

Comme estimateur pour σ^2

$S_n^2 \xrightarrow{\text{cvgce P}} \sigma^2$ (Th LGN - Slutsky)

(page 26 du Cours)

$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$

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 $E(S_n^2) = \sigma^2$

627 978 925 1105 1213

- le premier estimateur S_m^{2}

$$\frac{1}{5}\left[(627-950)^2+(878-950)^2+(925-950)^2+(1105-950)^2+(1213-950)^2\right]$$

$$\Rightarrow \boxed{\hat{\sigma}^2=40666} \Rightarrow \boxed{\hat{\sigma}=201}$$

- le second estimateur S'^{2}_m :

$$\hat{\sigma}^2=50832$$

$$\Rightarrow \boxed{\hat{\sigma}=225}$$

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

$$\boxed{S'^{2}_m=\frac{1}{m-1}\sum_{i=1}^{m}(X_i-\bar{X})^2}$$

Ex N°3 :

Modèle d'interférence :

- La population : Les adolescents agés 14/15 ans et scolarisés en classe de 3e dans la région

Scanné avec CamScanner

• Caractère étudié : surpoids / non
on note p la prévalence de surpoids
p : est un paramètre de valeur inconnue
$0 < p < 1$

• Echantillon prélevé : de taille 145 supposé tirée au hasard avec un taux de sondage inférieure à 10%

• La variable de modèle : Z défini par

$$Z = \begin{cases} 1 & \text{si surpoids} \\ 0 & \text{sinon} \end{cases}$$

$P(Z=1) = p$

2) on peut proposer comme estimateur p, la fréquence empirique $F = \frac{1}{n}\sum_{i=1}^{n} Z_i$
F est sans biais et convergent

Cet estimateur et les observations rapportées amènent l'estimation $\hat{p} = \frac{27}{145} = 18,6\%$

3) Voir l'exercice 19

Scanné avec CamScanner

EX N°4 :

T : estimateur sans biais donc $E(T) = \theta$

$\lim_{n \to +\infty} E(T) = \theta$, $\lim_{n \to +\infty} V(T) = 0$

$\lim_{n \to +\infty} E((T-\theta)^2) = 0$

$E((T-\theta)^2) = E(T^2 - 2T\theta + \theta^2)$

$= E(T^2) + E(\theta^2) - 2E(T\theta)$

$= \theta^2 - 2\theta E(T) + E(T^2)$

$= E(T)^2 - 2E(T)^2 + E(T^2)$

$= E(T^2) - E(T)^2$

$= V(T)$

or $\lim_{n \to +\infty} V(T) = 0$ donc $\lim_{n \to +\infty} E((T-\theta)^2) = 0$

par suite T converge vers θ en moyenne quadratique

EX N°5

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

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

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

- $P(X<0) = P(X=-1) = a$

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

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

$$E(T) = \alpha E(H) = \frac{\alpha}{n} \sum_{i=1}^{n} E(X_i^2) = \frac{\alpha}{n} \times n \times 2a$$

$$= 2\alpha a = a$$

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

* Test Convergent ?

T cv dans mq : sans biais ou asymp sans biais

$$V(T) \xrightarrow[n \to +\infty]{} 0$$

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

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

$$\Rightarrow V(T) = \frac{1}{4n^2} n(2a - 4a^2) = \frac{a}{2n}(1 - 2a) \xrightarrow[n \to +\infty]{} 0$$

Scanné avec CamScanner

$\Rightarrow$ T CV en mq Vers a, donc T CV en Pr vers a

3) $E(Y) = na$, $V(Y) = na(1-a)$

$Y \rightsquigarrow B(n,a)$

$W = \frac{Y}{n}$ est un estimateur sans biais?

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

$\frac{Y}{n} \xrightarrow{\text{en Pr}} a$ (en utilisant la LGN page 26-25 du Cours)

4) - Pour l'estimateur T: les observations amènent l'estimation:

$\tilde{a} = \frac{37+45}{2*100} = 0,27$

- Pour l'estimateur W: les observations amènent l'estimation

$\hat{a} = \frac{37}{100} = 0,24$

Scanné avec CamScanner

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

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

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

$$= \frac{a - 2a^2 - 2a + 2a^2}{2n} = \frac{-a}{2n} < 0$$

$$\Rightarrow V(T) < V(W)$$

On préfère T comme estimateur pour le paramètre a.

Ex6 :

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

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.

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

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

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

Scanné avec CamScanner

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

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

$\Rightarrow a n_2 = (1-a) n_1 \Rightarrow a(n_2 + n_1) = n_1$

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

Tableau des variations pour f :

a	0		$\frac{n_1}{n_1+n_2}$		1
$f'(a)$	$\|$	$+$	0	$+$	$\|$
$f(a)$	$\frac{1}{n_1^2}$	$\nearrow$	$\frac{1}{n_1+n_2}$	$\nearrow$	$\frac{1}{n_1}$

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

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

~~Vérification~~

Scanné avec CamScanner

Vérifions que S est meilleur que H :

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

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

$$V(S) = P(1-P)\left(\frac{1}{n_1+n_2}\right)$$

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

$$\frac{1}{4}\left(\frac{n_1+n_2}{n_1 n_2}\right) \overset{?}{\geq} \frac{1}{n_1 n_2}$$

$$(n_1+n_2)^2 \geq 4 n_1 n_2$$

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

$$S = H \overset{?}{\Longleftrightarrow} n_1 = n_2$$

" $\Rightarrow$ " $S = H \Rightarrow \frac{F_1}{2} + \frac{F_2}{2} = a F_1 + (1-a) F_2$

$$\Rightarrow a = \frac{1}{2} \Rightarrow \frac{1}{2} = \frac{n_1}{n_1+n_2}$$

$$\Rightarrow \boxed{n_1 = n_2}$$

" $\Leftarrow$ " $n_1 = n_2 \Rightarrow a = \frac{1}{2} \Rightarrow \boxed{S = H}$

Scanné avec CamScanner

Ex 4

1) $P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$

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

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

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

On cherche à maximiser la fct

$\lambda \longrightarrow L(x_1, \dots, x_n, \lambda)$

Ça revient à maximiser la fct $\lambda \longrightarrow \ln(L(x_1, \dots, x_n, \lambda))$

condition nécessaire :

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

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

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

$\frac{\partial}{\partial \lambda} \ln(L) = 0 \Leftrightarrow \boxed{\lambda = \bar{x}}$

Scanné avec CamScanner

Condition suffisante:

$\frac{\partial^2}{\partial \lambda^2} \ln(L) = -\frac{n\bar{x}}{\lambda^2} < 0 \quad \forall \lambda$

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

2) les estimateurs obtenus par la méthode de maximum HV sont toujours asymptotiquement sans biais

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

donc notre estimateur est sans biais

D'après LGN; $\bar{X} \xrightarrow{Pr.} E(X)$, donc $\bar{X}$ est un estimateur naturel pour λ.

λ

3) notre estimateur et les observation amènent l'estimation pour λ:

$$\hat{\lambda} = \frac{16\times 1 + 15\times 2 + 8\times 3 + 4\times 2}{52} = 1,48$$

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

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

on conclut que notre estimateur est précis

Exercice N°9 :

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

1/ Montrer que f définie bien une densité de probabilité

* $f \geq 0$
* 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

* $E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_{-1}^{0} ax\,dx + \int_{0}^{1} (1-a)x\,dx$

$$= a\frac{x^2}{2}\Big|_{-1}^{0} + (1-a)\frac{x^2}{2}\Big|_{0}^{1}$$

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

Scanné avec CamScanner

* $V(X) = \int_{-\infty}^{0} [\quad]$

$= \int_{-1}^{0} ax^2\,dx + \int_{0}^{1} (1-a)x^2\,dx - (E(X))^2$

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

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

* $P(X<0) = ?$

on a $P(X<0) = F(0)$

pour $x \in [-1, 0]$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-1}^{x} a\,dt$

$= at\Big|_{-1}^{x} = ax + a = a(x+1)$

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

2/ Calculons un estimateur pour a par la méthode des moments

$E(X) = \frac{1}{2} - a = \varphi(a)$, 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 $\boxed{T = \frac{1}{2} - \bar{X}}$

Scanné avec CamScanner

$E(T) = \frac{1}{2} - E(\overline{X}) = \frac{1}{2} - E(X) = \frac{1}{2} - (\frac{1}{2} - a) = a$

par conséquent notre estimateur est sans biais

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

$\Rightarrow T \xrightarrow{CV\ en\ Pr} a$

3/ On note Y l'effectif empirique de l'évènement $(X < 0)$ dans un n-échantillon de X.

si on a observé la valeur $(x_1, x_2, \dots x_n)$ pour $(X_1, X_2, \dots, X_n)$, la vrai semblance des observations :

$L(x_1, x_2, \dots x_n, a) = \prod_{i=1}^{n} f(x_i, a)$

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

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

$\Rightarrow L(x_1, \dots, x_n, a) = a^y (1-a)^{n-y}$

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

Scanné avec CamScanner

y : le nombre d'observations < 0 dans $x_1, x_2, \dots, x_n$

On cherche à maximiser la fonction

$a \longrightarrow a^y (1-a)^{n-y}$

ou encore la fonction : $a \longrightarrow \ln\left(a^y (1-a)^{n-y}\right)$

* CN : $\dfrac{\partial \ln(L)}{\partial a} = 0 \Rightarrow \dfrac{\partial}{\partial a}\left(y \ln(a) + (n-y) \ln(1-a)\right)$

$= \dfrac{y}{a} - \dfrac{(n-y)}{1-a} = 0$

$\Rightarrow (1-a)\, y = (n-y)\, a \Rightarrow y = na$

$\Rightarrow \boxed{a = \dfrac{y}{n}}$

* CS : $\dfrac{\partial^2 \ln(L)}{\partial a^2} = -\dfrac{y}{a^2} - \dfrac{n-y}{(1-a)^2} < 0$

$\forall\, a \in]0,1[$

par suite notre estimateur $\boxed{w = \dfrac{Y}{n}}$

- $E(w) = \dfrac{1}{n} E(Y) = \dfrac{1}{n} na = a$

$\Rightarrow w$ est un estimateur sans biais pour a.

Scanné avec CamScanner

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

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

5) on va comparer les 2 estimateurs T et W

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

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

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

Ça revient à comparer

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

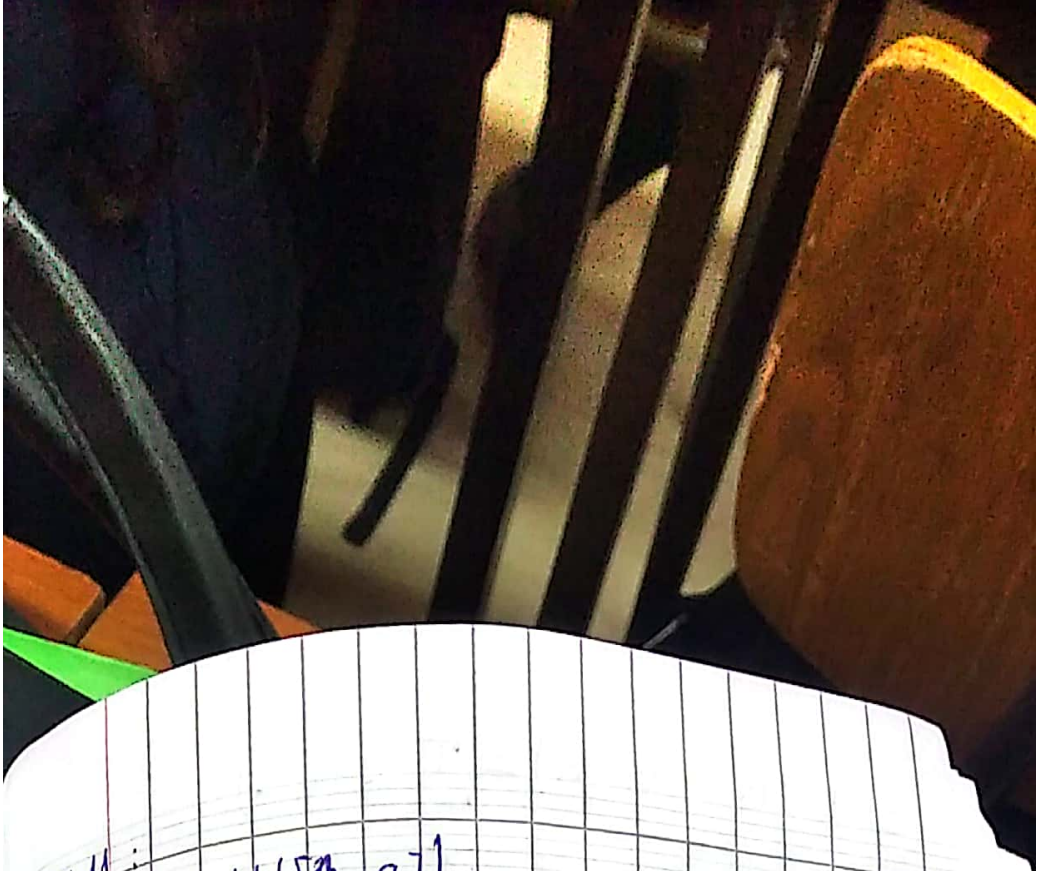
$$= \frac{1}{3} - \left(\frac{1}{4} + a^2 - a\right)$$

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

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

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

Scanné avec CamScanner



Exo 11 :

1/ $T \rightsquigarrow \mathcal{U}([\![1, a]\!])$

$E(T) = \frac{1+a}{2}$; $V(T) = \frac{(a-1)^2}{12}$

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 $\varphi^{-1}(x) = 2x - 1$

(calcul φ^{-1}

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

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 para.

A converge vers a en proba

$\bar{X} \xrightarrow{LGN} E(T)$

Scanné avec CamScanner

3) $L(t_1, \dots t_n, a) = \prod_{i=1}^{n} f(t_i, a)$

$= \prod_{i=1}^{n} \frac{1}{a-1}$

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

L est ↘ en a

donc $a = \min(t_1, \dots, t_n)$

Par suite votre estimateur $w = \min(T_1, \dots T_n)$

$P(w \leq u) = P(\min(T_1 \dots T_n) \leq u)$

$= P(\{T_1 < u\} \cap \{T_2 < u\} \dots \{T_n < u\})$

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

$= (P(T_i \leq u))^n$

$= (F(u))^n$

$u \in [1, a]$

$P(w \leq u) = (F(u))^n$ avec

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

Scanné avec CamScanner

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

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

$u \in [1,a]$

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

$$E(u) = \int_1^a u\, g(u)\, du$$

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

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

on pose $\quad u = u \longrightarrow u' = 1$

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

Scanné avec CamScanner

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

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

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

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

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

Converge vers a lorsque $n \to +\infty$

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

4)

A est un estimateur sans biais

W est un // Asymptotiquement sans biais.

don