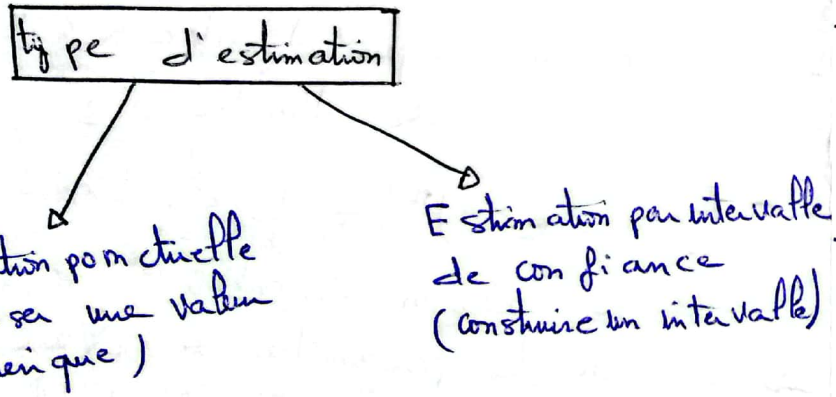
Estimation statistique

L'estimation statistique consiste à donner des valeurs approchées aux paramètres inconnus d'une population (m, σ^2 pour un caractère quantitatif ; p pour un caractère qualitatif à 2 modalités)

Pour toute v.a X et pour tout nombre réel x on prescrira :

$$f(x,\theta) = \begin{cases} f_\theta(x) & \text{densité de } x \text{ si } X \text{ v.a continue} \\ P_\theta(X = x) & \text{si } X \text{ v.a discrète.} \end{cases}$$



Estimation Ponctuelle :

* Propriétés des estimateurs :

•) Biais d'un estimateur :

X une v.a dont la loi donnée par $f(x,\theta)$ est connue, dont paramètre θ est inconnu.

Soient $(X_1, \dots, X_m)$ un échantillon de taille m de la v.a et $T_m = g(X_1, \dots, X_m)$ un estimateur de θ.

→ T_m est dit sans biais si $E(T_m) = \theta$.
→ θ est la paramètre que T est censé estimer.
→ on appelle biais de l'estimateur T_m la quantité $b_m = E(T_m) - \theta$.

→ L'estimateur T_m est dit asymptotiquement sans biais si $\lim_{m \to +\infty} b_m = 0 \Rightarrow \lim_{m \to +\infty} E(T_m) = \theta$.

Exps d'estimateurs sans biais :

* Fréquence empirique : $F_m = \dfrac{X_1 + X_2 + \dots + X_m}{m}$

avec $X_i = \begin{cases} 1 & \text{si cond vrai} \\ 0 & \text{sinon} \end{cases}$

$\Rightarrow E(F_m) = p = \theta \Rightarrow F_m$ estimateur sans biais de p.

Scanné avec CamScanner

⇒ plus n est grand, plus l'estimateur T_n a une forte probabilité de prendre des valeurs proches du paramètre inconnu θ.

• Estimateur Convergent :

- Un estimateur T_n du paramètre θ est dit convergent en probabilité si :

$$\forall \varepsilon > 0 \; ; \; \lim_{n \to +\infty} P(|T_n - \theta| < \varepsilon) = 1$$

ou

$$\lim_{n \to +\infty} P(|T_n - \theta| \geq \varepsilon) = 0$$

- Un estimateur T_n du paramètre θ est dit convergent en moyenne quadratique si :

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

- m.q ⟶ P.P.

- T_n est convergent en moyenne quadratique ssi { il est sans biais ou asymptotiquement sans biais ; sa variance tend vers 0.

- $(Y_n)_{n \geq 0} \xrightarrow{P.S.} Y$, g continue $\Rightarrow (g(Y_n)) \xrightarrow[n \to +\infty]{P.S.} g(Y)$

• Comparaison d'estimateurs :

- On appelle écart quadratique moyen (e.q.m.) la quantité $E\left[(T_n - \theta)^2\right]$

- Relation :

$$E\left((T_n - \theta)^2\right) = Var(T_n) + b_n^2$$

- [Si les deux estimateurs sont sans biais alors on compare les variances] si non l'e.q.m. sera faible, meilleur sera la précision de l'estimation.

• Construction d'estimateurs :

- Moment empirique d'ordre r : $E(X^r) = \frac{1}{n} \sum_{i=1}^{n} x_i^r$

- " " centré d'ordre r :

$$E\left[(X - E(X))^r\right] = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x_n})^r$$

- Méthode des moments :

⟶ cas où le paramètre θ est unidimensionnel

① Soit $(x_1, \dots, x_n)$ un n-échantillon de X

② LGN (loi des grands nombres)

$$\overline{x_n} \xrightarrow{P.S.} E(X) = \varphi(\theta)$$

③ On pose $T_n = \varphi^{-1}(\overline{x_n})$

④ si $E(X) = \varphi(\theta)$ avec φ est bijective sur $\mathbb{R}$ donc inversible

Scanné avec CamScanner

donc $T_m = \varphi^{-1}(\overline{x_m})$

⑤ si $E(x)$ n'est pas en fct de θ on passe au moment théorique d'ordre 2 qui sera en fct de θ et ainsi de suite.

donc $T_m = \varphi^{-1}\left(\frac{1}{m}\sum_{i=1}^{m} x_i^2\right)$

→ cas où le paramètre θ est multidimensionnel

si θ n'est pas un nombre mais un vecteur $\theta = (\theta_1, \dots, \theta_K)$

Il faut dire alors écrire K equations faisant intervenir les moments théoriques.

Le choix des moments sera fait que chacun des $\theta_1, \dots, \theta_K$ figure dans l'expression de l'un au moins de ces moments.

$\Rightarrow E_r(\theta_1, \dots, \theta_K) = E(x^r) \quad (r = 1, \dots, K)$

Il suffira alors de résoudre ce système par rapport aux inconnues $\theta_1, \dots, \theta_K$

- méthode de maximum de vraissemblance (ML)
- on appelle vraissemblance de θ la fonction

$\theta \longrightarrow L(x_1, \dots, x_m; \theta) = f(x_1, \theta) f(x_2, \theta) \dots f(x_m, \theta)$

on rappelle que $f(x, \theta)$ définit la loi de la v.a X par la formule :

$$f(x, \theta) = \begin{cases} f_\theta(x) \text{ densité de } x \text{ si } x \text{ v.a continue} \\ P_\theta(x = x_i) \quad \text{si } x \text{ v.a discrète} \end{cases}$$

- alors $L(x_1, \dots x_m, \theta)$ est la densité de m-uplet $(x_1, \dots x_m, \theta)$ sachant θ si x est une v.a continue et $P_\theta(x_1 = x_1, \dots x_m = x_m)$ si x est une v.a discrète.

les étapes :

① Définir la loi

② La fct de vraisemblance L

③ La log-vraisemblance $\ln$

$\theta \longrightarrow \ln L(x_1, \dots, x_m; \theta) = \sum_{i=1}^{m} \ln f(x_i, \theta)$

④ $\frac{\partial \ln}{\partial \theta} = 0 \Rightarrow \hat{\theta} = \dots = W_m$

Propriétés des estimateurs obtenus par M.V :

(1) L'estimateur W_m est asymptotiquement sans biais et convergent en probabilité.

Il est le meilleur parmi les autres estimateurs sans biais

(2) Lorsque m est grand, la loi de v.a $W_m - \theta$ est voisine de la loi Normale $N(0, \sigma_m)$

avec $\sigma_m = \frac{1}{\sqrt{\quad}}$ avec $I_m(\theta) = m.E\left(\left(\frac{\partial}{\partial \theta} \ln f(x, \theta)\right)^2\right)$ ②

Scanné avec CamScanner

Estimation par intervalle de confiances

I intervalle de confiance pour la moyenne d'une loi $\mathcal{N}(m, \sigma)$

a) L'écart type σ est connu

soit $\overline{X} = \frac{1}{m} \sum_{i=1}^{m} X_i \rightsquigarrow \mathcal{N}\left(m, \frac{\sigma}{\sqrt{m}}\right)$

donc soit la v.a est centrée réduite

$U = \frac{\overline{X} - m}{\frac{\sigma}{\sqrt{m}}} \rightsquigarrow \mathcal{N}(0,1)$

Par suite, si F_0 fonction de répartition de loi normale $\mathcal{N}(0,1)$ on a:

$P(-t \leqslant U \leqslant t) = F_0(t) - F_0(-t) = 2F_0(t) - 1$

et si le seuil de risque est α et on a

$F_0(t) = 1 - \frac{\alpha}{2}$ donc on a:

$m - \frac{t\sigma}{\sqrt{m}} \leqslant \overline{X} \leqslant m + \frac{t\sigma}{\sqrt{m}} \Rightarrow \overline{X} - \frac{t\sigma}{\sqrt{m}} \leqslant m \leqslant \overline{X} + \frac{t\sigma}{\sqrt{m}}$

et par la suite:

$P(-t \leqslant U \leqslant t) = P\left(\overline{X} - \frac{t\sigma}{\sqrt{m}} \leqslant m \leqslant \overline{X} + \frac{t\sigma}{\sqrt{m}}\right)$

$= 1 - \alpha$

D'où: lorsque σ est connu, l'intervalle de confiance de niveau $1-\alpha$ pour m est $\left[\overline{X} - \frac{t\sigma}{\sqrt{m}}, \overline{X} + t.\frac{\sigma}{\sqrt{m}}\right]$

Rqs:

- si on a $\overline{X}$ donc on a la fonchette d'estimation pour m au risque α suivante:

$\overline{X} - t\frac{\sigma}{\sqrt{m}} / \leqslant m \leqslant \overline{X} + t\frac{\sigma}{\sqrt{m}}$

- plus α ↗ plus t ↘ et donc la longueur de l'intervalle ↘

b) L'écart type σ est inconnu

((sm) ne biais ?)

soit $s'^2 = \frac{1}{m-1} \sum_{i=1}^{m} (X_i - \overline{X})^2$ tel que s'^2 estimateur sans biais et convergent de $\sigma^2 \Rightarrow$ on remplace σ par s' de façon que voici là l'intervalle de confiance de m: $\left[\overline{X} - \frac{t s'}{\sqrt{m}}, \overline{X} + \frac{t s'}{\sqrt{m}}\right]$

qui est dû à un calcul précédent de $T = \frac{\overline{X} - m}{\frac{s'}{\sqrt{m}}}$

tel que si $m \geq 30$; $T \rightsquigarrow \mathcal{N}(0,1)$

et si $m < 30$; T ne loi de student à $\nu = m-1$

et par suite; $P(-t \leqslant T \leqslant t) = 1 - \alpha$.

D'où:

lorsque σ est inconnue, l'intervalle de confiance de niveau $1-\alpha$ pour m est:

$\left[\overline{X} - \frac{t s'}{\sqrt{m}}, \overline{X} + \frac{t s'}{\sqrt{m}}\right]$

Scanné avec CamScanner

si on a $\bar{x}$ donc on a la fonchette d'estimation pour m au risque α suivante :

$$\bar{x} - t\frac{s'}{\sqrt{m}} \leqslant m \leqslant \bar{x} + t\frac{s'}{\sqrt{m}}$$

* Intervalle de confiance pour la moyenne d'une loi quelconque :

- même principe que dans $N(m, \sigma)$

Lo σ connu ; $U = \dfrac{\bar{x} - m}{\frac{\sigma}{\sqrt{m}}}$

Lo σ inconnu ; $T = \dfrac{\bar{x} - m}{\frac{s'}{\sqrt{m}}}$

* Intervalle de confiance pour une proportion

soit la suite $(x_1, \dots, x_m)$ un m-échantillon de la v.a X no loi de Bernoulli de paramètre p donc on a la fréquence empirique $F_m = \dfrac{x_1 + \dots + x_m}{m}$ un estimateur sans biais et convergent de p

et $\begin{cases} E(F_m) = p \\ V(F_m) = \dfrac{p(1-p)}{m} \end{cases}$

- Pour m assez grand :

$F_m \rightsquigarrow N\left(p, \sqrt{\frac{pq}{m}}\right)$ d'après TCL

alors soit $U = \dfrac{F_m - p}{\sqrt{\frac{pq}{m}}} \rightsquigarrow N(0,1)$

et $P(|U| < t) = 2\Phi_0(t) - 1 = 1 - \alpha$.

$$\Rightarrow P\left(-t\sqrt{\frac{pq}{m}} \leqslant F_m - p \leqslant t\sqrt{\frac{pq}{m}}\right) = 1 - \alpha$$

$$\Rightarrow P\left(F_m - t\sqrt{\frac{pq}{m}} \leqslant p \leqslant F_m + t\sqrt{\frac{pq}{m}}\right) = 1 - \alpha$$

- On est dans le cas où l'écart type de F_m inconnue :

Méth 1 : estimer pq par $F_m(1 - F_m)$

donc l'intervalle de confiance pour p est :

$$\left[F_m - t\sqrt{\frac{F_m(1-F_m)}{m}} ; F_m + t\sqrt{\frac{F_m(1-F_m)}{m}}\right]$$

Méth 2 : majorer la quantité pq par $\frac{1}{4}$

on a donc $\sqrt{\frac{pq}{m}} \leqslant \frac{1}{2\sqrt{m}}$

donc l'intervalle de confiance par excès sera :

$$\left[F_m - \frac{t}{2\sqrt{m}}, F_m + \frac{t}{2\sqrt{m}}\right]$$

Scanné avec CamScanner

Détermination de la taille d'échantillon pour obtenir une précision et une crédibilité fixées :

* Estimation de la moyenne d'une V.a normale :

$X \rightsquigarrow N(m, \sigma)$

et dans le cas où σ connu

on a l'intervalle de confiance est :

$$\left[\overline{x_m} - \frac{t\sigma}{\sqrt{m}} ; \overline{x_m} + \frac{t\sigma}{\sqrt{m}}\right]$$

donc pour obtenir la précision :

on a $\frac{t\sigma}{\sqrt{m}} \leqslant K|m| \Rightarrow m \geqslant \left(\frac{t\sigma}{K|m|}\right)^2$

Si on a $A_1 \leqslant m \leqslant A_2 \Rightarrow m \geqslant A_1$

$$\Rightarrow \frac{t\sigma}{A_1 K} \geqslant \frac{t\sigma}{K|m|}$$

d'où $m \geqslant \left(\frac{t\sigma}{A_1 K}\right)^2 = \dots$

* Estimation d'une proportion :

soit $m > 100$

donc l'intervalle de confiance est :

$$\left[F_m - t\sqrt{\frac{pq}{m}}, F_m + t\sqrt{\frac{pq}{m}}\right]$$

donc pour obtenir la précision :

on a $t\sqrt{\frac{pq}{m}} \leqslant Kp \Rightarrow m \geqslant \frac{t^2}{K^2}\left(\frac{1}{p} - 1\right)$

Rq :

la crédibilité est mesurée par le niveau de confiance $1-\alpha$

Scanné avec CamScanner

Exo:

1) Rappel (cours échantillonnage pas)

La loi des grandes nombres (LGN)

$(\bar{x}) \xrightarrow{p} E(x) = m$

⇒ on peut utiliser $\bar{x}$ comme estimateur pour le paramètre m.

$E(\bar{x}) = m$

⇒ cet estimateur est sans biais

$(\bar{x}) \xrightarrow{p} m$

$V(\bar{x}) = \frac{V(x)}{m} = \frac{\sigma^2}{m} = \frac{250^2}{5} = 12500$

• cet estimateur et les observations amènent à l'estimation

$\hat{m} = \frac{1}{m} \sum x_i$

$= \frac{1}{5}(627 + 878 + 925 + 1105 + 1213) = 950.$

2) $X \sim N(950, 250)$

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

$P(-r \leq T \leq r) = 1 - \alpha = 0,95$

$\Rightarrow r = 1,96$

$\Rightarrow -1,96 \leq T \leq 1,96$

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

$460 \leq x \leq 1440$

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

⇒ Les taxis chinois bouchent la France

$2000 - 2000 \times \frac{2}{3} = \frac{1}{3} 2000$

$= 667$

⇒ Non, l'article du journal "Le monde" n'a pas précisé.

Exo 2:

1) S_m^2 est un estimateur biaisé, pour σ^2

$S_m^2 = \frac{1}{m} \sum_{i=0}^{m} (x_i - \bar{x})^2$

$= \frac{1}{m} \sum_{i=0}^{m} x_i^2 - \bar{x}^2$

ça converge en proba vers σ^2

(LGN et Slutsky)

2) $S_m'^2$ est un estimateur sans biais pour σ^2.

3) estimateur pour σ^2 en utilisant le premier estimateur S_m^2.

$\sigma^2 = \frac{1}{5}((627 - 950)^2 + (878 - 950)^2 + (925 - 950)^2 + (1105 - 950)^2 + (1213 - 950)^2)$

$= 40548.40666 \Rightarrow \boxed{\sigma = 202}$

* estimateur pour σ^2 en utilisant le second estimateur $S_m'^2$

• $\sigma^2 = \frac{5}{4}(40666) = 50833.$

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

4) ~~S_m et le meilleur~~

$S_m = \sqrt{\frac{1}{m-1} \sum_{i=0}^{m} (x_i - \bar{x})^2}$

est le meilleur estimateur pour σ.

[On essaye comparer l'eq. m].

Scanné avec CamScanner

Ex 3

1) * Le modèle d'urne :
- population : les élèves de 16/17 ans de seconde en classe de 3ème dans la région
- caractère étudié : sont parents nés en métropole ou à l'étranger
- proportion inconnue de parents nés à l'étranger : θ
- échantillon prélevé : Taille 146 représentative au hasard avec un tirage de vendeur (iid)
- variables du modèle : X_i définies par $\begin{cases} 1 \text{ si parents nés à l'étranger} \\ 0 \text{ sinon} \end{cases}$
On a $X_i \sim \mathcal{B}(\theta)$

4) On peut proposer comme estimateur pour θ la fréquence empirique $f = \frac{\sum_{i=1}^{n} X_i}{n}$. Cet estimateur est sans biais et convergent.
* Cet estimateur et la réalisation approchée avec cette estimation $\hat{f} = \frac{29}{146} \approx 19,9\%$

3) a revoir dans l'exercice 19.

Ex 4

- T_n est un estimateur sans biais $E(T_n) = \theta$
- $\lim_{n \to +\infty} V(T_n) = 0$

$V(T_n) = V(T_n - 0)$
$= E\left((T_n - 0)^2\right) - \left(E(T_n - 0)\right)^2$
car $E(T_n - 0) = E(T_n) - E(0)$
$= \theta - \theta = 0$

$\lim_{n \to +\infty} V(T_n) = \lim_{n \to +\infty} E\left((T_n - 0)^2\right) = 0$

2) $\lim_{n \to +\infty} V(T_n) = 0$

$\lim_{n \to +\infty} b_n = 0 \quad (b_n = E(T_n) - \theta)$

on veut montrer que $\lim_{n \to +\infty} E\left((T_n - \theta)^2\right) = 0$

$E\left((T_n - \theta)^2\right) = E\left(T_n^2 - 2T_n\theta + \theta^2\right)$
$= E(T_n^2) - 2\theta E(T_n) + \theta^2$
$= E(T_n^2) - 2E^2(T_n) + E^2(T_n) + E^2(T_n) - 2\theta E(T_n) + \theta^2$
$= E\left(\left(T_n - E(T_n)\right)^2\right) + \left(E(T_n) - \theta\right)^2$
$= V(T_n) + b_n^2$

$\lim_{n \to +\infty} E\left((T_n - \theta)^2\right)$
$= \lim_{n \to +\infty} \left(V(T_n) + b_n^2\right)$
$= \lim_{n \to +\infty} V(T_n) + \lim_{n \to +\infty} b_n^2 = 0$

Donc T_n converge en moyenne quadratique vers θ.

Scanné avec CamScanner

Ex3:

1) * Le modèle d'inférence :

- population : Ado, agés de 14/15 ans et scolarisé en classe de 3ème dans la région R.
- caractère étudié : Surpoids/non on note p la prévalence de surpoids p est un paramètre de valeur inconnue $0 \leq p \leq 1$
- Echantillon prélevé : Taille 145 supposé tiré au hasard avec un taux de sondage $< 10\%$.
- La variable de modèle :

Z définie par $\begin{cases} Z = 1 \text{ si surpoids} \\ Z = 0 \text{ si non.} \end{cases}$

On a évidemment $\mathbb{P}(Z=1) = p$

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

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

3) a revoir, dans l'exercice 19.

Ex4:

- T_n est un estimateur sans biais

$E(T_n) = \theta$.

- $\lim_{n \to +\infty} V(T_n) = 0$.

On veut montrer que $\lim_{n \to +\infty} E(|T_n - \theta|^2) = 0$

$V(T_n) = V(T_n - \theta)$

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

or $E(T_n - \theta) = E(T_n) - E(\theta)$

$= \theta - \theta = 0$

$\lim_{n \to +\infty} V(T_n) = \lim_{n \to +\infty} E\left((T_n - \theta)^2\right) = 0$

2)

$\lim_{n \to +\infty} V(T_n) = 0$

$\lim_{n \to +\infty} b_n = 0$. $\left(b_n = E(T_n) - \theta\right)$

On veut montrer que $\lim_{n \to +\infty} E\left((T_n - \theta)^2\right) = 0$

$E\left((T_n - \theta)^2\right) = E\left(T_n^2 - 2T_n\theta + \theta^2\right)$

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

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

$+ E^2(T_n) - 2\theta E(T_n) + \theta^2$.

$= E\left((T_n - E(T_n))^2\right) + \left(E(T_n) - \theta\right)^2$

$= V(T_n) + b_n^2$.

$\lim_{n \to +\infty} E\left((T_n - \theta)^2\right)$

$= \lim_{n \to +\infty} \left(V(T_n) + b_n^2\right)$

$= \lim_{n \to +\infty} V(T_n) + \lim_{n \to +\infty} b_n^2 = 0$.

Donc T_n converge en moyenne quadratique vers θ.

Scanné avec CamScanner

$V(W) > V(T)$

donc on préfère T comme estimateur pour a

Ex 6

1) $E(m) = \frac{1}{2}\left(E(F_1) + E(F_2)\right)$

$= \frac{1}{2}(p+p) = p$.

Donc m est un estimateur sans biais pour p.

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

$= pa + (1-a)p = p$

Donc S est un estimateur sans biais pour p.

2) $a = ?$, s'il soit de variance minimale

$S' = a F_1 + (1-a) F_2$

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

$= a^2 \frac{p(1-p)}{m_1} + \frac{(1-a)^2 p(1-p)}{m_2}$

$= p(1-p)\left[\frac{a^2}{m_1} + \frac{(1-a)^2}{m_2}\right)$

on prend $f(a) = \frac{a^2}{m_1} + \frac{(1-a)^2}{m_2}$

f est dérivable sur $]0,1[$

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

$f'(a) = \frac{2a}{m_1} - \frac{2(1-a)}{m_2}$

$= 2\left(\frac{a}{m_1} - \frac{(1-a)}{m_2}\right)$

donc $f'(a) = 0 \Rightarrow \frac{a}{m_1} - \frac{(1-a)}{m_2} = 0$

$\Rightarrow a m_2 = m_1(1-a)$

$\Rightarrow a(m_1 + m_2) = m_1$

$\Rightarrow a = \frac{m_1}{m_1 + m_2}$.

* le tableau de variations :

a	0		$\frac{m_1}{m_1+m_2}$		1
$f'(a)$	$\|$	$-$	0	$+$	
$f(a)$	$\frac{1}{m_2}$	$\searrow$	$\frac{1}{m_1+m_2}$	$\nearrow$	$\frac{1}{m_1}$

$f\left(\frac{m_1}{m_1+m_2}\right) = \frac{\frac{m_1^2}{(m_1+m_2)^2}}{m_1} + \frac{\left(\frac{m_2}{m_1+m_2}\right)^2}{m_2}$

$= \frac{m_1}{(m_1+m_2)^2} + \frac{m_2}{(m_1+m_2)^2}$

$= \frac{1}{m_1+m_2}$

* montrer que l'estimateur S' est meilleur que m

$V(S') = p(1-p)\left(\frac{a^2}{m_1} + \frac{(1-a)^2}{m_2}\right)$

$V(m) = \frac{1}{4}\left(\frac{p(1-p)}{m_1} + \frac{p(1-p)}{m_2}\right)$.

$= \frac{p(1-p)}{4}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$.

$\frac{a^2}{m_1} + \frac{(1-a)^2}{m_2} \quad \frac{1}{4}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$

$\frac{1}{m_1+m_2} \leq \frac{1}{4}\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$

$\frac{4}{m_1+m_2} \leq \frac{1}{m_1} + \frac{1}{m_2}$

$4\frac{m_1 m_2}{m_1+m_2} \leq m_1 + m_2$

$4 m_1 m_2 \leq (m_1+m_2)^2$; $0 \leq (m_1+m_2$

$V(m) \leq V(S')$

$\Rightarrow S'$ est le meilleur estimateur pour a.

Scanné avec CamScanner

Exs:

1) a = ?

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

- $E(x) = \sum_{i=1}^{3} x_i p_i$

$= (-1)\,a + 0 \times (1-2a) + 1 \times a$

$= 0$

- $V(x) = \sum_{i=1}^{m} x_i^2 p_i - \underbrace{[E(x)]^2}_{0} = 2a$

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

2) on pose $m = \frac{1}{m}\sum_{i=1}^{m} x_i^2$

on cherche α / $T = \alpha m$ soit un estimateur sans biais pour a

- $E(T) = \alpha E(m) = \frac{\alpha}{m}\sum_{i=1}^{m} E(x_i^2)$

$= \frac{\alpha}{m} \cdot m \cdot 2a = 2\alpha a = a$.

$\Rightarrow \boxed{\alpha = \frac{1}{2}}$

$\Rightarrow \boxed{T = \frac{1}{2} m}$

* montrons que T est c.V

c.V en m.q $\Rightarrow$ c.V en Pr.

- $V(T) = \frac{1}{4} V(m) = \frac{1}{4m^2}\sum_{i=1}^{m} V(x_i^2)$

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

$= (a+a) - (2a)^2 = (2a - 4a^2)$

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

$= \frac{a}{2m}(1-2a) \xrightarrow[m \to \infty]{} 0$

$\Rightarrow$ T c.V en m.q vers a, par.

Conséquent T c.V en Pr vers a.

3) $z_i = \ldots \Leftrightarrow \ldots$

$z_i = 0$ sinon.

on pose $y = \sum_{i=1}^{m} z_i$

- $E(y) = ma$; $V(y) = ma(1-a)$

y suit la loi Binomiale $B(m,a)$

- $W = \frac{y}{m}$, estimateur sans biais ?

Comme $E(W) = a$, alors W est un estimateur sans biais.

* La convergence de $W = \frac{y}{m}$

En utilisant la L. c. V

on a: $W \xrightarrow{Pr.} a$ (page 26 du cours d'échantillonnage)

4) Pour l'estimateur $T = \frac{1}{2} m$:

8 observations amènent

$\hat{a} = \frac{1}{2}\left(\frac{37}{150} + \frac{45}{150}\right)$

$\Rightarrow \hat{a} = 0{,}27$

Pour l'estimateur $W = \frac{y}{m}$:

8 observations amènent :

$\hat{a} = \frac{37}{150} = 0{,}24$.

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

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

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

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

$= \frac{2a - 2a^2 - a + 2a^2}{2m} = \frac{a}{2m} > 0$

Scanné avec CamScanner

$= \frac{F_1 + F_2}{2}$

$S = a F_1 + (1-a) F_2$

$m = S' \iff m_1 = m_2$?

"$\Rightarrow$"

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

$\Rightarrow a = \frac{1}{2} \Rightarrow \frac{1}{2} = \frac{m_1}{m_1 + m_2}$

$\Rightarrow 2 m_1 = m_1 + m_2 \Rightarrow m_1 = m_2$

"$\Leftarrow$" $m_1 = m_2 \Rightarrow a = \frac{1}{2} \Rightarrow S' = m$.

Ex 7 :

1°/ un estimateur W par la m.m.v pour λ :

$$L(x_1, x_2, \dots, x_n, \lambda) = \prod_{i=1}^{n} P_\lambda(X = x_i)$$

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

On cherche à maximiser la fonction qui a $\lambda \longrightarrow L(x_1, x_2, \dots, x_n, \lambda)$ ce qui est équivalent à maximiser la fonction $\ln(L(x_1, x_2, \dots, x_n, \lambda))$

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

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

Condition nécessaire :

$$\frac{\partial \ln(L)}{\partial \lambda} = 0 \Rightarrow -n + \frac{n\bar{x}}{\lambda} = 0.$$

$$\Rightarrow \lambda = \bar{x}$$

Condition suffisante :

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

$\Rightarrow$ elle est vérifiée pour tout λ.

Donc, l'estimateur du m.v est la statistique $W = \bar{X}$

2) L'estimateur du m.v est toujours asymptotiquement sans biais et converge en probabilité, de plus $W = \bar{X}$

$E(\bar{X}) = E(X) = \lambda$.

$\Rightarrow$ W est sans biais

W est estimateur "naturel" pour $E(X)$

(LGN : $\bar{X} \xrightarrow{P} E(X) = \lambda$).

3)

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

3) une estimation pour λ fournie par l'estimateur W et les observations est : $\hat{\lambda} = \frac{16 \times 2 + 15 \times 2 + 5 \times 3 + 6 \times 4}{52}$

$$\boxed{\hat{\lambda} = 1{,}5}$$

$P(X=0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\hat{\lambda}} = e^{-1{,}5}$

$$\boxed{P(X=0) = 0{,}22}$$

La fréquence observée pour l'estimation $(X=0)$ est : $\frac{12}{52} = 0{,}23$

on remarque que notre estimation est précise relativement.

Scanné avec CamScanner

Ex9:

1) f est une densité de probabilité ssi : • f est positive
• f est continue p.p.

$$\int_{\mathbb{R}} f(x)\,dx = 1$$

$$\int_{\mathbb{R}} f(x)\,dx = \int_{-1}^{0} a\,dx + \int_{0}^{1} (1-a)\,dx$$

$$= [ax]_{-1}^{0} + [(1-a)x]_{0}^{1}$$

$$= +a + (1-a) = 1.$$

$\Rightarrow f$ est une densité de probabilité

$$E(X) = \int_{\mathbb{R}} x f(x)\,dx$$

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

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

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

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

$$V(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \int_{\mathbb{R}} x^2 f(x)\,dx$$

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

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

$$= \frac{a}{3} + \frac{(1-a)}{3} = \frac{1}{3}$$

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

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

si $x \in [-1, 0]$

$$F(x) = \int_{-1}^{x} a\,dt = [at]_{-1}^{x} = a(x+1)$$

$$\Rightarrow P(X<0) = F(0)$$

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

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

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

* calculons un estimateur T pour a par la méthode des moments

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

avec $\varphi(a) = \frac{1}{2} - a$

φ est bijective sur $]0,1[$, donc inversible et $\varphi^{-1}(a) = \frac{1}{2} - a$

Donc l'estimateur T pour a par la méthode des moments est

$$T = \frac{1}{2} - \bar{X}$$

on a : $E(T) = \frac{1}{2} - E(\bar{X}) = \frac{1}{2} - E(X)$

$= a$

Par suite T est un estimateur sans biais et c.v en proba vers a.

3) on note y l'effectif empirique de l'événement $(X<0)$ dans un échantillon $(x_1, x_2, \dots, x_n)$ de X

• $Y \sim B(n, a)$

si on a observé la valeur $(x_1, x_2, \dots, x_n)$ pour $(X_1, \dots, X_n)$ la vraisemblance des observations $x_1, x_2, \dots, x_n$ sachant a est :

Scanné avec CamScanner

$(x_1, x_2, \dots, x_m, a) = \prod_{i=1}^{m} f(x_i, a)$

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

et

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

$\Rightarrow L(x_1, x_2, \dots, x_m, a) = a^{y} (1-a)^{m-y}$

où y est l'effectif de l'evenement $(x<0)$

c'est à dire $y = mb$ d'observations <0 dans $x_1, x_2, \dots, x_m$.

4) calculons aussi un estimateur W pour a par la méthode de maximum de vraissemblance:

on cherche à maximiser la fonction $a \longrightarrow L(x_1, x_2, \dots, x_n, a)$ on encore la fonction :

$a \longrightarrow \ln(L(x_1, x_2, \dots, x_n, a))$

$\ln(L(x_1, x_2, \dots, x_m, a))$

$= y \ln(a) + (m-y) \ln(1-a)$

* C.N : $\frac{\partial \ln(L)}{\partial a} = 0$.

$\Rightarrow \frac{y}{a} - \frac{m-y}{1-a} = 0 \Rightarrow \frac{y}{a} = \frac{m-y}{1-a}$

$\Rightarrow (1-a)y = (m-y)a \Rightarrow$

$\Rightarrow y - ay = ma - ya$

$\Rightarrow \boxed{a = \frac{y}{m}}$

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

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

Donc notre estimateur W obtenu par la méthode du M.V est $W = \frac{Y}{m}$

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

* Comparaison de deux estimateurs T et W :

$\bar{T} = \frac{1}{2} - \bar{X}$

$\Rightarrow V(T) = V(\bar{X}) = \frac{V(X)}{m}$

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

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

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

$= \frac{1}{3} - \frac{1}{4} = \frac{1}{12} > 0$

$\Rightarrow V(T) > V(W)$ donc W est le meilleur estimateur pour a

Proposition :

si T est un estimateur sans biais obtenu par la méthode du M.V il est le meilleur parmi les estimateurs sans biais.

Exercice :

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

calculons un estimateur A par méthode des moments

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

$\varphi(A) = \frac{1+a}{2}$

Scanné avec CamScanner

φ est bijective sur $]0,+\infty[$ donc $\varphi(a)$ admet une valeur bien $\Rightarrow \varphi(a) = \gamma$

$\Rightarrow \frac{1+a}{2} = \gamma \Rightarrow a = 2\gamma - 1$

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

donc notre estimateur est $A = 2\bar{X} - 1$

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

$E(\bar{X}) = E(T) = \frac{1+a}{2}$

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

$\Rightarrow$ notre estimateur A est sans biais $A \xrightarrow[LGN]{cvp} a$

3) calcul d'un estimateur W par la méthode de MV:

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

$= \prod_{i=1}^{m} \left(\frac{1}{a-1}\right) = \left(\frac{1}{a-1}\right)^m$

L est $\searrow$ en $a \Rightarrow$ donc L est maximal lorsque a est minimal

$a = \min(x_1, x_2, \dots, x_m)$

par conséquent, notre estimateur $W = \text{minimum}(x_1, x_2, \dots, x_m)$

La loi de W

$P(W < u) = P(\text{minimum}(x_1, \dots x_2, x_n) < u)$

$\Rightarrow P(\{x_1 < u\} \cap \{x_2 < u\} \dots \{x_m < u\}) = \prod_{i=1}^{m} P(x_i < u)$

$= [P(x_i < u)]^m = \left(\frac{u-1}{a-1}\right)^m$

ou $P(x_i < u) = \int_1^u \frac{1}{a-1}\,dt = \frac{u}{a-1}$

* la fonction densité g

$g(u) = G'(u) = \frac{m}{a-1}\left(\frac{u-1}{a-1}\right)^{m-1}, a \in [1,a]$.

Donc

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

$E(W) = \int_1^a \frac{mu}{a-1}\left(\frac{u-1}{a-1}\right)^{m-1} du$

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

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

$= u\left(\frac{u-1}{a-1}\right)^m \Big|_1^a - \frac{a-1}{m+1}$

$= a - \left(\frac{a-1}{m+1}\right) \xrightarrow{m \to +\infty} a$

W est un estimateur asymptotiquement sans biais

$E\left((A-a)^2\right) = \text{Var}(A) = V(2\bar{X} - 1)$

$= 4V(\bar{X}) = \frac{V(T)}{m} = \frac{(a-1)^2}{12m}$

$= \frac{4(a-1)^2}{12m} = \frac{(a-1)^2}{3m}$

$E\left((W-a)^2\right) = V(W) + b_m^2$

$V(W) = E(W^2) - [E(W)]^2$

$E(W^2) = \frac{m}{a-1}\int_1^a u^2\left(\frac{u-1}{a-1}\right)^{m-1} du$

$u = u^2 \qquad u' = 2u$

$v' = \left(\frac{u-1}{a-1}\right)^{m-1} \qquad v = \frac{a-1}{m}\left(\frac{u-1}{a-1}\right)^m$.

Scanné avec CamScanner

$$E(W^2) = \frac{m}{a-1}\left[\frac{u^2(a-1)}{m}\left(\frac{u-1}{a-1}\right)^m\right]_1^a - \frac{2m}{a-1}\int_1^a \frac{u(a-1)}{m}\left(\frac{u-1}{a-1}\right)^m$$

$u = u \longrightarrow u' = 1$

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

$$\Rightarrow E(W^2) = a^2 - 2\left[\frac{u(a-1)}{m+1}\left(\frac{u-1}{a-1}\right)^{m+1}\right]_1^a + 2\int_1^a \frac{a-1}{m+1}\left(\frac{u-1}{a-1}\right)^{m+1} du$$

$$= a^2 - 2\left(\frac{a(a-1)}{m+1}\right) + \left[2\frac{(a-1)^2}{(m+1)(m+2)}\left(\frac{u-1}{a-1}\right)^{m+2}\right]_1^a$$

$$= a^2 - 2\frac{a(a-1)}{m+1} + \frac{2(a-1)^2}{(m+1)(m+2)}$$

$$\Rightarrow V(W) = a^2 - 2\frac{a(a-1)}{m+1} + \frac{2(a-1)^2}{(m+1)(m+2)} - \left(a - \frac{(a-1)}{m+1}\right)^2$$

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

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

$$\Rightarrow E\left((W-a)^2\right) = \frac{2(a-1)^2}{(m+1)(m+2)} - \left(\frac{a-1}{m+1}\right)^2 + \left(\frac{a-1}{m+1}\right)^2$$

$$= \frac{2(a-1)^2}{(m+1)(m+2)}$$

$$\frac{1}{3m} - \frac{2}{(m+1)(m+2)}$$

$$= \frac{(m+1)(m+2) - 6m}{3m(m+1)(m+2)}$$

$$= \frac{m^2 + 2m + m + 2 - 6m}{3m(m+1)(m+2)}$$

$$= \frac{m^2 - 3m + 2}{3m(m+1)(m+2)}$$

$$= \frac{(m-1)(m-2)}{3m(m+1)(m+2)} > 0 \quad \forall m > 2$$

$\Rightarrow E\left((A-a)^2\right) > E\left((W-a)^2\right)$

$\Rightarrow$ On préfère W.

Scanné avec CamScanner

"Echantillonnage"

*Moyenne empirique :

$$\bar{x} = \frac{x_1 + \dots + x_m}{m} = \frac{1}{m} \sum_{i=1}^{m} x_i$$

- propriétés :

$E(\bar{x}) = m \quad ; \quad V(\bar{x}) = \frac{V}{m}$

preuve :

- $E(\bar{x}) = E\left(\frac{1}{m} \sum_{i=1}^{m} x_i\right) = \frac{1}{m} E\left(\sum_{i=1}^{m} x_i\right) = \frac{1}{m} \sum_{i=1}^{m} E(x_i)$

$= \frac{1}{m} (m - 1 + 1)\, m = m.$

- $V(\bar{x}) = V\left(\frac{1}{m} \sum_{i=1}^{m} x_i\right) = \frac{1}{m^2} V\left(\sum_{i=1}^{m} x_i\right)$

$= \frac{1}{m^2} \sum_{i=1}^{m} V(x_i) = \frac{1}{m^2} (m - 1 + 1) \cdot V = \frac{V}{m}.$

*Variance empirique :

$$s^2 = \frac{1}{m} \sum_{i=1}^{m} (x_i - \bar{x})^2 = \frac{1}{m} \sum_{i=1}^{m} x_i^2 - (\bar{x})^2$$

*Variance empirique corrigée :

$$s'^2 = \frac{1}{m-1} \sum_{i=1}^{m} (x_i - \bar{x})^2 = \frac{m}{m-1} s^2.$$

- propriétés :

$E(s^2) = \frac{(m-1)V}{m} \quad ; \quad E(s'^2) = V.$

preuve :

- $E(s^2) = E\left(\frac{1}{m} \sum_{i=1}^{m} x_i^2 - (\bar{x})^2\right) = \frac{1}{m} \sum_{i=1}^{m} E(x_i^2) - E(\bar{x}^2)$

on $E(x_i) = 0$ et $E(x_i^2) = V(x_i) = V.$

x_i sont centrées

$\Rightarrow E(\bar{x}) = 0$ et $E(\bar{x}^2) = \frac{V}{m}.$

donc

$E(s^2) = \frac{1}{m} \left((V \dots) (m - 1 + 1) - \frac{V}{m} = V - \frac{V}{m} \right).$

$E(s'^2) = \frac{m}{m-1} s^2 = \frac{m}{m-1} \left(\frac{(m-1)V}{m} \right) = V.$

Scanné avec CamScanner

* Fréquence empirique :

• $F = \frac{Y}{m}$: fréquence empirique, elle représente la proportion dans l'échantillon de taille m.

• $Y = \underbrace{X_1 + \dots + X_m}$ sont des v.a de Bernoulli de paramètre p et indépendantes ; X_i prend $\begin{cases} 1 \\ on \\ 0 \end{cases}$

• $Y = mF$: suit une loi binomiale de paramètre m et p.

- propriétés :

$$E(F) = p \quad ; \quad V(F) = \frac{1}{m} p(1-p)$$

* Si $m \geq 50$:

• Convergence en probabilité d'une suite de v.a. :

Soit (X_m) suite de v.a et un nbre réel a.

• $X_m \xrightarrow[m \to +\infty]{\text{converge en probabilité } (P_1)} a$ ou $\lim_{m \to +\infty} X_m = a \ (P_1)$

ssi $\lim_{m \to +\infty} P(|X_m - a| > \varepsilon) = 0$

ou

$\lim_{m \to +\infty} P(|X_m - a| < \varepsilon) = 1$.

• si $\begin{cases} E(X_m) \xrightarrow[m \to +\infty]{} a \\ \text{et} \\ V(X_m) \xrightarrow[m \to +\infty]{} 0 \end{cases}$ alors $X_m \xrightarrow[m \to +\infty]{P_1} a$

• $X_m \xrightarrow[m \to +\infty]{P_1} X$ ssi $(X_m - X) \xrightarrow[m \to +\infty]{P_1} 0$

• Th de Slutsky :

→ conséquence 1 :

$$\lim_{m \to +\infty} (Y_m + Z_m) = Y + Z \ (P_1)$$

$$\lim_{m \to +\infty} (Y_m Z_m) = Y.Z \ (P_1)$$

$$\lim_{m \to +\infty} \frac{Y_m}{Z_m} = \frac{Y}{Z} \ (P_1).$$

Scanné avec CamScanner

→ on obtient :

$\lim_{m\to+\infty} g(X_m) = g(x)$ (P_1) avec g est une fct continue d'une seule intervalle

- La loi des grands nombres (LGN) :

si X v.a telle que $m = E(X)$ existe, alors pour toute suite infinie $(X_m)_{m\in\mathbb{N}}$ de v.a deux à deux indépendantes et de même loi que X :

$$\overline{X_m} = \frac{1}{m}\sum_{i=1}^{m} X_i \xrightarrow[m\to+\infty]{(P_1)} m = E(X)$$

- convergence en probabilité de la variance empirique et de l'écart type empirique :

$S^2 \xrightarrow[m\to+\infty]{(P_1)} V$; $S'^2 \xrightarrow[m\to+\infty]{(P_1)} V$

preuve :

on a $\lim_{m\to+\infty} \frac{1}{m}\sum_{i=1}^{m} X_i^2 = E(X^2)$ (P_1) (d'après LGN)

et $\lim_{m\to+\infty} (\overline{X_m})^2 = m^2 = (E(X))^2$ (P_1) (d'après ... corollaire)

donc $S^2 = \frac{1}{m}\sum_{i=1}^{m} X_i^2 - (\overline{X_m})^2 \xrightarrow[m\to+\infty]{(P_1)} V = E(X^2) - (E(X))^2$

de même $S'^2 = \frac{1}{1-\frac{1}{m}} S \xrightarrow[m\to+\infty]{(P_1)} V$

* Convergence en moyenne quadratique :

- $X_m \xrightarrow[m\to+\infty]{\text{converge en moyenne quadratique (mq)}} X$ ssi

ssi $\left(E\left(X_m - X\right)^2\right) \xrightarrow[m\to+\infty]{} 0$

- si $E(X_m) \xrightarrow[m\to+\infty]{} a$ et $V(X_m) \xrightarrow[m\to+\infty]{} 0$ } alors $X_m \xrightarrow[m\to+\infty]{mq} a$.

- $X_m \xrightarrow[m\to+\infty]{mq} X$ ssi $(X_m - X) \xrightarrow[m\to+\infty]{mq} 0$

Bq :

$X_m \xrightarrow[m\to+\infty]{mq} X \Rightarrow X_m \xrightarrow[m\to+\infty]{(P_1)} X$

%

Scanné avec CamScanner

* Convergence en loi :

•) Soit (X_n) une suite de variables aléatoires, on note G_n la fonction de répartition de X_n, on dit que X_n converge en loi vers X de fonction de répartition G ssi $\forall x \in \mathbb{R}$, G continue et $G_n \xrightarrow[n \to +\infty]{} G$ dans $\mathbb{R}$ on note $X_n \xrightarrow[n \to +\infty]{\mathcal{L}} X$

Rq :

$$X_n \xrightarrow[n \to +\infty]{mq} X \Rightarrow X_n \xrightarrow[n \to +\infty]{P} X \Rightarrow X_n \xrightarrow[n \to +\infty]{\mathcal{L}} X$$

•) convergence en loi de la moyenne empirique :

→ Th. centrale limite :

il s'applique si la taille n de l'échantillon est assez grande si n est grande ($n \geq 50$), la loi de $U_n = \dfrac{\overline{X_n} - m}{\frac{\sigma}{\sqrt{n}}}$ est assimilée ou n est petite

à la loi $\mathcal{N}(0,1)$

Donc le TcL ; La moyenne empirique $\overline{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i$ suit la loi Normale $\mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$ où $m = E(X)$, $\sigma^2 = V(X)$

•) convergence de la loi binomiale vers la loi normale :

→ Th. de Moivre - Laplace :

soit (X_n) une suite de variables des lois binomiales $B(n,p)$ alors quand $n \to +\infty$, La suite $U_n = \dfrac{X_n - np}{\sqrt{np(1-p)}}$

converge en loi vers la loi normale $\mathcal{N}(0,1)$

d'où, la loi binomiale $B(n,p)$ est assimilée par une loi normale $\mathcal{N}\left(np, \sqrt{np(1-p)}\right)$

→ seconde version du TH. centrale limite,

soit X_i^* la v.a centrée réduite associée à X_i, tel que $X_i^* = \dfrac{X_i - m}{\sigma}$ alors $T_n = \dfrac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i^*$.

Converge en loi vers $\mathcal{N}(0,1)$

Scanné avec CamScanner

Donc le TCL, $S_n = \sum_{i=1}^{n} x_i$ suit la loi normale $N(nm, \sigma\sqrt{n})$

* si $n < 50$:

X désigne une v.a normale, $E(X) = m$ et σ

- densité de X : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-m)^2}{2\sigma^2}\right)}$

- $T = \frac{X-m}{\sigma}$ suit la loi normale centrée réduite $N(0,1)$ de densité $f_0(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t}{2}}$ et fonction de répartition $F_0(t) = \int_{-\infty}^{t} f_0(u)\,du$ et $F_0(t) = 1 - F_0(t)$

- si $(X_1, \dots, X_n)$ est normal, la moyenne empirique suit une loi normale

c.à.d $\overline{X_n} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right)$

- soit $(X_1, \dots, X_n)$ un échantillon normal,

on

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 \quad \text{et} \quad s'^2 = \frac{n}{n-1} S^2$$

alors

$$\frac{n S^2}{\sigma^2} = \frac{(n-1) s'^2}{\sigma^2} \sim \text{La loi } \chi^2_{n-1}.$$

Les remarques par les exercices

- La fonction densité atteint son maximum au point m donc le plus petit intervalle contenant $\bar{x}$ sera centré en m

$I = [m - R ; m + R]$

Scanné avec CamScanner

TD echantillonnage

Exercice n°18

loi de X

Valeurs de x	1	2	3
Probabilité	0,3	0,5	0,2

x_1	x_2	Probabilité	Moy.
1	1	0,09	1
1	2	0,15	1,5
1	3	0,06	2
2	1	0,15	1,5
2	2	0,25	2
2	3	0,1	2,5
3	1	0,06	2
3	2	0,1	2,5
3	3	0,04	3

Loi de M

Valeurs de M	1	1,5	2	2,5	3
Probabilité	0,09	0,3	0,37	0,2	0,04

2)

$E(X) = \sum_{k=1}^{3} x_i p_i = 1\times0,3 + 2\times0,5 + 3\times0,2 = 1,9$

$E(\bar{X}) = \sum_{k=1}^{5} m_i p_i$

$= 1\times0,09 + 1,5\times0,3 + 2\times0,37 + 2,5\times0,2 + 3\times0,04 = 1,9$

$V(X) = E(X^2) - E^2(X)$

$E(X^2) = \sum_{i=}^{5} x_i^2 p_i$

$= 1\times0,3 + 4\times0,5 + 9\times0,2 = 4,1$

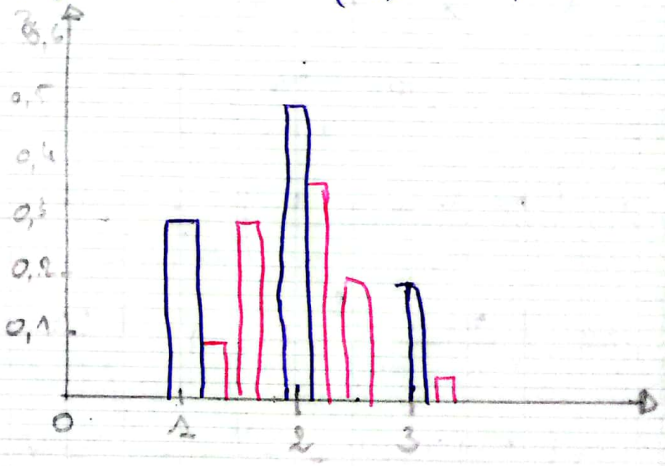
$V(X) = 4,2 - (1,9)^2 = 0,49$

$V(\bar{X}) = E(\bar{X}^2) - E^2(\bar{X})$

$E(\bar{X}^2) = \sum_{K=1}^{5} m_K^2 p_K$

$= 1\times0,09 + (1,5)^2\times0,3 + 4\times0,37 + (2,5)^2\times0,2 + 9\times0,04 = 3,855$

$V(\bar{X}) = 3,855 - (1,9)^2 = 0,245.$



- Loi de X
- Loi de M

On remarque que $E(M) = E(X)$ et $V(M) \leq V(X)$ Par loi de M on a mais dispersée que par loi X mais ils ont le même moyen

Scanné avec CamScanner

- Loi discrète
- Loi continue
- Echantillonnage.

Ex 2:

$\overline{X_1}$ et $\overline{X_2}$ sont les moyennes empiriques $\Rightarrow E(\overline{X_1} - \overline{X_2}) = E(\overline{X_1}) - E(\overline{X_2}) = 0$

$V(\overline{X_1} - \overline{X_2}) = V(\overline{X_1}) + V(-\overline{X_2}) = V(\overline{X_1}) + (-1)^2 V(\overline{X_2})$

$= V(\overline{X_1}) + V(\overline{X_2}) = \frac{V}{n_1} + \frac{V}{n_2} = V\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$

$E\left(\frac{\overline{X_1} + \overline{X_2}}{2}\right) = \frac{1}{2} E(\overline{X_1}) + \frac{1}{2} E(\overline{X_2}) = \frac{1}{2} m + \frac{1}{2} m = m.$

$E\left(\frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}\right) = \frac{n_1}{n_1 + n_2} E(\overline{X_1}) + \frac{n_2}{n_1 + n_2} E(\overline{X_2})$

$= \frac{n_1}{n_1 + n_2} m + \frac{n_2}{n_1 + n_2} m = m\left(\frac{n_1 + n_2}{n_1 + n_2}\right) = m.$

$\Rightarrow E\left(\frac{\overline{X_1} + \overline{X_2}}{2}\right) = E\left(\frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}\right)$

$V\left(\frac{\overline{X_1} + \overline{X_2}}{2}\right) = \frac{1}{4} V(\overline{X_1}) + \frac{1}{4} V(\overline{X_2}) = \frac{V}{4n_1} + \frac{V}{4n_2}$

$= V\left(\frac{n_1 + n_2}{4 n_1 n_2}\right)$

$V\left(\frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}\right) = \left(\frac{n_1}{n_1 + n_2}\right)^2 V(\overline{X_1}) + \left(\frac{n_2}{n_1 + n_2}\right)^2 V(\overline{X_2})$

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

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

$= V\left(\frac{(n_1 + n_2)^2 - 4 n_1 n_2}{4 n_1 n_2 (n_1 + n_2)}\right) = V\left(\frac{(n_1 - n_2)^2}{4 n_1 n_2 (n_1 + n_2)}\right)$

$\Rightarrow \frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}$ a la plus petite variance.

Scanné avec CamScanner

Ex 3 :

1) Les hypothèses satisfaites par les variables $x_1, \dots, x_n$

- les v.a sont I.I.D (indépendantes et identiquement distribuées)

$E(\bar{x}) = E(x) = m$

$V(\bar{x}) = \frac{V(x)}{n}$

la loi de $\bar{x}$ pour $n = 9$.

$\bar{x}$ suit la loi normale $\mathcal{N}\left(m, \sqrt{\frac{V(x)}{n}}\right) \Rightarrow \bar{x}$ suit la loi normale $\mathcal{N}(2000, 110)$

La fonction densité atteint son maximum au point m donc notre intervalle sera centré en m et il sera sous la forme

$I = [m - r, m + r]$ donc on va chercher r de telle sorte que

$P(\bar{x} \in I) = 0,95$

$P(\bar{x} \in I) = P(|\bar{x} - m| \leq r) = 0,9.$

$\bar{x}$ suit la loi normale $\mathcal{N}(2000, 110)$

donc $T = \frac{\bar{x} - m}{\sigma}$ suit la loi normale centré reduite $\mathcal{N}(0,1)$

$\Rightarrow P\left(|T| \leq \frac{r}{\sigma}\right) = 0,9$

$\Rightarrow P\left(|T| \leq \frac{r}{110}\right) = 0,9 \Rightarrow 2F_0\left(\frac{r}{110}\right) - 1 = 0,9$

$\Rightarrow F_0\left(\frac{r}{110}\right) = 0,95 \Rightarrow \frac{r}{110} = F_0^{-1}(0,95)$

$\Rightarrow r = 110\, F_0^{-1}(0,95) = 110 \times 1,6449 = 181$

$\Rightarrow r = 181$

3) on cherche n minimale de sorte que le rayon de cet intervalle n'excède pas $\frac{r}{2}$.

$P\left(|\bar{x} - m| \leq \frac{r}{2}\right) = 0,9$

$\bar{x}$ suit la loi normale $\mathcal{N}\left(2000, \frac{330}{\sqrt{n}}\right)$

donc $T = \frac{\bar{x} - m}{\frac{330}{\sqrt{n}}} = \frac{\bar{x} - 2000}{\frac{330}{\sqrt{n}}}$ suit la loi $\mathcal{N}(0,1)$.

Scanné avec CamScanner

~~$\Rightarrow 2F_0\left(\frac{\pi}{\frac{660}{\sqrt{n}}}\right) - 1 = 0,9$~~

$$P\left(|T| \leq \frac{\pi}{\frac{660}{\sqrt{n}}}\right) = 0,9 \Rightarrow 2F_0\left(\frac{\pi}{\frac{660}{\sqrt{n}}}\right) - 1 = 0,9$$

$$\Rightarrow F_0\left(\frac{\pi}{\frac{660}{\sqrt{n}}}\right) = 0,95.$$

$$\Rightarrow \frac{\pi}{\frac{660}{\sqrt{n}}} = F_0^{-1}(0,95) \Rightarrow \sqrt{n} = \frac{660}{\pi} F_0^{-1}(0,95)$$

$$\Rightarrow n = \left(\frac{660}{180} \times 1,6449\right)^2 \Rightarrow n = 36.$$

Rq:

Ex4:

1) $Y_n = \frac{X_n}{n}$ est la fréquence empirique elle représente la proportion dans l'échantillon de taille n

* $X_n \sim$ la loi binomiale de paramètres $p = 0,1$ avec $E(X_n) = np$. $V(X_n) = np(1-p)$ et $p = 0,1$.

* $E(Y_n) = E\left(\frac{X_n}{n}\right) = \frac{1}{n} E(X_n) = \frac{1}{n} \cdot n \cdot p = p = 0,1.$

$$V(Y_n) = V\left(\frac{X_n}{n}\right) = \frac{1}{n^2} V(X_n) = \frac{1}{n^2} n \cdot p(1-p) = \frac{0,09}{n}$$

2) Pour $n = 10 < 50$

$E(Y_{10}) = 0,1$

$V(Y_{10}) = 0,009$

Valeurs de Y_{10} $\quad Y_{10} = \frac{X_{10}}{10} = \frac{k}{10}$, où $k \in \{0, \dots, 10\}$.

Scanné avec CamScanner

$P\left(Y_{10} = \frac{k}{10}\right) = P\left(\frac{X_{10}}{10} = \frac{k}{10}\right) = P(X_{10} = k) = C_{10}^{k} (0,1)^{k} (0,9)^{10-k}$

Valeurs	0	0,1	0,2	0,3	0,4	0,5	0,6
Probabilité	0,3487	0,3874	0,1937	0,0574	0,0112	0,0015	0,0001

$P(Y_n > 0,2) = 1 - P(Y_n \leq 0,2) = 1 - P(X_n \leq 2)$

$= 1 - \sum_{k=0}^{2} C_{10}^{k} (0,1)^{k} \times (0,9)^{10-k} = 0,07.$

autre méthode :

$P(Y_n > 0,2) = 1 - P(Y_n \leq 0,2) = 1 - (0,3487 + 0,3874 + 0,1937)$

$= 0,07.$

3) $P(Y_n > 0,2) = P(X_n > n \times 0,2)$ et $X_n \sim B(n, 0,1)$

n	$n \times 0,2$	$P(Y_n > 0,2) = P(X_n > n \times 0,2) = 1 - F(n \times 0,2$
20	4	$1 - F(4) = 1 - 0,9568 = 0,0432.$
40	8	$1 - F(8) = 1 - 0,9845 = 0,0155.$
50	10	$1 - F(10) = 1 - 0,9906 = 0,0094.$

4) On peut remarquer que la probabilité diminue quand n augmente et ceci est normale parce que la loi de Y_n s'approche de 0,1 et sa variance diminue quand n augmente.

Exo :

Y_n : nombre de boules blanches. $F = \frac{Y_n}{n}$.

* Y_n suit la loi binomiale de paramètre n et p.

$P(F > 0,5) = P\left(\frac{Y_n}{n} > 0,5\right) = P\left(\frac{Y_n}{n} > 0,5n\right)$

$= 1 - P(Y_n \leq 0,5n)$

Scanné avec CamScanner

pour $n = 20$

$P(F > 0,5) = 1 - P(Y_n \leq 10) = 1 - 0,8725 = 0,1275$

2) pour $n = 150$.

Y_{150} suit la loi Binomiale $B(150, 0,4)$

$E(Y_{150}) = 150 \times 0,4 = 60$

$V(Y_{150}) = np(1-p) = 150 \times 0,4 \times 0,6 = 36$

$\sigma(Y_{150}) = \sqrt{V(Y_{150})} = 6$

pour n assez grand on peut approximer la loi binomiale $B(150, 0,4)$ par par une loi normale $N(n \times p, \sqrt{np(1-p)})$

$N(60, 6)$

$P(F > 0,5) = 1 - P(Y_n \leq 75)$

$T = \frac{Y_n - m}{\sigma}$ suit la loi normale centrée reduit $N(0,1)$

$P(Y_n \leq 75) = P\left(T \leq \frac{75-60}{6}\right) = P(T \leq 2,5) = F_0(2,5) = 0,9938.$

$P(F > 0,5) = 1 - 0,9938 = 0,0062.$

Autre méthode (TCL)

$F = \sum \frac{z_i}{n}$ ou z_i sont des v.a qui suivent la loi de bernoulli de parametre 0,4.

avec $z_i = 0$ si la boule obtenu est blanche et $z_i = 1$ sinon

les z_i sont indépendants (tirage avec remise)

On en appliquant le TCL, $F = \frac{\sum_{i=0}^{150} z_i}{150}$ sont de façon approché, la loi Normale $(0,4, 0,04)$

$P(F_{150} > 0,5) = 1 - 0,9938 = 0,0062.$

$B(150, 0,4)$

Scanné avec CamScanner

Ex6: deux échantillons de taille n → F_1 : la proportion de boules blanches dans E_1
F_2 : prop de boules blanches dans E_2.

1) $n = 75$

$P(|F_1 - F_2| > 0,1)$

$F_1 = \frac{Y_1}{n}$; $F_2 = \frac{Y_2}{n}$

Y_1 et $Y_2 \sim B(75, 0,4)$ ⇒ d'après le th de Moivre - Laplace : on peut approximer $B(75, 0,4)$ en loi de la normale :

$N(np, \sqrt{np(1-p)})$ qui est $N(30, 3\sqrt{2})$

soit $Y = Y_1 - Y_2$

$E(Y) = E(Y_1) - E(Y_2) = 0$

$V(Y) = V(Y_1) + V(Y_2) = 2 \times (3\sqrt{2})^2 = 36$.

⇒ $Y \sim N(0, 6)$

soit $T = \frac{Y - m}{\sigma} = \frac{Y}{6} \sim N(0,1)$

$P(|Y| > 7,5) = P(|T| > 1,25) = 2(1 - F_0(1,25))$

$= 2(1 - 0,8944) = 2 \times 0,1056 = 0,2112$.

2)

$P\left(\frac{|Y|}{n} > 0,1\right) \leqslant 0,1 \iff P(|Y| > n \times 0,1) \leqslant 0,1$

$\iff P\left(|T| > \frac{n \times 0,1}{6}\right) \leqslant 0,1 \iff 2\left(1 - F_0\left(\frac{n \times 0,1}{6}\right)\right) \leqslant 0,1$

$\iff 1 - F_0\left(\frac{n \times 0,1}{6}\right) \leqslant 0,05 \iff -F_0\left(\frac{n \times 0,1}{6}\right) \leqslant -0,95$

$\iff F_0\left(\frac{n \times 0,1}{6}\right) \geqslant 0,95$

$\Rightarrow \frac{n \times 0,1}{6} \geqslant 1,6449 \Rightarrow n = \frac{1,6449 \times 6}{0,1} = 99$

Scanné avec CamScanner

On cherche m tq $P(|F_1 - F_2| > 0{,}1) < 0{,}1$

$F_i \rightsquigarrow N\left(0{,}4, \sqrt{\frac{0{,}4 \times 0{,}6}{m}}\right)$

$F_1 - F_2 \rightsquigarrow N\left(0, \sqrt{\frac{2 \times 0{,}4 \times 0{,}6}{m}}\right)$

$\Rightarrow F_1 - F_2 \rightsquigarrow N\left(0, \sqrt{\frac{0{,}48}{m}}\right)$

soit $|T| = \frac{|F_1 - F_2|}{\sqrt{\frac{0{,}48}{m}}} \rightsquigarrow N(0,1) \Rightarrow P(|F_1 - F_2| > 0{,}1)$

$= P\left(|T| > \frac{0{,}1}{\sqrt{\frac{0{,}48}{m}}}\right)$

$= 2\left(1 - F_0\left(\frac{0{,}1 \times \sqrt{m}}{\sqrt{0{,}48}}\right)\right) < 0{,}1$

$\Leftrightarrow F_0\left(\frac{0{,}1\sqrt{m}}{\sqrt{0{,}48}}\right) \geq 0{,}95 \Rightarrow \frac{0{,}1 \times \sqrt{m}}{\sqrt{0{,}48}} \geq 1{,}6449$

$\Rightarrow \sqrt{m} \geq \frac{1{,}6449 \times \sqrt{0{,}48}}{0{,}1} \Rightarrow m \geq \left(\frac{1{,}6449 \times \sqrt{0{,}48}}{0{,}1}\right)^2$

$\Rightarrow m \geq 129{,}8. \Rightarrow m = 130.$

Ex 7 :

1) $E(U) = \sum_{i=1}^{m} E\left(\frac{x_i - m}{\sigma}\right) = \sum_{i=1}^{m} \frac{E(x_i) - m}{\sigma} = 0.$

$V(U) = V\left(\sum_{i=1}^{m} \frac{x_i - m}{\sigma}\right) = \sum_{i=1}^{m} \frac{1}{\sigma^2} V(x_i) = \frac{m\sigma^2}{\sigma^2} = m.$

soit $T = \frac{x_i - m}{\sigma} \rightsquigarrow N(0,1)$

$\Rightarrow V \rightsquigarrow$ Khi-deux de m de gré de liberté $= m.$

$\Rightarrow E(V) = m$ et $V(V) = 2m.$

2) $U \rightsquigarrow N(0, \sqrt{m})$

$V \rightsquigarrow$ Khi deux de m de gré de liberté $= m.$

$W \rightsquigarrow \chi^2_{(m-1)}$ loi de Khi-deux de nbre de degré de liberté $m-1$).

W : donnée.

Scanné avec CamScanner

Ex 9 :

1) f est une densité de probabilité $\cdot k \geq 0$ $\cdot f$ est continue p.p.

$\cdot \int_{\mathbb{R}} f(t)\,dt = 1$ donc $\int_0^a k t\,dt = \left[\frac{kt^2}{2}\right]_0^a = \frac{ka^2}{2} = 1$.

$\Rightarrow \boxed{k = \frac{2}{a^2}}$

$E(T) = \int_{\mathbb{R}} t f(t)\,dt = \int_0^a k t^2\,dt = \left[\frac{kt^3}{3}\right]_0^a = \frac{ka^3}{3} = \frac{2}{3}a.$

$V(T) = E(T^2) - E(T)^2$ or $E(T^2) = \int_0^a k t^3\,dt = \left[\frac{kt^4}{4}\right]_0^a$

$= \frac{k a^4}{4} = \frac{a^4}{4} \times \frac{2}{a^2} = \frac{a^2}{2}.$

$V(T) = \frac{a^2}{2} - \left(\frac{2}{3}a\right)^2 = \frac{1}{18}a^2.$

Fct $\int_{-\infty}^x f(t)\,dt = \int_0^x f(t)\,dt = \int_0^x kt\,dt = k\left[\frac{t^2}{2}\right]_0^x = \frac{kx^2}{2} = \frac{x^2}{a^2}.$

donc $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x^2}{a^2} & \text{si } x \in [0,a]. \\ 1 & \text{si } x > a. \end{cases}$

2) $M_n = \max(T_1, T_2, \dots, T_n)$

soit G la fonction de repartition de M_n.

$G(x) = P(M_n < x) = P(\max(T_1, T_2, \dots, T_n) < x)$

$= P(\{T_1 < x\} \cap \{T_2 < x\} \dots \{T_n < x\}) = \prod_{k=1}^n P(T_i < x)$

$= [P(T_i < x)]^n \Rightarrow G(x) = \begin{cases} 0 & \text{si } x < 0 \\ \left(\frac{x^2}{a^2}\right)^n & \text{si } x \in [0,a]. \\ 1 & \text{si } x > a. \end{cases}$

soit g la fonction densité de M_n

$g(x) = G'(x) = \begin{cases} \frac{2n x^{2n-1}}{a^{2n}} & \text{si } x \in [0,a]: \\ 0 \text{ si non.} \end{cases}$

Scanné avec CamScanner

$$E(M_n) = \int_{-\infty}^{+\infty} x\, g(x)\, dx = \int_0^a \frac{2n}{a^{2n}}\, x^{2n}\, dx$$

$$= \frac{2n}{a^{2n}} \left[\frac{x^{2n+1}}{2n+1}\right]_0^a = \frac{2n\, a^{2n+1}}{(2n+1)\, a^{2n}} = \frac{2n}{(2n+1)}\, a$$

$$V(M_n) = \int_{-\infty}^{+\infty} x^2\, g(x)\, dx - [E(M_n)]^2$$

$$= \int_0^a 2n\, \frac{x^{2n+1}}{a^{2n}}\, dx - \left(\frac{2n}{(2n+1)}\right)^2 a^2$$

$$= \frac{2n}{a^{2n}} \left[\frac{x^{2n+2}}{2n+2}\right]_0^a - \left(\frac{2n}{2n+1}\right)^2 a^2$$

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

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

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

* $\lim_{n \to +\infty} E(M_n) = \lim_{n \to +\infty} \frac{2n}{2n+1} \cdot a = a.$

* $\lim_{n \to +\infty} V(M_n) = \lim_{n \to +\infty} \left(\frac{2n}{2n+2} - \left(\frac{2n}{2n+1}\right)^2\right) a^2 = 0.$

Conclusion

La v.a M_n C. M. Q. vers a $\left(E(M_n) \xrightarrow[n \to +\infty]{} a,\ V(M_n) \xrightarrow[n \to +\infty]{} 0\right)$

Ex 11:

On cherche le nbr maximum de conteneurs pour qu'une benne pourra relever au cours d'une tournée

soit X: le poids des ordures d'un conteneur

$E(X) = 50$

$\sigma(X) = 30$

On cherche $P(S_n \leq 5500) = 0,95$ avec $S_n = \sum_{i=1}^{n} X_i$

On pourra approximer S_n par une loi normale $N(n50, \sqrt{n}30)$.

Scanné avec CamScanner

$T = \frac{S_m - 50m}{30\sqrt{m}}$, T suit la loi normale $N(0,1)$

$P(S_m \leqslant 5500) = 0,95$

$\Rightarrow P\left(T \leqslant \frac{5500 - 50m}{30\sqrt{m}}\right) = 0,95$

$\Rightarrow \frac{5500 - 50m}{30\sqrt{m}} = F_0^{-1}(0,95) = 1,6449.$

$\Rightarrow 5500 - 50m = 30 \times 1,6449\sqrt{m} \Rightarrow 5500 = 49,347\sqrt{m} + 50m.$

$\Rightarrow 50m + 49,347\sqrt{m} - 5506 = 0$ on pose $x = \sqrt{m}$

$\Rightarrow 50x^2 + 49,347x - 5500 = 0. \Rightarrow x_1 = 10, x_2 = -10,99$ à rejeter.

Donc $x_1 = 10 \Rightarrow \boxed{m = 100 t}$

2) on cherche m tq $P(S_m \leqslant 6000) = 0,95$

$\Rightarrow P\left(T \leqslant \frac{6000 - 50m}{30\sqrt{m}}\right) = 0,95$

$\Rightarrow \frac{6000 - 50m}{30\sqrt{m}} = F_0^{-1}(0,95) = 1,6449.$

$\Rightarrow 50m + 49,342\sqrt{m} - 6000 = 0.$ on pose $x = \sqrt{m}$

$\Rightarrow 50x^2 + 49,347x^2 - 6000 = 0$

$\Rightarrow x_1 = 10,47$, $x_2 = -11,45$ à rejeter

$\Rightarrow x_1 = 10,47$ donc $\boxed{m = 110 t}$

E x 12

soit x le nombre de parking à construire pour les 425 logements $X = \sum_{i=1}^{425} Z_i$, où $Z_i = \begin{cases} 1 & \text{si parking.} \\ 0 & \text{si non.} \end{cases}$

$P(Z_i = 1) = \frac{3}{5}$.

$E(X) = 425 \times \frac{3}{5} = 255.$

$V(X) = \frac{3}{5} \times \frac{2}{5} \times 425 = 102.$

$\sigma = 10,1$

soit x_0 nombre de parking à construire

$P(X_1 < x_0) = 0,9.$

Scanné avec CamScanner

on pourra approximer la loi de X par une loi Normale $N(255, 10,1)$ soit $T = \frac{X - 255}{10,1}$ une loi normale $N(0,1)$

$$P(X_1 < x_0) = P\left(T_1 < \frac{x_0 - 255}{10,1}\right) = 0,9$$

$$\frac{x_0 - 255}{10,1} = F_0^{-1}(0,9) = 1,2816.$$

$$x_0 = 1,2816 \times 10,1 + 255 = 268.$$

Ex 14:

au $i^{ème}$ cycliste on associe la variable Z_i $Z_i = 1$ (non dopé)

$Z_i = 0$ sinon.

$P(Z_i = 0) = 0,4$

soit Y_m le nbre de cycliste non dopés parmi m convoqués

$E(Y_m) = m \times 0,4$

$V(Y_m) = m \times 0,4 \times 0,6$

pour $m = 125$ ~~$\times 0,4 = 50$~~

$E(Y_{125}) = 125 \times 0,4 = 50$

$\sigma(Y_{125}) = \sqrt{30} = 5,48$

Le T.M.L

$Y_{125} \underset{app}{\sim} N(50, 5,48)$ cherchons $P(Y_m \geq 50) = 1 - P(Y_m < 50)$

soit $T = \frac{Y_m - 50}{5,48}$ on $N(0,1)$

$P(Y_m < 50) = P(T < 0) = F_0(0) = 0,5.$

$\Rightarrow P(Y_m \geq 50) = 0,5$

2) on cherche m tq $P(Y_m > 50) \geq 0,95$

$m > 150$ donc T.C.L $Y_m \underset{app}{\sim} N(m \times 0,4, \sqrt{m \times 0,4 \times 0,6})$

$$\Rightarrow T = \frac{Y_m - m \times 0,4}{\sqrt{m \times 0,4 \times 0,6}}$$

$$\Rightarrow P\left(T > \frac{50 - m \times 0,4}{\sqrt{m \times 0,4 \times 0,6}}\right) \geq 0,95. \quad \text{donc } 1 - P\left(T < \frac{50 - m \times 0,4}{\sqrt{m \times 0,4 \times 0,6}}\right) \leq 0,05$$

Scanné avec CamScanner

Ch I

$$F_0\left(\frac{50 - m \times 0{,}4}{\sqrt{m \times 0{,}4 \times 0{,}6}}\right) \leqslant 0{,}01 \Rightarrow \frac{50 - m \times 0{,}4}{\sqrt{m \times 0{,}4 \times 0{,}6}} \leqslant F_0^{-1}(0{,}05) = 1{,}6449$$

$$\Rightarrow 50 - m \times 0{,}4 \leqslant 1{,}6449 \sqrt{m \times 0{,}4 \times 0{,}6}$$

$$\Rightarrow \sqrt{m} = 10{,}22 \Rightarrow m = 105.$$

Scanné avec CamScanner

"Rappels et compléments sur les lois de Probabilités"

Ω se appelle événement certain

ϕ " " événement impossible

Algèbre d'événement : toute partie A de $\mathcal{P}(\Omega)$ vérifiant :

1) $\Omega \in A$

2) si $A \in \mathcal{A}$, alors $\bar{A} \in \mathcal{A}$

3) si $(A_i)_{i \in I}$ $I \subset \mathbb{N}$, I fini ou non) alors $\bigcup_{i \in I} A_i \in \mathcal{A}$

- **2 événement incompatible :**

$$A \cap B = \phi$$

- **système complet d'événement :**

$(A_i)_{i \in I}$ une famille d'événements 2 à 2 incompatibles et tq

$$\bigcup_{i \in I} A_i = \Omega$$

Ld $P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$

- si $(A_i)_{i \in I}$ une famille quelconque d'événements ona : $P\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} P(A_i)$
- si $(B_i)_{i \in I}$ forment un sys complet d'événement ona :

$$P(A) = \sum_{i \in I} P(A \cap B)$$

- **Probabilité conditionnelles :**

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

- A et B sont indépendant ssi : $P(A \cap B) = P(A) \cdot P(B)$

N.B : on désigne par une majuscule X, l'app et par la lettre minuscule x toute valeur prix par x.

X 2 une V.a. discrète si $E = x(\Omega)$ 2 fini ou dénombrable

loi discrète : x_i, p_i

Scanné avec CamScanner

- Fonction de répartition :

$F(x) = P(X \leq x)$ $x \in \mathbb{R}$; 1) $F \in [0,1]$

2) $F(-\infty) = 0$ $F(+\infty) = 1$ F est $\nearrow$

3) F est continue P.P

- Probabilité d'un intervalle :

$P(a < X \leq b) = F(b) - F(a) = \sum_{a < x \leq b} p_i$

- Moment d'une V.a discrète : $a < x \leq b$

Espérance :

$E(x) = \sum_{i \in I} x_i p_i$

1) $E(a) = a$ $a \in \mathbb{R}$

2) $E(ax) = a E(x)$

3) $E(a + x) = E(x) + a$

4) $E(x + y) = E(x) + E(y)$

Variance :

$V(x) = E\left((x - E(x))^2\right)$

Ou $V(x) = E(x^2) - (E(x))^2$

$= \sum_{i \in I} (x_i - E(x))^2 p_i$

1) $V(ax) = a^2 V(x)$

- $V(a + x) = V(x)$
- $V(x + y) = V(x) + V(y) + 2\,Cov(x, y)$

où $Cov(x, y) = E(x.y) - E(x)\,E(y)$

- $V(ax + by) = a^2 V(x) + b^2 V(y) + 2ab\,Cov(x, y)$.

Ecart-type :

$\sigma(x) = \sqrt{V(x)}$

Loi discrète usuelle :

Loi uniforme : $X \in U = [1 \dots n]$

$P(X = k) = \frac{1}{n}$ $E(X) = \frac{n+1}{2}$ $V(X) = \frac{n^2 - 1}{12}$; $F(k) = \frac{k}{n}$.

si $X \in U = [0 \dots n] \Rightarrow E(X) = \frac{n-1}{2}$.

$F(k) = \frac{k+1}{n}$

$p = \begin{cases} 1 & \text{si } X \in U \\ 0 & \text{sinon.} \end{cases}$

Scanné avec CamScanner

Loi de Bernoulli : $X \sim B(p)$

→ La réalisation d'une expérience n'ayant que deux issues possible → $X(\Omega) = \{0, 1\}$.

$P(X=1) = p$; $P(X=0) = 1-p$.

$E(X) = p$; $V(X) = pq$.

Loi Binomiale : $X \sim B(n,p)$, $X(\Omega) = \mathbb{N}$

→ La réalisation de n essais indépendants d'une expérience à deux issues possibles dont la probabilité de succès est p

X est le nbre de succès (loi Binomiale = somme de loi de Bernoulli)

$P(X=k) = C_n^k p^k (1-p)^{n-k}$.

$E(X) = np$; $V(X) = npq$

Loi géométrique : $X(\Omega) = \mathbb{N}^*$

$\forall k \in \mathbb{N}^* \quad P(X=k) = p(1-p)^{k-1}$

$E(X) = \frac{1}{p}$; $V(X) = \frac{q}{p^2}$.

Loi de Poisson : $X(\Omega) = \mathbb{N}$

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

$E(X) = V(X) = \lambda$

Lois continues :

→ $X(\Omega) = I$ où I : intervalle de $\mathbb{R}$.

- La loi de X est alors définie par sa f° de répartition F définie par : $F(x) = P(X(x) \quad x \in \mathbb{R}$.
- F est dérivable p.p.
- f : densité de probabilité. ; $f \geq 0$; continue

$$F(x) = \int_{-\infty}^{x} f(x)\,dx \qquad F(x) = \begin{cases} 0 \text{ si } x \notin X(\Omega) \\ x \text{ si } x \in X(\Omega) \\ 1 \text{ si mon.} \end{cases}$$

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1.$$

Rq : Pour une variable continue les probabilités ponctuelles sont nulles ce pour tout $x \in X(\Omega)$ on a : $P(X=x) = 0$.

Scanné avec CamScanner

- Probabilité d'un intervalle :

$$P(a < x \leq b) = F(b) - F(a) = \int_a^b f(u)\,du$$

$$P(a < x \leq b) = P(a \leq x < b) = P(a \leq x \leq b) = P(a < x < b)$$

- Moments d'une variable aléatoire continue :

$$E(x) = \int_{-\infty}^{+\infty} x f(x)\,dx$$

$$V(x) = E\left((x - E(x))^2\right) = E(x^2) - E^2(x) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - [E(x)]^2$$

- moment centré d'ordre k :
  le nbre $\mu_k = E\left((x - E(x))^k\right)$ $\quad \mu_1 = 0 ; \mu_2 = V(x)$
- moment non centré d'ordre k :
  le nbre $m_k = E(x^k)$; $m_2 = \mu_2 + m_1^2$.

- si Y la v.a définie par $Y = \varphi(x)$ où x est un v.a continue
soit G la fonction de répartition et g la densité de Y

Si φ est monotone :
Le φ est inversible et $G(y) = P(Y \leq y) = P(x \leq \varphi^{-1}(y))$ $y \in \mathbb{R}$
$= F[\varphi^{-1}(y)]$ $y \in \mathbb{R}$

et en dérivant on obtient la densité g de la variable Y.

Si φ est $\searrow$ et alors Y est inversible et son inverse φ^{-1} est $\searrow$.
on a alors : $G(y) = P(Y < y) = P(x < \varphi^{-1}(y)) = 1 - F(\varphi^{-1}(y))$

- Lois continues usuelles :

Loi uniforme sur $[0,a]$

$$f(x) = \begin{cases} \frac{1}{a-0} & \text{si } x \in [0,a] \\ 0 & \text{si non} \end{cases} \quad ; \quad E(x) = \frac{a+0}{2} \; ; \; V(x) = \frac{(a-0)^2}{12}.$$

$E(x) = \frac{a+b}{2}$ $\quad V(x) = \frac{(b-a)^2}{12}$

- Loi exponentielle $\mathcal{E}(a)$, $a > 0$

$$f(x) = \begin{cases} a e^{-ax} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

$$E(x) = \frac{1}{a}, \quad V(x) = \frac{1}{a^2}.$$

Scanné avec CamScanner

• Loi normale $N(0,1)$: $\mathbb{R}$

$f(x) = f_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$ $\quad E(x) = 0 \quad V(x) = 1$

• Loi normale $N(m, \sigma)$ $\quad m \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right)$ $\quad E(x) = m \quad V(x) = \sigma^2$

Rq :

1) $F_0(0) = 0,5$

2) $F_0(-t) = 1 - F_0(t)$; $\forall t \in \mathbb{R}$

3) $P(t > |T|) = P(t > T) - P(-t > T)$
$= 2F_0(t) - 1$, $(t \in \mathbb{R})$

4) $P(t < |T|) = P(t < T) + (T < -t) = 2(1 - F(t))$

Valeurs remarquables :

• $P(-1,64 < T < 1,64) = 0,90$

• $P(-1,96 < T < 1,96) = 0,95$

• $P(-3,05 < T < 3,05) = 0,998$

La Normale (m, σ)

1) X suit une $N(m, \sigma)$ ssi $X = \sigma T + m$ où T suit une loi $N(0,1)$

2) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right)$

3) Soit X_1 et X_2 2 v.a indépendantes $\sim N(m_1, \sigma_1); N(m_2, \sigma_2)$ et a et b deux nombres réels :

• $a X_2 \sim N(a m_1, |a| \sigma_1)$

• $a + X_2 \sim N(m_1 + a, \sigma_1)$

• $aX_2 + bX_1 \sim N(am_1 + bm_2, \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$

Scanné avec CamScanner

Lois dérivées de Lois normales :

1) Loi de Khi - deux

$T_1, T_2, \dots, T_m$ sont des v.a de loi $N(0,1)$ indépendantes

$\Rightarrow$ v.a $X_m = T_1^2 + \dots + T_m^2$ suit loi de Khi - deux de paramètre m, ce que l'on note $X_m = X_m^2$

m : le nbre de degré de liberté de la loi du χ^2.

- $E(X_m) = m$; $V(X_m) = 2m$. $X_m \geq 0$.

La somme de deux v.a indépendantes qui suivent la loi du χ^2 respectivement à m_1 et m_2 degrés de liberté. suit la loi $\chi^2_{m_1+m_2}$

- si m est assez grand ($m \geq 100$) la loi χ_m^2 peut être approchée par $N(m, \sqrt{2m})$
- $m \geq 30$ on a : soit Z_m une v.a de loi χ_m^2 ; lorsque m tend à ∞

Par suite des v.a $\sqrt{2Z_m} - \sqrt{2m-1}$ tend en loi vers $N(0,1)$

2) Loi de Student :

soit U une v.a de loi $N(0,1)$; X et U sont indépendantes.

et $X \sim \chi_m^2$

$\Rightarrow T = \dfrac{U}{\sqrt{\frac{X}{m}}}$ suit une loi de Student à m degré de liberté.

- $E(T) = 0$ car la loi de T est symétrique
- $F(-t) = 1 - F(t)$, $(t \in \mathbb{R})$ pour étudier la partition
- $m \geq 30$ $T \sim N(0,1)$.

Scanné avec CamScanner

"Correction des exercices"

X suit la loi uniforme sur $\{1, \dots, 6\}$.

$X(\Omega) = \{1, 2, 3, 4, 5, 6\}$

$P(X = k) = \frac{1}{6}$; $k \in \{1, \dots, 6\}$

On

$F(k) = P(X \leq k) = \frac{k}{6}$; $k \in \{1, \dots, 6\}$

a) $E(X) = \sum_{i=1}^{6} x_i p_i = \frac{7}{2} = 3,5$

$V(X) = E(X^2) - [E(X)]^2 = \frac{35}{12} = 2,9$.

$\sigma(X) = \sqrt{V(X)} = 1,7$

3) Fct de rép est en escalier.

$F(x) = 0$ si $x < 1$

$F(x) = 1$ si $x > 6$.

$F(x) = \frac{[x]+1}{6}$ sinon.

Ex2

1) X_1 et X_2 sont 2 V.a indépendantes ssi tout événement lié à l'une est indépendante de tout événement lié à l'autre. $\forall i, j \in \{1, 2, 3\}$,

$P(X_1 = x_i \cap X_2 = x_j) = P(X_1 = x_i) P(X_2 = x_j)$

2) $E(A) = E\left(\frac{1}{2}(X_1 + X_2)\right)$

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

car X_1 et X_2 sont de même loi alors:

$E(X_1) = E(X_2)$

$V(A) = \frac{1}{4}(V(X_1) + V(X_2))$

$= 0,25$.

en effet.

$V(X_1) = E(X_1^2) - (E(X_1))^2$

• $E(X_1^2) = 1 \times 0,2 + 4 \times 0,5 + 9 \times 0,3$

$= 4,9$

$V(X_1) = 4,9 - (2,1)^2 = 0,49 = V(X_2)$

$\Rightarrow V(A) = \frac{0,49}{2} = 0,245$

• $E = E(2X_1 - 1) = 2E(X_1) - 1$

$= 2 \times 2,1 - 1 = 3,1$

$E(C) = 2E(X_1) - 1 = 3,1$

$V(C) = 4V(X_1) = 1,96$.

2)

X_1 \ X_2	1	2	3
1	1	3/2	2
2	3/2	2	5/2
3	2	5/2	3

b.

Loi du couple (X_1, X_2)

X_1 \ X_2	1	2	3	$p_{X_2}(\cdot)$
1	0,04	0,1	0,06	0,2
2	0,1	0,2	0,15	0,5
3	0,06	0,15	0,09	0,3
$p_{X_1}(\cdot)$	0,2	0,5	0,3	1

• Loi de V. a. A:

La valeur de A	1	1,5	2	2,5	3	T
Probabilité	0,04	0,2	0,37	0,3	0,09	1

• Loi de B V. a. B:

La valeur de B	1	2	3	T
Probabilité	0,36	0,55	0,09	1

Scanné avec CamScanner

Loi de v.a.c:

Les valeurs de c	1	3	5	7
Probabilité	0,6	0,5	0,3	1

Espérance de B:

$E(B) = \sum b_i p_i = 1 \times 0,36 + 2 \times 0,27$

$+ 3 \times 0,09$

- Variance de B:

$V(B) = \sum b_i^2 p_i - (E(B))^2 \Rightarrow$

$= 1 \times 0,36 + 4 \times 0,27 + 9 \times 0,09$

$- (1,73)^2 = 0,37.$

Ex7:

$(\forall a, b) \in \mathbb{R}^2 ; a < b.$

f, la fonction de densité

G, ——— de répartition

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

$$G(x) = \int_{-\infty}^{x} f(t)\,dt.$$

$$= \int_a^x \frac{1}{b-a}\,dt = \frac{x-a}{b-a}$$

$$\Rightarrow G(x) = \begin{cases} 0 \text{ si } x < a \\ \frac{x-a}{b-a} \text{ si } x \in [a,b] \\ 1 \text{ sinon.} \end{cases}$$

$$E(x) = \frac{a+b}{2} \qquad V(x) = \frac{(b-a)^2}{12}$$

$$\sigma(x) = \frac{b-a}{2\sqrt{3}}.$$

$X(\Omega) = a + (b-a)\Omega$

$\Rightarrow X = a + (b-a)U$

- $F(x) = \begin{cases} 0 \text{ si } x < 0 \\ x \text{ si } x \in [0,1] \\ 1 \text{ sinon.} \end{cases}$

$G(x) = F\left(\frac{x-a}{b-a}\right), \quad x \in \mathbb{R}$

- $a \leq x \leq b$

$0 \leq \frac{x-a}{b-a} \leq \frac{b-a}{b-a} = 1.$

$\Rightarrow F\left(\frac{x-a}{b-a}\right) = \frac{x-a}{b-a}$

- si $x < a \quad x - a < 0.$

$\Rightarrow \frac{x-a}{b-a} < 0 \Rightarrow F\left(\frac{x-a}{b-a}\right) = 0.$

- si $x > b \Rightarrow x - a > b$

$\Rightarrow \frac{x-a}{b-a} > 1 \Rightarrow F\left(\frac{x-a}{b-a}\right) = 1.$

- $g(x) = G'(x) = \left(F^*\left(\frac{x-a}{b-a}\right)\right)'$

$= \frac{1}{b-a} F'\left(\frac{x-a}{b-a}\right) = \frac{1}{b-a} f\left(\frac{x-a}{b-a}\right).$

3) $X \sim$ loi uniforme sur $[-1,1]$.

$Y = |X|$ ne prend valeur sur $[0,1]$.

$P(Y < y) = P(|X| < y) \; ; \; y \in [0,1)$

$= P(-y < X < y)$

$= G(y) - G(-y) = \frac{y+1}{2} - \frac{-y+1}{2}$

$= y.$

Donc Y suit la loi uniforme sur $[0,1)$.

Scanné avec CamScanner

ce n°3:

$= E($

$E(x^2) = \sum_{i \in I} x_i^2 p_i$

$= \sum_{i \in I} \left(x_i^2 - [E(x)]^2 + [E(x)]^2 \right) p_i$

$= \sum_{i \in I} \left(x_i^2 - [E(x)]^2 \right) p_i + \sum_{i \in I} [E(x)]^2 p_i$

$= V(x) + [E(x)]^2$

Ex 4:

1) $E(y) = 10 E(x_1) + E(x_2)$

$E(x_1) = 0 \times \frac{1}{10} + \frac{1}{10} + \frac{2}{10}$

$+ \frac{3}{10} + \frac{4}{10} + \frac{5}{10} + \frac{6}{10} + \frac{7}{10}$

$+ \frac{8}{10} + \frac{9}{10} = 4,5 = E(x_2)$

~~$V(y) = E(y^2) - [E(y)]^2$~~

$V(y) = V(10 x_1 + x_2) = 100 V(x_1) + V(x_2)$

$= 10 V(x_1)$

$E(x_1^2) = \frac{1}{10} + \frac{4}{10} + \frac{9}{10} + \frac{16}{10}$

$+ \frac{25}{10} + \frac{36}{10} + \frac{49}{10} + \frac{64}{10} + \frac{81}{10}$

$= 28,5$

$V(x_1) = E(x_1^2) - [E(x_1)]^2$

$= 28,5 - (4,5)^2 = 8,25$

$\Rightarrow V(y) = 10 \times 8,25 = 833,25.$

~~m.q:~~

a) $m \leq x_i, \; i = 1, 2.$

$\Rightarrow E(m) \leq E(x_i), \; i = 1, 2$

$\Rightarrow E(m) \leq \min [E(x_i), \; i = 1, 2].$

b) y est la loi uniforme sur $\{0, 1, \dots, 99\}$

les valeurs prises par $M \in \{0, \dots, 99\}$

$F(k) = \frac{k+1}{10}$

$P(m > k) = P(x_1 > k \cap x_2 > k)$

$= P(x_1 > k) \cdot P(x_2 > k)$

$P(x > k) = 1 - P(x \leq k) = F(k)$ (fonction de répartition)

$= \left(1 - \frac{k+1}{10}\right)^2$

$P(m = k) = P(m > k-1) - P(m > k)$

$= \left(1 - \frac{k}{10}\right)^2 - \left(1 - \frac{k+1}{10}\right)^2$

pour $k = 0$: $P(m = 0)$

$= 1 - \left(1 - \frac{1}{10}\right)^2$

$= 1 - \left(\frac{9}{10}\right)^2 = \frac{19}{100}$

pour $k = 1$:

$P(m = 1) = \left(1 - \frac{1}{10}\right)^2 - \left(1 - \frac{2}{10}\right)^2$

$= \frac{81}{100} - \frac{64}{100} = \frac{17}{100}$

pour $k = 2$:

$P(m = 2) = \left(1 - \frac{2}{10}\right)^2 - \left(1 - \frac{3}{10}\right)^2$

$= \frac{64}{100} - \frac{49}{100} = \frac{15}{100}.$

pour $k = 3$:

$P(m = 3) = \frac{13}{100}$

$P(m = 4) = \frac{11}{100}$

$P(m = 5) = \frac{9}{100}$

$P(m = 6) = \frac{7}{100}$

$P(m = 7) = \frac{5}{100}$

$P(m = 8) = \frac{3}{100}$

$P(m = 9) = \frac{1}{100}$

(e)

Scanné avec CamScanner



$$E(m) = \sum_{i=0}^{9} m_i P(m=i)$$

$$= 0 \times \frac{19}{100} + 1 \times \frac{17}{100} + 2 \times \frac{15}{100}$$

$$+ 3 \times \frac{13}{100} + 4 \times \frac{11}{100} + 5 \times \frac{9}{100}$$

$$+ 6 \times \frac{5}{100} + 8 \times \frac{3}{100} + 9 \times \frac{1}{100}$$

$$V(m) = \sum_{i=0}^{2} m_i^2 P(m=i) - [E(m)]^2$$

EX 5:

La probabilité que l'un au moins présente un symptôme.

A_i : L'événement l'adolescent n°i présente un symptôme

On va noter $\overline{A_i}$ = l'événement contraire

$P(A_i) = 0,2 \Rightarrow P(\overline{A_i}) = 0,8$

soit E l'événement

$$E = \bigcup_{i=1}^{5} A_i$$

$$\overline{E} = \bigcap_{i=1}^{5} \overline{A_i}$$

$$\Rightarrow P(\overline{E}) = P\left(\bigcap_{i=1}^{5} \overline{A_i}\right) = \left(P(\overline{A_1})\right)^5$$

$$= (0,8)^5$$

$$P(E) = 1 - P(\overline{E}) = 1 - (0,8)^5$$

$$= 0,67.$$

2) On effectue n épreuves successives indépendantes et on observe à chaque fois la réalisation d'un certain événement A et on note le nombre X de réalisation de A. On définie ainsi une variable aléatoire X qui suit la loi binomiale de paramètre n et p avec $n = 30$ et $p = 0,2$

Au ième élève de la classe on associe la variable Z_i définie par $Z_i = 1$ si l'individu i présente un symptôme

$Z_i = 0$ sinon.

Alors $X = \sum_{i=1}^{30} Z_i$

$E(X) = n \cdot p = 30 \times 0,2 = 6$

$V(X) = n \cdot p \cdot q = 30 \times 0,8 \times 0,2 = 4,8$

$F(4) = 0,2552$ } données.

$F(3) = 0,1227$

$F(8) = 0,8713.$

$P(X \geq 5) = 1 - P(X < 5)$

$= 1 - P(X \leq 4) = 1 - F(4)$

$= 1 - 0,2552 = 0,74$

$|X - E(X)| \leq 2$

$\Rightarrow -2 \leq X - E(X) \leq 2$

$-2 + E(X) \leq X \leq 2 + E(X)$

$\Rightarrow 4 \leq X \leq 8.$

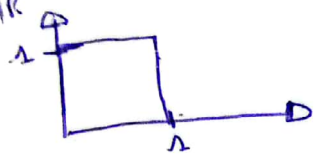
Scanné avec CamScanner

$P(3 < X < 8) = F(8) - F(3)$

$= 0,8713 - 0,1227 = 0,74.$

Ex 6:

1) F une densité de probabilité ssi $\begin{cases} \cdot f \geq 0 \\ \cdot f \text{ cont p.p} \\ \cdot \int_{\mathbb{R}} f(t)\,dt = 1 \end{cases}$



$f = \begin{cases} 1 & \text{si } x \in [0,1] \\ 0 & \text{sinon} \end{cases}$

2) Fonction de répartition

$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0,1) \\ 1 & \text{sinon} \end{cases}$

* $P(U < 0,25) = F(0,25) = 0,25$
* $P(0,2 < U < 0,8) = F(0,8) - F(0,2)$
* $P(U \geq 0,6) = 0,6$

$= 1 - P(U < 0,6) = 1 - F(0,6)$

$= 0,4$

3)

$E(U) = \frac{0+1}{2} = \frac{1}{2}$

$V(U) = \frac{1-0}{12} = \frac{1}{12}$ ✗

$\sigma(U) = \sqrt{V(U)} = \frac{1}{2\sqrt{3}}.$

Ex 9:

1) $F(x) = \int_{-\infty}^{x} f(t)\,dt$

$= \int_0^x \frac{1}{\alpha} \exp\left(-\frac{1}{\alpha}t\right) dt$

$= \frac{1}{\alpha} \left[-\alpha \exp\left(-\frac{1}{\alpha}t\right)\right]_0^x$

$= -\left[\exp\left(-\frac{1}{\alpha}x\right) - 1\right]$

$F(x) = 1 - \exp\left(-\frac{1}{\alpha}x\right)$

$E(X) = \frac{1}{\frac{1}{\alpha}} = \alpha$

$V(X) = \alpha^2$

2) $\alpha = 2500$

$P(2500 < x < 3500) = F(3500) - F(2500)$

$= 1 - \exp\left(-\frac{3500}{2500}\right) - 1 + \exp\left(-\frac{2500}{2500}\right)$

$= \exp(-1) - \exp\left(\frac{3500}{2500}\right) = 0,12.$

3)

$P(x > 3500 \mid x > 2500) = \frac{P(x > 3500 \cap x > 2500)}{P(x > 2500)}$

$= \frac{1 - P(x < 3500)}{1 - P(x < 2500)}$

$= \frac{1 - F(3500)}{1 - F(2500)} = 0,67.$

Ex 10:

$P(0 < T < 1,5) = F_0(1,5) - F_0(0)$

$= 0,9332 - 0,5$

$= 0,4332.$

Scanné avec CamScanner

$P(T < -0,5) = 1 - P(T < 0,5)$

$= 1 - F_0(0,5)$

$= 1 - 0,6915 = 0,3085$

$P(|T| > 2,5) = 2(1 - F_0(2,5))$

$= 2(1 - 0,9938) = 0,0124.$

• $P(T < a) = 0,95 = 1,6449$

• $P(|T| < b) = 0,6.$

$= 2F_0(b) - 1$

$\Rightarrow F_0(b) = \frac{1,6}{2} = 0,8$

$\Rightarrow b = 0,8416$

$P(T < c) = 0,25$

$\Rightarrow c = -0,6745$

3) $f_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$; $x \in \mathbb{R}$

$\max(f_0(x)) = \frac{1}{\sqrt{2\pi}}$

Exercice 2

X suit une loi normale $N(2,5; 5^2)$

~~$P(0 \leq X < 7,5)$~~

donc $T = \frac{X - 2,5}{5}$ suit la loi normale centrée réduite $N(0;1)$

$P(0 \leq X < 7,5) = P(-\frac{1}{2} < X < 1)$

$F_0(1) - F_0(-\frac{1}{2})$

$= 0,8413 - 0,3085 = 0,5328.$

* $P(|X - 2,5| > 7,5)$

$|X - 2,5| > 7,5$

$\Rightarrow \left|\frac{X - 2,5}{5}\right| > \frac{7,5}{5}$

$= |T| > 1,5$

$\Rightarrow P(|X - 2,5| > 7,5)$

$= P(|T| > 1,5)$

$= 2(1 - F_0(1,5))$

$= 2(1 - 0,9332)$

$= 0,1336$

Exercice

1) X suit une loi normale

on a: $P(X > 2) = 0,6$; $P(X \leq -1,5) = 0,1$

$T = \frac{X - m}{\sigma}$ suit une loi normale $N(0,1)$

$P\left(\frac{X - m}{\sigma} > \frac{2 - m}{\sigma}\right) = 0,6$

$= P\left(T > \frac{2 - m}{\sigma}\right)$

$= 1 - P\left(T < \frac{2 - m}{\sigma}\right)$

$\Rightarrow P\left(T < \frac{2 - m}{\sigma}\right) = 0,4.$

$\frac{2 - m}{\sigma} = -0,2533$

* $P\left(T \leq \frac{-1,5 - m}{\sigma}\right) = 0,1$

$\frac{-1,5 - m}{\sigma} = -1,2816$

$\begin{cases} \frac{2 - m}{\sigma} = -0,25533 \\ \frac{-1,5 - m}{\sigma} = -1,2816. \end{cases}$

$\begin{cases} m - 2 = 0,2533\,\sigma & (1) \\ 1,5 + m = 1,2816\,\sigma. & (2) \end{cases}$

Scanné avec CamScanner

Ex 13 :

$T \sim N(0,1)$

pour $n=1$ $E(T)=0$; $V(T)=1$

pour $n=2$ $E(T^2)=1$

$V(T^2)=2$.

(car soit $X_m = T_1^2 + T_2^2 + \dots + T_m^2$ sont la loi du Khi-deux de paramètre n $\Rightarrow E(X_m)=n$, $V(X_m)=2n$).

$$E(T^3) = \int_{\mathbb{R}} x^3 \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx.$$

soit $u(x) = x^2$

$v'(x) = x \exp\left(\frac{-x^2}{2}\right) \Rightarrow \begin{cases} u'(x) = 2x \\ v(x) = -\exp\left(\frac{-x^2}{2}\right) \end{cases}$

IPP $\Rightarrow \int_{\mathbb{R}} x^3 \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx$

$= \left[-x^2 \exp\left(\frac{-x^2}{2}\right)\right]_{-\infty}^{+\infty} + \int_{\mathbb{R}} 2x \exp\left(\frac{-x^2}{2}\right)$

$= -2 \int_{\mathbb{R}} -x \exp\left(\frac{-x^2}{2}\right)$

$= 2 E(T) = 0$

1) $E(X) = E(2T-1) = 2E(T) - 1 = -1$.

$V(X) = V(2T-1) = 4V(T) = 4$.

- $E(Z) = E(T^2) = 1$.
  $V(Z) = V(T^2) = 2$

2) X suit la loi Normale $N(-1,2)$.

Z suit la loi du Khi-deux de paramètre 1.

3) $P(Y_1 \le y) = P(|T| \le y)$

$= P(-y \le T \le y)$

$= F(y) - F(-y) = F(y) - (1 - F(y))$

$= 2F_0(y) - 1$. $y \in \mathbb{R}_+$

$$F(y) = \begin{cases} 2F_0(y) - 1 & \text{si } y \in \mathbb{R}_+ \\ 0 & \text{si non.} \end{cases}$$

$$f(y) = \begin{cases} 2f(y) & \text{si } y \in \mathbb{R}_+ \\ 0 & \text{si non.} \end{cases}$$

$$E(Y) = \int_{-\infty}^{+\infty} y f(y) dy.$$

$$= \int_0^{+\infty} 2y f(y) dy.$$

$$= \int_0^{+\infty} \frac{2y}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right) dy$$

$$= -\frac{2}{\sqrt{2\pi}} \left[\exp\left(\frac{-y^2}{2}\right)\right]_0^{+\infty}$$

$$= \frac{2}{\sqrt{2\pi}}$$

Exo 4 :

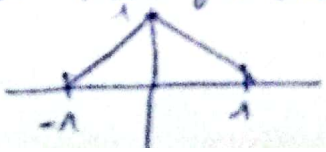
1) f est continue p.p. positive

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-1}^{0} (1+x) dx + \int_0^1 (1-x) dx$$

$$= \left[x + \frac{x^2}{2}\right]_{-1}^{0} + \left[x - \frac{x^2}{2}\right]_0^1$$

$$= \frac{3}{2} - \frac{1}{2} = 1.$$

d'où f est une densité de probabilité



Scanné avec CamScanner

2) $E(x) = \int_{-\infty}^{+\infty} x f(x)\, dx$

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

$= \left[\frac{x^2}{2}\right]_{-1}^{0} + \left[\frac{x^3}{3}\right]_{-1}^{0} + \left[\frac{x^2}{2}\right]_{0}^{1} - \left[\frac{x^3}{3}\right]_{0}^{1}$

$= -\frac{1}{2} + \frac{1}{3} + \frac{1}{2} - \frac{1}{3} = 0$

si $x \in]0,1]$

$F(x) = \int_{0}^{x} (x - x^2)\, dx$

$= \left[\frac{x^2}{2}\right]_{0}^{x} - \left[\frac{x^3}{3}\right]_{0}^{x}$

$= \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$

$V(x) = E(x^2) - E^2(x)$

$E(x^2) = \int x^2 f(x)\, dx$

$= \int_{-1}^{0} (x^2 + x^3)\, dx + \int_{0}^{1} (x^2 - x^3)\, dx$

$= \left[\frac{x^3}{3}\right]_{-1}^{0} + \left[\frac{x^4}{4}\right]_{-1}^{0} + \left[\frac{x^3}{3}\right]_{0}^{1} - \left[\frac{x^4}{4}\right]_{0}^{1}$

$= \frac{1}{3} + \frac{1}{4} + \frac{1}{3} - \frac{1}{4} = \frac{1}{3}$

3) $F(x) = \int_{-\infty}^{x} f(t)\, dt$

- si $x \leq -1$: $F(x) = 0$.
- si $x \in]-1, 0]$:

$F(x) = \int_{-1}^{x} (t + t^2)\, dt$

$= \left[\frac{t^2}{2}\right]_{-1}^{x} + \left[\frac{t^3}{3}\right]_{-1}^{x}$

$= \frac{x^2}{2} - \frac{1}{2} + \frac{x^3}{3} + \frac{1}{3}$

$= \frac{3x^2 + 2x^3}{6} - \frac{1}{6}$

Scanné avec CamScanner

$\Phi(1,5) = (1,2816 - 0,2533)$

$\sigma = 3,4$

$\begin{cases} \sigma = 3,4 \\ m = 2,86 \end{cases}$

a) $P(-0,5 \leqslant X < 5) = P\left(\frac{-0,5-2,86}{3,4} \leqslant T < \frac{6-2,86}{3,4}\right)$

~~$= P(-0,99)$~~ -

$= P(-0,99 \leqslant T < 0,92)$

$= F_0(0,92) - F_0(-0,99)$

$= 0,8212 - 1 + 0,8389$

$= 0,6601.$

Ex 15:

1) pour que f soit une densité de probabilité, il faut que f soit continue p.p

$f \geqslant 0 \Rightarrow K \geqslant 0$

$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \Rightarrow \int_a^{+\infty} \frac{K}{x^3}\,dx = 1$

$\Rightarrow \left[\frac{-K}{2x^2}\right]_a^{+\infty} = \frac{K}{2a^2} = 1.$

$\Rightarrow K = 2a^2$

$E(x) = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_a^{+\infty} \frac{2a^2}{x^2}\,dx$

$= 2a^2 \left[-\frac{1}{x}\right]_a^{+\infty} = 2a.$

$V(x) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - 4a^2$

$= 2a^2 \int_a^{+\infty} \frac{dx}{x} - 4a^2$: n'a pas convergente

$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_a^x \frac{2a^2}{t^3}\,dt$

$= 2a^2 \left[-\frac{1}{2t^2}\right]_a^x$

$= 2a^2 \left[-\frac{1}{2x^2} + \frac{1}{2a^2}\right]$

$= 1 - \frac{2a^2}{2x^2} = 1 - \frac{a^2}{x^2}$

$\Rightarrow F(x) = \begin{cases} 1 - \frac{a^2}{x^2} & \text{si } x \geqslant 0 \\ 0 & \text{sinon.} \end{cases}$

$F(m_e) = \frac{1}{2}$

$\Rightarrow \frac{1 - a^2}{m_e^2} = \frac{1}{2} \Rightarrow \frac{a^2}{m_e^2} = \frac{1}{2}$

$\Rightarrow m_e = a\sqrt{2}.$

- $P(X \leqslant m_e) = F(m_e)$

$= \frac{1 - a^2}{m_e^2}$

$P(X \geqslant m_e) = 1 - P(X < m_e)$

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

$P(X \leqslant m_e) = P(X \geqslant m_e)$

$\Rightarrow 1 - \frac{a^2}{m_e^2} = \frac{a^2}{m_e^2}.$

$\Rightarrow \frac{2a^2}{m_e^2} = 1 \Rightarrow m_e^2 = 2a^2$

$\Rightarrow m_e = a\sqrt{2}.$

Scanné avec CamScanner

Ex 8 :

1) • X prend valeurs dans $[0,1]$ ($X(\Omega)=[0,1]$)
• Y prend valeurs dans $[\frac{1}{2},1]$ ($Y(\Omega)=[\frac{1}{2},1]$)
• Z ——— dans $[0,+\infty[$, (...

2) soit F la fonction de répartition de X.

$F(x) = P(X \leq x) = P(\sqrt{U} \leq x) = P(U \leq x^2)$

$= x^2$

$\Rightarrow F(x) = \begin{cases} 0 & \text{si } x<0 \\ x^2 & \text{si } x \in]0,1[\\ 1 & \text{sinon.} \end{cases}$

$f(x) = \begin{cases} 2x & \text{si } x \in]0,1[\\ 0 & \text{sinon.} \end{cases}$

• soit G la fonction de répartition de Y si $y \in [\frac{1}{2},1]$.

$\Rightarrow P(Y \leq y) = P\left(\frac{1}{2U+1} \leq y\right)$

$= P\left(2U+1 \geq \frac{1}{y}\right) = P\left(U \geq \frac{\frac{1}{y}-1}{2}\right)$

$= 1 - P\left(U \leq \frac{\frac{1}{y}-1}{2}\right)$

$= 1 - \frac{\frac{1}{y}-1}{2} = \frac{3-\frac{1}{y}}{2}$

$= \frac{3}{2} - \frac{1}{2y}$.

$G(y) = \begin{cases} 0 & \text{si } y < \frac{1}{2} \\ \frac{3}{2} - \frac{1}{2y} & \text{si } \frac{1}{2} \leq y \leq 1 \\ 1 & \text{si non} \end{cases}$

donc $g(y) = \begin{cases} \frac{1}{2y^2} & \text{si } y \in [\frac{1}{2},1] \\ 0 & \text{si non.} \end{cases}$

• soit H la fonction de répartition de Z.

$P(Z \leq y) = P(-\ln(U) \leq y)$

$= P(\ln(U) \geq -y) = P(U \geq e^{-y})$

$= 1 - P(U < e^{-y}) = 1 - e^{-y}$.

donc Z n'a pas de paramètre.

$V(Z) = 1$.

3) $E(X) = E(\sqrt{U}) = \int_{-\infty}^{+\infty} x f(x)\,dx$

$= \int_0^1 2x^2\,dx = \frac{2x^3}{3}\Big]_0^1 = \frac{2}{3}$.

$V(X) = V(\sqrt{U}) = E(X^2) - E(X)^2$

$E(X^2) = \int_{\mathbb{R}} x^2 f(x)\,dx = \int_0^1 2x^3\,dx$

$= \int_0^1 2x^3\,dx = \frac{x^4}{2}\Big]_0^1$

$= \frac{1}{2}$

$V(X) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$.

$E(Y) = \int_{\mathbb{R}} y f(y)\,dy = \int_{1/3}^1 y \times \frac{1}{2y^2}\,dy$.

$= \int_{1/3}^1 \frac{1}{2y}\,dy = \frac{1}{2}\left[\ln(y)\right]_{1/3}^1$

$= \frac{1}{2}\ln\left(\frac{1}{1/3}\right)$

$= \frac{1}{2}\left(\ln(1) - \ln(3)\right)$

$= \frac{1}{2}\left(\ln(\sqrt{3}^2) = \frac{2}{2}\ln(\sqrt{3})\right.$

$= \ln(\sqrt{3})$

$V(Y) = E(Y^2) - E(Y)^2$.

$E(Y^2) = \int_{1/3}^1 y^2 \times \frac{1}{2y^2}\,dy$

$= \int_{1/3}^1 \frac{1}{2}\,dy = \frac{1}{2}\left[1 - \frac{1}{3}\right] = \frac{1}{2} \times \frac{2}{3}$

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

$\Rightarrow V(Y) = \frac{1}{3} - \left(\ln(\sqrt{3})\right)^2$.

• Z n'a pas de paramètre.

$E(Z) = V(Z) = 1$.

Scanné avec CamScanner