Statistique

TD n°2

1/2

Statistique TD n°2

Exercice 2

$X_n \rightsquigarrow B(n, \lambda/n)$

On a $P(X_n = x) = C_n^k \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$

$$= \frac{n!}{(n-x)!\,x!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)\cdots(n-(x-1))(n-x)!}{(n-x)!\,x!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{\lambda^x}{x!} \times \frac{n \cdot n\left(1-\frac{1}{n}\right)\cdots n\left(1-\frac{1-x}{n}\right)}{n^x} \cdot \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{\lambda^x}{x!} \cdot \frac{\cancel{n^x}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{\cancel{n^x}} \cdot \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}$$

d'où

$$\lim_{x \to +\infty} P(X_n = x) = \frac{\lambda^x}{x!} \lim_{n \to +\infty} \left(1-\frac{\lambda}{n}\right)^n$$

$$= e^{\lambda} \cdot \frac{\lambda^x}{x!}$$

d'où $F_n(x) = \sum_{k=0}^{\infty} P(X_n = x) = P(X \leq n)$

$\Rightarrow F(x) = e^{\lambda} \cdot \frac{\lambda^x}{x!}$ qui est la fonction de répartition de la loi de Poisson.

Exercice 3

X est une va presque sûrement (presque partout) vers a (cte sauf sur un intervalle donc la proba est nulle.

(1) Mq. $X_n \xrightarrow{mq} X \overset{\text{rappel}}{\Longleftrightarrow} \lim\limits_{n \to \infty} E(X_n - X)^2 = 0$

On a :

$$E(X_n - X)^2 = E[(X_n - E[X_n]) + (E[X_n] - X)]^2$$
$$= E(X_n - E[X_n])^2 + E(E[X_n] - X)^2$$
$$+ 2\underbrace{E[(X_n - E[X_n])(E[X_n] - X)]}_{= 0}$$

$E[X_n E[X_n]]$
$= E[X_n]E[X_n]$
$= (E[X_n])^2$

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

avec $X = a$ et d'après les hypothèses :

$$\lim_{n \to \infty} E(X_n - X)^2 = 0 + (E(\lim_{n \to +\infty} E(X_n) - a))^2$$
$$= 0$$

(2) $X_n \xrightarrow{P} X \Longleftrightarrow \forall \varepsilon > 0 \ \lim\limits_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$

Soit $\varepsilon > 0$. On a:

$$P(|X_n - X| > \varepsilon) = P(|X_n - X|^2 \geq \varepsilon^2)$$

inégalité de markov $\rightarrow \leq \dfrac{E[X_n - X]^2}{\varepsilon^2} \xrightarrow[n \to \infty]{} 0$ d'après (1)

$$\Rightarrow \lim_{n \to +\infty} P(|X_n - X| > a) = 0$$

(3) $X_n \xrightarrow{P} X \not\Rightarrow V(X_n) \not\to 0$

ou $E(X_n) \not\to X = a$

Exemple : soit (X_n) une suite de v.a. définie sur $(0,1)$

$P(X_n = 0) = 1 - \frac{1}{n}$ et $P(X_n = n) = \frac{1}{n}$

x_n	0	n	TOTAL
$P(X_n = x_n)$	$1-\frac{1}{n}$	$\frac{1}{n}$	1

X_n v.a. discrète

On prend $X = a = 0$

On a bien $X_n \xrightarrow{P} a = 0$ En effet :

$$P(|X_n| > \varepsilon) = 1 - P(|X_n| \leq \varepsilon)$$
$$= 1 - P(X_n = 0)$$
$$= 1 - (1 - \frac{1}{n}) = 1/n$$

d'où $\lim\limits_{n \to \infty} P(|X_n| > \varepsilon) = \lim\limits_{n \to \infty} \frac{1}{n} = 0$

Par contre

$$E(X_n) = \sum x_n \, P(X_n = x_n)$$
$$= 0 \cdot \left(1 - \frac{1}{n}\right) + n \times \frac{1}{n} = 1$$

$$V(X_n) = (E(X_n))^2 = 0 + n^2 \times \frac{1}{n} - 1^2 = n \longrightarrow +\infty$$

Exercice 4

Les X_i sont considérés comme des vaiid
indépendantes et de
même loi

avec $E[X_i] = d$
$V[X_i] = 4$

Soit n le nombre d'épreuves nécessaires

et $S_n = X_1 + \ldots + X_n$

et $\overline{X}_n = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$ la moyenne empirique qui est

une estimation la vraie moyenne.

L'erreur de calcul est $\varepsilon = |\overline{X}_n - d|$. On cherche n tq :

$$P(\varepsilon \leq 1/2) = 0,95$$

$$\text{TCL} : Z_n = \frac{S_n - E(S_n)}{\sigma_{S_n}} \xrightarrow[n \to \infty]{P} \mathcal{N}(0, 1)$$

$S_m = \sum_{i=1}^{m} X_i \qquad E(S_m) = \sum_{i=1}^{m} E(X_i) = m\,E(X) = m\mu$

$V(S_m) = \sum_{i=1}^{m} V(X_i) = m\,V(X) = 4m$

$S_m = m\overline{X}_m \qquad \sigma_{S_m} = 2\sqrt{m}$

Donc $Z_m = \dfrac{S_m - m\mu}{2\sqrt{m}} \sim \mathcal{N}(0,1)$

On veut

$0{,}95 = P(|\overline{X}_m - \mu| < 0{,}5)$

$= P(-1/2 < \overline{X}_m - \mu < 1/2)$

$= P(-m/2 < m\overline{X}_m - m\mu < m/2)$

$= P\left(\dfrac{-\frac{m}{2}}{2\sqrt{m}} < Z_m < \dfrac{m/2}{2\sqrt{m}}\right)$

$= P\left(-\dfrac{\sqrt{m}}{4} < Z_m < \dfrac{\sqrt{m}}{4}\right)$

$= F\left(\dfrac{\sqrt{m}}{4}\right) - F\left(-\dfrac{\sqrt{m}}{4}\right) = 2F\left(\dfrac{\sqrt{m}}{4}\right) - 1 = 0{,}95$

$\Rightarrow F\left(\dfrac{\sqrt{m}}{4}\right) = \dfrac{1{,}95}{2} = 0{,}975$

La table de $\mathcal{N}(0,1)$ donne $\dfrac{\sqrt{m}}{4} = 1{,}96$

$\Rightarrow m \simeq 62$

Exercice 5

Il s'agit d'expériences de Bernoulli.

(1) $X \sim \mathcal{B}(m,p)$

$p = \dfrac{84}{233}$ et $m = 233$

$E(X) = mp = 84$

$V(X) = mp(1-p) = 233 \cdot \dfrac{84}{233}\left(1 - \dfrac{84}{233}\right) = 53{,}71$

Statistique TD n°2 2/2

(2) $n = 100$

On pose $Y = \frac{X}{n}$ $\quad X = \sum_{i=1}^{n} X_i \quad X_i \sim \mathcal{B}er(p)$

$$E(Y) = \frac{1}{n} E(X) = \frac{100 \cdot 0{,}36}{100} = 0{,}36$$

$$V(Y) = \frac{1}{n^2} V(X) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n} = \frac{100 \times 0{,}36(1-0{,}36)}{100}$$

$$= 0{,}0023$$

$$Y = S_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$n = 100 \gg 30 \Rightarrow$ on peut utiliser le TCL

on peut approximer la loi de Y par une loi normale $\mathcal{N}(\mu, \sigma)$

où $\mu = E(Y) = E(S_n)$

$\sigma^2 = V(S_n)$

$Y \sim \mathcal{N}(0{,}36 ; 0{,}0023)$

$P\left(\frac{X}{n} \geq -0{,}44\right) = 0{,}03$

$\Leftrightarrow P(Y > 0{,}44) = 0{,}03$

$\Leftrightarrow P\left(\frac{Y - E(Y)}{\sigma_Y} > \frac{0{,}44 - E(Y)}{\sigma_Y}\right) = 0{,}03$

$\frac{0{,}44 - 0{,}36}{\sqrt{\frac{0{,}36(1-0{,}36)}{n}}}$

$$E(Y) = \frac{E(X)}{n} = \frac{np}{n} = p$$

$$V(Y) = \frac{V(X)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

$\Leftrightarrow P\left(Z_n > \frac{0{,}08}{0{,}45} \sqrt{n}\right) = 0{,}03$

La table donne $\frac{0{,}08}{0{,}48} \sqrt{n} = 1{,}88 \Rightarrow n \simeq 128$