Corrigé de l'épreuve de Probabilités 1 du jeudi 9 Juin 2011.

Exercice 1 : 1) Sachant que $X=n$, Y est de loi binomiale de paramètre $(n, p=0{,}1)$ (en effet Y = le nombre de succès au cours de n épreuves répétées, si succès = être candidat à l'emploi). Donc

$$P(Y=k|X=n)=\begin{cases} C_n^k p^k(1-p)^{n-k} & \text{si } 0\leq k\leq n \\ 0 & \text{si } n<k \end{cases}$$

2) Les événements $[X=n]$ $(n\in\mathbb{N})$ forment un système complet d'événements donc d'après la formule de la probabilité totale :

$$\mathbb{P}(Y=k)=\sum_{n=0}^{+\infty}\mathbb{P}(Y=k|X=n)\,\mathbb{P}(X=n)$$

$$=\sum_{n=k}^{+\infty}C_n^k p^k(1-p)^{n-k}e^{-\lambda}\frac{\lambda^n}{n!}$$

$$=e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{n=k}^{+\infty}\frac{(\lambda(1-p))^{n-k}}{(n-k)!}=e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{\lambda(1-p)}$$

$$=e^{-\lambda p}\frac{(\lambda p)^k}{k!}\quad (k\in\mathbb{N}).$$

Y suit donc la loi de Poisson de paramètre $\lambda p=1000\cdot 0{,}1=100$ et $\mathbb{E}(Y)=100$ et $\operatorname{Var}Y=100$ (car on sait que pour une loi de Poisson, l'espérance et la variance sont égales à la valeur du paramètre.)

Exercice 2 : $f(x)=kxe^{-x/2}\mathbb{1}_{[0,+\infty[}(x)$

1) $f\geq 0$ (clair) f est une densité de probabilité si et seulement si $\int_{-\infty}^{\infty}f(x)dx=1$ i.e. $\int_0^{+\infty}kxe^{-\frac{x}{2}}dx=1$. Or on a :

$$\int_0^{\infty}\underbrace{x}_{v}\underbrace{e^{-x/2}}_{u'}dx=\underbrace{\left[\frac{e^{-x/2}}{-1/2}x\right]_0^{\infty}}_{=0}+2\int_0^{\infty}e^{-x/2}dx=2\left[\frac{e^{-x/2}}{-1/2}\right]_0^{\infty}=4$$

D'où $k=\frac{1}{4}$.

2) Soit $f_{(X,Y)}(x,y)$ la densité du couple (X,Y). Comme X et Y sont indépendantes, on a $f_{(X,Y)}(x,y)=f_X(x)f_Y(y)=\frac{1}{16}xye^{-\frac{x}{2}}e^{-\frac{y}{2}}\mathbb{1}_{[0,+\infty[^2}(x,y)$

Soit h une fonction borélienne bornée sur $\mathbb{R}$. On calcule :

$$\mathbb{E}(h(X+Y))=\iint_{\mathbb{R}^2}h(x+y)f_{(X,Y)}(x,y)\,dxdy= (*)$$

$$\frac{1}{16}\int_0^{\infty}\int_0^{\infty}h(x+y)xy\,e^{-\frac{1}{2}(x+y)}dxdy.$$ Posons $u=x+y$, $v=y$

$$\frac{D(u,v)}{D(x,y)}=\begin{vmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}1 & 1\\ 0 & 1\end{vmatrix}=1.$$ D'où

$$(*)=\frac{1}{16}\iint_{0\leq v\leq u<+\infty}h(u)(u-v)v\,e^{-\frac{1}{2}u}du\,dv=\frac{1}{16}\int_0^{+\infty}\left(\int_0^u(u-v)v\,dv\right)h(u)e^{-\frac{1}{2}u}du$$

$$=\frac{1}{16}\int_0^{+\infty}h(u)\frac{u^3}{6}e^{-\frac{1}{2}u}du=\int_0^{+\infty}h(u)\frac{1}{96}u^3e^{-\frac{1}{2}u}du.$$ Ceci montre que $X+Y$ a une densité égale à $f_{X+Y}(u)=\frac{1}{96}u^3e^{-\frac{1}{2}u}\mathbb{1}_{\mathbb{R}_+}(u)$.

$$\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y)=2\,\mathbb{E}(X)=\frac{1}{2}\int_0^{+\infty}\underbrace{x^2}_{u}\underbrace{e^{-x/2}}_{v'}dx=$$

$$\frac{1}{2}\left(\underbrace{\left[x^2\frac{e^{-x/2}}{-1/2}\right]_0^{\infty}}_{=0}-\int_0^{\infty}2x\frac{e^{-x/2}}{-1/2}dx\right)=2\int_0^{\infty}xe^{-x/2}dx=2\times 4=8.$$

$\operatorname{Var}Z=\operatorname{Var}X+\operatorname{Var}Y$ (indépendance) donc :

$\operatorname{Var} Z = 2 \operatorname{Var} X = 2\left(\mathbb{E}(X^2) - (\mathbb{E}(X))^2\right)$. Calculons $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \frac{1}{4}\int_0^\infty x^3 e^{-x/2}\,dx = \frac{1}{4}\left[\left[x^3 \frac{e^{-x/2}}{-1/2}\right]_0^\infty - \int_0^\infty 3x^2 \frac{e^{-x/2}}{-1/2}\,dx\right]$$

$$= \frac{3}{2}\int_0^\infty x^2 e^{-x/2}\,dx = 3\left(2\mathbb{E}(X)\right) = 6\,\mathbb{E}(X) = 6\times 4 = 24.$$

D'où $\operatorname{Var} Z = 2(24-16) = 2\times 8 = 16$.

Exercice 3 : 1) Soit X le nombre de votants pour A sur les 998000 indifférents. X est de loi binomiale $\mathcal{B}(n = 998000, p = 1/2)$.

a) $\mathbb{P}(A \text{ est élu}) = \mathbb{P}(X > 498000)$. Comme n est grand et $p = 1/2$ on utilise l'approximation normale :

$$\mathbb{P}(X > 498000) = \mathbb{P}\left(\frac{X-np}{\sqrt{np(1-p)}} > \underbrace{\frac{498000 - 499000}{499{,}5}}_{-2}\right) = \int_{-2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt$$

$$= 0{,}9772.$$

b) D'après le cours, on sait que

$$\mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq x\sqrt{\frac{p(1-p)}{n}}\right) = \int_{-x}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt = F(x) - F(-x) \quad (*)$$

(sinon on le rétablit en écrivant :

$$\mathbb{P}\left(\left|\frac{X-np}{\sqrt{n(1-p)}}\right| \leq x\right) \cong \int_{-x}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt \quad \text{(approximation normale)}$$

Si $x = 2{,}58$, on a $(*) = 0{,}99$ et l'intervalle de confiance de $\frac{X}{n}$ est $\left[p - x\sqrt{\frac{p(1-p)}{n}}, p + x\sqrt{\frac{p(1-p)}{n}}\right]$

$= [0{,}4987, 0{,}5013]$. Mais attention la

proportion de votants pour A est

$$\frac{X + 2000}{10^6} = \frac{X}{n}\cdot 0{,}998 + 0{,}02 \quad (n = 998000).$$

Le véritable intervalle de confiance est en fait égal à :

$$[0{,}4987 \times 0{,}998 + 0{,}02 \;;\; 0{,}5023 \times 0{,}998 + 0{,}02]$$

$$= [0{,}4997 ; 0{,}5023].$$

2) La v.a. Y = le nombre de votants pour B, suit la loi binomiale $\mathcal{B}(n = 998000, p)$. On cherche p pour que

$\mathbb{P}(Y > 500.000) \geq 0{,}99$. Autrement dit

$$\mathbb{P}\left(\frac{Y-np}{\sqrt{np(1-p)}} > \frac{500.000 - 998.000\times p}{\sqrt{998.000\times p(1-p)}}\right) \geq 0{,}99. \quad (*)$$

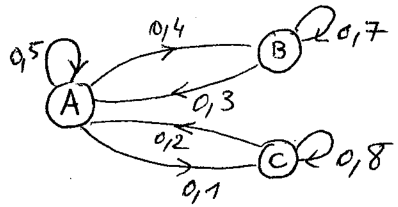
Si on note $x_p = \dfrac{500.000 - 998.000\times p}{\sqrt{998.000\times p(1-p)}}$ et $N = \dfrac{Y-np}{\sqrt{np(1-p)}}$

et qu'on utilise l'approximation normale pour N, la condition $(*)$ s'écrit

$$\mathbb{P}(N > x_p) = \int_{x_p}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt = 1 - F(x_p) \geq 0{,}99.$$

D'où $x_p \leq -2{,}33$. Compte tenu de l'expression de x_p, on obtient une inéquation (ou équation : $x_p = -2{,}33$) qui se ramène à une équation du second degré en p qu'on peut facilement déterminer avec une calculette. On trouve en fait : $p \geq 0{,}5022$ (mais cette valeur n'était pas demandée).

Exercice 4 : 1) La chaîne de Markov qui décrit les changements dus à la campagne de publicité est :



La matrice des transitions vaut :

$$P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,7 & 0 \\ 0,2 & 0 & 0,8 \end{pmatrix} \end{array}$$

La loi initiale de la chaîne est $\vec{C_0} = (0,3\,;\,0,5\,;\,0,2)$.

L'état du marché après la 1re campagne de publicité est l'état à l'instant $n=1$ de la chaîne de Markov i.e.

$$\vec{C_1} = \vec{C_0}\,P = (0,34\,;\,0,47\,;\,0,19)$$

i.e. 34% sont abonnés à A, 47% à B et 19% à C.

Après deux campagnes de publicité l'état du marché est donné par le vecteur $\vec{C_2} = \vec{C_1}\,P = (0,349\,;\,0,465\,;\,0,186)$

La matrice P étant ergodique régulière, la chaîne de Markov admet une loi stationnaire $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ telle que

$$\vec{\alpha}P = \vec{\alpha} \iff \begin{cases} 5\alpha_1 + 3\alpha_2 + 2\alpha_3 = 10\alpha_1 \\ 4\alpha_1 + 7\alpha_2 = 10\alpha_2 \\ \alpha_1 + 8\alpha_3 = 10\alpha_3 \\ \alpha_1 + \alpha_2 + \alpha_3 = 1 \end{cases} \iff \begin{array}{l} \alpha_1 = 2\alpha_3 \\ \alpha_2 = \frac{4}{3}\alpha_1 = \frac{8}{3}\alpha_3 \\ 2\alpha_3 + \frac{8}{3}\alpha_3 + \alpha_3 = 1 \end{array}$$

D'où $\alpha_3 = \frac{3}{17} = 0,1764\,;\ \alpha_2 = \frac{8}{17} = 0,4706\,;\ \alpha_1 = 0,3529$.

L'état stationnaire du marché est 35,29 % pour A, 47,06% pour B et 17,64% pour C.