L2 sciences économiques **Année universitaire 2014 – 2015**

Université Paris 1

Partiel de statistiques
Mardi 5 mai – Durée 2 heures
Sujet 1

Les exercices sont indépendants. Au sein d'un exercice, il est conseillé de traiter les questions dans l'ordre. Le barème est indicatif. L'utilisation de documents, calculatrices, téléphones portables ou tout autre appareil électronique, est interdite. Les réponses devront être soigneusement argumentées et justifiées. Vous pouvez laisser les résultats sous la forme de fractions, sans évaluer les fonctions usuelles comme exp (aussi noté e) ou ln. Si les notations C_n^p, A_n^p et/ou $\binom{n}{p}$ sont utilisées, leurs définitions en fonction de la factorielle devront être rappelées.

Exercice 1 (*8 points*)

On étudie une urne contenant 5 jetons numérotés de 1 à 5. On tire dans cette urne deux jetons, successivement et sans remise.

Question 1 Donner l'univers Ω et la probabilité $\mathbb{P}$ de l'expérience aléatoire.

Correction

On définit l'ensemble des jetons par :

$$J = \{J_1, J_2, J_3, J_4, J_5\},$$

où l'indice désigne le numéro du jeton. Comme on a un tirage sans remise, l'univers est celui des arrangements de 2 éléments parmi 6, soit

$$\Omega = \{(j_1, j_2) \in J^2 | j_1 \neq j_2\}.$$

On a $|\Omega| = A_5^2 = 5 \times 4 = 20$. Sans mention contraire dans l'énoncé, on peut poser que chaque événement élémentaire est équiprobable, c'est-à-dire que chaque couple de jetons a autant de chances de sortir que n'importe quel autre. On choisit donc la probabilité uniforme sur Ω, soit donc $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ pour tout $A \subset \Omega$.

On définit les deux variables aléatoires X et Y de la façon suivante. X vaut 1 si on obtient au moins un numéro pair sur l'un des deux jetons tirés et 0 sinon. Y vaut la valeur absolue de la différence entre les numéros des deux jetons. Par exemple, si on tire d'abord le jeton 2 puis le jeton 3, X vaut alors 1 (car 2 est pair) et Y vaut 1 (c'est-à-dire $|2 - 3|$).

Question 2 Déterminer $X(\Omega)$ et $Y(\Omega)$.

Correction

Par définition de X, $X(\Omega) = \{0, 1\}$. D'autre part $Y(j_1, j_2) = |j_1 - j_2|$ en identifiant un jeton à son numéro. Or j_1 et j_2 prennent avec cette convention toutes valeurs entières entre 1 et 5, avec toujours $j_1 \neq j_2$. On en déduit donc que $Y(\Omega) = \{1, 2, 3, 4\}$.

Question 3 Donner la loi jointe du couple (X, Y). Pour simplifier le calcul de la loi jointe, il est vivement conseillé de construire pour chaque variable un tableau donnant la **valeur** prise par la variable en fonction du résultat du tirage.

Correction

On commence par construire les tableaux conseillés. Pour X, on a ainsi :

(j_1,j_2)	J_1	J_2	J_3	J_4	J_5
J_1		1	0	1	0
J_2	1		1	1	1
J_3	0	1		1	0
J_4	1	1	1		1
J_5	0	1	0	1	

et pour Y :

(j_1,j_2)	J_1	J_2	J_3	J_4	J_5
J_1		1	2	3	4
J_2	1		1	2	3
J_3	2	1		1	2
J_4	3	2	1		1
J_5	4	3	2	1	

Chaque couple de jetons du tableau à la probabilité $\frac{1}{20}$ d'être tiré. Pour obtenir $\mathbb{P}(X = x, Y = y)$ il suffit de compter le nombre de couples dont l'image par (X, Y) est (x, y) puis de diviser ce nombre par 20. On obtient la loi jointe suivante :

$\mathbb{P}(X = x, Y = y)$	$y = 1$	$y = 2$	$y = 3$	$y = 4$
$x = 0$	$\frac{0}{20}$	$\frac{4}{20}$	$\frac{0}{20}$	$\frac{2}{20}$
$x = 1$	$\frac{8}{20}$	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{0}{20}$

Question 4 Donner les lois marginales de X et Y.

Correction

On utilise le fait que $\mathbb{P}(X = x) = \sum_{y\in Y(\Omega)} \mathbb{P}(X = x,Y = y)$. On trouve ainsi pour X

x	0	1
$\mathbb{P}(X = x)$	$\frac{6}{20}$	$\frac{14}{20}$

et pour Y

y	1	2	3	4
$\mathbb{P}(Y = y)$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{4}{20}$	$\frac{2}{20}$

Question 5 Démontrer que X et Y ne sont pas indépendantes.

Correction

Il suffit de trouver x et y tels que $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x)\mathbb{P}(Y = y)$. Or, $\mathbb{P}(X = 0, Y = 1) = 0$ alors que $\mathbb{P}(X = 0)\mathbb{P}(Y = 1) = \frac{6}{20} \times \frac{8}{20} \neq 0$.

Question 6 Calculer la loi conditionnelle de Y sachant $X = 0$.

Correction

Il suffit de calculer les valeurs de $\frac{\mathbb{P}(X=0,Y=y)}{\mathbb{P}(X=0)}$. On obtient le tableau suivant :

y	1	2	3	4
$\mathbb{P}(Y = y \mid X = 0)$	0	$\frac{2}{3}$	0	$\frac{1}{3}$

Exercice 2 (*3,5 points*)

Un fabricant produit de fausses pièces en métal pour un jeu de société. Il annonce au concepteur du jeu que le diamètre des pièces produites suit une loi normale de moyenne $\mu = 2$ cm et d'écart-type $\sigma = 0{,}1$ cm.

Question 1 Le concepteur du jeu a prévu une boîte de diamètre 2,2 cm. Calculer le pourcentage des pièces produites par le fabricant qui ne pourront pas rentrer dans la boîte. On supposera qu'une pièce de diamètre inférieur ou égal à 2,2 cm rentre dans la boîte.

Correction

Soit D la variable aléatoire représentant le diamètre des pièces. On sait que $D \sim \mathcal{N}(2{,}0{,}1^2)$. Donc la variable aléatoire $X = \frac{D-2}{0{,}1}$ suit une loi normale centrée et réduite $\mathcal{N}(0{,}1^2)$.
On cherche à déterminer $\mathbb{P}(D > 2{,}2)$. Or

$$D > 2{,}2 \Leftrightarrow D - 2 > 0{,}2 \Leftrightarrow \frac{D-2}{0{,}1} > 2 \Leftrightarrow X > 2.$$

Donc $\mathbb{P}(D > 2{,}2) = \mathbb{P}(X > 2)$. Or, $\mathbb{P}(X > 2) = 1 - F_X(2)$, par définition de la fonction de répartition. La table de la loi normale centrée réduite donne $F_X(2) = 0{,}9772$ et donc $\mathbb{P}(D > 2{,}2) = 0{,}0228$. Il y aura donc environ 2,3 % des pièces qui ne pourront pas rentrer dans la boîte.

Question 2 Le concepteur du jeu souhaite qu'au moins 90 % des pièces aient un diamètre compris entre 1,9 cm et 2,1 cm. Est-ce le cas ?

Correction

On veut connaître la probabilité $\mathbb{P}(D \in [1{,}9; 2{,}1])$ (sachant que l'inclusion ou non des bornes de l'intervalle n'importe pas car la variable aléatoire est continue). Or

$$\begin{aligned} 1{,}9 \leq D \leq 2{,}1 &\Leftrightarrow -0{,}1 \leq D - 2 \leq 0{,}1 \\ &\Leftrightarrow -1 \leq \frac{D-2}{0{,}1} \leq 1 \\ &\Leftrightarrow -1 \leq X \leq 1. \end{aligned}$$

De ce fait, on doit donc calculer $\mathbb{P}(X \in [-1, 1]) = F_X(1) - F_X(-1)$ par définition de la fonction de répartition (et continuité de X). On lit dans la table de la loi normale centrée réduite que $F_X(1) = 0{,}8413$. De plus, par symétrie de la loi normale, on sait que $F_X(-1) = 1 - F_X(1)$ et donc que $F_X(-1) = 0{,}1587$. On en déduit que

$$\mathbb{P}(X \in [-1, 1]) = 0{,}6826 = \mathbb{P}(D \in [1{,}9; 2{,}1]).$$

On constate donc que les exigences du concepteur ne sont pas satisfaites.

Question 3 Déterminer un intervalle de la forme $[2-t, 2+t]$ tel que 90 % des pièces aient un diamètre appartenant à l'intervalle.

Correction

On a

$$\begin{aligned} 2-t \leq D \leq 2+t &\Leftrightarrow -t \leq D-2 \leq t \\ &\Leftrightarrow -10t \leq \frac{D-2}{0{,}1} \leq 10t \\ &\Leftrightarrow -10t \leq X \leq 10t. \end{aligned}$$

Cherchons alors un intervalle de la forme $[-u,u]$ tel que $\mathbb{P}(X \in [-u,u]) = 0{,}9$. On sait que $\mathbb{P}(X \in [-u,u]) = F_X(u) - F_X(-u) = 2F_X(u) - 1$ la deuxième égalité venant de la symétrie de la loi normale centrée réduite. On cherche donc u tel que $F_X(u) = 0{,}95$. La lecture de la table donne $u = 1{,}65$ (avec $F_X(1{,}65) = 0{,}9505 \simeq 0{,}95$). Or,

$$-1{,}65 \leq X \leq 1{,}65 \Leftrightarrow -0{,}165 \times 10 \leq X \leq 0{,}165 \times 10.$$

Si on pose $t = 0{,}165$, on a donc

$$-u \leq X \leq u \Leftrightarrow -10t \leq X \leq 10t \Leftrightarrow 2-t \leq D \leq 2+t.$$

L'intervalle recherché est donc $[1{,}835; 2{,}165]$.

Exercice 3 (*6 points*)

Dans cet exercice, on pourra, si nécessaire, *utiliser les valeurs approchées suivantes :* $e^{-4} \simeq 0{,}02$, $e^{-2} \simeq 0{,}14$, $e^{-1} \simeq 0{,}37$, $e^{-\frac{1}{2}} \simeq 0{,}61$, $e^{-\frac{1}{4}} \simeq 0{,}78$, $e^{\frac{1}{4}} \simeq 1{,}28$, $e^{\frac{1}{2}} \simeq 1{,}65$, $e \simeq 2{,}71$, $e^2 \simeq 7{,}39$, $e^4 \simeq 54{,}60$.

Un investisseur s'intéresse aux créations d'entreprises innovantes. La durée d'activité d'une telle entreprise est modélisée par une variable aléatoire continue, D, qui est supposée suivre une loi exponentielle $\mathcal{E}(\lambda)$. D est mesurée en années. La durée de vie moyenne de ces entreprises est de 4 ans.

Question 1 Déterminer le paramètre λ de la loi suivie par D.

Correction

D, la variable aléatoire de durée d'activité de l'entreprise, suit une loi exponentielle, il nous reste à déterminer le paramètre λ. Or, on sait que $\mathbb{E}(D) = \frac{1}{\lambda}$ et comme par l'énoncé $\mathbb{E}(D) = 4$, on a $\lambda = \frac{1}{4}$.

Question 2 Calculer la probabilité qu'une entreprise de ce type ait une durée de vie inférieure (ou égale) à 4 ans.

Correction

$\mathbb{P}(D \leq 4) = F_D(4) = 1 - e^{-\frac{1}{4}\times 4} = 1 - e^{-1} = 1 - 0{,}37 = 0{,}63$

L'investisseur envisage le plan financier suivant : il investit un million d'euros dans une entreprise innovante qui se crée et au bout d'un an touche 1,7 millions d'euros en récupérant l'argent investi et des dividendes de 700 000 euros, à condition que l'entreprise soit encore en activité. Dans le cas contraire, c'est-à-dire si l'entreprise n'est plus en activité, l'investisseur perd l'intégralité de l'argent investi.

Question 3 Soit S la variable aléatoire qui vaut 1 si l'entreprise considérée est encore en activé au bout d'un an et 0 sinon. Donner la loi usuelle suivie par S en précisant la (les) valeur(s) de son (ou ses) paramètre(s).

Correction

Il s'agit d'une épreuve de Bernoulli : soit l'entreprise survit et peut rembourser et verser des dividences (succès, alors $S = 1$), soit elle fait faillite avant et $S = 0$. Le paramètre p de cette Benoulli est donnée par la probabilité que l'entreprise survive au moins jusqu'au moment du remboursement (c'est-à-dire un an). Or par continuité : $\mathbb{P}(D \geq 1) = \mathbb{P}(D > 1)$, et par complémentaire $\mathbb{P}(D > 1) = 1 - \mathbb{P}(D \leq 1) = 1 - F_D(1) = 1 - 1 + e^{-\frac{1}{4}} = 0{,}78$. Donc $S \sim \mathcal{B}(0{,}78)$.

Question 4 Soit M la variable aléatoire donnant le montant que l'investisseur touche au bout d'un an, exprimé en millions d'euros. Exprimer M comme une fonction de S.

Correction

M est une VA fonction de S, que l'on peut écrire comme : $M = S \times 1{,}7$. En effet, soit l'entreprise a survécu un an et l'investisseur touche 1,7 millions, soit l'entreprise n'a pas survécu et l'investisseur ne touche rien.

Question 5 Calculer l'espérance et la variance de M.

Correction

Par linéarité de l'espérance : $\mathbb{E}(M) = 1{,}7 \times \mathbb{E}(S)$ et comme S est une Bernoulli, on a $\mathbb{E}(S) = p = 0{,}78$. D'où $\mathbb{E}(M) = 1{,}7 \times 0{,}78$. Pour la variance, on sait que $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$.
On en déduit donc que : $\mathbb{V}(M) = 1{,}7^2\mathbb{V}(S) = 1{,}7^2 \times p(1-p) = 1{,}7^2 \times 0{,}78 \times 0{,}22$.

L'investisseur cherche à établir s'il ne serait pas plus judicieux de prêter moins d'argent mais à plusieurs entreprises. Il considère le cas où il investit 200 000 euros dans 5 entreprises différentes avec des dividendes de 140 000 euros au bout d'un an (selon un schéma similaire à celui des questions précédentes). En d'autres termes, pour chaque entreprise l'investisseur peut soit récupérer 340 000 euros si l'entreprise est encore en activité au bout d'un an, soit ne rien récupérer du tout. Les défaillances des entreprises sont considérées indépendantes.

Question 6 Soit A la variable aléatoire du nombre d'entreprises encore en activité au bout d'un an parmi les 5 dans lesquelles l'investisseur a placé son argent. Déterminer la loi de A et préciser la (les) valeur(s) de son (ou ses) paramètre(s).

Correction

Chaque entreprise représente une épreuve de Bernoulli de paramètre 0,78 (question précédente) et sont indépendantes par l'énoncé. A est donc la somme de 5 épreuves de Bernoulli de même paramètre (ou “identiquement distribuées”) et indépendantes, il s'agit donc d'une binomiale de paramètre 5 et de probabilité 0,78. Soit $A \sim \mathcal{B}(5; 0{,}78)$.

Question 7 Calculer la probabilité qu'une seule entreprise (parmi les 5) soit encore en activité au bout d'un an.

Correction

Dans le cas général pour $a \in A(\Omega) = \{0, 1, 2, 3, 4, 5\}$, on a

$$\mathbb{P}(A = a) = C_n^a p^a (1-p)^{n-a} = \frac{5!}{a!(5-a)!} \times 0{,}78^a \times 0{,}22^{5-a}$$

Dans le cas particulier de $a = 1$ cela donne :

$$\mathbb{P}(A = 1) = \frac{5!}{1!(5-1)!} \times 0{,}78^1 \times 0{,}22^{5-1} = 5 \times 0{,}78 \times 0{,}22^4.$$

Question 8 Soit N la variable aléatoire donnant le total que l'investisseur touche au bout d'un an, exprimé en millions d'euros. Calculer l'espérance et la variance de N.

Correction

N est une variable aléatoire fonction de A. En effet $N = 0{,}34A + 0 \times (5 - A) = 0{,}34A$. Or on sait que l'espérance de A est donnée par $\mathbb{E}(A) = np = 5 \times 0{,}78$ et par linéarité de l'espérance on a également $\mathbb{E}(N) = 0{,}34 \times \mathbb{E}(A) = 5 \times 0{,}78 \times 0{,}34 = 1{,}7 \times 0{,}78 = \mathbb{E}(S)$.
Pour la variance de N, on sait que A étant une binomiale on a $\mathbb{V}(A) = np(1-p) = 5 \times 0{,}78 \times 0{,}22$ et donc $\mathbb{V}(N) = 0{,}34^2 \mathbb{V}(A) = 0{,}34^2 \times 5 \times 0{,}78 \times 0{,}22$.

Question 9 En comparant les espérances et les variances de M et N, notamment au travers du ratio des variances, $\frac{\mathbb{V}(M)}{\mathbb{V}(N)}$, déterminer le plan d'investissement qui paraît le plus judicieux.

Correction

On remarque que N et S ont la même espérance, donc en moyenne les deux plans financiers correspondent aux mêmes gains.
Le ratio $\frac{\mathbb{V}(M)}{\mathbb{V}(N)}$ s'écrit :

$$\frac{\mathbb{V}(M)}{\mathbb{V}(N)} = \frac{1{,}7^2 \times 0{,}78 \times 0{,}22}{0{,}34^2 \times 5 \times 0{,}78 \times 0{,}22} = \frac{1{,}7^2}{0{,}34^2 \times 5} = \frac{(5 \times 0{,}34)^2}{0{,}34^2 \times 5} = 5$$

La variance de M est donc 5 fois plus grande que la variance de N. Les deux plans rapportent le même montant en moyenne mais la dispersion est beaucoup plus grande pour le plan M, et donc le « risque » plus grand. Il est probablement préférable de ce point de vue de choisir N plutôt que M.

Exercice 4 (*3 points*)

Soit f la fonction définie comme suit :

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

Avec e défini comme le réel tel que $e = e^1$ et $ln(e) = 1$, $e \simeq 2{,}71$, et a un nombre réel quelconque.

Question 1 Déterminer a de sorte que f soit la densité d'une variable aléatoire absolument continue.

Correction

Pour être la densité d'une variable aléatoire continue, f doit être positive (ce qui est le cas immédiatement) et intégrable sur $\mathbb{R}$ (ce qui est aussi le cas). Il reste à vérifier que $\int_{-\infty}^{\infty} f(x)dx = 1$.

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)dx &= \int_{-\infty}^{0} f(x)dx + \int_{0}^{a} f(x)dx + \int_{a}^{+\infty} f(x)dx \\ &= \int_{0}^{a} \frac{e^x}{e-1}dt = \frac{1}{e-1}\int_{0}^{a} e^x dx = \frac{1}{e-1}\Big[e^x\Big]_0^a = \frac{e^a-1}{e-1} \end{aligned}$$

Par respectivement la relation de Chasles, linéarité de l'intégrale et théorème fondamental.
D'où $\frac{e^a-1}{e-1} = 1$. D'où $e^a - 1 = e - 1 \Rightarrow a = 1$.

Question 2 Soit X une variable aléatoire de densité f (donc avec pour a la valeur calculée à la question précédente). Déterminer F_X la fonction de répartition de X.

Correction

On sait que F_X est donnée par la formule générale :

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

On obtient ainsi :

$$F_X(x) = \begin{cases} \int_{-\infty}^{x} 0dt = 0 & \text{si } x < 0 \\ \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0dt + \int_0^x \frac{e^t}{e-1}dt = \frac{1}{e-1}\int_0^x e^t dt \text{ (Chasles et linéarité)} & \text{si } x \in [0{,}1] \\ \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0dt + \int_0^1 \frac{e^t}{e-1}dt + \int_1^x 0dt = \int_0^1 \frac{e^t}{e-1}dt & \text{si } x > 1 \end{cases}$$

Il nous reste donc à calculer $\frac{1}{e-1}\int_0^x e^t dt$ (avec comme cas particulier $\int_0^1 \frac{e^t}{e-1}dt$).

$$\frac{1}{e-1}\int_0^x e^t dt = \frac{1}{e-1}\Big[e^x\Big]_0^x = \frac{1}{e-1}(e^x - e^0) = \frac{e^x-1}{e-1}$$

D'où, pour $x \in [0{,}1]$, $F_X(x) = \frac{e^x-1}{e-1}$ et pour $x > 1$, $F_X(x) = \frac{e-1}{e-1} = 1$.

Question 3 Calculer l'espérance de X. On pourra s'appuyer sur le calcul de la dérivée de la fonction $g(x) = xe^x - e^x$ pour obtenir une primitive utile pour le calcul de $\mathbb{E}(X)$.

Correction

On a $g(x) = xe^x - e^x$, soit $g'(x) = 1 \times e^x + x \times e^x - e^x = x \times e^x$. Autrement dit g est une primitive de la fonction $x \times e^x$. On calcule l'espérance de X. Par définition on a (et en suivant des raisonnements analogues à ceux précédents : relation de Chasles et linéarité de l'intégrale) :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 x\frac{e^x}{e-1}dx = \frac{1}{e-1}\int_0^1 xe^x dx$$

Or une primitive de xe^x est donnée par g, donc on a :

$$\mathbb{E}(X) = \frac{1}{e-1}xe^x dx = \frac{1}{e-1}\Big[xe^x - e^x\Big]_0^1 = \frac{1}{e-1}(1 \times e^1 - e^1 - 0 + e^0) = \frac{1}{e-1}$$