Examen final du 9 janvier 2020 (1ère session)

Durée : 2 heures. Tous documents interdits. Dans tout l'examen on travaille sur un espace de probabilité $(\Omega, \mathscr{A}, P)$. On utilisera l'abréviation 'v.a.' pour 'variable aléatoire'. La qualité et la rigueur de la rédaction seront prises en compte.

Exercice 1. Une inégalité de concentration. Soit X une v.a. de Poisson de paramètre $\theta > 0$.

a) Rappeler sans calcul l'espérance et la variance de X.

b) Montrer que pour tout $\varepsilon > 0$,

$$P(|X - \theta| \geq \varepsilon\sqrt{\theta}) \leq \frac{1}{\varepsilon^2}$$

Solution de l'exercice 1.

a) D'après le cours, $E(X) = \mathrm{Var}(X) = \theta$.

b) Application de l'inégalité de Bienaymé-Chebyshev :

$$P(|X - \mathbb{E}(X)| \geq \varepsilon\sqrt{\theta}) \leq \frac{\mathrm{Var}(X)}{(\varepsilon\sqrt{\theta})^2} = \frac{1}{\varepsilon^2}$$

Exercice 2. Une variable aléatoire à densité. Soit X une v.a. réelle positive de densité f donnée par

$$f(x) = cxe^{-x^2/2} \qquad x \geq 0,$$

où c est une constante positive.

a) Calculer c.

b) Calculer $E(X^2)$.

c) Donner la loi de X^2.

Solution de l'exercice 2.

a) La constante c est une constante de renormalisation, qui est choisie pour que l'intégrale de f soit égale à 1 :

$$1 = \int_0^\infty f(x)\,dx = c\int_0^\infty xe^{-x^2/2}dx = c[-e^{-x^2/2}]_0^\infty = c,$$

donc $c = 1$.

b) Par la formule de transfert puis le changement de variable $u = x^2$, on obtient

$$E(X^2) = \int_0^\infty x^2 f(x)\,dx = \int_0^\infty x^3 e^{-x^2/2}\,dx = \frac{1}{2}\int_0^\infty u e^{-u/2}\,du,$$

qui est l'espérance d'une variable aléatoire exponentielle de paramètre $1/2$, donc $E(X^2) = 2$.

c) Soit h une fonction continue bornée sur $\mathbb{R}_+$. Alors grâce au changement de variable $u = x^2$, on obtient

$$E(h(X^2)) = \int_0^\infty h(x^2) f(x)\,dx = \int_0^\infty h(u) f(\sqrt{u})\,\frac{du}{2\sqrt{u}} = \frac{1}{2}\int_0^\infty h(u) e^{-u/2}\,du.$$

Donc X^2 suit la loi exponentielle de paramètre $1/2$.

Exercice 3. Tirages des boules d'une urne. Soit $n \geq 2$ un entier. On se donne une urne contenant $n-1$ boules blanches et une boule noire. Nous supposons qu'une personne prélève uniformément au hasard une boule de l'urne, plusieurs fois de suite, jusqu'à tirer la boule noire. Les tirages se font soit tous avec remise (on remet la boule tirée dans l'urne avant le prochain tirage, question b) soit tous sans remise (question c).
Soit X_k la variable aléatoire égale à 0 si la boule tirée au k-ième tirage est blanche, égale à 1 si elle est noire. Nous nous intéressons à la loi de la variable aléatoire

$$T_n = \min\{k \geq 1 : X_k = 1\},$$

qui est le rang du premier tirage de la boule noire.

a) Soit U une v.a. uniforme sur $[0,1]$ et V une v.a. exponentielle de paramètre 1. Rappeler la densité de U et celle de V, et calculer leurs fonctions de répartition.

On traitera **au choix, soit la question b, soit la question c**.

b) On suppose dans cette question que les tirages ont lieu **avec remise**.
   i) Montrer que T_n suit une loi géométrique que l'on précisera.
   ii) Calculer à l'aide du paramètre $a = 1 - 1/n$ la fonction génératrice de T_n, son espérance et sa variance (puis donner le résultat final en fonction de n).
   iii) Calculer $P(T_n > k)$ pour tout entier k, puis donner la fonction de répartition de T_n/n.
   iv) En déduire que T_n/n converge en loi vers une v.a. que l'on précisera.

c) On suppose dans cette question que les tirages ont lieu **sans remise**, de sorte que l'on fait au maximum n tirages.
   i) Montrer que pour tout $k \in \{2, \ldots, n\}$

$$\begin{aligned} P(T_n \geq k) &= P(X_1 = \cdots = X_{k-1} = 0) \\ &= P(X_1 = 0) \prod_{i=2}^{k-1} P(X_i = 0 \mid X_1 = \cdots = X_{i-1} = 0), \end{aligned}$$

   où le produit est pris égal à 1 si $k = 2$.

ii) Montrer que

$$P(T_n \geq k) = \prod_{i=1}^{k-1} \left(1 - \frac{1}{n-i+1}\right).$$

iii) En déduire que

$$P(T_n \geq k) = \frac{n-k+1}{n},$$

puis que $P(T_n = k) = 1/n$ pour tout $k \in \{1, \ldots, n\}$.

iv) En utilisant la fonction de répartition de T_n, montrer que T_n/n converge en loi vers une v.a. que l'on précisera.

Solution de l'exercice 3.

a) Soit U une v.a. uniforme sur $[0,1]$ et V une v.a. exponentielle de paramètre 1. La densité de U vaut 1 sur $[0,1]$ et 0 ailleurs. Pour tout $t < 0$, $F_U(t) = P(U \leq t) = 0$. Pour tout $t > 1$, $F_U(t) = P(U \leq t) = 1$. Pour tout $t \in [0,1]$,

$$F_U(t) = P(U \leq t) = \int_0^t dx = t.$$

La densité de V vaut e^{-x} pour $x \geq 0$ et 0 ailleurs. Pour tout $t < 0$, $F_V(t) = P(V \leq t) = 0$. Pour tout $t \geq 0$,

$$F_V(t) = \int_0^t e^{-x} dx = 1 - e^{-t}.$$

b) On suppose dans cette question que les tirages ont lieu avec remise.

i) Montrer que T_n suit une loi géométrique que l'on précisera.

À chaque tirage indépendamment, la boule noire est tirée avec la même probabilité égale à $1/n$, donc les (X_k) sont des variables de Bernoulli indépendantes de paramètre $1/n$. Ainsi pour tout $k \geq 1$,

$$\begin{aligned} P(T_n = k) &= P(X_1 = \cdots = X_{k-1} = 0, X_k = 1) \\ &= \left[\prod_{i=1}^{k-1} P(X_i = 0)\right] P(X_k = 1) = \left(1 - \frac{1}{n}\right)^{k-1} \frac{1}{n} \end{aligned}$$

ii) Soit f_n la fonction génératrice de T_n, définie par

$$f_n(s) = E\left(s^{T_n}\right) = \sum_{k \geq 1} P(T_n = k) s^k = \frac{s}{n} \sum_{k \geq 1} (1 - 1/n)^{k-1} s^{k-1} = \frac{s/n}{1 - (1 - 1/n)s}$$

En notant $a = 1 - 1/n$, on obtient la formulation classique

$$f_n(s) = \frac{(1-a)s}{1-as} = \frac{1-a}{a}\left(-1 + \frac{1}{1-as}\right),$$

de sorte que

$$f_n'(s) = \frac{1-a}{(1-as)^2}$$

et

$$f_n''(s) = \frac{2a(1-a)}{(1-as)^3}.$$

Ainsi en prenant $s = 1$, on obtient

$$E(T_n) = f_n'(1) = \frac{1}{1-a} = n$$

et

$$E(T_n(T_n-1)) = f_n''(1) = \frac{2a}{(1-a)^2}.$$

Ainsi

$$\begin{aligned}\mathrm{Var}(T_n) &= E(T_n^2) - (E(T_n))^2 \\ &= E(T_n(T_n-1)) + E(T_n) - (E(T_n))^2 = \frac{2a}{(1-a)^2} + \frac{1}{1-a} - \frac{1}{(1-a)^2},\end{aligned}$$

ce qui donne

$$\mathrm{Var}(T_n) = \frac{a}{(1-a)^2} = n(n-1).$$

iii) Soit F_n la fonction de répartition de T_n/n. Noter d'abord que pour tout entier $k \geq 1$,

$$P(T_n > k) = \left(1 - \frac{1}{n}\right)^k$$

On a alors

$$F_n(t) = P(T_n/n \leq t) = P(T_n \leq tn) = 1 - P(T_n > tn) = 1 - \left(1 - \frac{1}{n}\right)^{[tn]},$$

où $[x]$ désigne la partie entière de x.

iv) Donc lorsque $n \to \infty$,

$$F_n(t) \longrightarrow 1 - e^{-t},$$

qui est la fonction de répartition de la v.a. V (voir première question). Donc T_n/n converge en loi vers une v.a. exponentielle de paramètre 1.

c) On suppose dans cette question que les tirages ont lieu sans remise.

i) Pour tout $k \in \{2, \ldots, n\}$, par définition du min,

$$P(T_n \geq k) = P(X_1 = \cdots = X_{k-1} = 0).$$

La formule demandée se démontre par récurrence sur k grâce à la formule des probabilités conditionnelles. En effet, pour tout $k \geq 2$, si la formule est vraie au rang k alors

$$\begin{aligned}P(X_1 = \cdots = X_k = 0) &= P(X_k = 0 \,|\, X_1 = \cdots = X_{k-1} = 0)\, P(X_1 = \cdots = X_{k-1} = 0) \\ &= P(X_k = 0 \,|\, X_1 = \cdots = X_{k-1} = 0) P(X_1 = 0) \prod_{i=2}^{k-1} P(X_i = 0 \,|\, X_1 = \cdots = X_{i-1} = 0),\end{aligned}$$

ce qui montre la formule au rang $k+1$.

ii) Pour montrer que

$$P(T_n \geq k) = \prod_{i=1}^{k-1} \left(1 - \frac{1}{n-i+1}\right),$$

il suffit de partir du résultat de la question précédente et de voir que ($P(X_1 = 0) = 1 - 1/n$ et que)

$$P(X_i = 0 \,|\, X_1 = \cdots = X_{i-1} = 0) = 1 - \frac{1}{n-i+1}.$$

En effet, sachant $\{X_1 = \cdots = X_{i-1} = 0\}$, la boule noire n'a pas encore été tirée, et il reste donc dans l'urne au i-ème tirage, $n - i + 1$ boules dont une (seule) noire. La probabilité de ne pas la tirer est donc $1 - 1/(n - i + 1)$.

iii) Reprenons l'égalité de la question précédente :

$$P(T_n \geq k) = \prod_{i=1}^{k-1} \left(1 - \frac{1}{n-i+1}\right) = \prod_{i=1}^{k-1} \frac{n-i}{n-i+1} = \frac{\prod_{i=1}^{k-1}(n-i)}{\prod_{j=1}^{k-1}(n-j+1)} = \frac{\prod_{i=1}^{k-1}(n-i)}{\prod_{i=0}^{k-2}(n-i)},$$

ce qui donne le résultat attendu :

$$P(T_n \geq k) = \frac{n-k+1}{n}.$$

On conclut en remarquant que

$$P(T_n = k) = P(T_n \geq k) - P(T_n \geq k+1) = \frac{1}{n}.$$

iv) Comme dans le cas des tirages avec remise, on calcule

$$P(T_n \leq k) = 1 - P(T_n > k) = \frac{k}{n}.$$

Ainsi, pour tout $t \in [0, 1]$,

$$P(T_n/n \leq t) = \frac{[tn]}{n},$$

qui converge lorsque $n \to \infty$ vers t, qui est la fonction de répartition de U (voir première question), v.a. uniforme sur $[0, 1]$.