BCPST $\boldsymbol{2_A}$

Devoir à préparer pour le lundi 31 mars 2025

Ne pas faire la partie 2 qui est très proche de ce que nous avons fait pour le concours blanc.

Théorème :

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires, on définit pour tout $n \in \mathbb{N}^*$: $M_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$

Si les X_n sont indépendantes de même loi, admettant une espérance μ et une variance σ^2 non nulle

alors $M_n^* = \dfrac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ converge en loi vers la loi normale centrée réduite.

(Autrement dit : $\forall x \in \mathbb{R}, \quad \lim\limits_{n\to+\infty} P(M_n^* \leqslant x) = \Phi(x)$)

où Φ est la fonction de répartition de la loi $\mathcal{N}(0,1)$

Ce théorème permet d'affirmer :

❶ Si les X_n sont indépendantes de même loi, admettant une espérance μ et une variance σ^2 non nulle, alors, pour n assez grand, la loi de $\dfrac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ peut être approchée par une loi normale centrée réduite.

❷ Si Y_n suit une loi binomiale de paramètres (n,p) alors :

pour n assez grand, la loi de $\dfrac{Y_n - np}{\sqrt{np(1-p)}}$ peut être approchée par une loi normale centrée réduite.

(Théorème de Moivre-Laplace).

Partie 1

Soient n et s des entiers supérieurs ou égaux à 2.

On considére une urne contenant des boules de couleurs distinctes $C_1, \dots, C_s$.

Pour chaque $i \in [\![1,s]\!]$, les boules de couleur C_i sont en proportion $p_i > 0$; on a donc $\displaystyle\sum_{i=1}^{s} p_i = 1$.

On effectue n tirages successifs d'une boule avec remise dans cette urne.

Pour tout i de $\{1,\dots,s\}$, on note X_i la variable aléatoire égale au nombre de boules de couleur C_i obtenues à l'issue des n tirages. On remarque que la variable X_i dépend de n.

On définit la variable aléatoire U_n par :

$$U_n = \sum_{i=1}^{s} \frac{(X_i - np_i)^2}{np_i}$$

A. Etude des variables X_i.

1) Soit $i \in \{1,\dots,s\}$, déterminer la loi de X_i, son espérance et sa variance.
2) Soit $(i,j) \in \{1,\dots,s\}^2$ tel que $i \neq j$.
   a) Déterminer la loi de $X_i + X_j$ et sa variance.
   b) En déduire que : $\operatorname{cov}(X_i, X_j) = -np_ip_j$.

B. On suppose dans cette partie que $s = 2$.

1) Montrer que : $U_n = {Z_1}^2$ où $Z_1 = \dfrac{X_1 - np_1}{\sqrt{np_1p_2}}$.
2) Par quelle loi peut-on approcher la loi de Z_1, lorsque n est grand ?

C. On suppose dans cette partie que $s = 3$ et que $p_1 = p_2 = \frac{1}{4}$ et $p_3 = \frac{1}{2}$.

On pose : $\quad Z_1 = \dfrac{2}{\sqrt{n}}\left(X_3 - \dfrac{n}{2}\right) \quad$ et $\quad Z_2 = \sqrt{\dfrac{2}{n}}\,(X_1 - X_2)$

1) Montrer que $U_n = {Z_1}^2 + {Z_2}^2$. *(On utilisera la relation : $X_1 + X_2 + X_3 = n$)*

2) Déterminer l' espérance et la variance de Z_1 et Z_2, puis la covariance de (Z_1, Z_2).

3) Par quelle loi peut-on approcher la loi de Z_1, lorsque n est grand ?

4) Pour i élément de $\{1, \ldots, n\}$, on définit la variable T_i par :
$T_i = 1$ si au i-iéme tirage on a obtenu une boule de couleur C_1,
$T_i = -1$ si au i-iéme tirage on a obtenu une boule de couleur C_2,
$T_i = 0$ si au i-iéme tirage on a obtenu une boule de couleur C_3.
 a) Exprimer $X_1 - X_2$ à l'aide des variables T_i.
 b) En déduire que l'on peut approcher la loi de Z_2, quand n est grand, par la loi normale centrée réduite.

5) a) Ecrire une fonction `Tirage()` permettant de simuler le tirage avec remise de 100 boules dans une urne contenant des boules de couleur C_1 ou C_2 ou C_3 en proportion respectivement $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$.
La fonction devra renvoyer une liste `C` où l'élément `C[i]` vaut 1,2 ou 3 et représente la couleur de la i-iéme boule tirée (C_1, C_2 ou C_3). On utilisera la fonction `random.randint(0,3)` retourne un entier aléatoire compris entre 0 et 3.
 b) Ecrire une fonction `Difference(C)` de paramétre la liste obtenu par la fonction `Tirage()` et qui retourne la valeur de $X_1 - X_2$.

D. On suppose désormais $s = 4$ et $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$.

Notation : Pour toute matrice A on note A^T sa matrice transposée.

On rappelle que pour toutes matrices A et B telles que le produit AB existe, $(AB)^T = B^T A^T$.

– –

Pour i élément de $\{1, 2, 3, 4\}$, on pose $\quad Y_i = \dfrac{X_i - np_i}{\sqrt{np_i}}$.

On note M la matrice carrée d'ordre 4 dont le coefficient en ligne i et colonne j vaut $\operatorname{cov}(Y_i, Y_j)$.

On définit N la matrice carrée d'ordre 4 dont tous les coefficients sont égaux à $\frac{1}{4}$.

1) Exprimer M en fonction de N et de I où I est la matrice unité.

2) On définit 4 vecteurs de $\mathscr{M}_{4,1}(\mathbb{R})$:

$$e_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \qquad e_2 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\\0\end{pmatrix}, \qquad e_3 = \frac{1}{2\sqrt{3}}\begin{pmatrix}1\\1\\1\\-3\end{pmatrix}, \qquad e_4 = \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

 a) Montrer que ces 4 vecteurs sont des vecteurs propres de N et qu'ils forment une base de $\mathscr{M}_{4,1}(\mathbb{R})$.

 On note Q la matrice de passage de la base canonique de $\mathscr{M}_{4,1}(\mathbb{R})$ à la base (e_1, e_2, e_3, e_4).

 b) Quelle est l'inverse de la matrice Q ?
 c) Expliciter la matrice $Q^T M Q$.

3) On définit les variables Z_1, Z_2, Z_3 et Z_4 par $\begin{pmatrix}Z_1\\Z_2\\Z_3\\Z_4\end{pmatrix} = Q^T \begin{pmatrix}Y_1\\Y_2\\Y_3\\Y_4\end{pmatrix}$.

a) Exprimer chaque Z_i en fonction de $Y_1, \dots, Y_4$ et montrer que $Z_4 = 0$.

b) Montrer que, pour $i = 1, 2, 3$, les variables Z_i sont centrées.

c) Montrer que $U_n = Z_1^2 + Z_2^2 + Z_3^2$. *(On pourra calculer* $(Y_1 \quad Y_2 \quad Y_3 \quad Y_4)\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}$ *)*

Partie 2 *(Cette partie a été faite dans le DS du concours blanc)*

1) Soit r un entier supérieur ou égal à 2 .

a) Donner la limite de $x^{\frac{r}{2}+1}e^{-\frac{x}{2}}$ quand x tend vers $+\infty$.

b) Justifier qu'il existe un réel $A > 0$ tel que : $\forall x > A, \quad x^{\frac{r}{2}-1}e^{-\frac{x}{2}} \leqslant \dfrac{1}{x^2}$

c) En déduire la convergence de l'intégrale $\displaystyle\int_0^{+\infty} x^{\frac{r}{2}-1}e^{-\frac{x}{2}}\,\mathrm{d}x$. On notera J_r sa valeur.

2) a) Pour tout réel t strictement positif, établir la convergence de l'intégrale $\displaystyle\int_t^{+\infty} x^{-\frac{1}{2}}e^{-\frac{x}{2}}dx$.

On note $G(t)$ sa valeur.

b) Montrer que, pour tout réel t strictement positif, $G(t) = 2\sqrt{2\pi}\,\left(1 - \Phi(\sqrt{t})\right)$

où Φ désigne la fonction de répartition de la loi normale centrée réduite.

Indication : on effectuera le changement de variable $u = \sqrt{x}$.

c) En déduire la convergence et la valeur J_1 de $\displaystyle\int_0^{+\infty} x^{-\frac{1}{2}}e^{-\frac{x}{2}}dx$

3) Soit r un entier naturel non nul. On définit la fonction f_r par :

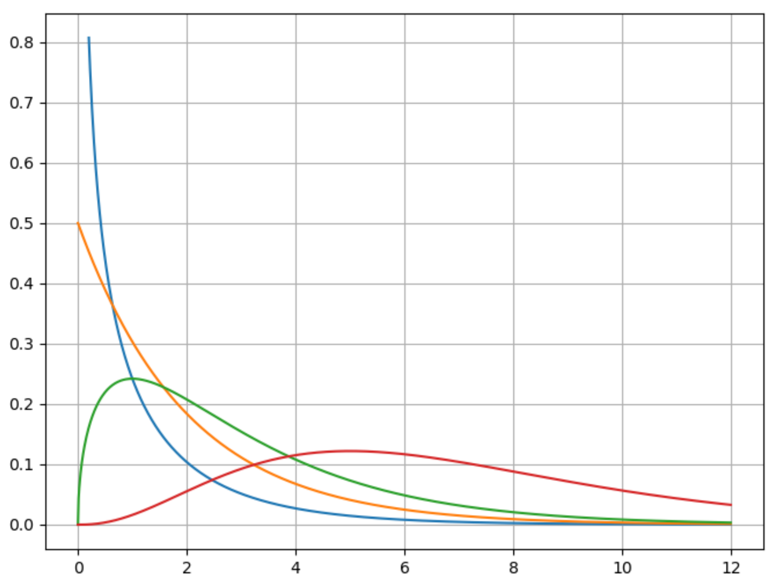
$$\forall x > 0, \quad f_r(x) = \frac{1}{J_r}\,x^{\frac{r}{2}-1}\,e^{-\frac{x}{2}} \quad \text{et} \quad \forall x \leqslant 0, \quad f_r(x) = 0$$

a) Montrer que f_r est une densité de probabilité.

On dit qu'une variable aléatoire X suit la loi du χ^2 (*lire "khi-deux"*) à r degrés de liberté si, et seulement si, X admet f_r pour densité.

b) Quelle loi usuelle reconnait-on pour $r = 2$?

c) Recopier le graphique ci dessous en identifiant les courbes représentatives des quatre fonctions f_1, f_2, f_3, f_7 en justifiant avec précision la réponse.



Partie 3

En reprenant les notations de la **Partie 1**, avec s entier quelconque supérieur ou égal à 2. On admet que :

n étant grand, la variable U_n suit la loi du χ^2 à $s-1$ degrés de liberté.

– –

On considère un algorithme générateur de nombres entiers aléatoires compris entre 1 et 4.

On utilise cet algorithme pour créer un échantillon de 10000 nombres compris entre 1 et 4.

On obtient alors :

2602 fois le nombre 1, 2534 fois le nombre 2, 2422 fois le nombre 3 et 2442 fois le nombre 4.

On se propose de tester la fiabilité de cet algorithme par l'examen de l'échantillon précédent.

On fait l'hypothèse que le nombre (aléatoire) fourni par le générateur suit une loi uniforme sur l'ensemble $\{1, 2, 3, 4\}$.

On donne les valeurs suivantes :

- $0,95 = F_3(7,81)$ où F_3 désigne la fonction de répartition de la loi du χ^2 à 3 degrés de liberté.
- $\dfrac{1}{2500}\left(102^2 + 34^2 + 78^2 + 58^2\right) = 8,4032$

En introduisant des variables convenables X_1, X_2, X_3, X_4 et la variable U_n associée, justifier le rejet de l'hypothèse de répartition uniforme des nombres fournis par le générateur.

Vous pourrez pour illustrer votre réponse recopier la courbe de la fonction f_3 vue à la question 3)c) de la partie 2.

∗ fin ∗