Loi khi-deux χ^2
Énoncé

On rappelle que la fonction Γ, définie par $x \longmapsto \int_0^{+\infty} t^{x-1}e^{-t}\,\mathrm{d}t$ sur $\mathbb{R}_+^*$, vérifie les relations

- $\forall x \in \mathbb{R}_+^* : \quad \Gamma(x+1) = x\Gamma(x)$
- $\forall n \in \mathbb{N}^* : \quad \Gamma(n) = (n-1)!$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Partie I: Loi Gamma et loi du χ^2

Dans toute cette partie, les variables aléatoires réelles, définies sur le même espace probabilisé (Ω, T, P).
On rappelle que si X est une variable aléatoire définie sur un espace probabilisé, alors sa fonction de répartition est la fonction F_X définie pour tout réel x par $F_X(x) = P(X \leqslant x)$.
Cette fonction de répartition caractérise la loi de la variable aléatoire réelle X.
On dit qu'une variable aléatoire X à valeurs réelles admet une densité f si sa fonction de répartition peut s'écrire sous la forme $F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$ où f est une fonction à valeurs réelles positives ayant un nombre fini de points de discontinuité et telle que $\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$
Une variable aléatoire Y suit une loi de Poisson de paramètre $\lambda > 0$, lorsqu'elle est déterminée pour tout entier naturel k par $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$.
Une variable aléatoire suit une loi normale de paramètres m et σ, lorsqu'elle admet une densité de probabilité définie pour tout réel x par $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}$

1. b et p sont deux paramètres réels strictement positifs. On considère l'application f définie sur $\mathbb{R}$ par :
$$\forall x \in \mathbb{R}, f(x) = \begin{cases} \dfrac{x^{p-1}e^{-\frac{x}{b}}}{b^p\Gamma(p)} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$$
Dans le cas où $b = 1$, donner les différentes allures du graphe de f suivant la position du paramètre p par rapport à 1 et à 2. On précisera les éventuelles demi-tangentes en 0.

2. Prouver que f est une densité de probabilité.
Si une variable aléatoire X admet cette densité, on dit que X suit la loi $\Gamma(b,p)$

3. Soit X une variable aléatoire suivant la loi $\Gamma(b,p)$.
Montrer que X admet une espérance et une variance et les calculer.

4. Dans le cas particulier où $p = 1$, reconnaître la loi de X et exprimer sa fonction de répartition.

On admet que si X et Y sont 2 variables aléatoires indépendantes de densités respectives f et g, alors la variable $X+Y$ possède une densité h telle que $h : x \longmapsto \int_{-\infty}^{+\infty} f(x-t)g(t)\,\mathrm{d}t$

5. Soient X_1 et X_2 , deux variables aléatoires réelles indépendantes, suivant les lois respectives $\Gamma(b, p_1)$ et $\Gamma(b, p_2)$. Soit $x, y \in \mathbb{R}_+^*$, X et Y deux variables indépendantes telles que $X \hookrightarrow \Gamma(x,1)$ et $Y \hookrightarrow \Gamma(y,1)$. Posons $S = X_1 + X_2$ et soit f_S sa fonction densité.
   (a) Soit $s \in \mathbb{R}_+^*$, montrer que $\int_0^s t^{x-1}(s-t)^{y-1}\,\mathrm{d}t = \beta(p_1,p_2)s^{p_1+p_2-1}$;
   (b) Montrer que:
$$\forall s \in \mathbb{R}, \quad f_S(s) = \begin{cases} \dfrac{\beta(p_1,p_2)}{b^{p1+p_2}\Gamma(p_1)\Gamma(p_2)}s^{p_1+p_2-1}e^{-\frac{s}{b}} & \text{si } s > 0 \\ 0 & \text{si } s \leqslant 0 \end{cases}$$

(c) En déduire que $\beta(p_1,p_2)=\dfrac{\Gamma(p_1)\Gamma(p_2)}{\Gamma(p_1+p_2)}$, puis déduire $X_1+X_2\hookrightarrow\Gamma\left(b,p_1+p_2\right)$

6. Soient $n\in\mathbb{N}^*$, $X_1,\cdots,X_n$ variables aléatoires réelles mutuellement indépendantes, suivant les lois respectives $\Gamma(b,p_i)$ pour tout $i\in[\![1,n]\!]$

   Déterminer la loi de la variable aléatoire $\displaystyle\sum_{i=1}^{n}X_i$

Soit s un entier strictement positif ; on appelle loi du χ^2 à s degrés de liberté ou loi $\chi^2(s)$ la loi $\Gamma\left(2,\frac{s}{2}\right)$

7. (a) Donner l'espérance et la variance d'une variable aléatoire suivant la loi $\chi^2(s)$.

   (b) Soit $n\in\mathbb{N}^*$. Si pour tout $i\in[\![1,n]\!]$, Les X_i sont des variables mutuellement indépendantes, X_i suivant la loi $\chi^2(s)$, donner la loi de $\displaystyle\sum_{i=1}^{n}X_i$

8. Pour $x>0$ et $n\in\mathbb{N}^*$, Justifier la formule : $e^x=\displaystyle\sum_{k=0}^{n-1}\frac{x^k}{k!}+\int_0^x e^{x-t}\frac{t^{n-1}}{(n-1)!}\,\mathrm{d}t$

9. Soient X_{2n} une variable aléatoire suivant la loi $\chi^2(2n)$ et λ un réel strictement positif.
   Démontrer que $P\left(X_{2n}>2\lambda\right)=P\left(Y_\lambda<n\right)$ où Y_λ désigne une variable aléatoire suivant la loi de Poisson de paramètre λ

10. (a) Soit X une variable aléatoire suivant une loi normale centrée réduite. Montrer que X^2 suit la loi du χ^2 à un degré de liberté.

    (b) $X_1,\cdots,X_n$ variables aléatoires réelles mutuellement indépendantes, suivant la loi normale centrée réduite. Quelle est la loi de $\displaystyle\sum_{i=1}^{n}X_i^2$

Partie II: De l'urne au χ^2

Une urne contient des boules de couleurs $C_1,\cdots,C_k$ où $k\in\mathbb{N}$ et $k\geqslant 2$. Les boules de couleur C_i sont en proportion p_i non nulle avec $\displaystyle\sum_{i=1}^{n}p_i=1$. Étant donné n entier naturel non nul, on tire successivement dans l'urne n boules, avec remise après chaque tirage.
Pour $i\in[\![1,k]\!]$, on appelle X_i la variable aléatoire désignant le nombre de boules de couleur C_i obtenues lors du tirage. X_i suit par conséquent une loi binomiale de paramètres n et p_i.

11. (a) Déterminer l'espérance et la variance de X_i

    (b) Pour $i\neq j$, déterminer la loi de X_i+X_j et en déduire que $\operatorname{cov}(X_i,X_j)=-np_ip_j$

12. Soit k entier supérieur ou égal à 2. On pose $Y_i=\dfrac{X_i-np_i}{\sqrt{np_i}}$ et on note M la matrice de covariance des variables aléatoires $Y_1,\cdots,Y_k$ définie par $M=\left(\operatorname{cov}(Y_i,Y_j)\right)_{1\leqslant i,j\leqslant k}$

    (a) Montrer que pour $i\neq j$, $\operatorname{cov}\left(Y_i,Y_j\right)=-\sqrt{p_ip_j}$ et que $\operatorname{cov}(Y_i,Y_i)=1-p_i$.
    En déduire l'expression de la matrice M.
    Dans la suite, $M_{n,p}(\mathbb{R})$ désigne l'ensemble des matrices à n lignes et p colonnes à coefficients réels. On notera I_k la matrice unite de $M_k(\mathbb{R})$

    (b) Préciser la matrice P telle que $M=I_k-P$, puis la matrice $C\in M_{k,1}(\mathbb{R})$ vérifiant $P=C{}^tC$, où ${}^tC$ désigne la transposée de la matrice C.

    (c) Déterminer le rang de P.

    (d) Calculer P^2. Préciser les valeurs propres de P et leur multiplicité.

13. Soit $J \in M_k(\mathbb{R})$ dont les coefficients $a_{i,j}$ vérifient: $a_{i,j} = \begin{cases} 1 & \text{si } i = j \leqslant k-1 \\ 0 & \text{sinon} \end{cases}$.

    Justifier l'existence d'une matrice $S \in M_k(\mathbb{R})$ vérifiant: ${}^tSS = I_k$ et ${}^tSMS = J$

14. On définit les variables aléatoires $Z_1,\, Z_2, \cdots, Z_k$ par la relation :

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_k \end{pmatrix} = {}^tS \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix}$$

    et on pose $S = (s_{ij})_{1 \leqslant i,j \leqslant k}$

    (a) Prouver que, pour tout entier i entre 1 et k, Z_i est centrée.

    (b) Déterminer la matrice de covariance de $Z_1, Z_2, \cdots, Z_k$. En déduire que Z_k est la variable certaine égale à zéro.

    (c) On pose $Q = \sum_{i=1}^{k} Y_i^2$. Calculer Q en fonction de $Z_1, Z_2, \cdots, Z_k$.

    (d) On suppose que n est grand, que les variables $Z_1, Z_2, \cdots, Z_{k-1}$ sont mutuellement indépendantes et qu'elles suivent des lois normales centrées réduites. Justifier le fait que Q suit la loi du χ^2 à $(k-1)$ degrés de liberté.