Bachelor académique en Sciences et Ingénierie
Probabilités et Statistique 2

Corrigé TD n° 7 2021

---

1. (*Loi du chi-carré*) Soient $X_1, \ldots, X_n$ indépendantes et de loi $N(0,1)$. Alors

$$U_n := X_1^2 + \ldots + X_n^2 \sim \chi_n^2$$

admet la densité suivante

$$f_{\chi_n^2}(t) = \frac{t^{n/2-1}}{2^{n/2}\Gamma(n/2)}\, e^{-t/2}\, \mathbb{1}_{\mathbb{R}_+}(t), \quad t \in \mathbb{R}.$$

(i) Dans le cas $n = 1$, on obtient la densité de X^2 avec $X \sim N(0,1)$ par les méthodes suivantes:

Méthode 1 (transformation de densité, changement de variable)

On a

$$\begin{aligned}
\mathbb{P}\left\{X^2 \leqslant t\right\} = \mathbb{P}\left\{|X| \leqslant \sqrt{t}\right\} &= \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{t}} e^{-x^2/2}\, \mathrm{d}x \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{t} e^{-u/2} \frac{1}{2\sqrt{u}}\, \mathrm{d}u \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{t} u^{\frac{1}{2}-1} e^{-u/2}\, \mathrm{d}u,
\end{aligned}$$

où on a utilisé le changement de variable $u = x^2$ et $\mathrm{d}u = 2x\,\mathrm{d}x = 2\sqrt{u}\,\mathrm{d}x$. Donc, la densité est donnée par

$$f_{X^2}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}\left\{X^2 \leqslant t\right\} = \frac{t^{1/2-1}}{2^{1/2}\Gamma(1/2)}\, e^{-t/2}\, \mathbb{1}_{\mathbb{R}_+}(t), \quad t \in \mathbb{R},$$

car $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. (D'abord, on utilise le changement de la variable $t = u^2$ dans la définition de la fonction gamma et ensuite on passe aux coordonnées polaires).

Méthode 2 (fonction de répartition)

On trouve

$$\begin{aligned}
F_{X^2}(t) = \mathbb{P}\left\{X^2 \leqslant t\right\} &= \mathbb{P}\left\{|X| \leqslant \sqrt{t}\right\} \\
&= \mathbb{P}\left\{-\sqrt{t} \leqslant X \leqslant \sqrt{t}\right\} \\
&= N(\sqrt{t}) - N(-\sqrt{t}) = 2N(\sqrt{t})
\end{aligned}$$

et donc pour la densité

$$f_{X^2}(t) = \frac{\mathrm{d}}{\mathrm{d}t}F_{X^2}(t) = \frac{1}{\sqrt{t}} f_X(\sqrt{t}) = \frac{1}{\sqrt{t}}\frac{1}{\sqrt{2\pi}} e^{-t/2}, \quad t > 0.$$

(ii) On procède ensuite par récurrence:
Supposons le résultat soit vrai pour un $n \in \mathbb{N}$. Montrons qu'il est vrai pour $n+1$: On a $U_{n+1} = U_n + X_{n+1}^2$ et évidemment U_n et X_{n+1}^2 sont indépendantes. Par conséquent,

$$f_{U_{n+1}} = f_{U_n} * f_{X_{n+1}^2}.$$

Par définition du produit de convolution, on obtient

$$\begin{aligned}
f_{U_{n+1}}(x) &= \int_0^x f_{U_n}(u) f_{U_1}(x-u)\ \mathrm{d}u \\
&= \int_0^x \frac{u^{n/2-1}e^{-u/2}}{2^{n/2}\Gamma(n/2)} \frac{(x-u)^{-1/2}e^{-(x-u)/2}}{2^{1/2}\Gamma(1/2)}\ \mathrm{d}u \\
&= \frac{e^{-x/2}}{2^{(n+1)/2}} \int_{u=0}^x \frac{u^{n/2-1}(x-u)^{-1/2}}{\Gamma(n/2)\Gamma(1/2)}\ \mathrm{d}u, \quad u = xt, \quad \mathrm{d}u = x\,\mathrm{d}t \\
&= \frac{e^{-x/2}}{2^{(n+1)/2}} \frac{x^{(n+1)/2-1}}{\Gamma(\frac{n+1}{2})} \underbrace{\int_0^1 \frac{\Gamma(\frac{n+1}{2})}{\Gamma(n/2)\Gamma(1/2)} t^{n/2-1}(1-t)^{-1/2}\ \mathrm{d}t}_{=\,1},
\end{aligned}$$

en utilisant que

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\ \mathrm{d}t \equiv \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$$

est la fonction bêta.

Remarque 1. Il y a une preuve plus directe. Par définition, la densité de probabilité commune de $(X_1, \ldots, X_n)$ est

$$\rho(x_1, ..., x_n) = \prod_{i=1}^n f_{X_i}(x_i) = \frac{1}{(2\pi)^{n/2}\sigma_1 \ldots \sigma_n} \exp\left(-\sum_{i=1}^n \frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right),$$

où $\mu_1 = ... = \mu_n = 0$ et $\sigma_1 = ... = \sigma_n = 1$. Pour une fonction intégrable $f(x_1, ..., x_n)$ de n variables $x_1, ..., x_n$, on a

$$\int_{|x|^2<r} f(x_1, ..., x_n)\,\mathrm{d}x = \int_0^{\sqrt{r}} \mathrm{d}s \int_{|x|=s} f(x)\,\mathrm{d}S(x).$$

S'il f est une fonction à symétrie radiale, $f(x) = f(|x|)$, alors

$$\int_{|x|=r} f(x)\mathrm{d}S(x) = \sigma_{n-1} f(r) r^{n-1},$$

avec $\mathrm{d}S(x)$ l'élément de surface de la sphère $\{|x| = r\}$ et $\sigma_{n-1} = \dfrac{2\pi^{n/2}}{\Gamma(n/2)}$ l'aire de la $(n-1)$-sphère. D'où on obtient la densité de U par dérivation:

$$f_U(t) = \frac{1}{2\sqrt{t}} \int_{|x|=\sqrt{t}} \rho(x)\,\mathrm{d}S(x),$$

et en plus

$$\begin{aligned}
\frac{1}{2\sqrt{t}} \int_{|x|=\sqrt{t}} \rho(x)\mathrm{d}S(x) &= \frac{1}{2\sqrt{t}} \frac{2\pi^{n/2}}{\Gamma(n/2)} \frac{e^{-t/2}}{(2\pi)^{n/2}} t^{(n-1)/2} \\
&= t^{-1/2} \frac{t^{(n-1)/2}}{2^{n/2}\Gamma(n/2)} e^{-t/2} = \frac{t^{n/2-1}}{2^{n/2}\Gamma(n/2)} e^{-t/2}.
\end{aligned}$$

Remarque 2 (cf. Définition 6.11). Soit G une variable aléatoire de densité f_G. On dit que G suit une loi gamma de paramètres ℓ et ϑ, si

$$f_G(t) = \frac{t^{\ell-1}}{(\ell-1)!\vartheta^\ell}\, e^{-t/\vartheta}\, \mathbb{1}_{\{t\geqslant 0\}}.$$

On observe que $\Gamma(n) = (n-1)!$ et si G est une variable de loi gamma avec des paramètres $\ell = \frac{n}{2}$ et $\vartheta = 2$, alors $G \sim \chi_n^2$. En particulier, la fonction génératrice des moments pour une variable G de loi gamma, resp. pour N^2 avec N de loi normale centrée réduite, est donnée par

$$M_G(t) = \frac{1}{(1-\vartheta t)^\ell}, \quad M_{N^2}(t) = \frac{1}{(1-\vartheta t)^{1/2}}.$$

On pose maintenant $\ell = n/2$, $\vartheta = 1/2$ et on utilise que la fonction génératrice d'une somme des variables aléatoires indépendantes est le produit des fonctions génératrices (cf. Compléments, Proposition 3.4), ce qui montre que la densité est bien donnée par $f_{\chi_n^2}$.

2. (*Vecteur gaussien et densité*) Par définition, $\langle a, X\rangle$ est gaussien, et donc on a pour tout $a \in \mathbb{R}^n$:

$$\mathbb{E}e^{it\langle a,X\rangle} = e^{it\mu(a) - \frac{1}{2}t^2\sigma^2(a)}, \quad t \in \mathbb{R},$$

où $\mu(a)$ la moyenne et $\sigma^2(a)$ la variance de $\langle a, X\rangle$. Ensuite

$$\begin{aligned}\mu(a) &= \mathbb{E}\langle a, X\rangle = \langle a, \mathbb{E}X\rangle = \langle a, \mu\rangle,\\ \sigma^2(a) = \mathrm{var}\langle a, X\rangle = \mathrm{var}\left(\sum_{k=1}^n a_k X_k\right) &= \sum_{i,j=1}^n \mathrm{cov}(a_i X_i, a_j X_j)\\ &= \sum_{i,j=1}^n a_i a_j \,\mathrm{cov}(X_i, X_j) = \langle a, Ca\rangle,\end{aligned}$$

où $\mu := \mathbb{E}[X] \in \mathbb{R}$ et $C := (\mathrm{cov}(X_i, X_j))_{i,j=1,\dots,n}$. Apparemment, la matrice C est symétrique et strictement définie positive, car $\langle a, Ca\rangle = \sigma^2(a) > 0$ pour tout $a \in \mathbb{R}^n$.

À cause de l'unicité de la fonction caractéristique (Remarque 8.5), il suffit de montrer que

$$\int f_X(x) e^{i\langle a,x\rangle}\mathrm{d}x = e^{i\langle \mu,a\rangle - \frac{1}{2}\langle a,Ca\rangle}$$

avec

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C}} \exp\left(-\frac{\langle x-\mu, C^{-1}(x-\mu)\rangle}{2}\right), \quad x \in \mathbb{R}^n.$$

Car C est strictement définie positive, il existe une racine $\sqrt{C}$ uniquement déterminée, symétrique et positive, c.à.d. une $n\times n$-matrice telle que $\sqrt{C}\sqrt{C} = C$ et $|\sqrt{C}a|^2 = \langle a, Ca\rangle$ pour tout $a \in \mathbb{R}^n$. De plus C^{-1} est aussi strictement définie positive et on a $\sqrt{C^{-1}} = \sqrt{C}^{-1}$. En utilisant le changement de la variable $\sqrt{C}z = x - \mu$, $\mathrm{d}z = (\det\sqrt{C})^{-1}\mathrm{d}x$, on obtient

$$\begin{aligned}\int f_X(x) e^{i\langle a,x\rangle}\mathrm{d}x &= \int e^{i\langle a,x\rangle}\frac{1}{\sqrt{\det C}}\frac{1}{(2\pi)^{n/2}} e^{-\langle x-\mu, C^{-1}(x-\mu)\rangle/2}\mathrm{d}x\\ &= e^{i\langle a,\mu\rangle}\int e^{i\langle a,\sqrt{C}z\rangle}\frac{1}{(2\pi)^{n/2}}e^{-\langle\sqrt{C}z,\sqrt{C}^{-1}z\rangle/2}\mathrm{d}z\\ &= e^{i\langle a,\mu\rangle}\int e^{i\langle z,\sqrt{C}a\rangle}\frac{1}{(2\pi)^{n/2}}e^{-|z|^2/2}\mathrm{d}z\\ &= e^{i\langle a,\mu\rangle}e^{-|\sqrt{C}a|^2/2} = e^{i\langle a,\mu\rangle}e^{-\langle a,Ca\rangle/2}.\end{aligned}$$

Si la matrice C est symétrique et positive, alors il existe une matrice

$$A \in \mathrm{Matr}(n \times n; \mathbb{R}) \quad \text{telle que} \quad C = AA^t.$$

Supposons maintenant C non inversible, alors A n'est pas inversible et

$$\det A = \sqrt{\det C} = 0.$$

Par conséquent,

$$H := \mathrm{Im}\ A + \mu$$

est un sous-espace affin de $\mathbb{R}^n$ de dimension $\dim H = \mathrm{rang}\ A < n$. En particulier,

$$\text{la mesure de Lebesgue de } H = \lambda^n(H) = 0.$$

Avec l'hypothèse que la loi de X soit absolument continue de densité f_X, on obtient une contradiction:

$$\begin{aligned} 1 &= \mathbb{P}\{AZ + \mu \in H\} = \mathbb{P}\{X \in H\} \\ &= \int_{\mathbb{R}^n} \mathbb{1}_H(x) f_X(x) \mathrm{d}x \\ &= \int_H f_X(x) \lambda_n(\mathrm{d}x) = 0. \end{aligned}$$

Donc, la loi de X n'est pas absolument continue.

3. (*Les billes métalliques*) Soit $(X_1, ..., X_n)$ un n-échantillon de $N(0, \sigma^2)$. L'intervalle de confiance est donnée par

- σ^2 *connu*:

$$\left[\bar{X} - z_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}} \ , \ \bar{X} + z_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}\right],$$

  où $z_{(1-\frac{\alpha}{2})}$ est le $(1-\alpha/2)$-quantile de la loi normale.

- σ^2 *inconnu*:

$$\left[\bar{X} - t_{(1-\frac{\alpha}{2};n-1)} \frac{S}{\sqrt{n}} \ , \ \bar{X} + t_{(1-\frac{\alpha}{2};n-1)} \frac{S}{\sqrt{n}}\right],$$

  où $t_{(1-\frac{\alpha}{2};n-1)}$ le $(1-\alpha/2)$-quantile de la loi de Student à $n-1$ degrés de liberté.

(a) On calcule la moyenne $\bar{X}$ de l'échantillon:

$$\bar{X} = \frac{19{,}6 + 20 + 20{,}2 + 20{,}1 + 20 + 19{,}9 + 20 + 2{,}03 + 20{,}1 + 19{,}8}{10} = 20.$$

Calculons la variance puis l'écart-type de l'échantillon à partir de la moyenne de l'échantillon:

$$S^2 = \frac{10}{9} \left(\frac{19{,}6^2 + 20^2 + \ldots + 19{,}8^2}{10} - 20^2 \right) = 0{,}04,$$

puis

$$S = \sqrt{0{,}04} = 0{,}2.$$

Dans la table de la loi de Student, pour 9 ($= 10 - 1$) degrés de liberté, on trouve

$$\mathbb{P}\{|T| > 2{,}26\} = 0{,}05 \quad \text{où} \quad \mathbb{P}\{|T| < 2{,}26\} = 0{,}95.$$

L'intervalle de confiance pour le poids moyen est donc:

$$\left[20 - 2{,}26 \times \frac{0,2}{\sqrt{10}},\ 20 + 2{,}26 \times \frac{0,2}{\sqrt{10}}\right] \approx [19{,}86,\ 20{,}14]\,.$$

(b) Si l'écart-type de la population est connu, on utilise la loi normale:

$$\mathbb{P}\{|U| > 1{,}96\} = 0{,}05 \quad \text{où} \quad \mathbb{P}\{|U| < 1{,}96\} = 0{,}95.$$

L'intervalle de confiance pour le poids moyen est alors:

$$\left[20 - 1{,}96 \times \frac{0,2}{\sqrt{10}},\ 20 + 1{,}96 \times \frac{0,2}{\sqrt{10}}\right] \approx [19{,}88,\ 20{,}12]\,.$$