CaMUS 1 (2010), 1–6

Caractérisation de la loi normale

Djilali Ait Aoudia

Résumé. Dans ce travail, nous nous intéressons à deux des caractérisations les plus célèbres de la loi normale : soient le théorème de Bernstein et celui de Geary concernant l'indépendance de la moyenne et de la variance expérimentales. Nous présenterons succinctement une approche différente qui dégage la structure des moments qui découlent des hypothèses de ces théorèmes.

1. Introduction

La loi normale est l'une des principales lois de distributions de probabilité formulée par le mathématicien français Abraham de Moivre en 1733 et elle a été mise en évidence par Gauss au XIXe siècle. L'expérience montre qu'un grand nombre de caractères physiques, biométriques, peuvent être modélisés avec succès par une loi normale. Une des explications de ce phénomène est fournie par le théorème centrale limite. En effet, de nombreux phénomènes sont dûs à l'addition d'un grand nombre de petites perturbations aléatoires. Dans cette note, nous revisitons deux des caractérisations les plus célèbres de la loi normale, notamment celle de Bernstein [**1**] concernant l'indépendance de $X-Y$ et $X+Y$ lorsque X et Y sont des variables aléatoires indépendantes, ainsi que celle de Geary [**3**] concernant l'indépendance de la moyenne et de la variance expérimentales. Ce papier a pour objectif de présenter des approches un peu différentes pour établir ces résultats en s'appuyant sur le fait qu'une loi normale est caractérisée par ses moments, ce qui n'est pas vrai en général (voir [**2**], exemple 2.3.5 page 64).

2. Le théorème de Bernstein

Le théorème ci-dessous dû à Bernstein [**1**] est l'un des résultats fondamentaux qui caractérise la loi normale par l'indépendance linéaire, important en particulier

Je tiens à remercier très chaleureusement les Professeurs Éric Marchand et François Perron pour leurs conseils et leurs commentaires.

par son extension et son emploi à des résultats plus généraux.

Théorème 1 (S. Bernstein, 1941)

Soient X et Y deux variables aléatoires indépendantes, alors

$$(2.1)\quad (U = X - Y \text{ indépendant de } V = X + Y) \Longleftrightarrow \begin{cases} X \sim \mathcal{N}\left(\mathbb{E}(X), \sigma^2\right) \\ \quad\quad \text{et} \\ Y \sim \mathcal{N}\left(\mathbb{E}(Y), \sigma^2\right) \end{cases}$$

Démonstration.

La réciproque est un exercice élémentaire. En effet, puisque le vecteur (dans $\mathbb{R}^2$) (U, V) est gaussien (car toute combinaison linéaire de ses composantes est de loi gaussienne), pour montrer l'indépendance de U et V, il suffit de démontrer que $\mathbb{C}ov(U, V) = 0$. Pour cela, on vérifie que

$$\mathbb{C}ov(U, V) = \mathbb{C}ov(X - Y, X + Y) = \mathbb{V}ar(X) - \mathbb{V}ar(Y) = \sigma^2 - \sigma^2 = 0.$$

Nous proposons deux méthodes pour l'implication : la première s'appuyant sur le fait qu'une loi normale est caractérisée par ses moments et la deuxième en exploitant la structure induite sur les fonctions caractéristiques de X et de Y. Notons d'abord que les conditions du théorème implique que tous les moments de X et Y existent (voir par exemple [4], Lemme 5.3.2, page 65) et sans perte de généralité, nous supposons que $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ et $\mathbb{E}[X^2] = \sigma^2$.

Première méthode

Commençons par montrer que $\mathbb{E}\left[X^n\right] = \left[X^n\right]$ pour tout $n \geq 0$. Nous allons procéder par récurrence sur n. C'est vrai pour $n = 0, 1$. Supposons le vrai pour $n = 0, 1, \cdots, m-1$ et montrons le pour $n = m$. Un calcul simple nous donne

$$\begin{aligned}
&\mathbb{E}\left[X^m - Y^m\right] \\
&= \mathbb{E}\left[(X - Y)(X^{m-1} + X^{m-2}Y + \cdots + Y^{m-1})\right] \\
&= \mathbb{E}\left[(X - Y)\{(X + Y)^{m-1} - (X + Y)^{m-1} + \sum_{i=0}^{m-1} X^{m-1-i}Y^i\}\right] \\
&= \mathbb{E}\left[(X - Y)\{(X + Y)^{m-1} - \sum_{i=0}^{m-1}[\binom{m-1}{i} - 1]X^{m-1-i}Y^i\}\right] \\
&= \mathbb{E}\left[(X - Y)(X + Y)^{m-1}\right] - \mathbb{E}\left[(X - Y)\sum_{i=0}^{m-1} A_{i,m}X^{m-1-i}Y^i\right]
\end{aligned}$$

avec

$$A_{i,m} = [\binom{m-1}{i} - 1] = A_{m-i-1,m}.$$

Mais d'une part, par l'indépendance de $X - Y$ et $X + Y$, on a

$$\mathbb{E}\left[(X - Y)(X + Y)^{m-1}\right] = \left[\mathbb{E}(X) - \mathbb{E}(Y)\right]\mathbb{E}\left[(X + Y)^{m-1}\right] = 0.$$

D'autre part, par l'hypothèse de récurrence, on a bien

$$\begin{aligned}
&\mathbb{E}[(X-Y)(\sum_{i=0}^{m-1} A_{i,n}X^{m-1-i}Y^i)] \\
&\qquad = \mathbb{E}\left[\sum_{i=1}^{m-1} A_{i,m}X^{m-i}Y^i\right] - \mathbb{E}\left[\sum_{i=0}^{m-1} A_{i,m}X^{m-i-1}Y^{i+1}\right] \\
&\qquad = \sum_{i=0}^{m-1} \mathbb{E}\left[A_{i,m}X^{m-i}Y^i\right] - \sum_{i=0}^{m-1} \mathbb{E}\left[A_{m-i-1,m}X^iY^{m-i}\right] \\
&\qquad = \sum_{i=0}^{m-1} \mathbb{E}\left[A_{i,n}X^{m-i}Y^i\right] - \sum_{i=0}^{m-1} \mathbb{E}\left[A_{i,m}X^iY^{m-i}\right] \\
&\qquad = \sum_{i=0}^{m-1} A_{i,m}\left(\mathbb{E}\left[X^{m-i}\right]\mathbb{E}\left[Y^i\right] - \mathbb{E}\left[X^i\right]\mathbb{E}\left[Y^{m-i}\right]\right) \quad \{\text{indépendance}\} \\
&\qquad = 0 \qquad \{\text{ hypothèse de récurrence}\}.
\end{aligned}$$

Ainsi, on conclut que

$$\mathbb{E}\left[X^m - Y^m\right] = 0 \quad \text{et} \quad \mathbb{E}\left[X^n\right] = \mathbb{E}\left[Y^n\right] = 0 \quad \text{pour tout} n \geq 0.$$

Maintenant, soient Z_1 et Z_2 deux variables aléatoires indépendantes de loi normale centrée et $\mathbb{E}\left[Z_i^2\right] = \sigma^2$. Donc pour montrer le théorème, il suffit de montrer que $\mathbb{E}\left[X^n\right] = \mathbb{E}\left[Z_1^n\right]$ pour tout $n \geq 0$. C'est vrai pour $n = 0, 1$. Supposons le vrai pour $n = 0, 1, \cdots, m$ et montrons le pour $n = m+1, m \geqslant 1$. On a

$$\begin{aligned}
\mathbb{E}\left[X^{m+1} + Y^{m+1}\right] &= \mathbb{E}\Big[(X+Y)^{m-1}\left\{(X-Y)^2 + 4XY\right\} \\
&\qquad - \left\{(X+Y)^{m+1} - (X^{m+1} + Y^{m+1})\right\}\Big] \\
&= \mathbb{E}\left[(X+Y)^{m-1}(X-Y)^2 - \sum_{k=1}^{m} a_{mk}X^kY^{m+1-k}\right] \\
&= \mathbb{E}\left[(X+Y)^{m-1}\right]\mathbb{E}\left[(X-Y)^2\right] \\
&\qquad - \sum_{k=1}^{m} a_{mk}\mathbb{E}\left[X^k\right]\mathbb{E}\left[Y^{m+1-k}\right] \\
&= \mathbb{E}\left[(Z_1+Z_1)^{m-1}\right]\mathbb{E}\left[(Z_1-Z_2)^2\right] \\
&\qquad - \sum_{k=1}^{m} a_{mk}\mathbb{E}\left[Z_1^k\right]\mathbb{E}\left[Z_2^{m+1-k}\right] \\
&= \mathbb{E}\left[(Z_1+Z_1)^{m-1}(Z_1-Z_2)^2 - \sum_{k=1}^{m} a_{mk}Z_1^kZ_2^{m+1-k}\right] \\
&= \mathbb{E}\left[Z_1^{m+1} + Z_2^{m+1}\right],
\end{aligned}$$

avec $a_{mk} = \binom{m+1}{k} - 4\binom{m-1}{k-1}$ pour $1 \leq k \leq m$.
Et le théorème est démontré, en s'appuyant sur le fait que la loi normale est

caractérisée par ses moments (voir par exemple [4], corollary 2.3.3).

Deuxième méthode
Soient

$$\phi(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[e^{itY}],\ V = X + Y \text{ et } W = (X - Y)^2.$$

Il vient

$$\begin{aligned}(a)\ \mathbb{E}[We^{itV}] &= \mathbb{E}[(X^2 - 2XY + Y^2)e^{it(X+Y)}] \\ &= \underbrace{\mathbb{E}[X^2e^{itX}]}_{-\phi''(t)}\underbrace{\mathbb{E}[e^{itY}]}_{\phi(t)} - 2\underbrace{\mathbb{E}[Xe^{itX}]}_{-i\phi'(t)}\underbrace{\mathbb{E}[Ye^{itY}]}_{-i\phi'(t)} \\ &\qquad + \underbrace{\mathbb{E}[Y^2e^{itY}]}_{-\phi''(t)}\underbrace{\mathbb{E}[e^{itX}]}_{\phi(t)} \\ &= -2\phi^{''}(t)\phi(t) + 2(\phi^{'}(t))^2,\end{aligned}$$

$$\begin{aligned}(b)\ \mathbb{E}[We^{itV}] &= \mathbb{E}[W]\mathbb{E}[e^{itV}] \qquad \text{(indépendance)} \\ &= \mathbb{E}[(X - Y)^2]\mathbb{E}[e^{it(X+Y)}] \\ &= \mathbb{E}[(X^2 - 2XY + Y^2)]\mathbb{E}[e^{it(X+Y)}] \\ &= [\mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y^2)]\mathbb{E}[e^{itX}]\mathbb{E}[e^{itY}] \\ &= 2\sigma^2(\phi(t))^2,\end{aligned}$$

$$(a) + (b) \Longrightarrow \begin{cases} -\phi^{''}(t)\phi(t) + (\phi^{'}(t))^2 = \sigma^2(\phi(t))^2 \\ \qquad \text{et} \\ \phi(0) = 1, \quad \phi^{'}(0) = 0 \end{cases} \tag{2.2}$$

$$\Longrightarrow \begin{cases} \left(\frac{\phi^{'}}{\phi}\right)^{'} = -\sigma^2 \\ \qquad \text{et} \\ \phi(0) = 1, \quad \phi^{'}(0) = 0 \end{cases} \Longrightarrow \quad \phi(t) = e^{\frac{-\sigma^2t^2}{2}} \quad \square \tag{2.3}$$

3. Le théorème de Geary

La caratérisation de la loi normale par l'indépendance de la moyenne et de la variance est donnée par le théorème suivant dû à R.C. Geary [**3**].
Théorème 2 (R.C. Geary, 1936)
Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes de même loi et de variances finies. On pose

$$\overline{X} = \frac{\sum_{i=1}^n X_i}{n} \quad \text{et} \quad S^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}.$$

Alors

$$\{\overline{X} \quad \text{indépendant de} \quad S^2\} \iff \{X_i \sim \mathcal{N}(\mu, \sigma^2), i = 1, 2, \cdots, n\}.$$

Démonstration.
Notons que notre démonstration se compare à celle de Quine [**7**], mais elle est un peu plus directe. L'approche que nous présentons est analogue à la deuxième démonstration ci-dessus du résultat de Bernstein. En effet, sans perte de généralité, nous supposons d'abord (voir Remarque à la suite de la démonstration) que

$$\mathbb{E}[X_i] = 0 \quad \text{et} \quad \mathbb{E}[X_i^2] = \sigma^2, \quad i = 1, 2, \cdots, n.$$

On a d'abord $\mathbb{E}[X_i X_j] = 0, \forall i \neq j$ et

$$\begin{aligned} nS^2 &= n\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n} = \sum_{i=1}^n X_i^2 - n\overline{X}^2 \\ &= \sum_{i=1}^n X_i^2 - \frac{1}{n}(\sum_{i=1}^n X_i)^2 = (1 - \frac{1}{n})\sum_{i=1}^n X_i^2 - \frac{2}{n}\sum_{i<j} X_i X_j \end{aligned}$$

et ainsi le résultat bien connu

$$\mathbb{E}[nS^2] = (n-1)\sigma^2.$$

On a aussi

$$\begin{aligned} (c) \quad \mathbb{E}[nS^2 e^{itn\overline{X}}] &= \mathbb{E}[nS^2]\mathbb{E}[e^{itn\overline{X}}] \\ &= \mathbb{E}[nS^2]\mathbb{E}[e^{it\sum_{j=1}^n X_j}] \\ &= (n-1)\sigma^2 \underbrace{\mathbb{E}[e^{itX_1}]}_{\phi(t)} \underbrace{\mathbb{E}[e^{itX_2}]}_{\phi(t)} \cdots \underbrace{\mathbb{E}[e^{itX_n}]}_{\phi(t)} \\ &= (n-1)\sigma^2 \phi^n(t), \end{aligned}$$

$$\begin{aligned} (d) \quad \mathbb{E}[nS^2 e^{itn\overline{X}}] =& \mathbb{E}\left[\left((1-\frac{1}{n})\sum_{j=1}^n X_j^2 - \frac{2}{n}\sum_{j<k} X_j X_k\right) e^{it\sum_{j=1}^n X_j}\right] \\ &= (1-\frac{1}{n})\sum_j \underbrace{\mathbb{E}\left[X_j^2 e^{itX_j} e^{it\sum_{k\neq j} X_k}\right]}_{-\phi''(t)\phi^{n-1}(t)} \\ &\quad - \frac{2}{n}\sum_{j<k} \underbrace{\mathbb{E}\left[X_j e^{itX_j} X_k e^{itX_k} e^{it\sum_{h\neq j,k} X_h}\right]}_{-\phi'^2(t)\phi^{n-2}(t)} \\ &= -(1-\frac{1}{n})n\phi''(t)\phi^{n-1}(t) + \frac{2}{n}\frac{n(n-1)}{2}\phi'^2(t)\phi^{n-2}(t), \end{aligned}$$

et donc

$$(c) + (d) \implies -\phi''\phi + (\phi')^2 = \sigma^2\phi^2 \implies \phi(t) = e^{\frac{-\sigma^2 t^2}{2}}.$$

La réciproque est mieux connue et élémentaire. Soient $X_1, \cdots, \cdots, X_n$ des variables aléatoires indépendantes de même loi normale avec $\mathbb{V}(X_1) = \sigma^2$. Pour montrer que $\overline{X}$ et S^2 sont indépendantes, il suffit de montrer que $\overline{X}$ est indépendante du vecteur $M = (X_1 - \overline{X}, \ldots, X_n - \overline{X})$, puisque S^2 est M-mesurable. Or le vecteur $(\overline{X}, X_1 - \overline{X}, \cdots, X_n - \overline{X})$ est gaussien (car toute combinaison linéaire de ses composantes est une combinaison linéaire des X_i), et ainsi pour démontrer l'indépendance de $\overline{X}$ et M, il suffit de vérifier que $\mathbb{C}ov(\overline{X}, X_i - \overline{X}) = 0$ pour tout i. Pour cela, on calcule

$$\begin{aligned} \mathbb{C}ov(\overline{X}, X_i) &= \mathbb{C}ov(\frac{1}{n}\sum_{i=1}^{n} X_i, X_i) \\ &= \frac{1}{n}\mathbb{C}ov(X_i, X_i) \\ &= \frac{\sigma^2}{n}, \end{aligned}$$

et donc

$$\begin{aligned} \mathbb{C}ov(\overline{X}, X_i - \overline{X}) &= \mathbb{C}ov(\overline{X}, X_i) - \mathbb{V}ar(\overline{X}) \\ &= \frac{\sigma^2}{n} - \frac{\sigma^2}{n} \\ &= 0, \end{aligned}$$

d'où le résultat. □

Remarques

1) Si $\mathbb{E}[X_i] = m \neq 0$, on peut poser $L_i = X_i - m$ et $\overline{L} = \overline{X} - m$. Alors, on a que

$$S^2 = \frac{\sum_{i=1}^{n}(L_i - \overline{L})^2}{n},$$

que l'indépendance de S^2 et $\overline{X}$ équivaut à l'indépendance de S^2 et $\overline{L}$, et on peut donc prendre $\mu = 0$ sans perte de généralité.

2) Le théorème de Geary est une généralisation de celui de Bernstein car, pour $n = 2$, $\overline{X} = X_1 + X_2$ et $S^2 = (X_1 - X_2)^2/2$.

Références

[1] S. Bernstein. *On a characteristic property of the normal law.* Trud. Leningrad Poly. Inst. 3. (1941)

[2] Casella, G and Berger, R. L. *Statistical inference.* Wadsworth & Brooks/Cole. (1990)

[3] R.C. Geary. *Distribution of Student's ratio for non-normal samples.* J.Roy. Statist. Soc. Suppl 3. (1936)

[4] W. Bryc. *The normal distribution. Characterizations with applications.* Lecture notes in Statistics, 100. Springer, New York. (2005)

[5] E. Lukacs. *A characterisation of the normal distribution.* Ann. math. Statist. 13 (1942)

[6] N.L. Johnson. S. Kotz and N. Balakrishnan. *Continuous univariate distributions, volume 1.* Wiley, New York, second edition. (1994)

[7] M.P Quine. *On three characterizations of the normal distribution.* Prob. and Math. Stat. Vol 14. (1993)

DJILALI AIT AOUDIA, DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUES, UNIVERSITÉ DE MONTRÉAL