La loi log-normale

Connaitre une variable alléatoire, c'est avant tout connaître sa loi. Dans le cas d'une variable aléatoire réelle à densité X, cette loi est entièrement déterminée aussi bien par sa fonction de répartition $F_X : x \mapsto P(X \leqslant x)$ que par sa fonction de densité $f_X := F'_X$. Chacune de ses représentation de la loi de X se montre plus ou moins adaptée à la résolution de certains problèmes. On est donc souvent amené à voyager d'une représentation à l'autre. L'étude de la loi log-normale sera l'occasion d'illustrer différentes dynamiques qui se jouent entre fonctions de répartitions et fonctions de densités dans l'étude des lois de variables aléatoires. Elle sera aussi l'occasion de montrer à quel point il est important de connaître intimement les lois de variables alléatoires usuelles.

On dit qu'une variable aléatoire X suit une loi log-normale de paramètre (μ, σ^2), ce qu'on note $X \sim \text{Log}-\mathcal{N}(\mu,\, \sigma^2)$, si son logarithme suit une loi normale de paramètre (μ, σ^2). Autrement dit, si elle est de la forme $X = e^T$ où $T \sim \mathcal{N}(\mu, \sigma^2)$; son logarithme est alors bien normal, autrement dit $\ln X \sim \mathcal{N}(\mu, \sigma^2)$. Nous allons montrer comment étudier une telle loi en nous ramenant à des lois normales.
Pour cela, notons Φ la fonction de répartition de la loi normale centrée réduite et φ sa fonction de densité.
Comme on connait la relation entre X et T, il est aisé d'établir une relation entre leurs fonctions de répartitions, donc de déterminer la fonction de répartition de X :

$$
\begin{aligned}
F_X(x) : x &\mapsto P(X \leqslant x) \\
&\mapsto P\left(e^T \leqslant x\right) \\
&\mapsto \begin{cases} P(T \leqslant \ln x) & si\ x > 0 \\ 0 & si\ x \leqslant 0 \end{cases} \\
&\mapsto P(T \leqslant \ln x) \cdot \mathbb{1}_{]0,+\infty[}(x) \\
&\mapsto F_T(\ln x) \cdot \mathbb{1}_{]0,+\infty[}(x) \\
&\mapsto \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \cdot \mathbb{1}_{]0,+\infty[}(x)
\end{aligned}
$$

La fonction indicatrice d'un ensemble E, notée $\mathbb{1}_E$ est définie par :

$$
\forall x \in \mathbb{R}, \mathbb{1}_E = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \notin E \end{cases}
$$

Elle permet une écriture plus compacte, et donc plus commode, des fonctions ayant des formules algébriques différentes selon les parties de $\mathbb{R}$ où elles sont définies.

Un calcul direct nous permet alors de trouver une expression de la densité de X:

$$
\boxed{f_X : x \mapsto F'_X(x) = \frac{1}{\sigma x} \cdot \varphi\left(\frac{\ln x - \mu}{\sigma}\right) \cdot \mathbb{1}_{]0,+\infty[}(x)}
$$

Si la relation entre F_T et Φ ne vous apparaît pas clairement, voyons comment la retrouver en trouvant une relation entre f_T et φ. Le calcul ci-contre vise à *reconnaitre* une loi normale centrée réduite dans l'*écriture* d'une loi normale quelconque. Mais nous ne savons *reconnaitre* une telle loi que sur sa densité, c'est pourquoi nous écrivons la fonction de répartition comme une intégrale de la fonction de densité. On utilisera pour ce faire le changement de variable :

$$
\tau = \frac{t - \mu}{\sigma} \quad (\text{d'où } dt = \sigma d\tau)
$$

$$
\begin{aligned}
F_T : u &\mapsto \int_{-\infty}^{u} f_T(t)\, dt \\
&\mapsto \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt \\
&\mapsto \int_{-\infty}^{\frac{u-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}}\, d\tau \\
&\mapsto \int_{-\infty}^{\frac{u-\mu}{\sigma}} \varphi(\tau) d\tau \\
&\mapsto \Phi\left(\frac{u-\mu}{\sigma}\right)
\end{aligned}
$$

De même, s'il ne vous a pas échappé qu'on a pas pris garde à traiter correctement le cas de la dérivée de F_X en 0, voici un moyen direct de démontrer qu'on a bien $F'_X(0) = 0$. Comme $\lim_{x \to 0^+} \frac{\ln x - \mu}{\sigma} = -\infty$, on peut affirmer l'existence d'un α pour lequel $\forall x \in]0; \alpha[, \forall \tau < \frac{\ln x - \mu}{\sigma}, -\tau > \frac{1}{\sqrt{2\pi}}$.
Dés lors, le calcul suivant achève la démonstration :

$$
\forall x \in\,]0; \alpha[\,, \frac{F_X(x) - F_X(0)}{x} = \frac{1}{x} \cdot \Phi\left(\frac{\ln x - \mu}{\sigma}\right) = \frac{1}{x} \cdot \int_{-\infty}^{\frac{\ln x - \mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}}\, d\tau
$$

$$
< e^{-\ln x} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} = e^{-\frac{(\ln x)^2 + 2(\sigma^2 - \mu)\ln x + \mu^2}{2\sigma^2}} = \frac{1}{x} \cdot \int_{-\infty}^{\frac{\ln x - \mu}{\sigma}} -\tau . e^{-\frac{\tau^2}{2}}\, d\tau = \frac{1}{x} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \xrightarrow[x \to 0^+]{} 0
$$

Moments d'une loi log-normale

L'inégalité de Markov prétend que pour toute variable alléatoire positive et pour tout $a > 0$, on a $P(A \geqslant a) \leqslant \frac{1}{a} \cdot E(A)$. En prenant A de la forme $|X - E(X)|^n$ avec $n \in \mathbb{N}$ et en posant $s := \sqrt[n]{a}$, on obtient l'inégalité $P(|X - E(X)| \geqslant s) \leqslant \frac{1}{s^n} \cdot E\left(|X - E(X)|^n\right)$ valable pour tout n. En particulier, dans le cas où $n = 2$, en posant $k := \frac{s}{\sigma}$, on obtient l'inégalité de Bienaymé-Tchebychev $P(|X - E(X)| \geqslant k\sigma) \leqslant \frac{1}{(k\sigma)^2} \cdot V(X) = \frac{1}{k^2}$. De telles inégalités permettent de majorer de façon probable l'écart d'une variable alléatoire à sa valeur moyenne. Elle sont à ce titre fondamentale pour réaliser des estimations, par exemple en statistiques. C'est la raison pour laquelle, outre la valeur de $E(X)$, on est intéréssé plus générallement par la valeur de $E(X^n)$ pour tout $n \in \mathbb{N}^*$, cette quantité est appelée le moment d'ordre n de X.

En remarquant que :

$$f_X \equiv 0 \text{ sur }]-\infty;0]$$

et en effectuant le changement de variable :

$$t = \frac{\ln x - \mu}{\sigma} \quad (\text{d'où } dx = \sigma e^{\sigma t+\mu} dt)$$

nous sommes maintenant à même de calculer l'espérance de X. La clef de notre calcul tient dans la remarque suivante : la fonction définie sur $\mathbb{R}$ par :

$$t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\sigma)^2}{2}}$$

n'est autre que la fonction de densité d'une loi normale $\mathcal{N}(\sigma, 1)$, d'où :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\sigma)^2}{2}} dt = 1$$

$$\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x \cdot f_X(x)\, dx \\
&= \int_{0}^{+\infty} \frac{1}{\sigma} \cdot \varphi\left(\frac{\ln x - \mu}{\sigma}\right) dx \\
&= \int_{-\infty}^{+\infty} \varphi(t)\, e^{\sigma t+\mu} dt \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} e^{\sigma t+\mu} dt \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2-2\sigma t-2\mu}{2}} dt \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\sigma)^2-\sigma^2-2\mu}{2}} dt \\
&= e^{\frac{\sigma^2+2\mu}{2}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\sigma)^2}{2}} dt \\
&= e^{\mu+\frac{\sigma^2}{2}}
\end{aligned}$$

Un moyen direct de déterminer les moments de X serait de faire appel au lemme de transfert, en calculant $\int_{-\infty}^{+\infty} x^n \cdot f_X(x)\, dx$ pour chaque $n \in \mathbb{N}^*$. Nous allons nous montrer plus futé ! Nous allons montrer que la variable aléatoire Y définie par $Y = X^n$ suit une loi log-normale dont on va chercher les paramètres. On pourra alors appliquer la formule de l'espérance que nous venons de démontrer en utilisant les paramètres ainsi trouvés.
Pour démontrer que Y suit une loi log-normale, il suffit de la mettre sous la forme $e^{T'}$ où T' suit une loi normale.

$$Y = X^n = \left(e^T\right)^n = e^{nT}$$

Il s'agit donc de montrer que $T' := nT$ suit une loi normale et d'en trouver les paramètres.

$$\begin{aligned}
F_{T'} : t &\mapsto P(T' \leqslant t) \\
&\mapsto P(nT \leqslant t) \\
&\mapsto P\left(T \leqslant \frac{t}{n}\right) \\
&\mapsto F_T\left(\frac{t}{n}\right)
\end{aligned}$$

D'où

$$\begin{aligned}
f_{T'} : t &\mapsto F'_{T'}(t) \\
&\mapsto \frac{1}{n} f_T\left(\frac{t}{n}\right) \\
&\mapsto \frac{1}{\sqrt{2\pi} n\sigma} e^{-\frac{\left(\frac{t}{n}-\mu\right)^2}{2\sigma^2}} \\
&\mapsto \frac{1}{\sqrt{2\pi} n\sigma} e^{-\frac{(t-n\mu)^2}{2(n\sigma)^2}}
\end{aligned}$$

Nous avons ainsi $T' \sim \mathcal{N}(n\mu, (n\sigma)^2)$. Pour arriver à ce résultat, nous avons manipulé la fonction de densité de T' dans le but de *faire apparaître* la densité d'une loi normale. C'est d'ailleurs pour cette raison que nous sommes passé par la densité, comme quand nous avons cherché une relation entre F_T et Φ, c'est le moyen le plus simple de reconnaître une loi normale.
Nous pouvons maintenant conclure fièrement :

$$\boxed{E(X^n) = e^{n\mu + \frac{n^2}{2}\sigma^2}}$$