# Théorème de Cramér pour des variables de Bernoulli

## Mon développement

Pour une probabilité $\mu$ sur $\mathbb{R}$, on note $M : t \mapsto \mathbb{E}(e^{tX})$, où $X$ est une v.a. de loi $\mu$, sa transformée de Laplace, $\Lambda : t \mapsto \ln M(t)$ sa log-Laplace, et $\Lambda^* : a \mapsto \sup_{t\geq 0}(at - \Lambda(t))$ sa transformée de Cramér.

**Lemme (Inégalité de Chernoff).** Soit $X$ une v.a. réelle, dont la transformée de Cramér $\Lambda^*$ est définie. Pour tout $a \in \mathbb{R}$, on a :

$$\mathbb{P}(X \geq a) \leq \exp(-\Lambda^*(a))$$

Par l'inégalité de Markov, pour tout $t \geq 0$, on a :

$$\mathbb{P}(X \geq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \leq e^{-ta}\,\mathbb{E}(e^{tX}) = \exp(-ta + \Lambda(t)) \ .$$

Ceci étant vrai pour tout $t \geq 0$, on a donc

$$\mathbb{P}(X \geq a) \leq \exp\left(\inf_{t\geq 0}(-ta + \Lambda(t))\right) = \exp(-\Lambda^*(a)) \ .$$

$\square$

**Proposition.** Pour la loi de Bernoulli de paramètre $p \in ]0,1[$, on a :

$$\forall a \in [p, 1[,\ \Lambda^*(a) = a \ln\left(\frac{a}{p}\right) + (1-a)\ln\left(\frac{1-a}{1-p}\right)$$

Pour la loi de Bernoulli de paramètre $p \in ]0,1[$, on a $\forall t \in \mathbb{R}$, $M(t) = 1 - p + pe^t$.
Soit $a \in [p, 1[$. La fonction $f_a : t \mapsto at - \ln(1-p+pe^t)$ est dérivable sur $\mathbb{R}_+$ et

$$\forall t \geq 0,\ f_a'(t) = a - \frac{pe^t}{1-p+pe^t} = a - \frac{p}{(1-p)e^{-t}+p}$$

donc

$$\begin{aligned} f_a'(t) \geq 0 \quad &\Leftrightarrow \quad (1-p)e^{-t} + p \geq \frac{p}{a} \\ &\Leftrightarrow \quad (1-p)e^{-t} \geq p\frac{1-a}{a} \\ &\Leftrightarrow \quad t \leq -\ln\left(\frac{p(1-a)}{(1-p)a}\right) = \ln\left(\frac{(1-p)a}{p(1-a)}\right) \ . \end{aligned}$$

Puisque $p \leq a$, on a $\frac{a-ap}{p-ap} \geq 1$ donc son logarithme est positif et on a :

$$\begin{aligned} \Lambda^*(a) &= f_a\left(\ln\left(\frac{(1-p)a}{p(1-a)}\right)\right) \\ &= a\ln\left(\frac{a(1-p)}{(1-a)p}\right) - \ln\left(1-p+p\frac{a(1-p)}{(1-a)p}\right) \\ &= a\ln\left(\frac{a}{p}\right) - a\ln\left(\frac{1-a}{1-p}\right) - \ln\left(\frac{1-p}{1-a}\right) \end{aligned}$$

ce qui est la formule annoncée. $\square$

**Théorème.** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. i.i.d. de loi de Bernoulli de paramètre $p \in ]0,1[$. On note pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{i=1}^{n} X_i \ .$$

Pour tout $\varepsilon \in ]0, 1-p[$, on a :

$$\lim_{n\to+\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{S_n}{n} \geq p+\varepsilon\right) = -\Lambda^*(p+\varepsilon)$$

On note resp. $M$, $\Lambda$, $\Lambda^*$ les transformées de Laplace, log-Laplace et de Cramér de la loi de Bernoulli de paramètre $p \in ]0,1[$.

D'une part, pour tout $n \in \mathbb{N}^*$, la transformée de Laplace de la loi de $\frac{S_n}{n} - p$ est la fonction qui à $t$ associe

$$\mathbb{E}\left(e^{t(S_n/n-p)}\right) = e^{-tp}\,\mathbb{E}\left(\prod_{i=1}^{n} e^{tX_i/n}\right) = e^{-tp} M\left(\frac{t}{n}\right)^n$$

puisque les $X_i$ sont i.i.d. Sa log-Laplace est donc $t \mapsto -pt + n\Lambda\left(\frac{t}{n}\right)$ et sa transformée de Cramér est $a \mapsto n\Lambda^*(p+a)$. Par l'inégalité de Chernoff, on a donc, pour tout $\varepsilon \in ]0, 1-p[$,

$$\mathbb{P}\left(\frac{S_n}{n} - p \geq \varepsilon\right) \leq \exp(-n\Lambda^*(p+\varepsilon)) \ .$$

Donc

$$\forall n \in \mathbb{N}^*,\ \frac{1}{n} \ln \mathbb{P}\left(\frac{S_n}{n} \geq p+\varepsilon\right) \leq -\Lambda^*(p+\varepsilon) \ . \tag{1}$$

D'autre part, pour tout $n \in \mathbb{N}^*$, on note $k_n = 1 + \lfloor n(p+\varepsilon) \rfloor$. On a

$$\mathbb{P}\left(\frac{S_n}{n} \geq p+\varepsilon\right) \geq \mathbb{P}(S_n = k_n) \ . \tag{2}$$

Or,

$$\mathbb{P}(S_n = k_n) = \frac{n!}{k_n!(n-k_n)!} p^{k_n} (1-p)^{n-k_n}$$

car $S_n$ suit la loi binomiale de paramètres $n$ et $p$. Comme $k_n$ et $n - k_n$ tendent vers $+\infty$ quand $n \to +\infty$, d'après la formule de Stirling, on a, quand $n \to +\infty$ :

$$\begin{aligned}
\mathbb{P}(S_n = k_n) \quad &\sim \quad \frac{e^{-n} n^n \sqrt{2\pi n}}{e^{-k_n} k_n^{k_n} \sqrt{2\pi k_n} e^{k_n - n} (n-k_n)^{n-k_n} \sqrt{2\pi(n-k_n)}} p^{k_n} (1-p)^{n-k_n} \\
&= \quad \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1}{k_n/n(1-k_n/n)}} \left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1-p)}{n-k_n}\right)^{n-k_n}
\end{aligned}$$

donc, $\mathbb{P}(S_n = k_n)$ ne tendant pas vers 1 et $\frac{k_n}{n}$ étant équivalent à $p + \varepsilon$ :

$$\begin{aligned}
\ln \mathbb{P}(S_n = k_n) \quad &\sim \quad -\frac{1}{2}\ln(2\pi n) - \frac{1}{2}\ln\left(\frac{k_n}{n}\left(1 - \frac{k_n}{n}\right)\right) + k_n \ln\left(\frac{np}{k_n}\right) + (n-k_n)\ln\left(\frac{n(1-p)}{n-k_n}\right) \\
&\sim \quad -\frac{1}{2}\ln(2\pi n) - \frac{1}{2}\ln\left((p+\varepsilon)(1-p-\varepsilon)\right) + k_n \ln\left(\frac{p}{p+\varepsilon}\right) + (n-k_n)\ln\left(\frac{1-p}{1-p-\varepsilon}\right) \\
&\sim \quad k_n \ln\left(\frac{p}{p+\varepsilon}\right) + (n-k_n)\ln\left(\frac{1-p}{1-p-\varepsilon}\right)
\end{aligned}$$

donc

$$\frac{1}{n}\ln \mathbb{P}(S_n = k_n) \sim (p+\varepsilon)\ln\left(\frac{p}{p+\varepsilon}\right) + (1-p-\varepsilon)\ln\left(\frac{1-p}{1-p-\varepsilon}\right) = -\Lambda^*(p+\varepsilon)$$

On en déduit à l'aide de (2) que

$$\liminf_{n\to+\infty} \frac{1}{n}\ln \mathbb{P}\left(\frac{S_n}{n} \geq p+\varepsilon\right) \geq -\Lambda^*(p+\varepsilon) \tag{3}$$

Finalement, en regroupant (1) et (3), on obtient la limite souhaitée. □

## Références

J'ai utilisé [GS01, p. 220], [Les01, p. 16]. On peut aussi consulter [CDGCM11, p. 163].

## Leçons correspondantes

J'utilise ce développement pour les leçons 242, 249. On peut également l'utiliser pour les leçons 250, 251.

## Remarques

– Dans le cas d'une loi de Bernoulli, $M$ et $\Lambda$ sont bien définies sur $\mathbb{R}$ mais $\Lambda^*$ n'est définie que sur $]-\infty, 1]$.
– Avec ce théorème, on fait un premier (petit) pas dans la théorie des grandes déviations. Si on n'est pas à l'aise, on peut ne pas mentionner les noms des différents objets et se contenter de faire les calculs (c'est ce que j'avais décidé de faire).
– À l'inverse, si on est très à l'aise et familier avec la théorie, on peut proposer en développement une partie de la démonstration du théorème de Cramér dans $\mathbb{R}$ par exemple.
– L'inégalité de Chernoff et la log-Laplace jouent un rôle important lorsqu'on s'intéresse au phénomène de concentration de la mesure. La transformée de Cramér joue elle un rôle important dans la théorie des grandes déviations.
– Le théorème de Cramér est en quelque sorte le troisième résultat important, venant à la suite de la loi des grands nombres et du théorème central limite, il peut donc éventuellement trouver sa place dans la leçon 250. Par contre, avec le recul, il n'est pas forcément très approprié pour la leçon 251.

## Questions possibles

1. Quelles sont les généralisations possibles de ce résultat ?
2. Calculer $\Lambda^*(a)$ pour $a \notin [p, 1[$.
3. Calculer la transformée de Cramér d'une loi normale.
4. Montrer que, pour une loi de Bernoulli, $\Lambda^*$ est croissante et convexe.

# Références

[CDGCM11] Marie COTTRELL, Christian DUHAMEL, Valentine GENON-CATALOT et Thierry MEYRE : *Exercices de probabilités*. Cassini, 2011.

[GS01] Geoffrey GRIMMETT et David STIRZAKER : *Probability and random processes*. Oxford University Press, 2001.

[Les01] Emmanuel LESIGNE : *Pile ou face*. Ellipses, 2001.