

LOI NORMALE ET ESTIMATION

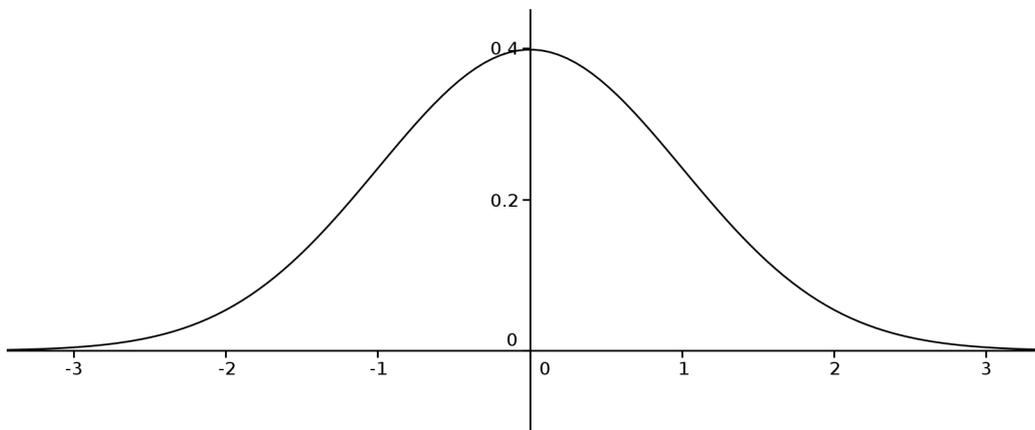
1. Théorème de Moivre-Laplace

Si X_n est une variable aléatoire suivant une loi binomiale $B(n; p)$ et si on note Z_n la variable aléatoire $\frac{X_n - np}{\sqrt{np(1-p)}}$ (On a donc $Z_n = \frac{X_n - E(X_n)}{\sigma(X_n)}$), alors pour tous réels a et b tels que $a \leq b$, $\lim_{n \rightarrow +\infty} p(X_n \in [a; b]) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$.

2. Loi normale centrée réduite

a. Définition

La loi normale centrée réduite est une loi de probabilités dont la densité est donnée par la fonction $x \rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. On la note $N(0; 1)$.



Remarque : La définition est justifiée par le fait que $\lim_{a \rightarrow +\infty} \int_{-a}^a \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$.

b. Propriétés

Si une variable aléatoire X suit une loi normale centrée réduite, alors l'espérance de X est $E(X) = 0$ et la variance de X est $V(X) = E((X - E(X))^2) = 1$.

c. Remarque

Dans la pratique, si X_n suit une loi binomiale $B(n; p)$, on considère que l'on peut remplacer la loi de $Z_n = \frac{X_n - E(X_n)}{\sigma(X_n)}$ par la loi normale centrée réduite si $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$.

3. Loi normale

a. Définition

On dit que la variable aléatoire X suit une loi normale de paramètres μ et σ^2 (notée $N(\mu; \sigma^2)$) si la variable aléatoire $\frac{X-\mu}{\sigma}$ suit la loi normale centrée réduite.

b. Propriétés

Si X suit une loi normale $N(\mu; \sigma^2)$, alors $E(X) = \mu$ et $V(X) = \sigma^2$.

c. Quelques valeurs

Si X suit la loi normale $N(\mu; \sigma^2)$ alors :

$$p(X \in [\mu - \sigma; \mu + \sigma]) \approx 0,683$$

$$p(X \in [\mu - 2\sigma; \mu + 2\sigma]) \approx 0,954$$

$$p(X \in [\mu - 3\sigma; \mu + 3\sigma]) \approx 0,997$$

4. Échantillonnage

a. Théorème

Si X est une variable aléatoire suivant la loi normale centrée réduite, alors pour tout réel α appartenant à $]0; 1[$, il existe un unique réel positif u_α tel que

$$p(-u_\alpha \leq X \leq u_\alpha) = 1 - \alpha.$$

On a notamment les valeurs usuelles $u_{0,05} \approx 1,96$ et $u_{0,01} \approx 2,58$.

b. Intervalle de fluctuation asymptotique

Si X_n suit une loi binomiale $B(n; p)$ alors, si on note $F_n = \frac{X_n}{n}$ la fréquence du

succès, on a $\lim_{n \rightarrow +\infty} p\left(p - u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq F_n \leq p + u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right) = 1 - \alpha$.

Dans la pratique, pour $\alpha = 0,05$ et si $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$,

l'intervalle $I = \left[p - 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}}; p + 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$ est appelé intervalle de

fluctuation asymptotique au seuil de 95 % de F_n . On a : $p(F_n \in I) \approx 0,95$.

Remarque : L'intervalle $J = \left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$ est une assez bonne approximation

de cet intervalle. (plus précisément, $I \subset J$ donc $p(F_n \in J) > 0,95$)

5. Estimation

a. Situation

Si X_n suit une loi binomiale de $B(n; p)$, n étant connu, il s'agit de déterminer un intervalle contenant les valeurs « probables » de p à partir d'une valeur observée x de X_n .

b. Intervalle de confiance

Si $X_{n,p}$ suit une loi binomiale $B(n; p)$, on note $F_{n,p} = \frac{X_{n,p}}{n}$ la fréquence du succès et $f = \frac{x}{n}$ une fréquence observée.

On dit que I est un intervalle de confiance au niveau $1-\alpha$ ($\alpha \in]0 ; 1[$) si pour tout $p \in I$, $p(f \in J_p) \geq 1-\alpha$, où J_p est l'intervalle de fluctuation asymptotique au seuil $1-\alpha$ de $F_{n,p}$.

c. Intervalle de confiance à 95 %

Si $n \geq 30$, $nf \geq 5$ et $n(1-f) \geq 5$, alors $\left[f - \frac{1}{\sqrt{n}}; f + \frac{1}{\sqrt{n}} \right]$ est un intervalle de confiance au niveau de 95 % de p .