Lois de probabilité usuelles (rappels)

Généralités

Fonction de répartition d'une loi discrète

Si X est une variable aléatoire telle que $X(\Omega) = \{x_1, \ldots, x_n\}$, sa fonction de répartition est égale à

$$F_X(x) = P(X \leqslant x) = \sum_{\substack{1 \leqslant i \leqslant n \\ x_i \leqslant x}} P(X = x_i)$$

Fonction de répartition d'une loi continue

Si X est une variable aléatoire de densité f, sa fonction de répartition est égale à

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

On a alors $P(X > x) = 1 - F_X(x)$
et sa densité vaut $f(x) = F'_X(x)$

Probabilités du min et du max

Si les variables T_i sont indépendantes,

$$P(\max T_i \leqslant x) = \prod_{i=1}^{n} P(T_i \leqslant x)$$
$$P(\min T_i \leqslant x) = 1 - \prod_{i=1}^{n} [1 - P(T_i \leqslant x)]$$

Espérance et variance dans le cas discret

Si X est une variable aléatoire discrète,

$$E(X) = \sum_{k=0}^{+\infty} k P(X = k)$$
$$E(X^2) = \sum_{k=0}^{+\infty} k^2 P(X = k)$$
$$V(X) = E(X^2) - E(X)^2$$

Espérance et variance dans le cas continu

Si X est une variable aléatoire continue de densité f,

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$$
$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x$$
$$V(X) = E(X^2) - E(X)^2$$

Propriétés de l'espérance et de la variance

Si X et Y sont deux variables aléatoires et a un réel,

$$E(aX) = aE(X)$$
$$E(X + Y) = E(X) + E(Y)$$

Important : toujours calculer à l'intérieur de l'espérance avant de séparer les termes. Par exemple,

$$E((X - a)^2) = E(X^2 - 2aX + a^2) = E(X^2) - 2aE(X) + a^2$$

Si les X_i sont des variables aléatoires,

$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

et si elles sont indépendantes,

$$V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i)$$

Principales lois discrètes

Loi uniforme

$X(\Omega) = \{x_1, \ldots, x_n\}$
$P(X = x_i) = 1/n$

Loi de Bernoulli $\mathcal{B}(p)$

$X(\Omega) = \{0, 1\}$, paramètre p
$P(X = 1) = p$, $P(X = 0) = 1 - p$
$E(X) = p$, $V(X) = p(1 - p)$

Loi binomiale $\mathcal{B}(n, p)$

$X(\Omega) = \{0, \ldots, n\}$, paramètre p
$P(X = k) = \binom{n}{k}\, p^k\, (1 - p)^{n-k}$
$E(X) = np$, $V(X) = np(1 - p)$

Loi hypergéométrique $\mathcal{H}(N, n, N_A)$

On effectue n tirages sans remise dans une urne contenant N objets dont $N_A \geqslant n$ objets de type A. X est le nombre d'objets de type A obtenus.
$X(\Omega) = \{1, \ldots, n\}$, paramètres N, n et N_A $(p = N_A/N)$
$P(X = k) = \frac{\binom{N_A}{k}\binom{N-N_A}{n-k}}{\binom{N}{k}}$
$E(X) = np$, $V(X) = np(1 - p)(N - n)/(N - 1)$

Loi de Poisson $\mathcal{P}(\lambda)$

$X(\Omega) = \mathbb{N}$, paramètre λ
$P(X = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$
$E(X) = \lambda$, $V(X) = \lambda$

Loi géométrique

$X(\Omega) = \mathbb{N}^*$, paramètre p
$P(X = k) = (1 - p)^{k-1} p$
$E(X) = 1/p$, $V(X) = (1 - p)/p^2$

Principales lois continues

Loi uniforme $\mathcal{U}(a, b)$

$X(\Omega) = [\,a\,;b\,]$, paramètres a et b
$$f(x) = \begin{cases} 1/(b - a) & \text{si } x \in [\,a\,;b\,] \\ 0 & \text{sinon} \end{cases}$$
$E(X) = (a + b)/2$, $V(X) = (b - a)^2/12$

Loi exponentielle $\mathcal{E}(\lambda)$

$X(\Omega) = \mathbb{R}^+$, paramètre λ
$$f(x) = \begin{cases} 1/\lambda\ e^{-x/\lambda} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$$
$E(X) = 1/\lambda$, $V(X) = 1/\lambda^2$

Loi normale $\mathcal{N}(m, \sigma)$

$X(\Omega) = \mathbb{R}$, paramètres m (moyenne) et σ (écart-type)
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right) \quad \text{si } x \in \mathbb{R}$$
$E(X) = m$, $V(X) = \sigma^2$

Loi du khi-deux χ^2_n

$X(\Omega) = \mathbb{R}^+$, paramètre n (degré de liberté)
$E(X) = n$, $V(X) = 2n$

Loi de Student T_n
$\mathrm{X}(\Omega) = \mathbb{R}$, paramètre n (degré de liberté)
$\mathrm{E}(\mathrm{X}) = 0$ pour $n > 1$, $\mathrm{V}(\mathrm{X}) = n/(n-2)$ pour $n > 2$

Relations entre les principales lois

Propriétés

Si les variables X_i suivent une loi $\mathcal{B}(p)$ et sont indépendantes, alors la variable $\sum_{i=1}^{n} \mathrm{X}_i$ suit une loi $\mathcal{B}(n, p)$.

Si les variables X_i suivent une loi $\mathcal{P}(\lambda_i)$ et sont indépendantes, alors la variable $\sum_{i=1}^{n} \mathrm{X}_i$ suit une loi $\mathcal{P}(\sum \lambda_i)$.

Si la variable X suit une loi $\mathcal{N}(m, \sigma^2)$, la variable $\mathrm{Y} = a\mathrm{X} + b$ suit une loi $\mathcal{N}(am + b, a^2\sigma^2)$.

Si la variable X suit une loi $\mathcal{N}(m, \sigma^2)$, alors la variable $\mathrm{Y} = (\mathrm{X} - m)/\sigma$ suit une loi $\mathcal{N}(0, 1)$. En particulier, $\mathrm{P}(\mathrm{X} \leqslant u) = \mathrm{P}(\mathrm{Y} \leqslant (u - m)/\sigma)$.

Approximations (voir chapitre 2)

Si $n \geqslant 30$ et $np < 5$, on peut approcher une loi $\mathcal{B}(n, p)$ par une loi $\mathcal{P}(np)$.

Si $n \geqslant 30$, $np \geqslant 5$ et $n(1 - p) \geqslant 5$, alors on peut approcher une loi $\mathcal{B}(n, p)$ par une loi $\mathcal{N}(np, np(1 - p))$.

Si $\mathrm{N} \geqslant 10n$, on peut approcher une loi $\mathcal{H}(\mathrm{N}, n, p\mathrm{N})$ par une loi $\mathcal{B}(n, p)$.

Si λ est assez grand, on peut approcher une loi $\mathcal{P}(\lambda)$ par une $\mathcal{N}(\lambda, \sqrt{\lambda})$.

Si n est assez grand, on peut approcher une loi χ^2_n par une loi $\mathcal{N}(n, \sqrt{2n})$.

Si n est assez grand, on peut approcher une loi T_n par une loi $\mathcal{N}(0, \sqrt{1})$.

Lois normale, du χ^2 et de Student (voir chap. 1)

Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes et $\mathrm{X}_i \rightsquigarrow \mathcal{N}(0, 1)$ pour tout $i \in \{1, \ldots, n\}$, alors ${\mathrm{X}_1}^2 + \cdots + {\mathrm{X}_n}^2 \rightsquigarrow \chi^2_n$.
Si $\mathrm{X} \rightsquigarrow \mathcal{N}(0, 1)$, Y suit une loi de χ^2 à n degrés de liberté et X et Y sont indépendantes, alors $\mathrm{Z} = \sqrt{n}\,\mathrm{X}/\sqrt{\mathrm{Y}}$ suit une loi de Student à n degrés de liberté.

Cas particuliers importants (moments empiriques) :
Moyenne empirique

$$\overline{\mathrm{X}}_n = \frac{1}{n} \sum_{i=1}^{n} \mathrm{X}_i$$

Variance empirique

$$\mathrm{S}^2_n = \frac{1}{n-1} \sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}}_n)^2$$

Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes et $\mathrm{X}_i \rightsquigarrow \mathcal{N}(0, 1)$ pour tout $i \in \{1, \ldots, n\}$, alors $\mathrm{Y}_n = \sum_{i=1}^{n} {\mathrm{X}_i}^2 \rightsquigarrow \chi^2_n$.
Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes et $\mathrm{X}_i \rightsquigarrow \mathcal{N}(m, \sigma)$ pour tout $i \in \{1, \ldots, n\}$, alors $\mathrm{Z}_n = \sum_{i=1}^{n} (\mathrm{X}_i - m)^2/\sigma^2 \rightsquigarrow \chi^2_n$.
Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes et $\mathrm{X}_i \rightsquigarrow \mathcal{N}(m, \sigma)$ pour tout $i \in \{1, \ldots, n\}$, alors

$$(n-1){\mathrm{S}_n}^2/\sigma^2 \rightsquigarrow \chi^2_{n-1}$$

(car les $\mathrm{X}_i - \overline{\mathrm{X}_n}$ sont liées par une relation : leur somme vaut 0 puisque $\overline{\mathrm{X}_n} = \sum \mathrm{X}_i/n$).
Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes et $\mathrm{X}_i \rightsquigarrow \mathcal{N}(m, \sigma)$ pour tout $i \in \{1, \ldots, n\}$, alors $\mathrm{T}_n = \sqrt{n}\frac{\overline{\mathrm{X}_n} - m}{\mathrm{S}_n} \rightsquigarrow \mathrm{T}(n-1)$.

Utiliser les tables statistiques

Loi normale centrée réduite

La table de la fonction de répartition de la loi normale centrée réduite donne les valeurs de $\mathrm{P}(\mathrm{X} \leqslant u)$ pour u positif donné. Pour déterminer d'autres valeurs, on utilise les formules suivantes :

$$\mathrm{P}(\mathrm{X} > u) = 1 - \mathrm{P}(\mathrm{X} \leqslant u)$$
$$\mathrm{P}(u < \mathrm{X} < v) = \mathrm{P}(\mathrm{X} < v) - \mathrm{P}(\mathrm{X} \leqslant u)$$
$$\mathrm{P}(|\mathrm{X}| \leqslant u) = \mathrm{P}(-u \leqslant \mathrm{X} \leqslant u)$$
$$\mathrm{P}(|\mathrm{X}| \geqslant u) = 2\mathrm{P}(\mathrm{X} \geqslant u)$$
$$\mathrm{P}(\mathrm{X}^2 \leqslant u) = \mathrm{P}(|\mathrm{X}| \leqslant \sqrt{u})$$

Pour u négatif, on utilise

$$\mathrm{P}(\mathrm{X} \leqslant u) = \mathrm{P}(\mathrm{X} > -u) = 1 - \mathrm{P}(\mathrm{X} \leqslant -u)$$

On peut également utiliser la table « à l'envers », pour déterminer u tel que $\mathrm{P}(\mathrm{X} \leqslant u) = p$ pour p donné.

Loi de χ^2

La table de la loi de χ^2 donne la valeur u telle que $\mathrm{P}(\mathrm{X} \geqslant u) = p$ pour p donné. Pour déterminer u telle que $\mathrm{P}(\mathrm{X} \leqslant u) = p$ pour p donné, on utilise la formule

$$\mathrm{P}(\mathrm{X} > u) = 1 - \mathrm{P}(\mathrm{X} \leqslant u) = 1 - p$$

et on cherche dans la table la valeur u correspondant à $1 - p$.

Lorsque l'on cherche deux valeurs u_1 et u_2 telles que $\mathrm{P}(u_1 \leqslant \mathrm{X} \leqslant u_2) = p$, on considère un intervalle symétrique et on cherche u_1 et u_2 tels que $\mathrm{P}(\mathrm{X} \geqslant u_2) = p/2$ et $\mathrm{P}(\mathrm{X} \geqslant u_1) = 1 - \mathrm{P}(\mathrm{X} < u_1) = 1 - p/2$.

Loi de Student

La table de la loi de Student donne la valeur de u telle que $\mathrm{P}(|\mathrm{X}| > u) = p$ pour p donné. Pour trouver la valeur de u telle que $\mathrm{P}(-u \leqslant \mathrm{X} \leqslant u) = p$ pour p donné, on cherche la valeur de u telle que $\mathrm{P}(|\mathrm{X}| > u) = 1 - p$.