Formulaire de Probabilités et Statistiques

AD+JS

1 Rappels de combinatoire

Arrangements avec répétitions Nombre d'applications d'un ensemble à k éléments dans un ensemble à n éléments :

$$n^k$$

Arrangements (sans répétitions) Nombre d'arrangements de k objets choisis parmi n:

$$A_n^k = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1)$$

Permutations Nombre de permutations de n objets :

$$A_n^n = n!$$

Combinaisons (sans répétitions) Nombre de combinaisons de k objets choisis parmi n (ou nombre de sous-ensembles à k éléments dans un ensemble à n éléments):

$$C_n^k = \binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \qquad \text{(coefficient binomial)}$$

Combinaisons avec répétitions Nombre de répartitions de n boules indiscernables dans k urnes discernables:

$$C_{n+k-1}^n = C_{n+k-1}^{k-1} = \frac{(n+k-1)!}{n!(k-1)!}$$

Coefficients multinomiaux Nombre de répartitions de n boules discernables dans k urnes discernables avec n_1 boules dans la 1ère urne, n_2 boules dans la 2ème urne, …, n_k boules dans la kème urne:

$$C_n^{n_1} C_{n-n_1}^{n_2} \cdots C_{n-n_1-\cdots-n_{k-1}}^{n_k} = \begin{pmatrix} n \\ n_1, \ldots, n_k \end{pmatrix} = \frac{n!}{n_1! n_2! \cdots n_k!} \qquad \text{(coefficient multinomial)}$$

Propriétés élémentaires des coefficients binomiaux

$$\begin{aligned}
C_n^{k-1} + C_n^k &= C_{n+1}^k && \text{(triangle de Pascal)} \\
C_{n+m}^k &= \sum_{j=0}^k C_m^j C_n^{k-j} && \text{(formule de Vandermonde)} \\
(x+y)^n &= \sum_{k=0}^n C_n^k x^k y^{n-k} && \text{(formule du binôme)} \\
(1-x)^{-n} &= \sum_{k=0}^\infty C_{n+k-1}^k x^k && \text{(série géométrique généralisée)}
\end{aligned}$$

2 Probabilités élémentaires

Modelè probabiliste L'ensemble fondamental Ω est l'ensemble des résultats possibles d'une expérience. Un événement A est un sous-ensemble de Ω. Une distribution de probabilités P sur Ω associe à tout événement A un nombre $P(A)$, tel que $0 \le P(A) \le 1$, appelé probabilité de A.

Ci-dessous, soient A, B, A_1, A_2, …, A_n des événements. On note A^c le complémentaire de A dans Ω.

Une partition de Ω (un système complet d'événements) $\{B_i\}_{i=1}^n$ a les propriétés que l'union des événements est Ω et qu'ils sont disjoints deux á deux:

$$\bigcup_{i=1}^n B_i = \Omega, \qquad B_j \cap B_j = \emptyset, \quad i \neq j.$$

Propriétés élémentaires de P

$$\begin{aligned}
P(\Omega) &= 1 \\
P(A^c) &= 1 - P(A) \\
P(A \cup B) &= P(A) + P(B) - P(A \cap B)
\end{aligned}$$

Formule d'inclusion-exclusion à n termes

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} P(A_{i_1} \cap \cdots \cap A_{i_k})$$

Événements indépendants A et B sont indépendants ssi

$$P(A \cap B) = P(A)P(B)$$

Événements indépendants deux à deux

$$P(A_i \cap A_j) = P(A_i)P(A_j) \qquad i \neq j$$

Événements totalement indépendants

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}) \quad 1 \le i_1 < \cdots < i_k \le n, k = 1, \ldots, n$$

Probabilités conditionnelles Si A est un événement quelconque et si l'événement B vérifie $P(B) > 0$ alors la probabilité conditionnelle de A sachant B est

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Conditionnement multiple

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1) \cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Formule des probabilités totales Soit $(B_i)_{1 \le i \le n}$ une partition de Ω, avec $P(B_i) > 0$, $\forall i$. Alors pour tout événement A,

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n).$$

Formule de Bayes Soit $\{B_i\}_{i=1}^n$ une partition de Ω, avec $P(B_i) > 0$, $\forall i$. Alors pour tout événement A de probabilité non nulle

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n)}.$$

3 Variables aléatoires réelles

Variables discrètes Une variable aléatoire réelle X est une fonction $X : \Omega \to E$ ou E est un sous-ensemble de nombres réels. Si E est discrèt, alors X est appelé variable aléatoire discrète. Dans ce cas la distribution de probabilités de X est la donnée des nombres : $P(X = x_k)$. La fonction $f_X(x) = P(X = x)$ est appelée fonction de masse.

Fonction de répartition $F_X(t) = P(X \le t)$

Quantile Pour une probabilité $0 < p < 1$, la p quantile de F_X est $x_p = \inf\{x : F_X(x) \ge p\}$. Très souvent on a $F_X(x_p) = p$.

Variables continues X est une variable aléatoire continue si sa fonction de répartition s'écrit sous la forme

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \qquad \text{donc} \qquad f_X(x) = \frac{dF_X(x)}{dx},$$

où $f_X(x)$ est une fonction non négative, appelée densité de X.

Couples de variables aléatoires On considère des événements relatifs à deux variables aléatoires X et Y.

Fonction de répartition conjointe $F_{X,Y}(x,y) = P(X \le x, Y \le y)$. Pour un couple (X,Y) conjointement continu

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} ds \int_{-\infty}^{y} dy\, f_{X,Y}(s,t) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(s,t)\, dsdt$$

et $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$.

Fonction de répartition marginale $F_X(x) = P(X \le x) = \lim_{y\to\infty} F_{X,Y}(x,y)$. Pour un couple (X,Y) conjointement continu

$$F_X(x) = \int_{-\infty}^{x} f_X(s)\, ds$$

où $f_X(x)$ est la densité marginale de X donnée par $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$.

Variables aléatoires indépendantes Deux variables aléatoires X et Y sont indépendantes si pour tous ensembles A et B,

$$P(X \in A \text{ et } Y \in B) = P(X \in A)\, P(Y \in B).$$

En particulier:

$$F_{X,Y}(x,y) = F_X(x)\, F_Y(y).$$

Dans le cas de deux variables discrètes

$$P(X = x_k \text{ et } Y = y_l) = P(X = x_k)\, P(Y = y_l), \qquad \forall x_k, y_l \in \mathbb{R},$$

et dans le cas de deux variables continues

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y), \qquad \forall x, y \in \mathbb{R}.$$

Loi de probabilité d'une fonction de variables aléatoires

Changement de variables à une dimension Si X est une variable aléatoire continue de densité $f_X(x)$ et si $Y = g(X)$ pour une fonction g réelle et inversible, on a

$$F_Y(y) = \int_{g^{-1}(-\infty,y)} f_X(x)dx.$$

Si $g(x)$ est strictement monotone et continûment dérivable, alors la variable aléatoire $Y = g(X)$ est une variable aléatoire continue, avec densité

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\left|\frac{d\, g^{-1}(y)}{dy}\right|, & \text{si } y = g(x) \text{ pour un } x \text{ quelconque,} \\ 0, & \text{si } y \ne g(x) \text{ pour tout } x. \end{cases}$$

Changement de variables à deux dimension Si (X_1, X_2) sont deux variables aléatoires conjointement continues de densité $f_{X_1,X_2}(x_1,x_2)$ et si $(Y_1,Y_2) = (g_1(X_1,X_2), g_2(X_1,X_2))$ pour une fonction $g = (g_1, g_2)$ inversible, on a

$$F_{Y_1,Y_2}(y_1,y_2) = \iint_{\substack{g_1(x_1,x_2) \le y_1 \\ g_2(x_1,x_2) \le y_2}} f_{X_1,X_2}(x_1,x_2)\, dx_1 dx_2$$

En particulier, si $g(x)$ est une bijection continûment dérivable avec son inverse noté par $h = (h_1, h_2)$, alors le couple aléatoire (Y_1, Y_2) est conjointement continu, avec densité

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}\left(h_1(y_1,y_2), h_2(y_1,y_2)\right) |J(y_1,y_2)|$$

où $|J(y_1,y_2)|$ est le jacobien de h, et

$$J(y_1,y_2) = \frac{\partial h_1}{\partial y_1}\frac{\partial h_2}{\partial y_2} - \frac{\partial h_1}{\partial y_2}\frac{\partial h_2}{\partial y_1}.$$

4 Espérance mathématique

Définition Soit X une variable aléatoire. On désigne par

$$\mathrm{E}[X] = \int x dF_X(x)$$

l'espérance mathématique de X, definie quand $\mathrm{E}(|X|) < \infty$.

Espérance d'une fonction de variables aléatoires Soit g une fonction réelle. L'espérance de $g(X)$ se calcul par $\mathrm{E}[g(X)] = \int g(x) dF_X(x)$. En particulier,

pour une variable discrète $\mathrm{E}[g(X)] = \sum_{k \in K} g(x_k) P(X = x_k)$

pour une variable continue $\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

pour un couple (X, Y) $\mathrm{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\, dxdy$

Propriétés de l'espérance

Espérance d'une constante $\mathrm{E}[c] = c$ pour toute constante c réelle

Linéarité $\mathrm{E}[a + bX + cY] = a + b\mathrm{E}[X] + c\mathrm{E}[Y]$ pour tous réels a, b, c

Positivité $\mathrm{E}[X] \geq 0$ si $X \geq 0$

Variance et covariance Soient X et Y variables aléatoires telles que $\mathrm{E}(X^2), \mathrm{E}(Y^2) < +\infty$. On définit les quantités suivantes:

Variance de X

$$\mathrm{var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

Covariance de X et Y

$$\mathrm{cov}[X, Y] = \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

Coefficient de corrélation de X et Y

$$\mathrm{corr}[X, Y] = \frac{\mathrm{cov}[X,Y]}{\sqrt{\mathrm{var}[X]}\sqrt{\mathrm{var}[Y]}}$$

Variance et covariance — propriétés: $\mathrm{var}[X] = \mathrm{cov}[X, X]$, et $\mathrm{var}[X] \geq 0$.

(Bi-)Linéarité de la covariance Pour a, b, c réels,

$$\mathrm{cov}[a + bX_1 + cX_2, Y] = b\ \mathrm{cov}[X_1, Y] + c\ \mathrm{cov}[X_2, Y].$$

Variance et covariance

$$\mathrm{var}[X + Y] = \mathrm{var}[X] + \mathrm{var}[Y] + 2\mathrm{cov}[X, Y]$$

Homogénéité de la variance Pour tout a réel, $\mathrm{var}[aX] = a^2\ \mathrm{var}[X]$.

Transformée de Laplace et moments Soit X une variable aléatoire et $t \in \mathbb{R}$. On appelle fonction génératrice des moments (FGM) (aussi transformée de Laplace) la fonction définie par

$$\mathcal{M}_X(t)(\equiv (\mathcal{L}_X(t)) = \mathrm{E}[e^{tX}],$$

si elle est finie.

Moment d'ordre k Pour tout entier positif k, le moment d'ordre k est $\mathrm{E}[X^k]$.

Moment centré d'ordre k Pour tout entier positif k, le moment centré d'ordre k est $\mathrm{E}[(X - \mathrm{E}[X])^k]$. En particulier, si $k = 2$, on a $\mathrm{var}[X]$.

Fonction génératrice des cumulants Soit X une variable aléatoire et $t \in \mathbb{R}$. On appelle fonction génératrice des cumulants la fonction définie par

$$\mathcal{K}_X(t) = \log \mathcal{M}_X(t) = \log \mathrm{E}[e^{tX}],$$

si elle est finie.

Cumulant d'ordre k Pour tout entier positif k, le cumulant d'ordre k est κ_k donné par

$$\mathcal{K}_X(t) = \sum_{k=1}^{\infty} \kappa_k \frac{t^k}{k!};$$

on a également $\kappa_k = d^k \mathcal{K}_X(t)/dt^k|_{t=0}$. En particulier, $\mathrm{E}[X] = \kappa_1$, $\mathrm{var}[X] = \kappa_2$.

Quelques inégalités classiques

Inégalité de Cauchy–Schwarz $|\mathrm{E}[XY]| \leq \sqrt{\mathrm{E}[X^2]} \cdot \sqrt{\mathrm{E}[Y^2]}$

Inégalité de Jensen Si g est une fonction convexe, i.e. pour $t \in [0, 1]$: $g(t\,x + (1-t)\,y) \leq t\,g(x) + (1-t)\,g(y)$, alors

$$g(\mathrm{E}[X]) \leq \mathrm{E}[g(X)]$$

Inégalité de Markov Si $X \geq 0$ et si $a > 0$, alors

$$P(X \geq a) \leq \mathrm{E}[X]/a.$$

Inégalité de Bienaymé–Chebychev Pour tout $a > 0$, on a

$$P(|X - \mathrm{E}[X]| \geq a) \leq \mathrm{var}[X]/a^2.$$

5 Variables aléatoires — Lois usuelles

5.1 Lois discrètes

Nom	Paramètres	$P(X=k)$	$\mathrm{E}[X]$	$\mathrm{var}[X]$	$\mathcal{M}_X(t)$
Bernoulli	$0 \le p \le 1$	$p^k(1-p)^{1-k}$ $k=0,1$	p	$p(1-p)$	pe^t+1-p $\forall t$
Binomiale	n $0 \le p \le 1$	$C_n^k p^k (1-p)^{n-k}$ $k \in \{0,1,\ldots,n\}$	np	$np(1-p)$	$(pe^t+1-p)^n$ $\forall t$
Géométrique	$0 < p \le 1$	$p(1-p)^{k-1}$ pour $k \in \{1,2,\ldots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$ si $(1-p)e^t < 1$
Binomiale négative	$n \in \{2,3,\ldots\}$ $0 \le p \le 1$	$C_{k-1}^{n-1} p^n (1-p)^{k-n}$ $k \in \{n, n+1, \ldots\}$	$\frac{n}{p}$	$n \cdot \frac{1-p}{p^2}$	$\left(\frac{pe^t}{1-(1-p)e^t}\right)^n$ si $(1-p)e^t < 1$
Hyper-géométrique	n, N_1, N_2 $n \le N_1+N_2$	$\frac{C_{N_1}^k \cdot C_{N_2}^{n-k}}{C_{N_1+N_2}^n}$ $\max(0, n-N_2) \le k \le \min(n, N_1)$	$\frac{N_1 \cdot n}{N_1+N_2}$	$\frac{nN_1N_2(N_1+N_2)}{(N_1+N_2)^2(N_1+N_2-1)}$	
Poisson	$\lambda > 0$	$e^{-\lambda}\frac{\lambda^k}{k!}$ $k \in \{0,1,\ldots\}$	λ	λ	$e^{\lambda(e^t-1)}$ $\forall t$

5.2 Lois continues

Nom	Paramètres	Densité $f(x)$	$\mathrm{E}[X]$	$\mathrm{var}[X]$	$\mathcal{M}_X(t)$
Uniforme	$a < b$	$1/(b-a)$ si $x \in [a,b]$ (0 sinon)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$ $\forall t \neq 0$
Exponentielle	$\lambda > 0$	$\lambda e^{-\lambda x}$ pour $x > 0$ (0 sinon)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$ si $t < \lambda$
Gamma	$\lambda > 0, n \ge 1$	$\frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}$ pour $x > 0$ (0 sinon)	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$(\frac{\lambda}{\lambda-t})^n$ si $t < \lambda$
Gaussienne	$\mu \in \mathbb{R}, \sigma^2 > 0$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ pour tout x	μ	σ^2	$\exp(\mu t + \frac{\sigma^2}{2}t^2)$ $\forall t$
Student (t_ν)	$\nu \ge 1$	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})(1+x^2)^{(\nu+1)/2}}$ pour tout x	0 si $\nu \ge 2$	$\frac{1}{\nu-2}$ si $\nu \ge 3$	
Chi deux (χ^2_ν)	$\nu \ge 1$	$\frac{x^{\nu/2-1}e^{-x/2}}{2^{\nu/2}\Gamma(\nu/2)}$ pour $x > 0$ (0 sinon)	ν	2ν	$(1-2t)^{-\frac{\nu}{2}}$ si $t < \frac{1}{2}$

5.3 Théorèmes limites

Loi hypergéométrique vers une loi binomiale Soit (X_m) une suite de variables hypergéométriques de paramètres $N_1(m), N_2(m)$ et n, avec $\frac{N_1(m)}{N_1(m)+N_2(m)} \underset{m\to\infty}{\longrightarrow} p$. Alors pour chaque valeur $0 \le k \le n$:

$$P(X_m = k) \underset{m\to\infty}{\longrightarrow} C_n^k p^k (1-p)^{n-k}.$$

Loi binomiale vers une loi de Poisson Soit X_n une variable aléatoire binomiale de paramètres (n,p). Si $n \to \infty$ et $p \to 0$ tel que $np \to \lambda$, alors

$$P(X_n = k) \underset{n\to\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,2,\ldots$$

Loi géométrique vers une loi exponentielle Soit T une variable aléatoire géometrique de paramètre p. Si $n \to \infty$ et $p \to 0$ tel que $n \cdot p \to \lambda > 0$, alors

$$P(T/n > t) \underset{n\to\infty}{\longrightarrow} e^{-\lambda t}, \quad t > 0.$$

Loi binomiale vers une loi gaussienne Soit X_n une variable aléatoire binomiale de paramètres (n,p).

Théorème de Moivre-Laplace Si $n \to \infty$ et $k \sim np$, alors

$$P(X_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}}$$

Théorème central limite Si $n \to \infty$,alors pour tout t réel

$$P(X_n \le np + \sqrt{np(1-p)} \cdot t) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

6 Variables indépendantes et théorèmes limites

Indépendance de plusieurs variables aléatoires Les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes si et seulement si elles vérifient l'une des conditions équivalentes suivantes:

(i) Les fonctions de répartition respectives $F_X, F_{X_1}, \ldots, F_{X_n}$ de $X = (X_1, \ldots, X_n), X_1, \ldots, X_n$ vérifient pour tous x_i réels , $i = 1, \ldots, n$:

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \cdots F_{X_n}(x_n).$$

(ii) Pour toutes fonctions $g_1, \ldots, g_n$, on a

$$E[g_1(X_1) \cdots g_n(X_n)] = \mathrm{E}[g_1(X_1)] \cdots \mathrm{E}[g_n(X_n)]$$

Une ensemble de variables $X_1, \ldots, X_n$ indepéndantes et identiquement distribuées (iid) est appelé un échantillon aléatoire. On note, par exemple $X_1, \ldots, X_n \overset{\text{iid}}{\sim} F$, ou $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f$.

Cas particuliers de (ii) Si $X_1, \dots, X_n$ sont des variables aléatoires indépendantes:

$$E[X_1 \cdots X_n] = \mathrm{E}[X_1] \cdots \mathrm{E}[X_n], \quad \mathcal{M}_{X_1+\cdots+X_n}(t) = \mathcal{M}_{X_1}(t) \cdots \mathcal{M}_{X_n}(t)$$

Indépendance et covariance Si $X_1, \dots, X_n$ sont indépendantes, alors

$$\mathrm{Cov}[X_i, X_j] = 0, \quad \forall i \neq j.$$

Maximum et minimum de variables aléatoires indépendantes Soient

$$X_{(n)} = \max(X_1, \dots, X_n) \quad \text{et} \quad X_{(1)} = \min(X_1, \dots, X_n).$$

$X_{(n)}$ et $X_{(1)}$ ont pour fonctions de répartition respectives

$$F_{X_{(n)}}(x) = F_{X_1}(x) \cdots F_{X_n}(x), \quad F_{X_{(1)}}(x) = 1 - (1 - F_{X_1}(x)) \cdots (1 - F_{X_n}(x))$$

Théorèmes limites pour les valeurs extrêmes de variables iid Soit $X_1, \dots, X_n, \dots \overset{\text{iid}}{\sim} F$, et soient $\tau > 0$, et $(x_n)_{n \geq 1}$, une suite de nombres réels vérifiant

$$\lim_{n\to\infty} n[1 - F(x_n)] = \tau.$$

Alors

$$P(X_{(n)} \leq x_n) \underset{n\to\infty}{\longrightarrow} e^{-\tau}.$$

Si une telle suite existe, alors il existent des suites $(b_n)_{n\geq 1}$, $(a_n > 0)_{n \geq 1}$, telles que

$$P\left(\frac{X_{(n)} - b_n}{a_n} \leq y\right) \underset{n\to\infty}{\longrightarrow} H(y) = \begin{cases} \exp\left[-\{1 + \xi(y-\eta)/\tau\}_+^{-1/\xi}\right], & \xi \neq 0, \\ \exp\left[-\exp\{-(y-\eta)/\tau\}\right], & \xi = 0, \end{cases}$$

où $u_+ = u$ pour $u > 0$ et $u_+ = 0$ pour $u \leq 0$, et $\eta, \xi \in \mathbb{R}$, $\tau > 0$.

Statistique d'ordre et vecteur des rangs Soient $X_1, \dots, X_n \overset{\text{iid}}{\sim} f$, ceci étant une fonction de densité. Les variables ordonnées $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ sont appelées statistique d'ordre associé à $X_1, \dots, X_n$. On appelle R_j le rang de la valeur X_j et $(R_1, \dots, R_n)$ le vecteur des rangs associé à $X_1, \dots, X_n$. Alors le vecteur des rangs et la statistique d'ordre sont indépendantes et $(R_1, \dots, R_n)$ a une distribution uniforme dans l'ensemble des permutations de $\{1, \dots n\}$. La statistique d'ordre a pour densité

$$f_{X_{(1)},\dots,X_{(n)}}(x_1, \dots, x_n) = \begin{cases} n! f(x_1) \dots f(x_n), & x_1 \leq \cdots \leq x_n, \\ 0, & \text{sinon.} \end{cases}$$

Loi limite des statistiques d'ordre centrales Soient $0 < p < 1$, $X_1, \dots, X_n \overset{\text{iid}}{\sim} F$, F une loi continue avec densité f, et $x_p = F^{-1}(p)$. Alors si $f(x_p) > 0$,

$$\frac{X_{(\lceil np \rceil)} - x_p}{[p(1-p)/\{n f(x_p)^2\}]^{1/2}} \xrightarrow{D} N(0,1) \quad \text{lorsque } n \to \infty.$$

En particulier la mediane, avec $p = 0.5$, a une loi limite normale, de même que les quartiles, avec $p = 0.25, 0.75$: pour n grand,

$$X_{(\lceil np \rceil)} \quad \dot{\sim} \quad N\left(x_p, \frac{p(1-p)}{n f(x_p)^2}\right).$$

Sommes de variables indépendantes Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et $S_n = X_1 + \cdots + X_n$.

Somme de deux variables indépendantes discrètes

$$P(S_2 = y) = \sum_{(j,k): y = x_j^{(1)} + x_k^{(2)}} P(X_1 = x_j^{(1)}) P(X_2 = x_k^{(2)})$$

Somme de deux variables indépendantes continues La densité f_{S_2} de $S_2 = X_1 + X_2$ est donnée par la convolution de f_{X_1} et f_{X_2}, i.e.

$$f_{S_2}(y) = f_{X_1} * f_{X_2}(y) = \int_{-\infty}^{+\infty} f_{X_1}(y-x) f_{X_2}(x) dx$$

Somme de n variables indépendantes continues La densité f_{S_n} de $S_n = X_1 + \cdots + X_n$ est donnée par la convolution des densités f_{X_j} : $f_{S_n} = f_{X_1} * \cdots * f_{X_n}$.

Théorèmes de stabilité

Loi binomiale Soient S_m et S_n deux variables aléatoires binomiales indépendantes de paramètres (m, p) et (n, p). Alors $S_m + S_n$ suit une loi binomiale de paramètres $(m + n, p)$.

Loi de Poisson Soient X_1 et X_2 deux variables aléatoires de Poisson indépendantes de paramètres λ_1 et λ_2. Alors $X_1 + X_2$ suit une loi Poisson de paramètres $\lambda_1 + \lambda_2$.

Loi normale Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes avec $X_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$, et soient $a, b_1, \dots, b_n$ des constantes. Alors la combinaison linéaire

$$a + b_1 X_1 + \cdots + b_n X_n \sim \mathcal{N}(a + b_1\mu_1 + \cdots + b_n\mu_n, b_1^2\sigma_1^2 + \cdots + b_n^2\sigma_n^2).$$

Lois des grands nombres Soit $X_1, \dots, X_n, \dots$ une suite de variables aléatoires réelles iid d'espérance μ et variance σ^2. Alors, la moyenne $S_n/n = (X_1 + \cdots + X_n)/n$ converge en probabilité vers μ (**loi faible**). La moyenne converge aussi presque sûrement vers $\mathrm{E}[X]$, i.e. la probabilité de l'événement " S_n/n converge vers $\mathrm{E}[X]$ " est égale 1 (**loi forte**).

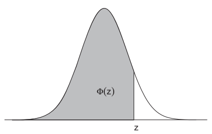
Théorème central limite Soit $X_1, \dots, X_n, \dots$une suite de variables aléatoires réelles iid d'espérance μ et variance σ^2. La suite des sommes partielles centrées et réduites

$$Z_n = \tfrac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

converge en distribution vers une variable $\mathcal{N}(0,1)$ (normale standard):

$$P(Z_n \leq z) \underset{n\to\infty}{\longrightarrow} \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2}\, dy = \Phi(z), \quad \forall z \in \mathbb{R}.$$

Distribution normale standard $\Phi(z)$



Pour $z < 0$ on utilise symmetrie: $P(Z \leq z) = \Phi(z) = 1 - \Phi(-z)$, $z \in \mathbb{R}$.

z	0	1	2	3	4	5	6	7	8	9
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56750	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84850	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92786	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

Définition 1

- *Une variable aléatoire réelle Z est dite* **gaussienne centrée réduite** *si elle admet pour densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ la fonction :*

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)}$$

On note $Z \rightsquigarrow \mathcal{N}(0,1)$.

- *Une variable aléatoire réelle X est dite* **gaussienne** *s'il existe $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ et $Z \rightsquigarrow \mathcal{N}(0,1)$ tels que $X = \mu + \sigma Z$. La densité de X est alors*

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

On note $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$. Quand $\sigma = 0$, on dit que X est une variable gaussienne dégénérée.

Une variable gaussienne est caractérisée par sa fonction caractéristique, donnée par la proposition suivante :

Théorème 1
La fonction caractéristique de $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ est donnée par

$$\boxed{\forall t \in \mathbb{R}, \quad \varphi_X(t) = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)}$$

Définition 2
Un vecteur aléatoire X à valeurs dans $\mathbb{R}^d$ est dit **gaussien** *si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne.*
Si $X = {}^t(X_1, \ldots, X_d)$ est un vecteur gaussien, on définit son **vecteur moyenne** *$\mathbb{E}(X)$ par*

$$\mathbb{E}(X) = {}^t\left(\mathbb{E}(X_1), \ldots, \mathbb{E}(X_d)\right)$$

et sa **matrice de variance-covariance** *$Var(X)$ par*

$$Var(X) = \mathbb{E}\left((X - \mathbb{E}(X)) \times {}^t(X - \mathbb{E}(X))\right)$$

Théorème 2
Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. On a que X admet pour fonction caractéristique la fonction

$$\boxed{\forall u \in \mathbb{R}^d, \quad \varphi_X(u) = \mathbb{E}\left[\exp(i\,{}^t uX)\right] = \exp(i\,{}^t um - {}^t u\Sigma u)}$$

La loi de X est donc entièrement déterminée par m et Σ. On note $X \rightsquigarrow \mathcal{N}(m, \Sigma)$.

Corollaire 1 (Propriété de linéarité)
Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. On a pour toute matrice A possédant d colonnes et pour tout vecteur $b \in \mathbb{R}^d$,

$$AX + b \rightsquigarrow \mathcal{N}(Am + b\ ,\ A\Sigma\,{}^t A)$$

Corollaire 2 (Propriété pour l'indépendance)
Soit $X = {}^t(X_1, \ldots, X_d)$ un vecteur gaussien. Pour tout $(i, j) \in \{1, \ldots, d\}^2$ tel que $i \neq j$, X_i et X_j sont indépendantes si et seulement si $cov(X_i, X_j) = 0$.

Théorème 5
Soit $X = {}^t(X_1, \dots, X_d)$ un vecteur gaussien. On note $m = \mathbb{E}(X)$ et $\Sigma = Var(X)$. X admet une densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$ si et seulement si $\det(\Sigma) \neq 0$.

- *Si* $\det(\Sigma) = 0$, *la loi de $X - m$ est presque sûrement portée par un espace vectoriel engendré par les vecteurs propres associés aux valeurs propres non nulles de Σ.*
- *Si* $\det(\Sigma) \neq 0$,

$$\boxed{\forall x \in \mathbb{R}^d, \quad f(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^d \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{{}^t(x-m)\Sigma^{-1}(x-m)}{2}\right)}$$

Fonction caractéristique

La fonction caractéristique d'une variable aléatoire à densité est la transformée de Fourier de sa densité. C'est un outil pratique pour caractériser les lois.

(Fonction caractéristique)
Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. La fonction caractéristique de X est la quantité

$$\varphi_X : \mathbb{R} \to \mathbb{C} \tag{4.1}$$
$$t \mapsto \varphi_X(t) = \mathrm{E}\left[e^{itX}\right] \tag{4.2}$$

- Si X est une variable aléatoire discrète, on a
$$\varphi_X(t) = \sum_{k \in \mathbb{N}} P(X = k) e^{itk}$$
- Si X admet une densité f_X, alors pour tout $t \in \mathbb{R}$
$$\varphi_X(t) = \int_{\mathbb{R}} f_X(x) e^{itx} \, \mathrm{d}x$$

La fonction caractéristique caractérise (!) parfaitement la loi d'une variable aléatoire :

Soit X et Y deux variables aléatoire sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$, de loi respective P_X et P_Y, alors

$$P_X = P_Y \Leftrightarrow \varphi_X = \varphi_Y$$

(Propriétés générales)
Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$, de loi P (discrète ou continue).

1. $\varphi_X(0) = 1$
2. Une fonction caractéristique est uniformément continue et bornée
$$\forall t \in \mathbb{R} |\varphi_x(t)| \leq 1$$
3. Une fonction caractéristique possède la symétrie hermitienne
$$\forall t \in \mathbb{R} \; \varphi_X(-t) = \overline{\varphi_X(t)}$$
4. $P_X = P_{-X} \Leftrightarrow \varphi_X$ est paire $\Leftrightarrow \varphi_X$ est réelle
5. Soit $a, b \in \mathbb{R}$, alors $\forall t \in \mathbb{R} \; \varphi_{aX+b}(t) = e^{ibt} \varphi_X(at)$

Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. Si X, de densité f_X, admet des moments jusqu'à l'ordre $p \geq 1$, alors φ_X est de classe C^p au moins et

$$\varphi_X^{(k)}(t) = \int_{\mathbb{R}} (ix)^k e^{itx} f_X(x)\,\mathrm{d}x$$

De plus, on a

$$\varphi_X^{(k)}(0) = i^k \mathrm{E}\left[X^k\right].$$

En particulier, on notera la moyenne $\mathrm{E}[X] = -i\varphi'_X(0)$ et le moment d'ordre 2 $\mathrm{E}\left[X^2\right] = \varphi''_X(0)$. On obtient de plus un développement limité de φ

$$\varphi_X(t) = 1 + it m_1 + \frac{(it)^2}{2} m_2 + \ldots + \frac{(it)^p}{p!} m_p + o(|t|^p)$$

où m_k désigne le moment d'ordre k : $m_k = \mathrm{E}\left[X^k\right]$

Une des intérêts de la fonction caractéristique, est qu'elle permet de calculer la loi d'une somme de variables aléatoires.

(Fonction caractéristique d'une somme de variable aléatoire)
Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. Soit $S_n = X_1 + X_2 + \ldots + X_n$, la variable aléatoire définie comme la somme des n variables aléatoires X_i. Alors

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \varphi_{X_k}(t)$$

Ce théorème est valable aussi bien dans le cas continue que dans le cas discret. Pour le montrer, il faut admettre le résultat suivant

(Loi d'une somme de deux variables aléatoires indépendantes)
Soit X et Y deux variables aléatoires indépendantes sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.

1. Si X et Y sont discrètes de loi $P_X = (p_k = P(X = k))_{k \in \mathbb{N}}$ et $P_Y = (q_k = P(Y = k))_{k \in \mathbb{N}}$. Alors

$$P_{X+Y}(X + Y = n) = \sum_{k \in \mathbb{N}} p_k q_{n-k}$$

ie

$$P_{X+Y} = P_X \star P_Y$$

2. Si X et Y admettent pour densité respectives f_X et f_Y, alors

$$f_{X+Y}(t) = \int_{\mathbb{R}} f_X(x) f_Y(t-x)\,\mathrm{d}x$$

ie

$$f_{X+Y} = f_X \star f_Y$$