

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):

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

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 1^{ère} urne, n_2 boules dans la 2^{ème} urne, ..., n_k boules dans la $k^{\text{ème}}$ urne:

$$C_n^{n_1} C_{n-n_1}^{n_2} \cdots C_{n-n_1-\cdots-n_{k-1}}^{n_k} = \binom{n}{n_1, \dots, n_k} = \frac{n!}{n_1!n_2! \cdots n_k!} \quad (\text{coefficient multinomial})$$

Propriétés élémentaires des coefficients binomiaux

$$C_n^{k-1} + C_n^k = C_{n+1}^k \quad (\text{triangle de Pascal})$$

$$C_{n+m}^k = \sum_{j=0}^k C_m^j C_n^{k-j} \quad (\text{formule de Vandermonde})$$

$$(x+y)^n = \sum_{k=0}^n C_n^k x^k y^{n-k} \quad (\text{formule du binôme})$$

$$(1-x)^{-n} = \sum_{k=0}^{\infty} C_{n+k-1}^k x^k \quad (\text{série géométrique généralisée})$$

2 Probabilités élémentaires

Modèle 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 $0 \leq P(A) \leq 1$, appelé probabilité de A .

Ci-dessous, soient $A, B, A_1, A_2, \dots, A_n$ 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 à deux:

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

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

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

Événements indépendants A et B sont indépendants ssi

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

Événements indépendants deux à deux

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

Événements totalement indépendants

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

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

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

Conditionnement multiple

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

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

$$P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \dots + P(A | B_n)P(B_n).$$

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

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

3 Variables aléatoires réelles

Variables discrètes Une variable aléatoire réelle X est une fonction $X : \Omega \rightarrow E$ ou E est un sous-ensemble de nombres réels. Si E est discret, 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 $F_X(t) = P(X \leq t)$

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

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

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 $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$. 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 $F_X(x) = P(X \leq x) = \lim_{y \rightarrow \infty} F_{X,Y}(x, y)$. 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 $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$.

Variabes 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

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

et dans le cas de deux variables continues

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

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

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

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

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

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

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

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

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

l'espérance mathématique de X , définie quand $E(|X|) < \infty$.

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

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

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

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

Propriétés de l'espérance

Espérance d'une constante $E[c] = c$ pour toute constante c réelle

Linéarité $E[a + bX + cY] = a + bE[X] + cE[Y]$ pour tous réels a, b, c

Positivité $E[X] \geq 0$ si $X \geq 0$

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

Variance de X

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

Covariance de X et Y

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

Coefficient de corrélation de X et Y

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

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

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

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

Variance et covariance

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

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

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

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

si elle est finie.

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

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

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

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

si elle est finie.

Cumulant d'ordre k Pour tout entier positif k , le cumulante 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, $\mathbb{E}[X] = \kappa_1$, $\text{var}[X] = \kappa_2$.

Quelques inégalités classiques

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

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

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

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

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

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

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

5 Variables aléatoires — Lois usuelles

5.1 Lois discrètes

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

5.2 Lois continues

Nom	Paramètres	Densité $f(x)$	$E[X]$	$\text{var}[X]$	$\mathcal{M}_X(t)$
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	$\lambda > 0, n \geq 1$	$\frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}$ pour $x > 0$ (0 sinon)	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t}\right)^n$ 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)	$\nu \geq 1$	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi} \Gamma(\frac{\nu}{2}) (1+x^2)^{(\nu+1)/2}}$ pour tout x	0 si $\nu \geq 2$	$\frac{1}{\nu-2}$ si $\nu \geq 3$	
Chi deux (χ_ν^2)	$\nu \geq 1$	$\frac{x^{\nu/2-1} e^{-x/2}}{2^{\nu/2} \Gamma(\nu/2)}$ pour $x > 0$ (0 sinon)	ν	2ν	$(1-2t)^{-\nu/2}$ 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

$\frac{N_1(m)}{N_1(m)+N_2(m)} \xrightarrow{m \rightarrow \infty} p$. Alors pour chaque valeur $0 \leq k \leq n$:

$$P(X_m = k) \xrightarrow{m \rightarrow \infty} C_n^k p^k (1-p)^{n-k}.$$

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

$$P(X_n = k) \xrightarrow{n \rightarrow \infty} e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

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

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

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

Théorème de Moivre-Laplace Si $n \rightarrow \infty$ 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 $n \rightarrow \infty$, alors pour tout t réel

$$P(X_n \leq np + \sqrt{np(1-p)} \cdot t) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx$$

6 Variables indépendantes et théorèmes limites

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

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

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

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

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

Un ensemble de variables X_1, \dots, X_n indépendantes et identiquement distribuées (iid) est appelé un échantillon aléatoire. On note, par exemple $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$, ou $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f$.

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

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

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

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

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 $X_1, \dots, X_n, \dots \stackrel{\text{iid}}{\sim} F$, et soient $\tau > 0$, et $(x_n)_{n \geq 1}$, une suite de nombres réels vérifiant

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

Alors

$$P(X_{(n)} \leq x_n) \xrightarrow{n \rightarrow \infty} e^{-\tau}.$$

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

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

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 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f$, ceci étant une fonction de densité. Les variables ordonnées $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ 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

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

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

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

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

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

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

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 $S_n = X_1 + \dots + X_n$ est donnée par la convolution des densités f_{X_j} : $f_{S_n} = f_{X_1} * \dots * f_{X_n}$.

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

$$a + b_1X_1 + \dots + b_nX_n \sim \mathcal{N}(a + b_1\mu_1 + \dots + b_n\mu_n, b_1^2\sigma_1^2 + \dots + b_n^2\sigma_n^2).$$

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 $S_n/n = (X_1 + \dots + X_n)/n$ converge en probabilité vers μ (**loi faible**). La moyenne converge aussi presque sûrement vers $E[X]$, i.e. la probabilité de l'événement " S_n/n converge vers $E[X]$ " 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

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

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

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

7 Éléments de Statistique

Biais et Risque d'un estimateur Si $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ est un estimateur du paramètre θ , on appelle *biais* de $\hat{\theta}$:

$$B_\theta(\hat{\theta}) = E_\theta[\hat{\theta}] - \theta$$

et *risque* (quadratique) ou *erreur moyenne quadratique* de $\hat{\theta}$ l'espérance du carré de $\hat{\theta} - \theta$:

$$R_\theta(\hat{\theta}) = E_\theta[(\hat{\theta} - \theta)^2]$$

$\hat{\theta}$ sera dit *consistant* si $\hat{\theta}$ converge en probabilité vers θ lorsque $n \rightarrow +\infty$, i.e. si

$$P(|\hat{\theta} - \theta| > \epsilon) \xrightarrow[n \rightarrow \infty]{} 0, \quad \forall \epsilon > 0$$

Intervalles de confiance (IC) et tests

(i) Cas d'un échantillon gaussien Dans ce qui suit on considère n observations x_1, \dots, x_n correspondant aux réalisations de n variables gaussiennes indépendantes $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$.

1. IC au niveau $1 - \alpha$ fixé (e.g. $1 - \alpha = 0,95 \equiv 95\%$, $\alpha = 0,99 \equiv 99\%$), pour l'espérance μ lorsque σ^2 est *connu*: soit $\bar{x} = n^{-1}(x_1 + \dots + x_n)$, alors l'IC est

$$I_\alpha = [\bar{x} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}; \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}],$$

où z_p est le p -quantile d'une variable gaussienne $Z \sim \mathcal{N}(0, 1)$, donc $P(Z \leq z_p) = p$.

2. IC au niveau $1 - \alpha$ pour l'espérance μ lorsque σ^2 est *inconnu*:

$$J_\alpha = [\bar{x} - t_{n-1}(1 - \alpha/2) \frac{s}{\sqrt{n}}; \bar{x} - t_{n-1}(\alpha/2) \frac{s}{\sqrt{n}}],$$

où $(n - 1)s^2 = \sum (x_i - \bar{x})^2$ et $t_\nu(p)$ est le p -quantile d'une variable de Student à ν degrés de liberté.

3. L'IC au niveau $1 - \alpha$ pour la variance σ^2 s'obtient à partir de la variance empirique s^2 comme:

$$[(n - 1)s^2 / \chi_{n-1}^2(1 - \alpha/2); (n - 1)s^2 / \chi_{n-1}^2(\alpha/2)]$$

où $\chi_\nu^2(p)$ est le p -quantile de la loi de khi-deux à ν degrés de liberté.

(ii) Estimateur de maximum de vraisemblance Soit y la réalisation d'une variable aléatoire Y dont la loi est $f(y; \theta)$, avec scalaire paramètre θ inconnu, alors la log vraisemblance est $\ell(\theta) = \log f(y; \theta)$ et l'information observée est

$$J(\theta) = -\frac{d^2 \ell(\theta)}{d\theta^2}.$$

Sous des conditions de régularité, un IC approximé au niveau $1 - \alpha$ pour θ est $\hat{\theta} \pm z_{1-\alpha/2} J(\hat{\theta})^{-1/2}$, où $\hat{\theta}$ est l'estimation de maximum de vraisemblance de θ , qui satisfait $\ell(\hat{\theta}) \geq \ell(\theta), \forall \theta$.

Souvent $y = (y_1, \dots, y_n)$ est supposé la réalisation d'un échantillon aléatoire Y_1, \dots, Y_n , et dans ce cas

$$\ell(\theta) = \sum_{j=1}^n \log f(y_j; \theta), \quad J(\theta) = -\sum_{j=1}^n \frac{d^2 \log f(y_j; \theta)}{d\theta^2},$$

$f(y_j; \theta)$ étant la densité de Y_j .

(iii) Tests d'hypothèses Soit x_1, \dots, x_n un échantillon correspondant aux réalisations $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} P_\theta$ (inconnue). On fait une hypothèse H_0 sur la distribution P_θ (par exemple: $\theta = \theta_0$) et l'on veut tester si les observations x_1, \dots, x_n permettent de rejeter l'hypothèse H_0 . On choisit une statistique de test T , construite de telle manière que des grandes valeurs de T suggèrent que H_0 soit fautive, et on calcule le niveau de signification

$$p_{\text{obs}} = \Pr_0(T \geq t_{\text{obs}}),$$

où t_{obs} est la valeur observée de T , c'est à dire sa valeur calculée avec les données x_1, \dots, x_n , et \Pr_0 représente une probabilité calculée sous H_0 . Plus p_{obs} est petite, plus on doute H_0 .

Une autre possibilité est de fixer un niveau α à l'avance (e.g. $\alpha = 0,05 \equiv 5\%$, $\alpha = 0,01 \equiv 1\%$ ou $\alpha = 0,001 \equiv 0,1\%$), de trouver le $(1 - \alpha)$ quantile de la loi de T sous H_0 , et de rejeter H_0 si t_{obs} excède ce quantile.

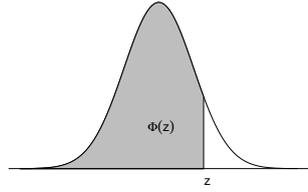
Dans le cas d'un échantillon x_1, \dots, x_n correspondant à des variables $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, avec σ^2 connue et μ inconnue, on pourra rejeter l'hypothèse $H_0 : \mu = \mu_0$ au niveau α si $\mu_0 \notin I_\alpha$, où I_α est l'intervalle de confiance de niveau $1 - \alpha$ pour la moyenne μ , autrement dit: on rejettera l'hypothèse $H_0 : \mu = \mu_0$ au niveau α si

$$\left| \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{1-\alpha/2}$$

Si σ^2 est inconnue on rejettera l'hypothèse $H_0 : \mu = \mu_0$ au niveau α si $\mu_0 \notin J_\alpha$, donc si

$$\left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| > t_{n-1}(1 - \alpha/2).$$

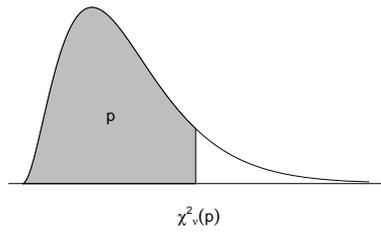
Distribution normale standard $\Phi(z)$



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

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

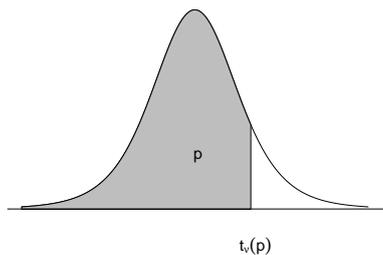
Distribution χ^2_ν



Des quantiles $\chi^2_\nu(p)$ pour la distribution en khi-carré à ν degrés de liberté.

ν	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995	.999
1	0	.0002	.010	.0039	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	.207	.297	.484	.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	73.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	86.7
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	99.6
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100.	104.	112.

Distribution t_ν de Student



Des quantiles $t_\nu(p)$ pour la distribution t de Student à ν degrés de liberté.

Pour $p < 0$ on utilise la symétrie: $t_\nu(p) = -t_\nu(1 - p)$.

ν	.75	.9	.95	.975	.99	.995
1	1.000	3.078	6.314	12.71	31.82	63.66
2	.8165	1.886	2.920	4.303	6.965	9.925
3	.7649	1.638	2.353	3.182	4.541	5.841
4	.7407	1.533	2.132	2.776	3.747	4.604
5	.7267	1.476	2.015	2.571	3.365	4.032
6	.7176	1.440	1.943	2.447	3.143	3.707
7	.7111	1.415	1.895	2.365	2.998	3.499
8	.7064	1.397	1.860	2.306	2.896	3.355
9	.7027	1.383	1.833	2.262	2.821	3.250
10	.6998	1.372	1.812	2.228	2.764	3.169
11	.6974	1.363	1.796	2.201	2.718	3.106
12	.6955	1.356	1.782	2.179	2.681	3.055
13	.6938	1.350	1.771	2.160	2.650	3.012
14	.6924	1.345	1.761	2.145	2.624	2.977
15	.6912	1.341	1.753	2.131	2.602	2.947
16	.6901	1.337	1.746	2.120	2.583	2.921
17	.6892	1.333	1.740	2.110	2.567	2.898
18	.6884	1.330	1.734	2.101	2.552	2.878
19	.6876	1.328	1.729	2.093	2.539	2.861
20	.6870	1.325	1.725	2.086	2.528	2.845
21	.6864	1.323	1.721	2.080	2.518	2.831
22	.6858	1.321	1.717	2.074	2.508	2.819
23	.6853	1.319	1.714	2.069	2.500	2.807
24	.6848	1.318	1.711	2.064	2.492	2.797
25	.6844	1.316	1.708	2.060	2.485	2.787
26	.6840	1.315	1.706	2.056	2.479	2.779
27	.6837	1.314	1.703	2.052	2.473	2.771
28	.6834	1.313	1.701	2.048	2.467	2.763
29	.6830	1.311	1.699	2.045	2.462	2.756
30	.6828	1.310	1.697	2.042	2.457	2.750
40	.6807	1.303	1.684	2.021	2.423	2.704
50	.6794	1.299	1.676	2.009	2.403	2.678
100	.6770	1.290	1.660	1.984	2.364	2.626
∞	.6745	1.282	1.645	1.960	2.326	2.576