

Variables aléatoires réelles : à retenir

Définition (probabilité)

(a) La *somme* d'une famille dénombrable $(a_i)_{i \in I}$ de réels positifs est :

$$\underbrace{\sum_{i \in I} a_i}_{\text{indépendant du choix de la numérotation}} := \lim_{n \rightarrow +\infty} \underbrace{(a_{i_0} + a_{i_1} + \dots + a_{i_n})}_{\text{croissant en } n} \in \mathbb{R}^+ \cup \{+\infty\} \quad \text{où } I = \underbrace{\{i_n\}}_{\text{distincts}}; n \in \mathbb{N}.$$

hors programme

(b) Une *probabilité* sur un ensemble Ω est une application \mathbb{P} qui à certaines parties A (il faudrait préciser) de Ω associe un réel $\mathbb{P}(A) \in [0, 1]$ de façon que :

$\mathbb{P}(\Omega) = 1$ et $\mathbb{P}(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ quand $A_n, n \in \mathbb{N}$, sont des événements deux à deux disjoints.

(c) La probabilité uniforme sur un ensemble fini non-vide Ω est la probabilité \mathbb{P} définie par :

$$\mathbb{P}(A) = \frac{\text{Card } A}{\text{Card } \Omega} \quad \text{pour } A \subseteq \Omega.$$

Dans la suite : on fixe une probabilité \mathbb{P} sur un ensemble Ω (muni d'« événements »).

Définition-Proposition (loi de probabilité d'une v. a. r. et fonction de répartition)

hors programme

(a) Une variable aléatoire réelle sur Ω est une application X de Ω dans \mathbb{R} telle que pour tous $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$ avec $a \leq b$, l'ensemble $\underbrace{\{\omega \in \Omega \mid a \leq X(\omega) \leq b\}}_{\{a \leq X \leq b\}}$ est un événement.

Sa *loi de probabilité* est la probabilité \mathbb{P}_X sur \mathbb{R} telle que :

$$\mathbb{P}_X([a, b]) = \mathbb{P}(a \leq X \leq b) \quad \text{dès que } -\infty \leq a \leq b \leq +\infty.$$

(b) On se donne une v. a. r. X sur Ω .

La *fonction de répartition* F_X de X , qui détermine la loi \mathbb{P}_X de X , est définie par :

$$F_X(x) = \mathbb{P}(X \leq x) \quad \text{pour tout } x \in \mathbb{R}.$$

Ainsi l'application $F_X : \mathbb{R} \rightarrow \mathbb{R}$ est continue à droite et croissante de $\underbrace{0}_{\text{en } -\infty}$ à $\underbrace{1}_{\text{en } +\infty}$ (*).

Définition-Proposition (loi discrète et loi à densité)

On se donne une v. a. r. X sur Ω .

(a) On suppose ici que Ω est fini ou dénombrable, de la forme :

$$\Omega = \underbrace{\{\omega_n\}}_{\text{distincts}}; n \in \mathbb{N}' \quad \text{avec } \mathbb{N}' = \underbrace{\{0, 1, \dots, N\}}_{\text{fixé dans } \mathbb{N}} \text{ ou } \mathbb{N}' = \mathbb{N}.$$

La donnée d'une probabilité Q sur $(\Omega, \mathcal{P}(\Omega))$ équivaut à la donnée d'une suite $(p_n)_{n \in \mathbb{N}'}$ de réels positifs vérifiant $\sum_{n \in \mathbb{N}'} p_n = 1$, au moyen des égalités suivantes : $p_n = Q(\{\omega_n\})$ pour tout $n \in \mathbb{N}'$.

(b) On dit que X est *discrète* s'il existe une suite de nombres réels de la forme $(x_n)_{n \in \mathbb{N}'}$ avec $\mathbb{N}' = \{0, 1, \dots, N\}$ ou $\mathbb{N}' = \mathbb{N}$, telle que $\text{Im } X \subseteq \underbrace{\{x_n\}}_{\text{fixé dans } \mathbb{N}}_{n \in \mathbb{N}'}$.

Dans ce cas \mathbb{P}_X est déterminée par la suite $(p_n)_{n \in \mathbb{N}'}$ suivante : $p_n = \mathbb{P}(X = x_n)$ pour $n \in \mathbb{N}'$.

(c) On dit que X *admet pour densité une fonction* (« mesurable ») $p : \mathbb{R} \rightarrow \mathbb{R}^+$ si :

$$\mathbb{P}(a \leq X \leq b) = \int_a^b p(x) dx \quad \text{quand } -\infty \leq a \leq b \leq +\infty.$$

Ceci équivaut à : $F_X(x) = \int_{-\infty}^x p(t) dt$ pour $x \in \mathbb{R}$. ← [Donc : $p(x_0) = F_X'(x_0)$ si p est continu en un $x_0 \in \mathbb{R}$.]

Une fonction mesurable $p : \mathbb{R} \rightarrow \mathbb{R}^+$ est associée à une telle X si et seulement si $\int_{-\infty}^{+\infty} p(x) dx = 1$.

(d) La v. a. r. X admet pour densité F_X' (définie sauf en a_1, \dots, a_p) quand il existe $a_1 < \dots < a_p$ dans \mathbb{R} tels que $F_X|_{]-\infty, a_1[}$, $F_X|_{]a_1, a_2[}$, ..., $F_X|_{]a_p, +\infty[}$ sont de classe C^1 .

(*) Réciproquement, toute $F : \mathbb{R} \rightarrow \mathbb{R}$ continue à droite et croissante de 0 à 1 est une fonction de répartition.

Définition-Proposition (v. a. r. indépendantes)

(a) On dit que des v. a. r. X_1, \dots, X_n sur Ω sont *indépendantes* si :

$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \dots \mathbb{P}(X_n \in B_n)$ pour tous évènements $B_1, \dots, B_n \subseteq \mathbb{R}$, c'est-à-dire $\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbb{P}(X_1 \leq x_1) \dots \mathbb{P}(X_n \leq x_n)$ pour tous $x_1, \dots, x_n \in \mathbb{R}$.

Dans ce cas, pour $1 \leq i_1 < \dots < i_k \leq n$ et pour toutes fonctions (mesurables) $h_1: \mathbb{R}^{i_1} \rightarrow \mathbb{R}, \dots, h_k: \mathbb{R}^{i_k - i_{k-1}} \rightarrow \mathbb{R}$, les v. a. r. $h_1(X_1, \dots, X_{i_1}), \dots, h_k(X_{i_{k-1}+1}, \dots, X_{i_k})$ sont aussi indépendantes.

(b) Quand X_1, \dots, X_n sont discrètes, l'indépendance de X_1, \dots, X_n équivaut à :

$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_n = x_n)$ pour tous $x_1 \in \text{Im}(X_1), \dots, x_n \in \text{Im}(X_n)$.

(c) Plus généralement, on dira qu'une suite $(X_n)_{n \geq 0}$ de v. a. r. sur Ω est formée de v. a. r. indépendantes si : X_1, \dots, X_n sont indépendantes pour tout $n \geq 0$.

Définition-Proposition (espérance)

On se donne deux v. a. r. X et Y sur Ω .

(a) On suppose X discrète.

On pose : $\mathbb{E}(|X|) = \sum_{n \in \mathbb{N}'} |x_n| \mathbb{P}(X = x_n) \in \mathbb{R}^+ \cup \{+\infty\}$. $\lim_{N \rightarrow +\infty} \sum_{n=0}^N x_n \mathbb{P}(X = x_n)$ quand $\mathbb{N}' = \mathbb{N}$

Lorsque $\mathbb{E}(|X|) < +\infty$, on définit *l'espérance de X* : $\mathbb{E}(X) = \sum_{n \in \mathbb{N}'} x_n \mathbb{P}(X = x_n)$.

on dit que « X a un moment d'ordre 1 »

(b) On suppose que X a une densité p .

On pose : $\mathbb{E}(|X|) = \int_{-\infty}^{+\infty} |x| p(x) dx \in \mathbb{R}^+ \cup \{+\infty\}$.

Lorsque $\mathbb{E}(|X|) < +\infty$, on définit *l'espérance de X* : $\mathbb{E}(X) = \lim_{A \rightarrow +\infty} \int_{-A}^A x p(x) dx$.

on dit que « X a un moment d'ordre 1 »

(c) On peut construire $\mathbb{E}(|X|) \in \mathbb{R}^+ \cup \{+\infty\}$, puis définir $\mathbb{E}(X) \in \mathbb{R}$ tel que $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$ lorsque $\mathbb{E}(|X|) < +\infty$, en généralisant (a) et (b) de façon que \mathbb{E} soit croissante et linéaire (unicité).
cas $X \geq 0$ cas où $\mathbb{E}(|X|) < +\infty$

Les v. a. r. X et Y ont même loi si et seulement si $\mathbb{E}(h(X)) = \mathbb{E}(h(Y))$ pour tout $h \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$.

(d) On suppose que $\mathbb{E}(|X|) < +\infty$ et $\mathbb{E}(|Y|) < +\infty$. Soient $\lambda, \mu \in \mathbb{R}$.

On a : $\mathbb{E}(1) = 1$ et $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.

Quand X et Y sont indépendantes, on a : $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Définition-Proposition (variance)

On se donne deux v. a. r. X et Y sur Ω .

(a) On suppose que $\mathbb{E}(X^2) < +\infty$. Donc $\mathbb{E}(|X|) < +\infty$.

on dit que « X a un moment d'ordre 2 »

On définit *la variance* $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ et *l'écart type* $\sigma(X) = \sqrt{\text{Var}(X)}$ de X .

$\mathbb{E}((X - \mathbb{E}(X))^2)$

(b) On suppose que $\mathbb{E}(X^2) < +\infty$ et $\mathbb{E}(Y^2) < +\infty$. Soit $\lambda \in \mathbb{R}$.

On a : $\text{Var}(\lambda X) = \lambda^2 \text{Var}(X)$ et $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))$.

Quand X et Y sont indépendantes, on a : $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

Proposition (théorème de transfert et retour sur l'indépendance)

On se donne deux v. a. r. X et Y sur Ω .

(a) On suppose que X est discrète et somme sur « $x \in \mathbb{R}$ » au sens de « $x \in \mathbb{R}, \mathbb{P}(X = x) \neq 0$ ». On a : $\mathbb{E}(h(X)) = \sum_{x \in \mathbb{R}} h(x) \mathbb{P}(X = x)$ pour $h: \mathbb{R} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ ou $h: \mathbb{R} \rightarrow \mathbb{R}$ avec $\mathbb{E}(|h(X)|) < +\infty$.

(b) Soit $p: \mathbb{R} \rightarrow \mathbb{R}^+$ mesurable telle que $\int_{-\infty}^{+\infty} p(x) dx = 1$.

On a : X a pour densité p si et seulement si $\mathbb{E}(h(X)) \stackrel{(*)}{=} \int_{-\infty}^{+\infty} h(x)p(x)dx$ pour $h \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$.

Dans ce cas on a $(*)$ pour $h: \mathbb{R} \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ mesurable ou $h: \mathbb{R} \rightarrow \mathbb{R}$ mesurable et $\mathbb{E}(|h(X)|) < +\infty$.

(c) Les v. a. r. X et Y sont indépendantes si et seulement si :

$$\mathbb{E}(h(X)k(Y)) = \mathbb{E}(h(X))\mathbb{E}(k(Y)) \text{ pour tous } h, k \in \mathcal{C}_b(\mathbb{R}, \mathbb{R}).$$

Vecteurs aléatoires : à retenir

On se donne une probabilité \mathbb{P} sur un ensemble Ω (muni d'évènements).

Définition-Proposition

On considère une v. a. $X = (X_1, \dots, X_n)$ à valeurs dans \mathbb{R}^n .

(a) La loi de X (« loi jointe de X_1, \dots, X_n ») est l'unique probabilité \mathbb{P}_X sur \mathbb{R}^n telle que :

$$\mathbb{P}_X([a_1, b_1] \times \dots \times [a_n, b_n]) = \mathbb{P}(a_1 \leq X_1 \leq b_1, \dots, a_n \leq X_n \leq b_n)$$

pour $-\infty \leq a_1 \leq b_1 \leq +\infty, \dots, -\infty \leq a_n \leq b_n \leq +\infty$.
ou « la loi de X est discrète »

(b) On dit que X est discrète s'il existe $D \subseteq \mathbb{R}^n$ finie ou dénombrable telle que $\mathbb{P}_X(D) = 1$.

Pour toute application $(x_1, \dots, x_n) \mapsto p_{x_1, \dots, x_n}$ de \mathbb{R}^n dans \mathbb{R}^+ nulle hors d'une partie finie ou dénombrable, la v. a. X est discrète avec $\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = p_{x_1, \dots, x_n}$ pour $(x_1, \dots, x_n) \in \mathbb{R}^n$ si et seulement si : $\mathbb{E}(h(X)) = \sum_{\substack{(x_1, \dots, x_n) \in \mathbb{R}^n \\ p_{x_1, \dots, x_n} \neq 0}} h(x_1, \dots, x_n) p_{x_1, \dots, x_n}$ pour toute $h: D \rightarrow \mathbb{R}$ bornée.
discrète ; tel que $p_{x_1, \dots, x_n} \neq 0$; ou pour toute $h: D \rightarrow \mathbb{R}$ positive

Une application $(x_1, \dots, x_n) \in \mathbb{R}^n \mapsto p_{x_1, \dots, x_n} \in \mathbb{R}^+$ nulle hors d'une partie finie ou dénombrable est associée à une telle X pour une probabilité \mathbb{P} si et seulement si $\sum_{(x_1, \dots, x_n) \in \mathbb{R}^n} p_{x_1, \dots, x_n} = 1$.

On appellera « fonction de masse sur \mathbb{R}^n » une telle fonction p .
ou « la loi de X admet pour densité »

(c) On dit que X admet pour densité une application $p: \mathbb{R}^n \rightarrow \mathbb{R}^+$ [mesurable] si :

$$\mathbb{P}(a_1 \leq X_1 \leq b_1, \dots, a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} p(x_1, \dots, x_n) dx_1 \dots dx_n$$

pour $-\infty \leq a_1 \leq b_1 \leq +\infty, \dots, -\infty \leq a_n \leq b_n \leq +\infty$.

Pour toute $p: \mathbb{R}^n \rightarrow \mathbb{R}^+$ [mesurable], la v. a. X admet pour densité p si et seulement si :

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} h(x_1, \dots, x_n) p(x_1, \dots, x_n) dx_1 \dots dx_n \quad \text{pour toute } h \in \mathcal{C}_b(\mathbb{R}^n, \mathbb{R}).$$

ou pour toute $h: \mathbb{R}^n \rightarrow \mathbb{R}$ [mesurable] positive

Dans ce cas, on a : $\mathbb{P}(X \in B) = \int \dots \int_B p(x_1, \dots, x_n) dx_1 \dots dx_n$ pour tout $B \subseteq \mathbb{R}^n$ (choisir $h = \mathbb{1}_B$).
événement

Donc deux telles applications [mesurables] p_0 et p_1 continues en $a \in \mathbb{R}^n$ vérifient $p_0(a) = p_1(a)$.

Une application $p: \mathbb{R}^n \rightarrow \mathbb{R}^+$ [mesurable] est associée à une telle X pour une certaine probabilité \mathbb{P} si et seulement si $\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p(x_1, \dots, x_n) dx_1 \dots dx_n = 1$.

On appellera « densité de probabilité sur \mathbb{R}^n » une telle fonction p .

Définition-Proposition

On considère une v. a. $X = (X_1, \dots, X_n)$ à valeurs dans \mathbb{R}^n .

(a) On suppose que X est discrète avec une fonction de masse $(x_1, \dots, x_n) \mapsto p_{x_1, \dots, x_n}$.

Les lois de X_1, \dots, X_n , appelées « lois marginales de X », sont discrètes avec :

$$\mathbb{P}(X_i = x) = \sum_{\substack{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathbb{R}^{n-1} \\ p_{x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n} \neq 0}} p_{x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n} \quad \text{pour } x \in \mathbb{R} \quad (1 \leq i \leq n).$$

(b) On suppose que X a une densité $p: \mathbb{R}^n \rightarrow \mathbb{R}^+$.

Les lois de X_1, \dots, X_n , appelées « lois marginales de X », ont des densités p_1, \dots, p_n définies par :

$$p_i(x) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n \quad \text{pour } x \in \mathbb{R} \quad (1 \leq i \leq n).$$

Définition

On considère deux v. a. r. $X, Y: \Omega \rightarrow \mathbb{R}$ telles que : $\mathbb{E}(X^2) < +\infty$ et $\mathbb{E}(Y^2) < +\infty$.

(a) On a : $|\mathbb{E}(|XY|)| \leq \sqrt{\mathbb{E}(X^2)} \sqrt{\mathbb{E}(Y^2)} < +\infty$ puis $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)} \sqrt{\mathbb{E}(Y^2)}$.

On définit : $\text{Cov}(X, Y) = \underbrace{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}_{\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))} \in \mathbb{R}$.
a un moment d'ordre 1

On a donc la formule suivante : $\text{Var}(X + Y) = \text{Var} X + \text{Var} Y + 2 \text{Cov}(X, Y)$.

(b) On dit que X et Y sont non-corrélées si $\text{Cov}(X, Y) = 0$ (par exemple X et Y indépendantes).

Quand $\text{Var}(X) \neq 0$ et $\text{Var}(Y) \neq 0$, le coefficient de corrélation linéaire de X et Y est défini par :

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \sigma(Y)} \in [-1, 1].$$

Théorème (« théorème de Fubini »)

(a) On considère une application $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ mesurable.

On a $\int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} |f(x, y)| dy) dx = \int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} |f(x, y)| dx) dy \leq +\infty$

et, quand le réel précédent est fini : $\int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} f(x, y) dy) dx = \int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} f(x, y) dx) dy$.

(b) On considère une suite d'applications $u_n : \mathbb{R} \rightarrow \mathbb{C}$ mesurables avec $n \geq 0$.

On a $\int_{-\infty}^{+\infty} (\sum_{n=0}^{+\infty} |u_n(x)|) dx = \sum_{n=0}^{+\infty} (\int_{-\infty}^{+\infty} |u_n(x)| dx) \leq +\infty$

et, quand le réel précédent est fini : $\int_{-\infty}^{+\infty} (\sum_{n=0}^{+\infty} u_n(x)) dx = \sum_{n=0}^{+\infty} (\int_{-\infty}^{+\infty} u_n(x) dx)$.

(c) On considère une suite double $(u_{p,q})_{p,q \geq 0}$ de nombres complexes.

On a $\sum_{p=0}^{+\infty} (\sum_{q=0}^{+\infty} |u_{p,q}|) = \sum_{q=0}^{+\infty} (\sum_{p=0}^{+\infty} |u_{p,q}|) \leq +\infty$

et, quand le réel précédent est fini : $\sum_{p=0}^{+\infty} (\sum_{q=0}^{+\infty} u_{p,q}) = \sum_{q=0}^{+\infty} (\sum_{p=0}^{+\infty} u_{p,q})$.

Théorème (« théorème de changement de variable »)

On considère une partie U de \mathbb{R}^n qui est ouverte et une application $\varphi = (\varphi_1, \dots, \varphi_n)$ de U dans \mathbb{R}^n qui a des dérivées partielles, toutes continues.

(a) Soit $x \in U$. On appelle *matrice jacobienne de φ en x* la matrice $J_\varphi(x) := \left(\frac{\partial \varphi_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$.

(b) On suppose que φ est injective et $\det J_\varphi(x) \neq 0$ pour tout $x \in U$.

Pour toute fonction continue bornée f de $\varphi(U)$ dans \mathbb{R} , on a :

$$\int_{\varphi(U)} f(y_1, \dots, y_n) dy_1 \dots dy_n = \int_U f(\varphi(x_1, \dots, x_n)) |\det J_\varphi(x_1, \dots, x_n)| dx_1 \dots dx_n$$

penser à $|dy_i|$ quand $n = 1$

« jacobien de φ en (x_1, \dots, x_n) », noté parfois $\frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)}$

Proposition

On considère une v. a. $X = (X_1, \dots, X_n)$ à valeurs dans \mathbb{R}^n .

(a) On considère des fonctions de masse p_1, \dots, p_n sur \mathbb{R} .

Les v. a. r. X_1, \dots, X_n sont discrètes de fonctions de masse p_1, \dots, p_n et sont indépendantes si et seulement si X est discrète de fonction de masse $p : (x_1, \dots, x_n) \mapsto p_1(x_1) \dots p_n(x_n)$.

(b) On considère des densités de probabilité p_1, \dots, p_n sur \mathbb{R} .

Les v. a. r. X_1, \dots, X_n admettent respectivement les densités p_1, \dots, p_n et sont indépendantes si et seulement si X a pour densité $p : (x_1, \dots, x_n) \mapsto p_1(x_1) \dots p_n(x_n)$.

Proposition

On considère des v. a. r. X et Y .

(a) On suppose que X et Y sont discrètes, et indépendantes.

La v. a. r. $X + Y$ est discrète et : $\mathbb{P}(X + Y = u) = \sum_{x \in \mathbb{R}} \underbrace{\mathbb{P}(X = x) \mathbb{P}(Y = u - x)}_{\text{lire 0 lorsque } \mathbb{P}(X = x) = 0}$ pour $u \in \mathbb{R}$.

(b) On suppose que X et Y ont des densités respectives p et q , et sont indépendantes.

La v. a. r. $X + Y$ a pour densité r avec : $r(u) = \int_{-\infty}^{+\infty} p(x) q(u - x) dx \leq +\infty$ pour tout $u \in \mathbb{R}$.

Définition-Proposition

(a) Soient X et Y deux v. a. r. discrètes, avec $Y \geq 0$ ou $\mathbb{E}(|Y|) < +\infty$.

Lorsque $x \in \mathbb{R}$ et $\mathbb{P}(X = x) \neq 0$, on introduit la « probabilité de $\{Y = y\}$ sachant $\{X = x\}$ » :

$$p_{Y|X=x}(y) = \frac{\mathbb{P}(\{Y=y\} \cap \{X=x\})}{\mathbb{P}(X=x)} \quad \text{pour } y \in \mathbb{R}.$$

Donc : $p_{Y|X=x}$ est une fonction de masse.

On a : $\mathbb{E}(Y) = \sum_{x \in \mathbb{R}} \underbrace{\mathbb{E}(Y|X = x) \mathbb{P}(X = x)}_{\text{lire 0 lorsque } \mathbb{P}(X = x) = 0}$ où $\mathbb{E}(Y|X = x) := \sum_{y \in \mathbb{R}} y p_{Y|X=x}(y)$.

(b) Soit (X, Y) une v. a. à valeurs dans \mathbb{R}^2 qui a une densité $p_{X,Y}$, avec $Y \geq 0$ ou $\mathbb{E}(|Y|) < +\infty$.

Lorsque $x \in \mathbb{R}$ et $p_X(x) \neq 0$, où $p_X(x) := \int_{-\infty}^{+\infty} p_{X,Y}(x, y) dy$, on pose :

$$p_{Y|X=x}(y) = \frac{p_{X,Y}(x, y)}{p_X(x)} \quad \text{pour } y \in \mathbb{R}$$

Donc : $p_{Y|X=x}$ est une densité de probabilité, appelée « densité conditionnelle de Y sachant $X = x$ ».

de la loi de probabilité sur \mathbb{R} qui envoie un événement A sur $\lim_{h \rightarrow 0^+} \mathbb{P}(Y \in A | x \leq X \leq x + h)$

On a : $\mathbb{E}(Y) = \int_{-\infty}^{+\infty} \underbrace{\mathbb{E}(Y|X = x) p_X(x)}_{\text{lire 0 lorsque } p_X(x) = 0} dx$ où $\mathbb{E}(Y|X = x) := \int_{-\infty}^{+\infty} y p_{Y|X=x}(y) dy$.

Fonctions caractéristiques : à retenir

On se donne une probabilité \mathbb{P} sur un ensemble Ω (muni d'évènements) et $n \in \mathbb{N}$.

Définition-Proposition

(a) Soit X une v. a. r. sur Ω .

La loi de X est déterminée par sa *fonction caractéristique* $\Phi_X: \mathbb{R} \rightarrow \mathbb{C}$ définie par :

$$\Phi_X(t) = \mathbb{E}(e^{itX}) \text{ pour } t \in \mathbb{R}.$$

(b) Soit $X = (X_1, \dots, X_n)$ une v. a. sur Ω à valeurs dans \mathbb{R}^n .

La loi de X est déterminée par sa *fonction caractéristique* $\Phi_X: \mathbb{R}^n \rightarrow \mathbb{C}$ définie par :

$$\Phi_X(t) = \mathbb{E}(e^{i(t_1 X_1 + \dots + t_n X_n)}) \text{ pour } t = (t_1, \dots, t_n) \in \mathbb{R}^n.$$

De plus les v. a. r. X_1, \dots, X_n sont indépendantes si et seulement si

$$\Phi_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \Phi_{X_1}(t_1) \cdots \Phi_{X_n}(t_n) \text{ pour tout } (t_1, \dots, t_n) \in \mathbb{R}^n. (*)$$

Exemples

Lorsque $X \sim \mathcal{N}(m, \sigma^2)$ avec $m \in \mathbb{R}$ et $\sigma > 0$, on a :

$$\Phi_X(t) = e^{imt} e^{-\frac{\sigma^2}{2}t^2} \text{ pour tout } t \in \mathbb{R}.$$

D'autres cas ($t \in \mathbb{R}$) :

- Si $X \sim \mathcal{B}(p)$ avec $p \in [0, 1]$, alors $\Phi_X(t) = 1 - p + pe^{it}$.
- Si $X \sim \mathcal{B}(n, p)$ avec $n \in \mathbb{N} \setminus \{0\}$ et $p \in [0, 1]$, alors $\Phi_X(t) = (1 - p + pe^{it})^n$.
- Si $X \sim \mathcal{P}(\lambda)$ avec $\lambda > 0$, alors $\Phi_X(t) = \exp(\lambda(e^{it} - 1))$.
- Si $X \sim \mathcal{G}(p)$ (resp. $\mathcal{G}(p)_{\mathbb{N} \setminus \{0\}}$) avec $p \in]0, 1[$, alors $\Phi_X(t) = \frac{p}{1 - (1-p)e^{it}}$ (resp. $\frac{pe^{it}}{1 - (1-p)e^{it}}$).
- Si $X \sim \mathcal{U}_{[-a, a]}$ avec $a > 0$, alors $\Phi_X(t) = \frac{\sin(at)}{at}$.
- Si $X \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$, alors $\Phi_X(t) = \frac{\lambda}{\lambda - it}$.
- Si $X \sim \Gamma(r, \lambda)$ avec $r > 0$ et $\lambda > 0$, alors $\Phi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^r. (**)$

Proposition

(a) Soit X une v. a. r. sur Ω .

La fonction Φ_X est continue et bornée avec $\Phi_X(0) = 1$ et $|\Phi_X| \leq 1$.

On a : $\left\{ \begin{array}{l} \text{si } \mathbb{E}(|X|) < +\infty, \text{ alors } \Phi_X \text{ est } C^1 \text{ et } \Phi'_X(0) = i\mathbb{E}(X); \\ \text{si } \mathbb{E}(X^2) < +\infty, \text{ alors } \Phi_X \text{ est } C^2 \text{ et } \Phi''_X(0) = -\mathbb{E}(X^2) \end{array} \right.$

et plus généralement, si $n \in \mathbb{N}$ et $\mathbb{E}(|X|^n) < +\infty$, alors Φ_X est C^n et $\Phi_X^{(n)}(t) = \mathbb{E}(i^n X^n e^{itX})$.

(Noter que la connaissance du développement limité de Φ_X en 0 permet ainsi de calculer $\mathbb{E}(X^n)$.)

(b) Soient X et Y deux v. a. sur Ω à valeurs dans \mathbb{R}^n .

Quand X et Y sont indépendantes, on a : $\Phi_{X+Y}(t) = \Phi_X(t) \Phi_Y(t)$ pour tout $t \in \mathbb{R}^n$.

Proposition (« formule d'inversion de Levy »)

Soit X une v. a. r. sur Ω .

Si $\int_{\mathbb{R}} |\Phi_X(t)| dt < +\infty$, alors X admet pour densité la fonction p_X définie par :

$$p_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \Phi_X(t) dt \text{ pour } x \in \mathbb{R}.$$

(*) Les fonctions $\Phi_{X_1}, \dots, \Phi_{X_n}$ se déduisent de Φ_X par : $\Phi_{X_k}(t_k) = \Phi_X(0, \dots, 0, \underbrace{t_k}_{\text{au } k^{\text{e}} \text{ emplacement}}, 0, \dots, 0)$.

(**) Soit $z \in \mathbb{C} \setminus \mathbb{R}^-$. On pose : $\text{Log}_{\mathbb{P}}(z) := \ln r + i\theta$ où z se décompose en $z = re^{i\theta}$ avec $r > 0$ et $\theta \in]-\pi, \pi[$. Cela permet de définir ici $z^\alpha := \exp(\alpha \text{Log}_{\mathbb{P}}(z))$ pour tout $\alpha \in \mathbb{R}$, en particulier pour $\alpha = r$.

Théorèmes limites : à retenir

On se donne une probabilité \mathbb{P} sur un ensemble Ω (muni d'évènements).

Proposition

Soit X une v. a. r. sur Ω vérifiant $\mathbb{E}(X^2) < +\infty$.

On a : $\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \text{Var}(X)$ pour tout $\varepsilon > 0$ « inégalité de Bienaymé-Tchebychev ».
 [Cela découle de l'« inégalité de Markov » : $\mathbb{P}(Z \geq a) \leq \underbrace{\frac{\mathbb{E}(Z)}{a}}_{\leq +\infty}$ pour Z v. a. r. positive et $a > 0$.]

Définition-Proposition

Soient $(X_n)_{n \geq 1}$ une suite de v. a. r. sur Ω , et, X et Y deux v. a. r. sur Ω .

(a) On dit que la suite $(X_n)_{n \geq 1}$ converge en probabilité vers X , et note « $X_n \xrightarrow[n \rightarrow +\infty]{P} X$ », si :

$$\mathbb{P}(|X_n - X| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0 \text{ pour tout } \varepsilon > 0.$$

De plus, si $X_n \xrightarrow[n \rightarrow +\infty]{P} X$ et $X_n \xrightarrow[n \rightarrow +\infty]{P} Y$ alors $\mathbb{P}(X = Y) = 1$ (unicité à l'égalité presque sûre près).

(b) On dit que $(X_n)_{n \geq 1}$ converge en loi vers X , et note « $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ », si elle vérifie l'une des trois propriétés équivalentes :

- (i) $\mathbb{E}(h(X_n)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(h(X))$ pour tout $h \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$;
- (ii) $F_{X_n}(x) \xrightarrow[n \rightarrow +\infty]{} F_X(x)$ pour tout $x \in \mathbb{R}$ en lequel F_X est continue ;
- (iii) $\Phi_{X_n}(t) \xrightarrow[n \rightarrow +\infty]{} \Phi_X(t)$ pour tout $t \in \mathbb{R}$.

D'après (i) : si $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ alors on a $f(X_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} f(X)$ pour tout $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$.

(c) Si $X_n \xrightarrow[n \rightarrow +\infty]{P} X$, alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$.

Proposition

Soit $(X_n)_{n \geq 1}$ une suite de v. a. r. indépendantes de même loi (« i. i. d. » pour « indépendantes et identiquement distribuées ») sur Ω , telle que $\mathbb{E}(X_1^2) < +\infty$.

en fait la condition « $\mathbb{E}(|X_1|) < +\infty$ » suffit

On a : $\overline{X}_n := \frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{P} \mathbb{E}(X_1)$ « loi faible des grands nombres ».

[Provient de la « convergence en moyenne quadratique » : $\mathbb{E}((\overline{X}_n - \mathbb{E}(X_1))^2) \xrightarrow[n \rightarrow +\infty]{} 0$.]
noté « $\overline{X}_n \xrightarrow[n \rightarrow +\infty]{L^2} \mathbb{E}(X_1)$ »

Compléments

Soient $(X_n)_{n \geq 1}$ une suite de v. a. r. sur Ω et X une v. a. r. sur Ω .

(a) On dit que la suite $(X_n)_{n \geq 1}$ converge presque sûrement vers X , et note « $X_n \xrightarrow[n \rightarrow +\infty]{p. s.} X$ », si :

$$\mathbb{P}(\{\omega \in \Omega \mid X_n(\omega) \xrightarrow[n \rightarrow +\infty]{} X(\omega)\}) = 0.$$

(b) Si $X_n \xrightarrow[n \rightarrow +\infty]{p. s.} X$, alors $X_n \xrightarrow[n \rightarrow +\infty]{P} X$.

(c) Soit $(X_n)_{n \geq 1}$ une suite de v. a. r. i. i. d. sur Ω telle que $\mathbb{E}(|X_1|) < +\infty$.

On a : $\overline{X}_n = \frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p. s.} \mathbb{E}(X_1)$ « loi forte des grands nombres ».

(d) Soit $(X_n)_{n \geq 1}$ une suite de v. a. r. i. i. d. sur Ω telle que $\mathbb{E}(X_1^2) < +\infty$.

On a : $\widehat{\sigma}_n^2 := \frac{(X_1 - \overline{X}_n)^2 + \dots + (X_n - \overline{X}_n)^2}{n} \xrightarrow[n \rightarrow +\infty]{p. s.} \text{Var}(X_1)$.

hors programme

Proposition (lemme de Scheffé)

On considère une suite $(X_n)_{n \geq 1}$ de v. a. r. sur Ω et une v. a. r. X sur Ω .

(a) On suppose que les X_n ($n \geq 1$) et X sont à valeurs dans \mathbb{Z} .

On a : $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \forall k \in \mathbb{Z} \quad \mathbb{P}(X_n = k) \xrightarrow[n \rightarrow +\infty]{} \mathbb{P}(X = k)$.

(b) On suppose que les X_n ($n \geq 1$) et X ont des densités $p_n: \mathbb{R}^n \rightarrow \mathbb{R}^+$ et $p: \mathbb{R}^n \rightarrow \mathbb{R}^+$.

Si $p_n(x) \xrightarrow[n \rightarrow +\infty]{} p_X(x)$ pour presque tout $x \in \mathbb{R}$, alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$.

Théorème

On note : $\Pi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$.

(a) Soit $(X_n)_{n \geq 1}$ une suite de v. a. r. i. i. d. sur Ω telle que $\mathbb{E}(X_1^2) < +\infty$ et $\sigma(X_1) \neq 0$.

On pose : $S_n = X_1 + \dots + X_n$ quand $n \geq 1$. On a :

$\underbrace{\frac{S_n - \mathbb{E}(S_n)}{\sigma(S_n)}}_{\text{somme réduite}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \sim \mathcal{N}(0, 1)$ « théorème central limite ».

Cela signifie que : $\mathbb{P}\left(a < \underbrace{\frac{S_n - \mathbb{E}(S_n)}{\sigma(S_n)}}_{\frac{\bar{X}_n - \mathbb{E}(X_1)}{\sigma(X_1)/\sqrt{n}}} < b\right) \xrightarrow[n \rightarrow +\infty]{} \underbrace{\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}_{\Pi(b) - \Pi(a)}$ quand $-\infty \leq a < b \leq +\infty$ (*).

(c) Soit $(X_n)_{n \geq 1}$ une suite de v. a. r. i. i. d. avec X_1 de loi de Bernoulli de paramètre $p \in]0, 1[$.

On pose : $S_n = X_1 + \dots + X_n$ quand $n \geq 1$. On a d'après le théorème central limite :

$\mathbb{P}\left(a < \frac{S_n - np}{\sqrt{np(1-p)}} < b\right) \xrightarrow[n \rightarrow +\infty]{} \Pi(b) - \Pi(a)$ quand $-\infty \leq a < b \leq +\infty$ « théorème de Moivre-Laplace ».

Remarque

Le théorème central limite est valable avec « \leq » au lieu de « $<$ » car $S'_n := \frac{S_n - \mathbb{E}(S_n)}{\sigma(S_n)}$ vérifie :

$\underbrace{\mathbb{P}\left(a < S'_n < b\right)}_{\geq \Pi(b) - \Pi(a) - \varepsilon \text{ si } n \geq N} \leq \mathbb{P}\left(a \leq S'_n \leq b\right) \leq \mathbb{P}\left(a - \frac{1}{k} < S'_n < b + \frac{1}{k}\right) \leq \underbrace{\Pi\left(b + \frac{1}{k}\right) - \Pi\left(a - \frac{1}{k}\right) + \frac{\varepsilon}{2}}_{n \geq N_k \leq \Pi(b) - \Pi(a) + \varepsilon \text{ si } k = K \text{ et } n \geq N_k}$.

Une construction admise (suite de v. a. r. indépendantes avec des lois imposées)

Soient U_n des v. a. r. sur des ensembles Ω_n munis chacun (d'évènements et) d'une probabilité \mathbb{P}_n sur Ω_n , où $n \geq 1$. Il existe une probabilité $\tilde{\mathbb{P}}$ sur l'ensemble $\Omega_{\mathbb{N}} = \prod_{n \geq 1} \Omega_n$ (muni de certains évènements) et une suite $(X_n)_{n \geq 1}$ de v. a. r. indépendantes sur $\Omega_{\mathbb{N}}$ telle que pour tout $n \geq 1$ la loi de X_n pour $\tilde{\mathbb{P}}$ est égale à la loi de U_n pour \mathbb{P}_n .

Exemple : la probabilité $\tilde{\mathbb{P}}$ du jeu de pile ou face infini s'obtient avec $\mathbb{P}_n = \mathcal{B}(p)$ et $U_n = \text{id}_{\mathbb{R}}$.

(*) Pour « n grand », on approxime en pratique $\mathbb{P}\left(\left|\frac{\bar{X}_n - \mathbb{E}(X_1)}{\sigma(X_1)/\sqrt{n}}\right| < a\right)$ par sa limite $2\Pi(a) - 1$ quand $n \rightarrow +\infty$, et cet abus fournit un intervalle de la forme $]\bar{X}_n - \varepsilon_n, \bar{X}_n + \varepsilon_n[$ dans lequel $\mathbb{E}(X_1)$ se trouve avec une probabilité $\geq 1 - \alpha$ où α est déduit de la table des valeurs prises par Π .

Vecteurs gaussiens : à retenir

On se donne une probabilité \mathbb{P} sur un ensemble Ω (muni d'évènements) et $n \in \mathbb{N} \setminus \{0\}$.

On note : $\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$ pour tous $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$.

Définition

Soient $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathfrak{M}(n, \mathbb{R})$ une matrice carrée d'ordre n et Z une v. a. r. sur Ω .

(a) On dit que A est *symétrique* si : $a_{j,i} = a_{i,j}$ pour tous $i, j \in \{1, \dots, n\}$.

Dans ce cas, on dit que A est *positive* si : $\langle Ax, x \rangle \geq 0$ pour tout $x \in \mathbb{R}^n$.

(b) On dit que Z est gaussienne s'il existe $m \in \mathbb{R}$ et $\sigma \geq 0$ tels que $\begin{cases} \sigma = 0 \text{ et } \mathbb{P}(Z = m) = 1 \\ \text{ou} \\ \sigma > 0 \text{ et } Z \sim \mathcal{N}(m, \sigma^2) \end{cases}$.

Définition-Proposition

Soient $X = (X_1, \dots, X_n)$ et $Y = (Y_1, \dots, Y_n)$ deux v. a. sur Ω à valeurs dans \mathbb{R}^n .

(a) Pour tous $\mu \in \mathbb{R}^n$ et $\Gamma \in \mathfrak{M}(n, \mathbb{R})$ symétrique positive, il existe une unique loi de probabilité $\mathcal{N}_n(\mu, \Gamma)$ sur \mathbb{R}^n dont la fonction caractéristique Φ vérifie : $\Phi(t) = e^{i\langle \mu, t \rangle} e^{-\frac{1}{2}\langle \Gamma t, t \rangle}$ pour $t \in \mathbb{R}^n$.

Ainsi : $\begin{cases} \text{si } X \sim \mathcal{N}_n(\mu, \Gamma) \text{ et } Y \sim \mathcal{N}_n(\nu, \Delta) \text{ sont indépendantes, alors } X+Y \sim \mathcal{N}_n(\mu+\nu, \Gamma+\Delta); \\ \text{si } X \sim \mathcal{N}_n(\mu, \Gamma) \text{ et } A \in \mathfrak{M}_{k,n}(\mathbb{R}) \text{ et } B \in \mathbb{R}^k, \text{ alors } AX+B \sim \mathcal{N}_k(A\mu+B, A\Gamma A). \end{cases}$

(b) On dit que la v. a. X est *gaussienne* s'il existe un vecteur $\mu \in \mathbb{R}^n$ et une matrice $\Gamma \in \mathfrak{M}(n, \mathbb{R})$ symétrique positive tels que $X \sim \mathcal{N}_n(\mu, \Gamma)$.

Dans ce cas X_1, \dots, X_n ont des moments d'ordre 2 avec :

$$\mu = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n)) \text{ et } \Gamma = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n}.$$

(c) La v. a. X est gaussienne si et seulement si $\sum_{i=1}^n \alpha_i X_i$ est gaussienne pour tous $\alpha_1, \dots, \alpha_n \in \mathbb{R}$.

Définition-Proposition

Soit X une v. a. à valeurs dans \mathbb{R}^n gaussienne d'espérance μ et de matrice de covariance Γ .

(a) On a : $\mathbb{P}(X \in \Gamma(\mathbb{R}^n) + \mu) = 1$.

En particulier, X n'a pas de densité lorsque $\det(\Gamma) = 0$.

(b) On suppose que $\det(\Gamma) \neq 0$.

La loi de X a pour densité la fonction p_X définie par :

$$p_X(x) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\Gamma)}} e^{-\frac{1}{2}\langle \Gamma^{-1}(x-\mu), x-\mu \rangle} \text{ pour } x \in \mathbb{R}^n.$$

Proposition

Soit $X = (X_1, \dots, X_r, X_{r+1}, \dots, X_n)$ une v. a. à valeurs dans \mathbb{R}^n , avec $1 \leq r < n$.

(a) Si X est gaussienne, alors (X_1, \dots, X_r) et (X_{r+1}, \dots, X_n) sont gaussiennes.

La réciproque est fautive.

(b) Si $(X_1, \dots, X_r) \sim \mathcal{N}_r(\mu', \Gamma')$ et $(X_{r+1}, \dots, X_n) \sim \mathcal{N}_{n-r}(\mu'', \Gamma'')$ sont indépendantes, alors $X \sim \mathcal{N}_n(\mu, \Gamma)$ avec $\mu = (\mu', \mu'')$ et $\Gamma = \begin{pmatrix} \Gamma' & 0 \\ 0 & \Gamma'' \end{pmatrix}$.

(c) Si X est gaussienne avec $\text{Cov}(X_i, X_j) = 0$ quand $1 \leq i \leq r$ et $r+1 \leq j \leq n$, alors (X_1, \dots, X_r) et (X_{r+1}, \dots, X_n) sont indépendantes.

Compléments de calcul intégral

1. On peut associer « canoniquement » un élément $\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \cdots dx_n$ de $\mathbb{R}^+ \cup \{+\infty\}$ (déterminé dans les cas usuels grâce au 3. ci-dessous) à chaque v. a. $f : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, $n \in \mathbb{N}$.

Soit $A \subseteq \mathbb{R}^n$ un évènement. Pour toute v. a. $f : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{+\infty\}$, on note :

(ou A , en prolongeant par 0 hors de A)

$$\int_A f(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_{\mathbb{R}^n} (f \times \mathbb{1}_A)(x_1, \dots, x_n) dx_1 \cdots dx_n \in \mathbb{R}^+ \cup \{+\infty\}.$$

Cas particulier : $\text{mes}(A) := \int_A dx_1 \cdots dx_n$ « mesure de A ».

Par exemple : $\text{mes}([a_1, b_1] \times \cdots \times [a_n, b_n]) = (b_1 - a_1) \times \cdots \times (b_n - a_n)$ pour $a_1 < b_1, \dots, a_n < b_n$.

2. déf : une v. a. r. $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est intégrable si et seulement si $\int_{\mathbb{R}^n} |f(x_1, \dots, x_n)| dx_1 \cdots dx_n < +\infty$.

Dans ce cas, on note :

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_{\{f \geq 0\}} f(x_1, \dots, x_n) dx_1 \cdots dx_n - \int_{\{f < 0\}} (-f(x_1, \dots, x_n)) dx_1 \cdots dx_n \in \mathbb{R}.$$

L'intégrale des v. a. r. intégrables est une application linéaire.

Soit $A \subseteq \mathbb{R}^n$ un évènement. Pour toute v. a. r. $f : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que $f \times \mathbb{1}_A$ est intégrable, on note :

(ou A , en prolongeant par 0 hors de A)

$$\int_A f(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_{\mathbb{R}^n} (f \times \mathbb{1}_A)(x_1, \dots, x_n) dx_1 \cdots dx_n \in \mathbb{R}.$$

(L'intégrale sur un évènement $A' \subseteq A$ avec $\text{mes}(A \setminus A') = 0$ coïncide avec l'intégrale sur A .)

3. Si $a < b$ dans \mathbb{R} et $f : [a, b] \rightarrow \mathbb{R}$ est continue, on a : $\int_{[a,b]} f(x) dx = \int_a^b f(x) dx$.

Si $-\infty < a < b \leq +\infty$ et $f : [a, b[\rightarrow \mathbb{R}^+$ (resp. \mathbb{R}) est continue (resp. continue et intégrable),

on a : $\int_{[a,b[} f(x) dx = \lim_{b' \rightarrow b^-} \int_a^{b'} f(x) dx$.

(Le cas $-\infty \leq a < b < +\infty$ est analogue. Pour le cas $a = -\infty$ et $b = +\infty$: $\int_{-\infty}^{+\infty} = \int_{-\infty}^0 + \int_0^{+\infty}$.)

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ (resp. \mathbb{R}) est une v. a. (resp. une v. a. r. intégrable) et $i \in \{1, \dots, n\}$, on a :

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \cdots dx_n = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f(x_1, \dots, x_n) dx_i \right) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \quad \text{« th. de Fubini »}$$

4. On se donne :

(i) une partie U de \mathbb{R}^n qui est réunion de parties de la forme

$$\{x \in \mathbb{R}^n \mid f_1(x) > a_1, \dots, f_p(x) > a_p \text{ et } g_1(x) < b_1, \dots, g_q(x) < b_q\}$$

avec $f_1, \dots, f_p, g_1, \dots, g_q : \mathbb{R}^n \rightarrow \mathbb{R}$ continues et $a_1, \dots, a_p, b_1, \dots, b_q \in \mathbb{R}$;

(ii) une injection $\varphi = (\varphi_1, \dots, \varphi_n) : U \rightarrow \mathbb{R}^n$ telle que

$$\frac{\partial \varphi}{\partial x_1}, \dots, \frac{\partial \varphi}{\partial x_n} \text{ existent et sont continues, avec } \frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)} := \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial x_1} & \cdots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix} \text{ qui ne s'annule pas.}$$

(Sous ces conditions, U et $\varphi(U)$ sont des évènements.)

Si $f : \varphi(U) \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ (resp. \mathbb{R}) est une v. a. (resp. une v. a. r. intégrable), on a :

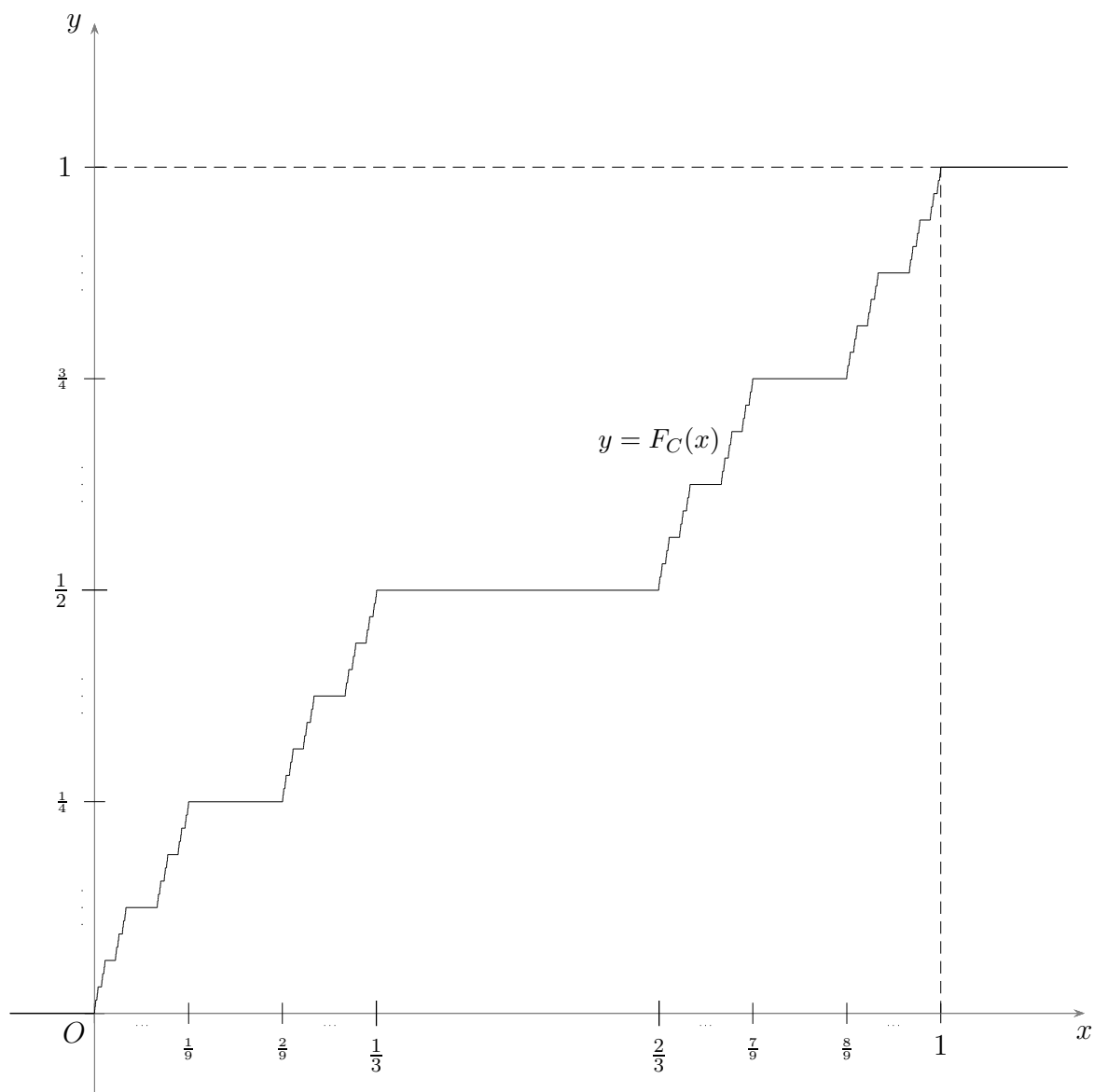
$$\int_{\varphi(U)} f(y_1, \dots, y_n) dy_1 \cdots dy_n = \int_U f(\varphi(x_1, \dots, x_n)) \left| \frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)} \right| dx_1 \cdots dx_n$$

« formule de changement de variable dans les intégrales multiples ».

Par exemple, dans le cas du « passage en coordonnées polaires », le changement de variable

« $x = r \cos \theta$ et $y = r \sin \theta$ » avec $(r, \theta) \in \underbrace{]0, +\infty[\times]-\pi, \pi[}_U$ donne « $dx dy = \underbrace{r}_{\left| \frac{D(x,y)}{D(r,\theta)} \right|} dr d\theta$ ».

L'escalier de Cantor (découvert en 1885 par Ludwig Scheeffer qui était un élève de Cantor) est le graphe suivant d'une certaine application continue $F_C : \mathbb{R} \rightarrow \mathbb{R}$ croissante de 0 à 1 :



Comme F_C est continue à droite et croissante de 0 à 1, il existe une variable aléatoire réelle C qui a pour fonction de répartition F_C .

L'application F_C est dérivable « presque partout »^(*) avec une dérivée nulle. En un certain sens, on a : $\int_{-\infty}^{+\infty} F'_C(t) dt = 0 < 1$, le problème venant du fait que F'_C n'est pas définie partout. [Toute application dérivable $F: [a, b] \rightarrow \mathbb{R}$ et de dérivée bornée vérifie $\int_a^b F'(t) dt = F(b) - F(a)$.]

Un résultat délicat permet d'en déduire que C n'a pas de densité.

[Si $p: \mathbb{R} \rightarrow \mathbb{R}$ est « intégrable au sens de Lebesgue », alors $p(x) = \frac{d}{dx}(\int_{-\infty}^x p(t) dt)$ presque partout.] On dispose d'ailleurs d'un **critère étonnant** : une variable aléatoire réelle X possède une densité si et seulement si $\int_{-\infty}^{+\infty} \underbrace{F'_X(t)}_{\text{défini presque partout et positif}} dt = 1$; dans ce cas, F'_X est une densité de X .

défini presque partout et positif

(*) On dit qu'une propriété portant sur $x \in \mathbb{R}$ est vérifiée « presque partout » si l'ensemble des points où elle n'est pas vérifiée est pour tout $\varepsilon > 0$ inclus dans une réunion d'intervalles ouverts deux à deux disjoints dont la somme des longueurs est strictement inférieure à ε .

Rappel des lois usuelles

Loi de Bernoulli $\mathcal{B}(p)$ de paramètre $p \in [0, 1]$:

$$\mathbb{P}(X = 0) = 1 - p \text{ et } \mathbb{P}(X = 1) = p.$$

Loi binomiale $\mathcal{B}(n, p)$ de paramètres $n \in \mathbb{N} \setminus \{0\}$ et $p \in [0, 1]$:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \text{ pour } k \in \mathbb{N}, \text{ où } \binom{n}{k} \stackrel{\text{convention}}{=} 0 \text{ quand } k > n.$$

Loi géométrique $\mathcal{G}(p)$ de paramètre $p \in]0, 1[$:

$$\mathbb{P}(X = k) = p(1 - p)^k \text{ pour } k \in \mathbb{N}. \text{ On note ensuite } \mathcal{G}_{\mathbb{N} \setminus \{0\}}(p) \text{ la loi de } X + 1.$$

moins utile $\left\{ \begin{array}{l} \text{Loi binomiale négative } \mathcal{B}^-(n, p) \text{ de paramètres } n \in \mathbb{N} \setminus \{0\} \text{ et } p \in]0, 1[: \\ \mathbb{P}(X = k) = \binom{n+k-1}{n-1} p^n (1 - p)^k \text{ pour } k \in \mathbb{N}. \end{array} \right.$

Loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$:

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \text{ pour } k \in \mathbb{N}.$$

Loi uniforme $\mathcal{U}_{[a,b]}$ sur un segment $[a, b]$, avec $a < b$:

$$\mathbb{P}_X = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x) dx.$$

Loi normale $\mathcal{N}(m, \sigma^2)$ d'espérance $m \in \mathbb{R}$ et d'écart type $\sigma > 0$:

$$\mathbb{P}_X = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$:

$$\mathbb{P}_X = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x) dx.$$

Loi Gamma $\Gamma(r, \lambda)$ de paramètres $r > 0$ et $\lambda > 0$:

$$\mathbb{P}_X = \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} \mathbb{1}_{\mathbb{R}^+}(x) dx, \text{ où } \Gamma(r) := \int_0^{+\infty} e^{-x} x^{r-1} dx.$$

Loi de Cauchy $\mathcal{C}(\lambda)$ de paramètre $\lambda > 0$:

$$\mathbb{P}_X = \frac{\lambda}{\pi(\lambda^2 + x^2)} dx.$$

Valeurs de la fonction de répartition Φ de la loi $\mathcal{N}(0, 1)$

x	0,0	0,010	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
x	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,8	4,0	4,5
$\Phi(x)$	0,99865	0,99903	0,99931	0,99952	0,99966	0,99977	0,999841	0,999928	0,999968	0,999997

On a aussi : $\Phi(-x) = 1 - \Phi(x)$ pour tout $x \in \mathbb{R}$.

Loi de Student à d degrés de liberté : $\mathbb{P}(|X| \geq ?) = \alpha$

$d \backslash \alpha$	0,50	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	1,000	3,078	6,314	12,706	31,821	63,656	127,32	318,29	636,58
2	0,816	1,886	2,920	4,303	6,965	9,925	14,089	22,328	31,600
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,214	12,924
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
100	0,677	1,290	1,660	1,984	2,364	2,626	2,871	3,174	3,390
$+\infty$	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,290

Loi du χ^2 à d degrés de liberté : $\mathbb{P}(X \geq ?) = \alpha$

$d \backslash \alpha$	0,50	0,20	0,10	0,05	0,02	0,01	0,005	0,001	0,0005
1	0,455	1,642	2,706	3,841	5,412	6,635	7,879	10,827	12,115
2	1,386	3,219	4,605	5,991	7,824	9,210	10,597	13,815	15,201
3	2,366	4,642	6,251	7,815	9,837	11,345	12,838	16,266	17,731
4	3,357	5,989	7,779	9,488	11,668	13,277	14,860	18,466	19,998
5	4,351	7,289	9,236	11,070	13,388	15,086	16,750	20,515	22,106
6	5,348	8,558	10,645	12,592	15,033	16,812	18,548	22,457	24,102
7	6,346	9,803	12,017	14,067	16,622	18,475	20,278	24,321	26,018
8	7,344	11,030	13,362	15,507	18,168	20,090	21,955	26,124	27,867
9	8,343	12,242	14,684	16,919	19,679	21,666	23,589	27,877	29,667
10	9,342	13,442	15,987	18,307	21,161	23,209	25,188	29,588	31,419
11	10,341	14,631	17,275	19,675	22,618	24,725	26,757	31,264	33,138
12	11,340	15,812	18,549	21,026	24,054	26,217	28,300	32,909	34,821
13	12,340	16,985	19,812	22,362	25,471	27,688	29,819	34,527	36,477
14	13,339	18,151	21,064	23,685	26,873	29,141	31,319	36,124	38,109
15	14,339	19,311	22,307	24,996	28,259	30,578	32,801	37,698	39,717
16	15,338	20,465	23,542	26,296	29,633	32,000	34,267	39,252	41,308
17	16,338	21,615	24,769	27,587	30,995	33,409	35,718	40,791	42,881
18	17,338	22,760	25,989	28,869	32,346	34,805	37,156	42,312	44,434
19	18,338	23,900	27,204	30,144	33,687	36,191	38,582	43,819	45,974
20	19,337	25,038	28,412	31,410	35,020	37,566	39,997	45,314	47,498
21	20,337	26,171	29,615	32,671	36,343	38,932	41,401	46,796	49,010
22	21,337	27,301	30,813	33,924	37,659	40,289	42,796	48,268	50,510
23	22,337	28,429	32,007	35,172	38,968	41,638	44,181	49,728	51,999
24	23,337	29,553	33,196	36,415	40,270	42,980	45,558	51,179	53,478
25	24,337	30,675	34,382	37,652	41,566	44,314	46,928	52,619	54,948
26	25,336	31,795	35,563	38,885	42,856	45,642	48,290	54,051	56,407
27	26,336	32,912	36,741	40,113	44,140	46,963	49,645	55,475	57,856
28	27,336	34,027	37,916	41,337	45,419	48,278	50,994	56,892	59,299
29	28,336	35,139	39,087	42,557	46,693	49,588	52,335	58,301	60,734
30	29,336	36,250	40,256	43,773	47,962	50,892	53,672	59,702	62,160