Cours 7 : Variables aléatoires réelles – Vraisemblance

Essentiel du cours 6

Définitions

- Loi de probabilité : $\sum_{x\in\mathcal{X}} p(x) = 1$
- Espérance : $\mathbb{E}(X) = \sum_{x\in\mathcal{X}} p(x)x$
- Variance :
  $\mathbb{V}(X) = \mathbb{E}\left((X-\mathbb{E}(X))^2\right) = \sum_{x\in\mathcal{X}} p(x)\,(x-\mathbb{E}(X))^2$
- Covariance :
  $\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right)$
- Corrélation : $\mathbb{C}\mathrm{or}(X,Y) = \dfrac{\mathbb{C}\mathrm{ov}(X,Y)}{\sqrt{\mathbb{V}(X)\,\mathbb{V}(Y)}}$
- Loi conditionnelle :
  $p_{Y|X=x}(y) = \mathbb{P}\{Y=y \mid X=x\} = \dfrac{p_{XY}(x,y)}{p_X(x)}$
- Espérance conditionnelle :
  $\mathbb{E}(Y \mid X=x) = \sum_{y\in\mathcal{Y}} p_{Y|X=x}(y)y$
- Variance conditionnelle : $\mathbb{V}(Y \mid X=x)$
  $= \mathbb{E}\left((Y-\mathbb{E}(Y \mid X=x))^2 \mid X=x\right)$
  $= \mathbb{E}\left(Y^2 \mid X=x\right) - (\mathbb{E}(Y \mid X=x))^2$
- X et Y indépendantes si et seulement si :
  $\forall(x,y),\, p_{XY}(x,y) = p_X(x)p_Y(y)$

Propositions

- Fonction de répartition :
  $F(x) = \mathbb{P}\{X \leqslant x\} = \sum_{y\leqslant x} p(y)$
- Formule alternative de la variance :
  $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$
- Transformation linéaire :
  $\mathbb{E}(a+bX) = a + b\,\mathbb{E}(X)$
  $\mathbb{V}(a+bX) = b^2\,\mathbb{V}(X)$
- Loi marginale : $p_X(x) = \mathbb{P}\{X=x\} = \sum_{y\in\mathcal{Y}} p_{XY}(x,y)$
- Formule alternative de la covariance :
  $\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$
- Probabilité conditionnelle totale :
  $\sum_{y\in\mathcal{Y}} p_{Y|X=x}(y) = 1$
- Propriétés des variables X, Y indépendantes :
  $p_{Y|X=x}(y) = p_Y(y)$
  $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$
  $\mathbb{C}\mathrm{ov}(X,Y) = 0$

On s'intéresse maintenant à une variable aléatoire X prenant ses valeurs dans un ensemble $\mathcal{X}$ infini, non dénombrable, par exemple

$$\mathcal{X} = [a, b], \qquad \mathcal{X} = \mathbb{R}^+, \qquad \mathcal{X} = \mathbb{R}.$$

Exemples. $X =$

- Temps d'attente avant le prochain autobus ($\mathcal{X} = \mathbb{R}^+$)
- Plus basse température observée l'an prochain en un lieu donné ($\mathcal{X} = \mathbb{R}$)
- Nombre réel tiré uniformément entre 0 et 1 ($\mathcal{X} = [0, 1]$)

Remarque importante. Pour une variable aléatoire X réelle de distribution continue, la probabilité $\mathbb{P}\{X = x\}$ est nulle pour tout $x \in \mathcal{X}$ donc la fonction $p(x)$ est donc nulle partout.

$\rightarrow$ La fonction $p(x) = \mathbb{P}\{X = x\}$ ne décrit donc pas la distribution de X.

Fonction de répartition

Définition 1 (Fonction de répartition)

La fonction de répartition F d'une variable aléatoire réelle X est la fonction

$$F : \begin{array}{ccc} \mathcal{X} & \mapsto & [0,1] \\ x & \to & F(x) = \mathbb{P}\{X \leqslant x\} \end{array}$$

La fonction de répartition vérifie notamment

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to +\infty} F(x) = 1.$$

Proposition 1 (Croissance de la fonction de répartition)

La fonction de répartition F est une fonction croissante :

$$x < y \qquad \Rightarrow \qquad F(x) \leqslant F(y).$$

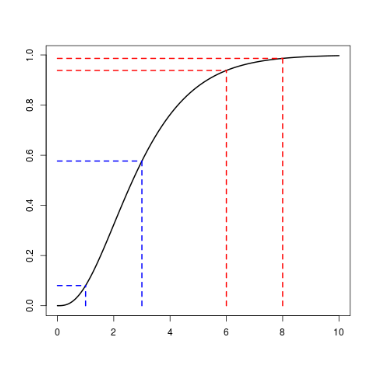
De plus

$$F(y) - F(x) = \mathbb{P}\{x < X \leqslant y\}.$$

Fonction de répartition

Fonction de répartition : exemple

Exemple de fonction de répartition : $\mathcal{X} = \mathbb{R}^+$



$$F(1) \simeq 0.0803, \qquad F(3) \simeq 0.5768, \qquad F(6) \simeq 0.938, \qquad F(8) \simeq 0.9862.$$

Les régions de $\mathcal{X}$ dans lesquelles il est le plus probable d'observer une réalisation de X correspondent aux régions dans lesquelles F augmente le plus vite.

Définition 2 (Densité de probabilité d'un variable réelle)

La densité de probabilité d'une variable X de fonction de répartition F est la dérivée de F :

$$\begin{array}{cccc} f: & \mathcal{X} & \mapsto & [0,1] \\ & x & \to & f(x) = F'(x). \end{array}$$

Remarques.

1. Par définition de la dérivée, F constitue une primitive de la fonction f : pour $a \leqslant b$,

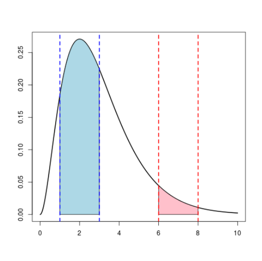
$$F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x.$$

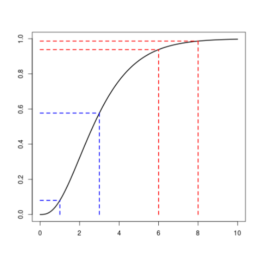
2. Par construction, f vérifie

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1, \qquad F(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u.$$

Exemple de densité : $\mathcal{X} = \mathbb{R}^+$

Densité de probabilité $f(x)$ Fonction de répartition $F(x)$





$$F(1) \simeq 0.0803, \quad F(3) \simeq 0.5768, \quad F(6) \simeq 0.938, \quad F(8) \simeq 0.9862.$$

$$\mathbb{P}\{1 < X \leqslant 3\} = \int_1^3 f(x)\,\mathrm{d}x = F(3) - F(1) \simeq 0.4965, \qquad \mathbb{P}\{6 < X \leqslant 8\} \simeq 0.04821.$$

Les intervalles $[1, 3]$ et $[6, 8]$ sont de même largeur, mais la probabilité de $[1, 3]$ est plus de 10 fois supérieure à celle de $[6, 8]$.

$\rightarrow$ La 'densité' de probabilité est plus grande dans $[1, 3]$ que dans le $[6, 8]$.

Définition 3 (Loi exponentielle)

La loi exponentielle (de paramètre λ) est la loi de la variable aléatoire X réelle positive ($\mathcal{X} = \mathbb{R}^+$) telle que

$$F(x) = \mathbb{P}\{X \leqslant x\} = 1 - \exp(-\lambda x).$$

On note alors $X \sim \mathcal{E}(\lambda)$.

Proposition 2 (Densité de la loi exponentielle)

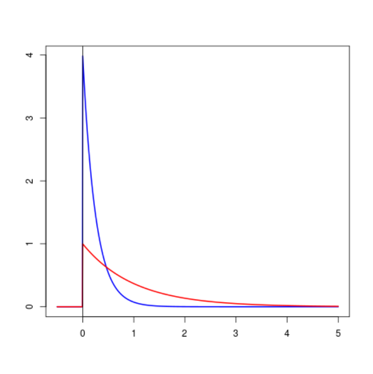
La densité de la loi exponentielle $\mathcal{E}(\lambda)$ est

$$f(x) = F'(x) = \lambda \exp(-\lambda x).$$

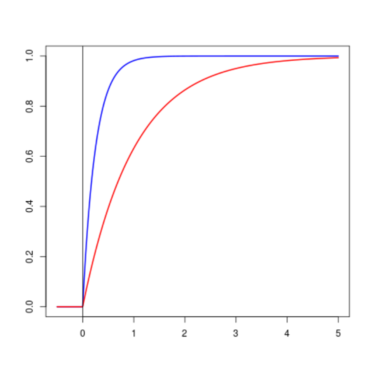
Loi exponentielle : exemple

Lois exponentielles $\mathcal{E}(1)$ et $\mathcal{E}(4)$

Densité de probabilité $f(x)$



Fonction de répartition $F(x)$



La loi exponentielle est souvent utilisée pour décrire la distribution d'un temps d'attente.

Loi exponentielle : exemple

Loi uniforme

Définition 4 (Loi uniforme)

La loi uniforme sur l'intervalle $[0,1]$ est la loi de la variable aléatoire X réelle à valeur dans $[0,1]$ telle que

$$f(x) = 1 \text{ si } x \in [0,1] \text{ et } 0 \text{ sinon.}$$

On note alors $X \sim \mathcal{U}([0,1])$.

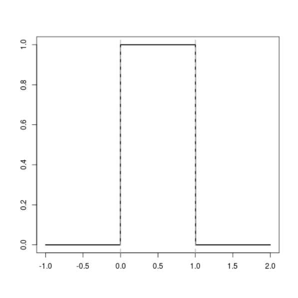
Proposition 3 (Fonction de répartition de la loi uniforme)

La fonction de répartition d'une variable $X \sim \mathcal{U}([0,1])$ vaut

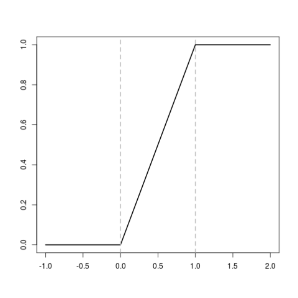
$$F(x) = \begin{cases} 0 & \textit{pour } x < 0, \\ x & \textit{pour } 0 \leqslant x \leqslant 1, \\ 1 & \textit{pour } x > 1. \end{cases}$$

Loi uniforme $\mathcal{U}([0,1])$.

Densité de probabilité $f(x)$



Fonction de répartition $F(x)$



La loi uniforme = loi élémentaire en matière de simulation numérique : la plupart des langages et systèmes proposent une fonction permettant de la simuler.
La simulation de variables issues d'autres lois de probabilités repose le plus souvent sur la simulation d'une ou plusieurs variables uniformes.

Loi exponentielle : exemple

Définition 5 (Loi normale)

La loi normale (de paramètre μ et σ^2) est la loi de la variable aléatoire X réelle ($\mathcal{X} = \mathbb{R}$) telle que

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

On note alors

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

Remarques.

1. Il n'existe pas de forme explicite pour la fonction de répartition de la loi normale :

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \mathrm{d}\, u.$$

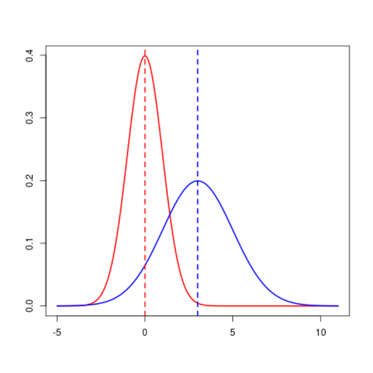
2. On peut remarquer que la fonction f est symétrique par rapport à μ :

$$f(\mu + x) = f(\mu - x) \qquad \Rightarrow \qquad F(\mu) = \frac{1}{2}.$$

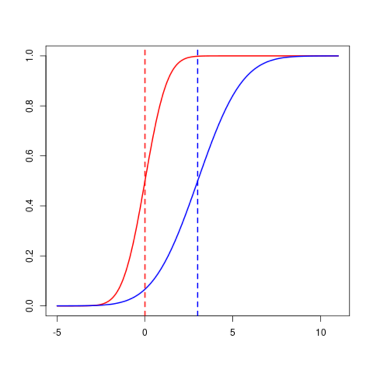
Loi normale : exemple

Lois normales $\mathcal{N}(0,1)$ et $\mathcal{N}(3,4)$

Densité de probabilité $f(x)$



Fonction de répartition $F(x)$



La loi normale est utilisée dans presque tous les domaines scientifiques et techniques.

Définition 6 (Espérance et variance d'une variable aléatoire réelle)

L'espérance (notée $\mathbb{E}$) et la variance (notée $\mathbb{V}$) d'une variable X réelle sont définies par

$$\mathbb{E}(X) = \int_{\mathcal{X}} f(x) x \, \mathrm{d}x, \qquad \mathbb{V}(X) = \int_{\mathcal{X}} f(x) \left(x - \mathbb{E}(X)\right)^2 \mathrm{d}x.$$

Remarques.

1. Là encore, il faut s'assurer que les intégrales existent (hors programme du cours).

2. On montre facilement que

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \left(\mathbb{E}(X)\right)^2.$$

   est également vraie pour une variable réelle.

Proposition 4 (Espérance et variance de la loi uniforme)

L'espérance et la variance d'une variable $X \sim \mathcal{U}([0,1])$ *valent*

$$\mathbb{E}(X) = \frac{1}{2}, \qquad \mathbb{V}(X) = \frac{1}{12}.$$

Proposition 5 (Espérance et variance de la loi normale)

L'espérance et la variance d'une variable $X \sim \mathcal{N}(\mu, \sigma^2)$ *valent*

$$\mathbb{E}(X) = \mu, \qquad \mathbb{V}(X) = \sigma^2.$$

Cette propriété n'est pas démontrée ici : la démonstration repose sur une intégration par partie.

Quantile d'une distribution

La distribution d'une variable continue nous renseigne sur les intervalles dans lesquels elle est le plus susceptible d'être observée.

$\rightarrow$ La notion de quantile permet de déterminer ces intervalles.

Définition 7 (Quantile)

Le *quantile* d'ordre u d'une variable X dont la fonction de répartition est strictement croissante et continue

$$\begin{array}{rccl} q: & [0,1] & \mapsto & \mathbb{R} \\ & u & \rightarrow & q(u) = F^{-1}(u). \end{array}$$

Par construction, le quantile $q(u)$ vérifie

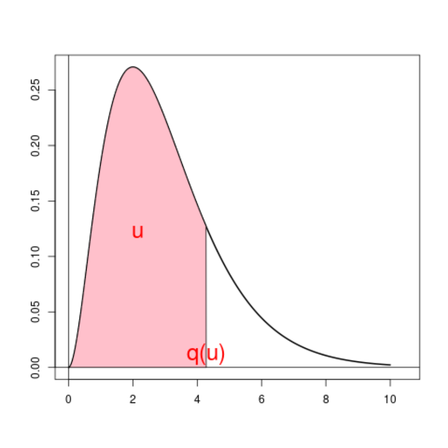
$$\mathbb{P}\{X \leqslant q(u)\} = u$$

puisque $\mathbb{P}\{X \leqslant q(u)\} = F\left(q(u)\right) = F\left(F^{-1}(u)\right) = u.$

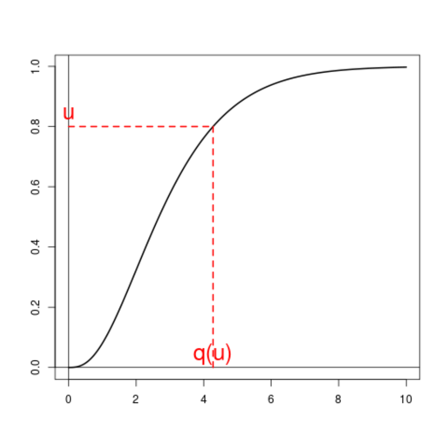
Quantile : exemple

Quantile d'une loi

Densité de probabilité $f(x)$



Fonction de répartition $F(x)$



Proposition 6 (Quantile de la loi exponentielle)

Le quantile d'ordre u d'une variable aléatoire exponentielle $X \sim \mathcal{E}(\lambda)$ vaut

$$q(u) = -\log(1-u)/\lambda.$$

Proposition 7 (Méthode de la fonction inverse)

Si $U \sim \mathcal{U}([0,1])$, alors

$$X = -\log(1-U)/\lambda$$

suit une loi exponentielle $\mathcal{E}(\lambda)$.

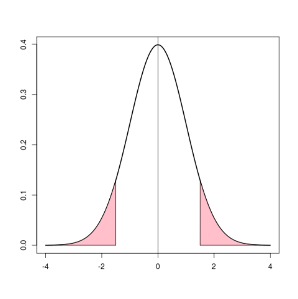
Remarque. La proposition précédente se généralise à toute variable aléatoire dont la fonction de répartition F est strictement croissante de 0 à 1 et continue :

$$U \sim \mathcal{U}([0,1]) \qquad \Rightarrow \qquad X = F^{-1}(U) \text{ a pour fonction de répartition } F.$$

Quantile et loi exponentielle

Quantile et loi exponentielle

Quantiles de la loi normale $\mathcal{N}(0, 1)$



La symétrie de la densité de la loi normale $\mathcal{N}(0, 1)$ implique la symétrie des quantiles :

$$q(1 - u) = -q(u).$$

u	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
$q(u)$	-1.960	-1.645	-1.282	-0.6745	0	0.6745	1.282	1.645	1.960

Distribution paramétrique. De façon générique, on note

- $\mathcal{F}$ = distribution (Bernoulli, exponentielle, ...)
- θ = paramètre (π, λ, ...).

Exemple :

Variables entières		
Loi	$\mathcal{F}$	θ
Bernoulli	$\mathcal{B}$	π
binomiale	$\mathcal{B}$	(n, π)

Variables réelles		
Loi	$\mathcal{F}$	θ
exponentielle	$\mathcal{E}$	λ
normale	$\mathcal{N}$	(μ, σ^2)

On note de même :

$$p(x; \theta), \qquad f(x; \theta) \qquad F(x; \theta).$$

Objectif : Estimer θ à partir de n réalisations $X_1 = x_1$, $X_2 = x_2$, $\dots X_n = x_n$ de la loi $\mathcal{F}(\theta)$.

Échantillon de variables aléatoires

Définition 8 (Échantillon i.i.d.)

Un ensemble de variables aléatoires $X_1, X_2, \ldots, X_n$ constitue un (n-)échantillon *i.i.d.* de la loi $\mathcal{F}(\theta)$ si et seulement si

- les variables $X_1, X_2, \ldots, X_n$ sont toutes indépendantes et
- chacune d'elles est distribuée selon la loi $\mathcal{F}(\theta)$:

$$\forall 1 \leqslant i \leqslant n : \qquad X_i \sim \mathcal{F}(\theta).$$

La notation i.i.d. signifie « indépendantes et identiquement distribuées ».

Exemples de réalisations d'échantillons de taille $n = 5$.

	x_1	x_2	x_3	x_4	x_5
Bernoulli $\mathcal{B}(1/2)$	0	0	1	1	0
Exponentielle $\mathcal{E}(1)$	0.458	0.2582	2.895	1.23	0.5397
Normale $\mathcal{N}(0,1)$	-0.005767	2.405	0.7636	-0.799	-1.148

Echantillon i.i.d.= ensemble d'expériences menées dans les mêmes conditions et indépendamment les unes des autres

Fonction de vraisemblance

Définition 9 (Vraisemblance d'un échantillon i.i.d.)

La fonction de vraisemblance de la réalisation d'un échantillon $(X_1, \dots X_n)$ i.i.d.est

- si les variables X_i sont discrètes :

$$V(x_1, \dots x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta),$$

- si les variables X_i sont réelles :

$$V(x_1, \dots x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta).$$

- Cas discret : puisque les X_i sont indépendants

$$V(x_1, \dots x_n; \theta) = \mathbb{P}_\theta\{X_1 = x_1, X_2 = x_2, \dots X_n = x_n\}.$$

- Cas continu : $V(x_1, \dots x_n; \theta) =$ *densité jointe* des variables $(X_1, \dots X_n)$.

Vraisemblance d'un échantillon de Bernoulli

On a $\theta = \pi$ et $p(x;\pi) = \pi^x(1-\pi)^{1-x}$, donc

$$V(x_1, \ldots x_n; \pi) = \prod_{i=1}^{n} \pi^{x_i}(1-\pi)^{1-x_i} = \pi^y(1-\pi)^{n-y} \qquad \text{avec } y = \sum_{i=1}^{n} x_i.$$

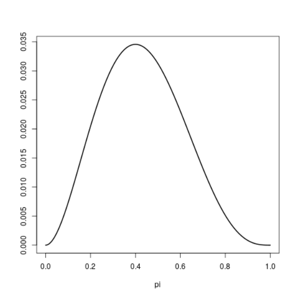
Échantillon de Bernoulli

Pour la réalisation

x_1	x_2	x_3	x_4	x_5
0	0	1	1	0

on a $y = 2$, donc

$$V(x_1, \ldots x_n; \pi) = \pi^2(1-\pi)^3$$



Vraisemblance d'un échantillon de loi exponentielle

Fonction de vraisemblance d'un échantillon de loi exponentielle $\mathcal{E}(1)$. On a $\theta = \lambda$ et $f(x; \lambda) = \lambda \exp(-\lambda x)$, donc

$$V(x_1, \ldots x_n; \pi) = \prod_{i=1}^{n} \lambda \exp(-\lambda x) = \lambda^n \exp(-\lambda y) \qquad \text{avec } y = \sum_{i=1}^{n} x_i.$$

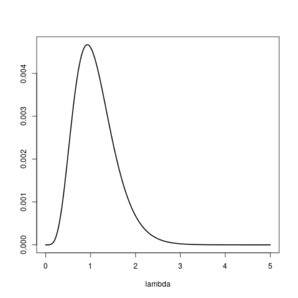
Echantillon de loi exponentielle

Pour la réalisation

x_1	x_2	x_3	x_4	x_5
0.458	0.2582	2.895	1.23	0.5397

on a $y = 5.38$, soit

$$V(x_1, \ldots x_n; \pi) = \lambda^n \exp(-5.38\ \lambda)$$



Fonction de vraisemblance d'un échantillon de loi normale $\mathcal{N}(0,1)$.
On a $\theta = (\mu, \sigma^2)$ et $f(x;\mu,\sigma^2) = \exp(-(x-\mu)^2/2\sigma^2)/(\sigma\sqrt{2\pi})$, donc

$$V(x_1,\ldots x_n;\pi) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x_i^2 - 2\mu x_i + \mu^2}{2\sigma^2}\right) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{z - 2\mu y + n\mu^2}{2\sigma^2}\right)$$

$$\text{avec } y = \sum_{i=1}^{n} x_i \text{ et } z = \sum_{i=1}^{n} x_i^2.$$

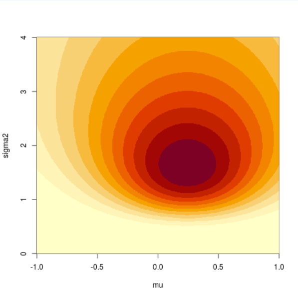
Echantillon de loi normale

Pour la réalisation

x_1	x_2	x_3	x_4	x_5
-0.00577	2.405	0.7636	-0.799	-1.148

on a $y = 1.216$ et $z = 8.321$, soit

$$V(x_1,\ldots x_n;\mu,\sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{8.321 - 2.432\,\mu + \mu^2}{2\sigma^2}\right)$$



Essentiel du cours 7

Définitions

- Fonction de répartition :
$$\begin{array}{rccl} F: & \mathcal{X} & \mapsto & [0,1] \\ & x & \to & F(x) = \mathbb{P}\{X \leqslant x\} \end{array}$$
- Densité de probabilité :
$$\begin{array}{rccl} f: & \mathcal{X} & \mapsto & [0,1] \\ & x & \to & f(x) = F'(x). \end{array}$$
- Espérance : $\mathbb{E}(X) = \displaystyle\int_{\mathcal{X}} f(x) x \, \mathrm{d}x$
- Variance : $\mathbb{V}(X) = \displaystyle\int_{\mathcal{X}} f(x) \left(x - \mathbb{E}(X)\right)^2 \, \mathrm{d}x$
- Quantile : si F continue et strictement croissante
$$\begin{array}{rccl} q: & [0,1] & \mapsto & \mathbb{R} \\ & u & \to & q(u) = F^{-1}(u) \end{array}$$
- Échantillon *i.i.d.* de variables $X_1, X_2, \ldots, X_n$: les variables sont toutes indépendantes et $\forall 1 \leqslant i \leqslant n : X_i \sim \mathcal{F}(\theta)$
- Vraisemblance d'un échantillon *i.i.d.* :
  - X_i discrètes : $V(x_1, \ldots x_n; \theta) = \prod_{i=1}^n p(x_i; \theta)$,
  - X_i réelles : $V(x_1, \ldots x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$.

Propositions

- La fonction de répartition F est croissante
- $F(y) - F(x) = \mathbb{P}\{x < X \leqslant y\}$
- Si $U \sim \mathcal{U}([0,1])$
  alors $X = F^{-1}(U)$ a pour fonction de répartition F

Lois

- Uniforme : $X \sim \mathcal{U}([0,1])$
  $f(x) = 1$ si $x \in [0,1]$ et 0 sinon
- Loi normale : $X \sim \mathcal{N}(\mu, \sigma^2)$
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$
- Exponentielle : $X \sim \mathcal{E}(\lambda)$
  $f(x) = \lambda \exp(-\lambda x)$ si $x \geqslant 0$ et 0 sinon