Probabilités continues

Julie Delon

UNIVERSITÉ
PARIS DESCARTES

Plan du cours

PART 1:	Introduction
PART 2:	Espérance, variance, quantiles
PART 3:	Lois usuelles
PART 4:	Loi normale et cie
PART 5:	Lois jointes, indépendance
PART 6:	Théorèmes limites

Première partie I

Introduction

Du discret au continu

Definition

Une variable aléatoire (abbr. v.a) réelle est une application **mesurable**

$$
\begin{aligned}
X \ : \quad \Omega &\longrightarrow \mathbb{R} \\
\omega &\longmapsto X(\omega)
\end{aligned}
$$

Une **variable aléatoire discrète** prend ses valeurs dans un ensemble fini ou dénombrable

- lancé de dé, $X(\Omega) = \{1, 2, 3, 4, 5, 6\}$
- nombre de photons émis par une source lumineuse pendant 1s, $X(\Omega) = \mathbb{N}$

Du discret au continu

Definition

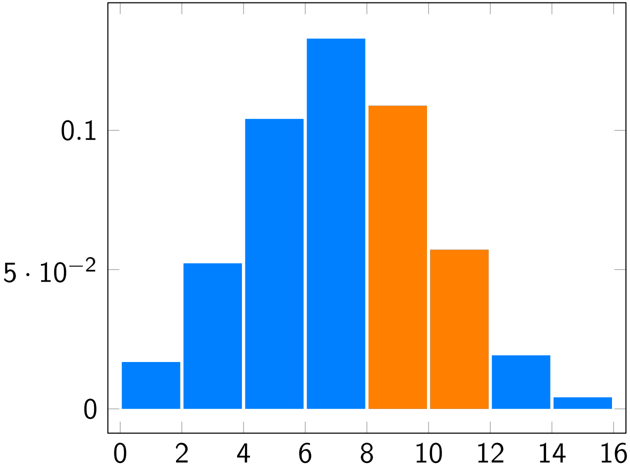
Une variable aléatoire (abbr. v.a) réelle est une application **mesurable**

$$\begin{array}{rcll} X & : & \Omega & \longrightarrow \mathbb{R} \\ & & \omega & \longmapsto X(\omega) \end{array}$$

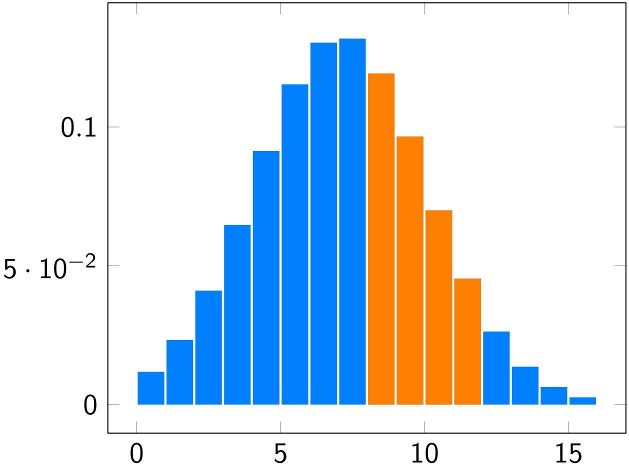
Une **variable aléatoire continue** peut prendre une infinité non dénombrable de valeurs, par exemple dans un intervalle ou sur tout $\mathbb{R}$.

- taille des individus d'une population, $X(\Omega) = [0, M]$
- temps d'attente à la poste, $X(\Omega) = \mathbb{R}^+$
- taux de cholestérol, $X(\Omega) = \mathbb{R}^+$
- poids à la naissance, $X(\Omega) = [0, m]$
- ...

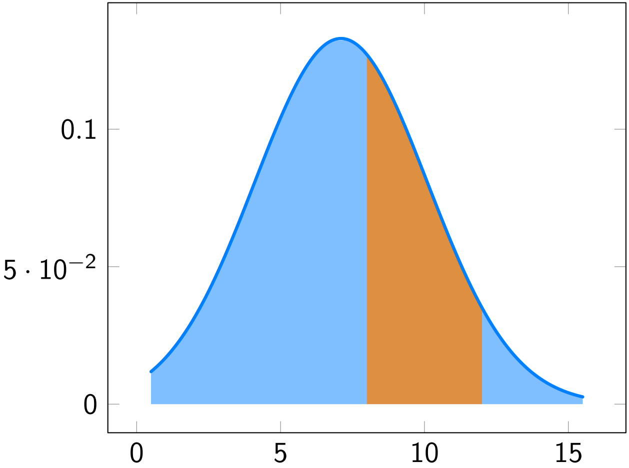
Loi d'une variable aléatoire : du discret au continu



Loi d'une variable aléatoire : du discret au continu



Loi d'une variable aléatoire : du discret au continu



Loi d'une variable aléatoire continue

- Si X a une loi continue, la probabilité que X prenne une valeur bien précise a est en général nulle.
- On ne peut donc pas définir la loi de X en se contentant de donner ses probabilités élémentaires $\mathbb{P}[X = a]$ pour tout a.

Si X désigne le taux de cholestérol d'un individu, alors $\mathbb{P}[X = 0.53969252982\ \mathrm{mg|l}] = 0$.

- On s'intéresse plutôt à la probabilité que X soit dans un intervalle donné $[a, b]$, ou qu'il soit inférieur à une valeur donnée a.

Du discret au continu

Exercice

On jette un stylo sur une table, et on note X l'angle (non signé, donc entre 0 et π) qu'il forme avec le bord de la table. Quelle est la loi de X ? Comment peut-on la représenter graphiquement ? Quelle est la probabilité de l'évènement $\{\frac{\pi}{3} \leq X \leq \frac{\pi}{2}\}$?

Densité de probabilité

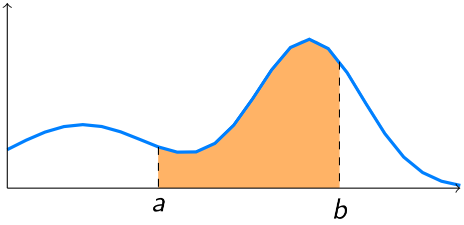
Definition

Une variable aléatoire X est dite **à densité** lorsqu'il existe une fonction positive $f_X : \mathbb{R} \to \mathbb{R}^+$ telle que

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx \quad \text{pour tous } a, b \in \mathbb{R}, \quad a \leq b.$$

Cette fonction f_X est appelée **densité** de X.

remarque : on peut prendre $a = -\infty$ ou $b = +\infty$ dans cette formule.



La probabilité $P(a \leq X \leq b)$ correspond à l'aire du domaine situé sous le graphe de f_X entre les abscisses a et b.

Densité de probabilité

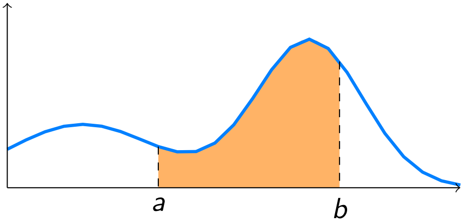
Definition

Une variable aléatoire X est dite **à densité** lorsqu'il existe une fonction positive $f_X : \mathbb{R} \to \mathbb{R}^+$ telle que

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx \quad \text{pour tous } a, b \in \mathbb{R}, \quad a \leq b.$$

Cette fonction f_X est appelée **densité** de X.

remarque : on peut prendre $a = -\infty$ ou $b = +\infty$ dans cette formule.



La probabilité $P(a \leq X \leq b)$ correspond à l'aire du domaine situé sous le graphe de f_X entre les abscisses a et b.

$\int_{-\infty}^{+\infty} f_X(x)dx = \mathbb{P}(X \in \mathbb{R}) = 1$. Il faut toujours penser à le vérifier !

Calculer la loi d'une variable à densité, c'est calculer sa densité !

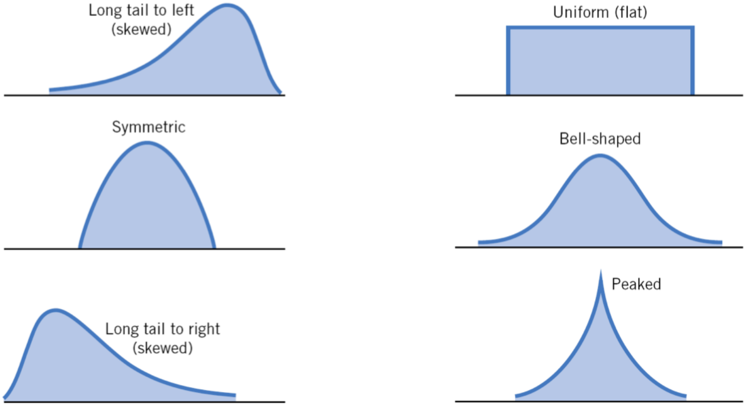
$\mathbb{P}[X = x]$

Si la variable aléatoire X a une densité f_X, alors pour toute valeur a, la probabilité que X prenne la valeur a est 0 ! ! !

$$\mathbb{P}(X = a) = \mathbb{P}(a \leq X \leq a) = \int_a^a f_X(x)dx = 0$$

On s'intéresse plutôt à la probabilité que X prenne ses valeurs dans un intervalle donné $[a, b]$

Décrire une loi



Même terminologie que pour des distributions discrètes : dyssymétrie (skewness), moyenne, variance, médian, mode, quantiles, etc.

Exercice

Exercice

Parmi les fonctions suivantes, lesquelles sont des densités de probabilité ?

1. $f(x) = \begin{cases} x+1 & \text{si } -1 \leq x \leq 0 \\ 1-x & \text{si } 0 \leq x \leq 11 \\ 0 & \text{sinon.} \end{cases}$
2. $f(x) = \begin{cases} x^2 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon.} \end{cases}$
3. $f(x) = \begin{cases} 2\cos(x) & \text{si } 0 \leq x \leq 2\pi \\ 0 & \text{sinon.} \end{cases}$
4. $f(x) = \begin{cases} \frac{3}{4}(1-x^2) & \text{si } -1 \leq x \leq 1 \\ 0 & \text{sinon.} \end{cases}$

Fonction de répartition

Definition

La **fonction de répartition** d'une variable aléatoire X est la fonction définie pour tout $t \in \mathbb{R}$ par

$$F_X(t) = \mathbb{P}(X \leq t).$$

Autrement dit, $F_X(t)$ est la probabilité de l'événement "la valeur de X est inférieure ou égale à t".

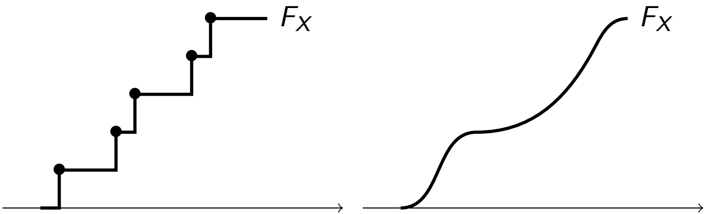


FIGURE: Exemples de fonctions de répartition d'une variable discrète et d'une variable continue.

Propriétés de la Fonction de répartition

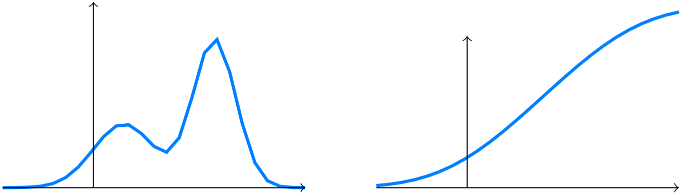
- $F_X(t) \in [0,1]$ pour tout $t \in \mathbb{R}$;
- F_X est une fonction **croissante**
- $\lim_{x\to-\infty} F_X(x) = 0$ et $\lim_{x\to+\infty} F_X(x) = 1$
- pour tout $a < b$ dans $\mathbb{R}$, $F_X(b) - F_X(a) = \mathbb{P}[a < X \leq b]$.

Densité de probabilité de fonction de répartition

Proposition

Si la fonction de répartition F_X est dérivable, alors X est une variable a densité et sa densité est la dérivée de F_X :

$$f_X = F'_X$$



Exercice

On jette un stylo sur une table, et on note X l'angle (non signé, donc entre 0 et π) qu'il forme avec le bord de la table. Quelle est la fonction de répartition de la loi de X ?

Deuxième partie II

Espérance, variance, quantiles

Espérance

Definition

Soit X une variable aléatoire continue de densité f_X, son espérance est

$$\mathbb{E}[X] = \int_{\mathbb{R}} t f_X(t) dt,$$

lorsque cette intégrale est bien définie.

Si l'intégrale précédente n'est pas convergente, alors l'espérance de X n'est pas définie.

$\mathbb{E}[X]$ est une moyenne pondérée des valeurs que peut prendre X.

Propriétés de l'espérance

Proposition

Soient X et Y deux variables aléatoires et $\lambda \in \mathbb{R}$ un nombre réel, on a alors

- $E(\lambda) = \lambda$
- $E(\lambda X) = \lambda E(X)$
- $E(X + \lambda) = E(X) + \lambda$
- $E(\lambda_1 X + \lambda_2 Y) = \lambda_1 E(X) + \lambda_2 E(Y)$

⚠ Attention cependant, meme si l'espérance admet beaucoup de propriétés qui la rendent agréable, elle ne respecte pas en général la multiplication ($E(XY) \neq E(X)E(Y)$), sauf pour des variables indépendantes.

Propriétés de l'espérance

Proposition

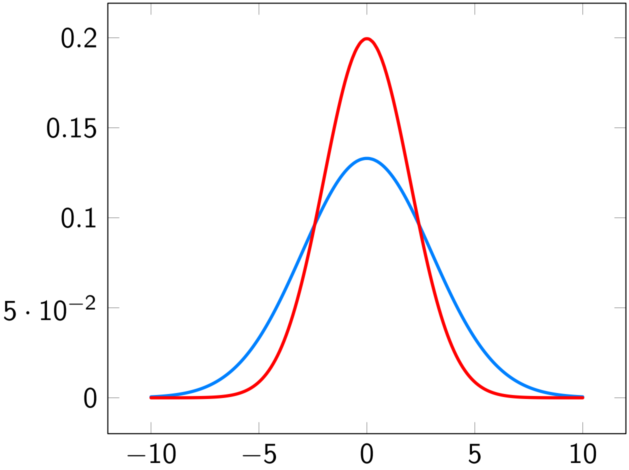
Soient X une variable aléatoire continue de densité f_X, et $g : \mathbb{R} \to \mathbb{R}$ une fonction quelconque. L'espérance de $g(X)$ se calcule ainsi :

$$E(g(X)) = \int_{\mathbb{R}} g(x) f_X(x) dx.$$

⚠Il se peut très bien que $E(g(X))$ n'existe pas alors que $E(X)$ existe !

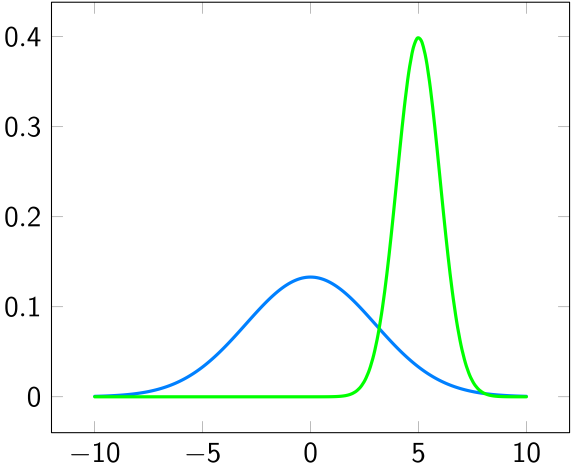
Espérance

Que pouvez-vous dire des espérances relatives de ces densités ?



Espérance

Que pouvez-vous dire des espérances relatives de ces densités ?



Variance

Definition

Soit X une variable aléatoire continue de densité f_X, sa variance est

$$
\begin{aligned}
Var[X] &= \mathbb{E}[(X-\mathbb{E}[X])^2] = \int_{\mathbb{R}} (t-\mathbb{E}[X])^2 f_X(t)dt \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_{\mathbb{R}} t^2 f_X(t)dt - \left(\int_{\mathbb{R}} t f_X(t)dt\right)^2,
\end{aligned}
$$

lorsque ces intégrales sont bien définies.

La variance est un nombre positif, qui peut être infini même si l'espérance existe.

Definition

L' **écart-type** d'une variable aléatoire X est la racine carrée de sa variance :

$$\sigma(X) = \sqrt{Var(X)}.$$

Propriétés de la variance

Proposition

Soit X une variable aléatoire et $\lambda \in \mathbb{R}$ un nombre réel, on a alors

- $Var(\lambda) =$
- $Var(X + \lambda) =$
- $Var(\lambda X) =$

Propriétés de la variance

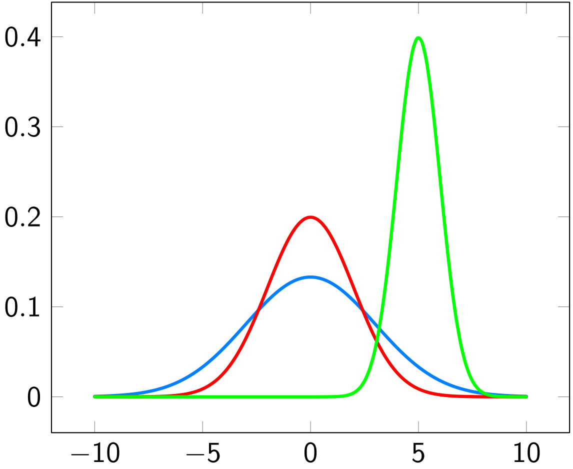
Proposition

Soit X une variable aléatoire et $\lambda \in \mathbb{R}$ un nombre réel, on a alors

- $Var(\lambda) = 0$
- $Var(X + \lambda) = Var(X)$
- $Var(\lambda X) = \lambda^2 Var(X)$

Variance

Que pouvez-vous dire des variances de ces densités ?



Quantile

Definition

Les **quantiles** d'une distribution f sont les valeurs permettant de diviser le support de la distribution en intervalles de poids égaux.

On parle de q-quantile lorsqu'on divise le poids de la distribution en q intervalles. Il y en a $q-1$
Par exemple

- 2-quantile = median
- 3-quantile = tercile
- 4-quantile = quartile
- 10-quantile = décile

Les q-quantiles de la distribution f_X de fonction de répartition F_X sont les valeurs

$$F_X^{-1}\left(\frac{i}{q}\right), \quad i \in \{1, \ldots, q-1\}$$

Exercice

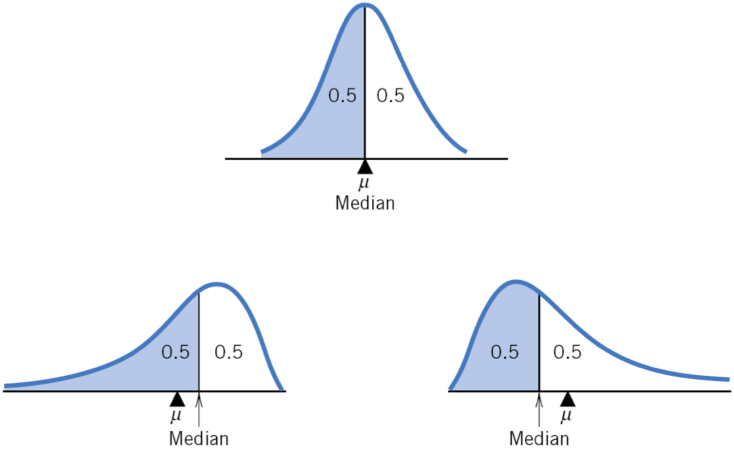
Exercice

Soit X une variable aléatoire de densité

$$f_X(x) = \begin{cases} \frac{4}{3}x^{1/3} & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon.} \end{cases}$$

1. Quelle est la fonction de répartition de X ?
2. Quelle est l'espérance de X ?
3. Quelle est la variance de X ?
4. Quelle est la probabilité de l'évènement $\{\frac{1}{3} \leq X \leq \frac{1}{2}\}$?
5. Quelles sont les terciles de la loi de X ?

Décrire une loi : relation entre moyenne et médian

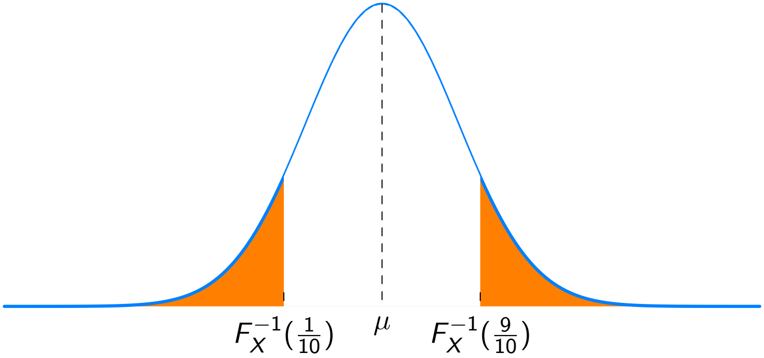


Cas des lois symétriques

Si une loi est symétrique par rapport à la valeur μ, sa moyenne et sa médiane coïncident et valent μ.

Dans ce cas, son i^{eme} q-quantile et son $(q-i)^{eme}$ quantile sont symétriques l'un de l'autre par rapport à a :

$$\mu - F_X^{-1}\left(\frac{i}{q}\right) = F_X^{-1}\left(\frac{q-i}{q}\right) - \mu$$



Troisième partie III

Lois usuelles

Loi uniforme

Definition

La loi uniforme sur un intervalle $[\alpha, \beta]$ est la loi de densité

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} \text{ si } x \in [\alpha, \beta], \\ 0 \text{ sinon.} \end{cases}$$

On note $X \sim \mathcal{U}([\alpha, \beta])$ ("X suit la loi uniforme sur $[\alpha, \beta]$").

Dans l'exemple du stylo qui tombe sur une table, il est raisonnable de supposer que l'angle X suit la loi uniforme sur $[0, \pi]$. On aura donc par exemple :

$$P\left(X \leq \frac{\pi}{2}\right) = \int_0^{\frac{\pi}{2}} \frac{1}{\pi} dx = \frac{1}{2},$$

ou encore :

$$P\left(\frac{\pi}{4} \leq X \leq \frac{\pi}{2}\right) = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{1}{\pi} dx = \frac{1}{4}.$$

Fonction de répartition d'une loi uniforme

Si X suit la loi uniforme sur $[\alpha, \beta]$, on a

$$f_X(x) = \begin{cases} \frac{1}{\beta-\alpha} \text{ si } x \in [\alpha, \beta], \\ 0 \text{ sinon}, \end{cases}$$

et donc

$$F_X(x) = \begin{cases} 0 \text{ si } x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha} \text{ si } x \in [\alpha, \beta], \\ 1 \text{ si } x \geq \beta \end{cases}$$

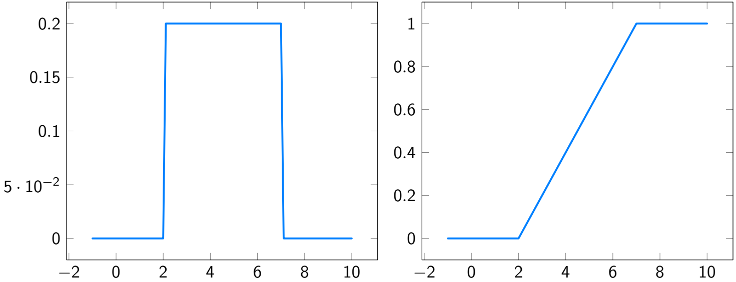


Figure: Densité (à gauche) et fonction de répartition (à droite) de la loi uniforme sur l'intervalle $[2, 7]$.

Loi uniforme

Exercice Soit X une variable aléatoire de loi uniforme sur $[\alpha, \beta]$. Calculez l'espérance et la variance de X.

Loi exponentielle

Definition

Soit $a > 0$ un réel. On dit que X suit la loi exponentielle de paramètre a si elle admet la densité

$$f_X(x) = ae^{-ax}1_{\mathbb{R}^+}(x) = \begin{cases} ae^{-ax} \text{ si } x \geq 0, \\ 0 \text{ sinon.} \end{cases}$$

Sa fonction de répartition est :

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 1 - e^{-ax} \text{ si } x \geq 0, \\ 0 \text{ sinon.} \end{cases}$$

Loi exponentielle

La loi exponentielle est souvent utilisée pour modéliser la loi de temps d'attente ou de durées de vie (a est l'inverse du temps d'attente moyen).

- durée de vie d'une ampoule, d'un appareil électrique
- temps jusqu'au prochain tremblement de terre
- temps d'attente à la poste...

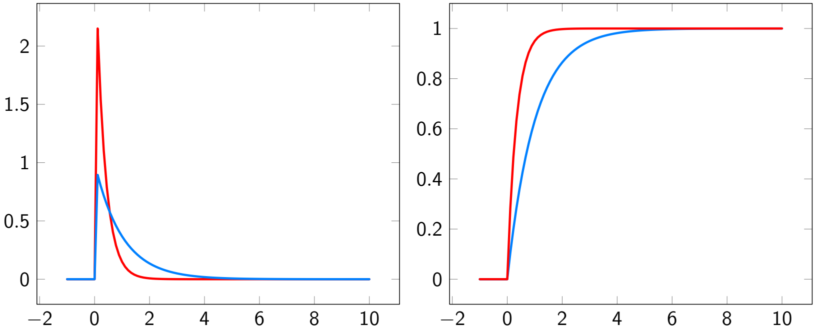


FIGURE: Densité (à gauche) et fonction de répartition (à droite) de la loi exponentielle. Courbe cyan pour $a = 1$ et rouge pour $a = 3$

Espérance et variance d'une variable de loi exponentielle

Exercice Soit X une variable aléatoire de loi exponentielle de paramètre $a > 0$. X admet la densité

$$f_X(x) = ae^{-ax}1_{\mathbb{R}^+} = \begin{cases} ae^{-ax} \text{ si } x \geq 0, \\ 0 \text{ sinon.} \end{cases} .$$

Calculez l'espérance et la variance de X.

Espérance de X

Calculons $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x) dx = \int_0^{+\infty} x a e^{-ax} dx.$$

On fait une intégration par parties avec $u = x$, $u' = 1$ et $v' = ae^{-ax}$, $v = -e^{-ax}$:

$$\begin{aligned}\mathbb{E}(X) &= \left[x(-e^{-ax})\right]_0^{+\infty} - \int_0^{+\infty} (-e^{-ax}) dx \\ &= 0 + \int_0^{+\infty} e^{-ax} dx = \left[-\frac{1}{a} e^{-ax}\right]_0^{+\infty} = \frac{1}{a}.\end{aligned}$$

Ainsi

$$\boxed{\mathbb{E}(X) = \frac{1}{a}}$$

Variance de X

Calculons $\mathrm{Var}(X)$:

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \int_{\mathbb{R}} x^2 f_X(x) dx - \frac{1}{a^2} = \int_0^{+\infty} x^2 a e^{-ax} dx - \frac{1}{a^2}.$$

On pose $u = x^2$, $u' = 2x$ et $v' = ae^{-ax}$, $v = -e^{-ax}$:

$$\begin{aligned}
\mathrm{Var}(X) &= \left[x^2(-e^{-ax})\right]_0^{+\infty} - \int_0^{+\infty} 2x(-e^{-ax}) dx - \frac{1}{a^2} \\
&= 0 + 2\int_0^{+\infty} x e^{-ax} dx - \frac{1}{a^2} \\
&= 2\left[x(-\frac{1}{a}e^{-ax})\right]_0^{+\infty} - 2\int_0^{+\infty} (-\frac{1}{a}e^{-ax}) dx - \frac{1}{a^2} \\
&= 0 + \frac{2}{a}\int_0^{+\infty} e^{-ax} dx - \frac{1}{a^2} \\
&= \frac{2}{a}\left[-\frac{1}{a}e^{-ax}\right]_0^{+\infty} - \frac{1}{a^2} = \frac{2}{a}\frac{1}{a} - \frac{1}{a^2} = \frac{1}{a^2}.
\end{aligned}$$

$$\boxed{\mathrm{Var}(X) = \frac{1}{a^2}}$$

Loi quelconque

⚠**Une variable aléatoire non discrète n'est pas nécessairement à densité !**

Exercice Une machine à remplir les bouteilles est défectueuse : elle verse dans chaque bouteille (de 75cL) une quantité aléatoire de boisson comprise entre 0 et 1 litre. Soit Y la quantité de boisson contenue dans la bouteille. Décrire sa loi.

Loi quelconque

⚠**Une variable aléatoire non discrète n'est pas nécessairement à densité !**

Exercice Une machine à remplir les bouteilles est défectueuse : elle verse dans chaque bouteille (de 75cL) une quantité aléatoire de boisson comprise entre 0 et 1 litre. Soit Y la quantité de boisson contenue dans la bouteille. Décrire sa loi.

Soit X la quantité de boisson versée par la machine. En l'absence d'autre précision, on peut considérer que X suit une loi uniforme sur $[0,1]$. On a

$$\begin{cases} Y = X \text{ si } X \leq 0.75 \\ Y = 0.75 \text{ si } X > 0.75 \end{cases}$$

Y n'est clairement pas une variable discrète (ses valeurs possibles correspondent à l'intervalle $[0,0.75]$), et n'est pas non plus une variable à densité car $P(Y=0.75) = P(X > 0.75) = 0.25 \neq 0$. On dit que la loi de Y possède un **atome** en $x = 0.75$.

On peut calculer facilement la fonction de répartition de Y :

$$F_Y(t) = \begin{cases} 0 \text{ si } t \leq 0, \\ t \text{ si } 0 < t < 0.75, \\ 1 \text{ si } t \geq 0.75. \end{cases}$$

Quatrième partie IV

Loi normale et cie

Loi normale centrée réduite (ou loi gaussienne)

Apparaît comme limite de certains processus.

Loi normale centrée réduite (ou loi gaussienne)

Definition

On dit que X suit une loi normale centrée réduite et on note $X \sim \mathcal{N}(0,1)$ si sa loi admet pour densité la fonction

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

La fonction de répartition correspondante n'a pas de formule simple. Dans les logiciels de calcul numérique (Matlab, R, etc) cette fonction est implémentée sous le nom de **normcdf** (Matlab) ou **pnorm** (R).

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

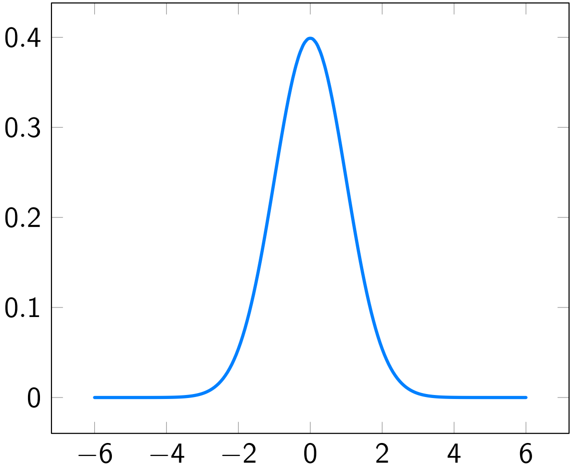
Proposition

Soit X une variable aléatoire suivant une loi normale centrée réduite, alors

$$\mathbb{E}[X] = 0 \;\text{ et }\; Var(X) = 1.$$

Loi normale

$\mathcal{N}(0,1)$



Exercice Soit X une variable aléatoire de loi normale centrée réduite. Vérifiez que $\mathbb{E}[X] = 0$.

Loi normale $\mathcal{N}(\mu, \sigma^2)$

Definition

Soit μ et σ deux réels, on suppose $\sigma \neq 0$. On dit qu'une variable aléatoire X suit une loi normale de moyenne μ et variance σ^2 si la variable aléaoire

$$Z = \frac{X - \mu}{\sigma}$$

suit une loi normale centrée réduite.
On note $X \sim \mathcal{N}(\mu, \sigma^2)$. X a pour densité la fonction

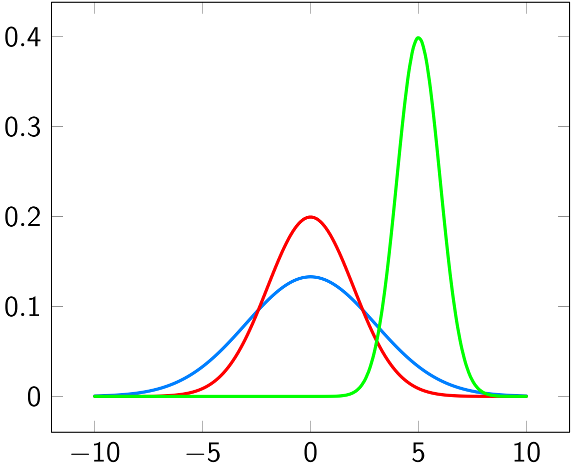
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

On a $\mathbb{E}[X] = \mu$ et $Var[X] = \sigma^2$.

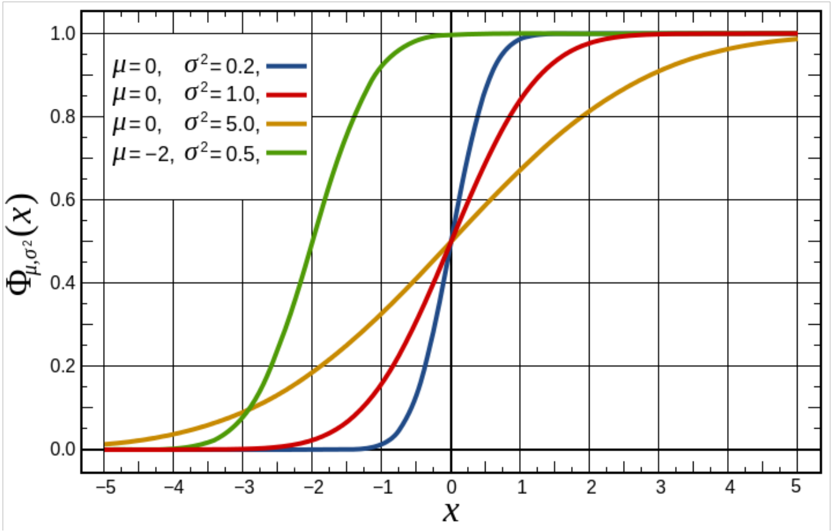
⚠Attention, dans certains ouvrages, il est noté $X \sim \mathcal{N}(\mu, \sigma)$ au lieu de $X \sim \mathcal{N}(\mu, \sigma^2)$.

Loi normale

$\mathcal{N}(0,3)$, $\mathcal{N}(0,2)$, $\mathcal{N}(5,1)$



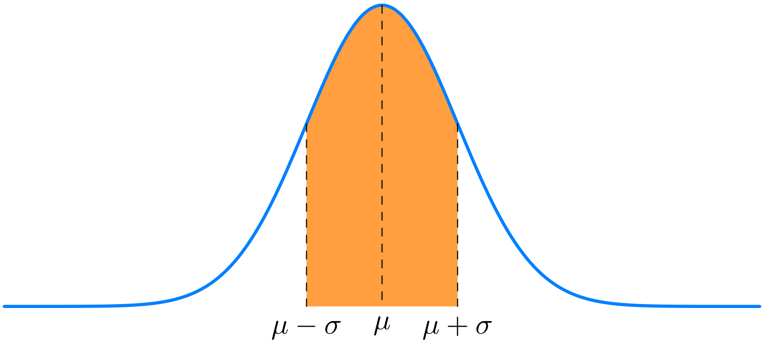
Fonction de répartition de la loi normale



Concentration autour de la moyenne

Dans l'intervalle $[\mu - \sigma, \mu + \sigma]$ centré autour de la moyenne μ, il y a 68% de la masse de la distribution $\mathcal{N}(\mu, \sigma^2)$

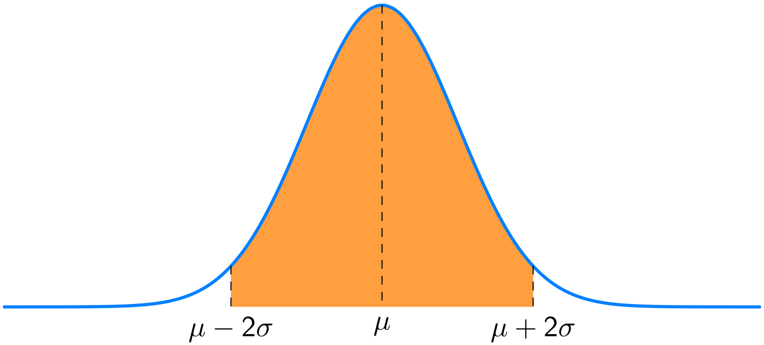
$$\mathbb{P}[\mu - \sigma \leq X \leq \mu + \sigma] \simeq 0.68.$$



Concentration autour de la moyenne

Dans l'intervalle $[\mu - 2\sigma, \mu + 2\sigma]$ (ou plus précisément dans $[\mu - 1.96\sigma, \mu + 1.96\sigma]$) centré autour de la moyenne μ, il y a 95% de la masse de la distribution $\mathcal{N}(\mu, \sigma^2)$

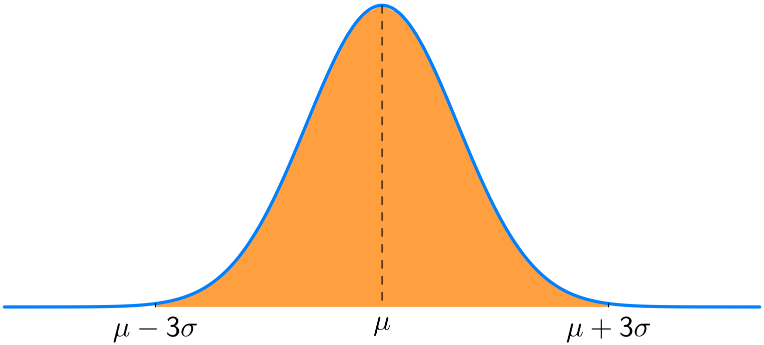
$$\mathbb{P}[\mu - 2\sigma \leq X \leq \mu + 2\sigma] \simeq 0.95.$$



Concentration autour de la moyenne

Dans l'intervalle $[\mu - 3\sigma, \mu + 3\sigma]$ centré autour de la moyenne μ, il y a $99,7\%$ de la masse de la distribution $\mathcal{N}(\mu, \sigma^2)$

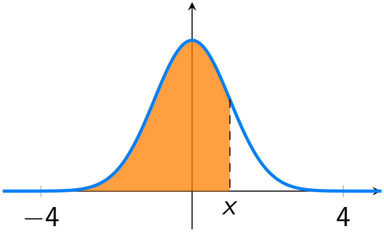
$$\mathbb{P}[\mu - 3\sigma \leq X \leq \mu + 3\sigma] \simeq 0.997.$$



Z-table pour $\mathcal{N}(0,1)$

$$\mathbb{P}[Z \leq x]$$

avec $Z \sim \mathcal{N}(0,1)$.



x	0.00	0.01	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	0.7703	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
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Exercice : poids à la naissance

Exercice Soit X la variable aléatoire représentant le poids d'un bébé à la naissance en France. On fait l'hypothèse que X suit une loi normale de moyenne $\mu = 3500$g et d' écart-type $\sigma = 500$g. Quelle est la probabilité qu'un enfant ait un poids inférieur à 3.1kg à la naissance ?

Exercice : poids à la naissance

Exercice Soit X la variable aléatoire représentant le poids d'un bébé à la naissance en France. On fait l'hypothèse que X suit une loi normale de moyenne $\mu = 3500$g et d' écart-type $\sigma = 500$g. Quelle est la probabilité qu'un enfant ait un poids inférieur à 3.1kg à la naissance ?

On commence par « standardiser » X en posant

$$Z = \frac{X - \mu}{\sigma}$$

On a alors

$$\begin{aligned}
\mathbb{P}[X \leq 3100] &= \mathbb{P}[\frac{X-\mu}{\sigma} \leq \frac{3100-\mu}{\sigma}] \\
&= \mathbb{P}[Z \leq \frac{3100-3500}{500}] \\
&= \mathbb{P}[Z \leq -0.8] = 1 - \mathbb{P}[Z \leq 0.8] \\
&\simeq 1 - 7881 = 0.2119.
\end{aligned}$$

Autres lois utiles pour les tests

On va s'intéresser maintenant à plusieurs lois utiles pour les tests statistiques et construites à partir de variables gaussiennes :

- la loi du χ^2
- la loi de Student
- la loi de Fisher.

Exercice

Exercice Soit X une v.a. normale centrée réduite, calculer la densité de la variable X^2.

Loi du χ^2

Definition

Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes, identiquement distribuées, de loi $\mathcal{N}(0,1)$. Alors la variable aléatoire

$$Z = X_1^2 + \cdots + X_n^2$$

suit une loi du Khi-deux à n degrés de liberté, notée $\chi^2(n)$ et de densité

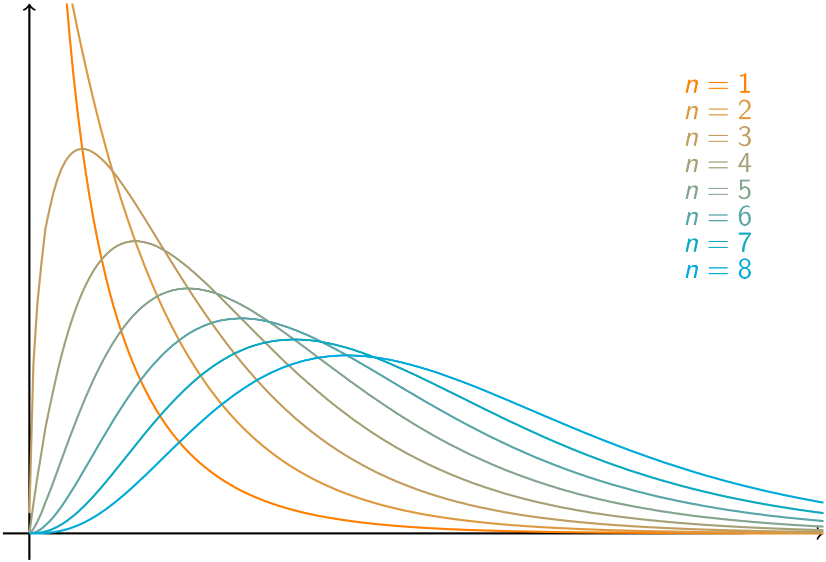
$$f_Z(x) = \frac{\left(\frac{1}{2}\right)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \mathbf{1}_{x \geq 0}.$$

- $\mathbb{E}[Z] = n$ et $Var(Z) = 2n$

Lorsque n est grand, cette loi peut être approchée par la loi $\mathcal{N}(n, 2n)$.

Cette loi est principalement utilisée dans le **test du** χ^2 en statistique, qui permet de tester l'adéquation entre des données et une famille de lois.

Loi du χ^2



Loi de Student

Definition

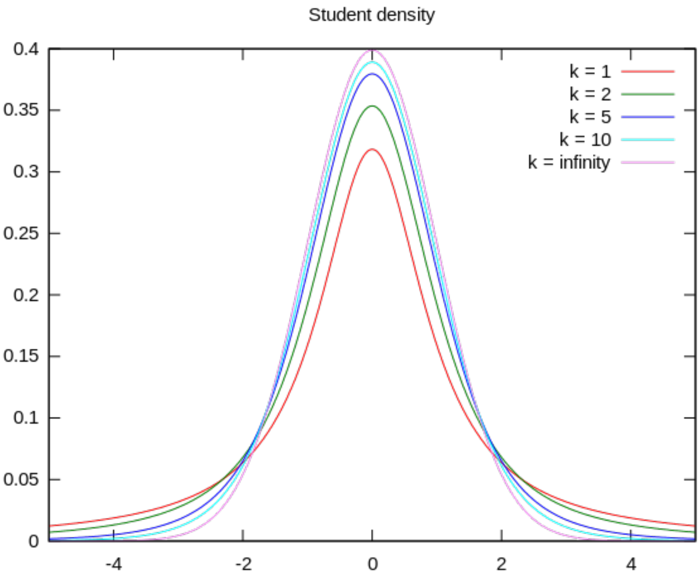
Soient $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(n)$ deux variables indépendantes. Alors $Z = \frac{X}{\sqrt{Y/n}}$ suit une loi de Student à n degrés de liberté. On note $Z \sim Student(n)$. Sa densité est

$$f_Z(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\,\Gamma(n/2)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}, \quad t \in \mathbb{R}$$

Lorsque n est grand, cette loi peut être approchée par la loi $\mathcal{N}(0,1)$.

Utilisation dans le test de Student, pour tester si la moyenne d'un échantillon qui suit une loi normale est égale à une valeur donnée, ou pour estimer un intervalle de confiance sur l'estimation de cette moyenne.

Loi de Student



Utilisation de la loi de Student

On observe une caractéristique X d'une population (la taille par exemple), et on suppose qu'elle suit une **loi normale** $\mathcal{N}(\mu, \sigma^2)$, avec μ et σ **inconnus**. On note X_i la valeur de cette caractéristique pour l'individu i.

On aimerait estimer la moyenne μ de la population.

Utilisation de la loi de Student

On observe une caractéristique X d'une population (la taille par exemple), et on suppose qu'elle suit une **loi normale** $\mathcal{N}(\mu, \sigma^2)$, avec μ et σ **inconnus**. On note X_i la valeur de cette caractéristique pour l'individu i.

On aimerait estimer la moyenne μ de la population.

Pour cela, on tire n individus dans la population et on calcule la moyenne empirique

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Utilisation de la loi de Student

On observe une caractéristique X d'une population (la taille par exemple), et on suppose qu'elle suit une **loi normale** $\mathcal{N}(\mu, \sigma^2)$, avec μ et σ **inconnus**. On note X_i la valeur de cette caractéristique pour l'individu i.

On aimerait estimer la moyenne μ de la population.

Pour cela, on tire n individus dans la population et on calcule la moyenne empirique

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

On aimerait savoir à quel point on peut faire **confiance** à cette estimation. Si on connaissait σ, on pourrait commencer par remarquer que

$$\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

et utiliser ce qu'on a vu sur la loi gaussienne pour dire par exemple que

$$\mathbb{P}[|\bar{X}_n - \mu| \leq 2\frac{\sigma}{\sqrt{n}}] \simeq 0.95$$

Mais on ne connaît pas σ...

Utilisation de la loi de Student

On peut à la place calculer l'écart-type empirique $\bar{\sigma}_n$ sur cet échantillon :

$$\bar{\sigma}_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$$

Utilisation de la loi de Student

On peut à la place calculer l'écart-type empirique $\bar{\sigma}_n$ sur cet échantillon :

$$\bar{\sigma}_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$$

Mais $\bar{\sigma}_n$ est seulement une estimation imparfaite de σ. Par contre, on peut montrer le résultat suivant :

Proposition

Si les X_i suivent tous une même loi gaussienne $\mathcal{N}(\mu, \sigma^2)$ et sont indépendants, alors la variable

$$Z = \frac{\bar{X}_n - \mu}{\bar{\sigma}_n / \sqrt{n-1}}$$

suit une loi de Student à $n-1$ degrés de liberté.

Utilisation de la loi de Student

On peut donc utiliser les quantiles de la loi de Student pour estimer un intervalle de confiance pour $\bar{X}_n$. Par exemple, si on connaît α tel qu'une variable Z suivant la loi de Student à $n-1$ degrés de liberté vérifie

$$\mathbb{P}[|Z| \leq \alpha] = 0.95,$$

on saura dire que

$$\mathbb{P}[|\bar{X}_n - \mu| \leq \alpha \frac{\bar{\sigma}_n}{\sqrt{n-1}}] = 0.95.$$

Test de Student

Une usine fabrique des comprimés qui doivent peser un gramme. Aucun comprimé ne pèse exactement $\mu = 1g$, mais on souhaite contrôler s'ils pèsent bien 1g en moyenne, en supposant que le poids des comprimés suit une loi normale.

Test de Student

Une usine fabrique des comprimés qui doivent peser un gramme. Aucun comprimé ne pèse exactement $\mu = 1g$, mais on souhaite contrôler s'ils pèsent bien 1g en moyenne, en supposant que le poids des comprimés suit une loi normale.
Imaginons qu'on extrait 17 comprimés, qu'on observe un poids moyen $\bar{X}_n = 0.995g$ et un écart-type empirique de $\bar{\sigma}_n = 0.01g$.

Test de Student

Une usine fabrique des comprimés qui doivent peser un gramme. Aucun comprimé ne pèse exactement $\mu = 1g$, mais on souhaite contrôler s'ils pèsent bien 1g en moyenne, en supposant que le poids des comprimés suit une loi normale.
Imaginons qu'on extrait 17 comprimés, qu'on observe un poids moyen $\bar{X}_n = 0.995g$ et un écart-type empirique de $\bar{\sigma}_n = 0.01g$.

Soit

$$Z = \frac{\bar{X}_n - \mu}{\frac{\bar{\sigma}_n}{\sqrt{n-1}}} = \frac{0.995 - 1}{0.01/\sqrt{16}} = -2.$$

Si les comprimés suivent bien une loi normale de moyenne μ, Z doit suivre une loi de Student à $n-1$ degrés de liberté, donc la probabilité que Z soit aussi faible est

Test de Student

Une usine fabrique des comprimés qui doivent peser un gramme. Aucun comprimé ne pèse exactement $\mu = 1g$, mais on souhaite contrôler s'ils pèsent bien 1g en moyenne, en supposant que le poids des comprimés suit une loi normale.

Imaginons qu'on extrait 17 comprimés, qu'on observe un poids moyen $\bar{X}_n = 0.995g$ et un écart-type empirique de $\bar{\sigma}_n = 0.01g$.

Soit

$$Z = \frac{\bar{X}_n - \mu}{\frac{\bar{\sigma}_n}{\sqrt{n-1}}} = \frac{0.995 - 1}{0.01/\sqrt{16}} = -2.$$

Si les comprimés suivent bien une loi normale de moyenne μ, Z doit suivre une loi de Student à $n-1$ degrés de liberté, donc la probabilité que Z soit aussi faible est

$$F_{Student}(-2) = 0.0314.$$

On a donc une probabilité inférieure à 3.14% d'observer ces valeurs si l'hypothèse est vérifiée. Cette quantité est appelée la **p-valeur** du test, elle indique la valeur du seuil de probabilité auquel l'hypothèse $\mu = \mu_0$ serait rejetée.

Loi de Fisher

Definition

Soient $X \sim \chi^2(n_1)$ et $Y \sim \chi^2(n_2)$ deux variables indépendantes. Alors $Z = \frac{X/n_1}{Y/n_2}$ suit une loi de Fisher de paramètres n_1, n_2. On note $Z \sim \mathcal{F}(n_1, n_2)$. Sa densité est

$$f_Z(x) = \frac{\Gamma((n_1+n_2)/2)(n_1/n_2)^{n_1/2} x^{n_1/2-1}}{\Gamma(n_1/2)\Gamma(n_2/2)[(n_1/n_2)x+1]^{(n_1+n_2)/2}} \qquad x > 0.$$

Remarque : si X suit une loi de Student à n degrés de libertés

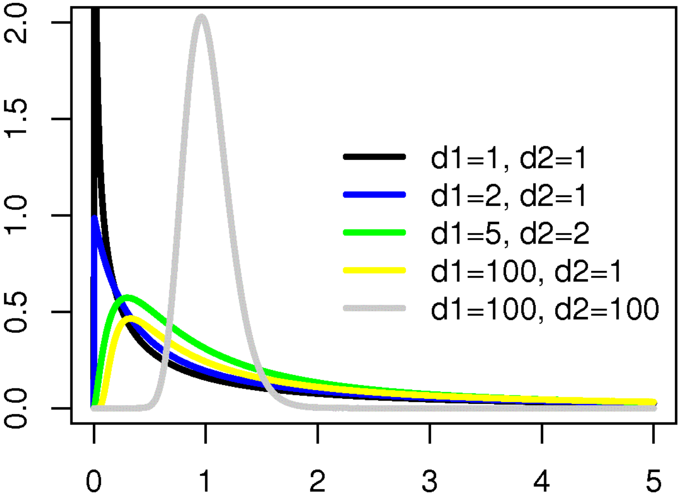
$$X \sim Student(n)$$

alors X^2 suit une loi de Fisher de paramètres $1, n$

$$X^2 \sim \mathcal{F}(1, n).$$

Utilisée en statistiques dans l'analyse de variance (ANOVA).

Loi de Fisher



Cinquième partie V

Lois jointes, indépendance

Vecteur aléatoire

Definition

Un **couple aléatoire** est un couple (X, Y) où X et Y sont des variables aléatoires. De même on parle de triplet aléatoire (X, Y, Z) si on considère trois variables X, Y, Z, et plus généralement de vecteur aléatoire, ou n-uplet aléatoire $(X_1, X_2, \ldots, X_n)$, en considérant n variables $X_1, X_2, \ldots, X_n$.

Si X et Y correspondent à des lancers de dés indépendants, l'ensemble des valeurs que peuvent prendre X et Y est $\{1, 2, 3, 4, 5, 6\}$. L'ensemble des valeurs que peut prendre le couple (X, Y) est $\{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$, ce qui se note aussi $\{1, 2, 3, 4, 5, 6\}^2$. Le cardinal de cet ensemble est donc $6^2 = 36$.

Exercice Avec 10 lancers de dé, on obtient 10 variables $X_1, \ldots, X_{10}$, et donc un vecteur aléatoire $(X_1, \ldots, X_{10})$. Combien de résultats différents peut-on obtenir pour ce vecteur ?

Loi jointe

Definition

Si un couple (X, Y) de variables aléatoires a une loi continue, on définit sa densité jointe comme la fonction $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\mathbb{P}[a \leq X \leq b; c \leq Y \leq d] = \int_a^b \int_c^d f_{X,Y}(x, y) dxdy$$

Exercice Supposons qu'un couple (X, Y) ait pour densité jointe la fonction $f_{X,Y}(x, y) = x + y$ pour $0 \leq x \leq 1$ et $0 \leq y \leq 1$. Est-ce bien une densité ? Calculez $\mathbb{P}[X \leq 0.5; Y \leq 0.5]$.

Marginales d'une loi jointe

Proposition

Si un couple (X, Y) a pour densité jointe la fonction $f_{X,Y}$, alors on peut retrouver les densités respectives des variables X et Y en intégrant

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy,$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dx.$$

Exercice Supposons qu'un couple (X, Y) ait pour densité jointe la fonction $f_{X,Y}(x, y) = x + y$ pour $0 \leq x \leq 1$ et $0 \leq y \leq 1$. Calculez $\mathbb{P}[X \leq 0.5]$.

Loi conditionnelle

Les probabilités conditionnelles peuvent être étendues aux variables aléatoires à densité :

$$f_Y(y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)},$$

$$f_Y(y|X = x)f_X(x) = f_{X,Y}(x, y) = f_X(x|Y = y)f_Y(y).$$

Variables indépendantes

Intuitivement, deux variables sont indépendants si la connaissance de l'une ne donne aucune information sur la valeur de l'autre.

On lance deux dés. Soit X le résultat du premier et Y le résultat du second. Alors X et Y sont indépendantes.

Variables indépendantes

Definition

Deux variables aléatoires X et Y sont dites **indépendantes** si pour tout a, b,

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b).$$

On note parfois $X \perp\!\!\!\perp Y$. ("X indépendant de Y").

Proposition

Une condition nécessaire et suffisante pour que deux variables soient indépendantes est que leur densité jointe s'écrive

Variables indépendantes

Definition

Deux variables aléatoires X et Y sont dites **indépendantes** si pour tout a, b,

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b).$$

On note parfois $X \perp\!\!\!\perp Y$. ("X indépendant de Y").

Proposition

Une condition nécessaire et suffisante pour que deux variables soient indépendantes est que leur densité jointe s'écrive

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

On a alors $f_Y(y|X=x) = f_Y(y)$ et $f_X(x|Y=y) = f_X(x)$ et $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Covariance

Definition

Soient X et Y deux variables aléatoires dont les variances sont définies. La **covariance** de X et Y est

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))).$$

La covariance est une mesure incomplète de l'indépendance de deux variables. En effet on a la résultat suivant :

Proposition

Soient X et Y deux variables indépendantes. Alors $\mathrm{Cov}(X, Y) = 0$.

Par contre la réciproque est fausse !

Matrice de covariance d'un vecteur aléatoire

Pour un vecteur aléatoire $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, on peut définir toutes les covariances $\mathrm{Cov}(X_i, X_j)$ pour $1 \leq i \leq n$ et $1 \leq j \leq n$ et les rassembler dans une matrice :

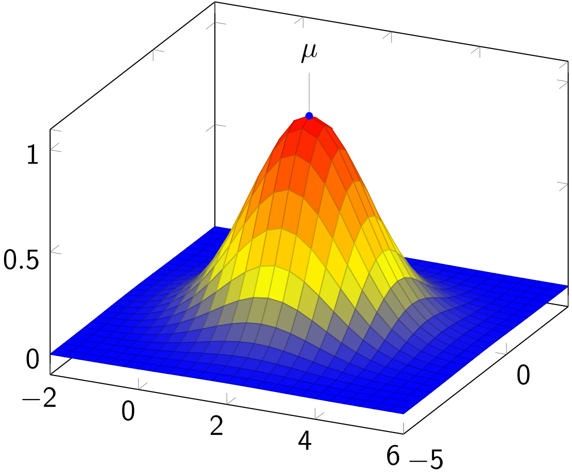
$$\Gamma(\mathbf{X}) = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \mathrm{Cov}(X_1, X_3) & \cdots & \mathrm{Cov}(X_1, X_n) \\ \mathrm{Cov}(X_1, X_2) & \mathrm{Var}(X_2) & \mathrm{Cov}(X_2, X_3) & \cdots & \mathrm{Cov}(X_2, X_n) \\ \mathrm{Cov}(X_1, X_3) & \mathrm{Cov}(X_2, X_3) & \mathrm{Var}(X_3) & \cdots & \mathrm{Cov}(X_3, X_n) \\ \vdots & & & \ddots & \vdots \\ \mathrm{Cov}(X_1, X_n) & \mathrm{Cov}(X_2, X_n) & \mathrm{Cov}(X_3, X_n) & \cdots & \mathrm{Var}(X_n) \end{pmatrix}.$$

Cette matrice est symétrique et à valeurs propres positives.

Exemple de loi jointe continue

Loi gaussienne

$$\mathcal{N}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \Sigma\right) \text{ avec } \Sigma \text{ matrice } 2 \times 2.$$



Loi d'une somme de variables discrètes

Proposition

Soient X et Y deux variables discrètes indépendantes, et $Z = X + Y$. Alors la loi de Z se calcule à partir de celles de X et Y via la formule : pour tout $z \in \mathcal{S}(Z)$,

$$P(Z = z) = \sum_{x \in \mathcal{S}(X)} P(X = x) P(Y = z - x),$$

où $\mathcal{S}(X)$ et $\mathcal{S}(Z)$ désignent respectivement les supports (ensemble des valeurs prises) des variables X et Z.

A l'aide de ce résultat (et de quelques calculs !), on peut montrer que la loi de deux variables de Poisson indépendantes de paramètres λ et μ est une loi de Poisson de paramètre $\lambda + \mu$.

Loi d'une somme de variables continues

Proposition

Soient X et Y deux variables indépendantes admettant des densités f_X et f_Y, et $Z = X + Y$. Alors Z est une variable à densité et f_Z se calcule à partir de f_X et f_Y via la formule : pour tout $z \in \mathbb{R}$,

$$f_Z(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x) dx.$$

On dit que f_Z est le **produit de convolution** de f_X et f_Y.

Somme de deux variables exponentielles

Exercice Soient X et Y deux variables indépendantes de lois exponentielles de paramètre a. Calculer la loi de $Z = X + Y$.

Somme de deux variables exponentielles

Exercice Soient X et Y deux variables indépendantes de lois exponentielles de paramètre a. Calculer la loi de $Z = X + Y$.

On a $f_X(x) = f_Y(x) = ae^{-ax}$ si $x \geq 0$, 0 sinon. La formule donne :

$$f_Z(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x) dx.$$

Or $f_X(x)f_Y(z-x)$ vaut $ae^{-ax}ae^{-a(z-x)}$ lorsque $x \geq 0$ et $z - x \geq 0$, soit $x \geq 0$ et $x \leq z$, soit encore $x \in [0, +\infty] \cap [-\infty, z]$. Sinon $f_X(x)f_Y(z-x)$ vaut 0. On a deux cas :

- Si $z < 0$ alors $[0, +\infty] \cap [-\infty, z] = \varnothing$ donc $f_Z(z) = 0$
- Si $z \geq 0$ alors $[0, +\infty] \cap [-\infty, z] = [0, z]$, donc

$$\begin{aligned} f_Z(z) &= \int_0^z ae^{-ax}ae^{-a(z-x)} dx = a^2 \int_0^z e^{-ax-az+ax} dx \\ &= a^2 \int_0^z e^{-az} dx = a^2 e^{-az} \int_0^z 1 dx = a^2 z e^{-az}. \end{aligned}$$

Ainsi la densité de Z est $f_Z(z) = a^2 z e^{-az}$ si $z \geq 0$, 0 sinon.

Somme de deux variables uniformes

Exercice Soient X et Y deux variables indépendantes de lois uniformes sur $[0, 1]$
Calculer la loi de $Z = X + Y$.

Somme de deux variables uniformes

Exercice Soient X et Y deux variables indépendantes de lois uniformes sur $[0,1]$
Calculer la loi de $Z = X + Y$.

X et Y sont des variables à densité et $f_X(x) = f_Y(x) = 1$ si $0 \leq x \leq 1$, 0 sinon. Z est donc une variable à densité d'après la proposition et sa densité est donnée par

$$f_Z(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x) dx.$$

Or $f_X(x)f_Y(z-x)$ vaut 1 lorsque $0 \leq x \leq 1$ et $0 \leq z - x \leq 1$, soit $0 \leq x \leq 1$ et $z - 1 \leq x \leq z$, soit encore $x \in [0,1] \cap [z-1, z]$. $f_X(x) = 0$ sinon. L'intersection $[0,1] \cap [z-1, z]$ est parfois vide. On a en fait plusieurs cas :

- Si $z < 0$ alors l'intersection est vide donc $f_Z(z) = 0$.
- Si $0 \leq z < 1$ alors $[0,1] \cap [z-1, z] = [0, z]$ donc $f_Z(z) = \int_0^z 1 dx = z$.
- Si $1 \leq z \leq 2$ alors $[0,1] \cap [z-1, z] = [z-1, 1]$ donc $f_Z(z) = \int_{z-1}^1 1 dx = 1 - (z-1) = 2 - z$.
- Si $z \geq 2$ alors l'intersection est vide donc $f_Z(z) = 0$.

Somme de deux gaussiennes

Proposition

Soient X et Y deux variables indépendantes suivant des lois gaussiennes $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$, alors $X + Y$ suit une loi gaussienne

$$\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

De même, si on somme n variables gaussiennes indépendantes $X_1, \ldots, X_n$ de même loi $\mathcal{N}(\mu, \sigma^2)$, alors $X_1 + \cdots + X_n$ aura une loi

Somme de deux gaussiennes

Proposition

Soient X et Y deux variables indépendantes suivant des lois gaussiennes $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$, alors $X + Y$ suit une loi gaussienne

$$\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

De même, si on somme n variables gaussiennes indépendantes $X_1, \ldots, X_n$ de même loi $\mathcal{N}(\mu, \sigma^2)$, alors $X_1 + \cdots + X_n$ aura une loi

$$\mathcal{N}(n\mu, n\sigma^2),$$

et donc la moyenne $\bar{X}_n = \sum X_i / n$ aura pour loi

Somme de deux gaussiennes

Proposition

Soient X et Y deux variables indépendantes suivant des lois gaussiennes $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$, alors $X + Y$ suit une loi gaussienne

$$\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

De même, si on somme n variables gaussiennes indépendantes $X_1, \dots, X_n$ de même loi $\mathcal{N}(\mu, \sigma^2)$, alors $X_1 + \cdots + X_n$ aura une loi

$$\mathcal{N}(n\mu, n\sigma^2),$$

et donc la moyenne $\bar{X}_n = \sum X_i / n$ aura pour loi

$$\mathcal{N}(\mu, \frac{\sigma^2}{n}).$$

Ce résultat est très important, il montre que la loi gaussienne est stable par moyenne, ce qui explique en partie le Théorème Central Limite.

Sixième partie VI

Théorèmes limites

Introduction

On lance un grand nombre de fois une pièce de monnaie équilibrée. On s'attend à observer en moyenne 50% de piles et 50% de faces.

Introduction

On lance un grand nombre de fois une pièce de monnaie équilibrée. On s'attend à observer en moyenne 50% de piles et 50% de faces.
Plus précisément, si on note X_i la v.a. de Bernoulli qui vaut

$$X_i = \begin{cases} 1 & \text{si on tombe sur pile} \\ 0 & \text{sinon} \end{cases}$$

alors le pourcentage de piles sur n tirages est représenté par

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i .$$

Introduction

On lance un grand nombre de fois une pièce de monnaie équilibrée. On s'attend à observer en moyenne 50% de piles et 50% de faces.
Plus précisément, si on note X_i la v.a. de Bernoulli qui vaut

$$X_i = \begin{cases} 1 & \text{si on tombe sur pile} \\ 0 & \text{sinon} \end{cases}$$

alors le pourcentage de piles sur n tirages est représenté par

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

On voudrait pouvoir dire que

$$\overline{X_n} \xrightarrow[n\to+\infty]{} 0.5 = \mathbb{E}[X_1]$$

en un certain sens. C'est ce qu'on appelle **la loi des grands nombres**.

Introduction

Exercice On lance n fois un dé à six faces et on note X_i le résultat du i^e lancer, puis on calcule la moyenne des n lancers $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Qu'obtient-on lorsque n devient grand ?

Introduction

Exercice On lance n fois un dé à six faces et on note X_i le résultat du i^e lancer, puis on calcule la moyenne des n lancers $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Qu'obtient-on lorsque n devient grand ?

L'intuition laisse penser que lorsque n devient grand, cette moyenne devient proche de l'espérance d'un des lancers X_i, c'est-à-dire de

Introduction

Exercice On lance n fois un dé à six faces et on note X_i le résultat du i^e lancer, puis on calcule la moyenne des n lancers $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Qu'obtient-on lorsque n devient grand ?

L'intuition laisse penser que lorsque n devient grand, cette moyenne devient proche de l'espérance d'un des lancers X_i, c'est-à-dire de

$$\mathbb{E}(X_i) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = 3.5$$

Introduction

Exercice On lance n fois un dé à six faces et on note X_i le résultat du i^e lancer, puis on calcule la moyenne des n lancers $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Qu'obtient-on lorsque n devient grand ?

L'intuition laisse penser que lorsque n devient grand, cette moyenne devient proche de l'espérance d'un des lancers X_i, c'est-à-dire de

$$\mathbb{E}(X_i) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = 3.5$$

On voudrait donc pouvoir dire que la limite de $\bar{X}_n$ lorsque $n \to +\infty$ vaut 3.5. C'est ce qu'on appelle la loi des grands nombres.

Introduction

Exercice On dispose d'un dé truqué : il a une probabilité p inconnue de tomber sur 6, et $\frac{1-p}{5}$ sur chacune des autres faces. Comment connaître la valeur de p simplement à partir des résultats obtenus en lançant n fois le dé ?

Introduction

Exercice On dispose d'un dé truqué : il a une probabilité p inconnue de tomber sur 6, et $\frac{1-p}{5}$ sur chacune des autres faces. Comment connaître la valeur de p simplement à partir des résultats obtenus en lançant n fois le dé ?

Si la loi des grands nombres est correcte, on peut l'utiliser ici et dire que la moyenne $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ des résultats est proche de l'espérance

Exercice On dispose d'un dé truqué : il a une probabilité p inconnue de tomber sur 6, et $\frac{1-p}{5}$ sur chacune des autres faces. Comment connaître la valeur de p simplement à partir des résultats obtenus en lançant n fois le dé ?

Si la loi des grands nombres est correcte, on peut l'utiliser ici et dire que la moyenne $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ des résultats est proche de l'espérance

$$\mathbb{E}(X_i) = (1+2+3+4+5) \times \frac{1-p}{5} + 6p = 3(1-p) + 6p = 3(p+1).$$

Introduction

Exercice On dispose d'un dé truqué : il a une probabilité p inconnue de tomber sur 6, et $\frac{1-p}{5}$ sur chacune des autres faces. Comment connaître la valeur de p simplement à partir des résultats obtenus en lançant n fois le dé ?

Si la loi des grands nombres est correcte, on peut l'utiliser ici et dire que la moyenne $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ des résultats est proche de l'espérance

$$\mathbb{E}(X_i) = (1+2+3+4+5) \times \frac{1-p}{5} + 6p = 3(1-p) + 6p = 3(p+1).$$

Or si $\bar{X}_n \simeq 3(p+1)$, c'est que $\frac{1}{3}\bar{X}_n \simeq p+1$, donc que $\frac{1}{3}\bar{X}_n - 1 \simeq p$. Ainsi une valeur approchée de p est donnée par $\frac{1}{3}\bar{X}_n - 1$. On a donc donné une estimation de p, et on dit que $\frac{1}{3}\bar{X}_n - 1$ est un **estimateur** de p car c'est une quantité que l'on peut calculer à partir des résultats des lancers, et qui donne une valeur approchée de la valeur inconnue p.

Une application en débruitage d'image (Gaussian additive noise, $\sigma = 20$)

Moyenne de 5 images

Moyenne de 10 images

Moyenne de 20 images

Moyenne de 40 images

Image non bruitée

Burst denoising

Figure 1.1: From left to right: (a) one long-exposure image (time=0.4 s, ISO=100), one of 16 short-exposure images (time=1/40 s, ISO=1600) and their average after registration. The long exposure image is blurry due to camera motion. (b) The middle short-exposure image is noisy. (c) The third image is about **four times** less noisy, being the result of averaging 16 short-exposure images. From [19].

From *Buadès et al., A note on multi-image denoising, 2009*.

Convergence d'une suite de variables aléatoires

Il existe plusieurs notions de convergence de variables aléatoires.

Convergence d'une suite de variables aléatoires

Il existe plusieurs notions de convergence de variables aléatoires.

De la même manière que l'on peut s'intéresser à la convergence d'une suite $(x_n)_{n\geq 0}$ de nombres réels déterminés, on peut aussi regarder la convergence d'une suite $(X_n)_{n\geq 0}$ de variables aléatoires, mais cette convergence aura forcément un caractère aléatoire puisque les X_n sont aléatoires.

Convergence d'une suite de variables aléatoires

Il existe plusieurs notions de convergence de variables aléatoires.

De la même manière que l'on peut s'intéresser à la convergence d'une suite $(x_n)_{n\geq 0}$ de nombres réels déterminés, on peut aussi regarder la convergence d'une suite $(X_n)_{n\geq 0}$ de variables aléatoires, mais cette convergence aura forcément un caractère aléatoire puisque les X_n sont aléatoires.
Dans ce cours, on se restreint à la convergence presque sûre

Definition

On dit que $(X_n)_{n\geq 0}$ **converge presque sûrement** vers X si $\mathbb{P}\left(X_n \xrightarrow[n\to\infty]{} X\right) = 1$. On l'écrit

$$X_n \xrightarrow[n\to\infty]{p.s.} X$$

Moyenne empirique

Definition

Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires. On appelle **moyenne empirique** (des n premiers termes de la suite) la variable aléatoire

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Proposition

Soit $(X_n)_{n\geq 0}$ une suite de variables indépendantes et toutes de même loi. On suppose que l'espérance et la variance communes à tous les X_n : $\mu = E(X_n)$, $\sigma^2 = \mathrm{Var}(X_n)$, sont bien définies. Alors pour tout $n \geq 0$,

$$E(\bar{X}_n) = \mu \qquad \text{et} \qquad \mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}.$$

Lorsque n tend vers $+\infty$, la variance de $\bar{X}_n$ tend vers 0, ce qui laisse penser que $\bar{X}_n$ converge vers un nombre non aléatoire !

Loi forte des grands nombres

Proposition

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles indépedantes et de même loi. On fait l'hypothèse que $\mathbb{E}[|X_1|] < \infty$. Alors, la suite des moyennes empiriques $(\bar{X}_n)$ converge presque sûrement vers $\mathbb{E}[X_1]$,

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[X_1].$$

Autres applications

- Vérifier si une pièce est équilibrée
- Estimer la proportion d'électeurs votant pour un candidat aux élections
- Estimer si un dé est truqué

Autres applications

- Vérifier si une pièce est équilibrée
- Estimer la proportion d'électeurs votant pour un candidat aux élections
- Estimer si un dé est truqué

Pour l'instant, on sait que

$$\bar{X}_n \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[X_1],$$

mais on ne sait pas à quelle **vitesse** la convergence a lieu. Donc combien de fois il faut lancer la pièce ou combien d'électeurs il faut interroger pour avoir une bonne estimation.

C'est le **théorème central limite** qui nous permet de répondre à cette question.

Théorème Central Limite

Theorème

Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes et identiquement distribuées, et Y une variable de loi $\mathcal{N}(0,1)$. On suppose que l'espérance $\mu = E(X_1)$ et la variance $\sigma^2 = \mathrm{Var}(X_1)$ sont bien définies. Alors pour tous a, b réels tels que $a < b$,

$$\mathbb{P}\left(a \leq \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \leq b\right) \xrightarrow[n\to\infty]{} \mathbb{P}(a \leq Y \leq b) = \int_a^b f_Y(t)dt = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt.$$

Remarque $\bar{X}_n$ a pour moyenne μ et variance $\frac{\sigma^2}{n}$. Et

$$\frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) = \frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}}.$$

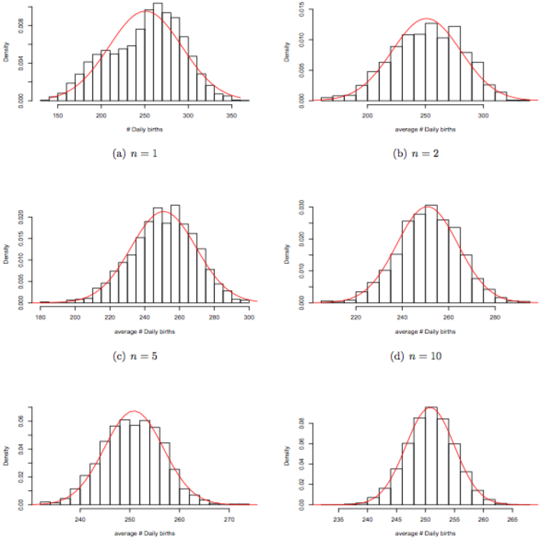
Donc ce théorème nous dit que lorsqu'on centre et qu'on réduit la variable $\bar{X}_n$, sa loi converge vers la loi $\mathcal{N}(0,1)$. En pratique, tout calcul de probabilité effectué sur cette moyenne de variables i.i.d. avec suffisamment d'échantillons peut donc être approché par un calcul de probabilité avec la loi normale centrée réduite.

Théorème Central Limite

Remarques

- La loi Normale a un caractère **universel** : la loi gaussienne est la seule loi continue de variance finie stable par moyenne, donc les distributions qui ne le sont pas ont tendance à converger vers elle.
- le $\sqrt{n}$ est indispensable...
- Ce résultat est crucial en statistique pour établir des tests et des intervalles de confiance

Exemple



(a) $n = 1$ (b) $n = 2$

(c) $n = 5$ (d) $n = 10$

(e) $n = 50$ (f) $n = 100$

Figure 7.4: Normal approximations to averages of n samples from the Quebec birth data.

Approximation de la loi binômiale par la loi normale

Proposition

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires i.i.d. de Bernoulli de paramètre p :
$P(X_i = 1) = p$ et $P(X_i = 0) = 1 - p$.
On note $S_n = X_1 + \cdots + X_n$. S_n suit une loi binômiale

$$S_n \sim \mathcal{B}(n, p).$$

D'après le théorème central limite, la loi de $\sqrt{\frac{n}{p(1-p)}}(\frac{S_n}{n} - p)$ tend vers la loi $\mathcal{N}(0, 1)$.
La loi de S_n peut donc être approchée en pratique par la loi

Approximation de la loi binômiale par la loi normale

Proposition

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires i.i.d. de Bernoulli de paramètre p :
$P(X_i = 1) = p$ et $P(X_i = 0) = 1 - p$.
On note $S_n = X_1 + \cdots + X_n$. S_n suit une loi binômiale

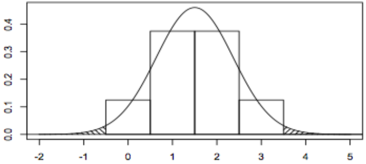
$$S_n \sim \mathcal{B}(n, p).$$

D'après le théorème central limite, la loi de $\sqrt{\frac{n}{p(1-p)}}(\frac{S_n}{n} - p)$ tend vers la loi $\mathcal{N}(0, 1)$.
La loi de S_n peut donc être approchée en pratique par la loi

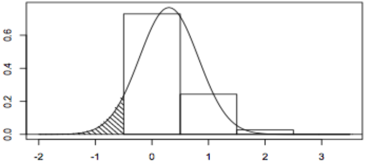
$$\mathcal{N}(np, np(1 - p)).$$

Pour que cette approximation soit bonne, il faut en particulier que la moyenne de la binômiale soit suffisamment grande devant son écart-type, donc $np >> \sqrt{np(1-p)}$, donc $np >> 1 - p$.

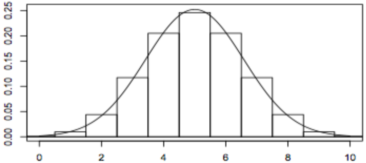
Approximation de la loi binômiale par la loi normale



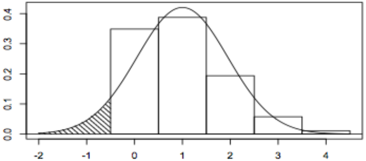
(a) $p = 0.5, n = 3$



(b) $p = 0.1, n = 3$



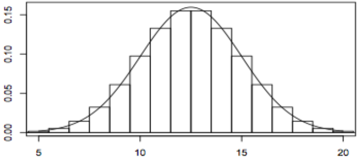
(c) $p = 0.5, n = 10$



(d) $p = 0.1, n = 10$

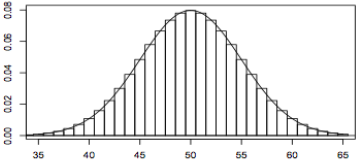
Si n est trop petit pour $p = 0.1$, l'approximation est mauvaise, car la probabilité que la variable gaussienne $\mathcal{N}(np, np(1-p))$ soit négative est trop importante pour bien représenter une somme de variables positives.

Approximation de la loi binômiale par la loi normale

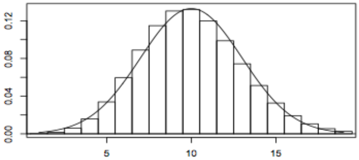


(e) $p = 0.5, n = 25$

(f) $p = 0.1, n = 25$



(g) $p = 0.5, n = 100$



(h) $p = 0.1, n = 100$

Si n est trop petit pour $p = 0.1$, l'approximation est mauvaise, car la probabilité que la variable gaussienne $\mathcal{N}(np, np(1-p))$ soit négative est trop importante pour bien représenter une somme de variables positives.

Exercice : the Aquarius Machine (Charles Tart, 1970s)

Certains humains peuvent ils prédire l'avenir ?

Exercice L'Aquarius machine possède 4 lampes qui s'allument dans un ordre aléatoire. Plusieurs personnes tentent de deviner quelle lampe va s'allumer à chaque fois.

- 15 personnes, 500 essais par personne = 7500 essais au total
- 2006 réponses correctes, 5494 erreurs

Ce résultat est-il étonnant ?

A chaque étape, on peut s'attendre à ce que chaque personne ait une probabilité $\frac{1}{4}$ de trouver la bonne réponse. Le nombre de réponses correctes X devrait donc suivre une loi binômiale

$$X \sim \mathcal{B}(7500, \frac{1}{4}).$$

On peut en déduire

$$\mathbb{P}[X \geq 2006] = \sum_{k=2006}^{7500} \binom{7500}{k} \frac{1}{4}^{k} \frac{1}{4}^{7500-k} = \ldots \simeq 0.000274$$

Exercice : the Aquarius Machine (Charles Tart, 1970s)

Cette loi peut aussi être approchée par une loi normale

Exercice : the Aquarius Machine (Charles Tart, 1970s)

Cette loi peut aussi être approchée par une loi normale

$$\mathcal{N}(\frac{7500}{4}, 7500 \times \frac{1}{4} \times \frac{3}{4})$$

Exercice : the Aquarius Machine (Charles Tart, 1970s)

Cette loi peut aussi être approchée par une loi normale

$$\mathcal{N}(\frac{7500}{4}, 7500 \times \frac{1}{4} \times \frac{3}{4})$$

On en déduit, en posant $Z = \frac{X - \frac{7500}{4}}{\sqrt{7500 \times \frac{1}{4} \times \frac{3}{4}}} = \frac{X - 1875}{1406.25}$

$$\mathbb{P}[X \geq 2006] \simeq \mathbb{P}[Z \geq \frac{2006 - 1875}{\sqrt{1406.25}}] \simeq \mathbb{P}[Z \geq 3.49] \simeq 2.4 \times 10^{-4}.$$

Intervalle de confiance lors d'élections

Exercice A et B sont les 2 candidats pour une élection. On veut estimer la probabilité qu'une personne vote pour A à l'aide d'un sondage sur n personnes. Pour estimer combien de personnes il faut interroger, calculer un intervalle de confiance dans lequel p doit se trouver avec une probabilité α.

Intervalle de confiance lors d'élections

Exercice A et B sont les 2 candidats pour une élection. On veut estimer la probabilité qu'une personne vote pour A à l'aide d'un sondage sur n personnes. Pour estimer combien de personnes il faut interroger, calculer un intervalle de confiance dans lequel p doit se trouver avec une probabilité α.

- Soit $X_i = 1$ si la personne i vote pour A et 0 sinon.
- Les X_i sont indépendantes et de même loi de Bernoulli de paramètre p. On a $\mathbb{E}[X_1] = p$ et $\mathrm{Var}[X_1] = p(1-p)$.

D'après le TCL, la loi de $\sqrt{\frac{n}{p(1-p)}}(\frac{\sum_{i=1}^n X_i}{n} - p)$ peut être approchée par la loi $\mathcal{N}(0,1)$. Donc, pour tout $\epsilon > 0$ on a

$$\mathbb{P}[\left|\sqrt{\frac{n}{p(1-p)}}(\frac{\sum_{i=1}^n X_i}{n} - p)\right| < \epsilon] \simeq \frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} e^{-\frac{t^2}{2}} dt.$$

Intervalle de confiance lors d'élections

Pour un niveau de confiance $0 \leq \alpha \leq 1$, on peut toujours trouver ϵ tel que $\frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} e^{-\frac{t^2}{2}} dt = \alpha$.
Par exemple, si on prend $\alpha = 0.95$, il faut choisir $\epsilon = 1,96$. La valeur de p a alors une probabilité α d'être dans l'intervalle

$$\left[\bar{X}_n - 1,96\sqrt{\frac{p(1-p)}{n}} \quad ; \quad \bar{X}_n + 1,96\sqrt{\frac{p(1-p)}{n}}\right].$$

En utilisant le fait que $p(1-p) \leq \frac{1}{4}$, on en déduit qu'il y a une probabilité de 95% pour que

$$p \in \left[\bar{X}_n - \frac{1,96}{2\sqrt{n}} \quad ; \quad \bar{X}_n + \frac{1,96}{2\sqrt{n}}\right].$$

Remarque : L'intervalle se resserre autour de p quand n croît. Si p est très proche de $\frac{1}{2}$ et qu'on veut prédire quel candidat va gagner, il faut prendre n suffisamment grand pour que $\frac{1}{2}$ n'appartienne pas à l'intervalle précédent.

Exercice

Exercice En octobre 2005, une étude portant sur 1012 américains choisis au hasard dans la population montrait que 30% d'entre eux possédaient une arme à feu. En déduire un intervalle de confiance à 95% sur la proportion d'américains possédant une arme à feu.