

Variables aléatoires continues et loi normale

H. Hocquard

HSE 2016-2017

université
de **BORDEAUX**



Définition

Une variable aléatoire est une application de l'univers Ω dans \mathbb{R}

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

Une variable aléatoire est généralement désignée par une lettre majuscule X , Y , etc. La variable aléatoire est dite continue si l'ensemble $X(\Omega)$ est un intervalle (ou une réunion d'intervalles) de \mathbb{R} .

Définition

Une variable aléatoire est une application de l'univers Ω dans \mathbb{R}

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

Une variable aléatoire est généralement désignée par une lettre majuscule X , Y , etc. La variable aléatoire est dite continue si l'ensemble $X(\Omega)$ est un intervalle (ou une réunion d'intervalles) de \mathbb{R} .

Exemple

X = taille d'un individu.

Problématique

La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X , la probabilité que X prenne une valeur bien précise x est nulle, $\mathbb{P}[X = x] = 0$. Il y a en effet une infinité de valeurs dans \mathbb{R} ou dans un intervalle, et au regard de toutes ces valeurs précises, le poids de la valeur particulière est tellement insignifiant qu'il en est nul !

Problématique

La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X , la probabilité que X prenne une valeur bien précise x est nulle, $\mathbb{P}[X = x] = 0$. Il y a en effet une infinité de valeurs dans \mathbb{R} ou dans un intervalle, et au regard de toutes ces valeurs précises, le poids de la valeur particulière est tellement insignifiant qu'il en est nul !

Par exemple, si $X = \text{taille d'un individu}$, alors

$$\mathbb{P}(X = 1,8245756) = 0$$

Problématique

La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X , la probabilité que X prenne une valeur bien précise x est nulle, $\mathbb{P}[X = x] = 0$. Il y a en effet une infinité de valeurs dans \mathbb{R} ou dans un intervalle, et au regard de toutes ces valeurs précises, le poids de la valeur particulière est tellement insignifiant qu'il en est nul !

Par exemple, si $X = \text{taille d'un individu}$, alors

$$\mathbb{P}(X = 1,8245756) = 0$$

Il n'est ainsi pas possible de définir la loi de X par la donnée des probabilités des événements élémentaires. Par contre, il est possible de déduire les probabilités que X prenne ses valeurs dans une partie de \mathbb{R} à partir de **la fonction de répartition** définie par :

$$F(x) = \mathbb{P}[X \leq x] = \mathbb{P}[X < x].$$

Propriétés

On a les propriétés suivantes :

- 1 F est une continue,
- 2 $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$,
- 3 F est une fonction croissante,
- 4 Pour tous $a, b \in \mathbb{R}$ et $a < b$,

$$F(b) - F(a) = \mathbb{P}[a < X \leq b].$$

Définition

Une variable aléatoire possède **une densité** si sa fonction de répartition F est dérivable. La dérivée notée f est appelée densité de probabilité de la variable aléatoire X .

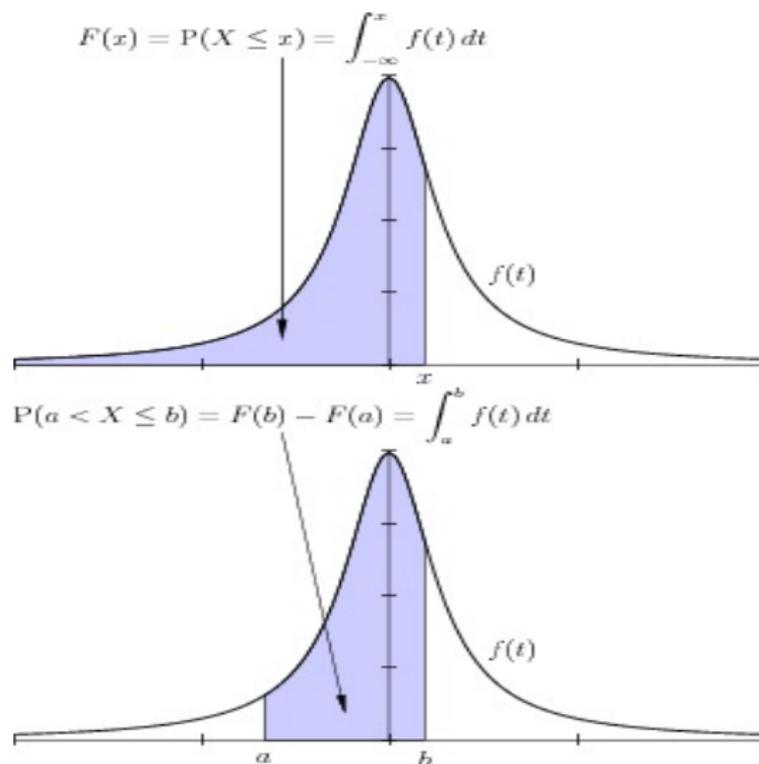
Propriétés

De ce fait,

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f(t)dt,$$

et la probabilité de trouver X dans un intervalle $[a, b]$ donné, apparaît comme l'aire d'une partie du graphique située entre la courbe de la densité f et l'axe des abscisses.

Représentation graphique



Propriétés

① $\forall x \in \mathbb{R}, f(x) \geq 0.$

②

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

③

$$\mathbb{P}[a < X \leq b] = F(b) - F(a) = \int_a^b f(x) dx.$$

Proposition

Soit X une variable aléatoire continue de Ω dans \mathbb{R} de densité f .
On calcule espérance et variance à l'aide des formules suivantes :

$$\mathbb{E}(X) = \int_{\mathbb{R}} t f(t) dt,$$

et

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

Proposition

Soit X une variable aléatoire continue de Ω dans \mathbb{R} de densité f .
On calcule espérance et variance à l'aide des formules suivantes :

$$\mathbb{E}(X) = \int_{\mathbb{R}} t f(t) dt,$$

et

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

Comparer ces formules avec le cas discret...

Cette loi modélise un phénomène uniforme sur un intervalle donné.

Définition

La v.a. X suit une loi uniforme sur l'intervalle borné $[a; b]$ si elle a une densité f constante sur cet intervalle et nulle en dehors. Elle est notée $\mathcal{U}([a; b])$. Sa densité est alors,

$$f(x) = \begin{cases} 1/(b - a) & \text{si } x \in [a; b], \\ 0 & \text{sinon} \end{cases}$$

Cette loi modélise un phénomène uniforme sur un intervalle donné.

Définition

La v.a. X suit une loi uniforme sur l'intervalle borné $[a; b]$ si elle a une densité f constante sur cet intervalle et nulle en dehors. Elle est notée $\mathcal{U}([a; b])$. Sa densité est alors,

$$f(x) = \begin{cases} 1/(b - a) & \text{si } x \in [a; b], \\ 0 & \text{sinon} \end{cases}$$

Cette loi est l'équivalent continu de la loi discrète équirépartie. Son espérance est $\mathbb{E}[X] = (b + a)/2$ et sa variance est $V(X) = (b - a)^2/12$.

Proposition

Si X est une v.a de loi uniforme sur $[a; b]$ alors pour tout intervalle \mathcal{I} de \mathbb{R} :

$$\mathbb{P}(X \in \mathcal{I}) = \frac{\ell([a; b] \cap \mathcal{I})}{\ell([a; b])},$$

où $\ell(J)$ désigne la longueur de l'intervalle J (ex : $\ell([a; b]) = b - a$).

Définition

Soit α un réel strictement positif. La v.a X suit une loi exponentielle de paramètre α , notée $\mathcal{E}(\alpha)$, si elle admet pour densité :

$$f(x) = \alpha e^{-\alpha x} \mathbf{1}_{[0; +\infty[}(x).$$

Définition

Soit α un réel strictement positif. La v.a X suit une loi exponentielle de paramètre α , notée $\mathcal{E}(\alpha)$, si elle admet pour densité :

$$f(x) = \alpha e^{-\alpha x} 1_{[0; +\infty[}(x).$$

Son espérance est $\mathbb{E}(X) = 1/\alpha$ et sa variance est $V(X) = 1/\alpha^2$.

Remarques

Les lois exponentielles sont souvent utilisées pour modéliser des **temps d'attente** ou des **durées de vie**.

Remarques

Les lois exponentielles sont souvent utilisées pour modéliser des **temps d'attente** ou des **durées de vie**.

Par exemple, les temps d'attente à partir de maintenant du prochain tremblement de terre, de la prochaine panne d'un appareil, de la prochaine désintégration dans un réacteur nucléaire suivent des lois exponentielles.

Remarques

Les lois exponentielles sont souvent utilisées pour modéliser des **temps d'attente** ou des **durées de vie**.

Par exemple, les temps d'attente à partir de maintenant du prochain tremblement de terre, de la prochaine panne d'un appareil, de la prochaine désintégration dans un réacteur nucléaire suivent des lois exponentielles.

Le paramètre α désigne alors l'inverse du temps d'attente moyen.

Introduction

La loi normale est apparue naturellement (d'où son nom) comme limite de certains processus.

Introduction

La loi normale est apparue naturellement (d'où son nom) comme limite de certains processus. C'est la loi la plus connue des probabilités, parfois sous le vocable **loi de Laplace-Gauss** et caractérisée par une célèbre "courbe en cloche".

Définition

La **loi normale centrée réduite** est une loi continue, d'une v.a. X à valeurs dans $X(\Omega) = \mathbb{R}$ tout entier, définie à partir de la densité

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

Définition

La **loi normale centrée réduite** est une loi continue, d'une v.a. X à valeurs dans $X(\Omega) = \mathbb{R}$ tout entier, définie à partir de la densité

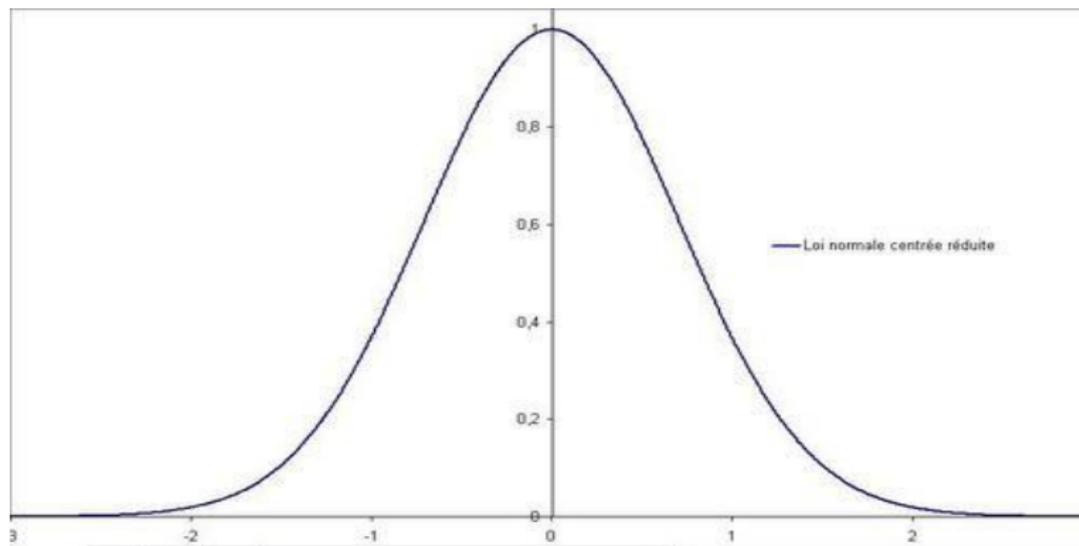
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Remarque

Il n'existe par contre pas d'expression simple de sa fonction de répartition autre que la formule intégrale

$$\forall a \in \mathbb{R}, F(a) = \int_{-\infty}^a f(t) dt$$

Allure de la densité normale centrée réduite



Remarque

Dans la pratique, les probabilités d'évènements de v.a. suivant une loi normale sont répertoriées dans des tables facilement manipulables.

Table de la loi normale centrée réduite

$$\Pi(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad \text{et} \quad \Pi(-t) = 1 - \Pi(t).$$

t	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	0.7704	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

Table de la loi normale centrée réduite

$$\Pi(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad \text{et} \quad \Pi(-t) = 1 - \Pi(t).$$

t	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	0.7704	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

On lit par exemple $\mathbb{P}(X \leq 0,64) = 0,7389$.

Espérance et variance

$$\begin{aligned}\mathbb{E}[X] &= 0, \\ V(X) &= 1.\end{aligned}$$

Espérance et variance

$$\begin{aligned}\mathbb{E}[X] &= 0, \\ V(X) &= 1.\end{aligned}$$

Définition

On dit que X suit une $\mathcal{N}(\mu, \sigma)$, si la densité est :

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

Définition

On dit que X suit une $\mathcal{N}(\mu, \sigma)$, si la densité est :

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

Remarque

L'usage d'un changement de variable $t = \frac{(x-\mu)}{\sigma}$ permet de se ramener à un calcul d'intégrale à partir de la loi $\mathcal{N}(0, 1)$, ce qui nous permettra de consulter les tables existant pour la loi standard précédente.

Définition

On dit que X suit une $\mathcal{N}(\mu, \sigma)$, si la densité est :

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

Remarque

L'usage d'un changement de variable $t = \frac{(x-\mu)}{\sigma}$ permet de se ramener à un calcul d'intégrale à partir de la loi $\mathcal{N}(0, 1)$, ce qui nous permettra de consulter les tables existant pour la loi standard précédente. On a le théorème suivant :

Théorème

Soit X une variable aléatoire de loi normale $\mathcal{N}(\mu, \sigma)$ et Z la variable aléatoire définie par

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

alors Z suit une loi normale centrée réduite $\mathcal{N}(0, 1)$.

Théorème

Soit X une variable aléatoire de loi normale $\mathcal{N}(\mu, \sigma)$ et Z la variable aléatoire définie par

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

alors Z suit une loi normale centrée réduite $\mathcal{N}(0, 1)$.

Attention



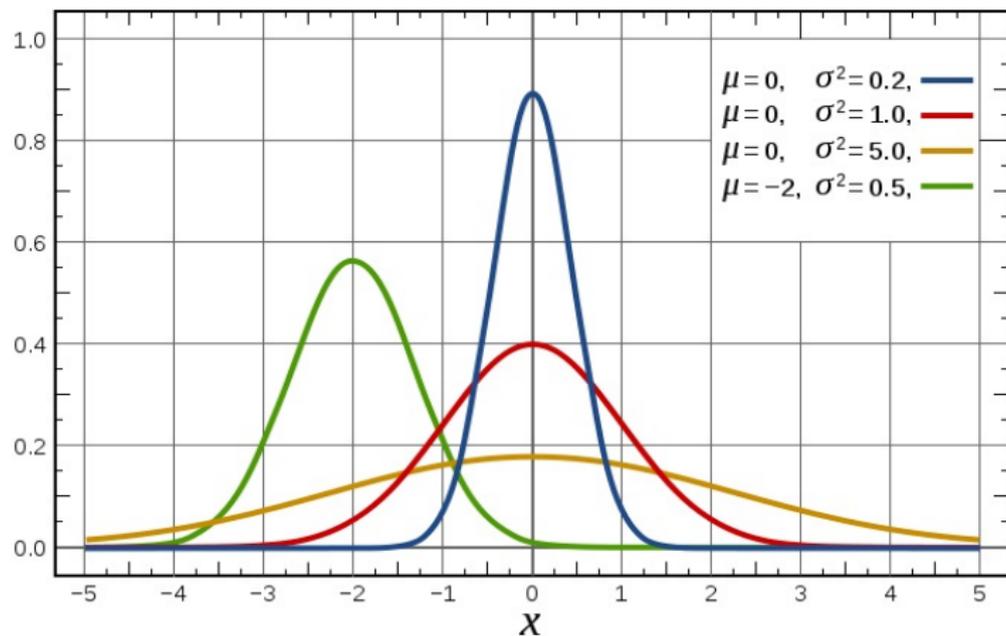
Certains auteurs utilisent la notation $\mathcal{N}(\mu, \sigma^2)$ et pas $\mathcal{N}(\mu, \sigma)$.

On a également :

Espérance et variance

$$\begin{aligned}\mathbb{E}[X] &= \mu, \\ V(X) &= \sigma^2.\end{aligned}$$

Allure de la densité en fonction de μ et σ



Exemple

Considérons X une v. a. qui suit une loi $\mathcal{N}(6, 2)$

Exemple

Considérons X une v. a. qui suit une loi $\mathcal{N}(6, 2)$
($\sigma(X)$ vaut donc ici 2 et $\mathbb{E}(X) = 6$)

Exemple

Considérons X une v. a. qui suit une loi $\mathcal{N}(6, 2)$
($\sigma(X)$ vaut donc ici 2 et $\mathbb{E}(X) = 6$)
Et soit Z une v.a. de loi $\mathcal{N}(0, 1)$,

Exemple

Considérons X une v. a. qui suit une loi $\mathcal{N}(6, 2)$
($\sigma(X)$ vaut donc ici 2 et $\mathbb{E}(X) = 6$)

Et soit Z une v.a. de loi $\mathcal{N}(0, 1)$, on a par exemple

$$\begin{aligned}\mathbb{P}[X \leq 7] &= \mathbb{P}\left[\frac{X - 6}{2} \leq \frac{7 - 6}{2}\right] \\ &= \mathbb{P}\left[Z \leq \frac{1}{2}\right] \\ &= 0.6915.\end{aligned}$$

Concentration autour de la moyenne

Dans l'intervalle $[m - \sigma, m + \sigma]$ de longueur 2σ et centré autour de la moyenne, on peut calculer qu'il y a 68% des individus, lorsque qu'une v.a. suit une loi $\mathcal{N}(m, \sigma)$:

$$\mathbb{P}[m - \sigma \leq X \leq m + \sigma] = 0.68$$

Concentration autour de la moyenne

Dans l'intervalle $[m - \sigma, m + \sigma]$ de longueur 2σ et centré autour de la moyenne, on peut calculer qu'il y a 68% des individus, lorsque qu'une v.a. suit une loi $\mathcal{N}(m, \sigma)$:

$$\mathbb{P}[m - \sigma \leq X \leq m + \sigma] = 0.68$$

On établit aussi que 95% d'un échantillon représentatif d'une loi normale $\mathcal{N}(m, \sigma)$ est approximativement situé entre $m - 2\sigma$ et $m + 2\sigma$. Plus exactement,

$$\mathbb{P}[m - 1.96\sigma \leq X \leq m + 1.96\sigma] = 0.95$$

Concentration autour de la moyenne

Dans l'intervalle $[m - \sigma, m + \sigma]$ de longueur 2σ et centré autour de la moyenne, on peut calculer qu'il y a 68% des individus, lorsque qu'une v.a. suit une loi $\mathcal{N}(m, \sigma)$:

$$\mathbb{P}[m - \sigma \leq X \leq m + \sigma] = 0.68$$

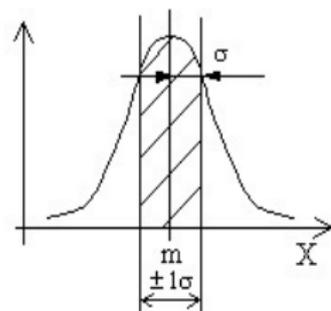
On établit aussi que 95% d'un échantillon représentatif d'une loi normale $\mathcal{N}(m, \sigma)$ est approximativement situé entre $m - 2\sigma$ et $m + 2\sigma$. Plus exactement,

$$\mathbb{P}[m - 1.96\sigma \leq X \leq m + 1.96\sigma] = 0.95$$

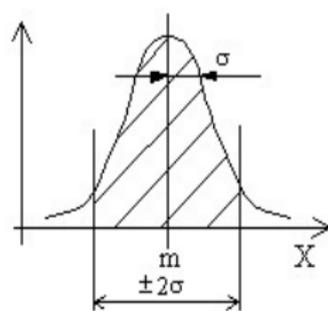
et on a même 99,7% des individus entre $m - 3\sigma$ et $m + 3\sigma$:

$$\mathbb{P}[m - 3\sigma \leq X \leq m + 3\sigma] = 0.997$$

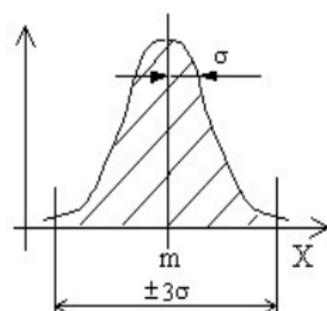
Concentration autour de la moyenne



environ 68% des individus
sont compris dans l'intervalle
 $m \pm 1\sigma$

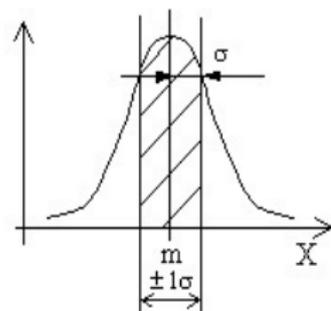


environ 95% des individus
sont compris dans l'intervalle
 $m \pm 2\sigma$

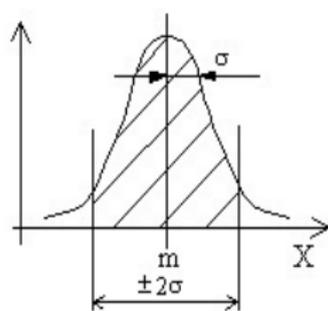


environ 99,8% des individus
sont compris dans l'intervalle
 $m \pm 3\sigma$

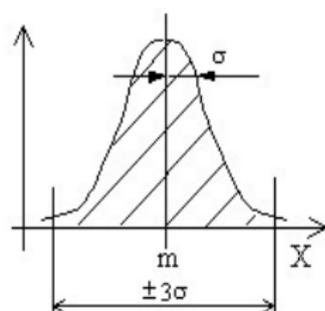
Concentration autour de la moyenne



environ 68% des individus
sont compris dans l'intervalle
 $m \pm 1\sigma$



environ 95% des individus
sont compris dans l'intervalle
 $m \pm 2\sigma$



environ 99,8% des individus
sont compris dans l'intervalle
 $m \pm 3\sigma$

Autrement dit, lorsque l'on a une variable aléatoire qui suit une loi normale $\mathcal{N}(m, \sigma)$, on est "pratiquement sûr" que la valeur se situera entre $m - 3\sigma$ et $m + 3\sigma$.

Théorème

Soit S_n binomiale $\mathcal{B}(n; p)$ et $U \sim \mathcal{N}(0; 1)$. On a :

$$\frac{S_n - np}{\sqrt{npq}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} U,$$

qui peut également s'écrire

$$S_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(np; \sqrt{npq}).$$

Dans la pratique, on considère que l'approximation est bonne lorsque $n \geq 30$, $p \geq 0.1$ et $n \cdot p > 15$.

Remarques

- Cette propriété est une conséquence du Théorème central limite que l'on abordera au chapitre suivant.

Remarques

- Cette propriété est une conséquence du Théorème central limite que l'on abordera au chapitre suivant.
- Ne pas confondre l'approximation de la loi de poisson par une binomiale et celle de la loi normale!!!

Définition

Soient X_1, \dots, X_n des v.a indépendantes de même loi $\mathcal{N}(0, 1)$.

Posons

$$Z = \sum_{i=1 \dots n} X_i^2,$$

par définition la v.a. Z suit une loi du khi-deux à n degré(s) de liberté (abréviation d.d.l.). On la note $\chi^2(n)$.

Définition

Soient X_1, \dots, X_n des v.a indépendantes de même loi $\mathcal{N}(0, 1)$.

Posons

$$Z = \sum_{i=1 \dots n} X_i^2,$$

par définition la v.a. Z suit une loi du khi-deux à n degré(s) de liberté (abréviation d.d.l.). On la note $\chi^2(n)$.

Propriétés

- $Z \geq 0$, cette loi n'est donc pas symétrique,
- Z admet une densité,
- $\mathbb{E}(Z) = n$ et $V(Z) = 2n$.

Allure de la densité d'un χ^2

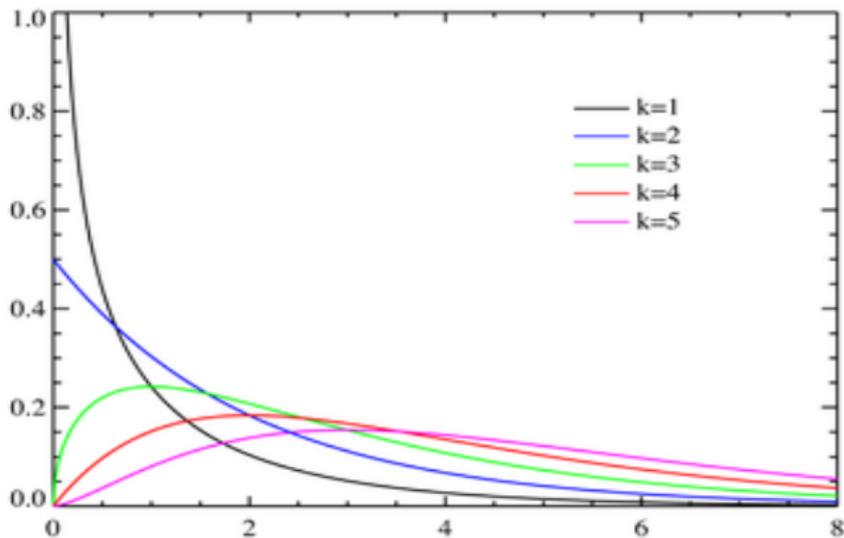


FIGURE – Densité de la loi $\chi^2(k)$.

Définition

Soient $X \sim \mathcal{N}(0, 1)$ et $Y \sim \chi^2(n)$. Posons $T = \frac{X}{\sqrt{Y/n}}$. Alors T suit une loi de Student à n degré(s) de liberté et on la note $\mathcal{T}(n)$ ou *Student*(n).

Allure de la densité de Student

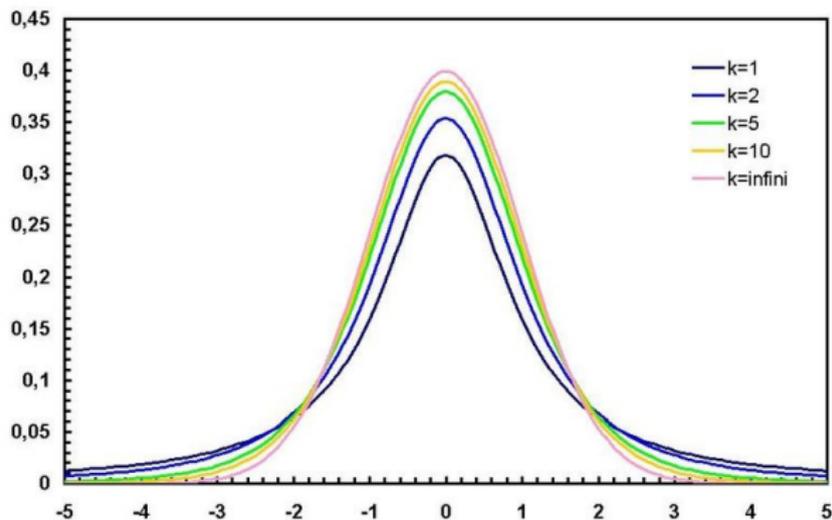


FIGURE – Densité de la loi de $Student(n)$.

Définition

Soient X et Y deux variables aléatoires indépendantes telles que $X \sim \chi^2(n)$ et $Y \sim \chi^2(m)$. Alors, on dit que la variable

$$Z = \frac{\frac{X}{n}}{\frac{Y}{m}}$$

suit une loi de Fisher-Snedecor(n, m). On la note $\mathcal{F}(n, m)$.

Allure de la densité de Fisher-Snedecor

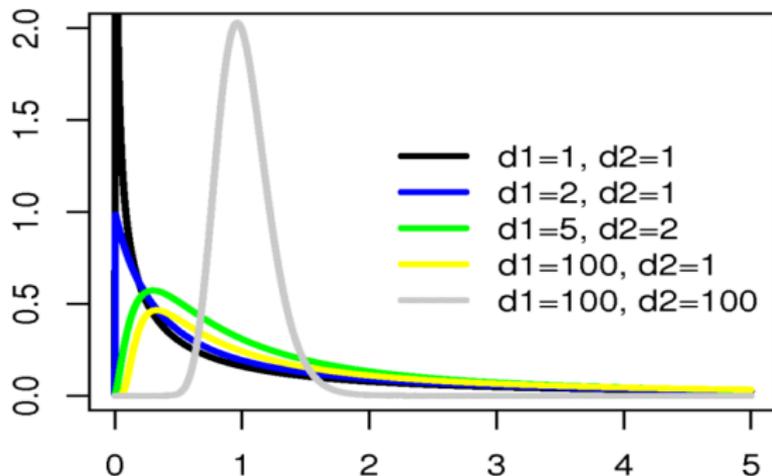


FIGURE – Densité de la loi de Fisher-Snedecor $\mathcal{F}(d_1, d_2)$.

Remarques

Ces trois dernières lois seront utiles dans la théorie des tests.

Remarques

Ces trois dernières lois seront utiles dans la théorie des tests. L'expression explicite des densités de ces lois n'est pas à connaître (sauf pour la loi normale). Des tables statistiques et des logiciels permettent de les manipuler.



Walter Apple

Probabilités pour les non-probabilistes

H&K, édition, 2013



Clément Rau

<http://www.math.univ-toulouse.fr/rau/>

Communication privée, 2015