MATH-H204 - Calcul des probabilités et statistiques
Yves De Smet
Résumé du cours

Rodrigue Van Brande

Dylan Gonzalez

24 juillet 2015

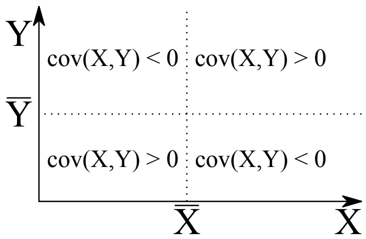
Table des matières

1 La statistique descriptive

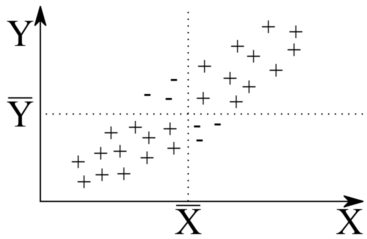
1.1 Statistique descriptive en 1D

1.2 Statistique descriptive en 2D

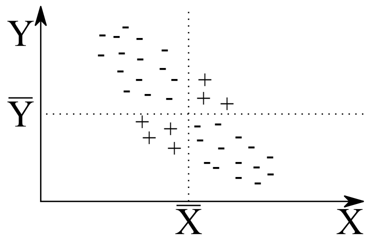
1.2.1 Covariance



Allure générale



$m_{11} > 0$



$m_{11} < 0$

La covariance $|m_{11}| \leq s_1 s_2$

La covariance est le moment d'ordre (1,1) :

$$m_{11} = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij} (x_i - \bar{x})(y_j - \bar{y})$$

$$\begin{aligned}
0 &\leq \alpha \\
&\leq \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij} \underbrace{(u(x_i - \bar{x})(y_j - \bar{y}))^2}_{^2 \text{ car toujours } \geq 0} \\
&\leq \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{q} n_{ij} (u^2 (x_i - \bar{x})^2 + 2u(x_i - \bar{x})(y_j - \bar{y}) + (y_j - \bar{y})^2) \\
&\leq u^2 s_1^2 + 2u\ m_{11} + s_2^2
\end{aligned}$$

Équation du second degré, on calcule son Δ :

$$\begin{aligned}
\Delta &\leq 0 \\
(2m_{11})^4 - 4 s_1^2 s_2^2 &\leq 0 \\
m_{11}^2 - s_1^2 s_2^2 &\leq 0 \\
m_{11}^2 &\leq s_1^2 s_2^2
\end{aligned}$$

$$\boxed{|m_{11}| \leq s_1 s_2}$$

La covariance maximale $|m_{11}| = s_1 s_2$

La valeur absolue de la covariance est maximale et vaut $|m_{11}| = s_1 s_2$.
Si les points observés se trouvent sur une droite $ax + bx + c = 0$, on a $ax_i + by_i + c = 0$.
On multiplie par $\frac{n_{ij}}{n}$ et on somme sur ij.

$$
\begin{aligned}
0 &= \sum_{i=1}^{p}\sum_{j=1}^{q}\frac{n_{ij}}{n}(ax_i + by_j + c) \\
&= a\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}x_i + b\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}y_j + c\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij} \\
&= a\bar{x} + b\bar{y} + c
\end{aligned}
$$

On soustrait $ax_i + by_j + c = 0$ par $a\bar{x} + b\bar{y} + c = 0$.

$$= a(x_i - \bar{x}) + b(y_j + \bar{y})$$

On utilise $u_0 = \frac{a}{b}$

$$
\begin{aligned}
&= u_0 b(x_i - \bar{x}) + \frac{a}{u_0}(y_j - \bar{y}) \\
&= u_0 b(x_i - \bar{x}) + \frac{u_0 b}{u_0}(y_j - \bar{y}) \\
&= u_0(x_i - \bar{x}) + (y_j - \bar{y})
\end{aligned}
$$

L'équation a la même forme que α, du coup...

$$
\begin{aligned}
0 &= \Delta \\
&= m_{11}^2 - s_1^2 s_2^2 \\
m_{11}^2 &= s_1^2 s_2^2
\end{aligned}
$$

$$|m_{11}| = s_1 s_2$$

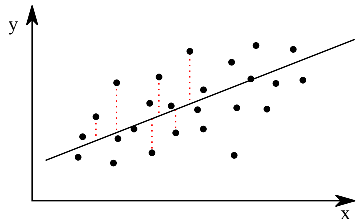
1.2.2 Le coefficient de corrélation

$$\boxed{r = \frac{m_{11}}{s_1 s_2}}$$

Propriétés

1. r sans dimensions ;
2. $r' = r$;
3. $-1 \leq r \leq 1$;
4. $|r| = 1$ car les points obervés se trouvent sur une droite non parall'ele aux axes.

1.2.3 Les droites de régression



La droite de régression de y en x est la droite qui minimise la somme des carrés des écarts des points observés à cette droite, les écarts étant pris parall'element à l'axe y. C'est donc la droite d'équation $y = ax + b$ qui minimise la quantité.

$$\boxed{g(a,b) = \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}(y_j - a\ x_i - b)^2}$$

Dérivée par rapport à a :

$$\begin{aligned}
0 &= g(a,b)|_a \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ 2(y_j - a\ x_i - b)(-x_i) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} -2n_{ij}\ x_i(y_j - a\ x_i - b) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i(-y_j + a\ x_i + b) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} -n_{ij}\ x_i\ y_j + \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ a\ x_i^2 + \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ b\ x_i \\
&= -1\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i\ y_j + a\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i^2 + b\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i \\
\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i\ y_j &= \underbrace{a\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i^2 + b\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i}_{\substack{\text{Il n'existe aucun terme en } j \\ \text{Donc on peut supprimer la somme } \sum_{j=1}^{q}}} \\
\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i\ y_j &= \overbrace{a\sum_{i=1}^{p} n_{i.}\ x_i^2 + b\sum_{i=1}^{p} n_{i.}\ x_i}
\end{aligned}$$

$$\boxed{\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i\ y_j = a\sum_{i=1}^{p} n_{i.}\ x_i^2 + b\sum_{i=1}^{p} n_{i.}\ x_i}$$

Dérivée par rapport à b :

$$\begin{aligned}
0 &= g(a,b)|_b \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ 2(y_j - a\ x_i - b)(-1) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} -2n_{ij}(y_j - a\ x_i - b) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}(-y_j + a\ x_i + b) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{q} -n_{ij}\ y_j + \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ a\ x_i + \sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ b \\
&= -1\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ y_j + a\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i + b\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij} \\
\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ y_j &= a\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}\ x_i + b\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij} \\
n\ \bar{y} &= a\ n\ \bar{x} + b\ n
\end{aligned}$$

$$\boxed{n\ \bar{y} = a\ n\ \bar{x} + b\ n}$$

On a obtenu ces deux réponses :

$$\begin{cases}
\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}x_iy_j = a\sum_{i=1}^{p} n_{i.}\ x_i^2 + b\sum_{i=1}^{p} n_{i.}\ x_i & (1) \\
n\bar{y} = a\ n\ \bar{x} + b\ n & (2)
\end{cases}$$

On soustrait le (1) par le double du (2) :

$$\begin{aligned}
\bar{x}.(2):\quad & n\bar{y}\bar{x} &=& \ a n\bar{x}^2 + bn\bar{x} \\
(1) - \bar{x}.(2):\quad & \left(\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}x_iy_j\right) - (n\bar{y}\bar{x}) &=& \left(a\sum_{i=1}^{p} -n_{i.}x_i^2 + b\sum_{i=1}^{p} n_{i.}x_i\right) - \left(a n\bar{x}^2 + bn\bar{x}\right) \\
& &\vdots&
\end{aligned}$$

Pour obtenir à la fin :

$$\boxed{a = \frac{m_{11}}{s_1^2}} \text{ et } \boxed{b = \bar{y} - \frac{m_{11}}{s_1^2}\bar{x}}$$

On remplace dans une droite :

$$\begin{aligned}
y &= ax + b \\
y &= \frac{m_{11}}{s_1^2}x + \bar{y} - \frac{m_{11}}{s_1^2}\bar{x}
\end{aligned}$$

Pour obtenir la droite de regression de y en x :

$$\boxed{y = \frac{m_{11}}{s_1^2}(x - \bar{x}) + \bar{y}}$$

Le raisonnement est symétrique pour le cas de la régression de x en y.

$$\boxed{x = \frac{m_{11}}{s_1^2}(y - \bar{y}) + \bar{x}}$$

1.2.4 Variances résiduelles

La variance résiduelle de y en x est :

$$\boxed{s_{21}^2 = s_2^2(1-r^2)}$$

Propriété et interprétation

- $s_{21}^2 = 0$ ssi $r = \pm 1$ ssi tous les points observés sont sur une droite ;
- $s_{21}^2 = s_2^2$ ssi $r = 0$ ssi les droites de régression sont parallèles aux axes ;
- $s_2^2 r^2$ représente une certaine proportion de s_2^2, d'autant plus grande que la dépendance linéaire entre x et y est forte : on peut donc la considérer comme la part de la variance de y qui est expliquée par le lien linéaire entre x et y, tandis que la variance résiduelle s_{21}^2 est la part de la variance de y qui n'est pas expliquée par ce lien linéaire (d'où son nom).

Démonstration

Nous partons de la variance de y en x :

$$\begin{aligned}
s_{21}^2 &= \underbrace{\frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}\left(y_i - ax_i - b\right)^2}_{\text{valeur minimum de g(a,b)}} \\
&\boxed{a = \frac{m_{11}}{s_1^2}} \text{ et } \boxed{b = \bar{y} - \frac{m_{11}}{s_1^2}\bar{x}} \\
&= \frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}\left(y_i - \frac{m_{11}}{s_1^2}x_i - \bar{y} + \frac{m_{11}}{s_1^2}\bar{x}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}\left((y_i - \bar{y}) - \frac{m_{11}}{s_1^2}(x_i - \bar{x})\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}(y_i-\bar{y})^2 - 2\frac{m_{11}}{s_1^2}(x_i-\bar{x})(y_i-\bar{y}) + \frac{m_{11^2}}{s_1^4}(x_i-\bar{x})^2 \\
&= \frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}(y_i-\bar{y})^2 - 2\frac{m_{11}}{s_1^2}\frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} + \frac{m_{11^2}}{s_1^4}\frac{1}{n}\sum_{i=1}^{p}\sum_{i=1}^{q} n_{ij}(x_i-\bar{x})^2 \\
&= s_2^2 - 2\frac{m_{11}}{s_1^2}m_{11} + \frac{m_{11^2}}{s_1^4}s_1^2 \\
&= s_2^2 - 2\frac{m_{11}^2}{s_1^2} + \frac{m_{11^2}}{s_1^2} \\
&= s_2^2 - \frac{m_{11}^2}{s_1^2}
\end{aligned}$$

On utilise le coefficient de corrélation $r = \dfrac{m_{11}}{s_1 s_2}$, donc $\dfrac{m_{11}}{s_1} = rs_2$:

$$\begin{aligned}
s_2^2 - \frac{m_{11}^2}{s_1^2} &= s_2^2 - (rs_2)^2 \\
&= s_2^2(1-r^2)
\end{aligned}$$

$$\boxed{s_{21}^2 = s_2^2(1-r^2)}$$

La variance résiduelle de x en y est similaire.

$$\boxed{s_{12}^2 = s_1^2(1-r^2)}$$

2 La théorie des probabilités

2.1 Probabilités

2.1.1 Axiomes de la théorie des probabilités

$$\begin{cases} P(E) & = 1 \\ P(A) & \geq 0 \qquad\qquad\qquad\qquad \forall A \subset E \\ A \cap B & = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B) \end{cases}$$

2.1.2 Probabilité indépendance

Si A et B sont indépendants alors
- A et $\bar{B}$ sont indépendants;
- $\bar{A}$ et $\bar{B}$ sont indépendants;
- $\bar{A}$ et B sont indépendants.

2.1.3 Probabilité conditionnelle et indépendance

Probabilité conditionnelle de A sous la condition B ("sachant B") :

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

Si A est indépendant de B : $\begin{cases} P(A \cap B) & = P(A)P(B) \\ P(A|B) & = P(A) \\ P(B|A) & = P(B) \end{cases}$

Une des propriétés des 3 suffit.

Démonstration

La probabilité conditionnelle respecte les axiomes de la théorie des probabilités

$$\begin{array}{ccc} P(E|A) & = & \frac{P(E \cap A)}{P(A)} = \frac{P(A)}{P(A)} & = 1 \\ P(B|A) & = & \frac{P(B \cap A)}{P(A)} & \geq 0 \\ X \cap Y \neq \emptyset & \Rightarrow & P(X \cup Y|A) \\ & & \frac{P((X \cup Y) \cap A)}{P(A)} \\ & & \frac{P((X \cap A) \cup (Y \cap A))}{P(A)} \\ & & \frac{P(X \cap A) + (Y \cap A)}{P(A)} \\ & & \frac{P(X|A)P(A) + P(Y|A)P(A)}{P(A)} \\ & & P(X|A) + P(Y|A) \end{array}$$

2.1.4 Probabilité conditionnelle inverse et indépendant

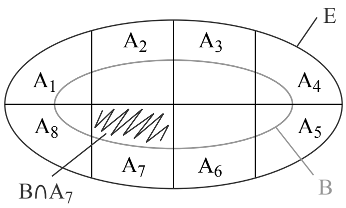
$$P(A|B) = P(A|\bar{B})$$

Démonstration

$$\begin{aligned}
P(A|B) &= P(A|\bar{B}) \\
\frac{P(A \cap B}{P(B))} &= \frac{P(A \cap \bar{B})}{P(\bar{B})} \\
\frac{P(A)P(B)}{P(B)} &= \frac{P(A)P(B)}{P(\bar{B})} \\
P(A) &= P(A)
\end{aligned}$$

2.1.5 Formule de Bayes

Cette formule est utilisée dans le cas où un évènement B peut survenir à cause d'évènement A_i incompatibles. Par exemple : une pièce défectueuse fabriquée par plusieurs machines différentes.



$$\begin{aligned}
B &= (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_m \cap B) \\
P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_m \cap B) \\
&= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_m)P(A_m)
\end{aligned}$$

$$\boxed{P(A_k|B) = \frac{A_k \cap B}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_{j=1}^{m} P(B|A_j)P(A_j)}}$$

2.2 Variables aléatoires

2.2.1 Distribution d'une fonction monotone d'une variable aléatoire

$$\boxed{W = G(V)}$$

Cas discret

$$\begin{aligned} F_W(x) &= P(W \leq x) \\ &= P(G(V) \leq x) \\ &= P(V \leq G^{-1}(x)) \\ &= F_V(G^{-1}(x) \end{aligned}$$

Cas continu

$$\begin{array}{rcc|c} & & \text{croissante} & \text{décroissante} \\ f_W(x) & = & \multicolumn{2}{c}{F'_W(x)} \\ & = & \multicolumn{2}{c}{(F_W(x))'} \\ & = & \multicolumn{2}{c}{(P(W \leq x))'} \\ & = & \multicolumn{2}{c}{(P(G(V) \leq x))'} \\ & = & (P(V \leq G^{-1}(x)))' & (P(V \geq G^{-1}(x)))' \\ & = & (F_V(G^{-1}(x)))' & (1 - F_V(G^{-1}(x)))' \\ & = & f_V(G^{-1}(x))\dfrac{1}{G'(G^{-1}(x))} & f_V(G^{-1}(x))\dfrac{-1}{G'(G^{-1}(x))} \end{array}$$

2.2.2 Distribution de la somme de deux variables aléatoires

$$\boxed{Z = V + W}$$

Cas discret

$$\begin{aligned} F_Z(x) &= \sum_i \sum_j p_{ij} \\ &= \sum_i \sum_j P(V \leq v_i, W \leq w_i) \end{aligned}$$

Cas continu

$$\begin{aligned} F_Z(x) &= P(Z \leq x) \\ F_{V+W}(x) &= P(V \leq x, W \leq x) \\ &= \iint\limits_{\xi+\eta \leq x} f_{(V,W)}(\xi, \eta)\ d\xi\ d\eta \end{aligned}$$

On remplace par $\begin{cases} \xi & = u \\ \eta & = v - u \\ J & = \begin{pmatrix} \dfrac{\delta\xi}{\delta u} & \dfrac{\delta\xi}{\delta v} \\ \dfrac{\delta\eta}{\delta u} & \dfrac{\delta\eta}{\delta v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.1 - 0.(-1) = 1 \end{cases}$

$$\begin{aligned} &= \iint\limits_{v \leq x} f_{(V,W)}(u, v-u)\, |1|\ du\ dv \\ &= \int_{-\infty}^{x} dv \int_{-\infty}^{+\infty} f_{(V,W)}(u, v-u)\ du \end{aligned}$$

$$\boxed{f_Z(x) = \frac{\delta F_z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_{(V,W)}(u, x-u)\ du}$$

ou si indépendant

$$\boxed{f_Z(x) = \frac{\delta F_z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_V\,(u) \,.\, f_W\,(x-u)\ du}$$

2.2.3 Distribution du produit de deux variables aléatoires

$$\boxed{Z = VW}$$

Cas continu

$$\begin{aligned} F_Z(x) &= P(Z \leq x) \\ F_{V.W}(x) &= P(V \leq x, W \leq x) \\ &= \iint\limits_{\xi.\eta \leq x} f_{(V,W)}(\xi, \eta)\ d\xi\ d\eta \end{aligned}$$

On remplace par $\begin{cases} \xi &= u \\ \eta &= \dfrac{v}{u} \\ J &= \begin{pmatrix} \dfrac{\delta\xi}{\delta u} & \dfrac{\delta\xi}{\delta v} \\ \dfrac{\delta\eta}{\delta u} & \dfrac{\delta\eta}{\delta v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{-1}{u^2} & \dfrac{1}{u} \end{pmatrix} = 1.\dfrac{1}{u} - 0.\dfrac{-1}{u^2} = \dfrac{1}{u} \end{cases}$

$$\begin{aligned} &= \iint\limits_{v \leq x} f_{(V,W)}(u, v-u)\ du\ dv \\ &= \int\limits_{-\infty}^{x} dv \int\limits_{-\infty}^{+\infty} f_{(V,W)}(u, v-u)\ du \end{aligned}$$

$$\boxed{f_Z(x) = \frac{\delta F_z(x)}{\delta x} = \int\limits_{-\infty}^{+\infty} f_{(V,W)}\left(u, \frac{x}{u}\right) . \left|\frac{1}{u}\right|\ du}$$

ou si indépendant

$$\boxed{f_Z(x) = \frac{\delta F_z(x)}{\delta x} = \int\limits_{-\infty}^{+\infty} f_V\left(u\right) . f_W\left(\frac{x}{u}\right) . \left|\frac{1}{u}\right|\ du}$$

2.2.4 Distribution de la somme des espérances

$$\boxed{Z = V + W}$$

Demonstration

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} x\ f_{(V+W)}(x)\ dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ f_{(V,W)}(u, x-u)\ du\ dv
\end{aligned}$$

On remplace par $\begin{cases} u &= \xi \\ x &= \xi - \eta \\ J &= \begin{pmatrix} \dfrac{\delta u}{\delta \xi} & \dfrac{\delta u}{\delta \eta} \\ \dfrac{\delta x}{\delta \xi} & \dfrac{\delta x}{\delta \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = 1. -1 - 0.1 = -1 \end{cases}$

$$\begin{aligned}
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\xi + \eta)\ f_{(V,W)}(\xi, \eta)\, |-1|\ \ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \xi\ f_{(V,W)}(\xi, \eta)\ d\xi\ d\eta + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \eta\ f_{(V,W)}(\xi, \eta)\ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty} \xi\ f_V(\xi)\ d\xi\ + \int_{-\infty}^{\infty} \eta\ f_W(\eta)\ d\eta \\
&= E(V) + E(W)
\end{aligned}$$

2.2.5 Distribution du produit des espérances

$$\boxed{Z = V.W}$$

Demonstration

Uniquement pour des variables indépendantes

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} x \ f_{(V.W)}(x) \ dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \ f_V(u) f_w\left(\frac{x}{u}\right) \left|\frac{1}{u}\right| \ du \ dv
\end{aligned}$$

On remplace par $\begin{cases} u &= \xi \\ x &= \xi\eta \\ J &= \begin{pmatrix} \frac{\delta u}{\delta \xi} & \frac{\delta u}{\delta \eta} \\ \frac{\delta x}{\delta \xi} & \frac{\delta x}{\delta \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \eta & \xi \end{pmatrix} = 1.\xi - 0.\eta = \xi \end{cases}$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\xi\eta) \ f_{(V,W)}(\xi, \eta) \, |\xi| \ d\xi \ d\eta \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi \ f_V(\xi) f_W(\eta) \ d\xi \ d\eta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta \ f_{(V,W)}(\xi, \eta) \ d\xi \ d\eta \\
&= \int_{-\infty}^{\infty} \xi \ f_V(\xi) \ d\xi \ \int_{-\infty}^{\infty} \eta \ f_W(\eta) \ d\eta \\
&= E(V).E(W)
\end{aligned}$$

2.3 Variables aléatoires particulières

2.3.1 Variable binomiale$\mathcal{B}(n,p)$

2.3.2 Variable de Poisson $\mathcal{P}_\lambda$

2.3.3 Variable exponentielle négative

2.3.4 Variable Normale $\mathcal{N}(\mu,\sigma)$

Propriété d'inverse

Soit $\bar{X}_1$ et $\bar{X}_2$ deux variables normales. Si $\bar{X}_2$ est l'inverse de $\bar{X}_1$ ($\bar{X}_2$ = -$\bar{X}_1$) alors

$$\boxed{\bar{X}_1 \sim N\left(\mu_1,\sigma\right) \text{ et } \bar{X}_2 \sim N\left(-\mu_2,\sigma\right)}$$

Propriété d'addition

Soit $N(\mu_1,\sigma_1)$ et $N(\mu_2,\sigma_2)$ deux variables normales indépendantes et $\psi_1(t)$ et $\psi_2(t)$ leurs fonctions génératrices.

$$\begin{aligned}\psi(t) &= \psi_1(t).\psi_2(t)\\ &= e^{\mu_1+t+\sigma_1^2t^2/2}.e^{\mu_2+t+\sigma_2^2t^2\frac{1}{2}}\\ &= e^{(\mu_1+\mu_2)t+\frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}\end{aligned}$$

On obtient bien une fonction génératrice d'une normale de paramètres $(\mu_1+\mu_2)$ et $\sqrt{\sigma_1^2+\sigma_2^2}$.

$$\boxed{N(\mu_1,\sigma_1)+N(\mu_2,\sigma_2) = N\left(\mu_1+\mu_2,\sqrt{\sigma_1^2+\sigma_2^2}\right)}$$

2.3.5 Variable Khi^2

2.3.6 Variable Student t_n

2.3.7 Variable Snedecor $\mathcal{F}_{(m,n)}$

2.4 Théorèmes fondamentaux

2.4.1 Inégalité de Bienaymé-Tchebycheff

La proportion d'individus s'écartant de la moyenne d'une distribution de plus k fois l'écart-type (σ) ne dépasse jamais $\frac{1}{k^2}$:

$$\boxed{\frac{1}{k^2} \geq P(|V-\mu| \geq k\sigma)}$$

Démonstration

Cas continu

$$\begin{aligned}
\sigma^2 &= \int_{-\infty}^{\infty} (x-\mu)^2 \ f(x) \ dx \\
&= \underbrace{\int_{-\infty}^{\mu-k\sigma}}_{\substack{x \leq \mu - k\sigma \\ x-\mu \leq -k\sigma \\ (x-\mu)^2 \geq (k\sigma)^2}} (x-\mu)^2 \ f(x) \ dx + \underbrace{\int_{\mu-k\sigma}^{\mu+k\sigma}}_{\geq 0} (x-\mu)^2 \ f(x) \ dx + \underbrace{\int_{\mu+k\sigma}^{\infty}}_{\substack{x \geq \mu + k\sigma \\ x-\mu \geq k\sigma \\ (x-\mu)^2 \geq (k\sigma)^2}} (x-\mu)^2 \ f(x) \ dx \\
\sigma^2 &\geq \int_{-\infty}^{\mu-k\sigma} (k\sigma)^2 \ f(x) \ dx + \int_{\mu+k\sigma}^{\infty} (k\sigma)^2 \ f(x) \ dx \\
&\geq (k\sigma)^2 \int_{-\infty}^{\mu-k\sigma} f(x) \ dx + (k\sigma)^2 \int_{\mu+k\sigma}^{\infty} f(x) \ dx \\
\frac{1}{k^2} &\geq P(V-\mu \leq -k\sigma) + P(V-\mu \geq k\sigma) \\
&\geq P(|V-\mu| \geq k\sigma) + P(|V-\mu| \geq k\sigma)
\end{aligned}$$

$$\boxed{\frac{1}{k^2} \geq P(|V-\mu| \geq k\sigma)}$$

Cas discret

$$\begin{aligned}
\sigma^2 &= \sum_i p_i(x_i-\mu)^2 \\
&= \sum_{i;x_i \leq \mu-k\sigma} p(x_i-\mu)^2 + \sum_{i;\mu-k\sigma<x_i<\mu+k\sigma} p(x_i-\mu)^2 + \sum_{i;\mu+k\sigma \leq x_i} p(x_i-\mu)^2 \\
\sigma^2 &\geq \sum_{i;x_i \leq \mu-k\sigma} p(x_i-\mu)^2 + \sum_{i;\mu+k\sigma \leq x_i} p(x_i-\mu)^2 \\
&\geq k^2\sigma^2 \sum_{i;x_i \leq \mu-k\sigma} p_i + \sum_{i;\mu+k\sigma \leq x_i} p_i \\
\frac{1}{k^2} &\geq \sum_{i;x_i \leq \mu-k\sigma} p_i + \sum_{i;\mu+k\sigma \leq x_i} p_i \\
&\geq P(V \leq \mu-k\sigma) + P(V \geq \mu+k\sigma) \\
&\geq P(V-\mu \leq -k\sigma) + P(V-\mu \geq k\sigma) \\
&\geq P(|V-\mu| \geq k\sigma) + P(|V-\mu| \geq k\sigma)
\end{aligned}$$

$$\boxed{\frac{1}{k^2} \geq P(|V-\mu| \geq k\sigma)}$$

2.4.2 Théorème de Bernouilli ou loi des grands nombres

Lors de n répétitions d'une expérience aléatoire, la fréquence relative $\frac{F}{n}$ d'un évènement tend vers sa probabilité p d'exister lorsque $n \to \infty$

$$\boxed{\frac{p(1-p)}{n\epsilon^2} \geq P\left(\left|\frac{F}{n} - p\right| \geq \epsilon\right) \overset{n\to\infty}{\to} 0}$$

Démonstration

On part avec le théorème de Bienaymé-Tchebycheff

$$\frac{1}{k^2} \geq P(|V-\mu| \geq k\sigma)$$

Et on considère une binomiale $V = B(n,p)$

$$\begin{aligned}
\frac{1}{k^2} &\geq P\left(|B(n,p) - np| \geq k\sqrt{np(1-p)}\right) \\
&\geq P\left(\frac{|B(n,p) - np|}{n} \geq \frac{k\sqrt{np(1-p)}}{n}\right) \\
&\geq P\left(\left|\frac{B(n,p)}{n} - p\right| \geq k\sqrt{\frac{np(1-p)}{n^2}}\right) \\
&\geq P\left(\left|\frac{B(n,p)}{n} - p\right| \geq k\sqrt{\frac{p(1-p)}{n}}\right)
\end{aligned}$$

On pose $k\sqrt{\frac{p(1-p)}{n}} = \epsilon$ et $B(n,p) = F$

$$\begin{cases}
k\sqrt{\frac{p(1-p)}{n}} & = \epsilon \\
k & = \epsilon\frac{\sqrt{n}}{\sqrt{p(1-p)}} \\
k^2 & = \frac{\epsilon^2 n}{p(1-p)} \\
\frac{1}{k^2} & = \frac{p(1-p)}{\epsilon^2 n}
\end{cases}$$

$$\boxed{\frac{p(1-p)}{n\epsilon^2} \geq P\left(\left|\frac{F}{n} - p\right| \geq \epsilon\right) \overset{n\to\infty}{\to} 0}$$

2.4.3 Théorème Central-Limite

Ce théorème stipule que si V est une somme de n variables aléatoires (quelconques) indépendantes

$$V = X_1 \ , \ X_2 \ , \ X_3 \ , \ \dots \ , \ X_n$$

alors sa variable réduite $\dfrac{V-\mu}{\sigma}$ tend vers une gausienne $N(0,1)$ lorsque $n \to \infty$.

$$\boxed{\frac{V-E(V)}{D(V)} = \frac{V-\mu}{\sigma} \overset{n\to\infty}{\to} N(0,1)}$$

Démonstration

On pose d'abord les calculs de l'espérance et de la variance de la somme de variable aléatoire V, pour un de ses éléments V_i, pour la somme de variable centrée W et pour un de ses éléments W_i :

$$\left\{\begin{array}{lcll}
V & = & \displaystyle\sum_{i=1}^{n} V_i & (1)\\
E(V_i) & = & \widehat{\mu} & (2)\\
D(V_i) & = & \widehat{\sigma} & (3)\\
E(V) & = & \displaystyle\sum_{i=1}^{n} E(V_i) & \\
 & = & n\widehat{\mu} & (4)\\
D(V) & = & \sqrt{D^2(V)} & \\
 & = & \sqrt{\displaystyle\sum_{i=1}^{n} D^2(V_i)} & \\
 & = & \sqrt{n\widehat{\sigma}^2} & \\
 & = & \sqrt{n}\ \widehat{\sigma} & (5)
\end{array}\right.
\qquad
\left\{\begin{array}{lcll}
W_i & = & V_i - \widehat{\mu} & (6)\\
W & = & \displaystyle\sum_{i=1}^{n} W_i & \\
 & = & \displaystyle\sum_{i=1}^{n} (V_i - \widehat{\mu}) & \\
 & = & \displaystyle\sum_{i=1}^{n} (V_i) - n\widehat{\mu} & \\
 & = & V - E(V) & (7)\\
E(W_i) & = & E(V_i - \widehat{\mu}) & \\
 & = & E(V_i) - \widehat{\mu} & \\
 & = & \widehat{\mu} - \widehat{\mu} = 0 & (8)\\
D^2(W_i) & = & D^2(V_i - \widehat{\mu}) & \\
 & = & D^2(V_i) & \\
 & = & \widehat{\sigma}^2 & (9)
\end{array}\right.$$

On commence avec la fonction génératrice des moments $\psi_V(t) = E(e^{tV})$

$$\begin{aligned}
\psi_Z(t) &= E\left(e^{tZ}\right)\\
&= E\left(e^{(t\frac{V-\mu}{\sigma})}\right)\\
&= E\left(e^{(\frac{t}{\sigma}(V-\mu))}\right)
\end{aligned}$$

$$\text{On remplace par} \begin{cases} V-\mu &= \sum_{i=1}^{n} V_i - \mu & \text{car propriété (1)} \\ &= \sum_{i=1}^{n} V_i - E(V) & \text{car propriété (4)} \\ &= \sum_{i=1}^{n} V_i - E\left(\sum_{i=1}^{n} V_i\right) & \text{car propriété (1)} \\ &= \sum_{i=1}^{n} V_i - \sum_{i=1}^{n} E\left(V_i\right) & \\ &= \sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \widehat{\mu} & \text{car propriété (2)} \\ &= \sum_{i=1}^{n} V_i - \widehat{\mu} & \end{cases}$$

$$= E\left(e^{(\frac{t}{\sigma}\sum_{i=1}^{n}(V_i-\widehat{\mu}))} \right)$$

$$= E\left(e^{(\frac{t}{\sigma}\sum_{i=1}^{n}(W_i))} \right) \qquad \text{car propriété (6)}$$

$$\text{Propriété : } Z = V + W (\text{indépendant}) \begin{cases} \psi_Z = E(e^{tZ}) \\ \quad = E(e^{t(V+W)}) \\ \quad = E(e^{tV}e^{tW}) \\ \quad = E(e^{tV})E(e^{tW}) \\ \quad = \psi_V \psi_W \end{cases}$$

$$= \psi_{W_1}\left(\frac{t}{\sigma}\right)\psi_{W_2}\left(\frac{t}{\sigma}\right) \dots \psi_{W_n}\left(\frac{t}{\sigma}\right)$$
$$= \left(\psi_{W_i}\left(\frac{t}{\sigma}\right)\right)^n$$

$$\boxed{\psi_Z(t) = E\left(e^{tZ}\right) = \left(\psi_{W_i}\left(\frac{t}{\sigma}\right)\right)^n}$$

Maintenant nous développons $\psi_{W_i}(t)$ par le théorème de Taylor :

$$\begin{cases} f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_n(x) \\ &= f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x) \end{cases}$$

$$\boxed{\begin{aligned} D^2(W_i) &= E(W_i^2) - E^2(W_i) \\ E(W_i^2) &= D^2(W_i) + E^2(W_i) \\ &= \widehat{\sigma}^2 + 0^2 \end{aligned}}$$

$$\begin{cases} \psi_{W_i}(t) &= 1 + E(W_i)t + E(W_i^2)\frac{t^2}{2} + R(t^3) \\ &= 1 + 0t + \widehat{\sigma}^2\frac{t^2}{2} + R(t^3) & \text{car propriété (8) et (9)} \end{cases}$$

$$\begin{aligned}
\lim_{n\to\infty} \psi_Z(t) &= \lim_{n\to\infty} \left(\psi_{W_i}\left(\frac{t}{\sigma}\right)^n\right) \\
&= \lim_{n\to\infty} \left(1 + \frac{t^2}{2n} + R(t^3)\right) \\
&= \lim_{n\to\infty} \left(1 + \frac{t^2}{2n}\right) \\
&= e^{\frac{t^2}{2}}
\end{aligned}$$

On a donc une allure d'une fonction génératrice d'une $N(0,1)$

$$\boxed{\psi_Z(t) = e^{\frac{t^2}{2}} \quad \text{lorsque } (n \to \infty)}$$

2.4.4 Théorème de De Moivre

C'est un cas particulier du théorème Central-Limite puisqu'une binomiale est bien une somme de variables quelconques de mêmes distributions (à savoir, des variables indicatrices). La variable binomiale est asymptotiquement normale lorsque $\to \infty$.

$$\boxed{\frac{B(n,p) - np}{\sqrt{np(1-p)}} \overset{n\to\infty}{\to} N(0,1)}$$

Démonstration

On part donc avec le théorème de Central-Limite :

$$\frac{V - E(V)}{D(V)} = \frac{V - \mu}{\sigma} \overset{n\to\infty}{\to} N(0,1)$$

Et on considère une binomiale $V = B(n,p)$.

$$\boxed{\frac{B(n,p) - np}{\sqrt{np(1-p)}} \overset{n\to\infty}{\to} N(0,1)}$$

3 L'inférence statistique

3.1 Distributions échantillonnées

3.2 Problèmes d'estimation

3.3 Intervalles de confiance et tests d'hypothèses

3.3.1 Comparaison des moyennes de deux populations normales de même écart-type σ connu

3.3.2 Comparaison des moyennes de deux populations normales de même écart-type σ inconnu

On commence avec deux moyennes des deux populations :

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma}{\sqrt{n_1}}\right)$$

$$\bar{X}_2 \sim N\left(\mu_2, \frac{\sigma}{\sqrt{n_2}}\right)$$

Maintenant nous devons faire une soustraction de deux normales $\bar{X}_1 - \bar{X}_2$, pour faire cela nous inversons une normale et nous les additionnons $\bar{X}_1 + (-\bar{X}_2)$ grace à la propriété d'inverse d'une normale 2.3.4 page 15 et la propriété d'addition 2.3.4 page 15.

$$\begin{aligned}
\bar{X}_1 + (-\bar{X}_2) &= N\left(\mu_1 - \mu_2, \sqrt{\left(\frac{\sigma}{\sqrt{n_1}}\right)^2 + \left(\frac{\sigma}{\sqrt{n_2}}\right)^2}\right) \\
&= N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right) \\
&= N\left(\mu_1 - \mu_2, \sigma\sqrt{\frac{n_1 + n_2}{n_1 n_2}}\right)
\end{aligned}$$

Ces deux variables étant indépendantes, on en déduit que

$$\begin{aligned}
T &= t_{n_1 - 1 + n_2 - 1} \\
&\sim \underbrace{\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{n_1 + n_2}{n_1 n_2}}}}_{N(0,1)} \underbrace{\frac{1}{\frac{1}{\sigma}\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}}}_{\chi^2_{n_1 - 1 + n_2 - 1}}
\end{aligned}$$

$$\boxed{T \sim \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}}\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}}}$$

4 Autres aides

4.1 Tableau du formulaire

	μ	σ^2	$\psi(t)$
$\mathcal{B}(n,p)$	np	$np(1-p)$	$(pe^t+q)^n$
$\mathcal{P}_\lambda$	λ	λ	$e^{\lambda(e^t-1)}$
Exp_λ	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$
Indicatrice(p)	p	$p(1-p)$	$1+p(e^t-1)$
Uniforme[a,b]	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{1}{t}\frac{e^{tb}-e^{ta}}{b-a}$
$\mathcal{N}(\mu,\sigma)$	μ	σ^2	$e^{\mu t+(\sigma^2 t^2)/2}$
$\chi^2_{(n)}$	n	$2n$	$(1-2t)^{-n/2}$
t_n	$0 \quad n>1$	$\frac{n}{n-2} \quad n>2$	aucun
$\mathcal{F}_{(m,n)}$	$\frac{n}{n-2} \quad n>2$	$\frac{2n^2(n+m-2)}{m(n-2^2(m-4)} \quad n>2$	aucun

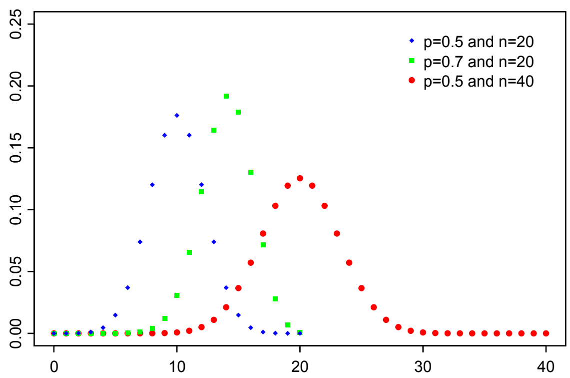
Tableau dans le formulaire disponible à l'examen écrit (en rouge à connaître)

4.2 Densité et répartition

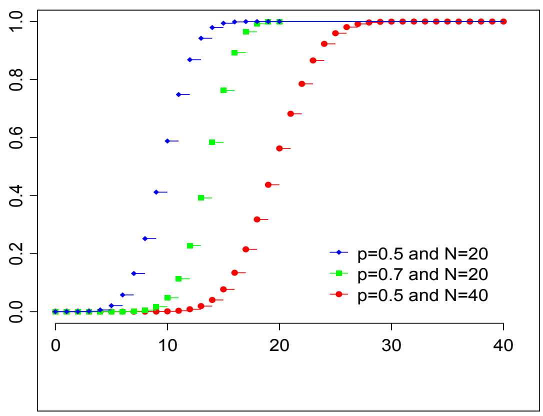
	Fonction de densité $f(x)$	Fonction de répartition $F(x)$
$\mathcal{B}(n,p)$	$P[B(n,p)=k]$	$\sum_{k=0}^{x} P[B(n,p)=k]$
$\mathcal{P}_\lambda$	$P[\mathcal{P}_\lambda=k]=\frac{\lambda^k}{k!}e^{-\lambda}$	$\sum_{k=0}^{x} P[\mathcal{P}_\lambda=k]$
Exp_λ	$\begin{cases}\lambda e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$	$\begin{cases}1-e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$
Indicatrice(p)	$V_A \Rightarrow \begin{cases}P(V_A=1)=p\\ P(V_A=0)=1-p\end{cases}$	$\begin{cases}0 & x<0\\ 1-p & 0\leq x<1\\ 1 & x\geq 1\end{cases}$
Uniforme[a,b]	$\begin{cases}\frac{1}{b-a} & a\leq x\leq b\\ 0 & sinon\end{cases}$	$\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\leq x<b\\ 1 & x\geq b\end{cases}$
$\mathcal{N}(\mu,\sigma)$	$\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{(u-\mu)^2}{2\sigma^2}}du$
$\chi^2_{(n)}$	$\begin{cases}\frac{1}{\sqrt{2\pi}}e^{\frac{x^2}{2}} & x>0\\ 0 & x\leq 0\end{cases}$	$\frac{1}{\sqrt{2\pi}}\int_0^x e^{\frac{x^2}{2}}du$
t_n	Densité indépendante de σ	
$\mathcal{F}_{(m,n)}$	Densité indépendante de σ	

4.3 Distributions

Distribution Binomiale $\mathcal{B}(n,p)$

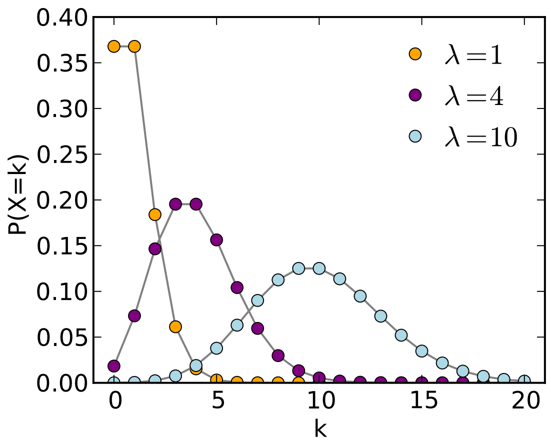


Densité de probabilité

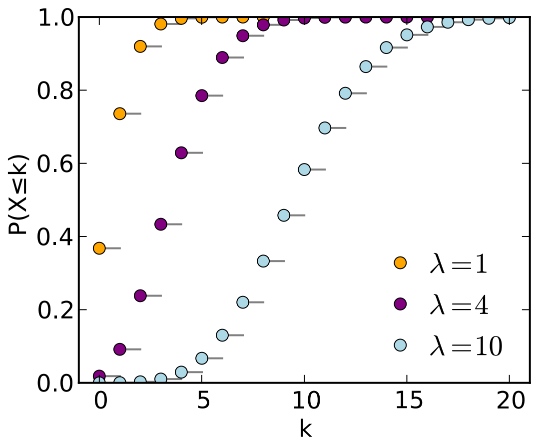


Fonction de répartition

Distribution Poisson $\mathcal{P}_\lambda$

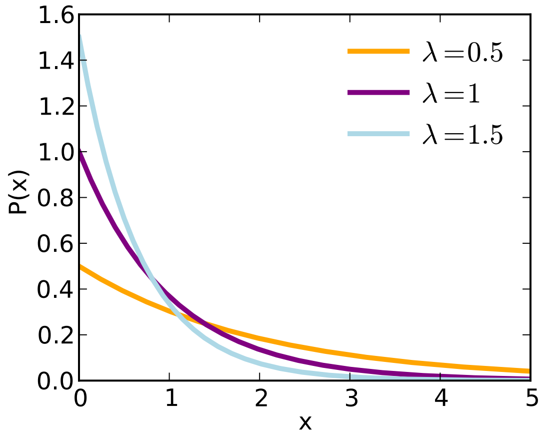


Densité de probabilité

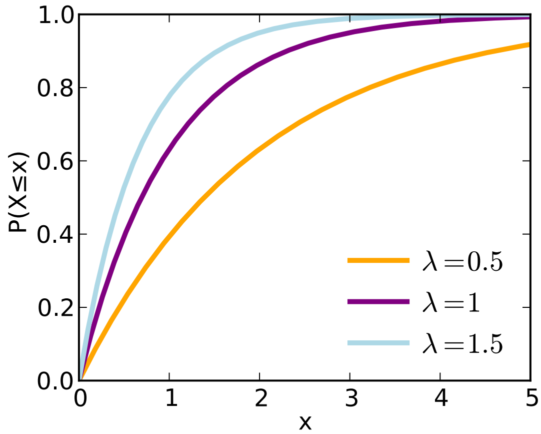


Fonction de répartition

Distribution Exponentielle Exp_λ

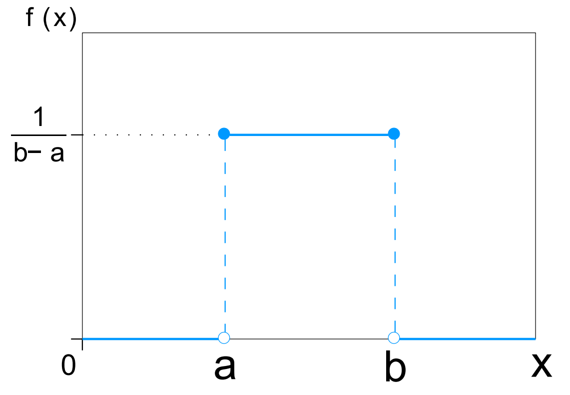


Densité de probabilité



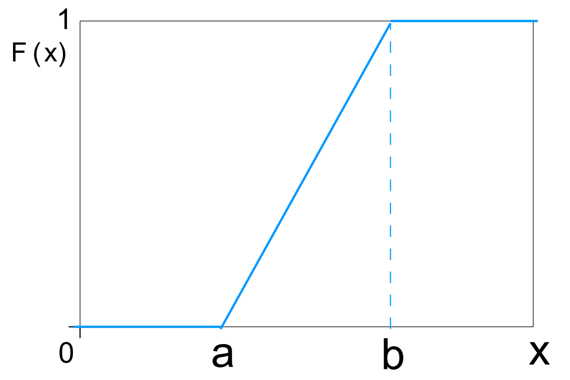
Fonction de répartition

<table>
<tr><th colspan="2">Distribution Uniforme $\mathcal{U}[a, b]$</th></tr>
<tr><td>



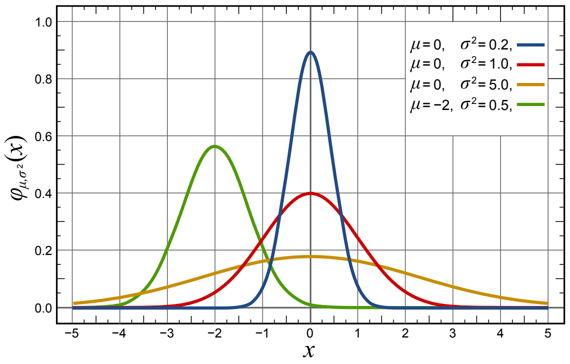
Densité de probabilité

</td><td>



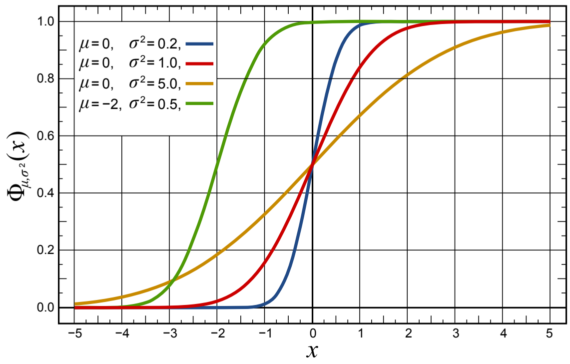
Fonction de répartition

</td></tr>
<tr><th colspan="2">Distribution Normale $\mathcal{N}(\mu, \sigma)$</th></tr>
<tr><td>



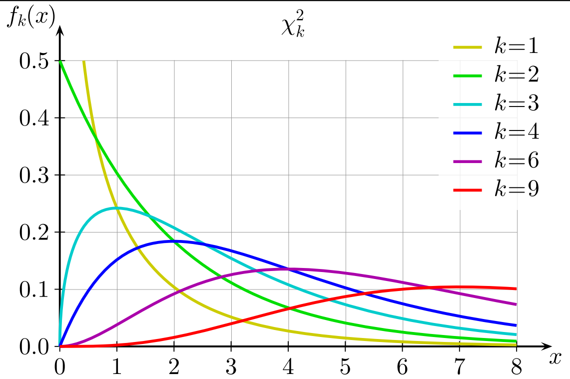
Densité de probabilité

</td><td>



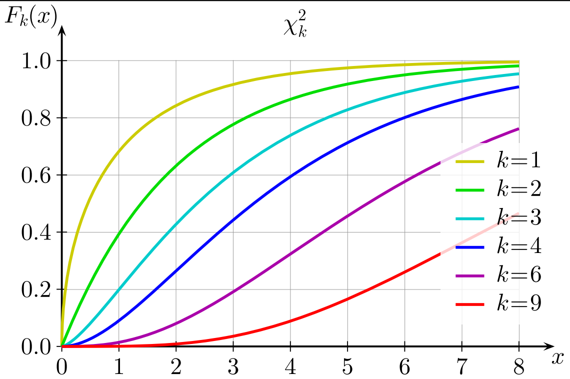
Fonction de répartition

</td></tr>
<tr><th colspan="2">Distribution Khi carré χ^2</th></tr>
<tr><td>



Densité de probabilité

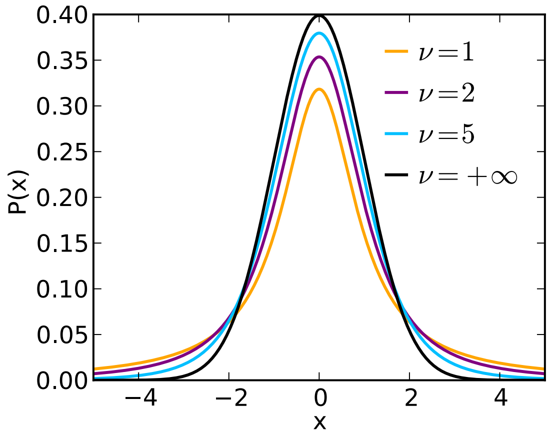
</td><td>



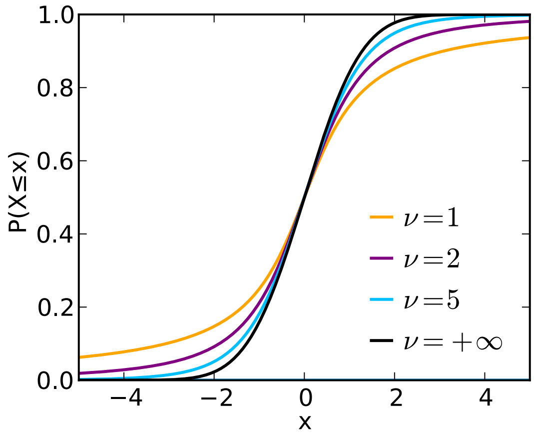
Fonction de répartition

</td></tr>
</table>

Distribution $\textsc{Student}$ t_n

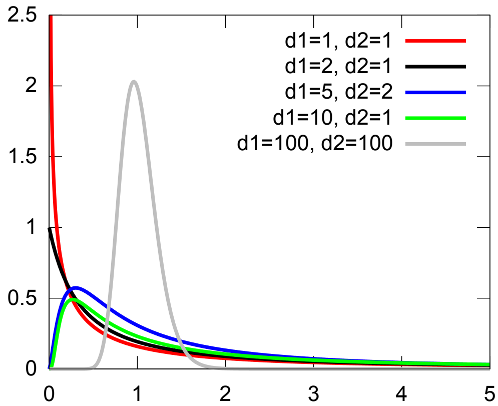


Densité de probabilité

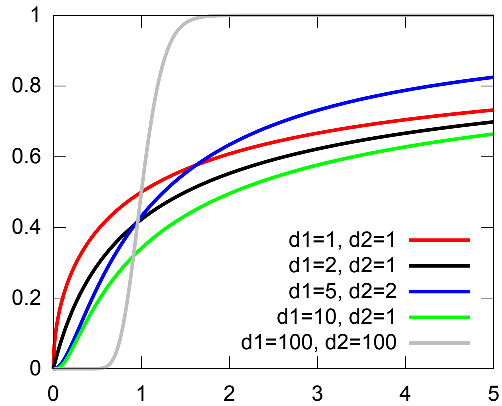


Fonction de répartition

Distribution $\textsc{Fisher-Snedecor}$ $\mathcal{F}_{m,n}$



Densité de probabilité



Fonction de répartition