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

Rodrigue Van Brande

Dylan Gonzalez

31 août 2015

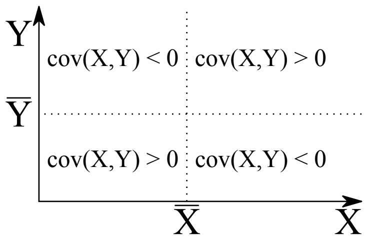
Table des matières

1 La statistique descriptive

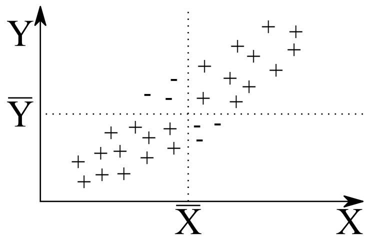
1.1 Statistique descriptive en 1D

1.2 Statistique descriptive en 2D

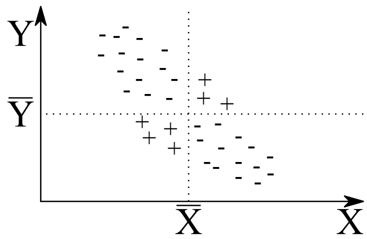
1.2.1 Covariance



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) :

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

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

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 si elle vaut $|m_{11}| = s_1 s_2$.
Si les points observés se trouvent sur une droite non parallèle aux axes $ax+by+c=0$, on a $ax_i + by_i + c = 0$.
On multiplie par $\frac{n_{ij}}{n}$ et on somme sur ij.

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

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

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

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

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

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

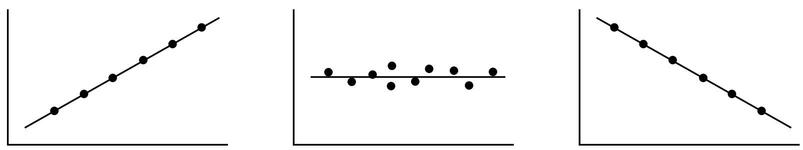
1.2.2 Le coefficient de corrélation

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

Propriétés

1. r sans dimensions ;
2. $r' = r$;
3. $-1 \leq r \leq 1$;
4. $|r| = 1$ ssi les points observés se trouvent sur une droite non parallèle aux axes.

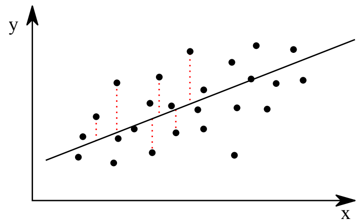
Représentation de la droite de régression en fonction du coefficient de corrélation



$r = 1$ $r = 0$ $r = -1$

- si $|r| = 1$, il y a une relation fonctionnelle linéaire entre X et Y ;
- si $r = 0$, Y est indépendante de X : la covariance est nulle et la droite de régression est horizontale ;
- la liaison entre X et Y est d'autant plus intime que |r| est voisin de 1, et d'autant plus faible que |r| est voisin de 0.

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 :

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

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

Dérivée par rapport à b :

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

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

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) :

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

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 :

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

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 :

$$\boxed{x = \frac{m_{11}}{s_2^2}(y - \bar{y}) + \bar{x}}$$

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 $\dfrac{m_{11}}{s_1} = r s_2$:

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

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

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

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

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

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

2.1.4 Probabilité conditionnelle inverse et indépendant

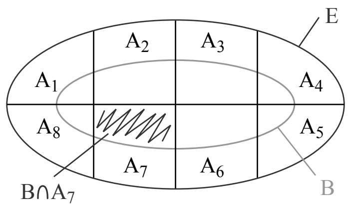
$$P(A|B) = P(A|\bar{B})$$

Démonstration

$$\begin{array}{ccc} P(A|B) & = & P(A|\bar{B}) \\ \dfrac{P(A\cap B)}{P(B)} & = & \dfrac{P(A\cap \bar{B})}{P(\bar{B})} \\ \dfrac{P(A)P(B)}{P(B)} & = & \dfrac{P(A)P(\bar{B})}{P(\bar{B})} \\ P(A) & = & P(A) \end{array}$$

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.



$$\begin{array}{ccccccccc} B & = & (A_1\cap B) & \cup & (A_2\cap B) & \cup & \ldots & \cup & (A_m\cap B) \\ P(B) & = & P(A_1\cap B) & + & P(A_2\cap B) & + & \ldots & + & P(A_m\cap B) \\ & = & P(B|A_1)P(A_1) & + & P(B|A_2)P(A_2) & + & \ldots & + & P(B|A_m)P(A_m) \end{array}$$

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

2.2 Variables aléatoires

2.2.1 Valeurs typiques

Aux distributions/densités marginales et conditionnelles, on peut associer les valeurs typiques :

Moyenne marginale μ_1	$= \sum_i p_{i.} x_i$	$= \int_{-\infty}^{\infty} x f_1$
Variance marginale σ_1^2	$= \sum_i p_{i.} (x_i - \mu_1)^2$	$= \int_{-\infty}^{\infty} (x - \mu_1)^2 f_1$
Moyenne conditionnelle	$\mu_{1/j} = \sum_i p_{i/j} x_i$	$\mu_{1/y_0} = \int_{-\infty}^{\infty} x f (x - \mu_1)^2 f_1$
Variance conditionnelle	$\sigma_{1/j}^2 = \sum_i p_{i/j} (x_i - \mu_{1/j})^2$	$\sigma_{1/j}^2 = \int_{-\infty}^{\infty} (x - \mu_{1/y_0})^2 f(x/y_0)$

2.2.2 Distribution d'une fonction monotone d'une variable aléatoire $G(V)$

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

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

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

$$\text{Rappel :} \left\{ \begin{array}{ccc} G\left(G^{-1}(x)\right) & = & x \\ \left(G\left(G^{-1}(x)\right)\right)' & = & (x)' \\ G'\left(G^{-1}(x)\right)\ \left(G^{-1}(x)\right)' & = & 1 \\ \left(G^{-1}(x)\right)' & = & \dfrac{1}{G'\left(G^{-1}(x)\right)} \end{array} \right.$$

$$= f_V\left(G^{-1}(x)\right)\frac{1}{G'\left(G^{-1}(x)\right)} \quad \Big| \quad f_V\left(G^{-1}(x)\right)\frac{-1}{G'\left(G^{-1}(x)\right)}$$

2.2.3 Distribution de la somme de deux variables aléatoires $V+W$

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

Cas discret

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

Cas continu

$$\begin{aligned} F_Z(x) &= P(Z \leq x) \\ F_{V+W}(x) &= P(V \leq \xi, W \leq \eta) \qquad\qquad \text{tel que } v_i + w_i \leq x \\ &= \iint\limits_{\xi+\eta \leq x} f_{(V,W)}(\xi, \eta)\ d\xi\ d\eta \\ &\text{On remplace par} \begin{cases} \xi &= u \\ \eta &= v - u \\ J &= \begin{pmatrix} \frac{\delta \xi}{\delta u} & \frac{\delta \xi}{\delta v} \\ \frac{\delta \eta}{\delta u} & \frac{\delta \eta}{\delta v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.1 - 0.(-1) = 1 \end{cases} \\ &= \iint\limits_{v \leq x} f_{(V,W)}(u, v-u)\,|1|\ du\ dv \\ &= \int_{-\infty}^{x} dv \int_{-\infty}^{+\infty} f_{(V,W)}(u, v-u)\ du \end{aligned}$$

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

ou si indépendant

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

2.2.4 Distribution du produit de deux variables aléatoires $V.W$

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

Cas continu

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

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

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

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

ou si indépendant

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

2.2.5 Distribution de la somme des espérances $E(V+W)$

$$\boxed{E(V+W) = E(V) + E(W)}$$

Demonstration

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} x \ f_{(V+W)}(x) \ dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \ f_{(V,W)}(u, x-u) \ du \ dv && \text{2.2.3 page 13} \\
\text{On remplace par } & \left\{ \begin{array}{ll} u & = \xi \\ x & = \xi - \eta \\ J & = \begin{pmatrix} \dfrac{\delta u}{\delta \xi} & \dfrac{\delta u}{\delta \eta} \\ \dfrac{\delta x}{\delta \xi} & \dfrac{\delta x}{\delta \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = 1.-1-0.1 = -1 \end{array} \right. \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\xi+\eta) \ f_{(V,W)}(\xi,\eta) \, |-1| \ d\xi \ d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \xi \ f_{(V,W)}(\xi,\eta) \ d\xi \ d\eta + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \eta \ f_{(V,W)}(\xi,\eta) \ d\xi \ d\eta \\
&= \int_{-\infty}^{\infty} \xi \ f_V(\xi) \ d\xi \ + \int_{-\infty}^{\infty} \eta \ f_W(\eta) \ d\eta \\
&= E(V) + E(W)
\end{aligned}$$

2.2.6 Distribution du produit des espérances $E(V.W)$

$$\boxed{E(V.W) = E(V).E(W)}$$

Demonstration

Uniquement pour des variables indépendantes

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} x\ f_{(V.W)}(x)\ dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ f_V(u) f_w\left(\frac{x}{u}\right)\left|\frac{1}{u}\right|\ du\ dv && \text{2.2.4 page 14} \\
\text{On remplace par} & \begin{cases} u &= \xi \\ x &= \xi\eta \\ J &= \begin{pmatrix} \frac{\delta u}{\delta \xi} & \frac{\delta u}{\delta \eta} \\ \frac{\delta x}{\delta \xi} & \frac{\delta x}{\delta \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \eta & \xi \end{pmatrix} = 1.\xi - 0.\eta = \xi \end{cases} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\xi\eta)\ f_{(V,W)}(\xi,\eta)\left|\frac{1}{u}\right| |\xi|\ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\xi\eta)\ f_{(V,W)}(\xi,\eta)\left|\frac{1}{u}\right| |u|\ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\xi\eta)\ f_{(V,W)}(\xi,\eta)\ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \xi\ f_V(\xi) f_W(\eta)\ d\xi\ d\eta \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \eta\ f_{(V,W)}(\xi,\eta)\ d\xi\ d\eta \\
&= \int_{-\infty}^{\infty} \xi\ f_V(\xi)\ d\xi\ \int_{-\infty}^{\infty} \eta\ f_W(\eta)\ d\eta \\
&= E(V).E(W)
\end{aligned}$$

2.2.7 Variance d'une variable aléatoire $D^2(aV+b)$

$$\boxed{D^2(aV+b) = a^2 D^2(V)}$$

Démonstration

D'après la propriété 2.2.1 page 11 : $\sigma_V^2 = D^2(V) = \int_{-\infty}^{\infty} (x-\mu)^2 f_V(x)dx = E\left((x-\mu)^2\right)$

$$\begin{aligned}
D^2(W) &= D^2(aV+b) \\
&= E\left((W-(a\mu+b))^2\right) \\
&= E\left((aV+b-a\mu-b)^2\right) \\
&= E\left(a^2(V-\mu)^2\right) \\
&= a^2 D^2(V)
\end{aligned}$$

2.2.8 Variance de la somme de deux variables aléatoires $D^2(V+W)$

$$\boxed{D^2(V+W) = D^2(V) + D^2(W) + 2\mu_{11}}$$

Démonstration

D'après la propriété 2.2.1 page 11 : $\sigma_V^2 = D^2(V) = \int_{-\infty}^{\infty} (x-\mu)^2 f_V(x)dx = E\left((x-\mu)^2\right)$

$$\begin{aligned}
D^2(Z) &= D^2(V+W) \\
&= E\left((Z-(\mu_V+\mu_W))^2\right) \\
&= E\left(((V+W)-(\mu_V+\mu_W))^2\right) \\
&= E\left(((V-\mu_V)+(W-\mu_W))^2\right) \\
&= E\left((V-\mu_V)^2\right) + E\left((V-\mu_W)^2\right) + 2E\left((V-\mu_V)(W-\mu_W)\right) \\
&= D^2(V) + D^2(W) + 2\mu_{11}
\end{aligned}$$

2.3 Variables aléatoires particulières

2.3.1 Schéma de relation entre les variables aléatoires

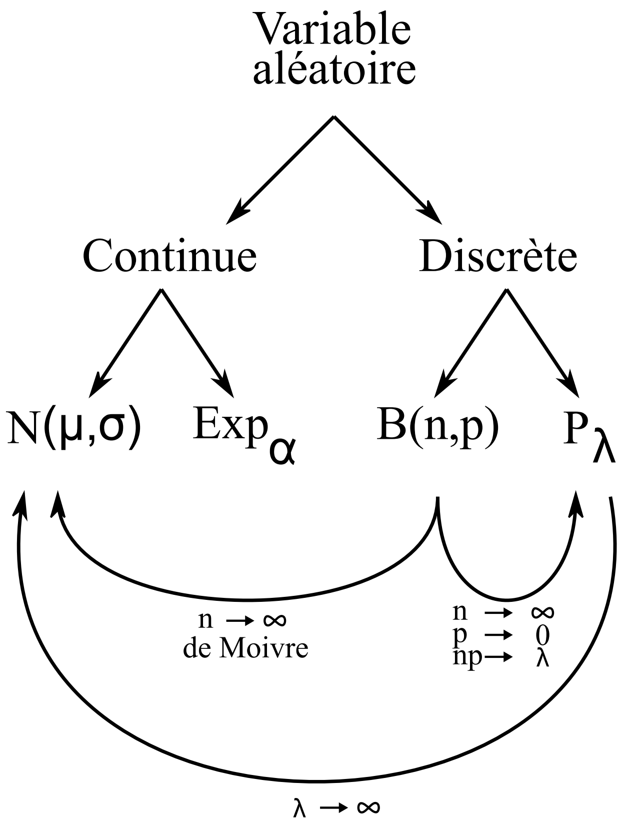


Schéma de relation entre les différentes variables aléatoires

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

Une variable binomiale est une expérience avec n répétitions. On veut savoir le nombre k de fois qu'un évènement A, de probabilité p, se produit :

$$\begin{aligned}
P(V=k) &= \left(\begin{array}{c} n \\ k \end{array}\right) & p(\text{obtenir } k \text{ fois } A) & \quad p(n-k \text{ fois ne pas obtenir } A) \\
&= \left(\begin{array}{c} n \\ k \end{array}\right) & p^k & \quad (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} & p^k & \quad (1-p)^{n-k}
\end{aligned}$$

La variable V est appelée variable binomiale de paramètre p et d'exposant n et est notée $B(n,p)$. C'est la somme de n de variable indicatrices.

Calcul de la moyenne $E(B(n,p))$

$$\begin{aligned}
\mu &= \sum_{k=0}^{n} x_i p_i \\
&= \sum_{k=0}^{n} k \left(\begin{array}{c} n \\ k \end{array}\right) p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}
\end{aligned}$$

On remplace $k-1$ par l :

$$\begin{aligned}
&= \sum_{l=0}^{n-1} \frac{n(n-1)!}{l!(n-l-1)!} p^{l+1} (1-p)^{n-l-1} \\
&= np \sum_{l=0}^{n-1} \left(\begin{array}{c} n-1 \\ l \end{array}\right) p^l (1-p)^{n-l-1} \\
&= np(p+(1-p))^{n-1} \\
&= np
\end{aligned}$$

Calcul de la variance $D^2(B(n,p))$

Calcul de la fonction génératrice $\psi(t)$

Calcul de la stabilité

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

Une variable de poisson est uniquement une binomiale dont la répétition n est "grand" ($n \geq 30$) et que la probabilité p est "petit" ($p \leq 0,1$).

$$\begin{aligned}
P(V=k) &= \begin{pmatrix} n \\ k \end{pmatrix} && p^k(1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} && p^k(1-p)^{n-k} \\
&= \frac{1}{k!}n(n-1)(n-2)\dots(n-k-1) && p^k(1-p)^{n-k} \\
&= \frac{1}{k!}n^k\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) && p^k(1-p)^{n-k} \\
&= \frac{1}{k!}n^k\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) && \frac{p^k\,(1-p)^n}{(1-p)^k} \\
&= \frac{1}{k!}n^k\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) && \frac{p^k\left(1-\frac{np}{n}\right)^n}{(1-p)^k}
\end{aligned}$$

Sous une autre forme :

$$\begin{aligned}
P(V=k) &= \frac{n^k p^k}{k!} && \left(1-\frac{np}{n}\right)^n && \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)}{(1-p)^k} \\
&= \frac{(np)^k}{k!} && e^{-\lambda} && \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)}{(1-p)^k}
\end{aligned}$$

Et donc quand $n \to \infty$:

$$P(V=k) \overset{n\to\infty}{=} \frac{\lambda^k}{k!}e^{-\lambda}\frac{(1-1)(1-1)\dots(1-1)}{(1-p)^k}$$

$$\boxed{P(V=k) \overset{n\to\infty}{=} \frac{\lambda^k}{k!}e^{-\lambda} = P_\lambda}$$

Calcul de la moyenne $E(P_\lambda)$

$$\begin{aligned}
\mu &= \sum_{k=0}^{\infty} x_i p_i \\
&= \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} \\
&= \sum_{k=0}^{\infty} \frac{\lambda^{k-1}\lambda}{(k-1)!}e^{-\lambda} \\
&= \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda}e^{\lambda} \\
&= \lambda
\end{aligned}$$

Calcul de la variance $D^2(P_\lambda)$

$$\begin{aligned}\sigma^2 &= \sum_{k=0}^{\infty} x_i^2 p_i - \mu^2 \\ &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \mu^2\end{aligned}$$

On remplace k^2 par $k(k-1)+k$

$$\begin{aligned}&= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} - \mu^2 \\ &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!} e^{-\lambda} + \mu - \mu^2 \\ &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^{k-2}\lambda^2}{k(k-1)(k-2)!} e^{-\lambda} + \mu - \mu^2 \\ &= \lambda^2 e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \mu - \mu^2 \\ &= \lambda^2 e^{-\lambda} e^{\lambda} + \mu - \mu^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda\end{aligned}$$

Calcul de la fonction génératrice $\psi(t)$

$$\begin{aligned}(t) &= \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^t} \\ &= e^{\lambda(e^t-1)}\end{aligned}$$

Calcul de la stabilité

$$\begin{aligned}\psi_{1+2}(t) &= \psi_1(t)\psi_2(t) \\ &= e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)} \\ &= e^{\lambda_1(e^t-1)(\lambda_1+\lambda_2)}\end{aligned}$$

La variable est bien stable.

2.3.4 Variable exponentielle négative

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

Propriété d'opposé

Soit $\bar{X}_1$ et $\bar{X}_2$ deux variables normales. Si $\bar{X}_2$ est l'opposé 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.

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

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

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

2.3.6 Variable Khi^2

2.3.7 Variable Student t_n

2.3.8 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 l'écart-type 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 :

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

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 :

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

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

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

$$\boxed{Z = \frac{W}{\sqrt{n}\widehat{\sigma}} = \frac{V-\mu}{\sigma} \overset{n\to\infty}{\to} N(0,1)}$$

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

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

$$\boxed{\psi_Z(t) \overset{n\to\infty}{\to} N(0,1)}$$

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 $\rightarrow \infty$.

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

Démonstration

On part donc avec le théorème de Central-Limite :

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

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

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

3 L'inférence statistique

3.1 Les valeurs

Échantillon de la normale $\frac{n_1 S_1^2}{\sigma^2}$

Variance de l'échantillon S_1^2

Variance théorique de l'ensemble de la population σ^2

3.2 Distributions échantillonnées

3.2.1 Distribution échantillonnées de la moyenne $\bar{X}$

Supposons que la distribution de la population ait une moyenne μ et un écart-type σ. Il en est donc de même pour les distributions échantillonnées de $X_1, X_2, \ldots, X_n$ et il résulte alors des propriétés de l'espérance mathématique que :

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ & \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\ & \frac{1}{n}\sum_{i=1}^{n} \mu \\ & \mu \end{aligned}$$

et il résulte des propriétés de la variance que :

$$D^2(\bar{X}) = D^2\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$\boxed{D^2(aV+b) = a^2 D^2(V) \qquad \text{2.2.7 page 17}}$$

$$\begin{aligned} &= \frac{1}{n^2}\sum_{i=1}^{n} D^2(X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \sigma \\ &= \frac{\sigma^2}{n} \end{aligned}$$

3.2.2 Distribution échantillonnées de la variance S^2

$$\begin{aligned} E(S^2) &= E\left(\frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^2\right) + 2E\left(\frac{1}{n}\sum_{i=1}^{n} X_i \bar{X}\right) + \frac{1}{n}\sum_{i=1}^{n} E\left(\bar{X}^2\right) \end{aligned}$$

3.2.3 Distribution échantillonnées d'une fonction de $\bar{X}$ et s^2

Il résulte de ce qui précède que lorsque la population a une distribution normale, la variable

$$\boxed{\sqrt{n-1}\frac{\bar{X}-\mu}{S} \sim t_{(n-1)}}$$

Démonstration

La définition d'une chi carré $\chi^2_{(n-1)}$ est une somme de normale au carré $\sum_{i=1}^{n-1}(N_i(0,1))^2$.

Or on a vu que $\chi^2_{(n-1)} = \frac{nS^2}{\sigma^2}$, donc

$$\begin{aligned}
\chi^2_{(n-1)} &= \sum_{i=1}^{n-1}(N_i(0,1))^2 \\
\frac{nS^2}{\sigma^2} &= \sum_{i=1}^{n-1}\left(\frac{N_i(0,\sigma)}{\sigma}\right)^2 \\
\frac{nS^2}{\sigma^2} &= \sum_{i=1}^{n-1}\frac{(N_i(0,\sigma))^2}{\sigma^2} \\
nS^2 &= \sum_{i=1}^{n-1}(N_i(0,\sigma))^2 \\
nS^2\frac{1}{n-1} &= \frac{1}{n-1}\sum_{i=1}^{n-1}(N_i(0,\sigma))^2 \\
\left(\frac{nS^2}{n-1}\right)^{\frac{-1}{2}} &= \left(\frac{1}{n-1}\sum_{i=1}^{n-1}(N_i(0,\sigma))^2\right)^{\frac{-1}{2}} \\
\frac{1}{\sqrt{\frac{nS^2}{n-1}}} &= \frac{1}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(N_i(0,\sigma))^2}} \\
\frac{N(0,\sigma)}{S\sqrt{\frac{n}{n-1}}} &= \frac{N(0,\sigma)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(N_i(0,\sigma))^2}} \\
\frac{N(0,\sigma)}{S\sqrt{\frac{n}{n-1}}} &= t_{(n-1)}
\end{aligned}$$

On obtient bien une student $t_{(n-1)}$ à $n-1$ degrés de liberté (2.3.7 page 22). On va donc remplacer la normale $N(0,\sigma)$ dans l'équation par :

$$\begin{cases}
\bar{X} & \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \\
\frac{\bar{X}-\mu}{\frac{1}{\sqrt{n}}} & \sim N(0,\sigma) \\
\sqrt{n}\left(\bar{X}-\mu\right) & \sim N(0,\sigma)
\end{cases}$$

$$t_{(n-1)} = \frac{N(0,\sigma)}{S\sqrt{\frac{n}{n-1}}} = \frac{\sqrt{n}\left(\bar{X}-\mu\right)}{S\sqrt{\frac{n}{n-1}}} = \frac{\bar{X}-\mu}{S\frac{1}{\sqrt{n-1}}} = \sqrt{n-1}\frac{\bar{X}-\mu}{S}$$

$$\boxed{t_{(n-1)} = \sqrt{n-1}\frac{\bar{X}-\mu}{S}}$$

3.3 Test d'hypothèse

Erreur de première espèce est de rejeter une hypothèse alors qu'elle est vraie. Le risque α de commettre cette erreur est ϵ;

Erreur de seconde espèce est d'accepter une hypothèse alors qu'elle est fausse. Si on accepte une moyenne μ_0 alors qu'elle vaut en réalité μ_1, alors le risque de commettre cette erreur est

$$\begin{aligned}
\beta &= P\left(\mu_0 - u_{\beta/2}\frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu_0 + u_{\beta/2}\frac{\sigma}{\sqrt{n}}\right) \\
&\left\{ \text{où } \bar{X} \sim N\left(\mu_1, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow \frac{\bar{X} - \mu_1}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \right. \\
&= P\left(\mu_0 - u_{\beta/2}\frac{\sigma}{\sqrt{n}} \leq N(0,1)\frac{\sigma}{\sqrt{n}} + \mu_1 \leq \mu_0 + u_{\beta/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\frac{\mu_0 - \mu_1}{\sigma}\sqrt{n} - u_{\beta/2} \leq N(0,1) \leq \frac{\mu_0 - \mu_1}{\sigma}\sqrt{n} + u_{\beta/2}\right)
\end{aligned}$$

Puissance d'un test est la valeur $1 - \beta$, c'est à dire la probabilité de rejeter l'hypothèse alors qu'elle est fausse.

3.4 Intervalles de confiance et tests d'hypothèses

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

On commence avec deux moyennes des deux populations :

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma_1}{\sqrt{n_1}}\right)$$

$$\bar{X}_2 \sim N\left(\mu_2, \frac{\sigma_2}{\sqrt{n_2}}\right)$$

Maintenant nous devons faire une soustraction de deux normales $\bar{X}_1 - \bar{X}_2$, pour faire cela nous prenons l'opposé d'une normale et nous les additionnons $\bar{X}_1 + (-\bar{X}_2)$ grace à la propriété d'opposé d'une normale 2.3.5 page 22 et la propriété d'addition 2.3.5 page 22.

$$\begin{aligned} \bar{X}_1 + (-\bar{X}_2) \quad & \sim N\left(\mu_1 - \mu_2, \sqrt{\left(\frac{\sigma_1}{\sqrt{n_1}}\right)^2 + \left(\frac{\sigma_2}{\sqrt{n_2}}\right)^2}\right) \\ & \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \end{aligned}$$

Ces deux variables étant indépendantes, on en déduit que

$$\boxed{T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)}$$

A tout $\epsilon > 0$, on peut donc associé un nombre tel que

$$P\left(-u_{\epsilon/2} \leq T \leq u_{\epsilon/2}\right) = 1 - \epsilon$$

d'où on peut donc déduire un intervalle de confiance pour $\mu_1 - \mu_2$. D'autre part on rejettera l'hypothèse $\mu_1 = \mu_2$ si

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_{\epsilon/2}$$

3.4.2 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 prenons l'opposé d'une normale et nous les additionnons $\bar{X}_1 + (-\bar{X}_2)$ grace à la propriété d'opposé d'une normale 2.3.5 page 22 et la propriété d'addition 2.3.5 page 22.

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

et avec un $\chi^2_{(n_1-1+n_2-1)}$ où $n =$ nombre de degré de libertés

$$\begin{aligned}
\frac{n_1 s_1^2 + n_2 s_2^2}{\sigma^2} &\sim \chi^2_{(n_1+n_2-2)} \\
\frac{1}{\sigma}\sqrt{n_1 s_1^2 + n_2 s_2^2} &\sim \sqrt{\chi^2_{(n_1+n_2-2)}}
\end{aligned}$$

Ces deux variables étant indépendantes, on en déduit que

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

$$\boxed{T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}} \sim t_{n_1-1+n_2-1}}$$

A tout $\epsilon > 0$, on peut donc associer un nombre tel que

$$P\left(-t_{\epsilon/2} \leq T \leq t_{\epsilon/2}\right) = 1 - \epsilon$$

d'où on peut donc déduire un intervalle de confiance pour $\mu_1 - \mu_2$. D'autre part on rejettera l'hypothèse $\mu_1 = \mu_2$ si

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}} > t_{\epsilon/2}$$

3.4.3 Comparaison des moyennes de deux populations quelconques

Si n_1 et n_2 sont suffisamment grands (au moins 20), alors $\bar{X}_1 - \bar{X}_2$ a une distribution approximativement normale :

$$\boxed{\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}$$

où on remplace σ_1 et σ_2 par s_1 et s_2 lorsqu'ils sont inconnus. On rejette alors l'hypothèse $\mu_1 = \mu_2$ lorsque

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > u_{\epsilon/2}$$

4 Autres

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	$e^{\mu t+(\sigma^2t^2)/2}$
$\chi^2_{(n)}$	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

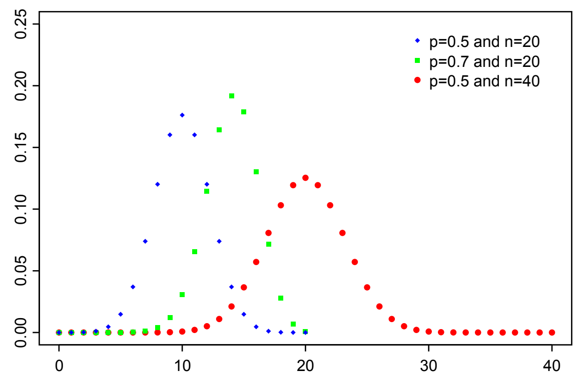
Tableau dans le formulaire disponible à l'examen écrit (en rouge à connaître)

4.2 Densité et répartition

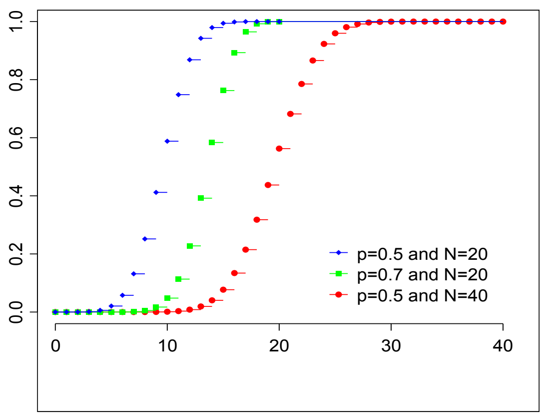
	Fonction de densité $f(x)$	Fonction de répartition $F(x)$
$\mathcal{B}(n,p)$	$P[B(n,p)=k]$	$\sum_{k=0}^{x} P[B(n,p)=k]$
$\mathcal{P}_\lambda$	$P[\mathcal{P}_\lambda=k]=\frac{\lambda^k}{k!}e^{-\lambda}$	$\sum_{k=0}^{x} P[\mathcal{P}_\lambda=k]$
Exp_λ	$\begin{cases}\lambda e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$	$\begin{cases}1-e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$
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 $\mathcal{B}(n, p)$

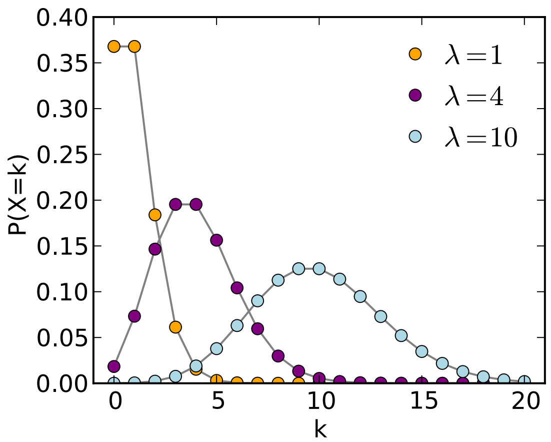


Densité de probabilité

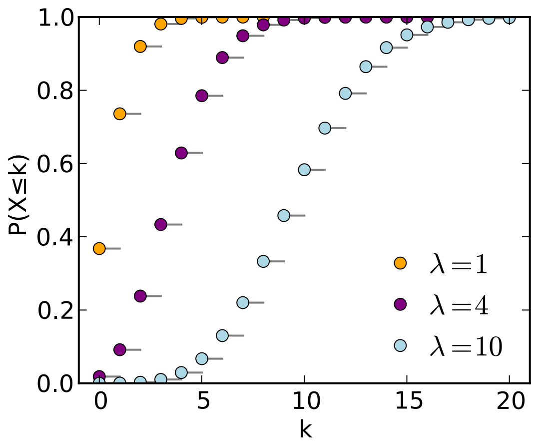


Fonction de répartition

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

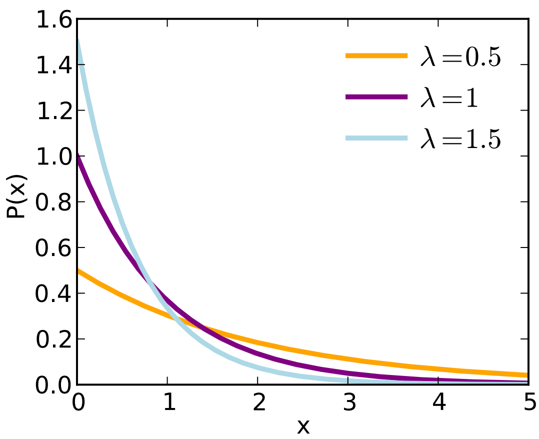


Densité de probabilité

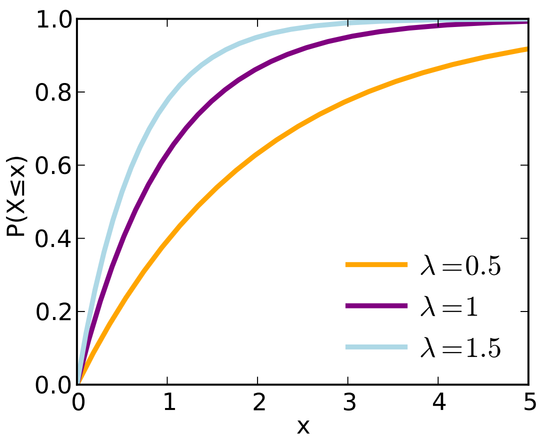


Fonction de répartition

Distribution Exponentielle Exp_λ

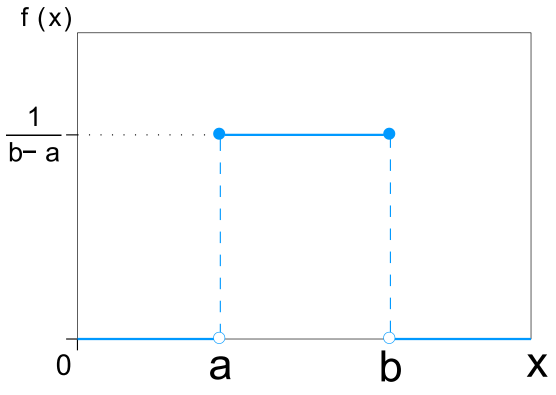


Densité de probabilité

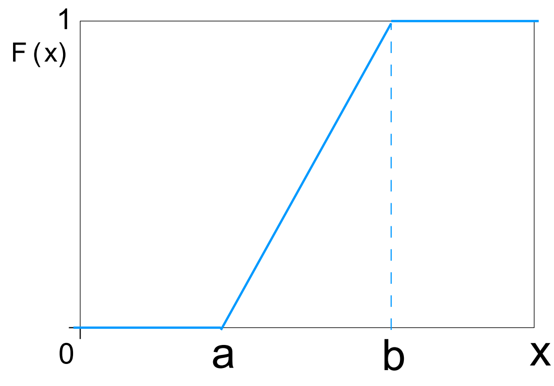


Fonction de répartition

Distribution Uniforme $\mathcal{U}[a,b]$

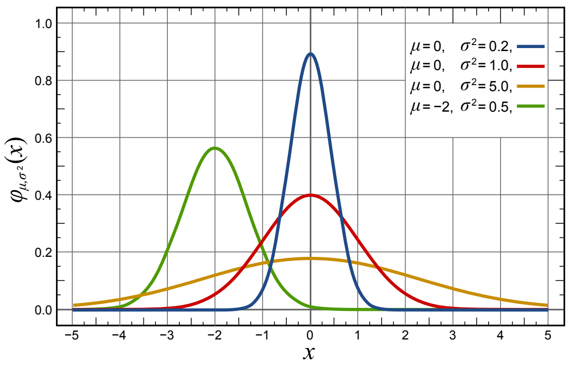


Densité de probabilité

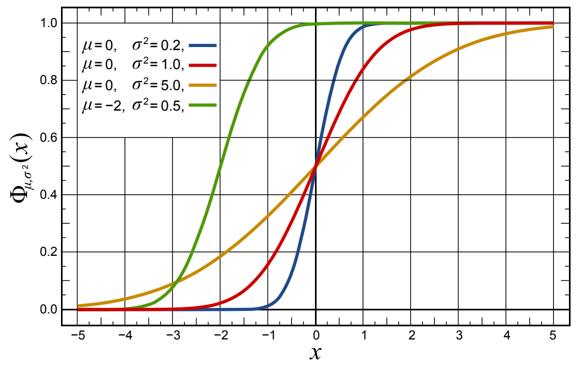


Fonction de répartition

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

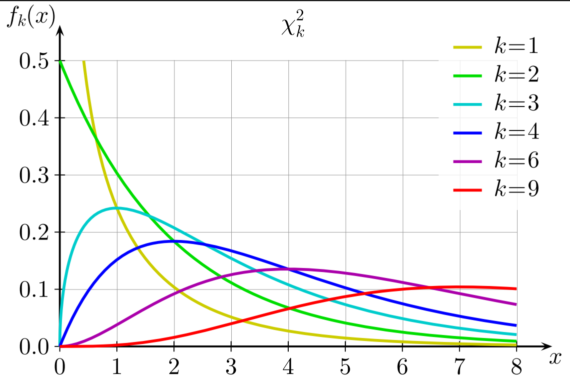


Densité de probabilité

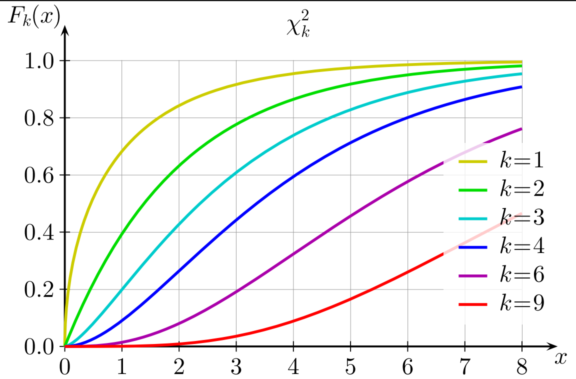


Fonction de répartition

Distribution Khi carré χ^2

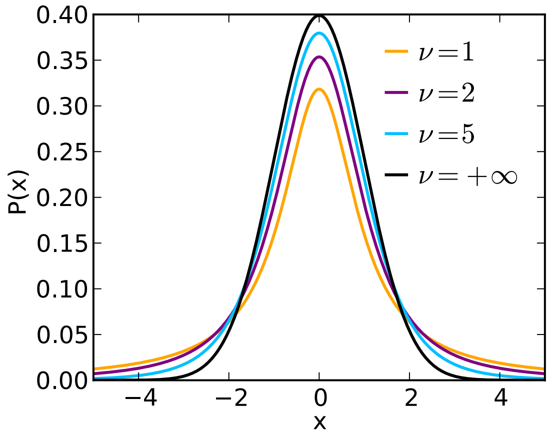


Densité de probabilité

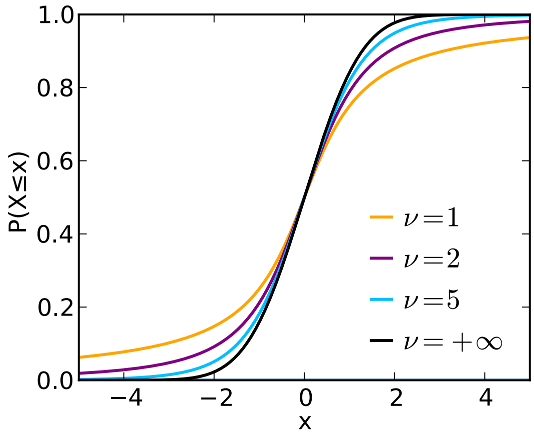


Fonction de répartition

Distribution Student t_n

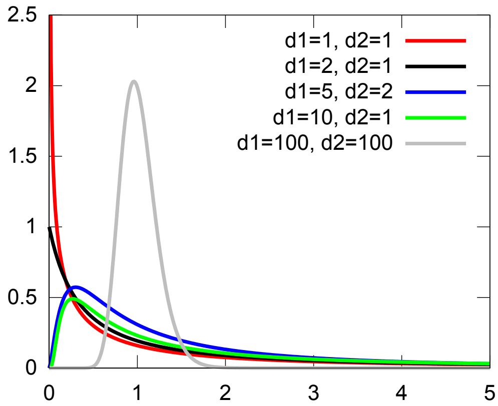


Densité de probabilité

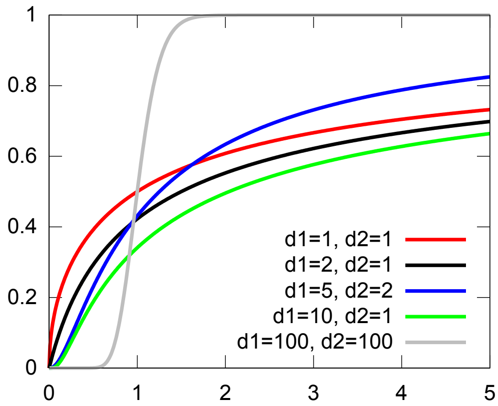


Fonction de répartition

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



Densité de probabilité



Fonction de répartition