

MATH-H204 - Calcul des probabilités et statistiques

Yves DE SMET

Résumé du cours

Rodrigue VAN BRANDE

19 juillet 2015

Table des matières

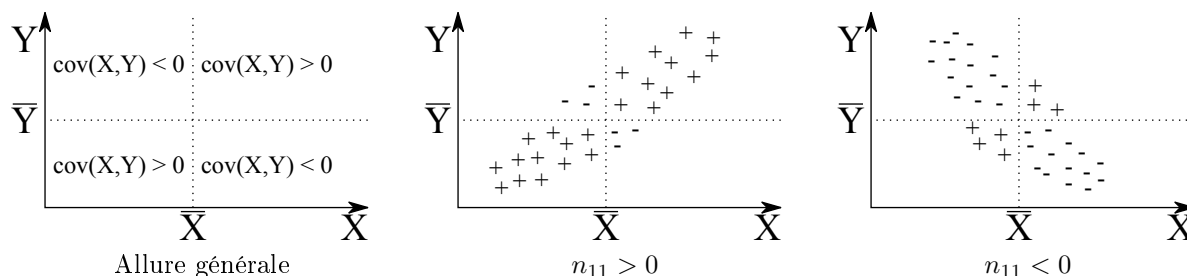
1	La statistique descriptive	3
1.1	Statistique descriptive en 1D	3
1.2	Statistique descriptive en 2D	3
1.2.1	Covariance	3
1.2.2	Le coefficient de corrélation	4
1.2.3	Les droites de régression	5
2	La théorie des probabilités	7
2.1	Probabilités	7
2.1.1	Axiomes de la théorie des probabilités	7
2.1.2	Probabilité conditionnelle et indépendance	7
2.1.3	Formule de Bayes	7
2.2	Variables aléatoires	7
2.2.1	Distribution d'une fonction monotone d'une variable aléatoire	7
2.2.2	Distribution de la somme de deux variables aléatoires	8
2.2.3	Distribution du produit de deux variables aléatoires	9
2.3	Variables aléatoires particulières	10
2.3.1	Variable binomiale $\mathcal{B}(n, p)$	10
2.3.2	Variable de Poisson \mathcal{P}_λ	10
2.3.3	Variable exponentielle négative	10
2.3.4	Variable Normale $\mathcal{N}(\mu, \sigma)$	10
2.3.5	Variable Khi^2	10
2.3.6	Variable Student t_n	10
2.3.7	Variable Snedecor $\mathcal{F}_{(m,n)}$	10
2.4	Théorèmes fondamentaux	10
2.4.1	Inégalité de Bienaymé-Tchebycheff	10
2.4.2	Théorème de Bernoulli ou loi des grands nombres	10
2.4.3	Théorème Central-Limite	11
2.4.4	Théorème de De Moivre	11
3	L'inférence statistique	12
4	Autres aides	13
4.1	Tableau du formulaire	13
4.2	Densité et répartition	13
4.3	Distributions	14

1 La statistique descriptive

1.1 Statistique descriptive en 1D

1.2 Statistique descriptive en 2D

1.2.1 Covariance



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

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

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

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

Équation du second degré, on calcule son Δ :

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

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 $ax + by + c = 0$, on a $ax_i + by_i + c = 0$.

On multiplie par $\frac{n_{ij}}{n}$ et on somme sur ij .

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

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

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

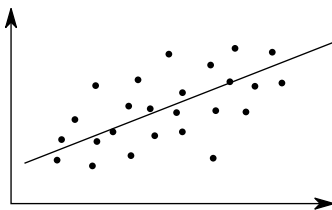
L'équation a la même forme que α , du coup...

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

1.2.2 Le coefficient de corrélation

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

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 (parallèles à l'axe y) des points observés à cette droite.

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

Dérivée par rapport à a .

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

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

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

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

$$\begin{aligned}
 \bar{x} (2) : n \bar{y} \bar{n} &= a n \bar{x}^2 + b n \bar{n} \\
 (1) - \bar{x} (2) : &
 \end{aligned}$$

2 La théorie des probabilités

2.1 Probabilités

2.1.1 Axiomes de la théorie des probabilités

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

2.1.2 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 :

$$P(A|B) = P(A)$$

alors

$$P(A \cap B) = P(A)P(B)$$

2.1.3 Formule de Bayes

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

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

2.2 Variables aléatoires

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

$$W = G(V)$$

	croissante	décroissante
$F_W(x)$	=	$\begin{aligned} &P(W \leq x) \\ &P(G(V) \leq x) \\ &P(V \leq G^{-1}(x)) \\ &F_V(G^{-1}(x)) \end{aligned}$
	=	$\begin{aligned} &P(V \geq G^{-1}(x)) \\ &1 - F_V(G^{-1}(x)) \\ &F'_V(G^{-1}(x)) \frac{-1}{G'(G^{-1}(x))} \end{aligned}$
$F'_W(x)$	=	$\begin{aligned} &f_W(x) = F'_V(G^{-1}(x)) \frac{1}{G'(G^{-1}(x))} \\ &= f_V(G^{-1}(x)) \frac{1}{G'(G^{-1}(x))} \end{aligned}$
	=	$\begin{aligned} &F'_V(G^{-1}(x)) \frac{-1}{G'(G^{-1}(x))} \\ &= f_V(G^{-1}(x)) \frac{-1}{G'(G^{-1}(x))} \end{aligned}$

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

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

On remplace par $\left\{ \begin{array}{l} \xi = u \\ \eta = v - u \\ J = \begin{pmatrix} \frac{\delta\xi}{\delta u} & \frac{\delta\xi}{\delta v} \\ \frac{\delta\eta}{\delta u} & \frac{\delta\eta}{\delta v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.1 - 0.(-1) = 1 \end{array} \right.$

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

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

ou si indépendant

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

2.2.3 Distribution du produit de deux variables aléatoires

$$\boxed{Z = VW}$$

Cas continu

$$\begin{aligned} F_Z(x) &= P(Z \leq x) \\ F_{V,W}(x) &= P(V \leq x, W \leq x) \\ &= \iint_{\xi, \eta \leq x} f_{(V,W)}(\xi, \eta) \, \delta\xi \, \delta\eta \end{aligned}$$

On remplace par $\left\{ \begin{array}{l} \xi = u \\ \eta = \frac{v}{u} \\ J = \begin{pmatrix} \frac{\delta\xi}{\delta u} & \frac{\delta\xi}{\delta v} \\ \frac{\delta\eta}{\delta u} & \frac{\delta\eta}{\delta v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{-1}{u^2} & \frac{1}{u} \end{pmatrix} = 1 \cdot \frac{1}{u} - 0 \cdot \frac{-1}{u^2} = \frac{1}{u} \end{array} \right.$

$$\begin{aligned} &= \iint_{v \leq x} f_{(V,W)}(u, v - u) \, \delta u \, \delta v \\ &= \int_{-\infty}^x \delta v \int_{-\infty}^{+\infty} f_{(V,W)}(u, v - u) \, \delta u \end{aligned}$$

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

ou si indépendant

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

2.3 Variables aléatoires particulières

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

2.3.2 Variable de Poisson \mathcal{P}_λ

2.3.3 Variable exponentielle négative

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

2.3.5 Variable Khi²

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

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

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

2.4.2 Théorème de Bernoulli 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 $n \rightarrow \infty$

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

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 $V = B(n, p)$

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

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

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

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 $\frac{V - \mu}{\sigma}$ tend vers une gaussienne $N(0, 1)$ lorsque $n \rightarrow \infty$.

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

Démonstration

???

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 $n \rightarrow \infty$.

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

Démonstration

On part donc avec le théorème de Central-Limite :

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

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

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

3 L'inférence statistique

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}_λ	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	$e^{\mu t + (\sigma^2 t^2)/2}$
$\chi_{(n)}^2$	n	$2n$	$(1 - 2t)^{-n/2}$
t_n	$0 \quad n > 1$	$\frac{n}{n-2} \quad n > 2$	aucun
$\mathcal{F}_{(m,n)}$	$\frac{n}{n-2} \quad n > 2$	$\frac{2n^2(n+m-2)}{m(n-2)(m-4)} \quad n > 2$	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}_λ	$P[\mathcal{P}_\lambda = k] = \frac{\lambda^k}{k!} e^{-\lambda}$	$\sum_{k=0}^x P[\mathcal{P}_\lambda = k]$
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]	$\begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$	$\begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$
$\mathcal{N}(\mu, \sigma)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$
$\chi_{(n)}^2$	$\begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$	$\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$
t_n	Densité indépendante de σ	
$\mathcal{F}_{(m,n)}$	Densité indépendante de σ	

4.3 Distributions

