Part I

Éléments de probabilité

Inférence

Stat	Proba
Observation	Réflexion (anticipation)
Moyenne $\sum x_i \times f_i$	espérance $\rightarrow \sum_i x_i P(X_i = x_i)$ $\rightarrow \int x f_X(x)\,dx$
fréquence	probabilité
Variance / écart-type / quantiles . . . (même dénomination)	

X variable aléatoire
$f_X(x)$ sa densité
$F_X(x)$ sa fonction de répartition.

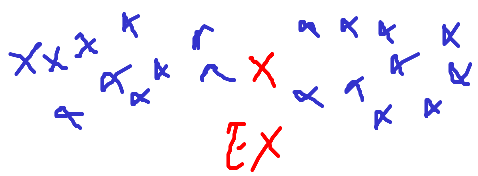
1 Rappels : variables aléatoires élémentaires

On rappelle que l'on peut définir les notions d'espérance de variable aléatoire comme la moyenne statistique ainsi que celles d'écart-type et de variance de manière analogue aux notions vues dans le cadre des statistiques descriptives. On rappelle les propriétés suivantes.

- La linéarité de l'espérance $E[aX+bY+c] = aEX+bEY+c$, avec X, Y des variables aléatoires et a, b, c des réels.

Sans condition sur X et Y !

La variance est un écart moyen entre les valeurs de X et EX.



Par translation des valeurs de X, la variance ne change pas

Par dilatation des valeurs $(\times a)$, la variance évolue

- La formule fondamentale : $Var(aX + b) = \underline{a^2} VarX$.
- Si X et Y sont indépendantes, alors $Var(X + Y) = Var(X) + Var(Y)$.

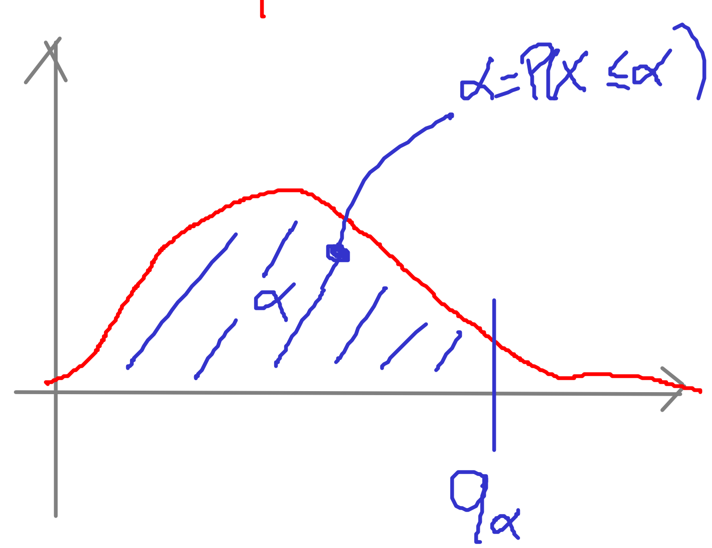
On rappelle qu'on définit la fonction de répartition d'une variable aléatoire par

$$F_X(t) = P(X \leq t).$$

Le quantile d'ordre α noté q_α est la valeur de t telle que

$$F_X(t) = \alpha$$

de manière équivalente, $F_X(q_\alpha) = P(X \leq q_\alpha) = \alpha$.



1.1 Loi de Bernouilli

X modélise une expérience à deux issues :

- 1 si réussite avec proba p
- 0 si échec avec proba $q = 1 - p$

$EX = p$ $\qquad$ $\text{Var}\, X = p(1-p)$

$X \sim \mathcal{B}(p)$

X compte le nombre de succès à n expériences aléatoires identiques et indépendantes à 2 issues = – réussite (avec proba p)
– échec

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

1.2 Loi binomiale

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$\xi_1 \dots \xi_n$ des lois $\mathcal{B}(p)$ indépendantes.

$$X = \sum_{i=1}^{n} \xi_i$$

indépendance

$$EX = E\left[\sum \xi_i\right] = \sum_{1}^{n} E\xi_i = np \quad // \quad Var(X) = Var\left(\sum \xi_i\right) = \sum_{i=1}^{n} Var(\xi_i)$$

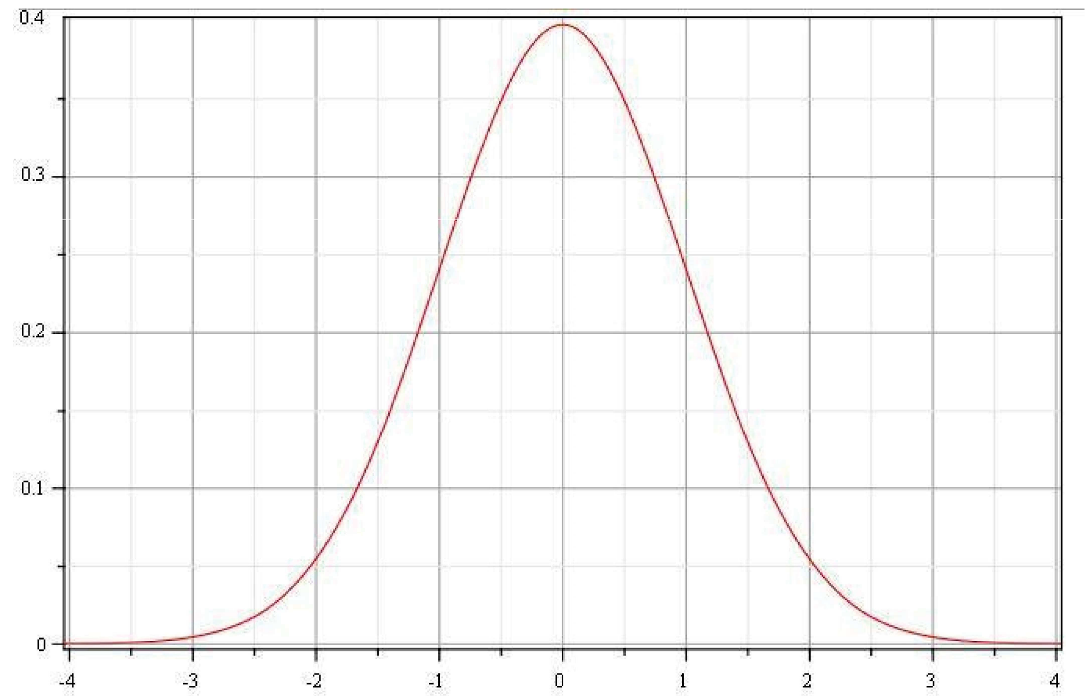
$$= np(1-p)$$

1.3 Loi normale

1.3.1 Loi normale centrée réduite

La loi normale centrée réduite, notée $\mathcal{N}(0,1)$ est la loi de référence. Elle a pour densité

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



Son espérance vaut 0 et son écart-type 1.

1.3.2 Loi normale d'espérance m, d'écart-type σ

Une telle loi, notée $\mathcal{N}(m, \sigma)$ ou plus rarement $\mathcal{N}(m, \sigma^2)$ est obtenue par translation et dilatation de la loi normale centrée réduite :

$$\mathcal{N}(m, \sigma) = \sigma \mathcal{N}(0, 1) + m.$$

Une loi normale est donc entièrement déterminée par son espérance m et son écart-type σ. On observera les densités suivantes.

$X \sim \mathcal{N}(m, \sigma^2)$, $\exists Y \sim \mathcal{N}(0,1)$ telle que $X = \sigma Y + m$

$EX = E[\sigma Y + m] = \sigma \overbrace{EY}^{0} + m = m$

$\mathrm{Var}\, X = \mathrm{Var}(\sigma Y + m) = \sigma^2 \mathrm{Var}\, Y = \sigma^2 \times 1 = \sigma^2$

On a $Y = \frac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$. On a centré-réduit Y.

$X \sim \mathcal{N}(m_X, \sigma_X^2)$

$Y \sim \mathcal{N}(m_Y, \sigma_Y^2)$

Theorem 1.1 *Soit X suivant une loi normale d'espérance m_X, et décart-type σ_X, et Y suivant une loi normale d'espérance m_Y, et décart-type σ_Y. En supposant les variables $\boxed{X \text{ et } Y \text{ indépendantes,}}$ on peut affirmer que $X + Y$ suit une loi normale d'espérance $m_X + m_Y$, et décart-type $\sqrt{\sigma_X^2 + \sigma_Y^2}$.*

$X+Y \sim \mathcal{N}(m, \sigma)$

$m = E(X+Y) = EX + EY = m_X + m_Y$

$\sigma = \sqrt{Var(X+Y)} = \sqrt{Var X + Var Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$

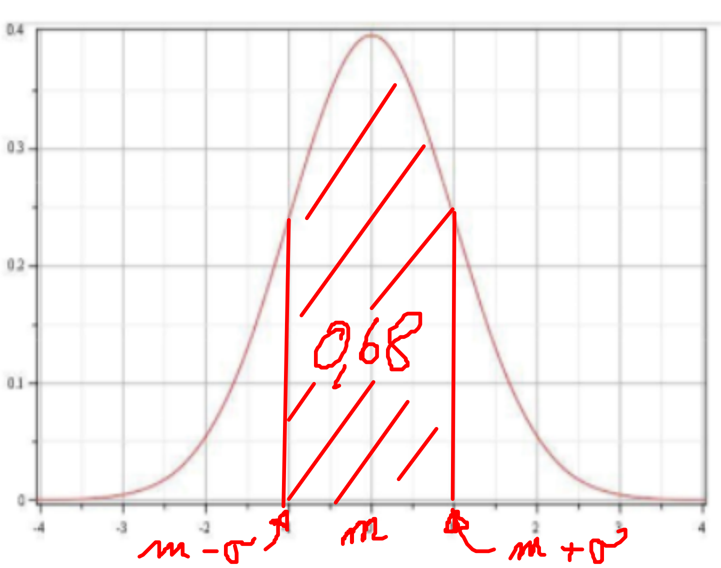
indépendance

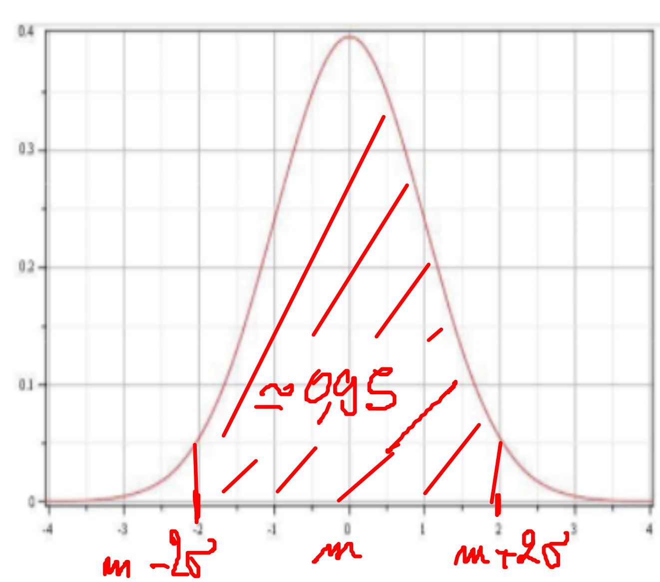
Ce résultat est vrai pour toutes combinaison linéaires de lois normales indépendantes.

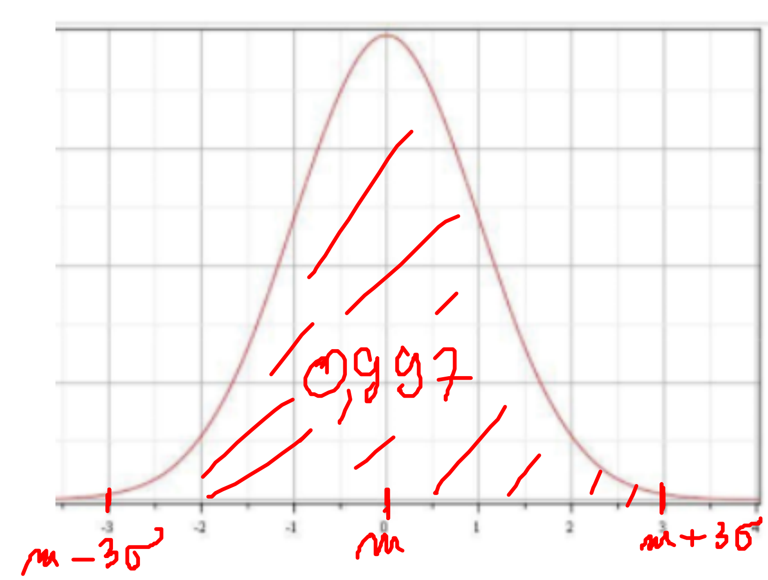
On aura pour ordre d'idée les intervalles remarquables :

$$
\begin{aligned}
P(X \in [m-\sigma, m+\sigma]) &\approx 0.68,\\
P(X \in [m-2\sigma, m+2\sigma]) &\approx 0.95,\\
P(X \in [m-3\sigma, m+3\sigma]) &\approx 0.997.
\end{aligned}
$$

vraie valeur : 1,96







2 Lois des statistiques

On propose, dans cette section, une sélection de lois très importantes en statistiques, définies à partir de sommes de lois normales centrées réduites.

2.1 Loi du χ^2

Si $X_1; \cdots ; X_n$ sont des variables aléatoires indépendantes de même loi $\mathcal{N}(0;1)$, alors

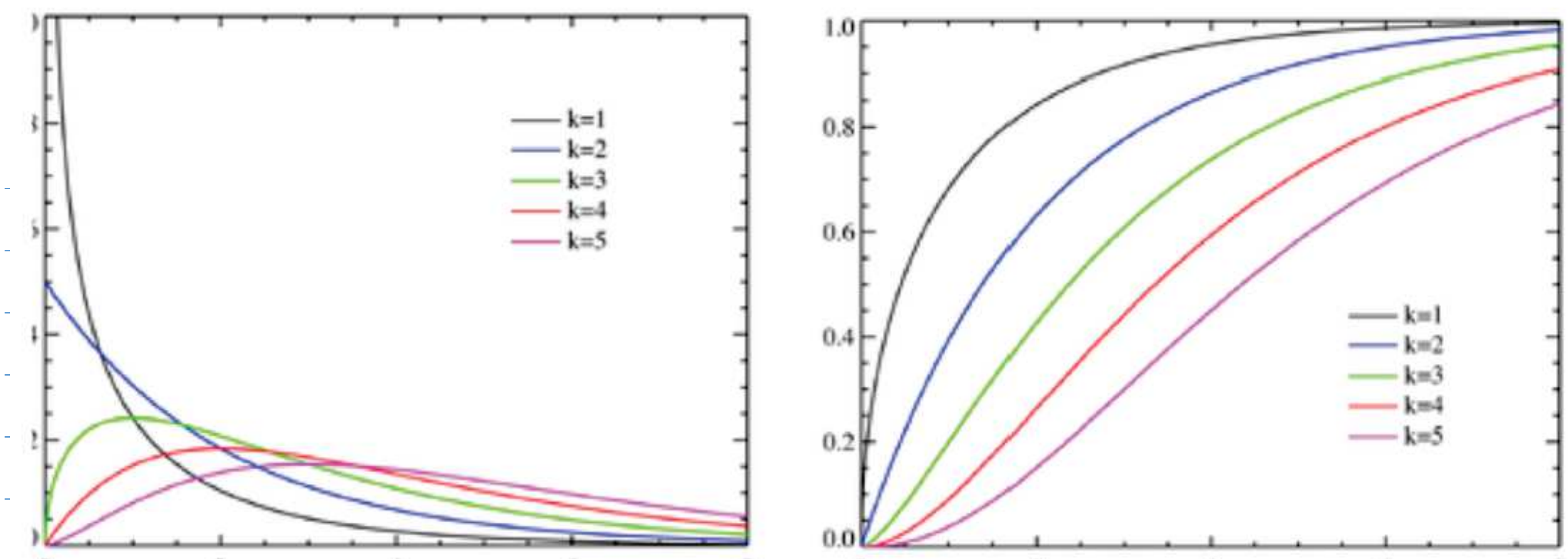
$$X_1^2 + \cdots + X_n^2$$

suit une loi du χ^2 à n degrés de libertés, notée $\chi^2(n)$.

la somme des carrés

On a alors

$$EX = n, \qquad VarX = 2n.$$



Proposition 2.1 *Si $X_1 \sim \chi^2(n)$ et $X_2 \sim \chi^2(m)$ sont deux variables aléatoires indépendantes, alors $X_1 + X_2 \sim \chi^2(n+m)$.*

$\xi_i \sim \mathcal{N}(0,1)$ indépendantes

↳ $X_1 = \xi_1^2 + \cdots + \xi_n^2$ $\quad X_2 = \xi_{n+1}^2 + \cdots + \xi_{n+m}^2$

$$X_1 + X_2 = \underbrace{\xi_1^2 + \cdots + \xi_n^2 + \xi_{n+1}^2 + \cdots + \xi_{n+m}^2}_{\mathcal{N}(0,1) \text{ indépendantes}} \sim \chi^2(n+m)$$

2.2 Loi de Student

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

Soient U et V deux variables aléatoires indépendantes, telles que $U \sim \mathcal{N}(0;1)$ et $V \sim \chi^2(k)$. La variable

$$X = \frac{U}{\sqrt{V/k}}$$

indépendantes.

suit une loi de Student à k degrés de libertés, notée $t(k)$.

ddl identique à $V \sim \chi^2(k)$

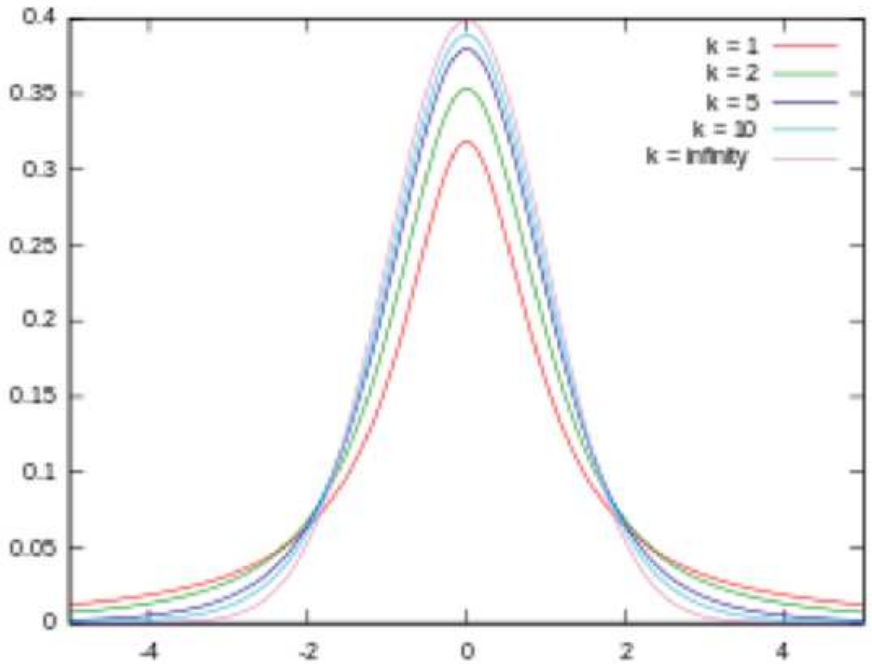
$\chi^2(k)$ ddl

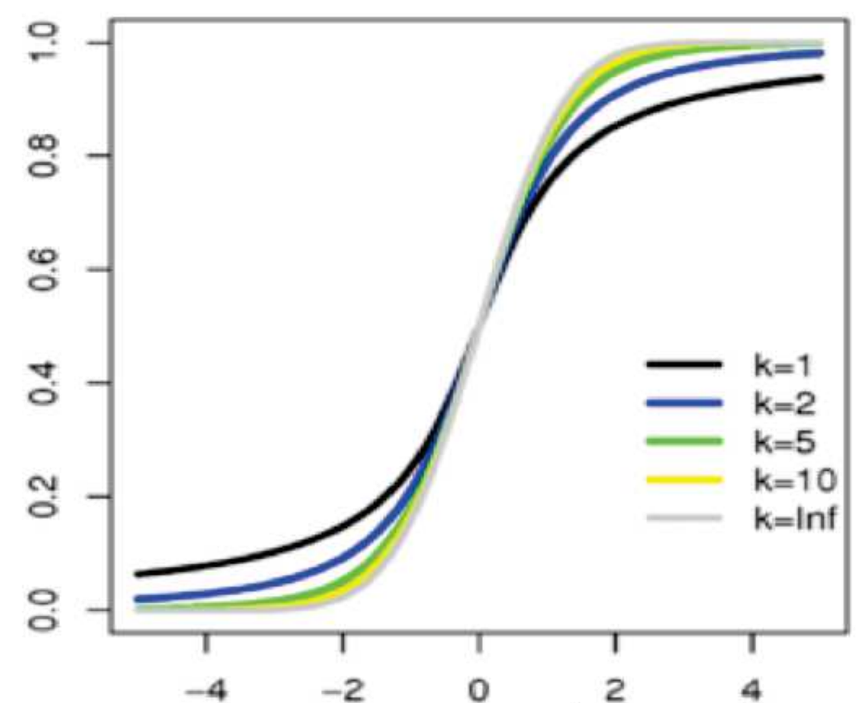
$X \sim t(k)$

On a alors

$$EX = 0, \qquad VarX = \frac{k}{k-2},.$$

pour $k > 2$.


k = 1
k = 2
k = 5
k = 10
k = infinity


k=1
k=2
k=5
k=10
k=Inf

2.3 Loi de Fisher

Soient U et V deux variables aléatoires indépendantes, telles que $U \sim \chi^2(d_1)$ et $V \sim \chi^2(d_2)$. La variable

$$F = \frac{U/d_1}{V/d_2}$$

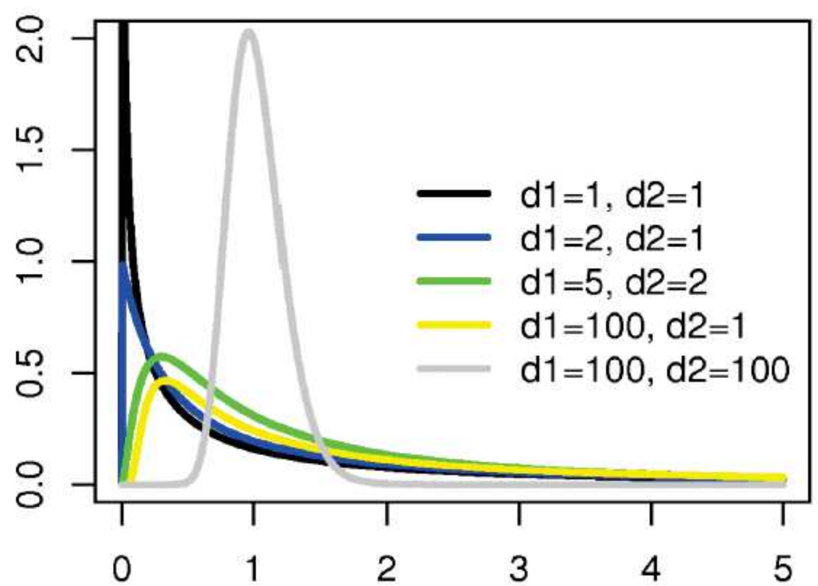
rapport de deux lois du χ^2 divisées par leur ddl.

suit une loi de Fisher–Snedecor de paramètres d_1 et d_2, notée $\mathcal{F}(d_1, d_2)$ ou F_{d_1,d_2}.

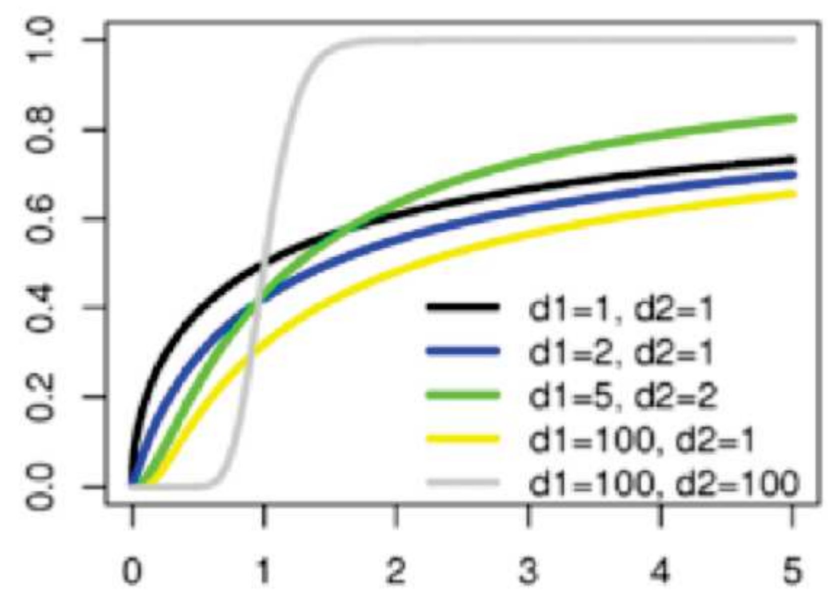
F positive.

On a alors

$$EX = \frac{d_2}{d_2 - 1}, \quad d_2 > 2; \qquad\qquad VarX = \frac{2d_2^2(d_1 + d_2 - 2)}{d_1(d_2 - 2)2(d_2 - 4)}, \quad d_2 > 4.$$



2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
d1=1, d2=1
d1=2, d2=1
d1=5, d2=2
d1=100, d2=1
d1=100, d2=100



1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
d1=1, d2=1
d1=2, d2=1
d1=5, d2=2
d1=100, d2=1
d1=100, d2=100

Proposition 2.2

$X \sim \mathcal{F}_{d_1,d_2}$

- *Si* $X \sim \mathcal{F}_{d_1,d_2}$ *alors* $1/X \sim \mathcal{F}_{d_2,d_1}$.
- *Si* $X \sim t_k$ *alors* $X^2 \sim \mathcal{F}_{1,k}$.
- *Si* $X \sim \mathcal{N}(0;1)$ *alors* $X^2 \sim \mathcal{F}_{1,\infty}$.

Indépendanc !

$X = \frac{U/d_1}{V/d_2}$ $\quad U \sim \chi^2(d_1) \quad V \sim \chi^2(d_2)$

$1/X = \frac{V/d_2}{U/d_1} \sim \mathcal{F}_{d_2 d_1}$

$X \sim t_k$ $\quad X = \frac{U}{\sqrt{V/k}}$ $\quad \mathcal{N}(0,1)$ $\quad V \sim \chi(k)$

$U^2 \sim \chi^2(1)$

$X^2 = \frac{U^2/1}{V/k} \sim \mathcal{F}_{1,k}$

Rappels

- Loi de Bernoulli : $\xi \sim \mathcal{B}(p)$ → 1 si réussite avec proba p, → 0 sinon (avec proba $1-p$)

  $E\xi = p$

  $Var(\xi) = p(1-p)$

- Loi binomiale $X \sim \mathcal{B}(n, p)$ $\xi_1 \dots \xi_n \sim \mathcal{B}(p)$ indépendantes

  alors $X = \sum_{i=1}^{n} \xi_i$ X compte le nombre de réussites à n expériences de Bernoulli indépendantes

  $EX = np$ $Var X = np(1-p)$

- Lois normales $\mathcal{N}(m, \sigma^2)$; loi normale centrée-réduite $\mathcal{N}(0,1)$

- $X_1, \dots, X_d$ des lois normales centrées-réduites indépendantes $\mathcal{N}(0,1)$

  $Y = \sum_{i=1}^{d} X_i^2 \sim \chi_2(d)$ degrés de liberté

- $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi_2(d)$ indépendantes $\frac{X}{\sqrt{Y/d}} \sim t(d)$ student, degrés de liberté.

- $Y_1 \sim \chi_2(d_1)$ $Y_2 \sim \chi_2(d_2)$ indépendantes $\frac{Y_1/d_1}{Y_2/d_2} \sim \mathcal{F}_{d_1,d_2}$ Fisher, degrés de liberté

Et les propriétés élémentaires :

- $Y_1 \sim \chi_2(d_1)$ et $Y_2 \sim \chi_2(d_2)$ indépendantes

$$Y_1 = \sum_{i=1}^{d_1} X_i^2 \qquad Y_2 = \sum_{i=1}^{d_2} X_{d_1+i}^2 \qquad Y_1 + Y_2 = \sum_{i=1}^{d_1+d_2} X_i^2 \sim \chi_2(d_1+d_2)$$

$\mathcal{N}(0,1)$ i.i.d.

$\mathcal{N}(0,1)$ i.id

- $X \sim t(d)$; $\exists\, U \sim \mathcal{N}(0,1)$ et $Y \sim \chi_2(d)$ indépendantes

$X = \dfrac{U}{\sqrt{Y/d}}$. On a alors $X^2 = \dfrac{U^2}{Y/d}$ et $U^2 \sim \chi_2(1)$

Donc $X^2 = \dfrac{U^2/1}{Y/d}$ est un rapport de χ^2 indépendantes

et donc $X^2 \sim \mathcal{F}_{1,d}$

- $X \sim \mathcal{F}_{d_1,d_2}$ - $\exists\, U \sim \chi^2(d_1)$ et $V \sim \chi^2(d_2)$ indépendantes

avec $X = \dfrac{U/d_1}{V/d_2}$. $\dfrac{1}{X} = \dfrac{V/d_2}{U/d_1} \sim \mathcal{F}_{d_2,d_1}$

3 Théorèmes limites

Il y a plusieurs types de convergences pour les suites de variables aléatoires. On rappelle la définition de la convergence en loi (suffisante pour tout ce dont nous auront besoin).

Fonction de répartition : $F_X(x) = P(X \leq x)$.

X_n suite de variable, $F_{X_n}(x) = P(X_n \leq x)$

3.1 Convergence en loi

Soit (X_n) une suite de variables aléatoires. On dit que (X_n) converge en loi vers X si

$$\lim_{x \to +\infty} F_{X_n}(x) = F_X(x), \qquad \forall x \in \mathbb{R}.$$

3.2 Loi des grands nombres

Soit (X_n) une suite de variables aléatoires identiques et indépendantes et qui admettent la même espérance μ et le même écart-type. Alors:

$$\bar{X}_n = \frac{1}{n}\sum_{i=0}^{n} X_i \to \mu.$$

$X_1 \ldots X_n$ i.i.d. (Jet de dé).

La moyenne des variables aléatoires $\frac{X_1 + \cdots + X_n}{n} \longrightarrow EX_1$.

En faisant la moyenne des résultats à une répétition d'expériences aléatoires, l'aléa tend à disparaître.

$X_1 \ldots X_n$ des lois $\mathcal{B}(p)$ indépendantes

$\hookrightarrow$ 1 pour opinion favorable (avec proba p.) → pas forcément connue

$\hookrightarrow$ 0 sinon

$$\frac{X_1 + \cdots + X_n}{n} \longrightarrow EX_1 = p.$$

Si on a fait n observation

Pour la personne 1 : $X_1 = x_1$ ← réponse 1 ou 0.

$\vdots$

personne n : $X_n = x_n$

les réponses

$\hookrightarrow f = \frac{x_1 + \cdots + x_n}{n} \simeq p$ ← probabilité.

fréquence observée = stat

$\Rightarrow$ Assimilation probabilité - fréquence.

3.3 Le théorème central limit (TCL)

Le théorème central limite établit la convergence en loi de la somme d'une suite de variables aléatoires vers la loi normale. Intuitivement, ce résultat affirme que toute somme de variables aléatoires indépendantes et identiquement distribuées (**i.i.d.** dans la suite) tend vers une variable aléatoire gaussienne.

Theorem 3.1 *Soit (X_n) une suite de variables aléatoires i.i.d.. Supposons que pour tout $i \in \mathbb{N}$, les variables X_i admettent une espérance μ et un écart-type σ. Alors la suite de variables aléatoires centrées réduites suivante converge en loi :*

$$\frac{\sum_{i=0}^{n} X_i - n\mu}{\sqrt{n}\sigma} \to \mathcal{N}(0,1).$$

À retenir : $X_1 \dots X_n$ i.i.d $EX_i = \mu$; $\sigma_{X_i} = \sigma$.

Pour n grand : $Y_n = \sum_{i=1}^{n} X_i \simeq \mathcal{N}(\mu_n, \sigma_n)$ avec

$$\begin{cases} \mu_n = EY_n = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} EX_i = n\mu \\ \sigma_n = \sqrt{Var\left(\sum X_i\right)} \underset{\text{indépendance}}{=} \sqrt{\sum_{i=1}^{n} \underbrace{Var(X_i)}_{\sigma^2}} = \sqrt{n\sigma^2} = \sqrt{n}\,\sigma \end{cases}$$

Formulation en terme de limite :

On centre et on réduit la somme :
$$\frac{\sum_{i=0}^{n} X_i - \mu_n}{\sigma_n} = \frac{\sum_{i=0}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \longrightarrow \mathcal{N}(0,1)$$

Si les $X_i \sim \mathcal{B}(p)$ indépendantes. $\quad EX_i = p \quad \sigma_{X_i} = \sqrt{p(1-p)}$

TCL : $$\frac{\sum_{i=0}^{n} X_i - np}{\sqrt{n}\sqrt{p(1-p)}} \longrightarrow \mathcal{N}(0,1).$$

$$Y_n = \sum_{i=1}^{n} X_i \simeq \mathcal{N}(\underbrace{np}_{EY_n}, \underbrace{\sqrt{np(1-p)}}_{\sigma_{Y_n}})$$

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

3.4 Approximation d'une loi binomiale par une loi normale

Theorem 3.2 *Soit X une loi binômiale $\mathcal{B}(n,p)$. Si n est grand, on peut approximer X par une loi normale dont l'espérance est celle de X, et d'écart-type celui de X.*

$= np \qquad = \sqrt{np(1-p)}$

$$\mathcal{B}(n,p) \simeq \mathcal{N}\left(np, \sqrt{np(1-p)}\right) \text{ si},$$

En pratique, on aura les critères $n \geq 50$ et :

- $0,4 \leq p \leq 0,6$;
- ou ou $npq \geq 18$;
- ou ou $np > 5$ et $nq > 5$.

$q = 1 - p$

Problème : $X \sim \mathcal{B}(n,p) \simeq Y \sim \mathcal{N}(np, \sqrt{np(1-p)})$ $\Rightarrow$ On a parlé de convergence des fonctions de répartition.

discrète : ne charge que les entiers

$P(Y=1)=0$

loi à densité

l'approximation est bonne en considérant

$P(X \leq x) \simeq P(Y \leq x)$

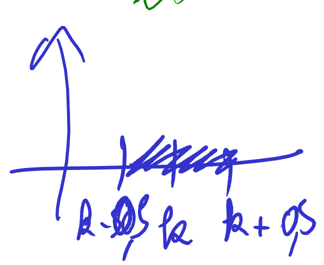
Soit Y la loi normale approximante. On a alors

$$P(a \leq X \leq b) \approx P(a \leq Y \leq b).$$

Si on souhaite faire l'approximation de $P(X = k)$, Y étant continue, on effectuera la correction de continuité

$$P(X = k) \approx P(k - 0,5 \leq Y \leq k + 0,5).$$

$$\underbrace{P(X=k)}_{\mathcal{B}(n,p)} = \underbrace{P(X \leq k+0,5) - P(X \geq k-0,5)}_{\text{fonctions de répartitions}} = P(k-0,5 \leq X \leq k+0,5)$$



$\simeq P(k-0,5 \leq Y \leq k+0,5)$

$\Rightarrow$ Correction de continuité.

3.5 Autres théorèmes limites

On a les limites suivantes.

Proposition 3.3

- *Si* $X \sim \mathcal{F}_{d_1;d_2}$ *alors*
$$lim_{d2\to\infty} d_1 X \to \chi^2(d_1).$$
- *La loi de Student converge vers la loi normale centrée réduite :*
$$t_k \xrightarrow[k\to\infty]{} \mathcal{N}(0;1).$$