Simulations des phénomènes aléatoires
3LM246 : CHAPITRE 2.

Raphaël KRIKORIAN
Irina KOURKOVA, 2017–2018
Université Paris 6

6 février 2018

Chapitre 1

Méthode de Monte Carlo

1.1 Introduction

La simulation de v.a indépendantes suivant une loi donnée (p.ex uniforme) a une application importante : la calcul approché d'integrales de fonctions en dimension grande. Pour fixer les idées, soit $f : [0,1]^d \to \mathbf{R}$ une fonction "régulière" définie sur le cube d-dimensionnel. On se propose de calculer l'intégrale $I(f) = \int_{[0,1]^d} f(x_1, \ldots, x_d) dx_1 \cdots dx_d$. Quand d égale 1 et que f est continue, les sommes

$$\frac{S_n(f)}{n} = \sum_{k=0}^{n-1} \frac{1}{n} f(\frac{k}{n}) \tag{1.1}$$

constituent une approximation de f dont on peut contrôler la qualité de la fa con suivante : Si $\delta(\cdot)$ est le module de continuité [1] de f on a

$$|I(f) - \frac{S_n(f)}{n}| \leq \delta(\frac{1}{n}).$$

Si f est plus régulière, on peut trouver des procédés plus efficaces. Par exemple, la méthode des trapèzes consiste à approcher le graphe de f par une fonction continue affine par morceaux : On pose

$$\frac{S_n}{n} = \frac{1}{2n}\Big((f_0 + f_1) + (f_1 + f_2) + \cdots + (f_{n-1} + f_n)\Big), \quad f_i = f(i/n)$$

Si f est C^2

$$|I(f) - \frac{S_n(f)}{n}| \leq \frac{C}{n^2}.$$

1. il est défini de la fa con suivante : $\delta(\epsilon) = \sup_{|x-y|\leq\epsilon}(|f(x) - f(y)|)$. En particulier, $|f(x) - f(y)| \leq \delta(|x - y|)$. Quand f est C^1, $\delta(\epsilon) \leq (\max_{[0,1]} |f'|) \cdot \epsilon$

La méthode de Simpson consiste à approcher f par une fonction continue et polynomiale de degré 2 par morceaux ; si n est pair on pose

$$\frac{S_n}{n} = \frac{1}{3n}\Big((f_0+4f_1+f_2)+(f_2+4f_3+f_4)+\cdots+(f_{n-2}+4f_{n-1}+f_n)\Big), \quad f_i = f(i/n);$$

si f est de classe C^4

$$|I(f) - \frac{S_n(f)}{n}| \leq \frac{C}{n^4}.$$

En dimension supérieure ($d \geq 2$) la généralisation de (1.1) est

$$\frac{S_n(f)}{n} = \frac{1}{n^d}\sum_{k_1=0}^{n-1}\cdots\sum_{k_d=0}^{n-1} f(\frac{k_1}{n},\ldots,\frac{k_d}{n})$$

et l'erreur que l'on commet dans l'évaluation de $I(f)$ est encore de la forme (si f est C^1)

$$|I(f) - \frac{S_n(f)}{n}| \leq C\frac{1}{n}.$$

Ainsi, pour obtenir une erreur de l'ordre de $C/n \approx \epsilon$ il faut calculer les valeurs de f en $n^d \approx (1/\epsilon)^d$ points. Quand d est grand (par exemple de l'ordre de 100) le temps de calcul est vite rédhibitoire. En outre, les méthodes que nous venons de décrire nécessitent de travailler avec une fonction f suffisamment différentiable.

Il est en fait possible de surmonter ces deux difficultés à condition d'accepter de travailler avec un algorithme probabiliste, la méthode de Monte Carlo.

1.2 Description de la méthode

Le principe repose sur la loi des grands nombres et le Théorème *Central Limit*. Si f est une fonction L^1 sur le pavé d-dimensionnel $[0,1]^d$ et si $X_1,\ldots,X_n,\ldots$ est une suite de *vecteurs* aléatoires uniformément distribués sur $[0,1]^d$ alors la suite $Y_1 = f(X_1),\ldots,Y_n = f(X_n),\ldots$ est une suite de variables aléatoires indépendantes et de même loi ($\mathbf{P}(f(X_i) \in A) = \mathbf{P}(f(X_1) \in A)$ pour tout intervalle ou borélien A de $\mathbf{R}$). La loi des grands nombres nous enseigne que pour $\mathbf{P}$-presque tout $\omega \in \Omega$

$$\lim_{n\to\infty}\frac{1}{n}\Big(f(X_1(\omega)) + \cdots + f(X_n(\omega))\Big) = \mathbf{E}(f(X_1)).$$

Calculons $\mathbf{E}(f(X_1))$. D'après la formule de transfert

$$\mathbf{E}(f(X_1)) = \int_{\mathbf{R}^d} f(x_1, \ldots, x_d)\rho(x_1, \ldots, x_d)dx_1 \cdots dx_d$$

où $\rho(x_1, \ldots, x_d) = \mathbf{1}_{[0,1]^d}(x_1, \ldots, x_d)$ est la densité de la loi uniforme sur $[0,1]^d$. On a donc

$$\mathbf{E}(f(X_1)) = \int_{[0,1]^d} f(x_1, \ldots, x_d)dx_1 \cdots dx_d = I(f)$$

et par conséquent $\mathbf{P}$-presque sûrement

$$\lim_{n\to\infty} \frac{1}{n}\Big(f(X_1(\omega)) + \cdots + f(X_n(\omega))\Big) = \int_{[0,1]^d} f(x_1, \ldots, x_d)dx_1 \cdots dx_d.$$

On est donc sûr qu'avec probabilité 1 la moyenne précédente

$$\frac{1}{n}S_n(f) := \frac{1}{n}\Big(f(X_1(\omega)) + \cdots + f(X_n(\omega))\Big)$$

converge vers l'intégrale

$$I(f) = \int_{[0,1]^d} f(x_1, \ldots, x_d)dx_1 \cdots dx_d.$$

La méthode du calcul approché de l'intégrale consiste donc à calculer $I(f)$ comme $\frac{1}{n}\Big(f(X_1(\omega)) + \cdots + f(X_n(\omega))\Big)$.

Il est important de savoir à quelle vitesse la convergence précédente a lieu si $n \to \infty$. Supposons f de carré intégrable. Le théorème Central Limit nous fournit la réponse suivante : Puisque $Y_1 = f(X_1), \ldots, Y_n = f(X_n), \ldots$ est une suite indépendante de v.a.r de même loi et de carré intégrable on sait que

$$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\Big(f(X_1) + \cdots + f(X_n)\Big) - \mathbf{E}(f(X_1))\right), \qquad \sigma^2 = \mathbf{Var}(f(X_1))$$

converge en loi vers une loi normale centrée réduite. En d'autres termes, si on pose

$$\begin{aligned}\sigma^2 &= \mathbf{E}(f(X_1)^2) - (\mathbf{E}(f(X_1)))^2 \\ &= I(f^2) - I(f)^2\end{aligned}$$

on a pour tout $a > 0$

$$\lim_{n\to\infty} \mathbf{P}\left(|\frac{1}{n}S_n(f) - I(f)| < \frac{\sigma a}{\sqrt{n}}\right) = \int_{-a}^{a} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx.$$

Observons que

$$\int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \approx 0.95, \qquad \text{pour } a = 1.96 \approx 2$$

et

$$\int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \approx 0.99, \qquad \text{pour } a = 2.6$$

En conclusion, si on sait évaluer (même grossièrement) σ on peut dire avec probabilité grande que $S_n(f)/n$ est une approximation de $I(f)$ à $\sigma a/\sqrt{n}$-près ($a = 2$ ou $a = 2.6$) pourvu que n soit assez grand. Par exemple, si ϵ est assez petit, $n = (2\sigma/\epsilon)^2$, $S_n(f)/n$ est une approximation de $I(f)$ à ϵ-près avec probabilité de 0.95. Remarquons que le calcul de $S_n(f)$ nécessite le calcul de $(2\sigma/\epsilon)^2$ valeurs de f et que ce nombre est *indépendant* de la dimension d de l'espace sur lequel on travaille. C'est un avantage considérable par rapport aux méthodes décrites dans la première section. En revanche, la faiblesse de la méthode réside dans le fait que :

1) le TCL n'est vrai qu'asymptotiquement : le "assez petit" (pour ϵ) ou "assez grand" (pour n) n'est *a priori* pas explicite ;
2) la méthode nécessite d'avoir une estimée raisonnable sur σ et suppose donc que l'on sache déjà calculer $I(f)$ et $I(f^2)$.

Ces deux problèmes admettent chacun une solution au moins d'un point de vue théorique. Leur solution donnée ci-dessous **ne sera pas demandée à l'examen.** C'est un complément de cours optionnel. La solution théorique du premier problème réside dans le résultat suivant :

Théorème 1.2.1 (Berry-Essen) *Soit $X_1, \ldots, X_n, \ldots$ une suite de v.a, i.i.d. centrées ($\mathbf{E}(X_n) = 0$) telles que $\mathbf{E}(X_1^2) = \sigma^2$ et $\rho = \mathbf{E}(|X_1|^3) < \infty$. Si on note F_n la fonction de répartition de $\frac{S_n}{\sigma\sqrt{n}}$ on a pour tout $x \in \mathbf{R}$*

$$\left| F_n(x) - \int_{-\infty}^{x} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \right| \leq \frac{3\rho}{\sigma^3\sqrt{n}}.$$

On a donc

Théorème 1.2.2 *Pour tout a*

$$\left| \mathbf{P}\left(|\frac{1}{n} S_n(f) - I(f)| < \frac{\sigma a}{\sqrt{n}} \right) - \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right| \leq 6 \left(\frac{|f|_{C^0}}{\sigma} \right)^3 \frac{1}{\sqrt{n}}.$$

Pour résoudre le second problème il est naturel de remplacer σ^2 par

$$\Sigma_n^2 = \frac{1}{n} S_n(f^2) - \left(\frac{1}{n} S_n(f) \right)^2.$$

Théorème 1.2.3

$$\lim_{n\to\infty} \mathbf{P}\left(|\frac{1}{n}S_n(f) - I(f)| < \frac{\Sigma_n a}{\sqrt{n}}\right) = \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Démonstration.— Notons

$$Z_n = \frac{\sqrt{n}}{\sigma}(\frac{S_n}{n} - I(f)), \qquad \tilde{Z}_n = \frac{\sqrt{n}}{\Sigma_n}(\frac{S_n}{n} - I(f)), \qquad Y_n = \frac{\sigma}{\Sigma_n}$$

de fa con que $\tilde{Z}_n = Y_n \cdot Z_n$. Puisque Σ_n converge p.s vers σ, il suffit de démontrer que si Z_n converge en loi vers Z et si Y_n converge p.s vers 1 alors $Z_n \cdot Y_n$ converge en loi vers Z. Démontrons donc le lemme suivant :

Lemme 1.2.1 *Si U_n converge en loi vers U et si V_n/U_n converge en loi vers 1 alors V_n converge en loi vers U*

Démonstration.— Comme

$$(V_n > t) \subset (U_n > \frac{t}{1+\epsilon}) \cup (\frac{V_n}{U_n} > 1+\epsilon), \qquad (U_n > t(1+\epsilon)) \subset (V_n > t) \cup (\frac{U_n}{V_n} > 1+\epsilon)$$

on a

$$F_{U_n}(\frac{t}{1+\epsilon}) - \mathbf{P}(\frac{V_n}{U_n} > 1+\epsilon) \leq F_{V_n}(t), \qquad F_{V_n}(t) \leq F_{U_n}(t(1+\epsilon)) + \mathbf{P}(\frac{U_n}{V_n} > 1+\epsilon)$$

et donc puisque $\frac{U_n}{V_n}$ et $\frac{V_n}{U_n}$ converge en loi vers 1 [2]et que U_n converge en loi vers U on a pour tout t et tout $\epsilon > 0$ tels que $t/(1+\epsilon)$, $t(1+\epsilon)$ soient points de continuité de F_U

$$F_U(\frac{t}{1+\epsilon}) \leq \liminf_{n\to\infty} F_{V_n}(t) \leq \limsup_{n\to\infty} F_{V_n}(t) \leq F_U(t(1+\epsilon)).$$

Soit t un point de continuité de F_U ; il est possible de trouver ϵ aussi petit qu'on veut de fa con que $t(1+\epsilon)^{\pm 1}$ soient points de continuité de F_U. Faisant $\epsilon \to 0$ sur ces ϵ on voit que $\lim_{n\to\infty} F_{V_n}(t)$ existe et vaut $F_U(t)$ pour tout t point de continuité de F_U : c'est la définition de la convergence en loi de V_n vers U.

□

Le théorème résulte alors du lemme et du fait que si la suite (Y_n) converge presque sûrement vers 1 alors elle converge aussi en loi vers 1.

□

2. Si X_n converge en loi vers 1 alors X_n^{-1} converge aussi en loi vers 1 : en effet pour tout $t > 0$, $\mathbf{P}(X_n > t) = \mathbf{P}(0 < X_n^{-1} < t^{-1})$, c'est-à-dire $1 - F_{X_n}(t) = F_{X_n^{-1}}(1/t) - \mathbf{P}(X_n^{-1} = 1/t)$. Donc pour $s = 1/t > 0$ en dehors d'un ensemble dénombrable $F_{X_n^{-1}}(s)$ converge quand $n \to \infty$ vers $1 - \mathbf{1}_{s^{-1}>1}(s) = \mathbf{1}_{s\geq 1}(s)$ et le même résultat est vrai pour $s < 0$ (on trouve 0). Il est facile de voir que cela signifie que la fonction de répartition de X_n^{-1} converge vers la fonction de répartition $\mathbf{1}_{t>1}$ de la v.a 1 en tout point $t \neq 0$.

1.3 Variante

Si l'on désire à présent calculer des intégrales sur $\mathbf{R}^d$ (et non plus sur des pavés compacts) on peut plus généralement procéder de la fa con suivante.

Soit $X_1, \ldots, X_n, \ldots$ une suite indépendante de vecteurs aléatoires de même loi donnée par une densité ρ qui est facile à simuler. On veut calculer $I(f) = \int_{\mathbf{R}^d} f(x)dx$

$$\begin{aligned} I(f) &= \int_{\mathbf{R}^d} f(x_1, \ldots, x_d)dx_1 \cdots dx_d \\ &= \int_{\mathbf{R}^d} \frac{f(x_1, \ldots, x_d)}{\rho(x_1, \ldots, x_d)} \rho(x_1, \ldots, x_d)dx_1 \cdots dx_d \\ &= \mathbf{E}(\phi(X)) \end{aligned}$$

où

$$\phi = \frac{f}{\rho}.$$

On supposera ρ non nulle ou plus généralement que son support contient celui de f. La loi des grands nombres nous dit que

$$\frac{S_n(\phi)}{n} = \frac{1}{n}\Big(\phi(X_1) + \cdots + \phi(X_n)\Big)$$

converge presque-sûrement vers $\mathbf{E}(\phi(X_1)) = I(f)$. En procédant comme dans la section précédente on peut démontrer

Théorème 1.3.1 *Si on pose*

$$\sigma_n^2 = \frac{1}{n}S_n(\phi^2) - \left(\frac{1}{n}S_n(\phi)\right)^2,$$

alors,

$$\lim_{n\to\infty} \mathbf{P}\left(I(f) \in [\frac{1}{n}S_n(\phi) - \frac{\sigma_n a}{\sqrt{n}}, \frac{1}{n}S_n(\phi) + \frac{\sigma_n a}{\sqrt{n}}]\right) = \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}dx.$$

Le choix de la densité ρ est assez arbitraire mais il est naturel de vouloir minimiser la variance

$$\sigma^2 = \mathrm{Var}(\phi(X)) = \int \frac{f^2}{\rho}dx - I(f)^2$$

la contrainte étant $\int \rho(x)dx = 1$. On pourrait utiliser la méthode des multiplicateurs de Lagrange, dont on ne donnera pas ici le détail.

Pour l'examen il faut savoir en quoi consiste la méthode approchée du calcul des intégrales par la méthode Monté-Carlo, le fait qu'elle se base sur la loi de Grands Nombres, que la vitesse de convergence est determinée par le thm de la limite centrale, mais on ne demandera pas d'autres détails sur la vitesse de convergence.

Chapitre 2

Rudiments de Statistiques

2.1 Sondages

Un sondeur veut déterminer dans une population de N individus (par exemple $N = 6.10^7$) le nombre de personnes appartenant à une catégorie A (donc connaitre le cardinal de l'ensemble A). Pour cela il effectue un sondage sur un échantillon de n individus (p.ex $n = 10^3$) tirés au hasard. Pour simplifier on supposera que le sondeur procéde de la façon suivante. Il réalise n tirages $X_1, \ldots, X_n$ qui sont autant de variables aléatoires $X_i : \Omega \to \{1, \ldots, N\}$ suivant une loi uniforme sur $\{1, \ldots, N\}$. On pose alors $Y_i = \mathbf{1}_A \circ X_i$ et on définit $Z_n = Y_1 + \cdots + Y_n$ le nombre de personnes interrogées qui appartiennent à la catégorie A. Il est facile de voir que les Y_i suivent une même loi de Bernoulli de paramètre p (où p est la proportion de personnes dans la population ayant l'opinion A) et sont indépendantes pourvu que les X_i le soient (**Exercice :** Prouver ces faits). Par conséquent, Z_n suit une loi de Binomiale (n, p). Dans la pratique, on utilisera l'approximation normale que donne le Théorème Central Limit : la suite de v.a $(\sqrt{n}/\sigma)(Z_n - np)$, $\sigma^2 = p(1-p)$, converge en loi vers une loi normale centrée réduite. On peut donc écrire

$$\mathbf{P}(\{\omega \in \Omega : p \in [\frac{Z_n(\omega)}{n} - \frac{c\sigma}{\sqrt{n}}, \frac{Z_n(\omega)}{n} + \frac{c\sigma}{\sqrt{n}}]\}) \approx \int_{-c}^{c} \frac{e^{-t^2/2}}{\sqrt{2\pi}}.$$

[On rappelle que quand $c \approx 2$ cette intégrale est à peu près égale à 0.95 et quand $c \approx 2.6$ elle vaut à peu près 0.99.] Dans la pratique cette approximation est bonne quand par exemple $np \approx 30$. Pour simplifier nous supposerons que nous sommes dans cette hypothèse asymptotique. Ainsi, si on connait l'écart type σ on dira que l'intervalle $I_n = [\frac{Z_n}{n} - \frac{2\sigma}{\sqrt{n}}, \frac{Z_n}{n} + \frac{2\sigma}{\sqrt{n}}]$ (resp. $[\frac{Z_n}{n} - \frac{2.6\sigma}{\sqrt{n}}, \frac{Z_n}{n} + \frac{2.6\sigma}{\sqrt{n}}]$) est un *intervalle de confiance* avec une fiabilité de 95% (resp. 99%) pour la valeur de p : la probabilité pour que p se trouve dans l'intervalle ainsi

déterminé [1] est de 0.95 (resp. 0.99). Malheureusement, on ne connait pas la valeur de σ (car sinon on connaitrait déjà celle de p puisque $\sigma^2 = p(1-p)$). En revanche, on a toujours l'inégalité $\sigma = \sqrt{p(1-p)} \leq 1/2$ si bien que l'on peut écrire dès que n est assez grand

$$\mathbf{P}(p \in [\frac{Z_n}{n} - \frac{c}{2\sqrt{n}}, \frac{Z_n}{n} + \frac{c}{2\sqrt{n}}]) \geq\sim \int_{-c}^{c} \frac{e^{-t^2/2}}{\sqrt{2\pi}}.$$

Un intervalle de confiance à 0.95 (resp. 0.99) est donc par exemple $[\frac{Z_n}{n} - \frac{1}{\sqrt{n}}, \frac{Z_n}{n} + \frac{1}{\sqrt{n}}]$ (resp.$[\frac{Z_n}{n} - \frac{1.3}{\sqrt{n}}, \frac{Z_n}{n} + \frac{1.3}{\sqrt{n}}]$)

2.2 Statistiques gaussiennes

Sondages gaussiens Dans l'exemple précédent, l'estimation *a priori* sur la variance des Y_i, rendue possible par le fait que les Y_i suivent une loi de Bernoulli, est l'élément clé pour obtenir un intervalle de confiance. Intéressons nous à présent au problème suivant.

On fait un sondage de N individus en leur demandant d'exprmier leur avis sur ce polycopié par un chiffre de $]-\infty, \infty[$. Lorsque ce polycopié ne leur semble ni bon, ni mauvais, ils émettent la note zéro, lorsqu'ils sont contents de ce poly, ils émettent une note positive (1000, 10^6, 10^{10} etc) et d'autant plus grande – plus ils sont ravis. Lorsqu'ils sont mécontents – ils donnent une note négative (-100, -10^5, -10^{10}, etc). Soit Y_i – la note donnée par le $i$ème individu.

On pourrait bien sûre critiquer cette modélisation, mais on le fera plus tard !

Supposons que les avis des individus $Y_1, Y_2, \ldots, Y_N$ sont indépendants les uns des autres et que ce sont des réalisations de variables aléatoires Gaussiennes indépendantes de même paramètres – l'espérance μ et la variance σ^2. Notre objectif est d'estimer l'avis moyen sur ce poly μ et la variance des avis σ^2 , plus précisement de construire des intervalles dans lesquelles ces paramètres se trouvent avec une grande probabilité.

On choisit un seuil de fiabilité s proche de 1, par exemple $s = 0.99$.

Notre objectif est de trouver des intervalles $I_s \subset \mathbf{R}$ et $J_s \subset \mathbf{R}^+$ tel que

$$P(\mu \in I_s) \geq s, \quad P(\sigma^2 \in J_s) \geq s$$

On va proceder de la manière suivante : on connait du cours de probabilités de base (ou un exercice sur les fonctions caractéristiques). que

1. Attention, cela signifie que $p = \#A/N$ étant fixé (mais inconnu) la probabilité qu'un sondage portant sur n personnes fournisse un intervalle I_n ne contenant pas p est inférieure à 0.05

$\frac{Y_1+Y_2+\cdots+Y_N-N\mu}{\sqrt{N\sigma^2}}$ est une réalisation d'une variable Gaussienne dont l'espérance est 0 et la variance est 1. Alors pour $c>0$

$$\mathbf{P}\Big(\frac{Y_1+Y_2+\cdots+Y_N-N\mu}{\sqrt{N\sigma^2}}\in[-c,c]\Big)=\frac{1}{\sqrt{2\pi}}\int_{-c}^{c}e^{-t^2/2}dt=1-\frac{2}{\sqrt{2\pi}}\int_{c}^{\infty}e^{-t^2/2}dt.$$

On cherche c_s tel que $1-\frac{2}{\sqrt{2\pi}}\int_{c_s}^{\infty}e^{-t^2/2}dt=s$, ce qui est équivalent au fait que $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c_s}e^{-t^2/2}dt=1-(1-s)/2=1/2+s/2$. La fonction de répartition $\Phi(c)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c}e^{-t^2/2}dt$ de la loi Gaussienne centrée réduite est tabulée. On peut trouver dans la table le nombre c_s tel que $\Phi(c_s)=1/2+s/2$. (Vous pouvez par exemple trouver sa table sur la page http ://homeomath2.imingo.net/tablelois1.htm) ; on voit dans cette table par exemple que $\Phi(0.25)=0.5987$, $\Phi(1.81)=0.9649$) On peut dire aussi que c_s est la quantile de la loi Gaussienne centrée réduite de niveau $1-(1/2+s/2)$, cf. exercice 1 de la feuille N2.

Dans ce cas

$$\mathbf{P}\Big(\frac{Y_1+Y_2\cdots+Y_N-N\mu}{\sqrt{N\sigma^2}}\in[-c_s,c_s]\Big)=s.$$

On obtient alors

$$\mathbf{P}\Big(\mu\in\Big[\frac{Y_1+\cdots Y_N-c_s\sqrt{N\sigma^2}}{N},\frac{Y_1+\cdots Y_N+c_s\sqrt{N\sigma^2}}{N}\Big]\Big)=s,$$

et

$$\mathbf{P}\Big(\sigma^2\in\Big[\frac{(Y_1+Y_2\cdots+Y_N-N\mu)^2}{Nc_s^2},\infty\Big]\Big)=s.$$

L'intervalle $\Big[\frac{Y_1+\cdots Y_N-c_s\sqrt{N\sigma^2}}{N},\frac{Y_1+\cdots Y_N+c_s\sqrt{N\sigma^2}}{N}\Big]$ est l'intervalle de confiance pour la moyenne des avis sur ce polycopié de niveau de fiabilité s. Remarquons que tous les ingrédients pour calculer cette intervalle ($Y_1,\ldots,Y_N$, N, c_s, σ^2) sont connus de notre sondage sauf la variance σ^2. Si on connaissait σ^2 (ou on pouvait la supposer) on pourrait calculer les valeurs numériques. On a produit ainsi l'intervalle de confiance pour la moyenne de variables aléatoires Gaussiennes indépendantes lorsque leur variance est connue.

De même $\Big[\frac{(Y_1+Y_2\cdots+Y_N-N\mu)^2}{Nc_s^2},\infty\Big]$ est l'intervalle de confiance pour la variance des avis sur ce polycopié de niveau de fiabilité s. Remarquons que tous les ingrédients dans cette intervalle sont connus de notre sondage sauf la moyenne μ. Si on connaissait μ (ou on pouvait la supposer) on pourrait produire les valeurs numériques. On a produit ainsi l'intervalle de confiance pour la variance de variables aléatoires Gaussiennes indépendantes lorsque leur moyenne est connue.

Dans les deux cas on peut parler aussi des intervalles de confiance de niveau de fiabilité à $100s\%$, comme dans la feuille d'exercice N2, ou encore de niveau de fiabilité $100(1-\alpha)\%$ avec $\alpha = 1-s$.

Mais comment produire l'intervalle pour un paramètre de v.a. Gaussiennes indépendantes (la moyenne ou la variance) lorsque l'autre paramètre est inconnu ? Pour cela on va introduire les lois de chi-deux et de Student.

Définition. Soient $X_1, \ldots, X_n$ des v.a. indépendantes de loi Gaussienne d'espérance 0 et de variance 1. La loi de la v.a. $X_1^2 + \cdots + X_n^2$ est dite *la loi de chi-deux de n degrés de liberté*. C'est aussi une loi gamma $\gamma_{1/2,n/2}$ de densité

$$\mathbf{1}_{]0,\infty[}(x)\frac{1}{2^{n/2}\Gamma(n/2)}\int_0^x t^{(n/2)-1}e^{-t/2}dt.$$

Définition. Soient X et Y deux variables aléatoires indépendantes. La v.a. X est de loi Gaussienne d'espérance 0 et de variance 1. La v.a. Y est de loi chi-deux de $n-1$ degrés de liberté. La loi de la v.a.

$$\frac{X\sqrt{n-1}}{\sqrt{Y}}$$

est dite la loi de Student de $n-1$ degrés de liberté. C'est la loi de densité

$$c_{n-1}\frac{1}{\left(1+\frac{t^2}{n-1}\right)^{n/2}}, \qquad c_{n-1} = \frac{\Gamma(\frac{n}{2})}{\sqrt{(n-1)\pi}\Gamma(\frac{n-1}{2})}$$

Cette densité tend vers la densité gaussienne réduite $(1/\sqrt{2\pi})e^{-t^2/2}$ quand n tend vers l'infini ; on a les approximations suivantes

$P(\lvert T_n\rvert \le a)$	$= 0.95$	0.99
si $n = 10$	pour $a = 2.26$	pour $a = 3.35$
20	2.09	2.86
30	2.04	2.76
∞	1.96	2.58

Par la suite nous introduisons la quantitÃ© qui s'appelle la moyenne empirique

$$\bar{Y_n} = \frac{Y_1 + \cdots + Y_n}{n}$$

et celle qui s'appelle variance empirique

$$V_n = \frac{(Y_1 - \bar{Y_n})^2 + \cdots (Y_n - \bar{Y_n})^2}{n}.$$

Théorème 2.2.1 *Si les v.a $Y_1, \ldots, Y_n, \ldots$ sont i.i.d. de méme loi normale $\mathcal{N}(\mu, \sigma^2)$ alors*

(a) la v.a

$$\frac{\sqrt{n}}{\sigma}(\bar{Y}_n - \mu)$$

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

(b) Les v.a. $\bar{Y}_n$ et V_n sont indépendantes et la v.a.

$$Z_n = \frac{\bar{Y}_n - \mu}{\sqrt{V_n/(n-1)}}$$

suit une loi de Student $\mathcal{T}(n-1)$ à $n-1$ degrés de liberté ;

(c) la v.a

$$\frac{nV_n}{\sigma^2}$$

suit la loi du chi-deux é $n-1$ degrés de liberté $\chi^2(n-1)$.

Démonstration. (a) est évident.

Notons que $Z_n = (\sqrt{n-1}\sqrt{n}(\bar{Y}_n - \mu)/\sigma)(\sqrt{nV_n/\sigma^2})^{-1}$. Il suffit de prouver que $S = \frac{\sqrt{n}}{\sigma}(\bar{Y}_n - \mu)$ et $T = \frac{nV_n}{\sigma^2}$ sont indépendantes, et que T soit de loi $\chi^2(n-1)$. Soient $W_i = (Y_i - \mu)/\sigma$, $i = 1, \ldots, n$. Les v.a. $W_1, \ldots, W_n$ sont indép., de loi Gaussienne d'espérance 0 et de variance 1. Soit $\vec{W} = (W_1, \ldots, W_n)$. Prenons A une matrice orthogonale dont la première ligne se compose de $\sqrt{n}$. Alors le vecteur $A\vec{W}$ est aussi Gaussien, d'espérance $(0, \ldots, 0)$ et de matrice de covariances $AIdA^T = AA^T = Id$ car A est orthogonale. Donc les coordonnés de $A\vec{W}$ sont des Gaussiennes indépendantes d'espérance 0 et de variance 1. Notamment $\sum_{i=2}^n ((A\vec{W})_i)^2$ est de loi de $\chi^2(n-1)$. Par le choix de la premiere ligne de A on a $(A\vec{W})_1 = S$. Alors S est indépendante de $(A\vec{W})_i$ pour $i = 2, \ldots, n$, et donc de $\sum_{i=2}^n ((A\vec{W})_i)^2$ qui est de loi de $\chi^2(n-1)$. Mais une matrcie orthogonale conserve les normes de vecteurs. Donc $S^2 + \sum_{i=2}^n ((A\vec{W})_i)^2 = \sum_{i=1}^n W_i^2$. Alors $\sum_{i=2}^n ((A\vec{W})_i)^2 = \sum_{i=1}^n W_i^2 - S^2$ ce qui est par un calcul direct $\frac{nV_n}{\sigma^2}$

Dans l'exemple qui nous intéresse on peut donc :

(i) Sans connaitre la valeur de σ^2, obtenir un intervalle de confiance pour la moyenne μ. La probabilité pour que l'intervalle

$$\left[\bar{Y}_n - c\sqrt{\frac{V_n}{n-1}}, \bar{Y}_n + c\sqrt{\frac{V_n}{n-1}}\right]$$

contienne le réel μ (inconnu) est égale à $\mathbf{P}(|Z_n| \leq c)$ (on a utilisé le (b) du théorème précédent) ; Il reste donc à chercher c_s tel que $\mathbf{P}(|Z_n| \leq c_s) \geq s$,

autrement dit $F_{Z_n}(c_s) = 1/2 + s/2$ où F_{Z_n} est la fonction de répartition de la loi de Student de $n-1$ degrés de liberté, cette fonction est tabulée (on peut chercher c_s dans les tables).

(ii) Sans connaitre la valeur de μ, obtenir un intervalle de confiance pour la moyenne σ^2. La probabilité pour que l'intervalle

$$\Big[\frac{nV_n}{c}, \infty\Big]$$

contienne σ^2 est égale à $\mathbf{P}(U_n \leq c)$ (on a utilisé le (c) du théorème précédent) ; Il reste donc à chercher c_s tel que $\mathbf{P}(U_n \leq c_s) \geq s$, autrement dit $F_{U_n}(c_s) \geq s$ où F_{U_n} est la fonction de répartition de la loi de chi-deux de $n-1$ degrés de libertés, cette fonction est tabulée Ã©galement.

Sondages Non-Gaussiens, intervalles de confiance assymptotiques. Dans la modélisation du paragraph précédent nous avons modélisé des avis de lecteurs de ce poly par des variables aléatoires Gaussiennes qui prennent des valeurs dans $]-\infty, \infty[$. Cette modélisation n'est bien sûre pas du tout raisonnable !!! On pourrait plutôt proposer aux lecteurs de donner une note à ce poly entre -10 et 10 et essayer de construire un intervalle de confiance pour la note moyenne. Nous allons construire un intervalle de confiance *assymptotique* dans ce cas.

Soient $Y_1, \ldots, Y_N$ des v.a. i.i.d. de second moment fini. Alors par le Thm de la limite Centrale $\frac{Y_1+\cdots+Y_N-N\mu}{\sqrt{N\sigma^2}}$ converge en loi vers la loi Gaussienne centrée réduite. Autrement dit,

$$\mathbf{P}\Big(\frac{Y_1+\cdots+Y_N-N\mu}{\sqrt{N\sigma^2}} \in [-c,c]\Big) \to \frac{1}{\sqrt{2\pi}}\int_{-c}^{c} e^{-t^2/2}dt = 1 - \frac{2}{\sqrt{2\pi}}\int_{c}^{\infty} e^{-t^2/2}dt,$$

quand $N \to \infty$. Nous pourrions alors construire un intervalle de confiance assymptotique :

$$\mathbf{P}\Big(\mu \in \Big[\frac{Y_1+\cdots Y_N - c_s\sqrt{N\sigma^2}}{N}, \frac{Y_1+\cdots Y_N + c_s\sqrt{N\sigma^2}}{N}\Big]\Big) \to s$$

quand $N \to \infty$ où c_s est choisi tel que $\Phi(c_s) = 1/2 + s/2$, $\Phi(\cdot)$ étant la fonction de répartition de la loi Gaussienne centrée réduite.

Cet intervalle a de nouveau un défaut de contenir une quantité inconnue σ^2.

Pour corriger ce défaut nous allons remplacer σ^2 par son "estimateur" V_N. Composons

$$V_N = \frac{1}{N}\sum_{i=1}^{N} Y_i^2 - \Big(\frac{1}{N}\sum_{i=1}^{N} Y_i\Big)^2.$$

Par la loi de grands nombres $\frac{1}{N}\sum_{i=1}^n Y_i^2 \to E(Y_1^2)$, $\frac{1}{N}\sum_{i=1}^n Y_i \to EY_1$ p.s., donc $V_N \to \sigma^2$ p.s., et donc $\sqrt{V_N/\sigma^2} \to 1$ p.s.

On va aussi admettre un résultat de probabilités de base : si une suite de v.a. ξ_n converge en loi vers ξ et une suite η_n converge vers une constante(!) c p.s., alors $\xi_n\eta_n$ converge en loi vers $c\xi$.

Il s'en suit que $\frac{Y_1+\cdots+Y_N-N\mu}{\sqrt{NV_N}}$ converge en loi vers la loi Gaussienne centrée réduite. Donc

$$\mathbf{P}\Big(\mu \in \Big[\frac{Y_1+\cdots+Y_N-c_s\sqrt{NV_N}}{N}, \frac{Y_1+\cdots Y_N+c_s\sqrt{NV_N}}{N}\Big]\Big) \to s.$$

On construit ainsi un intervalle de confiance asymptotique

$$\Big[\frac{Y_1+\cdots Y_N-c_s\sqrt{NV_N}}{N}, \frac{Y_1+\cdots Y_N+c_s\sqrt{NV_N}}{N}\Big]$$

de niveau de fiabilité s (ou de niveau de fiabilité à $100s\%$)

Tests statistiques Il est commode à ce stade d'introduire la notion de *test statistique*. Le type de probléme que l'on se propose d'étudier est le suivant : à partir d'un échantillon, c'est-à-dire une suite de v.a $X_1, \ldots, X_n$ de méme loi p qui en général est inconnue, décider si une hypothèse $\mathcal{H}_0$ (portant sur la loi de ces v.a) est raisonnable ou pas. Si elle ne l'est pas on la rejette, sinon on l'accepte, ce qui alors signifie seulement qu'elle est plausible. Dans ce type de discussion on introduit un *seuil* α qui est un réel compris entre 0 et 1 et quantifie le degré de certitude que l'on a lors de la procédure de rejet ou non de l'hypothèse $\mathcal{H}_0$. La procédure est la suivante : on considère une *statistique* c'est-à-dire une v.a $Z_{n,\mathcal{H}_0}$ qui est une fonction des v.a $X_1, \ldots, X_n$ et dont on connait la loi (et donc la fonction de répartition) *pourvu que* l'hypothèse $\mathcal{H}_0$ soit vérifiée. On décide de rejeter ou d'accepter l'hypothèse au seuil $\alpha \in (0,1)$ si par exemple, les données expérimentales donnent $|Z_{n,\mathcal{H}_0}| \geq t_\alpha$ où t_α est définie par $\mathbf{P}_{\mathcal{H}_0}(|Z_{n,\mathcal{H}_0}| \geq t_\alpha) = \alpha$. En d'autres termes on rejette l'hypothése $\mathcal{H}_0$ si la réalisation $\omega \in \Omega$ qui donne $|Z_{n,\mathcal{H}_0}(\omega)| \geq t_\alpha$ fait partie d'un événement très peu probable.

Dans l'exemple (ii) du paragraphe précédent l'échantillon est une suite de v.a de loi $\mathcal{N}(\mu, \sigma^2)$ qui dépend des deux paramètres μ, et σ *a priori* inconnus et on veut tester l'hypothèse suivante portant sur σ : $\mathcal{H}_0 : 10V_{10} \geq 20\sigma^2$. Pour cela on construit une statistique, en l'occurrence la v.a $Z_{10} = \frac{10V_{10}}{\sigma^2}$ dont on sait qu'elle suit une loi du chi-deux à 9 degrés de liberté et pour laquelle la fonction de répartition est tabulée. Au seuil $\alpha = 0.05$ on rejette l'hypothèse car

$$\mathbf{P}(Z_{10} \geq 20) \leq \mathbf{P}(Z_{10} \geq t_\alpha) = \mathbf{P}(\chi_9^2 \geq t_\alpha)$$

où $t_\alpha = 16.9$ et $\mathbf{P}(\chi^2_9 \geq t_\alpha) = 1 - F_{\chi^2(9)}(16.9) \approx 0.05$ est plus petite que le seuil $\alpha = 0.05$ que l'on s'était fixé.

Tests de Fisher et de Student. Ces deux tests ne seront pas demandés à l'examen et peuvent être omis à la première lecture. On considère deux échantillons gaussiens $X_1, \ldots, X_{n_x}$, $Y_1, \ldots, Y_{n_y}$ qui sont indépendants et suivent respectivement des lois $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\mathcal{N}(\mu_y, \sigma_y^2)$. On note

$$\bar{X} = \frac{X_1 + \cdots + X_{n_x}}{n_x}, \qquad \bar{Y} = \frac{Y_1 + \cdots + Y_{n_y}}{n_y}$$

les moyennes empiriques et

$$V_x = \frac{(X_1 - \bar{X})^2 + \cdots + (X_{n_x} - \bar{X})^2}{n_x}, \qquad V_y = \frac{(Y_1 - \bar{Y})^2 + \cdots + (Y_{n_y} - \bar{Y})^2}{n_y}$$

les variances empiriques. Le théorème suivant permet de tester l'hypothèse $\mathcal{H}_0 : \sigma_x^2 = \sigma_y^2$.

Théorème 2.2.2 *Le rapport*

$$\frac{\dfrac{n_x}{n_x - 1}\dfrac{V_x}{\sigma_x^2}}{\dfrac{n_y}{n_y - 1}\dfrac{V_y}{\sigma_y^2}}$$

suit une loi de Fisher $\mathcal{F}(n_x - 1, n_y - 1)$. *La densité d'une loi de Fisher* $\mathcal{F}(m, n)$ *est*

$$f_{m,n}(t) = C_{m,n} t^{(m/2)-1} (n + tm)^{-(m+n)/2} \mathbf{1}_{t>0}, \qquad C_{m,n} = \frac{m^{m/2} n^{n/2} \Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}.$$

Remarque : $n_x V_x / \sigma_x^2$ et $n_y V_y / \sigma_y^2$ suivent (d'après le Théorème 2.2.1) des lois de Student.

La statistique du test de Fisher $\mathcal{H}_0 : \sigma_x^2 = \sigma_y^2$ est donc

$$\frac{\dfrac{n_x}{n_x - 1} V_x}{\dfrac{n_y}{n_y - 1} V_y}$$

Sous l'hypothése d'égalité des variances $\sigma_x^2 = \sigma_y^2$ on peut tester l'égalité des moyennes $\mathcal{H}_0 : \mu_x = \mu_y$

Théorème 2.2.3 *Si $\sigma_x = \sigma_y$ la v.a*

$$\frac{\sqrt{n_x + n_y - 2}}{\sqrt{n_x^{-1} + n_y^{-1}}} \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{n_x V_x + n_y V_y}}$$

suit la loi de Student $\mathcal{T}(n_x + n_y - 2)$.

2.3 Test du chi-deux

Ce test est souvent demandé à l'examen.

Déterminer si un dé est pipé ou non : On veut savoir si un dé à 6 faces présente chacune de ses faces de façon équiprobable. Pour cela on jette le dé n fois (n grand). Le résultat du k-ème lancer est modélisé par une v.a X_k à valeurs dans $E = \{1, \ldots, 6\}$ et on suppose que les v.a X_k sont indépendantes et identiquement distribuées. Dire que le dé est pipé c'est dire que $(p_1, p_2, p_3, p_4, p_5, p_6) \neq (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$ (on a noté $p_i = \mathbf{P}(X_k = i)$). On veut donc tester l'hypothèse $\mathcal{H}_0$: "la loi $(p_i)_{i \in E}$ est uniforme". Pour fixer les idées supposons qu'on réalise 600 lancers et qu'on obtienne 60 fois le 1, 108 fois le 2, 108 fois le 3, 102 fois le 4, 108 fois le 5 et 114 fois le 6. Le dé est-il pipé ou non ? Pour répondre à cette question, on utilise le théorème suivant :

Théorème 2.3.1 *Soit $X_1, \ldots, X_n, \ldots$ une suite de v.a. i.i.d. à valeurs dans un ensemble fini $E = \{1, \ldots, r\}$. Nous noterons (p_i) leur loi commune avec $p_i = \mathbf{P}(X_n = i)$. Posons*

$$N_i^{(n)} = \sum_{k=1}^{n} \mathbf{1}_i \circ X_k,$$

la fréquence empirique de sortie de i. Alors,

(a) le vecteur aléatoire

$$\vec{Z}_n = \left(\frac{N_1^{(n)} - np_1}{\sqrt{np_1}}, \ldots, \frac{N_r^{(n)} - np_r}{\sqrt{np_r}}\right)$$

converge en loi vers un vecteur gaussien $\vec{Z}$ d'espérance $\vec{0}$ et de matrice de covariances B, dont les éléments $b_{i,i} = (1 - p_i)$ pour $i = 1, \ldots, r$, $b_{i,j} = -\sqrt{p_i p_j}$ pour $i \neq j$.

(b) la suite de v.a

$$T_n = \frac{(N_1^{(n)} - np_1)^2}{np_1} + \cdots + \frac{(N_r^{(n)} - np_r)^2}{np_r},$$

converge en loi quand $n \to \infty$ *vers un* χ^2 *à* $r-1$ *degrés de liberté.*

Démonstration. (a) est un corollaire direct du Thm Limite Central vectoriel appliqué aux sommes de vecteurs aléatoires inépendants $(\mathbf{1}_1 \circ X_k, \ldots, \mathbf{1}_r \circ X_k)$ pour $k = 1, 2, \ldots$. En effet, $E\mathbf{1}_i \circ X_k = p_i$ pour $i = 1, \ldots, r$. Calculons $cov(\mathbf{1}_i \circ X_k/\sqrt{p_i}, \mathbf{1}_i \circ X_k/\sqrt{p_j})$. Pour $i = j$ c'est $var(\mathbf{1}_i \circ X_k/\sqrt{p_i}) = p_i(1-p_i)/p_i = 1-p_i$. Pour $i \neq j$ on a $(\mathbf{1}_i \circ X_k) \cdot (\mathbf{1}_i \circ X_k) = 0$ car X_k ne peut pas prendre les valeurs i et j en même temps. Donc c'est $-E(\mathbf{1}_i \circ X_k/\sqrt{p_i})E(\mathbf{1}_i \circ X_k/\sqrt{p_j}) = -\sqrt{p_i p_j}$.

Comme $\vec{Z}_n$ converge en loi vers $\vec{Z}$, alors pour toute fonction g continue bornée $Eg(\vec{Z}_n) \to Eg(\vec{Z})$. En particulier pour toute f continue bornée $Ef(\|\vec{Z}_n\|) \to Ef(\|\vec{Z}\|))$ car $g = f \circ \|\cdot\|$ est continue bornée. Il s'en suit que $\|\vec{Z}_n\| \to \|\vec{Z}\|$ en loi.

Il reste à prouver que la loi de $\|\vec{Z}\|$ est la loi de $\chi^2(r-1)$. La matrice B est symétrique définie positive. Elle est donc diagonalisable en base orthonormée. Il existe une matrice A orthogonale telle que ABA^T est une matrice diagonale D qui se compose de valeurs propres de B. La combinaison linéaire des lignes de B avec les coefficients $\sqrt{p_1}, \ldots, \sqrt{p_r}$ est 0. Dons une valeur propre est 0. La matrice $B - Id$ a toutes les lignes proportionelles. Donc 1 est une valeur propre de B de multiplicité $r-1$. Donc D are 1 sur la diagonale $r-1$ fois et 0 pour le dernier élément. Le vecteur $A\vec{Z}$ est Gaussien d'espérance $\vec{0}$ et de matrice de covariances $ABA^T = D$. Donc la loi de $\|A\vec{Z}\|$ est la loi $\chi^2(r-1)$. Mais comme A est orthogonale, elle conserve la norme, par conséquent $\|\vec{Z}\| = \|\vec{A}\vec{Z}\|$ est donc de loi $\chi^2(r-1)$. Le Thm est démontré.

Ainsi, si l'hypothèse $\mathcal{H}_0$ est vérifiée la v.a

$$T_{600} = \frac{(N_1^{600} - 100)^2}{100} + \cdots + \frac{(N_6^{600} - 100)^2}{100},$$

doit suivre (approximativement) une loi du chi-deux é 5 $(= 6-1)$ degrés de liberté. Or, pour une telle loi $\mathbf{P}(\mathcal{T}_5 \geq 0.412) \leq 0.005$. Dans notre expérience on a obtenu

$$T_{600}(\omega) = \frac{(60-100)^2}{100} + \cdots + \frac{(114-100)^2}{100} = 3.92$$

Comme l'événement $3.92 > 0.412$ on décide au seuil $\alpha = 0.005$ de rejeter l'hypothèse $\mathcal{H}_0$ (le fait d'observer $\mathcal{T}_5 \geq 0.412$ est un événement rare).

De manière générale, supposons, on sait que $X_1, X_2, \ldots, X_k, \ldots$ sont des v.a. indépendantes et de même loi discrète prenant r valeurs $a_1, \ldots, a_r$. On veut tester l'hypothèse que ces valeurs sont prises avec probabilités $p_1, \ldots, p_r$. On observe la réalisation $X_1, X_2, \ldots, X_n$ pour n grand. On compose $N_i^n = \sum_{k=1}^n \mathbf{1}_{a_i} \circ X_k$ pour $i = 1, \ldots, r$. On calcule la valeur de $T_n = \frac{(N_1^{(n)} - np_1)^2}{np_1} + \cdots + \frac{(N_r^{(n)} - np_r)^2}{np_r}$. Sa loi est approximativement(!) de $\chi^2(r-1)$.

On se donne un seuil de confiance α et on trouve le quantile de la loi $\chi^2(r-1)$, cad $P(\chi > \chi_{\alpha,r-1}) = \alpha$. Ainsi $P(T_n > \chi_{\alpha,r-1}) \approx \alpha$.

S'il se trouve de notre calcul que $T_n > \chi_{\alpha,r-1}$, alors un événement de probabilité trop petite est réalisé, cet événement est jugé trop peu probable pour avoir lieu en réalité, par conséquant on se doute de notre hypothèse et on la rejette.

Si $T_n < \chi_{\alpha,r-1}$, ce test ne rejette pas notre hypothèse.

Dans la partie statistiques, pour l'examen, il faut savoir :

- Construire des intervalles de confiance pour l'espérance et la variance de variables aléatoires i.i.d. Gaussiennes lorsque
  - l'autre paramètre est connu (Paragraph 4.2 + Exercice 4, feuille N2)
  - l'autre paramètre est inconnu (Paragraph 4.2, Thm 4.2.1 + Exercice 4, feuille N2)
- Construire des intervalles de confiance assymptotiques pour la moyenne de variables aléatoires i.i.d. dans L^2, à la base du Thm de la limite centrale, remplacer la variance inconnue par son estimateur (Paragraph 4.2 + Exercice 5)
- Construire des intervalles de confiance pour des paramètres en se servant des inégalités de Markov et de Chebyshev (Paragraph 4.1, Exercices 3 et 6)
- Appliquer les test de Chi-deux (Paragraph 4.3 + Exercice 7).