UNIVERSITÉ DU LUXEMBOURG

Faculté des sciences,
Technologie et Communication

Experimental mathematics lab

Les Lois de probabilité

Auteurs :
Lynn Van der weken
Fa Zhu

Superviseur :
Prof. Ivan Nourdin

BASI Mathématiques
Année académique 2015-2016

Résumé

Le but de ce projet est de se familiariser avec les lois de probabilité. Nous verrons plusieurs méthodes qui permettent de simuler les lois de probabilité les plus populaires et nous allons expérimentellement calculer quelques propriétés clés de ces lois comme l'espérance, la médiane, les quantiles et les moments.

Table des matières

1 Cas discret **1**
1.1 Introduction 1
1.2 Loi uniforme discrète 3
1.3 Loi de Bernoulli 6
1.4 Loi binomiale 8
1.5 Loi géométrique 11
1.6 Loi de Poisson 14
1.7 Méthode des moments 17

2 Cas continu **20**
2.1 Loi uniforme 22
2.2 Loi exponentielle 24
2.3 Loi normale (ou gaussienne) 26
2.4 Loi du chi-carré 28
2.5 Loi de Student 30
2.6 Estimation de la loi par la méthode des moments 32
2.7 Estimation : Modèle OMMeR 34
2.8 Modèle OMMeR dans le cas normal partiellement fixé 36
2.9 Estimation d'une loi $L(\mu, \sigma^2)$ 38

Chapitre 1

Cas discret

1.1 Introduction

Commençons avec les définitions de base pour le cas discret.

Définition 1. *(i) Soit Ω l'espace des résultats possibles. Une* ***distribution de probabilité*** *sur Ω est une fonction*
$p : \Omega \to [0,1]$ telle que $\sum_{\omega\in\Omega} p(\omega) = 1$.

(ii) Nous appelons ***événement*** *toute partie A de Ω : $A \in \mathcal{P}(\Omega) := \{B : B \subset \Omega\}$.*

(iii) Soient p une distribution de probabilité sur Ω, $\mathcal{A} \subset \mathcal{P}(\Omega)$ et

$$\begin{aligned} \mathbb{P} : \mathcal{A} &\to [0,1] \\ A \in \mathcal{A} &\mapsto \sum_{\omega\in A} p(\omega). \end{aligned}$$

Alors nous appelons $(\Omega, \mathcal{A}, \mathbb{P})$ un ***espace de probabilité discrèt****. Nous remarquons que $\mathbb{P}(\{\omega\}) = p(\omega)$.*

(iv) Nous appelons ***variable aléatoire*** *une fonction $X : \Omega \to [-\infty, +\infty]$. La valeur de $X(\omega)$ dépend du hasard de $\omega \in \Omega$.*

(v) Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$. Alors $P_X(A) = \mathbb{P}\{X \in A\} = \mathbb{P}\{\omega \in \Omega | X(\omega) \in A\}$ définit une probabilité sur $X(\Omega)$ appelée ***la loi de X*** *ou* ***la distribution de X****.*

(vi) Nous appelons ***espérance de X*** *le nombre*

$$E[X] = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\}),$$

lorsqu'il est bien défini. Nous remarquons que $E[X]$ est toujours bien défini si $X \geq 0$ et $E[X] = +\infty$ n'est pas exclu. De plus l'espérance est linéaire.

(vii) Soient X une variable aléatoire telque $E|X| < +\infty$ et $n \in \mathbb{N}$. Nous appelons $E[X^n]$ ***moment d'ordre*** *n* ***de*** *X et $E[(X - E[X])^n]$* ***moment centré d'ordre*** *n* ***de*** *X. En particulier :*

$$Var(X) = \sigma^2(X) = E[(X - E[X])^2$$

est la ***variance de*** *X et $\sqrt{Var(X)} = \sigma^2(X)$ est* ***l'écart-type de*** *X.*

La dernière définition nous mêne au prochain lemme :

Lemme 2. *Soit X une variable aléatoire. La variance de X est égal à :*

$$Var(X) = E[X^2] - E[X]^2.$$

Démonstration. En effet,

$$\begin{aligned} Var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] \quad \text{par linéarité de l' espérance} \\ &= E[X^2] - 2E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2. \end{aligned}$$

□

Nous allons ensuite encore introduire quelques notions de la statistique.

Définition 3. *(i) Nous appelons **population**, un ensemble $(X_i)_{i \in I}$ de variables aléatoires qui suivent une loi L quelconque (partiellement) inconnue. Nous supposons que les X_i sont indépendantes et identiquement distribuées (iid).*

*(ii) Nous appelons **échantillon**, un tirage aléatoire d'ensembles de varibales aléatoires $(X_{i_1}, ..., X_{i_n})$ de la population.*

*(iii) Nous appelons **observations**, les valeurs prises par l'échantillon $(x_{i_1}, ..., x_{i_n})$.*

Le but de la statistique est de faire de l'inférence à partir des observations, c'est-à-dire, d'estimer les paramètres de la loi L.

1.2 Loi uniforme discrète

Définissons d'abord la loi uniforme discrète.

Définition 4. *Soit $X(\Omega) = \{1, ..., n\}$ avec $n \in \mathbb{N}$. Nous disons que X suit une* ***loi uniforme discrète*** *si $\mathbb{P}(X = k) = \frac{1}{n}$, pour tout $k \in X(\Omega)$.*

En particulier, pour $A \subset X(\Omega)$ nous avons

$$\mathbb{P}(A) = \frac{\#A}{\#X(\Omega)} = \frac{\#A}{n}.$$

Les exemples simples de la loi discrète sont : lancer d'un dé non biais, lancer d'une pièce, tirer une carte ou un jeu de roulette.

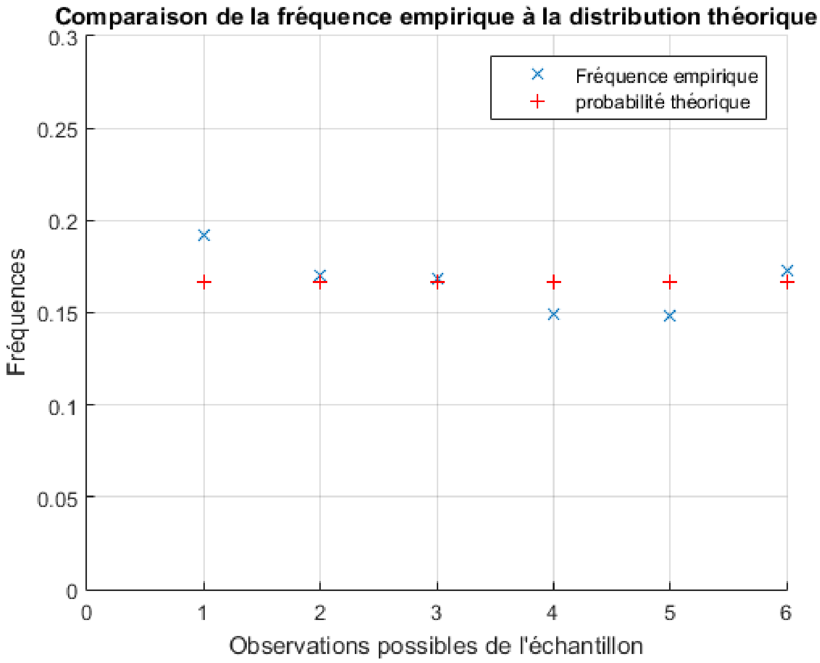
Calculons maintenant l'espérance et la variance de la loi uniforme discrète.

$$\begin{aligned} E[X] &= \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(X = X(\omega)) \\ &= \sum_{i=1}^{n} i\frac{1}{n} \\ &= \frac{1}{n}\frac{(n+1)n}{2} \\ &= \frac{n+1}{2}. \end{aligned}$$

Maintenant nous pouvons utiliser le lemme **2.** pour calculer la variance.

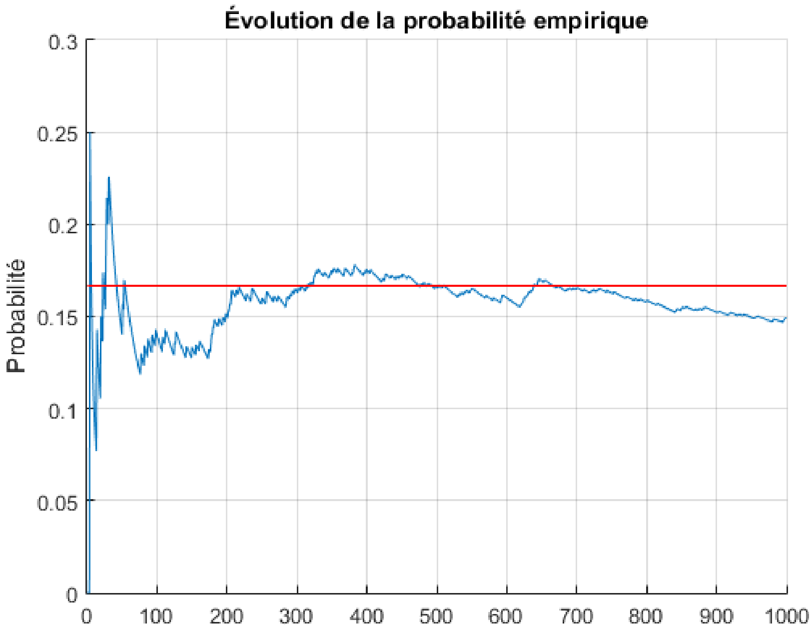
$$\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= \sum_{i=1}^{n} i^2\frac{1}{n} - \left(\frac{(n+1)}{2}\right)^2 \\ &= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} - \left(\frac{(n+1)}{2}\right)^2 \\ &= \frac{2n^2+3n+1}{6} - \frac{n^2+2n+1}{4} \\ &= \frac{n^2-1}{12}. \end{aligned}$$

Nous avons crée un échantillon aléatoire de taille 1000 de la loi uniforme, de paramètre 6. En calculant la fréquence empirique, nous obtenons ainsi le graphe suivant :

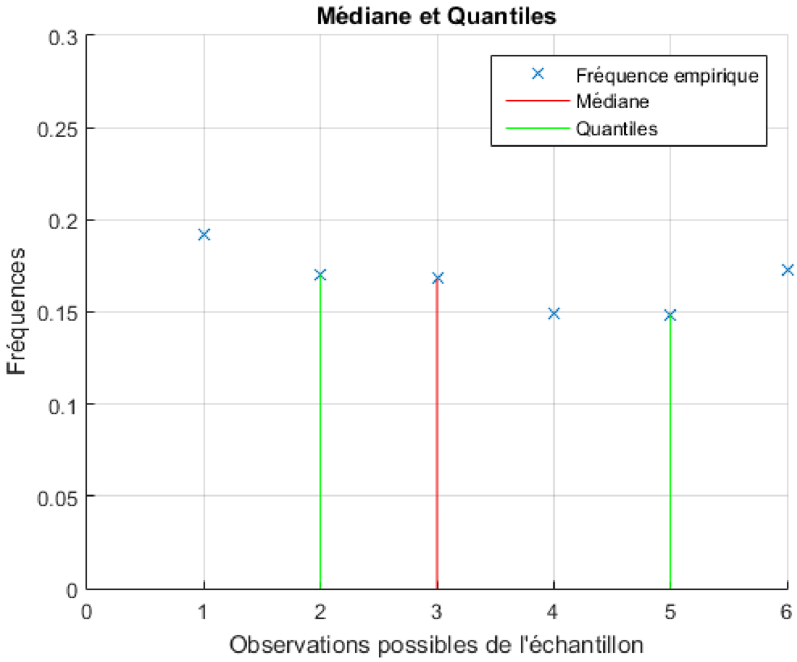


Observons maintenant un cas particulier. Nous nous intéressons à l'évolution de la fréquence d'un évènement particulier. Nous observons lar variation de la fréquence d'obtenir la valeur $k = 4$, lors des n premiers observations.

Cette fréquence se rapproche de la probabilité théorique $\frac{1}{6}$, mais semble ne pas être très stable.



Voici encore un graphique qui montre aussi la médiane avec les quantiles $\frac{1}{4}$ et $\frac{3}{4}$.



1.3 Loi de Bernoulli

Définition 5. *Soit $X(\Omega) = \{0,1\}$. La variable aléatoire X suit une* ***Loi de Bernoulli*** *si le succès de la variable X est de probabilité $p \in]0,1[$ et l'échec de la variable est de probabilité $p-1$, c'est-à-dire*

$$\begin{cases} P(X=1) = p, \quad p \in]0,1[\\ \mathbb{P}(X=0) = 1-p. \end{cases}$$

Si la variable aléatoire X suit une loi de Bernoulli de paramètre p nous écrivons $X \sim Bern(p)$.

Un exemple pour la loi de Bernoulli est la probabilité de tirer une boule blanche d'une urne, qui contient n boules, dont np boules sont blanches et $n(1-p)$ sont noires.

L'espérance pour la loi de Bernoulli.

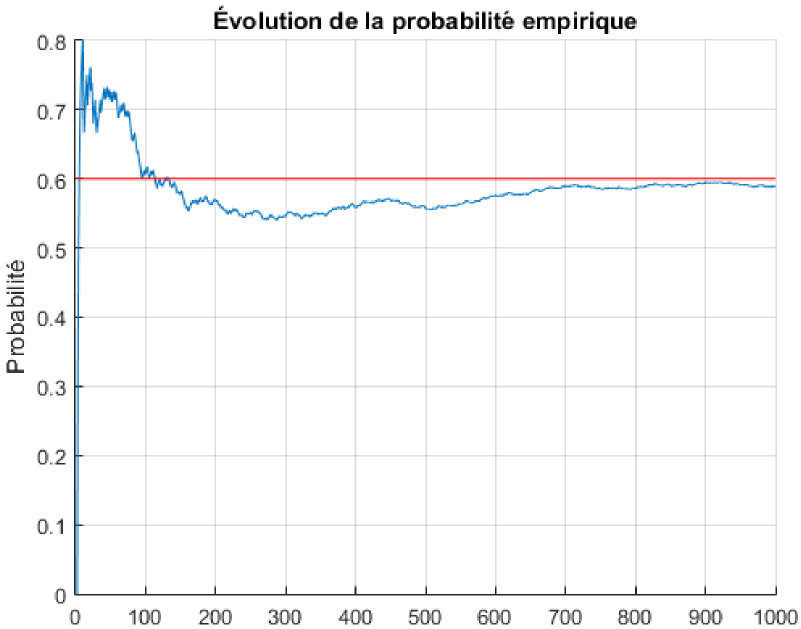
$$\begin{aligned} E[X] &= 1 \cdot \mathbb{P}(X=1) + 0 \cdot \mathbb{P}(X=0) \\ &= 1 \cdot p + 0 \cdot (1-p) \\ &= p. \end{aligned}$$

En utilisant de nouveau le lemme **2.** nous pouvons calculer la variance.

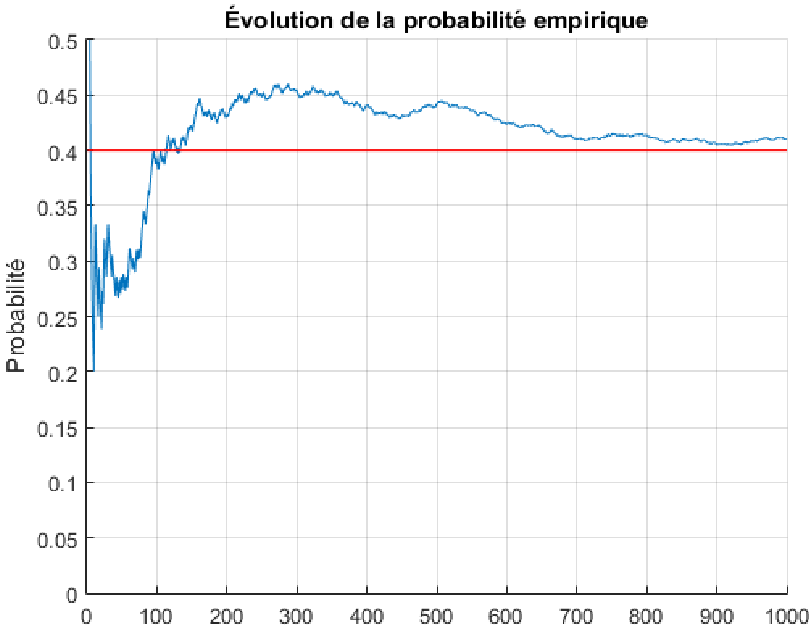
$$\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= 1 \cdot p + 0 \cdot (1-p) - p^2 \\ &= (1-p)p. \end{aligned}$$

Nous disposons d'un échantillon de taille 1000 d'une loi de Bernoulli de paramètre 0,4 et nous nous intéressons à l'évolution de la fréquence empirique de $\mathbb{P}(X=k)$.

Pour $k=0$, nous obtenons le graphique suivant :



Pour $k = 1$, nous obtenons :



1.4 Loi binomiale

Maintenant la loi binomiale.

Définition 6. *Soient X une variable aléatoire et $k, n \in \mathbb{N}$ tel que $k \leq n$. La variable aléatoire X suit une **loi binomiale**, et nous écrivons $X \sim Bin(n,p)$ si l'égalité suivante est vérifiée*

$$\mathbb{P}(X = k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k}.$$

La loi binomiale peut être engendrée par la loi de Bernoulli. En effet, la loi binomiale est la probabilité que le nombre de succès sur n lancers soit k.

Calculons maintenant l'espérance de la loi binomiale. Soient X_i des variables aléatoires tel que $(X_i)_{i=1,\dots,n} \sim B(p)$ où les X_I sont iid. Nous avons que $X = \sum_{i=1}^n X_i$ suit une loi binomiale$Bin(n,p)$. Nous obtenons donc

$$\begin{aligned} E[X] &= E\left[\sum_{i=1}^n X_i\right] \\ &= \sum_{i=1}^n E[X_i] \qquad \text{par l'indépendance des } X_i \\ &= \sum_{i=1}^n p \\ &= np. \end{aligned}$$

Maintenant nous pouvons calculer la variance de la loi binomiale.

$$\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= E\left[\sum_{i=1}^n X_i \sum_{j \neq i}^n X_j\right] - n^2p^2. \end{aligned}$$

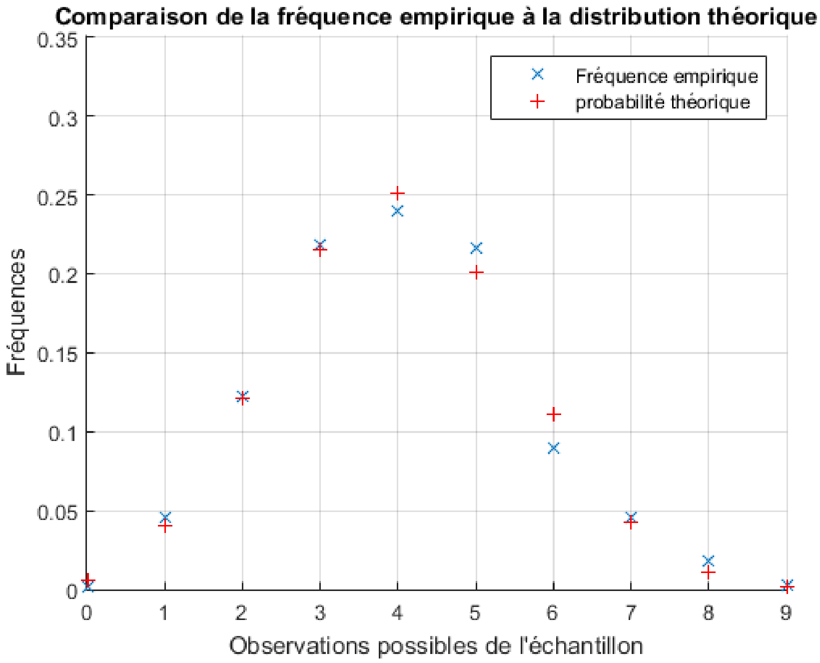
Pour calculer $E\left[\sum_{i=1}^n X_i \sum_{j \neq i}^n X_j\right]$, nous devons obserer deux cas différents. Si $i = j$ nous avons l'égalité $E\left[(\sum_{i=1}^n X_i)^2\right] = np$ et si $i \neq j$, nous obtenons

$$E\left[\sum_{i=1}^n X_i \sum_{j=1}^n X_j\right] = \sum_{\substack{i=1 \\ j \neq i}}^n E[X_i]E[X_j] = n(n-1)p^2,$$

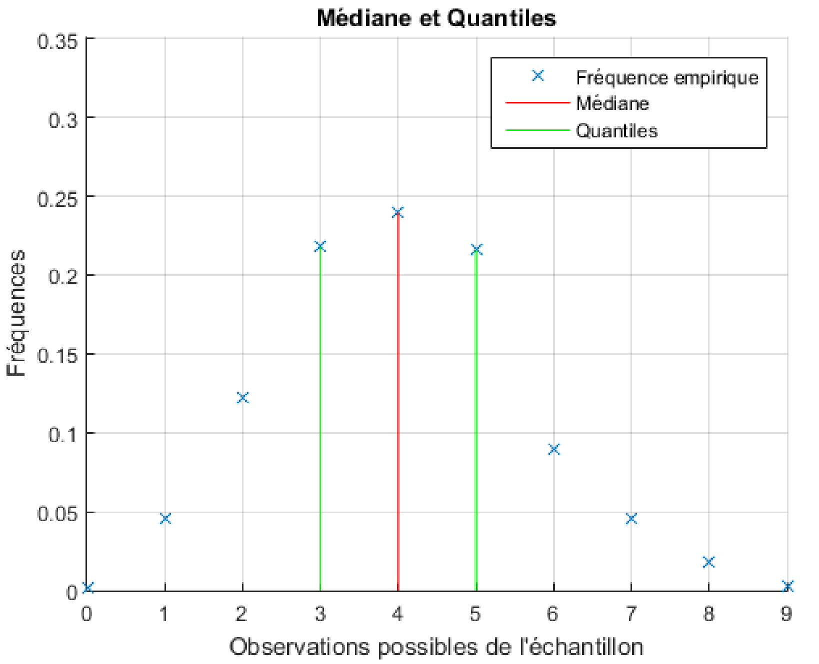
car les variables aléatoires sont iid. Nous avons donc

$$\begin{aligned} Var(X) &= np + n(n-1)p^2 - n^2p^2 \\ &= np - np^2 \\ &= n(1-p)p. \end{aligned}$$

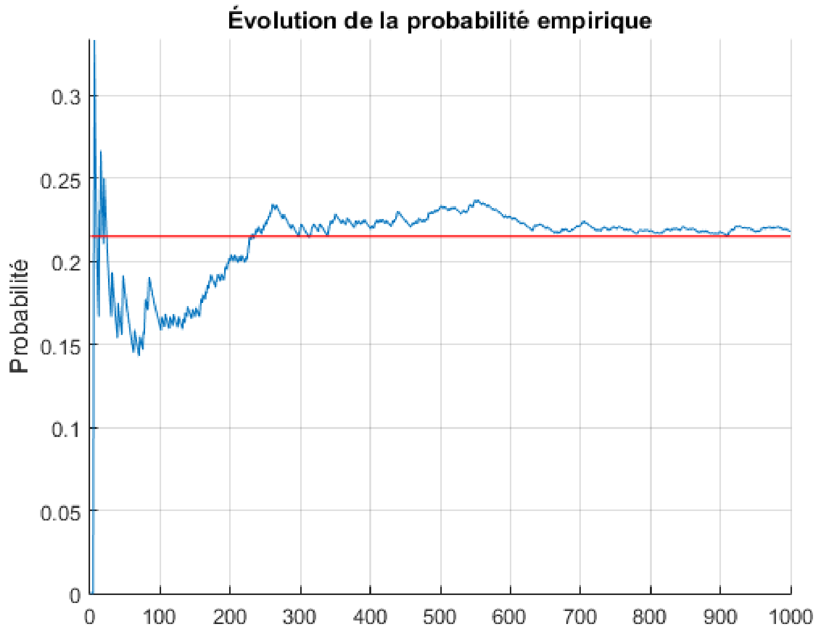
Nous disposons d'un échantillon de taille 1000 qui suit une loi binomiale de paramètres $n = 10$ et $p = 0,4$. Nous obtenons alors comme fréquence empirique :



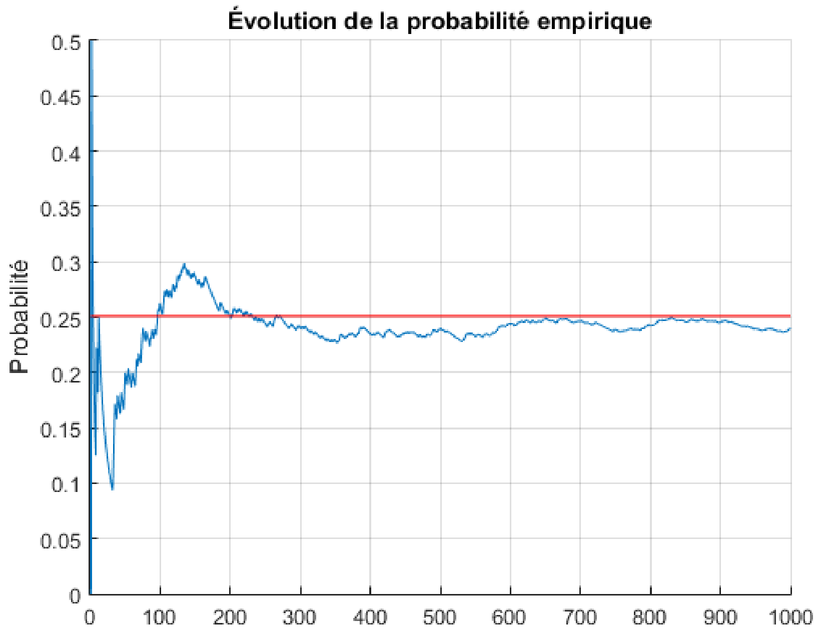
Voici encore le graphique avec la médiane pour les quantiles $\frac{1}{4}$ et $\frac{3}{4}$.



Regardons maintenant l'évolution de la fréquence empirique d'obtenir la valeur $k = 3$.



Pour $k = 4$, nous avons



1.5 Loi géométrique

Définition 7. *Soit X une variable aléatoire. X suit une* **loi géométrique** *si la probabilité d'obtenir k est la probabilité de nécessiter k essais pour obtenir le premier succès d'une loi de Bernoulli, c'est-à-dire*

$$\mathbb{P}(X = k) = (1-p)^{k-1}p,$$

avec $k \in \mathbb{N}^*$ et où $p \in]0,1[$ est la probabilité d'un succès de la loi de Bernoulli. Nous écrivons $X \sim Geo(p)$, lorsque X suit une loi géométrique de paramètre p.

Remarquons que la loi géométrique est définie sur tous les entier k, est ce que la définition de $\mathbb{P}$ donne bien un loi ? En effet, nous avons :

$$\begin{aligned}\sum_{k=1}^{+\infty} \mathbb{P}(X = k) &= p\sum_{k=0}^{+\infty}(1-p)^k \\ &= p\frac{1}{1-(1-p)} \\ &= \frac{p}{1-1+p} = 1.\end{aligned}$$

Calculons maintenant l'espérance de la loi geométrique. Remarquons tout d'abord que si $f(x) = \frac{1}{1-x}$ alors la série de Taylor est $f(x) = \sum_{k=0}^{+\infty} x^k$ et $f'(x) = \frac{1}{(1-x)^2} = \sum_{k=1}^{+\infty} kx^{k-1}$ pour $x \in]0,1[$. D'où

$$\begin{aligned}E[X] &= \sum_{k=1}^{+\infty} k\mathbb{P}(X = k) \\ &= \sum_{k=1}^{+\infty} kp(1-p)^{k-1} \\ &= p\sum_{k=1}^{+\infty} k(1-p)^{k-1} \\ &= p\frac{1}{(1-(1-p))^2} \\ &= p\frac{1}{p^2} = \frac{1}{p}.\end{aligned}$$

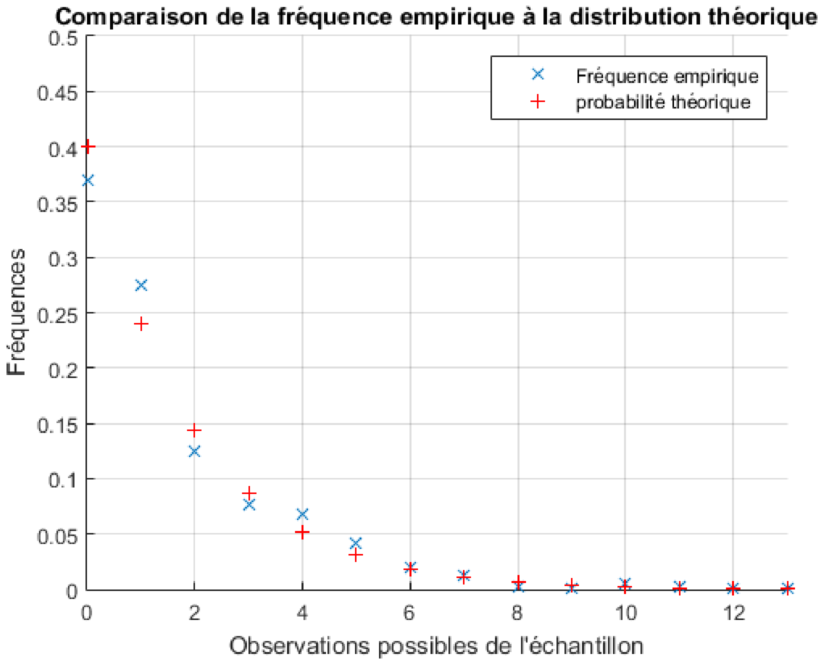
Remarquons que pour $f(x) = \frac{1}{1-x}$, la dérivée seconde est $f''(x) = \frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k(k-1)x^{k-2}$, et donc

$$\sum_{k=1}^{\infty} k^2x^{k-1} = x\sum_{k=1}^{\infty} k(k-1)x^{k-2} + \sum_{k=1}^{\infty} kx^{k-1} = \frac{2x}{(1-x)^3} + \frac{1}{(1-x)^2}.$$

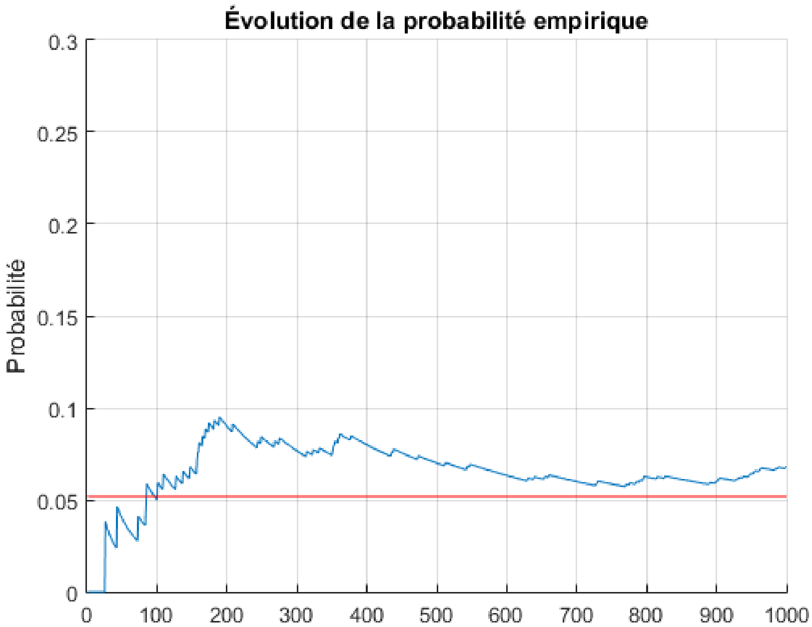
Ceci nous permet de calculer la variance

$$\begin{aligned}Var(X) &= E[X^2] - E[X]^2 \\ &= p\sum_{k=1}^{+\infty} k^2(1-p)^{k-1} - \frac{1}{p^2} \\ &= p\frac{2(1-p)}{p^3} + \frac{p}{p^2} - \frac{1}{p^2} \\ &= \frac{2-2p+p-1}{p^2} \\ &= \frac{1-p}{p^2}.\end{aligned}$$

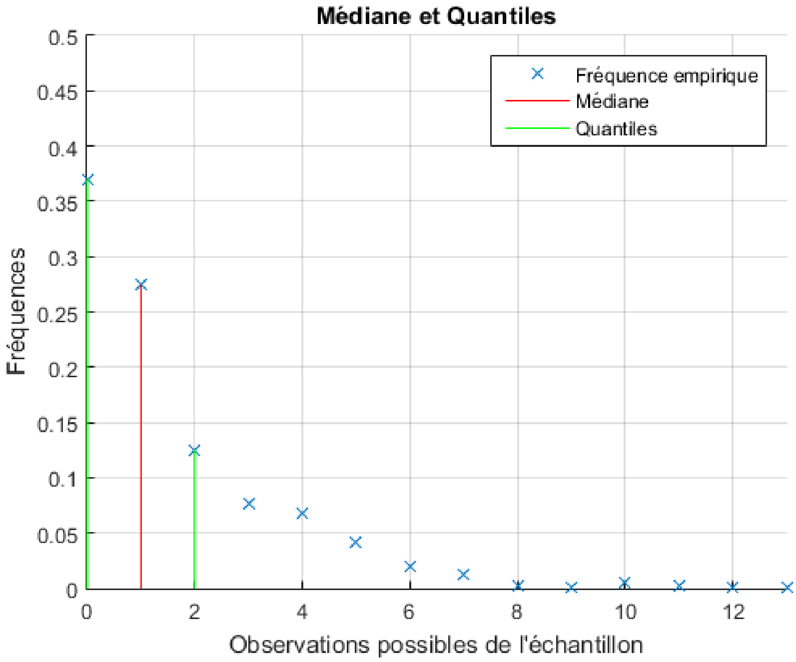
Nous disposon maintenant d'un échantillon de taille 1000 d'une loi géométrique de paramètre 0,4. En calculant la fréquence empirique nous obtenons le graphique suivant :



Regardons l'évolution de la fréquence de $\mathbb{P}(X = 4)$ lors du parcours de l'échantillon.



Voici encore un graphique qui montre aussi la médiane avec les quantiles $\frac{1}{4}$ et $\frac{3}{4}$.



1.6 Loi de Poisson

Définition 8. *Soit X une variable aléatoire. Nous disons que X suit une* ***loi de Poisson****, si*

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda},$$

pour $k \in \mathbb{N}$. Nous écrivons $X \sim P(\lambda)$.

Pour montrer que $P(\lambda)$ est bien une loi de probabilité remarquons d'abord que le développement de Taylor de $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + ... = \sum_{i=1}^{\infty} \frac{x^i}{i!}$ pour $x \in \mathbb{R}$. Ceci nous donne :

$$\begin{aligned}
\sum_{k=0}^{+\infty} \mathbb{P}(X = k) &= \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}e^{-\lambda} \\
&= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\
&= e^{-\lambda}e^{\lambda} \\
&= 1.
\end{aligned}$$

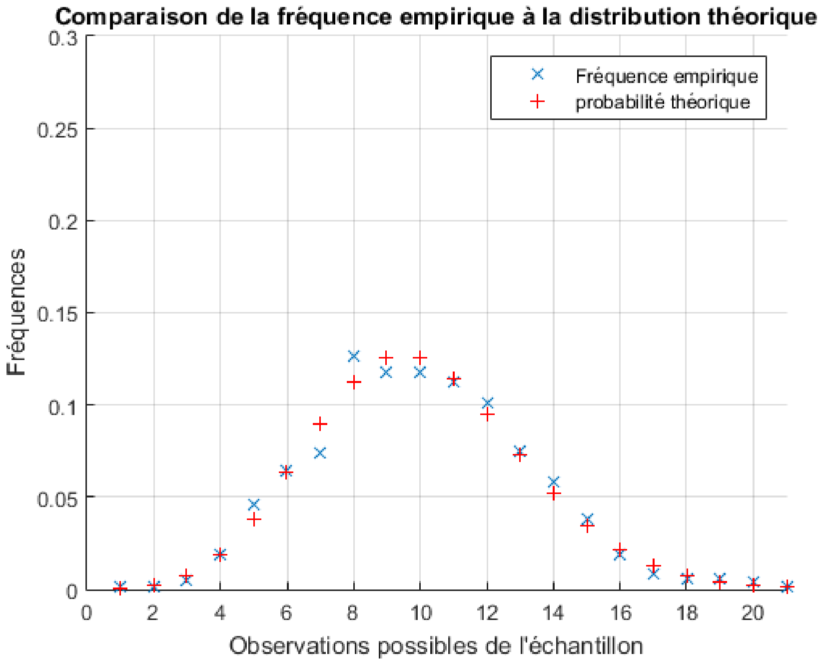
Calculons maintenant l'espérance pour la loi de Poisson.

$$\begin{aligned}
E[X] &= \sum_{k=0}^{+\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{+\infty} k\mathbb{P}(X = k) \\
&= \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda} \\
&= \frac{\lambda}{e^{\lambda}} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \frac{\lambda}{e^{\lambda}} \sum_{k=0}^{+\infty} \frac{\lambda^k}{(k)!} \\
&= \frac{\lambda}{e^{\lambda}}e^{\lambda} = \lambda.
\end{aligned}$$

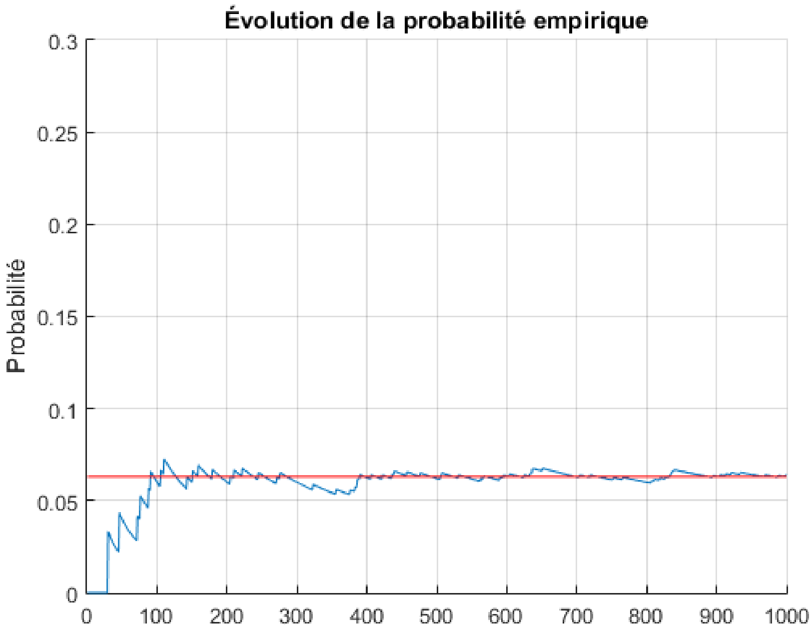
Ceci nous permet de calculer la variance de la loi de Poisson.

$$\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \sum_{k=1}^{+\infty} \left(k^2 \frac{\lambda^k}{(k-1)!}e^{-\lambda} \right) - \lambda^2 \\
&= \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \left(\frac{\lambda^k}{(k-1)!} \right) - \lambda^2 \\
&= \lambda e^{-\lambda}(1 + \lambda)e^{\lambda} - \lambda^2 \\
&= \lambda.
\end{aligned}$$

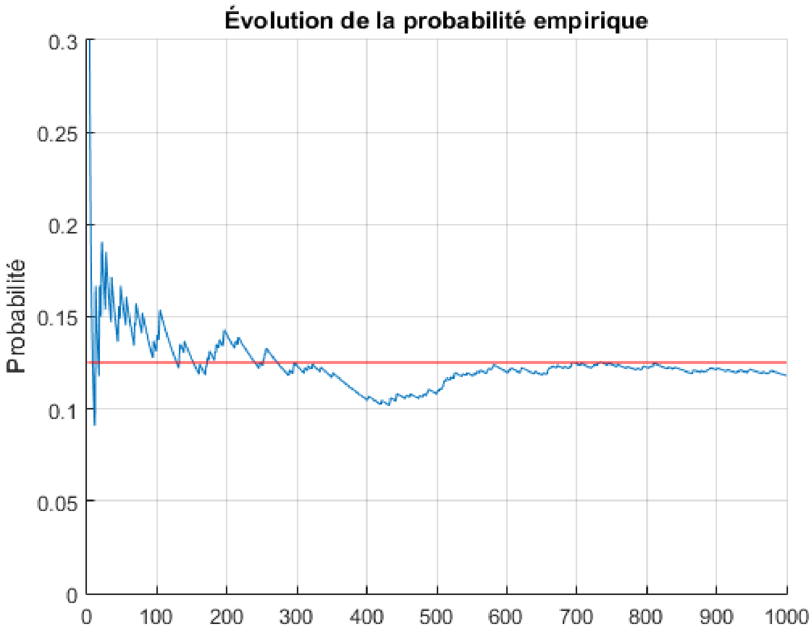
Analysons un échantillon de taille 1000 d'une loi de Poisson de paramètre 10. En calculant la fréquence empirique nous obtenons le graphique suivant :



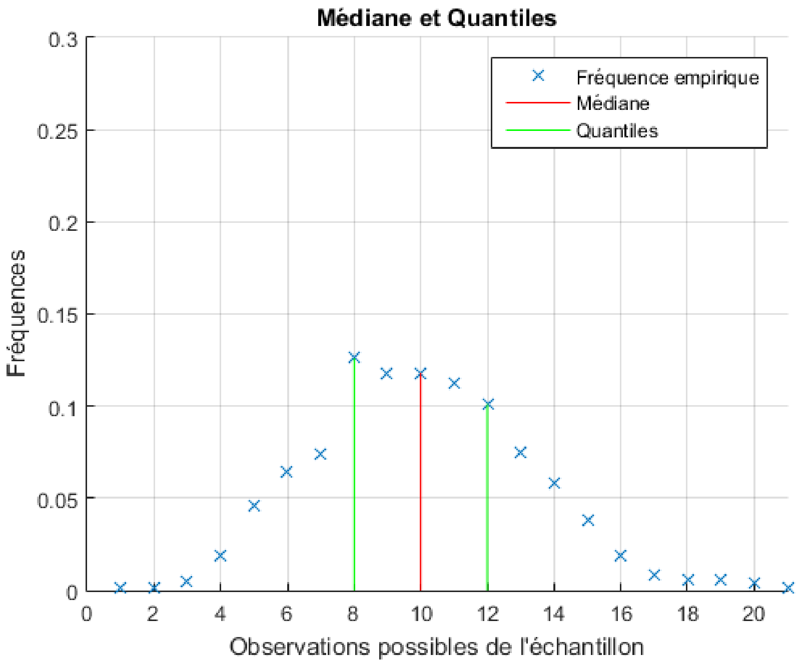
Nous nous intéressons à l'évolution de la fréquence d'obtenir $k = 6$.



Enusite pour $k = 10$.



Voici encore un graphique qui montre aussi la médiane avec les quantiles $\frac{1}{4}$ et $\frac{3}{4}$.



1.7 Méthode des moments

Supposons que nous disposons d'un nombre d'observations à valuers discrètes d'un échantillon i.i.d de loi inconnue. Ce que nous voulons déterminer est la loi que pourrait suivre notre variable aléatoire. Nous supposons ici, des variables aléatoires discrètes, de loi uniforme, géométrique, de Poisson ou binomiale (binomiale de paramètre n, le nombre de lancer total, connu, mais de paramètre p inconnu). (Nous ignorons la loi de Bernoulli car $k = 0$ ou $k = 1$.) Une première approximation des paramètres des différentes lois survient à l'aide de l'espérance. En effet, en connaissant l'espérance empirique, nous pouvons déduire les paramètres des différentes lois de manière empirique en assimilant la moyenne des valeurs d'espérance. Nous comparons ensuite la fonction de probabilité avec nos observations, pour déterminer éventuellement quelle loi serait plus probable.

Soit m, la moyenne des observations. Posons $E[X] = m$.

(i) Si la population suit une loi uniforme, alors

$$E[X] = \frac{n+1}{2} \Leftrightarrow m = \frac{n+1}{2}$$
$$\Leftrightarrow n = 2m - 1 \qquad n \text{ doit être entier.}$$

(ii) Si $X \sim Geo(p)$

$$m = \frac{1}{p} \Leftrightarrow p = \frac{1}{m} \quad (\in]0,1[).$$

(iii) Si $X \sim Bin(n,p)$ (avec n connu)

$$m = np \Leftrightarrow p = \frac{m}{n} \quad (\in]0,1]) \qquad \text{i.e. } m < n.$$

(iv) Si $X \sim P(\lambda)$, nous avons que $m = \lambda$.

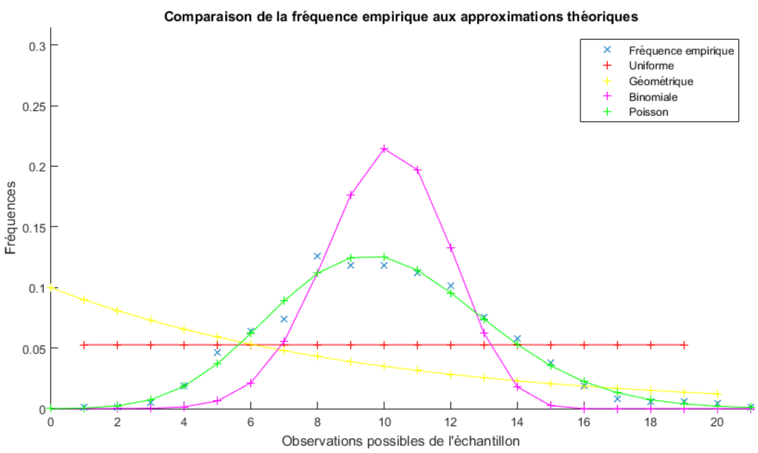
Exemple 9. *Nous disposons d'un échantillon de taille* 1000 *d'une loi de Poisson de paramètre* 10. *La moyenne de notre observation est* $m = 10,04$. *Les candidats possibles sont*

→ *La loi uniforme de paramètre* $n = 2m - 1 = 19$ *(arrondi pour être entier).*

→ *La loi geométrique de paramètre* $p = \frac{1}{m} = 0,0996$.

→ *La loi binomiale de paramètre* $n = 15$ *(choisi abitrairement, mais de façon à ce que* $n > m$ *pour que* $\frac{m}{n} < 1$*) et* $p = \frac{m}{n} = 0,6693$.

→ *La loi de Poisson de paramètre* $\lambda = m = 10,04$.



Nous constatons que la courbe de la loi de Poisson est celle qui approche le mieux notre fréquence empirique de nos observations, ce qui est tout a fait attendu puisque nous avons choisi un échantillon qui suit une loi de Poisson.

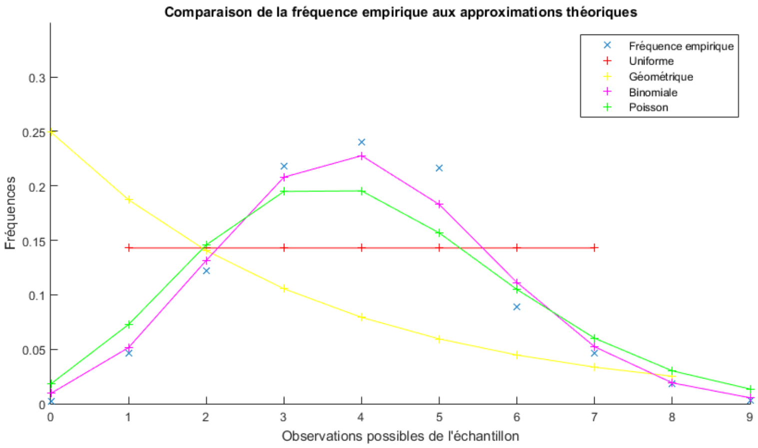
Refaisons maintenant la même expérience pour un échantillon de taille 1000 *d'une loi binomiale de paramètres* 10 *et* $0,4$. *La moyenne de notre observation est cette fois ci* $m = 4,011$. *Nous obtenons donc les candidats possibles suivants*

→ *La loi uniforme de paramètre* $n = 2m - 1 = 7$ *(arrondi pour être entier).*

→ *La loi geométrique de paramètre* $p = \frac{1}{m} = 0,2493$.

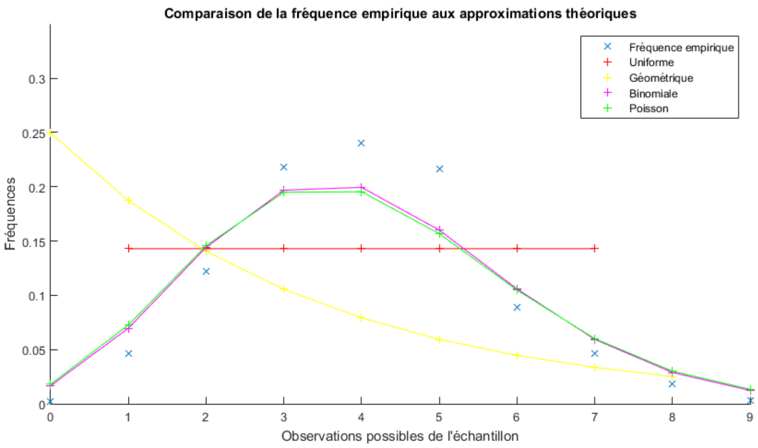
→ *La loi binomiale de paramètre* $n = 15$ *(choisi abitrairement, mais de façon à ce que* $n > m$ *pour que* $\frac{m}{n} < 1$*) et* $p = \frac{m}{n} = 0,2674$.

→ *La loi de Poisson de paramètre* $\lambda = m = 4,011$.



Ici, la courbe la plus proche est la binomiale de paramètres $(15; 0,2674)$. *Remaqruons que par notre choix arbitraire de l'estimation de* $n = 15$, *nous aurions aussi pu avoir qu'aucune courbe n'approxime la fréquence empirique.*

Voici un exemple avec les mêmes observations, mais avec $n = 1000$ *pour la loi binomiale.*



Remarquons aussi que la loi binomiale tend vers la loi de Poisson.

Théorème 10. *Si n est grand et p est petit, nous pouvons approximer une loi $Bin(n,p)$ par une loi $P(np)$.*

Chapitre 2

Cas continu

Commençons avec un exemple avant d'introduire les définitions importantes.

Exemple 11. *(Distribution uniforme sur* $[0,1]$*)*
Nous voudrions avoir $\mathbb{P}([a,b[) = b-a$ *pour* $a,b \in [0,1]$*. Mais nous ne pouvons pas définir la probabilité* $\mathbb{P}$ *sur* $\mathcal{P}(\Omega)$ *à cause du problème suivant :*

Soit $a \in [0,1[$*. Nous avons*

$$\mathbb{P}(\{a\}) \subseteq \mathbb{P}\left(\left[a, a+\frac{1}{n}\right[\right) = \frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 0.$$

Nous obtenons $\mathbb{P}(\{a\}) = 0$*. Et donc, pour* $A \subset [0,1[$

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{a\in A}\{a\}\right) \neq \sum_{a\in A}\mathbb{P}(\{a\}) = 0.$$

Définition 12. *Nous appelons* ***espace de probabilité****, tout triple* $(\Omega, \mathcal{A}, \mathbb{P})$ *tel que*

1. *Ω est un ensemle,*
2. *$\mathcal{A}$ est une tribu,*
3. *$\mathbb{P}$ es une mesure de probabilité.*

Nous considérons $\mathcal{A} = \mathcal{B}(\mathbb{R})$ *la tribu borélienne de* $\mathbb{R}$*.*

Définition 13. *Nous appelons* ***variable aléatoire****, toute fonction* $X : \Omega \to \mathbb{R}$ *mesurable, c'est-à-dire*

$$X^{-1}(B) = \{X \in B\} \in \mathcal{A}, \forall B \in \mathcal{B}(\mathbb{R}).$$

Définition 14. *Nous appelons* ***fonction de répartition*** *d'une loi* X

$$\begin{aligned} F_X : \mathbb{R} &\to [0,1] \\ b &\mapsto F_X(b) := \mathbb{P}_X(]-\infty, b]) = \mathbb{P}\{X \leq b\}. \end{aligned}$$

Cette fonction caractérise entièrement la loi de X*.*

Définition 15. *Nous disons que* X *admet une fonction de densité* f_X*, s'il existe une fonction* $f_X : \mathbb{R} \to \mathbb{R}_+$ *intégrable telle que*

$$F_X(a) = \int_{-\infty}^{a} f_x(t)dt.$$

Propriétés 16. *(i)* $\int_{-\infty}^{+\infty} f_X(t)dt = 1$*.*

(ii) Si $a < b$*, nous avons*

$$\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f_X(t)dt.$$

(iii) Si f_X est continue, alors F_X est dérivable et

$$\frac{d}{dt}F_X(t) = f_X(t).$$

Dans le cas continue nous pouvons toujours utiliser

$$E[X] = \int_{-\infty}^{+\infty} f_X(t)dt$$

pour calculer l'espérance.

Remarque 17.

$$E[h(X)] = \int_{-\infty}^{+\infty} h(t)f_X(t)dt,$$

et nous avons toujous

$$\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= \int_{-\infty}^{+\infty} t^2 f_X(t)dt - \left(\int_{-\infty}^{+\infty} t f_X(t)dt\right)^2. \end{aligned}$$

2.1 Loi uniforme

Nous disons que X suit une loi uniforme $X \sim U([a,b])$ avec $a < b$ si $f_X(t) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(t)$. Effectivement,

$$\int_{-\infty}^{+\infty} f_X(t)dt = \int_a^b \frac{1}{b-a}dt = 1.$$

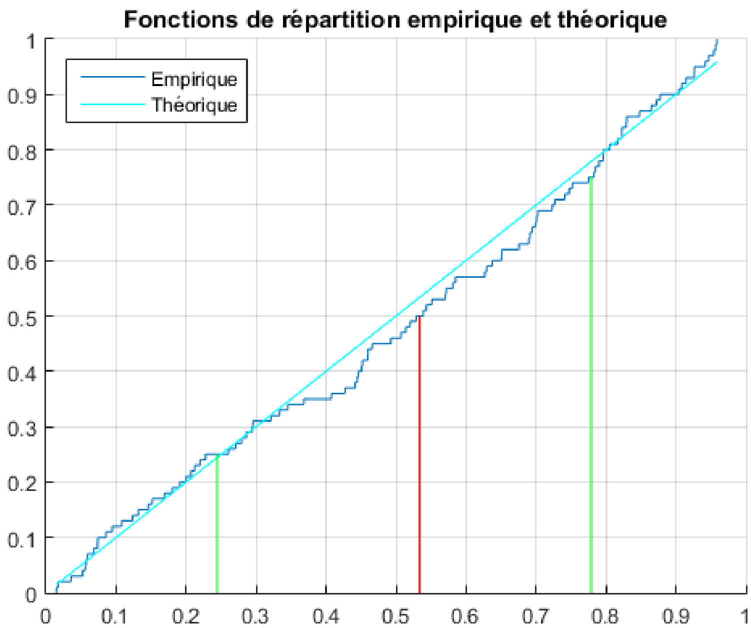
Nous pouvons donc calculer l'espérance et la variance.

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} t\frac{1}{b-a}\mathbb{1}_{[a,b]}dt \\ &= \frac{1}{b-a}\int_a^b tdt \\ &= \frac{1}{b-a}\left[\frac{1}{2}t^2\right]_a^b \\ &= \frac{b^2-a^2}{2(b-a)} \\ &= \frac{b+a}{2} \end{aligned}$$

et

$$\begin{aligned} Var(X) &= \int_{-\infty}^{+\infty} t^2\frac{1}{b-a}\mathbb{1}_{[a,b]}dt - \frac{(b+a)^2}{4} \\ &= \frac{b^2-a^3}{3(b-a)} - \frac{(b+a)^2}{4} \\ &= \frac{b^2+ab+a^2}{3} - \frac{b^2+a^2+2ab}{4} \\ &= \frac{(b-a)^2}{12}. \end{aligned}$$

Analysons un échantillon de taille 100 de loi uniforme de paramètre $a = 0$ et $b = 1$. Nous obtenons le graphique suivant pour la fonction de répartition.



2.2 Loi exponentielle

Nous disons que X suit une loi exponentielle $X \sim Exp(\lambda)$ avec $\lambda > 0$ si $f_X(t) = \lambda e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t)$. Nous remarquons que

$$\int_{-\infty}^{+\infty} f_X(t)dt = \int_0^{+\infty} \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t}\right]_0^{+\infty} = 1.$$

Calculons maintenant l'espérance.

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} t\lambda e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t)dt \\ &= \int_0^{+\infty} \lambda t e^{-\lambda t} dt. \end{aligned}$$

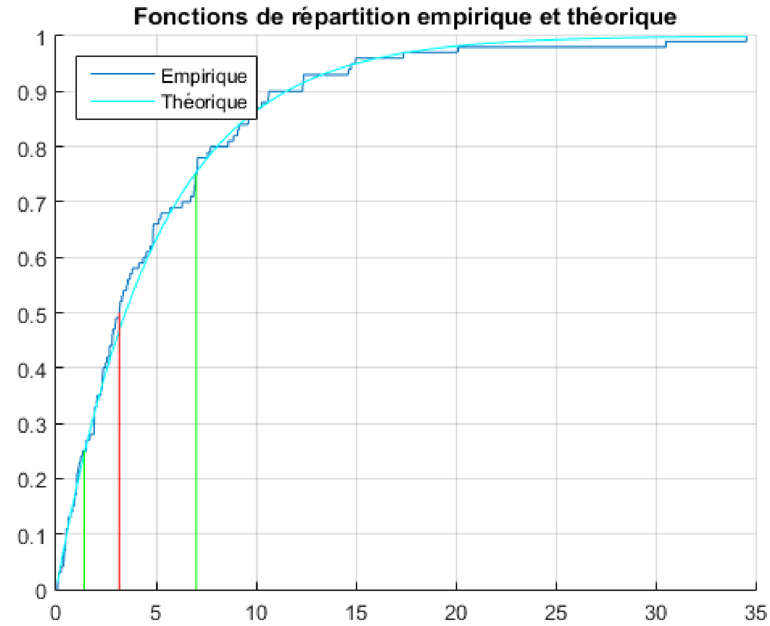
En faisant une intégration par parties nous obtenons

$$\begin{aligned} E[X] &= \left[-te^{-\lambda t}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda t} dt \\ &= 0 + \frac{1}{\lambda} \int_0^{+\infty} \lambda e^{-\lambda t} \\ &= \frac{1}{\lambda} \cdot 1 = \frac{1}{\lambda}. \end{aligned}$$

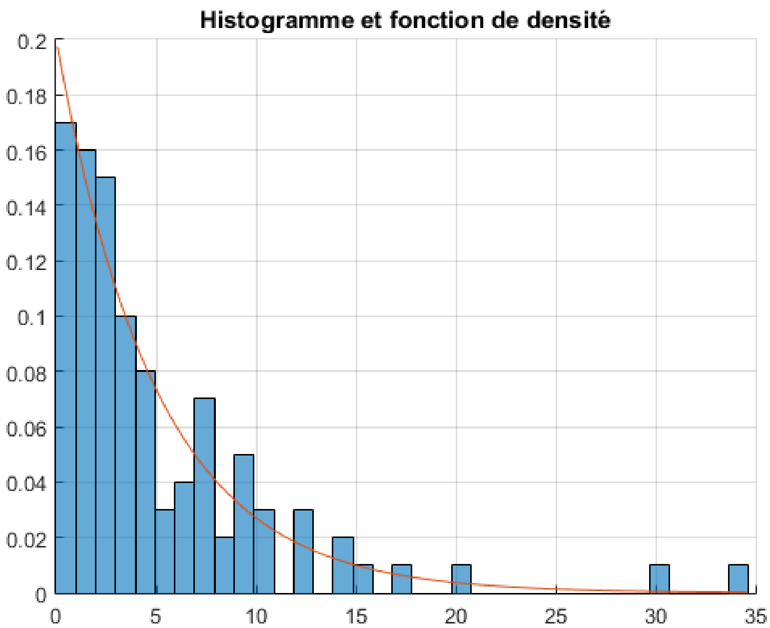
Ceci nous permet de calculer la variance.

$$\begin{aligned} Var(X) &= \int_0^{+\infty} t^2 \lambda e^{-\lambda t} dt - \frac{1}{\lambda^2} \\ &= \left[-t^2 e^{-\lambda t}\right]_0^{+\infty} + \int_0^{+\infty} 2te^{-\lambda t} dt - \frac{1}{\lambda^2} \qquad \text{, intégration par parties} \\ &= 0 + \frac{2}{\lambda} \int_0^{+\infty} \lambda t e^{-\lambda t} dt - \frac{1}{\lambda^2} \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}. \end{aligned}$$

Nous observons un échantillon de taille 100 de loi exponentielle de paramètre 5. La fonction de répartition ressemble à



et le graphique de la densité est



2.3 Loi normale (ou gaussienne)

Nous disons que X suit une loi normale $X \sim N(\mu, \sigma^2)$, si $f_X = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}$. Ceci est bien une probabilité, car $\int_{-\infty}^{+\infty} e^{\frac{-t2}{2}} dt = \sqrt{2\pi}$, et $\int_{-\infty}^{+\infty} e^{\frac{-t^2}{2\sigma^2}} dt = \sqrt{2\pi\sigma^2}$. Donc nous obtenons

$$\int_{-\infty}^{+\infty} f_X dt = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \sqrt{2\pi\sigma^2} = 1.$$

Calculons maintenant l'espérance.

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} \mu + (t-\mu) f_x(t) dt \\ &= \mu + \int_{-\infty}^{+\infty} (t-\mu) f_X(t) dt \\ &= \mu, \end{aligned}$$

car $(t-\mu) f_X(t)$ est une fonction impaire autour de μ, donc l'intégrale vaut 0.

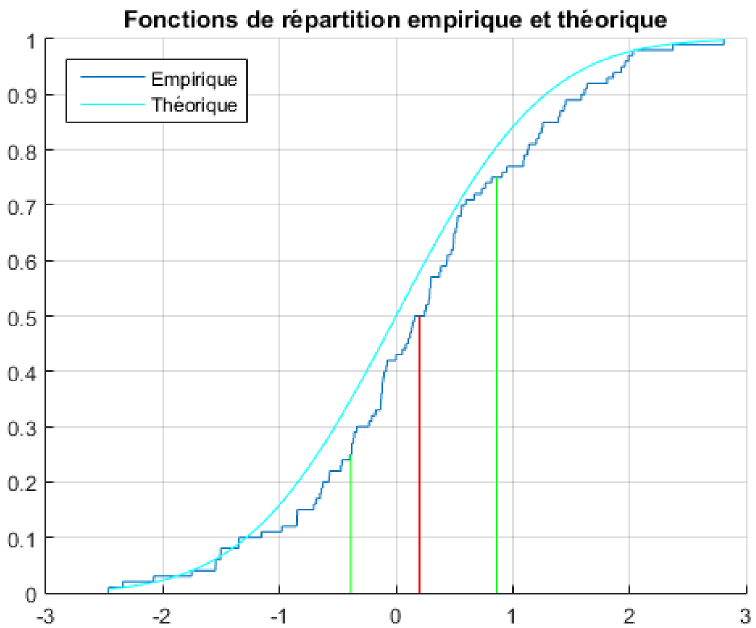
La variance est égale à

$$\begin{aligned} Var(X) &= \int_{0}^{+\infty} (t-\mu)^2 f_X(t) dt \\ &= \sigma^2, \end{aligned}$$

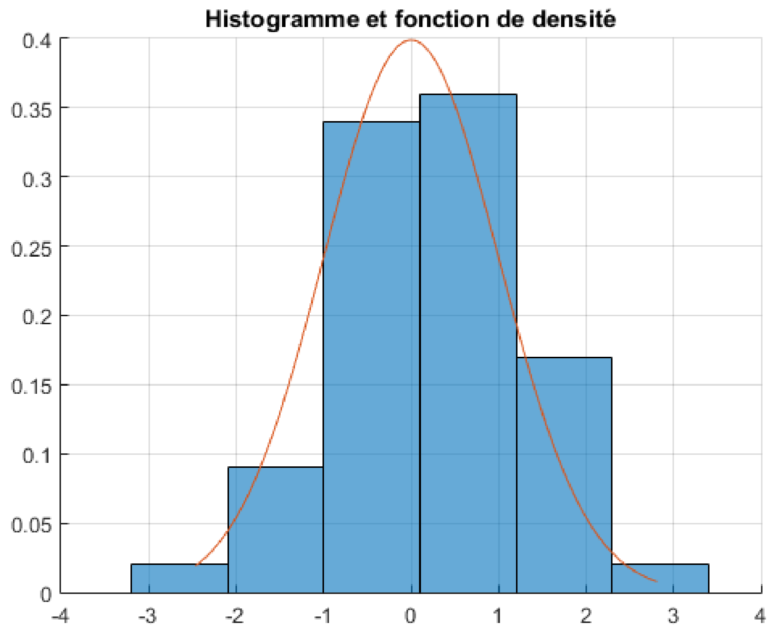
car

$$\int_{0}^{+\infty} t^2 e^{\frac{-t^2}{2\sigma^2}} dt = \sigma^2 \sqrt{2\pi\sigma^2}.$$

Analysons maintenant un échantillon de taille 100 de loi normale de paramètres $\mu = 0$ et $\sigma^2 = 1$. Le graphique de la fonction de répartition ressemble à ceci :



Et voici le graphique qui illustre la densité :



2.4 Loi du chi-carré

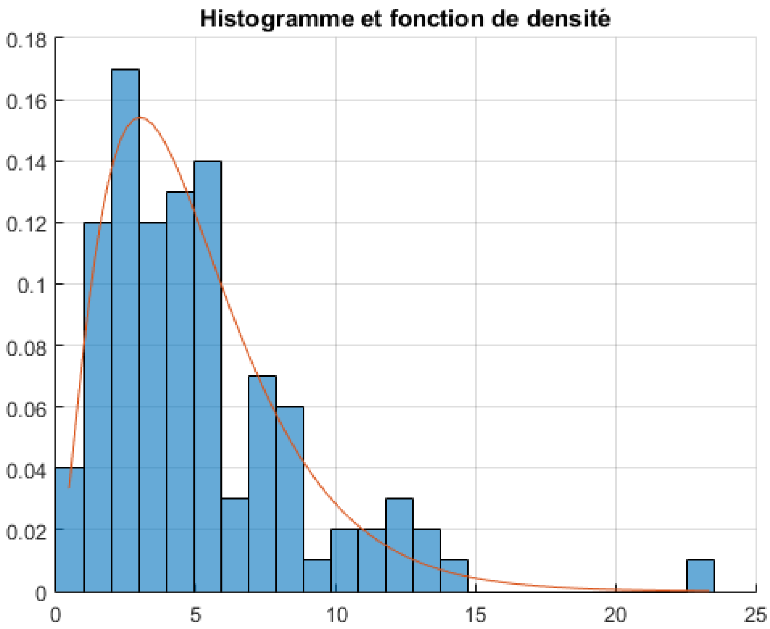
Soient $Z_1, ..., Z_n \simeq Z \sim N(0,1)$ des variables aléatoires gaussiennes centrées et réduites, c'est-à-dire $\mu = 0$ et $\sigma = 1$. Alors $X = \sum_{i=1}^{n} Z_i^2$ suit une loi de chi-carré à n degrès de liberté, $X \sim \mathcal{X}_n^2$. Sa fonction de densité est

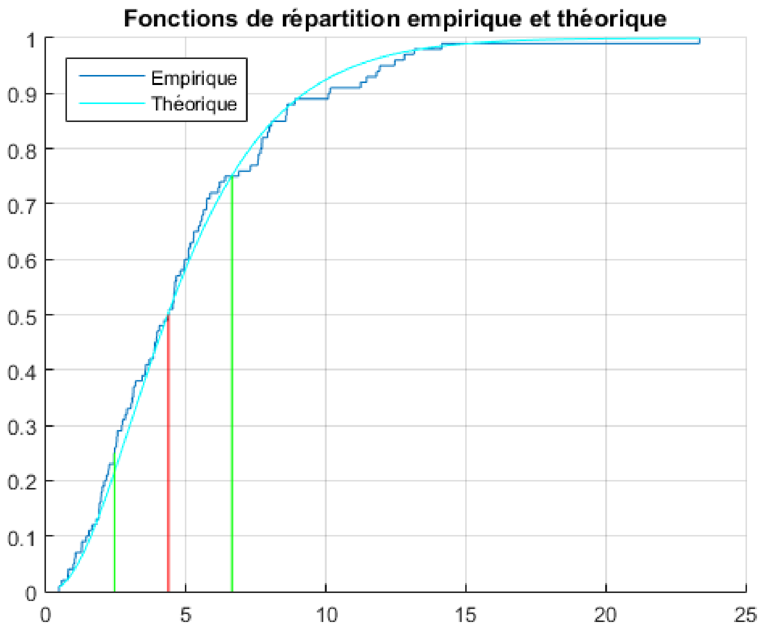
$$f_{\mathcal{X}_n^2}(t) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} \mathbb{1}_{\{\mathbb{R}_+\}}, n > 0$$

où $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$ est la fonction gamma.

L'espérance est $E[\mathcal{X}_n^2] = n$ et la variance vaut $Var(\mathcal{X}_n^2) = 2n$.

Les graphiques suivants montrent la densité d'une loi chi-carré de paramètre 5 et sur un échantillon de taille 100.




Fonctions de répartition empirique et théorique
Empirique
Théorique
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25

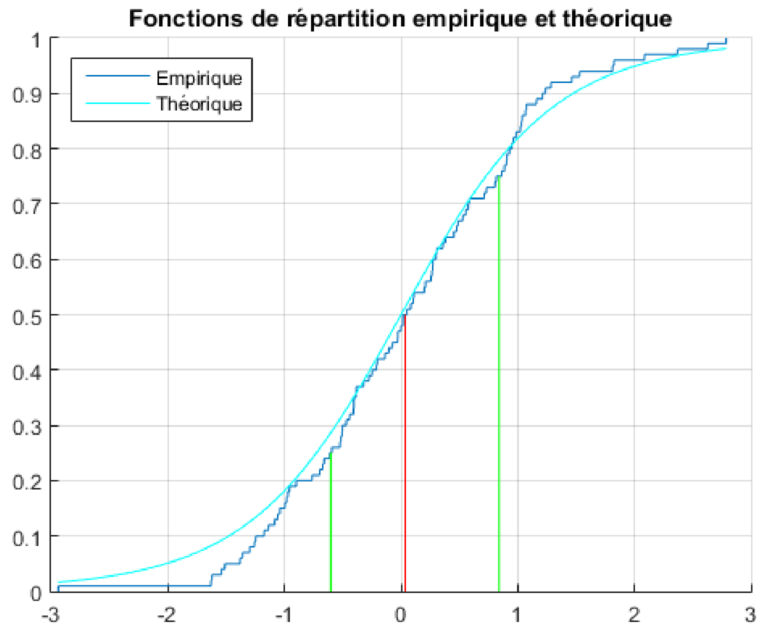
2.5 Loi de Student

Soient $Z \sim N(0,1)$ une gaussienne centrée, réduite et $X \sim \mathcal{X}_n^2$ une chi-carré à n degrès de liberté. Alors $T = \frac{Z}{\sqrt{\frac{X}{n}}}$ suit une loi de Student à n degrès de liberté, $T \sim t_n$. Sa fonction de densité est

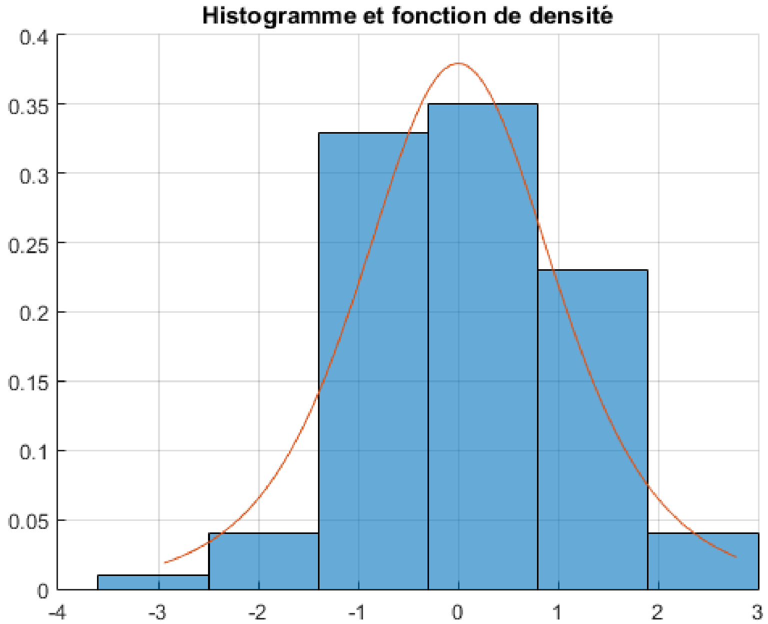
$$f_{t_n} = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, n > 0.$$

Son espérance n'est pas définie pour $n = 1$ et pour $n \geq 2$, $E[t_n] = 0$. Pour tout $n \leq 2$ la variance vaut $Var(t_n) = +\infty$ et pour tout $n \geq 3, Var(t_n) = \frac{n}{n-2}$.

Observons maintenant la fonction de répartition d'un échantillon de taille 100 et d'une loi de Student de paramètre 5.



Voici la densité du même échantillon.



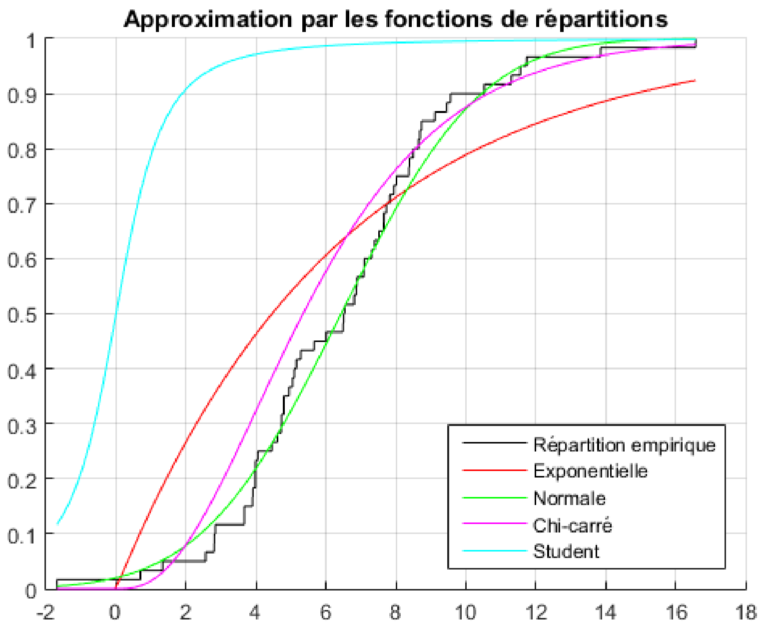
2.6 Estimation de la loi par la méthode des moments

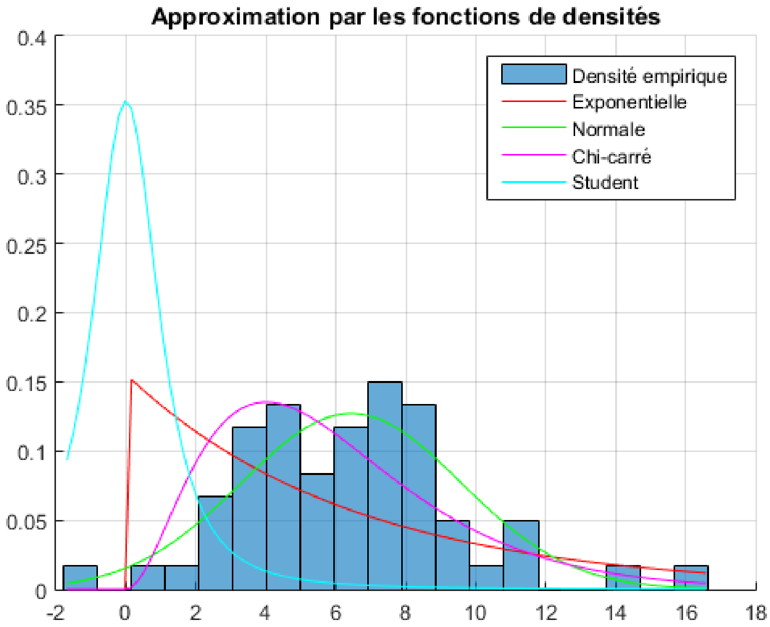
Similairement à la section 1.7 (Méthode des moments), nous voudrons approximer la loi d'une population à l'aide de l'information sur l'échantillon obtenue.

Nous allons comparer les observations avec la meilleure approximation des lois exponentielle, normale, chi-carré et de Student, par la méthode des moments, c'est-à-dire, estimer les paramètres à l'aide de la moyenne et de la variance.

Exemple 18. *Soit un échantillon aléatoire de taille* 60. *Nous avons comme moyenne des observations* $6,43$ *et comme variance* $9,84$. *Cela nous donne comme approximations*

1. *une loi exponentielle* $Exp(6,43)$,
2. *Une loi normale* $N(6,43; 9,84)$,
3. *une loi chi-carré* $\mathcal{X}_6$,
4. *une loi Student* t_2.





Nous constatons que la normale approche les observations au mieux.
Remarquons que la loi de Student n'est relevant que si la moyenne est presque nulle.

2.7 Estimation : Modèle OMMeR

Supposons que X suit une normale de paramètres inconnues $X \sim N(\mu, \sigma^2)$.

Objectif : Estimer les paramètres à l'aide des caractéristiques décrites par la loi $X \sim N(\mu, \sigma^2)$.

Modèle : Nous supposons avoir un modèle d'échantillonage, $(X_1, ..., X_n) \simeq X \sim N(\mu, \sigma^2)$. Nous choisissons un échantillon i.i.d. de taille n dans la population.

Méthode : Nous estimons les paramètres à l'aide de l'échantillon.

Estimateur :

$$\begin{aligned} \hat{\mu} &= \bar{X_n} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 &= \tilde{S_n}^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X_n})^2. \end{aligned}$$

Nous disposons des statistiques pivots suivant :

$$\begin{aligned} T &= \tfrac{\bar{X_n}-\mu}{\tilde{S_n}}\sqrt{n} &\sim t_{n-1} \\ Y &= (n-1)\tfrac{\tilde{S_n}^2}{\sigma^2} &\sim \mathcal{X}^2_{n-1}. \end{aligned}$$

Résultat : (a) Mesures d'incertitude de l'espérance μ :

Soit $c \in \mathbb{R}$, alors

$$\begin{aligned} \mathbb{P}(\hat{\mu} - c \leq \mu \leq \hat{\mu} + c) &= \mathbb{P}(\mu \leq \hat{\mu} + c) - \mathbb{P}(\mu \leq \hat{\mu} - c) \\ &= \mathbb{P}(\mu - c \leq \hat{\mu}) - \mathbb{P}(\mu + c \leq \hat{\mu}) \\ &= \mathbb{P}(\mu - c \leq \hat{\mu} \leq \mu + c) \\ &= 1 - \mathbb{P}(\hat{\mu} > \mu + c) - \mathbb{P}(\hat{\mu} < \mu - c) \\ &= 1 - \mathbb{P}\left(\sqrt{n}\frac{\hat{\mu}-\mu}{\tilde{S_n}} > \frac{\mu + c - \mu}{\tilde{S_n}}\sqrt{n}\right) - \mathbb{P}\left(\sqrt{n}\frac{\hat{\mu}-\mu}{\tilde{S_n}} > \frac{\mu - c + \mu}{\tilde{S_n}}\sqrt{n}\right) \\ &= 1 - \mathbb{P}\left(T \geq \frac{c}{\tilde{S_n}}\sqrt{n}\right) - \mathbb{P}\left(T \leq -\frac{c}{\tilde{S_n}}\sqrt{n}\right) \\ &= 1 - 2\mathbb{P}\left(T \geq \frac{c}{\tilde{S_n}}\sqrt{n}\right). \end{aligned}$$

(b) Intervalle de confiance pour μ :

Nous cherchons $[L, U]$ tel que $\mathbb{P}(L \leq \mu \leq U) \leq 1 - \alpha$. Or comme $\sqrt{n}\frac{\bar{X_n}-\mu}{\tilde{S_n}} \sim t_{n-1}$. Donc

$$\mathbb{P}\left(t_{n-1} \leq \frac{\bar{X_n} - \mu}{\tilde{S_n}}\sqrt{n} \leq -t_{n-1}\right) = 1 - \alpha. \tag{2.1}$$

Posons q_1 quantile $\frac{\alpha}{2}$ de t_{n-1} et q_2 quantile $\frac{\alpha-1}{2}t_{n-1}$ de de t_{n-1}. Nous avons $\frac{\alpha}{2} = t_{n-1} = -t_{n-1} = 1 - \frac{\alpha}{2}$ par symétrie de la loi de Student. Donc2.1 équivaut

$$\begin{aligned} & \mathbb{P}\left(-\tfrac{\tilde{S_n}}{\sqrt{n}}t_{n-1} - \bar{X_n} \leq -\mu \leq \tfrac{\tilde{S_n}}{\sqrt{n}}t_{n-1} - \bar{X_n}\right) &= 1 - \alpha \\ \Leftrightarrow \quad & \mathbb{P}\left(\bar{X_n} - \tfrac{\tilde{S_n}}{\sqrt{n}}t_{n-1} \leq \mu \leq \tfrac{\tilde{S_n}}{\sqrt{n}}t_{n-1} + \bar{X_n}\right) &= 1 - \alpha \\ \Leftrightarrow \quad & \mu \in \left[\bar{X_n} - \tfrac{\tilde{S_n}}{\sqrt{n}}; \bar{X_n} + \tfrac{\tilde{S_n}}{\sqrt{n}}t_{n-1}\right]. \end{aligned}$$

Ceci est l'intervalle de confiance de niveau de confiance $1 - \alpha$.

(c) Mesure d'incertitude pour σ^2 :

$$\begin{aligned}
\mathbb{P}(\sigma^2 \geq k\tilde{S_n}^2) &= \mathbb{P}\left(\frac{1}{k} \geq \frac{\tilde{S_n}^2}{\sigma^2}\right) \\
&= \mathbb{P}\left(\frac{n-1}{k} \geq (n-1)\frac{\tilde{S_n}^2}{\sigma^2}\right) \\
&= \mathbb{P}\left(Y \leq \frac{n-1}{k}\right).
\end{aligned}$$

(d) Intervalle de confiance pour σ^2 :

Nous cherchons $[L, U]$ tel que $\mathbb{P}(L \leq \sigma^2 \leq U) = 1 - \alpha$ en sachant que $\frac{n-1}{\sigma^2}\tilde{S_n}^2 \sim \mathcal{X}^2_{n-1}$.
Posons q_1 quantile $\frac{\alpha}{2}$ de $\mathcal{X}^2_{-1}$ et q_2 quantile $\frac{\alpha-1}{2}t_{n-1}$ de $\mathcal{X}^2_{-1}$.

$$\begin{aligned}
& \mathbb{P}\left(\mathcal{X}^2_{n-1} \leq \tfrac{(n-1)}{\sigma^2}\tilde{S_n}^2 \leq \mathcal{X}^2_{n-1}\right) && = 1 - \alpha \\
\Leftrightarrow \quad & \mathbb{P}\left(\tfrac{\mathcal{X}^2_{n-1}}{(n-1)\tilde{S_n}^2} \leq \tfrac{1}{\sigma^2} \leq \tfrac{\mathcal{X}^2_{n-1}}{(n-1)\tilde{S_n}^2}\right) && = 1 - \alpha \\
\Leftrightarrow \quad & \mathbb{P}\left(\tfrac{(n-1)\tilde{S_n}^2}{\mathcal{X}^2_{n-1}} \leq \sigma^2 \leq \tfrac{(n-1)\tilde{S_n}^2}{\mathcal{X}^2_{n-1}}\right) && = 1 - \alpha \\
\Leftrightarrow \quad & \sigma^2 \in \left[\tfrac{(n-1)\tilde{S_n}^2}{\mathcal{X}^2_{n-1}}, \tfrac{(n-1)\tilde{S_n}^2}{\mathcal{X}^2_{n-1}}\right].
\end{aligned}$$

2.8 Modèle OMMeR dans le cas normal partiellement fixé

a) Moyenne μ_0 connue

Objectif : Estimation de σ^2.

Modèle : $(X_1, ..., X_n) \simeq N(\mu_0, \sigma^2)$ iid

Méthode :

$$\text{Estimateur : } \hat{\sigma}^2 = S_n^2(\mu) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2,$$

$$\text{Pivot : } \frac{nS_n^2(\mu)}{\sigma^2} \sim \mathcal{X}_n^2.$$

b) Variance σ_0^2 connue

Objectif : Estimation de μ.

Modèle : $(X_1, ..., X_n) \simeq N(\mu, \sigma_0^2)$ iid.

Méthode :

$$\hat{\mu} = \bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$Z = \frac{\bar{X_n} - \mu}{\sigma_0}\sqrt{n} \sim N(0,1).$$

Exemple 19. *Soient $\sigma_0^2 = 7^2$ connu et $X \sim N(\mu, 7^2)$. Nous disposons d'une observation de taille 60 de la taille d'individus, et nous avons comme moyenne $\bar{X}_60 = 173,4cm$. Nous cherchons un intervalle de confiance de niveau $0,95$ pour le paramètre μ. Comme la normale est stable par addition, nous avons que $\bar{X}_60 \sim N(\mu, \frac{49}{60})$ et $\frac{\bar{X}_60-\mu}{\sqrt{\frac{49}{60}}} \sim N(0,1)$.*

En partageant le risque des deux côtés, nous avons

$$\mathbb{P}\left(q_{0,025} \leq \frac{\bar{X}_{60} - \mu}{\sqrt{\frac{49}{60}}} \leq q_{0,975}\right) = 0,95$$

$$\Leftrightarrow \mathbb{P}\left(\bar{X}_{60} - \frac{7}{\sqrt{60}}q_{0,975} \leq \mu \leq \bar{X}_{60} - \frac{7}{\sqrt{60}}q_{0,025}\right) = 0,95,$$

où $q_{0,025}$ (resp. $q_{0,975}$) est le quantile de niveau $0,025$ (resp. $0,975$) de la normale centrée réduite. Remarquons que par symétrie de la normale centrée, nous avons $q_{0,025} = -q_{0,975} = -1,96$.

Donc notre intervalle de confiance de niveau $0,95$ est

$$\left[173,4 - \frac{7}{\sqrt{60}}1,96; 173,4 + \frac{7}{\sqrt{60}}1,96\right] = [171,63; 175,17].$$

Exemple 20. *1. Soit $X \sim N(\mu, \sigma^2)$. Nous disposons de 10 observations de moyenne $\bar{X}_{10} = 30,2$ et de variance $\tilde{S}_{10}^2 = 11,49$. Nous avons $\bar{X}_{10} \sim N(\mu, \frac{\sigma^2}{10})$. Nous cherchons un intervalle de confiance de niveau $0,95$ pour μ. Nous savons que $\frac{\bar{X}_{10}-\mu}{\tilde{S}_10}\sqrt{10} \sim t_{10-1}$.*

Donc en partageant le risque des deux côtés, nous obtenons

$$\mathbb{P}\left(q_{0,025} \leq \frac{\bar{X}_{10} - \mu}{\tilde{S}_{10}}\sqrt{10} \leq q_{0,975}\right) = 0,95,$$

où $q_{0,025}$ et $q_{0,975}$ sont les quantiles de la Student à 9 degrès de liberté de niveau $0,025$ et $0,975$. Par symétrie de la Student, nous avons que $q_{0,025} = -q0,975 = -2,2622$.

Donc

$$\mathbb{P}\left(\bar{X}_{10} + \frac{\tilde{S}_{10}}{\sqrt{10}} q_{0,025} \leq \mu \leq \bar{X}_{10} - \frac{\tilde{S}_{10}}{\sqrt{10}} q_{0,025}\right) = 0,95.$$

Donc notre intervalle de confiance pour μ *de niveau* $0,95$ *est*

$$\left[30,2 - \sqrt{\frac{11,49}{10}} 2,2622; 30,2 + \sqrt{\frac{11,49}{10}} 2,2622\right] = [27,78; 32,62].$$

2. *Nous cherchons un intervalle de confiance pour* σ^2. *Nous savons que* $(10-1)\frac{\tilde{S}_{10}^2}{\sigma^2} \sim \mathcal{X}_{10-1}^2$.
Donc

$$\mathbb{P}\left(q'_{0,025} \leq (10-1)\frac{\tilde{S}_{10}^2}{\sigma^2} \leq q'_{0,975}\right) = 0,95,$$

où $q'_{0,025} = 2,7$ *et* $q'_{0,975} = 19,023$ *sont les quantiles de la chi-carré à* 9 *degrès de liberté.*
Donc

$$\mathbb{P}\left(\frac{(10-1)\tilde{S}_{10}^2}{q'_{0,975}} \leq \sigma^2 \leq \frac{(n-1)\tilde{S}_{10}^2}{q'_{0,025}}\right) = 0,95.$$

Donc notre intervalle de confiance pour σ^2 *est*

$$\left[\frac{9 \cdot 11,49}{19,023}; \frac{9 \cdot 11,49}{2,7}\right] = [5,44; 38,30].$$

2.9 Estimation d'une loi $L(\mu, \sigma^2)$

Dans cette section la loi L est connue et les paramètres μ et σ^2 sont inconnus.

a) Modèle de position

Nous supposons que la variance σ_0^2 de la loi est connue. Soit X une variable aléatoire telle que $X \sim L(\mu, \sigma_0^2)$.

Objectif : Estimation de μ.

Modèle : $(X_1, ..., X_n) \sim L(\mu, \sigma_0^2)$ iid.

Méthode : Estimateur $\hat{\mu} = \bar{X_n}$

Pivot : $Z = \frac{\bar{X_n} - \mu}{\sigma_0}\sqrt{n} \xrightarrow{L} N(0,1)$.

Nous disons que (X_n) **converge en loi vers** X si les fonctions de répartitions F_n de X_n convergent vers celle de X, i.e. $\lim_{n \to +\infty} F_n(x) = F(x)$ et nous notons $X_n \xrightarrow{L} X$.

Résultat : Estimation, intervalle de confiance et calculs de probabilité asymptotiques, c'est-à-dire, si n est grand, alors nous pouvons approximer en utilisant les mêmes calculs que dans le cas gaussien.

b) Modèle d'échelle :

Objectif : Estimation de σ^2.

Modèle : $(X_1, ..., X_n) \simeq X \sim L(\mu_0, \sigma^2)$ i.i.d. tel que le moment centré d'ordre 4 $\mu_4 = E[(X - E[X])^4]$ existe.

Méthode : Estimateur $\hat{\sigma}^2 = S_n^2(\mu_0) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$

Pivot : $\frac{S_n^2(\mu_0) - \sigma^2}{\sqrt{Var(S_n^2(\mu))}} = \frac{S_n^2(\mu) - \sigma^2}{\sqrt{\mu_4 - \sigma^4}} \xrightarrow{L} N(0,1)$.

Résultat : Estimation, intervalle de confiance et calculs de probabilité asymptotiques.

c) Modèle de position d'échelle :

Objectif : Estimation de μ et de σ^2.

Modèle : $(X_1, ..., X_n) \simeq X \sim L(\mu_0, \sigma^2)$ i.i.d. tel que le moment centré d'ordre 4 μ_4 existe.

Méthode : Estimateur $\hat{\mu} = \bar{X_n}$ et $\hat{\sigma}^2 = \tilde{S_n}^2$

Pivot :

$$\frac{\bar{X_n} - \mu}{\tilde{S_n}}\sqrt{n} \xrightarrow{L} N(0,1)$$

$$\frac{\tilde{S_n}^2 - \sigma^2}{\sqrt{Var(\tilde{S_n})^2}} \xrightarrow{L} N(0,1)$$

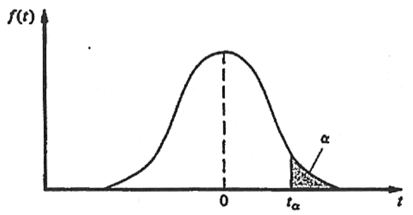
Résultat : Estimation, intervalle de confiance et calculs de probabilité asymptotiques.

Bibliographie

Auteur : Anton Thalmeier
Titre : Probabilités et Statistiques 1 et 2
Année : 2014-2015
Note : Université du Luxembourg, Bachelor in mathematics, 2*e* année
Auteur : Catherine Timmermans
Titre : Probabilité et Statistiques III
Année : 2015
Note : Université de Liège, Bachelier en sciences mathématiques, 5*e* semestre

Annexe

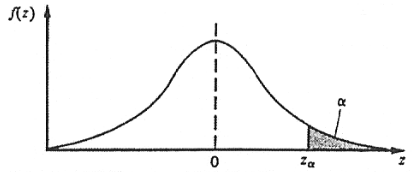
Table 8. *t*-distribution



Critical points (t_α) for different probability levels (α) and different number of degrees of freedom (ν). Example: For $\nu = 19$, $P(t > 2.0930) = 0.025$ and $P(|t| > 2.0930) = 0.05$.

ν \ α	0.4	0.25	0.15	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.3249	1.0000	1.9626	3.0777	6.3138	12.7062	31.8205	63.6567	318.3087	636.6189
2	0.2887	0.8165	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248	22.3271	31.5991
3	0.2767	0.7649	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409	10.2145	12.9240
4	0.2707	0.7407	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.2672	0.7267	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.2648	0.7176	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.2632	0.7111	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.2619	0.7064	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.2610	0.7027	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.2602	0.6998	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.2596	0.6974	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.2590	0.6955	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.2586	0.6938	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.2582	0.6924	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.2579	0.6912	1.0735	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.2576	0.6901	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.2573	0.6892	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.2571	0.6884	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.2569	0.6876	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.2567	0.6870	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	0.2566	0.6864	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	0.2564	0.6858	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	0.2563	0.6853	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.2562	0.6848	1.0593	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454
25	0.2561	0.6844	1.0584	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	0.2560	0.6840	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7066
27	0.2559	0.6837	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6896
28	0.2558	0.6834	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	0.2557	0.6830	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962	3.6594
30	0.2556	0.6828	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
35	0.2553	0.6816	1.0520	1.3062	1.6896	2.0301	2.4377	2.7238	3.3400	3.5911
40	0.2550	0.6807	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069	3.5510
45	0.2549	0.6800	1.0485	1.3006	1.6794	2.0141	2.4121	2.6896	3.2815	3.5203
50	0.2547	0.6794	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778	3.2614	3.4960
60	0.2545	0.6786	1.0455	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317	3.4602
70	0.2543	0.6780	1.0442	1.2938	1.6669	1.9944	2.3808	2.6479	3.2108	3.4350
80	0.2542	0.6776	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387	3.1953	3.4163
90	0.2541	0.6772	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316	3.1833	3.4019
100	0.2540	0.6770	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259	3.1737	3.3905
120	0.2539	0.6765	1.0409	1.2886	1.6577	1.9799	2.3578	2.6174	3.1595	3.3735
150	0.2538	0.6761	1.0400	1.2872	1.6551	1.9759	2.3515	2.6090	3.1455	3.3566
200	0.2537	0.6757	1.0391	1.2858	1.6525	1.9719	2.3451	2.6006	3.1315	3.3398
300	0.2536	0.6753	1.0382	1.2844	1.6499	1.9679	2.3388	2.5923	3.1176	3.3233
∞	0.2533	0.6745	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905

Table 5. Normal distribution (areas)

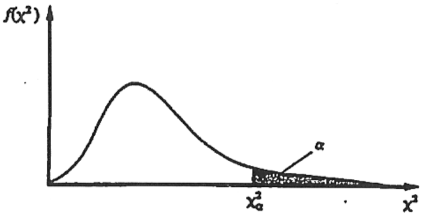


Area (α) in the tail of the standardised Normal curve, $N(0, 1)$, for different values of z. Example: Area beyond $z = 1.96$ (or below $z = -1.96$) is $\alpha = 0.02500$. For Normal curve with $\mu = 10$ and $\sigma = 2$, area beyond $x = 12$, say, is the same as area beyond $z = \frac{x - \mu}{\sigma} = \frac{12 - 10}{2} = 1$, i.e. $\alpha = 0.15866$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
1.0	.15866	.15625	.15386	.15150	.14917	.14686	.14457	.14231	.14007	.13786
1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07214	.07078	.06944	.06811
1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01254	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00509	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00403	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00263
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00085	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00009	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00004	.00003
4.0	.00003	.00003	.00003	.00003	.00003	.00002	.00002	.00002	.00002	.00002

α	0.4	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.001
z_α	.2533	.6745	.8416	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902

Table 7. χ^2 (Chi-squared)-distribution



Values of χ^2_α giving area (α) in the right-hand tail for different number of degrees of freedom (ν). **Example:** For $\nu = 15$ area beyond $\chi^2_{0.95} = 7.261$ is 0.950 and beyond $\chi^2_{0.10} = 22.307$ is 0.100.

ν \ α	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	0.0^43927*	0.0^31571*	0.0^39821*	0.0^23932*	0.01579	0.1015	0.4549	1.323	2.706	3.841	5.024	6.635	7.879
2	0.01003	0.02010	0.05065	0.1026	0.2107	0.5754	1.386	2.773	4.605	5.991	7.378	9.210	10.597
3	0.07172	0.1148	0.2158	0.3518	0.5844	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838
4	0.2070	0.2971	0.4844	0.7107	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860
5	0.4117	0.5543	0.8312	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.833	15.086	16.750
6	0.6757	0.8721	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548
7	0.9893	1.239	1.690	2.167	2.833	4.255	6.346	9.037	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	5.071	7.344	10.219	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	5.899	8.343	11.389	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	6.737	9.342	12.549	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	7.584	10.341	13.701	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	8.438	11.340	14.845	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	9.299	12.340	15.984	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	10.165	13.339	17.117	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	11.036	14.339	18.245	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	11.912	15.338	19.369	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	12.792	16.338	20.489	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	13.675	17.338	21.605	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	14.562	18.338	22.718	27.204	30.143	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	15.452	19.337	23.828	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	16.344	20.337	24.935	29.615	32.670	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	17.240	21.337	26.039	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.688	13.090	14.848	18.137	22.337	27.141	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	19.037	23.337	28.241	33.196	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	16.473	19.939	24.337	29.339	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	20.843	25.336	30.434	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	21.749	26.336	31.528	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	22.657	27.336	32.620	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	23.567	28.336	33.711	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	24.478	29.336	34.800	40.256	43.773	46.979	50.892	53.672
35	17.192	18.509	20.569	22.465	24.797	29.054	34.336	40.223	46.059	49.802	53.203	57.342	60.275
40	20.707	22.164	24.433	26.509	29.050	33.660	39.335	45.616	51.805	55.758	59.342	63.691	66.766
45	24.311	25.901	28.366	30.612	33.350	38.291	44.335	50.985	57.505	61.656	65.410	69.957	73.166
50	27.991	29.707	32.357	34.764	37.689	42.942	49.335	56.334	63.167	67.505	71.420	76.154	79.490
55	31.735	33.571	36.398	38.958	42.060	47.611	54.335	61.665	68.796	73.311	77.381	82.292	85.749
60	35.535	37.485	40.482	43.188	46.459	52.294	59.335	66.981	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	61.698	69.334	77.577	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	71.144	79.334	88.130	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	80.625	89.334	98.650	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	90.133	99.334	109.141	118.498	124.342	129.561	135.807	140.169
120	83.829	86.909	91.568	95.705	100.627	109.224	119.335	130.051	140.228	146.565	152.214	158.963	163.670
150	109.122	112.655	117.980	122.692	126.278	137.987	149.334	161.288	172.577	179.579	185.803	193.219	198.380
200	152.224	156.421	162.724	168.279	174.828	186.175	199.334	213.099	226.018	233.993	241.060	249.455	255.281
250	196.145	200.929	208.095	214.392	221.809	234.580	249.334	264.694	279.947	287.889	295.691	304.948	311.361

z_α	−2.5758	−2.3263	−1.9600	−1.6449	−1.2816	−0.6745	0.0000	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

* e.g. $0.0^43927 = 0.00003927$

Interpolation: For $\nu > 100$, $\chi^2_\alpha = \frac{1}{2}(z_\alpha + \sqrt{2\nu - 1})^2$ where z_α is the standardised Normal variable shown in the bottom line of the table.