Eléments de statistiques pour les data sciences

Cours 3 : probabilités continues

Dr. Ph. Müllhaupt

IGM - EPFL

Plan

densité de probabilité continue univariée

Soit X une variable aléatoire definie sur l'espace (Ω, F, P) avec la fonction de distribution cumulée

$$F(x) = P(X \leq x)$$

définition

S'il existe une fonction f telle que

$$F(x) = \int_{-\infty}^{x} f(y)\, dy$$

alors la fonction f est appelée la densité de probabilité (également appelée la distribution). Elle satisfait l'égalité

$$\int_{-\infty}^{+\infty} f(y)\, dy = 1$$

distribution uniforme

définition

La densité de probabilité uniforme (distribution uniforme) est définie par

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

Elle représente la possibilité qu'un nombre situé dans l'intervalle [a, b] puisse apparaître comme valeur de la variable aléatoire X avec autant de chance qu'un autre nombre de ce même intervalle.

densité de probabilité uniforme

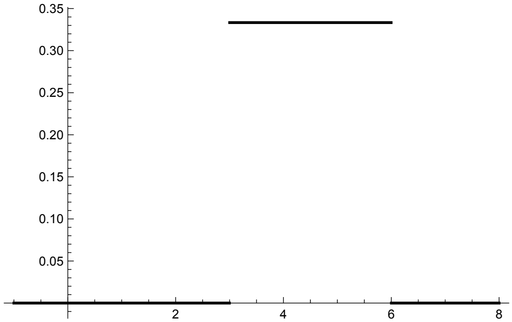


Figure – la distribution uniforme avec $a = 3$ et $b = 6$

la distribution normale (gaussienne,Laplace Gauss)

centrée à l'origine et de variance unité $\mathcal{N}(0,1)$

définition

La distribution normale de moyenne nulle et de variance unité notée $\mathcal{N}(0,1)$ est donnée par

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

Un moyen pour s'en souvenir est de prendre le logarithme naturel et on obtient la somme d'une constante et d'une forme quadratique négative standard $-\frac{1}{2}x^2$. Le facteur devant l'exponentiel est obtenu pour forcer $\int_{-\infty}^{+\infty} f(x)\,dx = 1$.

la distribution normale (gaussienne, Laplace-Gauss)

centrée à l'origine et de variance unité $\mathcal{N}(0,1)$

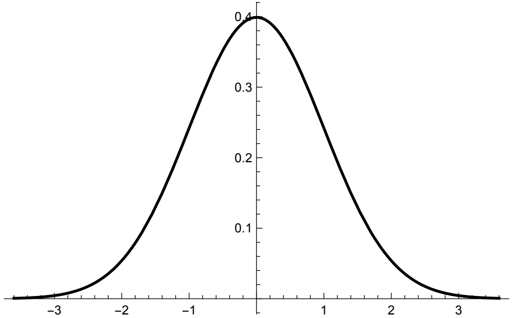


Figure – la distribution normale de moyenne nulle et de variance unité $\mathcal{N}(0,1)$

la distribution normale

non centrée et de variance quelconque

définition

La distribution normale de moyenne μ et de variance σ^2 $\mathcal{N}(\mu, \sigma^2)$ est donnée par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Le centrage est autour de la moyenne $\mu = \mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, f(x)\, dx$ et "l'épaisseur" est proportionnelle à l'écart type σ.

espérance mathématique

Soit une variable aléatoire continue X de densité de probabilité f et une fonction ψ de la varibale aléatoire X dans les nombre réels $\mathbb{R}$

$$\psi : X \to \mathbb{R}$$

L'espérance mathématique est définie par

$$\boxed{\mathbb{E}[\psi(X)] = \int_{-\infty}^{+\infty} \psi(x)\, f(x) dx}$$

moyenne

Soit une variable aléatoire continue X de densité de probabilité f. La moyenne μ est l'espérance mathématique de la fonction identité $\psi(X) = X$

définition

$$\boxed{\mu \triangleq \mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, f(x) dx}$$

variance

Soit une variable aléatoire continue X de densité de probabilité f. La variance σ^2 est l'espérance mathématique de la fonction $\psi(X) = (X-\mu)^2$

définition

$$\boxed{\sigma^2 \triangleq \mathbb{E}[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx}$$

Remarque : la définition $\mathbb{E}[(X-\mu)^2]$ est identique que X soit discrète ou X continue. Ce qui change est son calcul, par série lorsque X est discrète et par une intégrale lorsque X est continue.

la distribution normale

non centrée et de variance quelconque $\mathscr{N}(\sigma, \mu)$

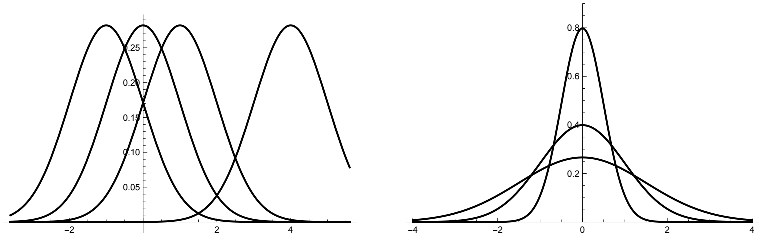


Figure – La distribution normale de moyenne nulle variable à gauche pour $\mu = -1, 0, 1, 4$ et de moyenne nulle mais de variance σ^2 variable, avec $\sigma = 0.5, 1, 1.5$

.

la distribution exponentielle

définition

La distribution exponentielle est définie par

$$f(x) = \begin{cases} \lambda\, e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Il est immédiat de vérifier que $\int_{-\infty}^{+\infty} \lambda e^{-\lambda x} = 1$. De plus, la valeur à l'origine donne le paramètre λ.

la distribution exponentielle

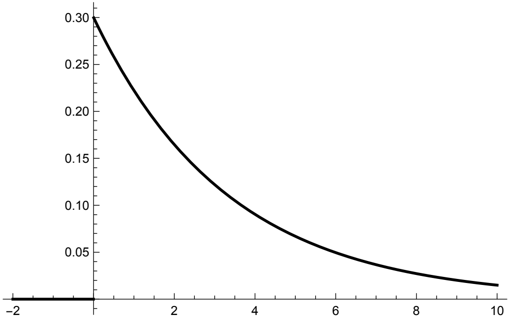


Figure – la distribution la distribution expoentielle de paramètre $\lambda = 0.3$.

la distribution d'Erlang

continue de paramètres $k \in \mathbb{N}$ et $\lambda \in \mathbb{R}, \lambda > 0$

définition

La distribution d'Erlang de paramètres $k \in \mathbb{N}$ et $\lambda \geq 0$ est donnée par

$$f(x; k, \lambda) \triangleq \frac{\lambda^k \, x^{k-1} \, e^{-\lambda x}}{(k-1)\,!} \qquad x, \lambda \geq 0$$

Remarques

intuitives, sans trop de rigueur...

1. En comparant les courbes, on constate que la distribution exponentielle est monotone décroissante.
2. Les distributions d'Erlang ont une tendance à croître initialement pour atteindre une valeur maximale et ensuite décroissent comme une exponentielle.
3. Il y a un 'combat' entre le facteur de croissance x^{k-1} la croissance monomiale et la décroissance exponentielle : pour de petite valeur de x le facteur x^{k-1}, croissance monomiale, l'emporte sur e^{-x}, la décroissance exponentielle. Pour de grande valeur de x, le facteur de décroissance exponentielle e^{-x} l'emporte sur la croissance monomiale x^{k-1}.

la fonction $\Gamma(\alpha)$

la fonction gamma généralise la factorielle

définition

La fonction $\Gamma(n)$ disrète est la fonction pour $n \in \mathbb{N}$

$$\Gamma(n) \triangleq (n-1)\,!$$

Pour une valeur $\alpha \in \mathbb{R}$, on a la définition

$$\Gamma(\alpha) \triangleq \int_0^\infty x^{\alpha-1} e^{-x} dx \qquad \alpha > 0$$

propriétés

Pour tout $\alpha > 0$, $\alpha \in \mathbb{R}$

1. $\int_0^\infty x^{\alpha-1}\, e^{-\lambda x}\, dx = \frac{\Gamma(\alpha)}{\lambda^\alpha} \qquad \lambda > 0$
2. $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$
3. $\Gamma(n) = (n-1)\,! \qquad n \in \mathbb{N}$
4. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

la distribution Γ (gamma)

C'est une autre façon de décrire un distribution du type de celle de Poisson où s'interpose l'effet de la croissance de la fonction d'un facteur x^α avec la décroissance de e^{-x}. L'exposant α évolue selon la fonction Γ.

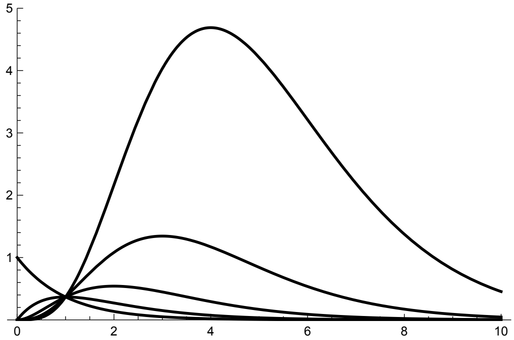


Figure – Graphiques de $x^{\alpha-1}\, e^{-x}$ pour $\alpha = 1, 2, 3, 4, 5$

la distribution gamma $\Gamma(\alpha, \lambda, x)$

définition

$\alpha > 0$, $\alpha \in \mathbb{R}$

$$\boxed{f(x) = \Gamma(\alpha, \lambda, x) = \begin{cases} \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x > 0 \\ 0 & x \leq 0 \end{cases}}$$

la distribution gamma $\Gamma(\alpha, \lambda, x)$

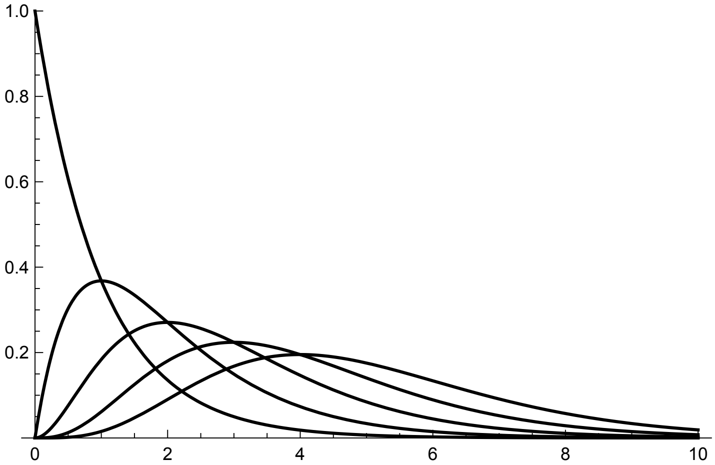


Figure – Représentation graphique de la densité de probabilité $\Gamma(\alpha, 1, x)$ pour les valeurs $\alpha = 1, 2, 3, 4, 5$. Remarque $\alpha = 1$ correspond à la distribution exponentielle.

cas particulier : la distribution du χ^2

$$X \sim \chi^2_\nu \Leftrightarrow X \sim \Gamma\left(\alpha = \frac{\nu}{2}, \lambda = \frac{1}{2}\right)$$

ν désigne les degrés de liberté (grosso modo le nombre de façons que possède le hasard de perturber de manière indépendante le résultat de l'expérience). On étudiera χ^2_ν à la leçon 4.

fonction d'une variable aléatoire

$Y = h(X)$

Soit X une variable aléatoire avec une densité de probabilité f et de fonction de répartition F. On aimerait déterminer la densité de probabilité de $Y = h(X)$ où h est une fonction réelle d'une variable réelle.

méthode recommendée

Il est commode de travailler avec la fonction de répartition

$$G(y) = P(Y \leq y)$$

à partir de la fonction de répartition $F(x) = P(X \leq x)$. Ensuite on utilise

$$g(y) = \frac{dG(y)}{dy}$$

exemple d'une fonction de variable aléatoire

$$Y = e^{-X}$$

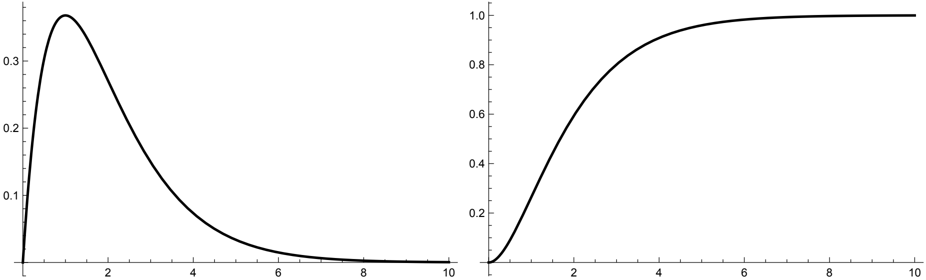


Figure – à gauche la densité de probabilité $f(x)$ et à droite la fonction de répartition $F(x)$

exemple d'une fonction de variable aléatoire

$Y = e^{-X}$

Soit X une variable aléatoire continue de densité de probabilité

$$f(x) = \begin{cases} x\,e^{-x} & x > 0 \\ 0 & \end{cases}$$

la fonction de répartion correspondante est

$$F(x) = \big(1 - (x+1)e^{-x}\big)$$

Etudions

$$Y = e^{-X}$$

exemple d'une fonction de variable aléatoire $Y = e^{-X}$

Soit G la fonction de répartition de Y.

$$\begin{aligned} G(y) &= P(Y \leq y) = P(e^{-X} \leq y) = P(-X \leq \ln y) \\ &= P(X \geq -\ln y) = 1 - P(X < -\ln y) \\ &= 1 - F(-\ln y) \end{aligned}$$

en substituant $x = -\ln y$ cela donne

$$G(y) = y(1 - \ln y)$$

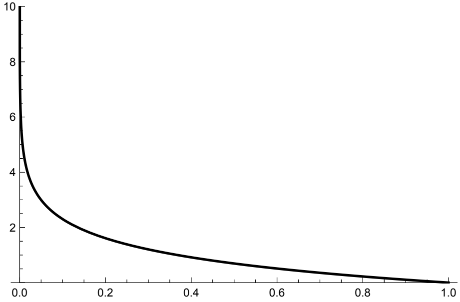
exemple d'une fonction de variable aléatoire $Y = e^{-X}$

On obtient g par dérivation

$$g(y) = \frac{dG}{dy} = \begin{cases} -\ln y & 0 < y < 1 \\ 0 & \text{sinon} \end{cases}$$

exemple d'une fonction de variable aléatoire

$Y = e^{-X}$



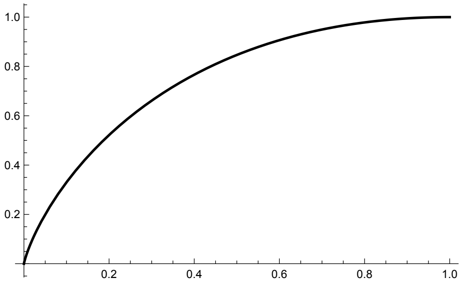


Figure – à gauche la densité de probabilité $g(y)$ et à droite la fonction de répartition $F(y)$

fonction monotone d'une variable aléatoire

croissante ou décroissante

Le résultat est général quel que soit le fonction monotone de $h(x)$ avec $Y = h(X)$.

théorème

Soit X une variable aléatoire continue de densité de probabilité $f(x)$ et Y une variable aléatoire continue de densité de probabilité $g(y)$. Soit également une fonction $Y = h(X)$ monotone dérivable (et donc continue) qui relie les deux variables aléatoires.On a le résultat

$$g(y) = f(x) \cdot \left| \frac{dx}{dy} \right|$$

$$x = h(y) \qquad y = h^{-1}(x) \qquad dx = \frac{dh}{dy} dy$$

densité de répartition jointe

Soit un couple de variables aléatoires X et Y.
A tout évènement exprimé par X et Y correspond une région A du plan. On a

$$P(X, Y) = P(A)$$

hypothèse

L'hypothèse est que $P(A)$ puisse se représenter par une intégrale double

$$P(A) \triangleq \int\int h(x, y)\, dx\, dy$$

alors la fonction $h(x, y)$ est appelée la densité de probabilité jointe.

fonction de répartition jointe

Soit l'évènement

$$\{\omega \in \Omega | X(\omega) \leq x\} \cap \{\omega \in \Omega | Y(\omega) \leq y\}$$

que l'on peut écrire avec un "abus de notation"

$$(X \leq x) \cap (Y \leq y)$$

définition

S'il existe une fonction $H(x, y)$ telle que

$$P(X \leq x) \cap P(Y \leq y) \triangleq H(x, y)$$

alors $H(x, y)$ est appelée la fonction de répartition jointe.

densité marginale

définition

Soit $h(x, y)$ la fonction de densité jointe des deux variables aléatoires X et Y. Pour une valeur de x donnée, on peut définir la fonction de densité de probabilité marginale

$$f_X(x) \triangleq \int_{-\infty}^{+\infty} h(x, y) dy$$

de même la fonction de densité marginale selon y

$$f_Y(y) \triangleq \int_{-\infty}^{+\infty} h(x, y) dx$$

somme de variables aléatoires

Soit la somme de deux variables aléatoires continues réelles

$$Z = X + Y$$

et une fonction $\phi(x, y)$ quelconque, on a

définition

$$\mathbb{E}[\phi(x, y)] \triangleq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi(x, y)\, h(x, y)\, dx\, dy$$

somme de variables aléatoires

en particulier pour $\phi(x, y) = x + y$

$$
\begin{aligned}
\mathbb{E}[Z] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\,h(x,y)\,dxdy \\
&= \int_{-\infty}^{+\infty} x\left(\int_{-\infty}^{+\infty} h(x,y)\,dy\right)\,dy + \int_{-\infty}^{+\infty} y\left(\int_{-\infty}^{+\infty} h(x,y)\,dx\right)\,dx \\
&= \int_{-\infty}^{+\infty} x\,f_X(x)\,dx + \int_{-\infty}^{+\infty} y\,f_Y(y)\,dy \\
&= \mathbb{E}(X) + \mathbb{E}(Y)
\end{aligned}
$$

corrélation

variables aléatoires X et Y continues

$$\boxed{\sigma_{XY} \triangleq \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_x)(y-\mu_y)\,h(x,y)\,dx\,dy}$$

corrélation

variables aléatoires X et Y discrètes

$$\boxed{\sigma_{XY} \triangleq \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty}(k-\mu_X)(l-\mu_Y)\,p(k,l)}$$

variable aléatoire discrète complexe

En introduisant une variable aléatoire discrète complexe

$$Z = X + j\,Y$$

avec $j = \sqrt{-1}$ la variable imaginaire et X et Y des variables aléatoires discrètes à valeurs dans les entiers naturels $\mathbb{N}$ on peut définir

$$\mathbb{E}[Z] \triangleq \mathbb{E}[X] + j\,\mathbb{E}[Y]$$

si, et seulement si, $\mathbb{E}[X] \in \mathbb{R}$ et $\mathbb{E}[Y] \in \mathbb{R}$ sont définis.

fonction génératrice

En introduisant le paramètre $s \in \mathbb{C}$, $|s| < 1$, on a comme cas particulier de variable aléatoire discrète complexe

$$Z = s^X$$

définition

Soit une seule variable aléatoire discrète X qui donne des nombres naturels $X \in \mathbb{N}$. La fonction génératrice est la fonction de $s \in \mathbb{C}$ et $|s| \leq 1$ qui associe à ce nombre (le paramètre $s \in \mathbb{C}$)) l'espérance mathématique de la variable aléatoire discrète complexe $Z = s^X$

$$\boxed{G_X(s) \triangleq \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k p_k \qquad \forall s \in \mathbb{C}, |s| \leq 1}$$

A l'intérieur du cercle unité, la série converge uniformément et de manière absolue pour s dans le cercle unité car $\sum_{k=1}^{\infty} p_k |s^k| \leq \sum_{k=1}^{\infty} p_k = 1.$

produit de convolution

entre distributions discrètes

définition

$$\boxed{h(k) \triangleq f * g \triangleq \sum_{l=0}^{k} f(l)g(k-l)}$$

produit de convolution

entre distributions discrètes

théorème

Soit $Z = X + Y$ la somme de deux variables aléatoires discrètes avec

- $f(k)$ la densité de probabilité discrète de X
- $g(k)$ la densité de probabilité discrète de Y
- h(k) la densité de probabilité discrète de Z

on alors les deux résultats suivants

$$\boxed{h = f * g}$$

$$\boxed{G_Z(s) = G_{X+Y}(s) = G_X(s)\, G_Y(s)}$$

produit de convolution

entre densité de probabilité (distributions continues)

Variables aléatoires X et Y et soit Z la somme des variables aléatoires

$$Z = X + Y$$

théorème de convolution

$$\boxed{f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)\, f_Y(z-x)\, dx = \int_{-\infty}^{+\infty} f_Y(y)\, f_X(z-y)\, dy}$$

On note de manière compact cette formule par le produit de convolution

$$\boxed{f_Z = f_X * f_Y}$$

fonction caractéristique d'une variable aléatoire

La définition de l'espérance mathématique peut-être étendue au cas où la variable X est une variable aléatoire complexe

$$X = X_1 + j\,X_2$$

où X_1 et X_2 sont des variables aléatoires rélles. On a

$$\mathbb{E}[X] = \mathbb{E}[X_1] + j\mathbb{E}[X_2]$$

définition

La fonction caractéristique d'une variable aléatoire X avec une densité de probabilité f_X est définie par

$$\boxed{\phi_X(u) \triangleq \mathbb{E}[e^{jux}] = \int_{-\infty}^{+\infty} e^{jux} f_X(x)dx}$$

C'est la transformée de Fourier de la densité de probabilité.

théroème des sommes de variables aléatoires

$$X = X_1 + X_2$$

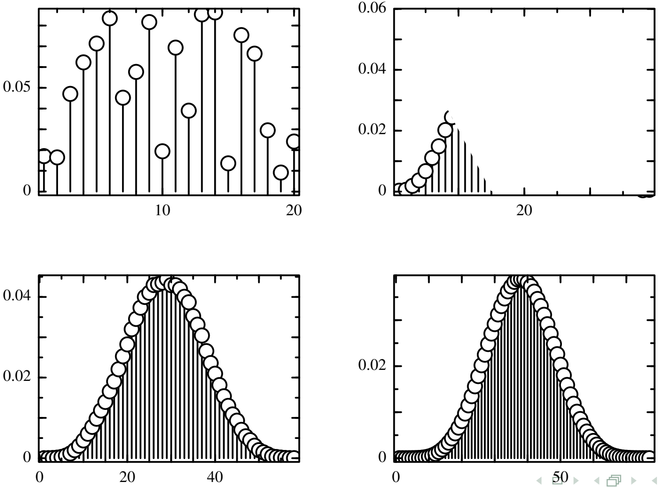
$$f_X(x) = \int_{-\infty}^{\tau} f_{X_1}(\tau) f_{X_2}(x-\tau) d\tau$$

théorème

$$\boxed{\Phi_X(u) = \Phi_{X_1}(u)\,\Phi_{X_2}(u)}$$

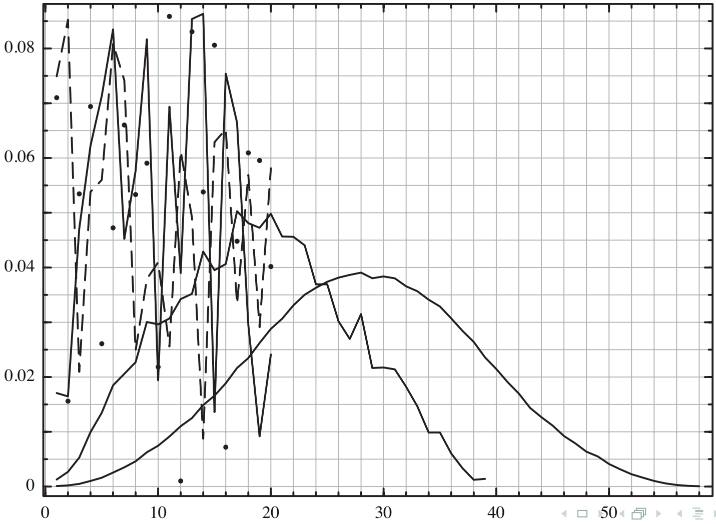
convolution

illustration



convolution

illustration



approximation de la distribution discrète binomiale

Théorème de De Moivre - Laplace

Si la probabilité de n expériences de Bernouilli p est constante, la probabilité d'avoir k succès parmi n durant ces expériences tend vers une distribution normale, plus précisément

$$\boxed{\sqrt{npq}\,\mathscr{B}in(n,p) : \frac{1}{2\pi} e^{-\frac{1}{2}x^2} \to 1}$$

lorsque $n \to \infty$ et de manière uniforme pour tout k pour lequel la variable x est contenue dans un certain intervalle fini. Ici la correspondance

$$x = \frac{k - np}{\sqrt{nkq}}$$

Rappel : $q = 1 - p$

approximation de la distribution discrète binomiale $\mathscr{B}in(n,p)$

Soit la distribution binomiale $\mathscr{B}in(n,p)$ donnée par

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n$$

On a l'approximation lorsque n est grand par

$$\boxed{P_n(k) \sim \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{np(1-p)}}\right)^2}}$$

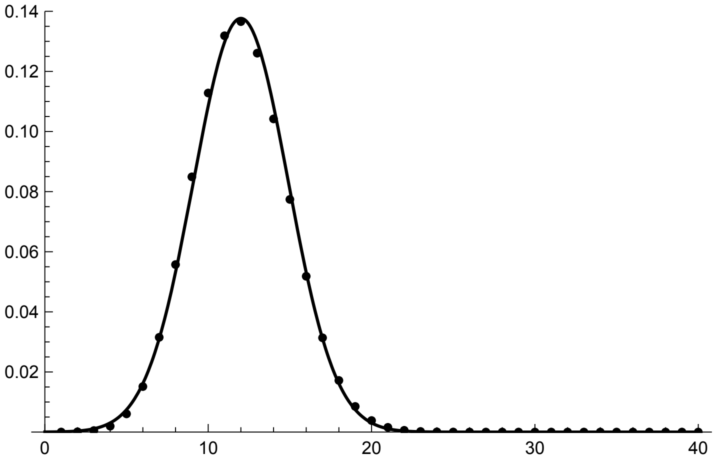


Figure – Illustration du théorème de Laplace–de Moivre pour $n = 40$ et $p = 0.3$.

Mathematica :

```
fct = 1/Sqrt[ 2 Pi n p (1 - p)] Exp[-1/ 2 ((k - n p)
                  /Sqrt[n p (1 - p)])^2] /. {n -> 40, p -> 0.3};
pp1 = Plot[fct, {k, 0, 40}, PlotRange -> {0, 0.14},
           PlotStyle -> Black];
pp2 = PDF[BinomialDistribution[40, 0.3], #] & /@
           Table[k, {k, 1, 40}] //
           ListPlot[#, PlotStyle -> Black] &;
Show[pp1,pp2]
```