Chapitre 2
Variables aléatoires

1. Définition. Variables qualitatives et quantitatives.

Une *variable aléatoire* est définie par référence à une expérience aléatoire comme une application dont la valeur dépend du résultat ω de cette expérience. Elle peut être *discrète* (lôensemble des valeurs possibles est discret, fini ou dénombrable), *continue* (ensemble non dénombrable), *mixte* ou *hybrides* (possédant des composantes discrète et continue).

Les variables discrètes peuvent être qualitatives ou quantitatives. Dans la majorité des situations pratiques pouvant être modélisées à lôaide dôexpériences aléatoires , on ne sôintéresse pas à proprement parler à lôespace des évènements $\Omega = \{\omega\}$ lui-même, ou aux évènements qui lui sont liés. Ces évènements peuvent être parfois complexes à décrire. Il est souvent plus simple de sôintéresser à certaines quantités variables liées à lôexpérience aléatoire considérée, et qui suffisent à résoudre les problèmes auxquels on sôintéresse. En fait, les variables sont du point de vue de lôinterface avec lôutilisateur le plus souvent qualitatives, mais on peut toujours ramener son étude à celle dôune variable quantitative. Ce passage du qualitatif au quantitatif est aussi appeler *codage*. Imaginons la situation suivante :

Exemple 1: Lors de lôexpérience qui consiste à lancer une pièce de monnaie, au lieu de considérer $\Omega = \{F, P\}$, on peut sôintéresser à la variable X qui prend la valeur 0 si le résultat ω est Pile , et $X = 1$ sinon. Une variable qui peut prendre deux valeurs possibles (on peut toujours se ramener aux valeurs 0 ou 1) sôappelle ***variable binaire*** ou de ***Bernoulli***. On peut ici, prendre $\Omega = \{0,1\}$.

Exemple 2: Si on observe lôétat de votre boîte e-mail, on peut considérer une variable qualitative à deux modalités (ou éventualités) : affectée par un virus ou non. On peut lui associer une variable de Bernoulli : $X = 0$ ou 1. Côest encore un codage binaire.

Exemple 3 : Une image peut être qualifiée de bonne , mauvaise ou moyenne. Une première approche serait de quantifier lôimage par une variable pouvant prendre trois valeurs possibles $0,1,2$. Dans les applications réelles de Vision et traitement dôimages, une image $2D$ (deux dimensions) peut être décrite par une variable multidimensionnelle indiquant la luminosité en un certain point du plan (voir chapitre 4).

Exemple 4 : Le nombre dôadmis à lôépreuve dôAlgorithmique est une variable aléatoire entière pouvant prendre les valeurs dans lôensemble $\{0,1,2,...,N\}$, où N est le nombre dôétudiants sôétant présentés à cette épreuve.

Exemple 5 : Le nombre de buts marqués au cours de la rencontre JSK-MCA.

Lôensemble des valeurs possibles est encore {0,1,2,..}. Revenons un peu sur la notion de hasard. Avant la fin du match, on ne peut prévoir le score, et encore moins le nombre de buts marqués. Une autre idéalisation des conditions de lôexpérience est cette manière de ne pas borner lôespace des épreuves possibles, bien que lôon sache pertinemment quôil est « impossible » dôavoir en un temps fini (90 minutes) un nombre infini de buts. Cependant, toute modélisation et description cohérente avec les hypothèses donnera moins de poids au grand nombre de buts. Une description cohérente, indiquera certainement que la probabilité dôavoir plus de cinq buts est faible, les probabilités dôavoir 0 , 1 ou 2 buts seront intuitivement plus grandes.

Exemple 6 : Le nombre de séismes dans une région donnée au cours du mois.

Exemple 7: La durée de la connexion à un site donné.

Exemple 8: Le nombre de paquets arrivant à un routeur dôun réseau mobile.

Dans la suite, les variables considérées sont le plus souvent quantitatives.

2. Variables aléatoires discrètes.

La variable aléatoire X est appelée *variable aléatoire discrète* (v.a.d.) si lôensemble de ses valeurs possibles est discret (fini ou dénombrable). $E = \{ x_1, x_2, ..., x_N \}$, $(N \leq \infty)$. On peut toujours numéroter lôensemble des états et considérer que la variable est définie sur $E = IN = \{0,1,2,...\}$ lôensemble des entiers naturels (ou une de ses parties $E = \{1,2,...,N\}$; $E \subseteq IN$). La suite de nombres

$$p_1 = P(X_1 = x_1), \ p_2 = P(X_2 = x_2), ..., p_i = P(X_i = x_i), ...$$

vérifiant $\sum_{i=1}^{N} p_i = 1$ est appelée *fonction de densité* de la variable X .

Exemple 9: Supposons que X représente le numéro de la face qui sort lors du jet d'un dé "parfait": $E = \{1,2,3,4,5,6\}$. La densité dôune variable aléatoire peut être représentée par un tableau des probabilités pour pouvoir faciliter son stockage dans la mémoire dûun ordinateur :

x_i	1	2	3	4	5	6
$p_i = P(X = x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Exemple 9bis (suite): Supposons maintenant que le dé est pipé de telle sorte que la probabilité quûne face sorte soit proportionnelle au nombre de points. La fonction de densité devient la suivante

x_i	1	2	3	4	5	6
$p_i = P(X = x_i)$	1/21	2/21	3/21	4/21	5/21	6/21

Exemple 10: Soit la variable discrète de fonction de densité:

x_i	2	3	4
$P(X = x_i)$	0.3	0.2	0.5

Cette fonction peut être représentée sous forme du graphe (diagramme en bâtons) suivant

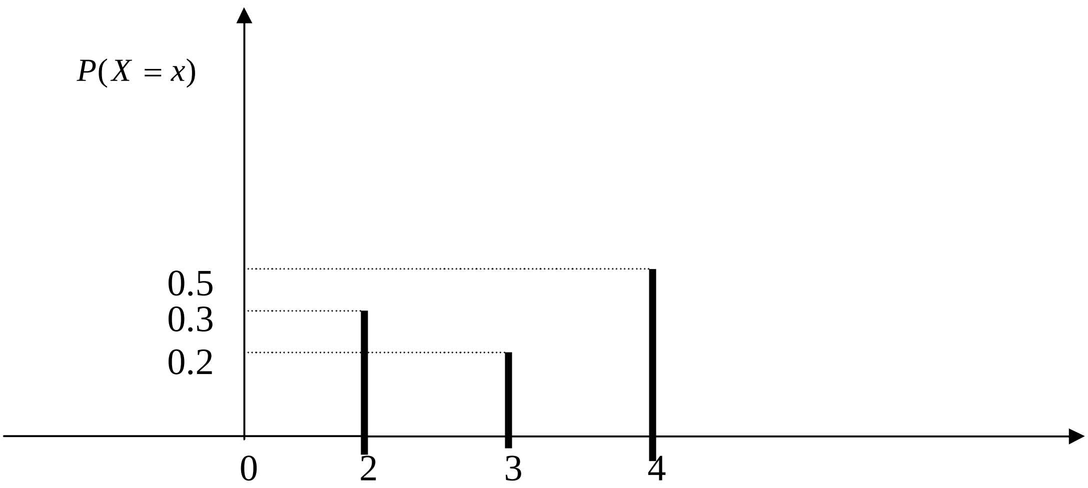


Figure 1. Fonction de densité discrète.

La fonction des probabilités cumulées
:

$$F(x) = P(X < x) = \begin{cases} 0 & si \quad x \leq 2 \\ 0.3 & si \quad 2 < x \leq 3 \\ 0.5 & si \quad 3 < x \leq 4 \\ 1 & si \quad x > 4 \end{cases}$$

sôappelle fonction de répartition de la variable discrète. Nous y reviendrons plus loin.

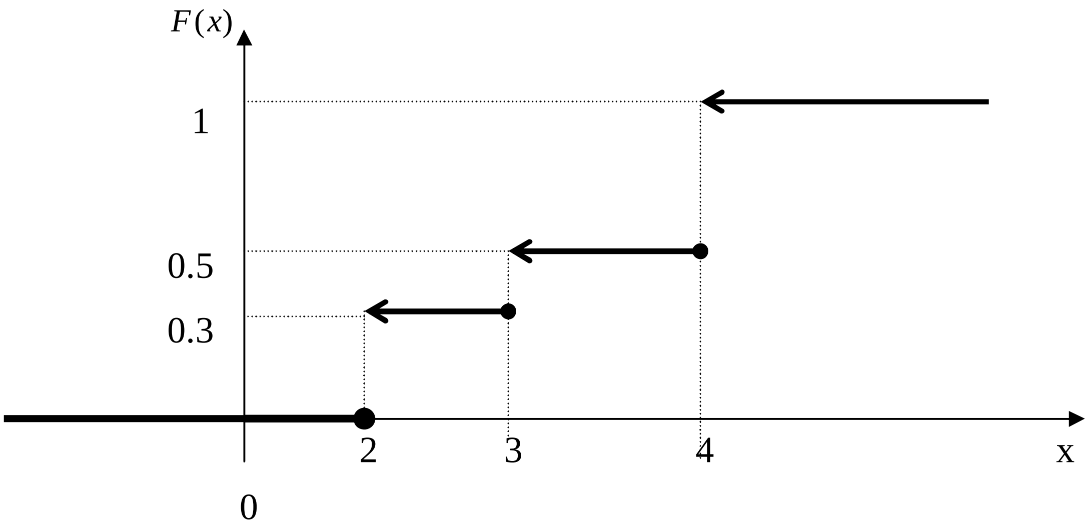
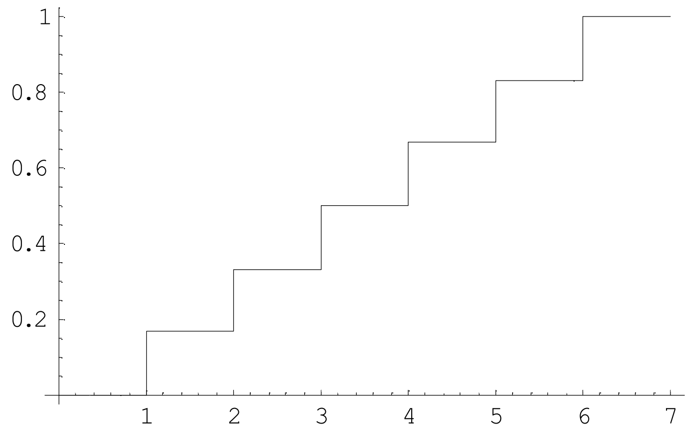


Figure 2. Fonction de répartition de la variable discrète.

Exemple 9 (suite) : Les fonctions de densité et de répartition correspondant à la densité de lôexemple 9 est la suivante

x_i	1	2	3	4	5	6
$p_i = P(X = x_i)$	1/6	2/6	3/6	4/6	5/6	6/6

x



4. Propriétés des fonctions de densité et de répartition.

La fonction de répartition $F(x)$, elle ñmesureò lœévènement $\{X < x\}\{\omega : X(\omega) < x\}$={ensemble des valeurs de la variable X inférieures à la quantité x}. Notons enfin que la fonction de répartition caractérise la loi de la variable aléatoire. En particulier, deux variables aléatoires ayant même loi ont même fonction de répartition et inversement. Cependant, le fait que deux variables aient la même loi ne signifie pas que cœst la même variable, du point de vue de lœinterprétation physique par exemple.

Ainsi, $F(x) = P\{X < x\}$ peut être définie par $F(x) = \sum_{i<x} p_i$

Notons que dans certains ouvrages, la fonction de répartition est définie par $F(x) = P\{X \leq x\}$. Pour les variables discrètes par contre $P\{X \leq x\} = P(X < x) + P(X = x)$, le second terme pouvant être non nul.

(i) F est non décroissante (croissante au sens large) i.e. si $a < b$, alors $F(a) \leq F(b)$.

(ii) $F(+\infty) = \lim_{x\to\infty} F(x) = 1$

(iii) $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$

(iv) F est continue à gauche ;

(v) Le nombre de points de discontinuité de F est au plus dénombrable.

(vi) $F(b) - F(a) = P(a \leq X < b)$ pour tous $a < b$.

(vii) Les propriétés (i) et (ii) fournissent une interprétation intéressante de la densité. En effet, la densité est une « mesure » locale, alors que la fonction de distribution est une mesure « spatiale ».

Lois de probabilités discrètes usuelles.

Loi de Bernoulli (ou loi 0-1). Cœst la loi dœune variable aléatoire binaire X à valeurs dans $E=\{0,1\}$. On note $X \in B(p)$. Une telle variable ne prend deux valeurs 0 ou 1 avec les probabilités: $P(X=1)=p$, $P(X=0)=1-p$, $0<p<1$. On peut écrire la densité sous la forme: $P(X=i)=p^i(1-p)^{n-i}$, $i=0,1$.Ses caractéristiques moyennes valent: $m=E(X)=p;\ Var(X)=p(1-p)$.

Interprétation : La nature binaire de la variable permet dœenglober une variété de modèles où seules deux éventualités possibles sont prises en compte. Considérons par exemple, un évènement A lié à une expérience aléatoire telle que sa probabilité dœoccurrence soit égale à $p=P(A)$ On peut introduire la variable

$$X=\begin{cases}1 & \text{si A s'est réalisé dans l'expérience (succés)}\\ 0 & \text{si A ne s'est pas réalisé (échec)}\end{cases}$$

La variable associée à une loi de Bernoulli est assimilée à la fonction indicatrice dœun événement (ou un ensemble). Elle est ainsi très pratique pour le codage binaire, en particulier pour la communication avec les ordinateurs. Par exemple:

(i) Au cours du jet dœune pièce de monnaie, on peut considérer lœévènement A=$\{$occurrence de Pile$\}$.

(ii) Considérons la transmission de chiffres binaires (bits) à travers un canal de communication ; A=$\{$le chiffre est transmis correctement$\}$.

Loi binomiale $B(n,p)$.

On dit que X suit une *loi binomiale* de paramètre $n\in\{1,2,...\}$ et $0<p<1$, si son ensemble des valeurs possibles est $E=\{0,1,2,...,n\}$ et si sa densité de probabilité est donnée par

$$p_i=P(X=i)=C_n^i p^i(1-p)^{n-i},\ i=0,1,2,...,n$$

Interprétation : Considérons une série de n expériences de Bernoulli indépendantes, dans des conditions identiques (par exemple n jets d'une même pièce de monnaie), et dans chacune desquelles on peut observer la réalisation (ou la non réalisation) d'un certain événement A (le résultat du jet est Pile). Pour chaque expérience (la i ème disons), on introduit la variable aléatoire de Bernoulli $X_i = 1$ si l'événement A se réalise dans la ième expérience, $X_i = 0$ si A ne se réalise pas dans cette ième expérience $(i = 1,2,...,n)$. La variable représente le nombre d'occurrences de l'événement A (succès) au cours de ces n expériences. Un exemple de modèle simple est lorsque X = nombre de boules blanches tirées d'une urne (avec remise) qui en contient $n = a + b$. Un autre exemple décrit est la loi du nombre de défectueux dans un lot d'articles.

Exemple 1: On sait que la proportion de lampes défectueuses est de 20%. Quelle est la probabilité que parmi 10 lampes, 2 soient défectueuse. Cette probabilité vaut: $p_2 = C_{10}^2 (0.2)^2 (0.8)^8 \approx 0.30199$.

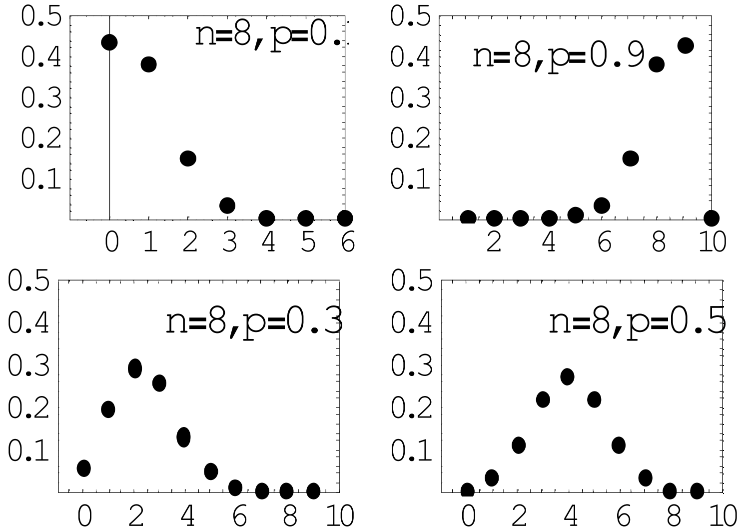
Exemple 2: Les graphes de 4 distributions binomiales de paramètres différents sont représentés sur les figures ci-dessous :

(i) $n = 8, p = 0.1$ $m = 0.8$, graphe décroissant ($p = 0.1 < 8/9$;

(ii) $n = 8, p = 0.9$ $m = 7.2$, graphe croissant ($p = 0.9 > 8/9$;

(ii) $n = 8, p = 0.3$ $m = 2.4$, mode=2 ;

(iv) $n = 8, p = 0.5$ $m = 4 = \text{mod}e$.



Loi géométrique ou de Pascal G(p). On dit que X suit une loi géométrique de paramètre p partant de 1 $(0 < p < 1)$, si $E = IN^* = \{1,2,é\ \}$ et $p_i = P(X = i) = (1-p)^{i-1}p,\ i \in E$. On peut vérifier que $E(X) = 1/p$; $Var(X) = (1-p)/p^2$.

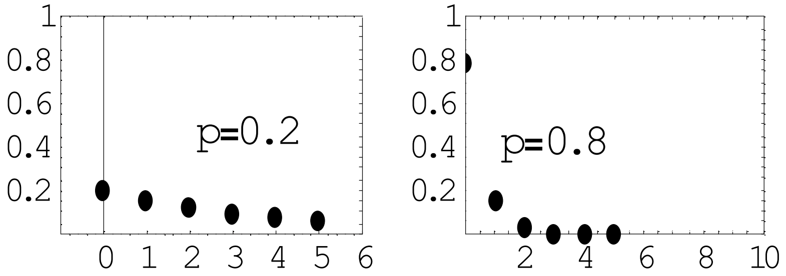


Figure 1. Densité de la loi géométrique de paramètre p=0.2 et 0.8.

Interprétation : X peut-être interprétée comme le nombre dôépreuves de Bernoulli pour obtenir un succès pour la première fois. Par exemple, le nombre de jets d'une pièce de monnaie pour obtenir "Pile" pour la première fois. La quantité p_i représente la probabilité dôobtenir un premier succès à la i-ème tentative (y compris la dernière tentative qui sôachève par un succès).
Notons que la loi géométrique satisfait la propriété dôabsence de mémoire côest-à-dire

$$X \in Géo(p) \Leftrightarrow P(X \geq i+j) = P(X \geq i).P(X \geq j),\ \forall i, j \in IN$$

Côest la seule loi discrète possédant cette propriété.

Exemple 4: On transmet un message en code binaire par un canal de communication. Le message est donc une suite de « 0 » et de « 1 » qui se succèdent avec des probabilités identiques et indépendantes les unes des autres, par exemple 001110111000é .Considérons un bloc quelconque de chiffres qui se répètent, disons, 0000 ou 111 (le bloc peut contenir un seul chiffre 0 ou 1). On choisit un bloc au hasard et on définit la variable X égale au nombre de chiffres dans le bloc. Trouver la probabilité pour que dans un bloc choisi au hasard le nombre de chiffres quôil contient soit supérieur à une valeur donnée i .

La variable X=nombre de chiffres dans le bloc suit une loi géométrique de paramètre ½ : $P(X = i) = \left(\frac{1}{2}\right)^i, i = 1,2,...$ Donc, $E(X) = \frac{1}{0.5} = 2$; $Var(X) = \frac{0.5}{0.5^2} = 2$

$$P(X \geq i) = \sum_{k=i}^{\infty} 0.5^k = (0.5)^{k-1}.$$

Remarque: Parfois, on considère plutôt une modification du modèle précédent représenté par la variable $Y = X - 1$ qui représente le nombre d'échecs avant d'obtenir un succès; dans ce cas, $p_i = (1-p)^i p, i = 0,1,2,\ldots$ $E(X) = (1-p)/p$; ; $Var(X) = (1-p)/p^2$. On lôappellera loi géométrique partant de 0.

Loi de Poisson $\mathcal{P}(\lambda)$.

Une variable X suit une loi de Poisson de paramètre $\lambda > 0$, si $E = IN$ et

$$p_i = \frac{\lambda^i}{i!} e^{-\lambda}, \qquad i = 0,1,2,\ldots$$

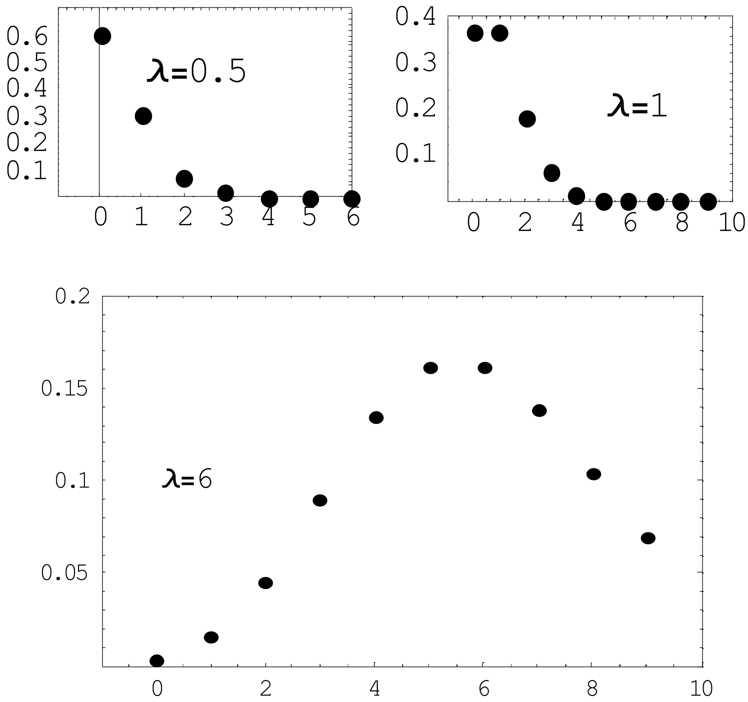


Figure 2. Densité de la loi de Poisson de paramètres $\lambda = 0.5, 1 \, et \, 6$.

Soit Y_n le nombre de succès dans n expériences de Bernoulli où **P**(succès)= p_n (ici $p = p_n$ dépend du nombre dôexpérience). Cette variable suit donc une loi binomiale

B(n,pn) : $P(Y_n = i) = P_n(i, p_n) = C_n^i p_n^i (1-p_n)^{n-i}$. Supposons que $p = p_n = \lambda_n / n$ avec $\lambda_n \to \lambda > 0$ lorsque $n \to \infty$, alors

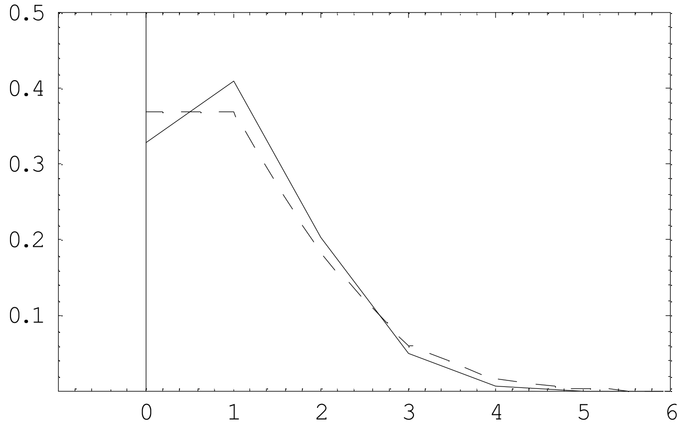
$$\lim_{n\to\infty} P_n(i, p_n) = \frac{\lambda^i}{i!} e^{-\lambda} , \ i = 0,1,2,...$$

Ce résultat (connu sous le nom de théorème de Poisson) stipule que le modèle de loi binomial $B(n, p)$ peut être approché par un modèle de loi de Poisson lorsque le nombre dœxpériences n est suffisamment grand, et que la probabilité de succès p est suffisamment petite, le produit np restant approximativement constant égal à λ . Cœst pourquoi, la loi de Poisson sœppelle parfois « *loi des évènements rares* ». En pratique, lœpproximation est satisfaisante dès que $n > 200$, $p \le 0.1$(ce nœst pas une règle !!!, on trouvera dœutre indications sur la manière de choisir n et p dans la littérature selon la nature de lœpplication).

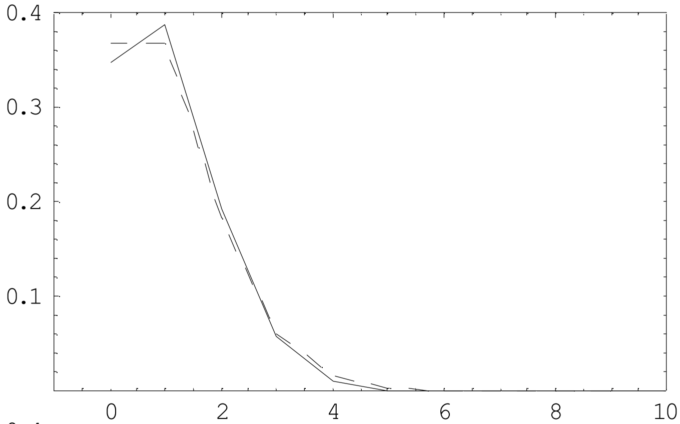
Pour illustrer lœrreur dœpproximation, le lecteur peut sœnspirer des petites simulations ci-dessous

i	$n = 5$ $p = 0.2$	$n = 10$ $p = 0.1$	$n = 20$ $p = 0.01$	$\lambda = 1$
0	0.32768	0.348678	0.358486	0.367889
1	0.4096	0.38742	0.377354	0.367879
2	0.2048	0.19371	0.188677	0.18394
3	0.0512	0.0573956	0.0595821	0.00613132
4	0.0064	0.0111603	0.0133276	0.0153283
5	0.00032	0.00148803	0.00224465	0.00306566
6		0.000137781	0.000295348	0.000510944
7		$8.748.10^{-6}$	0.0000310893	0.000072992
8		$3.645.10^{-7}$	0.00000	0.00000
9		9.10^{-9}	0.00000	0.00000
10		1.10^{-10}		0.00000

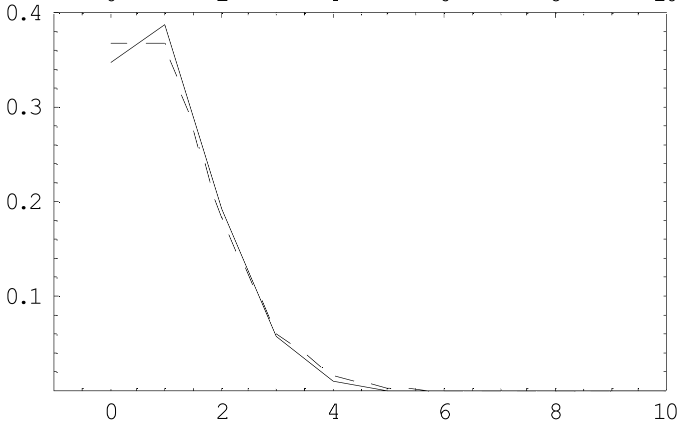
La courbe de la loi de Poisson est représentée en pointillés.



n=5
p=0.2



n=10
p=0.1



n=20
p=0.01

Sous certaines hypothèses de nature mathématique, la loi de Poisson est utilisée pour décrire le nombre dôapparitions dôun certain événement dans un intervalle de temps fixé ou dans un domaine spatial fixé. Ainsi, moyennant une traduction des hypothèses mathématiques vers les hypothèses physiques correspondantes, la loi de Poisson a été utilisée pour modéliser par exemple :

- la loi du nombre dôappels à un central téléphonique,
- la loi du nombre de visites à un site Web durant une période donnée,
- la loi du nombre de particules radioactives émise dans une direction ,
- le nombre de défauts ou de pannes dôune machine,
- la loi du nombre de paquets au niveau du routeur dôun réseau mobile,
- nombre de points jetés sur une cible circulaire du plan ou dôun autre espace etcé .

Exemple 6 : Le nombre dôétudiants qui arrivent au restaurant universitaire (durant la période dôouverture) suit une loi de Poisson de paramètre $\lambda = 2$ étudiants par minute. La probabilité dôarrivée de plus de 6 étudiants durant la première minute dôouverture est

$$P(X \geq 10) = e^{-2} \sum_{i=10}^{\infty} \frac{\lambda^i}{i!} = 1 - e^{-2} \sum_{i=0}^{10} \frac{\lambda^i}{i!} = 0.00453381$$

Le nombre moyen dôétudiants par minute est justement $E(X) = \lambda = 2/\min ute$, par conséquent, le nombre moyen dôétudiants durant les deux heures dôouverture sera $2 \times 2 \times 60 = 240$ étudiants.

Exemple 6(bis) : Si lôévènement consiste en lôarrivée de lôétudiant au restaurant universitaire , on doit avoir :

(i) P(un évènement durant Δt)=$\lambda \Delta t + O(\Delta t)$,

(ii) P(2 évènement durant Δt)=$O(\Delta t)$,

(iii) Les nombres dôévènements durant des intervalles disjoints sont des variables aléatoires indépendantes et de même loi Poisson (λ).

Loi uniforme discrète.

Côest la loi dôune variable discrète qui prend chacune des valeurs entières de lôintervalle $[0, N]$ avec des probabilités équiprobables

$$P(X = i) = 1/N, \qquad i = 1,2,...,N.$$