Probabilités et Statistique pour Informatique et Communications

http://stat.epfl.ch

3 Variables Aléatoires slide 87

3.1 Idées de Base slide 88

Variables aléatoires

Souvent on considère des quantités aléatoires numériques.

Exemple 62. *Deux dés équilibrés sont lancés, un rouge et un vert. Soit X le total des faces supérieures. Trouver les valeurs possibles de X, et les probabilités correspondantes.*

Définition 63. *Soit $(\Omega, \mathcal{F}, \Pr)$ un espace de probabilité. Une* **variable aléatoire (va)** *$X : \Omega \mapsto \mathbb{R}$ est une application de l'ensemble fondamental Ω dans $\mathbb{R}$.*

Définition 64. *L'ensemble des valeurs prises par X,*

$$D_X = \{x \in \mathbb{R} : \exists \omega \in \Omega \text{ tel que } X(\omega) = x\}$$

s'appele le **support** *de X. Si D_X est dénombrable, alors X est une* **variable aléatoire discrète**.

La va X associe des probabilités à des sous-ensembles S inclus dans $\mathbb{R}$, données par

$$\Pr(X \in S) = \Pr(\{w \in \Omega : X(w) \in S\}).$$

En particulier, on pose $A_x = \{\omega \in \Omega : X(\omega) = x\}$. Il est à noter qu'il faut que $A_x \in \mathcal{F}$, pour pouvoir calculer $\Pr(X = x)$. Si ce n'est pas le cas, on dit que X n'est pas **mesurable** par rapport à $(\Omega, \mathcal{F}, \Pr)$.

Exemples

Exemple 65. *On jette une pièce plusieurs fois et indépendamment. Soit X la variable aléatoire représentant le nombre de lancés nécessaires jusqu'à ce qu'on obtienne face. Calculer*

$$\Pr(X = 3),\ \Pr(X = 15),\ \Pr(X \leq 3.5),\ \Pr(X > 1.7),\ \Pr(1.7 \leq X \leq 3.5).$$

Exemple 66. *Un ensemble Ω naturel quand je joue aux fléchettes est le mur sur lequel la cible est fixée. La fléchette attérit à un point $\omega \in \Omega \subset \mathbb{R}^2$. Mon score $X(\omega) \in D_X = \{0, 1, \ldots, 60\}$.*

Jacob Bernoulli (1654–1705)

Ars Conjectandi, Basel (1713)

Variable aléatoire de Bernoulli

Définition 67. *Une variable aléatoire qui prend comme valeurs seulement 0 et 1 est appelée une* **variable indicatrice**, *ou une* **variable aléatoire de Bernoulli**, *ou un* **essai de Bernoulli**.

Exemple 68. *Supposons que n pièces identiques sont lancées indépendamment, soit F_i l'événement 'on obtient face pour la $i$ème pièce', et soit $I_i = I(F_i)$ l'indicatrice de cet événement. Alors*

$$\Pr(I_i = 1) = \Pr(F_i) = p, \quad \Pr(I_i = 0) = \Pr(F_i^c) = 1 - p,$$

où p est la probabilité d'obtenir face. Si $n = 3$ et $X = I_1 + I_2 + I_3$, décrire Ω, D_X et les ensembles A_x. Que représentent

$$X = I_1 + \cdots + I_n, \quad Y = I_1(1 - I_2)(1 - I_3), \quad Z = \sum_{j=2}^{n} I_{j-1}(1 - I_j)?$$

Fonction de masse

Une variable aléatoire X associe des probabilités à des sous-ensembles de $\mathbb{R}$. En particulier lorsque X est discrète, nous avons $A_x = \{\omega \in \Omega : X(\omega) = x\}$, et nous pouvons définir :

Définition 69. *La* **fonction de masse (fm)** *de X est*

$$f_X(x) = \Pr(X = x) = \Pr(A_x), \quad x \in \mathbb{R}.$$

Elle a deux propriétés clés :

(i) $f_X(x) \geq 0$, *et est positif seulement pour $x \in D_X$, où D_X est l'ensemble image de X, c'est à dire le* **support** *de f_X ;*

(ii) *la probabilité totale $\sum_{\{i:x_i \in D_X\}} f_X(x_i) = 1$.*

Quand il n'y a pas de risque de confusion, notons $f_X \equiv f$ et $D_X \equiv D$.

En anglais la fonction de masse est appelée **probability mass function (PFM)** ou **probability density function (PDF)**.

Variable aléatoire binomiale

Exemple 70. *Donner les fm et supports de I_i, de Y et de X.*

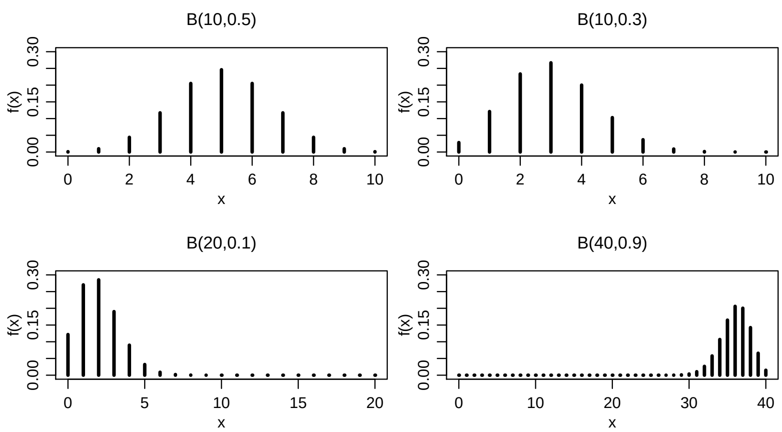
Définition 71. *Une variable aléatoire* **binomiale** *X a une fm*

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n, \quad n \in \mathbb{N}, 0 \leq p \leq 1.$$

On note $X \sim B(n, p)$, et appelle n le **dénominateur** *et p la* **probabilité de succès***. Avec $n = 1$, c'est une variable de Bernoulli.*

Remarque : on utilise $\sim$ comme abréviation de 'a pour distribution'.
Le modèle binomial est utilisé quand on considère le nombre de 'succès' d'une épreuve répétée de façon indépendante un nombre fixé de fois, et que chaque essai a la même probabilité de succès.

Fonctions de masse binomiale



Exemples

Exemple 72. *Un test contient 20 questions, pour chacune d'elles il faut choisir la bonne réponse parmi 5 réponses possibles. La moyenne est obtenue avec 10 réponses justes. Un étudiant choisit ses réponses au hasard. Donner la loi de son nombre de réponses justes. Quelle est la probabilité qu'il réussisse le test ?*

Exemple 73. *Certains traits physiques sont déterminés par une paire de gènes représenté par 2 allèles : dominant d et récessif r. Une personne porteur du génotype dd est dominant pur, du génotype dr est un hybride, et du génotype rr est récessif. Les génotypes dd et rd déterminent le même phénotype, et donc ne peuvent pas être différenciés physiquement. Un enfant reçoit au hasard un gène de chaque parent. Si les parents sont tous les 2 hybrides et ont 4 enfants, quelle est la probabilité que seulement 3 des enfants ont le phénotype dominant ? Quelle est la probabilité qu'au plus 3 d'entre eux ont le phénotype dominant ?*

Loi géométrique

Définition 74. *Une variable aléatoire* **géométrique** X *a pour fm*

$$f_X(x) = p(1-p)^{x-1}, \quad x = 1, 2, \ldots, \quad 0 \leq p \leq 1.$$

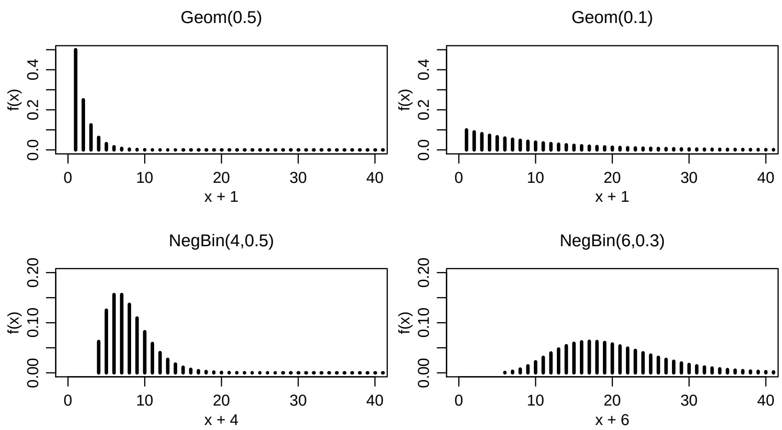
On note $X \sim \mathrm{Geom}(p)$, *et on appelle* p *la* **probabilité de succès**.

Elle modélise le temps d'attente jusqu'à un premier événement, dans une série d'essais indépendants ayant la même probabilité de succès.

Exemple 75. *Pour commencer un jeu de société, des joueurs jettent un dé chacun à leur tour. Le premier qui obtient six commence. Donner les probabilités que le 3ème joueur commence, et d'attendre au moins 6 lancés de dé avant le début du jeu.*

Théorème 76 (Manque de mémoire)**.** *Si* $X \sim \mathrm{Geom}(p)$, *alors* $\Pr(X > n + m \mid X > m) = \Pr(X > n)$.

FMs géométrique et binomiale négative



Loi binomiale négative

Définition 77. *Une variable aléatoire* **binomiale negative** X *de paramètres* n *et* p *a pour fonction de masse*

$$f_X(x) = \binom{x-1}{n-1} p^n (1-p)^{x-n}, \quad x = n, n+1, n+2, \dots, \quad 0 \le p \le 1.$$

On note $X \sim \mathrm{NegBin}(n, p)$. *Lorsque* $n = 1$, $X \sim \mathrm{Geom}(p)$.

Elle modélise le temps d'attente jusqu'au $n$ème succès dans une série d'essais indépendants ayant la même probabilité de succès.

Exemple 78. *Deux joueurs lancent successivement une pièce. Quelle est la probabilité que 2 faces apparaissent avant 5 piles ?*

Exemple 79 (Problème des allumettes de Banach)**.** *Un mathématicien fumeur de pipe a une boîte d'allumettes dans chacune des poches de sa veste, une du côté droit et une du côté gauche. Les deux boîtes contiennent initiallement* m *allumettes. Chaque fois qu'il allume sa pipe, il choisit une boîte d'allumettes au hasard, et jette l'allumette utilisée. Après un moment il constate que la boîte qu'il a choisie est vide. Quelle est alors la distribution du nombre d'allumettes dans l'autre boîte ?*

Stefan Banach (1892–1945)

Théorie des opérations linéaires (1932)

Szkocka Café, Lvov

Lieu de rencontre pour beaucoup de mathématiciens, tels que Banach, Steinhaus, Ulam, Mazur et Kac. Des problèmes ont été écrits dans un livre tenu par le proprietaire, avec des prix offerts pour des solutions.

Loi binomiale négative : version alternative

Parfois on écrit les variables géométriques et binomiale negatives sous une forme plus générale, prenant $Y = X - n$, et alors la fonction de masse est

$$f_Y(y) = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)y!} p^\alpha (1-p)^y, \quad y = 0, 1, 2, \ldots, \quad 0 \le p \le 1, \alpha > 0,$$

où

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u}\, du, \quad \alpha > 0$$

est la **fonction Gamma**. Ses propriétés principales sont :

$$\begin{aligned} \Gamma(1) &= 1; \\ \Gamma(\alpha+1) &= \alpha\Gamma(\alpha), \quad \alpha > 0; \\ \Gamma(n) &= (n-1)!, \quad n = 1, 2, 3, \ldots; \\ \Gamma(\tfrac{1}{2}) &= \sqrt{\pi}. \end{aligned}$$

Distribution hypergéométrique

Définition 80. *On tire sans remise un échantillon de m boules d'une urne contenant b blanches et n noires. Soit X le nombre de boules blanches tirées. Alors*

$$\Pr(X = i) = \frac{\binom{b}{i}\binom{n}{m-i}}{\binom{b+n}{m}}, \quad i = \max(0, m-n), \ldots, \min(b, m),$$

et la loi de X est **hypergéométrique**.

Exemple 81. *J'ai six boîtes, dont 2 contiennent du fruit. Si je choisis 3 des 6 au hasard, trouver la loi du nombre de boîtes de fruit parmi les 3.*

Exemples

Exemple 82. *Un électricien achète des composants par paquets de 10. Il examine trois composants choisis au hasard dans un paquet, et il l'accepte seulement si les trois choisis sont parfaits. Si 30% des paquets contiennent 4 mauvais composants et les autres 70% contiennent juste un mauvais composant, quelle proportion de paquets va-il rejeter ?*

Exemple 83. *Dans le but d'estimer le nombre de poissons N dans un lac, nous attrapons tout d'abord r poissons, les marquons, et les relâchons. Après avoir attendu assez longtemps pour que la population de poissons soit bien mélangée, nous prélevons un autre échantillon de taille s, composé de $0 \leq m \leq s$ poissons marqués. Trouver la loi du nombre de poissons marqués, M, dans cet échantillon. Montrer que la valeur de N qui maximise $\Pr(M = m)$ est $\lfloor rs/m \rfloor$, et calculer la meilleure estimation de N lorsque $s = 50$, $r = 40$, et $m = 4$.*

L'idee à la base de cet exemple est utilisée pour estimer les tailles des populations des espèces en danger, le nombre de toxicomanes, de sans-papiers, etc., dans les populations humaines, etc. Un problème pratique souvent rencontré est que certains individus deviennent plus difficile à recapturer, alors que d'autres l'aiment ; ainsi les probabilités de reprise sont hétérogènes.

Fonction de répartition

Définition 84. *La* **fonction de répartition (fr)** *d'une variable aléatoire X est*

$$F_X(x) = \Pr(X \leq x), \quad x \in \mathbb{R}.$$

Si X est discrète, on peut écrire

$$F_X(x) = \sum_{\{x_i \in D_X : x_i \leq x\}} \Pr(X = x_i),$$

c'est une fonction en escalier avec des sauts aux points du support D_X de $f_X(x)$.

Quand il n'y a pas de risque de confusion, on note $F \equiv F_X$.

Exemple 85. *Donner le support et les fonctions de masse et de répartition d'une variable aléatoire de Bernoulli.*

Exemple 86. *Donner la fonction de répartition d'une variable aléatoire géométrique.*

Exemples

La définition suivante généralise le résultat d'un jet de dé.

Définition 87. *Une variable aléatoire* **discrète uniforme** *X a pour fm*

$$f_X(x) = \frac{1}{b-a+1}, \quad x = a, a+1, \ldots, b, \quad a < b, \quad a, b \in \mathbb{Z}.$$

La loi de Poisson apparaît partout dans la probabilité et les statistiques.

Définition 88. *Une variable aléatoire de* **Poisson** *X a pour fm*

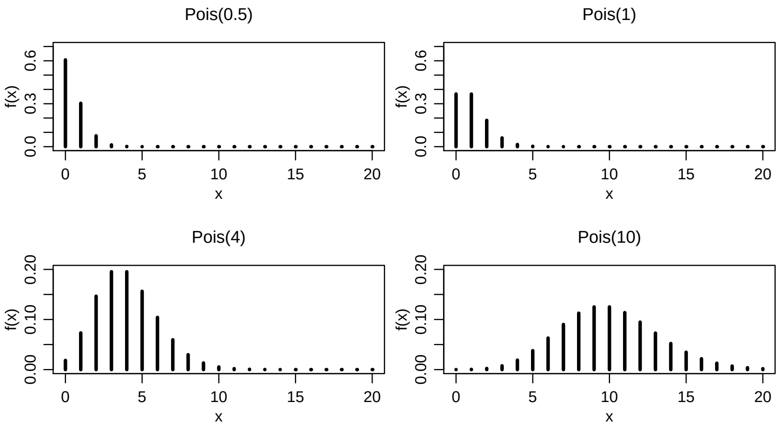
$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, \ldots, \quad \lambda > 0.$$

On note $X \sim \mathrm{Pois}(\lambda)$.

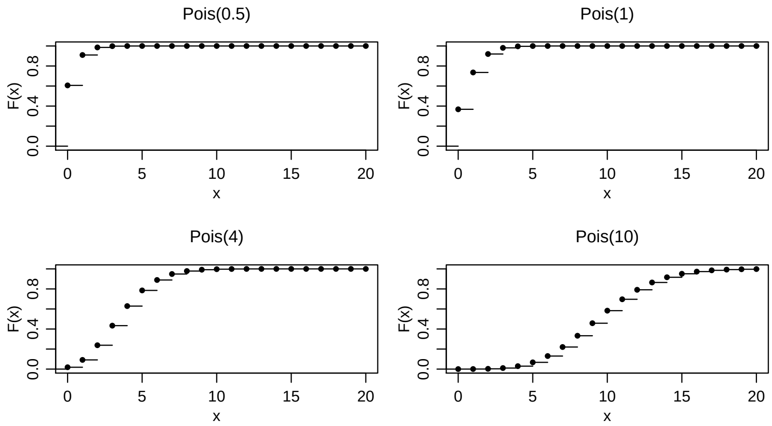
Siméon-Denis Poisson (1781–1840)

'La vie n'est bonne qu'à deux choses : à faire des mathématiques et à les professer.'

Fonctions de masse de Poisson



Fonctions de répartition de Poisson



Propriétés d'une fonction de répartition

Théorème 89. *Soit* $(\Omega, \mathcal{F}, \Pr)$ *un espace de probabilité et* $X : \Omega \mapsto \mathbb{R}$ *une variable aléatoire. Sa fonction de répartition* F_X *satisfait :*

(a) $\lim_{x\to-\infty} F_X(x) = 0$ *;*

(b) $\lim_{x\to\infty} F_X(x) = 1$ *;*

(c) F_X *est non-décroissante, ainsi* $F_X(x) \leq F_X(y)$ *pour* $x \leq y$ *;*

(d) F_X *est continue à droite, ainsi*

$$\lim_{t\downarrow 0} F_X(x+t) = F_X(x), \quad x \in \mathbb{R};$$

(e) $\Pr(X > x) = 1 - F_X(x)$ *;*

(f) si $x < y$*, alors* $\Pr(x < X \leq y) = F_X(y) - F_X(x)$*.*

Remarques

– On peut obtenir la fonction de masse à partir de la fonction de répartition par

$$f(x) = F(x) - \lim_{y\uparrow x} F(y).$$

Dans de nombreux cas X ne prend que des valeurs entières, $D_X \subset \mathbb{Z}$, et alors $f(x) = F(x) - F(x-1)$ pour $x \in \mathbb{Z}$.

– Dorénavant nous ignorerons la plupart du temps l'espace de probabilité $(\Omega, \mathcal{F}, \Pr)$ sous-jacent quand on a affaire à une variable aléatoire X. Nous penserons plutôt en termes de X, $F_X(x)$, et $f_X(x)$. On peut légitimer cet 'oubli' mathématiquement.

– On peut specifier la loi d'un variable aléatoire de manière équivalente en disant (par ex.) :
  – X suit une loi de Poisson avec paramètre λ ; ou
  – $X \sim$ Pois(λ) ; ou
  – en donnant la fonction de masse de X ; ou
  – en donnant la fonction de répartition de X.

Transformations de variables aléatoires discrètes

Des fonctions à valeurs réelles de variables aléatoires sont elles-même des variables aléatoires, elles ont donc aussi des fonctions de masse et de répartition.

Théorème 90. *Si X est une variable aléatoire et $Y = g(X)$, alors Y a pour fonction de masse*

$$f_Y(y) = \sum_{x:g(x)=y} f_X(x).$$

Exemple 91. *Calculer la fonction de masse de $Y = I(X \geq 1)$ lorsque $X \sim \mathrm{Pois}(\lambda)$.*

Exemple 92. *Soit Y le reste de la division par quatre du total de 2 lancés indépendants d'un dé. Calculer la fm de Y.*

3.2 Espérance slide 113

Espérance

Définition 93. *Soit X une variable aléatoire discrète pour laquelle $\sum_{x\in D_X} |x| f_X(x) < \infty$, où D_X est le support de f_X.* **L'espérance** *de X est*

$$\mathrm{E}(X) = \sum x \Pr(X = x) = \sum_{x\in D} x f_X(x).$$

- $\mathrm{E}(X)$ est parfois appelée la "moyenne de X". Nous limiterons l'utilisation du mot "moyenne" aux quantités empiriques.
- L'espérance est analogue en mécanique à la notion de **centre de gravité** d'un objet dont la masse est distribuée selon f_X.

Exemple 94. *Calculer l'espérance d'une va de Bernoulli de probabilité p.*

Exemple 95. *Calculer l'espérance de $X \sim B(n, p)$.*

Exemple 96. *Calculer l'espérance des variables aléatoires de fms suivantes*

$$f_X(x) = \frac{4}{x(x+1)(x+2)}, \quad f_Y(x) = \frac{1}{x(x+1)}, \quad x = 1, 2, \ldots.$$

Espérance d'une fonction

Théorème 97. *Soit X une variable aléatoire de fonction de masse f, et soit g une fonction à valeurs réelles de $\mathbb{R}$. Alors*

$$\mathrm{E}\{g(X)\} = \sum_x g(x)f(x),$$

lorsque $\sum_x |g(x)|f(x) < \infty$.

Exemple 98. *Soit $X \sim \mathrm{Pois}(\lambda)$. Calculer les espérances de*

$$X, \quad X(X-1), \quad X(X-1)\cdots(X-r+1), \quad \cos(\theta X).$$

Propriétés de l'espérance

Théorème 99. *Soit X une variable aléatoire d'espérance finie $\mathrm{E}(X)$, et soit $a, b \in \mathbb{R}$ des constantes. Alors*

(a) $\mathrm{E}(\cdot)$ est un opérateur linéaire, i.e. $\mathrm{E}(aX+b) = a\mathrm{E}(X) + b$;

(b) si $\Pr(X = b) = 1$, alors $\mathrm{E}(X) = b$;

(c) si $\Pr(a < X \le b) = 1$, alors $a < \mathrm{E}(X) \le b$;

(d) si $g(X)$ et $h(X)$ ont des espérances finies, alors

$$\mathrm{E}\{g(X) + h(X)\} = \mathrm{E}\{g(X)\} + \mathrm{E}\{h(X)\};$$

(e) $\{\mathrm{E}(X)\}^2 \le \{\mathrm{E}(|X|)\}^2 \le \mathrm{E}(X^2)$.

Remarque : La linéarité de l'espérance, (a), est très utile en pratique.

Exemples

Exemple 100. *Soit $X = I_1 + \cdots + I_n$, où $I_1, \ldots, I_n$ sont des variables de Bernoulli indépendantes de probabilité p. Calculer $\mathrm{E}(X)$. L'indépendance des I_i est nécessaire ?*
Dans l'exemple 51, soit X le nombre d'hommes qui s'en vont avec le correct chapeau. Montrer que $\mathrm{E}(X) = 1$, pour tout n.

Exemple 101. *Soit $I_A, I_B, \ldots$ les indicatrices des événements $A, B, \ldots$. Montrer que*

$$I_{A\cap B} = I_A I_B, \quad I_{A\cup B} = 1 - (1 - I_A)(1 - I_B), \quad \mathrm{E}(I_A) = \Pr(A).$$

et en déduire la formule d'inclusion-exclusion

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{r=1}^n (-1)^{r+1} \sum_{1\le i_1<\cdots<i_r\le n} \Pr(A_{i_1} \cap \cdots \cap A_{i_r}).$$

Moments d'une distribution

Définition 102. *Si X a une fm $f(x)$ telle que $\sum_x |x|^r f(x) < \infty$, alors*

(a) le $r$ème **moment** *de X est $\mathrm{E}(X^r)$;*

(b) le $r$ème **moment centré** *de X est $\mathrm{E}[\{X - \mathrm{E}(X)\}^r]$;*

(c) le $r$ème **moment factoriel** *de X est $\mathrm{E}\{X(X-1)\cdots(X-r+1)\}$;*

(d) la **variance** *de X est $\mathrm{var}(X) = \mathrm{E}[\{X - \mathrm{E}(X)\}^2]$.*

Remarque : De tous ces moments l'espérance et la variance sont les plus importants, car ils mesurent la localisation et la dispersion de f_X. La variance est analogue en mécanique au **moment d'inertie**.

Exemple 103. *Calculer la variance du score quand on lance un dé.*

Propriétés de la variance

Théorème 104. *Soit X une variable aléatoire dont la variance existe, et soient a, b des constantes. Alors*

$$\begin{aligned} \mathrm{var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \mathrm{E}\{X(X-1)\} + \mathrm{E}(X) - \mathrm{E}(X)^2; \\ \mathrm{var}(aX+b) &= a^2\mathrm{var}(X); \\ \mathrm{var}(X) = 0 &\Rightarrow X \text{ est constante de probabilité 1.} \end{aligned}$$

Exemple 105. *Calculer les divers moments d'une variable aléatoire Poissonienne.*

Propriétés de la variance II

Théorème 106. *Si X prend ses valeurs dans $\{0, 1, \ldots\}$, $r \geq 2$, et $\mathrm{E}(X) < \infty$, alors*

$$\begin{aligned} \mathrm{E}(X) &= \sum_{x=1}^{\infty} \mathrm{Pr}(X \geq x), \\ \mathrm{E}\{X(X-1)\cdots(X-r+1)\} &= r\sum_{x=r}^{\infty}(x-1)\cdots(x-r+1)\mathrm{Pr}(X \geq x). \end{aligned}$$

Exemple 107. *Soit $X \sim \mathrm{Geom}(p)$. Calculer $\mathrm{E}(X)$ et $\mathrm{var}(X)$.*

Exemple 108. *Chaque paquet d'un certain produit a d'égales chances de contenir un des n différents types de bons de réduction, indépendamment de chaque autre paquet. Quel est le nombre espéré de paquets que vous devez acheter pour obtenir au moins un de type de chaque bon ?*

Lois conditionnelles

Définition 109. *Soit $(\Omega, \mathcal{F}, \Pr)$ un espace de probabilité, sur lequel on définie une variable aléatoire X, et soit $B \in \mathcal{F}$ avec $\Pr(B) > 0$. Alors la* **fonction de masse conditionnelle** *de X sachant B est*

$$f_X(x \mid B) = \Pr(X = x \mid B) = \Pr(A_x \cap B)/\Pr(B),$$

où $A_x = \{\omega \in \Omega : X(\omega) = x\}$.

Théorème 110. *La fonction $f_X(x \mid B)$ satisfait*

$$f_X(x \mid B) \geq 0, \quad \sum_x f_X(x \mid B) = 1,$$

et est ainsi une fonction de masse bien définie.

Exemple 111. *Calculer les fm conditionnelles de $X \sim \mathrm{Geom}(p)$, (a) sachant que $X > n$, (b) sachant que $X \leq n$.*

Espérance conditionnelle

Définition 112. *Supposons que $\sum_x |g(x)| f_X(x \mid B) < \infty$. Alors l'espérance conditionnelle de $g(X)$ sachant B est*

$$\mathrm{E}\{g(X) \mid B\} = \sum_x g(x) f_X(x \mid B).$$

Théorème 113. *Soit X une variable aléatoire d'espérance $\mathrm{E}(X)$ et soit B un événement avec $\Pr(B), \Pr(B^c) > 0$. Alors*

$$\mathrm{E}(X) = \mathrm{E}(X \mid B)\Pr(B) + \mathrm{E}(X \mid B^c)\Pr(B^c).$$

Plus généralement, lorsque $\{B_i\}_{i=1}^{\infty}$ est une partition de Ω, $\Pr(B_i) > 0$ pour tout i, et que la somme est absolument convergente, alors

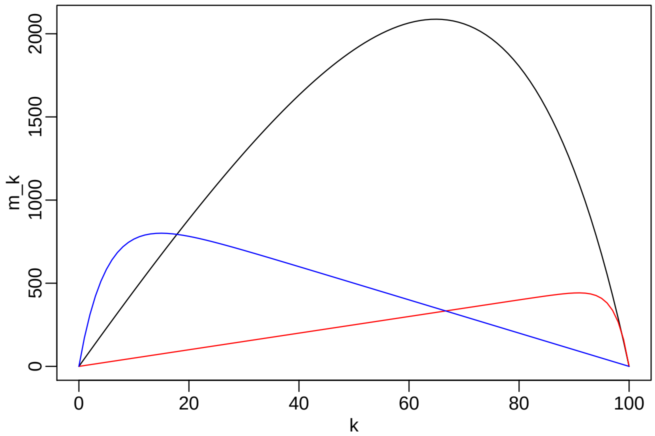
$$\mathrm{E}(X) = \sum_{i=1}^{\infty} \mathrm{E}(X \mid B_i)\Pr(B_i).$$

Exemples

Exemple 114. *La distribution de Poisson tronquée est définie en posant $X \sim \mathrm{Pois}(\lambda)$ et $B = \{X > 0\}$. Calculer la fonction de masse conditionnelle et l'espérance de cette distribution.*

Exemple 115 (The Hobbit)**.** *Bilbon le hobbit et Smaug le dragon ont respectivement b et de s pièces d'or. Ils jouent à une série de jeux indépendants à l'issue duquel le perdant donne une pièce d'or au gagnant. Ils s'arrêtent de jouer lorsque l'un d'entre eux n'a plus de pièce. Si Bilbon gagne chaque jeu avec une probabilité p (et $p \neq q = 1 - p$), calculer le nombre espéré de jeux avant qu'ils ne s'arrêtent.*

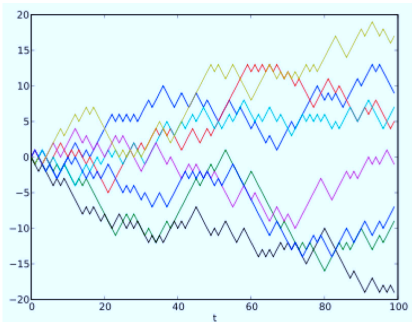
Temps moyen d'attente



Le nombre espéré de jeux dans l'exemple 115 pour $p = 0.49$ (noir), $p = 0.4$ (rouge), $p = 0.55$ (bleu).

Marche aléatoire

Les exemples 115 et 52 sont des exemples d'un modèle probabiliste qui s'appelle la **marche aléatoire**. Ceci décrit les mouvements d'une particule sur un ensemble $\mathcal{S}$. Dans ce cas $\mathcal{S} \subset \mathbb{Z}$, mais on peut généraliser $\mathcal{S}$ à des sous-ensembles de $\mathbb{Z}^p$ ou aux reseaux tel que le web. Les conditionnements sont essentiels pour obtenir ses propriétés. On peut utiliser une marche aléatoire sur un réseau de taille inconnue pour estimer le nombre de sommets du réseau, etc. Ici 20 réalisations d'une marche aléatoire avec $p = q = 1/2$, à partir de l'origine, sur $\mathbb{Z}$:



(Wikipedia)

Convergence des distributions

On veut souvent approcher une distribution par une autre. La base mathématique pour le faire est la convergence des distributions.

Définition 116. *Soient $\{X_n\}$, X des variables aléatoires dont les fonctions de répartitions sont $\{F_n\}$, F. Alors on dit que les variables aléatoires $\{X_n\}$* **convergent en distribution** *(ou* **en loi***) vers X, si, pour tout $x \in \mathbb{R}$ où F est continue, on a*

$$F_n(x) \to F(x), \quad n \to \infty.$$

On écrit alors $X_n \xrightarrow{D} X$.

Si $D_X \subset \mathbb{Z}$, alors $F_n(x) \to F(x)$ si $f_n(x) \to f(x)$ pour tout x, $n \to \infty$.

Lemme 117. $n^{-r}\binom{n}{r} \to 1/r!$ *pour tout $r \in \mathbb{N}$, quand $n \to \infty$.*

Loi des petits nombres

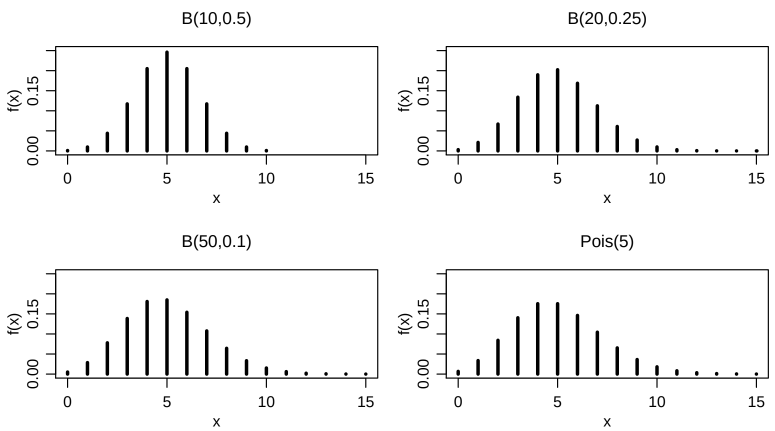
Théorème 118 (Loi des petits nombres)**.** *Soit $X_n \sim B(n,p)$, et supposons que $np \to \lambda > 0$ lorsque $n \to \infty$, alors la fonction de masse limite de X_n est $\mathrm{Pois}(\lambda)$.*

Exemple 119. *Dans l'Exemple 51 on a vu que la probabilité d'avoir exactement r points fixes dans une permutation au hasard de n objets est*

$$\frac{1}{r!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} \to \frac{e^{-1}}{r!} \quad \text{as } n \to \infty.$$

Ainsi le nombre de points fixes a une distribution limite de $\mathrm{Pois}(1)$.

Loi des petits nombres



Fonctions de masse de trois lois binomiales la loi de Poisson, toutes avec espérance 5.

Comparaison numérique

Exemple 120 (Loi binomiale et loi Poisson). *Comparer* $\Pr(X \leq 3)$ *pour* $X \sim B(20, p)$, *avec* $p = 0.05, 0.1, 0.2, 0.5$ *avec les resultats d'une approximation de Poisson, en utilisant les fonctions* `pbinom` *et* `ppois` *du logiciel* `R` *— voir*

http://www.r-project.org/

Ainsi par exemple, on a :

```
> pbinom(3,size=20,prob=0.05)
[1] 0.9840985
> ppois(3,lambda=20*0.05)
[1] 0.9810118
```


People versus Collins

Exemple 121. *En 1964 un sac-à-main a été volé à Los Angeles par une jeune femme avec des cheveux blonds en queue de cheval. La voleuse a disparu, mais peu après on l'a aperçu dans une voiture jaune avec un noir barbu avec moustache. La police a ensuite arrêté une femme appellée Janet Collins, qui ressemblait à la description, et avait un ami noir barbu avec moustache, conducteur d'une voiture jaune.*

Puisque il manquait de preuves et de témoins fiables, le procureur a essayé de convaincre le jury que Mme Collins et son ami étaient le seul couple à Los Angeles qui aurait pu commettre le délit. Il a trouvé une probabilité $p = 1/(12 \times 10^6)$ *qu'un couple tiré au hasard suivre la description, et ils ont été condamné.*

Dans un tribunal supérieur on a argumenté que le nombre de couples X *suivant la description devrait suivre une loi de Poisson avec* $\lambda = np$, *où* n *est la taille de la population à laquelle le couple appartient. Pour être certain que ce couple soit coupable, il faut que* $\Pr(X > 1 \mid X \geq 1)$ *soit très petite. Mais avec* $n = 10^6$, 2×10^6, 5×10^6, 10×10^6, *ces probabilités sont* 0.041, 0.081, 0.194, 0.359 *: c'était donc loin d'être certain qu'ils soient les coupables. Ils ont finalement été blanchi.*

Exemple 122. *Soit X_N une variable hypergéométrique, alors*

$$\Pr(X_N = i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}, \quad i = \max(0, m+n-N), \ldots, \min(m, n).$$

Ceci est la distribution du nombre de balles blanches obtenues quand on prélève un échantillon aléatoire de taille n sans remise d'une urne contenant m balles blanches et $N-m$ balles noires. Montrer que lorsque $N, m \to \infty$ de façon à ce que $m/N \to p$, où $0 < p < 1$,

$$\Pr(X_N = i) \to \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, \ldots, n.$$

Ainsi la distribution limite de X_N est $B(n, p)$.

3.5 Fonctions Génératrices de Moments

Définition

Définition 123. *On définit la* **fonction génératrice des moments** *d'une variable aléatoire X par*

$$M_X(t) = \mathrm{E}(e^{tX})$$

pour $t \in \mathbb{R}$ tel que $M_X(t) < \infty$.

- $M_X(t)$ est aussi appelé la **transformée de Laplace** de $f_X(x)$.
- La FGM est utile comme sommaire de toutes les propriétés de X, on peut écrire

$$M_X(t) = \mathrm{E}(e^{tX}) = \mathrm{E}\left\{\sum_{r=0}^{\infty} \frac{t^r X^r}{r!}\right\} = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mathrm{E}(X^r),$$

  dont on peut obtenir tous les moments $\mathrm{E}(X^r)$ par différentiation.

Exemple 124. *Calculer $M_X(t)$ lorsque : (a) X est une variable indicatrice ; (b) $X \sim B(n, p)$; (c) $X \sim \mathrm{Pois}(\lambda)$.*

Théorèmes importants

Théorème 125. *Soit $M(t)$ la FGM d'un variable aléatoire X, alors*

$$\begin{aligned}
M_X(0) &= 1;\\
M_{a+bX}(t) &= e^{at} M_X(bt);\\
\mathrm{E}(X^r) &= \left.\frac{\partial^r M_X(t)}{\partial t^r}\right|_{t=0};\\
\mathrm{E}(X) &= M'_X(0);\\
\mathrm{var}(X) &= M''_X(0) - M'_X(0)^2.
\end{aligned}$$

Théorème 126 (Pas de preuve)**.** *Il existe une bijection entre les fonction de répartitions $F_X(x)$ et les fonctions génératrices des moments $M_X(t)$.*

Théorème 127 (Continuité, pas de preuve)**.** *Soient $\{X_n\}$, X des variables aléatoires avec fonctions de répartitions $\{F_n\}$, F, dont les FGMs $M_n(t)$, $M(t)$ existent pour $0 \le |t| < b$. S'il exist un $0 < a < b$ tel que $M_n(t) \to M(t)$ pour $|t| \le a$ quand $n \to \infty$, alors $X_n \xrightarrow{D} X$, c'est à dire, $F_n(x) \to F(x)$ en chaque $x \in \mathbb{R}$ où F est continue.*

Exemple 128. *Soit $X_n \sim B(n,p)$ et $X \sim \mathrm{Pois}(\lambda)$. Montrer que lorsque $n \to \infty, p \to 0$ de façon à ce que $np \to \lambda$, $X_n \xrightarrow{D} X$.*

4 Variables Aléatoires Continues slide 137

4.1 Notions de Base slide 138

Variables aléatoires continues

Dans beaucoup de situations, on veut travailler avec des variables continues :
– le temps jusqu'à la fin du cours $\in (0, 45)$ min ;
– la paire (hauteur, poids) $\in (0, \infty)^2$.
Jusqu'à présent nous avions supposé que le support

$$D_X = \{x \in \mathbb{R} : X(\omega) = x, \omega \in \Omega\}$$

de X est dénombrable, ainsi X est une variable aléatoire discrète. On suppose maintenant que D_X n'est pas dénombrable, ce qui implique aussi que Ω lui-même n'est pas dénombrable.

Définition 129 (Rappel)**.** *Soit* $(\Omega, \mathcal{F}, \Pr)$ *un espace de probabilité. La* **fonction de répartition** *d'une va* X *définie sur* $(\Omega, \mathcal{F}, \Pr)$ *est*

$$F(x) = \Pr(X \leq x) = \Pr(\mathcal{B}_x), \quad x \in \mathbb{R},$$

où $\mathcal{B}_x = \{\omega : X(\omega) \leq x\} \subset \Omega$.

Fonction de densité

Définition 130. *Une variable aléatoire* X *est* **continue** *s'il existe une fonction* $f(x)$*, appelée* **la densité** *de* X*, telle que*

$$\Pr(X \leq x) = F(x) = \int_{-\infty}^{x} f(u)\, du, \quad x \in \mathbb{R}.$$

Les propriétés de F impliquent (i) $f(x) \geq 0$, et (ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

Remarques :
– Evidemment, on a

$$f(x) = \frac{dF(x)}{dx}.$$

– Comme $\Pr(x < X \leq y) = \int_x^y f(u)\, du$ pour $x < y$, pour tout $x \in \mathbb{R}$,

$$\Pr(X = x) = \lim_{y \downarrow x} \Pr(x < X \leq y) = \lim_{y \downarrow x} \int_x^y f(u)\, du = \int_x^x f(u)\, du = 0.$$

– Si X est discrète, alors sa fm $f(x)$ est aussi appelée sa fonction de densité.

Motivation

On étudie des variables aléatoires pour plusieurs raisons :

– elles paraîssent dans les modèles simples mais puissants—par exemple, la loi **exponentielle** est la loi du temps d'attente dans un processus où des événements se passent de manière aléatoire ;

– elles fournissent des approximations simples mais très utiles pour des problèmes complexes—par exemple, la loi **normale** apparaît comme approximation pour la loi d'une moyennes, sous des conditions assez générales ;

– elles sont à la base de la modélisation de problèmes complexes soit en probabilité soit en statistiques—-par exemple, la loi de **Pareto** est souvent une bonne approximation pour les données à queues lourdes, dans la finance et avec l'internet.

On va parler de quelques lois très connues, mais il en a plein d'autres . . .

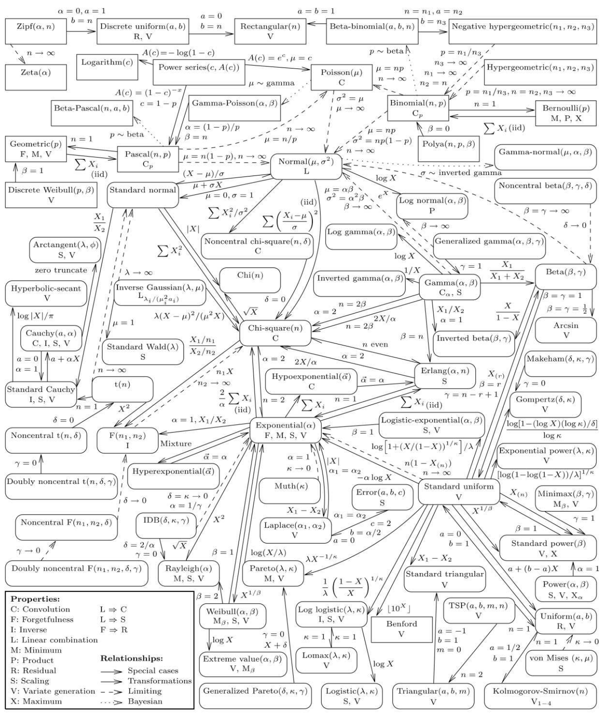


Figure 1. Univariate distribution relationships.

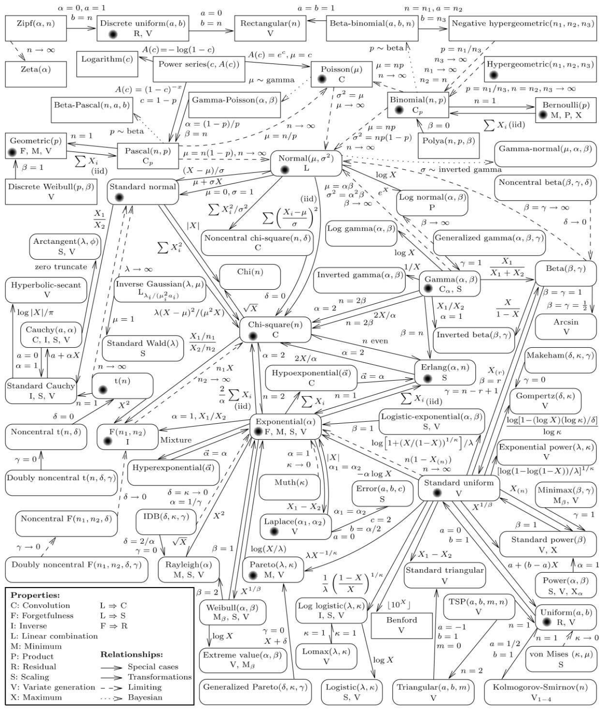


Figure 1. Univariate distribution relationships.

Lois de base

Exemple 131 (Uniforme). *La variable aléatoire U de densité*

$$f(u) = \begin{cases} \frac{1}{b-a}, & a < u < b, \\ 0, & \text{sinon}, \end{cases} \qquad a < b,$$

est appelée une variable aléatoire uniforme. On la note $U \sim U(a, b)$. Trouver sa fonction de répartition.

Exemple 132 (Exponentielle). *La variable aléatoire X de densité*

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

est appelée une variable aléatoire exponentielle de paramètre $\lambda > 0$. On la note $X \sim \exp(\lambda)$. Trouver sa fonction de répartition, et établir la propriété de **manque de mémoire** *de X :*

$$\Pr(X > x + t \mid X > t) = \Pr(X > x), \quad t, x > 0.$$

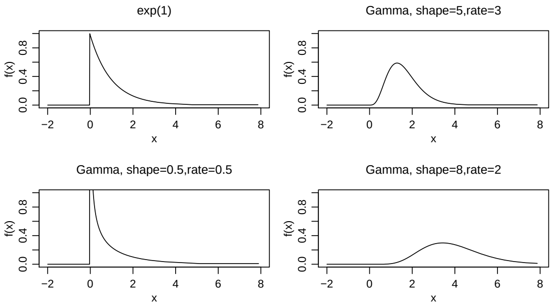
En pratique les vas sont presque toujours soit discrètes ou soit continues, avec quelques exceptions, tel que la loi de probabilité de la pluie journalière.

Loi de gamma

Exemple 133 (Gamma). *La variable aléatoire X de densité*

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

est appelée une variable aléatoire gamma de paramètres $\alpha, \lambda > 0$. Ici α s'appelle le **paramètre de forme (shape)**, *et λ s'appelle le* **taux (rate)**, *avec λ^{-1} le* **paramètre d'échelle (scale)**. *En posant $\alpha = 1$ on obtient la densité exponentielle, et quand $\alpha = 2, 3, \ldots$ on a la densité de* **Erlang**.



Loi de Laplace

Exemple 134 (Laplace). *La variable aléatoire X de densité*

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x-\eta|}, \quad x \in \mathbb{R}, \quad \eta \in \mathbb{R}, \lambda > 0,$$

est appelée une variable aléatoire de Laplace. Trouver sa fonction de répartition.

Pierre-Simon Laplace (1749–1827) : **Théorie Analytique des Probabilités** (1814)
Selon Napoléon Bonaparte : 'Laplace ne traitait aucune question d'un bon point de vue : il cherchait des subtilités de partout, il avait seulement des idées problématiques et enfin il portait l'esprit de l'infiniment petit jusque dans l'administration.'

Loi de Pareto

Exemple 135 (Pareto)**.** *La variable aléatoire X de fonction de répartition*

$$F(x) = \begin{cases} 0, & x < \beta, \\ 1 - \left(\frac{\beta}{x}\right)^{\alpha}, & x \geq \beta, \end{cases} \quad \alpha, \beta > 0,$$

est appelée une variable aléatoire de Pareto. Trouver sa fonction de densité.

Vilfredo Pareto (1848–1923) : Professeur à l'Université de Lausanne, père de l'économie scientifique.

Moments

Définition 136. *Soient $g(x)$ une fonction à valeurs réelles, et X une variable aléatoire continue de densité $f(x)$. Alors on définie l'***espérance** *de $g(X)$ comme*

$$\mathrm{E}\{g(X)\} = \int_{-\infty}^{\infty} g(x) f(x)\, dx,$$

*en supposant que $\mathrm{E}\{|g(X)|\} < \infty$. En particulier l'***espérance** *et la* **variance** *de X sont*

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x)\, dx, \\ \mathrm{var}(X) &= \int_{-\infty}^{\infty} \{x - \mathrm{E}(X)\}^2 f(x)\, dx = \mathrm{E}(X^2) - \mathrm{E}(X)^2. \end{aligned}$$

Exemple 137. *Calculer l'espérance et la variance des distributions : (a) $U(a, b)$; (b) exp(λ) ; (c) gamma ; (d) Laplace ; (e) Pareto.*

Fonction génératrice des moments

Définition 138. *La* **fonction génératrice des moments (fgm)** *de X est*

$$M_X(t) = \mathrm{E}\{\exp(tX)\},$$

pour $t \in \mathbb{R}$ tel que $M_X(t) < \infty$.

Exemple 139. *Calculer la fgm de la distribution $\exp(\lambda)$, et ainsi trouver $\mathrm{E}(X)$ et $\mathrm{var}(X)$.*

Densités conditionelles

On peut aussi calculer les fonctions de répartitions et densités conditionelles : pour des ensembles $\mathcal{A} \subset \mathbb{R}$ raisonnables on a

$$F_X(x \mid X \in \mathcal{A}) = \Pr(X \leq x \mid X \in \mathcal{A}) = \frac{\Pr(X \leq x \cap X \in \mathcal{A})}{\Pr(X \in \mathcal{A})} = \frac{\int_{\mathcal{A}_x} f(y)\, dy}{\Pr(X \in \mathcal{A})},$$

où $\mathcal{A}_x = \{y : y \leq x, y \in \mathcal{A}\}$, et

$$f_X(x \mid X \in \mathcal{A}) = \begin{cases} \frac{f_X(x)}{\Pr(X \in \mathcal{A})}, & x \in \mathcal{A}, \\ 0, & \text{sinon.} \end{cases}$$

Avec $I(X \in \mathcal{A})$ la variable indicatrice de l'événement $X \in \mathcal{A}$, on peut écrire

$$\mathrm{E}\{g(X) \mid X \in \mathcal{A}\} = \frac{\mathrm{E}\{g(X)\ I(X \in \mathcal{A})\}}{\Pr(X \in \mathcal{A})},$$

Exemple 140. *Soit $X \sim \exp(\lambda)$. Trouver la densité et la fonction de répartition de X, sachant que (a) $X < 3$, (b) $X > 3$.*

Exemple

Exemple 141. *Pour obtenir un visa pour un pays lointain, vous appelez chaque matin à son consulat à 10.00. Le fonctionnaire ne répond pas aux appels un jour sur deux, et quand il répond, il laisse l'appareil sonner pendant un temps aléatoire T (min) dont la loi est*

$$F_T(t) = \begin{cases} 0, & t \leq 1, \\ 1 - t^{-1}, & t > 1. \end{cases}$$

(a) Si vous appelez un matin et ne raccrochez pas, quelle est la probabilité que vous ecoutez la tonalité pendant au moins s minutes ?
(b) Vous décidez d'appeler une fois chaque jour, mais de raccrocher s'il n'y a pas eu de réponse après s minutes. Trouver la valeur de s qui minimise votre temps d'écoute à la tonalité.

Quantiles

Définition 142. *Soit $0 < p < 1$. On définie le $p$ième* **quantile** *de la fonction de répartition $F(x)$ par*

$$x_p = \inf\{x : F(x) \geq p\}.$$

Pour la plupart des variables aléatoire continues, x_p est unique et vaut $x_p = F^{-1}(p)$, où F^{-1} est la fonction inverse de F. Ainsi x_p est la valeur pour laquelle $\Pr(X \leq x_p) = p$. En particulier, on appelle le 0.5ème quantile la **médiane** *de F.*

Exemple 143. *Soit $U \sim U(0,1)$. Montrer que $x_p = p$.*

Exemple 144. *Soit $X \sim \exp(\lambda)$. Montrer que $x_p = -\lambda^{-1}\log(1-p)$.*

Exemple 145. *Trouver le $p$ième quantile de la loi de Pareto.*

Transformations

On considère souvent $Y = g(X)$, où g est une fonction connue, et on veut calculer F_Y et f_Y à partir de F_X et f_X.

Exemple 146. *Soit $X \sim \exp(\lambda)$ et $Y = \exp(X)$, trouver F_Y et f_Y.*

Exemple 147. *Soit $Y = -\log(1-U)$, où $U \sim U(0,1)$. Calculer $F_Y(y)$ et discuter. Calculer aussi la densité et la fonction de répartition de $W = -\log U$. Expliquer.*

Exemple 148. *Soit $Y = \lceil X \rceil$, où $X \sim \exp(\lambda)$ (ainsi Y est le plus petit entier plus grand que X). Calculer $F_Y(y)$ et $f_Y(y)$.*

Exemple 149. *Soient $Y = X^2$, où X suit la loi de Laplace avec $\lambda = 1$, $\eta = 0$. Calculer F_Y et f_Y.*

Transformation générale

Définition 150. *Soient* $g : \mathbb{R} \mapsto \mathbb{R}$ *une fonction et* $\mathcal{B} \subset \mathbb{R}$ *un sous-ensemble quelconque de* $\mathbb{R}$*. Alors* $g^{-1}(\mathcal{B}) \subset \mathbb{R}$ *est l'ensemble pour lequel* $g\{g^{-1}(\mathcal{B})\} = \mathcal{B}$*.*

Théorème 151. *Soit* $Y = g(X)$ *une variable aléatoire et* $\mathcal{B}_y = (-\infty, y]$*. Alors*

$$F_Y(y) = \Pr(Y \le y) = \begin{cases} \int_{g^{-1}(\mathcal{B}_y)} f_X(x)\,dx, & X \text{ continue}, \\ \sum_{x \in g^{-1}(\mathcal{B}_y)} f_X(x), & X \text{ discrète}, \end{cases}$$

où $g^{-1}(\mathcal{B}_y) = \{x \in \mathbb{R} : g(x) \le y\}$*. Lorsque* g *est monotone croissante et a pour fonction inverse* g^{-1}*, on a*

$$F_Y(y) = F_X\{g^{-1}(y)\}, \quad f_Y(y) = \frac{dg^{-1}(y)}{dy} f_X\{g^{-1}(y)\}, \quad y \in \mathbb{R}.$$

On obtient un résultat similaire pour g *monotone décroissante.*

Exemple 152. *Soient* $Y = \beta X^{1/\alpha}$*, où* $X \sim \exp(1)$ *et* $\alpha, \beta > 0$*. Trouver les fonctions de répartition et de densité de la variable aléatoire de* **Weibull** Y*.*

'Probability integral transform'

Ecrivons $\stackrel{D}{=}$ pour 'a la même loi que'.

Lemme 153 ('Probability integral transform (PIT)')**.** *Soient* $X \sim F$ *une va continue avec fonction de répartition* F*,* $U \sim U(0,1)$*, et*

$$F^{-1}(p) = \min\{x : F(x) \ge p\}, \quad 0 < p < 1.$$

Alors $F^{-1}(U) \stackrel{D}{=} X$*, et* $F(X) \stackrel{D}{=} U \sim U(0,1)$*.*

Exemple 154. *Si* $X \sim \exp(\lambda)$*, montrer que*

$$X \stackrel{D}{=} -\lambda^{-1} \log U.$$

La simulation : méthode de Monte Carlo

Beaucoup de problèmes probabilistes sont trop difficile à résoudre par la route analytique, et on utilise la **simulation** de variables pseudo-aléatoires, générées sur ordinateur. Ceci s'appelle aussi la **méthode de Monte Carlo**, très souvent utilisée pour approximer les integrales dans dimension élévée, ...

Ici on génère des variables aléatoires $U_1, \ldots, U_n$ selon la loi $U(0,1)$:

```
n <- 50
u <- runif(n) #  generate n U(0,1) variables
u[1:50]         # show the first 50 variables
EDF <- function(x,n=length(x)) list(x=sort(x),y=c(1:n)/n)  # EDF of u
par(mfrow=c(1,2),pty="s")  # set up graphics
lim <- c(0,1)    # we'll fix x-axis of graphs
hist(u,prob=T,nclass=20,xlim=lim)  # estimated density
rug(u)          # rug showing values of the u's
plot(EDF(u),type="s",panel.first={abline(0,1,col="grey")},xlim=lim) # CDF of u
```

Observer ce qui se passe quand on prend $n = 500, 5000, 50000, \ldots$

La méthode d'inversion

Le lemme 153 donne un moyen de simuler des variables $X_1, X_2, \ldots, \sim F$ par la **méthode d'inversion** :

$$X_1 = F^{-1}(U_1), X_2 = F^{-1}(U_2), \ldots, \text{ où } U_1, U_2, \ldots \sim U(0,1).$$

```
n <- 5000; lambda <- 3
u <- runif(n) #  generate n U(0,1) variables
x <- -log(u)/lambda  # transform the u's to exp(lambda)
x[1:50]         # show the first 50 variables
par(mfrow=c(2,2),pty="s")  # set up graphics
lim <- c(0,3)  # we'll fix x-axis of graphs
hist(x,prob=T,nclass=20,xlim=lim)  # estimated density
rug(x)
plot(EDF(x),type="s",xlim=lim) # CDF of x
y <- 1/u  # transform the u's to Pareto
y[1:50]         # show the first 50 variables
hist(y,prob=T,nclass=20)  # estimated density
rug(y)
plot(EDF(y),type="s") # CDF of y
```

Observer ce qui se passe quand on prend $n = 500, 5000, 50000, \ldots$

Loi normale

Définition 155. *Une variable aléatoire X de densité*

$$f(x) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \sigma > 0,$$

est une **variable aléatoire normale** *d'espérance μ et de variance σ^2 : on écrit $X \sim \mathcal{N}(\mu, \sigma^2)$. Quand $\mu = 0$, $\sigma^2 = 1$, la variable aléatoire correspondante Z est* **normale centrée réduite**, *$Z \sim \mathcal{N}(0,1)$, de densité*

$$\phi(z) = (2\pi)^{-1/2} e^{-z^2/2}, \quad z \in \mathbb{R}.$$

Alors

$$F_Z(x) = \Pr(Z \le x) = \Phi(x) = \int_{-\infty}^{x} \phi(z)\, dz = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} e^{-z^2/2}\, dz.$$

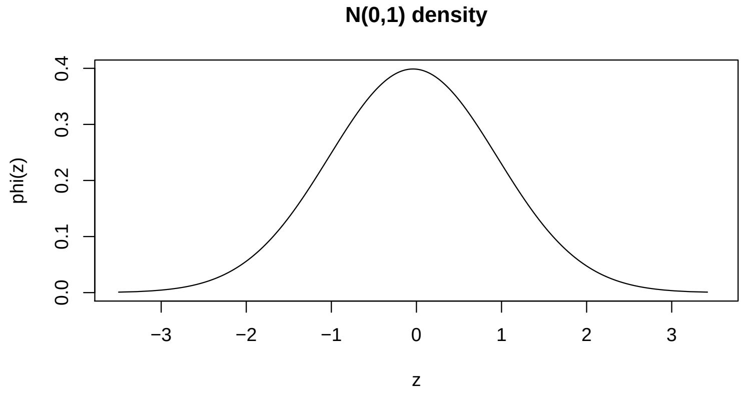
Cette intégrale est tabulée dans le formulaire.

Il est à noter que $f(x) = \sigma^{-1}\phi\{(x-\mu)/\sigma\}$.

Johann Carl Friedrich Gauss (1777–1855)

La loi normale est souvent appellée la **loi gaussiénne**. Gauss l'a utilisée pour la combination de mesures astronomiques et topographiques.

Densité normale centrée réduite



La fameuse courbe en cloche :

$$\phi(z) = (2\pi)^{-1/2} e^{-z^2/2}, \quad z \in \mathbb{R}.$$

Propriétés

Théorème 156. *La densité* $\phi(z)$*, la fonction de répartition* $\Phi(z)$*, et les quantiles* z_p *de* $Z \sim \mathcal{N}(0,1)$ *satisfont, pour tout* $z \in \mathbb{R}$ *:*

(a) la densité est symétrique par rapport à $z = 0$*, i.e.,* $\phi(z) = \phi(-z)$ *;*

(b) $\Pr(Z \leq z) = \Phi(z) = 1 - \Phi(-z) = 1 - \Pr(Z \geq z)$ *;*

(c) les quantiles normaux centrés réduits z_p *satisfont* $z_p = -z_{1-p}$*, pour tout* $0 < p < 1$ *;*

(d) $z^r \phi(z) \to 0$ *lorsque* $z \to \pm\infty$*, pour tout* $r > 0$*. Ceci implique que les moments* $\mathrm{E}(Z^r)$ *existent pour tout* $r \in \mathbb{N}$ *;*

(e) on a

$$\phi'(z) = -z\phi(z), \quad \phi''(z) = (z^2 - 1)\phi(z), \quad \phi'''(z) = -(z^3 - 3z)\phi(z), \quad \ldots$$

Ceci implique que $\mathrm{E}(Z) = 0$*,* $\mathrm{var}(Z) = 1$*,* $\mathrm{E}(Z^3) = 0$*, etc.*

(f) Si $X \sim \mathcal{N}(\mu, \sigma^2)$*, alors* $Z = (X - \mu)/\sigma \sim \mathcal{N}(0, 1)$*.*

Noter que si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors on peut écrire $X = \mu + \sigma Z$, où $Z \sim \mathcal{N}(0, 1)$.

Valeurs de la fonction $\Phi(z)$

z	0	1	2	3	4	5	6	7	8	9
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56750	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84850	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92786	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169

Remarque : Une table plus detaillé se trouve dans le *Formulaire*. Vous pouvez également utiliser la fonction `pnorm` du logiciel `R` : $\Phi(z) =$ `pnorm(z)`.

Exemple 157. *Calculer*

$$\Pr(Z \le -1.86), \quad \Pr(Z \le 0.53), \quad \Pr(-1.86 < Z < 0.53), \quad z_{0.95}, \quad z_{0.025}, \quad z_{0.5}.$$

Exemples et calculs

Exemple 158. *La durée en minutes d'un cours de math est $\mathcal{N}(47, 4)$, mais devrait être de 45. Donner la probabilité que (a) le cours se termine tôt, (b) le cours se termine avec un retard de plus de 5 minutes.*

Exemple 159. *Montrer que l'espérance et la variance de $X \sim \mathcal{N}(\mu, \sigma^2)$ sont μ et σ^2, et trouver le p quantile de X.*

Exemple 160. *Calculer la fonction de répartition et la densité de $Y = |Z|$ et $W = Z^2$, où $Z \sim \mathcal{N}(0, 1)$.*

Exemple 161. *Si $X \sim \mathcal{N}(\mu, \sigma^2)$, montrer que*

$$M_X(t) = \exp(t\mu + t^2\sigma^2/2), \quad t \in \mathbb{R},$$

et en déduire que $\mathrm{E}(X) = \mu$, $\mathrm{var}(X) = \sigma^2$.

Approximation normale de la distribution binomiale

La loi normale est une loi centrale en probabilité, en partie car elle peut être utilisée pour approcher les probabilités des autres lois. Un des résultats de base est :

Théorème 162 (de Moivre–Laplace). *Soit* $X_n \sim B(n, p)$, *où* $0 < p < 1$, *posons*

$$\mu_n = \mathrm{E}(X_n) = np, \quad \sigma_n^2 = \mathrm{var}(X_n) = np(1-p),$$

et soit $Z \sim \mathcal{N}(0, 1)$. *Alors quand* $n \to \infty$,

$$\Pr\left(\frac{X_n - \mu_n}{\sigma_n} \leq z\right) \to \Phi(z), \ z \in \mathbb{R}; \textit{ c'est à dire que, } \frac{X_n - \mu_n}{\sigma_n} \xrightarrow{D} Z.$$

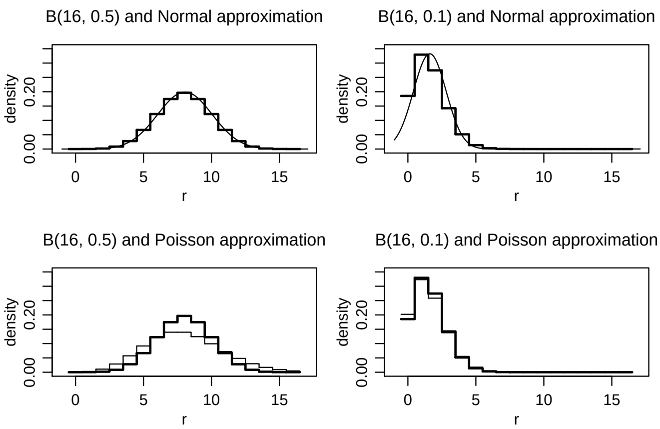
Ceci nous donne une approximation de la probabilité que $X_n \leq r$:

$$\Pr(X_n \leq r) = \Pr\left(\frac{X_n - \mu_n}{\sigma_n} \leq \frac{r - \mu_n}{\sigma_n}\right) \doteq \Phi\left(\frac{r - \mu_n}{\sigma_n}\right),$$

ce qui correspond à $X_n \dot{\sim} \mathcal{N}\{np, np(1-p\}$.
En pratique l'approximation est mauvaise quand $\min\{np, n(1-p)\} < 5$.

Approximations normale et poissoniènne de la binomiale

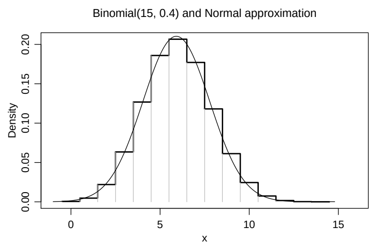
On a dêjà rencontré l'approximation poissoniènne de la loi binomial, valable pour grand n et petite p. L'approximation normale est valable pour n grand et $\min\{np, n(1-p)\} \geq 5$. A gauche : un cas où l'approximation normale est valable. A droite : un cas où l'approximation poissoniènne est valable.



Correction de continuité

Une meilleure approximation de $\Pr(X_n \le r)$ est donnée en remplaçant r par $r+\frac{1}{2}$; le $\frac{1}{2}$ est connu sous le nom de **correction de continuité** . Donc un meilleure approximation est

$$\Pr(X_n \le r) \doteq \Phi\left(\frac{r+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right).$$



Exemple 163. *Soit $X \sim B(15, 0.4)$. Calculer les valeurs exactes et approchées de $\Pr(X \le r)$ pour $r = 1, 8, 10$, avec et sans la correction de continuité. Commenter.*

Résultats numéRiques

```
pbinom(c(1,8,10),15,prob=0.4)
[1] 0.005172035 0.904952592 0.990652339

pnorm(c(1,8,10),mean=15*0.4,sd=sqrt(15*0.4*0.6))
[1] 0.004203997 0.854079727 0.982492509

pnorm(c(1,8,10)+0.5,mean=15*0.4,sd=sqrt(15*0.4*0.6))
[1] 0.008853033 0.906183835 0.991146967
```


Exemple

Exemple 164. *Le nombre total d'étudiants dans une classe est 100.*
(a) Chaque étudiant assiste indépendamment à un cours de mathématiques avec une probabilité de 0.6. Quelle est la taille de la plus petite salle de cours adaptée au nombre d'étudiants assistant aux cours, avec une probabilité 0.95 ?
(b) Il y a 14 cours par semestre, et les étudiants décident indépendamment d'assister à chaque cours. Quelle est maintenant la taille nécessaire de la plus petite salle de cours ?

Quantile-quantile (Q-Q) plots

Une manière pour comparer un échantillon $X_1, \ldots, X_n$ avec une loi théorique F :

– on ordonne les X_j, donnant
$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)},$$
puis on les graphe contre $F^{-1}\{1/(n+1)\}, F^{-1}\{2/(n+1)\} \ldots, F^{-1}\{n/(n+1)\}$.

– L'idée : dans un cas idéale $U_1, \ldots, U_n \sim U(0,1)$ devrait couper l'intervalle $(0,1)$ en $n+1$ sous-intervalles de largeurs $1/(n+1)$, donc on devrait grapher les U_j contre $1/(n+1)$, ..., $n/(n+1)$, et ainsi les $X_j \stackrel{D}{=} F^{-1}(U_j)$ contre $F^{-1}\{1/(n+1)\}$, ..., $F^{-1}\{n/(n+1)\}$;

– plus le graphe se rapproche d'une droite, plus les données ressemblent à un échantillon issu de F ;

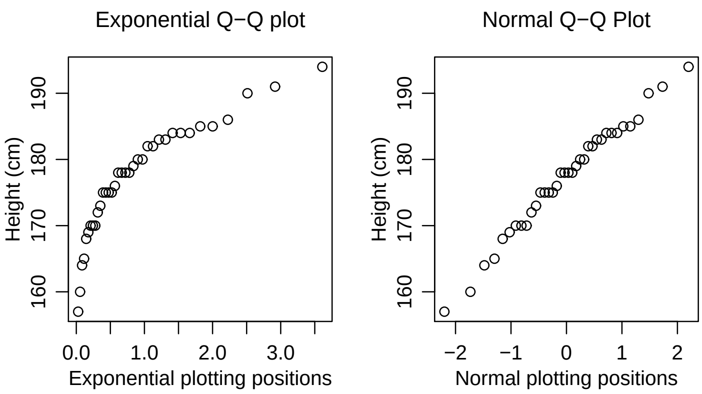
– le plus souvent on prend une version standard de F (e.g., $\exp(1)$, $\mathcal{N}(0,1)$), et alors les $F^{-1}(r/(n+1)\}$ s'appellent des **plotting positions** de F—alors la pente donne une estimation du paramètre de dispersion de la loi, et l'intercepte donne une estimation du paramètre de position ;

– pours les lois $\exp(1)$ et $\mathcal{N}(0,1)$ on a respectivement
$$F^{-1}\{j/(n+1\} = -\log\left(1 - \frac{j}{n+1}\right), \quad F^{-1}\{j/(n+1\} = \Phi^{-1}\left(\frac{j}{n+1}\right);$$

– il est difficile de tirer des conclusions fortes d'un tel graphique pour n petit, car la variabilité est alors grande—on a tendance à le sur-interpréter.

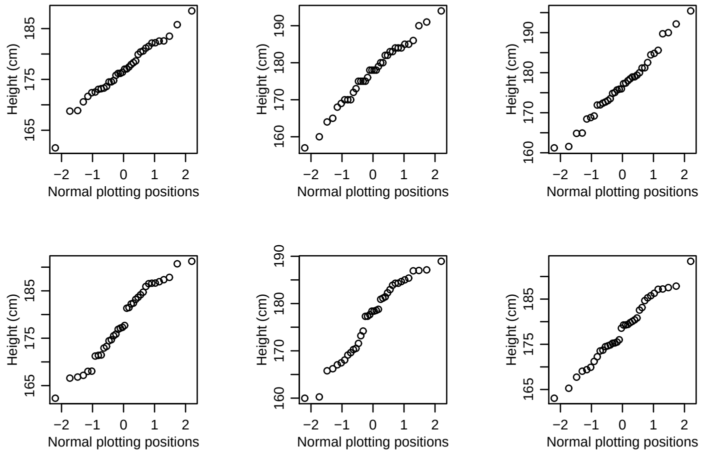
Hauteurs d'étudiants

Q-Q plots pour les hauteurs de $n = 36$ étudiants en SSC, pour les lois exponentielle et normale.



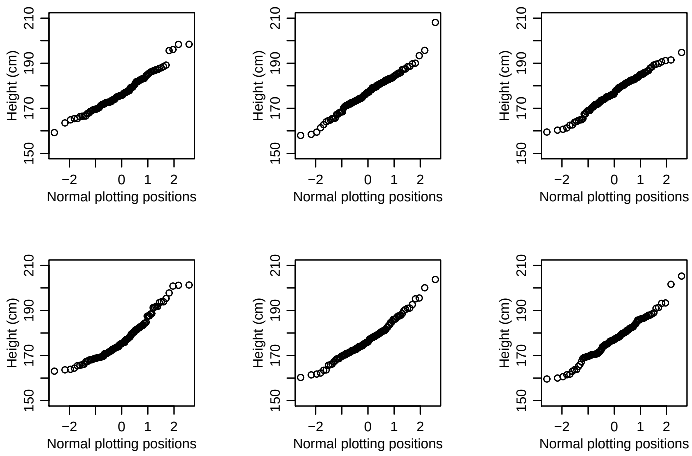
$n = 36$: Quel échantillon n'est pas normal ?

Il y a cinq échantillons de variables normales simulées, et des vrais données.



$n = 100$: Quel échantillon n'est pas normal ?

Il y a cinq échantillons de variables normales simulées, et un échantillon gamma simulé.



$n = 500$: Quel échantillon n'est pas normal ?

Il y a cinq échantillons de variables normales simulées, et un échantillon gamma simulé.