Random variables

Florence Perronnin

Univ. Grenoble Alpes, LIG, Inria

September 28, 2018

Outline

Variables aléatoires

Une variable aléatoire X est une application $X : \Omega \to F$ ou F est un ensemble ordonné, t.q.

$$\forall x \in F, \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$$

Human langage translation

A random variable is a function $X : \Omega \to \mathbb{R}$ (or $\mathbb{N}$) characterizing the possible outcomes in Ω such that intervals in $\mathbb{R}$ have a probability.

- X est une v.a. discrète si F est dénombrable (typ. $\mathbb{N}$)
- X est une v.a. continue si F est indénombrable (typ. $\mathbb{R}$)

Random variables are usually written as capital letters: X,. . .

Some examples of random variables

- $X =$result of a single die
- $Y =$sum of two dice
- (coin tossing) $X = 1$ if *heads*, $X = 0$ if *tails*
- $Z =$ lifetime of a device
- $A_n = n$-th measurement (collection of random variables)
- $\frac{1}{k}\sum_{n=1}^{k} A_n$ is also a random variable (mean)

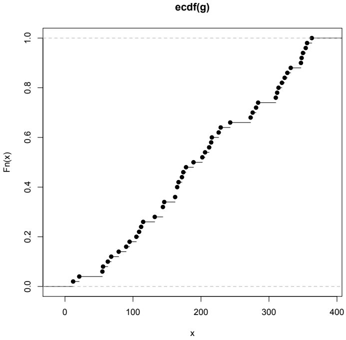
Cumulative Distribution function (C.D.F)

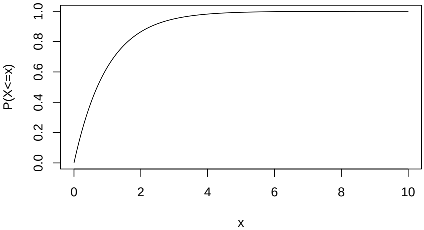
Pour toute v.a. X on définit sa fonction de répartition (CDF en anglais):

$$\forall x, \quad F_X(x) = \mathbb{P}[X \leq x]$$

F_X donne les probabilités *cumulées*.

F_X décrit la loi ou distribution de probabilité de X.





Properties:

- $0 \leq F_X(x) \leq 1 \qquad \forall x \in \mathbb{R}$
- F_X is nondecreasing

Cas particulier d'une variable aléatoire discrète

La fonction de répartition (C.D.F) F_X d'une variable aléatoire discrète X est une fonction en escalier (sommes cumulées).

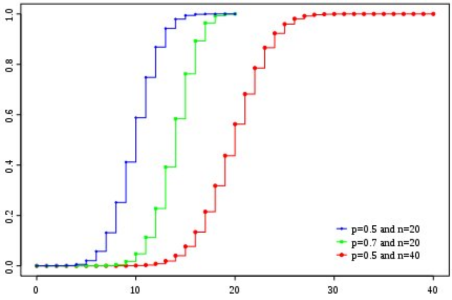


Figure : Example of CDF : Binom(n,p)

Quantiles

A quantile is the inverse of the cumulative distribution function.

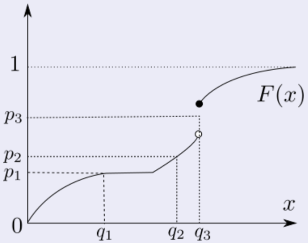
Median

The median is the value m such that $\mathbb{P}[X \leq m] = \frac{1}{2}$. (the value that splits Ω into two equiprobable parts.) It is therefore the 1/2-quantile or the 0.5-quantile.

q-quantile

The q-quantile is the value x_q such that :

$$F_X(x_q) = \mathbb{P}[X \leq x_q] = q$$



By Erzbischof - Own work, CC0,

https://commons.wikimedia.org/w/index.php?curid=18

Probability density function (P.D.F)

La distribution d'une variable aléatoire peut également être définie de façon plus "élémentaire" par :

Cas d'une variable discrète

Fonction de distribution :

$$f_X(k) = \mathbb{P}[X = k]$$

pour toute valeur k possible.

$$\sum_{n \in \mathbb{N}} f(n) = 1$$

Cas d'une variable continue (réelle)

Densité de probabilité (PDF en anglais) $f(x)$

$$\mathbb{P}[x \leq X \leq y] = \int_x^y f_X(u)du$$

attention aux bornes d'intégration...

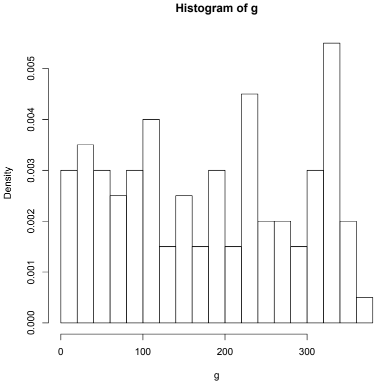
$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

PDF pour les variables aléatoires discrètes

Pour une v.a. discrète (à valeurs entières par exemple) on définit la fonction de distribution (ou loi) par

$$f_X(k) = \mathbb{P}[X = k]$$

. C'est la fonction qui donne la probabilité de chaque valeur.



- $0 \leq f_X(k) \leq 1$
- $F_X(k) = \sum_{m=-\infty}^{k} f_X(m)$
- f_X peut être approchée par l'histogramme d'un échantillon

PDF for continuous random variables

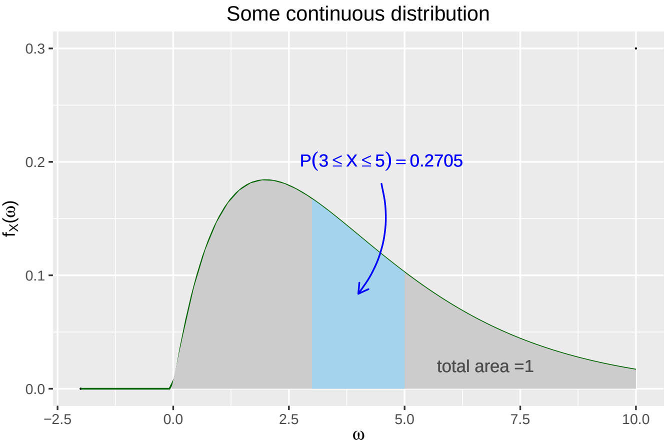
The probability density function f can be defined as the derivative of the CDF $F(x)$

$$F_X(x) - F_X(a) = \mathbb{P}[a \leq X \leq x] = \int_a^x f_X(u)du => f_X(x) = \frac{\partial F_X}{\partial x}(x)$$

Intuitively:

- The area below the curve of f between a and b gives the probability that the r.v. falls into the x-axis interval $[a, b]$
- The density is the derivative of F_X, so: $f_X(x)dx = dF_X(x)$

PDF and CDF

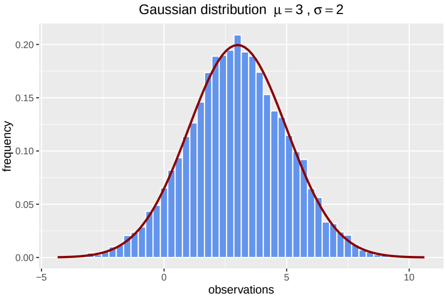


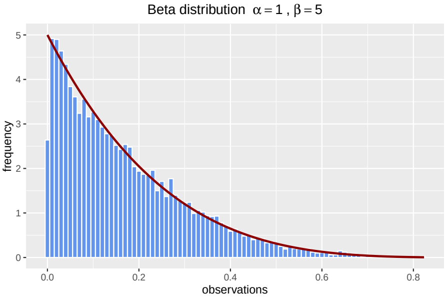
$$\int_{-\infty}^{\infty} f_X(u)du = 1$$

$$F_X(x) = \int_{-\infty}^{x} f_X(v)dv$$

Density and histograms

A probability density function f_X can also be viewed as the "limit of the histogram". Indeed, each "box" of the histogram gives $P(x_n \leq X \leq x_{n+1}) = F_X(x_{n+1}) - F_X(x_n)$.





Outline

Expected value

Espérance

La moyenne (mean) ou espérance (expectation en anglais) d'une v.a. discrète X à valeurs dans I (dénombrable) est:

$$\mathbb{E}[X] = \sum_{x \in I} x \mathbb{P}[X = x]$$

Example

Soit $0 < p < 1$. Une variable aléatoire X à valeurs dans $\{0, 1\}$ suit une loi de Bernoulli $\mathcal{B}(p)$ ssi $\mathbb{P}[X = 1] = p$ et $\mathbb{P}[x = 0] = 1 - p$.

Son espérance est $\mathbb{E}[X] = 0 \times (1 - p) + 1 \times p = p$.

Exercise

What is the expected value of the sum of two fair dice?

Expected value of a continuous r.v.

Expectation of a continuous random variable X

Let X be a continuous r.v. with values in $\mathbb{R}$. Then

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

(or equivalently : $\mathbb{E}[x] = \int_{-\infty}^{\infty} x dF(x)$)

Exercise

let Y be an exponential r.v. with parameter λ. The PDF is $f(x) = \lambda e^{-\lambda x}\mathbb{1}_{\{x\geq 0\}}$. Compute $\mathbb{E}[Y]$.

$$\begin{aligned} \mathbb{E}[Y] &= \int_0^\infty x f(x) dx & (1)\\ &= \int_0^\infty x\lambda e^{-\lambda x} dx & (2)\\ &= \left[x(-e^{-\lambda x})\right]_0^\infty - \int_0^\infty (-e^{-\lambda x}) dx & (3)\\ &= \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^\infty = \frac{1}{\lambda} & (4) \end{aligned}$$

Expectation of a function of X

Let X be an $\mathbb{R}$-valued r.v. and a function $g : \mathbb{R} \to \mathbb{R}$. Then

$$\mathbb{E}\left[g(x)\right] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Exercise

Let Y be a uniform r.v. over $[a, b]$ (with $0 < a < b$). Compute $\mathbb{E}\left[-log(Y)\right]$

Existence

If the r.v. x has a infinite number of possible values, the expected value is not always finite.

St Petersburg paradox

How much should the bank ask for the following game?

- flip a coin until it lands on *tails*.
- If *tails* happens at the
  - 1st round: reward=2$
  - 2nd round: reward=4$
  - 3nd round: reward=8$
  - and so on...

Let X be the reward and Y the number of coin tosses until *tails*.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{n=1}^{\infty} 2^n \mathbb{P}[Y=n] = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} \qquad (\mathrm{Y} \sim \mathsf{Geom}(1/2)) \\ &= (1+1+\ldots) = \infty\end{aligned}$$

However in reality no bank has an infinite amount of money. Let's assume the bank cannot pay more than $1,000,000\$$. Then if $Y \geq \log_2(10^6) \geq 19$, the player will only earns $10^6\$$ after 19 *heads* events (when he should earn more).

$$\begin{aligned}\mathbb{E}[X] &= \sum_{n=1}^{19} 2^n \mathbb{P}[Y=n] + 10^6 \mathbb{P}[Y>19] = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} + 10^6 \frac{1}{2^{19}} \\ &= 19 + \frac{10^6}{524288} = 20,9\$\end{aligned}$$

So if the bank can't pay more than a million \$, the expected reward of the player is only 20, 9\$. A lot less than infinity...

Existence

The example of the St Petersburg paradox shows that a finite random variable can have an infinite expected value.

Linearity of the Expectation operator

Linearity

Given two random variables X and Y we have:

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

even if X and Y are dependent variables.

Also for any constant c we have:

$$\mathbb{E}[cX] = c\mathbb{E}[X]$$

More generally, if we have n r.v. $X_1, X_2, \ldots X_n$ then:

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Exercise (using the linearity of $\mathbb{E}[.]$)

Initially n objects are correctly placed. But someone came by and rearranged all the objects. *"On average, 1 object is correctly placed."*

let X_i be the Bernoulli r.v. such that

- $X_i = 1$ of object i is correctly placed
- $X_i = 0$ otherwise

Let Y be the total number of correctly placed objects. Then clearly $Y = \sum_{i=1}^{n} X_i$.

Expectation exercises

- Compute the expectation of a binomial random variable Z.
- Let $Y = X - \mathbb{E}[X]$. Compute $\mathbb{E}[Y]$ as a function of the moments of X.

Conditional expectation

$$\mathbb{E}\left[Y|Z=z\right]=\sum_{y} y\mathbb{P}[Y=y|Z=z]$$

Law of total probability:

$$\mathbb{E}\left[Y\right]=\mathbb{E}\left[\mathbb{E}\left[Y|Z\right]\right]$$

where $\mathbb{E}\left[Y|Z\right]$ is a random variable ($f(Z)$).

Variance

Définition

La variance d'une v.a. est définie par :

$$var(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$$

Exercise : prove the above equality.

- Cas discret : $var(X) = \sum_{x \in I} x^2 \mathbb{P}[X = x] - \left(\sum_{x \in I} x \mathbb{P}[X = x]\right)^2$
- Cas continu : $var(X) = \int_{x \in I} x^2 f(x) dx - \left(\int_{x \in I} x f(x) dx\right)^2$

Question

Pourquoi ne pas définir $\mathbb{E}[(X - \mathbb{E}[X])]$ au lieu de l'espérance du carré des écarts?

Lois jointes

Deux variables aléatoires X et Y définies sur un même espace de probabilités ont une fonction de répartition jointe :

$$F(x,y) = \mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[(X \leq x) \cap (Y \leq y)]$$

Les distributions de X et Y respectives sont appelées lois marginales: $F_X(x) = \mathbb{P}[X \leq x] = \lim_{y \to +\infty} F(x,y)$.

La densité jointe se calcule comme précédemment : $f_{XY} = \dfrac{\partial F_{XY}}{\partial x \partial y}$ soit $\mathbb{P}[X \leq x, Y \leq y] = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v) du dv$

Densité conditionnelle : $f_{X|Y}(x,y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$

Variables aléatoires multiples

Espérance conditionnelle

Cas discret : $\mathbb{E}\left[X|Y=y\right] = \sum_{x \in I} x\mathbb{P}[X=x|Y=y]$.

Cas continu: $\mathbb{E}\left[X|Y=y\right] = \int_{x \in I} x f_{X|Y}(x,y)dx$

Loi de l'espérance conditionnelle: $\mathbb{E}\left[X\right] = \sum_{y \in J} \mathbb{E}\left[X|Y=y\right] \mathbb{P}[Y=y]$

Soit en continu : $\mathbb{E}\left[X\right] = \int_{\mathbb{R}} \mathbb{E}\left[X|Y=y\right] f_Y(y)dy$

Indépendance

Indépendance

Deux v.a. X et Y sont indépendantes ssi
$\mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[X \leq x]\,\mathbb{P}[Y \leq y]$.

Si deux v.a. *indépendantes* ont une espérance alors :
$\mathbb{E}[XY] = \mathbb{E}[Y]\,\mathbb{E}[Y]$ Attention : la réciproque n'est pas vraie.

Outline

Bernoulli trials

Bernoulli

A binary random variable $X \in \{0, 1\}$ such that $\mathbb{P}[X = 1] = p$ is a Bernoulli random variable $\mathcal{B}(p)$.

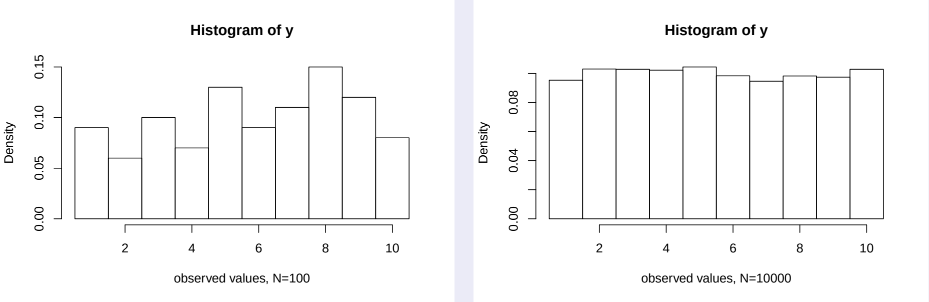
Typical example : coin tossing, test result (pass/fail)...

Uniform distribution

Une variable $X \in \{n_1, n_2, n_3, \ldots, n_k\}$ pouvant prendre k valeurs est uniforme si chacune de ces valeurs a la même probabilité de se réaliser :

$$\mathbb{P}[X = n_i] = \frac{1}{k}$$

```
y=sample.int(10,size=N,replace=TRUE)
hist(y,breaks=br,freq=FALSE)
```



Binomial distribution

Repeating bernoulli trials...

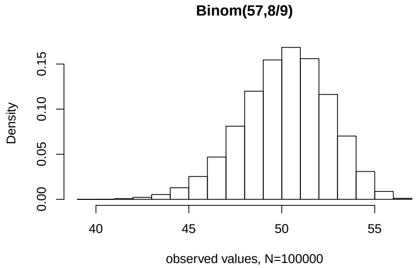
Binomial

La somme de n v.a. de Bernoulli indépendantes de même paramètre p est une v.a. binomiale $Binom(n, p)$. Sa distribution est:

$$\mathbb{P}[X_1 + ... + X_n = k] = C_n^k p^k (1-p)^{n-k}$$

Exercice :

1. Preuve de la distribution?
2. Montrer que $\sum_{i=0}^{n} \mathbb{P}[X = k] = 1$
3. Montrer que l'espérance est $\mathbb{E}[X] = np$



Geometric distribution (1)

Sequence of independent Bernoulli trials $\mathcal{B}(p)$ until first success:

$$\mathbb{P}[X = k] = p \times (1 - p)^{k-1}$$

Warning

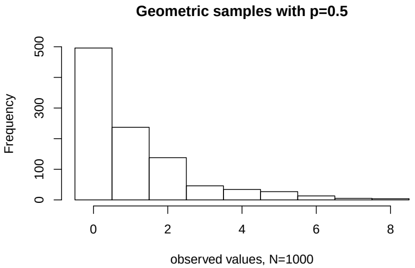
The geometric distribution in the R language (`rgeom`) is slightly different (number of failures before first success)

Exercice:

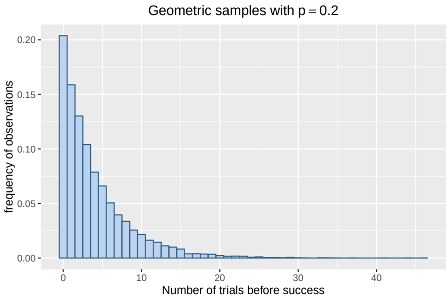
- Prove that $\sum_{i=0}^{\infty} \mathbb{P}[X = k] = 1$
- Prove that $\mathbb{E}[X] = \frac{1}{p}$

Geometric distribution (2)

Note that the geometric r.v. has infinite support, i.e. can take an infinite number of values.




```
y=
br=seq(from=min(y)-0.5,to=max(y)+0.5,by=1)
hist(y,xlab="observed values,
N=1000",main='Geom(0.5)',breaks=br)
```



Poisson distribution

Limite de la binomiale pour de grandes valeurs de n (et petites valeurs de p)

Paramètre: λ

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda x}$$

- Prove that $\mathbb{E}[X] = \lambda$
- Link with binomial ? (hint: use $\lambda = np$)

```
rpois(n,lambda)
```

Other classical distributions that you should know about

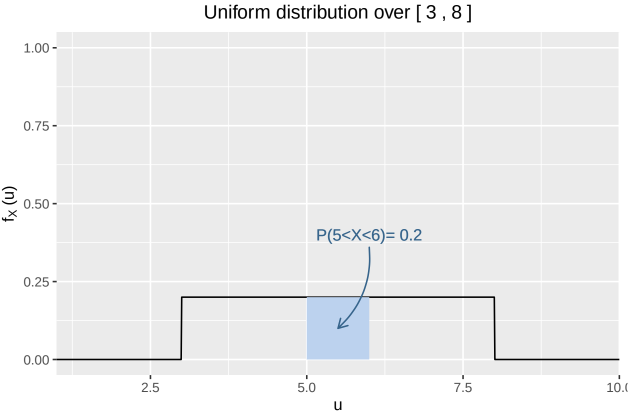
- Hypergeometric
- Negative binomial distribution
- Multinomial distribution

Outline

Uniform r.v

Let X be a continuous uniform random variable over an interval I.

- Why can't I be $\mathbb{R}$?

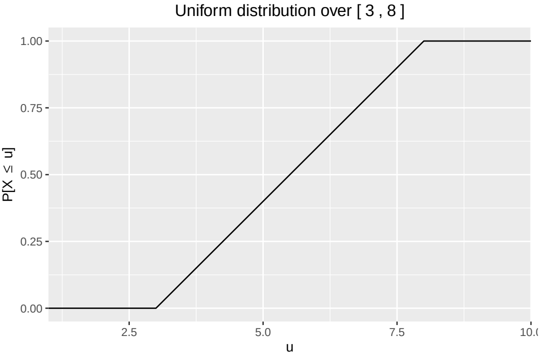


If $I = [a, b]$ then: $$f(x) = \frac{1}{b-a}\mathbb{1}_{\{x\in[a,b]\}}$$

Uniform distribution

Exercise

What is the CDF of a uniform r.v. U over $[a, b]$?

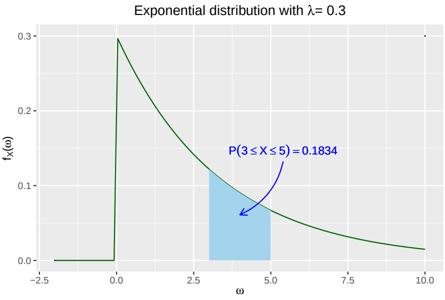


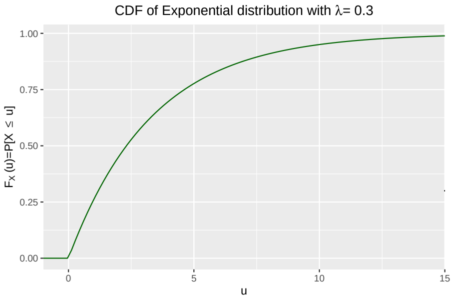
$$
\begin{aligned}
\mathbb{P}[U \leq x] &= \int_{-\infty}^{\infty} \frac{\mathbb{1}_{\{x \in [a,b]\}}}{b-a} dx \\
&= \int_{a}^{x} \frac{1}{b-a} dx \\
&= \left[\frac{x}{b-a} \right]_a^x \\
&= \frac{x-a}{b-a}
\end{aligned}
$$

Exponential r.v

Let X be a continuous exponential random variable over $\mathbb{R}+$ with parameter λ

$$
\begin{aligned}
f(x) &= \lambda e^{-\lambda x} \\
F(x) &= 1 - e^{-\lambda x}
\end{aligned}
$$





Playing with exponential distribution

Memoryless

Prove that an exponential r.v. has no memory:

$$\mathbb{P}[X > x + y | x > x] = \mathbb{P}[X > y]$$

The first event

Let X and Y be two exponential r.v.s with respective parameters λ and μ. (for instance, the time before failure of two devices). Let $Z = \min(X, Y)$ (the time of first failure). What is the distribution of Z?

Gaussian r.v $\mathcal{N}(\mu, \sigma)$

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