STATISTICS

Module 2 *of*
MATHEMATICS AND STATISTICS

Marina Santacroce

marina.santacroce@unibocconi.it
Department of Decision Sciences, Bocconi University

L9. Continuous random variables. Notable examples: uniform, exponential, gamma and normal.

Outline

We have studied discrete random variables, for which the set of possible values is *finite* or *countable*.

In many practical applications, we have to deal with random variables whose set of possible values is *uncountable*.

- *Waiting for a bus*. The time (in minutes) which elapses between arriving at a bus stop and a bus arriving can be modeled as a random variable X taking values in $[0, \infty)$.
- *Share price*. The values of one share of a specific stock at some given future time can be modeled as X taking values in $[0, \infty)$.
- 

  Weight. The weight of a randomly chosen individual can be modeled as a r.v. X taking values in $[0, 650)$ (in kilos)[1].

[1] Jon Brower Minnoch (1941-1983), U.S. citizen, weight 635 Kg (the largest ever documented)

Probability Density function

Consider a real random variable X taking values in a *continuous subset of* $\mathbb{R}$.

Def. We say that X is an *(absolutely) continuous* random variable if there exists a *nonnegative function*

$$f : \mathbb{R} \longrightarrow [0, +\infty) \qquad \text{such that} \qquad \int_{-\infty}^{+\infty} f(x)\, dx = 1.$$

Moreover, for all the sets $B \subset \mathbb{R}$ such that the integral is well defined

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(x)\, dx.$$

The function f is said *probability density function* of X.

Thus, $\boxed{F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)\, du}$ and $\boxed{f(x) = \frac{dF(x)}{dx}}$

For $B = [a, b]$, $\quad \mathbb{P}(a \leq X \leq b) = \int_a^b f(x)\, dx = F(b) - F(a)$

Rem. In the continuous case the probability $\mathbb{P}(a \leq X \leq b)$ corresponds to the area under the curve $f(x)$ between a and b.

Consider a small interval $(x, x + \Delta x]$, then we have

$$\mathbb{P}(x < X \leq x + \Delta x) = \int_x^{x+\Delta x} f(u)\, du \cong f(x)\Delta x$$

so $\mathbb{P}(x < X \leq x + \Delta x)$ is proportional to $f(x)$ for Δx small.

$\Rightarrow$ $f(x)$ is ***NOT*** *a* ***probability***

$\Rightarrow$ As $f(x)$ ***is NOT a probability***, *it is* ***NOT surprising*** *to have* $f(x) > 1$ for some values of $x \in \mathbb{R}$.

- For a continuous r.v., the probability of any single value is zero; i.e. for any $x \in \mathbb{R}$ we have $\mathbb{P}(X = x) = 0$.

Examples

1. Suppose that X is a continuous random variable whose probability density function is

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

   i) What is the value of C?
   ii) $\mathbb{P}(X > 1)$.

2. The lifetime of a certain battery has probability density function

$$f(x) = \begin{cases} \lambda e^{-x/100} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

   i) Find λ.
   ii) Compute $\mathbb{P}(X > 50)$.

Expected value for continuous random variables

Def. If X is a continuous random variable with probability density function f, then the expected value of X is

$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)\,dx$ (if the integral exists).

Rem X is said symmetric if $f(x) = f(-x)$. In this case $\Rightarrow \mathbb{E}(X) = 0$.

Prop. If X is non negative, then

$$\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > x)\,dx.$$

Indeed

$$\int_0^{+\infty} \mathbb{P}(X > x)\,dx = \int_0^{+\infty} \left(\int_x^{+\infty} f(y)\,dy \right) dx$$
$$= \int_0^{+\infty} \left(\int_0^{y} dx \right) f(y)\,dy = \int_0^{+\infty} yf(y)\,dy = \mathbb{E}(X)$$

$\mathbb{P}(X > x) = 1 - F(x) = R(x)$ *is known as survival function.*

Prop. If X is a continuous r.v. of density f and g is a real function, then

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx.$$

- (*linearity*) $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b \quad a,b \in \mathbb{R}$.

Recall that if X is a r.v. with $\mathbb{E}(X)$, then the *variance* of X is defined by

$$\mathrm{Var}(X) = \mathbb{E}[(X-E(X))^2].$$

▷ $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$;

- Se a e $b \in \mathbb{R} \Rightarrow \mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$.
- $\mathrm{Var}(X) \geq 0$;
- $\mathrm{Var}(X) = 0$ if and only if $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

- *Standardized random variable*

  Recall that the random variable Z defined by

$$Z = \frac{X-\mathbb{E}(X)}{\sigma_X},$$

  where $\mathbb{E}(X)$ and $\sigma_X = \sqrt{\mathrm{Var}(X)}$ are respectively the *expectation* and the *standard deviation* of X, has expected value zero and variance 1.

Example: Waiting times

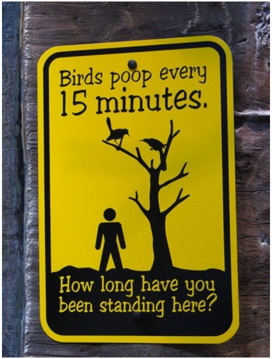
Suppose that the number of events occurring in a time interval $[0, t]$ follow the Poisson distribution with average λt. Show that the arrival time of the first event is a continuous r.v. with density

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

Moreover, find the distribution of the arrival time of the $n-th$ event.

Suppose 15 minutes is the average waiting time for the first event.

- which is the probability that waiting 5 minutes the event occurs?
- and that the event occurs within 10 minutes if nothing happened in the first 5?



Exponential random variable

Def. A continuous random variable X is said exponential of parameter λ, $\lambda > 0$, $X \sim$ exp(λ), if the density function is

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

- Verify that f is a density;
- $\boxed{\mathbb{E}(X) = \frac{1}{\lambda}}$; $\boxed{\mathrm{Var}\,(X) = \frac{1}{\lambda^2}}$
- verify $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$
  (property of *lack of memory* of the exponential r.v.);
- verify that the *hazard rate* defined by

$$\boxed{\lambda(x) = \frac{f(x)}{R(x)}}$$

  if $x > 0$, is constant and equal to λ.
- Show that the moment generating function is

$$\phi_X(t) = \frac{\lambda}{\lambda - t}, \qquad t < \lambda.$$

Gamma random variable

Def. A continuous random variable X is said gamma of parameters λ and α, with $\lambda, \alpha > 0$, $X \sim \text{gamma}(\alpha, \lambda)$, if its density function is

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The special function $\Gamma(\alpha)$ is defined as $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$.

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$;
- $\Gamma(1) = 1$;
- if $\alpha \in \mathbb{N}$ then $\Gamma(\alpha) = (\alpha - 1)!$

◇ Verify f is a density function;

◇ $\boxed{\mathbb{E}(X) = \frac{\alpha}{\lambda}}$ and $\boxed{\text{Var}\,(X) = \frac{\alpha}{\lambda^2}}$

◇ Show that the moment generating function is

$$\phi_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha, \qquad t < \lambda.$$

Uniform random variable U(0,1)

Def. A continuous random variable X is said to be uniform U(0,1) if the density

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

- $$\begin{cases} \mathbb{E}(X) = \int_0^1 x\,dx = 1/2 \\ \mathbb{E}(X^2) = \int_0^1 x^2\,dx = 1/3 \\ \mathrm{Var}(X) = 1/3 - (1/2)^2 = 1/12 \end{cases}$$
- The distribution function is $F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$
- The moment generating function is

$$\phi_X(t) = \frac{e^t - 1}{t}, \ t \in \mathbb{R}$$

Exercise: A stick of length 1 is split at a point U that is uniformly distributed over (0;1)). Determine the expected length of the piece containing the point p.

Gaussian or normal random variable

- A continuous random variable X is said normal or Gaussian, $X \sim \mathrm{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, if the density function is of the form

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.}$$

◇ $\boxed{\mathbb{E}(X) = \mu}$ and $\boxed{\mathrm{Var}(X) = \sigma^2}$.

- The moment generating function of X is

$$\phi_X(t) = \exp\left(t\mu + \frac{\sigma^2 t^2}{2}\right), \; t \in \mathbb{R}$$

If $\mu = 0$ and $\sigma^2 = 1$, the random variable is said standard normal and the density function is:

$$\boxed{F(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad z \in \mathbb{R}.}$$

- Let $X \sim \mathrm{N}(\mu, \sigma^2)$ and show that $Z = \frac{X-\mu}{\sigma}$ is a standard normal random variable.