Graph - Lezione 2 - 31/10/2019

sabato 2 novembre 2019 12:19

Lezione 2 -31/10/2019

Random variables,
We have seen the concept of probability space. Now we have in some way to represent the outcome of our experiment.

Definition: a real random variable (r. v)
Is a function X : Omega --> |R.
In general this mean that X is associating a number to outcome of our experiment.

It may happen that in reality the possible value that we consider is finite or countable.
Or
X: Omega --> |N U { 0 }
So only integers or a subset of |N. We consider also 0.

The outcome of the experiment maybe is not a number, but we can always associate 2 number to the two outcomes.

Example of events can be:
A = { X <= 4.7 } = { w app Omega such that X(w) <= 4.7 }
Taking to account that X is a function this notation mean this is the set of all the element of the probability space.

B { X is even integer } = { 0, 2 ,4 ,6 … }

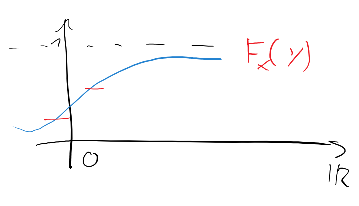
Def. We say that two random variable X and Y are independent
If any event Ax (related only with x and any related only with y) satisfy:
P(Ax inter By) = P(Ax) * P(By)
This is the definition of two independent events.

X1, X2, …, Xk are intendent random variable if called Ai an event related with Xi only, we have
A1, …, Ak are indepedent

Now we have to define what is a distribution of a r.v. We are never sure of the outcome of the experiment (they are random) what can make a congecture is the distribution of the random variable.

Def.
If X is a real r.v. then the distribution mux of X is the probability measure on |R satisfying:
Mu(Ax) = P(X app Ax) $\forall A_x$ related with X
[Mu of Ax is the probability that X belongs to AX for any Ax event related with X]
The distribution telling us the prob of any possible event relate with the r.v. can be realized.

Def.
The cumulative density function (cdf) of a r.v. X is the function
Fx (y) = P(X <= y ∀ y ∈ |R



Another Function is not always continuing

X: Omega --> S = { s1, s2, …, Sk}

μ= (μ1, …, μk)
μi = P(X = Si)

$$\sum_{i=1}^{k} \mu_i = 1$$

The prob of all possible outcome of an experiment, we just have to sums the probability we measure.
Discrete r.v.: assuming a finite or countable number of values

If X : Omega --> |R, X is a **CONTINUOUS RANDOM VARIABLE**
So this is the difference

The probability of all outcomes should sum up to 1.

Def.
A sequence X1, X2, X3, … of r.v is said to be independent identical distributed (IID) if:
1) They are independent
2) The CDF of the r.v is the same (Same distribution μ rappresented with the CDF)
   F_{xi} (y) = ^F(y) ∀y, ∀i

If they are IID we are in a good position because we have a lot of theorems.

Imagine X1, X2 … represent as a process {Xn} n app |N so n is a integer.
This sequence is called **STOCHASTIC PROCESSES** define by a continuing process.
{Xt} t app |R

When we observe the time series they are dependent but also not IID.

We will face the problem what is a stochastic Markov chain which a type of stochastic process. They are not dependent but conditional on the past.

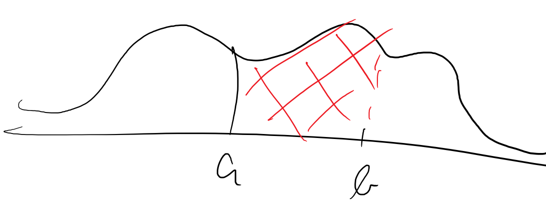
When we deal with finite r. v. we just list the probability of the outcome.
We say that X is a continuous r.v. not only if it takes value in |R but if exist a function fx(x)
Fx: |R --> [0, +infinity) such that the cumulative density fucntion of our r.v can be represent as the integral :

fx: Probabilty densisty function (PDF)

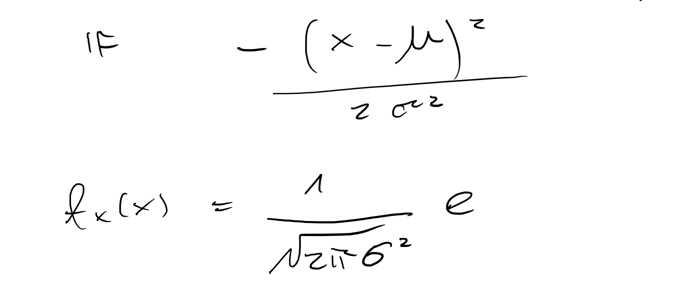
Ex.

PROB DENSITY FUNCTION

$$F_x(y) = \int_{-\infty}^{y} f_x(x)\,dx = P(X \le y)$$



$$P(X \in [a,b]) = F_x(b) - F_x(a) =$$
$$= P(X \le b) - P(X \le a) = \int_a^b f_x(x)\,dx$$

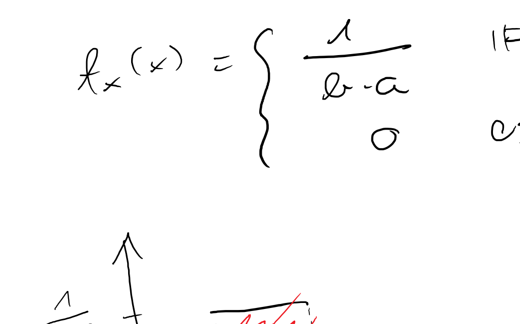


Ex.

a) X IS GAUSSIAN OR NORMAL (μ, σ^2)

IF

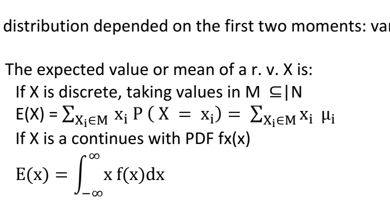
$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



b) X IS UNIFORMLY DISTRIBUTED IN [a, b]

IF ITS PDF IS

$$f_x(x) = \begin{cases} \frac{1}{b-a} & \text{IF } x \in [a,b] \\ 0 & \text{OTHERWISE} \end{cases}$$



$$\int_{-\infty}^{+\infty} f_x(x)\,dx = 1$$

ALWAYS 1 BY DEFINITION

Any distribution depended on the first two moments: variance an expected value.

Def. The expected value or mean of a r. v. X is:
If X is discrete, taking values in M ⊆|N
E(X) = $\sum_{X_i \in M} x_i P(X = x_i) = \sum_{X_i \in M} x_i \mu_i$
If X is a continues with PDF fx(x)
$E(x) = \int_{-\infty}^{\infty} x f(x) dx$

If we have X1,.., Xn iid with $E(X_i) = \mu$
We can estimate (approximate) μ with
$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} Xi$
Low of large numbers
If X1, X2, … is an iid sequence of r.v. having all the same mean (E(Xi) = μ and $\mu < +\infty$)
Then P($\bar{X}$n --> μ) = 1
n--> ∞
$\bar{X}$n --> μ almost surely

Def. The variance of a r.v X is defined as
Var(X) = E [(X - E(X))^2]
Then:
a) If X is discrete, taking values in M ⊆ N
$$Var(X) = \sum_{X_i \in M} (x_i - E(X))^2 P(X = xi)$$
b) if X is continuos with PDF fx(x)

μ_i ≥ 0

$$Var(x) = \int_{-\infty}^{\infty} (x - E(x))^2 f_x(x)\,dx \ge 0$$

X1, .. Xn iid with Var (Xi = sigma^2
Estimator of sigma^2: sample variance (unbiased)

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x}_n)^2$$

$$E(S_n^2) = \sigma^2$$

WHEN TRUE σ^2 UNBIASED
CORRECTNE

$$\tilde{S}_n^2 = \frac{1}{n}\sum (x_i - \bar{x}_n)^2 + E(\tilde{S}_n^2) = \frac{n-1}{n}\sigma^2$$

BIASED BUT ASYMPTOTICALLY UNBIASED

ALSO S_n^2 IS CONSISTENT FOR σ^2

$$S_n^2 \xrightarrow[n \to \infty]{} \sigma^2 \quad \text{ALMOST SURELY}$$

E AND σ ARE OPERATORS

PROPERTIES OF E(x)

E(·) IS LINEAR, THAT IS

a) $E(x_1, \dots, x_k) = E(x_1) + \dots + E(x_k)$

b) $E(cX) = c \cdot E(x)$

CONSTANT NOT RANDOM

$$\Rightarrow E(c_1 x_1 + c_2 x_2 + \dots c_k x_k) = c_1 E(x_1) + \dots + c_k E(x_k)$$

$X_1, \dots X_n$ IID WITH $E(x_i) = \mu$

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = E\left(\frac{1}{n}x_1 + \frac{1}{n}x_2 + \dots + \frac{1}{n}x_n\right) =$$
$$= \frac{1}{n}\underbrace{E(x_1)}_{\mu} + \dots + \underbrace{\frac{1}{n}E(x_n)}_{\mu} = \frac{1}{n}\cdot n\mu = \mu$$

PROPERTIES OF VAR(x)

a) $Var(cX) = c^2 Var(x)$

b) $Var(x_1 + \dots + x_k) = Var(x_1) + \dots + Var(x_n)$

ONLY IF $x_1 \dots x_n$ ARE INDEPENDENT

EXAMPLE 1 : BERNOULLI

$X: \Omega \to \{0, 1\}$

$$X = \begin{cases} 0 & 1-p = P(X=0) \quad \text{UNSUCCESS} \\ 1 & p = P(X=1) \quad \text{SUCCESS} \end{cases}$$

$$\mu = (\mu_0, \mu_1) = (1-p, p)$$

IT'S A PROB.
[0,1]

$$E(X) = 0\,(1-p) + 1\,p = p$$

E(x) · P(x=0)

$$Var(X) = (0-p)^2(1-p) + (1-p)^2 p = p^2(1-p) + p(1-p)^2 =$$
$$= p(1-p)[\cancel{p} + 1 - \cancel{p}] = p(1-p)$$

$Y = x_1 + \dots + x_n \quad x_i \sim \text{BERNOULLI}(p)$ INDEP.

$Y \sim \text{BERNOULLI}(n, p)$

$$E(Y) = E(x_1) + \dots + E(x_n) = np$$

$$VAR(Y) = Var(x_1) + \dots + Var(x_n) = np(1-p)$$

x_i IND.