Math 4501 - Probability and Statistics II

Preliminary material

(review from MATH 3501)

Random variable

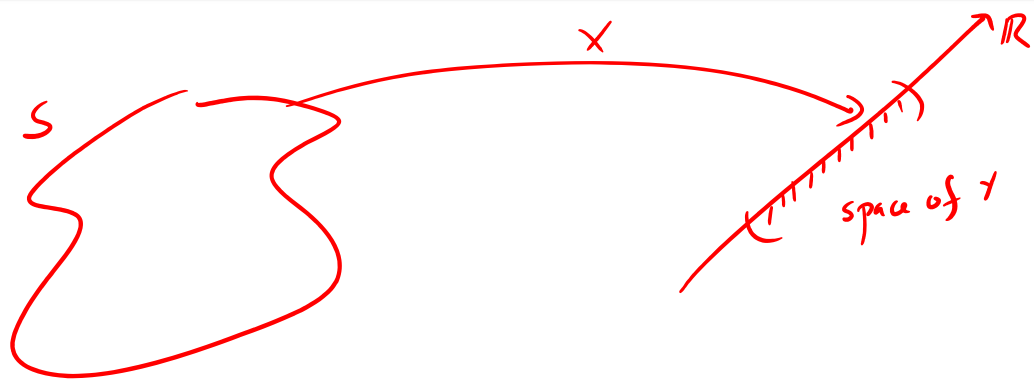
Definition

Given a random experiment with an outcome space S, a *random variable* is a real-valued function with domain S.

The *space of* X is the set of real numbers

$$\{x : X(s) = x, s \in S\} \ ,$$

where $s \in S$ means that the element s belongs to the set S.



Discrete random variable

Definition

An outcome space S is said to be *discrete* if it contains a countable number of points, that is, either:

i) S contains a finite number of points; or

ii) the points of S can be put into a one-to-one correspondence with $\mathbb{N}$.

Definition

A random variable defined on a discrete outcome space S is called a *discrete random variable*.

The probability distribution of a discrete random variable is said to be of the discrete type.

Note: A discrete random variable can assume at most a countable number of values.

Probability mass function

Definition

Let X be a discrete random variable with space S.

The *probability mass function* (pmf) of X , denoted $f(x)$ or $f_X(x)$, is given by

$$f(x) = P(X = x) \ , \qquad x \in \mathbb{R} \ .$$

Properties:

(a) $f(x) \geq 0$ for $x \in S$ and $f(x) = 0$ for $x \notin S$, that is S is the support of f

(b) $\sum_{x \in S} f(x) = 1$

(c) $P(X \in A) = \sum_{x \in A} f(x)$, where $A \subset R$

Cumulative distribution function

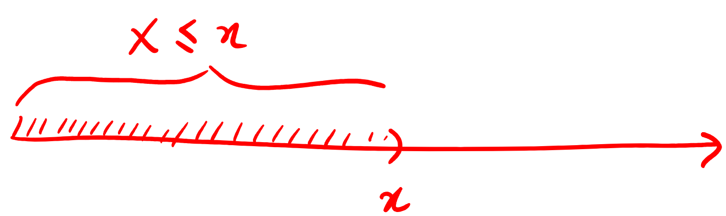
Definition

Let X be a discrete random variable with space S.

The *cumulative distribution function* (cdf) of X, denoted $F(x)$ or $F_X(x)$, is defined as

$$F(x) = P(X \leq x) , \qquad x \in \mathbb{R} .$$

Note: The cdf of X is sometimes referred to as the distribution function of the random variable X.



Random variable of continuous type

Definition

A random variable X is said to be of the *continuous type* if its space S is an interval or union of intervals and there exists an integrable function $f(x)$, called the *probability density function* (pdf) of X, for which the following conditions hold:

(a) $f(x) \geq 0$ if $x \in S$

(b) $\displaystyle\int_S f(x)\,\mathrm{d}x = 1$

(c) If $(a, b) \subseteq S$, then the probability of the event $\{a < X < b\}$ is given by

$$P(a < X < b) = \int_a^b f(x)\,\mathrm{d}x\ .$$

The distribution of probability of X is also said to be of the continuous type.

Cumulative distribution function

The cumulative distribution function (cdf) or distribution function of a random variable X of the continuous type, is related with the pdf of X via

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t\,, \quad -\infty < x < \infty$$

pdf of X

Note: From the fundamental theorem of calculus, we have that

$$\boxed{F'(x) = f(x)}$$

for values of x for which the derivative $F'(x)$ exists.

Mathematical Expectation

Definition

(VERY IMPORTANT !)

Let X be a random variable (discrete or continuous) with space S and pmf or pdf $f(x)$, and let u be a real-valued function of a single real variable.

The mathematical expectation or expected value of $u(X)$, denoted $E[u(X)]$, is given by

$$E[u(X)] = \begin{cases} \displaystyle\sum_{x \in S} u(x) f(x) \ , & \text{if } X \text{ is discrete} \\ \displaystyle\int_S u(x) f(x) \, \mathrm{d}x \ , & \text{if } X \text{ is continous} \end{cases}$$

provided the corresponding sum or integral is absolutely convergent.

Interpretation: $E[u(X)]$ may be regarded has a weighted mean of $u(x)$, $x \in S$, where the weights are given by the pmf/pdf $f(x)$, $x \in S$.

Properties of Mathematical Expectation

Theorem (VERY IMPORTANT!)

When it exists, the mathematical expectation satisfies the following properties:

(a) If c is a constant, then $E(c) = c$

(b) If c is a constant and u is a function, then

$$E[c \cdot u(X)] = c \cdot E[u(X)]$$

(c) If c_1 and c_2 are constants and u_1 and u_2 are functions, then

$$E[c_1 u_1(X) + c_2 u_2(X)] = c_1 E[u_1(X)] + c_2 E[u_2(X)]$$

linearity

Note: Property (c) can be extended to more than two terms

linearity →

$$E\left[\sum_{i=1}^{k} c_i u_i(X)\right] = \sum_{i=1}^{k} c_i E[u_i(X)]$$

IMPORTANT

Mean

Let X be a random variable with space S and pmf or pdf $f(x)$.

Definition (IMPORTANT !)

The mean of the random variable X (or of its distribution) is

$$\mu = E(X) \ ,$$

$= \begin{cases} \sum_{x \in S} x \cdot f(x) \ , & \text{if } X \text{ is discrete} \\ \int_S x \cdot f(x)\, dx \ , & \text{if } X \text{ is continuous} \end{cases}$

provided the expectation exists.

Notes:

1) the mean μ is also referred to as the *first moment of X about the origin*.

2) the first moment about the mean is always zero:

$$\underbrace{E[(X-\mu)] = E(X)}_{\text{linearity}} - E(\mu) = \mu - \mu = 0 \ .$$

Variance and standard deviation

Definition

The *variance of the random variable X* (or of its distribution), denoted Var(X), is the second moment of X about the mean:

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

whenever the expectation exists.

The positive square root of the variance is called the *standard deviation of X* and is denoted by the Greek letter σ (sigma)

$$\sigma = \sqrt{\mathrm{Var}(X)}\ .$$

Property:

$$\mathrm{Var}(X) = E\left(X^2\right) - (E(X))^2 = E\left(X^2\right) - \mu^2\ .$$

Efficient way to compute variance!

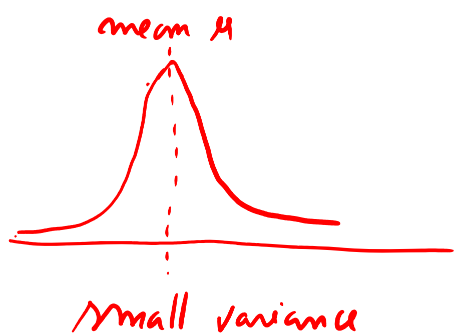
Interpretation

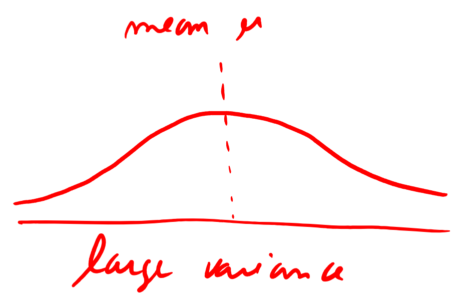
The mean $\mu = E(X)$ provides a measure of location:

- it gives the middle or center of the distribution of X relative to the pmf $f(x)$

Variance and the standard deviation are measures of dispersion or spread:

- they indicate how much the points in S spread out around the mean μ relative to the pmf $f(x)$





Higher order moments

Definition

Let r be a positive integer.

The *rth moment of X (or of its distribution) about the origin* is given by If

$$E(X^r) \ ,$$

provided the expectation exists.

Given $b \in \mathbb{R}$, the *rth moment of X (or of its distribution) about b* is given by

$$E[(X-b)^r] \ ,$$

provided the expectation exists.

very often we take $b = \mu = E[X]$, e.g.: for variance $Var(X) = E[(X-\mu)^2]$

$b = \mu = E[X]$

Moment generating function

very useful!!!

Definition

Let X be a random variable with space S and pmf/pdf $f(x)$.

If there is a positive number h such that the expectation $E\left(e^{tX}\right)$ exists and is finite for $-h < t < h$, then the function defined by

$$M(t) = E\left(e^{tX}\right)$$

function of t

is called the *moment-generating function of* X (or of the distribution of X).

We often use the abbreviation mgf.

Property: If the mgf exists, then for each positive integer r, we have

$$\boxed{M^{(r)}(0) = E\left(X^r\right) .}$$

Math 4501 - Probability and Statistics II

5.3 - Several random variables (review)

from Math 3501

Independent random variables

ONE OF THE KEY ASSUMPTIONS THROUGHOUT THIS COURSE

Definition

We say that that n random variables $X_1, X_2, \ldots, X_n$ are *independent* if

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \ldots P(X_n \in A_n)$$

for all subsets $A_1, A_2, \ldots, A_n$ of $\mathbb{R}$.

Consequence

Note: If $X_1, X_2, \ldots, X_n$ are independent random variables, the independence condition above is equivalent to

$$\boxed{F(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \ldots F_{X_n}(x_n) \ , \quad \text{for all } x_1, x_2, \ldots, x_n \in \mathbb{R}}$$

and to

$$\boxed{f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \ldots f_{X_n}(x_n) \ , \quad \text{for all } x_1, x_2, \ldots, x_n \in \mathbb{R} \ ,}$$

joint cdf/pdf/pmf equals product of marginals

where F and f denote, respectively the joint cdf and joint pmf/pdf of $X_1, X_2, \ldots, X_n$, while F_{X_i} and f_{X_i}, $i = 1, \ldots, n$, denote the corresponding marginal cfd and marginal pmf/pdf associated with X_i.

Random sample

same random experiment
repeated under the same conditions

sampling is always from the same "population"

Definition

We say that that n random variables $X_1, X_2, \ldots, X_n$ are *identically distributed* if all have the same distribution.

SECOND KEY ASSUMPTION FOR OUR COURSE

Definition (Alternative terminology: $X_1, X_2, \ldots, X_n$ iid r.v)

A collection of n independent and identically distributed random variables $X_1, X_2, \ldots, X_n$ is said to be a *random sample of size n*.

Note: The values of a random sample of size n may be regarded as measurements on the outcomes of n identical and independent random experiments.

The joint cfd and pmf/pdf of $X_1, X_2, \ldots, X_n$ are given by

joint cdf $$F(x_1)F(x_2)\ldots F(x_n) \quad \text{and} \quad f(x_1)f(x_2)\ldots f(x_n)$$ joint pdf/pmf

where F and f are, respectively, the marginal cfd and pmf/pdf of $X_1, X_2, \ldots, X_n$.

Example

Let $X_1, X_2, \ldots, X_{10}$ be 10 independent random variables whose values correspond to the respective outcomes of 10 rolls of a fair die.

The sequence $X_1, X_2, \ldots, X_{10}$ is a random sample of size $n = 10$ from a discrete uniform distribution over the integers $1, 2, \ldots, 6$.

outcome is a sequence of 10 numbers $(x_1, x_2, \ldots, x_{10})$

where each $x_i \in \{1, 2, \ldots, 6\}$

Example

Let X_1, X_2, X_3 be a random sample from a distribution with pdf

$$f(x) = e^{-x}, \quad x > 0 .$$

X_1, X_2, X_3 all have this distribution

Determine:

a) The joint pdf of these three random variables.

b) The probability $P(0 < X_1 < 1,\ 2 < X_2 < 4,\ 3 < X_3 < 7)$.

a) Let us denote $f_{joint}(x_1, x_2, x_3)$ the joint pdf of X_1, X_2, X_3

Since X_1, X_2, X_3 is a random sample (i.e. X_1, X_2, X_3 are iid)

$$f_{joint}(x_1, x_2, x_3) \underset{\text{independence}}{=} f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot f_{X_3}(x_3) \overset{\text{identically distributed}}{=} f(x_1) \cdot f(x_2) \cdot f(x_3)$$

$$= e^{-x_1} \cdot e^{-x_2} \cdot e^{-x_3} \quad \text{for } x_1, x_2, x_3 > 0$$

b) To evaluate the probability:

$$P(0 < X_1 < 1,\ 2 < X_2 < 4,\ 3 < X_3 < 7) =$$

$$= \int_3^7 \int_2^4 \int_0^1 f_{joint}(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3 =$$

$$= \int_3^7 \int_2^4 \int_0^1 e^{-x_1} \cdot e^{-x_2} \cdot e^{-x_3}\, dx_1\, dx_2\, dx_3$$

integrand function is product of a function of x_1 alone by a function of x_2 alone by a function of x_3 alone

$$= \left(\int_0^1 e^{-x_1}\, dx_1\right) \cdot \left(\int_2^4 e^{-x_2}\, dx_2\right) \cdot \int_3^7 e^{-x_3}\, dx_3) = \ldots$$

complete evaluate

Alternative approach.

Use independence right from the start:

independence

$$P(0 < X_1 < 1,\ 2 < X_2 < 4,\ 3 < X_3 < 7) =$$

$$= P(0 < X_1 < 1) \cdot P(2 < X_2 < 4) \cdot P(3 < X_3 < 7)$$

$$= \left(\int_0^1 e^{-x_1} \, dx_1 \right) \cdot \left(\int_2^4 e^{-x_2} \, dx_2 \right) \cdot \left(\int_3^7 e^{-x_3} \, dx_3 \right)$$

using the marginal pdf!