MA368

08/25/2020

Lecture 01

A discrete random variable (rv) X has probability mass function (PMF) given by $p(x)$,

$$p(x) := P(X = x) \text{ and the r.v is denoted } X \sim p(x)$$

where x is the "realized value" X, x

The cumulative distribution function (CDF) is

$$F(x) := P(X \le x)$$

define to be

and complementary CDF or "survival function" is

$$S(x) := P(X > x) = 1 - F(x)$$

This rv has "support" given by

$$Supp[X] := \{x : p(x) > 0, x \in \mathbb{R}\}$$

$$|Supp[X]| \le |\mathbb{N}| \text{ countably infinite at most.}$$

ie. finite or at most countably infinite.

elts in a set

Sets of this size are called "discrete" sets.

The support and the PMF are related by the following identity:

$$\sum_{x \in Supp[X]} p(x) = 1$$

Fall into pieces and regroup as a whole.

The most "fundamental" rv is the Bernouli:

$$X \sim \text{Bern}(p) := \underbrace{p^x (1-p)^{1-x}}_{p(x)} \text{ with Supp}[X] = \{0,1\}$$

Not in the Supp.

$p(7) = p^7 (1-p)^{-6}$ what should this be? Impossible. Should be 0 but it is not gonna happen.

Let's define the "indicator function"

$$\mathbb{1}_A := \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases} \Rightarrow X \sim \text{Bern}(p) := \underbrace{\overbrace{p^x (1-p)^{1-x}}^{p^{\text{old}}(x)} \mathbb{1}_{x \in \{0,1\}}}_{p(x)}.$$

$$\Rightarrow \sum_{x \in \mathbb{R}} p(x) = 1$$

this identity is 1 for all real #s.

What if $p = 1$? $= 1^x (1-1)^{1-x}$

$X \sim \text{Bern}(1) = 1^x 0^{1-x} \mathbb{1}_{x \in \{0,1\}}$.

$= \{1 \text{ w. prob. } 1$

$= \mathbb{1}_{x=1}$

$p(0) = 1^0 0^1 1 = 0$

$p(1) = 1^1 0^0 1 = 1$

plug in anyelse $\rightarrow$ back to 0.

$p(7) = 1^7 (1-1)^{1-7} = 0$

$X \sim \text{Deg}(1) = \{1 \text{ w. prob. } 1$

$X \sim \text{Deg}(c) := \mathbb{1}_{x=c}$

$= \{c \text{ w.p } 1$

$X \sim \text{Bern}(0) = \text{Deg}(0) = \{0 \text{ w.p } 1$

This is called a "degenerate" r.v. $X \sim \text{Deg}(1) = \{1 \text{ w.p. } 1$

broken down, buzzed

It contradicted itself.

The convention in this class is that parameter values (p is the parameter of the Bernoulli) that yield degenerate rv's are not part of the legal "parameter space":

$$\underbrace{p \in (0,1)}_{\text{parameter space of the Bernoulli}}$$ anything betw 0 and 1 but not including 0 and 1.

If we have more than one rv $X_1, X_2, \dots X_n$

we can group them together in a column vector:

$$\vec{X} := [X_1 \ X_2 \ \dots \ X_n]^T$$

and then define the "joint mass function" (JMF) as

$$P_{\vec{X}}(\vec{X}) = P_{X_1, \dots, X_n}(X_1, \dots, X_n) \text{ valid for } \vec{X} \in \mathbb{R}^n$$

and $\sum_{\vec{X} \in \mathbb{R}^n} p(\vec{X}) = 1$ Sum of the entire space.

If $X_1, X_2, \dots, X_n$ are independent rv's, then the JMF can be factored as

$$P_{\vec{X}}(\vec{X}) = P_{X_1}(X_1)\, P_{X_2}(X_2) \cdot \dots \cdot P_{X_n}(X_n)$$

$$= \prod_{i=1}^{n} P_{X_i}(X_i)$$ "Multiplication Rule"

If $X_1, X_2, \dots, X_n$ are identically distributed,

If $X_1 \overset{d}{=} X_2 \overset{d}{=} \dots \overset{d}{=} X_n$, this means "equal in distribution" meaning that their PMF's are the same. However, this offers no simplification of the JMF unless —

$$\forall_x \; p_{X_1}(x) = p_{X_2}(x) = \dots = p_{X_n}(x).$$

$X_1, X_2, \dots X_n$ iid (independent and identically distributed)

$$\Rightarrow p_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} p(x_i)$$

← shared PMF

use the multiplication rule.

Let $X_1, X_2 \overset{iid}{\sim} \text{Bern}(p)$, let $T_2 = f(X_1, X_2) = X_1 + X_2 \sim ?$

$\sim p_{T_2}(t)$

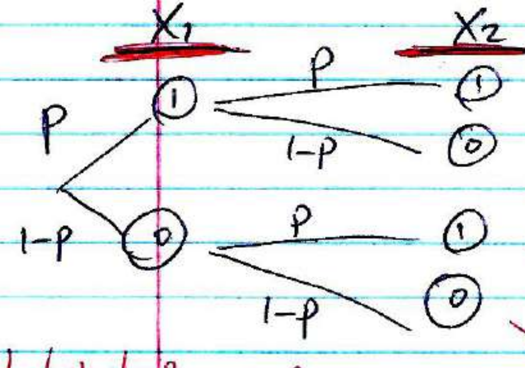
PMF the first. PMF the second

Denoted $p_{T_2}(t) = p_{X_1}(x) * p_{X_2}(x)$

↑ Convolution Operation.

CASES: 0+0, 0+1, 1+0, 1+1

$\text{Supp}[T_2] = \{0, 1, 2\}$ can't be 3, neg. fractional.



$p_{X_1, X_2}(x_1, x_2)$ ← JMF

p^2 (prob of ① & ① happen at the same time

$p(1-p)$ (" ① & ⓪ ")

$(1-p)p$ (" ⓪ & ① ")

$(1-p)^2$ (" ⓪ & ⓪ ").

Sum up = 1

Prob don't change b/c. they're iid events.