Lecture 01
08/26/2020
Math 621 : Prof. Adam Kapelner

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

$p(x) = P(X = x)$ and the rv is denoted $X \sim p(x)$

where x is the "realized value"

and cumulative distribution function (CDF) denoted $F(x)$: $F(x) := P(X \leq x)$
and complementary CDF also called survival function: $S(x) = P(X > x) = 1 - F(x)$

The rv has support
$Supp[X] = \{x : p(x) > 0, x \in \mathbb{R}\}$ and
$|Supp[X]| \leq |\mathbb{N}|$ i.e. finite or atmost countably infinite
↑
elements in sets

sets of this size are called "discrete"

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

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

The most "fundamental" rv is the Bernoulli

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

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

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

$P(7) = p^7(1-p)^{-6}$

To fix this, we introduce the "indicator function"

$$\mathbb{1}_A := \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases}$$

$$X \sim \text{Bern}(p) := \underbrace{\overbrace{p^x(1-p)^{1-x}}^{P^{old}(x)} \mathbb{1}_{x \in \{0,1\}}}_{P(x)}$$

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

What if $p=1$? $X \sim \text{Bern}(1) = 1^x \underbrace{(1-1)}_{0}^{1-x} \mathbb{1}_{x \in \{0,1\}}$

$P(0) = 1^0 0^1 1 = 0$ $\quad = \{1, P, 1 = \mathbb{1}_{x=1}$

$P(1) = 1^1 0^0 1 = 1$

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

$X \sim \text{Bern}(0) = \text{Deg}(0) = \{0$ with prob 1

Generally, $X \sim \text{Deg}(c) = \{c$ w/ prob 1
$= \mathbb{1}_{x=c}$

p is a "parameter" of the Bernoulli r.v.
What values of p are legal,
i.e, non-degenerate?

$p \in (0,1) \leftarrow$
the parameter space of the Bernoulli

If we have more than one
rv $X_1, X_2, \ldots, X_n$, we can group them
together in a column vector

$$\vec{X} = [X, X_2, \ldots, X_n]^T$$

which has a "joint mas function" (JMF)
defined as

$$p_{\vec{X}}(\vec{x}) = p_{X_1, \ldots X_n}(x_1, \ldots, x_n)$$

set,

$$\sum_{\vec{x} \in \mathbb{R}^n} p_{\vec{X}}(\vec{x}) = 1.$$

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

$$p_{\vec{X}}(\vec{x}) = p_{X_1}(x_1) \, p_{X_2}(x_2) \cdots p_{X_n}(x_n)$$
$$= \prod_{i=1}^{n} p_{X_i}(x_i)$$ the "multiplication rule"

If $X_1, \ldots, X_n$ are identically distributed
denoted as

$$X_1 \overset{d}{=} X_2 \overset{d}{=} \cdots \overset{d}{=} X_n,$$ then

$p_{X_1}(x) = p_{X_2}(x) = \cdots = p_{X_n}(x) \quad \forall x$ but this
offers no simplification of he
JMF unless . . .

$X_1, \dots, X_n \overset{iid}{\sim}$ denotes
"independent & identically distributed"

$$p_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} p(x_i) \leftarrow \text{shared PMF}$$

———— o ————

$0+0$
$0+1$
$1+1$

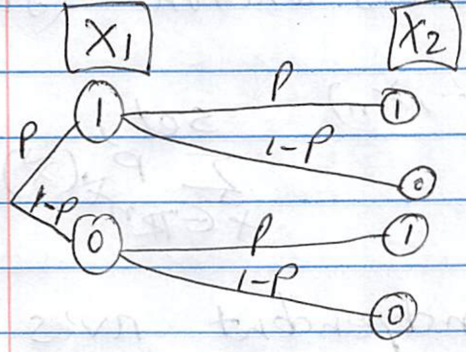
$Supp[T_2] = \{0, 1, 2\}$

$X_1, X_2 \overset{iid}{\sim} Bern(p)$

$$T_2 : f(X_1, X_2) = X_1 + X_2$$

$$p_{T_2}(t) = p_{X_1}(x_1) * p_{X_2}(x_2)$$

↑
Convolution Operation



$\boxed{p_{X_1, X_2}(x_1, x_2)}$

p^2

$p(1-p)$

$(1-p)p$

$(1-p)^2$

1

$$p^2 + 2p(1-p) + (1-p)^2 = \left((p) + (1-p)\right)^2 = 1^2 = 1 \checkmark$$