Lec 1

- Collect about 30 random variables ex: bernoulli, etc

A discrete random Variable (r.v.) has probability mass function (PMF)

which is defined as:

$p(x) := P(X = x)$ and the r.v. $X \sim p(x)$ where x is the "realized value"

$\underset{\text{r.v.}}{X}$, $\underset{\text{state it starts}}{x}$

The Cumulative distribution function (CDF) is:

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

and complementary CDF or "survival function" is

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

This rv has "support" given by $\mathrm{Supp}[X] :=$ (stuff that could be seen)

$\{x : p(x) > 0, x \in \mathbb{R}\}$

$|\mathrm{Supp}[X]| \leq |\mathbb{N}|$ countably infinite at most

elements in a set

Sets this size are called "discrete" sets.

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

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

The Most "fundamental" rv is the Bernoulli:

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

Although if $p(7) = p^7 (1-p)^{-6}$ this will not be within the support which is $\{0, 1\}$.

Let's define the "indicator function"

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

$p^{old}(x)$

use this to define a PMF for the Bern that is valid for all x

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

What if $p = 1$?

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

This rv just spits outs 1's which makes it a degenerate

Thus: $X \sim Deg(1) = \{1 \text{ w.p. } 1$ In General: $X \sim Deg(c) := \mathbb{1}_{x=c}$

If: $X \sim Bern(0) = Deg(0)$

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".

$p \in (0,1)$ (anything b/w 0, 1 but not exactly 0 or 1)

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

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

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

$p_{\vec{X}}(\vec{x}) = p_{X_1, \ldots, X_n}(x_1, \ldots, x_n)$ valid for $\vec{x} \in \mathbb{R}^n$ and

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

If $X_1, X_2, \dots, X_n$ are independent, then

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

If $X_1 \overset{d}{=} X_2 \overset{d}{=} \dots \overset{d}{=} X_n$ this means they are "equally ~~distri~~ distributed"

$\Updownarrow$

$\forall x \; p_{X_1}(x) = p_{X_2}(x) = \dots = p_{X_n}(x)$ meaning their PMFs are the same. However, this offers no simplification of the JMF unless...

$X_1, X_2, \dots X_n \overset{iid}{\sim}$ that means independent and identically distributed

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

Let $X_1, X_2 \overset{iid}{\sim} \text{Bern}(p)$, Let $T_2 = f(x_1, x_2) = X_1 + X_2 \sim ?$ is what the dist?

$= p_T(t)$

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

$*$ Convolution operator

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