8/26

A discrete r.v has probability mass function (pmf)
$p(x) := P(X=x)$ and the r.v $X \sim p(x)$ where x is the "realized value" X, x

The cumulative distribution function (cdf) is
$F(x) := P(X \le x)$

and complementary cdf or "survival function" is
$S(x) := P(X > x) = 1 - F(x)$

This rv has support given by $Supp[X] := \{x : p(x) > 0, x \in \mathbb{R}\}$

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

elements in a set

Sets this size are called "discrete sets"
The support and pmf are related by the following identity:

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

The most "fundamental" rv is the Bernoulli: (Brand name)
$X \sim Bern(p) := p^x (1-p)^{1-x}$ with support $Supp[X] = \{0,1\}$

not in support

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

Let's define the "indicator function"

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

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

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

with probability

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

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

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

$p \in (0,1)$

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

$p_{\vec{X}}(\vec{x}) = p_{X_1,\dots,X_n}(x_1,\dots,x_n)$ 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 $p_{\vec{X}}(\vec{x}) = p_{X_1}(x_1) p_{X_2}(x_2) \times \dots \times p_{X_n}(x_n)$

$= \prod_{i=1}^{n} p_{X_i}(x_i)$ "multiplication rule"

if $X_1 \overset{d}{=} X_2 \overset{d}{=} \dots \overset{d}{=} X_n$ this denotes "equal in distribution" meaning their pmfs are the same. However, this offers no simplification of the jmf unless.

$\forall x \ p_{X_1}(x) = p_{X_2}(x) = \dots = p_{X_n}(x)$

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

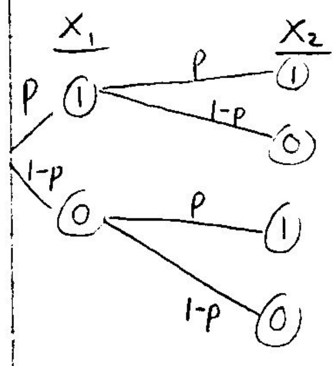
$\Rightarrow p_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} p(x_i)$

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

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

Convolution Operator

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



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

$$p^2$$
$$p(1-p)$$
$$(1-p)p$$
$$(1-p)^2$$