Math 621

Wed Aug. 26 (Lecture 01)

- A discrete random variable (r.v) has a probability mass function (pmf) $p(x) := p(X=x)$ and r.v $X \sim p(x)$ where x is the realized value. X is the random variable, x is the state the random variable realizes to.
- The cumulative distribution function (cdf) is $F(x) := p(X < x)$
- The survival function, denoted by S, is defined as $S(x) := 1 - F(X) = P(X > x)$ (called the survival function since it can be viewed as, "how long will it survive without being picked/occurring)
- This r.v (X) has support given by $supp[X] := \{x \mid x \in \mathbb{R},\ p(x) > 0\}$
- $|supp[X]| \leq \aleph_0$
  - Sets this size are called discrete sets
  - discrete r.v $\Rightarrow$ The number of things it can realize to is a discrete amount
- The pmf and the support are related by the following identity:

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

where each $x \in supp[X]$ is independent (The probability something happens is guaranteed)

The most fundamental r.v is the Bernoulli r.v

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

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

↳ 7 is not in the support. It is annoying to have to give 2 pieces of information (pmf and the support)

Let's define the indicator function: $\mathbb{1}_A := \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases}$

$$\Rightarrow X \sim Bern(p) = \underbrace{\overbrace{p^x (1-p)^{1-x}}^{p_{old}(x)} \mathbb{1}_{x \in \{0,1\}}}_{p(x)} \quad \Rightarrow \sum_{x \in \mathbb{R}} p(x) = 1$$

since every $x \in \mathbb{R}$ is now defined

(How this helps us instead of providing the support will be revealed next lecture)

What if $p = 1$?

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

calling this a random variable when it is determinent!?

$$X \sim Deg(1) = \{1 \text{ w.p } 1$$
$$X \sim Deg(0) = \{0 \text{ w.p } 1 = \mathbb{1}_{x=0}$$
$$X \sim Deg(c) := \mathbb{1}_{x=c}$$

The convention in this class is that paramater values (p is the paramater for the Bernoulli) that yield degenerate r.v's are not part of the legal paramater space.

↳ $p \in (0,1)$ (otherwise for to [not] happen)

If we have more than one r.v $X_1, X_2, \dots, X_n$, we can group them together in a column vector: $\vec{X} := [X_1, X_2, \dots, X_n]^T$ and define the joint mass function as:

$$p_{\vec{X}}(\vec{x}) = p_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n) \text{ valid for } \vec{x} \in \mathbb{R}^n \text{ and } \sum_{\vec{x} \in \mathbb{R}^n} p(\vec{x}) = 1$$

(multiple function pmf)

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

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

If $X_1 \overset{d}{=} X_2 \overset{d}{=} \dots \overset{d}{=} X_n$ ⇒ This denotes "equal distribution" meaning their pmf's are the same. However this offers no simplification of the pmf unless ... $X_1, X_2, \dots, X_n \overset{iid}{\sim}$ (independent and equally distributed)

⇕

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

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

Let $X_1, X_2 \overset{iid}{\sim} Bern(p)$. Let $T_2 = f(X_1, X_2) = X_1 + X_2 \sim p_T(t)$

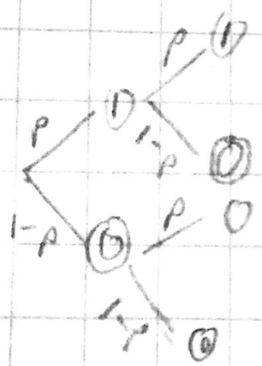
(function of 2 r.v and returns a r.v)

Sometimes denoted as $p_T(t) = p_{X_1}(x) * p_{X_2}(x)$ → convolution operator

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

0, 1 or 2 successes

(probability tree on the next page)



$p_{X_1,X_2}(x_1,x_2)$ (JmF)	T
p^2	2
$p(1-p)$	1
$p(1-p)$	1
$(1-p)^2$	0

$$T \sim \begin{cases} 2 & \text{w.p.} & p^2 \\ 1 & \text{w.p.} & 2p(1-p) \\ 0 & \text{w.p.} & (1-p)^2 \end{cases}$$