A discrete random variable (r.v) x has probability mass function (PMF) given by $p(x)$:

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

where x is the realized value

and cumulative distribution function (CDF) denoted $F(x)$:

$F(x) := F(X = x)$

And complementary CDF also called survival function:

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

The r.v. has support

$\text{Supp}[X] := \{x : p(x) > 0,\ x \in \mathbb{R}\}$ and $|\text{Supp}[X]| \leq |\mathbb{N}|$ i.e. finite or at most countably infinite.

# of elements in set

sets of this size are called "discrete"

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

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

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

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

$p(7) = p^7 (1-p)^{-6}$ → can't happen b/c $7 \notin \text{Supp}[X]$

To fix this, we introduce the <u>indicator function</u>

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

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

What if $p=1$?

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

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

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

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

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

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

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

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

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

which has a joint mass function (JMF) defined as:

$P_{\vec{X}}(\vec{x}) = P_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ s.t. $\sum_{\vec{x} \in \mathbb{R}^n} P_{\vec{X}}(\vec{x}) = 1$

If $X_1, \ldots, X_n$ are independent r.v.s then the JMF can be factored as:

$P_{\vec{X}}(\vec{x}) = P_{\vec{X}_1}(\vec{x}_1) \cdot \ldots \cdot P_{\vec{X}_n}(\vec{x}_n)$ (the multiplication rule)

If $X_1, \ldots, X_n$ are identically distributed denoted $X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} \ldots \stackrel{d}{=} X_n$ then $P_{X_1}(x) = P_{X_2}(x) = \ldots = P_{X_n}(x) \ \forall x$, but this offers no simplification of the JMF unless...

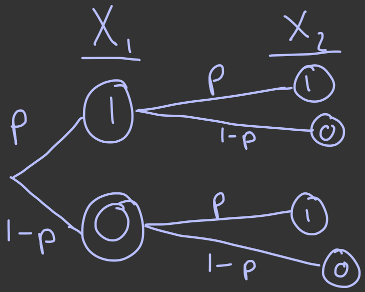
$X_1, \ldots, X_n \stackrel{iid}{\sim}$ denotes "independent and identically distributed"

$P_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} p(x_i) \rightarrow$ Shared PMF

$X_1, X_2 \stackrel{iid}{\sim} \text{Bern}(p)$, $T_2 \cdot F(X_1, X_2) = X_1 + X_2 \sim P_{T_2}(t)$, $\text{Supp}[T_2] = \{0, 1, 2\}$

$P_{T_2}(t) = \underbrace{P_{X_1}(x_1) * P_{X_2}(x_2)}$

$*$ ↓ convolution operator



$$\frac{p_{X_1,X_2}(x_1,x_2)}{\begin{array}{c} p^2 \\ p(1-p) \\ (1-p)p \\ (1-p)^2 \\ \hline 1 \end{array}}$$

$\boxed{\text{what happens in } X_1 \text{ does not determine what happens for } X_2 \text{ b/c they are indep.}}$

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