A discrete random variable (rv) X has probability mass function (PMF)

p(x) := P(X = x)

and cumulative distribution function (CDF)

F(x) := P(X ≤ x)

and complementary CDF / survival function

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

Note: F(x) + S(x) = 1

Discrete rvs have "support"

Supp[X] := {x : p(x) > 0 & x ∈ ℝ}
This set is "discrete" i.e. |Supp[X]| ≤ |ℕ| at most ctbly infinite

The PMF and support are related by the identity: $\sum_{x \in Supp[X]} p(x) = 1$

The most fundamanental rv is the "Bernoulli" rv. This is our first "brand name" rv.

X ~ Bernoulli(p) := $\overbrace{p^x (1-p)^{1-x}}^{p(x)}$ & Supp[X] = {0, 1}

$x \notin \{0,1\} \Rightarrow p(x) = 0, \quad p(7) = p^7 (1-p)^{1-7} = \frac{p^7}{(1-p)^6} \neq 0$ if $p \neq 0$

The problem is the PMF as stated is only valid for elements in the Supp[X]. We need to fix this. We want a function valid for all x ∈ ℝ.

Let's define the "indicator function" as follows:

$$\mathbb{1}_s := \begin{cases} 1 \text{ if s is true} \\ 0 \text{ if s is false} \end{cases}$$

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

$$p(x) = p_{old}(x) \mathbb{1}_{x \in Supp[X]}$$

Now we can update our humpty-dumpty formula: $\sum_{x \in \mathbb{R}} p(x) = 1$

$$\sum_{x \in \mathbb{R}} p_{old}(x) \mathbb{1}_{x \in Supp[X]} = \sum_{x \in Supp[X]} p_{old}(x) = 1.$$

What is p in the Bernoulli rv? What if p = 1?

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

This is a "degenerate rv". This is a nonrandom rv. It's an oxymoron.

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

What if p = 0 in the Bernoulli case?

$$X \sim Bern(0) := 0^x (1-0)^{1-x} \mathbb{1}_{x \in \{0,1\}} = \mathbb{1}_{x=0} = Deg(0)$$

What if p < 0 or p > 1? In these cases, the humpty-dumpty formula fails (i.e. does not sum to 1). Thus, it's an invalid PMF. So the valid cases which are non-degenerate are: p ∈ (0, 1) not inclusive. We call this set the "parameter space" and we call p the "parameter" of the rv. The "standard Bernoulli" means p = 0.5.

We have more than one discrete rvs $X_1, X_2, \ldots, X_n, n > 1$. We can group them together in a column vector, called a "vector rv". $\vec{X} = [X_1, X_2, \ldots, X_n]^T$ The PMF of a vector rv is called "joint mass function" (JMF) valid for all $\vec{x} \in \mathbb{R}^n$

$$p_{\vec{X}}(\vec{x}) = p_{X_1, \ldots, X_n}(x_1, \ldots, x_n), \quad \sum_{\vec{x} \in \mathbb{R}^n} p_{\vec{X}}(\vec{x}) = 1$$

If the rvs are "independent" denoted $X_1, X_2, \ldots, X_n \overset{ind}{\sim}$ then

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

If the rvs are "identically / equally distributed" denoted $X_1 \overset{d}{=} X_2 \overset{d}{=} \ldots \overset{d}{=} X_n$ this means by definition that

$$p_{X_1}(x) = p_{X_2}(x) = \ldots = p_{X_n}(x) = p(x)$$

but does not provide any simplificiation of the JMF computation.

If the rvs are independent and identically distributed (iid) denoted $X_1, \ldots, X_n \overset{iid}{\sim}$ then the JMF can be decomposed as:

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

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

$$p_{T_2}(t) = ?$$