Week 4

Def: the probability mass function (pmf) of a discrete RV X is

$$\boxed{f_X(x) := P_X(\{x\}) = P(\{X=x\}) = P_X(X=x) \quad \forall x}$$

Ex (Geometric Dist.): Let $X \sim Geo(p)$, $0<p<1$

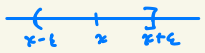
has $f_X(x) := P(X=x) := \begin{cases} (1-p)^{x-1}p, & x=1,2,\dots \\ 0, & \text{otherwise} \end{cases}$

- However, this concept is NOT useful for continuous RV's

indeed, if X is a continuous RV,
then $F_X(x)$ is a continuous function of x

Then,

$$P(X=x) \overset{AOP}{=} \lim_{\varepsilon \downarrow 0} P(X \in (x-\varepsilon, x+\varepsilon])$$



$$\overset{AOP}{=} \lim_{\varepsilon \downarrow 0} P(X \le x+\varepsilon) - P(X \le x-\varepsilon)$$

$$\overset{def}{=} \lim_{\varepsilon \downarrow 0} F_X(x+\varepsilon) - F_X(x-\varepsilon)$$

$$\overset{cont.}{=} F_X(x) - F_X(x)$$

$$= 0.$$

More useful is the cdf F_X and, for continuous RV's,
the notion of pdf:

Def: the probability density function (pdf), namely, $f_X(x)$, of a continuous RV X is a function that satisfies

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt}$$

Remarks: • If F_X is differentiable, then the pdf always exists, i.e. $f_X(x) = \frac{d}{dx}F_X(x)$

• F_X, or f_X contain all info there is abt the dist. of the RV X, we denote

$X \sim F_X$ or $X \sim f_X$ (equivalently)

Thm. A function $f_X(x)$ is a pdf (or pmf) of a RV i.f.f.

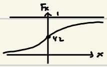
i) $f_X(x) \ge 0 \ \forall x$

ii) $\sum_x f_X(x) = 1$ (pmf)

$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$ (pdf)

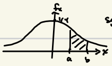
Ex: Recall the logistic cdf:

$F_X(x) = \frac{1}{1+e^{-x}}$



which gives

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$



symmetry since $f_X(x) = f_X(-x)$, i.e.

$$f(-x) = \frac{e^{x}}{(1+e^{x})^2} = \frac{e^{x}e^{-2x}}{(1+e^{x})^2 e^{-2x}} = \frac{e^{-x}}{(1+2e^{-x}+e^{2x})e^{-2x}} = \frac{e^{-x}}{e^{-2x}+2e^{-x}+1} = \frac{e^{-x}}{(1+e^{-x})^2} = f(x). \blacksquare$$

under which

$$P(X \in (a,b)) = \int_a^b \frac{e^{-x}}{(1+e^{-x})^2}\,dx$$

2.1: Functions of Random variables

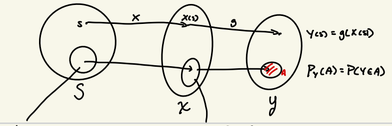
Recall (relevant ch. 1 context): A RV X is defined on a function from a sample space S to $\mathcal{X}$ (previously $\mathcal{K}$)
s.t. $X = X(s)$, $s \in S$.
information about the RV X is (completely) captured by its cdf:

$$F_X(x) = P(X \le x) = P(s \in S \mid X(s) \le x)$$

Take a function $g: \mathcal{X} \to \mathcal{Y}$ where $\mathcal{X}$: domain of RV X.

Then $Y = g(X)$ is also a RV taking values in $\mathcal{Y}$ because Y is in fact a function on S, i.e.

$$Y(s) = g(X(s)) \equiv g \circ X(s) \neq X \circ g$$



$Y^{-1}(A) = X^{-1}(g^{-1}(A))$ $\qquad$ $g^{-1}(A) := \{x \in \mathcal{X} \mid g(x) \in A\}$

Q: what is the prob. dist. of $Y = g(X)$?

$$\boxed{\begin{aligned} &\underline{\text{Prop.}}\ \text{By def., } \forall A \subset \mathcal{Y}, \\ &P(Y \in A) = P(g(X) \in A) \\ &\qquad = P(x \in g^{-1}(A)) \qquad = P(X(s) \in g^{-1}(A)) \\ &\qquad = P(s \in X^{-1}(g^{-1}(A))) = P(s \in X^{-1} \circ g^{-1}(A)) \\ &\qquad = P(s \in (g \circ X)^{-1}(A)) \\ &\qquad = P(s \in Y^{-1}(A)) \\ &\qquad = P(Y(s) \in A) \end{aligned}}$$

<u>Note</u>: we can 'create' new RVs by applying a function to an existing RV, instead of specifying the prob. dist. from scratch, i.e. via a sample space and sigma algebra

<u>Ex (Binomial Transformation)</u>: X is a binomial RV if its pmf

$$f_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

we write $X \sim \text{Binomial}(n, p)$

Let $Y := g(X)$ where $g(x) = n - x$,

i.e. $Y = n - x$ is also a RV

Q: what is the distribution of Y?

$Y \in \{0, 1, \dots, n\}$.

For $y \in \{0, 1, \dots, n\}$: $f_Y(y) := P(Y = y)$

$$= P(X = n - y) \quad , Y = n - x$$

$$= f_X(n-y) = \binom{n}{n-y} p^{n-y} (1-p)^y$$

$$= \binom{n}{y} (1-p)^y p^{n-y}$$ since $\binom{n}{y} = \binom{n}{n-y}$ by def. ★why?