Let X be a random variable
↳ data generating process

Let x be a realization (data).
$x \in Supp[X]$ "support" $= \{x : p(x) > 0\}$

Discrete r.v's X

$|Supp[X]| \le |\mathbb{N}|$ ← countable infinity.

size (# of elements)

Probability Mass Function (PMF)

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

two definitions

$p : \mathbb{R} \to [0, 1]$

$p : Supp[X] \to (0, 1]$

Cummulative Distribution Function (CDF)

$$F(x) := P(X \le x) = \sum_{\{y : y \in Supp[X] \text{ and } y \le x\}} p(y)$$

$F : \mathbb{R} \to [0, 1]$

Continuous r.v's

$|Supp[X]| = |\mathbb{R}|$ uncountable infinity

CDF is defined the same as for discrete r.v's

$f(x) := \frac{d}{dx}[F(x)]$ probability density function (PDF) $Supp[X] := \{x : f(x) > 0\}$

By Fundamental Thm of Calculus $a \le b$,

$$P(X \in [a,b]) = F(b) - F(a) = \int_a^b f(x)\, dx$$

Examples of Discrete r.v's

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

② $X \sim Binom(n, p) := \underbrace{\binom{n}{x} p^x (1-p)^{1-x}}_{p(x)}$ $Supp[X] = \{0, 1 \ldots n\}$

"Parameters" A tuning knob

For Bernoulli, p is the "probability of success" (if 1 is the definition of success).

$p \in (0, 1)$

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

"degenerate" are deterministic

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

$p=0$ or $p=1$ are degenerate cases. As a convention, we exclude them from the param space

Parameter Space the set of all parameter values which are not degenerate.

For Binomial $p \in (0,1)$

$n \in \{1, 2, \dots\} = \mathbb{N}$

Continuous r.v. Examples

$X \sim \text{Exp}(\lambda) := \underbrace{\lambda e^{-\lambda x}}_{f(x)}$

$\text{supp}[X] = (0, \infty)$

$\lambda \in (0, \infty)$

$X \sim N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(X-\mu)^2}$

$\text{Supp}[X] = \mathbb{R}$

$\mu \in \mathbb{R}$ $\sigma^2 = Var[X] := E[(X-\mu)^2]$

$\sigma^2 \in (0, \infty)$

Let θ be the unknown parameter.

And $\vec{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$ " parameters.

And Θ be the parameter space.

$X \sim \text{Bern}(\theta) := \theta^x (1-\theta)^{1-x}$

$X \sim N(\underbrace{\theta_1, \theta_2}_{\vec{\theta}})$

Parametric Model

$\mathcal{F} := \{ p(x; \theta) : \theta \in \Theta \}$

space of parametermodel

PMF/PDF

PMF "needs" the parameter value

If $X_1, X_2, \ldots, X_n$ are random variable $p(x_1, x_2, \ldots, x_n ; \vec{\theta})$ is the joint mass function / joint density function.

If $X_1, \ldots, X_n \overset{ind}{\sim}$ (independent)

$$p(x_1, x_2 \ldots x_n ; \vec{\theta}) = p_1(x_1 ; \vec{\theta}) \cdot p_2(x_2 ; \vec{\theta}) \cdot \ldots \cdot p_n(x_n ; \vec{\theta}) = \prod_{i=1}^{n} p_i(x_i ; \vec{\theta})$$

If $X_1, \ldots, X_n \overset{iid}{\sim}$ independent & identically distributed.

$$= \prod_{i=1}^{n} p(x_i ; \vec{\theta})$$

In the real world, you observe :

$x = \langle 0, 0, 1, 0, 1, 0 \rangle$

Step I Pick $\mathcal{F}$. Pick a model. Beyond the scope of the course.

Step II Inference (i.e figure out the value of $\vec{\theta}$)

Three Goals of Inference

(I) Point Estimation. Provide ~~a~~ best single guess of θ (denoted $\hat{\theta}$).

(II) Confidence set. Provide range of likely θ

(III) Theory Test. Test plausibility these θ is some value or belongs to some sets.