Continuos RV

(i) F_X is continuous then X is continuous

(ii) If F_X is differentiable then

$$\frac{dF_X}{dx} = f$$ density function

Discrete : $\mathbb{P}(a \le X \le b) = \sum_{x=a}^{b} p(x)$

Cont : $\mathbb{P}(a \le X \le b) = \int_a^b f(x)\,dx$

Def: $E(X) = \int x f(x) dx$

$$Var(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$$

$$= \int (x - \mathbb{E}(x))^2 f(x) dx$$

$\int (\quad F(x)) \; f(x)\,dx$

Ex: $f(x) = \dfrac{1}{b-a} I(a \le x \le b)$

$F(x) = \dfrac{x-a}{b-a}$

Ex: $f(x) = \lambda e^{-\lambda x} I(x \ge 0)$

$X \sim Exp(\lambda)$

$F(x) = 1 - e^{-\lambda x}$

Ex: $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \dfrac{(x-\mu)^2}{2\sigma^2} \right\}$

$X \sim N(\mu, \sigma^2)$

CDF – Computed numerically

Rmk: Gaussian closed under linear operations

Statistics

$$\begin{tikzcd} \boxed{\text{model}} \arrow[r, bend left, "\text{Prob.}"] & \boxed{\text{Data}} \arrow[l, bend left, "\text{Stat/Inference.}"] \end{tikzcd}$$

- We'll be assuming a parametric setting in most of this course
- End with nonparametric inference.

Ex: $X_1, \dots, X_n \overset{iid}{\sim} \text{Bern}(p)$

* Parametric inference attempts to estimate parameters in a certain model.

$$\hat{g}(X_1, \dots, X_n) = \hat{\theta} \leftarrow \text{random}$$

So the distribution of $\hat{\theta}$ is the sampling distribution

- $\hat{\theta}$ normally chosen to minimize some cost function

Ex: $\hat{p} = \frac{1}{n}\sum x_i \qquad X_i \sim \text{Bern}(p)$

$$\hat{p} \xrightarrow{a.s.} p \qquad \mathbb{E}(\hat{p}) = p$$

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{D} N(0,1)$$

Confidence Intervals

Def: A confidence set is a stat. $S(X_1, \ldots, X_n)$ s.t. $S(D) \ni \theta$ with high probability.

$$\mathbb{P}(S(D) \ni \theta) = 1 - \alpha \qquad (*)$$

Rmk: $S(D)$ usually an interval

(*) usually want entirely specify

$S(D)$

Ex: $X_1, \ldots, X_n \sim \text{iid } N(\mu, \sigma^2)$ known

$$\hat{\mu} = \frac{\sum X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

There always exists $z_{\alpha/2}$ s.t.

$$\mathbb{P}(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1 - \alpha$$

$$\Rightarrow \mathbb{P}\left(-z_{\alpha/2} \le \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right) = 1 - \alpha$$

$$\mathbb{P}\Big(\underbrace{\hat{\mu} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \hat{\mu} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\text{Classical Confidence}}\Big) = 1 - \alpha$$

Interview