Extra OH:

Thurs: 2-3
Mon: 2-3
Tues: 3-4

Defn: Random Sample

If $X_1, X_2, \dots, X_N$ are mutually independent all having marginal dist f

N ← sample size

then we say these Xs are a random sample (RS) from f.

i.e. $X_n \overset{iid}{\sim} f$

Notation:

$$\underset{\sim}{X} = (X_1, \dots, X_N)$$

← a mv-RV or a random vector

$$\underset{\sim}{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$$

Joint dist of a RS

$$f(\underset{\sim}{x}) = f(x_1, \dots, x_N)$$

$$= f(x_1) f(x_2) \cdots f(x_N) \quad \text{[by independent]}$$

$$= \prod_{n=1}^{N} f(x_n)$$

Ex. Let $X_n \overset{iid}{\sim} Exp(\lambda)$

$$f(x) = \lambda e^{-\lambda x} \text{ for } x > 0$$

more explicit

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

indicator

$$\mathbb{1}(\text{statement}) = \begin{cases} 1 & \text{statement true} \\ 0 & \text{statement false} \end{cases}$$

more compact

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}(x > 0)$$

What is the joint of X_ns?

$$f(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n) = \prod_{n=1}^{N} \lambda e^{-\lambda x_n} \mathbb{1}(x_n > 0)$$

$$= \lambda^N e^{-\lambda \sum_n x_n} \prod_n \mathbb{1}(x_n > 0)$$

$$e^a e^b = e^{a+b}$$

$$\prod_n e^{a_n} = e^{\sum_n a_n}$$

$$= \lambda^N e^{-\lambda \sum_n x_n} \prod_n \mathbb{1}(x_n > 0) \quad \prod_n e^{\dots} = e^{\dots}$$

$$= \lambda^N e^{-\lambda \sum_n x_n} \mathbb{1}(\text{all } x_n > 0)$$

$\mathbb{1}(A)\mathbb{1}(B) = \mathbb{1}(A \text{ and } B)$

$\prod_n \mathbb{1}(A_n) = \mathbb{1}(\text{all } A_n \text{ true})$

Defn: Statistic

Given a RS $X_n \overset{iid}{\sim} f$

and a function

$$T : \mathbb{R}^N \to \mathbb{R}^d$$

often $d \ll N$
e.g. $d = 1$

then $T(\underset{\sim}{X})$ is called a statistic.

Ex.

(1) Arithmetic Mean ($d=1$)

$$T(\underset{\sim}{X}) = \frac{1}{N} \sum_{n=1}^{N} X_n = \overline{X}_N$$

(2) Sample Variance

$$S^2 = \frac{1}{\quad} \sum^{N} (X_n - \overline{X}_N)^2$$

$_{N-1} \quad N-1 \quad n=1$

3. **Minimum:** $X_{(1)} = \min\{X_1, \dots, X_N\}$

4. **Maximum:** $X_{(N)} = \max\{X_1, \dots, X_N\}$

5. **Range:** $X_{(N)} - X_{(1)}$

6. **Order Statistics** $X_{(r)} = r^{th}$ smallest

Defn: Sampling Distribution

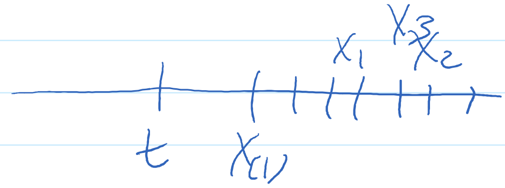
For a stat. T the sampling dist is simply its distribution.

Ex. what is the dist of $X_{(1)}$?

Lets assume $X_n \overset{iid}{\sim} f$ and f cts.

Let F be the CDF of X_n's.

I want the PDF of $X_{(1)}$.



$$\mathbb{P}(X_{(1)} \geq t)$$

$$= \mathbb{P}(X_1 \geq t, X_2 \geq t, \dots, X_N \geq t)$$ $\because$ Independence

$$= \mathbb{P}(X_1 \geq t, X_2 \geq t, \dots, X_N \geq t)$$

independence

$$= \mathbb{P}(X_1 \geq t)\,\mathbb{P}(X_2 \geq t) \cdots \mathbb{P}(X_N \geq t)$$

$$= \mathbb{P}(X_n \geq t)^N$$

$$= (1 - F(t))^N$$

$$F_{X_{(1)}}(t) = \mathbb{P}(X_{(1)} \leq t) = 1 - \mathbb{P}(X_{(1)} \geq t)$$

$$= 1 - (1 - F(t))^N$$

$$\boxed{f_{X_{(1)}}(t) = \frac{d F_{X_{(1)}}}{dt} = N\,(1 - F(t))^{N-1} f(t)}$$

Can play same game for $X_{(N)}$ and look at

$$\mathbb{P}(X_{(N)} \leq t)$$

... and get $\boxed{f_{X_{(N)}}(t) = N F(t)^{N-1} f(t)}$

Ex: Let $X_i \overset{iid}{\sim} \mathrm{Exp}(\lambda)$. | $f(x) = \lambda e^{-\lambda x}$

Ex: Let $X_n \overset{iid}{\sim} \text{Exp}(\lambda)$.

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

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

What is the dist of $X_{(1)}$?

$$f_{X_{(1)}}(t) = N(1-F(t))^{N-1} f(t)$$

$$= N(1-(1-e^{-\lambda t}))^{N-1} \lambda e^{-\lambda t}$$

$$= N(e^{-\lambda t})^{N-1} \lambda e^{-\lambda t}$$

$$f_{X_{(1)}}(t) = (N\lambda) e^{-(N\lambda)t}$$

PDF of $\text{Exp}(N\lambda)$

i.e. $\boxed{X_{(1)} \sim \text{Exp}(N\lambda)}$

so $\mathbb{E} X_{(1)} = \dfrac{1}{N\lambda}$