Lecture 1: Statistics

Defn: Random Sample

N = sample size

If $X_1, X_2, X_3, \dots, X_N$ are mutually independent all w/ marginal dist f then we say these X_ns are a random sample from f.

denoted: $X_n \overset{iid}{\sim} f$

indep and identically dist.

Notation:

a random vector

a MV rand var.

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

a vector in $\mathbb{R}^N$

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

Joint dist of a RS (rand. sample)

joint

$$\begin{aligned} f(\underset{\sim}{x}) &= f(x_1, x_2, \dots, x_N) \\ &= f(x_1) f(x_2) f(x_3) \cdots f(x_N) \\ &= \prod_{n=1}^{N} f(x_n) \end{aligned}$$

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

What is the joint of $\underset{\sim}{X}$.

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

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

$$= \lambda^N \prod_{n=1}^{N} \left(e^{-\lambda x_n}\right) \prod_{n=1}^{N} \mathbb{1}(x_n > 0)$$

$$= \lambda^N e^{-\lambda \sum_{n=1}^{N} x_n} \prod_{n=1}^{N} \mathbb{1}(x_n > 0)$$

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

$\text{Exp}(\lambda)$

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

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

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

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

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

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

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

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

Defn: Statistic

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

$$T: \mathbb{R}^N \to \mathbb{R}^d \qquad (\text{typically } d \ll N)$$

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

Ex. Arithmetic Mean : $(d=1)$

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

Sample Variance : $(d=1)$

$$S_{N-1}^2 = \frac{1}{N-1}\sum_{n=1}^{N}(X_n - \bar{X}_N)^2$$

Sample SD :

$$S_{N-1} = \sqrt{S_{N-1}^2}$$

Minimum :

$$X_{(1)} = \min_{n=1,\dots,N} X_n$$

Maximum :

$$X_{(N)} = \max_{n=1,\dots,N} X_n$$

Range: $R = X_{(N)} - X_{(1)}$.

Order Statistic: $X_{(r)} = r^{th}$ smallest among X_ns

Defn: Sampling Distribution

If $T = T(\underset{\sim}{X})$ is a statistic then the sampling dist of T is just its distribution.

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

Assume $X_n \overset{iid}{\sim} f$, f cts

F is the CDF

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

$$\mathbb{P}(X_{(1)} \geq t) = \mathbb{P}(X_1 \geq t, X_2 \geq t, X_3 \geq t, \dots, X_N \geq t)$$

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

(by independence)

$$= \prod_{n=1}^{N} \mathbb{P}(X_n \geq t)$$

$\mathbb{P}(X_n \geq t)$
$= 1 - \mathbb{P}(X_n < t)$
$= 1 - F(t)$

$$= \prod_{n=1}^{N} (1 - F(t))$$

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

$$\boxed{\mathbb{P}(X_{(1)} \geq t) = (1 - F(t))^N}$$

$$F_{X_{(1)}}(t) = \mathbb{P}(X_{(1)} \le t)$$
$$= 1 - \mathbb{P}(X_{(1)} \ge t)$$
$$= 1 - (1 - F(t))^N$$

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

Play similar game w/ maximum $X_{(N)}$

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

you get

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

Famous result from into stats:

$$X_n \overset{iid}{\sim} N(\mu, \sigma^2) \quad \text{then} \quad \overline{X}_N \sim N(\mu, \sigma^2/N)$$

Facts: Sums of RVs

Let $g: \mathbb{R} \to \mathbb{R}$ and $X_n \overset{iid}{\sim} f$

① $\mathbb{E}\left[\sum_{n=1}^{N} g(X_n)\right] = N\,\mathbb{E}[g(X_n)]$ any of them

Pf. $\mathbb{E}\left[\sum_n g(X_n)\right] = \sum_n \mathbb{E}[g(X_n)] = N\,\mathbb{E}[g(X_n)]$

② $\mathrm{Var}\left(\sum_n g(X_n)\right) = N\,\mathrm{Var}(g(X_n))$

Pf. Similar as above — NEED independence

<u>Theorem</u>: If $X_n \overset{iid}{\sim} f$ and

$\mathbb{E}X_n = \mu$ and $\mathrm{Var}(X_n) = \sigma^2$

then

① $\mathbb{E}[\bar{X}_N] = \mu$

② $\mathrm{Var}(\bar{X}_N) = \sigma^2/N$

③ $\mathbb{E}[S^2_{N-1}] = \sigma^2$