Lecture 1: Statistics

Defn: Random Sample

N = sample size

If $X_1, X_2, X_2, \dots, X_N$ are mutually independent random variables all w/ marginal dist. f then we say these X_ns are a **random sample** of size N

denote: $X_n \overset{iid}{\sim} f$ independent identically distributed

---

Notation:

$\underset{\sim}{X} = (X_1, \dots, X_N)$ is a random vector

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

Joint dist of a RS

$$f(\underset{\sim}{x}) = f(x_1, x_2, \dots, x_n)$$ joint dist

$$= f(x_1) f(x_2) f(x_3) \cdots f(x_n) \quad [\text{independent}]$$

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

$\sum_i a_i = a_1 + a_2 + \cdots$

$\prod_i a_i = a_1 a_2 a_3 \cdots$

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

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

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

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

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

$\mathbb{1}(x_1 > 0$ and $x_2 > 0$ and $x_3 > 0 \ldots)$

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

$\text{Exp}(\lambda)$

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

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

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

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

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

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

Defn: Statistic

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

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

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

Ex. Arithmetic Mean $(d=1)$

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

also random (univariate)

Sample Variance $(d=1)$

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

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_n.

Defn: Sampling Distribution

The sampling distribution of a statistic $T(\underline{X})$ is just its distribution.

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

Assume $X_n \overset{iid}{\sim} f$, f is continuous, F is the CDF

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

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

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

$F(t) = \mathbb{P}(X_n \le t)$

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

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

$F_Z(z) = \mathbb{P}(Z \le z)$

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

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

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

$$f_{X_{(1)}}(t) = \frac{\partial}{\partial t} F_{X_{(1)}}(t) = \frac{\partial}{\partial t}\left[1 - (1 - F(t))^N\right]$$

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

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

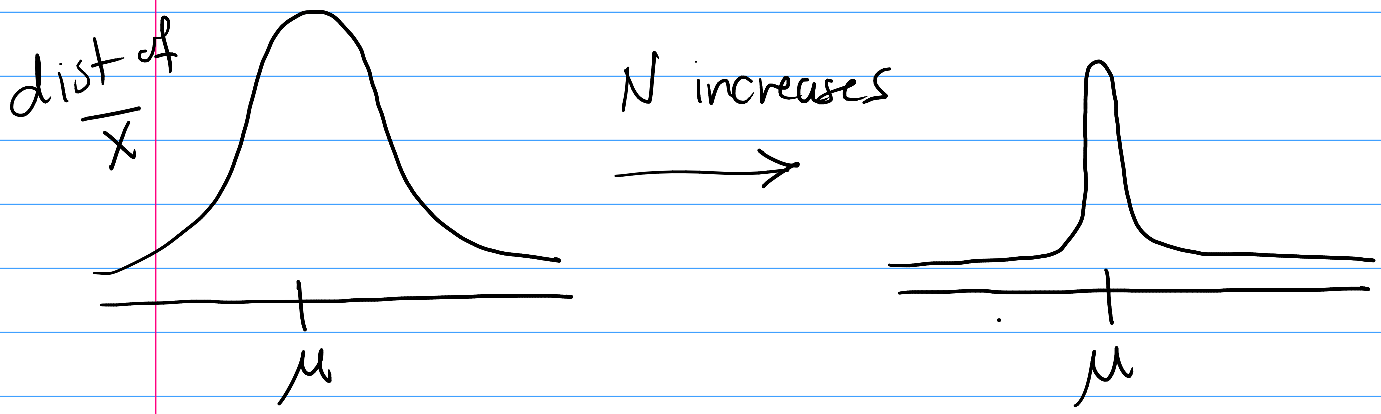
Can play a similar game for $X_{(N)}$

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

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

<u>Famous result (from intro. stats).</u>

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



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. $\mathrm{Var}\left(\sum_n g(X_n)\right) \underset{\text{NEED INDEPENDENCE}}{=} \sum_n \mathrm{Var}(g(X_n)) = N\,\mathrm{Var}(g(X_n))$

Theorem: 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) $\mathbb{E}[S^2_{N-1}] = \sigma^2$.