Defn: Random Sample

$N =$ Sample size

If $X_1, X_2, X_3, \ldots, X_N$ are mutually indep. RVs all w/ marginal dist f — then we say they are a random sample of size N.

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

Notation: $\underset{\sim}{X} = (X_1, \ldots, X_N)$ ← a radom vector

$\underset{\sim}{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ ← deterministic.

Joint dist of a RS:

$f(\underset{\sim}{x}) = f(x_1, \ldots, x_n)$

$$f(\underset{\sim}{x}) = f(x_1, \ldots, x_n)$$

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

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

$$a_1 + a_2 + \cdots = \sum_n a_n$$

$$a_1 \cdot a_2 \cdots = \prod_n a_n$$

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

What's the joint?

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

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

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

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

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

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

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

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

$= \lambda e^{\quad n} \left(\prod_n \mathbb{1}(x_n > 0)\right)$ $\quad e\, e = e$

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

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

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

$\prod \mathbb{1}(A_i) = \mathbb{1}(\text{all } A_i \text{ true})$

Defn: Statistic

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

$$T : \mathbb{R}^N \to \mathbb{R}^d \quad (\text{typ. } d \ll N)$$

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

Ex. Arithmetic Mean $(d=1)$

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

Ex, Sample Variance

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

Ex. Sample SD

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

Ex. Minimum

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

Ex. Maximum

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

Ex. Range :

$$R = X_{(N)} - X_{(1)}$$

Ex. **Order Statistic**

$$X_{(r)} = r^{th} \text{ smallest among } \mathbb{X}_n.$$

Defn: Sampling Dist

The sampling dist of a stat T is just its dist.

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

Assume $\mathbb{X}_n \overset{iid}{\sim} f$, f is cts, F is the CDF.

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

$$P(X_{(1)} > t) = P(\mathbb{X}_1 > t, \mathbb{X}_2 > t, \ldots, \mathbb{X}_N > t)$$

$$= P(\mathbb{X}_1 > t) P(\mathbb{X}_2 > t) \cdots P(\mathbb{X}_N > t)$$

[indep]

$$= \prod_{n=1}^{N} P(\cancel{X}_n > t)$$

$$= P(X_n > t)^N$$

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

$F(t) = P(\cancel{X}_n \le t)$

Defn

$$F_{X_{(1)}}(t) = \mathbb{P}(X_{(1)} \le t)$$

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

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

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

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

Can play similar game for maximum to get

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

Theorem:

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

① $E\left[\sum_{n=1}^{N} g(X_n)\right] = N E[g(X_n)]$

X_n = any of them

Pf. $E\left[\sum_n g(X_n)\right] = \sum_n E[g(X_n)]$

$= E[g(X_1)] + E[g(X_2)] + \cdots$

$= N E[g(X_n)]$

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

pf. $\mathrm{Var}\left(\sum_n g(X_n)\right)$

$\mathrm{Var}(A+B) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\,\mathrm{Cov}(A,B)$

$\geq \sum_n \mathrm{Var}(g(X_n))$

Need independence (really, un-corr.)

$= N\,\mathrm{Var}(g(X_n))$.

Theorem: If $X_n \overset{iid}{\sim} f$ and

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

then

① $E[\bar{X}_N] = \mu$.

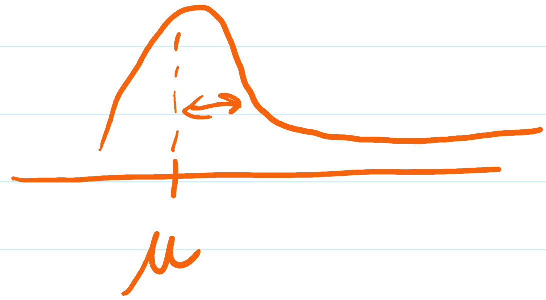
② $\mathrm{Var}(\bar{X}_N) = \sigma^2/N$

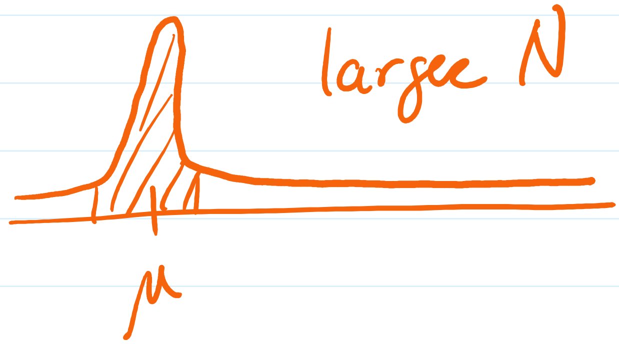
dist of $\bar{X}$

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

dist of $\bar{X}$

(3) $\mathbb{E}[S_{N-1}^2] = \sigma^2$.





pf.

(1) $\mathbb{E}\bar{X} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} X_n\right]$

$$= \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[X_n]$$

$$= \frac{1}{N}\sum_{n=1}^{N} \mu$$

$$= \frac{1}{N} N\mu \quad = \mu$$

(2) $\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{N}\sum_{n=1}^{N} X_n\right)$

$$② \ Var(\bar{X}) = Var\left(\frac{1}{N}\sum_{n=1} X_n\right)$$

$$= \frac{1}{N^2}\sum_{n=1}^{N} Var(X_n)$$

$$= \frac{1}{N^2}\sum_{n=1}^{N} \sigma^2$$

$$= \frac{1}{N^2} N\sigma^2 = \sigma^2/N$$