Defn: Random Sample

If $X_1, \dots, X_n \overset{iid}{\sim} f$ ← f is some distribution

then we call the $\{X_i\}$ a random sample (of size n) from f.

Fact: If X_i are a random sample then

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

$(x_1, \dots, x_n)$

$$= f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n) \quad \text{(b/c independence)}$$

$$= f(x_1) f(x_2) \cdots f(x_n) \quad \text{(b/c identical distributions)}$$

$$= \prod_{i=1}^{n} f(x_i)$$

All together: $\boxed{f(\underset{\sim}{x}) = \prod_{i=1}^{n} f(x_i)}$

Defn: Statistic

If $\{X_i\}_{i=1}^{n}$ are a random sample then if T is function $T : \mathbb{R}^n \to \mathbb{R}$, we say

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

Ex. 1) arithmetic mean

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots X_n)$$

Consider $\boxed{T : \mathbb{R}^n \longrightarrow \mathbb{R}}$

when $T(x_1, \ldots, x_n) = \frac{1}{n}(x_1 + \cdots x_n)$

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

2) Sample variance

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = T(x_1, \ldots, x_n)$$

3) Order Statistics

$$X_{(1)} = \text{minimum of } X_1, \ldots, X_n = \min_{i=1,\ldots,n} X_i$$

$$X_{(n)} = \text{maximum of } X_1, \ldots, X_n = \max_{i=1,\ldots,n} X_i$$

$X_{(r)} = r^{th}$ smallest value among $X_1, \ldots, X_n$

4) range: $X_{(n)} - X_{(1)} = R$

median:

$$M = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & n \text{ odd} \\ \dfrac{X_{(n/2)} + X_{(n/2+1)}}{2} & n \text{ even} \end{cases}$$

Defn: Sampling Distribution of a Statistic.

The sampling dist. of a Stat $T(\underset{\sim}{X})$ is simply the dist. of $T(\underset{\sim}{X})$.

univariate r.v.

Order Statistics

Henceforth: $\{X_i\}_{i=1}^n$ is a random sample

Minimum $\quad X_{(1)} = \min_{i=1,..,n} X_i$

Dist of $X_{(1)}$?

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

$$= P(X_1 > t) P(X_2 > t) P(X_3 > t) \cdots P(X_n > t) \quad \text{independence}$$

$$= \prod_{i=1}^n \underbrace{P(X_i > t)}_{1 - F(t)} \quad \text{same distribution w/ a CDF } F$$

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

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

$$P(X_{(1)} > t) = (1 - F(t))^n$$

$$\boxed{F_{X_{(1)}}(t) = 1 - P(X_{(1)} > t) = 1 - (1 - F(t))^n}$$

For continuous r.Vs

$$f_{X_{(1)}}(t) = \frac{d}{dt} F_{X_{(1)}}(t) = \frac{d}{dt}\left(1 - (1 - F(t))^n\right)$$

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

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

pdf for minimum

F CDF for X_i

f PDF of X_i

Maximum :

$$F_{X_{(n)}}(t) = P(X_{(n)} \le t) = P(X_1 \le t, X_2 \le t, \ldots, X_n \le t)$$

$$= \prod_{i=1}^{n} P(X_i \le t)$$

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

$$= F(t)^n$$

For continuous

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

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

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

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

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

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

$$= n\left(e^{-\lambda t}\right)^{n-1} \lambda e^{-\lambda t}$$

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

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

See: this is the pdf of $Exp(n\lambda)$

$$\boxed{X_{(1)} \sim Exp(n\lambda)}$$

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

$$\boxed{= n\left(1-e^{-\lambda t}\right)^{n-1} \lambda e^{-\lambda t}}$$

Theorem: $X_{(r)}$ = r^{th} smallest value among $X_1, \ldots, X_n$

If X_i are continuous

$$\boxed{f_{X_{(r)}}(t) = \frac{n!}{(r-1)!\,(n-r)!} F(t)^{r-1} (1-F(t))^{n-r} f(t)}$$

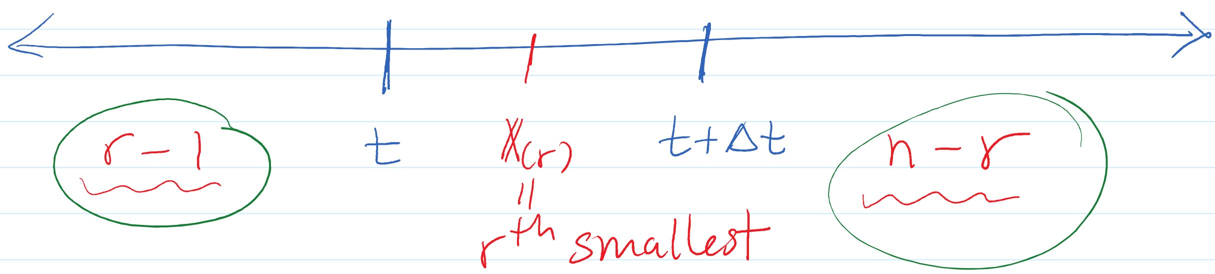
$X_{(r)}$ ... $(r-1)!(n-r)!$ ...

F CDF of $\{X_i\}$
f pdf of $\{X_i\}$

Notice: when $r=1$ or $r=n$
we get the previous formula.

Pf.

$$F(t)^{r-1} \cdot f(t) \cdot (1-F(t))^{n-r}$$



$$\boxed{\begin{aligned} f_{X_{(r)}}(t) &= \lim_{\Delta t \to 0} P(t \le X_{(r)} \le t+\Delta t) \\ &= \frac{n!}{(r-1)!(n-r)!} F(t)^{r-1} (1-F(t))^{n-r} f(t) \end{aligned}}$$

Ex. $X_i \overset{iid}{\sim} U(0,1)$

$f(t) = 1$ for $0 \le t \le 1$

$F(t) = t$ for $0 \le t \le 1$

So $f_{X_{(r)}}(t) = \frac{n!}{(r-1)!(n-r)!} F(t)^{r-1}(1-F(t))^{n-r} \underbrace{f(t)}_{1}$

$$\boxed{= \frac{n!}{(r-1)!(n-r)!} t^{r-1}(1-t)^{n-r}}$$

for $0 \le t \le 1$ PDF of a Beta$(r, n-r+1)$

So: $\boxed{X_{(r)} \sim \text{Beta}(r, n-r+1)}$

<u>Theorem:</u> <u>Joint Distribution of Order Statistics</u>

If $r < s$ then $X_{(r)} = r^{th}$ smallest

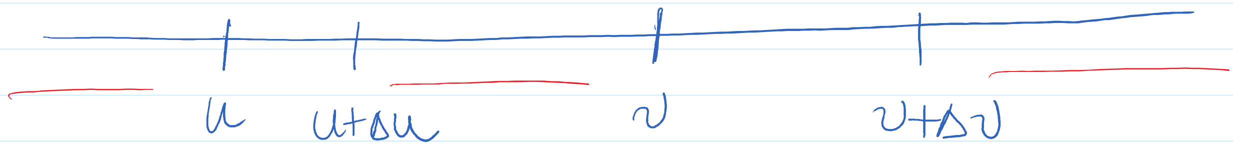
$X_{(s)} = s^{th}$ smallest

$$X_{(r)} < X_{(s)}$$

$$\boxed{f_{X_{(r)},X_{(s)}}(u,v) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \cdot F(u)^{r-1}(F(v)-F(u))^{s-r-1}(1-F(v))^{n-s} f(u) f(v).}$$

<u>Pf.</u> $r-1$ $\quad \underset{1}{X_{(r)}}$ $\quad s-r-1$ $\quad \underset{1}{X_{(s)}}$ $\quad n-s$

PP: $r-1$ 1 $s-r-1$ 1



$$\boxed{\frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}\,F(u)^{r-1}\left(1-F(v)\right)^{n-s}\left(F(v)-F(u)\right)^{s-r-1}}$$

Ex. Let $X_i \overset{iid}{\sim} U(0,1)$ then

$$f_{X_{(r)},X_{(s)}}(u,v)=\frac{n!}{(r-1)!\,(n-s)!\,(s-r-1)!}\,u^{r-1}(v-u)^{s-r-1}(1-v)^{n-s}$$

recall: $f(t)=1$ for $0\le t\le 1$

$F(t)=t$

Ex. $R = X_{(n)} - X_{(1)}$.

Q: what is the dist of R?

If $X_i \overset{iid}{\sim} U(0,1)$ Bivariate Transformation!

Let $\boxed{U = R = X_{(n)} - X_{(1)}}$

$V = X_{(1)}$

$X_{(1)} = V$ so $g_1^{-1}(u,v) = v$

$X_{(n)} = U+V$ $\quad g_2^{-1}(u,v) = u+v$

$$J = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \text{ so } \underline{|\det J| = 1}$$

So:

$$\boxed{f_{U,V}(u,v) = \underbrace{f_{X_{(1)},X_{(n)}}(v, u+v)}}$$

Aside: $r=1$ and $s=n$

$$\boxed{f_{X_{(1)},X_{(n)}}(a,b)} = \frac{n!}{(1-1)!(n-n)!(n-1-1)!} \cdot F(a)^{1-1}\left(F(b)-F(a)\right)^{n-1-1}\left(1-F(b)\right)^{n-n} f(a)f(b)$$

Simplify!

$$= \boxed{n(n-1)\left(F(b)-F(a)\right)^{n-2} f(a)f(b)}$$

plug in:

$n-2$

$$f_{U,V}(u,v) = n(n-1)\left(F(u+v) - F(v)\right) f(u+v) f(v)$$

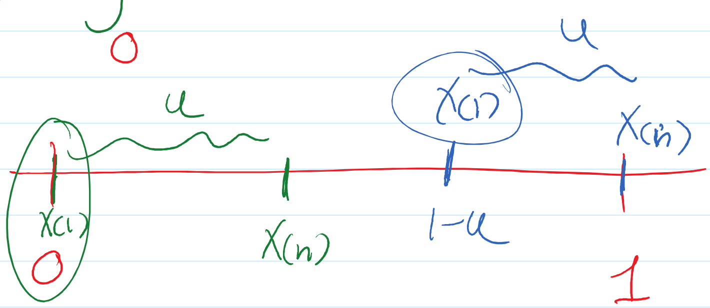
Since: $X_i \overset{iid}{\sim} U(0,1)$

$f(t) = 1 \quad F(t) = t$

$$\boxed{= n(n-1)\, u^{n-2}}$$

joint of $\begin{cases} U = R = X_{(n)} - X_{(1)} \\ \boxed{V = X_{(1)}} \end{cases}$

$$f_U(u) = \int_0^{1-u} n(n-1)\, u^{n-2}\, dv = n(n-1)\, u^{n-2} \int_0^{1-u} dv$$

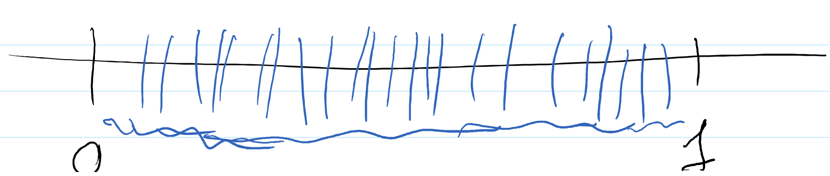


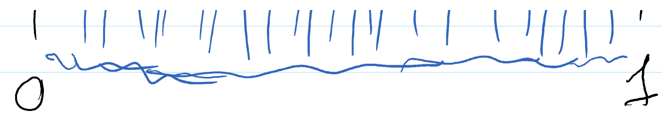
$U = R = X_{(n)} - X_{(1)}$

$$= n(n-1)\, u^{n-2} (1-u)$$

$$\boxed{R \sim \text{Beta}(n, 2)}$$

Fact: $E[R] = \dfrac{n-1}{n+1} \xrightarrow{n \to \infty} 1$





Theorem: Joint Dist. of All Order Stats.

$$X_{(1)}, X_{(2)}, X_{(3)}, \ldots, X_{(n-1)}, X_{(n)}$$

$$\boxed{f_{X_{(1)},X_{(2)},X_{(3)},X_{(4)},\ldots,X_{(n)}}(u_1,u_2,u_3,\ldots,u_n) = n! \prod_{i=1}^{n} f(u_i)}$$

basically $n! \cdot$ Joint PDF of my random Sample.