Defn: Random Sample

If $X_1, X_2, X_3, \dots, X_n \overset{iid}{\sim} f$ ← f pmf/pdf of same univariate dist.

then we call

$\{X_i\}_{i=1}^{n}$ a random sample from f.
← of size n

Facts: If $\{X_i\}$ are a random sample

then
$$\begin{cases} \underset{\sim}{X} = (X_1, X_2, \dots, X_n)^T \\ \underset{\sim}{x} \in \mathbb{R}^n \end{cases}$$ notation

then

$$\boxed{f_{\underset{\sim}{X}}(\underset{\sim}{x})} = f_{\underset{\sim}{X}}(x_1, x_2, x_3, \dots, x_n)$$

$$= f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) \cdots f_{X_n}(x_n) \quad \text{(independence)}$$

$$= f(x_1) f(x_2) f(x_3) \cdots f(x_n) \quad \text{(identical distributions)}$$

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

Defn: Statistic

If $\{X_i\}$ are a RS and $T: \mathbb{R}^n \to \mathbb{R}^d$

If $\{X_i\}$ are a RS and $T: \mathbb{R}^n \to \mathbb{R}^d$

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

Typically $d < n$ (e.g. $d = 1$).

<u>Ex.</u> 1 Arithmetic mean

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n) = T(\underset{\sim}{X})$$

$d = 1$

here $T(\underset{\sim}{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i$.

2 <u>Sample Variance</u>

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

3 <u>Order Statistics</u>

$X_{(1)}$ = minimum of $\{X_i\} = \min\limits_{i=1,\dots,n} X_i$

$X_{(n)}$ = maximum of $\{X_i\} = \max\limits_{i=1,\dots,n} X_i$

$X_{(r)}$ = r^{th} order statistic

= r^{th} smallest value among $\{X_i\}$

④ <u>range</u>: $R = X_{(n)} - X_{(1)}$

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

<u>Defn</u>: <u>Sampling Distribution of a Statistic</u>

The <u>sampling dist.</u> of a statistic $T(\underline{X})$ is simply the distribution of $T(\underline{X})$

$d=1$
univariate RV

<u>Ex,</u> <u>Order Statistics</u>

Henceforth: $\{X_i\}$ are a random sample.

<u>Minimum</u>: $X_{(1)} = \min\limits_{i=1,\dots,n} X_i$

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

$$\underline{P(X_{(1)} > t)} = P(X_1 > t, X_2 > t, \dots, X_n > t)$$
$$= P(X_1 > t) P(X_2 > t) \cdots P(X_n > t)$$
(independence)

(independence)

$$= \prod_{i=1}^{n} \underbrace{P(X_i > t)}_{1 - F(t)}$$

F = CDF of any X_i

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

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

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

$$\parallel$$

$$F_{X_{(1)}}(t)$$

For cts RVs

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

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

marginal CDF   marginal pdf

Maximum

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

$$= P(X_1 \le t) \cdots P(X_n \le t)$$

$$= F(t) \cdots F(t)$$

$$= F(t)^n$$

For cts:

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

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

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

$$F(x) = 1 - e^{-\lambda x} \text{ for } x > 0$$

Minimum?

$$f_{X_{(1)}}(t) = n\,(1 - F(t))^{n-1} f(t) \qquad \text{for } t > 0$$

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

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

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

this is a pdf of an exponential dist w/ rate $n\lambda$

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

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

e.g. $E[X_{(1)}] = n\lambda$

Maximum:

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

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

General Order Statistic:

$X_{(r)}$ = r^{th} order stat.

= r^{th} smallest among $X_1, \dots, X_n$

$X_{(1)} \neq X_1$

If X_i are continuous

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

CDF of X_i PDF of X_i

Pf.

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

$r-1$ smaller t $X_{(r)}$ $t+\Delta t$ $n-r$

$r-1$ smaller $\quad t \quad \underbrace{X_{(r)}} \quad t+\Delta t \quad$ $n-r$ larger

$$f_{X_{(r)}}(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(t \le X_{(r)} \le t+\Delta t)}{\Delta t} \longleftarrow$$

$$= \lim_{\Delta t \to 0} n \binom{n-1}{r-1} F(t)^{r-1} \overbrace{\left[\frac{F(t+\Delta t) - F(t)}{\Delta t}\right]}^{dF/dt} (1-F(t+\Delta t))^{n-r}$$

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

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

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

$f(t) = 1$ for $0 < t < 1$

$F(t) = t$ for $0 < t < 1$

So

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

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

$$= \boxed{\frac{n!}{(r-1)!\,(n-r)!}\, t^{r-1}(1-t)^{n-r}} \quad \text{for } 0<t<1$$

Notice: a Beta

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

Theorem: Joint Distribution of Order Stats.

If $r < s$ and consider $X_{(r)}$ and $X_{(s)}$

note: $X_{(r)} < X_{(s)}$.

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

pf.

$$F(u)^{r-1} \quad f(u) \quad (F(v)-F(u))^{s-r-1} \quad f(v) \quad (1-F(v))^{n-s}$$

$r-1$ $\quad u \quad$ $s-r-1$ $\quad v \quad$ $n-s$

smaller $X_{(r)}$ between $X_{(s)}$

$\underline{1}$ $\underline{1}$

Ex. $X_i \overset{iid}{\sim} U(0,1)$ $\quad$ $F(t)=t,\ f(t)=1 \quad 0<t<1$

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

Ex. U: $\boxed{R = X_{(n)} - X_{(1)}}$. $\quad$ $X_i \overset{iid}{\sim} U(0,1)$

what is the dist of R?

$\boxed{V = X_{(1)}}$ $\qquad$ $(X_{(1)}, X_{(n)}) \xrightarrow{g} (R, V)$

$\xleftarrow{g^{-1}}$

Inverse transf.

$X_{(1)} = \boxed{g_1^{-1}(R,V) = V}$ $\quad$ $X_{(n)} = \boxed{g_2^{-1}(R,V) = R+V}$

$$\frac{\partial g^{-1}}{\partial(R,V)} = J = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$|\det J| = 1$

$$f_{R,V}(r,v) = f_{X_{(1)},X_{(n)}}(v, r+v)\cdot 1$$

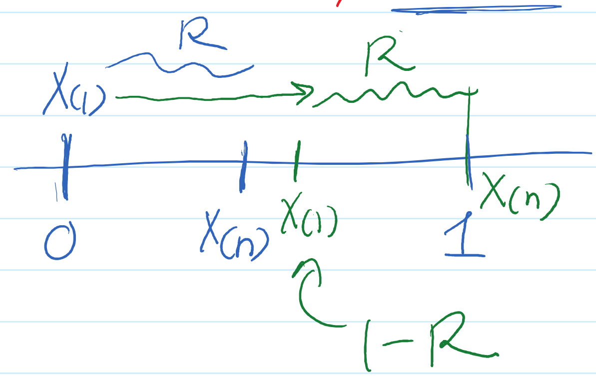
Aside:

$$f_{X_{(1)},X_{(n)}}(a,b) = n(n-1)\overset{(b-a)^{n-2}}{(F(b)-F(a))^{n-2}}\,\underset{1}{\cancel{f(a)}}\,\underset{1}{\cancel{f(b)}}$$

$$\boxed{f(r,v) = n(n-1)(r+v-v)^{n-2}}$$

joint of $R = X_{(n)} - X_{(1)}$, $V = X_{(1)}$.

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



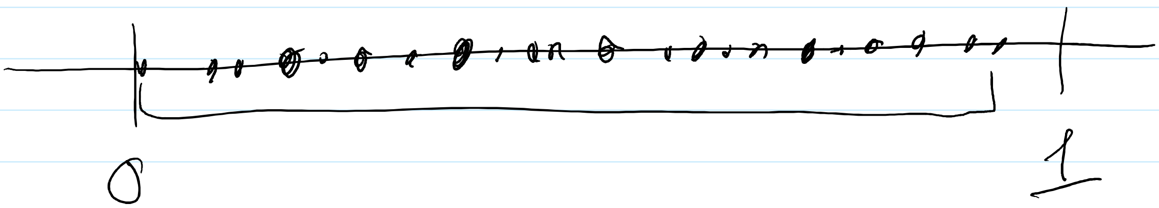
$$f_R(r) = \int_0^{1-r} n(n-1)r^{n-2}\,dv$$

$$= n(n-1)r^{n-2}\int_0^{1-r} dv$$

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

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

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



Theorem:

Joint dist of all order stats

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