STAT 712 fa 2022 Lec 13 slides

Order statistics

continuous random sample

$X_1, \dots, X_n \overset{\text{ind}}{\sim} f_X$

iid

pdf

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Order statistics

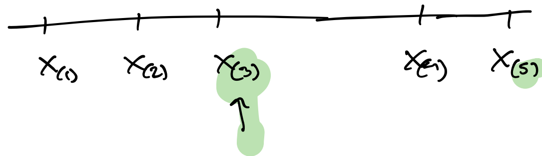
Given a random sample $X_1, \ldots, X_n$, define

$$\begin{aligned} X_{(1)} &= \text{ the least of } X_1, \ldots, X_n \\ X_{(2)} &= \text{ the next-to-least of } X_1, \ldots, X_n \\ &\vdots \\ X_{(n)} &= \text{ the greatest of } X_1, \ldots, X_n. \end{aligned}$$

Then $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ are called the *order statistics* of the rs.

Exercise: Define range, midrange, and median with order statistics.

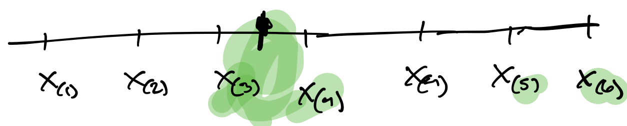
$R_n = X_{(n)} - X_{(1)}$



$M_n = \dfrac{X_{(1)} + X_{(n)}}{2}$

$$\text{Median}_n = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & n \text{ odd} \\[2em] \dfrac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} & n \text{ even} \end{cases}$$

Sampling dist. of these?



Theorem (pdf of kth order statistic)

Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of a rs with cdf F_X and pdf f_X.

Then the pdf of $X_{(k)}$ is given by

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}[F_X(x)]^{k-1}[1 - F_X(x)]^{n-k} f_X(x),$$

for $k = 1, \dots, n$.

Exercises:

1. Derive the above.
2. Let $U_1, \dots, U_n \stackrel{\text{ind}}{\sim} \text{Uniform}(0,1)$. Find pdf of $U_{(k)}$, find $\mathbb{E} U_{(k)}$ and $\text{Var}\, U_{(k)}$.
3. Draw samples $X_1, \dots, X_n \stackrel{\text{ind}}{\sim} \text{Uniform}(0,1)$ and record kth order statistic. Make histograms. Use $n = 10$ and consider $k = 1, 5, 9$.

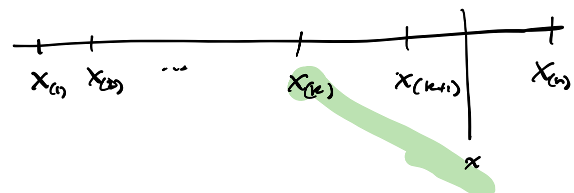
Pdf of $X_{(k)}$. Begin by finding cdf:

$$F_{X_{(k)}}(x) = P\left(X_{(k)} \le x\right)$$

$$= P\left(\text{at least } k \text{ of } X_1,\dots,X_n \text{ are less than or equal to } x\right)$$

$$= P\left(Y \ge k\right) \qquad Y = \#\ X_1,\dots,X_n \le x$$

$$= \sum_{y=k}^{n} \binom{n}{y} \left[F_X(x)\right]^{y} \left[1 - F_X(x)\right]^{n-y} \qquad Y \sim \text{Binomial}\left(n, F_X(x)\right)$$



$$f_{X_{(k)}}(x) = \frac{d}{dx} F_{X_{(k)}}(x)$$

$$= \frac{d}{dx} \sum_{y=k}^{n} \binom{n}{y} \left[F_X(x)\right]^{y} \left[1 - F_X(x)\right]^{n-y}$$

$$= \sum_{y=k}^{n} \binom{n}{y} \Big\{ y \left[F_X(x)\right]^{y-1} f_X(x) \left[1 - F_X(x)\right]^{n-y} - \left[F_X(x)\right]^{y} (n-y) \left[1 - F_X(x)\right]^{n-y-1} f_X(x) \Big\}$$

$y=k$

$$= \binom{n}{k} k \left[F_X(x)\right]^{k-1} \left[1 - F_X(x)\right]^{n-k} f_X(x)$$

$$\underbrace{- \left[F_X(x)\right]^{k} (n-k) \left[1 - F_X(x)\right]^{n-k-1} f_X(x) + \sum_{y=k+1}^{n} \binom{n}{y} \Big\{ y \left[F_X(x)\right]^{y-1} f_X(x) \left[1 - F_X(x)\right]^{n-y} - \left[F_X(x)\right]^{y} (n-y) \left[1 - F_X(x)\right]^{n-y-1} f_X(x) \Big\}}_{=0}$$

$$\boxed{= \frac{n!}{(k-1)!\,(n-k)!}\left[F_X(x)\right]^{k-1}\left[1-F_X(x)\right]^{n-k} f_X(x)}$$

② $U_1, \dots, U_n \overset{ind}{\sim} U(0,1)$. Find pdf of $U_{(k)}$.

$$f(u) = 1 \cdot \mathbb{1}(0<u<1)$$

$$F(u) = u \quad \text{for} \quad 0<u<1$$

$$f_{U_{(k)}}(u) = \frac{n!}{(k-1)!\,(n-k)!}\,[u]^{k-1}\,[1-u]^{n-k} \cdot 1 \cdot \mathbb{1}(0<u<1)$$

$$= \frac{\Gamma(k+n-k+1)}{\Gamma(k)\,\Gamma(n-k+1)}\, u^{k-1}\,(1-u)^{(n-k+1)-1}\,\mathbb{1}(0<u<1)$$

$$= \frac{\Gamma(n+1)}{\Gamma(k)\,\Gamma(n-k+1)}\, u^{k-1}\,(1-u)^{(n-k+1)-1}\,\mathbb{1}(0<u<1)$$

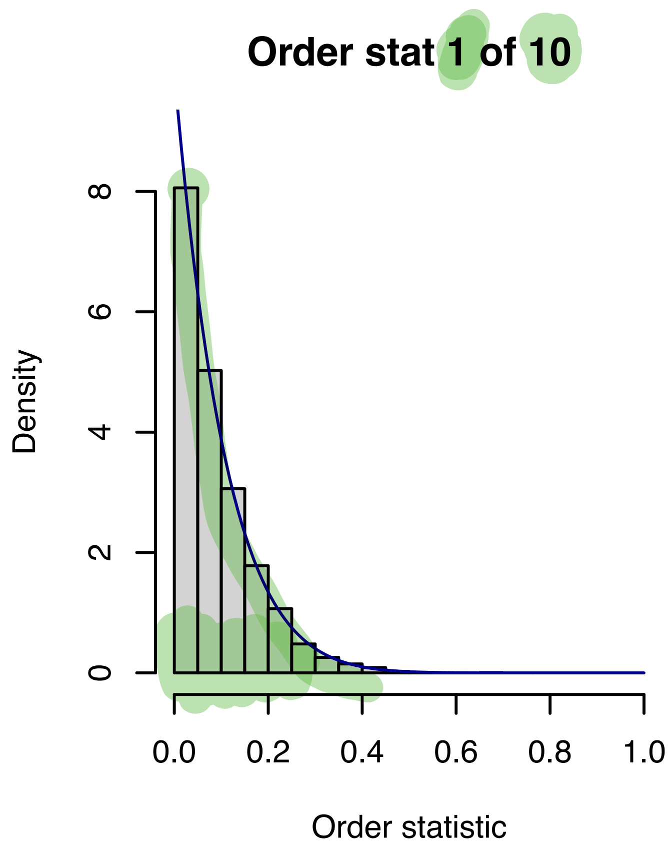
so

$$U_{(k)} \sim \text{Beta}(\alpha = k,\ \beta = n-k+1)$$

$\Rightarrow$

$$\mathbb{E}\, U_{(k)} = \frac{\alpha}{\alpha+\beta} = \frac{k}{k+n-k+1} = \frac{k}{n+1}$$

$$\text{Var}\, U_{(k)} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \dots .$$



Order stat 1 of 10
Order stat 5 of 10
Order stat 9 of 10
Density
Order statistic
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
1.0

$$F_X(x) = \frac{1}{1+e^{-x}}$$

To sample, find Quantile function, pass Unif(0,1) rvs through it.

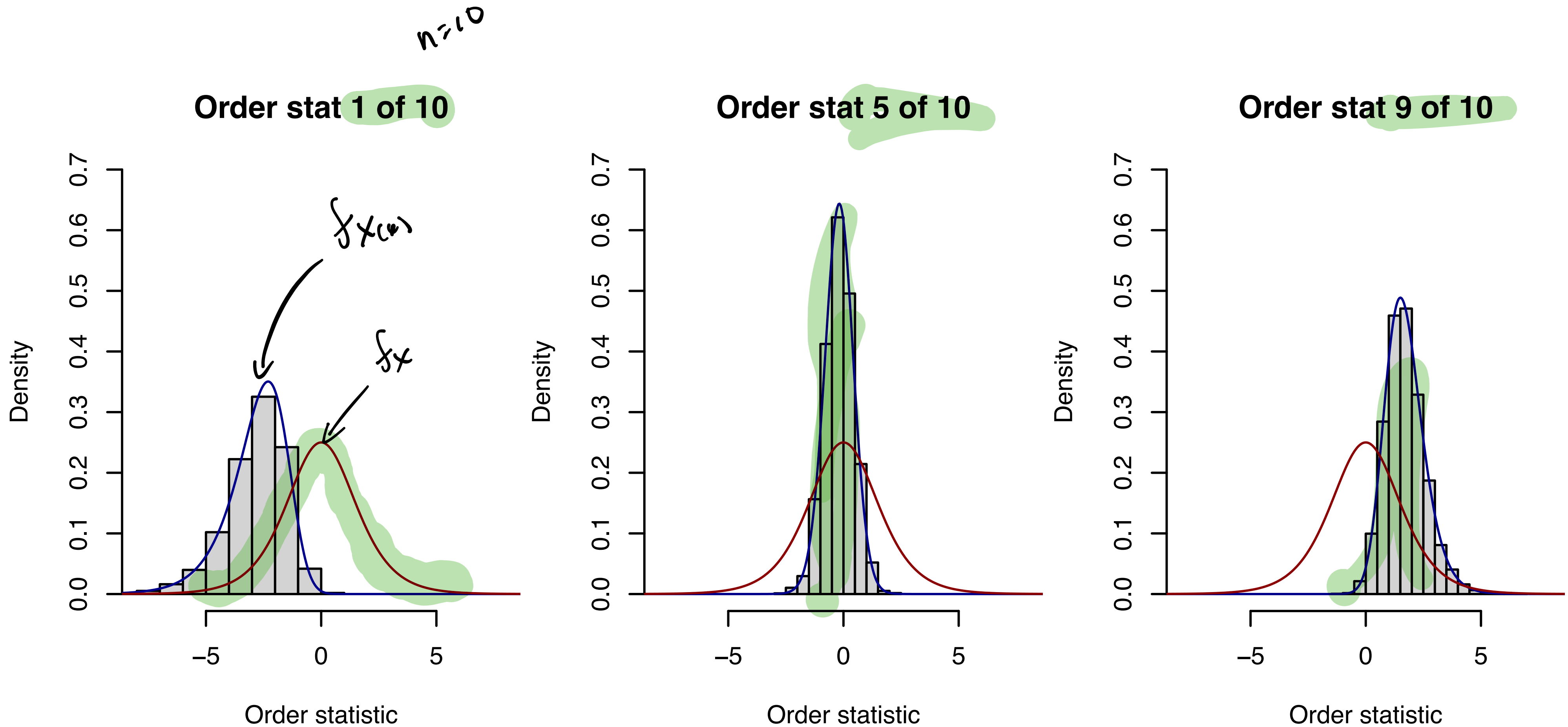
$$f_X(x) = \frac{d}{dx} F_X(x) = -\frac{1}{(1+e^{-x})^2} e^{-x} (-1) = \frac{e^{-x}}{(1+e^{-x})^2}$$

Exercise: Let $X_1, \ldots, X_n$ be independent rvs with cdf $F_X(x) = (1 + e^{-x})^{-1}$.

1. Find the pdf of the kth order statistic.
2. Draw samples $X_1, \ldots, X_n$ from F_X and record kth order statistic. Make histograms. Use $n = 10$ and consider $k = 1, 5, 9$.

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k+1)!} [F_X(x)]^{k-1} [1 - F_X(x)]^{n-k-1} f_X(x)$$

$$= \frac{n!}{(k-1)!(n-k+1)!} \left(\frac{1}{1+e^{-x}}\right)^{k-1} \left(1 - \frac{1}{1+e^{-x}}\right)^{n-k-1} \frac{e^{-x}}{(1+e^{-x})^2}$$


n=10
Order stat 1 of 10
Order stat 5 of 10
Order stat 9 of 10
$f_{X_{(1)}}$
f_X
Density
Order statistic
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
–5
0
5

$k=1$ or $k=n$

Corollary (pdf of minimum and maximum)

Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of a rs with cdf F_X and pdf f_X. Then

- *$X_{(1)}$ has cdf and pdf given by*

$$F_{X_{(1)}}(x) = 1 - [1 - F_X(x)]^n$$
$$f_{X_{(1)}}(x) = n[1 - F_X(x)]^{n-1} f_X(x)$$

- *$X_{(n)}$ has cdf and pdf given by*

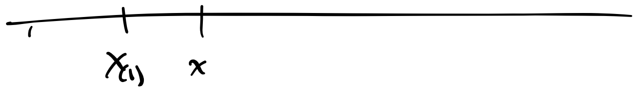
$$F_{X_{(n)}}(x) = [F_X(x)]^n$$
$$f_{X_{(n)}}(x) = n[F_X(x)]^{n-1} f_X(x)$$

* Derive the above

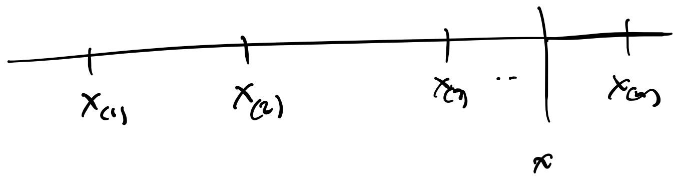
Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Exponential}(\lambda)$.

1. Find the pdf of $X_{(n)}$.
2. Find the pdf of $X_{(1)}$ and identify the distribution.

$$F_{X_{(1)}}(x) = P\left(X_{(1)} \le x\right)$$



$$= P\begin{pmatrix} \text{at least } 1 \\ \text{of } X_1, \ldots, X_n \\ \text{is } \le x \end{pmatrix}$$



$Y = \#\ X_1, \ldots, X_n \le x$

$Y \sim \text{Binom}(n, F_X(x))$

$P_Y(y) = \binom{n}{y}\left[F_X(x)\right]^y\left[1 - F_X(x)\right]^{n-y}$

$$= P(Y \ge 1)$$

$$= 1 - P(Y = 0)$$

$$= 1 - \underbrace{\binom{n}{0}}_{=1}\left[F_X(x)\right]^0\left[1 - F_X(x)\right]^{n-0}$$

$$= 1 - \left[1 - F_X(x)\right]^n$$

$$f_{X_{(1)}}(x) = -n\left[1 - F_X(x)\right]^{n-1}\left(-f_X(x)\right)$$

$$= n\left[1 - F_X(x)\right]^{n-1} f_X(x)$$

$$F_{X_{(n)}}(x) = P\left(X_{(n)} \le x\right)$$

$$= P\left(\text{all } X_1, \ldots, X_n \le x\right)$$

$$= P(Y = n)$$

$$= \binom{n}{n}\left[F_X(x)\right]^n\left[1 - F_X(x)\right]^{n-n}$$

$$= \left[F_X(x)\right]^n$$

$$f_{X_{(n)}}(x) = n\left[F_X(x)\right]^{n-1} f_X(x)$$

Let $\quad X_1, \dots, X_n \overset{\text{iid}}{\sim} f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \mathbb{1}(x>0)$

$$F_X(x) = \int_0^x \frac{1}{\lambda} e^{-t/\lambda} \, dt$$

$$= 1 - e^{-x/\lambda} \quad \text{for } x > 0.$$

So

$$f_{X_{(n)}}(x) = n \left[1 - e^{-x/\lambda} \right]^{n-1} \frac{1}{\lambda} e^{-x/\lambda} \; \mathbb{1}(x>0)$$

And

$$f_{X_{(1)}}(x) = n \left[1 - F_X(x) \right]^{n-1} f_X(x)$$

$$= n \left[1 - \left(1 - e^{-x/\lambda} \right) \right]^{n-1} \frac{1}{\lambda} e^{-x/\lambda} \; \mathbb{1}(x>0)$$

$$= \frac{n}{\lambda} \left[e^{-x/\lambda} \right]^{n-1} e^{-x/\lambda} \; \mathbb{1}(x>0)$$

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

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

$$= \frac{1}{(\lambda/n)} e^{-x/(\lambda/n)} \; \mathbb{1}(x>0).$$

So $\quad X_{(1)} \sim \text{Exponential}\left(\lambda/n \right).$

Theorem (Joint pdf of two order statistics)

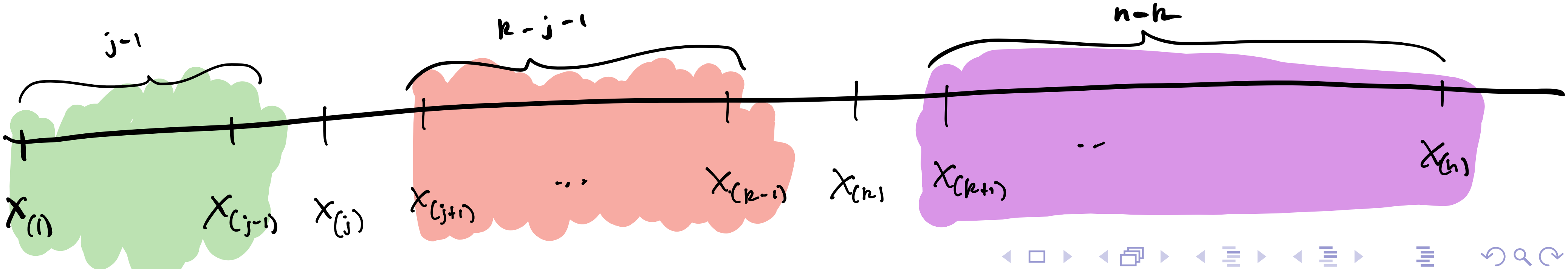
Let $X_{(1)}, \dots, X_{(n)}$ *be the order statistics of a rs with cdf* F_X *and pdf* f_X*.*

Then the joint pdf of $X_{(j)}$ *and* $X_{(k)}$*,* $1 \le j < k \le n$*, is given by*

$$f_{X_{(j)},X_{(k)}}(u,v) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!} f_X(u) f_X(v)$$
$$\times [F_X(u)]^{j-1} [F_X(v) - F_X(u)]^{k-j-1} [1 - F_X(v)]^{n-k}$$

$P(X \le x_{(j)})$ $\quad P(x_{(j)} < X < x_{(k)})$ $\quad P(X > x_{(k)})$

for $\boxed{-\infty < u < v < \infty.}$



$j < k$

$$f_{X_{(j)},X_{(k)}}(\overset{x_{(j)}}{u}, \overset{x_{(k)}}{v}) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!} f_X(u) f_X(v)$$
$$\times [F_X(u)]^{j-1}[F_X(v) - F_X(u)]^{k-j-1}[1 - F_X(v)]^{n-k}$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(0,1)$.

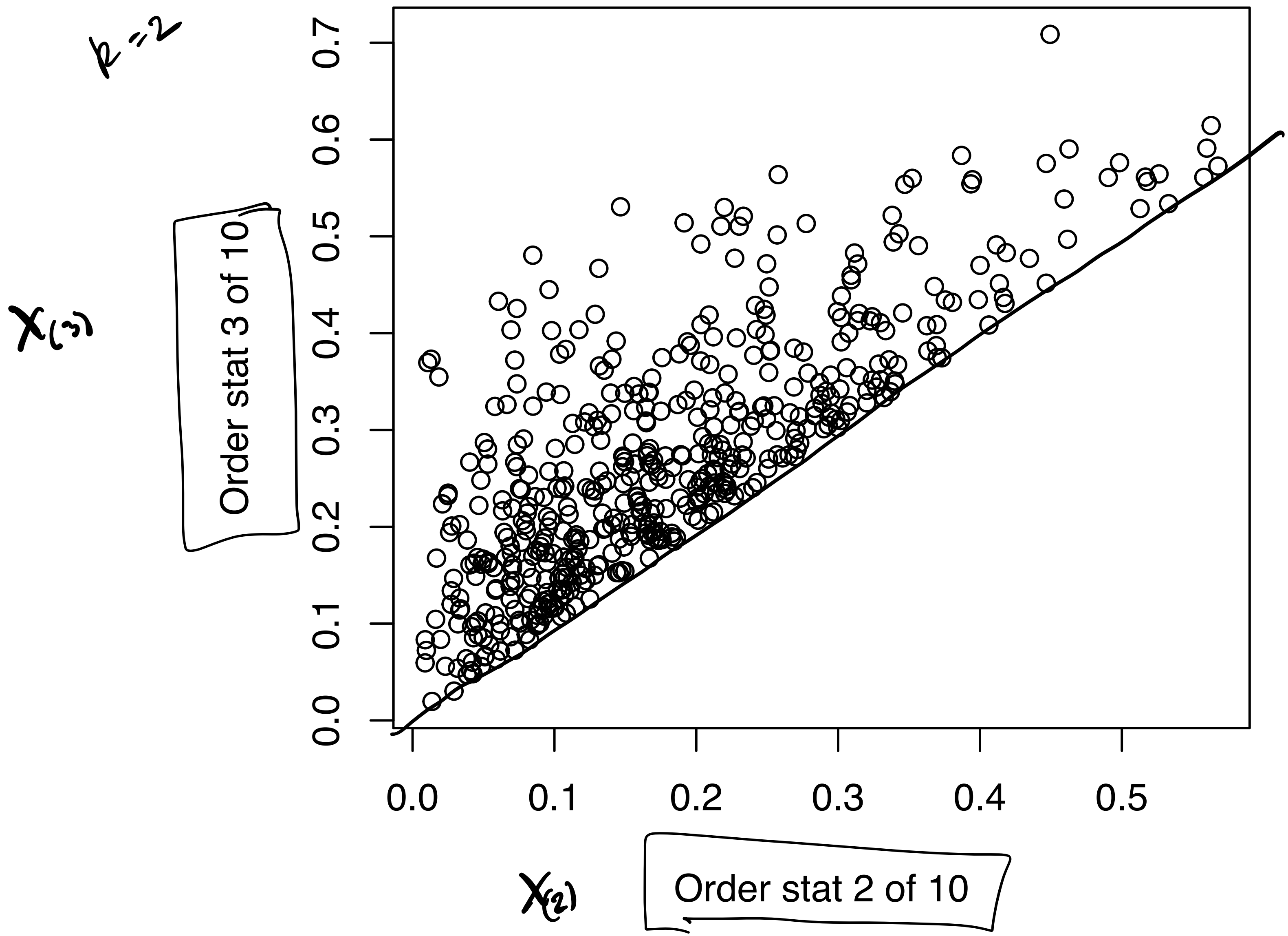
$F_X(x) = x$ for $x \in (0,1)$.

1. Find the joint pdf of the order statistics $U = X_{(k)}$ and $V = X_{(k+1)}$.
2. Draw samples $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(0,1)$ and record $(X_{(k)}, X_{(k+1)})$. Make a scatterplot of the values. Use $n = 10$, $k = 2$.

$$f_{U,V}(u,v) = \frac{n!}{(k-1)!\,(k+1-k-1)!\,(n-(k+1))!} 1 \cdot 1 \cdot [u]^{k-1} [v-u]^{k+1-k-1} [1-v]^{n-(k+1)} \mathbb{I}(0<u<v<1)$$

$$= \frac{n!}{(k-1)!\,(n-k-1)!} u^{k-1} (1-v)^{n-k-1} \mathbb{I}(0<u<v<1)$$

$k=2$



$X_{(3)}$
Order stat 3 of 10
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
0.1
0.2
0.3
0.4
0.5
$X_{(2)}$
Order stat 2 of 10

Corollary (Joint pdf of min and max)

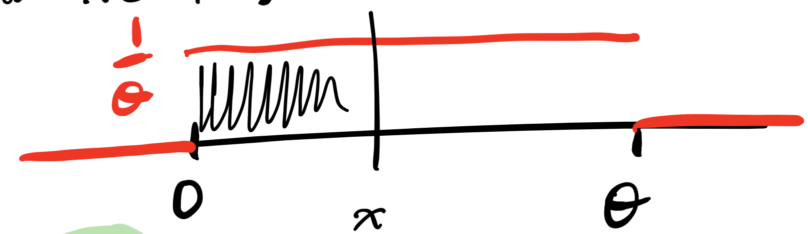
Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of a rs with cdf F_X and pdf f_X.

The joint pdf of $X_{(1)}$ and $X_{(n)}$ is given by

$$f_{X_{(1)},X_{(n)}}(u,v) = n(n-1)f_X(u)f_X(v)[F_X(v) - F_X(u)]^{n-2}$$

for $-\infty < u < v < \infty$.

$f_X(x) = \frac{1}{\theta} \mathbb{1}(0 < x < \theta)$ $\quad F_X(x) = \frac{x}{\theta}$ for $x \in (0, \theta)$



Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(0, \theta)$.

1. Find the joint pdf of $X_{(1)}$ and $X_{(n)}$.
2. Find the joint pdf of $R = X_{(n)} - X_{(1)}$ and $M = X_{(n)}$.
3. Find the marginal pdf of R.
4. Draw realizations of the range of Uniform$(0, 1)$ samples with $n = 4, 8, 16$. Make histograms and overlay densities.

$$f_{X_{(j)},X_{(k)}}(u,v) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!} f_X(u) f_X(v) \times \cancel{[F_X(u)]^{j-1}} [F_X(v)-F_X(u)]^{k-j-1} \cancel{[1-F_X(v)]^{n-k}}$$

$j=1, \quad k=n$

$$= \frac{n!}{(1-1)!\,(n-1-1)!\,(n-n)!} f_X(u) f_X(v) \left[F_X(v) - F_X(u)\right]^{n-2} \mathbb{1}(u<v)$$

$$= \frac{n!}{(n-2)!} = n(n-1)$$

$$f_X(x) = \frac{1}{\theta} \mathbb{1}(0<x<\theta) \qquad F_X(x) = \frac{x}{\theta} \text{ for } x \in (0,\theta)$$

① $$f_{X_{(1)},X_{(n)}}(u,v) = n(n-1) \frac{1}{\theta} \frac{1}{\theta} \left[\frac{v}{\theta} - \frac{u}{\theta}\right]^{n-2} \mathbb{1}(0<u<v<\theta)$$

$$= \frac{n(n-1)}{\theta^n} [v-u]^{n-2} \mathbb{1}(0<u<v<\theta).$$

② $R = \overset{v}{X_{(n)}} - \overset{u}{X_{(1)}} \qquad (R,M) \in \{(r,m) : \; 0<r<m<\theta\}$

$M = X_{(n)}$

$$\begin{aligned} r &= v-u = g_1(u,v) & & u = m-r = g_1^{-1}(m,r) \\ m &= v = g_2(u,v) & \iff & v = m = g_2^{-1}(m,r) \end{aligned}$$

$$J(m,r) = \begin{vmatrix} \frac{\partial}{\partial r} m-r & \frac{\partial}{\partial r} m \\ \frac{\partial}{\partial m} m-r & \frac{\partial}{\partial m} m \end{vmatrix} = \begin{vmatrix} -1 & 0 \\ 1 & 1 \end{vmatrix} = -1$$

$$f_{R,M}(r,m) = \frac{n(n-1)}{\theta^n} \left(m - (m-r)\right)^{n-2} \mathbb{1}(0 < r < m < \theta)$$

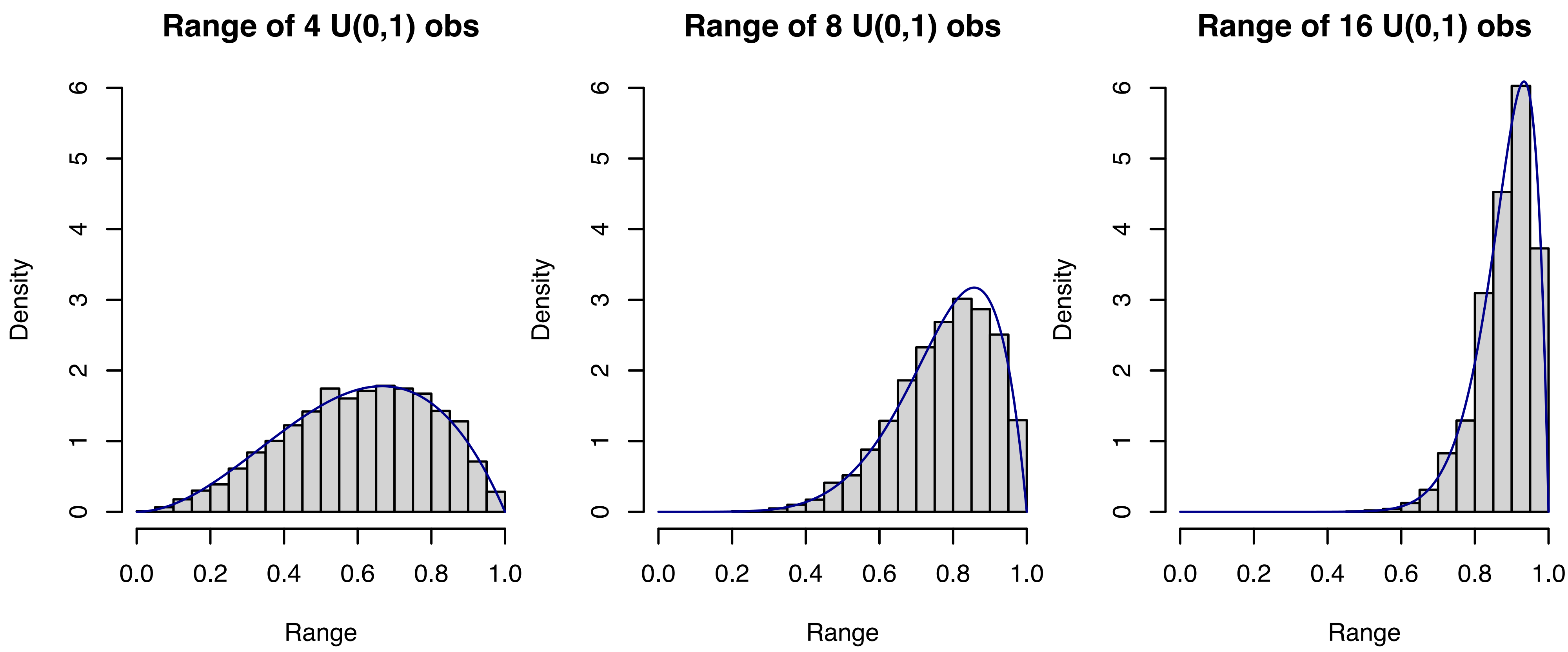
$$= \frac{n(n-1)}{\theta^n} r^{n-2} \mathbb{1}(0 < r < m < \theta)$$

③ Find marginal pdf of R. For $r \in (0,\theta)$

$$f_R(r) = \int_r^\theta \frac{n(n-1)}{\theta^n} r^{n-2} \, dm$$

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

$$f_R(r) = \frac{n(n-1)}{\theta} \left(\frac{r}{\theta}\right)^{n-2} \left(1 - \frac{r}{\theta}\right) \mathbb{1}(0 < r < \theta).$$



Range of 4 U(0,1) obs
Range of 8 U(0,1) obs
Range of 16 U(0,1) obs
Density
Range
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6