This document contain's solution's to Exercises 5.21 to 5.28 from Chapter 5, of statistical inference by Casella and Berger, focusing on Order Statistics

Some problem's are still in progress, and the document will be updated with their solution's soon.

5.21 what is the Probability that the larger of two continuous iid r.v's will exceed the the Population median? Generalize this result to samples of size m.

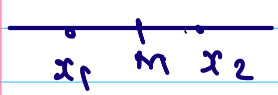
Sol^n $X_1, X_2 \overset{iid}{\sim} X$

let $X_{(1)}, X_{(2)}$ be the order statistic's then the Joint Pdf of $X_{(1)}, X_{(2)}$ is

$$M_{pop} = F_X(m) = 1 - F_X(m) = \frac{1}{2}$$

$$= \int_{-\infty}^{m} f_X(x)\,dx = \int_{m}^{\infty} f_X(x)\,dx = \frac{1}{2}$$

Scenario 1



Scenario 2

we can calculate $\mathbb{P}(X_{(2)} < M)$ which is Probability that Both $x_1, x_2 < m$

& subtract from 1 $\Rightarrow 1 - \mathbb{P}(X_{(2)} < M)$

$\Rightarrow$ This gives probability that $x_2 > M$

<u>therefore:</u>

$$\mathbb{P}(X_{(2)} < M) = \mathbb{P}(Y \geq 2) = \mathbb{P}(Y = 2)$$

$$= F_X(m)$$

$$\boxed{\Rightarrow \quad 1 - F_X(m)^2 \Rightarrow 1 - \left(\frac{1}{2}\right)^2}$$

<u>General n case:</u>

$x_1, x_2, \dots x_n$ iid X

$X_{(1)}, X_{(2)} \dots, X_{(n)}$ order statistics.

$\Rightarrow$ The largest $(X_{(n)})$ need's to be greater than median.

$$\Rightarrow \quad 1 - \mathbb{P}(X_{(n)} < m) = 1 - \mathbb{P}(Y \geq n)$$

$$= 1 - \mathbb{P}(Y = n)$$

$$\Rightarrow 1 - \mathbb{P}(X_{(n)} < m) = 1 - F_X(m)^n$$

$$\Rightarrow \quad \boxed{1 - \left(\frac{1}{2}\right)^n}$$

(5.22) Let X and Y be iid N(0,1) r.v, and define $Z = \min(X, Y)$. Prove that

$Z^2 \sim \chi_1^2$

Sol^n $X_1, X_2 \overset{iid}{\sim} N(0,1)$

$X_{(1)}, X_{(2)}$ order statistics

$$\Rightarrow \quad Z = \min(X_1, X_2) = X_{(1)}$$

$$\mathbb{P}(X_{(1)} < x) = \mathbb{P}(Y \geq 1)$$

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

$$f_{X_{(1)}}(x) = \frac{2!}{(1-1)!(2-1)!} f_X(x) \left[F_X(x)\right]^{1-1} \left(1 - F_X(x)\right)^{2-1}$$

$$\Rightarrow f_{X_{(1)}}(x) = 2 \cdot f_X(x) \left(1 - F_X(x)\right)$$

$$= 2 \cdot \phi(x) \cdot \left(1 - \Phi(x)\right)$$

$$f_Z(z) = 2 \cdot \phi(z)\,(1 - \Phi(z))$$

$$Y = Z^2 \Rightarrow \quad Z = -\sqrt{Y} \qquad -\infty < z < 0$$

$$Z = \sqrt{Y} \qquad 0 < z < \infty$$

$$f_Y(y) = 2 \cdot \phi(-\sqrt{Y})\,(1 - \Phi(-\sqrt{Y})) \cdot \frac{1}{2\sqrt{Y}}$$

$$+ 2 \cdot \phi(\sqrt{Y})\,(1 - \Phi(+\sqrt{Y})) \cdot \frac{1}{2\sqrt{Y}}$$

$$\boxed{\begin{array}{l} \Phi(-\sqrt{Y}) = 1 - \Phi(\sqrt{Y}) \\ \phi(-\sqrt{Y}) = \phi(\sqrt{y}) \end{array}}$$

$$f_Y(y) = \frac{1}{\sqrt{Y}}\left(\phi(\sqrt{y})\,\Phi(\sqrt{y}) + \phi(\sqrt{y})\,(1 - \Phi(\sqrt{y}))\right]$$

$$= \frac{1}{\sqrt{y}}\left(\phi(\sqrt{y}) \right)$$

$$\Rightarrow f_Y(y) = \frac{1}{\sqrt{y}} \phi(\sqrt{y})$$

$$= \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}$$

$$\Rightarrow \boxed{Y = Z^2 \sim \chi^2_1}$$

5.23 Let U_i, $i = 1, 2, \dots$, be independent uniform $(0,1)$ r.v, and X have distribution.

$$\mathbb{P}(X = x) = \frac{c}{x!} \qquad x = 1, 2, 3, \dots$$

where $c = \frac{1}{e-1}$. Find the distribution of $Z = \min\{U_1, U_2, \dots U_X\}$

Sol^n

$$Z = \min\{U_1, U_2, \dots, U_X\}$$

$$Z = U_{(1)}$$

where $U_{(1)}, U_{(2)}, \ldots U_{(x)}$ is
order statistic.

$$\Rightarrow \quad Z \mid X=x = U_{(1)} \quad (\text{for fixed } X=x)$$

$$f_U(u) = 1, \quad F_u(u) = u \qquad 0<u<1$$

$$f_{Z|X}(z|X=x) = \frac{x!}{(1-1)!\,(x-1)!}\, 1 \cdot (z)^{1-1} \cdot (1-z)^{x-1}$$

$$\boxed{f_{Z|X}(z|X=x) = x \cdot (1-z)^{x-1}}$$

$$f_Z(z) = \int_x f_{Z,X}(z,x)$$

$$= \int_{x} f_{Z|X}(z|X=x) \cdot \mathbb{P}(X=x)\, dx$$

$$f_Z(z) = \sum_{x=1}^{\infty} x(1-z)^{x-1} \cdot \frac{c}{x!}$$

$$= \sum_{x=1}^{\infty} \frac{(1-z)^{x-1} \cdot c}{(x-1)!}$$

$$= c \underbrace{\sum_{y=0}^{\infty} \frac{(1-z)^{y}}{y!}}_{\text{Possion}(1-z)}$$

$$= \frac{c}{e^{z-1}} = \frac{1}{(e-1)} \cdot \frac{1}{e^{z-1}}$$

$$\boxed{\Rightarrow \quad f_Z(z) = \frac{e^{1-z}}{e-1} \qquad 0<z<1}$$

$$F_Z(z) = \int_0^z \frac{e^{1-z}}{e-1}\, dz = \frac{1}{e-1} \int_{1-z}^{1} e^{p}\, dp$$

$1-z=p$
$-dz=dp$
$p: 1 \to 1-z$

5.24 Let $x_1, x_2, \ldots x_n$ be a random sample from a population with pdf

$$f_X(x) = \begin{cases} 1/\theta & \text{if } 0 < x < \theta \\ 0 & \text{o.w} \end{cases}$$

Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics. Show that $\frac{X_{(1)}}{X_{(2)}}$ and $X_{(n)}$ are independent r.v.

Sol^n

Joint pdf of $X_{(1)}, X_{(n)}$

$$f_{X_{(1)}, X_{(n)}}(x_1, x_n) = \frac{n(n-1)(x_n - x_1)^{n-2}}{\theta^n}$$

$0 < x_1 < x_n < \theta$

Now take $U = \frac{X_{(1)}}{X_{(n)}}$ $V = X_{(n)}$

$$\Rightarrow \left.\begin{array}{l} X_{(1)} = UV \\ X_{(n)} = V \end{array}\right\} \Rightarrow J = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

$$f_{U,V}(u,v) = \frac{n(n-1)}{\theta^n}(v-uv)^{n-2}$$

$$\Rightarrow f_{U,V}(u,v) = \frac{n(n-1)}{\theta^n} \cdot v^{n-2}(1-u)^{n-2}$$

$0 < u < 1$

$0 < v < \theta$

$$\Rightarrow f_{U,V}(u,v) \text{ is } = h(u)\,g(v)$$

$$\boxed{\Rightarrow U = \frac{X_{(1)}}{X_{(n)}}, \quad V = X_{(n)} \text{ are independent}}$$

5.25 As a generalization of the previous exercise, let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} X$

$$f_X(x) = \begin{cases} \frac{a}{\theta^a} x^{a-1} & 0 < x < \theta \\ 0 & \text{o.w} \end{cases}$$

Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistic.

Show that $\frac{X_{(1)}}{X_{(2)}}, \frac{X_{(2)}}{X_{(3)}}, \frac{X_{(3)}}{X_{(4)}} \cdots \frac{X_{(n-1)}}{X_{(n)}}$

and $X_{(n)}$ are mutualy independent r.v's. Find the distribution of each of them.

Sol^n

The Joint pdf of all the order statistics is

$$f_{X_{(1)}, X_{(2)} \cdots X_{(n)}}(x_1, x_2, \cdots x_n) = \begin{cases} n!\, f_X(x_1) f_X(x_2) \cdots f_X(x_n) & -\infty < x_1 < \cdots < x_n < \infty \\ 0 & o.w \end{cases}$$

$$\Rightarrow f_{X_{(1)}, X_{(2)} \cdots X_{(n)}}(x_1, x_2, \cdots x_n) = n!\left(\frac{a}{\theta^a}\right)^n x_1^{a-1} x_2^{a-1} \cdots x_n^{a-1}$$

$$0 < x_1 < x_2 < \cdots < x_n$$

$U_1 = \frac{X_{(1)}}{X_{(2)}}, U_2 = \frac{X_{(2)}}{X_{(3)}}, \cdots, U_{n-1} = \frac{X_{(n-1)}}{X_{(n)}}$

$U_n = X_{(n)}$

$$\Rightarrow \left.\begin{aligned} X_{(n)} &= U_n \\ X_{(n-1)} &= U_{n-1}U_n \\ X_{(n-2)} &= U_{n-2}U_n \\ &\vdots \\ X_{(2)} &= U_2U_3\cdots U_n \\ X_{(1)} &= U_1U_2U_3\cdots U_n \end{aligned}\right\} \quad J = \begin{bmatrix} \prod\limits_{i\neq 1} U_i & \prod\limits_{i\neq 2} U_i & \cdots & \prod\limits_{i\neq n} U_i \\ 0 & \prod\limits_{i\neq 1,2} U_i & \cdots & \prod\limits_{i\neq 1,n} U_i \\ 0 & 0 & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

upper triangle matrix

det J = Product of diagonal's.

$$|J| = U_2U_3^2U_4^3\cdots U_n^{n-1}$$

$$f_{U_1,U_2,\ldots U_n}(u_1,u_2,\ldots u_n) = \frac{n!\,a^n}{\theta^{an}}\cdot\left(u_1u_2^2\cdots u_n^n\right)^{a-1}\cdot u_2u_3^2\cdots u_n^{n-1}$$

$$= \frac{n!\,a^n}{\theta^{an}}\cdot u_1^{a-1}\cdot u_2^{2a-1}\cdot u_3^{3a-1}\cdots u_n^{na-1}$$

$$0 < u_1,u_2,\ldots u_{n-1} < 1$$

$$0 < u_n < \theta$$

$\Rightarrow$ as we can see we can divide $U_1, U_2, \dots U_n$ (seperate)

$\Rightarrow$ $\dfrac{X_{(1)}}{X_{(2)}}, \dfrac{X_{(2)}}{X_{(3)}}, \dots, \dfrac{X_{(n-1)}}{X_n}, X_{(n)}$

mutually independent.

5.26 Complete the proof of Theorem 5.4.6

Theorem: $X_{(1)}, \dots X_{(n)}$ order statistic

& $X_1, X_2, \dots X_n \overset{iid}{\sim} X$, $f_X(x)$, $F_X(x)$

Then the Joint pdf of $X_{(i)}$ & $X_{(j)}$, $1 \le i \le j < n$ is

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

$$\times \left[F_X(v) - F_X(u)\right]^{j-1-i} \left[1 - F_X(v)\right]^{n-j}$$

$-\infty < u < v < \infty$

a) Let U be a r.v that counts the number's of $X_1, X_2 \dots, X_n$ less than or equal to u, and let V be a random variable that counts the number's of $X_1, X_2, \dots X_n$ greater than u and less or equal to v. Show that $(U, V, n-U-V)$

is a multinomial random vector with n trails and cell Prob $\left(F_X(u), F_X(v) - F_X(u), 1 - F_X(v) \right)$

Sol^n

$X_1, X_2, \dots X_n$

① $U \equiv \#$ Success, if $X_i < u$, $\forall i$

$$\mathbb{P}(U = u) = \binom{n}{u} F_X(u)^{u} \left[1 - F_X(u)\right]^{n-u}$$

$u = 0, 1, 2, \dots n$

$U \sim \text{binomial}(n, F_X(u))$

② $V = \#$ success if $u < X_i < v$, $\forall i$

$$\mathbb{P}(V = v) = \binom{n}{v} \left(F_X(v) - F_X(u)\right)^{v} \left(1 - F_X(v) + F_X(u)\right)^{n-v}$$

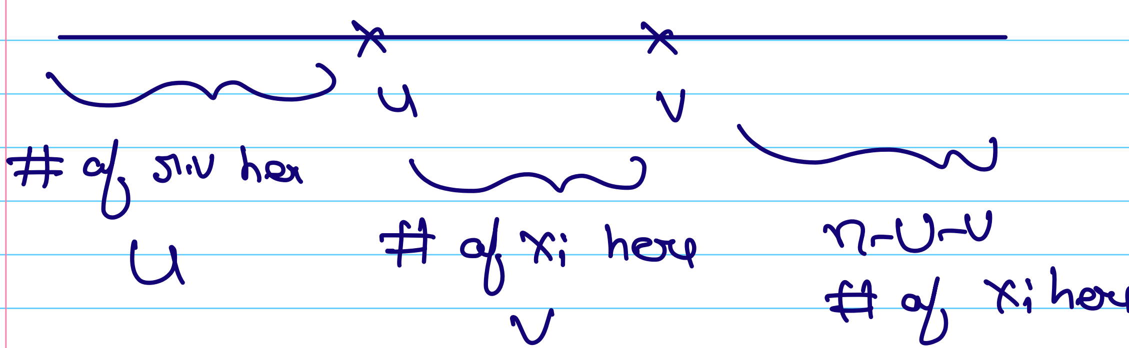
$v = 0, 1, 2, \dots n$

$\Rightarrow V \sim \text{binomial}(n, F_X(v) - F_X(u))$

(3) $n - U - V = \#$ success, $X_i > V$

$n - U - V \sim \text{binomial}\left(n, 1 - F_X(v)\right)$

therefore we can view this as



$$\Rightarrow f_{U,V,n-U-V}(q_1, q_2, n-q_1-q_2)$$

$$= \frac{n!}{k!\,m!\,(n-k-m)!} F_X(u)^{k} \left[F_X(v) - F_X(u)\right]^{m} \left[1 - F_X(v)\right]^{n-k-m}$$

$0 \le k \le n$
$0 \le m \le n$
$k + m \le n$

$$\Rightarrow U, V, n-U-V \sim \text{multinomial}\left(F_X(u), F_X(v) - F_X(u), 1 - F_X(v)\right)$$

(b) Show that Joint pdf of $X_{(i)}$ and $X_{(j)}$ can be expressed as.

$$\text{IF } (u,v) \Rightarrow \mathbb{P}\left(\underbrace{X_{(i)} < u}_{\text{mean's at least } i \text{ success}},\ \underbrace{X_{(j)} < v}_{\text{mean's at least } j \text{ success}}\right)$$

$X_{(i)}, X_{(j)}$

$\Downarrow$ $k \geq i$ $\qquad$ $\Downarrow$ $k+m \geq j$

$$\Rightarrow \mathbb{P}(U \geq i,\ U+V \geq j)$$

$$\{(k,m) \mid 0 \leq k \leq n,\ 0 \leq m \leq n,\ 0 \leq k+m \leq n,\ k \geq i,\ k+m \geq j\}$$

$$= \{(k,m) \mid 0 \leq k \leq n,\ 0 \leq m \leq n,\ 0 \leq k+m \leq n,\ k \geq i,\ m \geq j-k\}$$

$$= \{(k,m) \mid 0 \leq k \leq n,\ 0 \leq m \leq n,\ 0 \leq m \leq n-k,\ j-k \leq m,\ i \leq k\}$$

$$= \left\{ (k, m) \;\middle|\; i \le k \le n ,\; j - k \le m \le n - k \right\}$$

5.27

Let $X_1, X_2, \ldots, X_n$ iid $f_X(x)$ and cdf $F_X(x)$, and let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics.

(a) Find an expression for the conditional pdf of $X_{(i)}$ given $X_{(j)}$ in terms of f_X and F_X

Soln

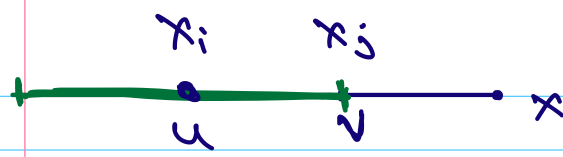
$$\mathbb{P}(X_{(i)} \mid X_{(j)} = x_j)\ ?$$

$$\mathbb{P}(X_{(i)}, X_{(j)}) = \mathbb{P}(X_{(i)} \mid X_{(j)} = x_j)\, \mathbb{P}(X_j = x_j)$$

$$\Rightarrow \quad \mathbb{P}(X_{(i)} \mid X_{(j)} = x_j) = \frac{\mathbb{P}(X_{(i)}, X_{(j)})}{\mathbb{P}(X_j = x_j)}$$

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

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



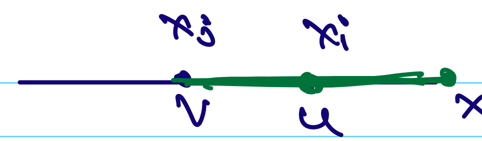
$\Rightarrow \quad i < j$

$\mathbb{P}(X_{(i)} \mid X_{(j)} = v)$

$\Rightarrow$ Given that $X_j = v$

$j-1$ Samples are below v

$\mathbb{P}(X_{(i)} \mid X_{(j)} = v) =$ j trails with $\frac{F_X(u)}{F_X(v)}$ success & $1 - \frac{F_X(u)}{F_X(v)}$ Failure



$\Rightarrow \quad i > j$

$\Rightarrow \quad \mathbb{P}(X_{(i)} \mid X_{(j)} = v)$

$\Rightarrow$ Given that $X_j = v$

$\Rightarrow \quad n-j$ Samples are above v

$\Rightarrow \mathbb{P}(X_{(i)} \mid X_{(j)} = v)$

$= (n-j)$ trails with success $\frac{F_X(u)}{1-F_X(v)}$

& Failure $1 - \frac{F_X(u)}{1-F_X(v)}$

CASE 1: $i < j$

$$\mathbb{P}(X_{(i)} \mid X_{(j)} = v) = \frac{(j-1)!}{(i-1)!\,(j-1-i)!} \, \frac{f_X(u)}{F_X(v)} \cdot \left[\frac{F_X(u)}{F_X(v)}\right]^{i-1} \left(1 - \frac{F_X(u)}{F_X(v)}\right)^{j-1-i}$$

CASE 2 $i > j$

$$f_{X_{(i)}|X_{(j)}=v}(u|v) = \frac{(n-j)!}{(i-1)!\,(n-j-i)!} \frac{f_X(u)}{1-F_X(v)} \left[\frac{F_X(u)}{1-F_X(v)}\right]^{i-1}$$

$$\times \left[1 - \frac{F_X(u)}{1-F_X(v)}\right]^{n-j-i}$$

5.27

(b) Find the pdf of $V|R=r$, where V and R are defined in Example 5.4.7

Sol^n

$$f_{R,V}(r,v) = \frac{n(n-1)\,r^{n-2}}{a^n}, \quad 0 < r < a, \quad \frac{r}{2} < v < a - r/2$$

$$f_{V|R}(v|R=r) = \frac{f_{R,V}(r,v)}{f_R(r)}$$

we derived $\quad f_R(r) = \dfrac{n(n-1)\, r^{n-2}(a-r)}{a^n}$

$0 < r < a$

$$\Rightarrow f_{V|R}(v|R=r) = \frac{\cancel{n}(\cancel{n-1})\, \cancel{r^{n-2}}}{\cancel{a^n}} \cdot \frac{1}{\dfrac{\cancel{n}(\cancel{n-1})\, \cancel{r^{n-2}}(a-r)}{\cancel{a^n}}}$$

$$= \frac{1}{a-r} \qquad \frac{r}{2} < v < a-\frac{r}{2}$$

$$\boxed{\begin{array}{l} \Rightarrow f_{V|R}(v|R=r) = \dfrac{1}{a-r} \;;\; \dfrac{r}{2} < v < a-\dfrac{r}{2} \\[2ex] \qquad V|R=r \sim \text{unif}\left(\dfrac{r}{2}, a-\dfrac{r}{2}\right) \end{array}}$$

(5.28) Let $x_1, x_2, \ldots x_n \overset{iid}{\sim} f_x(x)$ and CDF $F_x(x)$

and let $X_{(i_1)} < X_{(i_2)} < \cdots < X_{(i_L)}$ and

$X_{(j_1)} < X_{(j_2)} < X_{(j_3)} < \cdots < X_{(j_m)}$

be any two disjoint groups of order statistics. in term's of the pdf $f_x(x)$ and the CDF $F_x(x)$, find expression's

(a) The marginal CDF and PDF of

$X_{(i_1)}, \ldots, X_{(i_L)}$