Order Statistics

If we have a sample $X_1, \dots X_n$

$Y_k = k^{th}$ smallest of $X_1, \dots, X_n$

Ex. $Y_1 = \min(X_1, \dots, X_n)$ $\qquad Y_{\frac{n}{2}} = \text{median}(X_1, \dots, X_n)$ $n = \text{even}$

$Y_n = \max(X_1, \dots, X_n)$

We must have that $Y_1 < Y_2 < \cdots < Y_n$

Ex. - Arrival times
- Test scores (median, max, min)
- Weather (daily high & low)
- Stock prices

We will derive pdf's for $Y_1, Y_2, \dots, Y_n$

$Y_n = \text{Max } X_i$. Find the CDF $F_{Y_n}(x) = F_{\max X_i}(x)$

Lets assume $X_1, X_2, \dots, X_n$ of iid random variables w/ pdf $f(x)$ and CDF $F(x)$

$$\begin{aligned}
F_{\max X_i}(x) &= P(\max X_i \le x) && \text{(Defn of CDF)} \\
&= P(X_1 \le x, X_2 \le x, \dots, X_n \le x) && \text{(Defn of max } X_i\text{)} \\
&= P(X_1 \le x) P(X_2 \le x) \cdots P(X_n \le x) && \text{(independence)} \\
&= F_{X_1}(x) F_{X_2}(x) \cdots F_{X_n}(x) && \text{(Defn CDF for } X_i\text{)} \\
&= F(x) \cdots F(x) = [F(x)]^n && \text{(identical dist for } X\text{)}
\end{aligned}$$

$$f_{\max X_i}(x) = F'_{\max X_i}(x) = n[F(x)]^{n-1} f(x)$$

1) X_1, X_2, X_3, X_4 are an iid sample w/ pdf

$$f(x) = \frac{3}{x^4},\ x > 1 \Rightarrow F(x) = 1 - \frac{1}{x^3},\ x > 1$$

a) Find $P(Y_4 \leq 1.75) = P(\max X_i \leq 1.75)$

$$= P(X_1 < 1.75,\ X_2 < 1.75,\ X_3 < 1.75,\ X_4 < 1.75)$$

$$= P(X_1 < 1.75)\,P(X_2 < 1.75)\,P(X_3 < 1.75)\,P(X_4 < 1.75)$$

$$= F_{X_1}(1.75)\,F_{X_2}(1.75)\,F_{X_3}(1.75)\,F_{X_4}(1.75)$$

$$= [F(1.75)]^4$$

$$= \left(1 - \frac{1}{1.75^3}\right)^4 =$$

$$\approx .4378$$

$$=$$

b) $P(Y_4 > 2) = P(\max X_i > 2) = 1 - P(\max X_i \leq 2)$

$$= 1 - [F(2)]^4 = 1 - \left[1 - \frac{1}{2^3}\right]^4 = 1 - \left(\frac{7}{8}\right)^4$$

c) Find $f_{\max X_i}(x) = n[F(x)]^{n-1} f(x)$

$$= 4\left[1 - \frac{1}{x^3}\right]^3 \frac{3}{x^4},\ x > 1$$

d) $E(\max X_i) = \int_1^\infty x\, 12\left[1 - \frac{1}{x^3}\right]^3 \frac{1}{x^4}\, dx = \ldots$

e) $P(Y_4 > 2) = \int_2^\infty 12\left[1 - \frac{1}{x^3}\right]^3 \frac{1}{x^4}\, dx = 1 - \left(\frac{7}{8}\right)^4$

$Y_1 = \min(X_1, \dots, X_n)$. Find $F_{\min X_i}(x)$.

$$F_{\min X_i}(x) = P(\min X_i \le x) \qquad \text{(defn cdf)}$$

$$= 1 - P(\min X_i > x) \qquad \text{((Compliment))}$$

$$= 1 - P(X_1 > x, X_2 > x, \dots, X_n > x) \qquad \text{(defn } \min X_i\text{)}$$

$$= 1 - P(X_1 > x)P(X_2 > x)\cdots P(X_n > x) \qquad \text{(independence)}$$

$$= 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x))\cdots(1 - F_{X_n}(x)) \qquad ([P(X_i > x) = 1 - F_i(x)])$$

$$= 1 - (1 - F(x))\cdots(1 - F(x))$$

$$= 1 - [1 - F(x)]^n$$

$$f_{\min X_i}(x) = F'_{\min X_i}(x) = -n(1 - F(x))^{n-1}(-f(x)) = \underline{n(1 - F(x))^{n-1} f(x)}$$

f) Find $f_{\min X_i}(x)$ for $X_1, X_2, \dots, X_4$ $\quad f(x) = \frac{3}{x^4}, \; x > 1$

$$F(x) = 1 - \frac{1}{x^3}$$

$$= 4\left(\frac{1}{x^3}\right)^3 \frac{3}{x^4} = \frac{12}{x^{13}}, \; x > 1$$

g) $E(Y_1) = E(\min X_i) = \frac{12}{11} = \int_1^\infty x \frac{12}{x^{13}}\,dx = \frac{12}{11}$

h) $P(1.1 < Y_1 < 1.2) = P(\min X_i < 1.2) - P(\min X_i < 1.1)$

OR

$$= \int_{1.1}^{1.2} \frac{12}{x^{13}}\,dx = \left(\frac{5}{6}\right)^{12} \approx .206474$$

$Y_k = k^{th}$ smallest of $X_1, \ldots, X_n$

$F_{Y_k}(x) = P(Y_k \le x) = P(\text{at least } k \text{ observations are } \le x)$

$= P(k \text{ of } X_1, \ldots, X_n \le x) + P(k+1 \text{ of } X_1, \ldots, X_n \le x) + \cdots + P(n \text{ of } X_1, \ldots, X_n \le x)$

$= \sum_{i=k}^{n} P(i \text{ of } X_1, \ldots, X_n \le x)$

We need to place $X_1, X_2, \ldots, X_n$

Consider $i = k$

$\underbrace{\qquad\qquad}_{k \text{ values } \le x}^{P(X_i \le x) = F(x)}$ $\underbrace{\qquad\qquad}_{n-k\ X_i\text{'s } > x}^{P(X_i > x) = 1 - F(x)}$

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

the are $\binom{n}{k}$ ways of distributed $X_1, \ldots, X_n$ into $\le x$ and $> x$.

$\Rightarrow P(k \text{ of } X_1, X_2, \ldots, X_n \le x) = \binom{n}{k}[F(x)]^k[1-F(x)]^{n-k}$

$= \sum_{i=k}^{n} \binom{n}{i} [F(x)]^i [1-F(x)]^{n-i}$

$f_{Y_k}(x) = F'_{Y_k}(x) = \cancel{\frac{n!}{(k-1)!(n-k)!}} \sum_{i=k}^{n} \left[\binom{n}{i} i F(x)^{i-1}[1-F(x)]^{n-i} + (F(x))^i\right]$

$= \binom{n}{k} k [F(x)]^{k-1} [1-F(x)]^{n-k} f(x)$

Find $E(Y_2)$ $\quad f(x) = \frac{3}{x^4},\ x > 1,\quad F(x) = 1 - \frac{1}{x^3}$

$f_{Y_2}(x) = \binom{4}{2} 2 \left[1 - \frac{1}{x^3}\right] \left[\frac{1}{x^3}\right]^2 \frac{3}{x^4} = 36\left[1 - \frac{1}{x^3}\right]\frac{1}{x^{10}},\ x > 1$

$E(Y_2) = E(2^{nd} \text{ smallest } X_i) = \int_1^{\infty} x\, 36\left[1 - \frac{1}{x^3}\right]\frac{1}{x^{10}}\, dx = \frac{27}{22}$

$$f_{Y_k} = F'_{Y_k}(x) = \left(\sum_{i=k}^{n} \binom{n}{i} [F(x)]^i [1-F(x)]^{n-i}\right)'$$

$$= \sum_{i=k}^{n} \left[\binom{n}{i} i (F(x))^{i-1} f(x) [1-F(x)]^{n-i} + \binom{n}{i} [F(x)]^i (n-i) [1-F(x)]^{n-i-1} (-f(x)) \right]$$

1) pass summation to both terms
2) Seperate out the $i=k$ term from the 1st sum.
3) <u>notice</u> the remaining sums are equal

<u>Theorem 4.4.1</u> Joint pdf for $Y_1, Y_2, \ldots, Y_n$ $g(y_1, y_2, \ldots, y_n)$

Let $Y_1 < Y_2 < \cdots < Y_n$ be order statistics of a sample of iid random variables $X_1, \ldots, X_n$ from a cont. dist $f(x)$ on support (a,b). the joint pdf of $Y_1, \ldots, Y_n$ is:

$$g(y_1, y_2, \ldots, y_n) = n!\, f(y_1) f(y_2) \cdots f(y_n) \quad a < y_1 < y_2 < \cdots < y_n < b$$

* need a generalized version of the change of variables formula.

$S_1 : (X_1 = Y_1, X_2 = Y_2, \ldots, X_n = Y_n) \; - \; |J_1| = 1$

$S_2 : (X_1 = Y_2, X_2 = Y_1, \ldots, X_n = Y_n) \; - \; |J_2| = 1$

$\vdots$

$S_{n!} : (X_1 = Y_n, X_2 = Y_{n-1}, \ldots, X_n = Y_1) \; - \; |J_{n!}| = 1$

$$g(y_1, y_2, \ldots y_n) = \sum_{i=1}^{n!} |J_i| f(y_1) f(y_2) \cdots f(y_n) =$$

In the HW $X_1 \sim f_1(x)$; $X_2 \sim f_2(x)$ and independent

$$g(y_1, y_2) = f_1(x_1 = y_1)\, f_2(x_2 = y_2)\, |J_1| + f_1(x_1 = y_2)\, f_2(x_2 = y_1)\, |J_2|$$