AMS570

Order Statistics

1. Definition: Order Statistics of a sample.

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with p.d.f. f(x). Then,

$X_{(1)} = \min(X_1, X_2, \ldots, X_n)$
$X_{(n)} = \max(X_1, X_2, \ldots, X_n)$
and $X_{(1)} \le X_{(2)} \le \cdots \le X_{(k)} \le \cdots \le X_{(n)}$

2. p.d.f.'s for $X_{(1)}$ and $X_{(n)}$

W.L.O.G.(Without Loss of Generality), let's assume X is continuous.

$$P(X_{(1)} > x) = P(X_1 > x, X_2 > x, \ldots, X_n > x) = \prod_{i=1}^{n} P(X_i > x)$$

$$= \prod_{i=1}^{n} [1 - F_{X_i}(x)]$$

$$F_{X_{(1)}}(x) = 1 - \prod_{i=1}^{n}[1 - F_{X_i}(x)]$$

$f_{X_{(1)}}(x) = -\frac{d}{dx}\prod_{i=1}^{n}[1 - F_{X_i}(x)] = -\frac{d}{dx}\prod_{i=1}^{n}[1 - F(x)]^n =$
$n[1 - F(x)]^{n-1} f(x)$

$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x, X_2 \le x, \ldots, X_n \le x) =$
$\prod_{i=1}^{n} F_{X_i}(x) = [F(x)]^n$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1} f(x)$$

Example 1. Let $X_n \overset{i.i.d}{\sim} \exp(\lambda), i=1,\ldots,n$

Please (1). Derive the MLE of λ

(2). Derive the p.d.f. of $X_{(1)}$

(3). Derive the p.d.f. of $X_{(n)}$

Solutions.

(1).

$$L = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \left(\lambda e^{-\lambda x_i}\right) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

$$l = \ln L = n\ln\lambda - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{dl}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{\bar{X}}}$$

Is $\hat{\lambda}$ an unbiased estimator of λ? $E\left(\frac{1}{\bar{\bar{X}}}\right) = ?$

$$M_{X_i}(t) = \frac{\lambda}{\lambda - t}$$

$$M_{\sum_{i=1}^{n} x_i}(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

$$Y = \sum_{i=1}^{n} x_i \sim \text{gamma}(\lambda, n)$$

$$f_{Y-\sum_{i=1}^{n} x_i}(x) = \frac{\lambda}{\Gamma(n)} (\lambda y)^{n-1} e^{-\lambda y}$$

$$\text{Let } Y = \sum_{i=1}^{n} x_i$$

$$E\left(\frac{1}{Y}\right) = \int_0^{\infty} \frac{1}{y} \frac{\lambda}{(n-1)!} (\lambda y)^{n-1} e^{-\lambda y} dy$$

$$= \frac{\lambda}{n-1} \int_0^{\infty} \frac{\lambda}{(n-2)!} (\lambda y)^{n-2} e^{-\lambda y} dy$$

$$= \frac{\lambda}{n-1}$$

$$E\left(\frac{1}{\bar{\bar{X}}}\right) = n\left(\frac{\lambda}{n-1}\right) = \frac{n\lambda}{n-1} \neq \lambda$$

$\hat{\lambda}$ is not unbiased

(2). $X_{(1)} = \min(X_1, X_2, \dots, X_n)$

$$P(X_{(1)} > x) = \prod_{i=1}^{n} P(X_i > x) = \prod_{i=1}^{n} [1 - F(x)] = [1 - F(x)]^n$$

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$$

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1} f(x)$$

$$f(x) = \lambda e^{-\lambda x}, x > 0 \ (\textit{exponential distribution})$$

$$F(x) = \int_0^x f(u)du = \int_0^x \lambda e^{-\lambda u} du = \left[-e^{-\lambda u}\right]\Big|_0^x = 1 - e^{-\lambda x}$$

$$f_{X_{(1)}}(x) = n\lambda e^{-\lambda x}[1 - (1 - e^{-\lambda x})]^{n-1} = n\lambda e^{-\lambda x}(e^{-\lambda x})^{n-1}$$

$$= n\lambda(e^{-\lambda x})^n, x > 0$$

(3). $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$

$$F_{X_{(n)}}(x) = P\left(X_{(n)} \le x\right) = \prod_{i=1}^{n} P(X_i \le x) = [F(x)]^n$$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1} f(x) = n[1 - e^{-\lambda x}]^{n-1} \lambda e^{-\lambda x}, x > 0$$

3. <u>Order statistics are useful in deriving the MLE's.</u>

Example 2. Let X be a random variable with pdf.

$$f(x) = \begin{cases} 1, \text{if } x \in [\theta - \frac{1}{2}, \theta + 1/2] \\ 0, \text{otherwise} \end{cases}$$

Derive the MLE of θ.

Solution.

Uniform Distribution ⇒ important!!

$$L = \prod_{i=1}^{n} f(x_i) = \begin{cases} 1, \text{if all } x_i \in [\theta - \frac{1}{2}, \theta + 1/2] \\ 0, \text{otherwise} \end{cases}$$

MLE : max lnL -> max L

means => $\theta - \frac{1}{2} \le X_1 \le \theta + 1/2$

$\theta - \frac{1}{2} \le X_2 \le \theta + 1/2$

...

$\theta - \frac{1}{2} \le X_n \le \theta + 1/2$

Now we re-express the domain in terms of the order statistics as follows:

$$\theta - \frac{1}{2} \le X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)} \le \theta + 1/2$$

$$\theta \le X_{(1)} + \frac{1}{2}$$

$$\theta \ge X_{(n)} - \frac{1}{2}$$

Therefore,

If $\theta \in \left[X_{(n)} - \frac{1}{2}, X_{(1)} + \frac{1}{2}\right]$, then $L = 1$

Therefore, any $\hat{\theta} \in \left[X_{(n)} - \frac{1}{2}, X_{(1)} + \frac{1}{2}\right]$ is an MLE for θ.

4. The pdf of a general order statistic $X_{(j)}$

Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ denote the order statistics of a random sample, $X_1, X_2, \cdots, X_n$, from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{(j)}$ is

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

Proof: Let Y be a random variable that counts the number of $X_1, X_2, \cdots, X_n$ less than or equal to x. Then we have $Y \sim B\left(n, F_X(x)\right)$. Thus:

$$F_{X_{(j)}}(x) = P(Y \geq j) = \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [1-F_X(x)]^{n-k}$$

5. The Joint Distribution of Two Order Statistics $X_{(i)}$ *and* $X_{(j)}$

Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ denote the order statistics of a random sample, $X_1, X_2, \cdots, X_n$, from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the joint pdf of $X_{(i)}$ and $X_{(j)}$, $1 \leq i < j \leq n$, is

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

6. Special functions of order statistics

(1) Median (of the sample):

$$\begin{cases} X_{(k+1)}; & if\ n = 2k+1 \\ \dfrac{X_{(k)} + X_{(k+1)}}{2}; & if\ n = 2k \end{cases}$$

(2) Range (of the sample): $X_{(n)} - X_{(1)}$

7. More examples of order statistics

Example 3. Let X_1,X_2, X_3 be a random sample from a distribution of the continuous type having pdf f(x)=2x, 0<x<1, zero elsewhere.
(a) compute the probability that the smallest of X_1,X_2, X_3 exceeds the median of the distribution.
(b) If $Y_1 \le Y_2 \le Y_3$ are the order statistics, find the correlation between Y_2 and Y_3.

Answer:
(a)

$$F(x) = P(X_i < x) = x^2;$$

$$\int_0^t 2x dx = \frac{1}{2}; t = \frac{\sqrt{2}}{2}$$

$$P(\min(X_1, X_2, X_3) > t) = P(X_1 > t, X_2 > t, X_3 > t) = P(X_1 > t)P(X_2 > t)P(X_3 > t)$$

$$= [1 - F(t)]^3 = (1 - t^2)^3 = \frac{1}{8}$$

(b)
Please refer to the textbook/notes for the order statistics pdf and joint pdf formula. We have

$$f_{Y_3}(x) = 6 * x^5; 0 < x < 1$$

$$f_{Y_2}(x) = 12 * (x^3 - x^5); 0 < x < 1$$

$$E(Y_3) = 6/7,$$

$$E(Y_2) = 24/35;$$

$$f_{Y_2,Y_3}(y_2, y_3) = 24 * (y_2)^3 * y_3; 0 < \ y_2 \le y_3 < 1$$

$$E(Y_2Y_3) = \int_0^1 \left[\int_0^{y_3} y_2 * y_3 * 24 * (y_2)^3 * y_3 \, dy_2 \right] dy_3 = \frac{3}{5};$$

$$var(Y_3) = \frac{6}{8} - \left(\frac{6}{7}\right)^2 = \frac{6}{392};$$

$$var(Y_2) = \frac{1}{2} - \left(\frac{24}{35}\right)^2;$$

$$corr(Y_2, Y_3) = \frac{E(Y_2Y_3) - E(Y_2)E(Y_3)}{\sqrt{var(Y_2)var(Y_3)}} = 0.57$$

Example 4. Let $Y_1 \le Y_2 \le Y_3$ denote the order statistics of a random sample of size 3 from a distribution with pdf f(x) = 1, 0 < x < 1, zero elsewhere. Let Z = $(Y_1+Y_3)/2$ be the midrange of the sample. Find the pdf of Z.

From the pdf, we can get the cdf : F(x) = x, 0<x<1
Let

$$W = Y_1$$
$$Z = (Y_1 + Y_3)/2$$

The inverse transformation is:

$$Y_1 = W$$
$$Y_3 = 2Z - W$$

The joint pdf of Y_1 and Y_3 is:

$$f(y_1, y_3) = \begin{cases} 6(y_3 - y_1), & 0 < y_1 \le y_3 < 1 \\ 0, & o.w. \end{cases}$$

We then find the Jacobian: J= -2
Now we can obtain the joint pdf of Z, W:

$$f(z, w) = \begin{cases} |2|6(2z - w - w) = 24(z - w), & 0 < w \le 2z - w < 1 \\ 0, & o.w. \end{cases}$$

From $0 < w \le 2z - w < 1$, we have:

$$w > 0;$$
$$w > 2z - 1;$$
$$w \le z$$

Together they give us the domain of w as:

$$\max(0, 2z - 1) < w \le z$$

Therefore the pdf of Z (non-zero portion) is:

$$f(z) = \begin{cases} \int_{2z-1}^{z} 24(z - w)\, dw = 12(z-1)^2, & 2z - 1 > 0 \\ \int_{0}^{z} 24(z - w)\, dw = 12(z^2), & 2z - 1 \le 0 \end{cases}$$

We also remind ourselves that:

$$0 < z < 1$$

Therefore the entire pdf of the midrange Z is:

$$f(z) = \begin{cases} \int_{2z-1}^{z} 24(z-w)\,dw = 12(z-1)^2, & 1/2 < z < 1 \\ \int_{0}^{z} 24(z-w)\,dw = 12(z^2), & 0 < z \le 1/2 \\ 0, & otherwise \end{cases}$$

Example 5. Let $Y_1 \le Y_2 \le Y_3 \le Y_4$ be the order statistics of a random sample of size n = 4 from a distribution with pdf f(x) = 2x, 0 < x < 1, zero elsewhere.

(a) Find the joint pdf of Y_3 and Y_4.

(b) Find the conditional pdf of Y_3, given $Y_4 = y_4$.

(c) Evaluate $E[Y_3|y_4]$.

Solution:

(a)

$$f_{Y_1,Y_2,Y_3,Y_4,}(y_1, y_2, y_3, y_4) = 4!\,(2y_1)(2y_2)(2y_3)(2y_4)$$

for $0 < y_1 \le y_2 \le y_3 \le y_4 < 1$. We have:

$$f_{Y_3,Y_4}(y_3, y_4) = \int_{y_2=0}^{y_3} \int_{y_1=0}^{y_2} f_{Y_1,Y_2,Y_3,Y_4,}(y_1, y_2, y_3, y_4)\,dy_1 dy_2 = 48{y_3}^5 y_4$$

for $0 < y_3 \le y_4 < 1$

(Note: You can also obtain the joint pdf of these two order statistics by using the general formula directly.)

(b)

$$f_{Y_3|Y_4}(y_3|y_4) = \frac{f_{Y_3,Y_4}(y_3, y_4)}{f_{Y_4}(y_4)} = \frac{48{y_3}^5 y_4}{8{y_4}^7} = \frac{6{y_3}^5}{{y_4}^6}$$

for $0 < y_3 \le y_4$.

(c)

$$E[Y_3|Y_4 = y_4] = \int_0^{y_4} \frac{6{y_3}^6}{{y_4}^6}\,dy_3 = \frac{6y_4}{7}$$

Example 6. Suppose X1, . . . , Xn are iid with pdf f(x; θ) = $2x/\theta^2$, $0 < x \le \theta$, zero elsewhere. Note this is a nonregular case. Find:

(a) The mle $\hat{\theta}$ for θ.

(b) The constant c so that $E(c^*\hat{\theta}) = \theta$.

(c) The mle for the median of the distribution.

Solution:

$$\text{(a)}\, L(\theta; X) = \prod_{i=1}^{n} \frac{2x_i}{\theta^2} = \frac{2^n \prod_{i=1}^{n} x_i}{\theta^{2n}} \le \frac{2^n \prod_{i=1}^{n} x_i}{[\max(x_i)]^{2n}}$$

$$\text{So } \hat{\theta} = X_{(n)} \qquad X_{(n)} = \max(X_1, \cdots X_n)$$

Dear students: note that this is no typo in the above – the truth is that $\theta \ge X_{(n)}$ – and so the smallest possible value for θ is $X_{(n)}$

$$\text{(b)}\, F_X(x) = \int_0^x \frac{2t}{\theta^2} dt = \frac{x^2}{\theta^2} \qquad 0 < x \le \theta$$

$$\text{So } F_{X_{(n)}}(x) = \left(\frac{x^2}{\theta^2}\right)^n = \frac{x^{2n}}{\theta^{2n}} \qquad 0 < x \le \theta$$

$$f_{X_{(n)}}(x) = \frac{2nx^{2n-1}}{\theta^{2n}} \qquad 0 < x \le \theta$$

$$E(c\hat{\theta}) = cE(\hat{\theta}) = c\int_0^\theta x \frac{2nx^{2n-1}}{\theta^{2n}} dx = \frac{2nc}{2n+1}\theta = \theta$$

$$\text{So } c = \frac{2n+1}{2n}$$

$$\text{(c) Let } F(x) = \frac{x^2}{\theta^2} = \frac{1}{2}, \text{ then } x = \frac{\theta}{\sqrt{2}}$$

So the median of the distribution is $\frac{\theta}{\sqrt{2}}$

The mle for the median of the distribution is

$$\frac{\widehat{\theta}}{\sqrt{2}} = \frac{X_{(n)}}{\sqrt{2}} = \frac{\sqrt{2}}{2} X_{(n)}$$

Mean Squared Error (M.S.E.)

How to evaluate an estimator?

For unbiased estimators, all we need to do is to compare their variances, the smaller the variance, the better is estimator. Now, what if the estimators are not all unbiased? How do we compare them?

Definition: Mean Squared Error (MSE)
Let $T=t(X_1, X_2, \ldots, X_n)$ be an estimator of $\tau(\theta)$, then the M.S.E. of the estimator T is defined as :

$MSE_t(\tau(\theta)) = E\left[\left(T-\tau(\theta)\right)^2\right]$: average squared distance from T to $\tau(\theta)$

$$=E\left[\left(T-E(T)+E(T)-\tau(\theta)\right)^2\right]$$

$$=E\left[\left(T-E(T)\right)^2\right]+E\left[\left(E(T)-\tau(\theta)\right)^2\right]+2E\left[\left(T-E(T)\right)\left(E(T)-\tau(\theta)\right)\right]$$

$$=E\left[\left(T-E(T)\right)^2\right]+E\left[\left(E(T)-\tau(\theta)\right)^2\right]+0$$

$$=Var(T)+\left(E(T)-\tau(\theta)\right)^2$$

Here $|E(T)-\tau(\theta)|$ is "the bias of T"
If unbiased, $\left(E(T)-\tau(\theta)\right)^2=0$.
The estimator has smaller mean-squared error is better.

Example 1. Let $X_1, X_2, \ldots, X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$

M.L.E. for μ is $\hat{\mu}=\bar{X}$; M.L.E. for σ^2 is $\hat{\sigma}^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$

1. M.S.E. of $\hat{\sigma}^2$?
2. M.S.E. of S^2 as an estimator of σ^2

Solution.

1.

$$MSE_{\hat{\sigma}^2}(\sigma^2)=E\left[\left(\hat{\sigma}^2-\sigma^2\right)^2\right]=Var(\sigma^2)+(E(\hat{\sigma}^2)-\sigma^2)^2$$

To get $Var(\hat{\sigma}^2)$, there are 2 approaches.

a. By the first definition of the Chi-square distribution.

Note $X_i \sim N(\mu,\sigma^2)$; $W=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim {\chi_{n-1}}^2$, $Gamma(\lambda=\frac{1}{2}, S=\frac{n-1}{2})$

$$E(W)=\frac{S}{\lambda}=n-1;\ Var(W)=\frac{S}{\lambda^2}=2(n-1)$$

$$Var(\widehat{\sigma}^2)=Var(\frac{W}{n}\ \sigma^2)=\frac{\sigma^4}{n^2}Var(W)=\frac{\sigma^4}{n^2}2(n-1)$$

b. By the second definition of the Chi-squre distribution.

For $Z\sim N(0,1)$, $W=\sum_{i=1}^{n} {Z_i}^2$

$$Var(Z^2)=E\left[\left(Z^2-E(Z^2)\right)^2\right]$$

$$=E\left[\left(Z^2-\left(var(Z^2)+E(Z)\right)\right)^2\right]$$

$$=E[(Z^2-1)^2]$$

Since $Var(Z)=E(Z^2)-E(Z)=$

1 from $Z\sim N(0,1)$, $E(Z^2)=1=E[Z^4-2E(Z^2)+1]$

$$=E(Z^4)-1$$

Calculate the 4th moment of $Z\sim N(0,1)$ using the mgf of Z;

$$M_Z(t)=e^{t^2/2}$$

$$M'_Z(t)=te^{t^2/2}$$

$$M''_Z(t)=te^{t^2/2}+t^2e^{t^2/2}$$

$$M^{(3)}{}_Z(t)=3te^{t^2/2}+t^2e^{t^2/2}$$

$$M^{(4)}{}_Z(t)=3e^{t^2/2}+6t^2e^{t^2/2}+t^4e^{t^2/2}$$

Set $t=0$, $M^{(4)}{}_Z(0)=3=E(Z^4)$

$$Var(Z^2)=3-1=2$$

$$Var(W)=\sum_{i=1}^{n-1}Var({Z_i}^2)=2(n-1)$$

$$\widehat{\sigma}^2=\frac{\sigma^2}{n}W,\ Var(\widehat{\sigma}^2)=\frac{\sigma^4}{n^2}2(n-1)$$

$$MSE_{\widehat{\sigma}^2}(\sigma^2)=Var(\widehat{\sigma}^2)+[E(\widehat{\sigma}^2)-\sigma^2]^2$$

$$=\frac{2(n-1)}{n^2}\sigma^4+[E\left(\frac{n-1}{n}S^2\right)-\sigma^2]^2$$

$$=\frac{2(n-1)}{n^2}\sigma^4$$

$$+\left[\frac{n-1}{n}\sigma^2-\sigma^2\right]^2\ \left(\text{we know } E(S^2)\right.$$

$$\left.=\sigma^2\right)=\frac{2n-1}{n^2}\sigma^4$$

The M.S.E. of $\hat{\sigma}^2$ is $\frac{2n-1}{n^2}\sigma^4$

We know S^2 is an unbiased estimator of σ^2

$$E[(S^2-\sigma^2)^2] = Var(S^2) + 0 = Var\left(\frac{\sigma^2 W}{n-1}\right)$$

$$= \left(\frac{\sigma^2}{n-1}\right)^2 var(W) = \frac{2\sigma^4}{n-1}$$

Exercise:

Compare the MSE of $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$ and $\hat{\sigma}^2 = S^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$.
Which one is a better estimator (in terms of the MSE)?

Solution.

$$MSE_{\hat{\sigma}^2}(\sigma^2) = E\left[\left(\hat{\sigma}^2-\sigma^2\right)^2\right] = Var\left(\hat{\sigma}^2\right) + \left(E\left(\hat{\sigma}^2\right)-\sigma^2\right)^2$$

Then, we have M.S.E. of $\hat{\sigma}^2$ is $\frac{2n-1}{n^2}\sigma^4$.

$$MSE_{S^2}(\sigma^2) = E[(S^2-\sigma^2)^2] = Var(S^2) + (E(S^2)-\sigma^2)^2$$

$$MSE_{S^2}(\sigma^2) = Var(S^2) + 0 = Var\left(\frac{\sigma^2 W}{n-1}\right) = \left(\frac{\sigma^2}{n-1}\right)^2 var(W) = \frac{2\sigma^4}{n-1}$$

$$\because \frac{2n-1}{n^2}\sigma^4 - \frac{2\sigma^4}{n-1} = \frac{1-3n}{n^2(n-1)}\sigma^4 < 0$$

$$\therefore MSE_{\hat{\sigma}^2}(\sigma^2) = \frac{2n-1}{n^2}\sigma^4 < \frac{2\sigma^4}{n-1} = MSE_{S^2}(\sigma^2)$$

$$\therefore \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n} \text{ is the better estimator}$$

Homework : Read the following chapter/sections from our textbook: *Chapter1, 2, 3 (3.1, 3.2, 3.3), 4 (4.1, 4.2, 4.3, 4.5, 4.6),5 (5.1, 5.2, 5.3, 5.4), 7(7.1, 7.2.1, 7.2.2, 7.3.1). These are materials covered so far in our class.*