26-10-2020

Lecture 7

Order distributions Order statistics are sample values placed in ascending order. The study of order statistics deals with the applications of these ordered values and their functions

Let's say you had three weights,

$X_1 = 22$ kg, $X_2 = 44$ kg, and $X_3 = 12$ kg.

To get order statistic (Y_n), put the items in numerical increasing order.

$y_1 = 12$ kg

$y_2 = 22$ kg

$y_3 = 44$ kg. The kth smallest x value is normally called the kth order statistic.

More formally,

If $x_1, \dots, x_n$ are random iid observations taken from a pop^n with n continuous observations, the random variables

$y_1 < y_2 < \dots < y_n$ denote the sample order statistics.

Density of the kth Order Statistic:-

For $X_1, \dots, X_n$ iid continuous r.v with pdf f and cdf F, the density of kth order Statistic is

$$P(X_{(k)} \in [x, x+\epsilon]) = P\begin{pmatrix} \text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and} \\ \text{exactly } k-1 \text{ of the others} < x \end{pmatrix}$$

$$= \sum_{i=1}^{n} P\begin{pmatrix} X_i \in (x, x+\epsilon) \text{ and exactly } k-1 \\ \text{of the others} < x \end{pmatrix}$$

$$= n\, P\begin{pmatrix} X_1 \in (x, x+\epsilon] \text{ and exactly } k-1 \\ \text{of the others} < x \end{pmatrix}$$

$$= nP(X_1 \in [x, x+\epsilon])\, P(\text{exactly } k-1 \dots < x)$$

$$= nP(X_1 \in [x, x+\epsilon]) \left[\binom{n-1}{k-1} P(X<x)^{k-1} P(X>x)^{n-k}\right]$$

$$f_{(k)}(x) = nf(x) \binom{n-1}{k-1} F(x)^{k-1} (1-F(x))^{n-k}.$$

hence proved.

$= nf(x)(F(x) \cdots F(x))$

$= nf(x)\,(F(x))^{n-1}$

$$f_{(n)}(x) = nf(x)(F(x))^{n-1}$$

<u>Density of minimum:-</u> For $X_1, \dots, X_n$ iid continuous r.v with pdf f and cdf F the density of minimum is

$P(X_{(1)} \in (x, x+\varepsilon)) = P($one of $X's \in [x, x+\varepsilon]$ and all others $> x)$

$= \sum_{i=1}^{n} P(X_1 \in [x, x+\varepsilon]$, and all others $> x)$

$= nP(X_1 \in [x, x+\varepsilon]$ and all others $> x)$

$= nP(X_{(1)} \in [x, x+\varepsilon])\,P($all others $> x)$

$= n\,f(x)\,(1 - F(x))^{n-1}$

$$f_{(1)}(x) = nf(x)\,(1 - F(x))^{n-1}$$

hence proved.

$$Y_1 = X_1 = X_{(1)} = \min(X_1, \dots, X_n)$$

$$Y_n = X_n = X_{(n)} = \max(X_1, \dots, X_n).$$

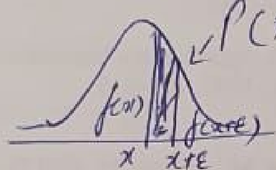
For a continuous random variable, we can see that

$$f(x)\varepsilon \approx P(x \le X \le x+\varepsilon) = P(X \in [x, x+\varepsilon])$$

$$\lim_{\varepsilon \to 0} f(x)\varepsilon = \lim_{\varepsilon \to 0} P(X \in [x, x+\varepsilon])$$

$$f(x) = \lim_{\varepsilon \to 0} P(X \in [x, x+\varepsilon])/\varepsilon$$

$$P(x \le X \cancel{+\varepsilon} \le x+\varepsilon) = P(X \in [x, x+\varepsilon])$$



<u>Density of maximum</u>: For $X_1, \dots, X_n$ iid continuous r.v with pdf f and cdf F the density of maximum is

$$P(X_{(n)} \in [x, x+\varepsilon]) = P(\text{one of the } X\text{'s} \in [x, x+\varepsilon] \text{ and all others} < x)$$

$$= \sum_{i=1}^{n} P(X_i \in [x, x+\varepsilon] \text{ and all others} < x)$$

$$= nP(X_1 \in [x, x+\varepsilon] \text{ and all others} < x)$$

$$= nP(X_1 \in (x, x+\varepsilon))\, P(\text{all other} < x)$$

$$= nP(X_1 \in (x, x+\varepsilon))\, P(X_2 < x) \cdots P(X_n < x)$$

$$\alpha_{(1)}^{n} = \frac{n!}{n!}\int_0^{\infty} e^{-nx}\,dx$$

$$= \int_0^{\infty} e^{-nx}\,dx = \frac{e^{-nx}}{-n}\Big|_0^{\infty}$$

$$\alpha_{(1)}^{n} = -\frac{1}{n}\left[\frac{1}{e^{n(\infty)}} - \frac{1}{e^{0}}\right] = -\frac{1}{n}[-1]$$

$$\boxed{\alpha_{(1)}^{n} = \frac{1}{n}}$$

Now

$$E[x_{(1)}^{n}]^2 = w_{11}^{(n)} = \frac{n!}{(n-1)!}\int_0^{\infty} x_{(1)}^2\,[f(x)_{(1)}]\,[1-F(x)_{(1)}]^{n-1}\,dx.$$

$$= n\int_0^{\infty} x_{(1)}^2\, e^{-x}\,(e)^{-(n-1)x}\,dx.$$

$$= n\int_0^{\infty} x_{(1)}^2\, e^{-nx}\,dx.$$

$$= n\left[x_{(1)}^2 \,\Big|\, \frac{e^{-nx}}{-n}\Big|_0^{\infty} - \int_0^{\infty} \frac{e^{-nx}}{-n}(2x)\,dx\right].$$

$$= n[0] + \frac{2}{n}\cdot n\int_0^{\infty} x\, e^{-nx}\,dx.$$

$$= \frac{2}{n}\cdot\frac{1}{n} = \frac{2}{n^2}.$$

$$Var[x_{(1)}^{n}] = w_{11}^{n} - [\alpha_1^{(n)}]^2 = \frac{2}{n^2} - \frac{1}{n^2} = \frac{1}{n^2}\ \text{A}$$

Q# Show that in a sample of 'n' observations from a dist^n $f(x) = e^{-x}, \quad 0 \le x \le \infty$

$= 0 \quad o\omega$

The variance of smallest observation is $\frac{1}{n^2}$

Sol First find the dist^n f^n

$$F(x) = \int_0^x e^{-x}\,dx = \left|\frac{e^{-x}}{-1}\right|_0^x = 1 - e^{-x}.$$

we know that $\cancel{= \int x f(x)(1-F(x))^{n-1}}$

$$E[X_{(1)}^{(n)}] = \alpha_1^{(n)} = \frac{n!}{(n-1)!}\int_0^\infty x_{(1)}\left[1-1+e^{-x}\right]^{n-1} e^{-x}\,dx.$$

$$\alpha_1^{(n)} = \frac{n\cancel{(n-1)!}}{\cancel{(n-1)!}}\int_0^\infty x_{(1)}\, e^{-x}\, e^{-x(n-1)}\,dx.$$

$$= n\int_0^\infty x_{(1)}\, e^{-xn}\,dx$$

integration by parts

$$= n\left[x_{(1)}\frac{e^{-nx}}{-n}\Big|_0^\infty - \int_0^\infty \frac{e^{-nx}}{-n}\,dx\right]$$

$$= n[0] + \frac{n}{n}\int_0^\infty e^{-nx}\,dx.$$

Lecture 8 :-

28-10-2020

Dist^n of max = $n f(x) (F(x))^{n-1}$

Dist^n of min = $n f(x) (1 - F(x))^{n-1}$

Density of kth order statistic = $n f(x) \binom{n-1}{k-1} F(x)^{k-1} (1-F(x))^{n-k}$

Joint dist^n of order statistics = $n! f(y_1) \cdots f(y_n)$

Example

Let $X_1, \ldots, X_n \overset{iid}{\sim} Exp(\lambda)$, we previously derived a more general result where the X's were not identically distributed and showed that $\min(X_1, \ldots, X_n) \sim Exp(\lambda_1, \ldots, \lambda_n) = Exp(n\lambda)$ in this more restricted case.

Sol density of min is $= n f(x) (1 - F(x))^{n-1}$

$$\left\{ \begin{array}{l} f(x) = \lambda e^{-\lambda x} \\ F(x) = 1 - e^{-\lambda x} \end{array} \right\}$$

$$f_{(1)}(x) = n\lambda e^{-\lambda x} (\cancel{1} - \cancel{1} + e^{-\lambda x})^{n-1}$$

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

$$f_{(1)}(x) = n\lambda (e^{-\lambda x})^{n \cancel{-1} \cancel{+1}}$$

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

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

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

$$= n \binom{n-1}{k-1} f(x) \left(F(x)\right)^{k-1} \left(1-F(x)\right)^{n-k}.$$

$$k = 1$$

$$= n \binom{n-1}{1-1} f(x) \left(F(x)\right)^{1-1} \left(1-F(x)\right)^{n-1}$$

$$= n \frac{(n-1)!}{0!\,(n-1)!} f(x) \left(1-F(x_1)\right)^{n-1}$$

$$= \frac{n \cancel{(n-1)!}}{\cancel{(n-1)!}} f(x) \left(1-F(x)\right)^{n-1}$$

which is the density of

$Exp(n\lambda)$

Maximum of exponentials : let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Exp(\lambda)$ then the density of $X_{(n)}$ is given by

$$f_{(n)}(x) = n f(x) (F(x))^{n-1}$$
$$= n(\lambda e^{-\lambda x})(1-e^{-\lambda x})^{n-1}$$

(i) Pdf of k^{th} order statistics:-

$$f_{(k)}(x) = n f(x) \binom{n-1}{k-1} \cancel{f(x)_k} (F(x))^{k-1} (1-F(x))^{n-k}$$

(ii) Mean or 1^{st} moment about mean

$$E[X_{(k)}^{(n)}] = \alpha_k^{(n)} = \frac{n!}{(k-1)!\,!!\,(n-k)!} \int_{-\infty}^{\infty} x_{(k)} [F(x)]^{k-1} f(x) [1-F(x)]^{n-k}$$

(iii) 2^{nd} moment about mean

$$E[X_{(k)}^{(n)}]^2 = W_{kk}^{(n)} = \frac{n!}{(k-1)!\,!!\,(n-k)!} \int_{-\infty}^{\infty} x_{(k)}^2 [F(x)]^{k-1} f(x) [1-F(x)]^{n-k}$$

(iv) Variance of k^{th} order statistics

$$V[X_{(k)}^{(n)}] = V_{kk}^{(n)} = W_{kk}^{(n)} - [\alpha_k^{(n)}]^2$$

Q Let there are 5^{3} observations from a dist^n $f(x) = e^{-x}$, $0 \le x \le \infty$

Find the dist^n of median.

<u>Sol</u> To find CDF,

$$F(x) = \int_0^x e^{-x}\,dx = \left|\frac{e^{-x}}{-1}\right|_0^x = -1(e^{-x} - e^{0})$$

$$= 1 - e^{-x}.$$ $\boxed{e^0 = 1}$

we know that k^{th} order statistics is

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

$$\text{median} = \frac{5+1}{2} = 3$$

$$f_{(3)}(x) = (5)\binom{5-1}{3-1} e^{-x}\,(1-e^{-x})^{5-1}\,(1-(1-e^{-x}))^{5-3}$$

$$= (5)\binom{4}{2} e^{-x}\,(1-e^{-x})^{4}\,(\cancel{1}-\cancel{1}+e^{-x})^{2}$$

$$= (5)\frac{4!}{2!\,(4-2)!}\, e^{-x}\,(1-e^{-x})^{4}\,(e^{-x})^{2}$$

$$= (5)\frac{1\cdot\cancel{2}\cdot 3\cdot \cancel{4}^{2}}{\cancel{2}!\,\cancel{2}!}\, e^{-x}\cdot e^{-2x}\,(1-e^{-x})^{4}$$

$$= 30\, e^{-3x}\,(1-e^{-x})^{4}.$$

i.e T_1 is a sample mean whereas T_2 is the weighted mean

which estimator should be preffered?

Sol: To check this we will find variance of both estimators

Let us first see if the two estimators are U/B

$$E(T_1) = E\left[\frac{X_1+X_2+X_3+X_4}{4}\right] = \frac{1}{4}E\left[X_1+X_2+X_3+X_4\right]$$

$$= \frac{1}{4}\left[E(X_1)+E(X_2)+E(X_3)+E(X_4)\right] = \frac{4\mu}{4} = \mu.$$

So T_1 is U/B estimator of μ.

$$E(T_2) = E\left[\frac{X_1+2X_2+3X_3+X_4}{7}\right]$$

$$= \frac{1}{7}\left[E(X_1)+2E(X_2)+3E(X_3)+E(X_4)\right]$$

$$= \frac{1}{7}\left[\mu+2\mu+3\mu+\mu\right] = \frac{7\mu}{7} = \mu$$

T_2 is also U/B estimator of μ.

Now we find their variances,

Question Consider a pop^n with density fn

$$f(x) = 2x \qquad 0 < x < 1/2$$

Then

$$F(x) = \int_0^x f(t)\,dt = x^2 \qquad ; 0 < x < 1/2$$

find density of $\min(X_1, \ldots, X_6)$ and $\max(X_1, \ldots, X_6)$

Sol

$$f_{(1)}(x) = n\, f(x)\,(1 - F(x))^{n-1}$$

$$= (6)\, 2y\, (1 - y^2)^{6-1}$$

$$\boxed{f_{(1)}(x) = 12y\,(1-y^2)^5}$$

$$f_{(6)}(x) = n\, f(x)\,(F(x))^{n-1}$$

$$= (6)(2y)(y^2)^{6-1}$$

$$= 12y\,(y^2)^5 = 12y\,(y^{10})$$

$$\boxed{f_{(6)}(x) = 12y^{11}}$$

2

the sample median is less efficient or sample mean is more efficient than the sample median as an estimator of μ.

<u>Minimum Variance U/B estimator</u>:- If an U/B estimator $\hat{\theta}$ has smaller ~~value~~ variance than any other U/B estimator, it is called the minimum variance U/B estimator. An U/B estimator having the minimum variance is called the best or most efficient estimator for θ.

Point to be noted that variance of $\hat{\theta}$ cannot becomes smaller than certain lower bound and a variance equal to this lower bound is called the minimum variance.

This result is given by Cramer-Rao Inequality

in other words, any estimator which attain the lower bound is called minimum variance.

29-10-2020

Lecture 9

Efficiency:- An U/B estimator is defined to be efficient if the variance of its sampling distⁿ is smaller than that of the sampling distⁿ of any other U/B estimators of the same parameter. Suppose there are two U/B estimators T_1 and T_2 of parameter θ, then T_1 will be efficient estimator than T_2 if $Var(T_1) < Var(T_2)$.

The relative efficiency of T_1 compared to T_2 is given by the ratio $\left[E_f = \frac{Var(T_2)}{Var(T_1)} > 1 \right]$

R.E or simple efficiency provides the way of comparing different U/B estimators.

Both sample mean and sample median of a normal distⁿ are U/B and consistent estimators of μ but the variance of the sampling distⁿ of sample means $\left(\frac{\sigma^2}{n}\right)$ is smaller than to the sampling distⁿ of sample medians $\left(\frac{\pi\sigma^2}{2n}\right)$

i.e $\frac{Var(med)}{Var(mean)} = \frac{\pi\sigma^2}{2n} \times \frac{n}{\sigma^2} = \frac{\pi}{2} > 1$. Hence

$$Var(T_1) = Var\left[\frac{X_1 + X_2 + X_3 + X_4}{4}\right] = \frac{1}{16}\left[\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2\right]$$

$$Var(T_1) = \frac{4\sigma^2}{16} = \frac{\sigma^2}{4}$$

$$Var(T_2) = Var\left[\frac{X_1 + 2X_2 + 3X_3 + X_4}{7}\right]$$

$$Var(T_2) = \frac{1}{49}\left[Var(X_1) + 4\,Var(X_2) + 9\,Var(X_3) + Var(X_4)\right]$$

$$Var(T_2) = \frac{1}{49}\left[\sigma^2 + 4\sigma^2 + 9\sigma^2 + \sigma^2\right] = \frac{15}{49}\sigma^2$$

Now $\dfrac{Var(T_1)}{Var(T_2)} = \dfrac{15}{49}\sigma^2 \times \dfrac{4}{\sigma^2} = \dfrac{60}{49} > 1$

hence it is proved that $Var(T_1) < Var(T_2)$

Hence T_1 is better estimator of μ than T_2.

<u>Example</u> :- Two different estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are being considered. To test their performance, 75 trials have been simulated, each with the true value set at $\theta = 2$. The following totals were obtain

6

$\sum_{j=1}^{75} \hat{\theta}_{1j} = 165; \quad \sum_{j=1}^{75} \hat{\theta}_{1j}^2 = 375$

$\sum_{j=1}^{75} \hat{\theta}_{2j} = 147; \quad \sum_{j=1}^{75} \hat{\theta}_{2j}^2 = 312$

Estimate the MSE for each estimator and determine the relative efficiency.

Sol

$$Var(\hat{\theta}_1) = \frac{\sum \hat{\theta}_{1j}^2}{75} - \left(\frac{\sum \theta_{1j}}{75}\right)^2$$

$$= \frac{375}{75} - \left(\frac{165}{75}\right)^2 = 0.16$$

$$Bias(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta \approx \frac{\sum \hat{\theta}_{1j}}{75} - \theta$$

$$= \frac{165}{75} - 2 = 0.2$$

$$MSE(\hat{\theta}_1) = Var(\hat{\theta}_1) + \{Bias(\hat{\theta}_1)\}^2$$

$$= 0.16 + (0.2)^2$$

$$= 0.2$$

on similar patter, find $MSE(\hat{\theta}_2)$ which is equal to 0.32

$$\text{relative efficiency} = \frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{0.2}{0.32} = 0.625$$

here $\hat{\theta}_2$ is more efficient than $\hat{\theta}_1$.

An estimator $\hat{\theta}$ that is linear, is U/B and has minimum variance among all linear U/B estimators of θ is called a Best Linear U/B Estimator or (BLUE).

As we know that MSE is equal to the variance and square of bias i.e.

$$MSE(\hat{\theta}) = E(\hat{\theta}-\theta)^2$$

$$= Var(\hat{\theta}) + (Bias(\theta))^2$$

now if $E(\hat{\theta}) = \theta$, i.e. $\hat{\theta}$ become U/B estimator of θ, then variance and MSE is concide.

Example Let X_1, X_2, X_3 and X_4 be a random sample of size $n=4$ from a $N(\mu, \sigma^2)$. A statistician want to estimate mean by using either ~~using~~ of the following two estimators of the mean μ

$$T_1 = \frac{X_1+X_2+X_3+X_4}{4} \quad \text{and} \quad T_2 = \frac{X_1+2X_2+3X_3+X_4}{\cancel{8}\ 7}$$