Assignment 4

Suraj - CS20BTECH11050

Download all python codes from

https://github.com/Suraj11050/Assignments–AI1103/tree/main/Assignment%204/Python%20codes

Download Latex-tikz codes from

https://github.com/Suraj11050/Assignments–AI1103/blob/main/Assignment%204/Assignment%204.tex

1 GATE 2021 (ST), Q.17 (STATISTICS SECTION)

If the marginal probability density function of the k^{th} order statistic of a random sample of size 8 from a uniform distribution on $[0, 2]$ is

$$f(x) = \begin{cases} \frac{7}{32}\, x^6\, (2 - x), & 0 < x < 2, \\ 0, & \text{otherwise,} \end{cases} \tag{1.0.1}$$

then k equals ______________

2 SOLUTION

Definition 2.1. For given statistical sample $\{X_1, X_2, \cdots X_n\}$, the order statistics is obtained by sorting the sample in ascending order. It denoted as $\{X_{(1)}, X_{(2)}, \cdots X_{(n)}\}$. The k^{th} smallest value $X_{(k)}$ is called k^{th} order statistic

Theorem 2.1. Let $\{X_1, X_2, \cdots X_n\}$ be n i.i.d random variables with common CDF = $F(x)$ and common PDF = $f(x)$, then the marginal probability distribution of k^{th} order statistic (CDF) is denoted by $F_{(k,n)}(x)$ and it is given by

$$F_{(k,n)}(x) = \sum_{j=k}^{n} {}^{n}\mathrm{C}_j \times (F(x))^j \times (1 - F(x))^{n-j} \tag{2.0.1}$$

Proof.

$$F_{(k,n)}(x) = \Pr\left(X_{(k)} \le x\right) \tag{2.0.2}$$

$$F_{(k,n)}(x) = \Pr\left(\text{At least k elements have value} \le x\right) \tag{2.0.3}$$

Since $\Pr(X \le x) = F(x)$, Let $Q \sim Bern(F(x))$

$$\Pr(Q = 1) = F(x) \tag{2.0.4}$$

$$\Pr(Q = 0) = 1 - F(x) \tag{2.0.5}$$

Let $P \sim B(n, F(x))$ taking n trails from $Bern(F(x))$

$$\Pr(P = i) = {}^{n}\mathrm{C}_i \Pr(Q = 1)^i \Pr(Q = 0)^{n-i} \tag{2.0.6}$$

$$\Pr(P = i) = {}^{n}\mathrm{C}_i\, F(x)^i\, (1 - F(x))^{n-i} \tag{2.0.7}$$

Equation (2.0.7) is probability of exactly i R.V of given sample have values $\le x$

$$F_{(k,n)}(x) = \Pr(P \ge k) = \sum_{j=k}^{n} \Pr(P = j) \tag{2.0.8}$$

$$\therefore F_{(k,n)}(x) = \sum_{j=k}^{n} {}^{n}\mathrm{C}_j\, (F(x))^j\, (1 - F(x))^{n-j} \tag{2.0.9}$$

□

Theorem 2.2. Let $\{X_1, X_2, \cdots X_n\}$ be n i.i.d random variables with common CDF = $F(x)$ and common PDF = $f(x)$, then the marginal probability density of k^{th} order statistic (PDF) is denoted by $f_{(k,n)}(x)$ and it is given by

$$f_{(k,n)}(x) = n\; {}^{n-1}\mathrm{C}_{k-1}\, f(x)\, (F(x))^{k-1}\, (1 - F(x))^{n-k} \tag{2.0.10}$$

Proof.

$$\frac{d}{dx} F_{(k,n)}(x) = \frac{d}{dx}\left(\sum_{j=k}^{n} {}^{n}\mathrm{C}_j\, (1 - F(x))^{n-j}\, F(x)^j\right) \tag{2.0.11}$$

$$f_{(k,n)}(x) = \sum_{j=k}^{n} {}^{n}\mathrm{C}_j\, (j)\, (1 - F(x))^{n-j}\, F(x)^{j-1}\, f(x) - \sum_{j=k}^{n} {}^{n}\mathrm{C}_j\, (n - j)\, (1 - F(x))^{n-j-1}\, F(x)^j\, f(x) \tag{2.0.12}$$

$$S_1 = \sum_{j=k}^{n} \frac{n!}{(n-j)!\,(j-1)!}\,(1-F(x))^{n-j}\,F(x)^{j-1} f(x) \tag{2.0.13}$$

$$S_2 = \sum_{j=k}^{n} \frac{n!}{(n-j-1)!\,j!}\,(1-F(x))^{n-j-1}\,F(x)^{j}\,f(x) \tag{2.0.14}$$

let $i = j+1$ change the limits for the summation in equation (2.0.14)

$$S_2 = \sum_{i=k+1}^{n} \frac{n!}{(n-i)!\,(i-1)!}\,(1-F(x))^{n-i}\,F(x)^{i-1}\,f(x) \tag{2.0.15}$$

$$f_{(k,n)}(x) = S_1 - S_2 \tag{2.0.16}$$

$$f_{(k,n)}(x) = \frac{n!\,f(x)\,(1-F(x))^{n-k}\,F(x)^{k-1}}{(n-k)!\,(k-1)!} \tag{2.0.17}$$

$$\therefore f_{(k,n)}(x) = n\,{}^{n-1}\mathrm{C}_{k-1}\;(1-F(x))^{n-k}\,F(x)^{k-1} f(x) \tag{2.0.18}$$

□

Method 1:
Let $X \in [0,2]$ be a random variable of uniform order statistic distribution of sample size 8 then

$$\int_0^2 \Pr(x)\;dx = 1 \tag{2.0.19}$$

$$\Pr(x) = \frac{1}{2}\;(\because \text{Uniform order}) \tag{2.0.20}$$

The PDF for X is

$$f(x) = \begin{cases} \frac{1}{2}, & 0 < x < 2, \\ 0, & \text{otherwise}, \end{cases} \tag{2.0.21}$$

The CDF for X is

$$F(x) = \begin{cases} 0, & x \le 0, \\ \frac{x}{2}, & 0 < x < 2, \\ 1, & x \ge 2 \end{cases} \tag{2.0.22}$$

Using theorem (2.2) PDF of k^{th} order statistic of given sample from equation (2.0.10)

$$f_{(k,n)}(x) = n\,{}^{n-1}\mathrm{C}_{k-1}\,\frac{1}{2}\left(\frac{x}{2}\right)^{k-1}\left(1-\frac{x}{2}\right)^{n-k} \tag{2.0.23}$$

$$f_{(k,8)}(x) = \frac{8}{2^{(1+(k-1)+(8-k))}} \times {}^{7}\mathrm{C}_{k-1}\,x^{k-1}\;(2-x)^{8-k} \tag{2.0.24}$$

$$f_{(k,8)}(x) = {}^{7}\mathrm{C}_{k-1}\,\frac{1}{32}\,x^{k-1}\;(2-x)^{8-k} \tag{2.0.25}$$

Comparing the PDF obtained in equation (2.0.25) with the equation (1.0.1)

$$\frac{1}{32}\,{}^{7}\mathrm{C}_{k-1}\;(2-x)^{8-k}\;x^{k-1} = \frac{7}{32}\;(2-x)\;x^{6} \tag{2.0.26}$$

$$\therefore k = 7 \tag{2.0.27}$$

Hence the marginal probability density given is 7^{th} order statistic and **the value of k is 7**

Definition 2.2. Uniform order statistics
Let $\{X_1, \cdots X_n\}$ be i.i.d form a uniform distribution on $[0,1]$ such that $f(x) = 1$ and $F(x) = x$, from theorem (2.2), equation (2.0.10)

$$f_{(k,n)}(x) = n\,{}^{n-1}\mathrm{C}_{k-1}\,x^{k-1}\;(1-x)^{n-k} \tag{2.0.28}$$

Since equation (2.0.28) is PDF

$$\int_0^1 n\,{}^{n-1}\mathrm{C}_{k-1}\,x^{k-1}\;(1-x)^{n-k}\;dx = 1 \tag{2.0.29}$$

$$\int_0^1 x^{k-1}\;(1-x)^{n-k}\;dx = \frac{(k-1)!\,(n-k)!}{n!} \tag{2.0.30}$$

$$\int_0^1 x^{k-1}\;(1-x)^{n-k}\;dx = \frac{\Gamma(k)\,\Gamma(n-k+1)}{\Gamma\left((n-k+1)+k\right)} \tag{2.0.31}$$

Definition 2.3. Beta function
From definition (2.2), equation (2.0.31) let $r = k$ and $s = n-k+1$ The **Beta function** is defined for r, s > 0

$$B(r,s) = \int_0^1 x^{r-1}\;(1-x)^{s-1}\;dx = \frac{\Gamma(r)\,\Gamma(s)}{\Gamma(r+s)} \tag{2.0.32}$$

Beta Distribution

The Beta distribution is a continuous distribution defined on the range $(0,1)$ whose PDF given by

$$f(x) = \frac{1}{B(r,s)} x^{r-1} (1-x)^{s-1} \tag{2.0.33}$$

where $\int_0^1 f(x) = 1$ as per definition (2.2)

CDF, Mean value and Variance of Beta distribution

$$F(x) = \frac{\int_0^x x^{r-1} (1-x)^{s-1}}{B(r,s)} = \frac{B_x(r,s)}{B(r,s)} \tag{2.0.34}$$

$$E(x) = \frac{r}{r+s} \tag{2.0.35}$$

$$Var(x) = \frac{rs}{(r+s)^2 (r+s+1)} \tag{2.0.36}$$

In Uniform order statistics on [0,1] the PDF of k^{th} order statistic follows Beta distribution with $r = k$, $s = n-k+1$ and PDF is given by

$$f(x) = \frac{1}{B(k, n-k+1)} x^{k-1} (1-x)^{(n-k+1)-1} \tag{2.0.37}$$

Method 2:

we know that, PDF of k^{th} order statistic of a uniform distribution on $[0,1]$ follows beta distribution

$$\int_0^2 f(x)\,dx = \int_0^2 \frac{7}{32} x^6 (2-x)\,dx \tag{2.0.38}$$

$$\int_0^2 f(x)\,dx = \int_0^2 56 \left(\frac{x}{2}\right)^6 \left(1 - \frac{x}{2}\right) d\left(\frac{x}{2}\right) \tag{2.0.39}$$

Let new random variable be t such that $t = x/2$, New sample be $\{T_1, \cdots T_8\}$ such that $T_i = X_i/2$.

$$f(t) = 56\,t^6 (1-t) \tag{2.0.40}$$

$$\int_0^2 f(x)\,dx = \int_0^1 f(t)\,dt = 1 \tag{2.0.41}$$

The Uniform distribution of new random sample is on $[0,1]$ such that PDF $= 1$ and CDF $= t$

$f(k,8)(x)$ in equation (1.0.1) (after conversion)

$$f_{(k,8)}(t) = \begin{cases} 56\,t^6 (1-t), & 0 < t < 1, \\ 0, & \text{otherwise,} \end{cases} \tag{2.0.42}$$

Since equation (2.0.42) is a Beta distribution with $r = k,\ s = n-k+1$

$$r - 1 = k - 1 = 6 \tag{2.0.43}$$

$$\therefore k = 7 \tag{2.0.44}$$

Hence the **value of k is 7**

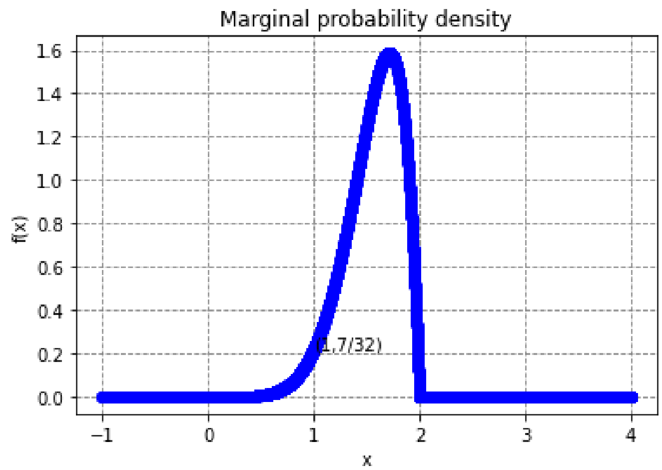

Fig. 1: PDF of $f_{(7,8)}(x)$

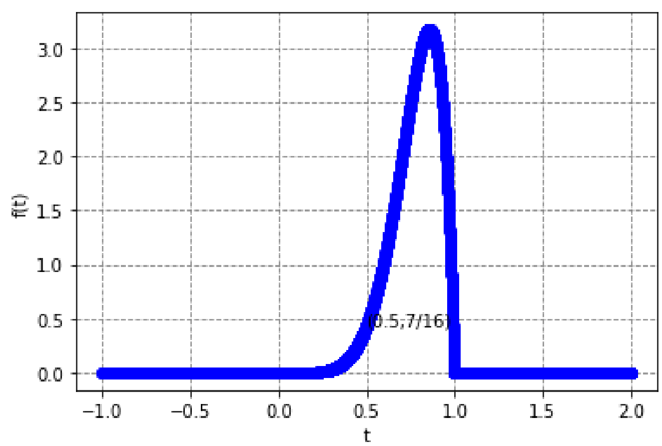

Fig. 2: PDF of $f_{(7,8)}(t)$

Presentation link:

https://github.com/Suraj11050/Assignments-AI1103/tree/main/Assignment4presentation