Ancillary Statistics

Def 6.2.16 :- A statistic $S(X)$ whose distribution does not depend on the parameter θ is called a ancillary statistic.

$\Rightarrow$ An ancillary statistic contain's no information about θ.

$\Rightarrow$ Observation on r.v whose distribution is fixed and known, unrelated to θ.

Example 6.2.17 (Uniform ancillary statistic)

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{unif}(\theta, \theta+1)$

$-\infty < \theta < \infty.$

Let $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ be the order statistics from the sample.

$R = X_{(n)} - X_{(1)}$ is an ancillary statistic

$$f(x|\theta) = \begin{cases} 1 & \text{if} \quad \theta < x_i < \theta+1 \\ 0 & \text{o.w} \end{cases}$$

$$F(x|\theta) = \begin{cases} 0 & \text{if} \quad x_i < \theta \\ x-\theta & \text{if} \quad \theta < x_i < \theta+1 \\ 1 & \text{if} \quad x_i > \theta+1 \end{cases}$$

The Joint pdf of two ordered statistis is

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

The Joint Pdf of $X_{(1)}$, $X_{(n)}$

$$f_{X_{(1)},X_{(n)}}(x_{(1)}, x_{(n)})$$

$$= \frac{n!}{(n-2)!} \cdot [x_{(1)} - \theta]^0 \cdot [x_{(n)} - \theta - x_{(1)} + \theta)^{n-2}$$

$$= \frac{n!}{(n-2)!} \cdot (x_{(n)} - x_{(1)})^{n-2}$$

$$\Rightarrow f_{X_{(1)},X_{(n)}}(x_{(1)}, x_{(n)} | \theta) = \begin{cases} n(n-1)(x_{(n)} - x_{(1)})^{n-2} & \theta < x_{(1)} < x_{(n)} < \theta + 1 \\ 0 & \text{o.w} \end{cases}$$

$$R = X_{(n)} - X_{(1)}, \quad M = \frac{X_{(1)} + X_{(n)}}{2}$$

$$\Rightarrow \quad X_{(1)} = \frac{2M - R}{2}$$

$$X_{(n)} = \frac{2M + R}{2}$$

$$\Rightarrow \quad J = \begin{vmatrix} -\frac{1}{2} & 1 \\ \frac{1}{2} & 1 \end{vmatrix} = 1$$

$$f_{R,M}(r,m) = f_{X_{(1)},X_{(n)}}\left(\frac{2m-R}{2}, \frac{2m+R}{2}\right)$$

$$= \begin{cases} n(n-1)\, r^{n-2} & 0<r<1 \\ & \theta+\frac{r}{2} < m < \theta+1-\frac{r}{2} \\ 0 & \text{o.w} \end{cases}$$

The pdf of R is

$$f_R(r|\theta) = \int_{\theta+r/2}^{\theta+1-\frac{r}{2}} n(n-1)\, r^{n-2}\, dm$$

$$= n(n-1)\, r^{n-2}(1-r) \qquad 0<r<1$$

$$\Rightarrow \quad f_R(r|\theta) = n(n-1)r^{n-2}(1-r) \qquad 0<r<1$$

$$R \sim \text{beta}(n-1, 2)$$

$\Rightarrow$ R does not depend on θ

$\Rightarrow$ R is a ancillary statistic.

Example 6.2.18: (Location family ancillary statistic)

$X_1, X_2, \ldots X_n \overset{iid}{\sim}$ location parameter family

with CDF $F(x-\theta)$ $-\infty<\theta<\infty$.

we will show that $R = X_{(n)} - X_{(1)}$ is a ancillary statistic.

$\Rightarrow \quad Z_1, Z_2, \ldots, Z_n \overset{iid}{\sim} F(x) \quad (\theta=0)$

$X_1 = Z_1+\theta, \quad \ldots, \quad X_n = Z_n+\theta.$

$$F_R(x|\theta) = P_\theta(R \leq x)$$

$$= \mathbb{P}_\theta\left(\max_i x_i - \min_i x_i \leq x\right)$$

$$= \mathbb{P}_\theta\left(\max(z_i+\theta) - \min(z_i+\theta) \leq x\right)$$

$$= \mathbb{P}\left(\max z_i - \min_i z_i \leq x\right)$$

independent of θ

$\Rightarrow$ R is ancillary statistic for location family distribution.