3. Properties of a Sufficient Statistic

Suppose $X_1, \dots, X_n$ is a random sample with a "density" $f(x, \theta)$, $\theta \in \Theta$

Remark 1

A sufficient statistic is not unique

Proof

Let $Y_1 = u_1(X_1, \dots, X_n)$ be a sufficient statistic for θ.

Let $Y_2 = g(Y_1)$, where $g : \mathbb{R} \xrightarrow{1-1} \mathbb{R}$. Then,

$$\prod_{i=1}^{n} f(x_i, \theta) = k_1[u_1(x_1, \dots, x_n), \theta] \cdot k_2(x_1, \dots, x_n)$$
$$= k_1(y_1, \theta) k_2(\underline{x}) = k_1(g^{-1}(y_2), \theta) k_2(\underline{x}).$$

By the factorization theorem Y_2 is also a sufficient statistic.

Lemma 1

If X_1 and X_2 are random variables such that ~~$Var X_1 \geq$~~ the variance of X_2 exist, then

$$\mathbb{E} X_2 = \mathbb{E}[\mathbb{E}[X_2 | X_1]] \quad \text{and} \quad Var X_2 \geqslant Var[\mathbb{E}[X_2 | X_1]].$$

Y_1 - sufficient statistic for θ

Y_2 - unbiased estimator of θ

$$\mathbb{E}[Y_2 | Y_1] = \varphi(Y_1),$$
$$\theta = \mathbb{E}[Y_2] = \mathbb{E}[\varphi(Y_1)],$$
$$Var Y_2 \geqslant Var[\varphi(Y_1)].$$

Theorem 1 (Rao-Blackwell)

Let $X_1, \dots, X_n$ be a random sample with the "density" $f(x;\theta)$, $\theta \in \Theta$. Let $Y_1 = u_1(X_1, \dots, X_n)$ be a sufficient statistic for θ, and let $Y_2 = u_2(X_1, \dots, X_n)$, not a function of Y_1, be an unbiased estimator of θ. Then $\mathbb{E}[Y_2 | Y_1] = \varphi(Y_1)$ (a function of the sufficient statistic) is an unbiased estimator of θ, and its variance is smaller than ~~be equal to~~ the variance of Y_2.

ok

Theorem 2

Let $X_1, \dots, X_n$ be a random sample with $f(x;\theta)$, $\theta \in \Theta$. If a sufficient statistic $Y_1 = u_1(X_1, \dots, X_n)$ for θ exists and a maximum likelihood estimator $\hat{\theta}$ of $\hat{\theta}$ exists and is unique, then $\hat{\theta}$ is a function of Y_1.

Proof

1. $f_{Y_1}(y_1;\theta)$ - density of Y_1.
2. We have

$$L(\theta) = L(\theta, x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i; \theta) = f_{Y_1}(y_1; \theta) \cdot H(x_1, \dots, x_n) =$$

$$f_{Y_1}[u_1(x_1, \dots, x_n); \theta] \cdot H(x_1, \dots, x_n), \text{ where } H(x_1, \dots, x_n)$$

does not depend on θ.

3. Thus L and f_{Y_1} as functions of θ are maximized simultaneously.
4. Since $\hat{\theta}$ is unique, it also maximized f_{Y_2} and must

be a function of $u_1(x_1, \dots, x_n)$.

Example 1

$X_1, \dots, X_n$ i.i.d $\quad X_i \sim f(x, \theta) = \theta e^{-\theta x} \mathbb{1}_{(0,+\infty)}(x), \ \theta > 0$

We want to find the MVUE of θ.

$$L(\theta) = \theta^n e^{-\theta \sum_{i=1}^n x_i}$$

$$Y_1 = \sum_{i=1}^n X_i \text{ - sufficient statistic}$$

$$\ell(\theta) = \log L(\theta) = n \log \theta - \theta \sum_{i=1}^n x_i$$

$$\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0 \Rightarrow \theta = \frac{1}{\bar{x}}$$

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0 \Rightarrow \hat{\theta} = \frac{1}{\bar{X}} \text{ MLE of } \theta.$$

Y_2 - asymptotically unbiased

$X_1 \sim \Gamma(1, \frac{1}{\theta})$, $Y_1 \sim \Gamma(n, \frac{1}{\theta})$, and

$$\mathbb{E}\left[\frac{1}{\bar{X}}\right] = n\,\mathbb{E}\left[\frac{1}{Y_1}\right] = n \int_0^\infty \frac{1}{x} \frac{\theta^n}{\Gamma(n)} x^{n-1} e^{-\theta x} dx =$$

$$n \int_0^\infty \frac{\theta^n}{\Gamma(n)} x^{n-2} e^{-\theta x} dx = n \frac{\Gamma(n-1)\,\theta}{\Gamma(n)} \int_0^\infty \frac{\theta^{n-1}}{\Gamma(n-1)} x^{n-2} e^{-\theta x} dx =$$

$$\frac{n}{n-1}\theta.$$

Thus $\dfrac{n-1}{\sum_{i=1}^n X_i}$ is the MVUE of the parameter θ.

Remark 2.

Y_2 - estimate of θ, $EY_2 = \theta$, Y_1 - sufficient statistic

$\varphi(Y_1) = \mathbb{E}[Y_2 | Y_1]$, $Var\, \varphi(Y_1) \leqslant Var\, Y_2$.

Y_3 - estimate of θ, $EY_3 = \theta$, Y_3 - is not a suff. stat.

$\psi(Y_3) = \mathbb{E}[\varphi(Y_1) | Y_3]$, $\mathbb{E}\psi(Y_3) = \theta$ and $Var\, \psi(Y_3) \leqslant Var\, \varphi(Y_1)$

Since Y_3 is not a sufficient statistic, the conditional distribution of Y_1 given Y_3 depends upon θ. Thus $\psi(Y_3)$ is not a statistic. because $\psi(Y_3)$ depends on θ.

Example 2.

X_1, X_2, X_3 i.i.d. $Exp(\theta)$, $\theta > 0$.

$$(X_1, X_2, X_3) \sim \frac{1}{\theta^3} e^{-(x_1+x_2+x_3)/\theta}, \quad x_i > 0, \; i = 1, 2, 3$$

The factorization theorem implies that $Y_1 = X_1 + X_2 + X_3$ is a sufficient statistic for θ.

$\mathbb{E}Y_1 = \mathbb{E}(X_1 + X_2 + X_3) = 3\theta$. Thus $\mathbb{E}\left[\frac{Y_1}{3}\right] = \theta$. $\bar{X} = \varphi(Y_1) = \frac{Y_1}{3}$

Let $Y_2 = X_2 + X_3$, $Y_3 = X_3$. The one-to-one transformation

$$x_1 = y_1 - y_2, \quad x_2 = y_2 - y_3, \quad x_3 = y_3$$

has Jacobian

$$J = \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{vmatrix} = 1,$$

and the joint distribution of (Y_1, Y_2, Y_3) is

$$g(y_1, y_2, y_3) = \frac{1}{\theta^3} e^{-y_1/\theta} \mathbb{1}(0 < y_3 < y_2 < y_1 < +\infty)$$

The joint distribution of (Y_1, Y_3)

$$g_{13}(y_1, y_3, \theta) = \int_{y_3}^{y_1} g(y_1, y_2, y_3)\, dy_2 = \int_{y_3}^{y_1} \left(\frac{1}{\theta}\right)^3 e^{-y_1/\theta}\, dy_2 =$$

$$= \frac{1}{\theta^3} e^{-y_1/\theta} (y_1 - y_3), \quad 0 < y_3 < y_1 < +\infty.$$

Obviously, $g_3(y_3, \theta) = \frac{1}{\theta} e^{-y_3/\theta}$, $0 < y_3 < +\infty$.

The condition distribution of Y_1 given $Y_3 = y_3$, is

$$g_{1|3}(y_1|y_3) = \frac{g_{13}(y_1, y_3, \theta)}{g_3(y_3, \theta)} = \frac{\frac{1}{\theta^3} e^{-y_1/\theta}(y_1 - y_3)}{\frac{1}{\theta} e^{-y_3/\theta}} =$$

$$\frac{1}{\theta^2}(y_1 - y_3) e^{-(y_1 - y_3)/\theta}, \quad 0 < y_3 < y_1 < +\infty$$

We have

$$\mathbb{E}\left[\frac{Y_1}{3} \middle| Y_3\right] = \mathbb{E}\left[\frac{Y_1 - Y_3}{3} \middle| Y_3\right] + \mathbb{E}\left[\frac{Y_3}{3} \middle| Y_3\right] =$$

$$\frac{1}{3}\int_{y_3}^{+\infty} \frac{1}{\theta^2}(y_1 - y_3)^2 e^{-(y_1 - y_3)/\theta}\, dy_1 + \frac{Y_3}{3} = \left\{ \begin{array}{l} z = y_1 - y_3 \\ dz = dy_1 \end{array} \right\} =$$

$$\frac{1}{3}\int_0^{+\infty} \frac{1}{\theta^2} z^2 e^{-z/\theta}\, dz + \frac{Y_3}{3} = \frac{1}{3}\int_0^{+\infty} \frac{1}{\theta^2} \frac{\Gamma(3)\theta^3}{\underbrace{\Gamma(3)\theta^3} z^2 e^{-z/\theta}}_{\Gamma(3,\theta)}\, dz + \frac{Y_3}{3} =$$

$$= \frac{1}{3}\Gamma(3)\theta + \frac{Y_3}{3} = \frac{2}{3}\theta + \frac{Y_3}{3}.$$

Put $\psi(Y_3) = \mathbb{E}\left[\frac{Y_1}{3} \middle| Y_3\right] = \frac{2}{3}\theta + \frac{Y_3}{3}$.

$\mathbb{E}\,\psi(Y_3) = \theta$ and $Var\,\psi(Y_3) \leq Var\left(\frac{Y_1}{3}\right)$,

but $\psi(Y_3)$ is not a statistic. (depends upon the parameter θ).