Chapter 6

Stat 562
2-14-19

(1)

let $X_1, \dots, X_n \sim$ iid $f(x|\theta)$

Then the joint pdf of $X_1, \dots, X_n$ is

$$f(\vec{x}\,|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

let $T(\vec{X})$ be a function of $X_1, \dots, X_n$,

but free of θ.

let $g(t|\theta)$ be the pdf of the r.v. $T(\vec{X})$ (2)

Then the conditional distribution of $\vec{X}$, given T

is $\dfrac{\prod_{i=1}^{n} f(x_i|\theta)}{g(t|\theta)}$

Defn: If this ratio is free of θ,

then T is a sufficient statistic for θ.

Example: $X_1, \dots, X_n \sim$ iid $N(\theta, \sigma^2)$ ← known

let $T = \sum_{i=1}^{n} X_i$

Show that T is sufficient for θ.

$T \sim N(n\theta, n\sigma^2)$

$$\frac{\prod_{i=1}^{n} f(x_i|\theta)}{g(t|\theta)} = \frac{\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma}\right)^2}}{\frac{1}{\sigma\sqrt{n}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-n\theta}{\sigma\sqrt{n}}\right)^2}}$$

$$= \frac{\sigma\sqrt{n}\sqrt{2\pi}}{\sigma^n(\sqrt{2\pi})^n} \frac{e^{-\frac{1}{2\sigma^2}(\sum x_i^2 - \cancel{2\theta\sum x_i} + \cancel{n\theta^2})}}{e^{-\frac{1}{2n\sigma^2}(t^2 - \cancel{2nt\theta} + \cancel{n^2\theta^2})}}$$

& this is free of θ.

Example: $X_1, \dots, X_n \sim$ iid $f(x|\theta) = e^{-(x-\theta)}$ $x > \theta$

let $T = X_{(1)}$

We know from a previous example that

T will have an exponential distribution with scale parameter n.

$$\frac{f(\vec{x}|\theta)}{g(t|\theta)} = \frac{\prod_{i=1}^{n} e^{-(x_i-\theta)}}{n e^{-n(t-\theta)}} = \frac{e^{-\Sigma x_i} \cancel{e^{n\theta}}}{n e^{-nt} \cancel{e^{n\theta}}}$$ ⑤

$\therefore T = X_{(1)}$ is sufficient for θ.

Theorem (Neyman): T is a sufficient statistic for θ iff

$$\prod_{i=1}^{n} f(x_i|\theta) = k(t|\theta)\underbrace{h(\vec{x})}_{t \text{ free of } \theta}$$

⑥

Proof: $(\Rightarrow)$ If T is sufficient,

then $$\frac{f(\vec{x}|\theta)}{g(t|\theta)} = \underbrace{h(\vec{x})}_{\text{free of } \theta}$$

So $f(\vec{x}|\theta) = g(t|\theta)h(\vec{x})$

$(\Leftarrow)$ Assume $\prod_{i=1}^{n} f(x_i|\theta) = k(t|\theta)h(\vec{x})$

Let $T_1 = T$

⑦

Create
$$\left.\begin{array}{l} T_2 = u_2(X_1, \dots, X_n) \\ \vdots \\ T_n = u_n(X_1, \dots, X_n) \end{array}\right\}$$
so that the system is invertible

So
$$\left.\begin{array}{l} X_1 = v_1(T_1, \dots, T_n) \\ X_2 = v_2(T_1, \dots, T_n) \\ \vdots \\ X_n = v_n(T_1, \dots, T_n) \end{array}\right]$$
Find J. It is free of θ.

⑧

$$g(t_1, \dots, t_n) = f(x_1, \dots, x_n)\,|J|$$
$$= k(t|\theta)\,\underbrace{h(\vec{x})\,|J|}_{\text{free of } \theta}$$

$$g(t_1) = \underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-1} k(t_1|\theta)\,h(\vec{x})\,|J|\,dt_2\cdots dt_n$$

$$= k(t_1|\theta)\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\overbrace{h(\vec{x})}^{\text{function of } t_1, \dots, t_n}|J|\,dt_2\cdots dt_n}_{\text{can only be a function of } t_1}$$

So $g(t) = k(t|\theta) \underbrace{m(t)}_{\text{free of } \theta}$

Then $\dfrac{f(\vec{x}|\theta)}{g(t|\theta)} = \dfrac{\cancel{k(t|\theta)}\, h(\vec{x})}{\cancel{k(t|\theta)}\, m(t)}$, free of θ

So T is sufficient.

Example: $X_1, \ldots, X_n \sim$ iid $f(x|\theta) = \theta x^{\theta - 1}$
$0 < x < 1$

Find a sufficient statistic for θ.

$$f(\vec{x}|\theta) = \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$$

$$= \underline{\theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}} \cdot \underline{\quad 1 \quad}$$

So $T = \prod_{i=1}^{n} x_i$ is sufficient for θ.

(11)

Note: the factorization isn't unique

$$f(\vec{x}|\theta) = \underline{\theta^n\left(\prod_{i=1}^n x_i\right)^\theta} \cdot \underline{\left(\prod_{i=1}^n x_i\right)^{-1}}$$

Defn. T is a minimal sufficient statistic for θ if, for any other sufficient statistic T^* of θ, T is a function of T^*.

(12)

Example: $X_1, \dots, X_n \sim \text{iid } N(\theta, \sigma^2)$

σ^2 known

$T = \bar{X}$ was sufficient for θ.

Let $T^* = (\bar{X}, S^2)$

T is a function of T^*,

but T^* is not a function of T.

So T^* is definitely not minimal

(13)

<u>theorem:</u> let $X_1, \ldots, X_n \sim$ iid $f(x|\theta)$

Suppose that a statistic $T(\vec{X})$ has the following property:

$$* \left\{ \begin{array}{l} \text{The ratio } \dfrac{f(\vec{x}|\theta)}{f(\vec{y}|\theta)} \text{ is free of } \theta \\ \text{iff } T(\vec{x}) = T(\vec{y}) \end{array} \right.$$

Then T is a minimal sufficient statistic for θ.

(14)

Proof: next time

HW #5 due 2/21

Chp6 *6
9
10
13

Stat 562 HW#5

6.6 Let $X_1, \ldots, X_n$ be a random sample from a gamma(α, β) population. Find a two-dimensional sufficient statistic for (α, β).

6.9 For each of the following distributions let $X_1, \ldots, X_n$ be a random sample. Find a minimal sufficient statistic for θ.

(a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$ (normal)

(b) $f(x|\theta) = e^{-(x-\theta)}, \quad \theta < x < \infty, \quad -\infty < \theta < \infty$ (location exponential)

(c) $f(x|\theta) = \frac{e^{-(x-\theta)}}{\left(1+e^{-(x-\theta)}\right)^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$ (logistic)

(d) $f(x|\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$ (Cauchy)

(e) $f(x|\theta) = \frac{1}{2} e^{-|x-\theta|}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$ (double exponential)

6.10 Show that the minimal sufficient statistic for the uniform$(\theta, \theta+1)$, found in Example 6.2.15, is not complete.

6.13 Suppose X_1 and X_2 are iid observations from the pdf $f(x|\alpha) = \alpha x^{\alpha-1} e^{-x^\alpha}, x > 0, \alpha > 0$. Show that $(\log X_1)/(\log X_2)$ is an ancillary statistic.