Sufficient Estimators

Gene Quinn

Sufficient Estimators

A more abstract property of considerable theoretical importance is *sufficiency*

Sufficient Estimators

A more abstract property of considerable theoretical importance is *sufficiency*

Losely speaking, an estimator $\hat{\theta}$ is said to be a **sufficient estimator** for the parameter θ if $\hat{\theta}$ contains all of the information relevant to θ that it is possible to get from the sample.

Sufficient Estimators

A more abstract property of considerable theoretical importance is *sufficiency*

Losely speaking, an estimator $\hat{\theta}$ is said to be a **sufficient estimator** for the parameter θ if $\hat{\theta}$ contains all of the information relevant to θ that it is possible to get from the sample.

A common example is the following: Very often the sample mean $\overline{x}$ is a sufficient estimator for the population mean μ.

If we know the sample mean $\overline{x}$, knowledge of the individual x_i values often does not provide any additional information about μ.

Sufficient Estimators

A formal definition of sufficiency is necessarily rather technical.

Sufficient Estimators

A formal definition of sufficiency is necessarily rather technical.

Suppose $x_1, \ldots, x_n$ is a random sample of size n from a distribution with density function $f(x; \theta)$.

The estimator $\hat{\theta} = h(x_1, \ldots, x_n)$ is said to be **sufficient** for θ if the likelihood function of the sample factors into the product of the density function of $\hat{\theta}$ and other factors that do not involve θ.

That is,

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) \;=\; f_{\hat{\theta}}(\hat{\theta}; \theta) \cdot b(x_1, \ldots, x_n)$$

Sufficient Estimators

Example: We saw that for a random sample $x_1, \ldots, x_n$ from an exponential distribution, the maximum likelihood and method of moments estimators are both equal to the sample mean $\overline{x}$,

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

Sufficient Estimators

The likelihood function of the sample is

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right) = \frac{1}{\theta^n} \exp\left(-\frac{n\overline{x}}{\theta}\right)$$

Note that, once we have written $L(\theta)$ in terms of $\overline{x}$, the individual x_i values no longer appear in the expression.

Sufficient Estimators

The likelihood function of the sample is

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right) = \frac{1}{\theta^n} \exp\left(-\frac{n\overline{x}}{\theta}\right)$$

Note that, once we have written $L(\theta)$ in terms of $\overline{x}$, the individual x_i values no longer appear in the expression.

This reflects the fact that $\overline{x}$ is a sufficient estimator for θ.

Sufficient Estimators

Of course, the likelihood function $L(\theta)$ does depend on the x_i values in the sense that $\overline{x}$ depends on the x_i values.

Sufficient Estimators

Of course, the likelihood function $L(\theta)$ does depend on the x_i values in the sense that $\overline{x}$ depends on the x_i values.

However, *any* random sample $z_1, \ldots, z_n$ with sample mean equal to $\overline{x}$ will produce the same likelihood function.

Sufficient Estimators

Of course, the likelihood function $L(\theta)$ does depend on the x_i values in the sense that $\overline{x}$ depends on the x_i values.

However, *any* random sample $z_1, \ldots, z_n$ with sample mean equal to $\overline{x}$ will produce the same likelihood function.

When we say that $L(\theta)$ does not depend on the individual x_i values, we mean that *for a given value of* $\overline{x}$, all random samples with sample mean equal to $\overline{x}$ have the same likelihood.

Sufficient Estimators

Example: Let $x_1, \ldots, x_n$ be a sequence of zeros and ones corresponding to a random sample consisting of n independent Bernoulli trials with probability of success p. The likelihood function of the sample is

$$L(p) \ = \ \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \ = \ p^{\sum x_i}(1-p)^{n-\sum x_i}$$

and the maximum likelihood estimator of p is $\hat{p} = \sum x_i / n = \overline{x}$

Sufficient Estimators

Example: Let $x_1, \ldots, x_n$ be a sequence of zeros and ones corresponding to a random sample consisting of n independent Bernoulli trials with probability of success p. The likelihood function of the sample is

$$L(p) \;=\; \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \;=\; p^{\sum x_i}(1-p)^{n-\sum x_i}$$

and the maximum likelihood estimator of p is $\hat{p} = \sum x_i / n = \overline{x}$

We can write the likelihood function in terms of $\overline{x}$ as

$$L(p) \;=\; p^{n\overline{x}}(1-p)^{n(1-\overline{x})}$$

and once again the x_i values disappear from the expression, indicating that $\overline{x}$ is a sufficient estimator for p.

Sufficient Estimators

As before, once $\overline{x}$ is fixed, the likelihood function $L(p)$ no longer depends on the individual x_i values.