STAT 703/J703
March 24th, 2005
-*Lecture 20*-

Instructor: Brian Habing
Department of Statistics
LeConte 203
Telephone: 803-777-3578
E-mail: habing@stat.sc.edu

Today

Some Properties of Estimators

- Efficiency
- Cramer-Rao Inequality
- Sufficiency
- Factorization Theorem
- Rao-Blackwell

8.6 Comparing Estimates and Tests

One of the standard tools for evaluating an estimate is the mean squared error:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta_0)^2$$
$$= Var(\hat{\theta}) + [E(\hat{\theta}) - \theta_0]^2$$

If two estimators are unbiased, then the efficiency of $\hat{\theta}$ relative to $\tilde{\theta}$ is

$$eff(\hat{\theta},\tilde{\theta}) = \frac{\mathrm{var}(\tilde{\theta})}{\mathrm{var}(\hat{\theta})}$$

For two tests T_1 and T_2 of the same H_0 and H_A with the same α-level, the relative efficiency of T_1 to T_1 is the ratio n_2/n_1 required so that they have the same power.

The <u>asymptotic relative efficiency</u> of the MWW to the *t*-test is:

- 0.955 if the populations are normal
- 1.0 if the populations are uniform
- 1.5 if the pops. are double-exp.
- 0.864 to infinity in general

assuming the populations differ only by location.

The asymptotic relative efficiency of the MWW to the median test is:

- 1.5 if the populations are normal
- 3.0 if the populations are uniform
- 0.75 if the pops. are double-exp.

assuming the populations differ only by location.

Cramer-Rao Inequality

Let $X_1, \ldots X_n$ be I.I.d. with density $f(x|\theta)$, and T be an unbiased estimate of θ. Then under appropriate smoothness assumptions on f

$$Var(T) \geq \frac{1}{nI(\theta)}$$

Recall that under smoothness conditions that the $1/n\,I(\theta)$ is the asymptotic variance of the MLE!

So, why isn't the MLE always best?

8.7 Sufficiency One of the key concepts in advanced mathematical statistics is that of sufficiency. Does a statistic summarize all of the information in the data about a parameter, or do we lose something by summarizing?

Defn A statistic $T(X_1, \dots X_n)$ is sufficient for θ if the conditional distribution of $X_1, \dots X_n$ given T=t does not depend on θ for any value of t.

Example: Consider a Poisson Distribution with parameter λ and a sample size of 2.

A) Consider $T=X_1 + X_2$

B) Consider $T=X_1 + 2X_2$

The Factorization Theorem

A necessary and sufficient condition for T to be sufficient for θ is that the joint p.d.f. factors in the form:

$$f(x_1,\ldots x_n \mid \theta) = g[T(x_1,..x_n),\theta]h(x_1,\ldots x_n)$$

STAT 703/J703 B.Habing Univ. of SC 13

Rao-Blackwell Theorem

Let $\hat{\theta}$ be an estimator of q with $E(\hat{\theta})^2<\infty$ for all q. If T is sufficient for q then $\tilde{\theta} = E(\hat{\theta} \mid T)$ satisfies

$$E(\tilde{\theta}-\theta)^2 \le E(\hat{\theta}-\theta)^2$$

for all θ.

STAT 703/J703 B.Habing Univ. of SC 14