22S:166 Computing in Statistics
Simulation studies in statistics
Lecture 12
October 2, 2006

Based on a lecture by Marie Davidian for
ST 810A - Spring 2005
Preparation for Statistical Research
North Carolina State University
`http://www4.stat.ncsu.edu/ davidian/st810a/`

Basics

- simulation studies are commonly done to evaluate the performance of a frequentist statistical procedure, or to compare the performance of two or more different procedures for the same problem
- enable us to see what happens "when many many samples of the same size are drawn from the same population"
- properties of estimators that are often evaluated by simulation
  - bias
  - mean squared error
  - coverage of confidence intervals
- properties of hypothesis tests also can be evaluated by simulation studies
  - size
  - power
- simulation studies are *experiments*, and the things you know about experimental design and sample size calculation apply

Terminology

- simulation: a numerical technique for conducting experiments on the computer
- Monte Carlo simulation: a computer experiment inolving random sampling from probability distributions
  - what statisticians usually mean by "simulations"

Rationale

- Properties of statistical methods must be established before the methods can safely be used in practice.
- But exact analytical derivations of properties are rarely possible
- *Large sample* approximations to properties are often possible
  - evaluation of the relevance of the approximation to (finite) sample sizes likely to be encountered in practice is needed
- Analytical results may require *assumptions* such as normality
  - What happens when these assumptions are violated? Analytical results, even large sample ones, may not be possible

Questions to be addressed regarding an estimator or testing procedure

- Is an estimator *biased* in finite samples? What is its *sampling variance*?
- How does it *compare* to competing estimators on the basis of bias, precision, etc.?
- Does a procedure for constructing a confidence interval for a parameter achieve the claimed *nominal level of coverage*?
- Does a hypothesis testing procedure attain the claimed *level* or *size*?
- If so, what *power* is possible against different alternatives to the null hypothesis? Do different test procedures deliver different power?

Role of Monte Carlo simulation

- Goal is to evaluate *sampling distribution* of an estimator under a particular set of conditions (sample size, error distribution, etc.)
- Analytic derivation of exact sampling distribution is not feasible
- Solution: Approximate the sampling distribution through simulation
  - Generate S independent data sets under the conditions of interest
  - Compute the numerical value of the estimator/test statistic $T(\text{data})$ for each data set, yielding $T_1, \ldots, T_S$
- If S is large enough, *summary statistics* across $T_1, \ldots, T_S$ should be good approximations to the true sampling properties of the estimator/test statistic under the conditions of interest

Simulation for properties of estimators

Simple example: Compare three estimators for the mean μ of a distribution based on i.i.d. draws $Y_1, \ldots, Y_n$

- Sample mean $T^{(1)}$
- Sample 20% trimmed mean $T^{(2)}$
- Sample median $T^{(3)}$

Remarks:

- If the distribution of the data is symmetric, all three estimators indeed estimate the mean
- If the distribution is skewed, they do not

Simulation procedure

For a particular choice of μ, n, and true underlying distribution

- Generate independent draws $Y_1, \ldots, Y_n$ from the distribution
- Compute $T^{(1)}$, $T^{(2)}$, $T^{(3)}$
- Repeat S times $\Rightarrow$

$$T_1^{(1)}, \ldots, T_S^{(1)}; \quad T_1^{(2)}, \ldots, T_S^{(2)}; \quad T_1^{(3)}, \ldots, T_S^{(3)}$$

- Compute for $k = 1, 2, 3$

$$\widehat{\text{mean}} = S^{-1} \sum_{s=1}^{S} T_s^{(k)} = \overline{T}^{(k)}, \ \widehat{\text{bias}} = \overline{T}^{(k)} - \mu$$

$$\widehat{\text{SD}} = \sqrt{(S-1)^{-1} \sum_{s=1}^{S} (T_s^{(k)} - \overline{T}^{(k)})^2},$$

$$\widehat{\text{MSE}} = S^{-1} \sum_{s=1}^{S} (T_s^{(k)} - \mu)^2 \approx \widehat{\text{SD}}^2 + \widehat{\text{bias}}^2$$

Relative efficiency

For a particular choice of μ,

Relative efficiency: For any estimators for which $E(T^{(1)}) = E(T^{(2)}) = \mu$

$$RE = \frac{\text{var}(T^{(1)})}{\text{var}(T^{(2)})}$$

is the relative efficiency of estimator 2 to estimator 1

- When the estimators are *not unbiased* it is standard to compute
$$RE = \frac{\text{MSE}(T^{(1)})}{\text{MSE}(T^{(2)})}$$
- In either case $RE < 1$ means estimator 1 is preferred (estimator 2 is inefficient relative to estimator 1 in this sense)

R code for example

```
> set.seed(3)

> S <- 1000

> n <- 15

> trimmean <- function(Y){mean(Y,0.2)}

> mu <- 1

> sigma <- sqrt(5/3)
```


Normal data:

```
> out <- generate.normal(S,n,mu,sigma)

> outsampmean <- apply(out$dat,1,mean)

> outtrimmean <- apply(out$dat,1,trimmean)

> outmedian <- apply(out$dat,1,median)

> summary.sim <- data.frame(mean=outsampmean,trim=outtrimmean,
+         median=outmedian)

> results <- simsum(summary.sim,mu)

> view(round(summary.sim,4),5)
First 5 rows

    mean   trim median
1 0.7539 0.7132 1.0389
2 0.6439 0.4580 0.3746
3 1.5553 1.6710 1.9395
4 0.5171 0.4827 0.4119
5 1.3603 1.4621 1.3452
```



```
> results
                      Sample mean Trimmed mean   Median
true value                  1.000        1.000    1.000
# sims                   1000.000     1000.000 1000.000
MC mean                     0.985        0.987    0.992
MC bias                    -0.015       -0.013   -0.008
MC relative bias           -0.015       -0.013   -0.008
MC standard deviation       0.331        0.348    0.398
MC MSE                      0.110        0.121    0.158
MC relative efficiency      1.000        0.905    0.694
```