# Monte Carlo Studies

A Monte Carlo study uses a computer experiment to evaluate procedures.

Often the procedure is a statistical method for inference.

It may be an estimation method or a statistical test.

The performance of a statistical method depends on the underlying distribution.

Performance:
- bias
- variance
- power
- significance level
- etc.

# Performance of Statistical Methods

In mathematical statistics, we work out such things as the bias, the variance, etc. of a statistical method, based on an assumed underlying distribution.

For complicated distributions, or for mixtures of distributions, or for distributions with outliers, the mathematical analysis may be extremely difficult or impossible.

There are many possible scenarios to investigate; that is, there are many possibilities for the underlying distributions.

Monte Carlo studies provide an alternative.

Monte Carlo studies use simulation.

# An Example: A Problem for a Research Statistician

Given a random sample from a mixture of two normal distributions, $p\mathrm{N}(\mu, \sigma^2)$ and $(1-p)\mathrm{N}(\mu, k^2\sigma^2)$. What is the "best" estimate of $\mu$?

Two possibilities would be $\bar{x}$, the average of the sample, and $\bar{x}_{W(\pi)}$, the Winsorized sample mean (set the largest $\pi$% of the values to the largest of the middle $(1-2\pi)$% and the smallest $\pi$% of the values to the smallest of the middle $(1-2\pi)$%).

How could we decide which is better, $\bar{x}$ or $\bar{x}_{W(\pi)}$?

If both estimators are unbiased, the one with smaller variance is the better one.

What are the variances?

They're hard to work out.

# Monte Carlo Experimentation

We estimate the variances by Monte Carlo methods.

Get a (pseudo-) random sample from the assumed distribution, compute $\bar{x}$ and $\bar{x}_{W(\pi)}$.

Do many times, so we get a sample of many $\bar{x}$'s and $\bar{x}_{W(\pi)}$'s.

We compute their sample variances. This is an estimate of the true variances of $\bar{x}$ and $\bar{x}_{W(\pi)}$ for the case in which the underlying population distribution is the one simulated.

# Monte Carlo Experimentation

A Monte Carlo study is an experiment.

The Monte Carlo study should adhere to the same high standards of any scientific experimentation:

- control
- reproducibility
- efficiency
- careful and complete documentation

# Monte Carlo Experimentation

In simulation, *control*, among other things, relates to the fidelity of a *nonrandom* process to a *random* process. The experimental units are only simulated using computer programs.

Questions about the setup of the computer experiment must be addressed. (How good are the pseudo-random number generators? etc.)

Likewise, *reproducibility* is predicated on good RNG's (or else on equally bad ones!)

The principles of good statistical design can improve the efficiency.

# Statistical Design and Analysis of Experiments

In experimentation, we have some *response* of interest.

We identify various *factors* that may affect the response.

Different values of the factors are called *levels*.

In statistical design, we often distinguish the factors as *blocking* factors or *treatment* factors.

The purpose may be to determine which treatment yields the best response for various blocking factors.

(Statistical design and analysis of experiments is one of the major fields of statistics. It is at the heart of analysis of variance (ANOVA), in which linear models are used to study the relationship of the factors to the response.)

# Monte Carlo Experimentation

A statistical experiment involves a number of *experimental units*. They are generally chosen randomly.

Each combination of levels of factors constitute a *cell* in the experiment, and the number of experimental units used in each cell determines the precision with which the response may be measured for that combination of factor levels.

Often the numbers are chosen the same for all cells, but sometimes it makes more sense to have more units in one cell than in others.

The number of experimental units in each cell in a Monte Carlo experiment is Monte Carlo sample size.

## Statistical Design and Analysis of Experiments

The principles of statistical design and analysis apply just as much to a Monte Carlo study as they do to any other scientific experiment.

Monte Carlo experimentation is one of the few chances that statisticians get to design an experiment, collect the data, and analyze the data.

# A Monte Carlo Experiment to Study the Variances

In the study to compare the two estimators of $\mu$ in the

$$p\mathrm{N}(\mu, \sigma^2) + (1-p)\mathrm{N}(\mu, k^2\sigma^2)$$

mixture distribution, the "treatments" are the estimators; the ordinary sample mean $\bar{x}$, and the Winsorized sample mean $\bar{x}_{W(\pi)}$. The Winsorized sample mean has a parameter, $\pi$.

The possible factors are $p$, $\mu$, $\sigma^2$, $k$, and $n$, the sample size.

Can these be standardized?

What other factors may be of interest?

The design of the experiment determines levels of the factors.

# Reporting a Monte Carlo Experiment

Sometimes, because of all of the factors involved in the experiment, it is difficult to present a clear description of experiment and the results.

The description of the experiment must include all relevant information, including exact descriptions of what was computed, and what software was used.

Tables and graphs can help to present the results. They are not substitutes for clear summary descriptions, however.

The numeric results are *estimates*. Reports of statistical estimates must always be accompanied by some indication of their variances. In tables showing Monte Carlo results, a good way of doing this is to provide the sample standard deviation in parentheses beside each reported value.