Quant II: Problem Set 1 – Estimators

Kenneth Benoit

Assigned: Wednesday, January 7, 2008
Due: Wednesday, January 14, 2008

1. R revision. Use the dataset from Benoit and Marsh (2008) for the following questions.

    (a) Obtain full descriptive statistics on the variables `votes1st`, `spend_total`, and `incumb`. Discuss these variables.

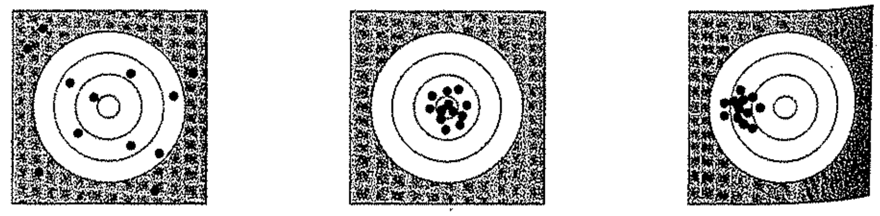
    (b) Make `incumb` into a factor for challengers and incumbents, with a 1 meaning incumbent.

    (c) Create a table of `wonseat` by `incumbent`, and do a χ^2 test on the resulting 2×2 table.

    (d) Perform a t-test for differences in `spend_total` by `incumbent`.

    (e) Regress `spend_total` on the variables `incumb`, `senator`, and `councillor`. Briefly discuss the results.

2. Bias, efficiecy, and MSE. Each of three guns is being tested by firing 12 shots at a target from a clamped position. Gun A was not clamped down hard enough, and wobbled. Gun B was clamped down in a position that pointed slightly to the left, due to a misaligned sight. Gun C was clamped down correctly.



    (a) Label each of the target patterns according to the gun that produced the pattern.

    (b) Which gun is biased?

    (c) Which gun has minimum variance?

    (d) Which gun has the largest mean squared errors?

    (e) Which gun is the least efficient?

3. When S successes occur in n trials, the sample proportion $P = S/n$ is generally used as an estimator of the probability of success of π. However, sometimes there are good reasons to use an alernative estimator $P^* = (S+1)/(n+2)$. Alternatively, P^* can be written as a linear combination of the familiar estimator P:

$$P^* = \frac{nP+1}{n+2} = \left(\frac{n}{n+2}\right)P + \left(\frac{1}{n+2}\right) \tag{1}$$

    (a) What is the MSE of P? Is it consistent?

    (b) What is the MSE of P^*? Is it consistent? (Hint: Calculate the mean and variance of P^*, in terms of the familiar mean and variance of P.)

    (c) To decide which estimator is better, P or P^*, does consistency help? What criterion would help?

    (d) Tabulate the efficiency of P^* relative to P, for example when $n = 10$ and $\pi = 0, .1, .2, \ldots, .9, 1.0$.

    (e) When might you prefer to use P^* instead of P to estimate π?

4. Extra credit: use R to plot or perform Monte Carlo simulations to show additional features of the answers to question 3 above.