Relative Efficiency and Coefficient of Variation

Marek Rychlik

Department of Mathematics
University of Arizona

February 16, 2009

The statement of the problem

Problem

In the table, the coefficients of variation and relative efficiencies (randomized complete block to completetely randomized) of the same experiment conducted at four locations are given. Each trial used a randomized complete block design.

Location	Coefficient of variation (%)	Relative Efficiency (%)
Tucson	10	100
Phoenix	10	150
Los Angeles	20	200
San Francisco	20	125

Notes on the problem

- Emphatically, the design is a Randomized Complete Block (RCB) design. The subjects are divided into blocks, according to some unspecified criteria. Subsequently, the treatments are then assigned at random to the subjects in the blocks, once in each block. Thus, in RCB the number of replications equals the number of blocks.
- In a completely randomized design, subjects are assigned to treatment groups at random.
- The relative efficiency is measured relative to a completely randomized design, and thus it cannot be smaller than 100%.

Part (a)

Problem

How many more replications of a completely randomized design would be necessary at Los Angeles to obtain the same precision as the randomized complete block design for estimating the treatment means? Explain your answer.

Notes on part (a)

- Los Angeles has relative efficiency of 200%. This means that

$$RE = \frac{I_1}{I_2} = 2$$

  where I_1 is the Fisher information of the design actually used to the hypothetical completely randomized design I_2.
- Fisher information $I = I(\mu) = 1/\sigma^2$ for a normal distribution, when σ is known. We note that the potential dependence on μ does not occur.
- When σ is estimated using sample variance s^2, with f degrees of freedom then the Fisher information is:

$$I = I(\mu) = \frac{f+1}{f+3}\frac{1}{s^2}.$$

Notes on part (a) - continued

- The formula for the number of replications in the case of known σ (two-sided test) is:

$$r \geq 2(z_{\alpha/2} + z_{\beta})^2 \left(\frac{\sigma}{\delta}\right)^2$$

  and it indicates that $r \sim (\sigma/\delta)^2$.
- Similarly, for a *ratio measurement* in the sense of Steven's

$$r \sim \left(\frac{\%CV}{\%\delta}\right)^2 .$$

Solution to part (a)

- The number of replications of a completely randomized design would have to be 2 times larger than that of the RCB.

Part (b)

Problem

If you were asked the quation in part (a) relative to San Francisco, would you require more or fewer replications in San Francisco than for Los Angeles? Explain your answer.

Solution of part (b)

- The San Francisco design is less efficient than the Los Angeles design, and the relative efficiency is:

$$RE = \frac{2}{1.25} = 1.6 = 160\%.$$

Hence, San Francisco requires 1.6 times as many replications.

Part (c)

Problem

Suppose four replications are required with the randomized complete block design at Tucson to detect differences of $\delta = 20\%$ with a test at the .05 level of significance and a probability (power) of .90. How many replications would be required in Phoenix with the same criteria for a randomized complete block design? Explain your answer.

Notes on part (c)

- The number of replications is so small (4 replications) that we must be careful about the number of degrees of freedom.
- The number of replications in a completely randomized design, when the population distribution is normal and σ is known, is indicated by:

$$r \geq 2(z_{\alpha/2} + z_{\beta})^2(\%CV/\%\delta)^2$$

  is 6 when $\%CV = 10\%$ and $\%\delta = 20\%$. When the power is $1 - \beta = .8$, we do get $r = 4$ ($\alpha = .05$). This suggests a typo in the problem: the power should be .8. Alternatively, if we use power .9, to achieve the replication number $r = 4$ we would need to decrease $\%CV$ from 10% to about 8% (the actual power will be about .95 not .9. This information can be obtained with an on-line replication number calculator.

Notes on part (c) - continued

- When we compare two designs, the formula:

$$RE = \frac{I_1}{I_2} = \frac{(f_1 + 1)(f_2 + 3)}{(f_1 + 3)(f_2 + 1)} \frac{s_1^2}{s_2^2}$$

may be used.

Solution to part (c)

- The Phoenix design is 1.5 times more efficient and it has the same coefficient of variation. That means, it should require approximately 1.5 times fewer replications. Rounding the number $4/1.5 = 2.67$ up we would get 3 replications.

Part (d)

Problem

Would you require more or fewer replications in Los Angeles than in Tucson with the same criteria for a randomized complete block design? Explain your answer.

Solution to part (d)

- Tucson and Los Angeles differ in both coefficient of variation and efficiency. Two-fold difference in the coefficient of variation translates into a 4 fold difference in the number of replication required. However, higher efficiency of the LA design cuts this number by a factor of 2, so the LA design requires in the end $4/2 = 2$ more replications than the Tucson design.
- The assumption we tacitly made is that $\%\delta$ is the same.