

is only an approximation to the true value, adopted from classical test theory. Taking the approximation

$$\tilde{\gamma}_i = \frac{g_i \phi_r + \gamma_i \phi_i}{\phi_r + \phi_i},$$

we find that

$$(3.7) \quad \varepsilon(\tilde{\gamma}_i - \gamma_i)^2 = \frac{\phi_r^2 \phi_i + \phi_i^2 (\gamma_i - \gamma_r)^2}{(\phi_r + \phi_i)^2}.$$

If $(\gamma_i - \gamma_r)^2 = \phi_r$, this expression becomes

$$(3.8) \quad \frac{\phi_r \phi_i (\phi_r + \phi_i)}{(\phi_r + \phi_i)^2} = \phi_i R_i,$$

the value discussed above.

The practical implication of this analysis is that much care should be exercised in forming homogeneous clusters of groups, thus giving a small value of ϕ_r . One can, of course, accept the Empirical Bayes position that this kind of work can be done for any cluster of groups, but the analysis presented above suggests that the casual clustering of groups will not be profitable. Also, as we shall see, the haphazard clustering of groups can raise much more serious problems.

Some idea of the average benefit to be gained from the Bayesian Model II analysis can be seen by looking at the within group sample size that would be required by the classical Model I method to obtain the same error standard deviation obtained by the Bayesian method. For large m the Bayesian Model II standard error is approximately $(4n_i + 2)^{-1/2} R_i^{1/2}$. The classical Model I procedure with a sample size of N_i would have the standard error $(4N_i + 2)^{-1/2}$. Equating these two expressions and solving for N_i we have

$$(3.9) \quad N_i = n_i R_i^{-1} + \frac{(1 - R_i)}{2R_i}.$$

For $R_i > .30$ and $n_i \geq 5$ say, the second term can be ignored for most purposes, and we can take $N_i = n_i R_i^{-1}$. For $R_i = \frac{1}{2}$, the Bayesian method will, at any specified level of precision, require fewer than half the number of observations required by the classical method. It is thus clear that the Bayesian method has much to offer. In Table I we tabulate N_i for selected values of n_i and R_i . Because of the substitution made in going from (3.7) to (3.8), the gain is only as much as that given in the table when $|\gamma_i - \gamma_r| < \sqrt{\phi_r}$. The discrepancy is serious (for γ_i more than two standard deviations from γ_r , say) when the reliability is small.

While the preceding analysis is informative, and we hope compelling, for practical work it will be more useful to study the saving as a function of a parameter that does not depend on the sample size. In the above analysis

TABLE I

Classical Sample Size N_i Required to Attain Any Specified Degree of Precision Attained by the Bayesian Method with Sample Size n_i when the Reliability is R_i

N_i - rounded to nearest integer with

R_i .5 rounded down and entry primed

.90	6	11	17	22	28	33
.80	6	13	19	25	31	38
.70	7	14'	22	29	35	43
.60	9	17	25	34	42	50
.50	10'	20'	30'	40'	50'	60'
.40	13	26	38	51	63	76
.30	18	34'	51	68	84	101
.20	27	52	77	102	127	152
.10	54'	104'	154'	204'	254'	304'
	$n_i = 5$	10	15	20	25	30

R_i does depend on n_i so that we are effectively entering a table both of whose arguments depend on n_i .

From (3.9) we can write

$$N_i = \frac{2n_i + (1 - R_i)}{2R_i} = \frac{2n_i + \{(4n_i + 2)^{-1}/[\phi_r + (4n_i + 2)^{-1}]\}}{2\phi_r[\phi_r + (4n_i + 2)^{-1}]^{-1}}$$

which after some algebra gives us

$$(3.10) \quad n_i = N_i - (4\phi_r)^{-1}$$

Assuming a very large value for m , equation (3.10) gives us the sample size n_i required in the i -th group by the Model II method to obtain the same error variance as obtained by the classical Model I method with sample size N_i .

Interestingly, if $\phi_r < (4N_i)^{-1}$, the Model II method with zero observations has a smaller error than the Model I method with sample size N_i . This simply means that the overall mean across groups is a better estimator than the within group sample mean for very small sample sizes. (The formal argument here is largely heuristic since the quoted formulas are not valid for small m . However, for n_i and ϕ_r sufficiently small, the statement is obviously true.) For values of ϕ_r equal to .01, .02, .03, and .04, the samples required for the within group mean to be a better estimator than the overall mean are 25, 13, 8, and 6.

We note that each γ_i lies in the interval $0 < \gamma_i < \pi/2$. Centering a