The variance of a linear combination of independent estimators using estimated weights

By DONALD B. RUBIN

Educational Testing Service, Princeton

AND SANFORD WEISBERG

Department of Applied Statistics, University of Minnesota, Twin Cities

SUMMARY

Consider k independent unbiased estimators of a common parameter τ. Let t^* be the linear combination of the k estimators with minimum variance, and let $\hat{t}$ be any linear combination whose weights sum to unity and are independent of the k estimators. Then $\hat{t}$ is unbiased for τ and the variance of $\hat{t}$ depends only on the variance of t^* and the mean squared error of the weights as estimates of the optimum weights.

Some key words: Combining information; Linear combination of estimators; Weighted mean.

In a variety of statistical problems, several independent unbiased estimators of the same parameter are available, and it is desirable to combine them in order to reduce variability. This situation will occur, for example, in utilizing inter- and intra-block information (Brown & Cohen, 1974; Roy & Shah, 1962), in combining determinations made by different technicians or apparatus (Cochran, 1937; Meier, 1953; Halperin, 1961) or in the combination of estimators in missing data problems. The minimum variance unbiased estimator of the common parameter can be obtained by weighting each estimator inversely proportional to its variance; if the variances are unknown, the appropriate weights must be estimated.

Suppose that $t_1, \ldots, t_k$ are independent and unbiased estimators of τ with variances $V_1, \ldots, V_k$, with $V_i > 0$ for all i. We let $W = \Sigma V_i^{-1}$ and $\alpha_i = (WV_i)^{-1}$, so that $t^* = \Sigma\alpha_i t_i$ is the minimum variance unbiased linear combination estimator of τ with variance $\text{var}(t^*) = \Sigma\alpha_i^2 V_i = 1/W = \alpha_i V_i$, for every i. Let $\hat{\alpha}_1, \ldots, \hat{\alpha}_k$ be a set of estimators independent of $t_1, \ldots, t_k$, with $\Sigma\hat{\alpha}_i = 1$. Consider $\hat{t} = \Sigma\hat{\alpha}_i t_i$. Then, letting the subscript $\hat{\alpha}$ refer to expectations over the distribution of the $\hat{\alpha}_i$, we have

$$E(\hat{t}) = E_{\hat{\alpha}}\{E(\hat{t}|\hat{\alpha}_1, \ldots, \hat{\alpha}_k)\} = E_{\hat{\alpha}}(\Sigma\hat{\alpha}_i\tau) = \tau,$$

since $\Sigma\hat{\alpha}_i = 1$, and hence $\hat{t}$ is unbiased for τ. Also,

$$\text{var}(\hat{t}) = E_{\hat{\alpha}}\{\text{var}(\hat{t}|\hat{\alpha}_1, \ldots, \hat{\alpha}_k)\} + \text{var}_{\hat{\alpha}}\{E(\hat{t}|\hat{\alpha}_1, \ldots, \hat{\alpha}_k)\} = E_{\hat{\alpha}}(\Sigma\hat{\alpha}_i^2 V_i).$$

Adding and subtracting $\text{var}(t^*) = \Sigma\alpha_i^2 V_i$ and then substituting $\text{var}(t^*)/\alpha_i$ for V_i, we find

$$\begin{aligned}\text{var}(\hat{t}) &= \text{var}(t^*) + E_{\hat{\alpha}}(\Sigma V_i\hat{\alpha}_i^2) - \Sigma\alpha_i^2 V_i \\ &= \text{var}(t^*)\left[1 + \Sigma E_{\hat{\alpha}}\left\{\frac{1}{\alpha_i}(\hat{\alpha}_i^2 - \alpha_i^2)\right\}\right] \\ &= \text{var}(t^*)\left[1 + \Sigma\alpha_i E_{\hat{\alpha}}\left\{\left(\frac{\hat{\alpha}_i - \alpha_i}{\alpha_i}\right)^2\right\}\right], \qquad (1)\end{aligned}$$

since $\Sigma(\hat{\alpha}_i^2 - \alpha_i^2)/\alpha_i = \Sigma\alpha_i\{(\hat{\alpha}_i - \alpha_i)/\alpha_i\}^2$. Note that $\text{var}(\hat{t})$ is always at least as large as $\text{var}(t^*)$. Also, $\text{var}(\hat{t})$ depends on the estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_k$ only through their individual mean squared errors, even though these estimates are dependent. Thus, from (1) if the $\hat{\alpha}_i$ are consistent estimators of the α_i, then $\hat{t}$ has the same asymptotic mean and variance as t^*.

In the important case of $k = 2$, letting $\alpha = \alpha_1 = V_2/(V_1 + V_2)$ and $\alpha_2 = 1 - \alpha$, and $\hat{\alpha} = \hat{\alpha}_1, \hat{\alpha}_2 = 1 - \hat{\alpha}$, then $\text{var}(t^*) = V_1 V_2/(V_1 + V_2)$ and

$$\text{var}(\hat{t}) = \text{var}(t^*)\left[1 + \frac{\alpha}{1-\alpha} E_{\hat{\alpha}}\left\{\left(\frac{\hat{\alpha} - \alpha}{\alpha}\right)^2\right\}\right]. \qquad (2)$$

If t_1 and t_2 are weighted inversely proportional to independent chi-squared estimates of variance, this result may be applied to give the result of Meier (1953). For combining inter- and intra-block information, where estimated weights are truncated to zero when estimates of variance components are negative, the lemma given by Roy & Shah (1962) is a special case of the result given here. The asymptotic variance and efficiency of an estimate for a simple missing data problem given by Lin & Stivers (1974) also follows immediately.

The work was partly supported by the U.S. National Institute of Education.

REFERENCES

BROWN, L. D. & COHEN, A. (1974). Point and confidence estimation of a common mean and recovery of inter-block information. *Ann. Statist.* **2**, 963–76.
COCHRAN, W. G. (1937). Problems arising in the analysis of series of similar experiments. *J. R. Statist. Soc., Suppl.* **4**, 102–18.
HALPERIN, M. (1961). Almost linearly-optimum combination of unbiased estimates. *J. Am. Statist. Assoc.* **56**, 36–44.
LIN, P.-E. & STIVERS, L. (1974). On difference of means with incomplete data. *Biometrika* **61**, 325–34.
MEIER, P. (1953). Variance of a weighted mean. *Biometrics* **9**, 59–73.
ROY, J. & SHAH, K. R. (1962). Recovery of interblock information. *Sankhyā* A **24**, 269–80.

[*Received November* 1974. *Revised January* 1975]

A note on empirical Bayes inference in a finite Poisson process

BY D. R. COX

Department of Mathematics, Imperial College, London

SUMMARY

A Poisson process of events is considered which terminates after a random time. Termination is not directly observable but has to be inferred from the nonoccurrence of events. The analysis of data from such processes is considered with particular emphasis on empirical Bayes methods, it being supposed that data on many similar processes are available. An application to inventory control is outlined.

Some key words: Empirical Bayes; Inventory control; Point process; Poisson process; Prediction.

1. INTRODUCTION

Nearly all work on the statistical analysis of point processes either assumes stationarity or deals with smooth trends in the rate of occurrence. Here we consider the following very special finite point process. Initially events occur in a Poisson process of rate λ. After a time that is exponentially distributed with with parameter ρ, independently of the Poisson process, and which is not directly observable, the process terminates, i.e. the rate of the Poisson process drops to zero.

Of course the initial point process, the termination mechanism and the relation between the two can be generalized in numerous ways.

Vit (1974) has examined the equivalence between alternative forms of termination mechanism and in his unpublished thesis has also studied significance tests based on observing for a finite time either a single realization or several independent realizations of the above process. For instance, for a single realization the possibility that the process has terminated is assessed by considering the null hypothesis that the data come from a simple Poisson process. Evidence against the null hypothesis, the gap from the last event to the end of the period of observation being excessive, is interpreted as evidence of termination. Note that the object of the inference is not so much an unknown parameter as the occurrence of a not directly observable random event, so that the problem is essentially one of prediction.

In the present paper an empirical Bayes approach is outlined. This leads to a stronger conclusion than the significance test in that we estimate the posterior probability that a particular series has terminated, given observations on it.

One possible application is to inventory control. Suppose that for each of a very large number of products a record is available of the times at which items have been requested. Termination of the process for a particular product occurs when that product has become obsolete; it is assumed useful to be