in the trial-monitoring setting, the usual maximum likelihood asymptotics hold. For projecting power, however, the previous study will usually have a smaller or similar size to that of the proposed study. Thus the distributional properties documented in Tables 1 and 2 are not unrealistic. The general message is to beware maximum likelihood estimation when the "parameter" depends on the sample size.

When maximum likelihood estimators perform poorly, sometimes a Bayesian analysis can offer some guidance. A pure Bayesian approach would not be concerned with hypothesis testing and power and therefore would be irrelevant. A partial Bayesian approach uses data from the previous study to update a prior distribution for the treatment effects. The derived posterior distribution can then be used to estimate the power of the proposed trial (Herson 1979; Spiegelhalter and Freedman 1986). A special case of this approach assumes a noninformative prior on the treatment means; it was used in the present setting by Brown, Herson, Atkinson, and Rozell (1987) and also in the trial monitoring setting (Choi, Smith, and Becker 1985; Hilsenbeck 1988). For the present application, a noninformative prior leads to the posterior distribution of the treatment means being multivariate normal (DeGroot 1970, pp. 196–198). This implies that the posterior distribution of $n\Delta^2$ is a noncentral chi square with noncentrality parameter $n\hat{\Delta}^2$. For reasonably large values of $n\hat{\Delta}^2$, the posterior median of $n\Delta^2$ is therefore given by $n\hat{\Delta}^2 + K - 1$. This leads to $\theta(\hat{\Delta}^2 + (K - 1)/n)$ being approximately the posterior median estimate of the power, suggesting that Bayesian estimates of power share the same problem of large variability that the maximum likelihood estimates have (Korn 1990). In fact, as noted by a referee, we would expect any point estimator of $n\Delta^2$ to lead to an estimated power with large variability. I conclude with a second message for students: In some applications it is not reasonable to give only a point estimate without some measure of variability, either via a confidence interval or a Bayesian posterior distribution.

[*Received April 1989. Revised December 1989.*]

REFERENCES

Brown, B. W., Herson, J., Atkinson, E. N., and Rozell, M. E. (1987), "Projection From Previous Studies: A Bayesian and Frequentist Compromise," *Controlled Clinical Trials*, 8, 29–44.

Choi, S. C., Smith, P. J., and Becker, D. P. (1985), "Early Decision in Clinical Trials When the Treatment Differences Are Small," *Controlled Clinical Trials*, 6, 280–288.

Cox, D. R., and Hinkley, D. V. (1974), *Theoretical Statistics*, London: Chapman & Hall.

DeGroot, M. H. (1970), *Optimal Statistical Decisions*, New York: McGraw-Hill.

Hardison, C. D., Quade, D., and Langston, R. D. (1986), "Nine Functions for Probability Distributions," in *SUGI Supplemental Library User's Guide, Version 5 Edition*, Cary, NC:SAS Institute Inc., pp. 385–393.

Herson, J. (1970), "Predictive Probability Early Termination Plans For Phase II Clinical Trials," *Biometrics*, 35, 775–783.

Hilsenbeck, S. G. (1988), "Early Termination of a Phase II Clinical Trial," *Controlled Clinical Trials*, 9, 177–188.

Korn, E. L. (1990), "Projection From Previous Studies: A Caution," *Controlled Clinical Trials*, 11, 67–69.

Lan, K. K. G., and Wittes, J. (1988), "The B-Value: A Tool for Monitoring Data," *Biometrics*, 44, 579–585.

McCune, E. D., and Gray, H. L. (1982), "Cornish–Fisher and Edgeworth Expansions," in *Encyclopedia of Statistical Sciences, Volume 2*, eds. S. Kotz and N. L. Johnson, New York: John Wiley, pp. 188–193.

Spiegelhalter, D. J., and Freedman, L. S. (1986), "A Predictive Approach to Selecting the Size of a Clinical Trial, Based on Subjective Clinical Opinion," *Statistics in Medicine*, 5, 1–13.

Stevens, J. (1986), *Applied Multivariate Statistics for the Social Sciences*, Hillsdale, NJ: Lawrence Erlbaum, pp. 140–143.

The Mean Is Within One Standard Deviation of Any Median

COLM ART O'CINNEIDE*

The title of this note refers to a well-known relation between the three most familiar parameters of a distribution. This relation appears to be due to Hotelling and Solomons (1932). My goal here is to present a neat proof of it. Although my methods are simple enough to lead one to expect that this is not the first time the result has been proved this way, I have not found this approach in the literature.

Let F denote the distribution function of a population. Let μ and σ^2 denote its mean and variance, assumed finite, and let m denote any median of F. Assume that $m \leq \mu$; the general case follows from this one by reversing signs in the population. Begin by splitting the population into two equal parts, each part concentrated on one side of m; that is, decompose F into two distributions, one assigning all of its mass to the interval $(-\infty, m]$ and the other assigning all of its mass to $[m, \infty)$ in a unique manner. To be precise, let F_1 and F_2 be distribution functions such that $F_1(m) = 1$, and $F_2(m^-) = 0$ ($^-$ refers to the left limit), for which

$$F(x) = \frac{1}{2} F_1(x) + \frac{1}{2} F_2(x). \qquad (1)$$

F_1 and F_2 are uniquely determined by these conditions; in fact, $F_1(x) = 2F(x)$ and $F_2(x) = 0$, for $x < m$, and $F_1(x) = 1$ and $F_2(x) = 2F(x) - 1$ for $x \geq m$. F, F_1, and F_2 all give the same point mass to m. Now let μ_i and σ_i^2 denote the mean and variance of F_i for $i = 1, 2$. Then, since F_1

*Colm Art O'Cinneide is Associate Professor of Statistics, Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR 72701. This note was written while the author was visiting the Department of Quantitative Business Analysis, Louisiana State University. The author thanks Jim Willis for providing the reference and Ron Dattero for comments on the first draft.

is concentrated on $(-\infty, m]$, we have

$$m \geq \mu_1. \tag{2}$$

Moreover, by (1),

$$\mu = \frac{1}{2}(\mu_1 + \mu_2). \tag{3}$$

Subtracting (2) and (3) then gives

$$(\mu_2 - \mu_1)/2 \geq \mu - m. \tag{4}$$

Now consider how the variances of the three distributions are related; by analysis of variance, we have

$$\begin{aligned} \sigma^2 &= \frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2 + \left(\frac{\mu_1 - \mu_2}{2}\right)^2 \\ &\geq \left(\frac{\mu_1 - \mu_2}{2}\right)^2 \\ &\geq (\mu - m)^2, \end{aligned} \tag{5}$$

the last inequality coming from (4). From this it follows that

$$|\mu - m| \leq \sigma, \tag{6}$$

as required.

Consider the distribution giving weight ½ to the two points $\mu \pm \sigma$, where μ is arbitrary and σ is nonnegative. If $\sigma = 0$, we interpret this as the point mass at μ. The mean and standard deviation of this distribution are indeed μ and σ, and the points $\mu \pm \sigma$ are both medians. Thus we see that equality holds in (6) for this distribution if we choose $m = \mu \pm \sigma$. Conversely, since equality in (6) implies equality in (5) and this in turn forces $\sigma_1^2 = \sigma_2^2 = 0$, it is not hard to deduce that the two-point (or one-point, if $\sigma = 0$) distribution seen previously is the only case in which equality holds in (6). Moreover, any m satisfying (6) is a median of this distribution. It follows that a triple (μ, m, σ) of real numbers represents the mean, median, and standard deviation of some distribution iff (6) holds.

A referee has pointed out that a similar analysis is possible for the pth percentile x_p, $0 < p < 1$, leading to the generalization

$$|x_p - \mu| \leq \sigma \max\left(\sqrt{\frac{1-p}{p}}, \sqrt{\frac{p}{1-p}}\right).$$

The proof relies on splitting the population at the point x_p, just as we did in (1) for $p = ½$. The details are omitted.

[*Received October 1989. Revised January 1990.*]

REFERENCE

Hotelling, H., and Solomons, L. M. (1932), "The Limits of a Measure of Skewness," *The Annals of Mathematical Statistics*, 3, 141–142.

A Note on the Estimation of Binomial Probabilities

RICHARD J. PULSKAMP*

I am concerned with the admissibility under quadratic loss of certain estimators of binomial probabilities. The minimum variance unbiased estimator is shown to be admissible for $\Pr(X = 0)$ and $\Pr(X = n)$, but it is inadmissible for $\Pr(X = k)$, where $0 < k < n$. An example is given of an admissible maximum likelihood estimator (MLE). It is conjectured that the MLE is always admissible.

KEY WORDS: Admissibility; Generalized Bayes estimators; Quadratic loss.

1. INTRODUCTION

Suppose that X is a random variable having a binomial distribution with parameters n and θ, $n > 1$ known and $0 < \theta < 1$. We wish to estimate the binomial probabilities

$$\begin{aligned} \Pr(X = k) &= p_{n,k}(\theta) \\ &= \binom{n}{k}\theta^k(1 - \theta)^{n-k}, \qquad k = 0, 1, \ldots, n. \end{aligned}$$

Under usual circumstances, a Bayes estimator is admissible. These estimators were considered by both Johnson (1971) and Ol'man (1984). The most frequently used estimators—the minimum variance unbiased estimator (MVUE) and the maximum likelihood estimator (MLE)—are generally not Bayes. In fact, the MVUE never can be. This fact has motivated the development of techniques that enable one to prove admissible a given estimator that is known not to be Bayes but that has a representation as a generalized Bayes estimator. One very fruitful procedure, having its genesis in Karlin (1958) and restricted to quadratic loss, requires that an estimator be expressed formally as the mean of the posterior distribution and that certain boundary conditions be satisfied. The study of estimators that do not have this representation, however, seems to have been given less emphasis. The MVUE and MLE of $p_{n,k}(\theta)$ are of precisely this type.

*Richard J. Pulskamp is Assistant Professor, Department of Mathematics, Xavier University, Cincinnati, OH 45207. The author is grateful for the changes suggested by a referee.