CR 113862

APPROXIMATIONS TO THE BAYES ESTIMATE FOR A QUANTAL ASSAY

WITH SIMPLE EXPONENTIAL TOLERANCE DISTRIBUTION

By

Andrés Petrásovits
Canada Department of Agriculture

and

Richard G. Cornell
Florida State University

September 1, 1970

CASE FILE COPY

Department of Statistics
Florida State University
Tallahassee, Florida
32306

Approximations to the Bayes Estimate for a Quantal Assay

with Simple Exponential Tolerance Distribution

By

Andrés Petrásovits
Canada Department of Agriculture

and

Richard G. Cornell
Florida State University

SUMMARY

Approximations to the computationally cumbersome Bayes estimate of the parameter λ in the exponential tolerance distribution $F(d) = 1 - \exp(-\lambda d)$, $d \geq 0$, are developed for dichotomous response quantal assays. A gamma prior distribution and a quadratic loss function for λ and a binomial distribution of positives at each dose d for fixed λ are assumed. The first approximation is for a single-dose assay, is easily computed and has average risk only slightly larger than the Bayes risk for small samples. The selection of d and n are considered when this approximation is to be used as an estimate of λ. The second approximation is a modification of the first, is somewhat more difficult to compute than the first and retains the relatively small average risk of the first for small samples but, unlike the first, is a consistent estimate of λ. The third approximation is an extension of the first to an assay with several doses.

1. INTRODUCTION

Approximations to the Bayes estimate are presented in this paper for a dichotomous response quantal assay with exponential tolerance distribution

$$F(d) = 1 - \exp(-\lambda d), \quad 0 < \lambda, \ d < \infty, \tag{1.1}$$

where d is the dose and λ the parameter to be estimated. Attention is focused on single-dose experiments for the most part, although an extension to several doses is suggested in Section 6. A single-dose experiment is reasonable when prior information is available (see Mantel (1967)). A table containing optimum doses is provided. At the dose utilized, n subjects are exposed with r of them responding positively. It is assumed that these n subjects react independently so that, for given λ, r is binomially distributed with expectation $nF(d)$, where $F(d)$ is given by (1.1). Hence the likelihood of r is

$$l(n|\lambda) = \binom{n}{r} e^{-\lambda d(n-r)} (1-e^{-\lambda d})^r . \tag{1.2}$$

The loss function for λ is assumed to be quadratic. The prior distribution of λ is taken to be gamma with probability density function

$$g(\lambda) = \Gamma^{-1} (\alpha+1)\beta^{\alpha+1}e^{-\lambda\beta}\lambda^{\alpha}, \quad \lambda > 0, \; \alpha > -1, \; \beta > 0 \tag{1.3}$$

and with mean and variance

$$E(\lambda) = (\alpha+1)/\beta;$$

$$V(\lambda) = E(\lambda)/\beta,$$

respectively. This distribution is in the family of extended natural conjugate priors, it has the desirable property that it does not depend on the design parameter d, and it is the natural conjugate prior for the corresponding counting, as opposed to quantal, response experiment based upon the same tolerance distribution. The maximum likelihood estimate of λ for a single-dose assay is developed by Cochran (1950) and called the "most probable number". Cochran also considers. the selection of dose level.

In Section 2 formulas for a Bayes analysis are presented for a single-dose assay. The Bayes estimate $\hat{\lambda}$ is difficult to compute, particularly if a digital computer is not available. Bounds for the Bayes estimate are developed in Section 3 which lead to an easily calculated approximation to the Bayes estimate given in Section 4. A table of average risks for this estimate are displayed and it is shown empirically that it nearly achieves the minimum (Bayes) risk for small sample sizes. Average risks along with optimum dosages given in the same table can be used for designing an assay, that is, for selecting n and d. In Section 5 a slightly more cumbersome approximation to the Bayes estimate for a single-dose assay is presented which retains the desirable small sample

properties of the first approximation but which, unlike the first, is a consistent estimator of λ. Then in Section 6 the first approximation is extended for use with multiple-dose assays.

2. BAYES RESULTS

Bayes results are presented here in order to display formulas used in calculations tabled in Section 4 and to show the need for a simple approximation to the Bayes estimate. Derivations for any number of doses are given by Petrásovits (1970).

The Bayes estimate of λ for a single-dose assay given the likelihood (1.2) and the prior density (1.3) is

$$\hat{\lambda} = [(\alpha+1)/d] \, S(\alpha+2)/S(\alpha+1), \tag{2.1}$$

where

$$S(\gamma) = \sum_{j=0}^{r} (-1)^j \binom{r}{j} (n-r+j+\beta/d)^{-\gamma}. \tag{2.2}$$

The posterior probability density function for λ is

$$h(\lambda|r) = d^{\alpha+1} \, \Gamma^{-1}(\alpha+1) e^{-\lambda d(n-r+\beta/d)} (1-e^{-\lambda d})^r e^{-\lambda\beta} \lambda^{\alpha}/S(\alpha+1), \quad \lambda > 0,$$

the ω-th posterior moment of λ about the origin is

$$\mu'_{\omega}(r) = d^{-\omega} \prod_{j=1}^{\omega} (\alpha+j) \, S(\alpha+\omega+1)/S(\alpha+1),$$

the marginal distribution of r is

$$m(r) = d^{-(\alpha+1)} \beta^{\alpha+1} \binom{n}{r} S(\alpha+1)$$

and the Bayes risk is

$$B(\hat{d}) = (\alpha+2)(\alpha+1)\beta^{-2} - d^{-1}(\alpha+1)^2\beta^{\alpha+1} \sum_{r=0}^{n} \binom{n}{r}S^2(\alpha+2)/S(\alpha+1).$$

Optimum dosages d_0 which minimize the Bayes risk are given in Table 1 in Section 4 for prior mean $\mu = 1$ and for prior coefficients of variation ranging from 0.1 to 2.0 and for sample sizes $n = 1$, 5 and 10. Corresponding Bayes risks evaluated at d_0, $B(d_0)$, are also given in that table and can be used to select n if an exact Bayes analysis is to be computed. Values of d_0 and $B(d_0)$ for prior mean $\mu \neq 1$ can be obtained by multiplying corresponding entries in Table 1 by μ^{-1} and μ^2, respectively.

3. BOUNDS FOR THE BAYES ESTIMATE

The Bayes estimate $\hat{\lambda}$ for a single-dose assay has been found to lie in the interval

$$\frac{r+\alpha+1}{dn+\beta} < \hat{\lambda} < \frac{r+\alpha+1}{d(n-r)+\beta} . \tag{3.1}$$

A derivation of these bounds for integral values of α is given in this section. In extensive calculations (3.1) has also been found to hold for nonintegral values of α. This analytical and empirical evidence is adequate for our purpose, namely, to propose approximations to $\hat{\lambda}$ in the interval given by (3.1). These approximations are presented in Section 4 and 5.

In order to see the rationale leading to (3.1), first define

$$T(h) = \sum_{j=0}^{r} (j+n-r+\beta/d)^{-(h+1)}. \tag{3.2}$$

By induction, for integral values of γ, it can be shown that

$$S(\gamma) = (\gamma-1)^{-1} \sum_{h=0}^{\gamma-2} S(\gamma-1-h) \, T(h), \quad \gamma=2,\ldots . \tag{3.3}$$

Moreover, $T(h)$ can be expressed as

$$T(h) = (n-r+\beta/d)T(h+1)+U(h+1) \qquad (3.4)$$

where

$$U(h+1) = \sum_{j=1}^{r} (j+n-r+\beta/d)^{-(h+2)} < rT(h+1)$$

so that

$$T(h) < (n+\beta/d)T(h+1). \qquad (3.5)$$

From (2.1), (3.3) and (3.5),

$$\hat{\lambda} = \alpha(\alpha+1)S(\alpha+2)/[d \sum_{h=0}^{\alpha-1} S(\alpha-h)T(h)]$$

$$> \alpha(\alpha+1)S(\alpha+2)/[d(n+\beta/d) \sum_{h=0}^{\alpha-1} S(\alpha-h)T(h+1)]$$

$$> \alpha(\alpha+1)S(\alpha+2)/\{d(n+\beta/d)[(\alpha+1)S(\alpha+2) - S(\alpha+1)T(0)]\} \qquad (3.6)$$

since

$$\sum_{h=0}^{\alpha-1} S(\alpha-h)T(h+1) = \sum_{k=0}^{\alpha} S(\alpha+1-k)T(k) - S(\alpha+1)T(0)$$

$$= (\alpha+1)S(\alpha+2) - S(\alpha+1)T(0).$$

Using (2.1), (3.6) can be rearranged to give

$$\hat{\lambda} > [\alpha/(dn+\beta)] + T(0)/d. \qquad (3.7)$$

Next the numerator of $\hat{\lambda}$ in (2.1) may be rewritten using (3.3) to obtain

$$\hat{\lambda} = \sum_{h=0}^{\alpha} S(\alpha+1-h)T(h)/[dS(\alpha+1)] = R + T(0)/d \tag{3.8}$$

where

$$R = [\sum_{h=1}^{\infty} S(\alpha+1-h)T(h)]/[dS(\alpha+1)].$$

From (3.3) and (3.4),

$$R = \sum_{k=0}^{\alpha-1} S(\alpha-k)T(k+1)/[dS(\alpha+1)]$$

$$= (\alpha/d)\{(n-r+\beta/d) + [\sum_{h=0}^{\alpha-1} S(\alpha-h)U(h+1)]/[\sum_{h=0}^{\alpha-1} S(\alpha-h)T(h+1)]\}^{-1}$$

$$< \alpha/[d(n-r+\beta/d)]. \tag{3.9}$$

Substitution of (3.9) into (3.8) yields

$$\hat{\lambda} < \alpha/[d(n-r) + \beta] + T(0)/d. \tag{3.10}$$

Now by setting the index of summation in each term of (3.2) equal to its upper and lower limits, it can easily be seen that

$$(r+1)(n+\beta/d)^{-1} < T(0) < (r+1)(n-r+\beta/d)^{-1}.$$

This inequality, together with (3.7) and (3.10), yields (3.1).

4. SMALL SAMPLE APPROXIMATION TO THE BAYES ESTIMATE

Because of the form of the bounds (3.1) on the Bayes estimate $\hat{\lambda}$, consideration is given in this section to the family of approximations to $\hat{\lambda}$ given by

$$\hat{\lambda}_1(D) = (r+\alpha+1)[d(n-r) + Dr+\beta]^{-1}, \quad 0 \leq D \leq d. \tag{4.1}$$

Now $\hat{\lambda}_1(D)$ is a continuous and monotonically decreasing function of D in

[0,d] and equals the lower and upper bounds given by (3.1) for D = d and 0, respectively. Therefore, for some unique D^*, where $0 < D^* < d$, $\hat{\lambda}(D^*) = \hat{\lambda}$.

One reasonable procedure would be to select D to minimize the average risk of $\hat{\lambda}_1(D)$ given by

$$R(D) = E_{\lambda,r}[\hat{\lambda}_1(D) - \lambda]^2,$$

where $E_{\lambda,r}$ denotes the expectation with respect to the joint distribution of λ and r. It can be shown that

$$R(D) = \sum_{r=0}^{n} \binom{n}{r} \hat{\lambda}_1(D)S(\alpha+2) + (\alpha+2)(\alpha+1)/\beta^2$$

where $S(\gamma)$ is defined in (3.3). The function R(D) is continuous in D on the interval [0,d] and attains a minimum on [0,d] at, say, D^*. Numerical evidence indicates that D^* is unique and is an interior point of that interval.

Values of D^* and corresponding $R(D^*)$ values have been computed for sample sizes 1, 5 and 10 and for prior coefficients of variation ranging from 0.1 to 1.0 in steps of 0.1 as well as 1.5 and 2.0. In each of these calculations d has been taken equal to d_0, the dose which minimizes the Bayes risk.

Two important empirical relationships were found. In every instance the ratio of D^* to d_0 was between 0.3 and 0.4 and the ratio of the Bayes (minimum)risk at d_0 to $R(D^*)$ at d_0 was greater than or equal to 0.969. The latter ratios show that $\hat{\lambda}_1(D^*)$ is a good alternative to the Bayes estimate for small samples. It has the advantage of being easily calculated if appropriate tables of D^* are available. The former ratios indicate that the disadvantage of requiring tables of D^* may be overcome by setting D equal to a constant proportion

Table 1. Values of the optimum dose (d_0), the average risk (R) of $\hat{\lambda}_1$ at $d = d_0$ and the ratio (E) at d_0 of the Bayes risk to R by prior coefficient of variation (ν) and sample size (n).

	n = 1			n = 5			n = 10		
ν	d_0	R	E	d_0	R	E	d_0	R	E
0.1	1.5850	0.0099	1.0000	1.5844	0.0097	0.9998	1.5838	0.0094	0.9993
0.2	1.5593	0.0390	1.0000	1.5524	0.0356	0.9995	1.5451	0.0321	0.9981
0.3	1.5179	0.0854	1.0000	1.4910	0.0712	0.9992	1.4674	0.0590	0.9968
0.4	1.4626	0.1468	1.0000	1.4019	0.1111	0.9989	1.3577	0.0860	0.9946
0.5	1.3961	0.2210	1.0000	1.2944	0.1533	0.9982	1.2320	0.1125	0.9912
0.6	1.3215	0.3066	1.0000	1.1792	0.1974	0.9971	1.1021	0.1396	0.9873
0.7	1.2418	0.4026	1.0000	1.0651	0.2439	0.9957	0.9824	0.1682	0.9830
0.8	1.1602	0.5087	0.9999	0.9572	0.2937	0.9943	0.8722	0.1987	0.9790
0.9	1.0789	0.6244	0.9999	0.8583	0.3470	0.9928	0.7737	0.2318	0.9753
1.0	1.0000	0.7502	0.9998	0.7693	0.4044	0.9915	0.6871	0.2677	0.9721
1.5	0.6717	1.5352	0.9991	0.4510	0.7615	0.9871	0.3969	0.4937	0.9610
2.0	0.4586	2.5961	0.9985	0.2921	1.2481	0.9833	0.2494	0.8028	0.9557

5. LARGE SAMPLE APPROXIMATION TO THE BAYES ESTIMATE

The estimate $\hat{\lambda}_1$ proposed in Section 4 has been found to be an excellent apporximation to the Bayes estimate for sample sizes of 10 and less, which are the sample sizes often used in assays. However, $\hat{\lambda}_1$ is not a consistent estimate of λ. An alternative family of estimates which could be used to construct estimates which approximate $\hat{\lambda}$ well for small samples and would be consistent is given by

$$\hat{\lambda}_2(D) = d^{-1} \sum_{j=0}^{r} (j+n-r+\beta/d)^{-1} + \alpha[d(n-r)+Dr+\beta]^{-1}, \quad 0 < D < d. \qquad (5.1)$$

For small samples D could be set equal, for instance, to 0.35d as in Section 4. The consistency of $\hat{\lambda}_2(D)$, which holds for all D in $[0,d]$, is gained at the expense of making the approximation of $\hat{\lambda}$ computationally more difficult. Of some help though is the fact that the sum in the first term on the right of (4.1), defined as $T(0)$ by equation (3.2), can be expressed as the difference of two digamma functions. Thus alternative expressions for $\hat{\lambda}_2(D)$ are

$$\hat{\lambda}_2(D) = d^{-1}T(0) + \alpha[d(n-r)+Dr+\beta]^{-1}, \quad 0 < D < d, \qquad (5.2)$$

$$= d^{-1}[\Psi(1+n+d/\beta)-\Psi(n-r+d/\beta)] + \alpha[d(n-r)+Dr+\beta]^{-1}, \quad 0 < D < d, \quad (5.3)$$

where the digamma function $\Psi(x) = d[\ln\Gamma(x)]/dx$ is tabled, for instance, by Abramowitz and Stegun (1964) and $\Gamma(x)$ is the gamma function $\Gamma(x) = \int_0^\infty e^{-t}t^x dt$.

The rationale for suggesting $\hat{\lambda}_2(D)$ is the same as that for $\hat{\lambda}_1(D)$ in Section 4 except that $\hat{\lambda}_2(D)$ is based on the bounds for $\hat{\lambda}$ given by (3.7) and (3.10) instead of the wider interval derived from them and given by (3.1). This relationship between $\hat{\lambda}_1(D)$ and $\hat{\lambda}_2(D)$ shows that $\hat{\lambda}_2(D)$ would perform well for small samples for suitably selected D and is the reason for suggesting that D be taken equal to 0.35 as in

the construction of $\hat{\lambda}_1$ in Section 4. In the remainder of this section it will be shown that, conditionally on λ for any D in $[0,d]$ as $n \to \infty$, the estimate $\hat{\lambda}_2(D)$ converges in probability to λ, that is, that $\hat{\lambda}_2(D)$ is a consistent estimator of λ.

The sum $T(0)$ in equation (5.2) can be expressed using the Maclaurin-Euler summation formula given by Jordan (1947, pp. 261-2) to obtain

$$T(0) = \ln(1+n+\beta/d) - \Psi(n-r+\beta/d) + (n-r+\beta/d)^{-1} + R_n \qquad (5.4)$$

where

$$R(n) = -2^{-1}(1+n+\beta/d)^{-1} - (12)^{-1}(1+n+\beta/d)^{-2} + \xi(120)^{-1}(1+n+\beta/d)^{-4},$$

$$0 < \xi < 1. \qquad (5.5)$$

But from Cramér (1946, p. 130) it follows that

$$\Psi(x) = \ln(x+1) - 2^{-1}(x+1)^{-1} - \int_o^\infty P_1(z)(x+1+z)^{-2} dz \qquad (5.6)$$

where $P_1(z) = [z] - z + 1/2$ and $[z]$ denotes the greatest integer $\leq z$. Noting that $|P_1(z)| \leq 1/2$, (5.6) can be rewritten as

$$\Psi(x) = \ln(x+1) - (1+\eta)2^{-1}(x+1)^{-1}, |\eta| \leq 1. \qquad (5.7)$$

Thus from (5.2), (5.4) and (5.7),

$$\hat{\lambda}_2(D) = d^{-1}\{\ln[(1/n) + 1 + (\beta/dn)]/[(1/n) + 1 - (r/n) + (\beta/dn)]\} +$$

$$(nd)^{-1} \{[1-(r/n) + (\beta/dn)]^{-1} + (1+\eta)2^{-1}[(1/n) + 1-(r/n) +\beta/(dn)]\}$$

$$+d^{-1}R_n + \partial n^{-1}[d(1-r/n) + D(r/n) + \beta/n]^{-1}, \qquad (5.8)$$

where $|\eta| \leq 1$ and R_n is given by (5.5). Noting that, as $n \to \infty$, $1-r/n$ converges

in probability to $\exp(-\lambda d)$, it follows readily from (5.8) that $\hat{\lambda}_2(D)$ converges

in probability to λ for fixed λ and any D in $[0,d]$.

6. EXTENSION TO SEVERAL DOSES

In this section a family of approximations to the Bayes estimate is proposed

for a k-dose assay with dose levels d_i when r_i subjects respond positively and

independently in samples of size n_i, $i = 1,2,\ldots, k$. The model described in

Section 1 is still assumed to hold at each dose level. The Bayes estimate may be

approximated by

$$\hat{\lambda}_3(D) = (\sum_{i=1}^{k} r_i + \alpha + 1)[\sum_{i=1}^{k} d_i(n_i-r_i) + D \sum_{i=1}^{k} r_i+\beta]^{-1},$$

$$0 < D < \max d_i. \tag{6.1}$$

This is the estimate, with $D = 0.35d_0$, that would be obtained if an experimenter were

to perform a sequence of k simple-dose assays at doses close to a stable optimum

d_0 determined from prior information using the approximate Bayes estimate $\hat{\lambda}_1(0.35d_0)$,

given by (4.1), after each assay. By a stable optimum d_0 is meant an optimum dose

which does not change appreciably when accumulated experimental information is used

after each successive single-dose assay to modify the prior distribution upon which

the optimum for the next assay is based. Since $\hat{\lambda}_1$ was found in Section 4 to be

robust relative to the Bayes estimate for departures of d from d_0, another reasonable

approximation to $\hat{\lambda}$ even when d_0 is not entirely stable could be constructed by

substituting $0.35\bar{d}$ for D in (6.1), where $\bar{d}$ is the mean of the k doses used in the

assay.

REFERENCES

Abramowitz, M. and Stegun, I. A. (Editors) (1964). Handbook of Mathematical Functions. Washington: National Bureau of Standards.

Cochran, W. G. (1950). Estimation of bacterial densities by means of the most probable number. Biometrics 16, 105-16.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton: Princeton University Press.

Jordan, C. (1947). Calculus of Finite Differences. New York: Chelsea.

Mantel, N. (1967). Adaption of Kärber's method for estimating the exponential parameter from quantal data, and its relationship to birth, death and branching processes. Biometrics 23, 739-46.

Petrásovits, A. (1970). Approximations to Bayes Procedures for Quantal Assays with Simple Exponential Tolerance Distributions. Tallahassee: Florida State University dissertation.