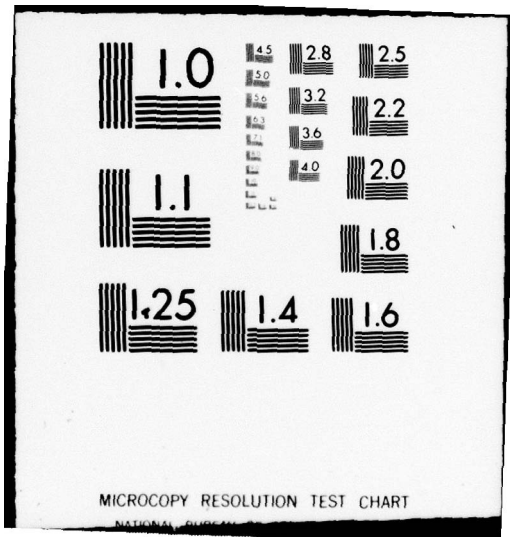
AD-A074 376 CASE WESTERN RESERVE UNIV CLEVELAND OHIO DEPT OF MA--ETC F/G 12/1
EMPIRICAL BAYES ESTIMATION OF CRITICAL DOSAGES HAVING SMALLEST --ETC(U)
AUG 79 O BIALIK, S ZACKS N00014-75-C-0529

UNCLASSIFIED TR-38 NL

| OF |
ADA
074376

END
DATE
FILMED
10-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MICROCOPY RESOLUTION TEST CHART

LEVEL II

EMPIRICAL BAYES ESTIMATION OF CRITICAL DOSAGES HAVING SMALLEST PREDICTIVE RISK

by

Ora Bialik and S. Zacks
Department of Research Medicine
University of Pennsylvania
and
Department of Mathematics and Statistics
Case Western Reserve University

TR-38

20 p.

Technical Report, No. 38
August 15, 1979

15 Aug 79

PREPARED UNDER CONTRACT
N 00014-75-C-0529 PROJECT NR 042-276
OFFICE OF NAVAL RESEARCH

Department of Mathematics and Statistics
CASE WESTERN RESERVE UNIVERSITY
Cleveland, Ohio

JAB

1. Introduction

Consider a monotone relationship of the expected response to the dosage in bioassay. A critical dosage is considered here to be the treatment level associated with an expected response threshold. The problem of determining critical dosages is known as the calibration or the inverse regression problem. Many studies can be found in the literature concerning the estimation of critical dosages for linear and non-linear regression models. (Krutchkoff (1967)). Methods of estimation of critical dosages in the single sample case provide either point estimates or confidence intervals (Fieller (1964)). Among the methods which involve multistage design approach we find the stochastic approximation methods and sequential search procedures (Eichhorn and Zacks (1973)). The present study considers the problem of determining critical dosages in a situation where a large number of regression lines are available from many related assays, but each regression line is based on a small number of observations. Thus, rather than estimating the critical dosage for each assay indivudually, the inter-block information provided by the various regression lines is intilized to increase the precision of the estimates. This is done within a Bayesian framework. More specifically, the model assumes that in each assay the dosage - response relationship is represented by a linear regression, with the same variance σ^2 around all regression lines and normal distribution of errors. The threshold, η, is the same for all assays. If α_k and β_k, $k = 1,\ldots,N$, are the true intercept and slope of the k-th regression line, the k-th critical dosage is $\xi_k = (\eta - \alpha_k)/\beta_k$, $\beta_k > 0$ and $\xi_k = \infty$ if $\beta_k \leq 0$. The

prior distribution of (α,β) is chosen to yield a negligible prior marginal probability for $\{\beta < 0\}$. The Bayesian model assumes that (α_k,β_k), $k = 1,\ldots,N$, are priorly independent and identically distributed vectors with a properly chosed prior bivariate normal distribution. Let (a_k,b_k) be the least-square estimators (LSE) of (α_k,β_k), based on the observations performed in the k-th assay. On the basis of this LSE a posterior bivariate normal distribution is determined for (α_k,β_k). This distribution yields a predictive normal distribution, given (a_k,b_k) for a response $Y(\xi)$ at a dosage ξ. (See Aitchison and Dunsmore 1975). The minimum predictive risk estimator of the critical dosage ξ_k is defined as the dosage, $\hat{\xi}_k$, which minimizes the predictive risk, i.e., $E\{(Y(\xi)-\eta)^2 | (a_k,b_k)\}$. The suggested estimator of ξ_k depends on the LSE (a_k,b_k) and on the parameters of the prior bivariate normal distribution of (α_k,β_k). When the number of assays, N, is large an empirical Bayes method can be employed for estimating the prior parameters. In the present study we develop the formulae for the empirical Bayes estimation of the critical dosages. The formula obtained resembles somewhat Stein-type estimators of a multivariate mean vector (Zacks (1971)). The procedure developed in the present study is applied for the determination of critical concentrations of benzo-soluble organic extracts in air samples taken in 1963 and 1964 from 53 and 54 different sites in the U.S.A. These organic extracts were tested for their toxicity in a series of photodynamic bioassays (Epstien et. al. (1965)). The toxicity of the benzo-soluble extracts from the pollutants is

a function of their chemical composition and concentration in the air. The chemical composition varies (at random) within a site and between the sites. The model developed in the present study was found suitable for the determination of critical air concentrations (dosages) for each site. These critical dosages can be compared with the actual concentrations of the organic extracts in the samples. Whenever an air sample contains organic extracts with concentration higher than the critical dosage evidence exists of undesirable toxicity of the air pollution. Similar applications can also be performed in other areas of the empirical sciences.

The present study consists of five sections. In section 2 we specify the statistical model and the Bayesian framework. The method of determining critical dosages by minimizing the predictive risk is provided in section 3. Section 4 is devoted to the empirical Bayes approach when the number of assays, N, is large. Finally, in section 5 we present the application to the analysis of the photodynamic bioassays, for the determination of critical concentration of benzo-soluble organic extracts in air samples.

2. The Statistical Model and the Baysian Framework.

Consider N sets of biological assays, having dose response relationship $Y(x_{ki}) = \alpha_k + \beta_k x_{ki} + \varepsilon_{ki}$, $k = 1,\ldots, N$, $i = 1,\ldots,n_k$, where ε_{ki} is a random variable normally distributed with expectation zero and varaince σ^2. The regressors $x_{ki}(i=1,..,n_k)$

are the log-dosage applied at the kth bioassay, and (α_k,β_k) are the linear regression parameters based on the observations (y_{ki},x_{ki}), $k = 1,\ldots,N$, $i = 1,\ldots,n_k$. The model assumes that the variance σ^2 is the same around all the N regression lines. Determine the common least square estimators (LSE) a_k and b_k of the linear regression parameters and the variance around the regression line s_k^2. Let s_p^2 denote the pooled estimator of this common variance, i.e.

$$(2.1) \qquad s_p^2 = \sum_{k=1}^{N} (n_k-2)s_k^2 / \sum_{k=1}^{N}(n_k-2).$$

The large number of assays considered in the present problem and the typically small error variance, σ^2 , provide estimates s^2 with samll standard error. Accordingly, we develop the following Bayesian model under the assumption that σ^2 is known and substitue s_p^2 for σ^2.

According to the theory of least-square estimation in normal models, $(a_k,b_k)'$ is a random vector having a conditional bivariate normal distribution, with an expectation vector $(\alpha_k,\beta_k)'$ and covariance matrix Σ_k, where

$$(2.2) \qquad \Sigma_k = \sigma^2 \begin{pmatrix} \frac{1}{n} + \frac{\bar{x}_k}{SDX_k} & \frac{-\bar{x}_k}{SDX_k} \\ \frac{-\bar{x}}{SDX_k} & \frac{1}{SDX_k} \end{pmatrix},$$

where

$\bar{x}_k$ designates the mean log-dosage at the k'th assay and $SDX_k = \sum_{k=1}^{N}(x_{ki} - \bar{x}_k)^2$. The Bayesian model assumes that each assay can be considered as a random sample from a larger population of assays. Accordingly, we assume that (α_k, β_k) follows a bivariate normal prior distribution with prior expectation $(\alpha_0, \beta_0)'$ and prior covariance matrix, T; i.e.,

$$(2.3) \qquad \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}, T \right) .$$

Given the estimates $(a_k, b_k)'$ and $\ddagger_k$, the posterior distribution of the regression parameters $(\alpha_k, \beta_k)'$ is also a bivariate normal distribution (Zacks 1971, Box and Tiao 1973) with expectation vector

$$(2.4) \qquad \begin{pmatrix} \alpha_k^* \\ \beta_k^* \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} + T(\ddagger_k + T)^{-1}\left[\begin{pmatrix} a_k \\ b_k \end{pmatrix} - \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \right]$$

and covariance matrix

$$(2.5) \qquad V_k = T - T'(\ddagger_k + T)^{-1}T.$$

3. *Bayesian Determination of Critical Dosages.*

The critical dosage, ξ_k, for the k-th bioassay, is defined as the value of x for which the expected response is η, i.e.,

$$(3.1) \qquad \xi_k = \frac{\eta - \alpha_k}{\beta_k} , \quad k = 1,\ldots,N .$$

We comment here that in practical applications of the model we assume that all $\beta_k > 0$. The Bayesian framework assumes a normal

marginal prior (posterior) distribution of β_k, which extends over negative values too. This problem is overcome in applications if the error variance σ^2 is relatively small and the dosages in the bio-assays are well designed. In such cases the posterior variance of β_k is often sufficiently small so that the posterior probability of negative β_k value is negligible.

In classical statistical analysis, ξ_k is frequently estimated by the least-squares inverse regression statistic

$$\tilde{\xi}_k = \frac{\eta - a_k}{b_k}, \qquad k = 1,\ldots,N. \tag{3.2}$$

Fieller's theorem (Fieller (1944)) is often applied to obtain classical confidence intervals for ξ_k. The application of Fieller's theorm in the Bayesian framework is not compatible with the definition (3.1) due to the interchange in the role of parameters and statistics. We consider therefore two types of Bayes point estimators of ξ_k. One is obtained by substituting in (3.1) the posterior estimates α_k^* and β_k^* of α_k and β_k, respectively. Accordingly, we obtain the (pseudo) Bayes estimator

$$\xi_k^* = \frac{\eta - \alpha_k^*}{\beta_k^*}, \qquad k = 1,\ldots,N. \tag{3.3}$$

Notice that ξ_k^* is not a Bayes estimator, since it does not minimize a prior (posterior) risk. We introduce a proper Bayes estimator of ξ_k by considering the value of x which minimizes the predictive risk $E\{(Y(x)-\eta)^2\}$. More specifically, we minimize the predictive expectation

(3.4) $$Q(x;\mathcal{F}_k) = E\{(Y(x)-\eta)^2 \mid \mathcal{F}_k\},$$

$k = 1,\ldots,N$. $E\{\cdot \mid \mathcal{F}_k\}$ designates the expectation with respect to the predictive distribution of $Y(x)$ in the k-th bioassay. In the present case the predictive distribution is the normal distribution with mean $\alpha_k^*+\beta_k^* x$ and variance $\sigma^2 + V_k(x) = \sigma^2 + V_{k11} + 2xV_{k12} + x^2V_{k22}$, where V_{kij}, $i,j = 1,2$ are the elements of the posterior covariance matrix V_k. We apply here the loss function $(Y(\hat{\xi}_k)-\eta)^2$ rather than $(\tilde{\xi}_k - \xi_k)^2$ since the posterior expectation of $(\eta-\alpha_k)/\beta_k$, given $\mathcal{F}_k$, does not exist. Thus

(3.5) $$Q(x;\mathcal{F}_k) = \sigma^2 + V_{k11} + 2xV_{k12} + x^2V_{k22} + (\alpha_k^*+\beta_k^* x-\eta)^2 .$$

The minimization of (3.4) with respect to x yields the estimator

(3.6) $$\hat{\xi}_k = \frac{\xi_k^* - V_{k12}/\beta_k^{*2}}{1 + V_{k22}/\beta_k^{*2}}, \quad k = 1,\ldots,N.$$

A $(1-\alpha)$ level predictive interval for $Y(\hat{\xi})$ is specified by the prediction limits

(3.7) $$P(\hat{\xi};\alpha) = \alpha^* + \beta^*\hat{\xi} \pm z_{1-\alpha/2}\sqrt{Q(\hat{\xi})} .$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ fractile of the standard normal distribution. It is not difficult to show that the three different estimators of ξ_k, namely $\tilde{\xi}_k$, ξ_k^* and $\hat{\xi}_k$ are consistant ones, as the number of observations n_k around the regression lines increase to infinity and SDX_k increase to infinity too. However, questions of consistency and asymptotic efficiency are irrelevant to our problem since generally we are concerned with cases of small number of obervations in each bioassay. For this reason we adopted the Bayesian approach to compenseate for the

lack of accuracy due to this deficiency. As we show in the next section, an empirical Bayes approach can utilize the information obtained from the large number of different assay to determine an adequate common prior distribution for the analysis of the individual assays.

4. An Empirical Bayes Approach for Large N.

Generally it is a difficult problem to determine that proper prior parameters for each regression line. However, if the analysis consists of a large number of regression lines from different assays, and if it is plausible to assume that the regression parameters (α_k,β_k), $k = 1,\ldots,N$, constitute a random sample from the sample bivariate (prior) normal distribution, one can estimate consistently the prior parameters. More specifically, under the assumption that the (true) regression parameters (α_k,β_k), $k = 1,\ldots,N$, are independent random vectors having the same bivariate normal distribution, with mean (α_0,β_0) and covariance matrix T then

$$(4.1) \qquad \begin{pmatrix} \hat{\alpha}_0 \\ \hat{\beta}_0 \end{pmatrix} = \frac{1}{N} \sum_{k=1}^{N} \begin{pmatrix} a_k \\ b_k \end{pmatrix}$$

is an unbiased, strongly consistent estimator of $(\alpha_0,\beta_0)'$ having a bivariate normal distribution with covariance matrix $\frac{1}{N}(T + \frac{1}{N}\sum_{k=1}^{N} \Sigma_k)$.

An unbiased and strongly consistant estimator of the total covariance matrix $T + \frac{1}{N}\sum_{k=1}^{N} \Sigma_k)$ is given by the sample covariance matrix

$$(4.2) \qquad C = \frac{1}{N-1} \begin{pmatrix} a' \\ b' \end{pmatrix} (I_N - \frac{1}{N} J_N)(a,b)$$

where $a' = (a_1,\ldots,a_N)$, $b' = (b_1,\ldots,b_N)$, I_N is the identity matrix of order N and J_N is an $N\times N$ matrix of 1's. Notice that the total covariance matrix $T + \frac{1}{N}\sum_{k=1}^{N} \Sigma_k$ is composed of the "within variance" component $\frac{1}{N}\sum_{k=1}^{N} \Sigma_k$ and the "between variance" component T. Thus as in the common components of variance model (see Graybill (1976)) and unbiased estimator of T is

$$(4.3) \qquad \hat{T} = C - \frac{1}{N} \sum_{k=1}^{N} \Sigma_k$$

We remark that if the design matrices of all the N assays are the same, i.e., $\Sigma_k = \Sigma$ for all $k = 1,\ldots,N$, the above formulae simplify. We further remark that (4.3) may be negative definite, if the "within variance" component is large and N is not sufficiently large. If this is the case, one has to apply a different approach, or use biased but consistent estimators of T. Finally, given the estimates $(\hat{\alpha}_0, \hat{\beta}_0)$ and the unbiased estimator $\hat{T}$ one can determine an estimate of V_k, namely

$$(4.4) \qquad \hat{V}_k = \hat{T} - \hat{T}\,(\Sigma_k + \hat{T}\,)^{-1}\,\hat{T}$$
$$= \begin{pmatrix} \hat{V}_{k11} & \hat{V}_{k12} \\ \hat{V}_{k12} & \hat{V}_{k22} \end{pmatrix}$$

and substitue its elements in (3.6) to obtain an empirical Bayes estimate of $\hat{\xi}_k$. This estimator is

$$(4.5) \qquad \hat{\xi}_k = \frac{\hat{\xi}_k^* - \hat{V}_{k12}/\hat{\beta}_k^2}{1 + \hat{V}_{k22}/\hat{\beta}_k^2} \quad ,$$

where $\hat{\alpha}_k$ and $\hat{\beta}_k$ are obtained from (2.4) by substituting $(\hat{\alpha}_0, \hat{\beta}_0)'$ for $(\alpha_0, \beta_0)'$ and $\hat{\xi}_k^* = (\eta - \hat{\alpha}_k)/\hat{\beta}_k$. Notice that the emperical Bayes estimator (4.5) is a shrinkage estimator whenever $-\hat{V}_{k12} \leq \hat{\xi}_k^* \hat{V}_{k22}$ This is the case, in particular when $\hat{V}_{k12} > 0$.

In the follwing section we provide a large scale application of the empirical Bayes approach described above.

5. An Application of the Model.

The model discussed in the previous sections was applied to the analysis of a large scale photodynamic bioassays, performed by Epstein et. al. (1965), for the purpose of evaluating the toxicity of organic extracts from atmospheric pollutatants. Air samples were collected in 1963 and 1964 from 53 and 54 different sites in the U.S., respectively. The benzo-soluble organic particles were chemically extracted from the air samples and the atmospheric concentrations [g/m^3 of air] were recorded. Proper solutions of the organic extracts (O.E.) were tested at three dilution levels $d = 10^{-4}, 10^{-5}, 10^{-6}$ [$g/m\ell$]. These preparations were applied in wells including 30 cells of Paramecia Caudatum. The measured response, called th LT90, was the time (in minutes) required to immobilise 90% of the cells under ultra-violet irradiation. The measurement of response was truncated at $t_0 = 90$ minutes. Response

values over 90 minutes are therefore unavailable, neither the propotion of living cells at the time of truncation. Four replicas were perfomed at each dose. Simultaneously, the LT90 was measured on a standard synthetic benzo-a-pyrene (BaP). 41 complete assays of the 1963 data and 54 of the 1964 data were available for analysis.

For the statistical analysis define $x = -\log_{10} d - 5$. The model assumes that $\ell n(LT90)$ is normally distributed with mean $\alpha + \beta x$ and variance σ^2. This model links the analysis described later to the theory developed in the previous sections. Bialik (1978) verified that the ℓn LT90 versus log-dose regression lines of the standard preparations, correspond to each year of test data, were not significaltly different. Accordingly, the analysis presented here does not have to adjust the regression line of each site for varying experimental conditions. In Table 1 we present the basic response statistics, the regression statistic and the expected LT90 corresponding to the actual concentation in the air for sites of the 1963 samples. The regression parameters (α, β) of different sites are not expected to be the same due to the different chemical composition of the O.E.. Since the toxicity of the organic pollutants is a combination of their chemical composition and atmospheric concentration, Bialik (1978) introduced a measure of toxicity, AIRLT90, which takes into account both factors. An equivalent air-dosage, AD, is defined as the atmospheric concentration of the O.E. in a given site solved in 1 mℓ of preparation. Let

(5.1) $$XAIR = -\log_{10}(AD) - 5.$$

Then the corresponding predicted LT90 is given by

(5.2) $$AIRLT90 = \exp(a + b.XAIR + \hat{\sigma}^2/2).$$

The XAIR and AIRLT90 of the various sites are given in Table 1. Notice that all the XAIR values in Table 1 fall in the experimental domain $(-1 \leq x \leq 1)$. Accordingly, the AIRLT90 values are not based on extrapolation. We also remark that intensive photodynamic activity is associated with low LT90 values.

In Table 2 we present. the Bayes estimator (α_k^*, β_k^*) and the corresponding $\tilde{\xi}_k$, ξ_k^* and $\hat{\xi}_k$ estimators of the inverse regression parameters corresponding to the threshold $\eta = 2.9$. This value of is the smallest $\ell n(AIRLT90)$ in Table 1. The Bayes estimators (α_k^*, β_k^*) were determined according to the empirical Bayes approach, described in Section 4, based on the 1963 and 1964 data. The LSE's (a_k, b_k) of the 1963 data yield the empirical Bayes estimates

(5.3) $$\begin{pmatrix} \hat{\alpha}_0 \\ \hat{\beta}_0 \end{pmatrix}_{63} = \begin{pmatrix} 3.8628 \\ 0.9241 \end{pmatrix}$$

The corresponding covariance matrix is

(5.4) $$(C)_{63} = \begin{pmatrix} .1645 & .0354 \\ .0354 & .0723 \end{pmatrix}$$

The corresponding covariance matrix for the 1963 data is

$$(5.5) \qquad (\bar{\Psi}_k)_{63} = \begin{pmatrix} .0018 & .0017 \\ .0017 & .0035 \end{pmatrix} .$$

Thus, the empirical Bayes estimate of the prior covariance matrix T is

$$(5.6) \qquad (T)_{63} = \begin{pmatrix} .1627 & .0337 \\ .0337 & .0683 \end{pmatrix} .$$

We computed similar estimates based on the 1964 data and obtained

$$(5.7) \qquad \begin{pmatrix} \hat{\alpha}_0 \\ \hat{\beta}_0 \end{pmatrix}_{64} = \begin{pmatrix} 3.5551 \\ 1.0673 \end{pmatrix}, \qquad (T)_{64} = \begin{pmatrix} .2437 & .0232 \\ .0232 & .0718 \end{pmatrix}$$

For the purpose of comparison we present in Table 2 for each site of the 1963 samples the empirical Bayes estimates corresponding to 1963 and 1964.

Observe that the values of $(\alpha_k^*, \beta_k^*)'$ are very close to $(a_k, b_k)' \pm (\text{S.D.}(a_k), \text{S.D.}(b_k))'$. As a result $\hat{\xi}_k$ and ξ_k^* have similar values. Also $.0001 \leq |\hat{\xi}_k - \xi_k^*| \leq .0623$ for every $k = 1,\ldots,41$. However, a visible effect on $\hat{\xi}_k$ and ξ_k^* is demonstrated by changing the prior distribution, corresponding to changes in the experimental conditions. This can be seen when the empirical prior distribution based on the 1964 data is applied to the 1963 data. Finally, the predicted response at XAIR, $Y^*(\text{XAIR}) = \alpha^* + \beta^* \text{XAIR}$, can be compared with the γ-fractile of the predictive distribution at $\hat{\xi}$, namely

$$(5.8) \qquad TL_\gamma = \alpha^* + \beta^* \hat{\xi} + z_\gamma \sqrt{Q(\hat{\xi})} .$$

In other words, if $Y^*(XAIR) > TL_\gamma$ we conclude, with predictive confidence γ, that the toxicity of the O.E. in the air is below the threshold.

REFERENCES

1. Aitchison, J. and Dunsmore, I.R. Statistical Predictive Analysis, Cambridge University Press, London, 1975

2. Bialik, O. (1977). Statistical Methods for the Analysis of Photodynamic Bioassay and Indexing Air Pollution Toxicity. Ph.D. Dissertation, Case Western Reserve University, Cleveland, Ohio.

3. Box, G.E.P. and Tiao, G. (1973). Bayesian Inference in Statistical Analysis. John Wiley and Sons, Inc., New York.

4. Cornfield, J. (1969). The Bayesian Outlook and its Application. Biometrics, 25, 4, 617-657.

5. Eichorn, B.E. and Zacks, S. (1973). Sequential Search of an Optimal Dosage I. Jour. Amer. Statist. Assoc., 68: No. 3. (September 1973).

6. Epstien, S.S., Sawicki, E. and Falk, H.L. (1965). Photodynamic Bioassay of Polycyclic Atmospheric Pollutants. Journ. Air Pollution Control, 15, No. 4, 174-176.

7. Friller, E.C. (1944). A Fundemental Formula in the Statitics of Biological Assay and Some Applications, Quarterly Jour. of Pharmacy and Pharmacology, 17: 117-213.

8. Graybill, F.A. Theory and Application of the Linear Model, Duxburg Press, North Scituate, Mass., 1976.

9. Krutchkoff, R.G. (1967). Classical and Inverse Regression Methods of Caliberation. Technometrics, 9: 425-439.

10. Wasan, M.T. Stochastic Approximation, CAmbridge Tracts in Mathematics and Mathematical Physics, No. 58. Cambridge University Press, London, 1969.

Table 1: Response Statistics of the Organic Extracts and the AIRLT90 for the 1963 Data

SITE	N	MEAN LT90			EXPECTED LT90			ST. DEV. LT90			A	B	$\hat{\sigma}$	SE{A}	SE{B}	R	AIRLT90
2	12	22.83	63.11	85.96	25.69	49.86	96.76	0.7	5.5	10.1	3.908	0.663	.037	.056	.068	0.954	54.37
3	8	28.50	59.59	*****	28.50	59.62	124.74	1.1	4.6	*****	4.087	0.738	.036	.095	.135	0.996	55.62
4	12	13.55	26.95	85.59	12.50	31.53	79.54	1.6	6.3	63.6	3.447	0.925	.093	.088	.108	0.984	31.48
5	8	11.47	22.79	*****	11.46	22.82	45.44	0.4	1.5	*****	3.126	0.689	.054	.116	.165	0.991	21.77
6	8	26.15	85.96	*****	26.13	86.08	283.51	2.0	21.6	*****	4.454	1.192	.054	.116	.165	0.997	91.90
7	8	14.55	38.98	*****	14.55	39.01	104.56	0.5	3.5	*****	3.663	0.986	.048	.109	.155	0.996	46.31
8	8	11.05	39.37	*****	11.04	39.40	140.59	0.2	2.7	*****	3.673	1.272	.042	.103	.146	0.998	28.66
9	12	15.41	38.35	83.70	15.75	36.71	85.57	0.1	0.6	2.8	3.603	0.846	.020	.041	.050	0.999	32.14
10	8	13.90	61.18	*****	13.90	61.22	269.64	0.3	5.7	*****	4.114	1.483	.039	.099	.139	0.999	61.03
11	12	14.68	40.49	89.69	15.21	37.65	93.15	0.5	3.5	17.0	3.627	0.906	.046	.062	.076	0.996	32.76
12	8	31.62	*****	89.82	31.61	53.30	89.88	1.3	13.0	10.5	3.975	0.523	.036	.067	.067	0.998	55.98
13	12	9.53	26.48	76.80	9.46	26.87	76.32	0.3	2.1	17.9	3.290	1.044	.055	.068	.083	0.998	25.33
14	8	30.61	64.62	*****	30.61	64.64	136.48	0.5	2.4	*****	4.168	0.747	.024	.078	.110	0.998	65.00
16	8	19.43	52.30	*****	19.44	52.31	140.74	0.9	6.3	*****	3.956	0.990	.048	.110	.155	0.996	59.06
17	8	28.25	72.73	*****	28.23	72.89	188.20	4.4	29.0	*****	4.286	0.949	.074	.136	.193	0.991	78.16
18	8	15.45	45.63	*****	15.70	45.67	132.85	15.9	138.8	*****	3.789	1.068	.254	.252	.356	0.925	50.51
19	8	16.91	30.31	*****	16.90	30.43	54.79	2.9	9.4	*****	3.410	0.588	.101	.159	.225	0.958	27.21
20	10	14.95	35.61	89.21	14.82	36.21	88.46	2.2	12.5	78.4	3.584	0.893	.099	.103	.133	0.992	34.26
22	8	36.57	79.85	*****	36.49	80.22	176.36	12.9	61.5	*****	4.380	0.788	.098	.156	.221	0.978	81.84
23	8	23.58	51.94	*****	23.59	51.94	114.40	0.6	2.8	*****	3.950	0.790	.032	.089	.126	0.998	57.38
30	8	11.11	21.11	*****	11.13	21.11	40.04	0.7	2.7	*****	3.047	0.640	.077	.139	.197	0.979	18.95
31	10	10.96	20.53	89.99	9.71	26.18	70.59	0.1	0.5	8.8	3.265	0.992	.033	.059	.077	0.964	25.19
32	8	12.23	43.77	*****	12.23	43.81	156.85	0.5	5.8	*****	3.778	1.276	.055	.117	.166	0.997	44.24
33	8	18.26	46.66	*****	18.26	46.73	119.60	1.5	9.8	*****	3.842	0.940	.067	.130	.183	0.992	46.72
34	8	11.09	40.51	*****	11.09	40.53	148.13	0.1	1.8	*****	3.701	1.296	.033	.091	.129	0.999	32.36
37	8	24.18	44.79	*****	24.19	44.80	82.96	1.1	3.7	*****	3.801	0.616	.043	.104	.147	0.993	44.98
43	8	21.96	69.45	*****	21.97	69.45	219.52	1.3	12.6	*****	4.239	1.151	.051	.113	.160	0.997	69.77
44	8	20.34	48.53	*****	20.35	48.53	115.77	0.2	0.9	*****	3.882	0.869	.020	.071	.101	0.999	52.21
45	12	14.29	36.76	74.32	14.87	33.92	77.38	0.1	0.8	3.2	3.524	0.825	.024	.045	.055	0.996	31.29
47	8	23.32	46.34	*****	23.34	46.33	91.99	1.4	5.4	*****	3.835	0.686	.050	.111	.157	0.992	41.73
48	8	28.81	73.76	*****	28.75	74.10	190.98	8.3	54.7	*****	4.300	0.947	.100	.150	.224	0.984	71.75
49	8	44.14	84.83	*****	44.13	84.87	163.20	2.3	8.3	*****	4.441	0.654	.034	.092	.131	0.996	186.99
50	8	28.70	72.93	*****	28.69	72.96	185.55	0.8	5.4	*****	4.289	0.933	.032	.089	.126	0.998	183.98
51	8	23.57	84.22	*****	23.58	84.24	300.93	1.1	13.7	*****	4.433	1.273	.044	.104	.148	0.908	99.37
52	8	18.41	40.69	*****	18.40	40.71	90.06	0.4	2.0	*****	3.706	0.794	.035	.009	.133	0.907	43.04
53	8	24.55	86.73	*****	24.54	86.85	307.35	1.7	21.2	*****	4.463	1.264	.053	.115	.162	0.997	64.60
54	9	21.87	34.31	93.61	20.59	38.85	73.30	3.1	7.6	56.3	3.656	0.635	.080	.106	.142	0.949	39.95
55	8	10.45	57.48	*****	10.49	57.82	318.86	4.0	120.2	*****	4.039	1.707	.189	.217	.308	0.982	33.56
56	12	10.93	28.75	71.18	11.02	28.30	72.67	3.9	27.2	166.8	3.327	0.943	.180	.122	.150	0.980	32.08
57	8	24.42	69.84	*****	24.39	70.05	201.19	3.8	31.3	*****	4.246	1.055	.080	.142	.200	0.991	90.38
59	8	35.35	85.64	*****	35.35	85.64	207.47	0.9	5.3	*****	4.450	0.885	.027	.081	.115	0.999	91.44

Notes:

1) N is the number of observations around each line.

2) Expected LT90 = $\exp\{A + BX + \hat{\sigma}^2/2\}$.

3) A and B are the least squares estimates, $\hat{\sigma}$ is the standard deviation around the regression lines.

4) AIRLT90 = $\exp\{A + BXAIR + \hat{\sigma}^2/2\}$, XAIR is given Table 2

Table 2: The Actual and the Critical Atmospheric Concentrations of Organic Extracts for 1963 Data

SITE	N	A	B	A^*_{63}	B^*_{63}	A^*_{64}	B^*_{64}	XAIR	$\tilde{\xi}_{63}$	ξ^*_{63}	$\hat{\xi}_{63}$	$\tilde{\xi}_{64}$	ξ^*_{64}	$\hat{\xi}_{64}$
2	12	3.908	0.663	3.908	0.664	3.908	0.664	0.1318	-1.5204	-1.5184	-1.5178	-1.5204	-1.5176	-1.5170
3	8	4.087	0.738	4.088	0.740	4.088	0.741	-.0929	-1.6084	-1.6053	-1.6040	-1.6084	-1.6039	-1.6026
4	12	3.447	0.925	3.449	0.923	3.447	0.927	0.0031	-0.5914	-0.5945	-0.5938	-0.5914	-0.5903	-0.5895
5	8	3.126	0.689	3.130	0.694	3.130	0.697	-.0663	-0.3280	-0.3315	-0.3320	-0.3280	-0.3306	-0.3311
6	8	4.454	1.192	4.450	1.186	4.451	1.189	0.0564	-1.3037	-1.3065	-1.3057	-1.3037	-1.3049	-1.3041
7	8	3.663	0.986	3.663	0.985	3.663	0.987	0.1754	-0.7738	-0.7746	-0.7743	-0.7738	-0.7733	-0.7730
8	8	3.673	1.272	3.671	1.268	3.672	1.269	-.2492	-0.6077	-0.6085	-0.6085	-0.6077	-0.6079	-0.6078
9	12	3.603	0.846	3.603	0.846	3.603	0.846	-.1564	-0.8310	-0.8310	-0.8310	-0.8310	-0.8308	-0.8307
10	8	4.114	1.483	4.111	1.477	4.111	1.478	-.0002	-0.8186	-0.8200	-0.8198	-0.8186	-0.8194	-0.8193
11	12	3.627	0.906	3.627	0.906	3.627	0.907	-.1521	-0.8024	-0.8028	-0.8026	-0.8024	-0.8018	-0.8015
12	8	3.975	0.523	3.977	0.527	3.977	0.528	0.0954	-2.0554	-2.0437	-2.0401	-2.0554	-2.0412	-2.0376
13	12	3.290	1.044	3.291	1.043	3.290	1.044	-.0552	-0.3736	-0.3751	-0.3750	-0.3736	-0.3738	-0.3737
14	8	4.168	0.747	4.168	0.748	4.168	0.748	0.0081	-1.6975	-1.6960	-1.6954	-1.6975	-1.6954	-1.6948
16	8	3.956	0.990	3.955	0.989	3.956	0.991	0.1241	-1.0667	-1.0672	-1.0665	-1.0667	-1.0657	-1.0650
17	8	4.286	0.949	4.284	0.948	4.285	0.952	0.0767	-1.4605	-1.4595	-1.4567	-1.4605	-1.4553	-1.4525
18	8	3.789	1.068	3.774	1.022	3.779	1.064	0.1248	-0.8324	-0.8548	-0.8475	-0.8324	-0.8264	-0.8200
19	8	3.410	0.588	3.424	0.611	3.426	0.620	-.1812	-0.8673	-0.8573	-0.8529	-0.8673	-0.8484	-0.8441
20	10	3.584	0.893	3.586	0.893	3.584	0.897	-.0562	-0.7660	-0.7680	-0.7668	-0.7660	-0.7629	-0.7617
22	8	4.380	0.788	4.380	0.797	4.383	0.802	0.0317	-1.8782	-1.8569	-1.8474	-1.8782	-1.8476	-1.8382
23	8	3.950	0.790	3.950	0.791	3.951	0.792	0.1268	-1.3291	-1.3279	-1.3272	-1.3291	-1.3269	-1.3262
30	8	3.047	0.640	3.056	0.652	3.057	0.658	-.1639	-0.2297	-0.2399	-0.2416	-0.2297	-0.2387	-0.2404
31	10	3.265	0.992	3.265	0.992	3.265	0.992	-.0380	-0.3679	-0.3685	-0.3685	-0.3679	-0.3680	-0.3680
32	8	3.778	1.276	3.775	1.268	3.776	1.272	0.0091	-0.6881	-0.6897	-0.6896	-0.6881	-0.6886	-0.6884
33	8	3.842	0.940	3.842	0.939	3.843	0.943	0.0023	-1.0021	-1.0025	-1.0013	-1.0021	-0.9995	-0.9983
34	8	3.701	1.296	3.700	1.293	3.700	1.294	-.1731	-0.6181	-0.6186	-0.6185	-0.6181	-0.6182	-0.6181
37	8	3.801	0.616	3.803	0.620	3.803	0.621	0.0082	-1.4627	-1.4561	-1.4538	-1.4627	-1.4536	-1.4513
43	8	4.239	1.151	4.236	1.147	4.237	1.149	0.0053	-1.1633	-1.1653	-1.1646	-1.1633	-1.1638	-1.1631
44	8	3.882	0.869	3.882	0.869	3.882	0.869	0.0844	-1.1300	-1.1299	-1.1297	-1.1300	-1.1296	-1.1294
45	12	3.524	0.825	3.524	0.825	3.524	0.825	-.0975	-0.7564	-0.7564	-0.7564	-0.7564	-0.7561	-0.7560
47	8	3.835	0.686	3.837	0.690	3.838	0.692	-.1504	-1.3630	-1.3574	-1.3551	-1.3630	-1.3545	-1.3523
48	8	4.300	0.947	4.296	0.945	4.298	0.952	-.0284	-1.4784	-1.4762	-1.4712	-1.4784	-1.4688	-1.4638
49	8	4.441	0.654	4.441	0.656	4.442	0.657	1.2091	-2.3563	-2.3488	-2.3464	-2.3563	-2.3472	-2.3448
50	8	4.289	0.933	4.289	0.933	4.289	0.934	0.9917	-1.4887	-1.4884	-1.4879	-1.4887	-1.4876	-1.4870
51	8	4.433	1.273	4.430	1.268	4.431	1.270	0.1305	-1.2042	-1.2064	-1.2060	-1.2042	-1.2054	-1.2050
52	8	3.706	0.794	3.707	0.795	3.707	0.796	0.0712	-1.0151	-1.0145	-1.0140	-1.0151	-1.0135	-1.0130
53	8	4.463	1.264	4.459	1.257	4.459	1.259	-.2328	-1.2366	-1.2398	-1.2391	-1.2366	-1.2383	-1.2376
54	9	3.656	0.635	3.658	0.642	3.658	0.644	0.0493	-1.1906	-1.1823	-1.1791	-1.1906	-1.1769	-1.1737
55	8	4.039	1.707	3.964	1.545	3.969	1.577	-.3079	-0.6673	-0.6889	-0.6877	-0.6673	-0.6781	-0.6770
56	12	3.327	0.943	3.337	0.935	3.329	0.949	0.1501	-0.4528	-0.4672	-0.4652	-0.4528	-0.4524	-0.4505
57	8	4.246	1.055	4.241	1.049	4.243	1.054	0.2448	-1.2758	-1.2785	-1.2764	-1.2758	-1.2744	-1.2723
59	8	4.450	0.885	4.450	0.885	4.450	0.886	0.0746	-1.7514	-1.7507	-1.7501	-1.7514	-1.7501	-1.7495

Notes:

1) A^*_{63}, A^*_{64}, B^*_{63} and B^*_{64} are the Bayes estimates of (α,β) based on the empirical Bayes of 1963 and 1964.

2) XAIR is the dosage equivalent of the actual air concentration of O.E.

3) $\tilde{\xi}, \xi^*, \hat{\xi}$ are the critical atmospheric concentrations (dosages) relative to $\eta = 2.9$

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 38	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EMPIRICAL BAYES ESTIMATION OF CRITICAL DOSAGES HAVING SMALLEST PREDICTIVE RISK.		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ora Bialik and S. Zacks		8. CONTRACT OR GRANT NUMBER(s) N 00014-75-C-0529 PROJECT NR 042-276
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics & Statistics Case Western Reserve University Cleveland, Ohio		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH ARLINGTON, VIRGINIA 22217		12. REPORT DATE August 15, 1979 13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear regression, Bayes estimates, empirical Bayes, critical dosages, calibration, predictive risk, photodynamic bioassays.

xi(eta) xi alpha beta

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An empirical Bayes procedure is developed for the estimation of critical dosages in the linear regression case. If $Y = \alpha + \beta x + e$ is the basic linear model, the critical dosage is defined as $\xi(\eta) = (\eta-\alpha)/\beta$, for $\beta > 0$. A new type of Bayes estimator of $\xi(\eta)$ is derived under the criterion of minimizing the predictive risk $E\{(\alpha + \beta\xi - \eta)^2 | \mathcal{F}_n\}$. The empirical Bayes procedure provides consistant estimators of the prior parameters when a large number of independent repetitions of the experiment is available. The methodology is developed to

eta

[20]
analyze a large set of photodynamic bioassays, for the determination of critical air concentrations of benzo-soluble organic extracts.