Statistica Sinica 4(1994), 581-601

THE USE OF HISTORICAL CONTROL DATA IN TESTING FOR TREND IN COUNTS

Brian G. Leroux, Karen Y. Fung, Daniel Krewski and Ross L. Prentice

Health Canada, University of Windsor, Carleton University and Fred Hutchinson Cancer Research Center

Abstract: A score statistic for testing for trend in count data based on the joint likelihood of historical control and current experimental counts is proposed using the negative-binomial distribution to represent between-study extra-Poisson variation. The score statistic is also derived using a Poisson likelihood for the current experiment and a negative-binomial likelihood for the historical controls. Similar statistics are derived using generalized estimating equations based on the first two moments of the data. The type I and type II error rates of these tests are evaluated using computer simulation, and compared with those of the Cochran-Armitage test for trend that does not make use of historical controls. Test statistics proposed by Tarone (1982) and Kikuchi and Yanagawa (1988) for use with historical controls are also considered.

Key words and phrases: Poisson distribution, negative-binomial distribution, extra-Poisson variation, overdispersion, score test, maximum likelihood, estimating equations.

1. Introduction

In many toxicological experiments, it is of interest to test for increasing response with increasing dose. In the Ames assay (Maron and Ames (1983)), for example, the number of mutant colonies of bacteria observed in replicate culture plates are counted at a series of increasing dose levels $d_0 < d_1 < \cdots < d_k$, including an unexposed control at $d_0 = 0$. Let X_{1ij} denote the number of mutant colonies observed in plate $j = 1, \ldots, n_{1i}$ at dose $i = 0, 1, \ldots, k$, and let $X_{1i} = \sum_j X_{1ij}$ denote the total number of mutant colonies at dose d_i. Suppose that the observations X_{1ij} follow independent Poisson distributions with means $\theta_1 \exp(\beta d_i)$, where θ_1 denotes the expected response in the control group. Then X_{1i} follows the Poisson distribution

$$Pr\{X_{1i} = x\} = e^{-\lambda_i} \lambda_i^x / x! \quad (x = 0, 1, 2, \ldots), \tag{1.1}$$

where $\lambda_i = n_{1i} \theta_1 \exp(\beta d_i)$.

To test for a linear trend in counts, Armitage (1955) proposed the statistic

$$T_0 = \sum_{i=0}^{k} X_{1i} d_i - \bar{X} \sum_{i=0}^{k} n_{1i} d_i, \tag{1.2}$$

with variance estimated by

$$\hat{V}(T_0) = \bar{X} \left[\sum_{i=0}^{k} n_{1i} d_i^2 - n_1^{-1} \left(\sum_{i=0}^{k} n_{1i} d_i \right)^2 \right], \tag{1.3}$$

where $\bar{X} = X_1/n_1$, $X_1 = \sum_i X_{1i}$ and $n_1 = \sum_i n_{1i}$. This statistic reduces to that proposed by Cochran (1954) in the case $n_{1i} \equiv n$, and is analogous to the statistic proposed by both Cochran and Armitage for testing for trend in binomial proportions. Tarone (1982) subsequently showed that the statistic T is asymptotically locally optimal for testing the null hypothesis of no trend against any smooth monotone increasing alternative. Under the null hypothesis, the standardized statistic $Z_{\text{CA}} = T_0/[\hat{V}(T_0)]^{1/2}$ is asymptotically normally distributed (see Appendix 2).

Suppose, now, that information on the response rate in the control groups from other experiments is available for analysis. Although the current control group represents the most appropriate group against which to compare the treated groups in the experiment at hand, the historical controls do provide some information on the spontaneous mutation rate. If properly utilized, this information may be of use in strengthening inferences based only on the current experimental data, particularly when the results are somewhat equivocal. However, it must be recognized that the experimental conditions of the historical and current experiments should be comparable. As well, care needs to be taken to accommodate any extra-Poisson variation that may be present in the historical controls due to between-study differences in experimental conditions.

Tarone (1982) proposed a formal statistical test for trend which makes use of the available historical control information. In the Ames assay, for example, let X_j denote the total number of mutant colonies in the n_j plates in historical control group $j = 2, \ldots, s$, and let θ_j denote the expected response rate in group j. Assuming that $\theta_1, \ldots, \theta_s$ are independent gamma random variates with mean μ and variance μ^2/ρ, it follows that $X_1, \ldots, X_s$ are independent negative-binomial random variates; in particular, for $j = 2, \ldots, s$,

$$Pr\{X_j = x\} = (n_j\mu/\rho)^x (x!)^{-1} (1+n_j\mu/\rho)^{-x-\rho} \prod_{i=0}^{x-1} (\rho+i) \quad (x = 0, 1, 2, \ldots), \tag{1.4}$$

with $E(X_j) = n_j\mu$ and $V(X_j) = n_j\mu(1 + n_j\mu/\rho)$; and a similar expression holds for the probability distribution of X_1, for which $E(X_1) = \mu \sum_{i=0}^{k} n_{1i} e^{\beta d_i}$ and

$V(X_1) = \mu \sum_{i=0}^{k} n_{1i} e^{\beta d_i} + (\mu^2/\rho)(\sum_{i=0}^{k} n_{1i} e^{\beta d_i})^2$. Here, the parameter $\rho > 0$ determines the degree of extra-Poisson variation. (The limiting case $\rho = \infty$ corresponds to Poisson variation.) Assuming the parameters μ and ρ to be known, Tarone (1982) showed that the score test for trend based on the negative-binomial likelihood of the data from the current experiment is

$$T = \sum_{i=0}^{k} X_{1i} d_i - \widetilde{X} \sum_{i=0}^{k} n_{1i} d_i, \tag{1.5}$$

where $\widetilde{X} = (X_1 + \rho)/\widetilde{n}$, with $\widetilde{n} = n_1 + \rho/\mu$. An estimate of the variance of T based on observed information is given by

$$\hat{V}(T) = \widetilde{X} \left\{ \sum_{i=0}^{k} n_{1i} d_i^2 - \widetilde{n}^{-1} \left(\sum_{i=0}^{k} n_{1i} d_i \right)^2 \right\}. \tag{1.6}$$

Note that Tarone's statistic $Z_T = T/[\hat{V}(T)]^{1/2}$ is of the same form as the Cochran-Armitage statistic except that $\bar{X}$ is replaced by $\widetilde{X}$, and n_1 is replaced by $\widetilde{n}$. Since the parameters μ and ρ are unknown in practice, Tarone (1982) proposed replacing them by their maximum-likelihood estimates $\hat{\mu}$ and $\hat{\rho}$ based on the historical control data. The asymptotic null distribution of the resulting statistic is standard normal under mild regularity conditions (Appendix 2); however, this statistic fails to account for estimation error in the historical control parameter values, which could affect its small-sample properties.

Noting that X_{10} is an ancillary statistic for the parameter β of interest, Kikuchi and Yanagawa (1988) showed that the score statistic based on the conditional likelihood is equivalent to that proposed by Tarone (1982), but with conditional variance

$$V(T|X_{10}) = [(X_{10} + \rho)/(n_{10} + \rho/\mu)]\hat{V}(T)/\widetilde{X}. \tag{1.7}$$

Kikuchi and Yanagawa (1988) demonstrated that the first six cumulants of the standardized conditional score statistic $Z_{\mathrm{KY}} = T/\sqrt{V(T|X_{10})}$ converge to those of a standard normal distribution.

In this paper, we develop score statistics based on the joint likelihood of the experimental and historical control data. Our approach is similar to that used by Prentice et al. (1992) for quantal response data. In addition to the negative-binomial model described above, we also consider a Poisson model for the experimental data coupled with a negative-binomial model for the historical control data, thereby retaining the original model for the current experiment even with the use of historical control information.

We also use an estimating-equation approach to derive tests for trend under the assumption of no extra-Poisson variation in the current experiment and also allowing for extra-Poisson variation in the current experiment. The estimating-equation approach is designed to avoid the parametric assumptions underlying the Poisson and negative-binomial models, with a possible sacrifice in efficiency under these models compared with the likelihood methods.

2. Score Tests for Trend with Historical Controls

2.1. The negative-binomial model

In this section we derive tests for trend under the negative-binomial model introduced in Section 1. When no overdispersion is apparent, each of the tests derived here reduces to the Cochran-Armitage test with the historical and current controls pooled to form one large control group.

2.1.1. Tests based on the likelihood

The likelihood function can be written as $L = L_1 L_2$, where

$$L_1 \propto \int_0^\infty \frac{\theta^{\rho-1} e^{-\theta\rho/\mu}}{\Gamma(\rho)(\mu/\rho)^\rho} \prod_{i=0}^{k} (n_{1i}\theta e^{\beta d_i})^{X_{1i}} e^{-n_{1i}\theta e^{\beta d_i}}\, d\theta \tag{2.1}$$

is the likelihood for $X_{10}, \ldots, X_{1k}$, and

$$L_2 \propto \prod_{j=2}^{s} (1 + n_j\mu/\rho)^{-X_j-\rho} (\mu/\rho)^{X_j} \prod_{\ell=0}^{X_j-1} (\rho + \ell) \tag{2.2}$$

is the likelihood for $X_2, \ldots, X_s$. Note that the likelihoods (2.1) and (2.2) would be unchanged if they were based on individual plate counts rather than total counts. The score statistic for testing $\beta = 0$ is given by

$$T = \frac{\partial \ell n L}{\partial \beta}\bigg|_{\beta=0,\mu=\hat{\mu}_0,\rho=\hat{\rho}_0} = \sum_{i=0}^{k} X_{1i} d_i - \frac{X_1 + \hat{\rho}_0}{n_1 + \hat{\rho}_0/\hat{\mu}_0} \sum_{i=0}^{k} n_{1i} d_i, \tag{2.3}$$

where $\hat{\mu}_0$ and $\hat{\rho}_0$ are maximum-likelihood estimates of μ and ρ under $\beta = 0$. To obtain these estimates the model is parametrized using $\phi = \mu/\rho$ in place of ρ to avoid numerical problems associated with the convergence of the negative-binomial likelihood to a Poisson likelihood as $\rho \to \infty$.

The variance of T can be estimated by

$$V_{\text{Obs}} = \iota_{\beta\beta} - \begin{bmatrix} \iota_{\beta\mu} \\ \iota_{\beta\rho} \end{bmatrix}^T \begin{bmatrix} \iota_{\mu\mu} & \iota_{\mu\rho} \\ \iota_{\mu\rho} & \iota_{\rho\rho} \end{bmatrix}^{-1} \begin{bmatrix} \iota_{\beta\mu} \\ \iota_{\beta\rho} \end{bmatrix}, \tag{2.4}$$

using the elements of the observed information matrix given in Appendix 1. If observed information is replaced by expected information, we obtain the simple variance estimator

$$V_{\text{Exp}} = \hat{\mu}_0 \left\{ \sum_{i=0}^{k} n_{1i} d_i^2 - \tilde{n}^{-1} \left(\sum_{i=0}^{k} n_{1i} d_i \right)^2 \right\}, \tag{2.5}$$

where

$$\tilde{n}^{-1} = \frac{\hat{\mu}_0}{\hat{\rho}_0 + n_1 \hat{\mu}_0} \left[1 + \frac{\hat{\rho}_0}{\hat{\rho}_0 + n_1 \hat{\mu}_0} \frac{1}{\sum_{j=1}^{s} n_j \hat{\mu}_0 / (\hat{\rho}_0 + n_j \hat{\mu}_0)} \right]. \tag{2.6}$$

We thus have two trend test statistics

$$Z_{\text{NBO}} = T/\sqrt{V_{\text{Obs}}} \tag{2.7}$$

and

$$Z_{\text{NBE}} = T/\sqrt{V_{\text{Exp}}}. \tag{2.8}$$

As shown in Appendix 2, Z_{NBO} is asymptotically normally distributed under the null hypothesis that $\beta = 0$. However, Z_{NBE} has a different limiting distribution because of the variation in the current control response rate θ_1. Conditional on θ_1, Z_{NBE} is asymptotically normally distributed with mean zero and variance θ_1/μ under the null hypothesis; averaging over the gamma density

$$f(\theta_1; \mu, \rho) = (\rho/\mu)^{\rho} \theta_1^{\rho-1} e^{-\theta_1 \rho/\mu} / \Gamma(\rho) \quad (\theta_1 > 0), \tag{2.9}$$

for θ_1 gives

$$P(Z_{\text{NBE}} > x) \approx \int_0^{\infty} [1 - \Phi(x\sqrt{\mu/\theta_1})] f(\theta_1; \mu, \rho) d\theta_1. \tag{2.10}$$

Note that the mixed-normal upper tail probability in (2.10) does not depend on μ and converges to the corresponding standard normal tail probability as $\rho \to \infty$. Table 1 gives the mixed-normal probability evaluated at standard normal critical values corresponding to significance levels 0.10, 0.05 and 0.01 for selected values of ρ. The mixed-normal and standard-normal tail probabilities are in reasonable agreement, especially when there is little overdispersion. This finding is similar to that reported by Krewski et al. (1991) using analogous models for quantal data. The mixed-normal tail probability tends to be higher than the normal at the 1% critical value and lower at the 10% critical value. This is consistent with a result of Shaked (1980) which implies that there is a unique positive critical value at which the two tail probabilities agree. In this case, it appears that this value is very close to the 5% critical value. Thus, in our simulation study we will assess the true error rates of all statistics, including

$Z_{\rm NBE}$, using standard normal critical values with a nominal significance level of 5%.

Table 1. The mixed-normal tail probability (2.10) evaluated at standard normal critical values

	$1 - E_{\theta_1}[\Phi(Z_{1-\alpha}\sqrt{\mu/\theta_1})]^*$		
ρ	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
0.4	0.070	0.046	0.022
1	0.082	0.049	0.019
2	0.088	0.049	0.016
6	0.095	0.049	0.013
20	0.098	0.050	0.011
40	0.099	0.050	0.010
100	0.100	0.050	0.010

* This probability does not depend on μ. It was evaluated numerically using the IMSL subroutine DQDAGS.

2.1.2. A test based on estimating equations

We now derive a trend test assuming only the moment structure implied by the negative-binomial model, viz.,

$$\begin{aligned}
\mu_1 &= E(\boldsymbol{X}_1) = (n_{10}\mu, \ldots, n_{1k}\mu e^{\beta d_k})^T, \\
\boldsymbol{V}_1 &= \operatorname{diag}(\mu_1) + \mu_1\mu_1^T\phi/\mu, \\
\mu_j &= E[X_j] = n_j\mu \quad (j = 2, \ldots, s), \quad \text{and} \\
V_j &= V(X_j) = n_j\mu(1 + n_j\phi) \quad (j = 2, \ldots, s),
\end{aligned} \tag{2.11}$$

where $\boldsymbol{X}_1 = (X_{10}, \ldots, X_{1k})^T$ and $\phi = \mu/\rho$. (The parametrization (μ, ϕ) is used here for mathematical convenience.)

The statistic is based on estimating equations of the form

$$\boldsymbol{D}_1^T \boldsymbol{V}_1^{-1}(\boldsymbol{X}_1 - \mu_1) + \sum_{j=2}^{s} \boldsymbol{D}_j^T V_j^{-1}(X_j - \mu_j) = \boldsymbol{0}, \tag{2.12}$$

where

$$\boldsymbol{D}_1^T = \begin{bmatrix} \partial\mu_1/\partial\mu \\ \partial\mu_1/\partial\beta \end{bmatrix} = \begin{bmatrix} n_{10} & n_{11}e^{\beta d_1} & \cdots & n_{1k}e^{\beta d_k} \\ 0 & n_{11}d_1\mu e^{\beta d_1} & \cdots & n_{1k}d_k\mu e^{\beta d_k} \end{bmatrix}, \tag{2.13}$$

and

$$\boldsymbol{D}_j^T = \begin{bmatrix} \partial\mu_j/\partial\mu \\ \partial\mu_j/\partial\beta \end{bmatrix} = \begin{bmatrix} n_j \\ 0 \end{bmatrix} \quad (j = 2, \ldots, s). \tag{2.14}$$

It is easily verified that $\boldsymbol{V}_1$ has inverse

$$\boldsymbol{V}_1^{-1} = \mu^{-1}\{\text{diag}(n_{10}^{-1}, \ldots, [n_{1k}e^{\beta d_k}]^{-1}) - (\phi^{-1} + \sum_{i=0}^{k} n_{1i}e^{\beta d_i})^{-1}\boldsymbol{J}\}, \tag{2.15}$$

where $\boldsymbol{J}$ is a matrix with all elements equal to unity. It follows that the estimating equations can be written as

$$\left[1 + \phi\sum_{i=0}^{k} n_{1i}e^{\beta d_i}\right]^{-1} \sum_{i=0}^{k}(X_{1i} - n_{1i}\mu e^{\beta d_i}) + \sum_{j=2}^{s}[1 + \phi n_j]^{-1}[X_j - n_j\mu] = 0, \tag{2.16}$$

and

$$\sum_{i=0}^{k} d_i\,(X_{1i} - n_{1i}\mu e^{\beta d_i}) - \left[\phi^{-1} + \sum_{i=0}^{k} n_{1i}e^{\beta d_i}\right]^{-1} \sum_{i=0}^{k}(X_{1i} - n_{1i}\mu e^{\beta d_i}) \sum_{i=0}^{k} n_{1i}d_i e^{\beta d_i} = 0. \tag{2.17}$$

Estimates $\hat{\mu}_0$ and $\hat{\phi}_0$ of μ and ϕ under $\beta = 0$ are obtained by iteratively solving the equations

$$\hat{\mu}_0 = \left[\sum_{j=1}^{s}(1 + n_j\hat{\phi}_0)^{-1}X_j\right] \Big/ \left[\sum_{j=1}^{s}(1 + n_j\hat{\phi}_0)^{-1}n_j\right], \tag{2.18}$$

and

$$\hat{\phi}_0 = \sum_{j=1}^{s} Z_j/s, \tag{2.19}$$

where $Z_j = [(X_j - n_j\hat{\mu}_0)^2/(n_j\hat{\mu}_0) - 1]/n_j$. Note that Equation (2.18) is derived from the first estimating equation (2.16).

The score-type statistic T for testing $\beta = 0$, obtained from the second estimating equation (2.17), has the same algebraic form as the score under the negative-binomial model given in (2.3). A variance estimate can be obtained using the negative of the derivative of the score-type vector $\boldsymbol{D}_1^T\boldsymbol{V}_1^{-1}[\boldsymbol{X}_1 - \mu_1]$ $+\sum_{j=2}^{s}\boldsymbol{D}_j^T V_j^{-1}[X_j - \mu_j]$, which is analogous to observed information in likelihood theory. This gives

$$V_{\text{EE-NBO}} = \hat{\mu}_0\frac{(1 + X_1\hat{\mu}_0^{-1}\hat{\phi}_0)}{(1 + n_1\hat{\phi}_0)}\left[\sum_{i=0}^{k} n_{1i}d_i^2 - \tilde{n}^{-1}\left(\sum_{i=0}^{k} n_{1i}d_i\right)^2\right], \tag{2.20}$$

where

$$\tilde{n}^{-1} = \frac{1}{(n_1 + \hat{\phi}_0^{-1})}\left(1 + \frac{1}{(1 + n_1\hat{\phi}_0)}\frac{1}{\sum_{j=1}^{s} X_j\hat{\mu}_0^{-1}/(n_j + \hat{\phi}_0^{-1})}\right). \tag{2.21}$$

If the expected "information" $(\boldsymbol{D}_1^T \boldsymbol{V}_1^{-1} \boldsymbol{D}_1 + \sum_{j=2}^s \boldsymbol{D}_j^T \boldsymbol{V}_j^{-1} \boldsymbol{D}_j)$ is used, the following variance estimate is obtained:

$$V_{\text{EE-NBE}} = \hat{\mu}_0 \left[\sum_{i=0}^{k} n_{1i} d_i^2 - \tilde{n}^{-1} \left(\sum_{i=0}^{k} n_{1i} d_i \right)^2 \right], \tag{2.22}$$

where

$$\tilde{n}^{-1} = \frac{1}{(n_1 + \hat{\phi}_0^{-1})} \left(1 + \frac{1}{(1 + n_1 \hat{\phi}_0)} \frac{1}{\sum_{j=1}^{s} n_j/(n_j + \hat{\phi}_0^{-1})} \right). \tag{2.23}$$

The latter variance estimate has the same algebraic form as the variance estimate based on expected information given by (2.5) and (2.6), although the parameter estimates take on different values.

The test statistics

$$Z_{\text{EE-NBO}} = T/\sqrt{V_{\text{EE-NBO}}} \tag{2.24}$$

and

$$Z_{\text{EE-NBE}} = T/\sqrt{V_{\text{EE-NBE}}} \tag{2.25}$$

have the same limiting distributions as the corresponding likelihood-based statistics, Z_{NBO} and Z_{NBE}, respectively (see Appendix 2).

2.2. The Poisson/negative-binomial model

In the previous section, we allowed for extra-Poisson variation from experiment to experiment. For the current experiment, however, extra variation due to differences between experiments may not be considered relevant. In fact, statistical analysis of data from the experiment at hand in the absence of historical control data would not allow for such overdispersion. Because of this, we consider the Poisson/negative-binomial model in which the observations in the current experiment are Poisson variates and the historical controls follow a negative-binomial distribution. Following Prentice et al. (1992), we assume that the mean μ of the historical control response rates is equal to the current control response rate θ_1. When the historical controls display no overdispersion, each of the tests derived here reduces to the Cochran-Armitage test with the historical and current controls pooled to form one large control group.

2.2.1. Tests based on the likelihood

The likelihood function can be written as $L = L_1 L_2$, where

$$L_1 \propto \prod_{i=0}^{k} \left(n_{1i} \mu e^{\beta d_i} \right)^{X_{1i}} e^{-n_{1i} \mu e^{\beta d_i}} \tag{2.26}$$

is the Poisson likelihood from the current study, and L_2 is the negative-binomial likelihood from the historical controls given in (2.2). The score statistic for testing $\beta = 0$ is given by

$$T = \frac{\partial \ell n L}{\partial \beta}\Big|_{\beta=0,\mu=\hat{\mu}_0,\rho=\hat{\rho}_0} = \sum_{i=0}^{k} X_{1i}d_i - \hat{\mu}_0 \sum_{i=0}^{k} n_{1i}d_i, \tag{2.27}$$

where $\hat{\mu}_0$ and $\hat{\rho}_0$ are maximum-likelihood estimates of μ and ρ under $\beta = 0$. As previously, these estimates are obtained using the parametrization (μ, ϕ).

The variance of T can be estimated by

$$V_{\text{Obs}} = \iota_{\beta\beta} - \iota_{\beta\mu}^2 \iota_{\rho\rho}/(\iota_{\rho\rho}\iota_{\mu\mu} - \iota_{\mu\rho}^2), \tag{2.28}$$

based on observed information (Appendix 1). Alternatively, using expected information the following simple formula results:

$$V_{\text{Exp}} = \hat{\mu}_0 \left\{ \sum_{i=0}^{k} n_{1i}d_i^2 - \left[n_1 + \hat{\rho}_0 \sum_{j=2}^{s} \frac{n_j}{(\hat{\rho}_0 + n_j\hat{\mu}_0)} \right]^{-1} \left(\sum_{i=0}^{k} n_{1i}d_i \right)^2 \right\}. \tag{2.29}$$

Note that this is similar in form to Tarone's variance estimate in (1.6).

The corresponding test statistics,

$$Z_{\text{PO}} = T/\sqrt{V_{\text{Obs}}} \tag{2.30}$$

and

$$Z_{\text{PE}} = T/\sqrt{V_{\text{Exp}}}, \tag{2.31}$$

are asymptotically normally distributed with mean zero and variance one under the null hypothesis (see Appendix 2).

2.2.2. A test based on estimating equations

We now derive a score-type test assuming only that the data has the moment structure implied by the model of the previous section. Specifically, we suppose

$$\begin{aligned} \mu_1 &= E[\boldsymbol{X}_1] = (n_{10}\mu, n_{11}\mu e^{\beta d_1}, \ldots, n_{1k}\mu e^{\beta d_k})^T \quad \text{and} \\ \boldsymbol{V}_1 &= \text{Cov}(\boldsymbol{X}_1) = \text{diag}(\mu_1). \end{aligned} \tag{2.32}$$

The moments of $X_2, \ldots, X_s$ are unchanged from Section 2.1.2 as are the derivatives $\boldsymbol{D}_j$.

The estimating equations can thus be written as

$$\sum_{i=0}^{k}(X_{1i} - n_{1i}\mu e^{\beta d_i}) + \sum_{j=2}^{s}(1 + n_j\phi)^{-1}(X_j - n_j\mu) = 0, \text{ and} \tag{2.33}$$

$$\sum_{i=0}^{k} d_i(X_{1i} - n_{1i}\mu e^{\beta d_i}) = 0. \tag{2.34}$$

Estimates $\hat{\mu}_0$ and $\hat{\phi}_0$ of μ and ϕ under $\beta = 0$ can be obtained by solving the equations

$$\hat{\mu}_0 = \left[X_1 + \sum_{j=2}^{s}(1 + n_j\hat{\phi}_0)^{-1}X_j\right] \Big/ \left[n_1 + \sum_{j=2}^{s}(1 + n_j\hat{\phi}_0)^{-1}n_j\right], \tag{2.35}$$

and

$$\hat{\phi}_0 = \sum_{j=2}^{s} Z_j/(s-1). \tag{2.36}$$

Note that Equation (2.35) is derived from the first estimating equation (2.33).

The second estimating equation evaluated at $\beta = 0$ and $\mu = \hat{\mu}_0$ gives the score-type statistic

$$T = \sum_{i=0}^{k} X_{1i}d_i - \hat{\mu}_0 \sum_{i=0}^{k} n_{1i}d_i. \tag{2.37}$$

Both the observed and expected "information" produce the same variance estimate for T, viz.,

$$V_{\text{EE-P}} = \hat{\mu}_0 \left\{ \sum_{i=0}^{k} n_{1i}d_i^2 - \left[n_1 + \sum_{j=2}^{s} n_j(1 + n_j\hat{\phi}_0)^{-1}\right]^{-1} \left(\sum_{i=0}^{k} n_{1i}d_i\right)^2 \right\}. \tag{2.38}$$

As with the previous model, this variance estimate has the same algebraic form as that in (2.29) based on expected information, although the parameter estimates are different. The test statistic

$$Z_{\text{EE-P}} = T/\sqrt{V_{\text{EE-P}}} \tag{2.39}$$

is asymptotically standard normal under the null hypothesis (see Appendix 2).

3. Examples

To illustrate the calculations of the above proposed statistics, consider the data in Table 2 previously analyzed by Tarone (1982). The results in Table 3 show that, with the exception of the Kikuchi-Yanagawa test, the use of historical control data in this example weakens the evidence for mutagenicity compared to the Cochran-Armitage test. This occurs because the historical control mutation rate (8.35 ± 0.36, standard error based on Fisher information) is much higher than that observed in the current experiment (3.33 ± 1.05), and the trend tests are based on the assumption of equal rates. The effect is most pronounced with tests based on the Poisson/negative-binomial model, which do not allow for extra-Poisson variation in the current experiment.

Table 2. Number of mutant colonies of Salmonella (TA1537) exposed to Benz(a)anthracene in the Ames assay†

Historical controls*
10(2), 13(1), 14(2), 15(2), 16(3), 17(4), 18(1), 19(4), 20(6), 21(6), 22(2), 23(4), 24(1), 26(2), 27(2), 28(2), 29(5), 31(1), 32(1), 33(1), 34(2), 35(2), 37(2), 38(1), 39(3), 40(1), 44(1), 46(1), 47(1)

Experimental data					
Dose:	0	0.3	1.0	3.3	10.0
Count:	10	18	21	16	35

† Each count is the sum of three replicate plate counts.
* Frequency of occurrence in parentheses.

Table 3. Tests for trend using the data in Table 2

Test statistic	Model parameters μ	ϕ	T	V	$Z = T/\sqrt{V}$	p-value
Z_{CA}			137.2	1387.0	3.684	0.00011
Z_{T}^* (M)	8.354	0.727	131.0	1489.5	3.394	0.00034
(EE)	8.354	0.711	130.9	1491.7	3.388	0.00035
(ML)	8.354	0.685	130.6	1495.4	3.378	0.00036
Z_{KY}^* (M)	8.354	0.727	131.0	1074.4	3.996	0.00003
(EE)	8.354	0.711	130.9	1080.8	3.981	0.00003
(ML)	8.354	0.685	130.6	1092.0	3.954	0.00004
Z_{PO}	8.040	0.671	77.0	2510.4	1.538	0.06205
Z_{PE}	8.040	0.671	77.0	2510.0	1.538	0.06205
$Z_{\text{EE-P}}$	8.015	0.770	78.1	2487.2	1.567	0.05862
Z_{NBO}	8.319	0.672	130.7	1493.0	3.382	0.00036
Z_{NBE}	8.319	0.672	130.7	1824.9	3.059	0.00111
$Z_{\text{EE-NBO}}$	8.320	0.709	131.0	1489.9	3.393	0.00034
$Z_{\text{EE-NBE}}$	8.320	0.709	131.0	1820.5	3.070	0.00107

* Estimates of the parameters μ and ϕ may be obtained by the method of moments (M), the estimating equations (2.18) and (2.19) (EE), or maximum likelihood (ML).

A more typical example, for which the application of the trend tests is reasonable, is given in Table 4. This data set was part of an interlaboratory study of the Ames assay (Claxton and Lewtas (1991)); the historical controls were obtained from different experiments performed in the same laboratory as part of this study. The current and historical controls appear to have equal response rates (44 ± 2 in the historical controls compared with 48 ± 5 in the current experiment). The results in Table 5 show that the use of historical controls by all tests strengthens the evidence for an increasing trend, particularly the tests based on the Poisson/Negative-Binomial model.

Table 4. Number of mutant colonies of Salmonella (TA98) exposed to Coal Tar solution in the Ames assay†

Historical controls*
52, 70, 74, 77, 85, 90, 97, 100, 107, 109, 113, 117, 123

Experimental data						
Dose:	0	1.25	2.5	5.0	7.5	10.0
Count:	96	120	115	104	109	132

† Each count is the sum of two replicate plate counts.

* Frequency of occurrence in parentheses.

Table 5. Tests for trend using the data in Table 4

Test statistic	Model parameters μ	ϕ	T	V	$Z = T/\sqrt{V}$	p-value
Z_{CA}			137.5	8362	1.504	0.0663
Z_{T}^* (M)	46.69	1.861	159.2	8851	1.692	0.0453
(EE)	46.69	1.679	161.4	8901	1.711	0.0435
(ML)	46.69	1.866	159.1	8850	1.692	0.0453
Z_{KY}^* (M)	46.69	1.861	159.2	7554	1.832	0.0335
(EE)	46.69	1.679	161.4	7598	1.852	0.0320
(ML)	46.69	1.866	159.1	7553	1.831	0.0336
Z_{PO}	54.22	3.359	248.4	9081	2.607	0.0046
Z_{PE}	54.22	3.359	248.4	10778	2.393	0.0084
$Z_{\text{EE-P}}$	53.64	2.297	278.8	11402	2.611	0.0045
Z_{NBO}	47.52	1.855	157.4	8722	1.685	0.0460
Z_{NBE}	47.52	1.855	157.4	7482	1.820	0.0344
$Z_{\text{EE-NBO}}$	47.53	1.648	159.7	8863	1.697	0.0448
$Z_{\text{EE-NBE}}$	47.53	1.648	159.7	7535	1.840	0.0329

* Estimates of the parameters μ and ϕ may be obtained by the method of moments (M), the estimating equations (2.18) and (2.19) (EE), or maximum likelihood (ML).

4. Small-Sample Evaluation of Test Statistics

In this section, the type I and type II error rates of the test statistics are evaluated using computer simulation under both the negative-binomial model and the Poisson/negative-binomial model. For the historical control series, gamma variates θ_j ($j = 2, \ldots, s$; $s = 6$ or 21) with mean $\mu = 8$, and $\rho = 6$ or 20 were generated using the IMSL (1987) subroutine RNGAM. These two values of ρ represent cases when the historical counts are moderately ($\rho = 20$) and highly ($\rho = 6$) overdispersed relative to Poisson variation. Three Poisson variates X_{j1}, X_{j2} and X_{j3} where then generated for each value of θ_j using IMSL subroutine RNPOI. For the current study, $k + 1 = 4$ dose levels ($d_0 = 0$, $d_1 = 1/4$, $d_2 = 1/2$ and $d_3 = 1$) were used, with $n_{1i} = 3$ ($i = 0, 1, 2, 3$) replicate counts for each dose. For the negative-binomial model, θ_1 was generated in the same way as the historical control response rates θ_j, whereas, for the Poisson/negative-binomial model, θ_1 took the value $\mu = 8$. Poisson responses for the different dose groups were then generated with means $\lambda_i = \theta_1 e^{\beta d_i}$.

Table 6 shows the type I error rates for the different test statistics based on 1000 simulations of the model described above with $\beta = 0$. The Cochran-Armitage test appears to adhere to the nominal 5% significance level. Tarone's test produced type I error rates only slightly above the desired 5% level with 20 historical control groups, but was subject to type I error rates of almost 10% with 5 historical controls. This inflation of the type I error rate indicates the need to take into account errors in parameter estimation, particularly with a small number of historical groups. The Kikuchi-Yanagawa test demonstrated error rates as high as 8% with 20 historical control groups, and over 10% with 5 historical controls. The different parameter estimation procedures (the method of moments, the estimating equations (2.18) and (2.19), and maximum likelihood) produced similar error rates for the Tarone and Kikuchi-Yanagawa tests. Tests based on likelihood methods or estimating equations derived under the Poisson/negative-binomial model (Z_{PO}, Z_{PE} and $Z_{\text{EE-P}}$) produced slightly inflated error rates when the assumed model is correct, but highly inflated error rates (13% – 19%) under the negative-binomial model with 20 historical groups. On the other hand, the tests based on the negative-binomial model (Z_{NBO}, Z_{NBE}, $Z_{\text{EE-NBO}}$ and $Z_{\text{EE-NBE}}$) adhered fairly closely to the nominal level under both models with 5 or 20 historical controls. For both models, the tests based on estimating equations had type I error rates similar to those of the corresponding likelihood-based tests.

The powers of the test statistics, based on 1000 simulations under $\beta = 0.5$, are presented in Table 7. The powers for those tests that produced unacceptable type I error rates (those as high as 0.07 — about three standard errors above

Table 6. Estimated type I error rates of tests for trend†

Test statistic	Negative-binomial model $\mu=8,\ \rho=20$	Negative-binomial model $\mu=8,\ \rho=6$	Poisson/negative-binomial model $\mu=8,\ \rho=20$	Poisson/negative-binomial model $\mu=8,\ \rho=6$
	$s-1=5$ *Historical control groups*			
Z_{CA}	0.054	0.052	0.048	0.056
Z_T (M)	0.082	0.070	0.060	0.073
(EE)	0.093	0.081	0.060	0.071
(ML)	0.096	0.083	0.064	0.076
Z_{KY} (M)	0.089	0.081	0.068	0.088
(EE)	0.104	0.089	0.069	0.082
(ML)	0.104	0.090	0.071	0.085
Z_{PO}	0.061	0.049	0.052	0.050
Z_{PE}	0.068	0.065	0.053	0.063
Z_{EE-P}	0.078	0.054	0.058	0.075
Z_{NBO}	0.078	0.069	0.057	0.070
Z_{NBE}	0.069	0.067	0.052	0.069
Z_{EE-NBO}	0.064	0.048	0.050	0.062
Z_{EE-NBE}	0.068	0.052	0.054	0.068
	$s-1=20$ *Historical control groups*			
Z_{CA}	0.055	0.052	0.066	0.056
Z_T (M)	0.064	0.050	0.056	0.052
(EE)	0.059	0.050	0.058	0.050
(ML)	0.067	0.053	0.056	0.051
Z_{KY} (M)	0.079	0.062	0.064	0.063
(EE)	0.080	0.061	0.063	0.061
(ML)	0.084	0.066	0.064	0.062
Z_{PO}	0.163	0.153	0.070	0.061
Z_{PE}	0.154	0.137	0.069	0.059
Z_{EE-P}	0.165	0.188	0.070	0.066
Z_{NBO}	0.059	0.050	0.058	0.051
Z_{NBE}	0.060	0.062	0.053	0.049
Z_{EE-NBO}	0.059	0.049	0.055	0.052
Z_{EE-NBE}	0.073	0.064	0.054	0.050

† Based on 1000 simulations of an experiment with nominal significance level .05, $d_0=0$, $d_1=1/4$, $d_2=1/2$, $d_3=1$, $n_{1i}=n_j=3$ $(i=0,1,\ldots,k;\ j=2,\ldots,s)$ and $s-1=5$ or 20 historical control groups. The standard errors of the estimated error rates range from 0.007 to 0.012.

Table 7. Estimated powers of tests for trend†

Test statistic	Negative-binomial model $\mu=8,\ \rho=20$	Negative-binomial model $\mu=8,\ \rho=6$	Poisson/negative-binomial model $\mu=8,\ \rho=20$	Poisson/negative-binomial model $\mu=8,\ \rho=6$
	$s-1=5$ *Historical control groups*			
Z_{CA}	0.664	0.640	0.649	0.663
Z_{T} (M)	0.707*	0.623*	0.781	0.732*
(EE)	0.710*	0.629*	0.782	0.750*
(ML)	0.713*	0.620*	0.794	0.735*
Z_{KY} (M)	0.724*	0.649*	0.800	0.766*
(EE)	0.727*	0.648*	0.804	0.784*
(ML)	0.722*	0.641*	0.808*	0.766*
Z_{PO}	0.647	0.551	0.741	0.673
Z_{PE}	0.681	0.579	0.777	0.730
$Z_{\mathrm{EE-P}}$	0.693*	0.618	0.801	0.769*
Z_{NBO}	0.700*	0.629	0.773	0.725*
Z_{NBE}	0.692	0.623	0.764	0.748
$Z_{\mathrm{EE-NBO}}$	0.685	0.616	0.762	0.722
$Z_{\mathrm{EE-NBE}}$	0.706	0.626	0.786	0.752
	$s-1=20$ *Historical control groups*			
Z_{CA}	0.672	0.621	0.672	0.672
Z_{T} (M)	0.726	0.613	0.790	0.720
(EE)	0.732	0.610	0.785	0.723
(ML)	0.731	0.607	0.791	0.722
Z_{KY} (M)	0.752*	0.648	0.815	0.761
(EE)	0.753*	0.643	0.815	0.761
(ML)	0.752*	0.644	0.820	0.763
Z_{PO}	0.758*	0.594*	0.907*	0.846
Z_{PE}	0.758*	0.586*	0.905	0.844
$Z_{\mathrm{EE-P}}$	0.761*	0.600*	0.908*	0.853
Z_{NBO}	0.727	0.610	0.791	0.720
Z_{NBE}	0.738	0.611	0.819	0.764
$Z_{\mathrm{EE-NBO}}$	0.727	0.607	0.786	0.717
$Z_{\mathrm{EE-NBE}}$	0.739*	0.613	0.821	0.764

† Based on 1000 simulations of an experiment with nominal significance level .05, $d_0=0$, $d_1=1/4$, $d_2=1/2$, $d_3=1$, $n_{1i}=n_j=3$ ($i=0,\ldots,k$; $j=2,\ldots,s$), $s-1=5$ or 20 historical control groups, and $\beta=0.5$. The standard errors of the power estimates range from 0.009 to 0.016.

* Note that tests with results marked with asterisks had inflated (≥ 0.07) type I error rates (see Table 6).

0.05) are indicated by asterisks. The use of historical control data produced more powerful tests than the Cochran-Armitage test in all cases except with only 5 historical control groups under the negative-binomial model with $\rho = 6$. The powers of all tests that use historical control data were higher with $\rho = 20$ than with $\rho = 6$ and were higher under the Poisson/negative-binomial model than under the negative-binomial model with the same parameter values.

With 20 historical control groups, the tests based on the negative-binomial model produced comparable power to Tarone's test under the negative-binomial model; however, Z_{NBE} and $Z_{\text{EE-NBE}}$ were more powerful than Tarone's test under the Poisson/negative-binomial model. With $\rho = 6$ in the negative-binomial model Z_{KY} was the most powerful test, whereas the highest powers under the Poisson/negative-binomial model were achieved by tests derived under this model.

With 5 historical control groups, the powers of the tests that use the historical control data were reduced only slightly in most cases from those obtained with 20 historical controls. $Z_{\text{EE-NBE}}$ is the most powerful test statistic under the negative-binomial model with $\rho = 20$ and also under the Poisson/negative-binomial model with $\rho = 6$.

5. Discussion and Conclusions

In this article, two different models were used to describe data from a current experiment as well as control data from historical experiments. The negative-binomial model is intended to accomodate the extra-Poisson variation in control response rates which arises, for instance, from inter-laboratory differences. Because extra-Poisson variation may not be considered pertinent in the analysis of the experiment of interest, a Poisson/negative-binomial model in which only Poisson variation is considered in the current experiment was introduced. Implementation of this latter model involved the assumption that the control response rates from the historical experiments are centered on that of the current experiment.

The tests proposed by Tarone and Kikuchi and Yanagawa both exhibit unacceptable type I error rates with 5 historical control groups. The trend tests derived here under the assumption of Poisson variation in the current experiment perform poorly in the presence of extra-Poisson variation. However, tests allowing for overdispersion in the experiment at hand achieve type I error rates close to the nominal 5% level, even under the Poisson/negative-binomial model. The use of historical control data can result in increased sensitivity in comparison with the Cochran-Armitage test used without historical controls. In particular, the statistic $Z_{\text{EE-NBE}}$ based on estimating equations appears to provide high sensitivity across a wide range of situations. The estimating-equation approach avoids the parametric assumptions of the likelihood-based approach, while producing a more powerful test. Tests based on estimating equations are also simpler to implement and require much less computing time.

The tests for trend with historical controls proposed in this article are applicable only when the historical control and current control data are compatible. Specifically, the tests based on the negative binomial model are derived under the assumption that the historical and current control response rates come from the same distribution. Tests based on the Poisson/negative binomial model assume that the expected historical control response rate is equal to the response rate in the current control group. Krewski et al. (1988) have shown that the type I error rates of these tests can be distorted when these assumptions are not satisfied. Krewski et al. (1987) considered the use of preliminary statistical tests to evaluate the compatibility of the historical and current controls: when the historical and current controls are dissimilar, the historical data are not used in testing for trend. A disadvantage of this latter approach is that the decision as to whether to use the historical data may be sensitive to slight perturbations in the experimental data. Further research on criteria to determine when historical data should be used in order to increase the efficiency of tests for trend is desirable.

The tests considered here all assume that the variation among observations obtained in an individual experiment corresponds to that of a Poisson distribution. Extensions to these procedures are required for application to situations where extra-Poisson variation is present in an individual experiment.

Appendix 1: Observed Information

The negative-binomial model

The elements of the observed information matrix used in Equation (2.4) are

$$\iota_{\beta\beta} = \hat{\mu}_0(\hat{\rho}_0 + X_1)(\hat{\rho}_0 + n_1\hat{\mu}_0)^{-1}\left[\sum_{i=0}^{k} n_{1i}d_i^2 - \hat{\mu}_0(\hat{\rho}_0 + n_1\hat{\mu}_0)^{-1}\left(\sum_{i=0}^{k} n_{1i}d_i\right)^2\right],$$

$$\iota_{\beta\mu} = \hat{\rho}_0(\hat{\rho}_0 + X_1)(\hat{\rho}_0 + n_1\hat{\mu}_0)^{-2}\sum_{i=0}^{k} n_{1i}d_i,$$

$$\iota_{\beta\rho} = -\hat{\mu}_0(X_1 - n_1\hat{\mu}_0)(\hat{\rho}_0 + n_1\hat{\mu}_0)^{-2}\sum_{i=0}^{k} n_{1i}d_i,$$

$$\iota_{\mu\mu} = \hat{\mu}_0^{-2}\sum_{j=1}^{s}\left[X_j - (X_j + \hat{\rho}_0)(n_j\hat{\mu}_0)^2(\hat{\rho}_0 + n_j\hat{\mu}_0)^{-2}\right],$$

$$\iota_{\mu\rho} = -\sum_{j=1}^{s} n_j(X_j - n_j\hat{\mu}_0)(\hat{\rho}_0 + n_j\hat{\mu}_0)^{-2}, \quad \text{and}$$

$$\iota_{\rho\rho} = \sum_{j=1}^{s}\left[\sum_{l=0}^{X_j-1}(\hat{\rho}_0+l)^{-2} - n_j\hat{\mu}_0\hat{\rho}_0^{-1}(\hat{\rho}_0+n_j\hat{\mu}_0)^{-1} - (X_j - n_j\hat{\mu}_0)(\hat{\rho}_0+n_j\hat{\mu}_0)^{-2}\right].$$

The Poisson/negative-binomial model

The elements of the observed information matrix used in Equation (2.28) are

$$\iota_{\beta\beta} = \hat{\mu}_0\sum_{i=0}^{k} n_{1i}d_i^2,$$

$$\iota_{\beta\mu} = \sum_{i=0}^{k} n_{1i}d_i,$$

$$\iota_{\mu\mu} = \hat{\mu}_0^{-2}\left\{X_1 + \sum_{j=2}^{s}\left[X_j - (X_j + \hat{\rho}_0)(n_j\hat{\mu}_0)^2(\hat{\rho}_0 + n_j\hat{\mu}_0)^{-2}\right]\right\},$$

$$\iota_{\mu\rho} = -\sum_{j=2}^{s} n_j(X_j - n_j\hat{\mu}_0)(\hat{\rho}_0 + n_j\hat{\mu}_0)^{-2}, \quad \text{and}$$

$$\iota_{\rho\rho} = \sum_{j=2}^{s}\left\{\sum_{l=0}^{X_j-1}(\hat{\rho}_0+l)^{-2} - n_j\hat{\mu}_0\hat{\rho}_0^{-1}(\hat{\rho}_0+n_j\hat{\mu}_0)^{-1} - (X_j - n_j\hat{\mu}_0)(\hat{\rho}_0+n_j\hat{\mu}_0)^{-2}\right\}.$$

Appendix 2: Asymptotic Distributions of Trend Test Statistics

The Cochran-Armitage statistic Z_{CA} is asymptotically standard normal provided $n_1 \to \infty$ and $\max_{0\le i\le k}(d_i - \bar{d})^2/s_d^2 \to 0$, where $s_d^2 = \Sigma n_{1i}(d_i - \bar{d})^2$ and $\bar{d} = \Sigma n_{1i}d_i/n_1$, by the Lindeberg Central Limit Theorem. Tarone's statistic Z_{T} is also asymptotically normal under the additional condition $\bar{d}/s_d \to 0$, either with a fixed set of historical experiments, or with an increasing number of historical controls provided that the estimates $\hat{\mu}$ and $\hat{\rho}$ are consistent and $\rho < \infty$.

The Poisson/negative-binomial model

The asymptotic normality of the test statistics based on the likelihood, Z_{PE} and Z_{PO}, can be proved using standard likelihood-theory methods. In particular, the results follow from a minor extension of Theorem 6.6.1 of Lehmann (1983), which concerns several iid samples of data simultaneously. In our situation the iid samples are indexed by the dose groups $i = 0, \ldots, k$ in the current experiment and the subsets of historical controls with $n_j = n$, for $n = 1, \ldots, N$, where $N = \max_{2\le j\le s} n_j$. The extension of the theorem results from the observation that the proof does not require the Fisher information corresponding to each sample to be positive-definite as assumed, but only that the total information, across all samples, be positive-definite. This extension is necessary for our purposes since the current experiment provides no information on ρ and the historical controls provide no information on β. The extended theorem applies to the Poisson/negative-binomial model under the following conditions.

(A) $n_1 \to \infty$,
(B) $n_{1i}/n_1 \to \lambda_i$,
(C) $\Sigma\lambda_i(d_i - \Sigma\lambda_i d_i)^2 > 0$,
(D) $s/n_1 \to c > 0$,
(E) $s^{-1}\#\{j : n_j = n\} \to p_n$, where $\Sigma_{n=1}^{N} p_n = 1$.

The test statistic ($Z_{\mathrm{EE-P}}$) based on estimating equations is easily shown to be asymptotically normal as well. This result does not require the distributional assumptions of the Poisson/negative-binomial model. Instead, we assume the first and second moments which were used in the derivation of the test, and the additional moment assumption

(F) $E(X_j^4) \le M < \infty, \quad (j = 2, \ldots, s)$.

Conditions (A), (D) and (E) above are also used, but in place of (B) and (C), the weaker assumptions given below are made:

(B′) $\max_{0\le i\le k} d_i^2/n_1 \to 0$,
(C′) $\bar{d} \to \mu_d > 0,\ s_d^2/n_1 \to \sigma_d^2 > 0$.

The first step in the proof establishes the consistency of $\hat{\mu}_0$, by showing that the estimating equation for μ has a consistent sequence of solutions. Noting that $\hat{\phi}_0$ can be written $\hat{\phi}(\hat{\mu}_0)$, the next step proves that $\hat{\phi}(\mu)$ is $\sqrt{s}$-consistent for fixed μ. The same asymptotic distribution for $Z_{\mathrm{EE-P}}$ holds with an arbitrary $\sqrt{s}$-consistent estimator $\hat{\phi}(\mu)$. This is not surprising given the general result of Liang and Zeger (1986) on the asymptotic distributions of estimators derived from estimating equations.

The negative-binomial model

For this model, the asymptotic distributions of the likelihood-based test statistics can be derived under the conditions (A), (B′), (C′), (D) and (E) given above. As before, results for the test statistics based on estimating equations require fewer assumptions; specifically, instead of the distributional assumptions, we assume the first and second moments used in the derivation of the test, and the additional moment assumptions (F) and

(G) $E(X_1^4) = O(n_1^4)$.

Each of the statistics derived under this model has numerator of the form

$$T = \sum X_{1i}d_i - (X_1 + \hat{\mu}_0/\hat{\phi}_0)(n_1 + \hat{\phi}_0^{-1})^{-1} \sum n_{1i}d_i.$$

It can be shown that $(T - T_0)/s_d \to_P 0$, where T_0 is the numerator of Z_{CA}. Note that T_0/s_d has an approximate $N(0, \theta_1)$ distribution, conditional on θ_1. Also, the estimators $\hat{\mu}_0$ and $\hat{\phi}_0$ (based on likelihood or estimating equations) are consistent under the conditions stated above.

The asymptotic distribution of a particular test statistic depends on the asymptotic behavior of the variance estimate used. Two cases arise. First, if $\hat{V}/s_d^2$ converges in probability to μ, as for Z_{NBE} and $Z_{\mathrm{EE-NBE}}$, then

$$\frac{T}{\hat{V}^{1/2}} = \frac{s_d}{\hat{V}^{1/2}} \left[\frac{T_0}{s_d} + \frac{T - T_0}{s_d}\right]$$

converges in distribution to Z where Z has the mixed-normal distribution given in (2.10). In the second case, which applies to Z_{NBO} and $Z_{\mathrm{EE-NBO}}$, $\hat{V}/V$ converges in probability to 1 unconditionally, where V/s_d^2 converges in probability to θ_1, conditional on θ_1. This implies that

$$\frac{T}{\hat{V}^{1/2}} = \left(\frac{V}{\hat{V}}\right)^{1/2} \left[\frac{s_d}{V^{1/2}}\frac{T_0}{s_d} + \frac{s_d}{V^{1/2}}\frac{T - T_0}{s_d}\right]$$

converges in distribution to the standard normal.

All of these test statistics, Z_{NBE}, $Z_{\mathrm{EE-NBE}}$, Z_{NBO} and $Z_{\mathrm{EE-NBO}}$, have the same distribution under the Poisson/negative-binomial model as under the negative-binomial model. Because X_1 has smaller variance under the model with no overdispersion in the current experiment, the same proofs apply. However, the limiting distributions (if they exist) of $Z_{\mathrm{EE-P}}$, Z_{PE} and Z_{PO} under the negative-binomial model are not easily determined. We cannot approximate the score $T = \Sigma X_{1i}d_i - \hat{\mu}_0 \Sigma n_{1i}d_i$ by T_0 as we did previously. The estimate $\hat{\mu}_0$ will not be consistent (or converge in probability to any constant), and we can expect there to be additional variability in the limiting distributions of the test statistics. The simulation results of Section 4 confirm this expectation.

Acknowledgements

This research was supported in part by grant no. A4649 from the Natural Sciences and Engineering Research Council of Canada to K. Y. Fung, by grant no. A8664 to D. Krewski, and by grant no. CA-53996 to R. L. Prentice from the National Cancer Institute (USA). This paper was presented at The First Conference on Recent Developments in Statistical Research held by the International Chinese Statistical Association in Hong Kong, in December, 1990.

References

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics* **11**, 375-386.

Claxton, L. D. and Lewtas, J. (1991). *A Summary of the EPA/NIST/IPCS Collaborative Study on Complex Mixtures Using the Salmonella Mutagenicity Assay.* Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park.

Cochran, W. G. (1954). Some methods for strengthening the common X^2 tests. *Biometrics* **10**, 417-451.

IMSL (1987). STAT/Library Users Manual. IMSL, Houston.

Kikuchi, Y. and Yanagawa, T. (1988). Incorporating historical information in testing for a trend in Poisson means. *Ann. Inst. Statist. Math.* **40**, 367-379.

Krewski, D., Smythe, R. T. and Colin, D. (1987). Tests for trend in binomial proportions with historical controls: A proposed two-stage procedure. In *Advances in the Statistical Sciences, Vol. V: Biostatistics* (Edited by I. B. MacNeill and G. J. Umphrey), 61-69. D. Reidel Publishing, Boston.

Krewski, D., Smythe, R. T., Dewanji, A. and Colin, D. (1988). Statistical tests with historical controls (with discussion). In *Carcinogenicity: The Design, Analysis and Interpretation of Long-Term Animal Studies* (Edited by H. Grice and J. L. Ciminera), 23-28, Springer-Verlag, New York.

Krewski, D., Smythe, R. T., Fung, K. and Burnett, R. (1991). Conditional and unconditional tests with historical controls. *Canad. J. Statist.* **19**, 407-423.

Lehmann, E. L. (1983). *Theory of Point Estimation.* John Wiley, New York.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

Maron, D. M. and Ames, B. N. (1983). Revised methods for the *Salmonella* mutagenicity test. *Mutation Research* **113**, 173-215.

Prentice, R. L., Smythe, R. T., Krewski, D. and Mason, M. (1992). On the use of historical control data to estimate dose response trends in quantal bioassay. *Biometrics* **48**, 459-478.

Shaked, M. (1980). On mixtures from exponential families. *J. Roy. Statist. Soc. Ser. B* **42**, 192-198.

Tarone, R. E. (1982). The use of historical control information in testing for a trend in Poisson means. *Biometrics* **38**, 457-462.

Health Protection Branch, Health Canada, Ottawa, ON K1A 0L2, Canada.

Department of Mathematics and Statistics, University of Windsor, Windsor, ON N9B 3P4, Canada.

Department of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada.

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98104, U.S.A.

(Received January 1992; accepted January 1994)