THE ANALYSIS OF COUNT DATA IN A ONE-WAY LAYOUT

Yuhua Wang

Dallas E. Johnson

Linda J . Young

The Analysis of Count Data in a One-Way Layout

Yuhua Wang
Quintiles Inc. P.O. Box 9708, Kansas City, MO 64134-0708
Dallas E. Johnson
Department of Statistics, Kansas State University
Linda J. Young
Department of Biometry, University of Nebraska-Lincoln

ABSTRACT

An efficient score statistic for testing the equality of the means of several groups of count data in the presence of a common dispersion parameter is introduced and a new approximation to its distribution is given. The performance of the efficient score statistic using this approximation, the original efficient score statistic approximated by $\chi^2(t-1)$, the likelihood ratio statistic and four more ANOVA methods based on raw data or transformed data are compared in terms of size and power by using Monte Carlo simulations. The efficient score statistic with its new approximation is recommended. An application is given.

1. Introduction

The two distributions commonly used to model count data, the Poisson and the negative binomial (with known dispersion parameter), are members of the exponential family. The generalized linear model (GLIM) presents an exciting alternative to the general linear model (GLM) in that the form of the distribution can be incorporated in the model. However, unless the population distribution is normal, the tests associated with the GLIM are asymptotic. One concern is whether we have sufficient data support for the asymptotic theory when dealing with standard experiments involving relatively small sample sizes.

Another concern with the generalized linear model is the role of the dispersion parameter. Two possibilities have been suggested in the literature for accounting for over-dispersion in the model. Cox (1983), in the context of maximum likelihood estimation, calls for the use of the negative binomial distribution as a detailed representation of over-dispersion in the Poisson case. In this case, Y_j can be considered a Poisson random variable with mean λ_j, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are iid gamma with mean μ and dispersion parameter τ. Then $Y_1, Y_2, \ldots, Y_n$ are iid negative binomial with parameters μ and τ. The use of the negative binomial distribution instead of the Poisson to model over-dispersed count data in the generalized linear model analysis follows naturally from this idea.

McCullagh and Nelder (1989) suggest a different approach for over-dispersed data. The data are modeled based on the anticipated distribution, such as the Poisson, and a over-dispersion parameter ϕ $(\phi > 1)$ is added to account for over-dispersion. Then the dispersion parameter ϕ is

estimated from the data by one of two methods. The first approach is to set ϕ equal to the deviance function divided by the degrees of freedom. The deviance function has the form

$$D(y, \mu) = 2\phi(l(y; y) - l(\hat{\mu}; y)),$$

where $l(y; y)$ and $l(\hat{\mu}; y)$ are the likelihoods as a function of the data and estimated parameters, respectively. The second method equates ϕ and Pearson's χ^2 statistic divided by the degrees of freedom. Pearson's χ^2 statistic is

$$\chi^2 = \sum (y - \hat{\mu})^2 / V(\hat{\mu}),$$

where $V(\hat{\mu})$ is the estimated variance function. Both procedures are too liberal when we use them to model over-dispersed count data and to test the null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_t$ (see Young, et al., (1999)). Therefore, we do not consider the approach of estimating ϕ here.

The classical statistical setting for hypothesis testing involves a sequence of independent random variables whose distribution depends on a *t*-dimensional parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_t)'$ belonging to a sample space Θ, an open subset of *t*-dimensional Euclidean space $\Re^t$. A null hypothesis H_0 usually involves restrictions of the parameter, $\theta = (\theta_1, \theta_2, \ldots, \theta_t)'$, $R_j(\theta) = 0$ for $j = 1, 2, \cdots, r, (r \le t)$. Now consider tests of the (composite) hypotheses

$$H_0 : R(\theta) = 0 \quad \text{vs.} \quad H_1 : R(\theta) \neq 0 \qquad (1.1)$$

where $R(\theta) = [R_1(\theta), \ldots, R_r(\theta)]'$ is a vector-valued function $R : \Re^t \to \Re^r$ such that the $(t \times r)$ matrix $W(\theta) = \left(\partial R_j / \partial \theta_i \right)$ exists and is continuous in θ and $\text{rank}(W(\theta)) = r$. More specifically, suppose $\tilde{\theta}$ is the MLE of θ under the restrictions imposed by the null hypothesis, $\hat{\theta}$ is the (unrestricted) MLE of θ and n is the sample size. Further, $I(\theta) = \left(-E\left[\frac{\partial^2 \log\{f(Y, \theta)\}}{\partial \theta_i \partial \theta_j} \right] \right)$ is Fisher's Information Matrix.

Tests of H_0 against H_1 have typically involved one of the three test statistics:

1. Neyman-Pearson's likelihood ratio statistic given by
   $\lambda = 2\left(L(\hat{\theta}) - L(\tilde{\theta})\right)$
   where L is the log likelihood;
2. Rao's efficient score statistic
   $\zeta = S'(\tilde{\theta})[I(\tilde{\theta})]^{-1} S(\tilde{\theta})$
   where $S(\theta)$ is the likelihood score function defined by $\partial L(\theta) / \partial \theta$; and
3. Wald's test statistic
   $\omega = n R'(\hat{\theta}) \left(W'(\hat{\theta}) \left(I(\hat{\theta}) \right)^{-1} W(\hat{\theta}) \right)^{-1} R(\hat{\theta}).$

Under H_0, each of the three statistics has an asymptotic $\chi^2(r)$ distribution. Rao's efficient score statistic depends only on the MLE for the restricted class of parameters under H_0, while Wald's

statistic depends on the MLE over the whole parameter space. Because Wald's test is complicated and does not have an explicit form, it is not considered here.

2. Testing equality of count means with a known common dispersion parameter τ

Consider the density function form for the negative binomial distribution that was proposed by Bliss & Owen (1958) in which the random variable Y follows a negative binomial distribution with mean μ and dispersion parameter τ, NB(μ, τ), if

$$\Pr(Y = y) = \frac{\Gamma(y+\tau^{-1})}{y!\,\Gamma(\tau^{-1})}\left(\frac{\tau\mu}{1+\tau\mu}\right)^{y}\left(\frac{1}{1+\tau\mu}\right)^{1/\tau}, \qquad (2.1)$$

for $y = 0, 1, 2, \ldots$; $\tau \geq 0$; and $\mu > 0$. For this parameterization $E(Y) = \mu$ and $Var(Y) = \mu + \mu\tau^2$. Now suppose

$$Y_{ij} \sim \text{ind NB}(\mu_i, \tau) \text{ for } j = 1, 2, \cdots, n_i \text{ and } i = 1, 2, \cdots, t. \qquad (2.2)$$

Consider testing $H_0 : \mu_1 = \mu_2 = \ldots = \mu_t$ versus H_1 : not all μ_i's are equal.

It follows from (2.1) and (2.2) that the log-likelihood function of the μ_i's is

$$L(\mu_1, \ldots, \mu_t) = \sum_{i=1}^{t}\left[\sum_{j=1}^{n_i}\log\left(\frac{\Gamma(Y_{ij}+\tau^{-1})}{Y_{ij}!\,\Gamma(\tau^{-1})}\right) + \left(\sum_{j=1}^{n_i} Y_{ij}\right)\log\left(\frac{\tau\mu_i}{1+\tau\mu_i}\right) - \frac{n_i}{\tau}\log(1+\tau\mu_i)\right]. \qquad (2.3)$$

2.1. Likelihood ratio test

Barnwal and Paul (1988) obtained Neyman-Pearson's likelihood ratio statistic under the assumption of a common dispersion parameter τ as

$$\begin{aligned}\lambda &= 2\left(L(\hat{\mu}) - L(\tilde{\mu})\right) \\ &= 2\sum_{i=1}^{t}\left[n_i \overline{Y}_i \log(\tau\overline{Y}_i)\right] - 2\sum_{i=1}^{t}\left[n_i(\overline{Y}_i + \tau^{-1})\log(1+\tau\overline{Y}_i)\right] - \\ &\quad 2\left(\sum_{i=1}^{t} n_i\right)\overline{Y}\log(\tau\overline{Y}) + 2\left(\sum_{i=1}^{t} n_i\right)\left[(\overline{Y} + \tau^{-1})\log(1+\tau\overline{Y})\right] \end{aligned} \qquad (2.4)$$

where $\overline{Y}_i$ is the (unrestricted) maximum likelihood estimator of μ_i under the full model

$$\hat{\mu}_i = \overline{Y}_i = Y_{i\bullet}/n_i = \sum_{j=1}^{n_i} Y_{ij}/n_i \quad (i = 1, \ldots, t). \qquad (2.5)$$

Let μ represent the common value of each μ_i under H_0. Then $\overline{Y}$ is the maximum likelihood estimator of each μ under H_0. That is, we have

$$\tilde{\mu} = \overline{Y} = \sum_{i=1}^{t}\frac{Y_{i\bullet}}{n}, \text{ where } n = \sum_{i=1}^{t} n_i. \qquad (2.6)$$

The likelihood ratio statistic λ has an asymptotic distribution that is chi-square with $(t-1)$ degrees of freedom.

2.2. Efficient score statistic

In this section, we develop Rao's efficient score statistic for testing the hypotheses

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_t \quad \text{vs.} \quad H_1 : \text{not all } \mu_i\text{'s are equal.}$$

Differentiating the log-likelihood function of the μ_i's with respect to μ_i yields

$$\frac{\partial L(\mu_1, \mu_2, \ldots, \mu_t)}{\partial \mu_i} = \frac{(n_i Y_{i\bullet})}{\mu_i(1+\tau\mu_i)} - \frac{n_i}{(1+\tau\mu_i)}. \tag{2.7}$$

Next replace μ_i in (2.7) by $\tilde{\mu} = \overline{Y}$. It follows from the definition of $S(\tilde{\mu})$ that

$$S'(\tilde{\mu}) = \left(\frac{n_1 \overline{Y}_1}{\overline{Y}(1+\tau\overline{Y})} - \frac{n_1}{(1+\tau\overline{Y})}, \ldots, \frac{n_t \overline{Y}_t}{\overline{Y}(1+\tau\overline{Y})} - \frac{n_t}{(1+\tau\overline{Y})} \right). \tag{2.8}$$

The (r, s)th element of Fisher's Information Matrix is given by

$$i_{r,s} = -E\left(\frac{\partial^2 L(\mu_1, \mu_2, \ldots, \mu_t)}{\partial \mu_r \partial \mu_s} \right). \tag{2.9}$$

With the help of the equation

$$\frac{\partial^2 L(\mu_1, \mu_2, \ldots, \mu_t)}{\partial \mu_i^{\,2}} = \frac{(Y_{i\bullet}) \cdot (-1 - 2\tau\mu_i)}{\mu_i^2 (1+\tau\mu_i)^2} + \frac{\tau n_i}{(1+\tau\mu_i)^2},$$

(2.9) becomes

$$i_{r,s} = \begin{cases} 0 & \text{if } r \neq s \\ \dfrac{n_i}{\mu_i(1+\tau\mu_i)} & \text{if } r = s \end{cases}. \tag{2.10}$$

Then, replacing each μ_i in (2.10) by $\overline{Y}$ we obtain

$$I(\tilde{\mu}) = \text{diagonal}\left(\frac{n_1}{\overline{Y}(1+\tau\overline{Y})}, \ldots, \frac{n_t}{\overline{Y}(1+\tau\overline{Y})} \right).$$

Thus, the efficient score statistic is

$$\zeta = S'(\tilde{\mu})[\, I(\tilde{\mu})\,]^{-1} S(\tilde{\mu})$$

$$= \sum_{i=1}^{t} \left(\frac{n_i \overline{Y}_i}{\overline{Y}(1+\tau\overline{Y})} - \frac{n_i}{1+\tau\overline{Y}} \right)^2 \left(\frac{n_i}{\overline{Y}(1+\tau\overline{Y}} \right)^{-1} = \sum_{i=1}^{t} \frac{n_i (\overline{Y}_i - \overline{Y})^2}{\overline{Y}(1+\tau\overline{Y})}. \tag{2.11}$$

Equation (2.11) is identical to Neyman's $C(\alpha)$ statistic when τ is known, as indicated by Barnwal and Paul (1988).

2.3. F-tests

One common way to analyze count data in a one-way layout data is to use analysis of variance and rely on the Central Limit Theorem and the robustness properties of the F test. If

variances are unequal, the impact on the F test can be significant for the one-way classification. Several variance stabilizing transformations have been suggested to address this concern.

(1) Anscombe (1949) suggested the transformation, $f(y) = \log(y+1)$, for the negative binomial distribution.

(2) The square root transformation, $f(y) = \sqrt{y}$, is commonly used when the populations are Poisson distributed.

(3) Beall (1942) suggested the following variance-stabilizing transformation for the negative binomial distribution provided that τ is known:

$$f(y) = \sqrt{\frac{1}{\tau}} \sinh^{-1} \sqrt{(y+0.5)\tau}.$$

Therefore, an analysis of variance could be performed on the raw data or (1) $f_1(y) = \log(y+1)$, (2) $f_2(y) = \sqrt{y}$, or (3) $f_3(y) = \sqrt{\frac{1}{\tau}} \sinh^{-1} \sqrt{(y+0.5)\tau}$.

If the resulting F statistic is significant then H_0 is rejected.

3. An approximation to the distribution of the efficient score statistic

Barnwal and Paul (1988) approximated the distribution of the efficient score statistic using $\chi^2(t-1)$. This section provides an alternative approach.

Let $X_1 = \sum_{i=1}^{t} \left(\frac{Y_{i\bullet}^2}{n_i} \right) - \frac{Y_{\bullet\bullet}^2}{n}$ and $X_2 = \overline{Y}(1+\tau\overline{Y}) = \frac{Y_{\bullet\bullet}}{n}\left(1 + \frac{\tau}{n} Y_{\bullet\bullet}\right)$. Then the efficient score statistic can be written as

$$\zeta = \frac{X_1}{X_2} \quad . \tag{3.1}$$

In this section, we find ν such that $\frac{\nu\zeta}{E(\zeta)}$ is approximately χ^2 with ν degree of freedom.

To accomplish this, we will need to find the mean and variance of ζ under H_0.

Theorem 3.1 Under H_0, the expected value of the efficient score statistic is

$$E(\zeta) = (t-1) \cdot \left(\frac{n}{n+\tau} \right). \tag{3.2}$$

Proof. For details see Wang (1999).

3.1 An approximation to the variance of ζ

In this section the variance of ζ is estimated by using a first-order Taylor approximation (Section 5.2.3, Mood, Graybill and Boes, 1974). We have

$$\operatorname{var}(\zeta) = \operatorname{var}\left(\frac{X_1}{X_2}\right) \approx \left(\frac{E(X_1)}{E(X_2)}\right)^2 \left(\frac{\operatorname{var}(X_1)}{[E(X_1)]^2} + \frac{\operatorname{var}(X_2)}{[E(X_2)]^2} - \frac{2\operatorname{cov}(X_1, X_2)}{E(X_1)E(X_2)}\right). \tag{3.3}$$

Theorem 3.2 The approximate expression of the (3.3) becomes

$$\begin{aligned}\operatorname{Var}(\zeta) = {} & \frac{n^2}{(n+\tau)^2(\mu+\tau\mu^2)^2}\left(\sum\frac{1}{n_i} - \frac{t^2}{n}\right)\mu + \\ & \frac{n^2}{(n+\tau)^3(\mu+\tau\mu^2)^2}\left\{\left(2nt - 2n - 5t^2\tau - 2\tau t + 7n\tau\sum\frac{1}{n_i} + 7\tau^2\left[\sum\frac{1}{n_i} - \frac{t^2}{n}\right]\right)\mu^2\right. \\ & + \left(4n\tau t - 4\tau n - 4\tau^2 t - 8\tau^2 t^2 + 12\tau^2 n\sum\frac{1}{n_i} + 12\tau^3\left[\sum\frac{1}{n_i} - \frac{t^2}{n}\right]\right)\mu^3 \\ & \left. + \left(2n\tau^2 t - 4\tau^3 t^2 - 2\tau^2 n - 2\tau^3 t + 6\tau^3 n\sum\frac{1}{n_i} + 6\tau^4\left[\sum\frac{1}{n_i} - \frac{t^2}{n}\right]\right)\mu^4\right\}. \end{aligned} \tag{3.4}$$

If $n_1 = n_2 = \cdots = n_t = m$, then (3.4) can be simplified.

Proof. For details see Wang (1999).

Corollary 3.1 Under the assumption of $n_1 = n_2 = \cdots = n_t = m$, (3.4) becomes

$$\operatorname{var}(\zeta) \approx \frac{2(t-1)n^2(n+\tau t)}{(n+\tau)^3}.$$

Proof. Under the assumption of $n_1 = n_2 = \cdots = n_t = m$, we have

$$n = mt \quad \text{and} \quad \sum\frac{1}{n_i} = \frac{t^2}{n}.$$

Therefore, the right side of (3.4) can be simplified as

$$\begin{aligned}& \frac{n^2}{(n+\tau)^3(\mu+\tau\mu^2)^2}\left\{2(t-1)(n+\tau t)\mu^2 + 4\tau(t-1)(n+\tau t)\mu^3 + 2\tau^2(t-1)(n+\tau t)\mu^4\right\} \\ & = \frac{2n^2(t-1)(n+\tau t)}{(n+\tau)^3}.\end{aligned}$$

3.2 A second approximation to the variance of ζ

In this section a better approximation to $\operatorname{var}(\zeta)$ is obtained. This approximation is obtained by using the fact that

$$\operatorname{var}(\zeta) = E(\operatorname{var}(\zeta \mid Y_{\bullet\bullet})) + \operatorname{var}(E(\zeta \mid Y_{\bullet\bullet})) \tag{3.5}$$

(Mood, Graybill and Boes, 1974).

Under H_0, $\text{var}(E(\zeta | Y_{\cdot\cdot})) = 0$, so

$$\text{var}(\zeta) = E(\text{var}(\zeta | Y_{\cdot\cdot})) \tag{3.6}$$

Now expression (3.6) is used to find the value of $\text{var}(\zeta)$.

Theorem 3.3 For $n_i = \frac{n}{t}$ for $i = 1, 2, \ldots, t$, the following is true.

$$\begin{aligned}\text{var}(\zeta) &= E(\text{var}(\zeta | Y_{\cdot\cdot})) \\ &= \frac{2n^3(t-1)(n+t\tau)}{(n+\tau)^2(n+2\tau)(n+3\tau)} - \frac{1}{(n+\tau)^2(n+2\tau)(n+3\tau)} \cdot E\left(\frac{1}{Y_{\cdot\cdot}(n+\tau Y_{\cdot\cdot})^2}\right) \\ &\quad - \frac{\tau}{(n+\tau)(n+2\tau)(n+3\tau)} \cdot E\left(\frac{1}{(n+\tau Y_{\cdot\cdot})^2}\right).\end{aligned} \tag{3.7}$$

Proof. For details see Wang (1999).

Note that the second and third terms on the right side of (3.7) are very small, so we have the following approximate expression

$$\text{var}(\zeta) \approx \frac{2n^3(t-1)(n+t\tau)}{(n+\tau)^2(n+2\tau)(n+3\tau)}. \tag{3.8}$$

Our simulation results showed that the approximations (3.8) and (3.2) are very accurate.

To approximate a critical point for the efficient score statistic we follow Satterthwaite (1946). We find ν such that

$$\frac{\nu\zeta}{E(\zeta)} \sim \chi^2(\nu) \text{ when } H_0 \text{ is true.}$$

This implies that ν must satisfy

$$\text{Var}\left(\nu\zeta / E(\zeta)\right) = 2\nu.$$

Hence, it follows from (3.2) and (3.8) that

$$\nu = \frac{2[E(\zeta)]^2}{\text{var}(\zeta)} = \frac{(t-1)(n+2\tau)(n+3\tau)}{n(n+t\tau)} \text{ when } n_i = \frac{n}{t} \text{ for } i = 1, 2, \ldots, t \tag{3.9}$$

4. Simulations

A Monte Carlo study was performed to compare the efficient score test and its sampling distribution approximated by a chi-square distribution with degrees of freedom equal to ν in (3.9) to several other previously introduced methods of testing for equal means for negative binomial distributions. The methods used were:

1. Efficient score test using the new Satterthwaite approximation to the critical point (RSCR).
2. Efficient score test using $\chi^2(t-1)$ to approximate the critical point (SCOR).

3. Likelihood ratio test (LR).
4. F-test (F) on raw data.
5. F-test with square root transformation (FSQ).
6. F-test with logarithm (of counts +1) transformation (FLG).
7. F-test with inverse hyperbolic transformation (FSH).

Two major criteria in evaluating these test methods are the robustness of

1. The observed significance level ($\hat{\alpha}$) which is the estimated probability of rejecting the null hypothesis when the population means are equal.
2. The observed power ($1-\hat{\beta}$) which is the estimated probability of rejecting the null hypothesis when the population means are not equal.

Empirical significance levels and powers of the seven tests described above are derived for $t = 4$ means based on 3000 samples from the negative binomial distribution for different values of μ_1, μ_2, μ_3 and μ_4, and τ. Data were simulated from negative binomial distributions with means of 0.25, 0.5, 0.75, 1, 2, 5, 10, 15, or 20, and values of τ of 4, 2, 4/3, 1 and 0.2. Only balanced designs were considered, with $n_1 = n_2 = n_3 = n_4 =$ 5, 10, 25 and 50 replications per treatment. Small means, small sample sizes, and large values of τ are emphasized because they are more representative of the situations most frequently encountered in biological studies. The estimated and nominal Type I error rates were compared at the 0.01, 0.05, and 0.1 levels for all tests. Some of the simulation results are presented in Tables 1–3.

Consider the case $\alpha = 0.05$. Simulation results show that for large sample sizes ($n_1 = n_2 = n_3 = n_4 = 50$), all of the test statistics hold their significance levels well even for small means. For moderately large sample sizes $n_1 = n_2 = n_3 = n_4 = 25$, the estimated Type I error rates are generally close to the nominal rates for all the test statistics, but LR tends to be liberal and SCOR and F conservative.

In general, the RSCR, FSQ, and FLG hold their significance levels well. The RSCR performs consistently well across all sample sizes, for different τ's and different μ's. Also, it gives the best performance when sample sizes are small τ is large and μ is small.

The simulation results show that for small sample sizes ($n_1 = n_2 = n_3 = n_4 = 5, 10$) the likelihood ratio chi-squared (LR) test is too liberal and gets worse as μ decreases and τ increases – that is, when the negative binomial distribution departs from the Poisson distribution. When τ is large (e.g. τ =4, 2, 4/3, 1) and sample sizes are small ($n_1 = n_2 = n_3 = n_4 = 5$), the estimated Type I error rate is smaller than the nominal rate for LR.

In general, the estimated Type I error rate is smaller than the nominal rate for F and SCOR. It only gets better when the sample size is large, τ is small (the negative binomial is close to the Poisson distribution), and μ is large.

When $\alpha = 0.01$ and $\alpha = 0.1$, we obtained results similar to those for $\alpha = 0.05$.

Using the formula $\alpha \pm z_{0.005}\sqrt{\alpha(1-\alpha)/3000}$, we were able to identify Type I error rates that are significantly different from α at the 0.01 significance level.

We constructed power curves for all seven statistics (RSCR, SCOR, LR, F, FSQ, FLG, and FSH) with $\tau = 4, 2, 1, 0.2$, $n_1 = n_2 = n_3 = n_4 = 5,\ 10,\ 25$, and some values of μ, with α levels fixed at 0.05. These power curves are plotted as functions of d. Larger values of d correspond to greater differences in the μ_i's. See Figures 1 – 4 for details.

In general, LR and RSCR are more powerful than the ANOVA methods. Among the ANOVA methods, F does not possess the robust properties with respect to the negative binomial distribution for small sample sizes, when τ is large, or when μ is small. The other three ANOVA methods that are based on the transformed data (FSQ, FLG and FSH) hold their significance levels well and their power curves are better than those of F (but not as good as for LR or RSCR). See Figures 1 – 4.

5. Example

The data in Table 4 (McCaughran & Arnold, 1976) refer to counts of embryonic deaths in a control group and two treatment groups ($n_1 = n_2 = n_3 = 10$).

Table 4: Counts of embryonic deaths

	Frequency		
Number of deaths	Control group	Dose Level 1	Dose Level 2
0	7	5	4
1	2	4	2
2	1	0	3
3	0	1	0
4	0	0	1

Suppose it is known from experience that $\tau = 0.25$, so we might assume that

$$Y_{ij} \sim \text{ind } NB(\mu_i, 0.25) \text{ for } i = 1, 2, 3 \text{ and } j = 1, 2, \cdots, 10.$$

To determine if there are differences in the mean counts of deaths among the groups we test the hypothesis

$$H_0 : \mu_1 = \mu_2 = \mu_3 \quad \text{versus} \quad H_1 : \text{at least two differ.}$$

Table 5 below shows the calculated p-values for the seven test statistics previously introduced.

Table 5: List of P – Values for Seven Methods (Example)

Test	RSCR	SCOR	LR	F	FSQ	FLG	FSH
P-value	0.345	0.194	0.192	0.227	0.279	0.257	0.254

Based on the p-values, all the tests indicate that the group means do not differ significantly from one another. This is in agreement with our simulation results.

6. SUMMARY

In general, RSCR and LR are more powerful tests than the other tests. However, while RSCR maintains the type I error rate, LR is too liberal for small values of μ and for small sample sizes (i.e., LR rejects the hypothesis more than it should when the population means are equal.). Hence it certainly will have greater power.

All tests have greater power for small τ than for large τ – that is, for small departures from the Poisson assumption, the tests are more powerful. Also all of the tests have greater power for large values of μ than they do for small values of μ.

Because LR is sometimes too liberal and the RSCR test is much more powerful than all of the other tests, we recommend the use of the RSCR test.

REFERENCES

Anscombe, F.J. (1949). The analysis of insect counts based on the negative binomial distribution. *Biometrics*, 5, 165-173.

Bliss, C.I. & Owen, A.R.G. (1958). Negative binomial distribution with a common k. *Biometrika*, 45, 37-58.

Barnwal, R. K. and Paul, S. R. (1988). Analysis of one-way layout of count data with negative binomial variation. *Biometrika*, 75, 215-222.

Beall, G. (1942), The Transformation from Entomological Field Experiments so that the Analysis of Variance Becomes Applicable. *Biometrika*, 32, 243-262.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models,* 2nd ed, London: Chapman and Hall.

McCaughran, D.A. & Arnold, D.W. (1976) Statistical models for numbers of implantation sites and embryonic deaths in mice. *Toxicol.Appl. Pharmacol.* 38, 325-333.

Mood A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill, Inc.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin 2: 110-14.*

Wang, Y. (1999). The analysis of count data in a one-way layout and a new bivariate negative binomial distribution. Ph.D dissertation, Department of Statistics, Kansas State University.

Young, L. J., Campbell, N. L., and Capuano, G. A. (1999). Analysis of overdispersed count data. *Journal of Agricultural, biological, and Environmental Statistics, volume 4, Number 3,* 258-275.

Table 1: Observed Type I Error Rates based on 3000 replications
Nominal level: $\alpha = 0.05$

$n_1=n_2$ $n_3=n_4$	τ	μ	RSCR	SCOR	LR	F	FSQ	FLG	FSH
5	4	0.25	0.032*	0.006*	0.012*	0.021*	0.027*	0.026*	0.027*
		0.50	0.042	0.012*	0.035*	0.028*	0.042	0.040*	0.043
		0.75	0.050	0.017*	0.053	0.027*	0.045	0.043	0.045
		1	0.055	0.020*	0.060*	0.030*	0.047	0.047	0.048
		2	0.051	0.016*	0.082*	0.030*	0.043	0.046	0.049
		5	0.053	0.018*	0.090*	0.028*	0.039*	0.046	0.049
		10	0.060	0.019*	0.089*	0.032*	0.040*	0.049	0.050
		15	0.054	0.019*	0.086*	0.032*	0.047	0.054	0.054
		20	0.051	0.020*	0.078*	0.028*	0.036*	0.045	0.046
	2	0.25	0.045	0.021*	0.028*	0.025*	0.034*	0.032*	0.034*
		0.50	0.057	0.036*	0.063	0.039*	0.050	0.047	0.049
		0.75	0.050	0.027*	0.062	0.033*	0.051	0.049	0.050
		1	0.042	0.028*	0.073	0.031*	0.043	0.040	0.043
		2	0.046	0.026*	0.078	0.037*	0.055	0.055	0.056
		5	0.049	0.031*	0.060	0.039*	0.051	0.055	0.056
		10	0.056	0.035*	0.067	0.036*	0.050	0.055	0.053
		15	0.047	0.028*	0.060	0.031*	0.047	0.053	0.052
		20	0.039*	0.023*	0.060	0.037*	0.048	0.050	0.050
5	1	0.25	0.033*	0.032*	0.034*	0.027*	0.038*	0.036*	0.036*
		0.50	0.041	0.034*	0.067*	0.041	0.052	0.049	0.050
		0.75	0.047	0.035*	0.079*	0.038*	0.049	0.046	0.047
		1	0.047	0.036*	0.072*	0.047	0.054	0.054	0.055
		2	0.046	0.037*	0.058	0.033*	0.045	0.046	0.046
		5	0.050	0.037*	0.060*	0.041	0.051	0.050	0.049
		10	0.045	0.033*	0.055	0.036*	0.039*	0.041	0.040*
		15	0.056	0.046	0.066*	0.048	0.056	0.056	0.055
		20	0.041	0.031*	0.052	0.037*	0.043	0.045	0.045
	0.2	0.25	0.038*	0.038*	0.053	0.024*	0.031*	0.030*	0.030*
		0.50	0.043	0.041	0.073*	0.040*	0.050	0.048	0.047
		0.75	0.051	0.047	0.080*	0.045	0.055	0.054	0.055
		1	0.048	0.044	0.070*	0.045	0.056	0.056	0.053
		2	0.045	0.041	0.053	0.040*	0.048	0.049	0.047
		5	0.056	0.053	0.061*	0.053	0.050	0.049	0.050
		10	0.050	0.046	0.052	0.050	0.052	0.049	0.052
		15	0.046	0.044	0.049	0.040*	0.047	0.047	0.048
		20	0.047	0.045	0.051	0.047	0.050	0.050	0.050

*Indicates empirically significantly different from the nominal level $\alpha = 0.05$. The 99% confidence interval is (0.04, 0.06)

Table 2: Observed Type I Error Rates based on 3000 replications
Nominal level: $\alpha = 0.05$

$n_1=n_2$ $n_3=n_4$	τ	μ	RSCR	SCOR	LR	F	FSQ	FLG	FSH
10	4	0.25	0.040*	0.020*	0.049	0.027*	0.036*	0.036*	0.035*
		0.50	0.050	0.032*	0.075*	0.038*	0.047	0.046	0.048
		0.75	0.045	0.024*	0.078*	0.033*	0.045	0.044	0.046
		1	0.040*	0.022*	0.066*	0.025*	0.040*	0.039*	0.042
		2	0.052	0.033*	0.070*	0.030*	0.046	0.047	0.045
		5	0.046	0.032*	0.068*	0.034*	0.047	0.053	0.055
		10	0.050	0.030*	0.062*	0.037*	0.048	0.053	0.053
		15	0.048	0.033*	0.062*	0.036*	0.047	0.047	0.048
		20	0.041	0.024*	0.058	0.030*	0.044	0.053	0.050
	2	0.25	0.045	0.038*	0.064*	0.033*	0.041	0.037*	0.038*
		0.50	0.042	0.034*	0.067*	0.038*	0.048	0.046	0.047
		0.75	0.043	0.034*	0.066*	0.038*	0.048	0.048	0.047
		1	0.045	0.034*	0.063*	0.037*	0.051	0.050	0.051
		2	0.051	0.041	0.056	0.033*	0.043	0.046	0.046
		5	0.049	0.038*	0.055	0.042	0.047	0.047	0.046
		10	0.045	0.035*	0.052	0.039*	0.046	0.049	0.050
		15	0.046	0.039*	0.060*	0.038*	0.051	0.050	0.052
		20	0.043	0.033*	0.057	0.029*	0.044	0.044	0.045
10	1	0.25	0.045	0.034*	0.084*	0.036*	0.040*	0.039*	0.039*
		0.50	0.050	0.043	0.067*	0.039*	0.043	0.043	0.044
		0.75	0.041	0.036*	0.055	0.038*	0.043	0.043	0.042
		1	0.050	0.046	0.063*	0.042	0.052	0.052	0.052
		2	0.043	0.040*	0.049	0.040*	0.049	0.051	0.050
		5	0.049	0.044	0.056	0.044	0.052	0.050	0.050
		10	0.048	0.042	0.055	0.042	0.051	0.051	0.051
		15	0.050	0.047	0.057	0.045	0.049	0.049	0.049
		20	0.058	0.054	0.061*	0.050	0.055	0.054	0.055
	0.2	0.25	0.047	0.047	0.077*	0.042	0.045	0.044	0.044
		0.50	0.048	0.047	0.067*	0.044	0.048	0.045	0.045
		0.75	0.050	0.048	0.060*	0.047	0.056	0.053	0.053
		1	0.051	0.050	0.059	0.047	0.057	0.053	0.052
		2	0.058	0.057	0.057	0.055	0.046	0.048	0.051
		5	0.054	0.054	0.053	0.051	0.052	0.054	0.053
		10	0.050	0.049	0.052	0.049	0.053	0.057	0.057
		15	0.048	0.047	0.049	0.046	0.044	0.042	0.042
		20	0.048	0.048	0.048	0.053	0.053	0.050	0.049

*Indicates empirically significantly different from the nominal level $\alpha = 0.05$. The 99% confidence interval is (0.04, 0.06)

Table 3: Observed Type I Error Rates based on 3000 replications
Nominal level: $\alpha = 0.05$

$n_1=n_2$ $n_3=n_4$	τ	μ	RSCR	SCOR	LR	F	FSQ	FLG	FSH
25	4	0.25	0.051	0.042	0.067*0.	0.041	0.048	0.047	0.048
		0.50	0.045	0.035*	0.055	0.035*	0.043	0.043	0.044
		0.75	0.053	0.042	0.062*	0.044	0.048	0.049	0.046
		1	0.050	0.040*	0.058	0.039*	0.045	0.046	0.048
		2	0.050	0.041	0.062*	0.042	0.055	0.053	0.055
		5	0.050	0.042	0.057	0.036*	0.043	0.045	0.046
		10	0.047	0.040*	0.054	0.043	0.048	0.052	0.055
		15	0.044	0.035*	0.048	0.038*	0.048	0.055	0.054
		20	0.052	0.041	0.062*	0.046	0.052	0.052	0.052
	2	0.25	0.051	0.043	0.063*	0.045	0.047	0.047	0.047
		0.50	0.046	0.043	0.053	0.038*	0.044	0.041	0.042
		0.75	0.046	0.042	0.051	0.038*	0.045	0.043	0.044
		1	0.051	0.047	0.059	0.042	0.048	0.049	0.049
		2	0.046	0.041	0.050	0.043	0.049	0.051	0.050
		5	0.048	0.044	0.054	0.043	0.050	0.051	0.051
		10	0.049	0.047	0.056	0.045	0.048	0.050	0.051
		15	0.054	0.050	0.059	0.048	0.055	0.056	0.054
		20	0.055	0.049	0.061*	0.047	0.048	0.053	0.052
	1	0.25	0.044	0.043	0.056	0.046	0.047	0.046	0.046
		0.50	0.047	0.047	0.055	0.048	0.051	0.052	0.053
		0.75	0.051	0.051	0.054	0.052	0.052	0.049	0.049
		1	0.052	0.051	0.055	0.053	0.055	0.055	0.055
		2	0.049	0.049	0.052	0.054	0.050	0.053	0.053
		5	0.052	0.051	0.052	0.049	0.055	0.052	0.052
		10	0.051	0.050	0.050	0.053	0.048	0.049	0.049
		15	0.049	0.049	0.049	0.043	0.048	0.048	0.050
		20	0.048	0.048	0.049	0.049	0.046	0.046	0.046
50	4	0.25	0.046	0.041	0.055	0.041	0.044	0.045	0.044
		0.50	0.050	0.046	0.052	0.046	0.050	0.049	0.051
		0.75	0.043	0.040	0.049	0.043	0.050	0.049	0.049
		1	0.052	0.045	0.050	0.044	0.049	0.048	0.048
		2	0.051	0.047	0.057	0.045	0.049	0.048	0.049
		5	0.051	0.046	0.053	0.038	0.046	0.046	0.045
		10	0.044	0.039	0.044	0.043	0.054	0.054	0.053
		15	0.052	0.047	0.055	0.042	0.050	0.050	0.050
		20	0.047	0.043	0.055	0.044	0.051	0.048	0.046

*Indicates empirically significantly different from the nominal level $\alpha = 0.05$. The 99% confidence interval is (0.04, 0.06)

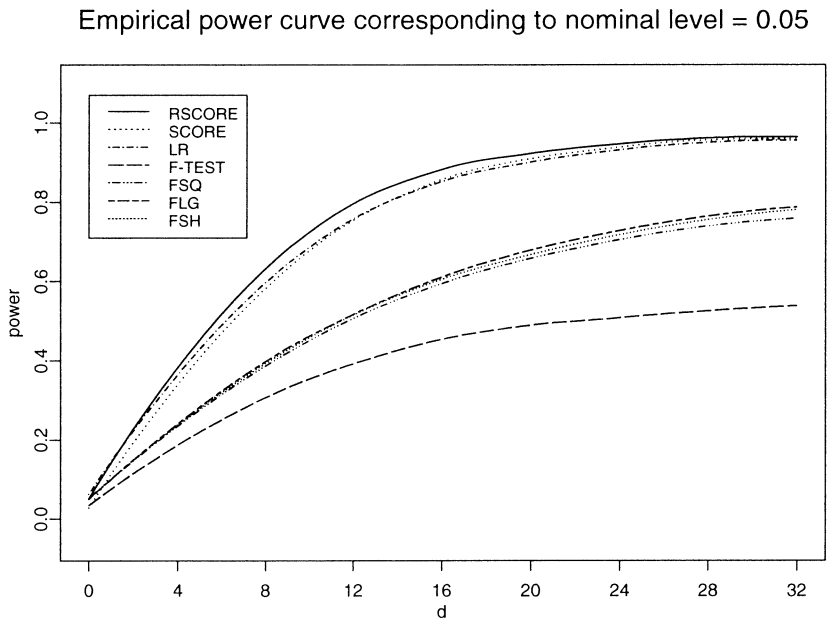

Figure 1 3000 replications; sample sizes $n_1 = n_2 = n_3 = n_4 = 5$; $\tau = 2$; $\mu_1 = \mu_2 = \mu_3 = 0.75$; $\mu_4 = \mu_1(1+d)$, $d > 0$

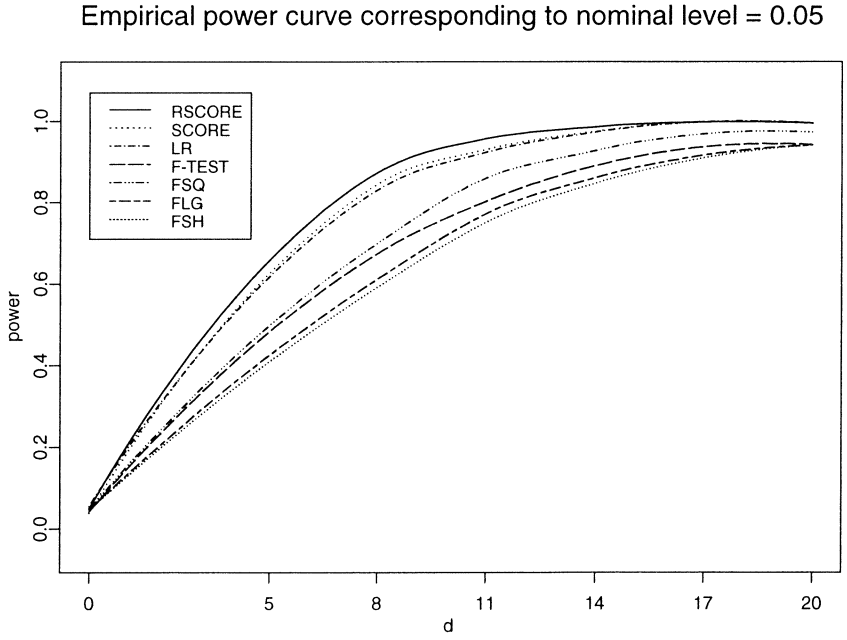

Figure 2 3000 replications; sample sizes $n_1 = n_2 = n_3 = n_4 = 10$; $\tau = 2$; $\mu_1 = \mu_2 = \mu_3 = 5$; $\mu_4 = \mu_1(1+d)$, $d > 0$

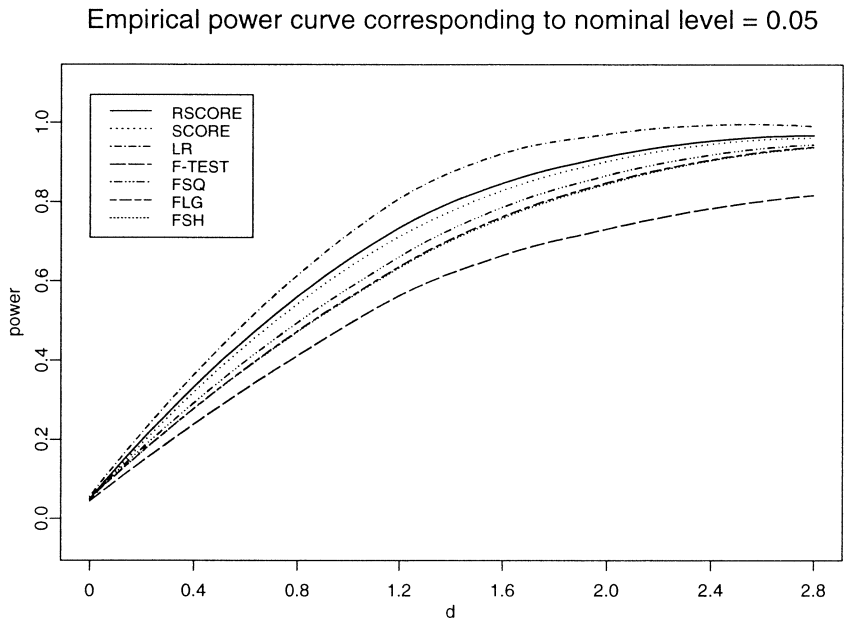

Figure 3 3000 replications; sample sizes $n_1 = n_2 = n_3 = n_4 = 10$; $\tau = 1$; $\mu_1 = 5$, $\mu_2 = \mu_1(1+d)$, $\mu_3 = \mu_1(1+2d)$, $\mu_4 = \mu_1(1+3d)$, $d > 0$

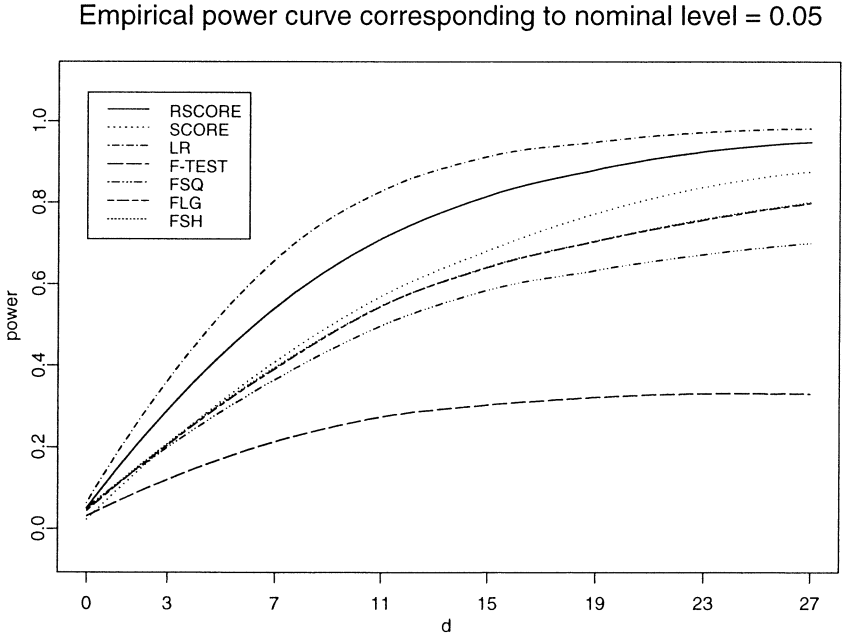

Figure 4 3000 replications; sample sizes $n_1 = n_2 = n_3 = n_4 = 5$; $\tau = 2$; $\mu_1 = \mu_2 = 0.75$; $\mu_3 = \mu_4 = \mu_1(1+d)$, $d > 0$