



Anscombe's Tests of Fit for the Negative Binomial Distribution

D. J. Best, *School of Mathematical and Physical Sciences,
University of Newcastle, NSW 2308, Australia.*
Email: John.Best@newcastle.edu.au

J. C. W. Rayner, *School of Mathematical and Physical Sciences,
University of Newcastle, NSW 2308, Australia.*
Email: John.Rayner@newcastle.edu.au

O. Thas, *Department of Applied Mathematics, Biometrics and
Process Control, Ghent University, B-9000 Gent, Belgium.*
Email: olivier.thas@UGent.be

Received: February 27, 2008 Revised: December 05, 2008

Abstract

The negative binomial model is an important and flexible two parameter distribution that models data from many application areas. Here we re-examine tests of fit for the negative binomial distribution that were introduced by Anscombe (1950); they are based on a dispersion statistic U and a third moment statistic T . Small sample power calculations are given for U and T . We are not aware that such powers have been given previously. We show Anscombe's tests are smooth tests in the sense of Rayner and Best (1989). Comparisons are made with an empirical probability generating function test suggested by Meintanis (2005). We suggest U not be used and that decisions on the fit of data to the negative binomial be made using bootstrap p -values rather than comparison with standard errors as suggested by Anscombe (1950). We show that tests based on a fourth moment component of a smooth test statistic have good power.

AMS Subject Classification: 62F03; 62G07.

Key-words: Bootstrap p -values; Empirical probability generating function test; Fourth moment based tests; Power study; Smooth tests.

1. Introduction

The negative binomial distribution describes many distributions of counts when the variance of the counts is greater than the mean. See, for example, Johnson, Kemp and Kotz

(2005, Chapter 5, section 10) for many references. One important application is the development of pest management sampling schemes for grain crops. Krebs (1998) describes such sampling schemes assuming a negative binomial distribution applies. Of course, when the mean and variance are about equal the Poisson model may be more appropriate, and when the mean of the counts is greater than the variance the binomial model may apply.

The form of the probability function of the negative binomial has been known since the early 1700s in connection with the number of coin tosses to achieve a fixed number of heads. If X is the random variable representing the number of independent trials necessary to obtain k occurrences of an event that has constant probability, p , of occurring at each trial, then

$$P(X = x) = \binom{k+x-1}{k-1} p^k q^x \text{ for } x = 0, 1, 2, \dots, \text{ in which } q = 1 - p \text{ and } 0 < p < 1.$$

These probabilities can be linked to a binomial expansion with a negative index and hence the name 'negative binomial'.

Another common genesis of the negative binomial follows from Poisson counts having a random mean. The probability of a count of x conditional on the value of the parameter λ being known is

$$P(X = x | \lambda) = e^{-\lambda} \lambda^x / x! \text{ for } x = 0, 1, 2, \dots, \text{ in which } \lambda > 0.$$

If we now assume the Poisson mean is in fact a random variable Λ with probability density function

$$f(\lambda) = \lambda^{\alpha-1} e^{-\lambda/\beta} / \{\beta^\alpha \Gamma(\alpha)\} \text{ for } 0 < \lambda < \infty, \text{ zero otherwise, in which } \alpha > 0 \text{ and } \beta > 0,$$

it then follows that the marginal probability function of X is negative binomial, since for $x = 0, 1, 2, \dots$,

$$\begin{aligned} P(X = x) &= \int_0^\infty \lambda^{\alpha-1} e^{-\lambda/\beta} (\lambda^x e^{-\lambda} / x!) d\lambda / \{\beta^\alpha \Gamma(\alpha)\} \\ &= \frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!} \left(\frac{\beta}{\beta+1}\right)^x \left(\frac{1}{\beta+1}\right)^\alpha. \end{aligned}$$

In the following we will define the negative binomial probability function π_x by

$$\pi_x = \frac{\Gamma(k+x)}{\Gamma(k)x!} p^k q^x \text{ for } x = 0, 1, 2, \dots, \text{ with } q = 1 - p, 0 < p < 1, \text{ and } k > 0.$$

It is common practice to test for goodness of fit of the negative binomial for a set of n data points by using the test statistic

$$X^2 = \sum_{j=0}^v (N_j - E[N_j])^2 / E[N_j]$$

in which N_j and $E[N_j]$ are respectively the count and expectation under the null hypothesis for the j th category, $j = 0, 1, \dots, v$, $N_0 + \dots + N_v = n$, and in which the number of classes, $v + 1$, is usually chosen so that the smallest class expectation is commonly 1, 5 or 10. This may involve pooling counts in either or both the upper and lower tails. To counter losing information due to this pooling, Anscombe (1950) suggested using tests based on the statistics

U and T which we define in the next section. Here we show Anscombe's tests are just special cases of smooth tests of fit for the negative binomial. Section 3 gives two examples while section 4 considers new tests based on the fourth moment. Section 5 gives a small sample size and power study for Poisson mixture, zero inflated Poisson and Neyman Type A alternatives. Section 6 is a brief conclusion. We are not aware of previous power studies for U and T although both statistics are still used in ecological and other applications. See, for example, Krebs (1998) or Jarvis (1989).

2. Definitions and estimation

Subsequently we confine ourselves to the common case where both k and $p = 1 - q$ are unknown. For the negative binomial the method of moments (MOM) estimators equate sample and population means and variances respectively, and so, for a random sample X_1, \dots, X_n , satisfy

$$\bar{X} = \tilde{k}\tilde{q}/\tilde{p} \quad \text{and} \quad S^2 = \tilde{k}\tilde{q}/\tilde{p}^2$$

in which $\bar{X} = \sum_{j=1}^v jN_j/n = \sum_{j=1}^n X_j/n$ and $nS^2 = \sum_{j=1}^v j^2N_j - n\bar{X}^2 = \sum_{j=1}^n X_j^2 - n\bar{X}^2$,

where now for an observed data set x_1, \dots, x_n , $v = \max(x_1, \dots, x_n)$. This gives MOM estimators

$$\tilde{k} = \bar{X}^2/(S^2 - \bar{X}) \quad \text{and} \quad \tilde{p} = \tilde{k}/(\tilde{k} + \bar{X}).$$

The maximum likelihood (ML) estimators \hat{k} and \hat{p} satisfy

$$\bar{X} = \hat{k}\hat{q}/\hat{p} \quad \text{and} \quad n \log \hat{p} = - \sum_{j=1}^v \left\{ (\hat{k} + j - 1)^{-1} \sum_{i=j}^v N_i \right\}.$$

Note that the left side of equation (5.48) of Johnson, Kemp and Kotz (2005, p. 223) is missing the multiplier n in the equation above giving $\log \hat{p}$. Iterative methods are needed to solve these equations. Iterative methods are also needed for the zero frequency and mean estimators suggested by Anscombe (1950).

To emphasise the link between U, T and the smooth tests, we will now define U and T in terms of the smooth tests. Rayner and Best (1989) and Rayner, Thas and Best (2009) give further information on smooth tests. To construct smooth test statistics we need polynomials $\{g_r(x)\}$ orthogonal on $\{\pi_x\}$. These can be constructed using a recurrence relation given by Lancaster (1975) and reported in Rayner and Best (1989, A.2.1). For the negative binomial case suppose $g_{-1}(x) = 0$ and $g_0(x) = 1$ for all x ; then it may be shown that for $r = 0, 1, 2, \dots$

$$g_{r+1}(t) = \{t - r(1 + q)/p\}g_r(t) - (rkq/p^2)\{1 + (r - 1)/k\}g_{r-1}(t)$$

where $t = (x - kq/p)$. Thus

$$g_1(t) = t, g_2(t) = t^2 - \{(q + 1)/p\}t - kq/p^2$$

and

$$g_3(t) = t^3 + (3 - 6/p)t^2 + (2 + 3k/p - 6/p - 3k/p^2 + 6/p^2)t + 2kq(1+q)/p^3.$$

The r th smooth test statistic is

$$V_r = \sum_{j=1}^n g_r(X_j) / \sqrt{nd_r} \quad \text{for } r = 1, 2, \dots,$$

where the d_r are normalising constants. Appendix A gives d_1, d_2, d_3 and d_4 . Write \tilde{V}_r for V_r with the unknown parameters estimated by MOM and write V_r^* to indicate the parameters are estimated by the zero frequency and mean approach. Similarly μ_r^* and $\tilde{\mu}_r$ denote r th population moments estimated under these same estimation methods. Then

$$U = V_2^* \sqrt{d_2^*/n} \quad \text{and} \quad T = \tilde{V}_3 \sqrt{\tilde{d}_3/n}$$

in which $U = \sum_j (X_j - \bar{X})^2 / n - \mu_2^*$ and $T = \sum_j (X_j - \bar{X})^3 / n - \tilde{\mu}_3$. Notice that d_2 and d_3 are not needed to define U and T , and that $U \equiv 0$ if MOM estimators are used. We also note that Anscombe (1950) gave

$$n \text{var}(U) = 2k(k+1)(q^2/p^4) \{1 - q^2 / (-q - \log p)\} + (q/p)^4 \{(p^{-k} - 1 - kq) / (-q - \log p)^2\}$$

and

$$n \text{var}(T) = n\sigma_T^2 = 2k(k+1)q^3(10 + 3k - 4p)/p^6.$$

Writing $\rho = q/p$, Meintanis (2005) suggested testing the negative binomial assumption using the statistic

$$M = n^{-1} \left[\bar{X}^2 \sum_{j=1}^n \sum_{k=1}^n I(X_{jk}^+ + 5) - 2\bar{X} \sum_{j=1}^n \sum_{k=1}^n X_j \{(1 + \tilde{\rho})I(X_{jk}^+ + 4) - \tilde{\rho}I(X_{jk}^+ + 5)\} \right. \\ \left. + \sum_{j=1}^n \sum_{k=1}^n X_j X_k \{(1 + \tilde{\rho}^2)I(X_{jk}^+ + 3) \tilde{\rho}^2 I(X_{jk}^+ + 5) 2\tilde{\rho}(1 + \tilde{\rho})I(X_{jk}^+ + 4)\} \right]$$

with $X_{jk}^+ = X_j + X_k$ and $I(\gamma) = (1 + \gamma)^{-1}$, $\gamma > -1$.

3. Two examples

Black bean aphid data. Krebs (1998, p.130) gives counts of black bean aphid (*Aphis fabae*) on bean stems shown in Table 3.1.

We find U takes the value -1.392 with $\sqrt{\text{var}(U)} = 2.303$, while T takes the value -10.812 with $\sqrt{\text{var}(T)} = 10.430$. Assuming $U/\sqrt{\text{var}(U)}$ and $T/\sqrt{\text{var}(T)}$ have approximate standard normal distributions neither U nor T appears to be significant. However it is shown in section 5 that the $N(0, 1)$ approximations are poor and so we recommend finding p -values using the parametric bootstrap approach described, for example, in Gürtler and Henze (2000). Some detail of this approach is given in Appendix B. For this example

Table 3.1. Counts of black bean aphids.

# aphids per stem	0	1	2	3	4	5	6	7	8	9
# stems	6	8	9	6	6	2	5	3	1	4

Table 3.2. Counts of bacterial cells from a milk smear.

# cells per square	0	1	2	3	4	5	6	7	8	9	10	19
frequency	56	104	80	62	42	27	9	9	5	3	2	1

the bootstrap p -values for U and T are 0.15 and 0.16. Notice that Krebs (1998, p.135) incorrectly calculates both $\text{var}(T)$ and $\text{var}(U)$ using ML estimators whereas we have used the zero frequency/mean estimators for U and the MOM estimators for T . We consider this example further at the end of section 4.

Bacterial Cells data. Jarvis (1989, p.37) gives the number of bacterial cells per microscope square from a milk smear. The counts are shown in Table 3.2.

Jarvis (1989, p.58) tests the fit of a negative binomial using a traditional chi squared (X^2) statistic approach where the frequencies for 8 or more cells per square were pooled. He obtained $X^2 = 5.38$ on 6 degrees of freedom and concluded that the negative binomial is a good model for the data. Of course, this pooling effectively ignores the one large count of 19 and if either U or T is calculated with this large count included then the negative binomial model would no longer be considered satisfactory. The microbiologist faced with such a large count may be advised, for example, to check the experimental protocol. Rayner and Best (1989) and Rayner, Thas and Best (2009) show that tests based on X^2 are often not as powerful as smooth tests or tests based on smooth test components such as U and T .

We also consider this example further at the end of section 4.

It is possible, using either ML or MOM, that the estimator of k is negative. In such cases we suggest that the negative binomial model is not suitable and that no test of fit is needed.

4. New tests based on the fourth moment

In this section we argue for testing the fit of the negative binomial distribution with Anscombe's T rather than U . However T alone will have poor power for alternatives with third moments close to that of the negative binomial. See, for example, the powers in Best and Rayner (2003), who consider testing fit for the geometric distribution when the alternative has similar variance to that of the geometric. Further, fourth moment based tests, such as the well-known kurtosis test of normality, have proved useful in testing for other distributions. For these reasons we now derive and examine some fourth moment based tests for the negative binomial.

Let W , our first proposed fourth moment test statistic, be given by

$$\begin{aligned}
 W &= m_4 - E[m_4], \text{ or} \\
 W &= m_4 - \tilde{k}\tilde{q}/\tilde{p}^2 \{6/\tilde{p}^2 - 6/\tilde{p} + 1\} - 3\tilde{k}^2\tilde{q}^2/\tilde{p}^4 \\
 &= m_4 - 6m_2^3/(m_1')^2 + 6m_2^2/m_1' - m_2 - 3m_2^2,
 \end{aligned}$$

where m_r is the r th central sample moment and m_1' is the sample mean. We see that W is of

the same form as $T = m_3 - \tilde{k}\tilde{q}(1 + \tilde{q})/\tilde{p}^3$. If we put $x = m'_1, y = m_2$ and $z = m_4$ then

$$w = f(x, y, z) = z - 6y^3/x^2 + 6y^2/x - y - 3y^2$$

and

$$\begin{aligned} \text{var}(W) &\approx \left(\frac{\partial f}{\partial x}\right)^2 \text{var}(X) + \left(\frac{\partial f}{\partial y}\right)^2 \text{var}(Y) + \left(\frac{\partial f}{\partial z}\right)^2 \text{var}(Z) \\ &+ 2\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right)\text{cov}(X, Y) + 2\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial z}\right)\text{cov}(X, Z) \\ &+ 2\left(\frac{\partial f}{\partial y}\right)\left(\frac{\partial f}{\partial z}\right)\text{cov}(Y, Z) \end{aligned}$$

using the so-called delta method formula given, for example, in Stuart and Ord (2005, p.350, equation 10.12). From Stuart and Ord (2005, Chapter 10) we also know that to order n^{-1}

$$\begin{aligned} \text{var}(X) &= \mu_2/n, \quad \text{var}(Y) = (\mu_4 - \mu_2^2)/n, \quad \text{var}(Z) = (\mu_8 - \mu_4^2 - 8\mu_3\mu_5 + 16\mu_2\mu_3^2)/n, \\ \text{cov}(X, Y) &= \mu_3/n, \quad \text{cov}(X, Z) = (\mu_5 - 4\mu_2\mu_3)/n, \quad \text{and} \quad \text{cov}(Y, Z) = (\mu_6 - \mu_2\mu_4 - 4\mu_3^2)/n. \end{aligned}$$

It follows that, to order n^{-1} ,

$$n \text{var}(W) = 24k(k+1)q^3\{105 + 41k + k^2 - 123p - 41kp - k^2p + 39p + 9kp^2 - 3p^3\}/p^8.$$

Subsequently we use $W^2/\text{var}(W)$ with parameters estimated by MOM as a test of fit. In section 5 size and power properties of this test are assessed.

Suppose we now define a statistic R in an analogous manner to U and T :

$$R = \tilde{V}_4 \sqrt{\tilde{d}_4/n}.$$

This results in a statistic involving m_4 as does W and m_3 as does T . Using the recurrence relation introduced previously we find

$$\begin{aligned} g_4(t) &= t^4 + (6 - 12/p)t^3 + (6k/p - 6k/p^2 + 36/p^2 - 36/p + 11)t^2 \\ &+ (14k/p - 42k/p^2 + 28k/p^3 - 24/p^3 + 36/p^2 - 24/p + 6)t \\ &+ (3k^2/p^4 - 6k^2/p^3 + 3k^2/p^2 - 18k/p^4 + 36k/p^3 - 24k/p^2 + 6k/p). \end{aligned}$$

As above, notice that d_4 is not needed to calculate R . We find

$$R = m_4 + (6 - 12/\tilde{p})m_3 - \tilde{k}\tilde{q}(3\tilde{k}\tilde{q} - 5\tilde{p}^2 - 18\tilde{q})/\tilde{p}^4.$$

Also, on using the delta method again, we find

$$\begin{aligned} n \text{var}(R) &= n\sigma_R^2 = 24k(k+1)q^4(3p^2 - 6p + k^2 + 5k + 9)/p^8, \quad \text{and} \\ n \text{cov}(T, R) &= n\sigma_{TR} = -24k(k+1)q^5/p^7. \end{aligned}$$

Now a smooth test statistic, \tilde{S}_2 say, can be given as $(R, T)L^{-1}(R, T)^T$ where L is the asymptotic covariance matrix of $(R, T)^T$. The statistic \tilde{S}_2 should be approximately χ_2^2 distributed. Note that \tilde{S}_2 here is a quadratic form with diagonal elements not one, whereas \tilde{S}_2 as discussed in Rayner and Best (1989), is an unweighted sum of squares.

To again have \tilde{S}_2 as a sum of squares we can use the Cholesky decomposition of the inverse of L . The Cholesky components of \tilde{S}_2 are $T^2/\text{var}(T)$ and, if $\Delta = \sqrt{\sigma_T^2\sigma_R^2 - \sigma_{TR}^2}$, $C = \{\sigma_TR/\Delta - \sigma_{RT}T/(\sigma_T\Delta)\}^2$. An important property of C is that it is asymptotically independent of $T^2/\text{var}(T)$, whereas R and T are *not* asymptotically independent. If T is not significant and C is significant this suggests that the fourth moment differences between the data and the fitted negative binomial distribution are important.

We have not considered new fourth moment tests based on ML estimators. Our experience when testing goodness of fit for other distributions where ML and MOM estimators differ is that there is little difference in the power of the ML and MOM based tests but the MOM based tests have far better interpretability. See, for example, Thas and Rayner (2005) and Meintanis (2005).

In summary, the smooth test based on $T^2/\text{var}(T)$ assesses third moment differences between the data and the fitted negative binomial distribution after the two negative binomial parameters have been fitted by MOM estimation, so the first two moments of the data and the negative binomial distribution agree. The test based on C assesses fourth moment effects orthogonal to those assessed by $T^2/\text{var}(T)$. The test based on $R^2/\text{var}(R)$ is not (asymptotically) independent of the test based on C .

For the aphid data $C = 0.704$ with p -value 0.25 while for the milk bacteria data $C = 5.140$ with p -value 0.006. The significance of both T and C suggests third and fourth moment departures of the data from the negative binomial. With the lone count of 19 deleted from the milk bacteria data $T = -0.045$, $\sqrt{\text{var}(T)} = 1.376$ with p -value 0.97 while $C = 0.925$ with p -value 0.20. The conclusions this analysis entails are consistent with the previous findings.

5. Size and power study

Table 5.1 examines whether or not the 95% critical values of $U^2/\text{var}(U)$ and $T^2/\text{var}(T)$ approach the χ_1^2 value of 3.841. It appears that both do approach 3.841 slowly, and that $T^2/\text{var}(T)$ does so more slowly than $U^2/\text{var}(U)$. We have noted above that $U \equiv 0$ if MOM estimators are used. This suggests U measures the difference between MOM and other estimates as well as deviation from a negative binomial model and it is not clear which of these possibilities is being examined.

Table 5.2 simulates, for various sample sizes, the 95% critical value of $W^2/\text{var}(W)$ and $R^2/\text{var}(R)$ with $k = \tilde{k}$ and $p = \tilde{p}$. As with $T^2/\text{var}(T)$, we see the convergence is extremely slow. Thus we suggest p -values for all the tests considered here should be found by using the parametric bootstrap. Our recommendation in regard to $T^2/\text{var}(T)$ is in conflict with the original suggestion of Anscombe (1950), and more recently in texts such as Jarvis (1989) and Krebs (1998). Although not shown here, a size study for various $(k, p), n = 100$ and nominal size 0.05 verified that actual test sizes calculated using the parametric bootstrap were close to this nominal size.

Tables 5.3 to 5.5 compare parametric bootstrap powers of the tests based on T, U and M , along with the other tests introduced in section 4. The parametric bootstrap method used 1,000 Monte Carlo samples in both the inner and outer loops. For the zero inflated alternatives of Table 5.5 we took a proportion ω of the zeros and proportion $(1 - \omega)$ of the Poisson (λ) values.

Table 5.1. Critical values of $U^2/\text{var}(U)$ and $T^2/\text{var}(T)$ for $\alpha = 0.05$, n as shown and 10,000 Monte Carlo simulations when $p = 2/3$ and $k = 2$.

n	$U^2/\text{var}(U)$	$T^2/\text{var}(T)$
20	2.42	1.07
50	2.83	1.37
100	3.41	1.71
500	3.63	2.71
1,000	3.70	3.05
10,000	3.74	3.69
100,000	3.76	3.75
∞	3.84	3.84

Table 5.2. Critical values of $W^2/\text{var}(W)$, $R^2/\text{var}(R)$ and C for $\alpha = 0.05$, n as shown and 10,000 Monte Carlo simulations when $p = 2/3$ and $k = 2$.

n	$W^2/\text{var}(W)$	$R^2/\text{var}(R)$	C
20	0.30	2.06	2.02
50	0.47	2.10	2.06
100	0.68	2.10	2.08
500	1.65	2.18	2.11
1,000	2.09	2.26	2.16
10,000	3.23	3.22	3.20
100,000	3.64	3.76	3.76
∞	3.84	3.84	3.84

Some powers for the usual Pearson X^2 test using MOM estimation and number of classes chosen so that all expectations are greater than five are given in Tables 5.3 to 5.5. These powers, and those for the tests based on $U^2/\text{var}(U)$ and $W^2/\text{var}(W)$, are not competitive.

The tests based on \tilde{S}_2 and $R^2/\text{var}(R)$ assesses third and fourth moment departures from the negative binomial, while the test based on C assesses fourth moment effects beyond those assessed by $R^2/\text{var}(R)$. These three tests had superior power in Tables 5.3 to 5.5. The tests based on M and $T^2/\text{var}(T)$ were less satisfactory.

6. Conclusion

Anscombe's statistics U and T have been considered. His test based on T has reasonable power for testing goodness of fit of the negative binomial but the test based on U does not. Moreover we suggest that parametric bootstrap p -values should be used rather than comparison of test statistics with their standard errors. We have noted that Anscombe's tests of fit for the negative binomial are special cases of smooth tests of fit and we have given extensions of Anscombe's tests involving fourth order smooth tests. In our small simulation study, apart from the test based on $W^2/\text{var}(W)$, these extensions provide powerful tests of fit for the negative binomial.

Table 5.3. Powers of various tests of fit for the negative binomial for 0.5 Poisson (λ_1) + 0.5 Poisson (λ_2) alternatives when $n = 100$ and size $\alpha = 0.05$.

λ_1	λ_2	M	$U^2/\text{var}(U)$	$T^2/\text{var}(T)$	$W^2/\text{var}(W)$	$R^2/\text{var}(R)$	C	\tilde{S}_2	X^2
0.3	3.7	0.89	0.69	0.85	0.55	0.92	0.92	0.90	0.64
0.5	3.5	0.51	0.32	0.57	0.29	0.66	0.65	0.62	0.30
0.7	3.3	0.18	0.11	0.28	0.10	0.31	0.31	0.29	0.12

Table 5.4. Powers of various tests of fit for the negative binomial for Neyman Type $A(\lambda_1, \lambda_2)$ alternatives when $n = 100$ and size $\alpha = 0.05$.

λ_1	λ_2	M	$U^2/\text{var}(U)$	$T^2/\text{var}(T)$	$W^2/\text{var}(W)$	$R^2/\text{var}(R)$	C	\tilde{S}_2	X^2
2.0	1.0	0.13	0.05	0.15	0.05	0.20	0.19	0.19	0.11
1.5	1.5	0.31	0.21	0.23	0.08	0.35	0.34	0.31	0.24
1.0	2.0	0.51	0.35	0.36	0.12	0.48	0.48	0.43	0.45

Table 5.5. Powers of various tests of fit for the negative binomial for zero inflated Poisson (λ, ω) alternatives when $n = 100$ and size $\alpha = 0.05$.

λ	ω	M	$U^2/\text{var}(U)$	$T^2/\text{var}(T)$	$W^2/\text{var}(W)$	$R^2/\text{var}(R)$	C	\tilde{S}_2	X^2
2	0.20	0.33	0.11	0.36	0.12	0.47	0.47	0.46	0.34
2	0.35	0.60	0.48	0.57	0.24	0.72	0.71	0.68	0.58
2	0.50	0.62	0.49	0.62	0.29	0.78	0.77	0.73	0.67

Appendix A. Normalization constants

This appendix gives $d_r = E[g_r^2(T)]$ for $r = 1, 2, 3$ and 4. As before, let μ be the population mean and μ_r the r th central population moment. We have $g_1(t) = t$ so that $d_1 = E[T^2] = E[(X - \mu)^2] = \mu_2$ and writing $g_2(t) = t^2 + a_2t + b_2$ where a_2 and b_2 can be read from $g_2(t)$ in section 2

$$d_2 = E[(X - \mu)^4 + 2a_2(X - \mu)^3 + 2a_2b_2(X - \mu) + 2b_2(X - \mu)^2 + a_2^2(X - \mu)^2 + b_2^2] = \mu_4 + 2a_2\mu_3 + (2b_2 + a_2^2)\mu_2 + b_2^2.$$

Similarly

$$d_3 = \mu_6 + 2a_3\mu_5 + (a_3^2 + 2b_3)\mu_4 + 2(a_3b_3 + c_3)\mu_3 + (b_3^2 + 2a_3c_3)\mu_2 + c_3^2 \text{ and}$$

$$d_4 = \mu_8 + 2a_4\mu_7 + (a_4^2 + 2b_4)\mu_6 + 2(a_4b_4 + c_4)\mu_5 + (b_4^2 + 2a_4c_4 + 2s_4)\mu_4 + 2(a_4s_4 + b_4c_4)\mu_3 + (c_4^2 + 2b_4s_4)\mu_2 + s_4^2$$

where $g_4(t) = t^4 + a_4t^3 + b_4t^2 + c_4t + s_4$ and a_4, b_4, c_4 and s_4 are given by the recurrence relation in section 2 above. To help calculate the d_i we list

$$\mu = \mu_1 = kq/p, \quad \mu_2 = kq/p^2, \quad \mu_3 = kq(1+q)/p^3,$$

$$\mu_4 = (kq/p^2)(6/p^2 - 6/p + 1) + 3k^2q^2/p^4,$$

$$\begin{aligned}
\mu_5 &= (kq/p^2)(24/p^3 - 36/p^2 + 14/p - 1) + 10k^2q^2(1+q)/p^5, \\
\mu_6 &= (kq/p^2)(120/p^4 - 240/p^3 + 150/p^2 - 30/p + 1) \\
&\quad + 15(k^2q^2/p^4)(6/p^2 - 6/p + 1) + 10k^2q^2(1+q)^2/p^6 + 15k^3q^3/p^6, \\
\mu_7 &= kq(720 + 1960kp^2 - 630kp^3 + 56kp^4 - 105k^2p^3 + 420k^2p^2 \\
&\quad - 2310kp - 1800p + 924k + 62p^4 - p^5 + 1560p^2 - 540p^3 \\
&\quad + 210k^2 - 525k^2p)/p^7, \text{ and} \\
\mu_8 &= kq\{5040 + 1806p^4 + 24080kp^2 - 126p^5 - 11620kp^3 - 21924kp \\
&\quad + p^6 + 7630k^2p^2 + 2275kp^4 - 15120p + 7308k + 16800p^2 \\
&\quad + 2380k^2 - 8400p^3 - 7140k^2p - 3360k^2p^3 - 119kp^5 \\
&\quad + 490k^2p^4 - 315k^3p + 315k^3p^2 - 105k^3p^3 + 105k^3\}/p^8.
\end{aligned}$$

Appendix B. P-values via the parametric bootstrap

Gürtler and Henze (2000, p. 223) suggest p -values can be obtained using an analogue of the parametric bootstrap. If W_n denotes a test statistic calculate $w_n := W_n(x_1, x_2, \dots, x_n)$ where x_1, x_2, \dots, x_n denotes, as usual, the data. Find estimates \hat{k} and \hat{p} (not necessarily the ML estimates) from the data and conditional on those estimates, generate $B = 10,000$ say pseudo-random samples of size n , each having the negative binomial (\hat{k}, \hat{p}) distribution. For $j = 1, \dots, B$ compute the value $W_{n,j}^*$ on each random sample. Then the parametric bootstrap p -value is the proportion of the $W_{n,j}^*$ values that are at least the observed w_n , namely $\sum_{j=1}^B I(W_{n,j}^* \geq w_n)/B$. To obtain p -values for two-tailed tests proceed as above and find the p -value, P say. If $P \leq 0.5$ the two-tailed p -value is $2P$. If $P > 0.5$ then the two-tailed p -value is $2(1 - P)$.

The above requires random negative binomial (k, p) values, which can be generated as suggested in the Introduction. If a random gamma $(k, q/p)$ value is taken as a Poisson mean then a random Poisson with this mean has a negative binomial (k, p) distribution. Many statistical packages have routines for gamma and Poisson random values. For example, the statistical package *R* has routines 'rpois' and 'rgamma', and these can be used to generate random negative binomial values.

References

- Anscombe, F.J., 1950. Sampling theory of the negative binomial and logarithmic distributions. *Biometrika*, 37, 358–382.
- Best, D.J., Rayner, J.C.W., 2003. Tests of fit for the geometric distribution. *Communications in Statistics, Simulation and Computation*, 32 (4), 1065–1078.
- Gürtler, N., Henze, N., 2000. Recent and classical goodness-of-fit tests for the Poisson distribution. *Journal of Statistical Planning and Inference*, 90, 207–225.
- Jarvis, B., 1989. *Statistical Aspects of the Microbiological Analysis of Foods*. Elsevier, Amsterdam.
- Johnson, N.L., Kemp, A.W., Kotz, S., 2005. *Univariate Discrete Distributions* (3rd edition). Wiley, New York.
- Krebs, C.J., 1998. *Ecological Methodology* (2nd ed.). Addison-Wesley-Longman, New York.
- Lancaster, H.O., 1975. Joint probability in the Meixner classes. *Journal of the Royal Statistical Society Series B*, 37, 434–443.

- Meintanis, S.G., 2005. Transform methods for testing for the negative binomial hypothesis. *Statistica*, LXV (3), 293-300.
- Rayner, J.C.W., Best, D.J., 1989. *Smooth Tests of Goodness of Fit*. Oxford University Press, New York.
- Rayner, J.C.W., Thas, O., Best, D.J., 2009. *Smooth Tests of Goodness of Fit: Using R* (2nd edition). Wiley, Singapore.
- Stuart, A., Ord, K., 2005. *Kendall's Advanced Theory of Statistics*. Vol. 1 (6th edition), Hodder Arnold, London.
- Thas, O., Rayner, J.C.W., 2005. Smooth tests for the zero inflated Poisson distribution. *Biometrics*, 61 (3), 808-815.