


11-1-2005

Testing Goodness Of Fit Of The Geometric Distribution: An Application To Human Fecundability Data

Sudhir R. Paul

University of Windsor, smjp@uwindsor.ca

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Paul, Sudhir R. (2005) "Testing Goodness Of Fit Of The Geometric Distribution: An Application To Human Fecundability Data," *Journal of Modern Applied Statistical Methods*: Vol. 4: Iss. 2, Article 8.
Available at: <http://digitalcommons.wayne.edu/jmasm/vol4/iss2/8>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized administrator of DigitalCommons@WayneState.

Testing Goodness Of Fit Of The Geometric Distribution: An Application To Human Fecundability Data

Sudhir R. Paul

Department of Mathematics and Statistics
University of Windsor

A measure of reproduction in human fecundability studies is the number of menstrual cycles required to achieve pregnancy which is assumed to follow a geometric distribution with parameter p . Tests of heterogeneity in the fecundability data through goodness of fit tests of the geometric distribution are developed, along with a likelihood ratio test statistic and a score test statistic. Simulations show both are liberal, and empirical level of the likelihood ratio statistic is larger than that of the score test statistic. A power comparison shows that the likelihood ratio test has a power advantage. A bootstrap p -value procedure using the likelihood ratio statistic is proposed.

Key words: Beta-geometric distribution; bootstrap p -value; fecundability data; geometric distribution; likelihood ratio test; score test.

Introduction

The geometric distribution is important in many real life data analyzes. For example, in fecundability studies (Weinberg & Gladen, 1986), the number of cycles required to achieve pregnancy would be distributed as a geometric distribution with parameter p . However, in real life data situations, the actual variation of the data may exceed that of the geometric distribution, as the parameter p may not remain constant in the course of the experiment. It is then useful to assume that the parameter p varies from observation to observation. One can assume one of many continuous distributions for p in the parameter space $0 < p < 1$. But, the most convenient and most sensible distribution for p is the beta distribution, because it is the natural conjugate prior distribution in the Bayesian sense.

It also produces a convenient mixed distribution, namely, the beta-geometric distribution. The parameters of this mixed distribution have practical interpretation. In some other analogous applications, such as in Toxicology, the beta-binomial distribution arises as a beta mixture of the binomial distribution (Weil, 1970; Williams, 1975; Crowder, 1978; Otake & Prentice, 1984).

It is assumed that $Y | p \sim$ geometric distribution. Let $q = 1 - p$. Then, the probability function of Y is $P(Y = y | q) = q^{y-1} p$.

In human reproduction the random variable Y may be the number of menstrual cycles required for conception in which the parameter p may be interpreted as the pre-cycle conception probability or a measure of fecundability (Weinberg & Gladen, 1986). It is assumed that the parameter p is fixed for a given couple, but across couples it varies according to some unspecified underlying distribution which is assumed to be beta with probability density function given by

$$f(p | \alpha, \beta) = \frac{p^{\alpha-1} (1-p)^{\beta-1}}{B(\alpha, \beta)}, 0 < p < 1,$$

Paul R. Sudhir is University Professor, Professor of Statistics, and Chair, Graduate Studies. His research interests are in correlated data, frailty models, generalized linear models, categorical data analysis, zero-inflated and over-dispersed count data regression models, & dose-response modeling. Email him at smjp@uwindsor.ca.

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

is the beta function and where $\Gamma(a)$ is the gamma function:

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx.$$

The mean and variance of the beta random variable p are $\mu = \frac{\alpha}{\alpha + \beta}$ and

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$
 respectively. The

marginal distribution of Y , then, is

$$\begin{aligned} P(Y = y) &= \int_0^1 P(Y = y | p) f(p | \alpha, \beta) dx \\ &= \frac{1}{B(\alpha, \beta)} \int_0^1 p^\alpha (1-p)^{y+\beta-2} dx \\ &= \frac{B(\alpha + 1, y + \beta - 1)}{B(\alpha, \beta)}. \end{aligned}$$

This distribution is known as the beta-geometric distribution. In the human reproduction literature $P(Y=y)$ is the probability that conception occurs at y for a randomly selected couple. The beta-geometric distribution can be written in terms of the parameter $\pi = \alpha/(\alpha + \beta)$ and $\theta = 1/(\alpha + \beta)$, where p is interpreted as the mean parameter and θ as the shape parameter (Weinberg & Gladen, 1986), which is given in what follows.

$$P(Y = y | n) = \frac{\pi \prod_{r=0}^{y-2} \{(1-\pi) + r\theta\}}{\prod_{r=0}^{y-1} \{1 + r\theta\}}.$$

The distribution has mean $\frac{1-\theta}{\pi-\theta}$ and variance

$\frac{\pi(1-\pi)(1-\theta)}{(\pi-\theta)^2(\pi-2\theta)}$. Obviously, $\theta=0$ corresponds to the geometric distribution with mean $\frac{1}{p}$ and variance $\frac{1-p}{p^2}$.

The purpose of this article is to develop tests of goodness of fit of the geometric distribution against the beta-geometric distribution. A score test and a likelihood ratio test are developed. The score test (Rao, 1947) is a special case of the more general $C(\alpha)$ test (Neyman, 1959) in which the nuisance parameters are replaced by their maximum likelihood estimates which are \sqrt{N} consistent estimates (N =number of observations used in estimating the parameters). The score or the $C(\alpha)$ class of tests (i) often maintain, at least approximately, a preassigned level of significance (Bartoo & Puri, 1967), (ii) require estimates of the parameters only under the null hypothesis, and (iii) often produce statistics which are simple to calculate.

These tests are robust in the sense that their optimality remain true whatever the form of the distribution assumed for the data under the alternative hypothesis - a property called robustness of optimality by Neyman and Scott (1966). The $C(\alpha)$ test has been shown by many authors to be asymptotically equivalent to the likelihood ratio test and to the Wald test (Moran, 1970; Cox & Hinkley, 1974). Potential drawbacks to the use of the likelihood ratio and Wald tests include the fact that both require estimates of the parameters under the alternative hypotheses and often show liberal or conservative behaviour. Examples of this may be found in Barnwal & Paul (1988), Paul (1989), Paul (1996), Paul & Banerjee (1998), and Paul and Islam (1995).

In the present context, although the score test statistic has a very simple form, both the score test and the likelihood ratio test have been found, by simulation, to be liberal. A power comparison, using the empirical quantiles derived from the corresponding size simulation to ensure that each test had approximately the nominal size, has been conducted. This comparison shows that the likelihood ratio test has power advantage over the score test. A

bootstrap likelihood ratio test is therefore proposed to test the fit of a geometric model against the over-dispersed geometric model. The bootstrap likelihood ratio test provides approximately correct p-value (Davison & Hinkley, 1998). McLachlan (1987) uses the bootstrap likelihood ratio test to test for the number of components in mixture of normal distributions. McLachlan notes that the bootstrap and the true null distribution of the likelihood ratio statistics are the same. The bootstrap likelihood ratio test was also used by others in similar contexts (Aitkin, Anderson & Hinde, 1981; Karlis & Xekalaki, 1999).

For the situation in which the data are found to be heterogeneous, maximum likelihood estimates of the parameters of the beta-geometric distribution and the elements of the exact Fisher information matrix are obtained. Two sets of data including one on human fecundability study from Weinberg & Gladen (1986) are analyzed.

Tests of Goodness of Fit

Estimation of the Parameters

Suppose data are available on n individuals as $y_i, i = 1, \dots, n$. The maximum likelihood estimate of the parameter p of the geometric distribution is $\hat{p} = 1/\bar{y}$, where

$$\bar{y} = \sum_{i=1}^n y_i / n.$$

The likelihood function for

the data based on the beta-geometric distribution is given as

$$L = \pi^n \prod_{i=1}^n \frac{\prod_{r=1}^{y_i-1} \{1 - \pi + (r-1)\theta\}}{\prod_{r=1}^{y_i} \{1 + (r-1)\theta\}}$$

and the corresponding log-likelihood can be written as

$$l = n \log(\pi) + \sum_{i=1}^n \left[\sum_{r=1}^{y_i-1} \log\{1 - \pi + (r-1)\theta\} - \sum_{r=1}^{y_i} \log\{1 + (r-1)\theta\} \right]$$

The maximum likelihood estimates $\hat{\pi}$ and $\hat{\theta}$ of the parameters π and θ are obtained by solving the maximum likelihood estimating equations

$$\frac{\partial l}{\partial \pi} = 0 \text{ and } \frac{\partial l}{\partial \theta} = 0 \text{ simultaneously. That is, by solving}$$

$$\frac{n}{\pi} - \sum_{i=1}^n \left\{ \sum_{r=1}^{y_i-1} \frac{1}{1 - \pi + (r-1)\theta} \right\} = 0$$

and

$$\sum_{i=1}^n \left\{ \sum_{r=1}^{y_i-1} \frac{r-1}{1 - \pi + (r-1)\theta} - \sum_{r=1}^{y_i} \frac{r-1}{1 + (r-1)\theta} \right\} = 0,$$

simultaneously subject to the constraints $0 < p < 1$ and $\theta > 0$. Note that there is no closed form solution for these equations. So these equations are to be solved using a numerical procedure such as the Newton-Raphson method or a numerical subroutine, such as the IMSL subroutine ZBRENT or NEQNF.

The Likelihood Ratio Test

The maximized log-likelihood under the geometric distribution is

$$l_0 = n \log(\hat{p}) + n(\bar{y} - 1) \log(1 - \hat{p}) \tag{1}$$

and that under the beta-geometric distribution is

$$l_1 = n \log(\hat{\pi}) + \sum_{i=1}^n \left[\sum_{r=1}^{y_i-1} \log\{1 - \hat{\pi} + (r-1)\hat{\theta}\} - \sum_{r=1}^{y_i} \log\{1 + (r-1)\hat{\theta}\} \right]$$

Then, the likelihood ratio statistic to test for $H_0 : \theta = 0$ against

$H_A : \theta > 0$ is $LR = 2(l_1 - l_0)$. Under standard conditions, the asymptotic null distribution of this likelihood ratio statistic would be chi-square with 1 degree of freedom. However, since the parameter θ is necessarily nonnegative, there is

a boundary problem and the regular asymptotic likelihood theory breaks down in this situation. In the course of a general discussion of asymptotic properties of likelihood procedures when some of the parameters are on the boundary, Self & Liang (1987) derive a representation for the asymptotic distribution of the likelihood ratio statistic. Since the parameter value under H_0 is on the boundary of the parameter space it can be easily seen from the results of Self & Liang (1987) that the correct distribution of the LR test is a 50:50 mixture of zero and chi-square with 1 degree of freedom provided $0 < p < 1$.

The Score Test

Define

$$S = \frac{\partial l}{\partial \theta} \Big|_{\theta=0},$$

$$I_{\pi\pi} = E \left(\frac{\partial^2 l}{\partial \pi^2} \Big|_{\theta=0} \right), I_{\pi\phi} = E \left(\frac{\partial^2 l}{\partial \pi \partial \phi} \Big|_{\theta=0} \right),$$

and

$$I_{\theta\theta} = E \left(\frac{\partial^2 l}{\partial \theta^2} \Big|_{\theta=0} \right).$$

Then, a score test statistic for testing $H_0 : \theta = 0$ against $H_A : \theta > 0$ is given by $Z = S / \sqrt{(I_{\theta\theta} - I_{\pi\phi}^2 / I_{\pi\pi})}$. If the nuisance parameter π is replaced by its maximum likelihood estimate under the null hypothesis, then, asymptotically, as $n \rightarrow \infty$, the distribution of Z is standard normal. Note, under the null hypothesis π becomes p . Then, the following is obtained

$$S = \frac{1}{1-p} \sum_{i=1}^n \sum_{r=1}^{y_i-1} (r-1) - \sum_{i=1}^n \sum_{r=1}^{y_i} (r-1) = 0, \quad (2)$$

$$I_{\pi\pi} = n / \{p^2(1-p)\}, I_{\pi\phi} = -n / p^2, \text{ and}$$

$$I_{\phi\phi} = \frac{n\{2-5p+p^2(4-p)-(p-1)(p-2)(1-p)^2\}}{\{p^3(1-p)^2\}}$$

It can be shown that $\text{Var}(S) = I_{\phi\phi} - I_{\pi\phi}^2 / I_{\pi\pi} = n / p^2$. Thus, the score test statistic for testing $H_0 : \theta = 0$ against $H_A : \theta > 0$ is given by $Z = S / \sqrt{(n / p^2)}$. If p is replaced by \hat{p} , where \hat{p} is the maximum likelihood estimate of the parameter p of the geometric distribution, in Z , then, under the null hypothesis $H_0 : \theta = 0$, the statistic Z will have an asymptotic standard normal distribution. Since this is a one-sided test the null hypothesis is rejected at $100(1-\alpha)\%$ level of significance if $Z > z_{\alpha}$, where, z_t is the $100(1-t)\%$ point of the standard normal distribution.

Simulations

A simulation experiment was conducted to study size properties of the likelihood ratio statistic LR and the score test statistic Z . Data have been generated from the geometric distribution with values of the geometric parameter $p = .1, .2, .3, .4, .5$, sample sizes, $n = 10, 20, 50$, and $\alpha = .05, .10$. Each simulation experiment was based on 5000 replications. Empirical size values are given in Table 1.

Table 1: Empirical sizes, in percent, for H_0 of score test statistics Z and the likelihood ratio statistic LR

n	α	Statistics	p				
			0.1	0.2	0.3	0.4	0.5
10	0.05	Z	8.0	6.9	7.2	6.5	6.6
		LR	12.0	10.6	10.6	10.5	10.0
		LR1	12.0	10.6	10.6	10.5	10.0
20		Z	11.2	10.2	10.2	11.5	13.3
		LR	13.0	11.4	11.4	12.7	15.0
		LR1	12.0	10.6	10.6	10.5	10.0
50		Z	13.3	12.6	12.3	13.8	16.4
		LR	14.2	13.3	13.2	14.6	16.6
		LR1	12.0	10.6	10.6	10.5	10.0
100		Z	13.3	12.6	12.3	13.8	16.4
		LR	14.2	13.3	13.2	14.6	16.6
		LR1	12.0	10.6	10.6	10.5	10.0
500		Z	13.3	12.6	12.3	13.8	16.4
		LR	14.2	13.3	13.2	14.6	16.6
		LR1	12.0	10.6	10.6	10.5	10.0
10	0.10	Z	14.0	12.6	12.4	12.8	12.8
		LR	19.0	17.1	16.6	18.0	18.3
		LR1	12.0	10.6	10.6	10.5	10.0
20		Z	17.9	16.7	16.6	17.9	21.8
		LR	20.0	18.2	18.2	19.7	23.0
		LR1	12.0	10.6	10.6	10.5	10.0
50		Z	21.6	20.2	19.9	21.9	25.5
		LR	21.2	20.6	20.0	22.5	25.6
		LR1	12.0	10.6	10.6	10.5	10.0
100		Z	13.3	12.6	12.3	13.8	16.4
		LR	14.2	13.3	13.2	14.6	16.6
		LR1	12.0	10.6	10.6	10.5	10.0
500		Z	13.3	12.6	12.3	13.8	16.4
		LR	14.2	13.3	13.2	14.6	16.6
		LR1	12.0	10.6	10.6	10.5	10.0

Table 2: Empirical powers, in percent, for H_0 , at $\alpha = 0.05$, of score test statistics Z and the likelihood ratio statistic LR. The extra-geometric variation is .01(.05)(.1)

n	Statistics	p		
		0.1	0.3	0.5
10	Z	6(32)(67)	8(20)(39)	5(9)(15)
	LR	7(38)(82)	9(25)(52)	5(10)(18)
20	Z	11(53)(88)	22(49)(70)	10(19)(39)
	LR	10(57)(96)	25(64)(86)	12(24)(46)
50	Z	15(81)(99)	53(93)(97)	8(38)(70)
	LR	16(84)(99)	54(97)(99)	13(44)(81)

From Table 1 it is seen that both the score test statistic and the likelihood ratio statistic are liberal. Empirical level of the likelihood ratio statistic is larger than that of the score test statistic. Also, empirical level increases as the sample size increases. A mean-variance correction of the score test statistic using Taylor series expansion (Paul, 1996) produces empirical levels that are too small compared with the nominal levels.

A power comparison of the two statistics was also conducted. The empirical 95% quantiles derived from the corresponding size simulation have been used to ensure that each test had approximately the nominal size of 0.05. Empirical quantiles were calculated based on 20,000 replications and empirical power calculations were based on 1000 replications. Empirical power values are given in Table 2. The likelihood ratio statistic, in general, shows power advantage, over the score test.

The Bootstrap Goodness of Fit Test

As seen from the simulation results in Section 3, both the likelihood ratio test and the test based on the score test statistic are liberal. However, the likelihood ratio test has some power advantage over the score test. So,

following Davison & Hinkley (1997), a bootstrap test of the null hypothesis $H_0 : \theta = 0$ against $H_A : \theta > 0$ is proposed. The bootstrap likelihood ratio test procedure proceeds according to the following steps:

Step 1. Obtain \hat{p} of the parameter p of the geometric distribution from the data. Calculate the value of the likelihood ratio statistic LR, say LR_0 , from the data.

Step 2. Generate n observations from the fitted null distribution, i.e., the geometric distribution with parameter $p = \hat{p}$ and calculate the likelihood ratio statistic LR_0^* .

Step 3. Repeat step 2 B times obtaining B values of the bootstrap likelihood ratio statistic, say, $LR_0^{(b)}$, $b=1,2,\dots,B$.

Step 4. Estimate the bootstrap p-value by

$$\hat{p}_{boot} = \frac{1 + \#\{LR_0^{(b)*} \geq LR_0\}}{B + 1}.$$

This gives the level at which to reject or not to reject H_0 . A typical value of B is 1000.

Elements of the Expected Fisher Information Matrix of the Beta-geometric Distribution

In this section, the elements of the expected Fisher Information matrix for the estimates of the parameters of the beta-geometric distribution are derived. The calculations are quite involved, so the details were omitted. The exact expressions are given in what follows.

$$I_{11} = E\left(\frac{-\partial^2 l}{\partial \pi^2}\right) = n/\pi^2 + n \sum_{r=2}^{\infty} \frac{P(Y \geq r)}{\{1 - \pi + (r-2)\theta\}^2},$$

$$I_{12} = E\left(\frac{-\partial^2 l}{\partial \pi \partial \phi}\right) = -n \sum_{r=3}^{\infty} \frac{(r-2)P(Y \geq r)}{\{1 - \pi + (r-2)\theta\}^2},$$

and

$$I_{22} = E\left(\frac{-\partial^2 l}{\partial \phi^2}\right) = n \left(\sum_{r=3}^{\infty} \frac{(r-2)^2 P(Y \geq r)}{\{1 - \pi + (r-2)\theta\}^2} - \sum_{r=2}^{\infty} \frac{(r-1)^2 P(Y \geq r)}{\{1 + (r-2)\theta\}^2} \right).$$

Calculations of the above terms do not pose any difficulty if ∞ in the upper limit of the summation is replaced by a sufficiently large number, say, 5000. Thus, the estimated variance of $\hat{\pi}$ and $\hat{\theta}$ are

$$\text{var}(\hat{\pi}) = \frac{\hat{I}_{22}}{(\hat{I}_{11}\hat{I}_{22} - \hat{I}_{12}^2)}$$

and

$$\text{var}(\hat{\theta}) = \frac{\hat{I}_{11}}{(\hat{I}_{11}\hat{I}_{22} - \hat{I}_{12}^2)}$$

respectively, where $\hat{I}_{11}, \hat{I}_{12},$ and \hat{I}_{22} are estimates of $I_{11}, I_{12},$ and I_{22} respectively obtained by replacing the parameter p by its maximum likelihood estimate.

Examples

Example 1: The data, given in the Table 3 from Weinberg & Gladden (1986), refer to times, taken by couples that were attempting to conceive, until pregnancy results.

Table 3: Data from Weinberg and Gladen (1986) on the number of menstrual cycles to pregnancy

Cycles	Number of Women
1	227
2	123
3	72
4	42
5	21
6	31
7	11
8	14
9	6
10	4
11	7
12	28

The data were obtained retrospectively, starting from a pregnancy in each case. Weinberg & Gladen (1986) analyzed fecundability data for a total of 586 women, contributing a total of 1844 cycles. See Weinberg & Gladen (1986) for more details regarding the data. For these data, the data for 12 or more cycles has been combined.

An estimate of the parameter p of the geometric distribution for these data is $\hat{p} = .3177874$. An estimate of the variance is $(1 - \hat{p}) / \hat{p}^2 = 6.76$. The observed variance, however, is 8.68 which is much larger than the variance predicted by the geometric distribution. This indicates that an over-dispersed geometric distribution may fit the data better than the geometric distribution. Now, the value of the likelihood ratio statistic is $LR = 14.97$ with a p -value (using the 50:50 mixture of 0 and chi-square with 1 degree of freedom) = 0.00000006 and the bootstrap p -value is 0.002. In calculating the bootstrap p -value $B = 500$ have been used. The data shows very strong evidence in favor of

the beta-geometric distribution. Note that in this example the p-value of the likelihood ratio statistic is much smaller than the corresponding bootstrap p-value. This is in line with the simulation results in Section 3 that the likelihood ratio test is liberal.

The maximum likelihood estimates of the parameters π and θ of the beta-geometric distribution are $\hat{\pi} = 0.36596$ and $\hat{\theta} = 0.0745$ and the standard errors of the estimates $\hat{\pi}$ and $\hat{\theta}$ are .0162 and .0204 respectively.

Example 2: In example 1 the data produce a value of 14.97 for the likelihood ratio statistic. This is rather large and therefore it is not surprising that both the ordinary likelihood ratio test and the bootstrap likelihood ratio test provide same conclusion. Moreover, the observed variance is about 28% larger than what is predicted by the geometric distribution. Thus, the data given in Table 4 was produced; it was obtained by modifying the data set in Table 3.

Table 4: Modified data of Table 3 on the number of menstrual cycles to pregnancy

Cycles	Number of Women
1	180
2	123
3	72
4	42
5	21
6	31
7	11
8	14
9	6
10	4
11	7
12	18

For these data an estimate of the variance predicted by the geometric distribution is $(1 - \hat{p}) / \hat{p}^2 = 6.88$ and the corresponding observed variance is 7.72. These two variances are much closer than the two corresponding variances for the data in Table 3. This indicates that the geometric distribution might fit these

data well. For these data the value of the likelihood ratio statistic is $LR = 2.51$ with a p-value (using the 50:50 mixture of 0 and chi-square with 1 degree of freedom) = 0.025 and the bootstrap p-value is 0.14. For these data, the bootstrap likelihood ratio procedure shows that the geometric distribution fits the data well at 5% level of significance which is contradicted by the ordinary likelihood ratio test. The reason for this is that the likelihood ratio test is liberal.

References

- Aitkin, M., Anderson, D., & Hinde, J. (1999). Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society, A*, 144, 419-461.
- Barnwal, R. K. & Paul, S. R. (1988). Analysis of one-way layout of count data with negative binomial variation. *Biometrika* 75, 215-222.
- Bartoo, J. B. & Puri, P. S. (1967). On optimal asymptotic tests of composite statistical Hypothesis. *The Annals of Mathematical Statistics*, 38, 1845-52.
- Crowder, M. J. (1978). Beta-binomial ANOVA for proportions. *Applied Statistics*, 27, 34-37.
- Cox, D. R. & Hinkley, D. V. (1974). *Theoretical Statistics*. London: Chapman and Hall.
- Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press.
- International Mathematical and Statistical Libraries (1994). *IMSL Manual*. The Numerical Solution Source, Houston, Texas.
- Karlis, D. & Xekalaki, E (1999). On testing for the number of components in a mixed Poisson model. *Annals of the Institute of Statistical Mathematics*, 51, 149-162.
- Otake, M. & Prentice, R. L. (1984). The analysis of chromosomally aberrant cells based on beta-binomial distribution. *Radiation Research*, 98, 456-470.
- McLachlan, G. J. (1987). On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. *Applied Statistics*, 36, 318-324.

Moran, P. A. P. (1970). On asymptotically optimal tests of composite hypothesis. *Biometrika*, 57, 47-55.

Neyman, J. (1959). *Optimal asymptotic tests for composite hypothesis*. In Probability and Statistics: The Harold Cramer Volume, U. Grenander (ed). New York: Wiley.

Neyman, J. & Scott, E. L. (1966). On the use of $C(\alpha)$ optimal tests of composite hypotheses. *Bulletin of the International Statistical Institute*, 41, 477-497.

Paul, S. R. (1989). Test for the equality of several correlation coefficients. *The Canadian Journal of Statistics*, 93, 217-227.

Paul, S. R. (1996). Score tests for intraclass correlation in familial data. *Biometrics* 52, 955-963.

Paul, S. R. & Banerjee, T. (1998). Analysis of two-way layout of count data involving multiple counts in each cell. *Journal of the American Statistical Association*, 93, 1419-1429.

Paul, S. R. & Islam, A. S. (1995). Analysis of proportions based on parametric and semi-parametric models. *Biometrics*, 51, 1400-1410.

Rao, C. R. (1947). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, 44, 50-57.

Self, S. G. & Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82, 605-610.

Weil, C. S. (1970). Selection of valid number of sampling units and a consideration of their combination in toxicological studies involving reproduction, teratogenesis or carcinogenesis reproduction, teratogenesis. *Food and Cosmetic Toxicology*, 8, 177-182.

Weinberg, P. & Gladen, B. C. (1986). The Beta-geometric distribution applied to comparative fecundability studies. *Biometrics*, 42, 547-560.

Williams, D. A. (1975). Analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics*, 31, 949-952.