ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 21, Number 2, December 2017
Available online at http://acutm.math.ut.ee

Tests of exponentiality against some parametric over/under-dispersed life time models

Rajibul Mian and Sudhir Paul

Abstract. We develop tests of goodness of fit of the exponential model against some over/under dispersion family of distributions. In particular, we develop 3 score test statistics and 3 likelihood ratio statistics. These are (S_1, L_1), (S_2, L_2), and (S_3, L_3) based on a general over-dispersed family of distributions, two specific over/under dispersed exponential models, namely, the gamma and the Weibull distributions, respectively. A simulation study shows that the statistics S_3 and L_3 have best overall performance, in terms of both, level and power. However, the statistic L_3 can be liberal in some instances and it needs the maximum likelihood estimates of the parameters of the Weibull distribution as opposed to the statistic S_3 which is very simple to use. So, our recommendation is to use the statistic S_3 to test the fit of an exponential distribution over any over/under-dispersed exponential distribution.

1. Introduction

There have been numerous studies of over-dispersion in discrete data (see [3, 9, 13, 18, 20]) As far as we can ascertain, very little work has been done on over-dispersed continuous data. The one-parameter exponential distribution has been extensively used to model data that arise in physical sciences and engineering applications. Exponential distribution has been vastly used in reliability theory and reliability engineering because of its memoryless property, which is an appropriate model for the constant hazard rate portion of the bathtub curve. In case of modelling the waiting time, exponential distribution is a good choice. It can be used to model the waiting time of cars exceeding certain speed limit within an interval, time for a radioactive

Received June 16, 2016.
2010 *Mathematics Subject Classification.* 62F03, 65C60.
Key words and phrases. Dispersion parameter, exponential model, score test, likelihood ratio test, two parameter gamma model, Weibull model.
http://dx.doi.org/10.12697/ACUTM.2017.21.14
Corresponding author: Sudhir Paul

particle to decay, waiting time for the next phone call to come. The heights of different molecules can also approximately follow exponential distribution, if the gas and pressure observed at a certain temperature and uniform gravitational field respectively. In exponential network, node's connectivity distribution follows exponential. Moreover, if the concerned variable is monthly or yearly maximum value of daily or weekly rainfall or river discharge volumes, in that case the research of these extreme cases, exponential distribution can be an appropriate choice. This distribution is a special case of a richer family of distributions, such as the Pareto distribution, the gamma distribution and the Weibull distribution, all of which are two-parameter distributions and can be expressed as a family of over-dispersed exponential models. The purpose of this paper is to develop tests of goodness of fit of the exponential model against the over-dispersed family of distributions. In particular, we develop score and likelihood ratio statistics to test goodness of fit of the exponential model against (i) general over-dispersion by using a over-dispersed generalized linear model proposed by [6] , (ii) against a over-dispersed model given by the two parameter gamma distribution, and (iii) against a over-dispersed model given by the Weibull distribution. A performance study in terms of level and power of these statistics, along with some 50:50 mixture counterparts is conducted.

The score test (see [14]) is a special case of the more general $C(\alpha)$ test (see [11]) in which the nuisance parameters are replaced by maximum likelihood estimates which are $\sqrt{N}$ (N is the number of observations used in estimating the parameters) consistent estimates. The score test is particularly appealing as it requires estimates of the parameters only under the null hypothesis, and often produces a statistic which is simple to calculate. Potential drawbacks 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 (see, for example, [2, 12, 19, 8]).

In Section 2, some two-parameter distributions are developed as over-dispersed exponential distributions. In Section 3, we develop the score and the likelihood ratio tests. A simulation study is conducted in Section 4. Some examples are analyzed in Section 5 and a discussion follows in Section 6.

2. Some over-dispersed exponential distributions

2.1. The Pareto distribution. An over-dispersed exponential distribution can be obtained as a compound distribution by mixing the exponential distribution with a gamma distribution for the exponential parameter ρ. Suppose that each individual survival time is exponentially distributed with rate parameter ρ as

$$f(y|\rho) = \rho e^{-\rho y}. \tag{1}$$

Suppose that the rate parameter ρ varies randomly between individuals according to a gamma (ρ_0, k) distribution with mean ρ_0 and index k having the density

$$f(\rho) = (k/\rho_0)(k\rho/\rho_0)^{k-1}e^{-k\rho/\rho_0}/\Gamma(k).$$

The unconditional distribution of y then is a Pareto distribution (see [7])

$$f(y) = \frac{k(k/\rho_0)^k}{(y + k/\rho_0)^{k+1}}. \tag{2}$$

Then, the unconditional mean of y is

$$E(y) = E(E(y|\rho)) = \frac{\rho_0/k}{k-1} = \psi$$

and the unconditional variance for y is

$$\mathrm{Var}\,(y)=E(\mathrm{Var}\,(y|\rho))+\mathrm{Var}\,(E(y|\rho))=\left(\frac{\rho_0/k}{k-1}\right)^2\left(\frac{k}{k-2}\right)=\psi^2\left(\frac{1}{1-2/k}\right).$$

For practical use a more convenient form of the distribution is obtained by using $c = 1/k$ in which case the above density can be written as

$$f(y) = \rho_0(1 + c\rho_0 y)^{-(1+c^{-1})} \tag{3}$$

with modified form of $\mathrm{Var}\,(y) = \psi^2(1/(1-2c))$. Clearly this is an over-dispersion distribution relative to the exponential distribution. This distribution tends to the exponential distribution as $k \to \infty$ or $c \to 0$.

2.2. A general over-dispersed exponential family of distributions. Suppose that for given ρ^*, y has the exponential family model with probability density function

$$f(y|\rho^*) = \rho^* e^{-\rho^* y},$$

where $\rho^* = \nu\rho$ with $E(\nu) = 1$ and $\mathrm{Var}\,(\nu) = \tau$. This allows modeling extra-exponential variation. Then following [6] and [5] we obtain the mixed model by expanding $f(y, \rho^*)$ in a Taylor series at ρ and taking expectations. The resulting over-dispersed exponential model can be written as

$$f_2(y;\rho,\tau) = f(y;\rho)\left\{1 + \sum_{r=2}^{\infty}\frac{\alpha_r}{r!}D_r(y;\rho)\right\},$$

where

$$D_r(y,\rho) = \left\{\frac{\partial^{(r)}}{\partial\rho^{*(r)}}f(y;\rho^*)|_{\rho^*=\rho}\right\}\{f(y;\rho)\}^{-1}$$

and $\alpha_r = E(\rho^* - \rho)^r$. Then it can be seen that

$$E(\rho^*) = \rho, \quad \mathrm{Var}\,(\rho^*) = \tau\rho^2 > 0.$$

Further, for small τ, we assume that $\alpha_r = o(\tau)$ for $r \geq 3$ and

$$f_2(y;\rho,\tau) = f(y;\rho)\left\{1 + \frac{\alpha_2}{2!}D_2(y;\rho)\right\} = f(y,\rho)\left\{1 + \frac{\tau}{2}y\rho(y\rho - 2)\right\}, \quad (4)$$

where $D_2(y;\rho) = (y\rho - 2)y\rho^{-1}$. It can be seen that $f_2(y;\rho,\tau)$ in equation (4) is a valid probability density function

$$E(Y) = (1+\tau)/\rho = \theta(1+\tau),$$

where $\theta = 1/\rho$, and

$$\mathrm{Var}\,(Y) = \theta^2(1 + 4\tau - \tau^2).$$

Obviously, the over-dispersed exponential model (4) reduces to the exponential model (1) for $\tau = 0$.

2.3. The gamma distribution. The two-parameter gamma (ρ, k) distribution with rate parameter ρ and shape parameter k has the density function (p.d.f.)

$$f(y) = \rho(\rho y)^{k-1}e^{-\rho y}/\Gamma(k). \quad (5)$$

It is easy to see that, for $k = 1$, the gamma (ρ, k) distribution reduces to the exponential (ρ) distribution. For further development it is convenient to reparameterize as $\theta = 1/\rho$ and $c = 1/k$. Then, it can be seen that

$$E(Y) = k\theta$$

and

$$\mathrm{Var}\,(Y) = k\theta^2 = \theta^2/c.$$

Thus, $c = 1$ corresponds to the exponential distribution with mean θ, $c > 1$ or $k < 1$ corresponds to an under-dispersed exponential model and $c < 1$ or $k > 1$ corresponds to an over-dispersed exponential model.

2.4. The Weibull distribution. The two-parameter Weibull distribution with scale parameter ρ and shape parameter k has p.d.f.

$$f(y) = k\rho(\rho y)^{k-1}exp[-(\rho t)^k]. \quad (6)$$

Note for $k = 1$, the Weibull distribution becomes an exponential distribution with rate parameter ρ. Hence we are interested in testing the null hypothesis $H_0 : k = 1$. As in Section 2.1 it will be more convenient to repameterize as $\theta = 1/\rho$ and $c = 1/k$. Then, it can be shown that

$$E(Y) = \theta\Gamma(1+c) = \mu \quad \text{and} \quad \mathrm{Var}\,(Y) = \mu^2[(\Gamma(1+2c)/\Gamma^2(1+c)) - 1].$$

Thus, $c = 1$ corresponds to the exponential distribution with mean θ, $c > 1$ or $k < 1$ corresponds to an over-dispersed exponential model and $c < 1$ or $k > 1$ corresponds to an under-dispersed exponential model.

3. The tests

3.1. Test for exponentiality against a general over-dispersed exponential family of distributions. *The score test.* A score test using the Pareto density of the forms in equation (2) or (3) runs into difficulty as the likelihood score evaluated as $k \to \infty$ or $c \to 0$ either becomes independent of the data or becomes unbounded. However, a score test using a general over-dispersed exponential model exists.

Let $Y_i,\ i = 1, \ldots, n$, be a sample of independent observations from (4) with $1/\rho_i = \theta_i$ a function of $p \times 1$ vector of covariates X_i and a vector of regression parameters β; that is, $\theta_i = \theta_i(X_i; \beta),\ i = 1, \ldots, n$. The log-likelihood can be written as

$$l(\theta, \tau; y) = \sum_{i=1}^{n} \left[-\log \theta_i - \frac{y_i}{\theta_i} + \log \left\{ 1 + \frac{\tau}{2} \frac{y_i}{\theta_i} \left(\frac{y_i}{\theta_i} - 2 \right) \right\} \right].$$

The score test statistic for testing the hypothesis $H_0 : \tau = 0$ against the alternative $H_A : \tau > 0$ is based on the score function evaluated under the null hypothesis, namely

$$\Psi = \left. \frac{\partial l}{\partial \tau} \right|_{\tau=0} = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{\theta_i^2} \left\{ (y_i - \theta_i)^2 - \theta_i^2 \right\}.$$

Now, define the expected mixed second order partial derivative matrices

$$I_{11} = E\left(- \left. \frac{\partial^2 l}{\partial \tau \partial \tau'} \right|_{\tau=0} \right),\ I_{12} = E\left(- \left. \frac{\partial^2 l}{\partial \tau \partial \beta} \right|_{\tau=0} \right),\ I_{22} = E\left(- \left. \frac{\partial^2 l}{\partial \beta \partial \beta'} \right|_{\tau=0} \right).$$

Then, it can be seen that the variance of Ψ is $V^2 = I_{11} - I_{12}' I_{22}^{-1} I_{12}$. Further, algebra shows that $I_{11} = 2n$, and the j^{th} element of I_{12} and the $(j, j')^{th}$ element of I_{22} are equal to

$$\sum_{i=1}^{n} \frac{1}{\theta_i} \frac{\partial \theta_i}{\partial \beta_j} \quad \text{and} \quad \sum_{i=1}^{n} \frac{1}{\theta_i^2} \frac{\partial \theta_i}{\partial \beta_j} \frac{\partial \theta_i}{\partial \beta_{j'}},$$

respectively. Then, under some conditions for the application of the central limit theorem to the score component Ψ and the regularity conditions of maximum likelihood estimates, the score test statistic for testing $H_0 : \tau = 0$ is

$$S_1 = \hat{\Psi} / \hat{V},$$

which, asymptotically, as $n \to \infty$, has a standard normal distribution, where, $\hat{\Psi} = \Psi(\hat{\beta})$, $\hat{V} = V(\hat{\beta})$ and $\hat{\beta}$ is the maximum likelihood estimate of β under the null hypothesis. Note, the maximum likelihood estimates of the regression parameters $\beta_j,\ j = 1, \ldots, p$, are obtained by solving the p estimating equations

$$\sum_{i=1}^{n}(y_i-\theta_i)\frac{1}{\theta_i^2}\frac{\partial\theta_i}{\partial\beta_j}=0. \tag{7}$$

Thus, the maximum likelihood estimate of θ_i is $\hat{\theta}_i=\theta_i(X_i;\hat{\beta})$. For the link function $\theta_i=X_i'\beta$, where X_i is a $p\times 1$ vector of regression variables, the j^{th} element of I_{12} and the $(j,j')^{th}$ element of I_{22} are, respectively,

$$\sum_{i=1}^{n}\frac{1}{\hat{\theta}_i}X_{ij}' \quad\text{and}\quad \sum_{i=1}^{n}\frac{1}{\hat{\theta}_i^{\,2}}X_{ij}'X_{ij'}.$$

When no covariate is involved, $\theta_i=\theta$, maximum likelihood estimate of θ is $\bar{y}$ and the score test statistic for testing $H_0:\tau=0$ against any over-dispersed family of exponential distributions reduces to

$$S_1=\frac{1}{2\sqrt{2n}}\left(\sum_{i=1}^{n}\frac{y_i^2}{\bar{y}^2}-2n\right).$$

The likelihood ratio test. The likelihood ratio test (LRT) statistic can be written as

$$\begin{aligned}L_1&=2(l(\tilde{\theta},\tilde{\tau};y)|H_A-l(\hat{\theta},\hat{\tau};y)|H_0)\\&=2(l(\tilde{\theta},\tilde{\tau};y)-l(\hat{\theta},0,;y)),\end{aligned}$$

which has, asymptotically, a $\chi^2_{(1)}$ distribution, as $n\to\infty$, where $\tilde{\theta}_i=\theta_i(X_i;\tilde{\beta})$, $\tilde{\beta}_j$, $j=1,\dots,p$, and $\tilde{\tau}$ are obtained by solving the estimating equations

$$\begin{aligned}\frac{\partial l}{\partial\beta_j}&=\sum_{i=1}^{n}\frac{(\frac{1}{\theta_i})(\frac{y_i}{\theta_i}-1)[1+\frac{\tau}{2}(\frac{y_i}{\theta_i})(\frac{y_i}{\theta_i}-4)]}{1+\frac{\tau}{2}(\frac{y_i}{\theta_i})(\frac{y_i}{\theta_i}-2)}\frac{\partial\theta_i}{\partial\beta_j},\\\frac{\partial l}{\partial\tau}&=\sum_{i=1}^{n}\frac{\frac{1}{2}(\frac{y_i}{\theta_i})(\frac{y_i}{\theta_i}-2)}{1+\frac{\tau}{2}(\frac{y_i}{\theta_i})(\frac{y_i}{\theta_i}-2)}\end{aligned}$$

simultaneously.

3.2. Test of exponentiality against a gamma alternative. *The score test.* Let $Y_i, i=1,\dots,n$, be a sample of independent observations from (5) with

$$(Y_i)=\mu_i=k\theta_i(X_i;\beta).$$

Our interest is to test $H_0:k=1$ against all alternatives. Now, the log-likelihood apart from a constant can be written as

$$l=-k\sum_{i=1}^{n}\log\theta_i+(k-1)\sum_{i=1}^{n}\log y_i-\sum_{i=1}^{n}\frac{y_i}{\theta_i}-n\,\log(\Gamma(k)).$$

Then, using this log-likelihood and a procedure similar to what was used in Section 3.1 it can be shown that the score test statistic for testing $H_0 : k = 1$ is

$$S_2 = \hat{T}/\hat{V},$$

where

$$\hat{T} = -\sum_{i=1}^{n} \log(\theta_i(\hat{\beta})) + \sum_{i=1}^{n} \log y_i - n\psi(1),$$

$\psi(1)$ is the value of a digamma function $\psi(k) = \dfrac{\partial}{\partial k}\log(\Gamma(k))$ under H_0, and $\hat{V}^2 = V^2(\hat{\beta})$ with $V^2 = I_{11} - I_{12}^{'} I_{22}^{-1} I_{12}$ and $I_{11} = n\psi^{'}(1)$. The expressions for the j^{th} element of I_{12} and the $(j, j^{'})^{th}$ element of I_{22} are the same as those obtained in Section 3.1. Further, the estimating equations for the p regression parameters here are the same as those given in equations (7). When no covariate is involved we have

$$\hat{T} = \sum_{i=1}^{n} \log(y_i/\bar{y}) - n\psi(1) = \sum_{i=1}^{n} \log(y_i/\bar{y}) - 0,5772156649\, n$$

and $\hat{V} = \sqrt{n\psi'(1)}$ with

$$\psi^{'}(1) = \int_0^\infty (t\exp(-t))/(1-\exp(-t))dt.$$

The likelihood ratio test. The LRT statistic can be written as

$$\begin{aligned} L_2 &= 2(l(\tilde{\theta}, \tilde{k}; y)|H_A - l(\hat{\theta}, \hat{k}; y)|H_0) \\ &= 2(l(\tilde{\theta}, \tilde{k}; y) - l(\hat{\theta}, 0; y)), \end{aligned}$$

which has, asymptotically, a $\chi^2_{(1)}$ distribution, as $n \to \infty$, where $\tilde{\theta}_i = \theta_i(X_i; \tilde{\beta})$, $\tilde{\beta}_j$, $j = 1, \ldots, p$, and $\tilde{k}$ are obtained by solving simultaneously the estimating equations

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{y_i}{\theta_i^2} - \frac{k}{\theta_i}\right) \frac{\partial \theta_i}{\partial \beta_j}, \quad \frac{\partial l}{\partial k} = \sum_{i=1}^{n} \left(\log \frac{y_i}{\theta_i} - \psi(k)\right),$$

where ψ is a digamma function.

3.3. Test of exponentiality against an Weibull alternative. *The score test.* Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from an Weibull distribution (6). Then the log-likelihood function apart from a constant can be written as

$$l = n \log k - k\sum_{i=1}^{n} \log \theta_i + (k-1)\sum_{i=1}^{n} \log y_i - \sum_{i=1}^{n} (y_i/\theta_i)^k .$$

Here θ_i=$\theta_i(X_i;\beta)$ which connects the mean of datum y with possible covariates X_i and regression coefficients β. Now, using this log-likelihood function and a procedure similar to what was used in Section 3.1, it can be shown that the score test statistic for testing $H_0 : k = 1$ is

$$S_3 = \hat{T}/\hat{V},$$

where

$$\hat{T} = n + \sum_{i=1}^{n} \left[\left(1 - \frac{y_i}{\hat{\theta}_i}\right) \log \frac{y_i}{\hat{\theta}_i} \right], \quad \hat{V}^2 = V^2(\hat{\beta})$$

with

$$V^2 = I_{11} - I_{12}^{'} I_{22}^{-1} I_{12} \quad \text{and} \quad I_{11} = n + \sum_{i=1}^{n} \frac{y_i}{\hat{\theta}_i} \left(\log \frac{y_i}{\hat{\theta}_i}\right)^2.$$

The j^{th} element of I_{12} and the $(j, j^{'})^{th}$ element of I_{22} are, respectively,

$$\sum_{i=1}^{n} \left(\hat{\theta}_i - y_i - y_i \log \frac{y_i}{\hat{\theta}_i}\right) \frac{1}{\hat{\theta}_i^{\,2}} \frac{\partial \theta_i}{\partial \beta_j} \quad \text{and} \quad \sum_{i=1}^{n} \frac{2y_i - \hat{\theta}_i}{\hat{\theta}_i^{\,3}} \frac{\partial \theta_i}{\partial \beta_j} \frac{\partial \theta_i}{\partial \beta_j{'}}.$$

Note that in order to avoid approximations to expected values here we have used observed information matrix to obtain asymptotic variance of the likelihood score T. The maximum likelihood estimates of the regression parameters β_j, $j = 1, \ldots, p$, are the same as those obtained in equation (7). If a linear link function $\theta_i = X_i'\beta$, where X_i is a $p \times 1$ vector of regression variables, is used, then the j^{th} element of I_{12} and the $(j, j')^{th}$ element of I_{22} are, respectively,

$$\sum_{i=1}^{n} \frac{\hat{\theta}_i - y_i - y_i \log(y_i/\hat{\theta}_i)}{\hat{\theta}_i^{\,2}} X_{ij}' \quad \text{and} \quad \sum_{i=1}^{n} \frac{2y_i - \hat{\theta}_i}{\hat{\theta}_i^{\,3}} X_{ij}' X_{ij'}.$$

In situations where there are no covariates, we have

$$T = n + \sum_{i=1}^{n} \left(1 - \frac{y_i}{\bar{y}}\right) \log \frac{y_i}{\bar{y}},$$

$$I_{11} = n + \sum_{i=1}^{n} \frac{y_i}{\bar{y}} \left(log \frac{y_i}{\bar{y}}\right)^2, \quad I_{12} = 0,$$

and hence $V = \sqrt{I_{11}}$. Thus the score statistic is

$$S_3 = \left(n + \sum_{i=1}^{n} \left(1 - \frac{y_i}{\bar{y}}\right) \log \frac{y_i}{\bar{y}}\right) \Bigg/ \sqrt{n + \sum_{i=1}^{n} \frac{y_i}{\bar{y}} \left(\log \frac{y_i}{\bar{y}}\right)^2}.$$

The statistic $U_{k0}(v_{kk})^{1/2}$ in [7] reduces to the statistic S_3 after simplifications, when there is no censoring or covariates.

The likelihood ratio tests. The LRT statistic can be written as

$$L_3 = 2(l(\tilde{\theta}, \tilde{k}; y)|H_A - l(\hat{\theta}, \hat{k}; y)|H_0)$$
$$= 2(l(\tilde{\theta}, \tilde{k}; y) - l(\hat{\theta}, 0; y)),$$

which has, asymptotically, a $\chi^2_{(1)}$ distribution, as $n \to \infty$, where $\tilde{\theta}_i = \theta_i(X_i; \tilde{\beta})$, $\tilde{\beta}_j$, $j = 1, \ldots, p$, and $\tilde{k}$ are obtained by solving the estimating equations

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{k}{\theta_i} \left[\frac{y_i}{\theta_i} \right]^k - 1 \right) \frac{\partial \theta_i}{\partial \beta_j}, \; \frac{\partial l}{\partial k} = \sum_{i=1}^{n} \left(\frac{1}{k} + \log \frac{y_i}{\theta_i} - \left[\frac{y_i}{\theta_i} \right]^k \log \frac{y_i}{\theta_i} \right)$$

simultaneously.

4. Simulation

A simulation study was conducted to study the performance, in terms of size and power of all the statistics. Note that in the case of the general over-dispersion test we deal with $H_0 : \tau = 0$ against the alternative $H_A : \tau > 0$. Thus, the value of the parameter under the null hypothesis is on the boundary of the parameter space. The distribution of the likelihood ratio test statistic and the score test then asymptotically have a 50 : 50 mixture of χ^2_0 and χ^2_1 distributions. That is, the distribution of each of these statistics is not χ^2_1, but rather $\frac{1}{2}\chi^2_1 + \frac{1}{2}\chi^2_0$, i.e., a 50 : 50 mixture of a point mass at zero and a chi square distribution with one degree of freedom. Note, point mass at zero is zero. So, the distribution is $\frac{1}{2}\chi^2_1$. So, to get a α level test, look up the 2α point of a χ^2_1. See, for example, [15, 16, 17].

Thus, in the simulation study we consider S_1, its 50 : 50 mixture counterpart S_{1F}, L_1, its 50 : 50 mixture counterpart L_{1F}, S_2, L_2, S_3 and L_3. For studying level properties of these statistics we simulated samples from the exponential distribution with mean $\mu = 1$. For power studies we simulated data from (a) the general over-dispersed exponential models, namely the gamma mixture of the exponential (the Pareto distribution) model, and the lognormal mixture of exponential model, (b) specific over-dispersed exponential models, namely, the gamma and the Weibull distributions. Empirical level and power results for data simulated from the gamma mixture of the exponential distribution, the lognormal mixture of exponential distribution, the gamma distribution, and the Weibull distribution are given in Table 1, Table 2, Table 3 and Table 4, respectively. In the simulation we considered nominal significant levels $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.10$. However, comparative results for data simulated from all the above distributions are similar for all the levels. So, in the tables we give results only for $\alpha = 0.05$.

Note that, the lognormal mixture of exponential distribution is used to generate over-dispersed data. Test statistics are applied to these over-dispersed data and their performances are checked to evaluate the robustness of the test statistics. Development of the test statistics based on the lognormal mixture of exponential distribution is difficult at best and so far seems infeasible and unnecessary, and hence was not considered in the study.

To do the simulation study, samples from the different distributions were generated using "stats" and "actuar" packages of the statistical software "R version 3.0.1". For solving the simultaneous equations, we used function "nleqslv()" of "nleqslv" package and "BBsolve()" function of "BB" package.

For power calculation, we computed absolute value of the test statistics ($|S|$) and compared it with absolute value the quantiles ($|Q|$) calculated from standard normal distribution based on α. We assign indicator 1 if $|S| \geq |Q|$, and 0 otherwise. The power of the test, in the simulation study, is then obtained from the average of these indicators. We calculate power for all test statistics under different values of dispersion parameter c. Thus, the results in the column under $c = 0$ in Table 1 and Table 2 refer to the empirical levels. However, the results in the column under $c = 1$ in Table 3 and Table 4 refer to the empirical levels. These levels are all the same as in all cases data are generated from the exponential distribution.

First we compare level properties of all the statistics. All the statistics are in general conservative. For small sample sizes ($n < 50$) the statistics S_1, S_{1F}, L_1 and L_{1F} are all very conservative. The statistic S_2 is very conservative for all sample sizes. The levels of the statistics S_3 and L_3 are closer to the nominal level as sample size increases. The performance of L_3 is liberal in terms of level as sample size increases compared to S_3.

For comparison of power among all the statistics, we find that in Table 1 and Table 2 power increases as the value of c increases from 0. However, in Table 3 and Table 4 power increases as the value of c increases from 0 and also increases as the value of c decreases from 0. In all cases power increases as the sample size increases.

The statistics S_3 and L_3 have the best overall performance, in terms of both, level and power. However, the statistic L_3 can be liberal for large sample sizes.

Based on the level and the power properties of the test statistics, we recommend to use the statistic S_3 to test the fit of an exponential distribution over any over/under-dispersed exponential distribution.

5. Examples

5.1. Example 1. In an attempt to fit a generalized Pareto distribution [4] used a set of fatique data originally reported by [1]. In [4] three different estimation methods are applied and compared to this data. According to their analysis they found that generalized Pareto distribution assumption is appropriate for fatigue data. The data are given in Table 5 which refers to lifetime (in hours) of Kevlar/Epoxy strand at a stress level 70%. For more details of the data, see [1]. We consider fatigue data as over-dispersed data and use this data to test the fit of the exponential distribution over some of the over-dispersed exponential models. For this data, the values of the test statistics S_3 and L_3 with p-values in the parenthesis are 17.46 (0.0000) and 26.45 (0.0000) respectively. It can be seen that both these test statistics S_3 and L_3 decisively reject the hypothesis of fit of an exponential distribution over an over-dispersed exponential model. These findings are in agreement with the results of the simulation study, i.e., the statistics S_3 and L_3 have best overall performance, in terms of both, level and power. Despite our recommendation about S_3 to test the exponentiality of the data, all developed score tests (S_1, S_2, and S_3), as well as likelihood ratio tests (L_1, L_2, and L_3) will facilitate researchers to competently identify the distributional assumption of the dispersed continuous data.

5.2. Example 2. The breakdown times data in Table 6 are from [10] . The data refer to the results of a life test experiment in which specimens of a type of electrical insulating fluid were subjected to different voltage stresses. There were seven groups of specimens in the data set and specimens were tested at voltages ranging from 26 to 38 kilovolts. We consider the breakdown times as over/under-dispersed data and use only a part of these data, namely, time to breakdown at stress level of 32 kilovolts, to test the fit of the exponential distribution over some of the over/under-dispersed exponential models. For this data the values of the test statistics, S_3 and L_3, with p-values in the parenthesis, are 10.41 (0.0012) and 8.21 (0.0041). Both these test statistics reject the fit of the exponential distribution in favour of an over-dispersed exponential distribution.

6. Discussion

We have developed three score tests statistics S_1, S_2 and S_3 and three likelihood ratio statistics L_1, L_2 and L_3 to test the fit of an exponential distribution over some over/under-dispersed exponential models. The score statistics S_1, S_2 and S_3 are simple and easy to use. These statistics use only the maximum likelihood estimates of the exponential distribution. The score statistics S_1 and S_2 are, in general, very conservative. The likelihood ratio statistics L_1, L_2 and L_3 use the maximum likelihood estimates of the parameters of the assumed over-dispersed exponential model. Moreover, as the score test statistics S_1 and S_2, the likelihood ratio statistics L_1 and L_2 are also very conservative. The statistics S_3 and L_3 have best overall performance, in terms of both, level and power. However, the statistic L_3 can be liberal for large sample sizes. Moreover, the statistic L_3 is based on the maximum likelihood estimates of the parameters of the Weibull distribution, the same time statistics S_3 is very simple to use. So, our recommendation is to use the statistic S_3 to test the fit of an exponential distribution over any over/under-dispersed exponential distribution.

Appendix: Tables

TABLE 1. Emperical power (%) of score test statistics S_1, S_{1F}, S_2, S_3, and LRT statistics L_1, L_{1F}, L_2, L_3 for testing over-dispersion with no covariates at $\alpha = 0.05$ when data are simulated from **gamma mixture of the exponential distribution** based on 20,000 replications ($c = 0$ indicates emperical level).

n	Statistic	c										
		0.00	0.05	0.10	0.20	0.30	0.40	0.50	1.00	2.00	3.00	4.00
5	S_1	0.00	0.00	0.03	0.07	0.09	0.29	0.62	04.80	20.10	34.81	45.12
	S_{1F}	0.06	0.08	0.12	0.25	0.58	1.03	2.00	09.18	28.10	43.04	53.65
	L_1	0.00	0.00	0.00	0.28	0.36	1.29	2.40	11.27	40.14	82.40	96.36
	L_{1F}	0.00	0.02	0.00	1.42	2.40	3.91	5.88	20.89	56.05	94.13	97.66
	S_2	0.49	0.67	0.84	1.18	1.80	2.76	4.00	14.43	41.93	62.74	75.55
	L_2	0.52	0.56	1.00	1.96	3.16	5.32	7.22	17.18	31.62	34.05	40.70
	S_3	2.38	2.39	2.66	3.17	3.84	5.24	7.09	19.25	48.03	67.31	79.14
	L_3	1.26	1.16	1.48	2.16	3.18	3.56	3.54	03.88	04.36	04.25	03.45
10	S_1	0.50	0.70	1.39	3.04	5.74	9.50	13.30	35.84	68.16	83.50	90.44
	S_{1F}	0.80	1.31	2.20	4.34	7.80	12.20	16.53	41.01	72.65	86.45	92.73
	L_1	0.07	0.04	0.10	0.24	1.40	3.51	4.90	28.18	71.35	99.41	99.80
	L_{1F}	0.38	0.67	0.91	2.70	5.76	10.22	14.43	49.00	85.90	99.70	99.80
	S_2	0.50	0.62	0.86	1.62	3.17	5.76	8.85	33.77	74.97	91.26	96.58
	L_2	1.11	1.64	2.40	4.14	7.34	11.96	16.12	45.08	69.70	77.20	84.55
	S_3	3.66	3.75	4.03	5.10	7.63	11.47	15.96	43.97	81.67	94.00	97.84
	L_3	2.96	3.30	3.72	3.68	5.22	6.96	11.74	49.02	88.36	95.45	98.40
20	S_1	0.64	1.58	3.43	8.52	15.56	24.20	33.17	70.59	95.16	98.96	99.83
	S_{1F}	1.25	2.44	5.11	11.30	19.30	28.72	38.20	74.86	96.26	99.28	99.89
	L_1	0.37	0.59	1.42	2.87	5.84	11.24	20.36	66.14	96.40	100	100.0
	L_{1F}	1.46	2.42	4.20	9.24	16.80	26.35	38.85	83.12	98.79	100	100.0
	S_2	0.35	0.55	0.82	2.43	5.99	11.33	18.48	62.41	95.92	99.55	99.94
	L_2	2.27	2.68	4.20	8.28	14.78	22.92	34.14	76.74	95.36	98.35	99.6
	S_3	3.84	3.89	4.88	8.31	15.20	23.64	33.74	75.95	98.00	99.80	99.98
	L_3	5.91	6.06	6.76	11.78	23.72	41.96	59.22	94.80	99.70	99.85	100.0
50	S_1	1.07	3.03	7.46	20.05	37.96	55.06	70.12	97.50	99.99	100	100
	S_{1F}	1.83	4.76	10.42	25.30	44.18	61.23	75.36	98.30	100	100	100
	L_1	2.00	3.23	6.10	15.09	31.24	50.71	67.57	98.88	100	100	100
	L_{1F}	4.80	8.34	13.98	29.97	50.12	68.44	82.12	99.74	100	100	100
	S_2	0.24	0.61	1.29	4.42	13.18	27.68	44.85	94.58	100	100	100
	L_2	4.08	4.84	7.44	16.06	32.24	50.26	66.36	98.44	100	100	100
	S_3	3.94	4.32	6.94	17.59	35.35	54.36	71.04	98.58	100	100	100
	L_3	8.25	7.28	11.98	34.10	61.30	80.60	90.22	99.86	100	100	100
100	S_1	1.00	4.23	11.70	36.40	63.43	82.48	92.95	99.97	100	100	100
	S_{1F}	1.83	6.72	16.53	44.05	70.17	86.92	95.19	99.99	100	100	100
	L_1	3.95	8.37	14.39	38.95	66.64	85.91	95.18	100	100	100	100
	L_{1F}	8.49	15.97	26.25	55.76	80.11	92.88	97.90	100	100	100	100
	S_2	0.24	0.57	1.60	8.46	26.76	53.12	75.70	99.87	100	100	100
	L_2	3.82	6.30	9.62	29.06	54.62	77.38	90.70	99.98	100	100	100
	S_3	3.77	5.20	10.73	32.82	61.77	82.98	93.92	100.00	100	100	100
	L_3	6.95	8.92	17.58	51.68	79.16	93.00	97.54	100.00	100	100	100

Table 2. Emperical power (%) of score test statistics S_1, S_{1F}, S_2, S_3, and LRT statistics L_1, L_{1F}, L_2, L_3 for testing over-dispersion with no covariates at $\alpha = 0.05$ when data are simulated from **lognormal mixture of the exponential distribution** based on 20,000 replications ($c = 0$ indicates emperical level).

n	Statistic	c										
		0.00	0.05	0.10	0.20	0.30	0.40	0.50	1.00	2.00	3.00	4.00
	S_1	0.00	0.01	0.00	0.00	0.00	0.02	0.03	0.64	9.03	22.80	35.49
	S_{1F}	0.06	0.03	0.02	0.05	0.08	0.12	0.24	2.29	16.47	33.00	45.18
	L_1	0.00	0.00	0.02	0.16	0.10	0.10	0.12	2.46	23.82	74.62	97.91
5	L_{1F}	0.00	0.02	0.02	0.65	0.55	0.86	1.07	7.03	43.49	94.58	98.64
	S_2	0.49	0.47	0.48	0.55	0.62	0.74	1.24	5.33	35.33	66.86	83.19
	L_2	0.52	0.50	0.70	1.34	1.44	1.48	2.50	9.38	42.70	61.40	83.65
	S_3	2.38	2.43	2.34	2.24	2.42	2.70	3.28	9.04	42.65	71.81	86.34
	L_3	1.26	1.04	0.84	1.50	1.70	1.96	2.28	2.36	4.56	11.20	36.65
	S_1	0.50	0.44	0.38	0.58	1.01	1.68	2.73	14.67	51.14	74.39	84.73
	S_{1F}	0.80	0.78	0.78	1.10	1.74	2.66	3.92	18.70	57.33	79.42	88.49
	L_1	0.07	0.06	0.08	0.14	0.16	0.16	0.72	8.04	57.84	99.54	99.61
10	L_{1F}	0.38	0.26	0.53	0.57	0.93	1.49	3.28	23.70	80.29	99.85	99.74
	S_2	0.50	0.40	0.43	0.61	0.72	0.98	1.55	12.67	69.38	94.38	98.94
	L_2	1.11	1.66	1.16	1.48	2.08	2.86	4.22	24.88	80.30	94.50	99.30
	S_3	3.66	3.68	3.50	3.55	3.70	4.05	4.81	21.97	78.22	96.43	99.41
	L_3	2.96	1.94	1.48	2.04	2.20	2.34	2.66	5.28	60.20	93.10	99.45
	S_1	0.64	0.94	0.89	1.48	2.40	4.71	7.72	37.54	86.61	97.58	99.52
	S_{1F}	1.25	1.55	1.54	2.28	3.74	6.49	10.48	43.66	90.03	98.38	99.73
	L_1	0.37	0.31	0.30	0.71	0.61	1.49	2.96	33.23	93.68	100	100
20	L_{1F}	1.46	1.57	1.60	2.01	3.15	5.38	9.70	57.43	97.99	100	100
	S_2	0.35	0.32	0.32	0.54	0.75	1.41	2.44	27.77	94.53	99.92	99.99
	L_2	2.27	2.62	2.42	2.92	3.26	5.02	7.12	50.62	97.92	99.95	100
	S_3	3.84	3.98	3.69	3.81	4.22	5.84	8.36	46.32	97.77	99.98	100.00
	L_3	5.91	3.20	4.30	3.50	4.64	5.14	7.02	52.78	99.06	100	100
	S_1	1.07	1.08	1.46	2.22	5.06	10.48	19.02	78.02	99.85	100	100
	S_{1F}	1.83	1.96	2.30	3.64	7.43	14.22	24.34	83.14	99.94	100	100
	L_1	2.00	2.16	2.08	3.11	4.91	9.28	18.07	89.64	100	100	100
50	L_{1F}	4.80	5.25	5.26	8.14	11.83	19.90	32.87	96.13	100	100	100
	S_2	0.24	0.25	0.32	0.43	0.89	2.00	4.71	65.59	100	100	100
	L_2	4.08	3.68	3.68	4.70	5.96	9.94	16.88	88.30	100	100	100
	S_3	3.94	3.95	3.82	4.05	5.54	10.05	18.11	86.81	100	100	100
	L_3	8.25	4.08	4.12	4.72	7.00	11.42	20.18	92.66	100	100	100
	S_1	1.07	1.19	1.43	3.15	8.12	18.14	33.57	97.06	100	100	100
	S_{1F}	1.83	2.17	2.53	5.24	12.21	24.21	41.64	98.31	100	100	100
	L_1	3.95	3.83	3.86	6.57	11.60	23.43	41.68	99.66	100	100	100
100	L_{1F}	8.49	8.25	8.40	13.28	22.63	38.31	57.64	99.90	100	100	100
	S_2	0.24	0.27	0.27	0.47	1.11	3.13	9.07	93.80	100	100	100
	L_2	3.82	4.02	4.34	5.38	8.60	15.50	30.22	99.20	100	100	100
	S_3	3.77	4.00	3.96	4.75	8.26	17.38	33.76	99.21	100	100	100
	L_3	6.95	3.94	3.98	5.86	10.90	21.40	39.38	99.82	100	100	100

TABLE 3. Emperical power (%) of score test statistics S_1, S_{1F}, S_2, S_3, and LRT statistics L_1, L_{1F}, L_2, L_3 for testing over-dispersion with no covariates at $\alpha = 0.05$ when data are simulated from **gamma distribution** based on 20,000 replications($c = 1$ indicates emperical level).

n	Statistic	c										
		0.70	0.80	0.90	0.95	1.00	1.10	1.20	1.50	2.00	3.00	4.00
5	S_1	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.10	0.56	3.12	6.65
	S_{1F}	0.00	0.00	0.04	0.03	0.06	0.10	0.19	0.60	2.05	6.93	12.59
	L_1	0.00	0.00	0.20	0.10	0.00	0.10	0.14	0.94	4.51	43.87	95.86
	L_{1F}	0.00	0.06	0.32	0.24	0.00	0.56	0.92	4.05	14.25	76.92	97.12
	S_2	0.02	0.09	0.24	0.33	0.49	0.98	1.80	5.69	16.55	42.06	61.19
	L_2	0.14	0.36	0.70	0.92	0.52	2.14	3.40	12.02	28.08	60.40	90.45
	S_3	3.31	2.84	2.37	2.13	2.38	2.90	3.77	8.51	21.24	47.24	65.99
	L_3	2.28	1.72	2.24	1.64	1.26	1.92	2.12	4.40	9.92	29.00	59.00
10	S_1	0.03	0.10	0.16	0.34	0.50	0.53	0.98	2.76	7.76	21.29	34.85
	S_{1F}	0.08	0.16	0.36	0.62	0.80	1.03	1.82	4.38	11.20	27.73	42.83
	L_1	0.00	0.02	0.08	0.36	0.07	0.68	1.06	5.66	19.96	88.71	95.61
	L_{1F}	0.04	0.16	0.44	0.81	0.38	2.39	3.97	15.14	40.10	95.16	96.49
	S_2	0.00	0.03	0.14	0.21	0.50	1.02	2.20	9.25	30.69	70.48	89.04
	L_2	0.18	0.24	0.50	1.08	1.11	2.96	5.82	20.82	55.84	90.70	99.35
	S_3	9.12	6.07	4.38	3.88	3.66	4.10	5.56	14.85	39.17	77.17	92.09
	L_3	2.66	2.56	2.00	2.18	2.96	2.02	2.96	6.28	24.96	64.25	93.65
20	S_1	0.03	0.16	0.39	0.54	0.64	1.40	2.44	6.86	19.47	49.34	70.23
	S_{1F}	0.08	0.34	0.66	1.01	1.25	2.32	3.94	10.05	25.98	58.13	77.20
	L_1	0.02	0.08	0.33	0.47	0.37	1.88	4.16	19.88	52.74	66.67	100
	L_{1F}	0.02	0.39	1.10	1.89	1.46	5.12	10.36	34.53	72.79	77.78	100
	S_2	0.00	0.02	0.06	0.17	0.35	1.21	3.01	16.81	54.48	93.70	99.28
	L_2	1.14	1.34	1.76	2.46	2.27	5.38	9.02	38.00	82.94	99.60	100
	S_3	17.73	9.75	5.41	4.30	3.84	4.73	7.77	27.45	66.70	96.40	99.67
	L_3	10.66	7.66	6.02	5.28	5.91	4.46	4.88	18.92	57.88	94.65	100
50	S_1	0.01	0.10	0.27	0.71	1.07	2.18	4.32	16.35	49.33	89.86	98.65
	S_{1F}	0.46	0.26	0.58	1.31	1.83	3.84	6.93	23.30	59.67	93.84	99.34
	L_1	0.02	0.04	0.62	1.17	2.00	6.10	13.71	58.18	94.87	100	100
	L_{1F}	0.04	0.35	1.98	3.19	4.80	11.14	21.56	65.00	97.72	100	100
	S_2	0.86	0.13	0.03	0.09	0.24	1.34	5.36	39.73	91.84	99.97	100.00
	L_2	45.04	22.10	9.06	5.34	4.08	6.56	17.46	69.98	99.26	100	100
	S_3	42.65	18.82	7.57	4.67	3.94	6.34	14.82	58.90	96.64	99.99	100.00
	L_3	53.28	29.00	13.28	7.56	8.25	5.48	12.50	53.36	96.48	100	100
100	S_1	0.76	0.10	0.22	0.53	1.00	3.18	6.70	32.87	81.42	99.70	100
	S_{1F}	8.92	1.92	0.82	1.06	1.83	5.27	10.95	43.77	88.61	99.87	100
	L_1	0.00	0.02	0.75	1.94	3.95	7.52	16.28	73.01	99.76	100	100
	L_{1F}	0.00	0.21	2.19	3.91	8.49	10.86	21.26	74.56	99.91	100	100
	S_2	13.83	1.88	0.14	0.12	0.24	2.20	10.14	72.21	99.72	100.00	100
	L_2	81.48	42.28	10.04	4.94	3.82	9.64	27.06	90.86	100.00	100	100
	S_3	73.82	35.03	10.71	5.73	3.77	9.70	26.51	87.39	99.96	100.00	100
	L_3	82.88	47.52	13.06	7.08	6.95	7.06	20.98	84.00	99.96	100	100

Table 4. Emperical power (%) of score test statistics S_1, S_{1F}, S_2, S_3, and LRT statistics L_1, L_{1F}, L_2, L_3 for testing over-dispersion with no covariates at $\alpha = 0.05$ when data are simulated from **Weibull distribution** based on 20,000 replications($c = 1$ indicates emperical level).

n	Statistic	c										
		0.70	0.80	0.90	0.95	1.00	1.10	1.20	1.50	2.00	3.00	4.00
5	S_1	0.00	0.00	0.00	0.00	0.00	0.04	0.04	0.50	3.33	14.48	27.77
	S_{1F}	0.00	0.00	0.02	0.04	0.06	0.14	0.38	1.84	7.91	23.82	38.02
	L_1	0.00	0.00	0.14	0.26	0.00	0.10	0.30	2.47	11.64	67.07	98.7
	L_{1F}	0.00	0.06	0.38	0.40	0.00	0.64	1.93	10.14	29.38	91.37	99.4
	S_2	0.00	0.04	0.22	0.26	0.49	1.23	2.55	10.45	31.70	68.36	85.86
	L_2	0.38	0.12	0.64	1.04	0.52	2.44	4.72	18.14	39.60	71.25	91.7
	S_3	5.05	3.55	2.63	2.31	2.38	3.03	5.13	14.84	38.22	73.14	88.56
	L_3	1.64	1.16	1.64	1.44	1.26	2.10	2.22	4.48	9.08	26.15	55.7
10	S_1	0.00	0.02	0.12	0.16	0.50	1.10	2.28	9.78	29.12	61.55	77.85
	S_{1F}	0.00	0.02	0.20	0.32	0.80	1.96	3.68	13.44	35.66	68.31	82.92
	L_1	0.00	0.02	0.10	0.14	0.07	0.66	2.17	13.73	40.53	98.43	100.0
	L_{1F}	0.02	0.06	0.26	0.52	0.38	2.82	7.61	32.36	66.00	99.48	100.0
	S_2	0.00	0.02	0.11	0.24	0.50	1.57	3.80	20.52	60.66	94.47	99.15
	L_2	1.50	0.44	0.50	0.84	1.11	3.40	8.86	36.14	73.76	96.20	99.7
	S_3	19.29	9.62	5.14	3.72	3.66	4.67	8.12	29.98	70.28	96.42	99.48
	L_3	6.40	3.36	2.54	2.22	2.96	1.64	2.90	8.64	31.06	86.95	99.6
20	S_1	0.00	0.02	0.16	0.36	0.64	2.40	6.24	25.60	63.58	93.19	98.75
	S_{1F}	0.00	0.03	0.25	0.62	1.25	3.93	8.87	32.14	70.56	95.55	99.29
	L_1	0.00	0.04	0.18	0.18	0.37	3.03	8.44	45.52	81.73	100	100
	L_{1F}	0.04	0.08	0.81	1.03	1.46	8.45	19.14	68.50	93.61	100	100
	S_2	0.00	0.00	0.04	0.12	0.35	1.70	6.64	40.76	89.69	99.91	100.00
	L_2	26.58	11.68	3.58	3.26	2.27	6.18	15.94	68.06	95.88	100	100
	S_3	42.8	19.32	7.45	4.73	3.84	5.88	14.26	57.03	94.70	99.97	100.00
	L_3	41.00	21.74	8.42	6.68	5.91	4.44	9.28	50.06	94.52	100	100
50	S_1	0.00	0.00	0.14	0.31	1.07	5.15	14.27	62.40	96.84	100	100
	S_{1F}	5.66	0.51	0.22	0.65	1.83	7.87	20.08	70.72	98.20	100	100
	L_1	0.00	0.02	0.29	0.66	2.00	12.28	36.64	94.01	99.83	100	100
	L_{1F}	0.00	0.08	0.68	2.02	4.80	21.04	49.00	96.96	99.98	100	100
	S_2	9.10	0.94	0.08	0.06	0.24	3.00	14.62	82.58	99.94	100	100
	L_2	86.56	54.08	16.58	7.70	4.08	12.06	37.88	96.92	99.94	100	100
	S_3	86.65	46.50	13.11	6.21	3.94	11.31	33.27	93.47	99.99	100	100
	L_3	92.54	65.30	22.04	9.90	8.25	9.66	31.04	94.34	99.98	100	100
100	S_1	25.72	1.35	0.05	0.16	1.00	8.04	28.07	90.85	99.98	100	100
	S_{1F}	69.66	15.13	1.30	0.70	1.83	12.60	37.38	94.64	99.98	100	100
	L_1	0.00	0.00	0.19	0.73	3.95	20.17	55.34	99.74	100	100	100
	L_{1F}	0.00	0.00	0.36	1.88	8.49	26.04	63.44	99.88	100	100	100
	S_2	68.79	12.68	0.76	0.18	0.24	5.76	32.36	98.83	100.00	100	100
	L_2	99.56	84.60	24.26	7.96	3.82	20.24	64.50	99.96	100	100	100
	S_3	99.44	77.79	23.42	9.03	3.77	19.59	60.34	99.87	100.00	100	100
	L_3	99.86	88.68	28.48	9.68	6.95	18.32	61.36	99.92	100	100	100

TABLE 5. Fatique data: Lifetime (in hours) for the Kevlar/Epoxy strand at 70% stress level.

1,051	1,337	1,389	1,921	1,942	2,322	3,629	4,006	4,012	4,063
4,921	5,445	5,620	5,817	5,905	5,956	6,068	6,121	6,473	7,501
7,886	8,108	8,546	8,666	8,831	9,106	9,711	9,806	10,205	10,396
10,861	11,026	11,214	11,362	11,604	11,608	11,745	11,762	11,895	12,044
13,520	13,670	14,110	14,496	15,395	16,179	17,092	17,568	17,568	

TABLE 6. Breakdown times (in minutes) at each seven levels of voltage.

Voltage levels	n_l	breakdown times							
26	3	5.79	1579.52	2323.7					
28	5	68.85	426.07	110.29	108.29	1067.6			
30	11	17.05	22.66	21.02	175.88	139.07	144.12	20.46	43.40
		194.90	47.30	7.74					
32	15	0.40	82.85	9.88	89.29	215.10	2.75	0.79	15.93
		3.91	0.27	0.69	100.58	27.80	13.95	53.24	
34	19	0.96	4.15	0.19	0.78	8.01	31.75	7.35	6.50
		8.27	33.91	32.52	3.16	4.85	2.78	4.67	1.31
		12.06	36.71	72.89					
36	15	1.97	0.59	2.58	1.69	2.71	25.50	0.35	0.99
		3.99	3.67	2.07	0.96	5.35	2.90	13.77	
38	8	0.47	0.73	1.40	0.74	0.39	1.13	0.09	2.38

Acknowledgement

This research was partially supported by the Natural Science and Engineering Research Council of Canada. Thanks to a referee for helpful comments which have improved the paper.

References

[1] D. F. Andrews and A. M. Herzberg, *Data. A Collection of Problems From Many Fields for the Student and Research Worker*, Springer Series in Statistics, Springer-Verlag, New York, 1985.

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

[3] N. E. Breslow, *Extra-Poisson variation in log-linear models*, Appl. Statist. **33** (1984), 38–44.

[4] E. Castillo and A. S. Hadi, *Fitting the generalized Pareto distribution to data*, J. Amer. Statist. Assoc. **92** (1997), 1609–1620.

[5] A. Chesher, *Testing of neglected heterogeneity*, Econometrika **52** (1984), 865–872.

[6] D. R. Cox, *Some remarks on overdispersion*, Biometrika **87** (1983), 269–274.

[7] D. R. Cox and D. Oakes, *Analysis of Survival Data*, Monographs on Statistics and Applied Probability, Chapman and Hall, London, 1984.

[8] D. Deng and S. R. Paul, *Score tests for zero-inflation and over-dispersion in generalized linear models*, Statist. Sinica **15** (2005), 257–276.
[9] B. Efron, *Why Isn't everyone a Bayesian?*, Amer. Statist. **40** (1986), 1–5.
[10] W. Nelson, *Applied Life Data Analysis*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, 1982.
[11] J. Neyman, *Optimal asymptotic tests for composite hypothesis*, in: Probability and Statistics: The Harold Cramér Volume, John Wiley & Sons, New York, 1959, pp. 213–234.
[12] S. R. Paul, *Test for the equality of several correlation coefficients*, Canad. J. Statist. **17** (1989), 217–227.
[13] R. L. Prentice, *A case-cohort design for epidemiologic cohort studies and disease prevention trials*, Biometrics **73** (1986), 1–11.
[14] C. R. Rao, *Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation*, Proc. Cambridge Philos. Soc. **44** (1947), 50–57.
[15] S. G. Self and K. Liang, *Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions*, J. Amer. Statist. Assoc. **82** (1987), 605–610.
[16] A. Shapiro, *Towards a unified theory of inequality constrained testing in multivariate analysis*, Internat. Statist. Rev. **56** (1988), 49–62.
[17] M. J. Silvapulle and P. Silvapulle, *A score test against one-sided alternatives*, J. Amer. Statist. Assoc. **90** (1995), 342–349.
[18] J. G. Skellam, *A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials*, J. Roy. Statist. Soc. Ser. B **10** (1948), 257–261.
[19] P. F. Thall, *Tests in the two-way layout of counts*, Comm. Statist. Theory Methods **21** (1992), 3017–3036.
[20] D. A. Williams, *The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity*, Biometrics **31** (1975), 949–952.

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

E-mail address: `smjp@uwindsor.ca`