CSISZAR DIVERGENCE FROM CONSTANT FAILURE RATE MODEL FOR GROUPED DATA

Bhaskar Bhattacharya
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901-4408
bhaskar@math.siu.edu

Key Words: Csiszar divergence; goodness-of-fit tests; Kullback-Leibler divergence; Satterthwaite approximation; Shannon entropy.

ABSTRACT

We propose a measure of divergence in failure rates of a system from the constant failure rate model for a grouped data situation. We use this measure to compare the divergences of several systems from the constant failure rate model and find the asymptotic distributions of the test statistics. Several applications are discussed to illustrate the procedure. In the context of testing the goodness-of-fit with the constant failure rate model, we conduct a simulation study which shows that this procedure compares favorably with the Pearson chi-square test and the likelihood ratio test procedures.

1. INTRODUCTION

Although the length of life of a system is a continuous random variable, the observations are often collected as grouped data due to limitations of monitoring devices. For example, if a system is monitored periodically, length of life can be defined as the maximum number of time periods successfully completed, or if a machine operates in cycles, length of life refers to the number of cycles successfully completed prior to the failure. Thus continuous type data are often collected at finite number of specified time points. Hence we assume that the time to failure T is discrete with k possible values labeled $t_1, t_2, \ldots, t_k$. We assume that all failures are observed; thus t_k could be large. For $1 \leq i \leq k$, let p_i denote the probability that an object chosen at random will fail at time t_i, where $p_i > 0$, $\sum_{i=1}^{k} p_i = 1$. Let $\boldsymbol{p} = (p_1, \ldots, p_k)^T$. Let θ_i be the conditional probability that an object will fail at time t_i given that it has not failed before; thus θ_i is the discrete failure rate at t_i and can be expressed as $\theta_i = p_i / \sum_{j=i}^{k} p_j$, $1 \leq i \leq k-1$. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{k-1})^T$.

Using a nonparametric approach, we propose a measure to detect deviations of the failure rates of a system from the constant failure rate (CF) model. This measure is based on the Csiszar divergence between the failure rate of the system and the constant failure rate of the system. When comparing between different models, we also provide test statistics to test hypotheses concerning divergences from the CF model.

Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_k)$ be two probability vectors (with positive coordinates which add to 1). The *Csiszar divergence* between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$D_\phi(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^{k} y_j \phi\left(\frac{x_j}{y_j}\right),$$

where ϕ is a nonincreasing twice continuously differentiable strictly convex function in $[0, \infty)$. It can be verified [2] that $D_\phi(\boldsymbol{x}, \boldsymbol{y}) \geq \phi(1)$ and equality holds if and only if $\boldsymbol{x} = \boldsymbol{y}$.

Let $v_j = \theta_j / \sum_{i=1}^{k-1} \theta_i$, $j = 1, \ldots, k-1$ and $v_0 = (k-1)^{-1}$. Let $\boldsymbol{v} = (v_1, v_2, \ldots, v_{k-1})^T$ and $\boldsymbol{v}_0$ be the vector of length $k-1$ with each coordinate being equal to v_0. Consider

$$U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = \frac{D_\phi(\boldsymbol{v}_0, \boldsymbol{v}) - \phi(1)}{\phi(v_0) - \phi(1)}. \tag{1.1}$$

When $\phi(t) = -\ln t$, writing D_{KL} for D_ϕ, we obtain

$$D_{KL}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^{k} y_j \ln \left(\frac{y_j}{x_j}\right),$$

which is also well-known as the *Kullback-Leibler divergence* between $\boldsymbol{x}$ and $\boldsymbol{y}$. Also, writing U_{KL} for U_ϕ, the expression in (1.1) simplifies as

$$U_{KL}(\boldsymbol{v}_0, \boldsymbol{v}) = 1 - \frac{H(\boldsymbol{v})}{\ln (k-1)}$$

where $H(\boldsymbol{v}) = -\sum_{j=1}^{k-1} v_j \ln v_j$ is the *Shannon entropy* of $\boldsymbol{v}$.

It can be shown that $0 \leq U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) < 1$. Clearly, a system has CF model if and only if $U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = 0$. Since $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$ is a divergence measure which provides a measure of 'distance' from uniformity, a higher value of divergence means farther 'distance' from uniformity. Thus the degree of departure from the CF model increases if and only if the value of $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$ increases.

Testing for constant failure rate in a grouped data situation may be considered as a test of discrete exponentiality which is also equivalent to a truncated geometric distribution. Tests of fit based on grouped data using likelihood ratio test or the Pearson's chi-square test is treated in [9, page 440]. Also using grouped data, [1] has considered testing the hypothesis of the constant failure rate against an increasing order restriction. The practical motivation for testing the constant failure rate model in a grouped data situation is that in many cases data are collected in this way (see examples in Section 4). It is desirable that the failure rates are preserved over the time intervals, otherwise some maintenance and/or improvements would be warranted to increase the efficiency of the system. The proposed tests in this paper provide a way to verify this. However in practice the failure rates may not be constant. So we provide a procedure to test if the 'distance' of the failure rates from the constant failure rate model in terms of Csiszar divergence is a specified amount (of course, smaller divergence is preferable). We also show how to compare among several systems based on their 'distances' from the constant failure rate model in terms of Csiszar divergence.

In Section 2, we discuss the asymptotic distribution of $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$ for different situations. The testing of divergences is treated in Section 3. In Section 4, several examples are illustrated along with a simulation study which shows that our procedure to test for CF model for a system using the KL divergence is no worse than the standard goodness-of-fit procedures of Pearson's chi-square and likelihood ratio tests.

2. ASYMPTOTIC DISTRIBUTION OF $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$

Let $n = \sum_{i=1}^k n_i$ items are put on test, where n_i denote the observed frequency for the ith time point. Assuming that the $\{n_i\}$ result from a full multinomial sampling, the sample version of the discrepancy $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$ is constructed by estimating p_i with $\hat{p}_i = n_i/n$ in the expression of $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$. Let $\hat{\boldsymbol{p}} = (\hat{p}_1, \ldots, \hat{p}_k)^T$.

Theorem 1. For $\boldsymbol{v} \neq \boldsymbol{v}_0$, the asymptotic distribution of $U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}})$ is given by

$$\sqrt{n}(U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) - U_\phi(\boldsymbol{v}_0, \boldsymbol{v})) \xrightarrow{\mathcal{L}} N(0, \sigma^2), \tag{2.1}$$

as $n \to \infty$ where

$$\begin{aligned} \sigma^2 &= \tfrac{1}{(\phi(v_0)-\phi(1))^2} \textstyle\sum_{j=1}^{k-1} \tfrac{\theta_j(1-\theta_j)}{P_j} h_i^2, \\ h_i &= -\tfrac{D_\phi}{T} + \tfrac{1}{T}\phi(b_i) - \tfrac{1}{(k-1)\theta_i}\phi'(b_i) + \tfrac{1}{(k-1)T} \textstyle\sum_{j=1}^{k-1} \phi'(b_j), \\ b_i &= \tfrac{T}{(k-1)\theta_i}, \ \ T = \textstyle\sum_{j=1}^{k-1} \theta_j, \ \ P_j = \textstyle\sum_{i=j}^{k} p_i. \end{aligned}$$

Proof. Let $\hat{\theta}_i = \hat{p}_i / \sum_{j=i}^{k} \hat{p}_j$, $1 \leq i \leq k-1$ and $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{k-1})^T$. Since $\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} = \boldsymbol{pI} - \boldsymbol{p^T p}$, it follows using the delta method [3] that $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is a diagonal matrix with jth diagonal entry is $\theta_j(1-\theta_j)/P_j$.

Using Taylor series expansion we obtain

$$U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) + \boldsymbol{h}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + R_1,$$

where $\boldsymbol{h} = (h_1, \ldots, h_{k-1})^T$ are the values given above, and $R_1 \to 0$ as $n \to \infty$. Thus it follows that $\sqrt{n}(U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) - U_\phi(\boldsymbol{v}_0, \boldsymbol{v})) \xrightarrow{\mathcal{L}} N(0, \boldsymbol{h}^T \Sigma_{\boldsymbol{\theta}} \boldsymbol{h})$. After algebraic manipulations, it follows that $\sigma^2 = \boldsymbol{h}^T \Sigma_{\boldsymbol{\theta}} \boldsymbol{h}$ which proves the theorem.

Remark 1. When $\phi(t) = -\ln t$, the distributional result of (2.1) simplifies to

$$\sqrt{n}(U_{KL}(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) - U_{KL}(\boldsymbol{v}_0, \boldsymbol{v})) \xrightarrow{\mathcal{L}} N(0, \sigma_{KL}^2)$$

where

$$\sigma_{KL}^2 = \frac{1}{T^2 \ln (k-1)} \sum_{j=1}^{k-1} \frac{\theta_j(1-\theta_j)}{P_j} \left(\ln \theta_j - \frac{\sum_{s=1}^{k-1} \theta_s \ln \theta_s}{T} \right)^2 .$$

Remark 2. Theorem 1 can be used to obtain the asymptotic confidence intervals. Let $\hat{\sigma}$ be the estimated divergence standard deviation obtained from Theorem 1 by estimating p_i with $\hat{p}_i$. Then $U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) \pm z_{\alpha/2}\hat{\sigma}/\sqrt{n}$ is an approximate $100(1-\alpha)\%$ confidence interval for $U_\phi(\boldsymbol{v}_0, \boldsymbol{v})$ where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point from the normal distribution. Also, $U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}_1) - U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}_2) \pm z_{\alpha/2}\sqrt{\hat{\sigma}_1^2/m_1 + \hat{\sigma}_2^2/m_2}$ is an approximate $100(1-\alpha)\%$ confidence interval for $U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_1) - U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_2)$, where $m_1,\ m_2$ are the sample sizes and $\hat{\sigma}_1,\ \hat{\sigma}_2$ are the estimated divergence standard deviations for populations 1 and 2, respectively, obtained from Theorem 1.

However, Theorem 1 is only valid when $\boldsymbol{v} \neq \boldsymbol{v}_0$. For the case of $\boldsymbol{v} = \boldsymbol{v}_0$, the following result can be used.

Theorem 2. For $\boldsymbol{v} = \boldsymbol{v}_0$, the asymptotic distribution of $U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}})$ is given by

$$2n(\phi(v_0) - \phi(1))^2 U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) \xrightarrow{\mathcal{L}} \sum_{j=1}^{k-1} \lambda_j \chi_j^2, \tag{2.2}$$

as $n \to \infty$ where χ_j^2 are independent chi-square with 1 degree of freedom, λ_j are eigen values of the matrix $H\Sigma_{\boldsymbol{\theta}}$, the elements of the matrix H are:

$$\begin{aligned} h_{ii} &= \tfrac{2}{T^2}(D_\phi - \phi(b_i) + b_i\phi'(b_i)) + \tfrac{(k-1)b_i-2}{(k-1)^2\theta_i^2}\phi''(b_i) \\ &\quad -\tfrac{2}{(k-1)T^2}\textstyle\sum_{s=1}^{k-1}\phi'(b_s) + \tfrac{1}{(k-1)^2T}\sum_{s=1}^{k-1}\tfrac{\phi''(b_s)}{\theta_s}, \\ h_{ij} &= \tfrac{2D_\phi}{T^2} - \tfrac{1}{T^2}(\phi(b_i) + \phi(b_j)) + \tfrac{1}{(k-1)T}\left(\tfrac{\phi'(b_i)}{\theta_i} + \tfrac{\phi'(b_j)}{\theta_j}\right) \\ &\quad -\tfrac{1}{(k-1)^2}\left(\tfrac{\phi''(b_i)}{\theta_i^2} + \tfrac{\phi''(b_j)}{\theta_j^2}\right) - \tfrac{2}{(k-1)T^2}\textstyle\sum_{s=1}^{k-1}\phi'(b_s) + \tfrac{1}{(k-1)^2T}\sum_{s=1}^{k-1}\tfrac{\phi''(b_s)}{\theta_s}, \end{aligned}$$

for $i \neq j$ and b_j's are defined in Theorem 1.

Proof. Using the Taylor series expansion of $U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}})$ at the point $\boldsymbol{v}$, we find that the first two terms are zero when $\boldsymbol{v} = \boldsymbol{v}_0$ and we obtain

$$(\phi(v_0) - \phi(1))^2 U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) = \frac{1}{2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T H (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + R_2$$

where $R_2 \to 0$ as $n \to \infty$ and H is the matrix of second derivatives obtained by differentiating D_ϕ as given above. The result now follows by using Theorem 4.4.4 of [4].

Remark 3. When $\phi(t) = -\ln t$, we obtain

$$\begin{aligned} h_{ii} &= \tfrac{1}{T^2}\left(\tfrac{2}{T}\textstyle\sum_{s=1}^{k-1}\theta_s \ln\theta_s - 2\ln\theta_i - 1 + \tfrac{T}{\theta_i}\right), \\ h_{ij} &= \tfrac{1}{T^2}\left(\tfrac{2}{T}\textstyle\sum_{s=1}^{k-1}\theta_s \ln\theta_s - \ln(\theta_i\theta_j) - 1\right), \end{aligned}$$

for $i \neq j$ corresponding to the Kullback-Leibler divergence.

Remark 4. The distribution $\sum_{j=1}^{k-1} \lambda_j \chi_j^2$ in (2.2) is difficult to calculate. We consider a version of Satterthwaite approximation [5] as follows:

$$\sum_{j=1}^{k-1} \lambda_j \chi_j^2 \sim a\chi_{k-1}^2 + b$$

where a and b are the solutions to the set of equations

$$\begin{aligned} tr(H\Sigma_{\boldsymbol{\theta}}) &= a(k-1) + b \\ tr(H\Sigma_{\boldsymbol{\theta}} H\Sigma_{\boldsymbol{\theta}}) &= a^2(k-1). \end{aligned}$$

Another simpler, but less efficient approach, suggested by [6] and [7], is to approximate the distribution in (2.2) by $\lambda_0 \chi_{k-1}^2$ where $\lambda_0 = tr(H\Sigma_{\boldsymbol{\theta}})$. In our case λ_0 can be expressed as

$$\lambda_0 = \sum_{j=1}^{k-1} \frac{h_{jj}\theta_j(1-\theta_j)}{P_j}. \tag{2.3}$$

We have used the Satterthwaite approximation as described above in the simulation study in Section 4.

3. COMPARING DIVERGENCES OF DIFFERENT SYSTEMS

The results obtained in the previous section can be used to compare divergences of different systems.

Result 1. Suppose we like to test the divergence to a given value. Thus consider test of

$$H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = U_0 \neq 0.$$

The statistic

$$V_1 = \frac{\sqrt{n}(U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}) - U_0)}{\hat{\sigma}} \sim N(0,1)$$

can be used where $\hat{\sigma}$ is the estimated divergence standard deviation given in Theorem 1.

Result 2. To test $H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = 0$, the statistic

$$V_2 = \frac{2n(\phi(v_0) - \phi(1))^2 U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}})}{\lambda_0} \sim \chi_{k-1}^2$$

can be used where λ_0 is defined in (2.3).

Result 3. Suppose we like to test g divergences to a given value. Thus consider test of

$$H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_1) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_2) = \cdots = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_g) = U_0 \neq 0.$$

The statistic

$$V_3 = \sum_{s=1}^{g} \frac{n_s(U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}_s) - U_0)^2}{\hat{\sigma}_s^2} \sim \chi_g^2$$

can be used where n_s is the sample size used and $\hat{\sigma}_s$ is the estimated divergence standard deviation for the sth sample.

Result 4. To test equality of g divergences, that is,

$$H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_1) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_2) = \cdots = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_g),$$

the following statistic can be used

$$V_4 = \sum_{s=1}^{g} \frac{n_s(U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}_s) - \overline{U})^2}{\hat{\sigma}_s^2} \sim \chi_{g-1}^2$$

where

$$\overline{U} = \frac{\sum_{s=1}^{g} \frac{n_s U_\phi(\boldsymbol{v}_0, \hat{\boldsymbol{v}}_s)}{\hat{\sigma}_s^2}}{\sum_{s=1}^{g} \frac{n_s}{\hat{\sigma}_s^2}}.$$

4. EXAMPLES AND SIMULATION

We consider two data sets, from [8], reproduced in Tables 1 and 2 below, to test if the CF model hold. Table 1 reports data concerning the number of thousand miles to failure for 104 bus motors.

Table 1: Frequency Distribution for Bus Motor Failure Data

# of Miles (000's)	0-20	20-40	40-60	60-80	80-100	100-120	$\geq$ 120
Observed frequency	19	13	13	15	15	18	11

We consider testing

$$H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = 0.$$

From Result 2, using $\phi(t) = -\ln t$, we obtain $V_2 = 14.9361$ with a p-value of 0.0368. Thus we would reject the hypothesis of constant failure rate for this data.

The data in Table 2, gives lifetimes of 100 V600 vacuum tubes.

Table 2: Lifetimes of V600 vacuum tubes

Lifetime (hours)	0-100	100-200	200-300	300-400	400-600	600-800	≥ 800
Observed frequency	29	22	12	10	10	9	8

When testing $H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}) = 0$, using $\phi(t) = -\ln t$, we obtain $V_2 = 3.2632$ with a p-value of 0.8596. Thus we would not reject the hypothesis of constant failure rate for this data.

We consider a survival data of cancer patients from [9], reproduced in Table 3 below. Time intervals represent months from treatment. The 285 patients are classified into four categories which correspond to two levels for each of two factors A and B. Thus there are four patient groups A_1B_1, A_1B_2, A_2B_1, A_2B_2 and the data give the survival experience of each of them.

Table 3: Survival Data for Cancer Patients Classified by Two Factors

Lifetime (months)	A_1B_1	A_1B_2	A_2B_1	A_2B_2
0-3	15	18	12	9
3-6	3	14	7	8
6-9	14	8	8	11
9-12	17	16	7	8
12-15	7	11	9	5
15-18	6	12	5	9
18-21	4	4	2	3
$\geq$21	1	2	2	4

To test whether the degree of divergence from constant failure rate is same for each of the four groups, we like to test

$$H_0 : U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_1) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_2) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_3) = U_\phi(\boldsymbol{v}_0, \boldsymbol{v}_4)$$

where $\boldsymbol{v}_i$ is from the ith group. For $\phi(t) = -\ln t$, we obtain $V_4 = 1.1217$ which has a p-value of 0.7718. Hence we fail to reject H_0, that is, these four groups do not differ with respect to the Kullback-Leibler divergence from the constant failure rate model.

To compare the effectiveness of the proposed measures with the standard chi-square tests and to investigate their small sample behavior a modest simulation study is performed. We consider the case of departures from the CF model using the measure $U_{KL}(\boldsymbol{v}_0, \boldsymbol{v})$ and the standard goodness-of-fit tests, the likelihood ratio and the Pearson chi-square tests [9], and also [10]. For the purpose of the simulation we have considered the result of Theorem 2 with $\boldsymbol{v}_0$ being a constant vector along with the Satterthwaite approximation given in Remark 4.

We have chosen θ_i's under the CF model and also different θ_i's to represent deviation from the CF model in a variety of ways. The p_i's are obtained from the θ_i's by using the transformation $p_i = \theta_i \prod_{j=1}^{i-1}(1-\theta_j),\ 1 \le i \le k-1,\ p_k = \prod_{j=1}^{k-1}(1-\theta_j)$. Samples are taken from multinomial distribution with these probabilities and three above tests are performed.

The likelihood ratio test (G^2) and the chi-square (X^2) tests are given by

$$G^2 = \sum_{i-1}^{k} n_i \ln \frac{n_i}{e_i} \quad \text{and} \quad X^2 = \sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i}$$

where $e_i = n\hat{p}_i^0$ are the expected values, $\hat{p}_i^0$ are obtained using the above transformations by replacing each θ_j by $\sum_{i=1}^{k-1} \hat{p}_i / \sum_{i=1}^{k-1} \sum_{j=i}^{k} \hat{p}_j$. We have considered $k = 5$ and $\alpha = .05$ with 5,000 repetitions in each case. We consider samples of sizes 30, 60, 100 and 200. The estimated powers are given in Table 4.

The powers increase as the sample size increases for all three tests. Within the scope of this simulation study, the test using the KL divergence compares favorably with the standard G^2 and X^2 tests as none of the three tests perform uniformly better over any other. The powers were computed at other values which produced similar results.

ACKNOWLEDGMENTS

The author thanks two referees for their comments which have improved this paper.

Table 4: Estimated power of different tests

n	theta	KL test	X^2 test	G^2 test	theta	KL test	X^2 test	G^2 test
30	(.1,.1,.1,.1)	0.057	0.045	0.076	(.8,.3,.8,.3)	0.086	0.531	0.533
60		0.046	0.044	0.057		0.216	0.886	0.884
100		0.044	0.050	0.055		0.372	0.987	0.986
200		0.043	0.046	0.051		0.760	1.000	1.000
30	(.2,.7,.7,.7)	0.540	0.968	0.963	(.7,.7,.2,.2)	0.112	0.325	0.375
60		0.814	1.000	1.000		0.406	0.717	0.724
100		0.975	1.000	1.000		0.647	0.922	0.916
200		1.000	1.000	1.000		0.966	0.997	0.998
30	(.6,.6,.6,.1)	0.039	0.077	0.120	(.5,.65,.8,.95)	0.334	0.185	0.247
60		0.458	0.195	0.370		0.748	0.428	0.577
100		0.627	0.415	0.647		0.845	0.722	0.850
200		0.946	0.870	0.929		0.944	0.983	0.993

BIBLIOGRAPHY

[1] Bhattacharya, B. Restricted tests for and against the increasing failure rate ortdering on multinomial parameters. Statist. Probab. Lett. **1995**, *25*, 309-316.

[2] Csiszar, I. Information type measures of difference of probability distributions and indirect observations. Studia. Sci. Math. Hungar. **1967**, *2*, 299-318.

[3] Serfling, R. J. *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, **1980**.

[4] Graybill, F. *Theory and application of the linear model*, Duxbury Press, Boston, **1976**.

[5] Satterthwaite, F. E. An approximate distribution of estimates of variance components. Biometrika **1946**, *2*, 110-114.

[6] Kotz, S., Johnson, N.L., Boyd, D.M. Series representation of distributions of quadratic forms in normal variables. I. Central case. A.M.S. **1967**, 823-837.

[7] Rao, J.N.K. and Scott, A.J. The analysis of categorical data from complex surveys: chi-square tests for goodness of fit and independence in two-way tables. J. Am. Statist. Assoc. **1981**, *76*, 221-230.

[8] Davis, D. J. An analysis of some failure data. J. Am. Statist. Assoc. **1952**, *47*, 113-150.

[9] Lawless, J. F. *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York, **1982**.

[10] Agresti, A. *Categorical data analysis*, John Wiley & Sons, Inc., New York, **1990**.