ANNALES DE L'I. H. P., SECTION B

M. A. STEPHENS

Components of goodness-of-fit statistics

Annales de l'I. H. P., section B, tome 10, nº 1 (1974), p. 37-54
<http://www.numdam.org/item?id=AIHPB_1974__10_1_37_0>

Ann. Inst. Henri Poincaré, Section B :
Vol. X, n° 1, 1974, p. 37-54. *Calcul des Probabilités et Statistique.*

Components of goodness-of-fit statistics

by

M. A. STEPHENS

SUMMARY. — A test that a random sample comes from a completely specified continuous distribution can be reduced to testing that a given set of n values x_i is uniformly distributed between 0 and 1. Suppose $F_n(x)$ is the E. D. F. (empirical distribution function) of the x_i, i. e. $F_n(x)$ = (number of $x_i \leqslant x)/n$, and consider $y_n(x) = F_n(x)-x$. Several well-known goodness-of-fit statistics are based on integrals involving $y_n(x)$. In a recent paper, Durbin and Knott suggested a partition of the Cramervon-Mises statistic W^2 into an infinite set of components. These are related to the coefficients when $y_n(x)$ is expanded as a Fourier sine series in the interval 0, 1. The components were compared, for asymptotic power, with the entire statistic W^2.

This idea is now extended in various ways. In particular the Watson statistic U^2 and the Anderson-Darling A^2 are partitioned and the components examined. U^2 depends on a Fourier Series expansion and A^2 on associated Legendre functions. Components of W^2 and U^2 have interesting distributions, and asymptotic power studies of a more general nature than those of Durbin and Knott bring out properties of these components and also of the parent statistics.

RÉSUMÉ. — Un test qu'un échantillon aléatoire provient d'une distribution continue donnée peut être réduit au test qu'un ensemble de n valeurs x_i est uniformément distribué entre 0 et 1. Supposons que $F_n(x)$ est la distribution empirique des x_i, c'est-à-dire $F_n(x)$ = (nombre de $x_i \leqslant x)/n$, et considérons $y_n(x) = F_n(x)-x$. Plusieurs statistiques bien connues d'adéquation sont basées sur des intégrales comprenant $y_n(x)$. Dans un article récent, Durbin et Knott ont suggéré une partition de la statistique W^2 de Cramer-von-Mises en une infinité de composantes. Celles-ci sont reliées aux coefficients de $y_n(x)$ en séries de Fourier en sinus sur l'intervalle 0, 1. Les

composantes ont été comparées avec la statistique W^2 asymptotiquement.

Cette idée est étendue dans cet article de diverses manières. En particulier les statistiques de Watson U^2 et de Anderson-Darling A^2 sont décomposées et leurs composantes examinées. La statistique U^2 dépend d'un développement en série de Fourier, tandis que A^2 est développée en fonction de Legendre associées. Les composantes de W^2 et de U^2 ont des distributions intéressantes et des formes asymptotiques d'une nature plus générale que celles de Durbin et Knott qui font apparaître les particularités de ces composantes et des statistiques apparentées.

INTRODUCTION

In a recent paper, Durbin and Knott (1972) partitioned the Cramer-von Mises statistic W^2, used for tests of goodness-of-fit, into an infinite set of components. For the case when the hypothesised distribution function is completely specified, these components have the interesting property that they all have the same distribution, to within a constant. Durbin and Knott (henceforth referred to by DK) suggested the use of some of the components as test statistics, and investigated their asymptotic powers, compared with W^2, for the case when the distribution tested on H_0 is N(0, 1), and when the family of alternatives has a different mean or a different variance. They also suggested how the Anderson-Darling statistic A^2 could be similarly partitioned and gave some power results for A^2 and the Watson statistic U^2. These statistics are all defined below. The purpose of this paper is to provide various extensions of the DK results with the intention of further illuminating the properties of the components and the entire statistics, particularly with respect to power. The extensions are

a) U^2 is also partitioned into components, which also have identical and well known distributions;

b) the asymptotic power studies for W^2, U^2 and A^2 are extended to the more general alternative of changes in both mean *and* variance of the normal distribution, and components of U^2 and A^2 are included;

c) power studies (this phrase always refers to asymptotic power) for the three statistics are given for a test for the exponential distribution.

d) A method of showing power on contour maps is introduced as a very straightforward way to compare statistics.

The results of these studies bring out the relative merits of W^2 and U^2 in detecting primarily mean shifts or variance shifts respectively, and also the overall superiority of A^2 (though results are given to suggest that the results for A^2 in DK Table 4, variance shift, are in error). They also reveal that, among the components, the first component of U^2 is better than any component of W^2 in a test for normality against the realistic alternative which allows for a shift in both mean and variance.

The extensions above involve coefficients which enter also into related work on tests for normality and exponentiality when parameters in the distribution must be estimated from the data; this is briefly described in Section 10.

1. THE TEST STATISTICS

Throughout the paper, we shall follow closely the DK notation. The definitions of the test statistics are:

$$W^2 = n \int (F_n(x) - x)^2 dx\,; \quad U^2 = n \int \left(F_n(x) - x - \int (F_n(x) - x)dx\right)^2 dx$$

$$A^2 = n \int ((F_n(x) - x)^2 / x(1 - x))dx$$

where the original observations are $y_1, y_2 \dots y_n$ from, on H_0, a distribution $F(y)$; $x_i = F(y_i)$, and $F_n(x)$ is the empirical distribution function (EDF) of the x_i. The integrals are all between limits 0 and 1; the statistics often have a suffix n, as in DK, but this will be omitted. For comments on these statistics see, e. g. DK, and also Pearson and Hartley (1972), Stephens (1970).

2. COMPONENTS OF U^2

Define $y_n(x) = \sqrt{n}(F_n(x) - x)$, $0 \leq x \leq 1$. For W^2, DK expand $y_n(x)$ as a Fourier sine series in the interval 0,1; if

$$F_n(x) = \sum_j b_j \sin(j\pi x), \quad \text{then} \quad b_j = \sqrt{2}\,\frac{z_{nj}}{j\pi},$$

where

$$z_{nj} = \sqrt{2}j\pi \int_0^1 y_n(x) \sin(j\pi x)dx = \sqrt{\frac{2}{n}} \sum_{i=1}^{n} \cos(j\pi x_i). \tag{1}$$

Further,

$$W^2 = \sum_{j=1}^{\infty} \frac{z_{nj}^2}{j^2\pi^2} \tag{2}$$

Define also

$$z_{nj}^* = 2j\pi \int_0^1 y_n(x) \cos(j\pi x)dx = -\sqrt{\frac{2}{n}} \sum_{i=1}^{n} \sin j\pi x_i. \tag{3}$$

The last equality in (1) is given by DK and the result (3) follows in the same way.

For U^2, expand $y_n(x)$ as a Fourier series into both sine and cosine components:

$$y_n(x) = \sum_{j=0}^{\infty} a_j \cos(2\pi jx) + \sum_{j=1}^{\infty} b_j \sin(2\pi jx);$$

then

$$a_0 = \int_0^1 y_n(x)dx = \sqrt{n} \int_0^1 (F_n(x) - x)dx;$$

$$a_j = 2 \int_0^1 y_n(x) \cos(2\pi jx)dx, \qquad j = 1, 2, \ldots$$

$$b_j = 2 \int_0^1 y_n(x) \sin(2\pi jx)dx. \qquad j = 1, 2, \ldots$$

It then follows that

$$U^2 = \int_0^1 \left(\sum_{j=1}^{\infty} a_j \cos(2\pi jx) + \sum_{j=1}^{\infty} b_j \sin(2\pi jx) \right)^2 dx$$

$$= \sum_{j=2}^{\infty}{}^{*} (z_{nj}^{*2} + z_{nj}^2)/j^2\pi^2 \tag{4}$$

where Σ^* will always indicate summation *over even values only* of the index. Il we now define *components*

$$S_j = \sum_{i=1}^{n} \cos(j\pi x_i), \quad T_j = \sum_{i=1}^{n} \sin(j\pi x_i) \tag{5}$$

and

$$R_j^2 = S_j^2 + T_j^2,$$

we have the parallel results:

$$W^2 = \frac{2}{n\pi^2} \sum_{j=1}^{\infty} \frac{S_j^2}{j^2} \quad \text{and} \quad U^2 = \frac{2}{n\pi^2} \sum_{j=2}^{\infty}{}^{*} \frac{R_j^2}{j^2} \tag{6}$$

DK observed that the S_j all have the same null distribution; all the R_j, for even j, also have the same null distribution. Further, both these distributions arise when n random unit vectors are summed.

3. REPRESENTATION ON A CIRCLE

This emerges clearly when the x_i are recorded around a circle. Suppose OS, OT are rectangular axes for coordinates s, t, and let the circle, centre 0, radius 1 cut OS in $A = (1, 0)$ and $B = (-1, 0)$. Mark points P_i on the circle so that the polar coordinate $\theta_i = \pi x_i$; θ is measured anticlockwise from OS as initial line. The arc length AP_i is then also πx_i; when the x_i are uniformly distributed between 0 and 1 the points P_i are uniform around the arc AB. Let the vectors $\underset{\sim}{OP}_i$, $i = 1, 2, \ldots, n$ have vector sum (resultant) $\underset{\sim}{R}_1$, length R_1. Then $S_1 = \Sigma \cos(\pi x_i)$ is the component of $\underset{\sim}{R}_1$ on the s-axis, and $T_1 = \Sigma \sin(\pi x_i)$ is the component on the t-axis.

Suppose now each θ_i is multiplied by an integer j, and then the points, now called P_i^*, marked on the circle. Let the resultant of vectors OP_i^* be called $\underset{\sim}{R}_j$, length R_j; clearly S_j and T_j are the components of $\underset{\sim}{R}_j$ on the two axes. If the x_i were originally uniform, the set P_i^* will be uniformly distributed around the appropriate portion of are length $j\pi$; it is then clear that all S_j for j even, have the same distribution, that of S_2; it is not so easily seen that S_j, for j odd, also has this distribution, but this is proved by DK. Also, T_j, j even, has the distribution of S_2, but not T_j when j is odd, because the original observations were on a semicircle and symmetry is lost. It follows that R_{2j}, all j, has the same distribution as R_2. The two basic distributions, those of R_2 and of S_1 (or S_2) are therefore those of the length of the resultant of n unit vectors randomly oriented around a circle, and of a *component* of this resultant on any arbitrary axis. These distributions have arisen in many applications, and that of R_2 in particular has been much studied. Significance points for R_2 are given in Greenwood and Durand (1955). Good approximate values for both R_2 and S_2 statistics are obtainable from tables in Stephens (1969). For reasonably large n, z_{nj} is well approximated by a standard normal distribution, and $2R_2^2/n$ by the χ^2 distribution, i. e. $u = R_2^2/n$ has the distribution $F(u) = 1 - e^{-u}$, $u \geqslant 0$.

Thus, to make a test based on the $j - th$ component of W^2 or U^2 one can calculate the related value of S_j or of R_j and refer it to the S_2 or R_2 distribution using significance points as given above. DK also gives exact points for $z_{nj} = (2/n)^{1/2}S_j$.

4. NON-NULL DISTRIBUTIONS. GENERAL

The circle diagram helps also in examining the distributions of components when H_0 is no longer true. In what follows we are expanding (for W^2) results put very succinctly in DK, § 5, and adding new results for U^2. Suppose the hypothesized distribution $F(y)$, in the W^2 and U^2 definition, is not the true distribution of y; let this be $F_T(y)$. When the mean of $F_T(y)$ is greater than that of $F(y)$, the values of x_i move towards 1, and the points P_i move round the circle towards B. The first component of W^2 will tend to be negative, so when testing H_0 against this alternative, S_1 is tested for significance in the lower tail. Conversely, for negative mean shift, S_1 is tested for significance in the upper tail. When angles θ_i are all doubled, it is clear that the points all tend to lie in the lower semicircle for positive mean shift and *vice versa*. In either case S_2 will not be large, though T_2 would be. When the variance of $F_T(y)$ is greater than that of $F(y)$, the values of x_i move towards 0 and 1; S_1 will not be large, but when the angles are doubled, the points tend to lie in the first and fourth quadrants, and S_2 will be large and positive. When the true variance is smaller than that of $F(y)$, S_2 will be negative. Thus S_2 tends to detect a change in variance. R_2, being a combination of S_2 and T_2, will grow larger for either a mean shift or a variance shift. These comments will be amplified later with numerical comparisons. We shall see that, for tests in the normal family, S_1 detects *only* mean shift and S_2 *only* variance shift, but R_2 detects a combination of both. The weighting of the components gives S_1 the main role in W^2, and R_2 the main role in U^2; thus W^2 tends to detect a difference in mean from that given by H_0, and U^2 a difference in variance. This confirms previous observation (DK; Stephens, 1970; Pearson and Hartley, 1972).

5. NON-NULL DISTRIBUTIONS: COMPONENTS OF W^2 AND OF U^2

DK show how asymptotic power may be calculated. They make the assumption that the tested distribution is of the form $F(y, \theta)$ with $\underset{\sim}{\theta}$ a vector of p parameters. On H_0, $\underset{\sim}{\theta}$ is specified equal to $\underset{\sim}{\theta}_0$, and for H_A, the alternative to H_0, $\underset{\sim}{\theta}$ takes the form $\underset{\sim}{\theta}_0 + \underset{\sim}{\gamma}/\sqrt{n}$, where $\underset{\sim}{\gamma}$ is a constant vector.

Thus for the alternative the sample is assumed to come always from the same *family* of distributions, with $\underset{\sim}{\theta}$ approaching its H_0 value as $n \to \infty$. This is restrictive compared with the situation usually envisaged in goodness-of-fit tests, but enables power studies to be made. Basically this is because DK show that in such cases, the z_{nj} are asymptotically normally distributed with variance 1 as before, but with non-zero means when H_0 is not true; thus the power of a component can easily be found. Let z_j be a random variable with the limiting distribution of z_{nj}, as $n \to \infty$. Let θ_i be the i-th component of $\underset{\sim}{\theta}$. Define $g_i(x)$ to be $\partial F(y, \theta)/\partial\theta_i$, evaluated at $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$, and written in terms of $x = F(y, \underset{\sim}{\theta}_0)$; let $g_i(x)$ be the i-th component of a vector $g(x)$. Let vector $\underset{\sim}{\delta}_j$ have the i-th component

$$\delta_{ji} = \sqrt{2}\pi j \int_0^1 \sin(\pi j x) g_i(x) dx, \quad j = 1, 2, \ldots \tag{7}$$

DK show that the mean of z_j is $\underset{\sim}{\gamma}'\underset{\sim}{\delta}_j$.

Illustrations. — They illustrate for the two cases where the distribution tested is the normal, and θ (with only one component) represents either the mean or the variance. Thus on H_0 the tested distribution if N(0, 1) and for the alternative called *Mean Shift* it is $N(\gamma_1/\sqrt{n}, 1)$ and for that called *Variance Shift* it is $N(0, 1 + \gamma_2/\sqrt{n})$. For the extension of their results to cover both a mean shift *and* a variance shift, let $\underset{\sim}{\theta}_0$ be (0,1) and let $\underset{\sim}{\gamma}$ be (γ_1, γ_2); then, on H_A, $\underset{\sim}{\theta}$ is $\underset{\sim}{\theta}_0 + \underset{\sim}{\gamma}/\sqrt{n}$, i. e. $\underset{\sim}{\theta} = (\gamma_1/n, 1 + \gamma_2/n)$. The corresponding components of $g(x)$ *are*

$$g_1(x) = -\frac{1}{\sqrt{2\pi}} \exp(-y^2/2) \quad \text{and} \quad g_2(x) = -\frac{y}{2\sqrt{2\pi}} \exp(-y^2/2) \tag{8}$$

where x is given by

$$x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-t^2/2) dt. \tag{9}$$

The δ_{ji} values, for $i = 1, 2$, are then found from (7), and asymptotically z_{nj} has then a normal distribution, with mean $\gamma_1\delta_{j1} + \gamma_2\delta_{j2}$, and variance 1.

Similarly, it is easy to show that z^*_{nj} will be asymptotically distributed $N(\underset{\sim}{\gamma}'\underset{\sim}{\delta}^*_j, 1)$, where vector δ^*_j has i-th component $(i = 1, 2)$

$$\delta^*_{ji} = \sqrt{2}\pi j \int_0^1 \cos(\pi j x) g_i(x) dx, \quad j = 1, 2, \ldots \tag{10}$$

Then $2R_j^2/n = z_{nj}^2 + z_{nj}^{*2}$ has asymptotically a non-central χ^2 distribution, d. f. 2, and non-centrality parameter

$$\lambda_j = (\underset{\sim}{\gamma}'\underset{\sim}{\delta}_j)^2 + (\underset{\sim}{\gamma}'\underset{\sim}{\delta}^*_j)^2 = (\gamma_1\delta_{j1} + \gamma_2\delta_{j2})^2 + (\gamma_1\delta^*_{j1} + \gamma_2\delta^*_{j2})^2. \tag{11}$$

6. MOMENTS OF W^2 AND U^2

We now turn to finding the moments of W^2 and U^2 on the alternative distribution. Suppose μ_0, σ_0^2 and $\varkappa_{30}$, with values respectively 0.1666, 0.02222 and 0.008466, are the asymptotic mean, variance and third cumulant of W^2 on H_0. On H_A, it is easily shown that these parameters become

$$\begin{aligned} \mu = E(W^2) &= \mu_0 + \sum_j (\underset{\sim}{\gamma}'\underset{\sim}{\delta}_j)^2/(j^2\pi^2) \\ &= \mu_0 + \sum_j Q_j/(j^2\pi^2) ; \\ \sigma^2 &= \sigma_0^2 + \sum_j Q_j/(j^4\pi^4) ; \\ \varkappa_3 &= \varkappa_{30} + \sum_j Q_j/(j^6\pi^6), \end{aligned} \tag{12}$$

where we write Q_j for $(\underset{\sim}{\gamma}'\underset{\sim}{\delta}_j)^2 = (\gamma_1\delta_{j1} + \gamma_2\delta_{j2})^2$. These results may be extended in an obvious way to higher cumulants.

For U^2, the null asymptotic cumulants are respectively $\mu_0 = 0.08333$, $\sigma_0^2 = 0.002777$ and $\varkappa_{30} = 0.000265$. Define now $Q_j^* = (\underset{\sim}{\gamma}'\underset{\sim}{\delta}_j^*)^2 = (\gamma_1\delta_{j1}^* + \gamma_2\delta_{j2}^*)$; then, on H_A:

$$\begin{aligned} \mu = E(U^2) &= \mu_0 + \sum_j (Q_j + Q_j^*)/(j^2\pi^2) \\ \sigma^2 &= \sigma_0^2 + \sum_j (Q_j + Q_j^*)/(j^4\pi^4) \\ \varkappa_3 &= \varkappa_{30} + \sum_j (Q_j + Q_j^*)/(j^6\pi^6). \end{aligned} \tag{13}$$

7. COMPARISONS WITH BEST TESTS

DK use the results for z_j to compare the asymptotic power of z_{nj} with that given by the best test when the shift is one of either mean only or variance only. For a test of mean shift, the best statistic is $\bar{y}$; for a 5 %

twotailed test, one rejects H_0 if $|\sqrt{n}\bar{y}| > 1.96$. Suppose γ_1 is positive; γ_2 is of course zero. The power, for given γ_1, is then

$$\begin{aligned}&\Pr(\sqrt{n}\bar{y} > 1.96) + \Pr(\sqrt{n}\bar{y} < -1.96)\\ &\quad = \Pr(\sqrt{n}(\bar{y} - \gamma_1/\sqrt{n}) > 1.96 - \gamma_1) + \Pr(\sqrt{n}(\bar{y} - \gamma_1/\sqrt{n}) < -1.96 - \gamma_1).\end{aligned}$$

Since $\sqrt{n}(\bar{y} - \gamma_1/\sqrt{n})$ is asymptotically N(0, 1), the power of the test based on $\bar{y}$ is 50 % when $\gamma_1 = 1.96$ (taking the second probability to be zero) and is 95 % when $\gamma_1 = 3.605$.

Since z_j is $N(\gamma_1\delta_{j1}, 1)$, the asymptotic power of the test based on z_{nj} can easily be found for these values of γ_1, and can be compared with the 50 % and 95 % powers of the best test. DK give some results in their Table 3; for example, when $\gamma_1 = 3.605$, so that $\bar{y}$ has 0.95 power, z_1 has power 0.928 and z_2 has power 0.05.

For the variance shift alternative ($\gamma_1 = 0$), the best test is based on sample variance s^2, and uses the result that the distribution of

$$s^2/(1 + \gamma_2/\sqrt{n}) \rightarrow N(1, 2/n).$$

The values of γ_2 for 50 % and 95 % best power are 2.772 and 5.098. Parallel calculations to those for the mean shift give the asymptotic powers of components z_{nj} against the best statistic s^2. Results are given in DK, Table 3.

8. MEAN AND VARIANCE SHIFT. PROPERTIES OF COMPONENTS

For this case, of course, no best alternative test exists. An orthogonality between the coefficients δ now leads to interesting results for the various statistics. For the test of normality, $g_1(x) - 0.5$ is an even function, and δ_{j1}, equation (7), is zero for all even values of j. Similarly, δ_{j2} is zero for all odd j, since $g_2(x) - 0.5$ is an odd function. Also, δ^*_{j1} will be zero for j odd, and $\delta^*_{j2} = 0$ for j even. Thus $\delta_{j1}\delta_{j2} = 0$ and $\delta^*_{j1}\delta^*_{j2} = 0$ for all j. Then z_1 will have mean $\gamma_1\delta_{j1}$ for any value of γ_2, so that the power of z_1 for a given change in mean remains the same no matter how large is the change in variance; in the example given above, when $\gamma_1 = 3.605$, z_1 has power 0.928 for all changes in variance. Similarly, z_2 has mean $\gamma_2\delta_{j2}$, and hence constant power for a given change in variance, for all changes in mean. These two statistics respectively detect changes in mean only or in variance only. The orthogonality also simplifies the non-centrality parameter for R_j; (11) becomes

$$\lambda_j = \gamma_1^2\delta_{j1}^{*2} + \gamma_2^2\delta_{j2}^2 \tag{14}$$

9. MEAN AND VARIANCE SHIFT. CUMULANTS OF W^2 AND U^2

The cumulants of W^2 and U^2 given in (12) and (13) also simplify considerably.

For $i = 1$ and 2, let

$$
\begin{aligned}
A_{1i} &= \sum_j \frac{\delta_{ji}^2}{j^2\pi^2}\,; & A_{2i} &= \sum_j \frac{\delta_{ji}^2}{j^4\pi^4}\,; & A_{3i} &= \sum_j \frac{\delta_{ji}^2}{j^6\pi^6}\,; \\
B_{1i} &= \sum_j{}^{*} \frac{\delta_{ji}^2}{j^2\pi^2}\,; & B_{2i} &= \sum_j{}^{*} \frac{\delta_{ji}^2}{j^4\pi^4}\,; & B_{3i} &= \sum_j{}^{*} \frac{\delta_{ji}^2}{j^6\pi^6}\,; \\
B^*_{1i} &= \sum_j{}^{*} \frac{\delta_{ji}^{*2}}{j^2\pi^2}\,; & B^*_{2i} &= \sum_j{}^{*} \frac{\delta_{ji}^{*2}}{j^4\pi^4}\,; & B^*_{3i} &= \sum_j{}^{*} \frac{\delta_{ji}^{*2}}{j^6\pi^6}
\end{aligned}
\tag{15}
$$

Where, as before, $\sum^*$ indicates the sum over even values only of j.
Then (12) becomes, for W^2.

$$
\begin{aligned}
\mu &= \mu_0 + \gamma_1^2 A_{11} + \gamma_2^2 A_{22} \\
\sigma^2 &= \sigma_0^2 + 4(\gamma_1^2 A_{21} + \gamma_2^2 A_{22}) \\
\kappa_3 &= \kappa_{30} + 24(\gamma_1^2 A_{31} + \gamma_2^2 A_{32})
\end{aligned}
\tag{16}
$$

and, for U^2

$$
\begin{aligned}
\mu &= \mu_0 + \gamma_1^2(B_{11} + B^*_{11}) + \gamma_2^2(B_{12} + B^*_{12}) \\
\sigma^2 &= \sigma_0^2 + 4\,\{\, \gamma_1^2(B_{21} + B^*_{21}) + \gamma_2^2(B_{22} + B^*_{22}) \,\} \\
\kappa_3 &= \kappa_{30} + 24\,\{\, \gamma_1^2(B_{31} + B^*_{31}) + \gamma_2^2(B_{32} + B^*_{32}) \,\}.
\end{aligned}
\tag{17}
$$

Note that this gives the interesting result that, for the test for normality, the increase in any cumulant for a mean and variance shift alternative is the sum of the increases for the corresponding mean shift only and variance shift only.

10. RELATED RESULTS

It happens that the values of δ_{ji} and δ^*_{ji}, necessary for the calculation of A_{ij}, B_{ij}, and B^*_{ij}, were known to the author because they arise again in the following connection. Suppose a test is to be made for the distribution $F(x;\theta)$, but θ is not known and must be estimated from the data. Under certain conditions, the asymptotic distributions of W^2 and U^2 are still expressible in the form of an infinite sum of weighted χ^2 variables, as given by (2) and (4), but the weights are no longer $(j\pi)^{-2}$. In the calculation of the new weights, one needs $g_i(x)$ and δ_{ij} as defined in § 5. These quantities were found in order to calculate the weights for the cases where $F(x;\theta)$ is the normal distribution, and θ is the vector (μ, σ^2), one or both of which must be estimated, and also where $F(x;\theta) = 1 - \exp(-\theta x)$, θ to be estimated. From the weights were calculated asymptotic percentage points for W^2, U^2 and A^2 for these cases (Stephens, 1971).

When a parameter is to be estimated, the asymptotic mean of W^2 or U^2 decreases. Let Δ_1, Δ_2 be the decreases for the cases where, in a test for normality, the parameter estimated is the mean or variance respectively, and let Δ_3 be the decrease when both are to be estimated. Then it may be shown that

$$\Delta_1 = \int g_1^2(x)dx = A_{11} \; ; \qquad \Delta_2 = 2\int g_2^2(x)dx = 2\,A_{12} \; ; \tag{18}$$
$$\Delta_3 = \Delta_1 + \Delta_2 \; ;$$

where $g_1(x), g_2(x)$ are given in (8) and A_{11}, A_{22} in (15). For U^2 the corresponding decreases are

$$\Delta_1 = B_{11} + B_{11}^* \; ; \qquad \Delta_2 = B_{12} + B_{12}^* \; ; \qquad \Delta_3 = \Delta_1 + \Delta_2.$$

It happens that the integrals in Δ_1 and Δ_2 for W^2 can be directly calculated, so that a check is available on the accuracy of A_{11} and A_{22} calculated from the δ's. Similarly, integrals exist for Δ_1 and Δ_2 for U^2, and parallel results hold for the statistic A^2.

11. THE STATISTIC A^2

This statistic may also be partitioned (DK), Section 6): let $u_i = 2x_i - 1$, and let

$$B = \sum_{i=1}^{n} u_1^2/n \; ;$$

then

$$A^2 = \sum_{j=1}^{\infty} z_{nj}^{*2}/(j^2 + j) \tag{19}$$

where

$$z_{nj}^* = -\,((2j+1)/n)^{1/2} \sum_{i=1}^{n} P_j(u_i) \tag{20}$$

and $P_j(.)$ are Legendre polynomials.

In particular

$$z_{n1} = -\sqrt{(3n)}\bar{u} \tag{21}$$

and

$$z_{n2} = -\frac{1}{2}(5/n)^{1/2}\sum_{i=1}^{n}(3u_1^2 - 1) = -\frac{3}{2}\sqrt{(5n)}\left(B - \frac{1}{3}\right) \tag{22}$$

where $\bar{u}$ is the mean of the u_i. The u_i are uniform $(-1, 1)$ and these statistics, particularly $\bar{u}$, have arisen in other connections. Tables of percentage points of $\bar{u}$ and of B are in Stephens (1966; B has the same distribution as T in that paper) so tests based on these components can be made. It would not be difficult to give good approximations to percentage points for higher order components also.

Asymptotically, each z_{nj} is N(0, 1) on H_0; non-null asymptotic theory is similar to that for W^2. A component z_{nj}^* will be asymptotically normal, with variance 1 and mean δ_j where δ_j has components δ_{ji} given by

$$\delta_{ji}(A^2) = (j(j+1))^{1/2}\int_0^1 g_i(x)P_j^1(2x-1)/(x-x^2)^{1/2}dx. \tag{23}$$

$P_j^1(2x-1)$ is a normalised Ferrar's associated Legendre function, and $g_i(x)$ is determined by H_0 as in Section 5. Thus again asymptotic powers of components may be found. For the normal case, with $g_i(x)$ given by (8), δ_{j1} (corresponding to mean shift alternatives) is 0 for j even, and $\delta_{j2} = 0$ for j odd, so, as with W^2, individual odd components will have constant power for a given mean shift, no matter how the variance changes, and *vice versa* for even components.

For the entire statistic, formulas (12) and (13) and later (15) (16) and (17), will hold for A^2, using the δ_{ji} of (22) and, in the denominators of the formulas, replacing $j^2\pi^2$ by $j(j+1)$, $j^4\pi^4$ by $j^2(j+1)^2$, and $j^6\pi^6$ by $j^3(j+1)^3$. The decreases in mean when a test is made with estimated parameters are now, using the notation of section 10,

$$\Delta_1 = \int(g_1^2(x)/(x-x^2))dx = A_{11}\,; \quad \Delta_2 = 2\int(g_2^2(x)/(x-x^2))^2dx = 2A_{12}\,;$$

$$\Delta_3 = \Delta_1 + \Delta_2\,;$$

the results correspond to those in (18) for W^2.

As for W^2 and U^2, the δ_{ji} were available from previous work on tests with estimated parameters, so non-null cumulants of A^2 could be found, and the distribution approximated.

12. POWER STUDIES: TEST FOR NORMALITY

Power studies have been made on W^2, U^2 and A^2, when the null hypothesis is that the distribution is N(0, 1) and on the alternative it is

$$N(\gamma_1/\sqrt{n},\ 1 + \gamma_2/\sqrt{n}).$$

These make use of the coefficients δ_{ji}, already available from previous work, as described above. The asymptotic power of a component S_j of W^2 can easily be found from normal tables, using the fact that asymptotically z_{nj} is $N(\mu, 1)$ where $\mu = \gamma'\delta_j = \gamma_1\delta_{ji} + \gamma_2\delta_{j2}$. Similarly the power of a component R_j of U^2 can be found using non-central χ_2^2 tables and (14). For the complete statistics W^2, U^2 and A^2, the cumulants of the non-null asymptotic distribution were found from (16) and (17), and these were used to fit an approximating distribution of the form $a + b\chi_p^2$, with constants a, b, p chosen so that the first three cumulants match the true values. A computer program giving the χ_p^2 integral was then used to find the power. Only two-sided tests have been considered, for components, and one-sided tests for the entire statistics. DK used another approximation; their calculations have been repeated using the $a + b\chi_p^2$ form, and results agree with their Table 4 very closely, certainly to the accuracy required for power comparisons, except at one point. These are the results for A^2, variance shift, for values of $\gamma_2 = 2.77$ and 5.095, the values used in their Table 4, columns 3 and 4 (γ_1 is then zero). Their values are .40 and .89, but our calculations give .15 and .52. Monte Carlo power studies were made to check this discrepancy using 1 000 samples of size 100 for $\gamma_2 = 2.77$ and 4 300 samples of size 100 for $\gamma_2 = 5.095$. Since we have to extrapolate from a finite sample size $n = 100$ to $n \to \infty$, we record also the Monte Carlo and theoretical results for W^2 and U^2. The following table gives the proportion of samples declared significant using a 5 % test.

		n	W^2	U^2	A^2
$\gamma_2 = 2.77$	Theory	∞	.09	.25	.15
	M. C.	100	.10	.21	.17
$\gamma_2 = 5.095$	Theory	∞	.26	.71	.52
	M. C.	100	.20	.52	.46

The results suggest that the DK values for A^2 are incorrect.

13. CONTOUR MAPS OF POWER

Graphical methods provide an effective way to compare the abilities of the statistics to detect combinations of changes in mean and variance. In Figures 1 and 2 values of γ_1^2 and γ_2^2 representing changes in mean and in variance respectively are recorded on the x-axis and y-axis. On the graphs, lines of constant power, called isodynes, have been drawn.

In Figure 1, they are shown for W^2, U^2 and for R_2, the first component of U^2 (called U_1^2 on the graph) and for power values of 20, 40, 60, 80 %.

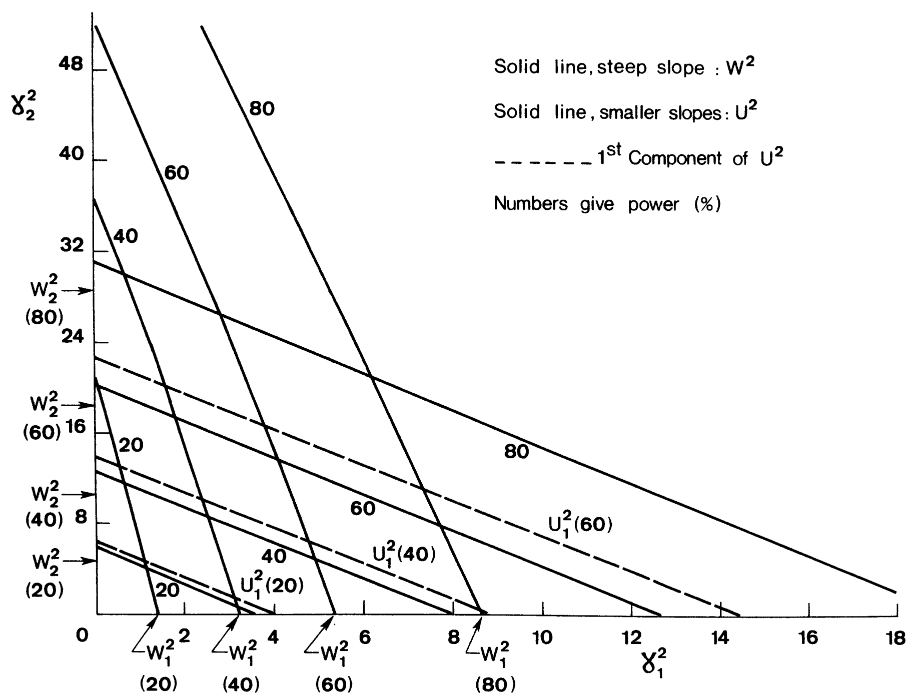

FIG. 1. — Test for Normality: W^2, U^2 and components.

In Figure 2, isodynes are drawn for A^2, W^2 and U^2. The complete set are shown for A^2, but only the 40 % and 80 % isodynes for W^2 and U^2. For comparison with DK Table 4, note that the four columns in that table correspond to points on the graphs with coordinates (γ_1^2, γ_2^2) equal to (3.84,0) (13.00,0) (0,7.67) (0,25.95).

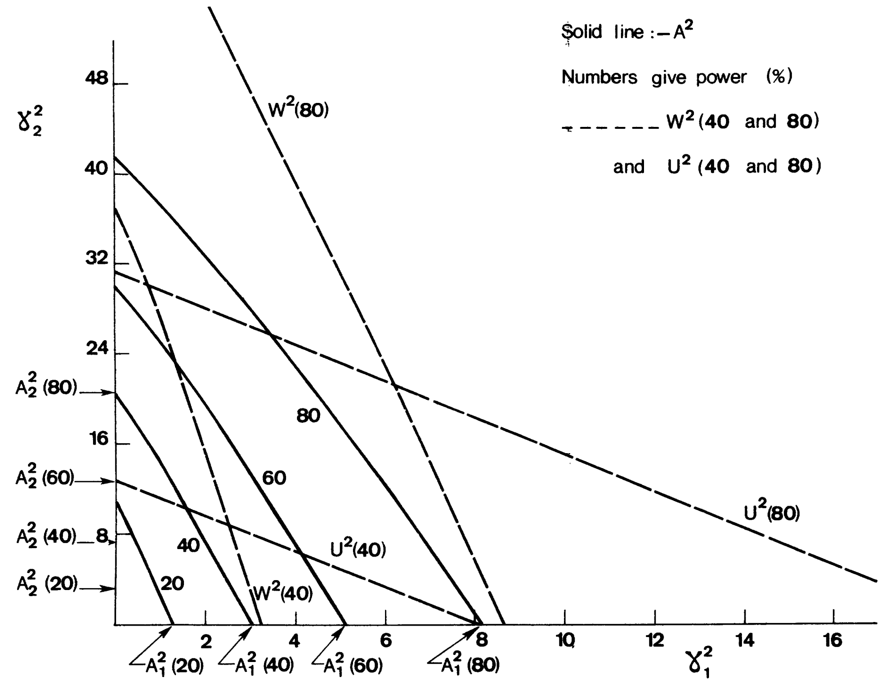

FIG. 2. — Test for Normality: W^2, U^2 and A^2.

Comments on the graphs. (*a*) A statistic will be more powerful than another if its isodyne lies nearer the origin. Thus we can see that for this situation, a test for N(0, 1) with a normal alternative, A^2 is everywhere more powerful than W^2. U^2 however, becomes better than A^2 as the variance shift, on the alternative, becomes larger. An interesting way to make comparisons is to calculate the areas enclosed by an isodyne and the axes, for a given statistic and given power. This cannot be absolute, since it depends on the parameters on the axes, but since the choice γ_1^2, γ_2^2 gives isodynes which are nearly straight lines—incidentally an interesting result—these areas have been calculated and are given in Table I. This table shows a clear overall superiority for A^2.

(*b*) Turning now to components, we know that S_j, the *j*-th component of W^2, will have isodynes parallel to the *y*-axis for *j* odd and to the *x*-axis for *j* even. Similar results hold for the components of A^2. The intercepts for the first two components, denoted by a subscript (e. g. W_1^2, W_2^2) are shown in Figures 1 and 2. The isodynes for W_1^2 and A_1^2 cut the *x*-axis almost at the same points as those of the entire statistic; but as γ_2^2 increases, they move parallel to the γ_2^2 axis while that of W^2 moves towards the axis. Thus there would be nothing gained in using the first component rather than the entire statistic. If one were more anxious to guard against a variance shift, however, the second components W_2^2 and A_2^2 will be considerably better than the parent statistic for small changes in mean. The area-measure for these components is infinite, reflecting the total insensitivity to one parameter.

(*c*) The first component of U^2 has much better properties; its isodynes, from (14) are straight lines, from axis to axis and it is nearly as powerful as U^2 itself. Area-measures are given in Table 1.

TABLE I. — *Comparison of* W^2, U^2 *and* A^2 *isodynes.*
The table gives the approximate area enclosed by the isodyne and the two axes, to the nearest unit.

Isodyne power	20 %	40 %	60 %	80 %
W^2	15	61	144	316
U^2	11	51	130	300
A^2	7	34	77	172
$R_2(U_1^2)$	13	62	162	

In the general goodness-of-fit situation, one would want a test statistic to be sensitive, ideally, to a change in the entire distribution; certainly, if one tests for N(0, 1) and holds to a normal distribution on the alternative, one would hope that a test statistic would be sensitive to changes in *both* parameters μ and σ^2. Thus in a comparison of *components* R_2 would seem to be the best, in this situation. Intuitively, this is appealing; if one plots the points on a circle, it seems more intuitive to use the entire resultant of the *n* unit vectors so formed, rather than a component on a specified axis.

14. POWER STUDIES: TEST FOR EXPONENTIALITY

The comparisons above may be extended to a test for exponentiality, where the null hypothesis H_0 is to test that the parent distribution is $F(y) = 1 - \exp(-\theta y)$ with $\theta = 1$, and, on H_A, the family remains exponential, with $\theta = 1 + \gamma_3/\sqrt{n}$. The power theory follows a pattern similar to that for the normal. The single parameter γ_3 gives a single

$$g(x) = -(1-x)\,ln(1-x);$$

δ_j is then found from (7) or (π) and δ_j^* from (10). Means of components, and cumulants of the entire statistics, are given as for the normal case, but with only one component in δ_j and δ_j^*. Values of these constants were again available to the author from work described in Section 10, and lead to power studies summarized in Figure 3 and Table II. The best test uses statistic $\bar{y}$; asymptotically, on H_0, $n(\theta\bar{y} - 1)$ is N(0, 1). For the best test to have 50 % asymptotic power, $\gamma_3 = 1.96$, i. e. $\gamma_3^2 = 3.84$, and for 95 % power $\gamma_3^2 = 13.00$. These points also are shown in Figure 3. In this Figure, powers are compared directly since only one parameter in involved. Results on components are given in Table II to avoid too much detail in Figure 3.

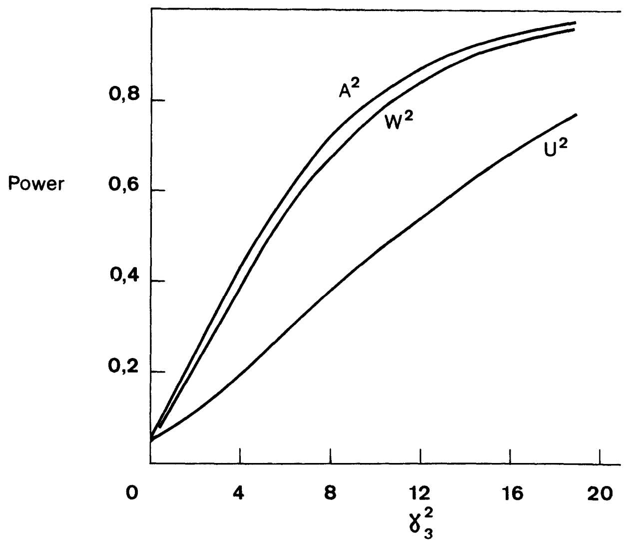

FIG. 3. Test for Exponentiality: W^2, U^2, A^2.

TABLE II. — *Test for exponentiality.*
Values of γ_3^2 to give power P(%) for A^2, W^2 and U^2, and their first components.

Statistic P (%)	40	80
A^2	3.60	9.7
A_1^2	3.9	10.2
W^2	4.0	10.7
W_1^2	4.2	11.2
U^2	8.6	20.6
U_1^2	9.5	23.9

Comments. A^2 is again the best statistic of the three, with W^2 a close second; U^2 is relatively weak. For all three statistics, the first components are almost as efficient as the entire statistic, but results for second components are considerably poorer.

Final remarks. The combined results suggest that there is very little gained in taking a component rather than the entire statistic, in the usual case of a very general alternative to the null. The technique of asymptotic power calculation introduced by DK and extended here is very useful to give measures of comparison for the various statistics and the isodyne graphs illustrate the results probably better than extensive tables. This particular study bears out results in Stephens (1972), where isodyne graphs are used for comparisons based on Monte Carlo studies, that A^2 is a very valuable statistic in a wide class of situations. On the theoretical side, it will be interesting to see extensions of the DK work to finding components for the case where parameters of the distribution are estimated before $F_n(x)$ is found.

ACKNOWLEDGEMENTS

The author attended the Conference on uses of the empirical distribution hosted by S. U. N. Y. at Buffalo in September 1971, where a first draft of the Durbin and Knott paper was presented, and expresses his gratitude to Professor E. Parzen for the invitation to participate and to N. S. F. for their support. This work was also supported by the U. S. office of Naval Research and by the National Research Council of Canada, and the author expresses his gratitude to these agencies: also to the French Government, and to Professor J. R. Barra, for a visiting Professorship and invitation to visit the University of Grenoble, 1971-1972, where this work was begun.

REFERENCES

J. DURBIN and M. KNOTT, Components of Cramer-von Mises Statistics I. *Jour. Royal Statist. Assn. B.*, t. **34**, 1972, p. 290-307.

J. A. GREENWOOD and D. DURAND, The distribution of length and components of the sum of *n* random unit vectors. *Ann. Math. Statist.*, t. **26**, 1955, p. 233-246.

E. S. PEARSON and H. O. HARTLEY, *Biometrika Tables for Statisticians*, vol. 2 : Cambridge : Cambridge University Press, 1972.

M. A. STEPHENS, Statistics connected with the uniform distribution; percentage points and application to tests for randomness of directions. *Biometrika*, t. **53**, 1966, p. 235-240.

M. A. STEPHENS, Tests for randomness of directions against two circular alternatives. *J. Amer. Statist. Assn.*, t. **64**, 1969, p. 250-289.

M. A. STEPHENS, Use of the Kolmogorov-Smirnov, Cramer-von Mises and related statistics without extensive tables. *Jour. Royal Statist. Assn. B.*, t. **32**, 1970, p. 115-122.

M. A. STEPHENS, Asymptotic results for goodness-of-fit statistics when parameters must be estimated. *Tech. Report.*, Department of Statistics, Stanford Univ., 1971.

M. A. STEPHENS, E. D. F. Statistics for Goodness-of-fit. Parts 1, 2, 3. *Tech. Report*, Department of Statistics, Stanford Univ., 1972.

(Manuscrit reçu le 18 octobre 1973).