e
W.

f

Journal of Applied Statistic

Edited by Gopal K Kanji

VOLUME 15 NUMBER 2 1988

SPECIAL ISSUE: Analysis of Circular Data

ISSN 0266–4763

carfax publishing company

Journal of Applied Statistics, Vol. 15, No. 2, 1988

Kuiper's *P*-value as a measuring tool and decision procedure for the goodness-of-fit test

H. ARSHAM, *Department of Information and Quantitative Sciences, University of Baltimore, Maryland*

SUMMARY *The need for computing* P-*values for the Kuiper statistic has been emphasised by Batschelet (1981). Some exact* P-*values, with useful interpretations for making inference about a probability model for circular data, are provided. Computation of the exact values are based on Durbin (1973) boundary crossing probabilities. A numerical example is used to demonstrate the usefulness of the results.*

1 Introduction

Suppose that a given random sample of n angular values $x1, x2, \ldots, xn$ represents directions in a plane, or time instants of a cyclic phenomenon. The problem is to test the null hypothesis ($H0$) that the population has a prescribed continuous cumulative distribution function (*cdf*), $F(x)$, based upon empirical *cdf*, $Fn(x)$ i.e.

$$Fn(x) = 1/n\,\{\text{number of observations} \leq x\}. \quad (1)$$

These types of problems arise in a number of scientific fields notably biology and geophysics. As a test statistic the Kuiper's statistic which is defined to be:

$$Kn = n^{\frac{1}{2}} \sup_x\,[Fn(x) - F(x)] + n^{\frac{1}{2}} \sup_x\,[F(x) - Fn(x)], \quad (2)$$

over all x's (see, e.g. D'Agostino & Stephens, 1986) has been widely applied. The value of Kn does not depend on the choice of origin for x and it is also distribution free (see, e.g. Durbin, 1973). Short tables of critical values of Kn for some selected tail probabilities are available, see, e.g. Stephens (1965). However, when the purpose of performing a test is to make an inference about a probability model for the data, a simple accept-reject dichotomy based on these tables and a preassigned level of significance (e.g. $\alpha = 0.05$) may not be sufficiently informative. There is a strong requirement to determine the *P*-value as a means for making more informative inferential statement (see, e.g. Batschelet, 1981, p. 53). This note provides a table of exact *P*-values for the Kuiper statistic useful to data analysts.

2 Computing the Kuiper statistic

Given a random sample of size n, suppose there are r different values of x for some $r \leq n$, and the ordered values are $x(1) < x(2) < \ldots < x(r)$, each observed with frequency at least 1. Then the computing formula for Kn based on (1) is:

$$Kn = n^{\frac{1}{2}} \max_i \{Fn[x(i)] - F[x(i)] + n^{\frac{1}{2}} \max_i \{F[x(i)] - Fn[x(i-1)]\}, \tag{3}$$

over all $1 \leq i \leq r$, with $Fn[x(0)] \equiv 0$. Since the calculated value of Kn by (3) is a measure of the agreement between the hypothesised distribution and the true distribution as reflected by the sample data, a large value of Kn tends to discredit the null hypothesis. Hence the appropriate measure of strength is a right tail probability (P-value) from the sampling distribution of Kn (see, e.g. Cox & Hinkley, 1974, p. 66).

3 The meaning and interpretation of P-values

The P-value, which directly depends on a given sample, attempts to provide a measure of the strength of the results, in contrast to a simple yes or no decision. If the null hypothesis is true and the chance of random variation is the only reason for sample differences, then the P-value is a quantitative measure to feed into the decision making process as evidence. Burdette & Gehan (1970, p. 9) provide a reasonable interpretation of P-values as follows:

P-value	Interpretation
$P < 0.01$	very strong evidence against $H0$
$0.01 \leq P < 0.05$	moderate evidence against $H0$
$0.05 \leq P < 0.10$	suggestive evidence against $H0$
$0.10 \leq P$	little or no real evidence against $H0$

This interpretation is widely accepted, and many scientific journals routinely publish papers using such an interpretation for the result of test of hypothesis. This interpretation would also be the principal criteria for the computation and tabulation of the P-values in this note. The general effect of sample size on the P-values is discussed by Lindley & Scott (1984, p.3). They show that the evidence represented by a P-value in a small sample size situation is stronger than with a larger sample size.

4 Algorithm

The distribution of Kn was first considered and tabulated by Stephens (1965) using a difference-equation method which may not be accurate due to rounding error. Durbin (1973) formulates the exact distribution of Kn as a 'boundary crossing probability' which has the advantage of being a 'computationaly probability' algorithm, an emerging discipline concerned with numerical solution of applied probability.

Let $U_{n-1}(x)$ be the empirical cdf of the Uniform [0, 1] random variable of size $n-1$, then from Durbin (1973) eq. (5.2.3) we have:

$$P[Kn \leq k] = n.\ \Pr[\text{path of } U_{n-1}(x) \text{ lies between the two lines, } (n-1)y1 = nx - 1 \text{ and } (n-1)y2 = nx + n^{\frac{1}{2}}\, k - 1] \tag{4}$$

It can be shown that the rhs of (4) is equal to

$$n \Pr[(i+1-n^{\frac{1}{2}} k)/n \leq U(i) \leq i/n] \tag{5}$$

where $U(i)$ is the ith order statistic of a random sample of size $(n-1)$ from uniform [0, 1]. The P-value is

$$P=1-n \Pr[(i+1-n^{\frac{1}{2}}k)/n \leq U(i) \leq i/n]. \tag{6}$$

5 Tabulation of the *P*-values

Several algorithms are available to compute P using (6). For a complete discussion of these algorithms see Shorack & Wellner (1987, chapter 9). Table 1 provides the P-values for some of the Kuiper statistic Kn selected values of k. This selection is dictated by the usefulness of the P-values within the limits discussed in Section 3 and space limitations. These results are based on (6) using Steck's and Noe's algorithms. Steck's algorithm performs well and is fast for small size ($n \leq 30$) samples, whereas Noe's algorithm can handle larger size samples. None of these algorithms can be used for asymptotic P-values. However, a closed formed expression for the limiting distribution can be found, e.g. in Durbin (1973):

$$\lim_{n\to\infty} P = 2 \sum_{j=1}^{\infty} (4 j^2k^2-1) \exp(-2 j^2k^2) \tag{7}$$

The last column in Table 1 is based on (7). Values in Table 1 are accurate to at least two digits after the decimal point.

6 Application

Batschelet (1981) considered the test of randomness at ($\alpha=0.05$) of the directions in which a sample of birds flew when released at a fixed point. The data is arranged in ascending order: 20, 135, 145, 165, 170, 200, 300, 325, 335, 350, 350, 350, 355. The sample size is $n=13$, with $r=11$ distinct x values. Randomness of directions is equivalent to hypothesis that the data is distributed at random on the circumference of the circle. That is, the sample comes from a Uniform distribution, i.e.

$$F(x)=\begin{cases} 0 & x \leq 0 \\ x/360 & 0 \leq x \leq 360 \\ 1 & x \geq 360 \end{cases}$$

Let $da=Fn[x(i)]-F[x(i)]$ and $db=F[x(i)]-Fn[x(i-1)]$. Table 2 shows the computation of da and db. Where da is the difference between the sample and Uniform probabilities on the same line and db is the difference between the Uniform probability on the same line and the sample probability on the preceding line.

The computational version of the Kuiper's statistic Kn is:

$$Kn=n^{\frac{1}{2}} \max_{1\leq i\leq 11} \{da\}+n^{\frac{1}{2}} \max_{1\leq i\leq 11} \{db\}. \tag{8}$$

From Table 2, the Kuiper's statistic for this example is calculated to be $Kn=3.6056$ $(0.0213+0.3718)\approx 1.417$. Interpoling from Table 1 with $Kn\approx 1.42$ for $n=13$, the P-value is:

$$P=\Pr[Kn \geq 1.42] \approx 0.1666 \approx 0.17$$

The probability of the Kn statistic being 1.42 or larger is approximately 17%, assuming the null hypothesis is true and the observed departure from randomness is due to chance variation only. This is a large probability, therefore there is little or no real evidence, based on these 13 observations against the null hypothesis. This conclusion is

TABLE 1. The P-value, P as a function of (k, n) for the Kuiper test

k	n=5	10	15	20	25	30	40	50	60	70	80	90	100	150	∞
1.40	0.138	0.173	0.190	0.200	0.207	0.212	0.219	0.224	0.228	0.231	0.234	0.236	0.238	0.244	0.271
1.42	0.124	0.157	0.173	0.182	0.189	0.194	0.201	0.206	0.210	0.212	0.215	0.217	0.218	0.224	0.250
1.44	0.111	0.142	0.157	0.166	0.172	0.177	0.184	0.188	0.192	0.195	0.197	0.199	0.200	0.206	0.231
1.46	0.099	0.128	0.142	0.151	0.157	0.161	0.168	0.172	0.175	0.178	0.180	0.182	0.183	0.188	0.212
1.48	0.088	0.116	0.129	0.137	0.143	0.147	0.153	0.157	0.160	0.162	0.164	0.166	0.167	0.172	0.194
1.50	0.078	0.104	0.116	0.124	0.129	0.133	0.139	0.143	0.146	0.148	0.150	0.151	0.152	0.157	0.178
1.52	0.068	0.093	0.105	0.112	0.117	0.120	0.126	0.129	0.132	0.134	0.136	0.137	0.139	0.143	0.162
1.54	0.060	0.083	0.094	0.101	0.105	0.109	0.114	0.117	0.120	0.122	0.123	0.125	0.126	0.130	0.148
1.56	0.052	0.074	0.084	0.091	0.095	0.098	0.103	0.106	0.108	0.110	0.112	0.113	0.114	0.118	0.134
1.58	0.045	0.066	0.075	0.081	0.085	0.088	0.093	0.095	0.098	0.099	0.101	0.102	0.103	0.106	0.122
1.60	0.039	0.058	0.067	0.073	0.076	0.079	0.083	0.086	0.088	0.090	0.091	0.092	0.093	0.096	0.110
1.62	0.033	0.052	0.060	0.065	0.068	0.071	0.075	0.077	0.079	0.080	0.082	0.083	0.084	0.086	0.100
1.64	0.028	0.046	0.053	0.058	0.061	0.063	0.067	0.069	0.071	0.072	0.073	0.074	0.075	0.078	0.090
1.66	0.024	0.040	0.047	0.051	0.054	0.056	0.060	0.062	0.063	0.065	0.066	0.067	0.067	0.070	0.081
1.68	0.020	0.035	0.042	0.045	0.048	0.050	0.053	0.055	0.057	0.058	0.059	0.059	0.060	0.062	0.073
1.70	0.017	0.031	0.037	0.040	0.043	0.045	0.047	0.049	0.050	0.052	0.052	0.053	0.054	0.056	0.065
1.72	0.015	0.027	0.032	0.036	0.038	0.040	0.042	0.044	0.045	0.046	0.047	0.047	0.048	0.050	0.058
1.74	0.012	0.023	0.028	0.031	0.033	0.035	0.037	0.039	0.040	0.041	0.041	0.042	0.043	0.044	0.052
1.76	0.010	0.020	0.025	0.027	0.029	0.031	0.033	0.034	0.035	0.036	0.037	0.037	0.038	0.039	0.046
1.78	0.009	0.017	0.022	0.024	0.026	0.027	0.029	0.030	0.031	0.032	0.033	0.033	0.033	0.035	0.041
1.80	0.007	0.015	0.019	0.021	0.023	0.024	0.026	0.027	0.027	0.028	0.029	0.029	0.030	0.031	0.037
1.82	0.006	0.013	0.016	0.018	0.020	0.021	0.022	0.023	0.024	0.025	0.025	0.026	0.026	0.027	0.033
1.84	0.005	0.011	0.014	0.016	0.017	0.018	0.020	0.020	0.021	0.022	0.022	0.023	0.023	0.024	0.029
1.86	0.004	0.010	0.012	0.014	0.015	0.016	0.017	0.018	0.019	0.019	0.020	0.020	0.020	0.021	0.025
1.88	0.003	0.008	0.011	0.102	0.013	0.014	0.015	0.016	0.016	0.017	0.017	0.017	0.018	0.019	0.022
1.90	0.003	0.007	0.009	0.010	0.011	0.012	0.013	0.014	0.014	0.015	0.015	0.015	0.016	0.016	0.020
1.92	0.002	0.006	0.008	0.009	0.010	0.010	0.011	0.012	0.012	0.013	0.013	0.013	0.014	0.014	0.017
1.94	0.002	0.005	0.007	0.008	0.008	0.009	0.010	0.010	0.011	0.011	0.011	0.012	0.012	0.012	0.015
1.96	0.001	0.004	0.006	0.007	0.007	0.008	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.013
1.98	0.001	0.003	0.005	0.006	0.006	0.007	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.012
2.00	0.001	0.003	0.004	0.005	0.005	0.006	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010

stronger than the conclusion of Batschelet (1981, p. 78) that 'randomness cannot be excluded.'

TABLE 2. The Kuiper test for randomness for the numerical example

i	*x*(*i*)	*Fn*[*x*(*i*)]	*F*[*x*(*i*)]	*da*	*db*
1	20	0.0769	0.0556	0.0213	0.0556
2	135	0.1538	0.3750	−0.2212	0.2981
3	145	0.2308	0.4028	−0.1720	0.2490
4	165	0.3077	0.4583	−0.1506	0.2275
5	170	0.3846	0.4722	−0.0876	0.1645
6	200	0.4615	0.5556	−0.0941	0.1710
7	300	0.5385	0.8333	−0.2948	0.3718
8	325	0.6154	0.9028	−0.2874	0.3643
9	335	0.6923	0.9306	−0.2383	0.3152
10	350	0.9231	0.9722	−0.0491	0.2799
11	355	1	0.9861	0.0139	0.0630

7 Summary and concluding remarks

A table of *P*-values of the Kuiper statistic has been presented. The *P*-value is a tool for measuring the strength of evidence against the null hypothesis and making more informative inferential statements rather than a simple accept-reject dichotomy based on existing tables. While directed originally to examining distributions on the circle, the Kuiper statistic is applicable, as noted by Durbin (1973), to distributions on the line—one has only to think of the line as turning full circle. When applied to distribution on the line, the Kuiper statistic is competitor to the well known usual Kolmogorov-Smirnov statistic (see, e.g. Koziol, 1980).

Acknowledgements

This work was supported by the Educational Foundation of the University of Baltimore, Summer 1988. I would also like to thank Dr R. Vemuganti for his support.

Correspondence: Professor H. Arsham, D.Sc. F.R.S., Department of Information & Quantitative Sciences, University of Baltimore, Baltimore MD21201, U.S.A.

REFERENCES

BATSCHELET, E. (1981) *Circular Statistics in Biology* (London, Academic Press).

BURDETTE, W.J. & GEHAN, E.A. (1970) *Planning and Analysis of Clinical Studies* (Springfield, Illinois, Charles C Thomas).

COX, D.R. & HINKLEY, D.V. (1974) *Theoretical Statistics* (London, Chapman & Hall).

D'AGOSTINO, R.B. & STEPHENS, M.A. (1986) *Goodness-of-fit Techniques* (New York, Marcel Dekker).

DURBIN, J. (1973) Distribution theory for tests based on the sample distribution function, *Regional Conferences Series in Applied Mathematics, Vol. 9* (Philadelphia, SIAM).

KOZIOL, J.A. (1980) Percentage points of the asymptotic distributions of one and two sample Kuiper statistic for truncated or censored data, *Technometrics*, 22, pp. 437–442.

LINDLEY, D.V. & SCOTT, W.F. (1984) *New Cambridge Elementary Statistical Tables* (Cambridge University Press).

SHORACK, G.R. & WELLNER, J.A. (1986) *Empirical Processes with Applications to Statistics* (New York, Wiley).

STEPHENS, M.A. (1965) The goodness-of-fit statistic *Vn:* distribution and significance points, *Biometrika*, 52, pp. 309–321.