Biometrika (1962), **49**, 1 *and* 2, *p.* 57
Printed in Great Britain

Goodness-of-fit tests on a circle. II

By G. S. WATSON
University of Toronto and Research Triangle Institute†

1. Introduction

In a previous paper (Watson, 1961, referred to hereafter as Part I) the statistic

$$U_N^2 = N\int_{-\infty}^{\infty}\left\{F_N(x)-F(x)-\int_{-\infty}^{\infty}[F_N(y)-F(y)]\,dF(y)\right\}^2 dF(x) \tag{1}$$

was proposed for testing the null hypothesis that the random sample of N with distribution function $F_N(x)$ has been drawn from a population with the continuous distribution function $F(x)$. It is useful for distributions on a circle since its value does not depend on the arbitrary point chosen to begin cumulating the probability density and the sample points. It was shown that

$$\lim_{N\to\infty} \operatorname{prob}(U_N^2 > v) = 2\sum_{m=1}^{\infty}(-1)^{m-1}\exp(-2m^2\pi^2 v). \tag{2}$$

The merit of this distribution as an approximation to the distribution of U_N^2 for finite N was not discussed.

The purpose of the present paper is twofold; (*a*) to show that the limiting distribution ($N_1, N_2 \to \infty$, $N_1/N_2 \to \lambda > 0$), of the two-sample version of (1), the Lebesgue–Stieljes integral,

$$U_{N_1,N_2}^2 = \frac{N_1N_2}{N_1+N_2}\int_{-\infty}^{\infty}\left\{F_{N_1}(x)-F_{N_2}(x)-\int_{-\infty}^{\infty}[F_{N_1}(y)-F_{N_2}(y)]\,dF^*(y)\right\}^2 dF^*(x), \tag{3}$$

where
$$F^*(x) = \frac{N_1F_{N_1}(x)+N_2F_{N_2}(x)}{N_1+N_2},$$

is again given by (2); (*b*) to show that the distribution (2) is adequate for the practical use, in small samples, of (1) and (3). Aim (*a*) is achieved in § 2 using methods due to Rosenblatt (1952) and Fisz (1960). To examine (*b*), one can either strive to find further terms in the asymptotic expansion for the distribution of U^2, (2) being the leading terms, by the method of Darling (1960) or proceed numerically. For statistic (1), the latter approach means a sampling study. For the statistic (3), one may either make a Monte Carlo study or, for small samples, compute the exact (discrete) distribution of U_{N_1,N_2}^2 by enumeration since it is easily shown that

$$U_{N_1,N_2}^2 = \frac{N_1N_2}{(N_1+N_2)^2}\left\{\sum_{i=1}^{N_1}[F_{N_1}(x_{(i)})-F_{N_2}(x_{(i)})]^2+\sum_{j=1}^{N_2}[F_{N_1}(y_{(j)})-F_{N_2}(y_{(j)})]^2 - \frac{\left[\sum_{i=1}^{N_1}\{F_{N_1}(x_{(i)})-F_{N_2}(x_{(i)})\}+\sum_{j=1}^{N_2}\{F_{N_1}(y_{(j)})-F_{N_2}(y_{(j)})\}\right]^2}{N_1+N_2}\right\}, \tag{4}$$

where $x_{(i)}$ ($i=1,\ldots,N_1$) and $y_{(j)}$ ($j=1,\ldots,N_2$) are the two ordered samples being compared, depends only on the relative ranks. Sampling studies were carried out for the writer by

† Research sponsored in part by the Office of Naval Research.

Mr M. A. Stephens and are reported in § 3. The exact method was undertaken by Mr E. J. Burr whose results, along with those of the related statistic $W^2_{N_1,N_2}$ are reported in Burr (1961).

Section 4 gives an interesting biological example of the use of statistic (4), which in fact, suggested its development.

Thus, in these two papers, one- and two-sample non-parametric goodness-of-fit tests have been provided that are useful on the circle where there is no natural point at which to start the cumulative distribution. Kuiper (1960) has also given tests applicable in this situation. Nothing is known of their relative merits.

2. THE TWO-SAMPLE TEST

It may first be shown that

$$R^2_{N_1,N_2} = \frac{N_1 N_2}{N_1+N_2}\int_{-\infty}^{\infty}\left\{F_{N_1}(x)-F_{N_2}(x)-\int_{-\infty}^{\infty}[F_{N_1}(y)-F_{N_2}(y)]\,dF(y)\right\}^2 dF(x), \tag{5}$$

has, as $N_1; N_2 \to \infty$, $N_1/N_2 \to \lambda > 0$, the limiting distribution (2). The expression (5) differs from (3) only by the use of F instead of F^*. Using $u = F(x)$ to transform to the unit interval, we may then define

$$D_N(u) = F_{N_1}(u)-F_{N_2}(u)-\int_0^1 [F_{N_1}(u)-F_{N_2}(u)]\,du. \tag{6}$$

Calculations show that

$$\operatorname{cov}[D_N(u), D_N(v)] = \left(\frac{1}{N_1}+\frac{1}{N_2}\right)\operatorname{cov}[Z_N(u), Z_N(v)], \tag{7}$$

where $Z_N(u)$ is given by equation (9) in Part I. Since.

$$U^2_N = \int_0^1 Z_N(u)^2\,du,$$

the required result follows immediately.

It is now necessary to show that, as $N_1, N_2 \to \infty$, $N_1/N_2 \to \lambda > 0$,

$$U^2_{N_1,N_2} - R^2_{N_1,N_2} \to 0 \quad \text{in probability.} \tag{8}$$

For result (8), combined with that of the previous paragraph, establishes that $U^2_{N_1,N_2}$ has asymptotically the distribution (2). Now

$$U^2_{N_1,N_2} - R^2_{N_1,N_2} = \frac{N_1 N_2}{N_1+N_2}\int_{-\infty}^{\infty}\{F_{N_1}(x)-F_{N_2}(x)\}^2\,d[F^*(x)-F(x)]$$
$$+\frac{N_1 N_2}{N_1+N_2}\left\{\left(\int_{-\infty}^{\infty}[F_{N_1}(x)-F_{N_2}(x)]\,dF(x)\right)^2-\left(\int_{-\infty}^{\infty}[F_{N_1}(x)-F_{N_2}(x)]\,dF^*(x)\right)^2\right\}. \tag{9}$$

The first term on the right-hand side of (9) arises in a similar discussion of the statistic W^2_N for

$$W^2_{N_1,N_2} = \frac{N_1 N_2}{N_1+N_2}\int_{-\infty}^{\infty}\{F_{N_1}(x)-F_{N_2}(x)\}^2\,dF^*(x) \tag{10}$$

is its two-sample version. It has been shown (Rosenblatt, 1952; Kiefer, 1959; Fisz, 1960) that this term tends to zero in probability. It remains to show that the same is true of the second term. This may be written as the product of the two factors,

$$L_+ = \sqrt{\frac{N_1 N_2}{N_1+N_2}}\int_{-\infty}^{\infty}[F_{N_1}(x)-F_{N_2}(x)]\,d[F(x)+F^*(x)], \tag{11}$$

and
$$L_- = \sqrt{\frac{N_1 N_2}{N_1+N_2}} \int_{-\infty}^{\infty} [F_{N_1}(x) - F_{N_2}(x)]\, d[F(x) - F^*(x)]. \tag{12}$$

But
$$L_+ = 2\sqrt{\frac{N_1 N_2}{N_1+N_2}} \int_{-\infty}^{\infty} [F_{N_1}(x) - F_{N_2}(x)]\, dF(x) - L_-, \tag{13}$$

and, writing $u_i = F(x_i)$, $v_j = F(y_j)$, $(i=1,\ldots,N_1, j=1,\ldots,N_2)$,
$$\sqrt{\frac{N_1 N_2}{N_1+N_2}} \int_{-\infty}^{\infty} [F_{N_1}(x) - F_{N_2}(x)]\, dF(x) = \sqrt{\frac{N_1 N_2}{N_1+N_2}}\, (\bar{u} - \bar{v}), \tag{14}$$

where
$$\bar{u} = \sum_1^{N_1} \frac{F(x_i)}{N_1}, \quad \bar{v} = \sum_1^{N_2} \frac{F(y_j)}{N_2}.$$

Thus asymptotically the expression (14) is normal with mean zero and variance $1/12$. Hence we only need to prove that L_- tends to zero in probability. Minus one times L_- may be written as
$$\sqrt{\frac{N_1 N_2}{N_1+N_2}} \int_{-\infty}^{\infty} \{[F_{N_1}(x) - F(x)] - [F_{N_2}(x) - F(x)]\}\, d\, \frac{N_1[F_{N_1}(x) - F(x)] + N_2[F_{N_2}(x) - F(x)]}{N_1+N_2},$$

which simplifies to
$$\frac{N_1^{\frac{1}{2}} N_2^{\frac{3}{2}}}{(N_1+N_2)^{\frac{3}{2}}} \int_{-\infty}^{\infty} [F_{N_2}(x) - F(x)]\, d[F_{N_2}(x) - F(x)] \\ - \frac{N_1^{\frac{3}{2}} N_2^{\frac{1}{2}}}{(N_1+N_2)^{\frac{3}{2}}} \int_{-\infty}^{\infty} [F_{N_1}(x) - F(x)]\, d[F_{N_1}(x) - F(x)], \tag{15}$$

since
$$N_1 \int_{-\infty}^{\infty} (F_{N_1} - F)\, d(F_{N_1} - F) = N_2 \int_{-\infty}^{\infty} (F_{N_2} - F)\, d(F_{N_2} - F) = 0.$$

Mapping onto the unit interval with $u = F(x)$, and writing
$$\left.\begin{aligned} z_1(u) &= \sqrt{N_1}[F_{N_1}(x) - F(x)], \\ z_2(u) &= \sqrt{N_2}[F_{N_2}(x) - F(x)], \end{aligned}\right\} \tag{16}$$

the expression (15) becomes
$$-L_- = \frac{N_2}{(N_1+N_2)^{\frac{3}{2}}} \int_0^1 z_1(u)\, dz_2(u) - \frac{N_1}{(N_1+N_2)^{\frac{3}{2}}} \int_0^1 z_2(u)\, dz_1(u). \tag{17}$$

Since $z_1(u)$ and $z_2(u)$ are independent functions with zero means, from (16), $E(L_-) = 0$. To find var (L_-), we need the well-known result,
$$E(z_i(u)\, z_i(v)) = \min(u, v) - uv \quad (i=1,2). \tag{18}$$

Then
$$\operatorname{var}(L_-) = \frac{N_2^2}{(N_1+N_2)^3} \int_0^1 \int_0^1 E[z_1(u)\, z_1(u')\, dz_2(u)\, dz_2(u')] \\ - \frac{2N_1 N_2}{(N_1+N_2)^3} \int_0^1 \int_0^1 E[z_1(u)\, z_2(u')\, dz_2(u)\, dz_1(u')] \\ + \frac{N_1^2}{(N_1+N_2)^3} \int_0^1 \int_0^1 E[z_2(u)\, z_2(u')\, dz_1(u)\, dz_1(u')]. \tag{19}$$

Thus var $(L_{-}) \to 0$, as $N_1, N_2 \to \infty$, $N_1/N_2 \to \lambda > 0$, provided the integrals in (19) are finite. This last point may be verified by using the transformation

$$z_i(u) = (u-1)\, Z_i\left(\frac{u}{1-u}\right) \quad (i = 1, 2), \tag{20}$$

employed by Rosenblatt for a similar purpose. We have thus demonstrated that the asymptotic distribution of $U^2_{N_1,N_2}$ is given by (2).

3. Sampling experiments

Mr M. A. Stephens programmed an I.B.M. 650 computor to draw two random samples of 10 members and to compute U^2_{10} for each and $U^2_{10,10}$ for the pair. As a matter of interest W^2_{10} and $W^2_{10,10}$ were also calculated, where W^2_N is defined by (1) in Part I and $W^2_{N_1,N_2}$ is its two-sample analogue. The common limiting distribution of these latter statistics is tabulated in Anderson & Darling (1952). Due to loss of cards, the numbers of calculated single and two-sample statistics do not match up.

There is some arbitrariness in reporting work of this kind. While the distributions of U^2_N and W^2_N are continuous, those of $U^2_{N_1,N_2}$ and $W^2_{N_1,N_2}$ are discontinuous and the number of samples used here was great enough to show this roughness. In the case of $W^2_{10,10}$, Mr E. J. Burr has found the exact distribution and the only interest in the sampling results in this case is that they showed some of his larger 'lumps'. The figures in the table below were obtained by plotting the upper part of the cumulative distribution of each of the four

Significance level (%)	U^2_{10} [5261]	$U^2_{10,10}$ [2671]	U^2_{∞} [∞]	W^2_{10} [5261]	$W^2_{10,10}$ [2429]	W^2_{∞} [∞]
50	0·070	0·077	0·069	0·122	0·129	0·119
30	·106	·104	·096	·187	·193	·184
20	·118	·124	·117	·241	·254	·241
15	·131	·137	·131	·284	·297	·284
10	·150	·156	·152	·343	·370	·347
7	0·165	0·172	0·170	0·395	0·424	0·405
6	·171	·178	·178	·418	·450	·430
5	·178	·185	·187	·445	·482	·461
4	·186	·194	·198	·478	·520	·499
3	·197	·206	·213	·525	·562	·549
2	·214	·224	·233	·580	·624	·620
1	·244	·250	·268	·715	·738	·743

samples; the significance points, corresponding to the levels listed, were read off from the graphs. For the U^2 statistics, the theoretical asymptotic significance points were calculated from (2). For the W^2 statistics, they were obtained from Anderson & Darling's table. This table gives the significance points, at the levels in the first column, for the four statistics in sampling experiments using the number of samples shown at the head of the column. The columns headed [∞] give the exact asymptotic points.

It is clear from the table that $U^2_{10,10}$ is stochastically larger than U^2_{10}. From the graphs of the sampling experiment the difference $P(U^2_{20} \leqslant u) - P(U^2_{10} \leqslant u)$ dropped from about 0·01, when the common value was about 0·90, to 0·005, when the common value was 0·98. From 50 % down to 7 %, the theoretical points follow those of U^2_{10} and below it are nearer those of $U^2_{10,10}$. In using the theoretical 5 % point for U^2_{10} one would be in fact working at the 4·5 %

level, as judged by those results. For $U^2_{10,10}$, no sensible error is made. At 1 %, the error for both is about 0·05 %. Since one may expect better agreement between the small and large sample significance points for samples bigger than 10, no serious error will be made in using the asymptotic distribution for the U^2 statistics. At this point it may be remarked that it is not necessary to have a table. For if

$$\alpha = P(U^2 > v),$$

then, from (2),

$$v \approx \frac{1}{2\pi^2}\left\{\log_e \frac{\alpha}{2} - \log_e [1 + (\tfrac{1}{2}\alpha)^3]\right\},$$

to excellent accuracy. In fact,

$$v \approx \frac{1}{19{\cdot}74}\log_e \frac{\alpha}{2},$$

will be all that is needed in practice.

From the graphs, $W^2_{10,10}$ is stochastically larger than W^2_{10} by about 1 % in the region of 95 %. The theoretical points follow the experimental points for W^2_{10} more than those of $W^2_{10,10}$ for levels down to 3 % after which the situation reverses. Again the use of the asymptotic approximation will certainly be adequate for practical tests.

4. An example

We will consider some data kindly supplied by Dr K. Schmidt-Koenig of the Department of Zoology, Duke University. In accordance with the theory of the sun-azimuth compass (Matthews, 1955; Kramer, 1957), displaced homing pigeons are misled in a predictable way if their 'internal clock' has been reset by exposure to a time-shifted sequence of day-night cycles. In this experiment Dr Schmidt-Koenig reset the 'clocks' of the experimental birds by 6 hr. clockwise. This would be expected from theory to give a left deviation of the experimental group of roughly 90°. The data are the bearings of the birds flying away from the release point, made just as they vanish in the distance. They are measured only to the nearest 5° and some account of this grouping is taken in the analysis.

Control group ($N_1 = 12$): 50, 290, 300, 300, 305, 320, 330, 330, 335, 340, 340, 355.

Experimental group ($N_2 = 14$): 70, 155, 190, 195, 215, 235, 235, 240, 255, 260, 290, 300, 300, 300.

To compute the statistic (4), it is necessary to set the samples down in two columns so that they are jointly ordered. If there are no ties, either within or between samples, the table will have $N_1 + N_2$ rows. In a third column the value of $F_{N_1}(x) - F_{N_2}(x)$ may be written as a difference of fractions, $i/N_1 - j/N_2$, and in a fourth this difference should be expressed as a decimal. The sum S_1 and sum of squares S_2 of this fourth column are then found and hence

$$U^2_{N_1,N_2} = \frac{N_1 N_2}{(N_1+N_2)^2}\left(S_2 - \frac{S_1^2}{N_1+N_2}\right).$$

In the present example there are a number of ties. Let both samples be ordered in one column, with each repeated value showing only once. Call these numbers $z_1, z_2, \ldots$ and suppose z_k is repeated in the data n_k times. Then (4) now reads

$$U^2_{N_1,N_2} = \frac{N_1 N_2}{(N_1+N_2)^2}\left\{\sum_k [F_{N_1}(z_k) - F_{N_2}(z_k)]^2 n_k - \frac{1}{N_1+N_2}[\sum_k \{F_{N_1}(z_k) - F_{N_2}(z_k)\} n_k]^2\right\}.$$

To find the values of $F_{N_1}(.)$ and $F_{N_2}(.)$, it is necessary to order jointly the two samples as before but not to list repeats *within* samples. An additional column for n_k must be added.

The computations are then similar to those above. In the present example, these calculations take following tabular form.

Z_k $x_{(i)}$	Z_k $y_{(j)}$	$\frac{i}{N_1}-\frac{j}{N_2}$	$= F_{N1}(Z_k)-F_{N2}(Z_k)$	n_k
50 (1)		1/12– 0/14	0·08333	1
	70 (1)	1/12– 1/14	·01190	1
	155 (1)	1/12– 2/14	−·05952	1
	190 (1)	1/12– 3/14	−·13095	1
	195 (1)	1/12– 4/14	−·20238	1
	215 (1)	1/12– 5/14	−·27381	1
	235 (2)	1/12– 7/14	−·41667	2
	240 (1)	1/12– 8/14	−·48810	1
	255 (1)	1/12– 9/14	−·55952	1
	260 (1)	1/12–10/14	−·63095	1
290 (1)	290 (1)	2/12–11/14	−·61905	2
300 (2)	300 (3)	4/12–14/14	−·66667	5
305 (1)		5/12– 1	−·58333	1
320 (1)		6/12– 1	−·50000	1
330 (2)		8/12– 1	−·33333	2
335 (1)		9/12– 1	−·25000	1
340 (2)		11/12– 1	−·08333	2
355 (1)		12/12– 1	0·00000	1
(N_1 = 12)	(N_2 = 14)			26

We find

$$S_1 = \sum_k n_k\{F_{N_1}(Z_k)-F_{N_2}(Z_k)\} = -9{\cdot}8214,$$

$$S_2 = \sum_k n_k\{F_{N_1}(Z_k)-F_{N_2}(Z_k)\}^2 = 5{\cdot}3179,$$

$$U^2_{12,14} = 0{\cdot}3996.$$

The asymptotic probability of U^2 exceeding this value on the null hypothesis is about 10^{-3}. From the trend of the sampling results shown in the table of § 3, the real probability on the null hypothesis, will be even smaller. The ties here are due of course to grouping, i.e. to coarse measurement. To see the effect of this suppose that the 290, 300, 300 in the control group had been 285, 295, 295, i.e. *all* between sample ties broken in favour of the null hypothesis. In this case, it is found that $U^2_{12,14} = 0{\cdot}3204$ which is still highly significant. Thus the data furnishes strong evidence that the control and experimental birds have been chosen from different populations. However, to test whether the data supports the theory, it would be more useful to add 90° to the vanishing angles of all the experimental birds before performing the test. With these figures $U^2_{12,14}$ become equal to 0·113, a value which is exceeded, on the null hypothesis, with a probability of approximately 20 %. Hence the data does support the theory. It would be possible to find a set of angular shifts which, like 90°, lead to non-significant values of U^2, at say the 5 % level. This would give an approximate 95 % confidence set for the angular displacement caused by the 'treatment' given the experimental birds. There seems to be no way of finding this set except by a rather tedious trial and error process.

I am indebted to Prof. E. S. Pearson for some helpful suggestions in the presentation of the paper.

References

Anderson, T. W. & Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Ann. Math. Statist.* **23**, 193–212.

Burr, E. J. (1961). To be submitted to *Biometrika.*

Darling, D. A. (1960). Sur les théorèmes de Kolmogorov–Smirnov. *Teor. Ver. i ee Prim.* **5**, 393–8.

Fisz, M. (1960). On a result by M. Rosenblatt concerning the von Mises–Smirnov test. *Ann. Math. Statist.* **31**, 427–9.

Kiefer, J. A. (1959). K sample analogues of the Kolmogorov–Smirnov and Cramér–von Mises tests. *Ann. Math. Statist.* **30**, 420–47.

Kramer, G. (1957). Experiments on bird orientation and their interpretation. *Ibis*, **99**, 196–227.

Kuiper, N. H. (1960). Tests concerning random points on a circle. *Proc. Koninkl. Nederl. Akad. Van Wetenschappen*, Series A, 383–97.

Matthews, G. V. T. (1955). *Bird Navigation.* Cambridge University Press.

Rosenblatt, M. (1952). Limit theorems associated with variants of the von Mises statistic. *Ann. Math. Statist.* **23**, 617–23.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika*, **48**, 109–14.