equal. Now

$$E(y_{i_1}^{\alpha_1} y_{i_2}^{\alpha_2} \cdots y_{i_k}^{\alpha_k}) = E(y_{i_1} y_{i_2} \cdots y_{i_k})$$

is the probability that $y_{i_1} = 1, y_{i_2} = 1, \cdots, y_{i_k} = 1$, simultaneously. This probability is either zero (for example, when $|i_2 - i_1| = 1, |i_3 - i_2| = 1$, etc. or when $i_1 = n$, etc.) or $\left(\frac{2}{n}\right)^k + O\left(\frac{1}{n^{k+1}}\right)$. Moreover, the ratio of the number of k-tuples $i_1, i_2, \cdots, i_k$ for which the probability is zero to the number of k-tuples for which the probability is $\left(\frac{2}{n}\right)^k + O\left(\frac{1}{n^{k+1}}\right)$ is $O\left(\frac{1}{n}\right)$. Let $Z_i (i = 1, \cdots, n)$ be independent stochastic variables each with the same distribution such that the probability that $Z_i = 1$ is $2/n$ and the probability that $Z_i = 0$ is $(n - 2)/n$. It follows readily that the limit, as $n \to \infty$, of the jth moment ($j = 1, 2, \cdots$, ad inf.) of y about the origin, is the same as the limit of the same moment of Z, where

$$Z = \sum_{i=1}^{n} Z_i.$$

Since the Z_i are independently distributed, and since each can take only the values 0 and 1, the probability of the value 1 being $2/n$, the jth moment of Z about the origin approaches, as $n \to \infty$,

$$\mu_j = e^{-2} \sum_{i=1}^{\infty} \frac{i^j 2^i}{i!},$$

which is the jth moment about the origin of the Poisson distribution with mean value 2. By the preceding paragraph, μ_j is also the limit of the jth moment of y about the origin. Now von Mises [2] has proved that if the jth moment ($j = 1, 2, \cdots$, ad inf.) of a chance variable X_n, ($n = 1, 2, \cdots$, ad inf.), approaches, as $n \to \infty$, the jth moment of a Poisson distribution, then the distribution of X_n approaches the Poisson distribution with corresponding mean value. From this it follows that y has in the limit the distribution (1). We have already shown that y and $n - W(R)$ have the same limiting distribution, so that the required result follows.

REFERENCES

[1] J. Wolfowitz, *Annals of Math. Stat.*, Vol. 13 (1942), p. 247.
[2] R. v. Mises, *Zeitschrift für die angewandte Math. und Mechanik*, Vol. 1 (1921), p. 298.

NOTE ON CONSISTENCY OF A PROPOSED TEST FOR THE PROBLEM OF TWO SAMPLES

By Albert H. Bowker
Columbia University

Certain tests for the hypothesis that two samples are from the same population assume nothing about the distribution function except that it is continuous. Since the power functions of these tests have not been obtained, optimum tests are not known. However, one desirable[1] test property, that of "consistency," has been introduced by Wald and Wolfowitz [1]. A test is called consistent if the probability of rejecting the null hypothesis when it is false (the power of the test) approaches one as the sample number approaches infinity. This is a logical extension of the familiar idea of consistency introduced by Fisher. It will be shown that a test recently proposed by Mathisen [2] is not consistent with respect to certain alternatives.

The test proposed by Mathisen [2] may be described briefly as follows: Given two samples, observe the number (m) of elements of the second sample whose values are less than the median of the first sample. The distribution of m is independent of the population distribution under the null hypothesis. Let $P\{m < a\}$ denote the probability of the relation in braces under the null hypothesis. If m_1 and m_2 are significance points ($m_1 > m_2$) such that

$$(1)\qquad \begin{aligned} P\{m > m_1\} &= \beta_1 \\ P\{m < m_2\} &= \beta_2 \\ \beta_1 + \beta_2 &= \beta < 1, \end{aligned}$$

the statistic m can be used to test the hypothesis at the significance level β. This is called the case of two intervals. The method is extended by using the two quartiles and the median of the first sample to define four intervals into which the elements of the second sample may fall. If the second sample is of size $4n$ and the number which actually falls in each interval is n_1, n_2, n_3, and n_4 respectively, the distribution of

$$(2)\qquad C = \frac{\sum_{i=1}^{4} (n_i - n)^2}{9n^2}$$

is also independent of the population distribution under the null hypothesis. Then if C^* is a significance point, such that

$$(3)\qquad P\{C > C^*\} = \beta' < 1,$$

C can be used as a test of the hypothesis at the level β'.

To show that Mathisen's test is not consistent, we shall consider first the case of two intervals. Let X and Y be two independent stochastic variables whose cumulative distribution functions $F(x)$ and $G(x)$ are continuous. Let $x_1 < x_2 \cdots < x_{2n+1}$ and $y_1 < y_2 \cdots < y_{2n}$ be sets of ordered independent observations on X and Y. Then m is such that

$$y_m < x_{n+1} < y_{m+1}.$$

Let m_1 and m_2 be the significance points of the distribution of m, defined by (1). Clearly m_1 and m_2 depend on n. We shall prove that the sequence

$$(4)\qquad \frac{m_1(n)}{2n} \qquad n = 1, 2, \cdots.$$

[1] For large samples.

converges to $\frac{1}{2}$. Since (4) is bounded, it has at least one limit point. Let h be such a limit point. If $h < \frac{1}{2}$ and $\frac{1}{2} - h = 3\delta$, then there exists a monotonically increasing subsequence of the integers $n_1, n_2, \cdots$ and a number N such that for $n_i > N$

$$\left| \frac{m_1(n_i)}{2n_i} - h \right| < \delta. \tag{5}$$

Clearly $m/2n$ converges stochastically to $\frac{1}{2}$. Hence if $0 < \epsilon < 1$ is any arbitrarily small number, we can select n so large that the probability is at least $1 - \epsilon$ that

$$\left| \frac{m}{2n} - \frac{1}{2} \right| < \delta. \tag{6}$$

Hence for n sufficiently large, $P\{m > m_1\}$ is at least $1 - \epsilon$, a contradiction with (1). A similar contradiction appears if $h > \frac{1}{2}$. Hence (4) has only one limit point, $\frac{1}{2}$. In the same way we can prove that the sequence

$$\frac{m_2(n)}{2n} \qquad n = 1, 2, \cdots. \tag{7}$$

also converges to $\frac{1}{2}$.

Let $0 < \delta \leq \frac{1}{6}$. Consider now two pairs of populations, A and B, described as follows:

A) $F(x) \equiv G(x) \equiv x \qquad (0 \leq x \leq 1)$

B) $F(x) \equiv x \qquad (0 \leq x \leq 1)$

$G(x) \equiv 0 \qquad (0 \leq x \leq \frac{1}{2} - 2\delta)$

$G(x) \equiv (x - \frac{1}{2} + 2\delta)(\frac{1}{2} - \delta)/\delta \qquad (\frac{1}{2} - 2\delta \leq x \leq \frac{1}{2} - \delta)$

$G(x) \equiv x \qquad (\frac{1}{2} - \delta \leq x \leq \frac{1}{2} + \delta)$

$G(x) \equiv \frac{1}{2} + \delta \qquad (\frac{1}{2} + \delta \leq x \leq 1 - \delta)$

$G(x) \equiv (\frac{1}{2} + \delta) + (x - 1 + \delta)(\frac{1}{2} - \delta)/\delta \qquad (1 - \delta \leq x \leq 1)$

For both A and B, $F(x) \equiv G(x) \equiv 0$ for $x < 0$ and $F(x) \equiv G(x) \equiv 1$ for $x > 1$. For B, it will be shown that there exist values of n greater than any preassigned arbitrarily large number, such that the probability of rejecting the hypothesis when it is false is less than $\beta_1 + \beta_2 + \epsilon$ where ϵ is an arbitrarily small positive number.

Let h_1, h_2, h_3 denote the number of observations on X which fall in the intervals $0 < x \leq \frac{1}{2} - \delta$, $\frac{1}{2} - \delta < x \leq \frac{1}{2} + \delta$, $\frac{1}{2} + \delta < x \leq 1$ respectively for a fixed value of n. Let h_1', h_2', h_3' be the corresponding numbers for Y. For a fixed n, the probability of a set $h_1, h_2, h_3, h_1', h_2', h_3'$ is the same whether the samples be drawn from A or B. From (4), (7), and the stochastic convergence of $m/2n$, it follows that we can find an N such that for all $n > N$ the probability is at least $1 - \epsilon/2$ of the occurrence of a set h_1, h_2, h_3, h_1', h_2', h_3' for which y_{m_1}, x_{n+1}, y_{m_2} will fall in the interval $(\frac{1}{2} - \delta, \frac{1}{2} + \delta)$. Furthermore, for fixed h_2, h_2' the distribution within the interval is the same whether the sample came from A or B. Hence, even when the sample is drawn from B, for n sufficiently large,

$$P\{m > m_1\} < \beta_1 + \frac{\epsilon}{2}$$

$$P\{m < m_2\} < \beta_2 + \frac{\epsilon}{2}.$$

That is, for samples of sufficiently large size from B, the probability of rejecting the null hypothesis is at most $\beta_1 + \beta_2 + \epsilon$. Since $\beta_1 + \beta_2 < 1$ and ϵ is arbitrarily small, the probability can be made less than one and the test is not consistent in the case of two intervals.

In the case of four intervals, the proof is similar. In this case, we assume that the second sample has size $4n$. Clearly, $n_1/4n$, $n_2/4n$, $n_3/4n$, and $n_4/4n$ converge stochastically to $\frac{1}{4}$. If C^* is the significance point defined by (3), the sequence

$$C^*(n) \qquad n = 1, 2, \cdots$$

converges to zero. Now consider two pairs, A and B, of populations. A is the same as before and B consists of one uniform distribution and one which is identical with the uniform distribution in small intervals containing $x = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, and 1, but is different everywhere else. As before, $F(x) \equiv G(x) \equiv 0$ for $x < 0$ and $F(x) \equiv G(x) \equiv 1$ for $x > 1$. Then for B, when n is large, the behavior of C, except for a probability arbitrarily near zero, will depend only on the intervals of coincidence. Hence for B

$$P\{C > C^*\} < \beta' + \epsilon$$

where ϵ is any arbitrarily small positive quantity.

Returning to the case of two intervals, if the samples are from different populations and if their cumulative distribution functions are identical in the neighborhood of their medians, the test is not consistent. If such a possibility is excluded from the class of admissible alternatives, we may expect that the test will be consistent. For example, if the class of alternatives is limited to those where $G(x) \equiv F(x + c)$, c a constant, the test will be consistent. A similar remark holds for the case of four intervals or for any fixed finite number of intervals. It appears, however, that if the number of intervals is a function of the sample size (say $\sqrt{n}$) and becomes infinite with sample size, a test of this kind will be consistent with respect to a general class of alternatives.

REFERENCES

[1] A. Wald and J. Wolfowitz, "On a test whether two samples are from the same population," *Annals of Math. Stat.*, Vol. XI (1940), pp. 147–62.

[2] H. C. Mathisen, "A method of testing the hypothesis that two samples are from the same population," *Annals of Math. Stat.*, Vol. XIV (1943), pp. 188–94.