LIMIT THEOREMS ASSOCIATED WITH VARIANTS OF THE VON MISES STATISTIC[1]

BY M. ROSENBLATT

University of Chicago

1. Summary. A multidimensional analogue of the von Mises statistic is considered for the case of sampling from a multidimensional uniform distribution. The limiting distribution of the statistic is shown to be that of a weighted sum of independent chi-square random variables with one degree of freedom. The weights are the eigenvalues of a positive definite symmetric function.

A modified statistic of the von Mises type useful in setting up a two sample test is shown to have the same limiting distribution under the null hypothesis (both samples come from the same population with a continuous distribution function) as that of the one-dimensional von Mises statistic. We call the statistics mentioned above von Mises statistics because they are modifications of the ω^2 criterion considered by von Mises [5].

The paper makes use of elements of the theory of stochastic processes.

2. Introduction. Let $X_i = (X_{i1}, \cdots, X_{ik})$, $i = 1, \cdots, n$, be a sample from a k-dimensional uniform distribution; that is, x_{ij}, $i = 1, \cdots, n$, $j = 1, \cdots, k$, are independent and uniformly distributed on [0, 1]. Let

$$\phi_t(x) = \begin{cases} 1 \text{ if } x \leqq t, \\ 0 \text{ if } x > t. \end{cases} \tag{1}$$

The sample distribution function is

$$S_n(\bar{t}) = S_n(t_1, \cdots, t_k) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{k} \phi_{t_j}(X_{ij}), \tag{2}$$

where $\bar{t} = (t_1, \cdots, t_k)$. Consider the process

$$Y_n(\bar{t}) = \sqrt{n}(S_n(t_1, \cdots, t_k) - t_1 \cdots t_k), \qquad 0 \leqq t_1, \cdots, t_k \leqq 1. \tag{3}$$

Clearly $EY_n(\bar{t}) = 0$. The covariance of the process is

$$\begin{aligned} E(Y_n(\bar{t})Y_n(\bar{t}')) &= r_n(\bar{t}, \bar{t}') \\ &= \frac{1}{n} E\left[\sum_{i=1}^{n}\left\{\prod_{j=1}^{k} \phi_{t_j}(X_{ij}) - t_1 \cdots t_k\right\} \sum_{i=1}^{n}\left\{\prod_{j=1}^{k} \phi_{t_j'}(X_{ij}) - t_1' \cdots t_k'\right\}\right] \\ &= \frac{1}{n} \sum_{i=1}^{n} E\left(\left\{\prod_{j=1}^{k} \phi_{t_j}(X_{ij}) - t_1 \cdots t_k\right\}\left\{\prod_{j=1}^{k} \phi_{t_j'}(X_{ij}) - t_1' \cdots t_k'\right\}\right) \\ &= \prod_{j=1}^{k} \min(t_j, t_j') - t_1 \cdots t_k t_1' \cdots t_k'. \end{aligned} \tag{4}$$

[1] Work done under contract with the Office of Naval Research.

Note that the covariance function $r_n(\bar{t}, \bar{t}')$ is independent of n and symmetric in $\bar{t}$ and $\bar{t}'$.

Consider the function $r_n(\bar{t}, \bar{t}') = r(\bar{t}, \bar{t}')$ as the kernel in the following eigenvalue problem

$$\int_0^1 r(\bar{t}, \bar{t}')\phi(\bar{t}')\, d\bar{t}' = \lambda\phi(\bar{t}),$$

where the integral is over all components of $\bar{t}'$. The kernel is positive definite (being a covariance function) and hence all its eigenvalues are positive. There are a denumerable number of eigenvalues. Denote the eigenvalues by $\lambda_1, \lambda_2, \cdots$ and the corresponding orthonormal eigenfunctions by

$$\phi_1(\bar{t}), \qquad \phi_2(\bar{t}), \qquad \cdots .$$

It is understood that each eigenvalue is repeated with the multiplicity of the linearly independent eigenfunctions corresponding to it. Now

$$r(\bar{t}, \bar{t}') = \sum_{j=1}^{\infty} \lambda_j \phi_j(\bar{t})\phi_j(\bar{t}') \tag{5}$$

with uniform convergence according to Mercer's theorem. The general theorem of Karhunen on representation of stochastic processes [3] then implies that

$$Y_n(\bar{t}) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\, \phi_j(\bar{t}) Y_{nj} \tag{6}$$

in the mean square, where

$$EY_{nj} = 0, \qquad EY_{nj}Y_{nk} = \delta_{jk} .$$

3. The limiting distribution. As $n \to \infty$, the joint distribution of $Y_n(\bar{t}_1), \cdots, Y_n(\bar{t}_m)$ approaches the joint distribution of $Y(\bar{t}_1), \cdots, Y(\bar{t}_m)$, where $Y(\bar{t})$ is a normal process with mean zero and covariance $r(\bar{t}, \bar{t}')$. Obviously the process

$$Y(\bar{t}) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\, \phi_j(\bar{t}) Y_j , \tag{7}$$

where the Y_j are independent normal random variables with mean zero and variance one.

THEOREM 1. *The von Mises statistic corresponding to a sample of n from a k-dimensional uniform distribution is*

$$\int_0^1 Y_n^2(\bar{t})\, d\bar{t} = \sum_{j=1}^{\infty} \lambda_j Y_{nj}^2 , \tag{8}$$

and the limiting distribution of (8) *as $n \to \infty$ is that of*

$$\int_0^1 Y^2(\bar{t})\, d\bar{t} = \sum_{j=1}^{\infty} \lambda_j Y_j^2 . \tag{9}$$

PROOF. Now

$$Y_n(\bar{t}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(\bar{t}),$$

where the random variables

$$Z_i(\bar{t}) = \prod_{j=1}^{k} \phi_{t_j}(X_{ij}) - t_1 \cdots t_k$$

are independent and identically distributed. Then

$$Y_{nj} = \sum_{k=1}^{n} \frac{1}{\sqrt{n}} Z_{kj},$$

where

$$Z_{kj} = \frac{1}{\sqrt{\lambda_j}} \int_0^1 Z_k(\bar{t})\phi_j(\bar{t})\, d\bar{t}.$$

The random vectors

$$(Z_{k1}, \cdots, Z_{kN}), \qquad k = 1, \cdots, n,$$

are independent and identically distributed. Moreover

$$EZ_{kj} = 0,$$

$$EZ_{kj}Z_{kl} = \delta_{jl}.$$

The multidimensional central limit theorem then implies that the random variables Y_{nj}, $j = 1, \cdots, N$, are asymptotically normal, independent random variables with mean zero and variance one as $n \to \infty$. $Y_{n1}^2, \cdots, Y_{nN}^2$ as $n \to \infty$ are asymptotically independent chi-square random variables with one degree of freedom and mean one. Note that

$$\int_0^1 r(\bar{t}, \bar{t})\, d\bar{t} = \sum_{j=1}^{\infty} \lambda_j. \tag{10}$$

Given any $\epsilon > 0$, let $N(\epsilon)$ be such that

$$\sum_{N(\epsilon)+1}^{\infty} \lambda_j < \epsilon^2.$$

Let

$$U_N = \sum_{j=1}^{N(\epsilon)} \lambda_j Y_{nj}^2,$$

$$V_N = \sum_{N(\epsilon)+1}^{\infty} \lambda_j Y_{nj}^2.$$

U_N asymptotically has the same distribution as

$$\sum_{1}^{N(\epsilon)} \lambda_j Y_j^2;$$

that is, for sufficiently large n

$$\left| P\{U_N \leqq x\} - P\left\{\sum_1^{N(\epsilon)} \lambda_j Y_{\cdot j}^2 \leqq x\right\}\right| < \epsilon.$$

The choice of $N(\epsilon)$ and Tchebycheff's inequality imply

$$\left| P\left\{\int_0^1 Y^2(\bar{t})\, d\bar{t} \leqq x + \epsilon\right\} - P\left\{\sum_1^{N(\epsilon)} \lambda_j Y_j^2 \leqq x\right\}\right| < \epsilon,$$

$$\left| P\{U_N \leqq x\} - P\left\{\int_0^1 Y_N^2(\bar{t})\, d\bar{t} \leqq x + \epsilon\right\}\right| < \epsilon.$$

Hence

$$\left| P\left\{\int_0^1 Y_n^2(\bar{t})\, d\bar{t} \leqq x + \epsilon\right\} - P\left\{\int_0^1 Y^2(\bar{t})\, d\bar{t} \leqq x + \epsilon\right\}\right| < 3\epsilon$$

for sufficiently large n. The distribution function of $\int_0^1 Y^2(\bar{t})\, d\bar{t}$ is continuous. Therefore the limiting distribution of $\int_0^1 Y_n^2(\bar{t})\, d\bar{t}$ as $n \to \infty$ is the same as that of $\int_0^1 Y^2(\bar{t})\, d\bar{t}$.

The distribution function of $\int_0^1 Y^2(\bar{t})\, d\bar{t}$ has been computed in the 1-dimensional case ($k = 1$). The eigenvalues of (4) are then $\lambda_j = 1/(\pi^2 j^2)$ $j = 1, 2, \cdots$ and hence the characteristic function of (9) is $\prod_{j=1}^{\infty} [1 - 2it/(\pi^2 j^2)]^{-\frac{1}{2}}$. One can invert the characteristic function by a contour integration and obtain the distribution function of (9) as given by Smirnov [5, 2]. It would be of great interest to find the eigenvalues of (4) when $k > 1$.

4. The two sample test. Let X_{1j}, $j = 1, \cdots, n$, and X_{2k}, $k = 1, \cdots, m$, be samples of n and m respectively from a population with some continuous distribution function $F(x)$. Let $S_1(t)$, $S_2(t)$ be the corresponding sample distribution functions. Various people [4] have suggested using

$$\frac{mn}{m+n} \int_{-\infty}^{\infty} (S_1(t) - S_2(t))^2\, d\left(\frac{S_1(t) + S_2(t)}{2}\right) \tag{11}$$

as a test statistic for the two sample problem.

THEOREM 2. *Statistic* (11) *has the same limiting distribution when* $n \to \infty$, $m/n \to \lambda > 0$ *as the one-dimensional von Mises statistic under the assumption that both samples come from the same continuous population.*

Consider computing the statistic for samples $F(X_{1j})$, $j = 1, \cdots, n$, $F(X_{2k})$, $k = 1, \cdots, m$, of n and m respectively from a population with the uniform distribution. The value of the statistic is the same as that obtained from the orig-

inal samples $\{X_{1j}\}$, $\{X_{2k}\}$ and consequently has the same distribution function as the latter. We need then only consider the statistic

$$\frac{mn}{m+n}\int_0^1 (S_1(t) - S_2(t))^2\, d\left(\frac{S_1(t)+S_2(t)}{2}\right) \tag{12}$$

when the samples are from a uniformly distributed population. It is obvious that

$$\frac{mn}{m+n}\int_0^1 (S_1(t) - S_2(t))^2\, dt \tag{13}$$

has the same limiting distribution as the one-dimensional von Mises statistic when $n \to \infty$, $m/n \to \lambda > 0$. It would then be sufficient to show that

$$\frac{mn}{m+n}\int_0^1 (S_1(t) - S_2(t))^2\, d\left(\frac{S_1(t)+S_2(t)}{2} - t\right) \tag{14}$$

converges to zero in probability when $n \to \infty$, $m/n \to \lambda > 0$. Now

$$\begin{aligned}
&\frac{mn}{m+n}\int_0^1 (S_1(t) - S_2(t))^2\, d\left(\frac{S_1(t)+S_2(t)}{2} - t\right) \\
&\quad = \frac{mn}{m+n}\int_0^1 (S_1(t) - S_2(t))^2\, d(S_2(t) - t) \\
&\quad = \frac{mn}{m+n}\int_0^1 (S_1(t) - t)^2\, d(S_2(t) - t) + \frac{mn}{m+n}\int_0^1 (S_2(t) - t)^2\, d(S_1(t) - t)
\end{aligned}$$

can be obtained by a series of integrations by parts. The proof is complete if one can show that both expressions directly above converge to zero in probability. By symmetry it is enough to consider one of the expressions.

Let

$$\begin{aligned}
x_1(t) &= n^{\frac{1}{2}}(S_1(t) - t), \\
x_2(t) &= m^{\frac{1}{2}}(S_2(t) - t).
\end{aligned} \tag{15}$$

Now

$$\begin{aligned}
Ex_i(t) &= 0, \\
Ex_i(\tau)x_i(t) &= \min(\tau, t) - \tau t, \qquad i = 1, 2.
\end{aligned} \tag{16}$$

We use the following transformation suggested by Doob [1]

$$x_i(t) = (t-1)Z_i\left(\frac{t}{1-t}\right), \qquad i = 1, 2. \tag{17}$$

The processes $Z_1(t)$, $Z_2(t)$ are independent of each other. Moreover, each of them is an orthogonal process with

$$\begin{aligned}
EZ_i(t) &= 0, \\
EZ_i(t)Z_i(\tau) &= \min(\tau, t), \qquad i = 1, 2.
\end{aligned} \tag{18}$$

A simple computation making use of (2), (15), (17) yields

$$EZ_1^2(t)Z_1^2(\tau) = \frac{1}{n}\left[\frac{\min(t,\tau)}{1+\min(t,\tau)} + \frac{\min(t,\tau)(\max(t,\tau)-\min(t,\tau))}{(1+t)(1+\tau)} + \frac{t^2\tau^2}{1+\max(\tau,t)}\right] + \frac{n-1}{n}t\tau + 2\frac{n-1}{n}(\min(t,\tau))^2 \tag{19}$$

and in particular

$$EZ_1^4(t) = \frac{1}{n}\frac{t+t^4}{1+t} + 3\frac{n-1}{n}t^2. \tag{20}$$

Now

$$\frac{mn}{m+n}\int_0^1 (S_1(t)-t)^2\,d(S_2(t)-t) = \frac{m^{\frac{1}{2}}}{m+n}\int_0^1 x_1^2(t)\,dx_2(t)$$

$$= \frac{m^{\frac{1}{2}}}{m+n}\int_0^1 (t-1)^2 Z_1^2\left(\frac{t}{1-t}\right)Z_2\left(\frac{t}{1-t}\right)dt + \frac{m^{\frac{1}{2}}}{m+n}\int_0^1 (t-1)^3 Z_1^2\left(\frac{t}{1-t}\right)dZ_2\left(\frac{t}{1-t}\right) \tag{21}$$

$$= \frac{m^{\frac{1}{2}}}{m+n}\int_0^\infty \frac{1}{(t+1)^4} Z_1^2(t)Z_2(t)\,dt \tag{22}$$

$$+ \frac{m^{\frac{1}{2}}}{m+n}\int_0^\infty \frac{1}{(t+1)^3} Z_1^2(t)\,dZ_2(t). \tag{23}$$

The random variables (22), (23) are the limits almost everywhere of

$$\frac{m^{\frac{1}{2}}}{m+n}\int_0^T \frac{1}{(t+1)^4} Z_1^2(t)Z_2(t)\,dt \tag{24}$$

and

$$\frac{m^{\frac{1}{2}}}{m+n}\int_0^T \frac{1}{(t+1)^3} Z_1^2(t)\,dZ_2(t), \tag{25}$$

respectively, as $T \to \infty$. The independence of the orthogonal processes $Z_1(t)$, $Z_2(t)$ implies that the second moments of (24), (25) are

$$\frac{m}{(m+n)^2}\int_0^T\int_0^T \frac{\min(\tau,t)E(Z_1^2(t)Z_1^2(\tau))}{(1+t)^4(1+\tau)^4}\,dt\,d\tau$$

and

$$\frac{m}{(m+n)^2}\int_0^T \frac{1}{(1+t)^6} E(Z_1^4(t))\,dt,$$

respectively.

Making use of (19), (20) one can see that (24), (25) converge in mean square as $T \to \infty$ to (22), (23) respectively. But then the second moments of (22), (23) exist and are given by

$$\frac{m}{(m+n)^2} \int_0^\infty \int_0^\infty \frac{\min(\tau, t) E(Z_1^2(t) Z_1^2(\tau))}{(1+t)^4 (1+\tau)^4} \, dt \, d\tau \tag{26}$$

and

$$\frac{m}{(m+n)^2} \int_0^\infty \frac{1}{(1+t)^6} E(Z_1^4(t)) \, dt, \tag{27}$$

respectively. The second moments (26), (27) converge to zero as $n \to \infty$, $m/n \to \lambda > 0$ and hence the random variables (22), (23) converge to zero in probability as $n \to \infty$, $m/n \to \lambda > 0$. This in turn implies that (21) converges to zero in probability. The same argument implies that

$$\frac{mn}{m+n} \int_0^1 (s_2(t) - t)^2 \, d(s_1(t) - t)$$

converges to zero in probability. Hence (14) converges to zero in probability as $n \to \infty$, $m/n \to \lambda > 0$ and the proof is complete.

REFERENCES

[1] J. L. Doob, "Heuristic approach to Kolmogorov-Smirnov theorems," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 393–403.

[2] M. Kac, "On some connections between probability theory and differential and integral equations," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 1951.

[3] K. Karhunen, "Uber lineare Methoden in der Warscheinlichkeitsrechnung," *Ann. Acad. Sci. Fennicae*, Series *A. I.*, no. 37 (1947), 79 pp.

[4] E. L. Lehmann, "Consistency and unbiasedness of certain nonparametric tests," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 165–179.

[5] N. Smirnoff, "Sur la distribution de ω^2," *Comptes Rendus de l'Academie des Sciences*, Vol. 202 (1936), p. 449.