JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 68B, No. 2, April–June 1964

On the Asymptotic Joint Normality of Quantiles From a Multivariate Distribution[1]

Lionel Weiss[2]

(December 18, 1963)

A simple proof is given of the asymptotic joint normality of sample quantiles from a multivariate population, under very mild conditions. The joint cumulative distribution function of the quantiles is studied, rather than the joint probability density function.

1. Introduction

Suppose $X_1, \ldots, X_k$ are jointly distributed random variables, with absolutely continuous cumulative distribution function $F(x_1, \ldots, x_k)$. $F_i(x_i)$ and $f_i(x_i)$ denote respectively the marginal cumulative distribution function and marginal density function for X_i. $F_{i,j}(x_i, x_j)$ denotes the joint marginal cumulative distribution function for X_i, X_j.

Let $b_1, \ldots, b_k$ be fixed quantities, each in the open interval $(0, 1)$. Suppose n independent observations are taken on the vector $(X_1, \ldots, X_k)$. Denote by $Y_i(n)$ the b_ith sample quantile of the n observed values of X_i. That is, $Y_i(n)$ is equal to that one of the n observed X_i's such that exactly $[nb_i]$ other X_i's are below it, where $[nb_i]$ denotes the largest integer contained in nb_i. We can ignore the possibility of ties, which occur with probability zero.

We assume that the equation $F_i(y)=b_i$ has a unique solution in y, which we denote by u_i $(i=1, \ldots, k)$, and that $f_i(x_i)$ is continuous and positive at $x_i=u_i$. Define

$$Z_i(n)=\sqrt{n}\, f_i(u_i)\frac{Y_i(n)-u_i}{\sqrt{b_i(1-b_i)}}.$$

We shall show that as n increases, the joint cumulative distribution function of $Z_1(n), \ldots, Z_k(n)$ approaches the k-variate normal cumulative distribution function with zero means, unit variances, and the covariance between the ith and jth variables $(i \neq j)$ given by

$$\frac{F_{i,j}(u_i, u_j)-b_ib_j}{\sqrt{b_ib_j(1-b_i)(1-b_j)}}.$$

This was proved by Mood [2] for the special case $b_i=\ldots=b_k=1/2$, and by Siddiqui [3] for the case $k=2$ and under the assumption that the first and second partial derivatives of $F(x_1, x_2)$ are continuous at u_1, u_2 and the joint probability density function is positive there. Siddiqui conjectured that the result holds for general k, and sketched a proof, assuming that the partial derivatives of $F(x_1, \ldots, x_k)$ of order k are continuous at $u_1, \ldots, u_k$. Both Mood and Siddiqui used the joint probability density function of $Z_1(n), \ldots, Z_k(n)$, and the resulting computations were lengthy. Below we give a simple proof, under less restrictive conditions. The proof uses the joint cumulative distribution function of $Z_1(n), \ldots, Z_k(n)$.

2. Proof of Asymptotic Normality

Fix a point $(z_1, \ldots, z_k)$ in $(x_1, \ldots, x_k)-$space. Define

$$y_i(n)=u_i+\frac{z_i\sqrt{b_i(1-b_i)}}{f_i(u_i)\sqrt{n}}.$$

Use the point $(y_1(n), \ldots, y_k(n))$ as the origin of the usual system of orthogonal axes in $(x_1, \ldots, x_k)-$space. These axes define 2^k "octants" in the space. We assume that these octants are labeled in some definite order, which will be held fixed. For example, when $k=2$, we have four quadrants, rather than "octants," which we assume are labeled in the conventional way. Let $N_j(n)$ denote the number of the n observed k-dimensional vectors which fall within the jth octant. Because of the continuity of $F(x_1, \ldots, x_k)$, we can ignore the possibility that one or more of the observed vectors will fall on the boundary of an octant. The quantities $N_j(n)$ $(j=1, \ldots, 2^k)$ have a multinomial distribution with parameters n, $p_1(n), \ldots, p_{2^k}(n)$, where $p_j(n)$ is the probability assigned to the jth octant by the distribution function $F(x_1, \ldots, x_k)$. Clearly, as n increases $p_j(n)$ approaches the probability assigned by $F(x_1, \ldots, x_k)$ to the jth octant determined by axes through the point $(u_1, \ldots, u_k)$. We denote this limiting value of $p_j(n)$ by p_j.

Let the symbol $\sum_j^{(i)}$ denote the operation of summing over all those values of j such that in the jth octant, the ith coordinate is greater than $y_i(n)$. There are 2^{k-1} such values of j. For example, when $k=2$,

[1] Sponsored by the United States Army under contract DA-11-022-ORD-2059.
[2] Invited paper; U.S. Army Mathematics Research Center and Cornell University.

using the conventional labeling of the quadrants, $\sum_j^{(1)}$ would sum over $j=1, 4$, and $\sum_j^{(2)}$ would sum over $j=1, 2$. Define $L_i(n)$ as $\sum_j^{(i)} N_j(n)$. Then the event $\{Z_i(n) < z_i\}$ is the same as the event $\{L_i(n) < n-[nb_i]\}$. Therefore

$$\begin{aligned} &P[Z_1(n) < z_1, \ldots, Z_k(n) < z_k] \\ &\qquad = P[L_1(n) < n-[nb_1], \ldots, L_k(n) < n-[nb_k]]. \end{aligned} \tag{1}$$

Next, we define

$$M_j(n) = \frac{N_j(n) - np_j(n)}{\sqrt{np_j(n)}} \qquad (j=1, \ldots, 2^k).$$

It is well known that as n increases, the joint distribution of $M_1(n), \ldots, M_{2^k}(n)$ approaches a normal distribution with zero means, variance of $M_j(n)$ equal to $(1-p_j)$, and covariance between $M_h(n)$ and $M_j(n)$ equal to $-\sqrt{p_h p_j}$. (See Cramer [1][3] p. 318.) We can write

$$L_i(n) = \sum_j^{(i)} [np_j(n) + \sqrt{np_j(n)}\, M_j(n)] \tag{2}$$

Clearly

$$\begin{aligned} \sum_j^{(i)} np_j(n) &= n[1 - F_i(y_i(n))] \\ &= n\left[1 - F_i\left(u_i + \frac{z_i}{\sqrt{n}\, f_i(u_i)} \sqrt{b_i(1-b_i)}\right)\right] \\ &= n(1-b_i) - \sqrt{n}\, z_i \sqrt{b_i(1-b_i)} - n\, g_i(n) \end{aligned} \tag{3}$$

where $\sqrt{n}\, g_i(n)$ approaches zero as n increases. Also,

$$n - [nb_i] = n(1-b_i) + c_i \tag{4}$$

where $|c_i| < 1$. Substituting (2), (3), and (4) in (1), and rearranging terms, we find

$$\begin{aligned} &P[Z_1(n) < z_1, \ldots, Z_k(n) < z_k] \\ &= P\left[\frac{\sum_j^{(i)} \sqrt{p_j(n)} M_j(n)}{\sqrt{b_i(1-b_i)}} < z_i + \frac{\sqrt{n} g_i(n)}{\sqrt{b_i(1-b_i)}} \right. \\ &\qquad \left. + \frac{c_i}{\sqrt{n}\sqrt{b_i(1-b_i)}};\ i=1, \ldots, k\right]. \end{aligned} \tag{5}$$

Since $p_j(n)$ approaches p_j as n increases, and $\sqrt{n} g_i(n)$ and $c_i/\sqrt{n}$ both approach zero as n increases, (5) shows that the joint asymptotic distribution of $Z_1(n), \ldots, Z_k(n)$ is the same as the joint asymptotic distribution of

$$\frac{\sum_j^{(1)} \sqrt{p_j} M_j(n)}{\sqrt{b_1(1-b_1)}}, \ldots, \frac{\sum_j^{(k)} \sqrt{p_j} M_j(n)}{\sqrt{b_k(1-b_k)}} \tag{6}$$

if this latter asymptotic joint distribution is absolutely continuous, which it is, since it is a nonsingular k-variate normal distribution. Therefore the asymptotic joint distribution of $Z_1(n), \ldots, Z_k(n)$ is normal, with zero means and variances and covariances the same as in the asymptotic joint distribution of the variables listed in (6). In computing these variances and covariances, we can look at just two of the variables in (6) at a time, and there is no loss of generality in looking at the first two. In doing this, we can assume that we are taking n observations on (X_1, X_2), ignoring $(X_3, \ldots, X_k)$. This reduces the problem to the case $k=2$, where $F_{1,2}(x_1, x_2)$ is the joint distribution of the basic variables, and we have four quadrants. Using the conventional labeling of the quadrants, we have

$$\frac{\sum_j^{(1)} \sqrt{p_j} M_j(n)}{\sqrt{b_1(1-b_1)}} = \frac{\sqrt{p_1} M_1(n) + \sqrt{p_4} M_4(n)}{\sqrt{b_1(1-b_1)}}$$

$$\frac{\sum_j^{(2)} \sqrt{p_j} M_j(n)}{\sqrt{b_2(1-b_2)}} = \frac{\sqrt{p_1} M_1(n) + \sqrt{p_2} M_2(n)}{\sqrt{b_2(1-b_2)}}$$

$$\begin{aligned} p_1 &= 1 - b_1 - b_2 + F_{1,2}(u_1, u_2) \\ p_2 &= b_1 - F_{1,2}(u_1, u_2) \\ p_4 &= b_2 - F_{1,2}(u_1, u_2). \end{aligned}$$

Using these facts, and the variances and covariances in the joint asymptotic distribution of $M_1(n)$, $M_2(n)$, $M_3(n)$, $M_4(n)$ given above, a straightforward computation shows that the variances in the asymptotic distribution of $Z_1(n)$ and $Z_2(n)$ are equal to unity, and the covariance is equal to

$$\frac{F_{1,2}(u_1, u_2) - b_1 b_2}{\sqrt{b_1 b_2 (1-b_1)(1-b_2)}}.$$

This completes the proof of the theorem.

3. Extensions

It is clear that the method above also proves the joint asymptotic normality for the case where we include several different quantiles for each X_i. In this case, we use several sets of axes, giving a collection of regions determined by the intersections of the octants defined by the various sets of axes. For the jth such region, we define $N_j(n)$ as the number of observed vectors falling in it, and proceed as above.

4. References

[1] Cramer, Harald, Mathematical methods of statistics (Princeton, Princeton University Press, 1951).

[2] Mood, A. M., On the joint distribution of medians in samples from a multivariate population, Ann. Math. Stat. **12**, 268–278 (1941).

[3] Siddiqui, M. M., Distribution of quantiles in samples from a bivariate population, J. Res. NBS **64B** (Math. and Math. Phys.) No. 3, 145–150. (1960).

(Paper 68B2–118)

[3] Figures in brackets indicate the literature references at the end of this paper.