

A NEW TEST OF COMPOUND SYMMETRY

by

S. N. ROY

INSTITUTE OF STATISTICS, UNIVERSITY OF NORTH CAROLINA

Institute of Statistics
Mimeograph Series No. 97

March 11, 1954

¹ Work sponsored by the Office of Naval Research under Contract No. N7-onr-28402.

A NEW TEST OF COMPOUND SYMMETRY

by

S. N. ROY

Institute of Statistics, University of North Carolina

1. Summary. If x_1 and x_2 have a bivariate normal distribution with a correlation coefficient ρ and the same standard deviations σ for both, then it is well known and easy to check that $x_1 + x_2$ and $x_1 - x_2$ are uncorrelated, which forms the basis of Pitman's well-known test of $H_0: \sigma_1 = \sigma_2$, for a bivariate normal population, in terms of the correlation coefficient r between $x_1 + x_2$ and $x_1 - x_2$ in a random sample of size, say, n , from this population. Starting from this test which has a number of reasonably good properties, and then using the union-intersection principle [1, 2], a test is obtained for compound symmetry, i.e., for $H_0: \sigma_{11} = \sigma_{22} = \dots = \sigma_{pp}$ and all σ_{ij} 's are equal ($i \neq j = 1, 2, \dots, p$), where σ_{ij} is any element of the covariance matrix Σ of a p -variate normal population.

2. Test Construction. Let $\rho(x_1, x_2)$ denote the population correlation coefficient between x_1 and x_2 and $r(x_1, x_2)$ the same in a random sample of size, say n , from that population. Then, for a set of stochastic variables $\underline{x}'(1 \times p) (= x_1, \dots, x_p)$ having a p -variate normal distribution, notice that for any arbitrary non-null vector $\underline{a}'(1 \times p)$, $\underline{a}'(1 \times p) \underline{x}(p \times 1)$

and $\sum_{i=1}^p x_i$ have a bivariate normal distribution and, now letting $\sum_{i=1}^p a_i = 0$,

consider the hypothesis: $\rho(\sum_{i=1}^p x_i, \underline{a}'\underline{x}) = 0 = H_{O\underline{a}}$ (say). Next notice that

(2.1) H_0 : all σ_{ii} 's are equal and all σ_{ij} 's are equal ($i \neq j = 1, \dots, p$)

$$= \bigcap_{\underline{a}} H_{O\underline{a}} = \bigcap_{\underline{a}} \left[\rho \left(\sum_{i=1}^p x_i, \underline{a}'\underline{x} \right) = 0 \right],$$

where $\underline{a}'(1 \times p)$ is any arbitrary row vector subject to $\sum_{i=1}^p a_i = 0$.

Now going back to $H_{O\underline{a}}$, we have for this hypothesis the Pitman critical region, say $W_{\underline{a}}(\alpha)$, of size α , given by

$$(2.2) \quad W_{\underline{a}}(\alpha): \quad r^2 \left(\sum_{i=1}^p x_i, \underline{a}'\underline{x} \right) \geq r_{\alpha}^2(n-2),$$

where $r_{\alpha}(n-2)$ is the upper $\alpha/2$ -point of the central r -distribution in random samples of size n .

Hence, by the union-intersection heuristic principle [1, 2] we have, for $H_0 (= \bigcap_{\underline{a}} H_{O\underline{a}})$, the critical region $W(\beta)$ of size β given by

$$(2.3) \quad W(\beta): \bigcap_{\underline{a}} \left[r^2 \left(\sum_{i=1}^p x_i, \underline{a}'\underline{x} \right) \geq r_{\alpha}^2(n-2) \right],$$

$$\text{i.e., } \sup_{\underline{a}} r^2 \left(\sum_{i=1}^p x_i, \underline{a}'\underline{x} \right) \geq r_{\alpha}^2(n-2)$$

$$\text{i.e., } \sup_{\underline{a}} r^2 \left(\sum_{i=1}^p x_i, \sum_{i=1}^{p-1} a_i (x_i - x_p) \right) \geq r_{\alpha}^2(n-2),$$

remembering that $\sum_{i=1}^p a_i = 0$, i.e., $a_p = -\sum_{i=1}^{p-1} a_i$.

It is easy to check that

$$(2.4) \quad \text{Sup}_{\underline{a}} r^2 \left(\sum_{i=1}^p x_i, \sum_{i=1}^{p-1} a_i (x_i - x_p) \right)$$

= square of the sample multiple correlation between $\sum_{i=1}^p x_i$ and the $(p - 1)$ -

set of variables $(x_1 - x_p), (x_2 - x_p), \dots, (x_{p-1} - x_p) =$

$$R^2 \left[\sum_{i=1}^p x_i \text{ and } (x_1 - x_p), (x_2 - x_p), \dots, (x_{p-1} - x_p) \right] = R^2 \text{ (say).}$$

Notice that, since the new p -set also has the p -variate normal distribution, this R has the well-known multiple correlation distribution with degrees of freedom $p - 1$ and $n - p$ and a non-centrality parameter which is the popula-

tion multiple correlation between $\sum_{i=1}^p x_i$ and the $(p - 1)$ -set above and

which let us call ρ^2 . It is easy to check that $\rho = 0$, i.e., that R has the central multiple correlation distribution, if and only if

$$\rho \left(\sum_{i=1}^p x_i, x_i - x_p \right) = 0 \quad (i = 1, 2, \dots, p - 1), \text{ i.e., if and only if}$$

$$\rho \left(\sum_{i=1}^p x_i, \underline{a}'\underline{x} \right) = 0 \quad (\text{for all non-null } \underline{a} \text{ subject to } \sum_{i=1}^p a_i = 0).$$

We have thus, for testing compound symmetry, the critical region $W(\beta)$ of size β given by

$$(2.5) \quad R \left[\sum_{i=1}^p x_i \text{ and } (x_1 - x_p), \dots, (x_{p-1} - x_p) \right] \geq R_\beta(p - 1, n - p),$$

when R_β is the upper β -point of the well-known central multiple correlation distribution.

It can be checked after some little algebra that in terms, respectively, of the elements of the sample and population covariance matrices S and Σ , R and ρ will be given by

$$(2.6) \quad R^2 = 1 - p \frac{\sum_{i=1}^p z_i z_i^1}{\left(\sum_{i=1}^p z_i\right) \left(\sum_{i=1}^p z_i^1\right)},$$

$$\rho^2 = 1 - p \frac{\sum_{i=1}^p \zeta_i \zeta_i^1}{\left(\sum_{i=1}^p \zeta_i\right) \left(\sum_{i=1}^p \zeta_i^1\right)},$$

when

$$z_i = \sum_{j=1}^p s_{ij}, \quad z_i^1 = \sum_{j=1}^p s^{ij},$$

$$\zeta_i = \sum_{j=1}^p \sigma_{ij} \quad \text{and} \quad \zeta_i^1 = \sum_{j=1}^p \sigma^{ij}.$$

Note that $s_{ij} = s_{ji}$, $s^{ij} = s^{ji}$, $\sigma_{ij} = \sigma_{ji}$ and $\sigma^{ij} = \sigma^{ji}$.

The power properties of this test will be discussed in a later note.

References

1. S. N. Roy, "On a heuristic method of test construction and its use in multivariate analysis," Annals of Mathematical Statistics, Vol. 24 (1953), pp. 220-38.
2. S. N. Roy and R. C. Bose, "Simultaneous confidence interval estimation," Annals of Mathematical Statistics, Vol. 24 (1953), pp. 513-36.