Tilburg University

Comparing means and variances of two simulations

Kleijnen, J.P.C.

Published in:
Simulation: Technical journal of the Society for Computer Simulation

Publication date:
1976

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Kleijnen, J. P. C. (1976). Comparing means and variances of two simulations. *Simulation: Technical journal of the Society for Computer Simulation*, *26*(3), 87-88.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 25. Sep. 2021

TECHNICAL COMMENT

Comparing means and variances of two simulations

by
Jack P. C. Kleijnen
Department of Business and Economics
Katholieke Hogeschool Tilburg
Hogeschoolaan 225
Tilburg, the Netherlands

ABSTRACT

Two systems are simulated using the same random-number stream. To compare their means one can use standard statistical techniques, e.g., the t*-test for paired comparisons. To compare their variances, however, the standard* F*-test fails. Three approaches are suggested. Approach 1 uses the traditional* F*-statistic and is conservative. Approach 2 splits the observations into two groups. For each group the* F*-statistic is exact. It is shown how to combine the two* F*-statistics in a very simple way. This approach is also conservative. Approach 3 was developed in 1946 by Wilks. It is slightly more complicated but exact.*

SIMULATION BACKGROUND OF PROBLEM

We simulate two variants of a system replicating the system a number of times (say n times). Each replication is independent of the other replications of the same variant. However, both variants use the same random-number stream to strengthen the comparison of the variants; see Kleijnen[1] for a detailed discussion of this variance-reduction technique.

STATISTICAL PROBLEM

Let system 1 yield responses $\underline{x}, \ldots, \underline{x}_n$ which are normally and independently distributed with common variance σ_x^2 and mean μ_x (stochastic variables are underlined)

$$\underline{x}_i \sim \text{NID}(\mu_x, \sigma_x^2) \tag{1}$$

In the same way system 2 yields responses $\underline{y}_i$ $(i = 1, \ldots, n)$, i.e.,

$$\underline{y}_i \sim \text{NID}(\mu_y, \sigma_y^2) \tag{2}$$

Since both systems are simulated using the same random-number streams, they are pairwise (positively) correlated:

$$\begin{aligned} \rho = \text{cor}(\underline{x}_i, \underline{y}_{i'}) &> 0 \qquad \text{if } i = i' \\ &= 0 \qquad \text{if } i \neq i' \end{aligned} \tag{3}$$

Question: How to compare the means μ and variances σ^2 of the two systems?

Solution: Comparing the *means* of paired observations is a standard statistical problem. Take the differences

$$\underline{d}_i = \underline{x}_i - \underline{y}_i \; ; \qquad i = 1, \ldots, n \tag{4}$$

and estimate its variance by

$$\underline{s}_d^2 = \sum_{i=1}^{n} (\underline{d}_i - \underline{\bar{d}})^2/(n-1) \tag{5}$$

to construct a confidence interval for $\delta = \mu_x - \mu_y$ using the t-statistic with $v = n-1$ degrees of freedom:

$$\underline{t}_v = \frac{\underline{\bar{d}} - \delta}{\underline{s}_d} \tag{6}$$

There also exist distribution-free tests like the sign test or Wilcoxon's rank test. However, the t-statistic is known to be quite insensitive to non-normality.[1] As far as we are aware, comparing the two *variances* σ_x^2 and σ_y^2 when the observations are pairwise correlated is a problem not discussed in statistical handbooks. We propose two different approaches.

(1) *Continue using the traditional* F*-statistic*, but be aware of its *conservative* character. As we know, the following definitions hold, if $\underline{s}_1^2$ and $\underline{s}_2^1$ are independent:

$$\underline{F}_{v_1, v_2} = \frac{\underline{s}_1^2}{\underline{s}_2^2} \tag{7}$$

where $\underline{s}_1^2$ is based on $v_1 (= n_1 - 1)$ degrees of freedom and $\underline{s}_2^2$ on $v_2 (= n_2 - 1)$ degrees of freedom. Further,

$$\alpha = P(\underline{F} > F^\alpha | H_0) \tag{8}$$

In our case $\underline{s}_1^2$ and $\underline{s}_2^2$ are not independent but positively correlated. Hence $\underline{F}$ is more concentrated around the value 1.

(*Example:* Suppose the random numbers happen to yield $x_1 \simeq x_2 \simeq \ldots \simeq x_n$ so that $s_x^2 \simeq 0$ and $s_x^2 < \sigma_x^2$. Then we expect $y_1 \simeq y_2 \simeq \ldots \simeq y_n$ so that $s_y^2 \simeq 0$ and $s_y^2 < \sigma_y^2$.

Example 2: If $\rho = 1$, then $x_i = y_i$ so $s_x^2 = s_y^2$ and $F = 1$.)

Consequently, denoting the resulting F-statistic by $\underline{\tilde{F}}$, we know that

$$P(\tilde{F} > F^\alpha | H_0) < \alpha \tag{9}$$

where F^α was defined in Equation 8. In other words, if $\underline{\tilde{F}}$ yields a significant value (i.e., exceeds F^α), then we certainly do have to reject the null-hypothesis H_0 since the probability of this event is extremely small (viz. smaller than α). So in our case the F-test is a *conservative* test since it tends not to reject the null-hypothesis (i.e., the power of the test is reduced). Another approach is as follows:

(2) *Split the observations into two groups:* the *first* group of observations on system 1 $(\underline{x}_1, \ldots, \underline{x}_m$ with $m = n/2)$ with the *second* group of observations on system 2 $(\underline{y}_{m+1}, \ldots, \underline{y}_n)$. Such an approach has been applied in other situations in Reference 1, pp. 150-151. Since these observations are independent, we can apply the traditional F-test of Equation 7 where, however, the degrees of freedom are reduced from n-1 to

$$v_1 = v_2 = \frac{n}{2} - 1 \tag{10}$$

Analogously, we can combine $\underline{x}_{m+1}, \ldots, \underline{x}_n$ with $\underline{y}_1, \ldots, \underline{y}_m$. Denote the F-statistic from the first combination by $\underline{F}_1$ and from the second combination by $\underline{F}_2$. What do we conclude if, say, $\underline{F}_1$ is significant and $\underline{F}_2$ is not? This can be solved as follows:

We know that $$P(F_1 \geq F^{\alpha}_{v_1,v_2} | H_0) = \alpha \quad (11)$$

$$P(F_2 \geq F^{\alpha}_{v_1,v_2} | H_0) = \alpha \quad (12)$$

It is easy to prove (Reference 1, p. 532) that the probability that one or both events happen equals

$$P(F_1 > F^{\alpha}_{v_1,v_2} \cup F_2 > F^{\alpha}_{v_1,v_2} | H_0) < \alpha + \alpha \quad (13)$$

Hence if we want to fix "the" probability of an α-error to the value α, then we compare each of the two F-statistics (F_1, F_2) with the F-value tabulated for $\alpha/2$ (and obviously v_1 and v_2 as in Equation 10). We reject the null-hypothesis if F_1 and/or F_2 are significant. This approach is again conservative since Equation 13 shows a $\leq$ sign, not an $=$ sign as in Equations 11 and 12. (Moreover, the power of this approach is decreased since the degrees of freedom are reduced.)

Note that we could devise other splittings. For instance we could combine $x_1, \ldots, x_p$ with $p = n/4$ with $y_{p+1}, \ldots, y_n$, and $x_{p+1}, \ldots, x_n$ with $y_1, \ldots, y_p$. Then F_1 would have as degrees of freedom

$$v_1 = \frac{n}{4} - 1 \quad (14)$$

$$v_2 = \frac{3n}{4} - 1 \quad (15)$$

and F_2

$$v_1 = \frac{3n}{4} - 1 \quad (16)$$

$$v_2 = \frac{n}{4} - 1 \quad (17)$$

Compared with Equation 10 we see that one of the degrees of freedom goes up and the other one goes down. Moreover we can also combine $x_{p+1}, \ldots, x_{2p}$ with $y_1, \ldots, y_p$, $y_{2p+1}, \ldots, y_n$, etc. In this way not two F-statistics result but four of them. Hence following Equation 13, we have to replace α by $\alpha/4$ when comparing the F-statistics with the tabulated values. Replacing α by $\alpha/4$ makes the statistic even more conservative.

(3) *Wilks's test:* As long ago as 1946 Wilks derived an exact procedure. Actually he investigated the more general problem of testing means and/or variances and/or covariances of k-variate normal distributions ($k \geqq 2$). For our problem his statistic is slightly more complicated than in the two approaches above. We refer to Wilks[2] for the definition of the statistic (see his Equation 1.5) and the required table of significance points (see his Table I).

ACKNOWLEDGEMENT
The original problem of this note was brought to my attention by B. Prakken, Free University, Amsterdam.

REFERENCES

1 KLEIJNEN, J.P.C.
Statistical Techniques in Simulation (in 2 parts)
Marcel Dekker, Inc. New York 1974/1975

2 WILKS, S.S.
Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Distribution
Annals of Mathematical Statistics vol. 17
1946 pp. 257-281