Testing Equality of Means of Correlated Variates with Missing Observations on Both Responses

Author(s): Dinesh S. Bhoj

Source: *Biometrika*, Vol. 65, No. 1 (Apr., 1978), pp. 225-228
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: http://www.jstor.org/stable/2335301
Accessed: 18-10-2017 18:51 UTC

Hochberg, Y. (1976). A modification of the T-method of multiple comparisons for a one-way layout with unequal variances. *J. Am. Statist. Assoc.* **71**, 200–4.
Miller, R. G. (1966). *Simultaneous Statistical Inference*. New York: McGraw-Hill.
Mudholkar, G. S., Davidson, M. L. & Subbaiah, P. (1974). Extended linear hypotheses and simultaneous tests in multivariate analysis of variance. *Biometrika* **61**, 467–77.
Pearson, E. S. & Hartley, H. O. (1966). *Biometrika Tables for Statisticians*, Vol. 1, 3rd edition. Cambridge University Press.
Roy, S. N. (1957). *Some Aspects of Multivariate Analysis*. New York: Wiley.

[*Received May* 1977. *Revised July* 1977]

Biometrika (1978), **65**, 1, *pp.* 225–8

Testing equality of means of correlated variates with missing observations on both responses

By DINESH S. BHOJ

Department of Statistics, Rutgers University, Camden, New Jersey

Summary

A statistic is proposed for testing the hypothesis of equality of the means of a bivariate normal distribution when observations are missing on both variates. The distribution of the statistic is approximated by a Student's t distribution under the null hypothesis. The expected squared lengths of the confidence intervals for the mean difference are used to measure the increased precision of our statistic relative to the conventional paired t test.

Some key words: Confidence interval comparison; Missing observations; Paired t test; Test for equality of means.

1. Introduction

Let $(x, y)'$ be a bivariate vector normally distributed with mean vector $(\mu_1, \mu_2)'$ and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Let $(x_i, y_i)'$ for $i = 1, \ldots, n$ be n pairs of observations on $(x, y)'$. In addition, there are n_1 observations on x only, and n_2 observations on y only. Without loss of generality, the data may be arranged as follows:

$$\begin{array}{lll} x_1, \ldots, x_n; & x_{n+1}, \ldots, x_{n+n_1}; & \\ y_1, \ldots, y_n; & & y_{n+1}, \ldots, y_{n+n_2}, \end{array} \qquad (1{\cdot}1)$$

where x_{n+j} for $j = 1, \ldots, n_1$ and y_{n+k} for $k = 1, \ldots, n_2$ are unpaired observations. It is assumed that (x_i, y_i), x_{n+j} and y_{n+k} are mutually independent for $i = 1, \ldots, n$, $j = 1, \ldots, n_1$ and $k = 1, \ldots, n_2$. We wish to use all available data to test the hypothesis H_0: $\mu_1 = \mu_2$ against the alternative of unequal means. Lin & Stivers (1974) proposed statistics for testing the equality of means when the incomplete data has pattern (1·1), and they carried out empirical investigations of the size and power of some of the tests. Ekbohm (1976) suggested two statistics and gave the results of some simulation studies for the above incomplete data pattern. In this paper we propose a test statistic which is a linear combination of the paired and unpaired t statistics.

2. Testing H_0: $\mu_1 = \mu_2$ with $\sigma_1^2 = \sigma_2^2$

2·1. *The test statistic*

Our statistic for testing H_0: $\mu_1 = \mu_2$ when $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is

$$T = \lambda \frac{\bar{x}-\bar{y}}{s/\sqrt{n}} + (1-\lambda)\frac{\bar{x}_1-\bar{y}_1}{s_1\sqrt{\{(1/n_1)+(1/n_2)\}}}, \tag{2·1}$$

where

$$n\bar{x} = \sum_{i=1}^{n} x_i, \quad n\bar{y} = \sum_{i=1}^{n} y_i, \quad n_1\bar{x}_1 = \sum_{j=1}^{n_1} x_{n+j}, \quad n_2\bar{y}_1 = \sum_{k=1}^{n_2} y_{n+k},$$

$$(n-1)s^2 = \sum_{i=1}^{n}(x_i - y_i - \bar{x} + \bar{y})^2, \quad (n_1+n_2-2)s_1^2 = \sum_{j=1}^{n_1}(x_{n+j}-\bar{x}_1)^2 + \sum_{k=1}^{n_2}(y_{n+k}-\bar{y}_1)^2,$$

and $\lambda\,(0 \leqslant \lambda \leqslant 1)$ is a constant at our choice. Without loss of generality, we assume that $\sigma^2 = 1$. The rationale behind using a statistic T is that it enables the investigator to use the missing observations in a simple manner, and T reduces to paired t when $\lambda = 1$. Under H_0, T is the weighted sum of two independent random variables which are distributed according to Student's t distribution with $f_1 = n-1$ and $f_2 = n_1+n_2-2$ degrees of freedom. By following Patil (1965), the null distribution of hT is approximated adequately by a t distribution with f degrees of freedom, where the constant h and degrees of freedom f are determined by equating the second and fourth moments of hT and t. More accurate percentage points of T can be obtained by using Ghosh's (1975) approximation.

2·2 *Expected squared confidence interval length*

The $(1-\alpha)$ confidence interval on the mean difference can be obtained by replacing $\bar{x}-\bar{y}$ and $\bar{x}_1-\bar{y}_1$ in (2·1) by $(\bar{x}-\bar{y})-(\mu_1-\mu_2)$ and $(\bar{x}_1-\bar{y}_1)-(\mu_1-\mu_2)$ respectively. Computation of expected squared length of the confidence interval for $0<\lambda<1$ involves a double integral. To evaluate that integral we use the transformation

$$(1-\lambda)u = r\cos^2\theta \quad (0 \leqslant r \leqslant \infty), \quad \lambda v = r\sin^2\theta \quad (0 \leqslant \theta \leqslant \tfrac{1}{2}\pi),$$

and integrate out r. Then the single integral is written in a convenient form for numerical computation by substituting $z = \sin^2\theta$. We finally have the following expression for the expected squared length of the confidence interval:

$$E(L_T^2) = \frac{8t_{\frac{1}{2}\alpha;f}^2(f_1+f_2)\,d^{f_2}}{h^2 b^2 \lambda^2 B(\frac{1}{2}f_1, \frac{1}{2}f_2)} \int_0^1 z^{f_2+1}(1-z)^{f_1+1}\{(1-z)^2 + d^2z^2\}^{-\frac{1}{2}(f_1+f_2+2)}\,dz,$$

where $b^2 = \frac{1}{2}n(n-1)/(1-\rho)$, $c^2 = n_1n_2(n_1+n_2-2)/(n_1+n_2)$, $d = (1-\lambda)c/(\lambda b)$, $t_{\frac{1}{2}\alpha;f}$ is the upper $\frac{1}{2}\alpha$ critical value of t with f degrees of freedom, $B(\frac{1}{2}f_1, \frac{1}{2}f_2)$ is the complete beta function and f and h are obtained from Patil (1965).

The expected squared lengths of the confidence intervals are used to measure the sensitivity of the T test over that of the usual t test based on n complete pairs of observations. Some values of the ratio $E(L_T^2)/E(L_t^2)$ for $\rho>0$ are given in Table 1, where $E(L_t^2)$ is the mean squared length of the paired t interval. The ratios are given for two values of λ: (i) $\lambda = 1-\lambda = \frac{1}{2}$, and (ii) λ is chosen such that the average value of $E(L_T^2)/E(L_t^2)$ over the various values of ρ used in Table 1 is a minimum. We note that when $\lambda = 1-\lambda$, both λ and $1-\lambda$ can be replaced by 1, and T becomes the unweighted sum of two independent random variables. As we might expect, Table 1 shows that for given n_1+n_2, the reduction in the ratios is larger for $n_1 = n_2$ than for $n_1 \neq n_2$ for all values of ρ. When both n_1 and n_2 are moderately large the gain in precision from the chosen value of λ is slight over that of $\lambda = \frac{1}{2}$. The precision of the missing value procedure can be considerably increased for large positive values of ρ by a suitable value of λ when n_1 and n_2 are both small compared to n or one of them is small and the other is moderately large. When ρ is negative or positive but small, the unweighted T is quite satisfactory even though a further reduction in the ratios is generally possible by choosing an appropriate value of λ.

Table 1. *Approximate ratios of expected squared lengths of the incomplete data and paired t 95% confidence intervals for positive* ρ

n	n_1	n_2	λ	$\rho = 0.10$	$\rho = 0.30$	$\rho = 0.50$	$\rho = 0.70$	$\rho = 0.90$
10	5	5	0·50	0·67	0·74	0·83	0·98	1·27
			0·68	0·66	0·70	0·76	0·83	0·97
10	5	15	0·50	0·52	0·59	0·67	0·81	1·09
			0·60	0·54	0·59	0·66	0·76	0·95
10	10	10	0·50	0·46	0·52	0·60	0·73	1·02
			0·58	0·48	0·53	0·60	0·70	0·92
10	10	20	0·50	0·39	0·44	0·52	0·65	0·93
			0·54	0·40	0·45	0·52	0·64	0·89
10	30	30	0·50	0·24	0·28	0·34	0·45	0·71
			0·46	0·23	0·27	0·33	0·44	0·73

3. Testing $H_0\colon \mu_1 = \mu_2$ with $\sigma_1^2 \neq \sigma_2^2$

Without loss of generality we assume that $n_1 \leqslant n_2$ and $\sigma_1^2 = 1$. The statistic we propose for testing $H_0\colon \mu_1 = \mu_2$ is

$$T' = \lambda \frac{\bar{x}-\bar{y}}{s/\sqrt{n}} + (1-\lambda)\frac{\bar{x}_1 - \bar{y}_1}{s_2/\sqrt{n_1}}, \tag{3·1}$$

where

$$(n_1-1)\,s_2^2 = \sum_{j=1}^{n_1} (w_j - \bar{w}_1)^2, \quad w_j = x_{n+j} - (n_1/n_2)^{\frac{1}{2}}\, y_{n+j}, \quad n_1 \bar{w}_1 = \sum_{j=1}^{n_1} w_j$$

and the other notation in (3·1) is the same as in (2·1). The second term in (3·1) without the multiplying constant $(1-\lambda)$ is Scheffé's (1943) statistic for testing the equality of means of uncorrelated variables and is distributed under H_0 as t with $n_1 - 1$ degrees of freedom. Computation of $E(L_{T'}^2)$ follows in exactly the same way as for $E(L_T^2)$. Some values of the ratio $E(L_{T'}^2)/E(L_t^2)$ for 95% confidence intervals are given in Table 2 for two values of λ as in § 2, $R = \sigma_2/\sigma_1$ and $\rho > 0$. We note that for moderately large values of n_1 and n_2 compared to n, the use of unweighted T' gives considerable gain over the paired t test. For $R > 1$, the ratio

Table 2. *Approximate ratios of expected squared lengths of the incomplete data and paired t 95% confidence intervals for* $\sigma_1 \neq \sigma_2$

n	n_1	n_2	R	λ	$\rho = 0.10$	$\rho = 0.30$	$\rho = 0.50$	$\rho = 0.70$	$\rho = 0.90$
10	6	10	0·6	0·50	0·65	0·71	0·79	0·91	1·11
				0·67	0·62	0·66	0·71	0·78	0·88
			4·0	0·50	0·54	0·56	0·60	0·63	0·68
				0·60	0·53	0·55	0·57	0·60	0·63
10	10	10	0·6	0·50	0·48	0·53	0·60	0·70	0·87
				0·58	0·49	0·53	0·59	0·67	0·80
			4·0	0·50	0·47	0·49	0·52	0·55	0·59
				0·53	0·47	0·49	0·52	0·55	0·58
10	10	30	0·6	0·50	0·43	0·48	0·55	0·68	0·82
				0·56	0·44	0·49	0·55	0·63	0·77
			4·0	0·50	0·27	0·29	0·31	0·33	0·36
				0·43	0·26	0·28	0·30	0·33	0·36

decreases as R increases, and for given n_1+n_2 the gain in precision is higher for $n_1 < n_2$ than for $n_1 = n_2$ when R is large, which, of course, is expected since more observations are available on the variable with larger variance.

References

Ekbohm, G. (1976). Comparing means in the paired case with incomplete data on both responses. *Biometrika* **63**, 299–304.

Ghosh, B. K. (1975). On the distribution of the difference of two t-variables. *J. Am. Statist. Assoc.* **70**, 463–7.

Lin, P. E. & Stivers, L. E. (1974). On the difference of means with incomplete data. *Biometrika* **61**, 325–34.

Patil, V. H. (1965). Approximation to the Behrens–Fisher distributions. *Biometrika* **52**, 267–71.

Scheffé, H. (1943). On solutions of the Behrens–Fisher problems, based on the t-distribution. *Ann. Math. Statist.* **14**, 35–44.

[*Received March* 1977. *Revised July* 1977]

Biometrika (1978), **65**, 1, *pp*. 228–31

Testing of hypotheses with trinomials generated by gauges

By TAJ HIRJI

Group Research & Development Centre, British–American Tobacco Co. Ltd, Southampton

AND

A. K. SHAHANI

Department of Mathematics, University of Southampton

Summary

Linear functions of trinomial frequencies, generated by gauges, provide simple statistics for testing hypotheses. A computationally easy method for determining the necessary parameters and coefficients, which is nearly optimal in the sense of maximizing power is suggested. Guidelines for the use of a normal approximation are given. Tests for the mean of a normal variate are discussed.

Some key words: Gauging; Hypothesis testing; Trinomial; Truncation.

1. Introduction

Over the past 30 or so years there has been an increasing interest in the use of inefficient statistics (Mosteller, 1946). Linear functions of order statistics have received special attention; see for example Lloyd (1952) and David (1970). Two of the attractive features of these inefficient statistics are simplicity (Quenouille, 1972) and robustness (Huber, 1972).

Stevens (1948) pointed out that we can have a wide variety of gauges. In statistical terms, a gauge set at a given level a determines whether or not the event $X \leqslant a$ occurs, X being a random variable. Gauges may be unidimensional or multidimensional; for example, a two-dimensional gauge will determine whether the event $(X \leqslant x_0, Y \leqslant y_0)$ has occurred. Thus a gauge creates a discrete classification usually, but not necessarily, by the division of a continuous scale. We consider gauges that partition the possible domain of a random vector X into three disjoint regions, A_1, A_2, A_3, so that $\mathrm{pr}\,(X \in A_1)+\mathrm{pr}\,(X \in A_2)+\mathrm{pr}\,(X \in A_3) = 1$. Let n independent observations be made, using the gauges, for testing an hypothesis, say H_0,