13

On the Sensitivity of the *t*-Statistic

JAN R. MAGNUS Tilburg University, Tilburg, The Netherlands

1. INTRODUCTION

There seems to be consensus among statisticians and econometricians that the t-statistic (unlike the F-statistic) is not sensitive (robust). If we define the t-statistic as $\tau = a'x/\sqrt{x'Bx}$, where x is a random vector with mean 0, a is a nonrandom vector, and B a positive semidefinite nonrandom matrix, then τ will in general not follow a Student distribution for three reasons. First, x may not be normally distributed; secondly, even if x is normally distributed, x may not be χ^2-distributed, and thirdly, the numerator and denominator may be dependent. The consensus is that nevertheless the Student distribution can be used to provide a good approximation of the distribution of τ. The purpose of this paper is to analyze some aspects of this situation.

We confine ourselves to the situation where the random vector x is normally distributed; some aspects of the non-normal case are discussed in Ullah and Srivastava (1994). Then all odd moments of τ which exist are 0. This chapter therefore concentrates on the even moments of τ and, in particular, on its variance and kurtosis. The special case where $x'Bx$ follows a χ^2-distribution, but where $a'x$ and $x'Bx$ are dependent, was considered by Smith (1992). He obtains the density for this special case, and, from the density, the moments, but not in an easy-to-use form. He concludes that the t-statistic is robust in most situations. Morimune (1989), in a simultaneous equation context, also concludes that the t-statistic is robust. Maekawa (1980) considered the t-statistic in the seemingly unrelated regression model and derived an Edgeworth expansion up to $\mathcal{O}(n^{-1})$, where n is the sample size.

We shall argue that, although it looks as if the t-statistic is robust because the moments of τ are close to moments of the Student distribution, in fact the conclusion is wrong. The reason for this apparent contradiction is the following. Clearly, the Student distribution is close to the standard normal distribution. Also, the t-statistic τ, properly scaled, is well approximated by the standard normal distribution. For this reason τ is close to a Student distribution since both are close to the standard normal distribution, and in this sense the t-statistic is robust. But in many cases, as we shall see, τ is better approximated by the standard normal distribution than by the appropriate Student distribution, and in this sense the t-statistic is not robust.

The paper is organized as follows. In Section 2 we define the problem and settle the notation. Theorem 1 in Section 3 gives a simple expression for the even moments of τ. This theorem is valid irrespective of whether $x'Bx$ follows a χ^2-distribution or whether $a'x$ and $x'Bx$ are independent. Theorem 2 gives precise conditions when these moments exist. Theorem 1 is a new result and has potential applications in many other situations. In Section 4 we obtain Theorem 3 as a special case of Theorem 1 by assuming that $x'Bx$, properly scaled, follows a χ^2-distribution. The even moments of τ then become extremely simple functions of one "dependence parameter" δ. We analyze the variance and kurtosis for this case and conclude that, if we use the standard normal distribution to approximate the Student distribution, then the approximation will be *better* with dependence than without. All proofs are in the Appendix.

2. SET-UP AND NOTATION

Let x be a normally distributed $n \times 1$ vector with mean 0 and positive definite covariance matrix $\Omega = LL'$. Let a be an $n \times 1$ vector and B a positive semidefinite $n \times n$ matrix with rank $r \geq 1$. We define the "t-type" random variable

$$\tau = \frac{a'x}{\sqrt{x'Bx}}. \tag{1}$$

In order to normalize B we introduce the matrix $B^* = (1/\text{tr}B\Omega)L'BL$, which satisfies $\text{tr}B^* = 1$. We denote by $\lambda_1, \ldots, \lambda_r$ the positive eigenvalues of B^* and by $\sigma_1, \ldots, \sigma_r$ the corresponding normalized eigenvectors, so that $B^*\sigma_i = \lambda_i\sigma_i$, $\sigma_i'\sigma_i = 1$, $\sigma_i'\sigma_j = 0\,(i \neq j)$.

We normalize the vector a by defining

$$\alpha_i = \frac{\sigma_i' L' a}{\sqrt{a'\Omega a}} \qquad (i = 1, \ldots, r). \tag{2}$$

An important role is played by the scalar

$$\delta = \sum_{i=1}^{r} \alpha_i^2. \tag{3}$$

It is easy to see that $0 \leq \delta \leq 1$. If $\delta = 0$, then $a'x$ and $x'Bx$ are independent and $L'a$ and $L'BL$ are orthogonal; if $\delta > 0$, they are dependent. If $\delta = 1$, then $L'a$ lies in the column space of $L'BL$ (or equivalently, a lies in the column space of B).

3. THE MOMENTS OF τ

All odd moments of τ which exist are 0. As for the even moments that exist, we notice that $\tau^2 = x'Ax/x'Bx$ with $A = aa'$. The exact moments of a ratio of quadratic forms in normal variables was obtained by Magnus (1986), while Smith (1989) obtained moments of a ratio of quadratic forms using zonal polynomials and invariant polynomials with multiple matrix arguments. In the special case where A is positive semidefinite of rank 1 and where the mean of x is 0, drastic simplifications occur and we obtain Theorem 1.

Theorem 1. We have, provided the expectation exists, for $s = 1, 2, \ldots,$

$$E\tau^{2s} = \left(\frac{a'\Omega a}{\text{tr}B\Omega}\right)^s \frac{\kappa_s}{(s-1)!} \int_0^\infty t^{s-1} \prod_{i=1}^{r} (1 - \mu_i(t))^{1/2} \left(1 - \sum_{i=1}^{r} \mu_i(t)\alpha_i^2\right)^s dt$$

where

$$\mu_i(t) = \frac{2t\lambda_i}{1 + 2t\lambda_i} \quad \text{and} \quad \kappa_s = 1 \times 3 \times \cdots \times (2s - 1).$$

From Theorem 1 we find the variance $\text{var}(\tau)$ and the kurtosis $\text{kur}(\tau) \equiv E\tau^4/(E\tau^2)^2$ as

$$\mathrm{var}(\tau) = \frac{a'\Omega a}{\mathrm{tr}B\Omega}\int_0^\infty f(t)g(t)dt \qquad \mathrm{kur}(\tau) = 3 \times \frac{\int tf(t)g(t)^2 dt}{(\int f(t)g(t)dt)^2} \tag{4}$$

where

$$f(t) = \prod_{i=1}^{r}(1-\mu_i(t))^{1/2}, \qquad g(t) = 1 - \sum_{i=1}^{r}\mu_i(t)\alpha_i^2. \tag{5}$$

Turning now to the existence of the even moments, we notice that for Student's t-distribution Et^{2s} is defined when $r > 2s$. We find the same condition for the moments of τ, except that when $\delta = 1$ all moments of τ exist.

Theorem 2 (existence):

(i) If $r \leq n-1$ and $\delta = 1$, or if $r = n$, then $E\tau^{2s}$ exists for all $s \geq 0$;
(ii) If $r \leq n-1$ and $\delta \neq 1$, then $E\tau^{2s}$ exists for $0 \leq s < r/2$ and does not exist for $s \geq r/2$.

The variance and kurtosis given in (4) can be evaluated for any given $\lambda_1, \ldots, \lambda_r$ and $\alpha_1, \ldots, \alpha_r$. To gain insight into the sensitivity of the t-statistic, we consider one important special case, namely the case where $x'Bx$, properly scaled, follows a $\chi^2(r)$ distribution, but where $a'x$ and $x'Bx$ are dependent.

4. SENSITIVITY FOR DEPENDENCE

When $x'Bx$, properly scaled, follows a $\chi^2(r)$ distribution, the only difference between τ and Student's t is that the numerator and denominator in τ are dependent, unless $\delta = 0$. In this section we investigate the impact of this dependence on the moments of τ. We first state Theorem 3, which is a special case of Theorem 1.

Theorem 3. *If the nonzero eigenvalues of $L'BL$ are all equal, then we have, provided the expectation exists, for $s = 1, 2, \ldots,$*

$$E\tau^{2s} = \left(\frac{a'\Omega a}{\mathrm{tr}B\Omega}\right)^s \frac{\kappa_s}{(s-1)!}\left(\frac{r}{2}\right)^s \sum_{j=0}^{s}\binom{s}{j}(-\delta)^j B\left(s+j, \frac{r}{2}-s\right),$$

where $B(\cdot,\cdot)$ denotes the Beta function.

It is remarkable that the even moments of τ now depend only on s, r and the "dependence parameter" δ (apart from a scaling factor $a'\Omega a/\mathrm{tr}B\Omega$). Evaluation of $E\tau^{2s}$ is very easy since no integral needs to be calculated. Under the conditions of Theorem 3 we obtain

$$\operatorname{var}(\tau) = \frac{r}{r-2}\left(1 - \frac{2\delta}{r}\right) \tag{6}$$

and

$$\operatorname{kur}(\tau) = 3 \times \frac{r-2}{r-4} \times \frac{r^2(r+2) - 8\delta r(r+2) + 24\delta^2 r}{r^2(r+2) - 4\delta r(r+2) + 4\delta^2(r+2)}. \tag{7}$$

For fixed r, both $\operatorname{var}(\tau)$ and $\operatorname{kur}(\tau)$ are monotonically decreasing functions on the $[0, 1]$ interval, and we find

$$1 \leq \operatorname{var}(\tau) \leq \frac{r}{r-2} \tag{8}$$

and

$$3\frac{r}{r+2} \leq \operatorname{kur}(\tau) \leq 3\frac{r-2}{r-4}. \tag{9}$$

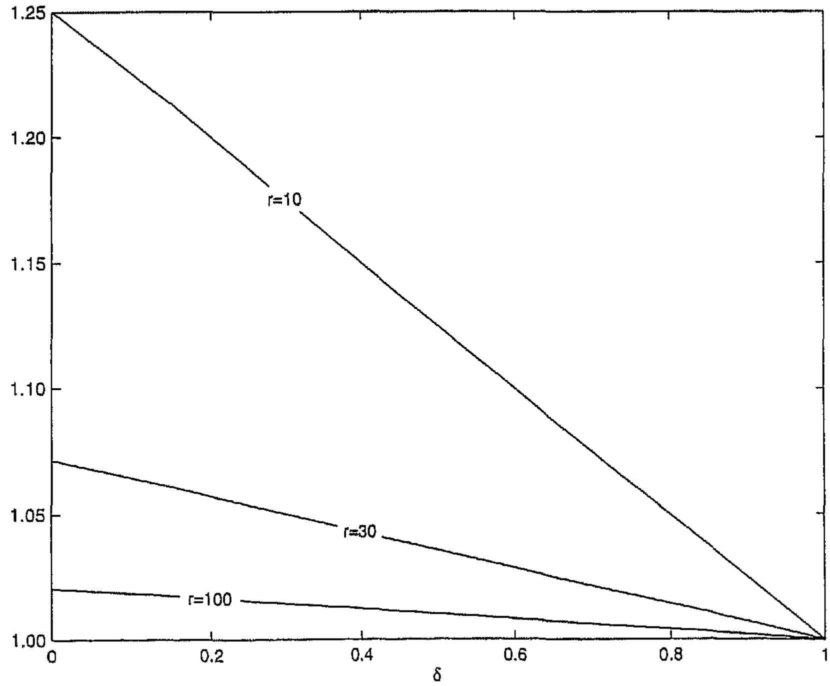

Figure 1. Variance of τ as a function of δ.

Figures 1 and 2 illustrate the behaviour of $\text{var}(\tau)$ and $\text{kur}(\tau)$ for $r = 10$, $r = 30$, and $r = 100$ (and, of course, $r = \infty$).

The variance of τ is linear in δ. For $\delta = 0$ we find $\text{var}(\tau) = r/(r-2)$, the variance of the $t(r)$-distribution. When $\delta > 0$, $\text{var}(\tau)$ decreases monotonically to 1. Hence, the more dependence there is (the higher δ is), the *better* is $\text{var}(\tau)$ approximated by the variance of the standard normal distribution.

The kurtosis of τ is *not* linear in δ, see (7), but in fact is almost linear, slightly curved, and convex on $[0, 1]$; see Figure 2. When $\delta = 0$ (independence) we find $\text{kur}(\tau) = 3(r-2)/(r-4)$, the kurtosis of the $t(r)$-distribution. When $\delta > 0$, the kurtosis decreases, becomes 3 (at some δ between 0.500 and 0.815), and decreases further to $3r/(r+2)$. The deviation of the kurtosis from normality is *largest* in the case of independence ($\delta = 0$).

In conclusion, if we use the standard normal distribution to approximate the t-distribution, then the approximation will be *better* with dependence than without. The t-statistic τ is thus better approximated by the standard normal distribution than by the appropriate Student distribution. In this sense the t-statistic is *not* robust.

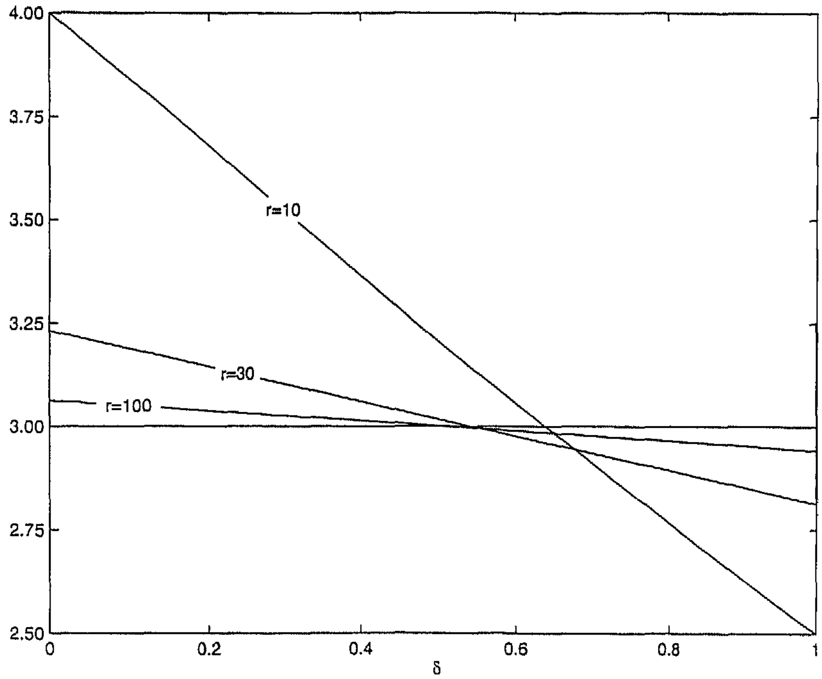

Figure 2. Kurtosis of τ as a function of δ.

APPENDIX: PROOFS

Proof of Theorem 1. Since $\tau^2 = x'Ax/x'Bx$ with $A = aa'$, the even moments of τ are the moments of a ratio of two quadratic forms. Using Theorem 6 of Magnus (1986), let P be an orthogonal $n \times n$ matrix and Λ^* a diagonal $n \times n$ matrix such that

$$P'L'BLP = \Lambda^*, \qquad P'P = I_n.$$

Then,

$$E\tau^{2s} = \frac{1}{(s-1)!} \sum_{\nu} \gamma_s(\nu) \int_0^{\infty} t^{s-1} |\Delta| \prod_{j=1}^{s} \left(\mathrm{tr} R^j\right)^{n_j} dt$$

where the summation is over all $1 \times s$ vectors $\nu = (n_1, n_2, \ldots, n_s)$ whose elements n_j are non-negative integers satisfying $\sum_{j=1}^{s} jn_j = s$,

$$\gamma_s(\nu) = s!\, 2^s \prod_{j=1}^{s} \left(n_j! (2j)^{n_j}\right)^{-1},$$

and

$$\Delta = (I_n + 2t\Lambda^*)^{-1/2}, \qquad R = \Delta P'L'aa'LP\Delta.$$

Now, R has rank 1, so that

$$\left(\mathrm{tr} R^j\right)^{n_j} = \left(a'LP\Delta^2 P'L'a\right)^{jn_j}$$

and hence, since $\sum_j jn_j = s$,

$$\prod_{j=1}^{s} \left(\mathrm{tr} R^j\right)^{n_j} = \left(a'LP\Delta^2 P'L'a\right)^{s}.$$

Also, using Lemma 3 of Magnus (1986) for the special case $n = 1$, we see that $\sum_{\nu} \gamma_s(\nu) = \kappa_s$. Hence,

$$E\tau^{2s} = \frac{\kappa_s}{(s-1)!} \int_0^{\infty} t^{s-1} |\Delta| \left(a'LP\Delta^2 P'L'a\right)^{s} dt.$$

Letting $\lambda_1^*, \ldots, \lambda_r^*$ denote the nonzero diagonal elements of Λ^*, we see that $\lambda_i^* = (\mathrm{tr} B\Omega)\lambda_i$. Letting $t^* = (\mathrm{tr} B\Omega)t$, we thus obtain

$$|\Delta| = \prod_{i=1}^{r} (1 + 2t\lambda_i^*)^{-1/2} = \prod_{i=1}^{r} (1 - \mu_i(t^*))^{1/2}$$

and

$$a'LP\Delta^2P'L'a = (a'\Omega a)\left(1 - \sum_{i=1}^{r}\mu_i(t^*)\alpha_i^2\right).$$

Proof of Theorem 2. Let Q be an $n \times (n-r)$ matrix of full column rank $n-r$ such that $BQ = 0$. Then, using Magnus (1986, Theorem 7) or Magnus (1990, Theorem 1) and noticing that $Q'a = 0$ if and only if $\delta = 1$, the result follows.

Proof of Theorem 3. Let $\lambda_i = \lambda = 1/r$ and $\mu = 2t\lambda/(1+2t\lambda)$. Theorem 1 implies that

$$E\tau^{2s} = \left(\frac{a'\Omega a}{\mathrm{tr}B\Omega}\right)^s \frac{\kappa_s}{(s-1)!}\int_0^\infty t^{s-1}(1-\mu)^{r/2}\left(1-\mu\sum_{i=1}^{r}\alpha_i^2\right)^s dt.$$

By making the change of variable

$$t = \frac{1}{2\lambda}\cdot\frac{\mu}{1-\mu}$$

we obtain

$$E\tau^{2s} = \left(\frac{a'\Omega a}{\mathrm{tr}B\Omega}\right)^s \frac{\kappa_s}{(s-1)!}\left(\frac{r}{2}\right)^s\int_0^1 \mu^{s-1}(1-\mu)^{r/2-s-1}(1-\mu\delta)^s d\mu.$$

Now,

$$\begin{aligned}\int_0^1 & \mu^{s-1}(1-\mu)^{r/2-s-1}(1-\mu\delta)^s d\mu \\ &= \sum_{j=0}^{s}\binom{s}{j}(-\delta)^j\int_0^1 \mu^{s+j-1}(1-\mu)^{r/2-s-1}d\mu \\ &= \sum_{j=0}^{s}\binom{s}{j}(-\delta)^j B(s+j, r/2-s)\end{aligned}$$

and the result follows.

REFERENCES

K. Maekawa. An asymptotic expansion of the distribution of the test statistics for linear restrictions in Zellner's SUR model. *Hiroshima Economic Review* 4:81–97, 1980.

J. R. Magnus. The exact moments of a ratio of quadratic forms in normal variables. *Annales d'Economie et de Statistique* 4:95–109, 1986.

J. R. Magnus. On certain moments relating to ratios of quadratic forms in normal variables: further results. *Sankhya, Series B* 52:1–13, 1990.

K. Morimune. *t* test in a structural equation. *Econometrica* 57:1341–1360, 1989.

M. D. Smith. On the expectation of a ratio of quadratic forms in normal variables. *Journal of Multivariate Analysis* 31:244–257, 1989.

M. D. Smith. Comparing the exact distribution of the *t*-statistic to Student's distribution when its constituent normal and chi-squared variables are dependent. *Communications in Statistics, Theory and Methods* 21(12):3589–3600, 1992.

A. Ullah, V. K. Srivastava. Moments of the ratio of quadratic forms in non-normal variables with econometric examples. *Journal of Econometrics* 62:129–141, 1994.