A RELATION BETWEEN t AND F-DISTRIBUTIONS

by
Theophilos Cacoullos

Technical Report #39

TR39

A RELATION BETWEEN t AND F-DISTRIBUTIONS

by Theophilos Cacoullos*

University of Minnesota

1. Let $X(\alpha)$ denote the upper α point of the distribution of a random variable X. It is well known that $t^2_n(\alpha/2) = F_{1,n}(\alpha)$, where t_n has the t-distribution with n degrees of freedom (d.f.) and $F_{m,n}$ has the F-distribution with m and n d.f. The purpose of this note is to point out a new relation between the percentage points of a t_n-distribution and those of an $F_{m,n}$ - distribution in the special case when m = n. Indeed, $F_{n,n}(\alpha)$ will be expressed as a function of $t_n(\alpha)$ and vice versa. It follows that the $F_{n,n}(\alpha)$ points in an F-table may be easily computed from the $F_{1,n}(\alpha)$ and conversely. It is therefore expected that these results, besides their theoretical interest, might be of some practical significance.

The relations between $F_{n,n}(\alpha)$ and $t_n(\alpha)$ given below in (2) and (5) are based on the following

Lemma

Let X and Y be two independent χ^2-distributed random variables each with n d.f. Then

$$Z = \frac{\sqrt{n}}{2} \frac{X-Y}{\sqrt{XY}} \qquad (1)$$

has the t_n-distribution.

Proof

Let $$G(u) = P[\frac{X-Y}{2\sqrt{XY}} \leq u] \quad .$$

* This work was supported in part by the National Science Foundation under Grant Number C-19126.

We have

$$G(u) = \frac{1}{2^n \Gamma^2(\frac{n}{2})} \iint_{(x-y) \leq 2\sqrt{xy}\, u} (xy)^{\frac{n}{2}-1} e^{-\frac{1}{2}(x+y)} dxdy \quad .$$

Setting first

$$x = \rho \cos^2\theta, \quad y = \rho \sin^2\theta,$$

and then

$$w = \cot 2\theta,$$

we readily get

$$G(u) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\,\Gamma(\frac{n}{2})} \int_{-\infty}^{u} \frac{dw}{(1+w^2)^{\frac{n+1}{2}}} \quad ,$$

from which the lemma follows.

Noting that Z can be written as

$$\frac{\sqrt{n}}{2}\left(\sqrt{F_{n,n}} - \frac{1}{\sqrt{F_{n,n}}}\right) \quad ,$$

which is an increasing function of $F_{n,n}$, we obtain, as a corollary of the lemma, the relation

$$t_n(\alpha) = \frac{\sqrt{n}}{2}\left(\sqrt{F_{n,n}(\alpha)} - \frac{1}{\sqrt{F_{n,n}(\alpha)}}\right) \quad . \tag{2}$$

In order to express $F_{n,n}(\alpha)$ as a function of $t_n(\alpha)$, we observe that $t_n = Z$, as a function of $F_{n,n}$ becomes equal to zero when $F_{n,n} = 1$. Therefore, for $0 < \alpha < 0.5$, we have

$$P[0 < t_n < t_n(\alpha)] = P[1 < F_{n,n} < C_\alpha] \tag{3}$$

where C_α is the root $F > 1$ of the equation

$$t_n(\alpha) = \frac{\sqrt{n}}{2}\left(\sqrt{F} - \frac{1}{\sqrt{F}}\right) ;$$

$$C_\alpha = 1 + \frac{2t_n^2(\alpha)}{n} + \frac{2t_n(\alpha)}{\sqrt{n}}\sqrt{1 + \frac{t_n^2(\alpha)}{n}} \quad . \tag{4}$$

By (3) and the fact that

$$P[F_{n,n} \leq 1] = \tfrac{1}{2} ,$$

we get

$$P[F_{n,n} \leq C_\alpha] = 1-\alpha ,$$

from which

$$F_{n,n}(\alpha) = C_\alpha . \tag{5}$$

This in conjuntion with (4), and (2) establish the one-to-one correspondence between $t_n(\alpha)$ and $F_{n,n}(\alpha)$ for $0 < \alpha < 0.5$. For $0.5 < \alpha < 1$ (corresponding to lower $1-\alpha$ points of the distributions) it is well known that

$$t_n(\alpha) = - t_n(1-\alpha) \quad \text{and} \quad F_{n,n}(\alpha) = F^{-1}_{n,n}(1-\alpha) .$$

2. Two applications of the lemma:

Let (x_i, y_i), $i = 1,\ldots,N$ be a sample from a bivariate normal distribution with mean (μ,ν) and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix} .$$

De Lury (1938) found the distribution of the sample correlation coefficient r' defined in this special case of equal variances by

$$r' = \frac{2\sum_{i=1}^{N} (x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N} [(x_i-\bar{x})^2 + (y_i-\bar{y})^2]} , \quad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i , \ \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i .$$

Using the distribution of r' it can be easily verified that

$$t = \frac{\sqrt{N-1}\,(r'-\rho)}{\sqrt{1-\rho^2}\sqrt{1-r'^2}} \tag{6}$$

has a t_{N-1}-distribution. A direct proof of this led to the expression of t in the form of Z in (1) with $n = N - 1$, and actually motivated the present investigation.

The reduction of t to Z is based on the following easily verified facts:

a) r' is distributed as

$$r' = \frac{\sum_{i=1}^{N-1} u_i v_i}{\sum_{i=1}^{N-1} (u_i^2 + v_i^2)} ,$$

where the u_i and v_i are jointly normally distributed with means zero and variances σ^2 and they are uncorrelated except for the pairs (u_i, v_i) for which $\mathrm{Cov}(u_i, v_i) = \rho\sigma^2$;

b) by a) also, the

$$u_i^* = \frac{u_i + v_i}{\sqrt{2\sigma(1+\rho)}} , \quad u_i^* = \frac{u_i - v_i}{\sqrt{2\sigma(1-\rho)}} \qquad i = 1, \ldots, N - 1$$

are completely independent standard normal variables;

c) t in (6) can be written as

$$t = \frac{\sqrt{N-1}}{2} \frac{\sum_{t=1}^{N-1} (u_i^{*2} - v_i^{*2})}{\sqrt{\Sigma u_i^{*2}} \sqrt{\Sigma v_i^{*2}}} .$$

ii) In testing the equality of variances of two normal populations when the sample size from each population is the same, N, say, instead of the usual F test based on S_1/S_2 , where S_1 and S_2 are the sample variances, we can use the equivalent t-test based on the statistic

$$t_{N-1} = \frac{\sqrt{N-1}\,(S_1 - S_2)}{2\sqrt{S_1 S_2}} .$$

REFERENCE

De Lury, D. B. (1938). Note on correlations Ann. Math. Statist., 9, 149-151.