An Application of Univariate Statistics to Hotelling's T^2

Dr. Athanase Polymenis

Department of Regional Economic Development, University of Central Greece, Greece

Abstract: Problem statement: Hotelling's T^2 statistic has been well documented in the existing literature and exact as well as asymptotic results have been obtained. In the present article, we focus on an important particular case of T^2 and we note that, there is no clear way to show, how univariate results, used in the theory of Student's t statistic, could be used to derive corresponding multivariate ones, for T^2. Therefore, our goal is to find an alternative method, which would be more useful than the usual one, in order to generalize directly univariate theory. **Approach:** At first, we used some matrix tools in order to obtain an equivalent algebraic form for T^2 and, then, we applied some univariate results concerning distributions which arise from the normal distribution. **Results:** We found an algebraic representation of T^2, which can be conceived as the natural extension of some results appearing in the literature and, we used our findings to show how standard univariate techniques can be applied in order to derive the exact and limiting distributions of T^2. **Conclusion:** Using the proposed representation of T^2 gives a better insight on the generalization of univariate results to multivariate analyses and indicates, at the same time, an alternative way to prove typical multivariate results. Furthermore, it allows for usual theoretical calculations to simplify.

Key words: Limiting distributions, multivariate theory, univariate results, independently distributed, null hypothesis, asymptotic results, orthogonal, comprehensive analysis

INTRODUCTION

The T^2 statistic was derived by Hotelling (1931), as a multivariate generalization of Student's t and has been very well documented over the years. Hsu (1938) derived the distribution of T^2 in the case where the null hypothesis is not true. Bowker (1960) derived a representation of T^2 as the ratio of two independent x_2 variables, the numerator being non central. This has, also, been described in details in Anderson (1984).

In the present study, we focus on a very important special case, which is as follows (see also Anderson (1984), Chapter 5). Let us first take a random vector sample $X_1, \ldots, X_r$, of dimension p, which is normally distributed and consider $T^2 = r(\bar{X} - \mu)^T M_r^{-1}(\bar{X} - \mu)$, where $\bar{X}$ is the mean vector and M_r the sample covariance matrix. It is well known that T^2 is used for testing hypotheses about the population mean vector μ. It is also interesting to note that, under the null hypothesis that $\mu = \mu_0$, it has been shown that $T^2 = T^2_0$ is a function of the likelihood ratio test λ_0, since $T_0^2 = (r-1)(\lambda_0^{-2/r} - 1)$. (Anderson (1984), U9).

We need to mention, at this point, the great use of the aforementioned statistic in diverse practical applications arising from many different fields. For instance, a recent article concerning marketing issues was published by Agarwall and Dey (2010) in order to assess the level of air travellers' satisfaction among Indian domestic airlines. Another recent example arises from the field of Insurance, where Momani *et al.* (2010) use simple and multiple regression techniques in order to examine the effect of different factors on the capital structure of insurance sector companies listed in the Amman stock market. Finally, a third indicative example concerns the use of a simple version of T^2, in an analysis of data from meteorological stations in order to examine climate change in Jordan (Hamdi *et al.*, 2009).

An important result is given by Corollary 5.2.1. of Anderson (1984), which gives the distribution of T^2 and which also allows for the case where the null hypothesis is not true. We now state Corollary 5.2.1., which we call Theorem 1, using our notation.

Theorem 1: Let $X_1, \ldots, X_r$ be a sample from $N(\mu, \Sigma)$ and let $T^2 = r(\bar{X} - \mu_0)^T M_r^{-1}(\bar{X} - \mu_0)$. The distribution of $[\frac{T^2}{r-1}][\frac{r-p}{p}]$ is non central F with p and r-p degrees of freedom and noncentrality parameter $r(\mu - \mu_0)^T \Sigma^{-1}(\mu - \mu_0)$. If $\mu = \mu_0$, then the F distribution is central.

In the present study, we need to extend the algebraic calculations appearing in Anderson (1984),

U12 and to show that, in the multivariate case, that is the case where p>1, T^2 can be put into a special algebraic form (representation), as a direct generalization of the univariate case p = 1. Under this new representation, we will revisit Theorem 1, for the multivariate case and we shall examine how univariate results can be used to infer about T^2. We shall focus on this methodology, since, according to our knowledge, it does not appear anywhere in the existing literature.

Furthermore, we consider the case where r→∞, using, once again, our algebraic representation for T^2. The corresponding asymptotic version of Theorem 1 is given by Theorem 5.2.3., of Anderson (1984), which we call Theorem 2. We now state briefly Theorem 2, using our notation see also (Polymenis, 2008).

Theorem 2: Let $\{X_i\}$, i = 1, 2,…, be a sequence of independently identically distributed random vectors with mean vector μ and covariance matrix Σ; then, the limiting distribution of T^2, as r→∞, is the $\chi^2(p)$ distribution if $\mu = \mu_0$.

In this case, we need to derive the limiting distribution of T^2 and to show, in the same spirit as before, how results for p=1 can be directly used to obtain results for p>1. Thus, we revisit Theorem 2, using our method. We, again, focus on using univariate results to obtain corresponding multivariate ones.

MATERIALS AND METHODS

A specific representation of T^2**:** We use here first, an algebraic approach similar to the one used by Anderson (1984), U15 in order to derive the distribution of T^2, at first under the null hypothesis. For reasons of simplicity, we first centre the X_i vectors, so that the null hypothesis becomes $\mu = \mu_0 = 0$. Furthermore, we choose the covariance matrix of $\sqrt{r}\bar{X}$ to be l, the Identity matrix, since this simplification will not affect the distribution of T^2 (see Anderson (1984)). The covariance matrix of X_i can, consequently, be chosen to be l, since it is equal to the covariance matrix of $\sqrt{r}\bar{X}$. We, then, proceed by considering an orthogonal (*pxp*) matrix $Q = (q_{ij})$ such that:

$$q_{1l} = \frac{\bar{X}_l}{\sqrt{\bar{X}^T\bar{X}}} \text{ for } l = 1,...,p$$

We now give the main result concerning the new representation of T^2.

Theorem 3: Hotelling's $T^2 = r\bar{X}^T M_r^{-1}\bar{X}$ can be written as:

$$T^2 = \frac{r\bar{X}^T\bar{X}}{\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2}$$

where, $V_i^{(1)}$ is the first element of the column vector $V_i = Q(X_i - \bar{X})$.

Proof of theorem 3: Let us set:

$$U = \sqrt{r}Q\bar{X} \text{ and } B = (r-1)QM_rQ^T$$

Where:

$$M_r = \frac{1}{(r-1)}\sum_{i=1}^{r}(X_i - \bar{X})(X_i - \bar{X})^T$$

Then we have:

$$u_1 = \sqrt{r}\sum_{i=1}^{p}q_{1l}\bar{X}_1 = \sqrt{r}(\frac{\bar{X}_1^T}{\sqrt{\bar{X}^T\bar{X}}}\bar{X}_1 + ...+$$

$$\frac{\bar{X}_p^T}{\sqrt{\bar{X}^T\bar{X}}}\bar{X}_p) = \frac{\sqrt{r}\bar{X}^T\bar{X}}{\sqrt{\bar{X}^T\bar{X}}} = \sqrt{r\bar{X}^T\bar{X}}$$

and, for j ≠ 1, we have:

$$u_j = \sqrt{r}\sum_{l=1}^{p}q_{jl}\bar{X}_l = \sqrt{r}\sum_{l=1}^{p}q_{jl}q_{1l}\sqrt{\bar{X}^T\bar{X}}$$

$$= \sqrt{r\bar{X}^T\bar{X}}\sum_{l=1}^{p}q_{jl}q_{1l} = 0$$

since Q is orthogonal.
Thus, we have that:

$$T^2 = r\bar{X}^T M_r^{-1}\bar{X} \text{ and } U^T B^{-1}U = \frac{r}{r-1}\bar{X}^T M_r^{-1}\bar{X} = \frac{T^2}{r-1}$$

where, $U = (u_1, 0,…,0)^T$ and:

$$B^{-1} = \begin{pmatrix} b^{11} & b^{12} & ... & b^{1p} \\ b^{21} & b^{22} & ... & b^{2p} \\ ... & ... & ... & ... \\ b^{p1} & b^{p2} & ... & b^{pp} \end{pmatrix}$$

Thus:

$$\frac{T^2}{r-1} = u_1^2 b^{11}$$

Set B to be the inverse matrix of B^{-1}, write $B = (b_{ij})$ and partition B as:

$$B = \begin{pmatrix} b_{11} & (b_{(1)})^T \\ b_{(1)} & B_{22} \end{pmatrix}$$

where, $b_{(1)} = (b_{12},\ldots, b_{1p})^T$ and:

$$B_{22} = \begin{pmatrix} b_{22} & \ldots & \ldots & b_{2p} \\ b_{32} & \ldots & \ldots & b_{3p} \\ \ldots & \ldots & \ldots & \ldots \\ b_{p2} & \ldots & \ldots & b_{pp} \end{pmatrix}$$

We partition B^{-1} in the same way:

$$B^{-1} = \begin{pmatrix} b^{11} & (b^{(1)})^T \\ b^{(1)} & B^{22} \end{pmatrix}$$

From Theorem 8.2.1, result 1(d), of Graybill (1969), we obtain:

$$b_{11} - (b_{(1)})^T B_{22}^{-1} b_{(1)} = (b^{11})^{-1}$$

This is a result from the theory of partitioned matrices.

If we write:

$$b_{11.2,\ldots,p} = \frac{1}{b^{11}} = b_{11} - (b_{(1)})^T B_{22}^{-1} b_{(1)}$$

Our statistic becomes:

$$T^2 = \frac{r\overline{X}^T \overline{X}}{\dfrac{b_{11.2,\ldots,p}}{(r-1)}}$$

On the other hand:

$$B = (r-1)QM_rQ^T = \sum_{i=1}^{r} (Q(X_i - \overline{X}))(Q(X_i - \overline{X}))^T$$

Setting then $V_i = Q(X_i - \overline{X})$, we obtain:

$$B = \sum_{i=1}^{r} V_i(V_i)^T$$

Now, partition X_i and V_i into two subvectors with 1 and (p-1) components, respectively, so that:

$$X_i = (X_i^{(1)}, (X_i^{(2)})^T)^T$$

and:

$$V_i = (V_i^{(1)}, (V_i^{(2)})^T)^T$$

with:

$X_i^{(1)} \in R$, $X_i^{(2)} = (X_{i,1}^{(2)},\ldots,X_{i,p-1}^{(2)})^T \in R^{p-1}$,

$V_i^{(1)} \in R$, $V_i^{(2)} = (V_{i,1}^{(2)},\ldots,V_{i,p-1}^{(2)})^T \in R^{p-1}$

Since $b_{11} \in R$ and $((V_1^{(1)})^2 + \ldots + (V_r^{(1)})^2) \in R$

$$E_2(E_2)^T = FV^{(2)}(V^{(2)})^T F^T = F(\sum_{i=1}^{r} V_i^{(2)}(V_i^{(2)})^T)F^T$$

We have:

$$b_{11} = \sum_{i=1}^{r}(V_i^{(1)})^2$$

We now compute the term $(b_{(1)})^T B_{22}^{-1} b_{(1)}$

For that purpose, setting:

$$(b_{(1)})^T B_{22}^{-1} = G \text{ and } B_{22} = H$$

We have:

$$(b_{(1)})^T B_{22}^{-1} b_{(1)} = ((b_{(1)})^T B_{22}^{-1})(B_{22})(B_{22}^{-1}b_{(1)}) = GHG^T$$

(since B_{22} is symmetric).

(Not that H is a ((p-1) × (p-1)) matrix and G a row vector of dimension p-1).

We have:

$$G = \sum_{i=1}^{r} V_i^{(1)}(V_i^{(2)})^T H^{-1} \text{ and } H = \sum_{i=1}^{r} V_i^{(2)}(V_i^{(2)})^T$$

Defining $V^{(1)} = (V_1^{(1)},\ldots,V_r^{(1)})$ and:

$$V^{(2)} = \begin{pmatrix} V_{1,1}^{(2)} & \ldots & V_{r,1}^{(2)} \\ \ldots & \ldots & \ldots \\ V_{1,p-1}^{(2)} & \ldots & V_{r,p-1}^{(2)} \end{pmatrix}$$

We then obtain $GHG^T = V^{(1)}(V^{(2)})^T H^{-1} V^{(2)}(V^{(1)})^T$.

Now, using Theorem A.4.1., of Anderson (1984), we can find a non-singular matrix F such that $FHF^T = l_{(p-1,p-1)}$, H being of full rank and, thus:

$$(F^T)^{-1}H^{-1}F^{-1} = I_{(p-1,p-1)}$$

Let $E_2 = FV^{(2)}$ so that $V^{(2)} = F^{-1}E_2$

We have:

$$E_2(E_2)^T = FV^{(2)}(V^{(2)})^T F^T = F(\sum_{i=1}^{r} V_i^{(2)}(V_i^{(2)})^T)$$

$$F^T = FHF^T = I_{(p-1,p-1)}$$

where, $l_{(p-1,p-1)}$ denotes the ((p-1) × (p-1)) identity matrix.

Thus, the p-1 rows of E_2 are orthogonal and their norm is equal to 1 (not that E_2 is a((p-1) × r) matrix.) Using lemma A.4.2, of Anderson (1984), we can find a ((r-p+1) × r) matrix E_1 such that $E = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}$ and E is an (r×r) orthogonal matrix.

Let us now set:

$$K_i = \sum_{\beta=1}^{r} e_{i,\beta} V_\beta$$

where the $e_{i,\beta}$'s are the entries of the matrix E and $K_i = (K_i^{(1)}, (K_i^{(2)})^T)^T$, with $K_i^{(1)} \in R$, and $K_i^{(2)} \in R^{p-1}$.

Then:

$$K_1^{(1)} = e_{1,1} V_1^{(1)} + ... + e_{1,r} V_r^{(1)}, ..., K_r^{(1)} = e_{r-p+1,1} V_1^{(1)} + ... + e_{r-p+1,r} V_r^{(1)}$$

So that, letting:

$K^{(1)} = (K_1^{(1)}, ..., K_r^{(1)})$, we obtain $(K^{(1)})^T = E(V^{(1)})^T$

Thus, $K^{(1)} = V^{(1)} E^T$, which is equivalent to $V^{(1)} = K^{(1)} E$ since E is orthogonal and $GHG^T = K^{(1)} E(E_2)^T (F^{-1})^T H^{-1} F^{-1} E^T (K^{(1)})T$.

However, on the one hand, $(F^{-1})^T H^{-1} F^{-1} = I_{(p-1,p-1)}$ and, on the other hand, $E_1(E_2)^T$ is the $((r-p+1) \times (p-1))$ matrix [0] because E is orthogonal and $E_2(E_2)^T = I^{(p-1,p-1)}$. For simplicity reasons, we write $I_{(p-1,p-1)} = I$. Thus, we obtain:

$$GHG^T = K^{(1)} E(E_2)^T (F^{-1})^T H^{-1} F^{-1} E_2 E^T (K^{(1)})^T$$

Where:

$$K^{(1)} ([0]^T, I)^T = (K_{r-p+2}^{(1)}, ..., K_r^{(1)})$$

and:

$$([0]^T, I)(K^{(1)})^T = (K_{r-p+2}^{(1)}, ..., K_r^{(1)})^T$$

We obtain finally:

$$GHG^T = (K_{r-p+2}^{(1)}, ..., K_r^{(1)})(K_{r-p+2}^{(1)}, ..., K_r^{(1)})^T$$
$$= (K_{r-p+2}^{(1)})^2 + ... + (K_r^{(1)})^2 = \sum_{i=r-p+2}^{r} (K_i^{(1)})^2$$

On the other hand, since $K_i = \sum_{\beta=1}^{r} e_{i\beta} V_\beta$, using Lemma 3.3.1., of Anderson (1984), we have:

$$\sum_{i=1}^{r} (V_i^{(1)})^2 = \sum_{i=1}^{r} (K_i^{(1)})^2$$

Thus, we obtain:

$$b_{11.2,...,p} = b_{11} - (b_{(1)})^T B_{22}^{-1} b_{(1)}$$
$$= \sum_{i=1}^{r} (K_i^{(1)})^2 - \sum_{i=r-p-2}^{r} (K_i^{(1)})^2 = \sum_{i=1}^{r-p+1} (K_i^{(1)})^2$$

and the T^2 statistic takes the from:

$$T^2 = \frac{r \overline{X}^T \overline{X}}{\frac{1}{r-1} \sum_{i=1}^{r-p+1} (K_i^{(1)})^2}$$

Since, as aforementioned:

$$\sum_{i=1}^{r} (V_i^{(1)})^2 = \sum_{i=1}^{r} (K_i^{(1)})^2$$

it results that:

$$\sum_{i=1}^{r-p+1} (V_i^{(1)})^2 = \sum_{i=1}^{r-p+1} (K_i^{(1)})^2$$

for all r>p-1 and, consequently:

$$T^2 = \frac{r \overline{X}^T \overline{X}}{\frac{1}{r-1} \sum_{i=1}^{r-p+1} (V_i^{(1)})^2}$$

Hence, Theorem 3 obtains.

Properties of the numerator and the denominator of T^2: We, first, present three useful preliminary lemmas.

Lemma 1: Under the null hypothesis, $\frac{1}{r-1} \sum_{i=1}^{r-p+1} (V_i^{(1)})^2$ is $\chi^2(r-p)$ distributed.

Proof of lemma 1: Since $V_i = Q(X_i - \overline{X})$, we obtain that:

$$V_i^{(1)} = q_{11}(X_i^{(1)} - \overline{X}_1) + q_{12}(X_i^{(2)} - \overline{X}_2) + ... + q_{1p}(X_i^{(p)} - \overline{X}_p)$$

with:

$$(X_i^{(j)} - \overline{X}_j) \sim N(0, 1 - \frac{1}{r})$$

and:

$$\tilde{0} = (0, ..., 0)^T, \overline{X} = (\overline{X}_1, ..., \overline{X}_p)^T$$

Then:

$$(X_i^{(j)} - \overline{X}_j) \sim N(0, 1 - \frac{1}{r}), \forall i = 1, ..., r-p+1 \text{ and } j = 1, ..., p$$

$$(\text{with } \overline{X}_j = \frac{1}{r} \sum_{i=1}^{r} X_i^{(j)})$$

This is a well-known univariate result, applied to $X_i^{(j)}$.

On the other hand, since q_{11}, q_{12},…, q_{1p} are the elements of the first row of the (p×p) matrix Q, which is orthogonal, it result that:

$$\sum_{j=1}^{p} q_{1j}^2 = 1, \text{ so that } E(V_i^{(1)}) = 0$$

and:

$$Var(V_i^{(1)}) = (1 - \frac{1}{r})\sum_{j=1}^{p} q_{1j}^2 = 1 - \frac{1}{r}$$

Thus:

$$r\overline{X}^T\overline{X} = r\overline{X}_1^2 + ... + r\overline{X}_p^2 \sim \chi^2(p)$$

which is, as aforementioned, the distribution of $(X_i^{(j)} - \overline{X}_j)$ and, so, the distribution of $\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$ is the same as that of $\sum_{i=1}^{r-p+1}(X_i^{(j)} - \overline{X}_j)^2$. Then, using theorem with proof (section 4.10, Grimmett and Stirzaker (2001)), applied to the independent, N(0,1) random variables $X_i^{(j)}$, we obtain that $\sum_{i=1}^{n}(X_i^{(j)} - \overline{X}_j)^2$ is $\chi^2(n-1)$ distributed. Finally, letting n = r – p+1, we obtain that $\sum_{i=1}^{r-p+1}(X_i^{(j)} - \overline{X}_j)^2$ is χ^2 (r-p) distributed.

We conclude that $\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$ is also χ^2 (r-p) distributed and independently of Q. N this is in accordance with a remark in the proof of Theorem 5.2.2, of Anderson (1984), stating that "since the conditional distribution of $b_{11.2,...p}$ does not depend on Q, it is unconditionally distributed as x^2".

Lemma 2: Under the null hypothesis, $r\overline{X}^T\overline{X}$ is $x^2(p)$ distributed.

Proof of Lemma 2: Since we know that:

$$\overline{X} = \begin{pmatrix} \overline{X}_1 \\ . \\ . \\ . \\ . \\ \overline{X}_p \end{pmatrix}, \text{ with } E(\overline{X}) = \begin{pmatrix} 0 \\ . \\ . \\ . \\ . \\ 0 \end{pmatrix} = \tilde{0}$$

and:

$$Var(\overline{X}) = \begin{pmatrix} 1/r & 0 & 0 & ... & 0 \\ 0 & 1/r & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & 0 & 1/r \end{pmatrix}$$

it results that $r\overline{X}^2 \sim \chi^2(1)$, so that all entries of the vector $\sqrt{r}\overline{X}$ are, independently, N(0,1) distributed and, as a result, $r\overline{X}^T\overline{X} = r\overline{X}_1^2 + ... + r\overline{X}_p^2 \sim \chi^2(p)$. Note that this is a direct generalization of the univariate case, p = 1, where $\overline{X} \sim N(0, \frac{1}{r})$ and, thus, $r\overline{X}^2 \sim \chi^2(1)$.

Lemma 3: $r\overline{X}^T\overline{X}$ and $\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$ are independently distributed random variables.

Proof of lemma 3: We first note that this lemma is equivalent to the application of Lemma 1 to expression (3.3), mentioned in Bowker (1960). We will now elaborate on that, using the result of Theorem 3.
Let us call V the vector:

$$\begin{pmatrix} V_1^{(1)} \\ V_2^{(1)} \\ . \\ . \\ . \\ V_{r-p+1}^{(1)} \end{pmatrix} =$$

$$\begin{pmatrix} q_{11}(X_1^{(1)} - \overline{X}_1) + q_{12}(X_1^{(2)} - \overline{X}_2) + ... + q_{1p}(X_1^{(p)} - \overline{X}_p) \\ q_{11}(X_2^{(1)} - \overline{X}_1) + q_{12}(X_2^{(2)} - \overline{X}_2) + ... + q_{1p}(X_2^{(p)} - \overline{X}_p) \\ . \\ . \\ q_{11}(X_{r-p+1}^{(1)} - \overline{X}_1) + q_{12}(X_{r-p+1}^{(2)} - \overline{X}_2) + ... + q_{1p}(X_{r-p+1}^{(p)} - \overline{X}_p) \end{pmatrix}$$

The vector $\begin{pmatrix} \overline{X} \\ V \end{pmatrix}$ is normally distributed. Indeed, every linear combination of its components is a linear combination of the $X_i^{(j)}$'s, which are independently distributed normal variables.

Thus, independency between $\overline{X}$ and V amounts to independency between $\overline{X}_j$ and $V_i^{(1)}$, $\forall i = 1,...,r-p+1$ and $j = 1,...,p$. Hence, in order to prove Lemma 3, it is sufficient to prove that $Cov(\overline{X}_j, \sum_{j=1}^{p}q_{1j}(X_i^{(j)} - \overline{X}_j)) = 0$, which, in turn, amounts proving that $Cov(\overline{X}_j, (X_i^{(l)} - \overline{X}_l)) = 0$, for all j, l = 1,…, p. In the case where j ≠ l, $E(\overline{X}_j X_i^{(l)}) = 0$ since $E(X_i^{(l)}X_k^{(j)}) = 0$, for all k = 1,...,r and, thus, $Cov(\overline{X}_j,$

$(X_i^{(l)} - \bar{X}_l))$ = 0. In the case where j = l, $E(\bar{X}_j X_i^{(j)}) = E(\bar{X}_j^2) = \frac{1}{r}$, so that, $Cov(\bar{X}_j, (X_i^{(j)} - \bar{X}_j))$=0, as well. The latter case is equivalent to the well known univariate result $Cov(\bar{X}, (X_i - \bar{X})) = 0$, which shows independency between numerator and denominator of student's t. (See, also, Grimmett and Stirzaker (2001), section 4.10, for a comprehensive analysis of the univariate case.)

We, now, present a useful lemma, concerning the limiting properties of the denominator of T^2.

Lemma 4: As r→∞, $\frac{1}{r-1}(\sum_{i=1}^{r-p+1} V_i^{(1)})^2 \to 1$ in probability.

Proof of Lemma 4: We first prove Lemma 4, for p = 1, since we will use its results in the proof of Lemma 4, for p>1. In this case, $T^2 = t^2$, the denominator of t^2 is $\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2$ and, we first need to show that, as r→∞, $\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2 \to 1$ in probability. This, in turn, amounts showing that $E(\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2) \to 1$ and $Var(\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2) \to 0$, as r→∞.

Thus, we first calculate $E(\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2)$. We know that $E((X_i - \bar{X})^2) = Var(X_i - \bar{X})$, since $E(X_i - \bar{X}) = 0$. On the other hand, $Var(X_i - \bar{X}) = 1 - \frac{1}{r}$, so that:

$$E(\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2) = \frac{1}{r-1}\sum_{i=1}^{r}Var(X_i - \bar{X})$$
$$= \frac{1}{r-1}(r-1) = 1$$

We, then, calculate $\frac{1}{(r-1)^2}Var(\sum_{i=1}^{r}(X_i - \bar{X})^2)$. In order to do that, we first calculate $\frac{1}{(r-1)^2}(\sum_{i=1}^{r}Var(X_i - \bar{X})^2)$, which we call (1). We show, first, that:

$$Var(X_i - \bar{X})^2 = Var(X_i^2) - 2Cov(X_i, \bar{X}) +$$
$$Var(\bar{X}^2) = 2 - \frac{2}{r} + \frac{2}{r^2}$$

Indeed, since $X_i \sim N(0,1)$, $Var(X_i^2) = 2$ and $Cov(X_i, \bar{X}) = \frac{1}{r}Cov(X_i, \sum_{i=1}^{r}X_i) = \frac{1}{r}Var(X_i) = \frac{1}{r}, \forall i = 1,...,r$. On the other hand:

$$Var(\bar{X}^2) = \frac{1}{r^4}Var[(X_1 + ... + X_r)^2] = \frac{1}{r^4}Var$$
$$[X_1^2 + ... + X_r^2 + 2\sum_{i \neq j}X_i X_j]$$
$$= \frac{1}{r^4}[\sum_{i=1}^{r}Var(X_i^2) + 4\sum_{i \neq j}Var(X_i X_j)]$$

This is a consequence of the independency between the random variables X_i and X_j, so that $Cov(X_i^2, X_j X_k) = 0$, for all j ≠ k and $Cov(X_i^2, X_j^2) = 0, \forall i \neq j$. Since $Var(X_i X_j) = E(X_i^2 X_j^2) = E(X_i^2)E(X_j^2) = 1$ and $Var(X_i^2) = 2$, it then results that :

$$Var(\bar{X}^2) = \frac{1}{r^4}(2r) + \frac{4}{r^4}((r-1) + (r-2) + ... + 1)$$
$$= \frac{1}{r^4}[2r + 4\frac{r(r-1)}{2}] = \frac{2}{r^2}$$

Thus:

$$Var(X_i - \bar{X})^2 = 2 - \frac{2}{r} + \frac{2}{r^2}$$

Then:

$$(1) = \frac{1}{(r-1)^2}(\sum_{i=1}^{r}Var(X_i - \bar{X})^2)$$
$$= \frac{1}{(r-1)^2}(r(2 - \frac{2}{r} + \frac{2}{r^2})) = \frac{2r - 2 + 2/r}{(r-1)^2}$$

Hence, (1)→0 as r→∞.

We, then, calculate $\frac{2}{(r-1)^2}\sum_{k \neq l}Cov[(X_k - \bar{X})^2, (X_l - \bar{X})^2]$, which we call (2). We have that:

$$Cov[(X_k - \bar{X})^2, (X_l - \bar{X})^2] = E[(X_k - \bar{X})^2$$
$$(X_l - \bar{X})^2] - E[(X_k - \bar{X})^2]E[(X_l - \bar{X})^2]$$

On one hand:

$$E[(X_k - \bar{X})^2] = Var(X_k - \bar{X}) = 1 - \frac{1}{r}$$

So that:

$$E[(X_k - \bar{X})^2]E[(X_1 - \bar{X})^2] = (1 - \frac{1}{r})^2$$

and on the other hand, we need to find $E[(X_k - \bar{X})^2(X_1 - \bar{X})^2]$. Then, simple calculations show that:

$$E(X_k^2 X_1^2) = 1, \ E(X_k^2 X_1 \bar{X}) = \frac{1}{r}, \ E(X_k^2 \bar{X}^2) = \frac{r+2}{r^2},$$

$$E(X_k \bar{X}^2 X_1) = \frac{2}{r^2}, \ E(X_k \bar{X}^3) = \frac{3}{r^3}, E(\bar{X}^4) = \frac{2r+1}{r^3}$$

So that:

$$E[(X_k - \bar{X})^2(X_1 - \bar{X})^2] = 1 - \frac{2}{r} + \frac{14}{r^2} - \frac{11}{r^3}$$

It, then, results that:

$$\sum_{k \neq 1} Cov[(X_k - \bar{X})^2, (X_1 - \bar{X})^2] =$$

$$\sum_{k \neq 1} [(1 - \frac{2}{r} + \frac{14}{r^2} - \frac{11}{r^3}) - (1 - \frac{1}{r})^2] =$$

$$\sum_{k \neq 1} (\frac{13}{r^2} - \frac{11}{r^3}) = \frac{r(r-1)}{2}(\frac{13}{r^2} - \frac{11}{r^3}) = \frac{13}{2} - \frac{12}{r} + \frac{11}{2r^2}$$

and, thus:

$$(2) = \frac{2}{(r-1)^2}(\frac{13}{2} - \frac{12}{r} + \frac{11}{2r^2})$$

$$= \frac{13}{(r-1)^2} - \frac{24}{r(r-1)^2} + \frac{11}{r^2(r-1)^2}$$

Hence, (2)→0 as r→∞.

Finally, $\frac{1}{(r-1)^2} Var(\sum_{i=1}^{r}(X_i - \bar{X})^2) = (1)+(2)$, which

tends to 0 as r→∞. This proves Lemma 4, for the case p = 1.

We now need to generalize that and show how univariate results can be used for obtaining corresponding multivariate ones. For that, let us consider p>1 and $T^2 = \dfrac{r\bar{X}^T \bar{X}}{\dfrac{1}{r-1}\sum\limits_{i=1}^{r-p+1}(V_i^{(1)})^2}$, where:

$$V_i^{(1)} = q_{11}(X_i^{(1)} - \bar{X}_1) + q_{12}(X_i^{(2)} - \bar{X}_2) + ... + q_{1p}(X_i^{(p)} - \bar{X}_p)$$

The general proof of Lemma 4, will be an immediate consequence of the results found for the case p = 1.

In order to prove Lemma 4, we need showing that:

$$E(\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2) \rightarrow 1, \text{ as } r \rightarrow \infty$$

and:

$$Var(\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2) \rightarrow 0 \text{ as } r \rightarrow \infty$$

We, now, digress to say that $\sum_{i=1}^{r}(X_i - \bar{X})^2$ is the univariate version of $\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$, in which case, p = 1, q_{11}=1 and q_{1j}= 0 for $j \neq 1$ as aforementioned Consequently, the limiting values of the expectations of $\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$ and $\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2$, must be equal and, those of the variances of $\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2$ and $\frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2$ must be equal, as well. This shows that:

$$E(\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2) \rightarrow 1$$

and:

$$Var(\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2) \rightarrow 0, \text{ as } r \rightarrow \infty$$

which concludes the general proof of Lemma 4.

RESULTS

The exact distribution of T^2: We, first, give two important remarks, based on previous calculations.

The algebraic representation of T^2, given in Theorem 3, extends Bowker (1960) expression (3.3), or Anderson (1984) expression (23), p. 162, for the particular case we are studying.

T^2, as presented in Theorem 3, is a generalization of the univariate case, p = 1, with $V_i^{(1)} = X_i - \bar{X}$, q_{11}=1 and $q_{1j} = 0$ for $j \neq 1$. In this case, T^2 equals t^2 (the square of student's t) and:

$$t^2 = \frac{r\bar{X}^2}{\dfrac{1}{r-1}\sum\limits_{i=1}^{r}(X_i - \bar{X})^2}$$

We now prove Theorem 1, using the result of Theorem 3.

Proof of Theorem 1: From Lemmas 1, 2, 3, under the null hypothesis H_0, T^2 is distributed as the ratio of two independent χ^2's. Then, it is well known, by definition of the *F* statistic (see, for instance, Grimmett and Stirzaker (2001), section 4.10), that:

$$\frac{r\bar{X}^T\bar{X} / (p)}{\sum_{i=1}^{r-p+1} (V_i^{(1)})^2 / (r-p)} = \frac{T^2}{r-1}\left(\frac{r-p}{p}\right) \sim F(p, r-p).$$

This result proves Theorem 1, when the null hypothesis is true. It is easy to note, that, for the univariate case $p = 1$, $T^2 = t^2 \sim F(1, r-1)$, under H_0.

In the case where the null hypothesis is not true, we use the same rationale as before. Let us consider:

$$T^2 = \frac{r(\bar{X}-\mu_0)^T(\bar{X}-\mu_0)}{\frac{1}{r-1}\sum_{i=1}^{r-p+1}(V_i^{(1)})^2}$$

with the denominator being same as before. Following notation as in Anderson (1984), p. 161, set $E[\sqrt{r}(\bar{X}-\mu_0)] = v^*$, so that $\sqrt{r}(\bar{X}-\mu_0)$ is $N(v^*, I)$ distributed and, consequently, the numerator of T^2 has a non central χ^2 distribution. The denominator of T^2 is the same as for the case where H_0 was true, so that Lemma 1 holds. Lemma 3 holds too, since, taking $\bar{X}-\mu_0$, instead of $\bar{X}$, does not affect independency between numerator and denominator of T^2. Finally, since the numerator of T^2 has a non central χ^2 distribution, $\frac{T^2}{r-1}\left(\frac{r-p}{p}\right)$ is non central F. This result proves Theorem 1, when H_0 is not true.

The limiting distribution of T^2: We now give two important remarks, concerning the limiting distribution of T^2:

- In the univariate case, we know that, from Lemma 2, $r\bar{X}^2$ is $x^2(1)$ distributed. Then, using Lemma 4, we conclude that, under H_0, as $r\to\infty$, t^2 tends in distribution to $x^2(1)$

- Since, from Lemma 2, $r\bar{X}^T\bar{X} \sim \chi^2(p)$, we use Lemma 4, in the same way as for $p = 1$ and we conclude that, under H_0, as $r\to\infty$, T^2 tends in distribution to $x^2(p)$. We also note that, if we replace p by 1, then $T^2 = t^2 \to \chi^2(1)$, in distribution, as $r\to\infty$, which coincides with the previous remark

We now prove Theorem 2, using the result of Theorem 3.

Proof of Theorem 2: The limiting distribution of T^2, mentioned in the above remarks, still remains valid, under H_0, even if the random vector sample $X_1,\ldots X_r$ is not normally distributed. Indeed, in this case, it is sufficient to apply a multivariate central-limit theorem (Theorem 3.4.3 anderson (1984)), stating that the limiting distribution, as $r\to\infty$, of $\sqrt{r}\bar{X}$ is $N(\tilde{0}, I)$ and, hence, we conclude that the limiting distribution of $r\bar{X}^T\bar{X}$ is $x^2(p)$. On the other hand, the asymptotic results of Lemma 4 still hold. Thus, using Proposition 6.3.8., of Brockwell and Davis (1991), for k = 1, we conclude that the limiting distribution of T^2 is $x^2(p)$. This result proves Theorem 2.

DISCUSSION

In the present article, we found an alternative form (representation) for an important particular case of Hotelling's T^2, which extends Bowker (1960) results, in order to prove usual multivariate exact and limiting theorems, using standard univariate statistics. We remark, that, for the univariate case $p = 1$, the usual form of T^2 and the new representation of T^2, given by Theorem 3, have exactly the same algebraic configurations and are both equal to the usual expression of t^2. In this case, exact and asymptotic properties of t^2 are easily derived. However, for the more complicated multivariate case, Theorem's 3 result has the obvious advantage, over the usual form of T^2, of using directly simple univariate results, obtained for p = 1, in order to derive exact and asymptotic statistical properties for T^2.

CONCLUSION

An important and very well documented univariate distribution, arising from the normal distribution, is Student's t. Our study showed how exact and asymptotic results, concerning t, can help deriving corresponding properties for a well known multivariate analogue of the square of t, namely Hotelling's T^2, by using a specific algebraic representation of T^2. In other words, in the present article, standard univariate material was used in a quite straightforward manner for explaining multivariate theory. As a result, usual multivariate theoretical calculations are considerably simplified.

REFERENCES

Agarwall, S. and A.K. Dey, 2010. Perception mapping of travellers: case of six indian domestic airlines. Am. J. Econ. Bus. Admin., 2: 141-146. DOI: 10.3844/ajebasp.2010.141.146

Anderson, T.W., 1984. An Introduction to Multivariate Statistical analysis. 2nd Edn., Wiley Series in Probability and Statistics, ISBN: 0471889873, pp: 69-598.

Bowker, A.H., 1960. A Representation of Hotelling's T^2 and Anderson's Classification Statistic W in Terms of Simple Statistics. In: Contributions to Probability and Statistics, Olkin, I., S.G. Ghurye, W. Hoeffding, W.G. Madow and H.B. Mann (Eds.). Stanford University, Stanford, California, pp: 142-149.

Brockwell, P.J. and R.A. Davies, 1991. Time Series: Theory and Methods. 2nd Edn., Springer-Verlag, ISBN: 0387974296, pp: 207.

Graybill, F.A., 1969. Introduction to Matrices with Applications in Statistics. Wadsworth Publishing Company, CA., pp: 164-165.

Grimmett, G. and D. Stirzaker, 2001. Probability and Random Processes. 3rd Edn., Oxford University Press, ISBN: 0198572220, pp: 119-121.

Hamdi, R., M. Abu-Allaban, A. Al-Shayeb, M. Jaber and N.M. Momani, 2009. Climate change in jordan: a comprehensive examination approach. Am. J. Environ. Sci., 5: 58-68. DOI: 10.3844/ajessp.2009.58.68

Hotelling, H., 1931. The generalization of student's ratio. Annals Math. Stat., 2: 360-378. DOI: 10.1214/aoms/1177692778

Hsu, P.L., 1938. Notes on Hotelling's generalized T. Annals Math. Stat., 9: 231-243. DOI: 10.1214/aoms/1177692778

Momani, G.F., M.A. Alsharayri and M.M. Dandan, 2010. Impact on firm's characteristics on determining the financial structure on the insurance sector firms in Jordan. J. Soc. Sci., 6: 282-286. DOI: 10.3844/jssp.2010.282.286

Polymenis, A., 2008. A note on a validity test using the stochastic em algorithm in order to assess the number of components in a finite mixture model. Statistics, 42: 261-274. DOI: 10.1080/02331880701829922