Vol. (49) \ No. 144 \ June \ Year 2024

DOI: https://doi.org/10.31272/jae.i144.1233
Available online at: https://admics.uomustansiriyah.edu.iq/index.php/admeco

Journal of Administration & Economics

Mustansiriyah University
College of Administration & Economics

P- ISSN: 1813-6729
E-ISSN: 2707-1359

Conducting Multivariate One- and Two-Sample Statistical tests using SPSS Syntax

Delshad Shaker Ismael Botani
Dep. of Statistics & Informatics, Adm. & Eco. College, Salahaddin University-Erbil, Erbil, Iraq
Email: first author's email, ORCID:/ https://orcid.org/0000-0002-5465-8989

Article Information

Article History:
Received: 1 / 2 / 2024
Accepted : 10 / 6 / 2024
Available Online: 25 / 7 / 2024
Page no : 1 – 13

Keywords:
Multivariate one-sample , Multivariate two-sample , SPSS syntax , Hotelling's T^2 , Custom dialog builder .

Correspondence:
Researcher name: Dr. Delshad Shaker Ismael Botani
Email: Delshd.botani@su.edu.krd

Abstract

The availability of multivariate one-sample and two-sample testing procedures is often restricted or merged with other statistical tests in many statistical applications. SPSS software, known for its user-friendly Graphical User Interface (GUI) for statistical analysis, also provides a specialised programming language called SPSS syntax. This syntax allows users to execute statistical procedures by writing commands, providing an alternative to the GUI interface. The purpose of this paper is to create new dialogs written in the SPSS syntax language using custom dialog builder for extensions. These dialogs are designed to test multivariate one- and two-sample data when ∑ is unknown by applying Hotelling's T2 test statistic, especially in situations where the GUI lacks these particular functionalities. In the case of two-sample data, both independent and paired multivariate testing methods are employed. Examples are supplied to demonstrate the key analyses performed using Hotelling's T2 tests. The syntax written for the analyses was performed using SPSS software v.27. The adoption of the developed dialogs is recommended for researchers requiring analysis of multivariate one- and two-sample data with limited or no programming experience, with potential future incorporation of the dialogs into forthcoming SPSS versions by IBM. Extending the scope of statistical analyses supported by SPSS dialogs through future research that encourages their use and supports wider acceptance of additional statistical methods not included in standard versions of SPSS.

1. Introduction

In statistics, data analysis is categorized as either univariate or multivariate. Univariate analyses, which focus on a single dependent variable, include univariate mean tests, ANOVA, simple linear regression, multiple linear regression, and so on. On the other hand, multivariate analyses look at correlations between several dependent variables or samples using sophisticated statistical techniques such multivariate one- and two-sample testing, sub-vector testing, MANOVA, multivariate multiple regression, etc.

In many statistical applications, the availability of multivariate one-sample and two-sample testing procedures is often constrained or merged with other statistical tests. To address this challenge, SPSS software offers a user-friendly Graphical User Interface (GUI) for statistical analysis, alongside a specialized programming language known as SPSS syntax. This syntax empowers users to execute additional statistical procedures through command-based operations, presenting an alternative to the GUI interface. In addition to writing syntax, it is possible to add new dialogs through "Extensions" menu to execute this written syntax, making them easily accessible and usable graphically.

This paper focuses on taking advantage of the capabilities of SPSS syntax to enhance multivariate data analysis. Specifically, the objective is to develop a new dialog using the "custom dialog builder for extensions" within SPSS. This new dialog is designed to facilitate the testing of multivariate one- and two-sample data when ∑ is unknown using Hotelling's T^2 test statistic, particularly when the GUI lacks these specific functionalities.

The approach addresses the needs of researchers who may have limited programming experience or prefer working with statistical dialogs. By providing examples and demonstrations of key analyses conducted using Hotelling's T2 tests, the effectiveness and flexibility of the proposed solution are illustrated. The new dialog presented was developed using SPSS 27 [3] and is compatible with SPSS version 24 or higher.

2. Multivariate Mean Vectors Tests

In many statistical analyses, researchers try to draw conclusions about population parameters such as mean and variance (μ,σ) depending on sample data. Consequently, they may seek to test hypotheses about the population mean of a random variable through one-sample or two-sample univariate tests, commonly known as t-tests. However, in the case of multiple variables, the test statistic employed is Hotelling's T2-test. The reason of using multivariate tests over univariate ones is as follows [9]:

1. Using p univariate tests raise the probability of α (Type I error), whereas in multivariate this probability remains constant.
2. Considering the correlation between p variables in multivariate tests.
3. The multivariate test is often more powerful (1-β) due to combined small effects of variables leads to test significance.
4. Multivariate means tests frequently provide linear combinations of variables, showing how the variables together reject the hypothesis.

2.1. One-Sample Testing on μ vector

In the univariate testing of one random variable (x) normally distributed for determining whether a specific value μ_0 is a reasonable value for the population mean, the hypothesis testing can be as follows: $H_0: \mu = \mu_0 \;\; and\; H_1: \mu \neq \mu_0$

To get a proper test statistic for a random sample of size n $(x_1, x_2, \dots, x_n)$ from a normal population with unknown σ, the following formula will be used [5]:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \tag{1}$$

$where\ S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}$ is the standard deviation of X.

Now, suppose that X is an n×p matrix represents a random sample of size n (X~$N_p(\mu, \sum)$). In the case of multivariate one-sample testing, when the variance-covariance matrix (∑) is unknown and there are p × 1 vector μ_0 that represent a reasonable value for the population mean, the appropriate test statistic is Hotelling's T^2 for testing the null hypothesis: $H_0: \mu = \mu_0 \;\; and\; H_1: \mu \neq \mu_0$

The formula for Hotelling's T2-test is as follows [5, 9]:

$$T^2 = (\bar{X} - \mu_0)' \left[\frac{S}{n}\right]^{-1} (\bar{X} - \mu_0) = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) \tag{2}$$

where, $\mu_0 = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{bmatrix}$, $\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix}$, and $S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix}_{p*p}$

If the following is noted, we reject H_0 and accepting H_1 at the significance level of α:

$$\frac{n-p}{(n-1)p}T^2 > F_{p,n-p}(\alpha)$$

$$\Rightarrow \alpha = P\left[T^2 > \frac{(n-1)p}{n-p}F_{p,n-p}(\alpha)\right] \qquad (3)$$

where, $F_{p,n-p}$ represents a random variable following an F-distribution with p and n-p degrees of freedom (d.f.).

Equations 2 and 3 can be applied to univariate one-sample testing when p=1, indicating a scenario where only one random variable is being tested.

2.2. Independent Two-Sample Testing (Two Mean Vectors)

Recalling the univariate two-sample testing, a statistical method used to compare the means of two independent groups (or two independent populations), aims to determine if there is a significant difference between them. It is assumed that the undertaken two samples are independent and the variances of samples are equal ($\sigma_1^2 = \sigma_2^2 = \sigma^2$), with unknown σ^2. The assumptions of independence and equal variances are necessary for the t statistic to test the hypothesis [9, 10]:

$$H_0: \mu_1 = \mu_2 \quad and\ H_1: \mu_1 \neq \mu_2$$

Then, the t statistic will be written as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (4)$$

where, $S_p^2 = \frac{SS_1 + SS_2}{n_1 + n_2 - 2} = \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1 + n_2 - 2}$ is the pooled variance.

The null hypothesis (H_0) is rejected if $|t| > t_{\alpha/2, n_1+n_2-2}$.

The above equation cannot be applied when there are multiple variables. Therefore, instead of using the t-test, Hotelling's T^2 is employed for testing multivariate independent two-samples, assuming $\sum_1 = \sum_2 = \sum$ with an unknown $\sum$. This test is appropriate for comparing responses from one set of data (population 1) with independent responses from another set of data (population 2) [5]. Consider the case where p variables are measured on each sampling unit of the two samples. The hypothesis for comparing the independent two samples is as follows [8, 9]:

$$H_0: \mu_1 = \mu_2 \quad and\ H_1: \mu_1 \neq \mu_2$$

The T^2-statistic for the above hypothesis will be written as follows

$$T^2 = \frac{n_1 n_2}{n_1 + n_2}(\bar{X}_1 - \bar{X}_2)' S_p^{-1} (\bar{X}_1 - \bar{X}_2) \qquad (5)$$

where, $\bar{X}_1 = \begin{bmatrix} \bar{X}_{11} \\ \bar{X}_{21} \\ \vdots \\ \bar{X}_{p1} \end{bmatrix}$, $\bar{X}_2 = \begin{bmatrix} \bar{X}_{12} \\ \bar{X}_{22} \\ \vdots \\ \bar{X}_{p2} \end{bmatrix}$,

and $S_p = \frac{1}{n_1+n_2-2}[(n_1-1)S_1 + (n_2-1)S_2]$ is the pooled variance-covariance matrix,

and S_1, S_2 are the variance-covariance matrices for the first and second samples respectively.

It is possible to write Eq. (5) as a constant (let's say k) times Mahalanobis distance (D^2) [1]:

$$T^2 = kD^2 \qquad (6)$$

where, $D^2 = (\bar{X}_1 - \bar{X}_2)' S_p^{-1} (\bar{X}_1 - \bar{X}_2)$

At the significance level of α and d.f. equals to $n_1 + n_2 - 2$, the null hypothesis (H_0) will be rejected if the observed: $T^2 > {T^2}_\alpha$

The T^2-statistic can be readily transformed to an F-statistic using the following [8, 9]:

$$\frac{n_1+n_2-p-1}{(n_1+n_2-2)p}T^2 > F_{p,n_1+n_2-p-1}(\alpha)$$
$$\Rightarrow T^2 > \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}(\alpha) \tag{7}$$

where, F_{p,n_1+n_2-p-1} represents a random variable following an F-distribution with the d.f. of p and n_1+n_2-p-1.

Equations 5 and 7 can serve in univariate independent two-sample testing when p=1, representing a scenario where only one random variable is being tested for each sample.

2.3. Paired Two-Sample Testing

The paired sample t-test, sometimes called the dependent sample t-test, is a statistical method used to determine whether the mean difference between two samples of observations is zero, where the two samples are not independent practically in many situations, such as when a measurement is taken twice to the same individual (before and after physical exercise) or when subjects are matched based on a certain factor, such as medical history or household history (Spouses, sisters and brothers).

In univariate hypothesis testing for paired samples, if $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$ are random samples from a normal population with $\sum$ unknown, the two samples can be reduced to one ($d_i = x_i - y_i$). The appropriate test statistic for testing the hypothesis:

$$H_0: \mu_x = \mu_y \ and\ H_1: \mu_x \neq \mu_y \qquad or \qquad H_0: \mu_d = 0\ and\ H_1: \mu_d \neq 0$$

is as follows: $t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$,

$where,\quad s_d^2 = \frac{\sum(d_i-\bar{d})^2}{n-1}$

As observed in the univariate case, the process will be extended to p variables on each sampling unit. Then, to testing the hypothesis: $H_0: \mu_d = 0\ and\ H_1: \mu_d \neq 0$,

the T^2-statistic will be written as follows [8, 9]:

$$T^2 = n\bar{d}' S_d^{-1}\bar{d}) \tag{8}$$

where, $\bar{d} = \frac{\sum d_i}{n}\ and\ S_d = \frac{\sum(d_i-\bar{d})(d_i-\bar{d})'}{n-1}$

At the α level of significance, we reject H_0 in favor of H_1 if the observed:

$$\frac{n-p}{(n-1)p}T^2 = F_{p,n-p}(\alpha) \qquad \Rightarrow \alpha = P\left[T^2 > \frac{(n-1)p}{n-p}F_{p,n-p}(\alpha)\right] \tag{9}$$

When p=1, or when just one random variable is being tested for each sample, Eq. (8 and 9) can be used in univariate paired two-sample testing.

3. Implementations of SPSS Syntax

The goal of this section is to provide an explanation of the SPSS syntax that will be used in the next section, based on the theoretical underpinnings of the multivariate one- and two-sample testing equations discussed earlier.

Firstly, the data should be prepared to commence to process of finding the required statistics for testing the mean vectors whether the data is independent or not. The “Matrix” and “End Matrix” commands [2] are used for preparing the data to be manipulated as a matrix. Next, using the "Get" command to retrieve the variables from the data editor window (let's say, X1 and X2) as seen below:

```
Matrix.
get x
   /variables= x1, x2
   /missing=accept
   /sysmis=omit.
End Matrix.
```

The subcommands “missing” and “sysmis” are used to accept missing values and reject system missing values from the dataset.

Secondly, after the variables are defined in the matrix, the process involves computing the statistics required for conducting the hypothesis test, such as the mean vectors, variance-covariance matrix, determinants, inverse of the matrix, etc. The main SPSS command used for this purpose is the "Compute" command, as shown below:

```
compute n=nrow (x).
compute p=ncol(x).
compute xbar=transpos(one)*x/n.
compute xbar_mat=one*xbar.
compute mean_dev=x-xbar_mat.
compute var_x=(transpos(mean_dev)* mean_dev)/(n-1).
compute inv_x=inv(var_x).
```

Certainly, there are several functions utilized in the above syntax including nrow, ncol, transpos, det, inv, etc., which are necessary for the computations. Additional information about these functions can be obtained from IBM Corporation [2].

In the third step of writing the SPSS syntax, The Hotelling's T2-statistic is computed, along with determining the probability value of the F distribution as outlined in the following syntax:

compute hotelling_t^2=(using Eq. (2) or (5) or (8) previous section).

compute p_value=SIG.F(using Eq. (3) or (7) or (9) previous section, p, n1+n2-p-1).

The function Sig.F(quantity, d.f.1, d.f.2) is used for finding upper-tail area, i.e. p-value of an F-distribution [6].

Finally, the results of the hypothesis testing of the mean vectors should be printed using the "Print" command, as described below:

```
Print Hotelling_t2
/Title "The value of Hotelling T2"
/Format f8.3.
```

The subcommand 'title' is used to display the title in the SPSS output window, while 'format' is utilized to specify how numeric values are displayed in the output, ensuring a total field width of 8 characters, which includes 3 decimal places.

The complete SPSS syntax for the three methods of multivariate one- and two-sample testing is outlined in the Appendix.

4. Custom Dialog Builder

IBM corporation built a new menu named "Extensions" that is available in SPSS versions 24 and later [7]. This menu enables users to perform additional tasks, extending the software's capabilities beyond its basic functionality through using the "Custom Dialog Builder for Extensions". The Custom Dialog Builder for Extensions assists users to create a GUI, known as a custom node dialog [4].

In this paper, three dialogs were designed to facilitate the multivariate one-sample test (means Vector), independent two-sample test, and paired two-sample test using the syntax provided in the Appendix. These three dialogs are located in the "Multivariate Sample Testing" submenu within the "Analyze" menu, as illustrated in the Figure below.

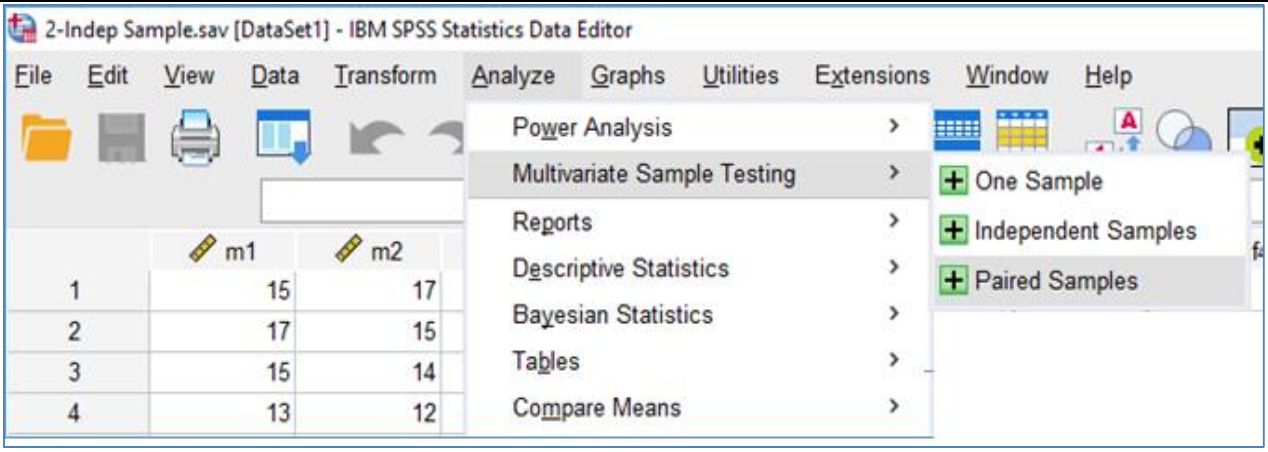

Figure (1): Custom Node Dialog for Multivariate Sample Testing

When clicking on the first dialog (node) labeled "One Sample," a new dialog window will appear with three boxes: Variables, Population Means (Mu), and Sample Variables. All variables will appear in the first box (Variables). Next, the variable with the supposed population means should be dragged or placed into the Population Means box, while the variables proposed to be tested should be placed into the Sample Variable box. Click on the run button to proceed Figure (2).

The syntax provided in the Appendix will be placed into the "Syntax Template" within the "Custom Dialog Builder for Extensions". Subsequently, the content of the boxes will be identified within the syntax by boxes Identifier names, which will be enclosed between two percentage symbols on each side (%%Identifier name%%), as shown below:

Get G1 /variables= %%First_group%%.

Get G2 /variables= %%Second_group%%.

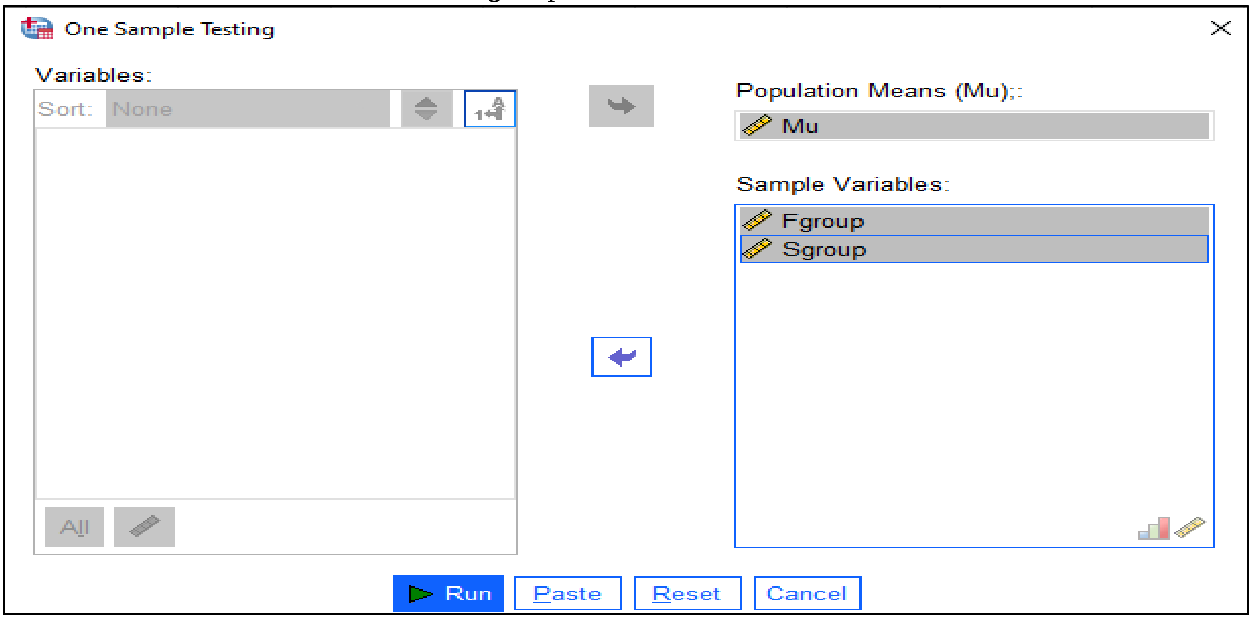

Figure (2): Dialog for Multivariate One Sample Testing

Regarding the second button or dialog (node) labeled "Independent Samples," clicking on it will open a new dialog window featuring three boxes: Variables, First Group, and Second Group. Initially, all variables will be displayed in the first box (Variables). Then, the variables intended for the first group should be dragged or moved to the First Group box, while those intended for the second group should be placed into the Second Group box. Finally, click on the Run button to proceed Figure (3).

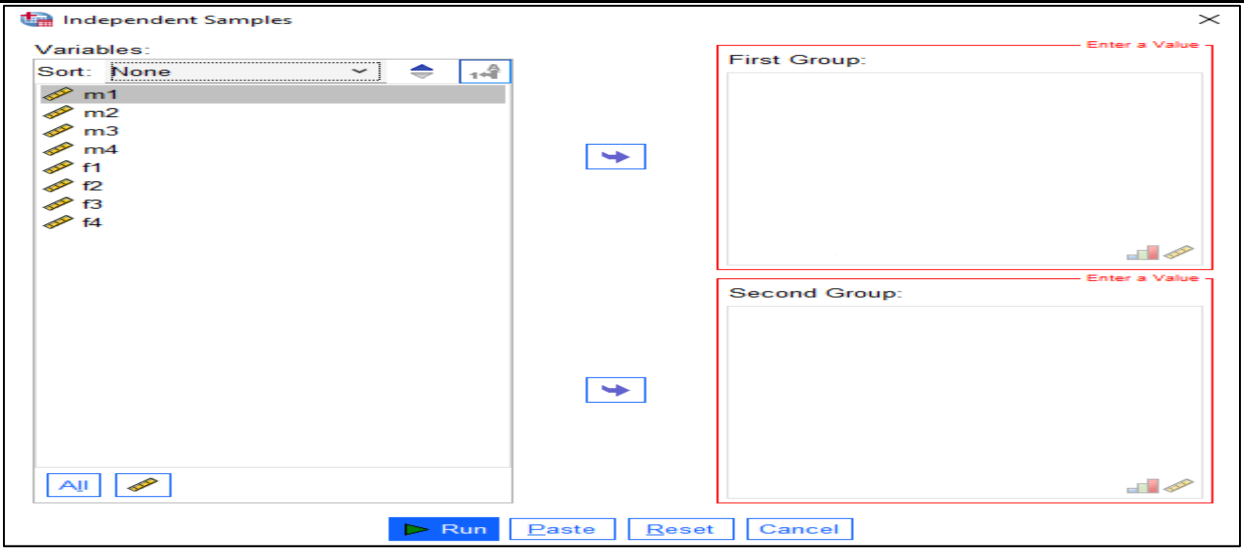

Figure (3): Dialog for Multivariate Independent Two-Sample Testing

The process is similar for the third button or dialog (node) labeled "Paired Samples," with the distinction that the samples are not independent; rather, they are related to each other over time or based on specific criteria.

5. Examples

It's now time to practically implement the dialogs by applying them to examples sourced from multivariate analysis literature.

5.1. Example of One-Sample Testing

The examples used are sourced from various multivariate statistics books. The first example is extracted from Srivastava's (2002) book, spanning pages 93 to 95. Upon utilizing the initial dialog, the outcomes are as follows:

X bar
31.250 -.750 3.125
Variance of X Matrix
1069.643 82.500 16.964
82.500 17.357 6.393
16.964 6.393 4.696
Inverse of Var-Covar Matrix
.002 -.011 .010
-.011 .194 -.223
.010 -.223 .482
Hotelling's T^2
79.064
P-Value of the T^2
0.004

The above results are the same as of the book.

The second example is taken from pages 213–214 of Johnson & Wichern's (2014) book. After using the first dialogue, the results are as follows:

X bar
8.000 6.000
Variance of X Matrix
4.000 -3.000
-3.000 9.000
Inverse of Var-Covar Matrix
.333 .111
.111 .148

Hotelling's T^2

.778

P-Value of the T^2

0.849

The results above correspond exactly to those in the book.

5.2. Example of Independent Two-Sample Testing

The initial example is drawn from the book of Rencher & Christensen (2012), covering pages 137 to 138. Upon using the second dialog, the results are as follows:

Mean vector of First Sample

15.969 15.906 27.188 22.750

Mean vector of Second Sample

12.344 13.906 16.656 21.938

VARINCE MATRIX of First Sample

5.193	4.545	6.522	5.250
4.545	13.184	6.760	6.266
6.522	6.760	28.673	14.468
5.250	6.266	14.468	16.645

VARINCE MATRIX of Second Sample

9.136	7.549	4.864	4.151
7.549	18.604	10.225	5.446
4.864	10.225	30.039	13.494
4.151	5.446	13.494	27.996

Pooled Variance Covariance Matrix

7.164	6.047	5.693	4.701
6.047	15.894	8.492	5.856
5.693	8.492	29.356	13.981
4.701	5.856	13.981	22.321

Inverse of Pooled Variance Covariance Matrix

.223	-.070	-.013	-.020
-.070	.098	-.014	-.002
-.013	-.014	.053	-.027
-.020	-.002	-.027	.067

Hotelling T^2 Test

97.601

P-Value of the Hotelling T^2

*10 ** -11 X 1.464 =0.000*

The second example is also taken from Rencher & Christensen's (2012) book, specifically from an exercise in pages 163-164. Following the use of the introductory conversation, the results are as follows:

Mean vector of First Sample

124.500 38.100 75.950 192.750 53.650 250.300

Mean vector of Second Sample

129.300 31.700 87.400 236.600 44.250 280.200

VARINCE MATRIX of First Sample

384.263	20.737	77.237	26.605	-15.921	82.895
20.737	68.200	10.374	26.868	-64.068	6.705
77.237	10.374	125.103	-33.118	4.139	-6.879
26.605	26.868	-33.118	1000.197	-24.355	-8.658
-15.921	-64.068	4.139	-24.355	322.450	-132.205
82.895	6.705	-6.879	-8.658	-132.205	470.221

VARINCE MATRIX of Second Sample

687.800	69.305	-52.179	76.968	-1.447	439.200

69.305 51.695 3.337 -46.968 -35.342 56.484
-52.179 3.337 110.884 43.853 -7.947 -111.874
76.968 -46.968 43.853 792.568 8.158 129.768
-1.447 -35.342 -7.947 8.158 173.671 116.211
439.200 56.484 -111.874 129.768 116.211 1253.747

Pooled Variance Covariance Matrix

536.032 45.021 12.529 51.787 -8.684 261.047
45.021 59.947 6.855 -10.050 -49.705 31.595
12.529 6.855 117.993 5.367 -1.904 -59.376
51.787 -10.050 5.367 896.383 -8.099 60.555
-8.684 -49.705 -1.904 -8.099 248.061 -7.997
261.047 31.595 -59.376 60.555 -7.997 861.984

Inverse of Pooled Variance Covariance Matrix

.002 -.002 -.001 .000 .000 -.001
-.002 .022 -.001 .000 .004 .000
-.001 -.001 .009 .000 .000 .001
.000 .000 .000 .001 .000 .000
.000 .004 .000 .000 .005 .000
-.001 .000 .001 .000 .000 .001

Hotelling T^2 Test

67.453

P-Value of the Hotelling T^2

.000

5.3. Example of Paired Two-Sample Testing

The first example is taken from Rencher & Christensen’s (2012) book, pages 148–149. Using the second dialogue, the outcomes are as follows:

d bar

8.000 3.067

Variance of d Matrix

121.571 17.071
17.071 21.781

Inverse of Var-Covar Matrix d

.009 -.007
-.007 .052

Hotelling's T^2

10.819

P-Value of the T^2

0.024

The second example is taken from Rencher and Srivastava's (2002) book, specifically pages 114–115. Using the third dialogue yields the following results:

d bar

1091.667 958.333

Variance of d Matrix

1416.667 83.333
83.333 4416.667

Inverse of Var-Covar Matrix d

.001 .000
.000 .000

Hotelling's T^2

6134.600000

P-Value of the T^2

*10 ** -7 X 6.632234201 = 0.000*

6. Conclusions

The developed SPSS dialogs provide an efficient means to conduct and compare univariate and multivariate one- and two-sample testing utilizing Hotelling's T2-test. These customized SPSS dialogs enable researchers to overcome certain limitations inherent in statistical software packages.

Therefore, the adoption of these dialogs is recommended for researchers lacking programming expertise, who require analysis of multivariate one- and two-sample data. It is possible that IBM may include these dialogs into next SPSS versions, which would improve the software's usability and accessibility. Moreover, increasing the range of statistical procedures that are available through SPSS dialogs would encourage researchers to use\ them, which would facilitate a wider acceptance and development of statistical analysis techniques.

References

[1] Fujikoshi, Y., Ulyanov, V. V., and Shimizu, R., (2010), Multivariate Statistics: High-Dimensional and Large-Sample Approximations, JOHN WILEY & SONS.

[2] IBM Corp., (2017), "IBM SPSS Statistics 25 Command Syntax Reference," IBM Corporation. [Online].Available: https://www.ibm.com/docs/SSLVMB_25.0.0/pdf/en/IBM_SPSS_Statistics_Command_Syntax_Reference.pdf.

[3] IBM Corp., (2020), *IBM SPSS Statistics for Windows, Armonk, NY. [Online] Available:* https://www.ibm.com/support/pages/node/1101369

[4] IBM Corp., (2023), "IBM SPSS Modeler Extensions,". [Online] Available: https://www.ibm.com/docs/en/SS3RA7_18.3.0/pdf/ModelerExtensions.pdf.

[5] Johnson, R. A., & Wichern, D. W., (2014), Applied Multivariate Statistical Analysis, Prentice Hall, 6th ed.

[6] Pardoe, I., (2020), Applied Regression Modeling, John Wiley & Sons, 3rd ed.

[7] Process, (2023), "Installing the PROCESS Custom Dialog in SPSS ,". [Online] Available: https://www.processmacro.org/uploads/3/4/8/5/34858384/dialoginstall.pdf.

[8] Raykov, T., and Marcoulides, G. A., (2008), Introduction to applied multivariate analysis, Taylor & Francis Group.

[9] Rencher, A., and Christensen, W. F., (2012), Methods of Multivariate Analysis, John Wiley & Sons, 3rd ed.

[10] Srivastava, M. S., (2002), Methods of multivariate Statistics, New York: Wiley-Interscience.

Appendix

```
TITLE  Test of Multivariate One Sample.
SET PRINTBACK=off.
Compute one=1.
Matrix.
get x
   /variables=%%First_group%%
   /missing=accept
   /sysmis=omit.
get mu
   /variables=%%Mu%%
   /missing=accept
   /sysmis=omit.
get one
   /variables=one
   /missing=accept
   /sysmis=omit.
compute n=nrow (x).
compute p=ncol (x).
```

```
compute xbar=transpos(one)*x/n.
compute xbar_mat=one*xbar.
compute x_xbar=x-xbar_mat.
compute var_x=(transpos(x_xbar)*x_xbar)/(n-1).
compute inv_var_x=inv(var_x).
compute hotelling_t2=n*transpos(transpos(xbar)-mu)*inv_var_x*(transpos(xbar)-mu).
compute p_value=SIG.F(((n-p)/((n-1)*p))*hotelling_t2,p,n-p).
   print mu / title "Mu zero".
   print n / title "Number of Rows".
   print p / title "Number of columns".
   print xbar / title "X bar" /format f8.3.
   print var_x/ title "Variance of X Matrix" /format f8.3.
   print inv_var_x/ title "Inverse of Var-Covar Matrix" /format f8.3.
   print hotelling_t2/ title "Hotelling's T2".
   print p_value/ title "P-Value of the T2".
End Matrix.
Delete variables one.
**********************************************************************
TITLE  Test of Multivariate Independent Two Samples.
SET PRINTBACK=off.
compute one=1.
Matrix.
get T1
   /variables=%%First_group%%
   /missing=accept
   /sysmis=omit.
get T2
   /variables=%%Second_group%%
   /missing=accept
   /sysmis=omit.
get one
   /variables=one
   /missing=accept
   /sysmis=omit.
COMPUTE n1=nrow(t1).
COMPUTE n2=nrow(t2).
COMPUTE p=ncol(t1).
COMPUTE T1bar=transpos(one)*t1/n1.
COMPUTE T2bar=transpos(one)*t2/n2.
COMPUTE T1_T2=T1bar-T2bar.
compute T1bar_mat=one*T1bar.
compute T2bar_mat=one*T2bar.
compute T1_T1bar=T1-T1bar_mat.
compute T2_T2bar=T2-T2bar_mat.
compute var_T1=(transpos(T1_T1bar)*T1_T1bar)/(n1-1).
compute var_T2=(transpos(T2_T2bar)*T2_T2bar)/(n2-1).
compute Sp=1/(n1+n2-2)*((n1-1)*var_T1+(n2-1)*var_T2).
compute inv_sp=inv(Sp).
compute hotelling_t2=((n1*n2)/(n1+n2))*(T1_T2*inv_sp*(transpos(T1_T2))).
compute p_value=SIG.F(((n1+n2-p-1) / ((n1+n2-2)*p))*hotelling_t2,p,n1+n2-p-1).
print n1 / title "Number of rows in sample 1".
print n2 / title "Number of rows in sample 2".
```

```
print p / title "Number of variables".
print T1bar /title"Mean vector of First Sample"/format f8.3.
print T2bar /title"Mean vector of Second Sample"/format f8.3.
print var_T1 / title "VARINCE MATRIX of First Sample"/format f8.3.
print var_T2 / title "VARINCE MATRIX of Second Sample"/format f8.3.
print Sp / title "Pooled Variance Covariance Matrix"/format f8.3.
print inv_sp/ title "Inverse of Pooled Variance Covariance Matrix"/format f8.3.
print hotelling_t2/ title "Hotelling T2 Test".
print p_value/ title "P-Value of the Hotelling T2".
End matrix.
Delete variables one.
*******************************************************************
TITLE  Test of Multivariate Paired Two Samples.
SET PRINTBACK=off.
compute one=1.
Matrix.
get T1
   /variables=%%First_group%%
   /missing=accept
   /sysmis=omit.
get T2
   /variables=%%Second_group%%
   /missing=accept
   /sysmis=omit.
get one
   /variables=one
   /missing=accept
   /sysmis=omit.
compute d = T1-T2.
compute n=nrow (d).
compute p=ncol (d).
compute dbar=t(one)*d/n.
compute dbar_mat=one*dbar.
compute d_dbar=d-dbar_mat.
compute var_d=(t(d_dbar)*d_dbar)/(n-1).
compute inv_var_d=inv(var_d).
compute hotelling_t2=n*dbar*inv_var_d*t(dbar).
compute p_value=SIG.F(((n-p)/((n-1)*p))*hotelling_t2,p,n-p).
print n / title "Number of Rows".
print p / title "Number of columns".
print dbar / title "d bar" /format f8.3.
print var_d/ title "Variance of d Matrix" /format f8.3.
print inv_var_d/ title "Inverse of Var-Covar Matrix d"/format f8.3.
print hotelling_t2/ title "Hotelling's T2" /format f8.3.
print p_value/ title "P-Value of the T2" /format f8.3.
End Matrix.
Delete variables one.
```

DOI: https://doi.org/10.31272/jae.i144.1233
Available online at: https://admics.uomustansiriyah.edu.iq/index.php/admeco

مجلة الادارة والاقتصاد
Journal of Administration & Economics

Mustansiriyah University
College of Administration & Economics

P- ISSN: 1813 - 6729
E-ISSN: 2707-1359

تنفيذ اختبارات إحصائية متعددة المتغيرات لعينة واحدة أو عينتين باستخدام SPSS Syntax

دلشاد شاكر اسماعيل بوتاني
قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة صلاح الدين-أربيل، العراق
Email: first author's email, ORCID:/ https://orcid.org/0000-0002-5465-8989

المستخلص

غالبًا ما تكون توفر اختبارات المتعددة المتغيرات للعينة الواحدة والعينتين مقيدة أو تتم دمجها مع الاختبارات الإحصائية الأخرى في العديد من التطبيقات الإحصائية. برنامج SPSS المعروف بواجهة المستخدم الصورية (GUI) سهلة الاستخدام للتحليل الإحصائي، ويوفر أيضًا لغة برمجة متخصصة تسمى SPSS Syntax. يتيح هذه اللغة تنفيذ الإجراءات الإحصائية عن طريق كتابة الأوامر للمستخدمين، مما يوفر بديلاً لواجهة المستخدم الصورية. الغرض من هذا البحث هو إنشاء مربعات حوار جديدة مكتوبة بلغة SPSS Syntax وباستخدام منشئ الحوار المخصص للملحقات. تم تصميم مربعات الحوار هذه لاختبار البيانات متعددة المتغيرات ذات العينة الواحدة أو العينتين عندما تكون ∑ غير معلومة من خلال تطبيق اختبار هوتيلينك T2، خاصة في حالة إفتقار واجهة المستخدم الصورية إلى هذه الوظائف المحددة. في حالة البيانات المكونة من عينتين، يتم استخدام طرق اختبارات متعددة المتغيرات المستقلة والمرتبطة أو والمزدوجة. تم تزويد أمثلة لتوضيح التحليلات الرئيسية التي تم إجراؤها باستخدام اختبارات T2. تم تنفيذ كتابة البرنامج لعمل التحليلات باستخدام برنامج SPSS v.27. يوصى باعتماد مربعات الحوار المطورة للباحثين الذين هم بحاجة إلى تحليل بيانات متعددة المتغيرات لعينة واحدة أو عينتين مع خبرة برمجية محدودة أو معدومة، مع إمكانية دمج مربعات الحوار بالمستقبل في إصدارات SPSS القادمة بواسطة IBM. توسيع نطاق التحليلات الإحصائية المدعومة بمربعات حوار SPSS من خلال عمل أبحاث مستقبلية تشجع استخدامها، وتدعم قبولًا أكبر لأساليب إحصائية أخرى غير موجودة في إصدارات SPSS الاعتيادية..

معلومات البحث

تواريخ البحث:

تاريخ تقديم البحث: 1 / 2 / 2024
تاريخ قبول البحث: 10 / 6 / 2024
عدد صفحات البحث 1 - 13

الكلمات المفتاحية:
متعدد المتغيرات لعينة واحدة ، متعدد المتغيرات لعينتين ، SPSS Syntax، Hotelling's T^2، منشئ الحوار المخصص.

المراسلة:
أسم الباحث: دلشاد شاكر اسماعيل بوتاني
Email: Delshd.botani@su.edu.krd