وزارة التعليم العالي والبحث العلمي

جامعة سبها- كلية العلوم

قسم- الإحصاء

بحث مقدم لاستكمال متطلبات الحصول على درجة البكالوريوس

بعنـــــــوان:

توزيع هوتلنج واستخداماته في اختبارات الفروض استخدام البرنامج (R)

إعداد الطالبة:

مني عبد الله عبد الوهاب عبد الله

تحت إشراف:

الدكتور/ السعيدي المهدي الطاهر

العام الجامعي

2017-2018

بسم الله الرحمن الرحيم

" قَالُوا سُبْحَانَكَ لاَ عِلْمَ لَنَا إِلاَّ مَا عَلَّمْتَنَا إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ"

سورة البقرة الآية 32

صدق الله العظيم

الإهـــــداء

ها هي الأيام قد مرت كلمح البصر لتصل بنا إلي نهاية هده المرحلة التعليمية بحيث لا يسعنا فيها إلا أن نهدي حصيلة جهدنا هدا إلي كل من كان سببا في مسيرتنا هده

فالي منارة العلم و المتعلمين إلي سيد التوابين والآخرين إلي الأمي الذي علم المتعلمين

(سيدنا محمد صلي الله عليه وسلم)

إلي من جرع كأسا فارغا ليسقيني قطرة حب إلي من كلت أنامله ليقدم لنا لحظه سعادة إلي من حصد الأشواك عن دربي ليمهد لي طريق العلم إلي القلب الكبير

(والدي حفضه الله)

إلي من ركع العطاء أمام قدميها وأعطتنا من دمها وروحها وعمرها حبا وتصميما ودفعا لغد أجمل إلي الغالية التي لا نرى الأمل إلا من عينيها

(إلي أمي الحبيبة)

إلي من هم اقرب إلي من روحي إلي من شاركني حضن الأمومة وبهم استمد عزوتي وإصراري

(أخواتي الأعزاء)

إلي من زرع التفاؤل في دربنا وقدم لنا الأفكار والمعلومات ربما دون إن يشعروا

(أصدقائي الأعزاء)

كلمة شكر وتقدير

لابد لنا ونحن نخطو خطواتنا الأخيرة في الحياة الجامعية من وقفة نعود بها إلي الأعوام التي قضيناها في رحاب الجامعة مع اساتذتنا الكرام الدين قدموا لنا الكثير باذلين بدلك جهودا كبير في بناء جيل الغد لتبعث الأمة من جديد

وقبل إن نمضي نتقدم باسمي آيات الشكر ولامتنان و التقدير والمحبة إلي الدين حملوا أقدس رسالة في الحياة إلي الدين مهدوا لنا طريق العلم والمعرفة إلي جميع

اساتذتنا الأفاضل

وكذلك نشكر كل من قدم لنا يد العون في إكمال هدا البحث ونخص بالذكر

الدكتور:- السعيدي المهدي الطاهر

إلي الذين كانوا عونا لنا في بحثنا هذا ونورا يضئ الظلمة التي كانت تقف أحيانا في طريقنا

أبي و أمي

فهرس المحتويات

ملخص البحث

في هدا البحت تم التطرق الي توزيع هوتلنج ودراسة خصائصه وكما بيان كيفية استخدام توزيع هوتلنج في اختبار فرضيات حول متوسط مجتمع واحد ومجتمعين في حالة متعدد المتغيرات.

تكوين جداول هوتلنج واستخدامها في اختبارات الفروض

الفصل الأول

المقدمة

1-1تمهيد:

يعتمد البحث العلمي للظواهر قيد الدراسة على تكرار محاولة دراستها بهدف التوصل إلى التفسير المنطقي لها ويتم ذلك بجمع البيانات عنها ودراستها وتحليلها. عند تحليل هذه البيانات قد يؤدي ذلك إلى تغيير فكرتنا وافتراضاتنا حول هذه الظواهر كما أن الدراسة المتكررة قد تؤدي إلى إضافة متغيرات جديدة لهذه الظواهر أو قد تؤدي إلى حذف بعض منها لذا فإن دراسة أغلب الظواهر تتطلب من الباحث جمع البيانات عن العديد من المتغيرات. فمن الصعوبات التي تواجهنا عند دراسة طرق التحليل متعدد المتغيرات هي الحاجة إلى فهم وتفسير العلاقة بين المتغيرات التي تشملها الظاهرة قيد الدراسة. كما إننا بحاجة إلى المزيد من أساليب المتقدمة لاستكشاف طرق إحصائية مطورة تساعدنا في التحليل الإحصائي للمتغيرات المتعددة.

إن جزءا كبيرا من التحليل الاحصائي للبيانات متعددة المتغيرات يهتم بالاستدلال الاحصائي أي بالتوصل إلى نتائج صحيحة بناءً على المعلومات المتاحة من العينة. وسنركز هنا على الاستدلال عن متجه متوسطات المجتمع وعناصره في حالة مجتمع واحد ومجتمعين. وعليه فإن هدفنا النهائي هو التوصل إلى تحليل إحصائي كامل حول متوسطات المجتمع باستخدام هوتلنج.

2-1أهداف البحث

يهدف هذا البحث إلى ما يلي:

1-التعريف بتوزيع هوتلنج وبيان خصائصه.

2- بيان كيفية استخدام توزيع هوتلنج في اختبار فرضيات حول متوسط مجتمع واحد ومجتمعين في حالة متعدد المتغيرات.

3-تكوين جداول هوتلنج واستخدامها في اختبارات الفروض.

3-1أهمية البحث

تكمن أهمية هذا البحث في التعرف على كفيه الاستفادة من البرامج الجاهزة في إجراء بعض استخدامات توزيع هوتلنج منها:

1-استخدام توزيع هوتلنج في اختبار الفرضية حول مجتمع واحد.

2-استخدام توزيع هوتلنج في اختبار تساوي متوسطي مجتمعين.

كما أن أهمية هذا البحث تكمن في تقديم أحد البرمجيات الإحصائية الحديثة (برنامجR)في التعامل مع التوزيع هوتلنج بدلا من استخدام الحلول اليدوية.

4-1أقسام البحث

قسم هذا البحث إلى أربعة فصول هي:

الفصل الأول: المقدمة -اهداف – اهمية البحث .

الفصل الثاني:مفاهيم أساسية في تحليل متعدد المتغيرات – توزيع وشارت – توزيع هوتلنج.

الفصل الثالث: التطبيق العملي لاختبارات الفروض حول متوسط مجتمع واحد ومجتمعين.

الفصل الرابع:الخلاصة والتوصيات.

الفصل الثاني

الاطار النظري

2-1 تمهيد

في هذا الفصل سوف نتطرق إلى بعض المفاهيم الأساسية في جبر المصفوفات - تحليل متعدد المتغيرات -التوزيع الطبيعي متعدد المتغيرات- توزيع وشارت –توزيع هوتلنج- استخدامات هوتلنج في اختبارات الفرضيات.

2-2 مفاهيم أساسية في جبر المصفوفة

- **تعريف المصفوفة**

هي عبارة عن تنظيم لمجموعة من البيانات في شكل صفوف وأعمدة (على شكل مستطيل) ويرمز لها بالحروف الكبيرة مثل A, B وقد تكون هذه البيانات رموزا أو أرقاما أو معادلات......الخ.

بعض أنواع المصفوفات

- **المصفوفة المربعة**

نقول بأن المصفوفة A مصفوفة مربعة إذا كان عدد الصفوف يساوي عدد الأعمدة فمثلا:

$$A = \begin{bmatrix} 6 & 4 & 1 \\ 2 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix}_{3*3}$$

- **المصفوفة القطرية**

المصفوفة القطرية هي مصفوفة مربعة جميع عناصرها أصفارا مع وجود على الأقل عنصراً واحدا غير صفري في القطر الرئيسي مثلا:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{3*3}$$

- **مصفوفة الوحدة**

هي مصفوفة قطرية جميع عناصر القطر الرئيسي فيها يساوي واحد فمثلاً:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3*3}$$

- **مصفوفة الوحدات**

هي مصفوفة جميع عناصرها تساوي واحد فمثلا ً:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}_{3*4}$$

- **المصفوفات المتساوية**

يقال أن المصفوفةA تساوي المصفوفة Bإذا كان كل عنصر من عناصرA يساوي العنصر المناظر له في المصفوفةB.

فمثلا ً A=B حيث

$$A = \begin{bmatrix} 5 & 4 \\ 2 & 1 \end{bmatrix}_{2*2} \quad , B = \begin{bmatrix} 5 & 4 \\ 2 & 1 \end{bmatrix}_{2*2}$$

- **المصفوفة المثلثية**

يقال أن المصفوفة A مصفوفة مثلثية عليا إذا كانت جميع العناصر أسفل القطر الرئيسي أصفارا مع وجود على الأقل عنصرا واحد ليس صفرا على القطر الرئيسي.

كمايقال أن المصفوفةAمصفوفة مثلثيه سفلى إذا كانت جميع العناصر أعلى القطر الرئيسي أصفار مع وجود علي الأقل عنصرا ليس صفرا أسفل القطر الرئيسي فمثلا:

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 5 & 0 & 0 \\ 7 & 0 & 8 \end{bmatrix}_{3*3} \quad B = \begin{bmatrix} 1 & 5 & 8 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}_{3*3}$$

مثلثية عليا　　　　　　مثلثية سفلى

- **محدد المصفوفة**

يستخدم للتعبير عن مصفوفة مربعة وهو قيمة أحاديه .فيما يلي توضيح طريقة حسابه عندما تكون المصفوفة من درجة 3*3 فمثلاً:

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}_{3*3}$$

$$\begin{bmatrix} 1 & 0 & 2 & \vdots 1 & 1 \\ 2 & 0 & 1 \vdots & 2 & 1 \\ 1 & 1 & 1 \vdots & 1 & 1 \end{bmatrix}$$

وبهذا يكون المحدد كما يلي:

$$\Delta=[(1*0*1)+(0*1*1)+(2*2*1)]$$

$$\Delta=-[(0*2*-1)+\ (1*1*1)+(2*0*1)]$$

$$[[\ 0+0+4]-[\ 0+1+0]]$$

4-1=3

- **محورة المصفوفة**

يمكن الحصول على محور المصفوفة A وذلك بجعل الصفوف أعمدة أو بجعل الأعمدة صفوف ويرمز لها بالرمز A^{T}فمثلاً

إذا كان

$$A=\begin{bmatrix} 3 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix}_{3*3}$$

فإن:

$$\mathrm{A}^{\mathbf{T}}=\begin{bmatrix} 3 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix}_{3*3}$$

- **المصفوفة المتماثلة**

يقال أن المصفوفةA متماثلة إذا كانت$A^T = A$

فمثلا:

$$A=\begin{bmatrix} 4 & 2 & 1 \\ 1 & 0 & 6 \\ 2 & 1 & 3 \end{bmatrix} \to A^T = \begin{bmatrix} 4 & 2 & 1 \\ 1 & 0 & 6 \\ 2 & 1 & 3 \end{bmatrix}_{3*3}$$

- **أثر المصفوفة**

يعرف أثر المصفوفة A المربعة على أنه مجموع عناصر القطر الرئيسي أن:

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

فمثلا إذا كانت

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 6 & -5 \\ 4 & 0 & 3 \end{bmatrix}$$

فإن أثر المصفوفة هو:

$$tr(A) = 1 + 6 + 3 = 10$$

- **معكوس المصفوفة**

إذا كان لدينا المصفوفةA فإن معكوس هذه المصفوفة يرمز لها بالرمز A^{-1} حيث:

$AA^{-1} = A^{-1}A = I$،حيث يشرط في وجود المعكوس أن يكون المحدد لا يساوي صفر (تامة الرتبة).

فإذا كانت المصفوفة من الدرجة (2*2) فإنه لإيجاد المعكوس علينا أولا إيجاد المحدد.مثلا لإيجاد معكوس المصفوفةAحيث إن:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

فإن المحدد $\Delta = (1)(1) - (-1)(2) = 1 + 2 = 3 \neq 0$ ومنها $A^{-1} = \frac{1}{\Delta} adj(A)$:

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{-2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

أما إذا كانت المصفوفة ذات رتبة (3*3)فإن المصفوفة تمثل كالآتي:

$$A = \begin{bmatrix} 2 & 6 & 4 \\ 5 & 30 & 15 \\ 6 & 3 & 9 \end{bmatrix}$$

علينا إيجاد المحدد أولا وباستخدام طريقة الأسهم فإن:

$$\begin{bmatrix} 2 & 6 & 4 & \vdots & 2 & 6 \\ 5 & 30 & 15 & \vdots & 5 & 30 \\ 6 & 3 & 9 & & 6 & 3 \end{bmatrix}$$

$$\Delta = [(2*30*9) + (6*15*6) + (4*5*3)]$$
$$-[(4*30*6) + (2*15*3) + (6*5*9)]$$

$$\Delta = [540 + 540 + 60] - [720 + 90 + 270] = 1140 - 1080 = 60$$

عليه فإن المعكوس كالتالي:

$$A^{-1} = \frac{1}{\Delta} adj(A)$$

$$A^{-1} = \frac{1}{60}\begin{pmatrix} +\begin{vmatrix} 30 & 15 \\ 3 & 9 \end{vmatrix} & -\begin{vmatrix} 5 & 15 \\ 6 & 9 \end{vmatrix} & +\begin{vmatrix} 5 & 30 \\ 6 & 3 \end{vmatrix} \\ -\begin{vmatrix} 6 & 4 \\ 3 & 9 \end{vmatrix} & +\begin{vmatrix} 2 & 4 \\ 6 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 6 \\ 6 & 3 \end{vmatrix} \\ +\begin{vmatrix} 6 & 4 \\ 30 & 15 \end{vmatrix} & -\begin{vmatrix} 5 & 15 \\ 6 & 9 \end{vmatrix} & +\begin{vmatrix} 2 & 6 \\ 5 & 30 \end{vmatrix} \end{pmatrix}$$

$$A^{-1} = \frac{1}{60}\begin{pmatrix} +(30*9-3*15) - (5*9-6*15) + (5*3-6*30) \\ -(6*9-3*4) + (2*9-6*4) - (2*3-6*6) \\ +(6*15-30*4) - (5*9-6*15) + (2*30-5*6) \end{pmatrix}$$

$$A^{-1} = \frac{1}{60}\begin{pmatrix} 225 & -45 & -165 \\ -42 & -6 & 30 \\ -30 & 45 & 30 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} \frac{15}{4} & -\frac{3}{4} & -\frac{11}{4} \\ -\frac{7}{10} & -\frac{1}{10} & \frac{1}{2} \\ -\frac{1}{2} & \frac{3}{4} & \frac{1}{2} \end{pmatrix}$$

- **الجذور المميزةو المتجهات المميزة**

إذا كان A مصفوفة مربعة فإن جذورهاالمميزة λ تحقق

$$|A-\lambda| = 0$$

$$A=\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

$$\Delta = \begin{vmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix}$$

$$(4-\lambda)(3-\lambda)\_(2)(1)=0$$

$$\lambda^2 -7\lambda +12-2=0$$

$$\lambda^2 -7\lambda +10=0$$

A=1 B=-7 C=10

$$\lambda = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

$$\lambda = \frac{(7)+\sqrt{7^2-40}}{2}$$

$$\lambda_1 = \frac{7+\sqrt{9}}{2} = \frac{7+3}{2} = 5$$

$$\lambda_2 = \frac{(7) - \sqrt{7^2 - 40}}{2} = \frac{7 - \sqrt{9}}{2} = \frac{7 - 3}{2} = 2$$

مجموع جذور المميزة يساوي أثر المصفوفة

$$\lambda_1 + \lambda_2 = 5 + 2 = 7$$

حاصل ضرب الجذور يساوي محدد المصفوفة

$$\lambda_1 * \lambda_2 = 5 * 2 = 10$$

إيجاد متجه المميزة 5 فمثلا:

$$\begin{bmatrix} 6-5 & 2 \\ 1 & 7-5 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$X_1 + 2X_2 = 0$$

$$3X_1 + 2X_2 = 0$$

نأخذ المعادلة الأولى

$$X_1 + 2X_2 = 0$$

نفرض أن $X_1 = 2$

$$X_1 = 2 \quad X_2 = -1$$

نجد متجه

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

نلاحظ عدد لانهائي من الحلول في متجهات المصفوفة المربعة لحساب ما يعرف بالمتجه المميز

$$\begin{bmatrix} \frac{2}{\sqrt{(2)^2+(-1)^2}} \\ \frac{-1}{(2)^2+(-1)^2} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} \end{bmatrix}$$

2-3مفاهيم أساسية في تحليل متعدد المتغيرات

- **مصفوفة البيانات(Data Matrix)**

بفرض أن لدينا عددnمن المشاهدات لكل متغير منp من المتغيرات $(p>1)$فإنه يمكن كتابة البيانات المتاحة لدينا كما يلي:

	المفردة رقم 1	المفردة رقم2		المفردة رقمj		المفردة رقمn
المتغير رقم 1 :	x_{11}	x_{12}	$\dots$	x_{1j}	$\dots$	x_{1n}
المتغير رقم 2 :	x_{21}	x_{22}	$\dots$	x_{2j}	$\dots$	x_{2n}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
المتغير رقم i	x_{i1}	x_{i2}	$\dots$	x_{ii}	$\dots$	x_{in}
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
المتغير رقم p :	x_{p1}	x_{p2}	$\dots$	x_i	$\dots$	x_{pn}

كذلك يمكن عرض البيانات في صورة تنظيم يرمز له برمزXويتكون من عددPمن الصفوف وعددnمن الأعمدة ويأخذ الشكل التالي:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2n} \\ & & \vdots & & & \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{in} \\ & & \vdots & & & \\ x_{p1} & x_{p2} & \dots & x_{pi} & \dots & x_{pn} \end{bmatrix}_{P*n}$$

- **متجه المتوسطات(Mean Vector)**

يرمز له بالرمز $(\overline{X})$ حيث أن: $\overline{X} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_p \end{pmatrix}_{p*1}$

حيث أن:

$$\overline{X}_1 = \frac{1}{n}\sum_{j=1}^{n} X_{1j}$$

$$\overline{X}_2 = \frac{1}{n}\sum_{j=1}^{n} X_{2j}$$

$$\vdots$$

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$$

$$\vdots$$

$$\overline{X}_p = \frac{1}{n}\sum_{j=1}^{n} X_{pj} \quad , j = 1,2, \dots . n \quad ; \ i = 1,2, \dots . p$$

ويمكن التعبير عن متجه المتوسطات بدلالة المصفوفة (X)كمايلي:

$$\bar{\mathbf{X}} = \frac{1}{\text{n}}\text{X}1 \Rightarrow \text{where} \quad 1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n*1}$$

- **مصفوفة التباينات والتباينات المشتركة (Covariance Matrix)**

في حالة المجتمع يرمز لهذه المصفوفة بالرمز(∑) حيث أن:

$$\sum = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}_{p*p}$$

أما في حالة العينة فإن هذه المصفوفة يرمز لها بالرمز(S)حيث أن:

$$\text{S} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix}_{p*p}$$

حيث أن:

$$\text{S}_{\text{ii}} = \frac{1}{\text{n}}\sum_{\text{j}=1}^{\mathbf{n}}(\text{x}_{\text{ij}} - \bar{\text{x}}\text{i})^2 \quad \text{i} = 1,2 \ldots\ldots \text{p}$$

$$\text{s}_{\text{ik}} = \frac{1}{\text{n}}\sum_{\text{j}=1}^{\text{n}}\left(\text{x}_{\text{ij}} - \bar{\text{X}}_{\text{i}}\right)(\text{x}_{\text{kj}} - \bar{\text{x}}_{\text{j}}), \qquad \text{i} = 1,2 \ldots\ldots \text{p} \quad \text{k} = 1,2, \ldots\ldots \text{p}$$

ومن خصائص هذه المصفوفة:

1- أنها مصفوفة مربعة أي من الدرجة $\text{P} * \text{P}$.

2- أنها مصفوفة متماثلة.

3- جميع عناصر القطر الرئيسي تكون موجبة.

وتكتب المصفوفة S بالشكل التالي :-

$$S = \frac{1}{n}XX' - \overline{X}\overline{X}'$$

$$S = \frac{1}{n}X'_{p*n}H_{n*n}X_{n*p}$$

حيث: $H = \left(I - \frac{1}{n}11'\right)$

- **مصفوفة الارتباط (Correlation Matrix)**

تعرف هذه المصفوفة رياضيا بالشكل $R = (r_{ik})$ حيث $i = 1,2,\ldots p \;\; , k = 1,2,\ldots .p$

وهي من الدرجة $P * P$ حيث r_{ik} تشير إلى معامل الارتباط البسيط بين المتغيرين X_i و X_k ويمكن أن تكتب كما يلي:

$$R = \begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1p} \\ r_{21} & r_{22} & \ldots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \ldots & r_{pp} \end{pmatrix}_{p*p}$$

ومن خصائص هذه المصفوفة هي

1- أنها مربعة أي من الدرجة $P * P$.

2- أنها متماثلة أي أن $R = R'$.

3-جميع عناصر القطر الرئيسي تساوي واحد.

ويمكن حساب معامل الارتباط r_{ik} كما يلي :

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}S_{kk}}} \quad = \frac{S_{ik}}{S_{ik}S_{ik}} \qquad i = 1,2 \ldots .. p \;\; , k = 1,2 \ldots . p$$

حيث أن:

S_{ii}: تمثل قيمة التباين للمتغير رقم i.

S_{kk}: تمثل قيمة التباين للمتغير رقم k.

S_{ik}: تمثل قيمة التباين المشترك بين المتغير iو k .

- **المتجه العشوائي(Random vector)**

يسمى التنظيم Xالمؤلف من p من المتغيرات العشوائية بالمتجه العشوائي ويكتب كما يلي:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_P \end{pmatrix}_{p*1}$$

توزيع هذا المتجه يسمى بالتوزيع المتعدد حيث أن متوسطة:

$$E(X) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_P) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_P \end{pmatrix} = \mu$$

أما تباينه يعرف كما يلي:

$$V(X) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1P} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{P1} & \sigma_{P2} & \dots & \sigma_{PP} \end{pmatrix}$$

وعند التعامل مع العينة فإن المتوسطات يرمز لها بالرمز له $\overline{X}$ ومصفوفة التباينات بالرمز S.

2-4التوزيع الطبيعي متعدد المتغيرات

يعتبر التوزيع الطبيعي متعدد المتغيرات من أشهر التوزيعات المتعددة لعدة أساب منها:

1- هذا التوزيع يعتبر تعميم للتوزيع الطبيعي الأحادي وهذا لا يحدث في التوزيعات الأخرى.

2-عدد المعالم التي نحتاج تقديرها في التوزيع الطبيعي متعدد المتغيرات أقل من عدد المعالم التي نحتاج لتقديرها في التوزيعات المتعددةالأخرى.

3-في حالة التوزيع الطبيعي متعدد المتغيرات إذا كان الارتباط بين المتغيرات يساوي صفر فهذا يعني وجود استقلالية تامة بين المتغيرات وهذه الخاصية غير موجودة في التوزيعات المتعددة الأخرى.

4-كل التوزيعات تؤول للتوزيع الطبيعي كلما كبر حجم العينة.

5-متوسط قيم العينات المتعددةيأخذ شكل التوزيع الطبيعي ولو لم يكن التوزيع لمتغير نفسه يتبع التوزيع الطبيعي .

في التوزيع الطبيعي متعدد المتغيرات تكون دالة كثافة الاحتمال كما يلي:

$$f(x) = \frac{1}{(2\Pi)^{p/2} \left| \Sigma \right|^{1/2}} e^{-1/2(x-\mu)'\Sigma^{-1}(x-\mu)}$$

ويكتب$X \sim N_P.(\mu, \Sigma)$..حيث μ هو متجه المتوسطات ، Σ هي مصفوفة التباينات وتسمى(μ, Σ) معالم التوزيع.

خواص التوزيع الطبيعي متعدد المتغيرات

1- إذا كان()()$X \sim N_P(\mu, \Sigma)$

$$E(X) = \mu$$

$$V(X) = \Sigma$$

2-الدله المولدة للعزوم تكتب بالصيغة التالية:

$$M_X(t) = e^{t\mu + \frac{1}{2}t'\Sigma t}$$

3-الدله المميزة تكتب بالصيغة التالية:

$$\emptyset_x(t) = e^{it'\mu - \frac{1}{2}t'\Sigma t}$$

2-5توزيع وشارت

إذا كان $X \sim N_P(\mu\ ,\Sigma)$وكان

$$\mathrm{M} = \mathrm{XHX'}$$

حيث أن H عباره عن مصفوفة متماثلة حيث $\mathrm{H} = \left(\mathrm{I} - \frac{1}{\mathrm{n}} 11'\right)$.

فإن المصفوفة M تتبع توزيع وشارت بالمعلمتين Σ, n وتكتب$M \sim W_P(\Sigma\ , n)$ ولها دالة كثافة احتمال معرفة بالشكل التالي:

$$\mathrm{f(M)} = \frac{|\mathrm{M}|^{\frac{\mathrm{n-p-1}}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1} M)}}{2^{\frac{\mathrm{np}}{2}} \pi^{\frac{\mathrm{p(p-1)}}{4}} |\Sigma|^{\frac{\mathrm{n}}{2}} \prod_{\mathrm{i=1}}^{\mathrm{p}} \Gamma\left(\frac{\mathrm{n+1-1}}{2}\right)}$$

حقائق ونظريات

1- إذا كان$X_{n*p} \sim \mathrm{N_p}(0, \Sigma)$ فان المصفوفة$\mathrm{M} = \mathrm{XX'}$ تتبع توزيع وشارت

$$\mathrm{M} \sim \mathrm{W_P}(\Sigma\ , \mathrm{n})$$

2-إذا كان$\mathrm{X_{n*p}} \sim \mathrm{N}(0, \mathrm{I})$فإن المصفوفة$\mathrm{M} = \mathrm{XX'}$تتبع توزيع وشارت$\mathrm{M} \sim \mathrm{W_P}(\mathrm{I}, \mathrm{n})$.

3-إذا كان$\mathrm{X_{n*p}} \sim \mathrm{N}(0, \mathrm{I})$ فإن المصفوفة$\mathrm{M} = (\mathrm{X} - \mu)(\mathrm{X} - \mu)$تتبع توزيع وشارت

$$\mathrm{M} \sim \mathrm{W_P}(\sum\ , \mathrm{n}).$$

4 - إذا كان $\mathrm{M} \sim \mathrm{W_P}(\Sigma, \mathrm{n})$ فإن $\Sigma^{-\frac{1}{2}} \mathrm{M} \Sigma^{-\frac{1}{2}} \sim \mathrm{W_P}(\mathrm{I}, \mathrm{n})$.

5- إذا كان متغيرين عشوائيين $M_1 \sim W_p(\Sigma, n_1)$ و $M_2 \sim W_P(\Sigma\ , n_2)$ نحصل على $M_1 +$ $M_2 \sim W_P(\Sigma\ , n_1 + n_2)$

6- الدالة المميزة لتوزيع وشارت $M \sim W_P(I, m)$.

$$\emptyset_{\mu}(t) = \frac{1}{|I-ik|^{\frac{n}{2}}} \quad ; K = (k_{ij})$$

$$i = \sqrt{-1}$$

$$k_{ij} = (1 + \delta_{ij})t_{ij}$$

$$\delta_{ij} = \begin{cases} 1 & ; i = j \\ 0 & ; i \neq j \end{cases}$$

2-6توزيع هوتلنج

تم اكتشاف هذا التوزيع من قبل العالم هوتلنج عام 1933وهذا التوزيع يناضر توزيع t في حالة التحليل الأحادي وله استخدماته عديدة.

تعريف:

إذا كان المتجه$\underline{d} \sim N_P(0, I)$ و$M \sim W_P(I, n)$ وكلاهما مستقل عن الآخر فإن المقدار $d\ nd'M^{-1}$له توزيع هوتلنج ويكتب$nd'M^{-1}d \sim T^2(p, n)$.

نظريات وحقائق:

1-إذا كان المتجه$M \sim W_P(\Sigma, n)$; $X \sim N_P(\mu, \Sigma)$فإن:

$$n(X - \mu)'M^{-1}(X - \mu) \sim T^2(p, n)$$

2-علاقة توزيع هوتلنج بتوزيعF

$$T^2(P, n) = \frac{np}{n - p + 1} F_{(P, n-p+1)}$$

ومنها فإن التوقع هو

$$E(T^2_{(P,n)}) = E(F_{(p,n-p+1)} \frac{np}{n+p+1}$$

والتباين

$$V(F_{(p,n-p+1)}) = V(T^2_{(p,n)}) = \frac{n^2 P^2}{(n - p + 1)^2}$$

3-معامل الالتواء$A\sim W_p(\Sigma, n)$

$$\frac{a'\Sigma^{-1}a}{a'A^{-1}a}\sim x^2(n-p+1)$$

حيث أنA, aمستقلان.

4- إذا كان$X\sim N_P(\mu, \Sigma)$فإن

$$\frac{n-p}{p}(\bar{X}-\mu)'S^{-1}(\bar{X}-\mu)\sim F_{(p,n-p)}$$

2-7استخدامات هوتلنج في اختبارات الفروض

1-اختبار الفرضية حول متوسط المجتمع

$$\mathbf{H_0}: \underline{\boldsymbol{\mu}} = \underline{\boldsymbol{\mu_0}}$$

إذا كان $X\sim N_P(\mu, \Sigma)$ وكانS_uهو التقدير غير متحيز ل Σ وكان $\bar{X}$ مقدر μ فإنه يمكن استخدام إحصاءه هوتلنج في اختبار الفرض أعلاه حيث تعرف الإحصاءة بالشكل التالي:

$$T_c^2 = n\left(\bar{X}-\mu_0\right)'S_u^{-1}(\bar{X}-\mu_0)$$

حيث:

$$T_c^2\sim T^2(p, n-1)$$

وتتم مقارنةT_c^2 بالجدولية المناظرة لها حيث:

$$T_{(p,n-1)} = \left[\frac{(n-1)p}{(n-p)}F_{(p,n-p,\alpha)}\right]$$

ويتم رفض H_0 إذاكان $T_c^2 \geq T^2_{(p,n-1,\alpha)}$

2-اختبار تساوي متوسطي مجتمعين:

$$H_0 : \underline{\boldsymbol{\mu}_1} = \underline{\boldsymbol{\mu}_2}$$

حيث أن:- μ_1متوسط المجتمع الأول،μ_2متوسط المجتمع الثاني.

وإحصاءه الاختبار المستخدمة كالتالي:

$$T_c^2 = \frac{n_1 n_2}{n_1 + n_2 - 2}(\bar{x}_1 - \bar{x}_2)' S_u^{-1} (\bar{x}_1 - \bar{x}_2) \sim T^2_{(p, n_{1+n_2-2})}$$

ولحساب القيمة الجدولية

$$T_c^2 = \frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - p - 1)} F_{p, n_1 + n_2 - p - 1, (\alpha)}$$

والقرار هو الرفض إذا كانت القيمة المحسوبة أكبر من القيمة الجدولية

$$S_u = \frac{n_1 s_1 + n_2 s_2}{n_{1+} n_2 - 2}$$

حيث أن:

$\bar{X}_1$:متجه متوسطات العينات المسحوبة من المجتمع الأول.

$\bar{X}_2$: متجه متوسطات العينات المسحوبة من المجتمع الثاني.

3- تحليل الملامح العامة:

يستخدم تحليل الملامح العامة "profile analysis"في المواقف التي يتم فيها تطبيقp من المعالجات (اختبارات ، أسئلة....) على مجموعتين أو أكثر من المفردات حيث يفترض أن استجابات المجموعات المختلفة مستقلة عن بعضها البعض كما يتطلب وجوب التعبير عن جميع الاستجابات بنفس وحدات القياس. من المعتاد هنا أن نضع السؤال التالي:هل جميع عناصر متجهات المتوسطات متساوية؟ في تحليل الملامح العامة هذا السؤال إلى عدد من إمكانات المحددة.

بالنسبة لمتجه المتوسطات$\mu_1 = [\mu_{11}, \mu_{12}, \mu_{13}, \mu_{14}]$الذي يمثل متوسط استجابات المجتمع الأول لأربعة معالجات. يمكننا توقع هذه المتوسط تباينيا وتوصلها بخطوط مستقيمة ويسمى شكل الخطوط المنكسرة هذا باسم خطوط الملامح العامة للمجتمع الأول.وبالطبع يمكن إنشاء خطوط الملامح العامة لكل مجتمع (مجموعة) وسنقوم هنا بالتركيز على مجموعتين فقط.

افترض أن المتجه$\mu'_1 = [\mu_{11}, \mu_{12}, \ldots . \mu_{1p}]$ والمتجه$\mu'_2 = [\mu_{21}, \mu_{22}, \ldots . \mu_{2p}]$

هما متجهي متوسطات استجابات المجتمعين الأول و الثاني على التوالي لعدد p من المعالجات.

في تحليل الملامح العامة للمجتمعين يمكن صياغة السؤال الخاص بالتساوي في خطوات متتالية كما يلي:

1-هل خطوط الملامح العامة للمجتمعين متوازية ؟أي هل فرض العدم

$$H_{01}: \mu_{1i} - \mu_{1i-1} = \mu_{2i} - \mu_{2i-1} \quad i = (2,3,4, \ldots\ldots p)$$

فرض مقبول؟

2-بافتراض إن خطوط الملامح العامة متوازية هل هي متطابقة أي هل هي فرض العدم؟

$H_{02}: \mu_{1i} = \mu_{2i}$ $(i = 1,2 \ldots p)$ فرض مقبول؟

3-بافتراض أن الخطوط متطابقة هل هي على استقامة واحدة بموازاة المحور الأفقي؟ أي هل جميع المتوسطات تساوي نفس المقدار الثابت ؟أي هل متغير فرض العدم

$H_{03}: \mu_{11} = \mu_{12} = \cdots = \mu_{1p} = \mu_{21} = \mu_{22} = \cdots = \mu_{2p}$ فرضا مقبولا؟

ويمكن كتابة فرض العدم الموجود

$$H_{01}: C_{\mu 1} = C_{\mu 2}$$

حيث تشيرcإلى مصفوفة المتغايرات.

$$C_{(p-1)(p)} = \begin{bmatrix} -1 & 1 & 0 & 0 & \ldots. & 0 & 0 \\ 0 & -1 & 1 & 0 & \ldots. & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots.. & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots.. & -1 & 1 \end{bmatrix}$$

اختبار توازى خطوط الملامح العامة لمجتمعين معتدلين

يرفض فرض العدم $H_{01}: C\mu_1 = C\mu_2$ توازي خطوط الملامح العامة عند مستوى معنويα إذا كان

$$T^2 = (\bar{x}_1 - \bar{x}_2)'C'\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) CS_{pooled}C'\right]^{-1} C(\bar{x}_1 - \bar{x}_2) > c^2$$

حيث

$$c^2 = \frac{(n_1 + n_2 - 2)(P - 1)}{n_1 + n_2 - p} F_{p-1.n_1+n_2-p}, (\alpha)$$

وعندما تكون خطوط الملامح العامة للمجتمعين متوازية. فإن خطوط الملامح العامة للمجتمع الأول قد تكون الخطوط المناظرة للمجتمع الثاني ($\mu_{1i} > \mu_{2i}$لجميع قيم i) وقد يحدث العكس.

ففي حالة تطبيق خطوط الملامح العامة للمجتمعين فقط إذا كان مجموع الارتفاعات

$$\mu_{11} + \mu_{22} + \cdots + \mu_{1p} = 1'\mu_1$$

مساويا لمجموع الارتفاعات$\mu_{11} + \mu_{22} + \cdots + \mu_{2p} = 1'\mu_2$ لذا يمكن كتابة فرض العدم في الخطوة الثانية كما في الصيغة التالية:

$$H_{02}: 1'\mu_1 = 1'\mu_2$$

وبالتالي يمكننا اختبار فرض العدم H_{02}باستخدام اختبار t لعينتين مستقلتين بناء على المشاهدات$1'x_{2j}$ n_2 .... ,$j = 1,2$والمشاهدات$1'x_{1j}$ n_1 .... ,$j = 1,2$.

اختبار تطابق خطوط الملامح العامة علما بأن ا متوازية

بالنسبة لمجتمعين معتدلين يرفض فرض العدم $H_{02}: 1'\mu_1 = 1'\mu_2$(خطوط متطابقة)عند مستوي معنويα إذا كانت

$$T^2 = 1'(\overline{X}_1 - \overline{X}_2)\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)1'S_{pooled}1\right]^{-1} 1'(\overline{X}_1 - \overline{X}_2)$$

$$= \left(\frac{1'\overline{X}_1 - \overline{X}_2}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)1'S_{pooled}1}}\right)^2 > t^2_{n_1+n_2-2}\left(\frac{\alpha}{2}\right) = F_{1.n_1+n_2-2}(\alpha)$$

وإذ تطابقت خطوط الملامح العامة فإن المشاهدات x_{11}،x_{12} ... x_{1n_1}

والمشاهداتx_{21}،x_{22} ... x_{2n_2} جميعها مسحوبة من المجتمع المعتدل.

وفي الخطوة التالية نختبر ما إذا كان لجميع المتغيرات نفس الوسط الحسابي. أي ما إذا كانت

خطوط الملامح العامة لهذا المجتمع المعتدل تقع على استقامة واحدة بموازاة المحور الأفقي ويمكن صياغة فرض العدم في خطوة التالية كما يلي:

$$H_{03}: C(\mu_1 + \mu_2) = 0$$

حيثCهي المصفوفة. وإذا قبلنا فرض العدم فإنه يتم تقدير متجه المتوسطات العامة باستخدام

$$\overline{X} = \frac{\sum_{j=1}^{n_1} x_{1j} + \sum_{j=1}^{n_2} x_{2j}}{n_1 + n_2} = \frac{n_1}{(n_1+n_2)}\overline{X}_1 + \frac{n_2}{(n_{1+}n_2)}\overline{X}_2$$

اختبار وقوع خطوط الملامح العامة على استقامة واحدة وبموازاة المحور الافقي علما بانها متطابقة.

بالنسبة لمجتمعين معتدلين يرفض فرض العدم $H_3: C(\mu_1 + \mu_2) = 0$ (خطوط علي استقامة واحدة وبموازاة المحور الافقي) عند مستوي معنوي α اذا كانت

$$T^2 = (n_1 + n_2)\overline{X}'C'\left[CS_{pooled}C'\right]^{-1}C\overline{X} > F_{p-1,n_1+n_2-p\prime(\alpha)}$$

الفصل الثالث

الجانب التطبيقي

3-1تمهيد:

يهدف هذا الفصل إلى بيان كيفية تطبيق توزيع هوتلنج واستخداماته في اختبارات الفروض المتعلقة بمتوسط مجتمع واحد وتساوي متوسطي مجتمعين وأخيرا الملامح العامة.

1.اختبار الفرضية المتعلقة بمتوسط مجتمع واحد

مثـال (1)

إذا كان لدينا البيانات الآتية اختبر الفرضية

$$H_0: \mu = (9,5)'$$

$$H_1: \mu \neq (9,5)'$$

X1	X2
22.0	**11.0**
24.5	**13.5**
23.1	**23.5**
24.5	**4.5**
12	**8.5**
14.5	**9.6**
10	**24**
14.6	**23**
13.6	**20**
27.0	**27**
20	**26**
23	**25.5**

الحل:

من واقع البيانات أعلاه يمكن الحصول على:

$$\overline{X}_1 = \frac{\sum_{i=1}^{12} X_{1i}}{12} = \frac{228.8}{12} = 19.06$$ وهو متوسط المجتمع الأول

$$\overline{X}_2 = \frac{\sum_{i=1}^{12} X_{2i}}{12} = \frac{216.1}{12} = 18.00$$ وهو متوسط المجتمع الثاني

متوسطات المجتمع هو

$$\overline{X} = \begin{pmatrix} 19.06 \\ 18.00 \end{pmatrix}$$

ولتكوين مصفوفة المجتمع نحتاج لحساب مايلي:

$$S_{11} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{12} X_{1i}^2\right) - \left((\sum_{i=1}^{12} X_{1i})^2 \big/ n\right)\right]$$

$$= \frac{1}{11}[4728.48 - (288.8)^2)/12] = 33.275$$

$$S_{22} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{12} X_{2i}^2\right) - \left((\sum_{i=1}^{12} X_{2i})^2/n\right)\right]$$

$$= \frac{1}{11}[4600.4 - (216.1)^2/12] = 64.437$$

$$S_{12} = S_{21} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{12} X_{1i}X_{2i}\right) - \left(\sum_{i=1}^{12} X_{1i}\right)\left(\sum_{i=1}^{12} X_{2i}\right)\big/ n\right]$$

$$= \frac{1}{11}[4150.35 - (288.8)(216.1)/12] = 2.731$$

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} 33.275 & 2.731 \\ 2.731 & 64.437 \end{bmatrix}$$

$$S^{-1} = \frac{1}{(33.275)(64.437) - (2.731)(2.731)}\begin{bmatrix} 64.437 & -2.731 \\ -2.731 & 33.275 \end{bmatrix}$$

$$= \frac{1}{2136.719}\begin{bmatrix} 64.437 & -2.731 \\ -2.731 & 33.275 \end{bmatrix}$$

$$S^{-1} = \begin{bmatrix} 0.03015 & -0.00127 \\ -0.00127 & 0.01557 \end{bmatrix}$$

$$T_c^2 = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0)$$

$$= 12[19.06 - 9, 18.00 - 5]' \begin{bmatrix} 0.03015 & 0.00127 \\ 0.00127 & 0.01557 \end{bmatrix} \begin{bmatrix} 19.06 - 9 \\ 18.00 - 5 \end{bmatrix}$$

$$= 12[10.06, 13]' \begin{bmatrix} 0.286765372 \\ 0.189591613 \end{bmatrix}$$

$$= 12(32.884859642 + 2.464690969) = 64.19460733$$

$$T_c^2 = 64.19460733$$

ولحساب القيمة الجد ولية

$$T^2 = \frac{P(n-1)}{n-P} F_{p,n-p} = \frac{2 * 11}{10} \; F_{2,10,0.05}$$

$$= \frac{22}{10} F_{4.10} = 9.02$$

نجد أن $T^2 = 64.19 > 9.02$ وبالتالي فإننا نرفض فرض العدم في مقابل الفرض البديل عند مستوى معنوية $\alpha = 0.05$.

الحــل:

تطبيق قاعدة توزيع هوتلنج واستخداماته في اختبار الفرضية باستخدام برنامج R

للحصول على نتائج الاختبار نستخدم الدالة T.test (انظر الملحق)

حيث تكون المدخلات X هي:

```
X=cbind(x1=c(22.0,24.5,23.1,24.5,12,14.5,10,14.6,13.6,27.0,20,23),x2=c(
11.0,13.5,23.5,4.5,8.5,9.6,24,23,20,27,26,25.5))

T.test(X,mu=c(9,5))
   T2        Fstat      df1      df2      p.value
 64.27859   29.21754    2        10       0.000066
```

إحصاءه هوتلنج تساوي 64.27859بينما عدد درجات الحرية (2,10)

ونلاحظ أن قيمة p-valve =0.000066 أقل من قيمة$\alpha = 0.05$.

إذاًالاختبار معنوي ذو دلالة إحصائية وقيمةp-valve أقل من α

∴بما أن القيمة المحسوبة أكبر من قيمة p-valve نرفض فرض العدم.

مثـال (2)

جمعت بيانات عن ثلاثة متغيرات خاصة بإفراز العرق X_1, X_2, X_3من 20 فتاة تتمتع بصحة

جيده حيث يشير X_1 إلى معدل إفراز العرق.X_2 إلى كمية الصوديوم.X_3 إلى كمية البوتاسيوم هذه البيانات سميت ببيانات إفراز العرق

بيانات إفراز العرق

رقم المشاهدة	معدل إفراز العرق	كمية الصوديوم	كمية البوتاسيوم
1	3.7	48.5	**9.3**
2	5.7	65.1	**8.0**
3	3.8	47.2	**10.9**
4	3.2	53.2	**12.0**
5	3.1	55.5	**9.7**
6	4.6	36.1	**7.9**
7	2.4	24.8	**14.0**

8	7.2	33.1	**7.6**
9	6.7	47.4	**8.5**
10	5.4	54.1	**11.3**
11	3.9	36.9	**12.7**
12	4.5	58.8	**12.3**
13	3.5	27.8	**9.8**
14	4.5	40.2	**8.4**
15	1.5	13.5	**10.1**
16	8.5	56.4	**7.1**
17	4.5	71.6	**8.2**
18	6.5	52.8	**10.9**
19	4.1	44.1	**11.2**
20	5.5	40.9	**9.4**

المطلوب اختبار فرض العدم ضد الفرض البديل

$$H_0 : \mu = [4,50,10]'$$

$$H_1 : \mu \neq [4,50,10]'$$

وذلك عند مستوى معنوي $\propto = 0.10$

الحـل:

متوسط المجتمع الأول $\overline{X}_1 = \frac{\sum_{i=1}^{20} X_{1i}}{20} = \frac{92.8}{20} = 4.640$

متوسط المجتمع الثاني $\overline{X}_2 = \frac{\sum_{i=1}^{20} X_{2i}}{20} = \frac{908}{20} = 45.400$

متوسط المجتمع الثالث $\overline{X}_3 = \frac{\sum_{i=1}^{12} X_{3i}}{20} = \frac{199.3}{20} = 9.965$

$$\overline{X} = \begin{bmatrix} 4.640 \\ 45.400 \\ 9.965 \end{bmatrix}$$

$$S_{11} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{1i}^2\right) - \left((\sum_{i=1}^{20} X_{1i})^2 \Big/ n\right)\right]$$

$$= \frac{1}{19}[485.3 - (92.8)^2/20] = 2.879$$

$$S_{12} = S_{21} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{1i}X_{2i}\right) - \left(\sum_{i=1}^{20} X_{1i}\right)\left(\sum_{i=1}^{20} X_{2i}\right) \Big/ n\right]$$

$$= \frac{1}{19}[4403.31 - (92.8)(908)/20] = 10.01$$

$$S_{13} = S_{31} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{1i}X_{3i}\right) - \left(\sum_{i=1}^{20} X_{1i}\right)\left(\sum_{i=1}^{20} X_{3i}\right) \Big/ n\right]$$

$$= \frac{1}{19}[4403.31 - (92.8)(908)/20] = -1.809$$

$$S_{22} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{2i}^2\right) - \left((\sum_{i=1}^{20} X_{2i})^2/n\right)\right]$$

$$= \frac{1}{19}[45019.18 - (908)^2/20] = 199.798$$

$$S_{23} = S_{32} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{2i}X_{3i}\right) - \left(\sum_{i=1}^{20} X_{2i}\right)\left(\sum_{i=1}^{20} X_{3i}\right) \Big/ n\right]$$

$$= \frac{1}{19}[8941.06 - (908)(199.3)/20] = -5.64$$

$$S_{33} = \frac{1}{n-1}\left[\left(\sum_{i=1}^{20} X_{3i}^2\right) - \left((\sum_{i=1}^{20} X_{3i})^2/n\right)\right]$$

$$= \frac{1}{19}[2054.95 - (199.3)^2/20] = 3.628$$

$$S=\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

$$S=\begin{bmatrix} 2.879 & 10.010 & -1.810 \\ 10.010 & 199.798 & -5.64 \\ -1.810 & -5.64 & 3.628 \end{bmatrix}$$

$$S=\begin{bmatrix} 2.879 & 10.002 & -1.810 & \vdots & 2.879 & 10.002 \\ 10.002 & 199.798 & -5.627 & \vdots & 10.002 & 199.798 \\ -1.810 & \_5.627 & 3.628 & \vdots & -1.810 & -5.627 \end{bmatrix}$$

$$S^{-1}=\begin{bmatrix} 0.586 & \_0.022 & 0.258 \\ \_0.022 & 0.006 & \_0.002 \\ 0.258 & \_0.002 & 0.402 \end{bmatrix}$$

ولحساب احصائية الاختبار T^2وجد أن:

$$T^2 = n(\bar{X}-\mu_0)' S^{-1}(\bar{X}-\mu_0)$$

$$T^2\begin{bmatrix} 0.586 & \_0.022 & 0.258 \\ \_0.042 & 0.006 & \_0.002 \\ 0.258 & \_0.002 & 0.402 \end{bmatrix}[4.640-4\ ,45.400-50,9.965-10]20=$$

$$=20[0.640-4.600,-0.035]\begin{bmatrix} 0.467 \\ \_0.042 \\ 0.160 \end{bmatrix}=9.74$$

نقارن قيمةT^2المحسوبة بالقيمة الجدولية:

$$T^2_{(p,n-1,\alpha)}=\frac{(n-1)p}{n-p}F_{P,n-p,0.01}=\frac{(19)3}{17}F_{3,17,0.01}=3.353(2.44)$$
$$=\ 8.18$$

نجد أن $T^2=9.74>8.18$وبالتالي فإننا نرفض فرض العدم في مقابل الفرض البديل عند مستوى معنوية $\alpha=0.10$.

ويجب ملاحظة أنه يتم رفض فرض العدم إذا اختلف أحد أو بعض عناصر متجه المتوسطات اختلافاً واضحا عن القيم التي يحددها فرض العدم لها في هذا المثال وهي'(4,50,10).وعند هذه

النقطة فإننا لانعرف القيم التي تعتبر سببا في رفض فرض العدم أي لا نعرف القيم التي لا تتفق مع بيانات العينة.

الحل باستخدام برنامج R

للحصول علىالنتائج نستخدم الدالة السابقة كما يلي، حيث تكون المداخلات هيX_1, X_2, X_3

```
X=cbind(x1=c(3.7,5.7,3.8,3.2,3.1,4.6,2.4,7.2,6.7,5.4,3.9,4.5,3.5,4.5,1.5,8.5,4.5,6.5,4.1,5.5),x2=c(48.5,65.1,47.2,53.2,55.5,36.1,24.8,33.1,47.4,54.1,36.9,58.8,27.8,40.2,13.5,56.4,71.6,52.8,44.1,40.9),x3=c(9.3,8.0,10.9,12.0,9.7,7.9,14.0,7.6,8.5,11.3,12.7,12.3,9.8,8.4,10.1,7.1,8.2,10.9,11.2,9.4))

T.test(X,mu=c(4,50,10))

      T2          Fstat          df1        df2          p.value

 9.73877     2.90454          3          17          0.06492
```

التفسير:

نجد أن $T^2 = 9.74 > 8.18$وبالتالي فإننا نرفض فرض العدم في مقابل الفرض البديل عند مستوى معنوية $\alpha = 0.10$.

ويجب ملاحظة أنه يتم رفض فرض العدم إذا اختلفت أحد أو بض عناصر متجه المتوسطات اختلافا واضحا عن القيم التي يحددها فرض العدم لها في هدا المثال وهي'(4,50,10).وعند النقطة فإننا لانعرف القيم التي تعتبر سببا في رفض فرض العدم أي لا نعرف القيم التي لا تتفق مع بيانات العينة**.**

.:بما أن القيمة المحسوبة أكبر من قيمة p-valveعليه نرفض فرض العدم.

2. اختبار تساوي متوسطي مجتمعين:

مثـال (3)

إذاكان لديك البيانات التالية:

اختبر الفرضية $H_0: \mu_1 = \mu_2$

المجتمع الثاني		المجتمع الاول	
X1	X2	X1	**X2**
15	112	10	**133**
12	111	17	**120**
10	120	12	**115**
13	140	15	**157**
17	123	20	**130**
20	115	33	**125**
30	102	47	**95**

الحل:

من واقع البيانات أعلاه فإن:

متوسط المجتمع الأول $\bar{X}_1 = \frac{\sum_{i=1}^{12} X_{1i}}{7} = \frac{117}{7} = 16.71$

متوسط المجتمع الأول $\bar{X}_{12} = \frac{\sum_{I=1}^{12} X_{2i}}{7} = \frac{823}{7} = 117.57$

متوسط المجتمع الثاني $\bar{X}_{21} = \frac{\sum_{i=1}^{12} X_{1i}}{7} = \frac{154}{7} = 22$

متوسط المجتمع الثاني $\bar{X}_2 = \frac{\sum_{i=1}^{12} X_{2i}}{7} = \frac{875}{7} = 125$

وبهذا فإن متجه المتوسطات للمجتمع الأول هو:

$$\bar{X}_1 = \binom{16.71}{117.57}$$

أما متجه المتوسطات للمجتمع الثاني هو:

$$\bar{X}_2 = \binom{22}{125}$$

لتكوين مصفوفة التباينات للمجتمع الأول نحتاج حساب:

$$S_{11} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{1i}^2 - \left((\sum_{i=1}^{7} X_{1i})^2 \Big/ n\right)\right] = \frac{1}{6}\left[2227 - \frac{(117)^2}{7}\right]$$

$$= 45.238$$

$$S_{22} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{2i}^2 - \left((\sum_{i=1}^{7} X_{2i})^2 \Big/ n\right)\right] = \frac{1}{6}\left[97623 - \frac{(823)^2}{7}\right]$$

$$= 143.619$$

$$S_{12} = S_{21} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{1i}X_{2i} - \left(\sum_{i=1}^{7} X_{1i}\right)\left(\sum_{i=1}^{7} X_{2i}\right)\Big/ n\right]$$

$$= [13483 - 13755.85] = -45.475$$

$$S_1 = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} 45.238 & -45.475 \\ -45.475 & 143.619 \end{pmatrix}$$

ولتكوين مصفوفة التباينات للمجتمع الثاني نحتاج لحساب:

$$S_{11} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{1i}^2 - \left((\sum_{i=1}^{7} X_{1i})^2 \Big/ n\right)\right] = \frac{1}{6}\left[4456 - \frac{(154)^2}{7}\right] = 178$$

$$S_{22} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{2i}^2 - \left((\sum_{i=1}^{7} X_{2i})^2 \Big/ n\right)\right] = \frac{1}{6}\left[111513 - \frac{(875)^2}{7}\right]$$

$$= 356.33$$

$$S_{12} = S_{21} = \frac{1}{n-1}\left[\sum_{i=1}^{7} X_{1i}X_{2i} - \left(\sum_{i=1}^{7} X_{1i}\right)\left(\sum_{i=1}^{7} X_{2i}\right)\Big/ n\right]$$

$$= \frac{1}{6}[18295 - 19250] = -159.16$$

$$S_2 = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} 178 & -159.16 \\ -159.16 & 356.33 \end{pmatrix}$$

$$S_u = \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2}$$

$$= \frac{6\begin{pmatrix} 45.238 & -45.475 \\ -45.475 & 143.619 \end{pmatrix} + 6\begin{pmatrix} 178 & -159.16 \\ -159.16 & 356.33 \end{pmatrix}}{12}$$

$$= \frac{\begin{pmatrix} 271.428 & -272.85 \\ -272.85 & 861.714 \end{pmatrix} + \begin{pmatrix} 1068 & -954.96 \\ -954.96 & 2137.98 \end{pmatrix}}{12}$$

$$= \frac{\begin{pmatrix} 1339.428 & -1227.81 \\ -1227.81 & 2999.694 \end{pmatrix}}{12}$$

$$S = \begin{pmatrix} 111.619 & -102.3175 \\ -102.3175 & 249.97 \end{pmatrix}$$

$$S^{-1} = \frac{1}{17432.52}\begin{pmatrix} 249.97 & 102.3175 \\ 102.3175 & 111.619 \end{pmatrix}$$

$$S^{-1} = \begin{pmatrix} 0.01433 & 0.005869 \\ 0.005869 & 0.0064029 \end{pmatrix}$$

$$T_c^{-1} = \frac{n_1 n_2}{n_1+n_2-2}(\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2)$$

$$\frac{7*7}{7+7-2}[16.71-22,117.57-125]' \begin{bmatrix} 0.01433 & 0.005869 \\ 0.005869 & 0.0064029 \end{bmatrix} \begin{bmatrix} 16.71-22 \\ 117.57-125 \end{bmatrix}$$

$$\frac{49}{12}[-5.29,-7.43]' \begin{bmatrix} -0.11 \\ -0.07 \end{bmatrix}$$

$$\frac{49}{12} * 1.102 = 4.499$$

لحساب القيمة الجدولية

$$C^2 = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p,n_1+n_2-p-1(\alpha)}$$

$$= \frac{(7+7-2)2}{(7+7-2-1)} \mathrm{F}_{2,7+7-2-1(0.05)}$$

2.1818*3.98=8.683

الحل باستخدام برنامجR

نستخدم الدالة T.test2 (انظر الملحق)

```
X=cbind(x1=c(15,12,1.13,17,20,30),x2=c(112,111,120,140,123,115,102)
 Y=cbind(x1=c(10,17,12,15,20,33,47),x2=c(113,120,115,157,130,125,95
T.test2(Y,X)
      T2          Fstat           df1    df2           p.value
 4.252207     1.948928         2       11          0.1885728
```

نجد أن $8.68 > 4.25 = T^2$وبالتالي فإننا نرفض فرض البديل في مقابل الفرض العدم عند مستوى معنوية $\alpha = 0.05$

تحليل الملامح العامة

مثـال(4)

في بحث أجري لدراسة العلاقات الزوجية بين المتزوجين حديثا بالولايات المتحدة الأمريكية اختيرت عينه عشوائية مكونة منn=30من الأزواج وعينة أخرى مكونة منn=30من الزوجات حيث وجهت إليهم أربعة أسئلة تتناول بعض أوجه العلاقة الزوجية افترض أنx_1 تشير إلى إجابات السؤال الأول المسجلة على مقياس ذي ثمان نقاط تأخذ القيم من 1إلى8 وأنx_2 تشير إلى إجابات السؤال الثاني المسجلة على نفس المقياس افترض أيضا أنx_3 تشير إلى إجابات السؤال الثالث التي سجلت على مقياس ذي خمس نقاط تأخذ القيم من1إلى 5 وأنx_4 تنتشر إلى إجابات السؤال الرابع المسجلة على نفس المقياس.

الحـل:

المجتمع الأول هو مجتمع الرجال المتزوجون.

المجتمع الثاني هو مجتمع النساء المتزوجات.

وباستخدام بيانات العينتين نجد أن متجهي متوسطات العينتين هما:

$$\bar{x}_1 = \begin{bmatrix} 6.833 \\ 7.033 \\ 3.967 \\ 4.700 \end{bmatrix} \quad \bar{x}_2 = \begin{bmatrix} 6.633 \\ 7.000 \\ 4.000 \\ 4.533 \end{bmatrix}$$

(ذكور) (إناث)

كما نجد أن مصفوفة التباينات والتغاير العامة هي:

$$S_{pooled} = \begin{bmatrix} 0.606 & 0.262 & 0.066 & 0.161 \\ 0.262 & 0.637 & 0.173 & 0.143 \\ 0.066 & 0.173 & 0.810 & 0.029 \\ 0.161 & 0.143 & 0.029 & 0.306 \end{bmatrix}$$

حيث أن حجم العينة كبير بدرجة معقولة. فإننا نستخدم الأساليب الإحصائية الناجمة من افتراض التوزيع المعتدل وذلك على الرغم من أنه من الواضح أن البيانات هي أعداد صحيحة لا تتبع التوزيع المعتدل ولاختبار التوازي$(C\mu_1 = C\mu_2)$ H_{01}: فإننا نقوم بحساب:

$$CS_{pooled}C' =$$

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0.606 & 0.262 & 0.066 & 0.161 \\ 0.262 & 0.637 & 0.173 & 0.143 \\ 0.066 & 0.173 & 0.810 & 0.029 \\ 0.161 & 0.143 & 0.029 & 0.306 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.719 & -0.268 & -0.125 \\ -0.268 & 1.101 & -0.751 \\ -0.125 & -0.751 & 1.058 \end{bmatrix}$$

كما نقوم بحساب

$$C(\bar{x}_1 - \bar{x}_2) = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0.200 \\ 0.033 \\ -0.033 \\ 0.167 \end{bmatrix} = \begin{bmatrix} -0.167 \\ -0.066 \\ 0.200 \end{bmatrix}$$

لذا فإن:

$$T^2 = [-.167\,, -.066\,, .200]\ (\frac{1}{30} + \frac{1}{30})^{-1} *$$

$$\begin{bmatrix} 0.719 & -0.268 & -0.125 \\ -0.268 & 1.101 & -0.751 \\ -0.125 & -0.751 & 1.058 \end{bmatrix}^{-1} \begin{bmatrix} -0.167 \\ -0.066 \\ 0.200 \end{bmatrix} = 15(0.067) = 1.005$$

أضف إلى ذلك انه عند $\alpha = 0.05$ فإن:

$$c^2 = \left[\frac{(30+30-2)(4-1)}{30+30-4}\right] F_{3,56}(0.05) = 3.11(2.8) = 8.7$$

وحيث أن$\mathrm{T}^2 = 1.005 < 8.7$ فإننا نقبل فرض العدم الخاص بتوازي خطوط الملامح العامة للرجال والسيدات.

نستطيع اختبار فرض تطابق هذه الخطوط ولاختبار $\mathrm{H}_{02}: 1'\mu_1 = 1'\mu_2$(خطوط الملامح العامةمتطابقة) فإننا نحتاج إلى مجموع عناصر المتجه $(\bar{\mathrm{X}}_1 - \bar{\mathrm{X}}_2)$ومجموع عناصر مصفوفة التباينات والتغاير العامة

$$1'(\bar{X}_1 - \bar{X}_2) = 0.367$$

$$1'S_{pooled}1 = 4.027$$

نجد أن:

$$\mathrm{T}^2 = \left(\frac{0.367}{\sqrt{(\frac{1}{30}+\frac{1}{30})4.027}}\right)^2 = 0.708$$

وعند $\alpha = 0.05$ نجد أن $\mathrm{F}_{1,58}(0.05) = 4.0$نجد أن

$$\mathrm{T}^2 = 0.708 < \mathrm{F}_{1,58}(0.05) = 4.0$$

عليه لا نستطيع رفض فرض العدم القائل بتطابق خطوط الملامح العامة.

الفصل الرابع
الخلاصة
والتوصيات

4-1خلاصة

في هذا البحث تم التعريف بتوزيع هوتلنج واستخداماته في اختبار الفرضيةالمتعلقة بمجتمع واحد وكذلك اختبار تساوي مجتمعين ، ثم تحليل الملامح العامة لمجتمعين حيث تم إيضاح طريقة اجراء هذه الاختبارات يدويا وباستخدام برنامج R.

4-2التوصيات

1-نوصي باستخدام برنامج R في اجراء الاختبارات المتعلقة باختبارات الفروض باعتباره احد البرامج المجانية ومرونته في التعامل مع البيانات متعددة المتغيرات.

2-نوصي بدراسة الاستخدامات الاخرى لإحصاء هوتلنج.

3-نوصي بدراسة مستقبلة يتم فيها استخدام بعض البرامج الاحصائية الاخرى

المراجع

المراجع العربية:

1-ريتشارد جونسون ،دين وشرن/تعريب : عبد المرضي حامد عزام ،مراجعة دكتور بوعلام بن جيلالي ، التحليل الإحصائي للمتغيرات المتعددة،(دار المريخ) المملكة العربية السعودية-الرياض،سنة1998

المراجع الأجنبية.

2-Anderson,T.W..An Introduction to Multivariate Statistical Analysis, New York: John Wiley,1985.

3-DasGupte,S.,and M.Perlman,''Power of the Non-central F-test :Effect of Additional Variables on Hotel ling s' T^2-tset, "Journal of the American Statistical Association,69,no 345(1974),174-180.

4-R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

URL https://www.R-project.org./

في حالة تكوين الجدول:

ملحق برنامج R

في حالة تكوين الجدول:

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

ومنها: m=n-1

بصورة العامة

$$T^2_{(p,m,\alpha)} = \frac{mp}{m-p+1} F_{(p,m-p+1,\alpha)}$$

بشرط أن تكون قمة M أكبر من P وأن تكون درجة المقام أكبر من الصفر

$T^2_{(p,m,0.05)}$

P / m	2	3	4	5	6	7	8	9	10

2	798	-	-	-	-	-	-	-	-
3	57	1941.366	-	-	-	-	-	-	-
4	25.4723	114.9858	3593.332	-	-	-	-	-	-
5	17.3607	46.3831	192.4679	5754.047	-	-	-	-	-
6	13.8867	29.6612	72.9375	289.4461	8423.496	-	-	-	-
7	12.0009	22.7197	44.7176	105.157	405.9202	11601.65	-	-	-
8	10.8284	19.0283	33.2299	62.5606	143.0503	541.8901	15288.49	-	-
9	10.0327	16.7663	27.2021	45.453	83.2023	186.6216	697.3557	19484	-
10	9.4589	15.2482	23.5446	36.5615	59.4035	106.6487	235.873	872.3172	24188.17
11	9.0262	14.1627	21.1082	31.2048	47.1225	75.0884	132.903	290.8059	1066.774
12	8.6887	13.3498	19.3765	27.6562	39.7643	58.8932	92.5117	161.967	351.421
13	8.4181	12.7191	18.0859	25.1453	34.9107	49.2316	71.8779	111.6757	193.842
14	8.1966	12.216	17.0886	23.2808	31.4884	42.8807	59.6116	86.0793	132.5818
15	8.0119	11.8057	16.2958	21.8446	28.9546	38.4154	51.5715	70.9073	101.4991
16	7.8556	11.4648	15.651	20.7058	27.0075	35.1172	45.9318	60.9863	83.1205
17	7.7217	11.1771	15.1166	19.7817	25.467	32.5879	41.7746	54.0412	71.1272
18	7.6056	10.9312	14.6667	19.0173	24.2192	30.5903	38.5919	48.9302	62.7456
19	7.5041	10.7186	14.2829	18.3749	23.1886	28.9745	36.0818	45.0231	56.5865
20	7.4145	10.533	13.9516	17.8276	22.3237	27.642	34.0543	41.9456	51.8841
21	7.3349	10.3696	13.6628	17.3559	21.5878	26.5249	32.3841	39.4626	48.1843
22	7.2638	10.2247	13.4089	16.9452	20.9543	25.5755	30.9854	37.419	45.2019
23	7.1997	10.0952	13.184	16.5846	20.4033	24.7589	29.7976	35.7092	42.7497
24	7.1418	9.979	12.9833	16.2653	19.9199	24.0494	28.7769	34.2585	40.6995
25	7.0892	9.8739	12.8032	15.9808	19.4923	23.4273	27.8906	33.0127	38.9611
26	7.0412	9.7786	12.6407	15.7257	19.1116	22.8775	27.1141	31.9317	37.4693
27	6.9972	9.6916	12.4933	15.4956	18.7704	22.3882	26.4283	30.985	36.1755
28	6.9567	9.612	12.359	15.2871	18.4629	21.95	25.8183	30.1495	35.0432
29	6.9194	9.5389	12.2362	15.0973	18.1844	21.5553	25.2722	29.4066	34.0442
30	6.8848	9.4715	12.1234	14.9238	17.931	21.198	24.7806	28.742	33.1565

$T^2_{(\text{p,m},0.10)}$

m \ p	2	3	4	5	6	7	8	9	10
2	198	-	-	-	-	-	-	-	-
3	27	482.3392	-	-	-	-	-	-	-
4	14.5664	54.9707	893.3274	-	-	-	-	-	-
5	10.8114	26.9539	92.4342	1431.002	-	-	-	-	-
6	9.0713	18.8589	42.7412	139.3894	2095.359	-	-	-	-
7	8.081	15.2018	28.7507	61.9402	195.8361	2886.392	-	-	-
8	7.4456	13.155	22.5293	40.5058	84.5557	261.7743	3804.095	-	-
9	7.0045	11.8571	19.0846	31.0768	54.1316	110.5901	337.2037	4848.464	-
10	6.681	10.9642	16.9173	25.8959	40.8541	69.6319	140.0446	422.1245	6019.498
11	6.4338	10.3138	15.4353	22.6548	33.6001	51.8656	87.0087	172.9199	516.5365
12	6.2389	9.8196	14.361	20.4484	29.0817	42.2024	64.1141	106.2631	209.2165
13	6.0814	9.4317	13.5477	18.8544	26.0163	36.2041	51.706	77.601	127.396
14	5.9514	9.1193	12.9115	17.6515	23.808	32.1457	44.0253	62.1126	92.3272
15	5.8424	8.8625	12.4005	16.7126	22.1452	29.2287	38.8402	52.5474	73.4234
16	5.7497	8.6476	11.9813	15.9601	20.85	27.0364	35.1204	46.1023	61.7717
17	5.6699	8.4653	11.6314	15.344	19.8137	25.3315	32.3292	41.4858	53.9333
18	5.6004	8.3086	11.3349	14.8303	18.9663	23.9692	30.1614	38.0264	48.3263
19	5.5394	8.1726	11.0805	14.3958	18.2608	22.8566	28.4313	35.3426	44.1295
20	5.4855	8.0534	10.86	14.0235	17.6647	21.9313	27.0197	33.2029	40.8769
21	5.4374	7.948	10.6669	13.701	17.1543	21.1501	25.8469	31.4586	38.2859
22	5.3943	7.8543	10.4966	13.4189	16.7127	20.482	24.8574	30.0105	36.1754
23	5.3555	7.7703	10.3451	13.1703	16.3268	19.9042	24.0118	28.7897	34.4244
24	5.3203	7.6947	10.2096	12.9494	15.9868	19.3997	23.281	27.747	32.9491
25	5.2882	7.6263	10.0876	12.7519	15.6849	18.9555	22.6432	26.8463	31.6898
26	5.2589	7.564	9.9773	12.5742	15.4152	18.5613	22.0819	26.0608	30.6026
27	5.232	7.5071	9.877	12.4136	15.1728	18.2093	21.5842	25.3697	29.6547
28	5.2072	7.455	9.7854	12.2677	14.9537	17.893	21.1398	24.7572	28.8212
29	5.1843	7.407	9.7015	12.1346	14.7547	17.6073	20.7408	24.2105	28.0827
30	5.1631	7.3626	9.6243	12.0126	14.5732	17.348	20.3804	23.7198	27.4239

$$T^2_{(\mathrm{p,m},0.01)}$$

m \ p	2	3	4	5	6	7	8	9	10
2	19998	-	-	-	-	-	-	-	-
3	297	48630.17	-	-	-	-	-	-	-
4	82.1774	594.9972	89993.33		-	-	-	-	-
5	45	147.2835	992.4937	144091.2	-	-	-	-	-
6	31.8574	75.1247	229.6792	1489.489	210923.5	-	-	-	-
7	25.4911	50.6518	111.8392	329.4326	2085.984	290489.4	-	-	-
8	21.8207	39.1182	72.9083	155.2186	446.5705	2781.979	382788.5	-	-
9	19.4605	32.5978	54.8898	98.7032	205.2927	581.1056	3577.472	487820.3	-
10	17.8256	28.4662	44.838	72.8825	128.0671	262.0758	733.0447	4472.464	605584.7
11	16.6308	25.637	38.5334	58.6177	93.1274	161.0149	325.5756	902.3918	5466.956
12	15.7216	23.5883	34.2511	49.7387	73.9687	115.6399	197.5548	395.7966	1089.149
13	15.0077	22.0411	31.1706	43.7446	62.1141	90.9068	140.4286	237.6916	472.7418
14	14.4328	20.8339	28.8568	39.4543	54.1499	75.6759	109.4408	167.4985	281.4285
15	13.9605	19.8671	27.0598	36.2455	48.4723	65.4834	90.4331	129.5759	196.853
16	13.5656	19.0762	25.6262	33.7623	44.2404	58.2414	77.7546	106.3911	151.3157
17	13.2307	18.4177	24.4576	31.7875	40.9749	52.8584	68.771	90.9692	123.5537
18	12.9433	17.8612	23.4874	30.1819	38.3846	48.7148	62.1094	80.0672	105.1308
19	12.6939	17.385	22.6697	28.8522	36.2831	45.4348	56.992	71.9994	92.1338
20	12.4755	16.973	21.9716	27.7339	34.5462	42.7788	52.9485	65.8126	82.5324
21	12.2828	16.6131	21.3688	26.7808	33.0879	40.5872	49.6794	60.9318	75.181
22	12.1113	16.296	20.8433	25.9593	31.847	38.7497	46.9856	56.9912	69.3891
23	11.958	16.0147	20.3812	25.2441	30.7789	37.1883	44.7301	53.7481	64.7187
24	11.8199	15.7634	19.9717	24.6161	29.8502	35.846	42.8156	51.0356	60.879
25	11.695	15.5376	19.6065	24.0604	29.0357	34.6801	41.1714	48.7355	57.6708
26	11.5814	15.3337	19.2787	23.5652	28.3157	33.6585	39.7447	46.762	54.9528
27	11.4778	15.1485	18.983	23.1214	27.6748	32.7563	38.4957	45.051	52.6224
28	11.3828	14.9797	18.7148	22.7212	27.1007	31.9539	37.3934	43.5541	50.6036
29	11.2954	14.8251	18.4706	22.3588	26.5837	31.2357	36.4138	42.2341	48.8387
30	11.2147	14.6831	18.2472	22.0289	26.1157	30.5894	35.5377	41.0618	47.2833

```
T.test<- function(X, mu=0){

X <- as.matrix(X)

n <- nrow(X)

p <- ncol(X)

df2 <- n - p

if(df2 < 1L) stop("Need nrow(X) >ncol(X).")

if(length(mu) != p) mu <- rep(mu[1], p)

xbar<- colMeans(X)

S <- cov(X)

T2 <- n * t(xbar - mu) %*% solve(S) %*% (xbar - mu)

Fstat<- T2 / (p * (n-1) / df2)

pval<- 1 - pf(Fstat, df1=p, df2=df2)

data.frame(T2=as.numeric(T2), Fstat=as.numeric(Fstat),

df1=p, df2=df2, p.value=as

.numeric(pval), row.names="")

}
```

```
T.test2

function(X, Y=NULL, mu=0, paired=FALSE, asymp=FALSE){

if(is.null(Y)){

# one-sample T^2 test: same code as before (omitted here)

} else {

if(paired){

# dependent two-sample T^2 test

X <- as.matrix(X)

Y <- as.matrix(Y)

if(!identical(dim(X),dim(Y))) stop("Need dim(X) == dim(Y).")

xx <- T.test(X-Y, mu=mu, asymp=asymp)

xx$type<- "dep-sample"

return(xx)

} else {

# independent two-sample T^2 test

X <- as.matrix(X)

Y <- as.matrix(Y)

nx<- nrow(X)

ny<- nrow(Y)

p <- ncol(X)

df2 <- nx + ny - p - 1

if(p != ncol(Y)) stop("Need ncol(X) == ncol(Y).")

if(min(nx,ny) <= p) stop("Need min(nrow(X),nrow(Y)) >ncol(X).")
```

```
Sp<- ((nx-1)*cov(X) + (ny-1)*cov(Y)) / (nx + ny - 2)

dbar<- colMeans(X) - colMeans(Y)

T2 <- (1/((1/nx) + (1/ny))) * t(dbar - mu) %*% solve(Sp) %*% (dbar -
mu)

Fstat<- T2 / ((nx + ny - 2) * p / df2)

if(asymp){

pval<- 1 - pchisq(T2, df=p)

} else {

pval<- 1 - pf(Fstat, df1=p, df2=df2)

}

return(data.frame(T2=as.numeric(T2), Fstat=as.numeric(Fstat),

df1=p, df2=df2, p.value=as.numeric(pval),

type="ind-sample", asymp=asymp, row.names=""))

} # end if(paired)

} # end if(is.null(Y))

}
```