استعمال توزيع t متعدد المتغيرات في دراسة بعض التغيرات الحيوية للمصابين بأمراض القلب

أيمان عبد علي داود[1] ، مناف يوسف[2]، رفعت عبد الرحمن الداغر[3] ، لمى سعد عبد الباقي[1]

[1] قسم الرياضيات ، كلية التربية للبنات ، جامعة تكريت ، تكريت ، العراق

[2] قسم الاحصاء ، كلية الادارة والاقتصاد ، جامعة بغداد ، بغداد ، العراق

[3] دكتوراه (بورد عربي) في الامراض الباطنية والصدرية والقلبية / م. تكريت التعليمي ، تكريت ، العراق

(تاريخ الاستلام: 27 / 10 /2013 ---- تاريخ القبول: 10 / 12 /2013)

الملخص

تم في هذا البحث استعمال توزيع t متعدد المتغيرات في دراسة بعض المتغيرات الحيوية التي تخص الرئتين واجزاء القلب (البطينين والاذينين) لمجموعة من المرضى المصابين بأحد انواع امراض القلب ، وقد تبين وجود فروق معنوية بين الاصحاء والمرضى بالنسبة لهذه المتغيرات الحيوية .

1- المقدمة

يعد توزيع t متعدد المتغيرات التعميم لتوزيع t الاحادي وله اهمية كبيرة جدا ضمن جانب الاستدلال الاحصائي اذ يعد توزيع t المتعدد البديل الامثل للتوزيع الطبيعي المتعدد وذلك لتحسس التوزيع الاخير للبيانات ذات الاطراف الطويلة (سميكة الذيل او الاطراف) او التي تحوي على قيم شاذة ، إضافة الى التقلطحات المشاهدة غالبا في التطبيقات الواقعية مثل الانترنت واشارات المرور والتأمين والتي من غير الممكن تحليلها من خلال التوزيع الطبيعي المتعدد.[11][14]

ويعد الباحث (Cornish,1954) أول من اشتق توزيع t المتعدد ومن ثم توالت الدراسات والأبحاث العديدة ومنها دراسة الباحث Pi-Erh Lin في عام 1971الذي بيّن فيها امكانية كتابة توزيع t المتعدد بصيغتين فضلا عن دراسته للخصائص المميزة له [7][10] .

ومن ثم استمرت البحوث حول هذا التوزيع الى السنوات الاخيرة ففي عام 2003 قام الباحث Gupta بدراسة توزيع t الملتوي متعدد المتغيرات والذي يتصف ببعض خصائص توزيع t المتعدد ، كما وبحث عزومه ايضا فضلا عن اثبات توزيع كوشي الملتوي متعدد المتغيرات كونه حالة خاصة منه [5].

اما في عام 2010 قدم الباحث Aeschliman وزملائه [3] خوارزمية جديدة لإيجاد قيم تقريبية لمعالم توزيع t المتعدد وتتسم هذه الخوارزمية بالسرعة والسهولة مقارنة بمقدرات الامكان الاعظم لكنها اقل دقة وهي بمثابة أسلوب نظري جديد لحل المشاكل التي تظهر ضمن معادلات الامكان الاعظم.

2- الدالة الاحتمالية لتوزيع t :

2-1 : توزيع t الاحادي The Univariate t-Distribution

يعرف المتغير العشوائي X بأنه متغير ذو توزيع t احادي المتغير بـ V من درجات الحرية اذا كانت دالة الكثافة الاحتمالية (pdf) لذلك المتغير تعرف بالصيغة الاتية [6][8]:

$$f_1(X)=\frac{\sqrt{\frac{v+1}{2}}}{\sqrt{\pi v}\,\Gamma\left(\frac{v}{2}\right)}\left[1+\frac{X^2}{v}\right]^{-\frac{v+1}{2}} \quad -\infty<X<+\infty \ldots\ldots(1)$$

ويرمز له t_v .

2-2:توزيع t القياسي ثنائي المتغيرات

The Standard Bivariate t-Distribution

يقال للتوزيع بانه توزيع t قياسي ثنائي المتغيرات اذا كان بالصيغة الاتية[6] :

$$f_2(X_1,X_2)=\frac{1}{2\pi}\left[1+\frac{1}{v}(X_1^2+X_2^2)\right]^{-(\frac{v}{2}+1)} \quad ,v>0 \ \ldots..(2)$$

ويمكن اثبات ان المتغيران X_1 و X_2 في (2) هما غير مرتبطين لكنهما غير مستقلين مالم يتحقق الشرط $v\to\infty$.

2-3: توزيع t الثنائي المرتبط *Correlated Bivariate t-Distribution*

ويعد هذا التوزيع حالة خاصة من توزيع t ثنائي المتغيرات اذ يمثل تراتيب متماثلة عندما تكون المركبات ذات متوسط $\underline{0}$ وتباين γ_2 ومعامل ارتباط ρ، وتكون صيغة هذا التوزيع كالاتي[6] :

$$f_3(X_1,X_2)=\frac{1}{2\pi\sqrt{1-\rho^2}}\left[1+\frac{X_1^2+X_2^2-2\rho X_1X_2}{v(1-\rho^2)}\right]^{-(\frac{v}{2}+1)} \qquad (3)$$

اذ ان

$$\gamma_2(v-2)=v \quad , \quad v>2$$

2-4: توزيع t متعدد المتغيرات *The Multivariate t-Distribution*

هناك صيغ عدة لتوزيع t متعدد المتغيرات مذكورة في كثير من البحوث وسوف نذكر في بحثنا هذا الصيغة الاتية :

ان دالة التوزيع الاحتمالي (*p.d.f*) الى p من المتغيرات تكون بالصيغة الاتية [9] :

$$f(X)=\frac{[\Sigma]^{-\frac{1}{2}}}{C(v,p)\pi^{\frac{p}{2}}}\left[1+\frac{1}{v}(X-\mu)\Sigma^{-1}(X-\mu)\right]^{-\frac{v+p}{2}} \ \ldots(4)$$

اذ ان

X :القيم الحقيقية ضمن المتجه العشوائي x ذو البعد (p * 1) .

μ : متجه المتوسط المجهول .

Σ : مصفوفة معرفة موجبة ذات بعد (p*p) وتحوي على معلمات القياس (scale parameters) .

C(v,p) : الثابت الطبيعي (Normalizing constant) .

اذ ان

$$\Gamma\left(\frac{v+p}{2}\right)C(v,p)=v^{\frac{p}{2}}\,\Gamma\left(\frac{v}{2}\right) \ \ldots..(5)$$

كما ان X (والذي يحوي على p من المتغيرات) يكون بمتوسط μ ومصفوفة تباين مشترك

$$\Sigma^* = v^*\Sigma \quad ....(6)$$

اذ ان $v^* = \frac{v}{v-2}$ ويمكن اعادة كتابتها بالشكل الاتي $T_p(\mu, v^*\Sigma)$ ويفترض ان تكون معلمة الشكل v $(v > 2)$ معلومة والحالة الخاصة لهذا التوزيع عندما تكون

$$\rho = 1 \ , \mu = 0 \ , \Sigma = 1$$

فانه تتحول الدالة متعددة المتغيرات في (4) الى توزيع t احادي المتغير، اما في حالة كون

$$\rho = 2 \ , \mu = 0 \ , \Sigma = I_2$$

فان صيغة دالة t متعدد المتغيرات في (4) سوف تتحول الى توزيع بيرسون الثنائي الاوجه(*Bivariate surface of pearson dist.*) .

ومن الجدير بالاشارة الى امكانية كتابة توزيع t وفق الصيغة الاتية :

$$f(x) = \int_0^{\infty} \frac{|w^2\Sigma|^{-\frac{1}{2}}}{(2\pi)^{\frac{\rho}{2}}} exp \frac{-(X-\mu)'(w^2\Sigma)^{-1}(X-\mu)}{2} h(w)d(w) \ldots (7)$$

و تمثل هذه الصيغة خليط بين كلا من التوزيع الطبيعي المتعدد $N_p(\mu, w^2\Sigma)$ و كذلك w التي تمثل توزيع كاما المعكوس ، اي ان

$$h(w) = \frac{2(\frac{v}{2})^{\frac{v}{2}}}{\Gamma(\frac{v}{2})} w^{-(v+1)} exp\left[\frac{-v}{2w^2}\right] \ldots .(8)$$

وتشير v هنا الى درجة حرية توزيع كاما المعكوس وبالمثل فان vw^{-2}تمثل توزيع مربع كاي وبدرجة حرية v . ولهذا فان لأيw يكون المتجه العشوائي X ذو توزيع طبيعي متعدد ، اي ان

$$(X|\Omega = w) \sim N_p(\mu, w^2\Sigma) \ldots \ldots (9)$$

اما في حالة $v \to \infty$ فان المتغير العشوائي Ω يتناقص مع كل القيم غير الصفرية اي التي تقع على نقطة الوحدة ، و من ثم تقترب الدالة الاحتمالية لتوزيع t المتعدد والمعرف بالصيغة (4) من التوزيع الطبيعي المتعدد $N_p(\mu, \Sigma)$ وهذا التقارب يكون نتيجة القاعدة الاتية:

$$C(v,p) \to 2^{\frac{p}{2}} \quad ....(10)$$

كذلك

$$\lim_{v \to \infty} \left(1 + \frac{u}{v}\right)^{-v} = e^{-u} \quad ....(11)$$

من المهم ايضا ملاحظة ان عدم الارتباط بين $X_1, X_2, \ldots \ldots, X_N$لا يعني انها مستقلة مالم يتحقق الشرط الاتي : $v \to \infty$.

3- دالة الانتروبي *Entropy*

تمثل دالة الـ Entropy للمتجه العشوائي المستمر X مقياس للتوزيعات متعددة المتغيرات اذ تشير الى مقياس لمدى ابتعاد بعض القيم عن بقية النقاط الاخرى اي انه يمثل مقياسا لمدىثقل الاطراف و يُشبه بمقياس التفرطح في حالة المتغير الواحد .

في بعض الاحيان يكون التفرطح غير موجود او خاطئا وربما يكون المتوسط غير موجود ولهذا السبب يتبين محدودية استعمال العزوم كمؤشرات لشكل التوزيع ، لكن مقياس الـ *Entropy* يكون دائما قابل للتطبيق ويمختلف المواضع التطبيقية اذ انه لا يخضع لقيود مثل افتراض التماثل وكذلك يمكن تطبيقه على التوزيعات ذات الاطراف الثقيلة [*12*] .

وبالامكان حساب الـ *Entropy* لـ x وفق الصيغة الاتية :

$$H(x) = E[-\log f(x)]$$

$$= -\int f(x) \log f(x)\, dx \quad ....(12)$$

وقد قام الباحثان *Cusumano & Guerrero* باشتقاق صيغة الانتروبي *Entropy* لتوزيع t المتعدد وفق الصيغة الاتية :

*حاولنا ترجمة الـ*Entropy* فوجدناها ضمن مترجمات كثيرة ومنها مترجم القوقل انها تعني " الكون " وهي بعيدة عن مضمون بحثنا هذا ولذلك تم كتابتها كما هي .

$$H(X;R) = \frac{1}{2}\log|R| + \log\left[\frac{(v\pi)^{\frac{p}{2}}}{\Gamma\left(\frac{P}{2}\right)} B\left(\frac{P}{2}, \frac{v}{2}\right)\right] + \frac{v+P}{2}\left[\psi\left(\frac{v+P}{2}\right) - \psi\left(\frac{v}{2}\right)\right] \ldots (13)$$

اذ ان

$$\psi(t) = \frac{d \log \Gamma(t)}{dt}$$

تمثل دالة كاما الثنائية (*digamma function*).

4- اختبار الفرضيات الاحصائية حول معلمات توزيع t المتعدد

The Statistical Hypotheses testing for the parameters of multivariate t-distribution

سيتم استعمال اختبار نسبة الامكان *likelihood Ratio* لغرض اختبار الفرضيات والذي يقوم على اساس تعظيم دالة الامكان تحت فرضية العدم H_0 وفي نفس الوقت تحت الفرضية البديلة H_1 للحصول على اكبر قدر ممكن من المعلومات المتاحة حول هاتين الفرضيتين تحت الدراسة ويمكن تنفيذ الاختباروفق الصيغة الرياضية الاتية[1]

$$\lambda = \frac{max\theta \in \Omega_0 f(y;\theta)}{max\theta \in \Omega f(y;\theta)} = \frac{L_0(\theta)}{L1(\theta)} \quad ....(14)$$

حيث ان

Ω_0 : فضاء المعلمات تحت الفرضية H_0

وغالبا ما يتم استعمال صيغة اخرى لاختبار نسبة الترجيح وهي

$$-2\ln(\lambda) = 2(L_1 - L_0) \ldots..(15)$$

4 - 1: اختبار الفرضيات حول متجه المتوسط للمجتمع الواحد

4 – 1 – 1: عندما تكون مصفوفة التباين Σ غير معلومة

لاختبار الفرضية

$$H_0: \mu = \mu_0$$

ضد الفرضية

$$H_1: \mu \neq \mu_0 \quad .....(16)$$

اذ ان

μ_0: متجه معلوم القيم .

من المعلوم ان مقدرات الامكان الاعظم لـ μ , Σ بصورة عامة تكون مساوية الى

$$\hat{\mu} = \bar{X}$$
$$\hat{\Sigma} = \frac{1}{n} S \ldots . (17)$$

علما ان

S : مصفوفة التباين المقدرة

ويفرض ان

$$d = (\bar{X} - \mu_0)$$

وبعمل عدة خطوات للاشتقاق نحصل على

$$-2\ln(\lambda) = n\ln\left[|S|\left(I + \acute{d}S^{-1}d\right)\right] - n\,ln|S|$$
$$= n\ln\left(I + \acute{d}S^{-1}d\right) \ldots . (18)$$

من الواضح ان هذه الاحصائية تعتمد على الحد

$$T = \acute{d}S^{-1}d$$

وهو يمثل صيغة اختبار T^2 ومن المعلوم ان توزيع الاحصاءة $\left(\frac{m}{P}\right)T^2$ تتبع توزيع F اللامركزي بمعلمة لامركزية مقدارها

$$\delta = (\mu - \mu_0)' \Sigma^{-1}(\mu - \mu_0) \ldots (19)$$

ويعبر عنه بالصيغة الاتية :

$$\left(\frac{m}{P}\right)T^2 \sim F_{(P,n-P+1)}(\delta) \ldots . (20)$$

ومع ذلك عندما تكون H_0 صحيحة فان المعلمة اللامركزية δ لتوزيع F اللامركزي تساوي صفر وبذلك يتحول الى توزيع F المركزي بدرجات حرية $(P, n-P+1)$ ولهذا فان معيار رفضنا او قبولنا لـH_0 هو اننا نرفض H_0عند مستوى معنوية (α) اذا كان

$$\frac{m}{P}\acute{d}S^{-1}d \geq F_{(\alpha,P,m)} \ldots \ldots (21)$$

اذ ان $m = n - P + 1$. وتقبل H_0 فيما عدا ذلك .

4 – 1 – 2: عندما تكون مصفوفة التباين Σمعلومة .

بالرجوع الى الفرضية (16) وبعد سلسلة من الاشتقاقات نجد ان معيار الرفض الى H_0 يتم اذا كان

$$n(\bar{X} - \mu_0)' \Sigma^{-1}(\bar{X} - \mu_0) \geq \frac{1}{P}F_{\alpha,P,\nu} \ldots \ldots (22)$$

وتقبل الفرضيةH_0 فيما عدا ذلك .

2-4:اختبار الفرضيات حول متجهي متوسطين لمجتمعي t المتعدد

1-2-4 : مصفوفة التباين لمجتمعي توزيع t متساوية وغير معلومة.

لاختبار الفرضية

$$H_0: \mu_1 = \mu_2$$

ضد الفرضية

$$H_1: \mu_1 \neq \mu_2 \ldots . . (23)$$

وقبل اجراء الاختبار المذكور انفا يجب ايجاد مقدرات الامكان الاعظم للمعالم μ , Σ تحت الفرضية H_0 اذ تكون هذه المقدرات مقيدة بالفرضية H_0 وهي

$$\hat{\mu} = \bar{X} = \frac{1}{n_1 + n_2}\sum_{1}^{2}\sum_{j=1}^{n_j} X_{ij}$$

$$\hat{\Sigma} = \frac{1}{n_1 + n_2} R$$
$$= \frac{1}{n_1 + n_2}\sum_{i=1}^{2}\sum_{j=1}^{\cdots j}(X_{ij} - \bar{X})(X_{ij} - \bar{X})' \ldots . (24)$$

وبعد سلسلة من الخطوات نجد ان

$$\left(\frac{m}{P}\right)T^2 \sim F_{(P,n_1+n_2-P+1)} \ldots . . (25)$$

وعليه ترفض الفرضيةH_0 عند مستوى معنوية (α) اذا كان

$$\frac{m}{P}(\bar{X}_2 - \overline{X_1})' S^{-1}(\bar{X}_2 - \overline{X_1}) \geq F_{\alpha,P,m} \ldots \ldots (26)$$

اذ ان

$$m = n_1 + n_2 - \rho + 1$$

وتقبل الفرضيةH_0 في الحالات الاخرى .

2-2-4:مصفوفة التباين لمجتمعي توزيع t متساوية ومعلومة .

بفرض اننا نريد اختبار الفرضية الاتية :

$$H_0: \mu_1 = \mu_2 = \acute{\mu}$$

ضد الفرضية

$$H_1: \mu_1 \neq \mu_2 \neq \acute{\mu} \ldots \ldots . (27)$$

اذ ان

$\acute{\mu}_t$: متجه ثابت معلوم القيمة .

وبعد سلسلة من الاشتقاقات يكون معيار الرفض للفرضيةH_0 بالشكل

$$n_1(\bar{X}_1 - \mu_1)' \Sigma^{-1}(\bar{X}_1 - \mu_1) + n_2(X_2 - \mu_2)' \Sigma^{-1}(X_2 - \mu_2) \geq 2pF_{\alpha,2p,\nu} \ldots \ldots \ldots (28)$$

وتقبل H_0 في الحالات الاخرى .

5– الجانب التطبيقي

5–1: المقدمة

لقد اجتمعت كل الدول التي تهتم بجمع الاحصاءات الحيوية على ارتفاع نسبة الوفيات خلال السنوات الاخيرة نتيجة الاصابة بأمراض القلب والشرايين، فقد اصبحت السبب الرئيس للوفاة في كثير من بلاد العالم .

وقد انتشر هذا النوع من الامراض في العراق كما هو الحال في مختلف انحاء العالم، ويمكن تصنيف هذه الامراض الى [2]:

– امراض القلب الولادية المنشأ *(Congenital Heart Disease)*

– امراض القلب المكتسبة *(Acquired Heart Disease)*

و ان الدراسات والبحوث الاحصائية حول هذا المرض ربما قد تساهم في مجال تقليل انتشاره، كما ان المعلومات الافضل عن امراض القلب وعلاقتها بكيمياء الجسد تفتح تدريجيا افاقا مشرقة تبشر بتخفيف وطأة هذه الامراض الشديدة الفتك في يومنا هذا .

2-5 : وصف البيانات

تم دراسة خمس متغيرات هي :

X_1 : FVC (السعة الحيوية) ويمثل حجم الهواء المقذوف من الرئتيين خلال الزفير بقوة، بعد شهيق كامل. ويعد اهم فحوصات وظائف الرئة [4] .

X_2 :$FEV1$(حجم الزفير خلال الثانية الاولى) ويمثل حجم الهواء المقذوف من الرئتين خلال اول ثانية من الزفير. وهو من فحوصات الرئة المهمة ويميز بين النوعين الرئيسين لاضطراب وظائف الرئة وهما الانسدادي والانكماشي [4] .

X_3 :$Mvpeake$ ويمثل كمية الدم المقذوفة من الاذين الايسر الى البطين الايسر خلال المرحلة الاولى من الطور الانبساطي، وهو يقاس بجهاز الايكو باستعمال تقنية الدوبلر النبضي ويمثل الجزء الاول من الوظيفة الانبساطية لعضلة القلب[13] .

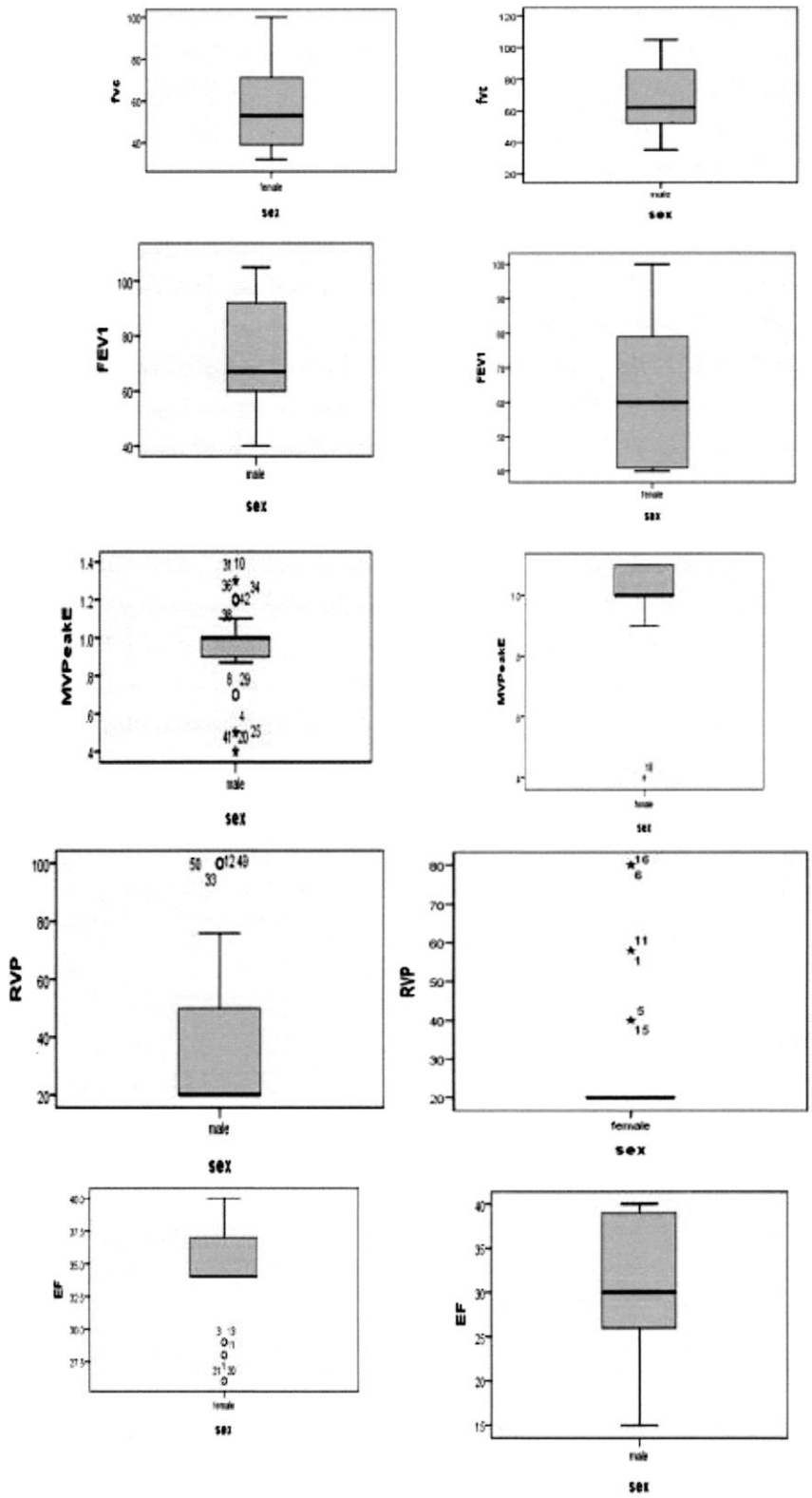
X_4 : EF (نسبة المقذوف) ويمثل نسبة الدم المقذوفة من البطين الايسر خلال الطور الانقباضي ويقاس بعدة طر ائق اهمها الايكو والنظائر المشعة.وهو يمثل وظيفة القلب الانقباضية ويحدد نوع عجز القلب، و يستعمل كذلك في مراقبة تطور امراض القلب الاخرى[13].

X_5 : RVP ويمثل ضغط الدم داخل البطين الايمن ويقاس بواسطة الايكو وباستعمال تقنية الدوبلر المستمر. وهو يمثل مقياسا لوظيفة الشرايين الرئوية ومقاومتها لضغط الدم[13] .

وقد تم جمع البيانات من مستشفى تكريت التعليمي وكان عدد المشاهدات هو 80 مريض (53 من الذكور و27 من الاناث) من المصابين بأحد امراض القلب اضافة الى 40 من الاشخاص غير المصابين بهذه الامراض (20 من الذكور و20 من الاناث) .

كما وتم استعمال برنامج *Matlab R2008a* لحساب قيم الاحصاءات المستعملة في الاختبارات قيد البحث.

وقبل البدء بتحليل هذه البيانات كان لابد من الكشف عن وجود مشاهدات شاذة من عدم وجودها وتوجد طرائق مختلفة للكشف عن مثل هذه المشاهدات في البيانات المتعددة، وفي بحثنا هذا استعملت طريقة الرسم الصندوقي *Box Plot* اذ يتم فحص مشاهدات كل متغير من متغيرات المتجه العشوائي بإحدى طرائق الكشف عن المشاهدات الشاذة للمتجه العشوائي الاحادي المتغير فاذا ظهرت مشاهدة واحدة (او اكثر) شاذة ضمن المتغير الواحد فان هذا يدل على ان مشاهدات العينة المتعددة المتغيرات تكون شاذة [1] .



شكل (1) الكشف عن المشاهدات الشاذة في عينة الدراسة باستعمال الرسم الصندوقي

3-5 : اختبار الفرضيات

***1-3-5* :اختبار فرضية تساوي متوسطي مجتمعي الذكور الاصحاء مع الاناث الاصحاء**

عند اختبار الفرضية (*23*)، في حالة كون مصفوفة التباين العامة للمجتمعين غير معلومة، ولغرض ايجاد المختبر الاحصائي الخاص بالفرضية السابقة وحسب الصيغة (*26*) لابد لنا من حساب قيمة μ_1 (ويمثل متوسطات المتغيرات المدروسة للذكور الاصحاء) وقيمة μ_2 (ويمثل متوسطات المتغيرات المدروسة للاناث الاصحاء) وكذلك S حيث تكون بالمصفوفة الاتية :

$$\mu_1 = \begin{bmatrix} 95.4 \\ 106.7 \\ 1.04 \\ 70.1 \end{bmatrix} \quad , \quad \mu_2 = \begin{bmatrix} 103.5 \\ 113.3 \\ 1.09 \\ 69.6 \end{bmatrix}$$

$$S = \begin{bmatrix} 434.831 & 476.6 & 0.599 & 4.063 \\ 476.6 & 663.389 & 1.047 & 7.947 \\ 0.599 & 1.047 & 0.021 & -0.009 \\ 4.063 & 7.947 & -0.009 & 8.979 \end{bmatrix}$$

نلاحظ ان كلا من μ_1 و μ_2 لا يحوي على قيمة المتغير RVP حيث انه لا يحسب للأشخاص الاصحاء ولهذا لم نأخذه في مقارناتنا الحالية واللاحقة الخاصة بالأصحاء.

عند حساب قيمة المختبر الاحصائي وجد انه يساوي (2.947) وعند مقارنته مع القيمة الجدولية فأننا نقبل فرضية العدم اي لا يوجد فروق معنوية بين المتوسطات المدروسة للذكور الاصحاء والاناث الاصحاء.

2-3-5 : اختبار فرضية تساوي متوسطي مجتمعي الذكور المصابين والاناث المصابات بالمرض

هنا الفرضية تخص المصابين (الذكور والاناث) لهذا سنأخذ المتغير RVP ايضا ضمن المتوسطات المقارنة، وعند القيام بنفس الحسابات السابقة نجد ان قيمة المختبر الاحصائي تساوي (5.607) وبمقارنتها مع القيمة الجدولية فأننا نرفض فرضية العدم ونقبل بالفرضية البديلة، اي وجود فروق معنوية بين متوسطات مجتمعي الذكور المصابين والاناث المصابات بالمرض بالنسبة للمتغيرات المدروسة.

3-3-5: اختبار فرضية تساوي متوسطي مجتمعي الذكور المصابين والذكور الاصحاء

بإعادة نفس الخطوات السابقة نجد ان القيمة المحسوبة للمختبر الإحصائي تساوي (434.968) وبمقارنتها مع القيمة الجدولية فأننا نرفض فرضية العدم ونقبل بالفرضية البديلة، اي وجود فروق معنوية بين الذكور المصابين والاصحاء بالنسبة للمتغيرات المدروسة .

المصادر:

1- اللهيبي ، معتزصلاح غانم ، التقدير وبناء الاختبارات الاحصائية حول معلمات توزيع t المتعدد المتغيرات ، رسالة ماجستير في علوم الاحصاء ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل ، 2009 .

***4-3-5* : اختبار فرضية تساوي متوسطي مجتمعي الاناث المصابات مع الاناث غير المصابات بالمرض.**

بأجراء نفس الخطوات السابقة نجد ان القيمة المحسوبة للمختبر الاحصائي تساوي (705.25) وبمقارنتها مع القيمة الجدولية فأننا نرفض فرضية العدم ونقبل الفرضية البديلة، اي وجود فروق معنوية بالنسبة للمتغيرات المدروسة بين الاناث المصابات مع الاناث غير المصابات بالمرض.

***5-3-5* : اختبار فرضية تساوي متوسطي مجتمعي الاشخاص الاصحاء (اناث وذكور) والمصابين بالمرض (اناث وذكور).**

بإعادة نفس الخطوات السابقة نجد ان القيمة المحسوبة للمختبر الاحصائي تساوي (822.75) وبمقارنتها مع القيمة الجدولية فأننا نرفض فرضية العدم ونقبل الفرضية البديلة، اي وجود فروق معنوية بين متوسطات المتغيرات المدروسة بالنسبة للأصحاء (اناث وذكور) والمصابين بالمرض (اناث وذكور).

6- الاستنتاجات

من خلال هذا البحث تم التوصل الى :

1- عدم وجود فروق معنوية بين الذكور الاصحاء والاناث غير المصابات بالنسبة لمتوسطات المتغيرات المدروسة

2- وجود فروق معنوية عالية بين متوسط المتغيرات المدروسة للذكور المصابين والاناث المصابات.

3- وجود فروق معنوية عالية بين متوسط المتغيرات المدروسة للذكور المصابين والذكور الاصحاء.

4- وجود فروق معنوية عالية بين متوسط المتغيرات المدروسة للإناث المصابات والغير مصابات.

5- وجود فروق معنوية عالية بين متوسط المتغيرات المدروسة للأصحاء (ذكور واناث) مقارنة مع المصابين (ذكور واناث).

7: التوصيات

وفقا لما تم بحثه نوصي :

1- استعمال توزيع t المتعدد في حالة احتواء المشاهدات على قيم شاذة وحالة تعذر استعمال التوزيع الطبيعي المتعدد.

2- اجراء دراسة موسعة تشمل كل ما يتعلق ويتأثر بالإصابة بأحد أمراض القلب اخذين بنظر الاعتبار الجنس لأنه حسب ما استنتجناه انه ذو تأثير كبير جدا ويكون باستطاعة الطبيب من خلال هذه الدراسة الاستدلال على التأثيرات السلبية للإصابةعلى هذه الفعاليات الحيوية وهذا الاستدلال يساعده في اتخاذ الاجراءات اللازمة لمحاولة التقليل من خطورة التأثيرات السلبية للاصابة على هذه الفعاليات.

2- داود ، ايمان عبدعلي ، استعمال طريقة الاتحاد – التقاطع في تحليل البيانات المصنفة ومقارنة نتائجها مع الطرق الاخرى ، رسالة ماجستير في علوم الاحصاء ، كلية الادارة والاقتصاد ، جامعة بغداد ، 1994 .

9-Kibria B.M.G., Joarder A.H.,(2006), "On the Multivariate t-Distribution and Some of its Applications", King Fahd University of Petroleum and Minerals. Technical Report Series, TR 359 .
10-Lin,Pi-Erh, (1971) "some characterizations of the Multivariate t-Dist.", J Statistics ReportM200 .
11-Mclachlan,G.J.,Bean, R.W. , Jones L.B., (2009),"Extension of the Mixture of Factor Analyzers Model to Incorporate the Multivariate t-Distribution", Brisbane 4072, Australia .
12-Nadarajah,S., Kotz,S.,(2005) "Mathematical Properties of the Multivariate t Distribution", Acta Applicandae Mathematicae , 89:53-84.
13- Scott, D. Solomorechocardiography. (2007), Humana press.
14-Yuan M., Huang J.Z., (2009), "Regularized Parameter estimation of high dimensional t Distribution", Journal of Statistical Planning and Inference, 139,2284-2292.
3-Aeschliman,C.,Park, J. ,Kak , A.C.,(2010),"A Novel Parameter Estimation Algorithm for the Multivariate t-Distribution and Its Application to Computer Vision", European Conference on Computer Vision.
4-Amercan thoracic society –standardization of spirometry, (1994) update –American J. of resp. & crit care med (1995),152;1107-1136.
5-Gupta A.K., (2003), "Multivariate Skew t-Distribution", A Journal of Theoretical and Applied Statistics , Volume 37,Issue 4.
6-Joarder,A.H.,"An Introduction to the Bivariate T-Distribution", AMS: 60_ 01, 60E10 .
7-Joarder, A.H., Kibria, B.M.G., (2006), "A Modest Introduction to Uncorrelated T-Model", King Fahd University of Petroleum and Minerals, Technical Report Series , TR 354 .
8-Joarder, A.H.,(2007), "On Some characteristics of the Bivariate T-Distribution", International Journal of Mathematics,2(2),191-204.

The use of multivariate – t distribution in studying some vital variations for heart diseases patients

EmanA.Dawood[1], Munaf Yusif [2], Rifaat A. Al-Daghir[3], Luma S. Abd_Albaqi[1]

[1] *Department of mathematics/college of education for women /university of Tikrit.*
[2] *Department of statistics/college of administration and economy/university of Baghdad.*
[3] *PHD in Medicine /Tikrit teaching hospital.*

(Received: 27 / 10 / 2013 ---- Accepted: 10 / 12 / 2013)

Abstract

In this article, we use multivariate - t distribution to study some vital variables due to lungs and parts of the heart (ventricles & atriums) for patients suffer from some heart diseases, The results showed significant differences among the groups(control and patients).

Keywords: Multivariate –t dist.