تحليل تصميم القياسات المكررة للنموذج العام باتجاه واحد باستخدام بعض الطرائق المعلمية واللامعلمية*

المستخلص

ان نموذج القياسات المكررة العام هو النموذج الذي فيه كل مفردة (individual) تستلم عدد من المعالجات بحيث ان المشاهدات في هذا النموذج غير مستقلة وان القياسات لكل وحدة تجريبية تتوزع متعدد متغيرات ذو بعد r . مع مصفوفة تباين وتباين مشترك $\sum$ غير محددة . ان الغرض من هذا البحث هو تلخيص نتائج دراسة المحاكاة لمقارنة بعض الطرائق اللامعلمية بالاضافة الى الطريقة المعلمية لتصميم القياسات المكررة للنموذج العام باتجاه واحد لاختبار عدم وجود فرق او اختلاف بين تاثيرات المعالجات في حالة ثلاثة معالجات فاكثر وكذلك توضيح تلك الطرائق بتطبيق بيانات تجربة حقيقية عليها بالاضافة الى عرض الطريقة المقترحة من قبل الباحثين ومقارنة نتائجها .

Abstract

The generalized repeated measures model is the model in which each individual receive a number of treatments such that the observations in these models are dependent and the Measurements for each subject distributed as a multivariate with r-dimensions with unspecified variance-covariance matrix. The purpose of this method is to summarize the result of simulation study to compare some of nonparametric Methods as well as the parametric method for repeated measures designs for one-way generalized model to test the null hypothesis that there is no differences between treatments effects for three treatments and more and to clarify these methods by applying data of an actual experiment on them as well as present the suggested method and compare their results .

المقدمة والهدف

في الكثيرمن الدراسات الطبية والتربوية والسيكولوجية (علم النفس)وعلم الاجتماع نشاهد بان المفردات (individuals) تتكرر تحت مختلف الشروط التجريبية(conditions) او المعالجات (treatments) . ويطلق مصطلح تصاميم القياسات المكررة على هذا النوع من البيانات من حيث ان الاستجابة لكل وحدة تجريبية او مفردة تشاهد عند فرص او شروط متعددة . ويمكن ان يكون متغير الاستجابة بصورة مفردة او متعددة الا ان اهتمامنا سوف يكون في حالة متغير الاستجابة المفرد(Univariate) مقاسا عند شروط تجريبية متعددة .

ان الهدف من هذه الدراسة هو مقارنة بعض الاختبارات للطرائق اللامعلمية بالاضافة الى الطريقة المعلمية لغرض اختبار فرضية عدم وجود فرق او تاثير بين المعالجات المختلفة في حالة ثلاث معالجات فاكثر لتصميم القياسات المكررة للنموذج العام باتجاه واحد عندما لايفترض شكل محدد لمصفوفة التباين والتباين المشترك وذلك من خلال تطبيق بيانات تجربة حقيقية على تلك الطرائق واستعمال اسلوب المحاكاة الذي استند على معياري احتمال الخطا من النوع الاول وقوة الاختبار. كذلك عرض اسلوب الطريقة المقترحة من قبل الباحثين ومقارنتها مع الطرائق الاخرى وتحليل نتائجها .

***بحث مستل من الاطروحة (1)**

من اهم المؤلفات الاحصائية التي اهتمت بموضوع القياسات المكررة باتجاه واحد كتاب Winer في عام(1971) و كذلك نجده في الكتب المخصصة لدراسة حالات متعددة المتغيرات مثل كتاب Timmفي سنة (1975) و Morrison في سنة (1976) . ومن اهم الدراسات التي اهتمت بهذا الموضوع دراسة كل من Akritas& Arnold في عام (1994) حيث اثبتا بان طريقة Rt هي مناسبة لاختبار الفرضيات عندما تكون لامعلمية بصورة كاملة بدلالة دوال التوزيع distribution function وفي حالة تصاميم القياسات المكررة باتجاه واحد وباتجاهين وافترضا بان $\sum$ غير محددة بصورة كاملة. ودراسة كل من Chen&Cheng&Wang في عام (1995) حول استخدام الرتب لمقارنة مجموعة من المعالجات مع معالجة السيطرة في تصميم القياسات المكررة باتجاه واحد . ودراسة كل من Rashid & Nandram في عام (1995) حول استخدام الرتب لتحليل تصاميم القياسات المكررة باتجاه واحد باستخدام مصفوفة تباين وتباين مشترك متغييرة المقادير. ودراسة كل من Shah&Maddenفي عام (2004) حول التحليل اللامعلمي للبيانات الترتيبية في تصاميم التجارب العاملية متضمنا تصاميم القياسات المكررة باتجاه واحد . في عام (2004) نشر Antoniou بحثا يتعامل مع التصاميم العاملية بحيث ان كل وحدة قد شوهدت تحت مجموعة من نقاط الزمن وفيها بيانات مفقودة وقد تم استخدام الطريقة اللامعلمية في تقدير دالة التوزيع التجميعية الحدية عند كل نقطة من الزمن واستخدمت لاختبار تاثير العوامل والتفاعل باستخدام طرائق الـ Kernel المفردة .

الجانب النظري

يتناول هذا المبحث فكرة موجزة عن نموذج القياسات المكررة العام باتجاه واحد واهم الاختبارات المعلمية واللامعلمية لهذا النوع من التصاميم :

نموذج القياسات المكررة العام باتجاه واحد

ان نموذج القياسات المكررة العام (generalized repeated measures model) هو النموذج الذي فيه كل مفردة (individual) تستلم عدد من المعالجات r . اذ لايمكن اعتبار المشاهدات في هذه النماذج بانها مستقلة . وان كل وحدة تستلم نفس العدد r من المعالجات وان القياسات لكل وحدة تجريبية تتوزع متعدد التغيرات ذو بعد r . مع مصفوفة تباين وتباين مشترك $\sum$ بحيث لايفترض شكل محدد لمصفوفة التباين والتباين المشترك أي عندما لايفترض لها بان تكون كل القياسات لها نفس التباين وكل زوج من القياسات له نفس التباين المشترك أي بالشكل التالي:

$$\sum = \sigma^2 A(\rho),\ A(\rho) = \begin{bmatrix} 1 & \rho & . & . & \rho \\ \rho & 1 & . & . & \rho \\ . & . & . & . & . \\ \rho & . & . & . & 1 \end{bmatrix}$$

انما مصفوفة غير محددة .

لتكن البيانات لـ n من الوحدات المستقلة مرتبة كما في الجدول التالي:

Sampling unit	Conditions			
	1	2	…	r
1	y_{11}	y_{12}	…	y_{1r}
2	y_{21}	y_{22}	…	y_{2r}
3	y_{31}	y_{32}	…	y_{3r}
.	.			
.	.			
.	.			
n	y_{n1}	y_{n2}	…	y_{nr}

لتكن y_{ik} هي المشاهدات للمعالجة i^{th} للوحدة k^{th} .

لتكن $\underline{y_k} = (y_{1k}, y_{2k}, ..., y_{rk})', \ldots \quad k = 1,2,\ldots,n)$ هو متجه المشاهدات للوحدة k_{th} .

فان نموذج القياسات المكررة العام باتجاه واحد (One – way generalized repeated measures model) في هذه الحالة يكون بالشكل التالي :

$$y_{ik} = \mu + \alpha_i + e_{ik} \qquad \ldots(1)$$

$$\sum_i \alpha_i = 0$$

$$\underline{e_k} = (e_{1k}, e_{2k}, \ldots, e_{rk}) \propto N_r(\underline{O}, \Sigma)$$

حيث ان :

μ : المتوسط العام لكل المشاهدات

α_i : قياس التأثير المحدد للشرط i^{th}

e_{ik} : الحد العشوائي

وان $\underline{e_k}$ مستقلة.

اما اهم الاختبارات لهذا النوع من التصاميم فهي كما يلي :

1 - اختبار Arnold,S.F.

اقترح الباحث Arnold,S.F. (4) اختبارا لتحليل تصاميم القياسات المكررة للنموذج العام باتجاه واحد عندما تتوزع البيانات توزيعا طبيعيا ووفق النموذج (1) وعندما لايفترض شكل محدد لمصفوفة التباين والتباين المشترك .لغرض الاختبار بان

$$\alpha_i = 0 \qquad \ldots(2)$$

فان الاختبار الامثل لرفض مثل هذه الفرضية هو:

$$H_1(y) = c(n, r-1)(A\underline{\bar{y}})'(ASA')^{-1} A\underline{\bar{y}} > F^{\alpha}_{r-1,\, n-r+1} \qquad \ldots(3)$$

حيث ان :

$$c(n,s) = n(n-s)/s(n-1)$$

وان $\underline{\bar{y}}$: متجه المتوسطات للعينة اذ ان $\bar{y} = \dfrac{\sum_{k=1}^{n} y_k}{n}$

S: هي مصفوفة التباين والتباين المشترك للعينة حيث ان عناصرها:

$$S(K,K') = \frac{\sum_{k=1}^{n}(y_k - \bar{y})(y_k - \bar{y})'}{n-1}$$

وان:

$$S(K,K) = \frac{\sum_{k=1}^{n}(y_k - \bar{y})^2}{n-1}$$

وان: A=(I-1) هي مصفوفة تقابلات (contrast) حيث ان صفوفها مستقلة ومجموعها الى الصفر وان I هي مصفوفة احادية ببعد (r-1)*(r-1) . وان 1 هو متجه من الواحدات ببعد (r-1)*1.

2 - اختبار Akritas& Arnold

ان الباحثين(1994) Akritas& Arnold (3) اثبتا بان طريقة RT. مناسبة دائما لتصاميم القياسات المكررة لاختبار الفرضيات التي هي لامعلمية بصورة كاملة (Fully non parametric) بمعنى انها معرفة بمصطلح دالة التوزيع(Distribution function) . وذلك من خلال مقترحهما التالي:

ان النموذج باتجاه واحد one-way اللامعلمي يفترض فقط بان y_{ik} لها توزيع F_i .

لتكن

$$F_i(u) = \mu(u) + A_i(u) \qquad ...(4)$$

$$\sum_i A_i(u) = 0$$

في النموذج اللامعلمي (4) ان فرضية الاختبار هي :

$$A_i(u) = 0 \quad ...(5)$$

أي ان : $F_i(u)$ كلها متساوية.

احصاءةالاختبار

لتكن $k = 1,2,3,...,n$, $\underline{y_k} = (y_{1k}, y_{2k}, ..., y_{rk})'$ هي متجهات عشوائية تتوزع بصورة مستقلة ومتماثلة ولها دوال توزيع مستمر ولم يفترض متجه متوسطات اومصفوفة تباين و تباين مشترك محددة.

ولغرض اختبار فرضية العدم والتي تنص على تساوي F_i يكون وفق مايلي:

$$(r-1)H_1(\varsigma) = (r-1)c(n,r-1)(A\underline{\bar{Z}})'(AT\,A')^{-1}A\underline{\bar{Z}} \approx \chi^2(r-1) \quad ...(6)$$

حيث ان :

$$c(n,s) = n(n-s)/s(n-1)$$

$\underline{\bar{Z}}$ و T هما متجه متوسطات العينة ومصفوفة التباين المشترك لـ $\underline{Z_k}$

$$\underline{Z_k} = (Z_{1k}, Z_{2k}, ... Z_{rk})'$$

$$Z_{ik} = G(y_{ik})$$

$$Z_{ik}^{*} = \hat{G}(y_{ik}) \qquad ... \}(7)$$

$$\hat{G}(u) = \sum_i \frac{\hat{F_i}(u)}{r}$$

$$= \sum_{i=1}^{r}\sum_{k=1}^{n} \frac{I(y_{ik} \le u)}{rn}$$

الرتب لـ y_{ik} في المجموعة $\{y_{jm}\}$ هي:

$$nr\,\hat{G}(y_{ik}) = nrZ^*_{ik} \qquad ...(8)$$

A :هي مصفوفة التقابلات بدرجة r*(r-1) المعرفة في الفقرة (2)
ولهذا سوف ترفض فرضية العدم في حالة $(r-1)\,H_1(\varsigma\ ) > \chi^{2\alpha}_{r-1}$

3- الطريقة المقترحة

لتكن المشاهدة y_{ik} حيث ان $y_{ik} \sim F_i$ وان F_i تشير الى دالة التوزيع لـ y_{ik}

وان $$F_i(x) = \frac{1}{2}\left[F_i^+(x) + F_i^-(x)\right]$$

بحيث ان : $F_i^+(x)$ التوزيع المستمر من جهة اليمين.

$F_i^-(x)$ التوزيع المستمر من جهة اليسار.

وان

$$F_i^+(x) = p(y_{ik} \le x)$$

$$F_i^-(x) = p(y_{ik} < x)$$

ليكن النموذج اللامعلمي العام

$$F_i(x) = \mu(x) + A_i(x) \qquad ...(9)$$

$$, \sum A_i = 0$$

فلاختبار الفرضية

$$H_0 = A_i(x) = 0 \qquad ..(10)$$

وذلك بان جميع الـ F_i متساوية.

لتكن $F = (F_1, F_2, ..., F_r)$ متجه لدوال التوزيع F_i وان الفرضية اللامعلمية للفرضية اعلاه تكتب :

$$H_0 = CF = 0 \qquad ...(11)$$

C : مصفوفة التقابلات (contrast)

ان دالة التوزيع التجريبي .(Empirical distribution function) لـ F_i يرمز لها بـ

$$\hat{F}_i = \frac{1}{2}\left[\hat{F}_i^{+}(x) + \hat{F}_i^{-}(x) \right]$$

$$= \frac{1}{n}\sum_{k=1}^{n} c(x - y_{ik})$$

اذ ان

$$c(u) = \frac{1}{2}\left[c^{+}(u) + c^{-}(u)\right]$$

تشير الى دالة الحساب حيث ان $c^{+}(u) = 0 \text{ or } 1$ طبقا الى $u < 0 or \geq 0$ و
$c^{-}(u) = 0 \text{ or } 1$ طبقا الى $u \leq 0 or > 0$
وان الرتبة للمشاهدة y_{ik} بين كل من N من المشاهدات تعطى بواسطة

$$R_{ik} = \frac{1}{2} + \sum_{j=1}^{r}\sum_{m=1}^{n} c(y_{ik} - y_{jm}) \qquad ...(12)$$

وان ½ تضاف الى من اجل الحصول على موقع الاعداد للمشاهدات المرتبة في حالة عدم وجود تكرار (ties) بما ان c(0)=1/2 وكذلك فان R_{ik} هي متوسط الرتب في حالة وجود تكرار (ties).

ان الاحصاءة المقترحة هي مماثلة لاحصاءة T-Hotteling لتصاميم القياسات المكررة . ولكن باعطاء الرتب للمشاهدات وابدال مصفوفة التقابلات A المذكورة سابقا بمصفوفة التقابلات التالية:
لتكن C هي مصفوفة (تقابلات) بدرجة (r*r) والتي تساوي:

$$C = I_i - (1/i)J$$

حيث ان : $J_i = 1_i 1_i'$

$1_i = (1,1,\ldots 1)'$

و I_i : مصفوفة الوحدة بدرجة i*i
وبواسطة المصفوفة C سوف نختبر الفرضية (10) و (11).

لتكن

$$\overline{R_{i.}} = \frac{\sum_{k=1}^{n} R_{ik}}{n}$$

$$\underline{\overline{R}} = \left[\overline{R_{1.}}, \overline{R_{2.}}, \ldots, \overline{R_{r.}}\right]'$$

لتكن مصفوفة $\hat{V}$ تشير الى مصفوفة التباين والتباين المشترك مع العناصر $\hat{\sigma}_{ik}^{2}$. حيث ان :

$$\hat{\sigma}_{ik}^{2} = \left[(n-1)\right]^{-1}\sum_{k=1}^{n}(R_{ik} - R_{i.})(R_{ik} - R_{i.})'$$

لهذا فان احصاءة الاختبار سوف تكون بالشكل التالي:

$$Su = n\underline{\bar{R}}'\, C'\,(C\hat{V}\,C')^{-1}\,\underline{\bar{R}}C \qquad ...(13)$$

فان Su تتوزع χ^2_{r-1} (r : عدد المعالجات).

اختبار Mardia (1975)

لتكن X1,X2,…,Xn هي عينه عشوائية من مجتمع بمتجه متوسطات μ ومصفوفة تباين وتباين مشترك Σ افترض بان $\overline{X}$ و S تشير الى متجه متوسطات العينة ومصفوفة التباين والتباين المشترك على التوالي. ولغرض اختبار فرضية العدم بان مشاهدات العينة العشوائية تتبع التوزيع الطبيعي متعدد المتغيرات فقد اعتمد Mardia على القياسات التالية :

لغرض اختبار فرضية العدم

$$H_0 : b_{1,p} = 0 \qquad ...(14)$$

فان مقياس الالتواء المتعدد يحسب كما يلي:

$$b_{1,p} = \frac{1}{n^2}\sum_{i,j=1}^{n} r_{ij}^3 \qquad ...(15)$$

$$r_{ij} = (X_i - \overline{X})' S^{-1} (X_j - \overline{X})$$

ولاختبار الفرضية

$$H_0 : b_{2,p} = p(p+2) \qquad ...(16)$$

يستخدم مقياس التفرطح المتعدد حيث ان:

$$b_{2,p} = \frac{1}{n}\sum_{i=1}^{n} r_{ij}^2 \qquad ...(17)$$

وتقارن القيمة المستخرجة لـ $b_{1,p}$ و $b_{2,p}$ مع القيم الجدولية فاذا كانت القيمة المستخرجة اقل من الجدولية لاترفض فرضية العدم بان البيانات تتوزع توزيع طبيعي متعدد المتغيرات .

الجانب التجريبي

لقد تم استعمال اسلوب المحاكاة لغرض المقارنة بين طرائق الاختبارات المختلفة لتصميم القياسات المكررة للنموذج العام باتجاه واحد لاختبار فرضية عدم وجود فرق في تاثير المعالجات المختلفة وبالتالي الوصول الى الطريقة المثلى . فقد تم الاعتماد على طريقة مونت كارلو في توليد البيانات التي تتبع النموذج المذكور اعلاه وكما يلي :

وصف تجربة المحاكاة

ان تجربة المحاكاة لتصميم القياسات المكررة للنموذج العام باتجاه واحد جاءت ايضا للاجابة عن التساؤلات المذكورة نفسها في الحالة السابقة والتي تتضمن عملية المقارنة بين مجموعة من الطرائق لتصميم القياسات المكررة العامة باتجاه واحد من خلال تغيير مجموعة من المؤشرات التي تؤثر على النموذج وكما هي في الحالة الأولى وهذه المؤشرات هي:

- نوع التوزيع للمتغير العشوائي
- عدد المعالجات الداخلة في العينة

- حجم العينة
- تاثير المعالجات الداخلة في العينة
- مصفوفة التباين والتباين المشترك

تم الاجابة على تلك التساؤلات من خلال استخدام معياري احتمال الخطأ من النوع الأول وقوة الاختبار للطرائق المذكورة في الفصل السابق اذ تم توليد بيانات لمجموعة من التجارب المختلفة والتي تتبع توزيعا معينا وتختلف من حيث عدد المعالجات وعدد المشاهدات بعدد يمكننا من تطبيق جميع الاختبارات وبالتالي المقارنة بينها وعند توليد البيانات تم الاخذ بنظر الاعتبار عدد المعالجات فقد تم افتراض اعداد المعالجات $k=3,4$ لمعرفة سلوك الاختبار عند تغير اعداد المعالجات الداخلة في البحث.

لكل حالة مفترضة لعدد المعالجات فقد افترضنا حجوما مختلفة للعينات وهي $n=10,20,30$ وذلك لقياس اثر تغيير حجوم العينات على سلوك الاختبارات . وكذلك تم تحديد قيم تأثير المعالجات (τ_j's)وبمستويات مختلفة اذ تم اختيار مجاميع مختلفة لتأثير المعالجات اذ إن المجموعة التي فيها جميع τ_j's مساوية للصفر تستخدم لملاحظة سيطرة الطرائق المختلفة على احتمال الخطأ من النوع الأول. والمجاميع الاخرى المختلفة من التأثيرات للمعالجات تستخدم لدراسة قوة الاختبار لاختبار الطرائق عندما تكون فرضية العدم غير صحيحة . و كانت تلك التأثيرات مرتفعة و منخفضة و متغيرة لمعرفة سلوك احصاءات الاختبار وتاثيرها على قوة الاختبارتحت الدراسة.

جدول (1) تأثيرات المعالجات حسب عدد المعالجات

عدد المعالجات / تاثير المعالجة	K=3	K=4
$\tau 1$	(0,0,0) '	(0,0,0,0) '
$\tau 2$	(-15,0,15) '	(0,0,-20,20) '
$\tau 3$	(-6.5,-6.5,13) '	(-6,-6,-6,18) '
$\tau 4$	(-6,-4,10) '	(-17,-9,9,17) '

كذلك افترضت مصفوفات تباين وتباين مشترك وبارتباطات مختلفة بين المشاهدات كما يظهر في الجدول التالي لتبيان اثر تغيير الارتباطات بين المشاهدات على معياري احتمال احتمال الخطا من النوع الاول وقوة الاختبار.وفيما يلي مصفوفات التباين والتباين المشترك المستخدمة في البحث :

1 – عندما يكون عدد المعالجات $k=3$ فان المصفوفات المستخدمة هي:

$$\Sigma_1 = \begin{bmatrix} .945 & .184 & .190 \\ & .945 & .188 \\ & & .945 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} .945 & .809 & .745 \\ & .945 & .754 \\ & & .945 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} .957 & .475 & .242 \\ & .957 & .478 \\ & & .957 \end{bmatrix}$$

$$\Sigma_4 = \begin{bmatrix} .942 & .190 & .194 \\ & .942 & -.177 \\ & & .942 \end{bmatrix}, \Sigma_5 = \begin{bmatrix} .955 & -.766 & -.783 \\ & .955 & .771 \\ & & .955 \end{bmatrix}$$

$$\Sigma_6 = \begin{bmatrix} .943 & -.470 & .234 \\ & .943 & .472 \\ & & .943 \end{bmatrix}$$

2- عندما يكون عدد المعالجات $k=4$ فان المصفوفات المستخدمة هي:

$$\Sigma_1 = \begin{bmatrix} .945 & .184 & .190 & .174 \\ & .945 & .188 & .188 \\ & & .945 & .182 \\ & & & .945 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} .945 & .809 & .745 & .770 \\ & .945 & .754 & .760 \\ & & .945 & .754 \\ & & & .945 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} .957 & .475 & .242 & .335 \\ & .957 & .478 & .283 \\ & & .957 & .479 \\ & & & .957 \end{bmatrix}$$

$$\sum_4 = \begin{bmatrix} .942 & -.190 & .-194 & .188 \\ & .942 & .177 & .185 \\ & & .942 & .192 \\ & & & .942 \end{bmatrix}, \sum_5 = \begin{bmatrix} .955 & -.766 & -.783 & -.794 \\ & .955 & .771 & .768 \\ & & .955 & .759 \\ & & & .955 \end{bmatrix}$$

$$\sum_6 = \begin{bmatrix} .943 & -.470 & .234 & .119 \\ & .943 & .472 & .237 \\ & & .943 & .472 \\ & & & .943 \end{bmatrix}$$

وتم دراسة الحالات المذكورة في حالة التوزيع الطبيعي متعدد المتغيرات وتوزيع Multivariate –t وتوزيع Multivariate –cauchy

وكان عدد التجارب التي اجريت (120) تجربة ولكل التوافيق لـ n,k وعلى كل توزيع من التوزيعات الثلاثة المذكورة وتم سحب 1000 عينة لكل حالة. وفي جميع الحالات المذكورة تم افتراض مستويين للمعنوية هما 01. و05.

سير عملية المحاكاة

في البدء تم صياغة انموذج محاكاة للصيغة (1) وللتوزيعات المذكورة اعلاه يليه توليد المتغير العشوائي ϵ_{ij} وحسب التوزيعات المذكورة اعلاه. و يتم تحديد عدد المعالجات(k) وحجوم العينات(n) وقيم تأثيرات المعالجات τ_j's وحسب الحالة المفترضة .وتضاف تلك التأثيرات بعضها الى بعض وقد تم اضافة قيمة المتوسط العام μ والتي افترضت مساوية لـ100 وبذلك سوف نحصل على قيمة المشاهدة X_{ij} .

وبعد ان يتم توليد جميع مشاهدات بيانات التجربة لجميع العينات العشوائية المستقلة ولكل التوزيعات المذكورة سابقا وللحالتين تبدا عملية المقارنه بين الطرائق المدروسة ولكافة التشكيلات المدروسة المختارة على اساس احتمال الخطأ من النوع الأول وقوة الاختبار للمقارنة بين الاختبارات للوصول الى افضل اختبار.

توليد متغيرات ذات توزيع طبيعي متعدد المتغيرات

يتم توليد القيم Zi التي تتوزع توزيعا طبيعيا قياسيا بمتوسط 0 وتباين 1 حسب طريقة Box-Muller المذكورة سابقا بحيث ان $Z_i \sim N(0,1)$ ثم توليد المصفوفة Z من القيم حيث

$$Z \approx MN_r(\underline{0}, I)$$

وللحصول على مصفوفة المتغيرات X بمتجه متوسطات $\underline{\mu}$ ومصفوفة تباين $\sum$ يكون باستخدام الصيغة نفسها في التوزيع الطبيعي المنفرد . حيث تصبح هنا

$$X = \underline{\mu} + ZC$$

بافتراض ان $\sum = CC'$

ومصفوفة C هي مصفوفة مثلثية سفلى برتبة (r*r) وتمثل الجذر التربيعي للمصفوفة Σ ويتم استخراجها باستخدام التجزئة بواسطة طريقة choelski حسب المعادلات الاتية:

$$t_{11} = a_{11}^{1/2} \quad t_{1j} = a_{1j} = a_{1j}\, t_{11}^{-1} \quad (1 \le j \le n)$$

$$t_{ii} = (a_{ii} - \sum_{k=1}^{i-1} t_{ki}^2)^{1/2} \qquad (1 < i \le n)$$

$$t_{ij} = (a_{ij} - \sum t_{ki}\, t_{kj}) t_{i1}^{-1} \qquad (1 < i < j \le n)$$

$$t_{ij} = 0 \qquad (1 < j < i \le n)$$

حيث ان t_{ij} هو العنصر i,j من المصفوفة Σ المفترضة.

وبهذا نحصل على المصفوفة $X \approx MN_r(\underline{\mu}, \Sigma)$

<u>توليد متغيرات عشوائية ذات توزيع Multivariate –t</u>

لغرض توليد متغيرات عشوائية ذات توزيع Multivariate –t نتبع ما يلي:

1) توليد المتغيرات $\underline{X}$ والتي تتبع التوزيع متعدد المتغيرات الطبيعي .

2) توليد المتغير y والذي يتبع التوزيع $\chi 2$ بدرجة حرية V والذي يكون عبارة عن حاصل مجموع مربع المتغير الطبيعي القياسي أي ان :

$$y \approx \chi^2_{\nu} = \sum_{i=1}^{\nu} Z_i^2$$

وان Z_i هو المتغير الطبيعي القياسي المتولد حسب طريقة Box-muller.

ايجاد قيم المتغير u_i باستخدام الصيغة التالية :

$$g_i = \frac{x_i}{\sqrt{y/\nu}} \qquad ...(18)$$

وبذلك يتم الحصول على المصفوفة g والتي تتوزع توزيع Multivariate –t

<u>توليد متغيرات عشوائية ذات توزيع Multivariate – cauchy</u>

عندما تكون درجة الحرية v=1 في الصيغة (18) عند ئذ نتبع نفس الاجراءات المتبعة في توليد متغيرات عشوائية تتوزع Multivariate –t ولكن بدرجة حرية v=1

<u>تحليل نتائج المحاكاة</u>

من خلال ملاحظة النتائج الخاصة باحتمال الخطأ من النوع الأول وقوة الاختباروحسب الترتيب يمكن ابداء الملاحظات التالية :

<u>التوزيع الطبيعي متعدد المتغيرات</u>

من ملاحظة نتائج احتمال الخطأ من النوع الأول يتبين بان الاختبار المعلمي AR لتصميم القياسات المكررة العامة باتجاه واحد هو افضل الاختبارات وذلك لاقتراب احتمال الخطأ من النوع الأول من قيمة (α) الاسمية يليه الاختبار اللامعلمي المقترح Su ونلاحظ ابتعاد احتمال الخطأ من النوع الأول للاختبار اللامعلمي Ak من مستوى المعنوية (α) . وا ن اعلى قوة للاختبار بين مجموعة الاختبارات الداخلة في المقارنة لهذا التوزيع عندما عدد المعالجات k=3,4 كانت للاختبار AR على الاغلب وذلك لامتلاكها اعلى قوة اختبار بين الاختبارات الاخرى وكان بالمرتبة الثانية الاختبار Su لامتلاكه ثاني اكبر قوة اختبار لاغلب الحالات وعند جميع مصفوفات التباين والتباين المشترك وعند مختلف تأثيرات المعالجات وعند حجوم العينات المختلفة.

التوزيع T – متعدد المتغيرات

من ملاحظة نتائج احتمال الخطأ من النوع الأول يظهر بان الاختبار اللامعلمي المقترح Su هو افضل الاختبارات وذلك لاقتراب احتمال الخطأ من النوع الأول من قيمة (α) الاسمية يليه الاختبار المعلمي AR .

ان الطريقة المقترحة Su قد سجلت اعلى قوة للاختبار على الاغلب بين مجموعة الطرائق الاخرى لاغلب الحالات ولجميع مصفوفات التباين والتباين المشترك ولمختلف تأثيرات المعالجات و حجوم العينات تلاها بالمرتبة الثانية الاختبار AR الذي يمتلك على الاغلب ثاني اعلى قوة اختبار بين الاختبارات الاخرى وفي الحالتين عندما يكون k= 3 , 4 .

التوزيع Cauchy- متعدد المتغيرات

من ملاحظة نتائج احتمال الخطأ من النوع الأول يظهر بان الاختبار اللامعلمي المقترح Su هو افضل الاختبارات وذلك لاقتراب احتمال الخطأ من النوع الأول من قيمة (α) الاسمية يليه الاختبار للاختبار المعلمي AR وجاء ثالثا الاختبار اللامعلمي Ak .

ان الطريقة المقترحة Su قد سجلت وفي حالة عدد المعالجات k= 3 اعلى قوة للاختبار بين مجموعة الطرائق الاخرى لاغلب الحالات ولجميع مصفوفات التباين والتباين المشترك وعند مختلف تأثيرات المعالجات وعند حجوم العينات المختلفة تلاها بالمرتبة الثانية الاختبار Ak الذي يمتلك على الاغلب ثاني اعلى قوة اختبار بين الاختبارات الاخرى .

اما في حالة عدد المعالجات k=4 فقد كان تسلسل قوة الاختبار لاغلب تاثيرات المعالجات ولجميع مصفوفات التباين والتباين المشترك ومختلف حجوم العينات بصورة متغيرة من حالة الى اخرى.

من خلال ملاحظة نتائج احتمال الخطأ من النوع الأول عندما α=.01 وعندما يكون عدد المعالجات k=3,4 نجد بان الاختبار AR احتل المرتبة الأولى من حيث اقتراب احتمال الخطأ من النوع الأول من قيمة مستوى المعنوية α في حالة التوزيع الطبيعي المتعدد المتغيرات. يليه في المرتبة الثانية الاختبار المقترح Su اما في حالة التوزيعين T - المتعدد المتغيرات و Cauchy - المتعدد المتغيرات فنجد بان الاختبارSu ثم الاختبار AR حلا بالمرتبتين الأولى والثانية من حيث اقتراب احتمال الخطأ من النوع الأول من قيمة مستوى المعنوية α ومن ملاحظة قوة الاختبار عندما α=.01 وفي حالة k=3,4 سوف نجد بان الاختبار AR كان اعلى قوة للاختبار في حالة التوزيع الطبيعي تليه كل من الاحصاءتين Su ثم Ak . اما بالنسبة للتوزيع T - المتعدد المتغيرات فقد كان الاختبار المقترح Su اكثر قوة من جميع الاختبارات الاخرى يليه كل من الاختباران AR و Ak على التوالي. اما بالنسبة الى قوة الاختبارللتوزيع Cauchy - المتعدد المتغيرات فقد حل الاختبار المقترحSu المرتبة الأولى لامتلاكه اكثر قوة من الاخرين وفي حالة عدد المعالجات k=3,4 تلاه الاختبارAk وبالمرتبة الثانية هذا في حالة k=3 اما في حالة k=4 فان المرتبة الثانية تناوب عليها الاختباران الاخران وبصورة متغيرة بين تأثيرات المعالجات ومصفوفات التباين والتباين المشترك.

وكمثال على النتائج التي تم الحصول عليها فان الجدول (3) يبين نتائج قوة الاختبار عندما k=3 ولـ (-15,0,15) =2 τ ولمستويي المعنوية .05, .01=α.

الجانب التطبيقي

خصص هذا المبحث الى دراسة بيانات تجربة حقيقة لتطبيقها على الطرائق التي تم عرضها في الجانب النظري لتوضيح الفرق بين نتائج الطرائق المدروسة وقد تمت كتابة البرامج الخاصة بالطرائق المذكورة بلغة Qbasic.

التجربة

تبين هذه التجربة نتائج تجربة اوزان الحملان والتي عددها (28) حملا اذ وزنت جميع الحملان الحديثة الولادة خلال الساعات الاربعة والعشرون الاولى من عمرها واعتبر هذا ممثلا لوزن الحمل الابتدائي . عزلت جميع الحملان التي ظهرت عليها اعراض مرض الكولي باسيللوسز والتي عولجت بمادة الابرالان عن طريق الفم في اليوم الاول لظهور المرض وعددها (28) حملا. ثم اعطيت الجرعات الاخرى في اليومين التاليين ثم اخذت اوزان الحملان عند اليوم السابع والرابع عشر من عمر الحيوان كما يظهر في الجدول (2) . اذ سوف يتم تطبيق كافة الاختبارات عليها وكما يلي:

اختبار التوزيع الطبيعي للبيانات

لاختبار البيانات من حيث اتباعها للتوزيع الطبيعي ام لا يتم استخدام اختبار الالتواء واختبار التفرطح ، اذ تم احتساب الاحصاءة $b_{1,3}$ و$b_{2,3}$ من الصيغتين (15) و(17) على التوالي وكانت قيمة$b_{1,3}$ =6.3825 وقيمة $b_{2,3}$=18.4168 وكانت القيم المحسوبة اكبر من القيمة الجدولية اذ كانت $b_{1,3}$ = 3.5 لمستوى 0.05 وتساوي 4.7 لمستوى 0.01 . وبمقارنة قيمة $b_{2,3}$ مع القيمة الجدولية وللمستويين 0.05 و0.01 اذ كانت تساوي 16.7 و18.3 على التوالي وبما ان القيمة المستخرجة اكبر من القيمة الجدولية لذا فان البيانات لاتتبع التوزيع الطبيعي.

-الاختبارAR

لاختبار الفرضية (2) بعدم وجود فروق معنوية بين اوزان الحملان لاسابيع التجربة. فقد تم احتساب الاحصاءة (3) اذ كانت قيمة $H_1(y)$ تساوي 316.789 وعند مقارنتها مع قيمة F(2,26) عند المستويين 0.05 و0.01 حيث كانت تساوي 3.37 و 5.53 على التوالي لذا نجد بان الفروق جدا معنوية اذ ترفض فرضية العدم .

-الاختبار اللامعلمي Akritas &Arnold (Ak)

لاختبار الفرضية (5) يتم استخراج Z_{ik} حسب الصيغة (7) وذلك بدمج قيم مشاهدات التجربة بعينة واحدة يرمز لها بـ Y_{jm} ثم يتم حساب لكل مشاهدة Y_{ik} في التجربة عدد القيم التي هي اقل منها او تساويها من قيم العينة المدمجة Y_{jm} فتنتج قيم Z_{ik} لكل مشاهدة وتستبدل القيمة الاصلية للمشاهدة بتلك القيمة .ويتم استخراج الاحصاءة (6) التي كانت قيمتها تساوي H1(ς)=298.7119 وبمقارنتها مع قيمة مربع كاي$\chi^2_{(2)}$ عند المستويين 0.05 و0.01 التي تساوي 5.991 و9.2103 على التوالي . لذا سنرفض فرضية العدم بمعنى وجود فروق معنوية بين اوزان الحملان للاسابيع الثلاثة .

-الطريقة المقترحة (Su)

لاختبار الفرضية (10) يتم حساب الرتبة للمشاهدة y_{ik} حسب الصيغة (12) ثم تستخرج احصاءة الاختبار المقترحة حسب الصيغة (13) اذ كانت القيمة المستخرجة Su =736.666 وبمقارنتها مع قيمة$\chi^2_{(2)}$ المذكورة في الفقرة السابقة وعند المستوين المذكورين يتم رفض فرضية العدم بمعنى وجود فروق بين اوزان الحملان تحت الاسابيع الثلاثة

.

جدول (2)*

اوزان حملان المجموعة العلاجية التي اعطيت المضاد الحيوي (ابرامايسين) عن طريق الفم لعلاج مرض الكولي باسيللوسز (الاصابة بالعصية القولونية) في الحملان

تسلسل الحمل	وزن الحمل بعمر يوم	وزن الحمل بعمر اسبوع	وزن الحمل بعمر اسبوعين
1	5	6.75	9
2	3.5	5.25	7.5
3	5	7	9.5
4	4	6	8.5
5	4	6.75	8.5
6	4.5	7.5	13
7	5	7.75	9.5
8	5	7.5	10.5
9	4.5	6.75	9
10	5	7.5	10
11	4.5	6.75	9.5
12	3.5	8	8
13	6	8	9
14	4.5	5.5	8.5
15	5	6.25	8.25
16	4.5	6.5	9
17	4.5	5.75	8.5
18	3.75	6	8
19	5	6.5	8
20	4.5	7.5	12
21	5.5	6.5	8
22	5.5	7.75	9.5
23	4.5	7.5	8.5
24	5	5.5	6.5
25	4	7.5	8.5
26	4.5	7.5	9.5
27	4	5	7.5
28	4.5	7	8.5

* المصدر (1)

يتضح من النتائج اعلاه بان جميع نتائج الاختبارات الثلاثة كانت متوافقة من حيث رفض فرضية العدم .

الاستنتاجات

1-اظهرت النتائج بان اعلى قوة للاختبار كانت للاختبار AR الذي جاء بالمرتبة الاولى في حالة التوزيع الطبيعي متعدد المتغيرات عند اغلب الحالات ولجميع مصفوفات التباين والتباين المشترك وعند مختلف تاثيرات المعالجات ومختلف حجوم العينات اما بالنسبة للتوزيعين T – متعدد المتغيرات و Cauchy - متعدد المتغيرات فقد كانت الاختبارات اللامعلمية هي الاكثر قوة وبصورة متباينة ومختلفة وخاصة عند التوزيع و Cauchy - متعدد المتغيرات.

2-اظهرت نتائج قوة الاختبار ولجميع التوزيعات ومصفوفات التباين والتباين المشترك وتأثيرات المعالجات المختلفة بانها تزداد وعلى العموم بازدياد حجوم العينات.

3- كانت نتائج قوة الاختبار عندما α=.01 مماثلة في ترتيبها وعلى الاغلب مع نتائج α=.05 ولجميع التوزيعات .

المصادر

1- الهاشمي، سجى محمد حسين (2005) " مقارنة بعض الطرائق المعلمية واللامعلمية لبعض تصاميم القياسات المكررة " اطروحة دكتوراه في الاحصاء ، كلية الادارة والاقتصاد ،جامعة بغداد

2-Antoniou, E.S.(2004)"Nonparametric Imputation And (MiD-)Rank Tests For Mixed Effects Models With Missing Data"A PH.D thesis in Statistics ,The Pennsylvania State University .

3-Akritas,M.G.and Arnold, S.F.(1994)"Fully nonparametric hypotheses for factorial designs I:multivariate repeated measures designs".JASA,Vol 89 ,no 425.

4-Arnold,S.F.(1981)"Theory of linear models and Multivariate analysis ",john wiley &sons.

5-Brunner, edgar &Osgood,D.Wayne (2002),"Rank Procedures for repeated measures with missing values" Sociological Methods &research , vol. 30,N0. 3

6-Chen,Yuh-Ing,Cheng,Sheng-shu,Wang,Hong-long(1996)"Rank-based comparisons of treatments with acontrol for repeated measures designs.

7-Mardia,K.V. (1970)"Measures of multivariate Skeweness and kurtosis with applications " Biometrika vol.57, no.1.

8-Mardia,K.V. (1974)" Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies". Sankhya, vol. 36, series B,pt.2.

9-Mardia,K.V. (1975)"assessment of multinormality and the robustness of Hottellings T2 test" Applied Statistics .vol.24, no.2.

10-Mardia,K.V., Bibby,J.M., Kent,J.T. (1979) " multivariate analysis" Academic press inc.(London).ltd

11-Morrison,D.R.,(1976)"Multivariate Statistical Methods" Mc Graw-Hill , New york.

12-Rashid,M.M. & Nandram,B.(1995)"Arank-based analysis of one-way repeated measures designs with achanging covariate" J.stat.plann. & infer.

13-Shah,D.A.& madan l.V.(2004) "Nonparametric Analysis of ordinal Data in Designed Factorial Experiments",The American Phytopathological Society , vol,94,No.1

14-Timm,N.H. (1975)" multivariate Analysis with Applications education and Psychology "

15-Winer,B.J.(1971)" Statistical Principles in Experimental Design" Mc Graw-Hill ,inc.

جدول (3)

قوة الاختبار في حالة عدد المعالجات K= 3 وعندما (-15,0,15) =2 τ ولمستويي المعنوية

مستوى المعنوية			α = .05			α = .01		
مصفوفة التباين والتباين المشترك	حجم العينة	الاختبار / التوزيع	AR	Ak	Su	AR	Ak	Su
Σ1	N=10	Mul-N	.746	.472	.577	.625	.289	.295
		Mul-t	.013	.037	.2	.004	.005	.090
		Mul-c	.91	.976	.974	.807	.877	.871
	N=20	Mul-N	.965	.823	.895	.848	.602	.838
		Mul-t	.103	.044	.248	.038	.012	.100
		Mul-c	.943	.987	.981	.842	.862	.872
	N=30	Mul-N	.994	.955	.939	.975	.827	.912
		Mul-t	.113	.1	.901	.801	.020	.248
		Mul-c	.988	.995	.992	.974	.990	.982
Σ2	N=10	Mul-N	.152	.078	.267	.068	.024	.151
		Mul-t	.098	.1	.185	.016	.024	.100
		Mul-c	.07	.985	.836	.065	.875	.630
	N=20	Mul-N	.355	.112	.224	.29	.039	.162
		Mul-t	.183	.04	.195	.035	.014	.198
		Mul-c	.898	.989	.994	.188	.980	.781
	N=30	Mul-N	.797	.605	.769	.695	.422	.398
		Mul-t	.703	.988	.318	.055	.382	.108
		Mul-c	.986	.984	.992	.788	.991	.786
Σ3	N=10	Mul-N	.288	.081	.096	.186	.032	.022
		Mul-t	.095	.032	.223	.022	.007	.113
		Mul-c	.879	.99	.994	.553	.58	.893
	N=20	Mul-N	.51	.009	.50	.343	.064	.178
		Mul-t	.113	.078	.263	.034	.024	.125
		Mul-c	.855	.993	.993	.681	.884	.684
	N=30	Mul-N	.654	.429	.599	.405	1	.393
		Mul-t	.2	.229	.267	.049	.094	.163
		Mul-c	.915	.993	.996	.910	.991	.993

تكملة جدول (3)

قوة الاختبار في حالة عدد المعالجات K= 3 وعندما (-15,0,15) =2 τ ولمستويي المعنوية

مستوى المعنوية			α = .05			α = .01		
مصفوفة التباين والتباين المشترك	حجم العينة	الاختبار / التوزيع	AR	Ak	Su	AR	Ak	Su
Σ4	N=10	Mul-N	.638	.359	.47	.483	.219	.191
		Mul-t	.073	.044	.176	.012	.019	.089
		Mul-c	.912	.962	.973	.805	.946	.944
	N=20	Mul-N	.856	.623	.857	.725	.366	.628
		Mul-t	.132	.043	.192	.040	.015	.095
		Mul-c	.98	.994	.997	.979	.989	.981
	N=30	Mul-N	.976	.83	.909	.895	.608	.840
		Mul-t	.184	.123	.237	.065	.033	.117
		Mul-c	.977	.996	.998	.933	.992	.990
Σ5	N=10	Mul-N	.256	.057	.121	.142	.013	.038
		Mul-t	.169	.039	.209	.057	.010	.119
		Mul-c	.992	.978	.996	.476	.192	.891
	N=20	Mul-N	.285	.071	.246	.159	.020	.087
		Mul-t	.123	.052	.235	.033	.022	.123
		Mul-c	.62	.362	.979	.589	.270	.565
	N=30	Mul-N	.35	.109	.184	.24	.042	.120
		Mul-t	1	1	.994	.694	.742	.735
		Mul-c	.898	.994	.986	.773	.886	.594
Σ6	N=10	Mul-N	.122	.056	.064	.119	.025	.089
		Mul-t	.183	.055	.256	.055	.009	.134
		Mul-c	.542	.951	.958	.385	.533	.514
	N=20	Mul-N	.54	.118	.347	.211	.033	.129
		Mul-t	.163	.126	.219	.043	.041	.133
		Mul-c	.817	.99	.999	.809	.686	.614
	N=30	Mul-N	.299	.213	.562	.275	.075	.371
		Mul-t	.343	.059	.37	.144	.025	.238
		Mul-c	.978	.999	.996	.862	.898	.848