مجلة ابن الهيثم للعلوم الصرفة والتطبيقية المجلد23 (3) 2010

بعض الاختبارات الحصينة لمعامل الارتباط / دراسة مقارنة

نادية هاشم النور، هدى عبد الله رشيد، اقبال جبار سلطان
قسم الرياضيات ، كلية العلوم ،الجامعة المستنصرية

الخلاصه

اهتم البحث بدراسة اداء احصاءة اختبار فرضية استقلالية المتغيرين (الفرضية القائلة بعدم وجود ارتباط بين المتغيرين قيد الدراسة) في حالة استيفاء البيانات لشرط التوزيع الطبيعي وفي حالة الابتعاد عن تلك التوزيع نتيجة لوجود القيم الشاذة (الملوثة) ومقارنته مع اداء بعض الطرائق المحورة والمقترحة الاخرى .

المقدمة

يعد تقدير معامل الارتباط بين متغيرين عشوائيين ثم اختبار معنوية الارتباط من الأساليب الإحصائية التي تستخدم في الكثير من العلوم . ان الاساس النظري لاختبار معنوية الارتباط بين متغيرين عشوائيين من خلال اختبار فرضية استقلالية المتغيرين يقتضي توزيع البيانات توزيعاً طبيعياً ثنائياً [1] ، الا انه غالباً مايواجه الباحثون مشكلة ابتعاد البيانات عن التوزيع المفترض وذلك من خلال كون البيانات تتبع توزيعاً احتمالياً ثنائياً غير التوزيع الطبيعي او احتوائها على القيم الشاذة الامر الذي يؤدي الى انخفاض كفاءة الاختبار من خلال ارتفاع حجم الخطأ من النوع الأول ، أو انخفاض قوة الاختبار . استحوذ موضوع اختبار معنوية معامل الارتباط على اهتمام العديد من الباحثين فعلى سبيل المثال في عام 1981 ناقش *Conover* و *Iman* [2] موضوع التحويلات الرتبية بوصفها حلقة وصل بين الاحصاءات المعلمية واللامعلمية واوضحا امكانية استخدام اختبار t لاختبار معنوية معامل ارتباط سبيرمان الذي يعد احد التقديرات الحصينة لتقدير معامل الارتباط .

في عام 2001 اهتم *Wilcox* [3] بموضوع اكتشاف العلاقات اللاخطية فضلاً عن طرح بعض التعليقات المتعلقة بموضوع اختبار الفرضيات حول معامل الارتباط .

في عام 2006 اقترح *Ebd El-Salam* [4] مقياس بديل لمعامل الارتباط كطريقة حصينة لتقليل اثر وجود القيم الشاذة في البيانات ، اذ اشار الى ان وجود تلك القيم في البيانات يؤثر في تقدير معالم أنموذج الانحدار والاحصاءات المرتبطة بها ومنها قيمة معامل الارتباط .

ان البحث الحالي يهدف الى ايجاد الطرائق المثلى التي يمكن اعتمادها في اختبار معنوية معامل الارتباط عند ابتعاد البيانات عن التوزيع الطبيعي من خلال وجود بعض القيم الشاذة ، وذلك من خلال اقتراح بعض الطرائق التي تعتمد على التقديرات الحصينة لمعلمة الموقع فضلاً عن تطبيق بعض التحويلات الرتبية ومن ثم دراسة سلوك كلاً منها ومقارنتها مع بعضها باستخدام المحاكاة من خلال اعتماد الخطأ من النوع الاول وقوة الاختبار .

أختبار معامل الارتباط

اذا كان $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ ازواجاً من المشاهدات المستقلة التي تتبع التوزيع الطبيعي الثنائي (*Bivariate Normal Distribution*) مع دالة كثافة احتمالية بالصيغة :

$$......(1)\ f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right\}\right]$$

اذ ان :

$\mu_x, \mu_y, \sigma_x, \sigma_y$: الوسط الحسابي والانحراف المعياري للمتغيرين x ، y على التوالي .

ρ : معلمة الارتباط بين المتغيرين x ، y .

فأن تقدير معلمة الارتباط على وفق الصيغة الرياضية التي وضعها كارل بيرسون تكون معرفة على النحو الآتي **:[7 ،5,6]**

$$r_p = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad ......(2)$$

اذ ان :

$\bar{x}$ ، $\bar{y}$: الوسط الحسابي لقيم المتغير x وقيم المتغير y على التوالي .

ولاختبار معنوية الارتباط ، من خلال اختبار فرضية العدم (H_0) مقابل الفرضية البديلة (H_1) الموضحتين ادناه ،

$$\left.\begin{array}{l} H_0 : \rho = 0 \\ H_1 : \rho \neq 0 \end{array}\right\} \quad ......(3)$$

يتم اعتماد احصاءة t للاختبار والمتمثلة بالصيغة :

$$......(4)\ t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

اذ ان احصاءة الاختبار (t) تتوزع توزيع t بدرجة حرية $(n-2)$.

اذا وقعت قيمة احصاءة الاختبار في منطقة الرفض فيتم رفض فرضية العدم القائلة بعدم وجود علاقة بين المتغيرين او الظاهرتين ، $H_0 : \rho = 0$ ، وتتحدد المنطقة الحرجة حسب الفرضية البديلة ، $H_1 : \rho \neq 0$، كما يأتي :

$|t| > t_{\alpha/2,(n-2)}$ ، بمعنى يتم رفض H_0 اذا كانت $|t|$ المحسوبة على وفق الصيغة [4] اعلاه اكبر من t الجدولية بمستوى معنوية $(\alpha/2)$ ودرجة حرية $(n-2)$.

تقدير بديل عن معامل ارتباط بيرسون ، r_p ، الذي يمكن اعتماده في حالة ابتعاد البيانات عن التوزيع الطبيعي هو معامل الارتباط الرتبي (*Ordinal Correlation Coefficient*) لسبيرمان ، r_s ، المبني على اساس رتب البيانات والمتمثل بالصيغة :

$$......(5)\ r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

اذ ان :

d : الفرق بين رتب مستويات المتغير الاول x ومستويات المتغير الثاني y ، $d = R_x - R_y$.

الطرائق الحصينة

ان الطرائق الحصينة هي الطرائق التي تقاوم حالة عدم تحقق (اختراق) شرط واحد أو اكثر من شروط التقدير او الاختبار وقد ينتج ذلك من وجود بضع مشاهدات تنحرف بشكل ملحوظ عن بقية المشاهدات ، بمعنى آخر تكون غير متسقة معها ، فان مثل هذه المشاهدات تدعى بالقيم الشاذة (*Outliers*) التي غالباً ماتنشأ من توزيعات ثقيلة (متينة) الذيل (*Heavy Tailed Distributions*)، او من التوزيع المختلط (*Mixture Distribution*) ، او ناتجة عن اخطاء في الحسابات او القراءات او التسجيل وما الى ذلك .

اشار *Huber* [8] الى ان وجود قيمة شاذة واحدة فقط قد تفقد المقدرات التقليدية مزاياها الجيدة ، بمعنى ان وجود قيمة شاذة واحدة ضمن مجموعة المشاهدات قد يؤثر في بعض مقدرات الطرائق التقليدية ، اذ تكون هذه المقدرات الى حد ما حساسة لمثل هذه المشاهدات مما قد يجعلها اقل كفاءة لما هي عليه عند عدم وجودها وربما تفقدها الكثير من الخصائص الجيدة التي تتمتع بها . وفي ضوء ذلك ظهرت الحاجة الى ابتداع مقدرات/اساليب اكثر موضوعية تنطوي جميعها تحت مسمى الحصانة (*Robustness*) .

ان مصطلح الحصانة استعمل ولاول مرة [9] من *Box* عام (1953) ليشير الى ان الطريقة الاحصائية تعد من الطرائق الحصينة اذا كان الاستدلال لا يتأثر جداً نتيجة لاختراق اي من الشروط الاساسية لتلك الطريقة . اما فيما يتعلق بالاختبارات فان هناك رأيين حول هذا المفهوم [10] ، اذ يرى الرأي الاول ان الاختبار الاحصائي يعد حصيناً لاختراق اي من الشروط الاساسية لذلك الاختبار اذا لم يترتب على ذلك الاختراق اي تغيير واضح في اي من احتمالي الخطأ من النوع الاول والخطأ من النوع الثاني " الذي ينعكس بدوره على قوة الاختبار " . اما الرأي الثاني وتجنباً للتعقيد فيرى ان مفهوم الحصانة ينطبق على مفهوم الخطأ من النوع الاول من دون توسيعه ليشمل قوة الاختبار .

الطرائق المحورة والمقترحة

بغية الوقوف على اداء اختبار معنوية معامل بيرسون للارتباط وللتعرف على مدى تغير درجة حصانته تجاه القيم الشاذة عند اجراء بعض التحويرات ، لذا نقترح هنا بعض التحويرات التي غالباً مااعتمدت على المقدر الحصين لمعلمة الموقع بدلاً من الوسط الحسابي .

تمثلت الطرائق المقارن بينها والطرائق المحورة والمقترحة بالآتي :

1. TP: اعتماد معامل بيرسون للارتباط (الصيغة (2)
2. TM: اعتماد الوسيط (M) بدلاً من الوسط الحسابي عند تطبيق معامل بيرسون للارتباط.
3. TG: اعتماد مقدر *Gastwirth* ، المبني على أساس الوسيط واثنين من مشاهدات العينة المرتبة بدلاً من الوسط الحسابي عند تطبيق معامل بيرسون للارتباط ، ويكون على وفق الصيغة الاتية :

$$G = 0.3\ x_{[n/3+1]} + 0.4\ M + 0.3\ x_{(n-[n/3])} \quad ......(6)$$

اذ ان :

[n/3+1] : الجزء الصحيح من العدد الحقيقي (n/3+1) .

[n/3] : الجزء الصحيح من العدد الحقيقي (n/3) .

يعد الوسيط ومقدر *Gastwirth* [11] أحد المقدرات الخطية (*Linear Estimators* (*L.E.*)) الحصينة لمعلمة الموقع وتعرف تلك المقدرات بكونها تراكيب خطية موزونة لمشاهدات العينة المرتبة فاذا كانت $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ مشاهدات مرتبة لعينة بحجم (n) و $a_1, a_2 \ldots\ldots a_n$ اعداد حقيقية إذ ($0 \leq a_i \leq 1;\ i=1,\ldots\ldots,n$) تمثل اوزان تقترن مع المشاهدة $x_{(i)}$ ، فالصيغة العامة للمقدرات الخطية لمعلمة الموقع :

$$\hat{\mu} = \sum_{i=1}^{n} a_i\, x_{(i)} \qquad \ldots\ldots(7)$$

وعليه يشترط في الاوزان (a_i) عادةً ان تحقق القيد ($\sum_{i=1}^{n} a_i = 1$) لضمان تمتع المقدر بخاصية ثبات الموقع (*Location Invariance*) وان تكون $a_i = a_{(n+1-i)}$ عند افتراض ان توزيع المتغير العشوائي (x) متماثل حول معلمة الموقع μ كما تكون صغيرة للقيم المتطرفة (التي يمكن ان تكون ضمنها القيم الشاذة) بينما تكون كبيرة للقيم الأخرى وذلك للتقليل من تأثير القيم الشاذة ومن ثم ضمان الحصول على مقدر حصين .

4. TSP: اعتماد معامل سبيرمان للارتباط (الصيغة (5)).
5. TP*: استبدال القيم الواقعة خارج المجال $\bar{x} \pm s$ بالقيمة $\bar{x} \pm s$ لكل متغير ومن ثم تطبيق معامل بيرسون للارتباط ، اذ ان (s: الانحراف المعياري).
6. TM*: اعتماد الوسيط بدلاً من الوسط الحسابي عند تطبيق معامل بيرسون للارتباط للبيانات الموضحة في الفقرة (5).
7. TG*: اعتماد صيغة مقدر *Gastwirth* بدلاً من الوسط الحسابي عند تطبيق معامل بيرسون للارتباط للبيانات الموضحة في الفقرة (5) .
8. TSP*: اعتماد معامل سبيرمان للارتباط للبيانات الموضحة في الفقرة (5) .
9. TRP: تحويل البيانات الى الرتب الخاصة بكل متغير ثم تطبيق معامل بيرسون للارتباط على الرتب.

الجانب التجريبي

لدراسة سلوك كل من الطرائق التسعة المذكورة آنفاً فقد تم اعتماد طريقة (*Monte Carlo*) للمحاكاة بتكرار مقداره (*I=1000*) وذلك من خلال المقارنة بين اداء احصاءة الاختبار، t ، لتلك الطرائق عند الحالات الاربعة ادناه :

- الحالة القياسية لتوزيع المتغيرين : $x \sim N(0,1)$ ، $y \sim N(0,1)$.
- التوزيع الطبيعي الملوث من جهة واحدة : $x \sim N(0,1)$, $y \sim 0.8\ N(0,1) + 0.2\ N(5,1)$ ، بمعنى ان 80% من مشاهدات المتغير العشوائي y تتبع التوزيع الطبيعي القياسي اما بقية المشاهدات المتمثلة بنسبة 20% فكانت تتبع التوزيع الطبيعي بمتوسط مقداره (5) وتباين مقداره (1) ، اي تلويث المتغير العشوائي y من جهة واحدة لمعلمة الموقع .
- التوزيع الطبيعي الملوث من جهة واحدة مع زيادة تباين المتغير العشوائي y ليساوي 10 :
  $x \sim N(0,1)$, $y \sim 0.8\ N(0,10) + 0.2\ N(5,10)$.
- التوزيع الطبيعي الملوث من جهة واحدة مع زيادة تباين المتغير العشوائي y ليساوي 20 :
  $x \sim N(0,1)$, $y \sim 0.8\ N(0,20) + 0.2\ N(5,20)$.

وقد تم تطبيق الدراسة لاحجام عينات مختلفة تراوحت مابين الصغيرة ، المتوسطة والكبيرة :

n = 8 , 20 , 50 , 100 ، اما قيم معامل الارتباط فكانت : ρ = 0 , 0.9 .

بعد توليد الحالات على وفق احجام العينات وقيم معامل الارتباط الموضحة اعلاه ، تتم المقارنة بين سلوك الطرائق قيد البحث على اساس معدلات الخطأ من النوع الاول ($\hat{\alpha}$) وقوة الاختبار ($\hat{\tau}$) اللتين يتم ايجادهما باعتماد الصيغتين ادناه :

$$\hat{\alpha} = \frac{R_0(t)}{I} \qquad \ldots\ldots(8)$$

$$\hat{\tau} = \frac{R_1(t)}{I} \qquad \ldots\ldots(9)$$

اذ ان :

$R_0(t)$: عدد مرات رفض الاحصاءة (t) للفرضية (H_0) الصحيحة .

$R_1(t)$: عدد مرات رفض الاحصاءة (t) للفرضية (H_0) الخاطئة .

مجلة ابن الهيثم للعلوم الصرفة والتطبيقية المجلد23 (3) 2010

I : عدد مرات تكرار التجربة .

ومن الجدير بالذكر ان الفترة التي تعد ملائمة أن تقع ضمنها معدلات الخطأ من النوع الاول للطرائق الحصينة هي الفترة المتمثلة بـ (0.036،0.063) والناتجة من تطبيق صيغة *Salter* و *Fawcett* [12] الواردة ادناه : ($\alpha = 0.05$)

$$\alpha \pm 2\sqrt{\frac{\alpha(1-\alpha)}{I}} \qquad \text{...... (10)}$$

الاستنتاجات والتوصيات

استناداً الى نتائج أسلوب المحاكاة ، الجداول (1).....(8) ، يمكن التوصل الى الاستنتاجات الاتية :-

1. من خلال مراجعة الجدولين (1) ، (2) المتضمنين نتائج الحالة المثلى لكلا المتغيرين اي حالة التوزيع الطبيعي الثنائي يتضح ان احتمال معدلات الخطأ من النوع الاول لجميع الطرائق قد وقعت ضمن الفترة المحددة المقبولة ولجميع احجام العينات اما قيم قوة الاختبار فكانت مساوية الى الواحد الصحيح لجميع الطرائق قيد الدراسة عند ($n \geq 20$) اما عند ($n = 8$) فنجدها قد تراوحت بين (0.836،0.958) .
2. بالرجوع الى الجدول (3) ، (4) وعند ($n=8$) يتبين ان مجرد تلويث المتغير العشوائي y من جهة واحدة ولمعلمة الموقع قد اثر سلباً في اداء معامل اختبار t لمعامل ارتباط بيرسون (TP) اذ ان احتمال الخطأ من النوع الاول قد اصبح خارج الفترة المقبولة كذلك الحال بالنسبة اليه عند اعتماد القيم المستبدلة (TP*) الا ان ذلك التأثر سرعان مايختفي عند زيادة حجم العينة اما بقية الطرائق المحورة فقد اثبتت حصانتها تجاه القيم الشاذة عند هذا النوع من التلويث ، اذ حافظت احتمالات معدلات الخطأ من النوع الاول على بقائها ضمن الفترة المحددة ماعدا تأثر كل من (TSP ، TSP* ، TRP) بوجود القيم الشاذة عند ($n = 20$).
3. من خلال مراجعة جدول (5) وعند ($n=8$) يتضح التأثر السلبي وللمرة الثانية في اداء معامل اختبار t لمعامل ارتباط بيرسون سواء باعتماد القيم الاصلية ام القيم المستبدلة (TP) ، (TP*) الا ان ذلك التأثر سرعان مايختفي عند زيادة حجم العينة اما بقية الطرائق المحورة فقد اثبتت حصانتها تجاه القيم الشاذة عند هذه الحالة من التلويث ، اذ حافظت احتمالات معدلات الخطأ من النوع الاول على بقائها ضمن الفترة المحددة ولجميع حجوم العينات. وبالرجوع الى الجدول (6) وعند ($n=8$،20) يتبين ارتفاع قوة الاختبار لطريقة TG ، TG* تليهما طريقتي TM ، TM* ، في حين كانت قوة الاختبار مساوية الى الواحد الصحيح لجميع الطرائق عند ($n=50$،100).
4. من خلال مراجعة جدول (7) وعند ($n=8$) يتضح التأثر السلبي لجميع الطرائق ماعدا المبنية على اعتماد الوسيط ومقدر *Gastwirth* فقد اثبتت هاتان الطريقتان حصانتهما تجاه القيم الشاذة ، اذ حافظت احتمالات معدلات الخطأ من النوع الاول على بقائها ضمن الفترة المحددة ولجميع حجوم العينات . ويتبين من الجدول (8) ارتفاع قوة الاختبار لطريقتي TG،TG* تليهما طريقتي TM،TM*.
5. عند جميع الحالات المدروسة ، اثبتت كل من TM ، TG ، TM* ، TG* حصانتها تجاه القيم الشاذة وربما يعزى ذلك الى كون كل من الوسيط ومقدر *Gastwirth* من المقدرات الحصينة لمعلمة الموقع .
6. ان قوة الاختبار لبعض الطرائق تكون اعلى من غيرها لكن بمجرد الرجوع الى معدلات الخطأ من النوع الاول لتلك الطرائق نجدها قد وقعت خارج الفترة المحددة المقبولة ويمكن ملاحظة ذلك عند مراجعة الجداول (3) و (4) ، (5) و (6) ، (7) و (8).
7. عند مقارنة اداء TG مع TG* ، TM مع TM* و TSP مع TSP* يلاحظ الاداء نفسه عند الحالة الثالثة والرابعة "حالتي التوزيع الطبيعي الملوث من جهة واحدة مع زيادة تباين المتغير العشوائي y ليساوي 10 و20 على التوالي" وهذا يعني ان اعتماد القيم الاصلية او المستبدلة لم يؤثر في ادائها عند الحالتين اعلاه ويمكن مراجعة الجداول (5) و (6) ، (7) و (8).
8. بأعتماد الرأي الاول القائل بأن الاختبار الاحصائي يعد حصيناً لاختراق اي من الشروط الاساسية لذلك الاختبار اذا لم يترتب على ذلك الاختراق اي تغيير واضح في اي من احتمال الخطأ من النوع الاول وقوة الاختبار ، نرى امكانية تلخيص افضل الطرائق بمقارنة بعضها مع بعض للحالات الاربعة المدروسة ولحجوم العينات قيد المقارنة بالاتي :

n	8	20	50	100
الحالة الاولى	TG, TM	TG, TM	جميع الطرائق كانت متقاربة في الاداء مع بعض التذبذب البسيط في احتمال الخطأ من النوع الاول	
الحالة الثانية	TSP	TG*,TM*	TG,TM	TG,TP
الحالة الثالثة	TG,TM,TG*,TM*	TM,TM*	TSP,TSP*	TP,TP*
الحالة الرابعة	TG,TM,TG*,TM*	TM,TM*	TG,TM, TG*,TM*	TG,TG*

في ضوء ماورد آنفاً نوصي بدراسة سلوك الطرائق عند حجوم العينات الصغيرة والمتوسطة بشكل منفرد ، كما نوصي باعتماد المقدرات الحصينة لمعلمة الموقع بدلاً من الوسط الحسابي عند تطبيق معامل بيرسون للارتباط .

المصادر

1. Pitman, E.J.G. (1939). Biometrika, 31, 9-12.
2. Conover, W.J. and Iman, R.L. (1981). Am. Statistician, 35, 9-124.
3. Wilcox, R.R. (2001). Journal of Educational and Behavioral Statistics, 26, 73 - 83.
4. Ebd El-Salam, M.E. (2006). J. Kin Saud Univ., Vol. 19, Admin. Sci. (1), pp. 41-50.
5. Chase, W. and Brown, F. (1992) "General Statistics". 2nd Edition. John Wiley and Sons. New York.
6. Maronna, R.A.; Martin, R.D. and Yahai, V.J. (2006) "Robust Statistic, Theory and Methods " . John Wiley & Sons, Ltd, England.
7. Schefler, W.C. (1988) " Statistics: Concepts and Applications ". The Benjamin/ Cummings Publishing Co. Inc.
8. Huber, P.J. (1973) .Ann. Statist. , 1, 799-821.
9. Rupert, D. (1988). Encyclopedia of Statistical Sciences, 8, 176-181.
10. Scheffe, H. (1959) "The Analysis of Variance ". John Wiley and Sons. New York.
11. Gastwirth, J.L. (1966).J.Amer.Statist.Assn. 61, 929-948.
12. Salter, K.C. and Fawcett, R.F. (1985). Comm. Statist., 14, 807-828.

جدول (1) احتمال الخطأ من النوع الاول : $y \sim N(0,1)$ ، $x \sim N(0,1)$ ، $\rho=0$

n	8	20	50	100
TP	0.062	0.055	0.054	0.055
TM	0.051	0.051	0.056	0.055
TG	0.051	0.049	0.054	0.053
TSP	0.054	0.056	0.047	0.048
TP*	0.062	0.055	0.051	0.057
TM*	0.052	0.054	0.049	0.054
TG*	0.056	0.054	0.050	0.058
TSP*	0.054	0.056	0.046	0.049
TRP	0.059	0.058	0.048	0.050

جدول (2) قوة الاختبار : $y \sim N(0,1)$ ، $x \sim N(0,1)$ ، $\rho=0.90$

n	8	20	50	100
TP	0.958	1	1	1
TM	0.940	1	1	1
TG	0.946	1	1	1
TSP	0.856	1	1	1
TP*	0.944	1	1	1
TM*	0.920	1	1	1
TG*	0.935	1	1	1
TSP*	0.853	1	1	1
TRP	0.836	1	1	1

مجلة ابن الهيثم للعلوم الصرفة والتطبيقية   المجلد23 (3) 2010

جدول(3)احتمال الخطأ من النوع الاول : $x \sim N(0,1)$ ، $y \sim 0.8\ N(0,1) + 0.2\ N(5,1)$ ، ρ ρ =0

n	8	20	50	100
TP	0.065	0.059	0.060	0.041
TM	0.051	0.054	0.047	0.050
TG	0.054	0.053	0.049	0.040
TSP	0.053	0.066	0.060	0.047
TP*	0.067	0.060	0.061	0.047
TM*	0.053	0.050	0.058	0.052
TG*	0.059	0.046	0.059	0.046
TSP*	0.055	0.064	0.058	0.045
TRP	0.056	0.065	0.063	0.047

جدول (4) قوة الاختبار : $x \sim N(0,1)$ ، $y \sim 0.8\ N(0,1) + 0.2\ N(5,1)$ ، ρ =0.90

n	8	20	50	100
TP	0.122	0.421	0.815	0.994
TM	0.077	0.361	0.777	0.989
TG	0.090	0.365	0.792	0.992
TSP	0.245	0.787	0.997	1
TP*	0.159	0.536	0.886	0.999
TM*	0.087	0.468	0.844	0.992
TG*	0.115	0.490	0.852	0.995
TSP*	0.242	0.755	0.994	1
TRP	0.238	0.752	0.992	1

جدول(5) احتمال الخطأ من النوع الاول : $x \sim N(0,1)$ ، $y \sim 0.8\ N(0,10) + 0.2\ N(5,10)$ ، ρ =0

n	8	20	50	100
TP	0.069	0.059	0.059	0.054
TM	0.060	0.056	0.060	0.061
TG	0.062	0.057	0.060	0.058
TSP	0.061	0.057	0.056	0.059
TP*	0.073	0.060	0.060	0.056
TM*	0.060	0.056	0.060	0.061
TG*	0.062	0.057	0.060	0.058
TSP*	0.061	0.057	0.056	0.059
TRP	0.063	0.057	0.058	0.060

جدول (6)قوة الاختبار : $x \sim N(0,1)$ ، $y \sim 0.8\ N(0,10) + 0.2\ N(5,10)$ ، ρ =0.90

n	8	20	50	100
TP	0.575	0.990	1	1
TM	0.538	0.983	1	1
TG	0.565	0.986	1	1
TSP	0.470	0.974	1	1
TP*	0.588	0.986	1	1
TM*	0.538	0.983	1	1
TG*	0.565	0.986	1	1
TSP*	0.470	0.974	1	1
TRP	0.468	0.973	1	1

جدول(7)احتمال الخطأ من النوع الاول :$(5,20)N$ 0.2 + $(0,20)N$ 0.8 ~ y ، $x \sim N(0,1)$ ، ρ =0

n	8	20	50	100
TP	0.071	0.059	0.061	0.063
TM	0.058	0.057	0.055	0.061
TG	0.063	0.063	0.055	0.055
TSP	0.064	0.058	0.057	0.060
TP*	0.071	0.061	0.057	0.061
TM*	0.058	0.057	0.055	0.061
TG*	0.063	0.063	0.055	0.055
TSP*	0.064	0.058	0.057	0.060
TRP	0.067	0.058	0.056	0.062

جدول(8)قوة الاختبار :$(5,20)N$ 0.2 + $(0,20)N$ 0.8 ~ y ، $x \sim N(0,1)$ ، ρ ρ =0.90

n	8	20	50	100
TP	0.761	0.998	1	1
TM	0.693	0.995	1	1
TG	0.730	0.996	1	1
TSP	0.638	0.993	1	1
TP*	0.748	0.997	1	1
TM*	0.693	0.995	1	1
TG*	0.730	0.996	1	1
TSP*	0.638	0.993	1	1
TRP	0.633	0.993	1	1

IBN AL- HAITHAM J. FOR PURE & APPL. SCI. VOL.23 (3) 2010

Some Robust Tests for the Correlation Coefficient: A Comparative Study

N. H. Al-Noor, H. A. Rasheed, A. J. Sultan .

Department of Mathematics, College of Science ,University of AL-Mustanseriya

Abstract

This paper is Interested with studying the performance of statistic test the hypothesis of independence of the two variables (the hypothesis that there is no correlation between the variables under study) in the case of data to meet the requirement of normal distribution in the case away from the distribution due to the presence of outliers (contaminated values) and compared with the performance of some of the other methods proposed and modified.