استخدام المتغيرات المساعدة لدراسة خواص المقدر المقترح بالمعاينة العشوائية المرتبة(RSS)

ريكان عبد العزيز احمد*

الملخص:

في هذا البحث تم استخدام المتغيرات المساعدة لدراسة خواص المقدر المقترح بالمعاينة العشوائية المرتبة(RSS). وقد وتم إثبات أن خصائص المقدر المقترح ذو كفاءه نسبيه عالية مقارنة مع مقدر الانحدار.النسبة والمتوسط العام نسبياً إلى مدى معامل الارتباط بين المتغير المساعد والمتغير الرئيسي.

المقدمة:

إن أحد مقومات نمو الإحصاء النظري هو ظهور قدر كبير من الاساليب النظرية التي تناقش كيفية صنع تقديرات جيدة من بيان إحصائي باستخدام نظرية المعاينة الإحصائية ، ومن تلك الاساليب استخدام المتغيرات المساعدة في التقدير وقد تم الاعتماد على هذا الاسلوب لعدة اسباب من اهمها:

1-أن معظم المسوح الاحصائية التي تحتوي على عدد كبير من المفردات تسهل عملية التقدير بأسلوب المتغيرات المساعدة وذلك كون الحسابات التي يتم إجرائها تكون نوعا ما سهلة وغير معقدة ،بينما تضطرنا طرق التقدير المتفوقة في الإحصاء النظري مثل طريقة الإمكان الأعظم إلى سلسلة من التقديرات المتتالية قبل إن نتمكن من أيجاد التقدير.

2-معظم طرق التقدير في الإحصاء النظري تفترض إننا نعرف الشكل الدالي للتوزيع الاحتمالي الذي تتبعه المعلومات الإحصائية في العينة وتكون طريقة التقدير معدة بعناية لهذا النوع من التوزيعات على خلاف الطريقة المستخدمة في هذا البحث لا تفترض ذلك.

* مدرس/كلية الادارة والاقتصاد/قسم الاحصاء/جامعة البصرة

ولقد استخدم كل من Cochran و Jessen أسلوب التقدير بـالمتغيرات المساعدة بالمعاينة العشوائية البسيطة Simple Random Sample(SRS), لغرض حساب تقديرات ذات اهمية عالية في الجانب التطبيقي نظراً لوجود ارتباط قوي بين المتغير الرئيسي والمتغيرات المساعدة ومن الامثلة على ذلك تقدير متوسط المجتمع , تقدير النسبة Ratio Estimators , وتقدر الانحدار Regression Estimators.

هدف البحث:

إن أستخدام أسلوب المعاينة في السحب يراعى فية عدة امور من أهمها زيادة الدقة في التقدير وتقليل في الوقت والكلفة إلى حد كبير. وان طريقة المعاينة المرتبة (RSS) هي احدى تلك الاساليب التي تعطي تقديرات جيدة وذلك نتيجة التمثيل الكامل لمفردات المجتمع داخل العينة المرتبة وعلية تم الاعتماد عليها في الوقت الحاضر في العديد من البحوث الإحصائية مثل تقدير النسبة باستخدام المعاينة المرتبة وكذلك تقدير معادلة الانحدار باستخدام المعاينة المرتبة كما سيبين ذلك لاحقا. يهدف البحث إلى مناقشة خواص تلك المقدرات السابقة الذكر التي تم دراستها بالمعاينة المرتبة مع خواص المقدر المقترح في هذا البحث باستخدام اسلوب المتغيرات المساعدة في التقدير بالمعاينة المرتبة Rank Set Sampling (RSS) لغرض معرفة افضل تلك المقدرات من حيث الدقة.

الرموز والمصطلحات المستخدمة:

ليكن X_j , Y_j هي قيم إلى المتغيرات X , Y على التوالي من مجتمع يحتوي على N من الوحدات ليكن μ_x يمثل متوسط المجتمع بالنسبة إلى المتغير المساعد X وهو معلوم ,وليكن μ_Y يمثل متوسط المجتمع بالنسبة إلى المتغير الرئيسي Y وهو غير معلوم. وليكن $\bar{X}_{(n)}$, $\bar{Y}_{(n)}$ يمثلان متوسط العينة بالنسبة إلى المتغير المساعد والمتغير الرئيسي على التوالي وبالاعتماد على المعاينة المرتبة بحجم $n = rm$ حيث أن

$$\bar{X}_{(n)} = \frac{\sum_{j=1}^{r}\sum_{i=1}^{m} X_{(i)j}}{n} \ , \ \bar{Y}_{(n)} = \frac{\sum_{j=1}^{r}\sum_{i=1}^{m} Y_{(i)j}}{n}$$

وكذلك سيتم استخدام بعض الصيغ في هذا البحث والتي تعرف بما يلي:

$$\mu_y = E(Y) \quad ,\mu_x = E(X) \quad ,\mu_{y(i)} = E(Y_{(i)}) \quad ,\mu_{x(i)} = E(X_{(i)}) \quad ,\sigma_y^2 = E(Y-\mu_y)^2,$$
$$\sigma_x^2 = E(Y-\mu_x)^2 \quad ,\sigma_{y(i)}^2 = E(Y_{(i)}-\mu_{y(i)})^2, \sigma_{x(i)}^2 = E(X_{(i)}-\mu_{x(i)})^2, T_{y(i)} = (\mu_{y(i)}-\mu_y)$$
$$T_{x(i)} = (\mu_{x(i)}-\mu_x), T_{yx(i)} = (\mu_{y(i)}-\mu_y)(\mu_{x(i)}-\mu_x) \quad ,\sigma_{yx(i)} = E(\mu_{y(i)}-\mu_y)(\mu_{x(i)}-\mu_x)$$

حيث أن

$$\sum_{i=1}^{n} T_{y(i)} = \sum_{i=1}^{n} T_{x(i)} = 0 \quad , \quad \sum_{i=1}^{n} \sigma_{y(i)}^2 = n\sigma_y^2 - \sum_{i=1}^{n} T_{y(i)}^2 \; ,$$
$$\sum_{i=1}^{n} \sigma_{x(i)}^2 = n\sigma_x^2 - \sum_{i=1}^{n} T_{x(i)}^2 , \quad \sum_{i=1}^{n} \sigma_{yx(i)}^2 = n\sigma_{yx} - \sum_{i=1}^{n} T_{yx(i)}$$

<u>المقدر المقترح:</u>

في هذا البحث تم اقتراح المقدر $\overline{Y}_{gen}$ باستخدام المعاينة المرتبة والذي يعرف حسب العلاقة التالية.

$$\overline{Y}_{gen} = \alpha\, \overline{Y}_{(n)} + \beta\left(\mu_x - \overline{X}_{(n)}\right) \qquad \cdots(1)$$

حيث أن كلا من β , α اوزان ويحققان الشرط $\alpha + \beta = 1$

وسيتم دراسة خواص المقدر المقترح $\overline{Y}_{gen}$ ومقارنة تلك الخواص مع خواص المقدرات الاخرى التي تم دراسة خواصها مسبقاً وباستخدام نفس أسلوب المعاينة المرتبة والتي تعتمد في تقديرها على المتغير المساعد والتي قام بدراستها كل من Stokest و Sammawi وهي كما يلي:

$$Simple \qquad \overline{Y}_{sim} = \overline{Y}_{(n)} \qquad \cdots(2)$$

$$Ratio \qquad \overline{Y}_{rat} = \overline{Y}_{(n)} \frac{\mu_{(x)}}{\overline{X}_{(n)}} \qquad \cdots(3)$$

$$\text{Regression} \qquad \overline{Y}_{reg} = \overline{Y}_{(n)} + b\,(\mu_x - \overline{X}_{(n)}) \qquad \cdots(4)$$

حيث إن:

$$b=\frac{m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}}{m\sigma_{x[i]}^{2}-\sum_{i=1}^{m}T_{x[i]}^{2}}$$

ولدراسة خواص المقدر المقترح $\overline{Y}_{gen}$ تم الاعتماد على النظرية الأولى والثانية المثبته عند (Cochran) لغرض اشتقاق متوسط وتباين المقدر المقترح وبالاعتماد على التقريب من مرتبة $\frac{1}{n}$ وعلية يكون مقدار التوقع والتباين للمقدر المقترح $\overline{Y}_{gen}$ على التوالي.

$$E\left(\overline{Y}_{gen}\right)=\alpha\,\overline{Y}_{(n)}\qquad\cdots(5)$$

$$Var(\overline{Y}_{gen})=\alpha^{2}\,Var(\overline{Y}_{(n)})-2\alpha\beta Cov\left(\overline{Y}_{(n)},\overline{X}_{(n)}\right)+\beta^{2}\,Var(\overline{X}_{(n)})\cdots(6)$$

ولكي نجعل هذا التباين اقل ما يمكن يجب ان نختار قيم الى الاوزان β,α بحيث تجعل مقدار التباين للمقدر المقترح $\overline{Y}_{gen}$ اصغر مايمكن وافضل تقدير لهما تم استخراجة باستخدام طريقة لاكرانك وهما على التوالي.

$$\hat{\alpha}=\frac{\left[m\sigma_{x[i]}^{2}-\sum_{i=1}^{m}T_{x[i]}^{2}\right]+\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]}{\left[m\sigma_{y(i)}^{2}-\sum_{i=1}^{m}T_{x(i)}^{2}\right]+2\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]+\left[m\sigma_{x[i]}^{2}-\sum_{i=1}^{m}T_{x[i]}^{2}\right]}\cdots(7)$$

والقيمة المقدرة إلى β تعرف حسب العلاقة التالية.

$$\hat{\beta}=1-\hat{\alpha}\qquad\cdots(8)$$

وعليه يكون اقل تباين ممكن إلى المقدر المقترح $\overline{Y}_{gen}$ نستطيع الحصول علية بعد تعويض المعادلتين رقم (7) و (8) في المعادلة رقم (6) فنحصل على الصيغة التالية:

$$V_{\min}(\bar{Y}_{gen}) = \frac{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right]\left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right] - \left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right]^2}{rm^2\left\{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right] + \left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right] + 2\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right]\right\}} \quad \cdot\cdot(9)$$

ومن الملاحظ إذا كانت قيمة $\alpha = 1$ فان المعادلة رقم (7) نستطيع إعادة كتابتها على الصـــيغة التالية.

$$\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right] = -\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right] \qquad \cdots(10)$$

وبتعويض المعادلة رقم (10) في المعادلة رقم (9) وبحلها نحصل على المعادلة التالية.

$$V_{\min}(\bar{Y}_{gen}) = \frac{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right]}{rm^2} \qquad \cdots(11)$$

والمعادلة رقم (11)هي عبارة عن تباين الوسط الحسابي المقدر للعينة المرتبة الذي يعتبر تقــدير غير متحيز إلى متوسط المجتمع الذي اثبتة Stokest. مما يدل على ان المقدر المقترح $\bar{Y}_{gen}$ هو اعم من المقدر $\bar{Y}_{Sim}$

المقارنات النظرية:

سيتم في هذا المبحث مقارنة الكفاءة النسبية بالنسبة لتباين المقدر المقترح مع تباين المقدرات الأخرى التي تم الإشارة إليها في كل من المعادلات(4,3,2) والتي تم حسابها مـــن قبــل البـــاحثين Stokest و Sammawi وان صيغ التباين إلى تلك المقدرات هي على الترتيب كما يلي:

$$Simple \qquad Var(\bar{Y}_{sim}) = \frac{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right]}{rm^2} \qquad \cdots(12)$$

$$Ratio\,Var(\bar{Y}_{rat}) = \frac{\mu_y^2}{rm^2}\left\{\frac{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right]}{\mu_y^2} + \frac{\left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right]}{\mu_x^2} - 2\frac{\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right]}{\mu_x\mu_y}\right\} \cdot\cdot(13)$$

$$\text{Regression}\quad Var\left(\overline{Y}_{reg}\right)=\frac{1}{rm^2}\left\{\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]-\frac{\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]^2}{\left[m\sigma^2_{x(i)}-\sum_{i=1}^{m}T^2_{x[i]}\right]}\right\}\cdots(14)$$

يلاحظ إن المقدر المقترح $\overline{Y}_{gen}$ يكون ذو كفاءة نسبية عالية من مقدر الانحدار $\overline{Y}_{reg}$ إذا كان.

$$Var\left(\overline{Y}_{gen}\right)\leq Var\left(\overline{Y}_{reg}\right)\qquad\cdots(15)$$

إذا تم تعويض المعادلة رقم (9) والمعادلة رقم (14) في المعادلة رقم (15) نحصل على.

$$\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]+2\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]\quad>0\qquad\cdots(16)$$

إن المعادلة رقم (16) دائماً تكون صحيحة إذا كان معامل الارتباط المرتب بين المتغيــرين X و Y موجب إما إذا كان معامل الارتباط المرتب سالب فأننا نحصل على النتيجة التالية.

$$\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]>-\frac{\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]}{2}\qquad\cdots(17)$$

أن الكفاءة النسبية للمقدر البسيط $\overline{Y}_{sim}$ إلى المقدر المقترح $\overline{Y}_{gen}$ نستطيع تعريفها حسب العلاقة التالية.

$$RE\,(Simple/Gen)=\frac{Var(\overline{Y}_{Sim})}{Var(\overline{Y}_{gen})}$$

$$RE\ (Simple\,/\,Gen)=\frac{2\left[m\sigma^2_{x[i]}-\sum_{i=1}^{m}T^2_{x[i]}\right]}{\left[m\sigma^2_{x[i]}-\sum_{i=1}^{m}T^2_{x[i]}\right]-\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]}\quad\cdots(18)$$

$$\text{Regression } Var(\bar{Y}_{reg}) = \frac{1}{rm^2}\left\{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right] - \frac{\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right]^2}{\left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right]}\right\} \cdots (14)$$

يلاحظ إن المقدر المقترح $\bar{Y}_{gen}$ يكون ذو كفاءة نسبية عالية من مقدر الانحدار $\bar{Y}_{reg}$ إذا كان.

$$Var(\bar{Y}_{gen}) \le Var(\bar{Y}_{reg}) \qquad \cdots (15)$$

إذا تم تعويض المعادلة رقم (9) والمعادلة رقم (14) في المعادلة رقم (15) نحصل على.

$$\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right] + 2\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right] > 0 \qquad \cdots (16)$$

إن المعادلة رقم (16) دائماً تكون صحيحة إذا كان معامل الارتباط المرتب بين المتغيـرين X و Y موجب إما إذا كان معامل الارتباط المرتب سالب فأننا نحصل على النتيجة التالية.

$$\left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right] > -\frac{\left[m\sigma^2_{y(i)} - \sum_{i=1}^{m} T^2_{y(i)}\right]}{2} \qquad \cdots (17)$$

أن الكفاءة النسبية للمقدر البسيط $\bar{Y}_{sim}$ إلى المقدر المقترح $\bar{Y}_{gen}$ نستطيع تعريفها حسب العلاقة التالية.

$$RE\,(Simple/Gen) = \frac{Var(\bar{Y}_{Sim})}{Var(\bar{Y}_{gen})}$$

$$RE\,(Simple/Gen) = \frac{2\left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right]}{\left[m\sigma^2_{x[i]} - \sum_{i=1}^{m} T^2_{x[i]}\right] - \left[m\sigma_{yx} - \sum_{i=1}^{m} T_{yx[i]}\right]} \cdots (18)$$

$$\text{Regression}\quad Var\left(\bar{Y}_{reg}\right)=\frac{1}{rm^2}\left\{\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]-\frac{\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]^2}{\left[m\sigma^2_{x(i)}-\sum_{i=1}^{m}T^2_{x[i]}\right]}\right\}\cdots(14)$$

يلاحظ إن المقدر المقترح $\bar{Y}_{gen}$ يكون ذو كفاءة نسبية عالية من مقدر الانحدار $\bar{Y}_{reg}$ إذا كان.

$$Var\left(\bar{Y}_{gen}\right)\leq Var\left(\bar{Y}_{reg}\right)\qquad\cdots(15)$$

إذا تم تعويض المعادلة رقم (9) والمعادلة رقم (14) في المعادلة رقم (15) نحصل على.

$$\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]+2\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]\quad>0\qquad\cdots(16)$$

إن المعادلة رقم (16) دائماً تكون صحيحة إذا كان معامل الارتباط المرتب بين المتغيـرين X و Y موجب إما إذا كان معامل الارتباط المرتب سالب فأننا نحصل على النتيجة التالية.

$$\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]>-\frac{\left[m\sigma^2_{y(i)}-\sum_{i=1}^{m}T^2_{y(i)}\right]}{2}\qquad\cdots(17)$$

أن الكفاءة النسبية للمقدر البسيط $\bar{Y}_{sim}$ إلى المقدر المقترح $\bar{Y}_{gen}$ نستطيع تعريفها حسب العلاقة التالية.

$$RE\,(Simple/Gen)=\frac{Var(\bar{Y}_{Sim})}{Var(\bar{Y}_{gen})}$$

$$RE\ (Simple\,/\,Gen)=\frac{2\left[m\sigma^2_{x[i]}-\sum_{i=1}^{m}T^2_{x[i]}\right]}{\left[m\sigma^2_{x[i]}-\sum_{i=1}^{m}T^2_{x[i]}\right]-\left[m\sigma_{yx}-\sum_{i=1}^{m}T_{yx[i]}\right]}\quad\cdots(18)$$

الاستنتاجات:

يلاحظ من النتائج التي تم التوصل أليها من الجدولين أعلاه وكذلك النتائج التي تم التوصل أليها في المبحث السابق مايلي.

1- أن المقدر المقترح $\overline{Y}_{gen}$ هو افضل من كافة المقدرات السابقة وذلك كونه يعطي اقل تباين وأعلا كفاءة نسبية مقارنة مع بقية المقدرات الأخرى.

2- يلاحظ اختلاف النتائج بين الجدول الأول والثاني راجع إلى اختلاف الترتيب حيث أن في الجدول الأول تم الاعتماد على المتغير الرئيسي Y في الترتيب والمتغير المساعد X يتبع المتغير الرئيسي, أما في الجدول الثاني تم الاعتماد على المتغير المساعد X في الترتيب والمتغير الرئيسي Y يتبع المتغير المساعد.

3-يلاحظ أن النتائج التي تم التوصل أليها في الجدول الثاني هي افضل من النتائج التي تم التوصل أليها في الجدول الأول وذلك كون بيانات المتغير المساعد X ترتيبها وتقاربها افضل من ترتيب وتقارب المتغير الرئيسي Y.

4- يلاحظ تطابق النتائج النظرية التي تم التوصل أليها في المبحث الرابع مع النتائج العملية التي تم التوصل أليها في المبحث الخامس مما يدل على أن الطريقة المستخدمة في التقدير والحساب طريقة جيدة.

5- المقدر المقترح $\overline{Y}_{gen}$ افضل من مقدر الانحدار $\overline{Y}_{\mathrm{Re}g}$ ومقدر النسبة $\overline{Y}_{rat}$ في عملية التقدير وتوظيف كافة بيانات المتغير المساعد لتوضيح المتغير الرئيسي تحت الدراسة وذلك للأسباب أعلاه.

التوصيات:

1- يوصي الباحث باستخدام أسلوب السحب بالمعاينة المرتبة وذلك كونها تعطي معلومات دقيقة عن متغيرات الدراسة وتمثيل كافة مفردات المجتمع في العينة المسحوبة.

2- يوصي الباحث باستخدام المعاينة المرتبة لدراسة خواص التوزيعات الاحتمالية والتي تم دراستها باستخدام المعاينة العشوائية البسيطة وذلك لسبيل المقارنة ومعرفة دقة المعاينتين.

3- يوصي الباحث باستخدام المقدر المقترح $\overline{Y}_{gen}$ كمؤشر إحصائي يتم المقارنة به كونه أفضل من كافة المقدرات الأخرى السابقة الذكر.

المصادر

1-AL- Samarraie,A.Y.(1999).Regression estimators using extreme ranked set sampling Biom.J.Vol. 30,PP 412-419.

2-Cochran,W.G.(1977).Sampling Techniques,third edition.John Wiley & Sons

Jessen,R.J.(1978). Statistical Survey Techniques. John Wily & Sons

3-Moltammad S. Ahmed & Walid Abu-Dayyeh. (1999). Estmation the population mean using extreme ranked set sampling. Biom..J 38

Muttlak.H.A. (1998)Median ranked set sampling with concomitant variables. Envtronmetrics,Vol.9,PP255-267..

4-Samawi,H.M.(1996). Estimation of ratio using rank set sampling. Biom.J.Vol.38, PP 753-764.

5-Stokest,S.L.(1977). Ranked set sampling with concomitant variables. Comm.Stat-Theory Math.Vol.A6(12),PP1207-1211

....

....

....