مقدر الاختبار الأولي لأنموذج الانحدار الخطي للبيانات غير التامة المشاهدات[+]

PRETEST ESTIMATION OF LINEAR REGRSSION MODEL FOR INCOMPLETE DATA

ايدن حسن حسين الكناني[*]

المستخلص

في التحليل الاحصائي لاي مشكلة الانموذج الانحدار للخط البياني تكون غير تامة اذا اعطيت بيانات غير دقيقة وتكون النتيجة غير جيدة. لحل هذه المشكلة وجد الباحثون العديد من النظريات ولكن اثنين منها كانت الاكثر فعالية

بالنسبة للبيانات المتاحة. اخر ماتوصل له الباحثون هو دمج هاتين النظريتين للحصول على افضل النتائج.

Abstract

In statistical analysis, the problem of estimating the parameters of simple linear model is when we have incomplete data and the resulting estimators are inefficient. Many methods are found by researchers to treat such problem. In this work, we study two of these methods when using incomplete data namely, complete case analysis and available case analysis. A new procedure is suggested called shrinkage estimator which connects between estimate parameters of linear model and those two methods by using constant function. The bias, MSE and RE for the estimators of the parameters of simple linear model are found also, the values of these estimators are calculated.

المقدمـــــة

ان اغلب طرائق التحليل الاحصائي تفترض توافر بيانات تامة المشاهدات للعينات المدروسة وقد بينت جميع الطرائق الاحصائية على هذا الاساس،ولكن في الكثير من الظواهر الطبيعية والاقتصادية والاجتماعية وغيرها تتعرض جزء من بيانات هذه الظواهر الى الفقدان او عدم المشاهدة (not observed) وتختلف اسباب الفقدان فمنها ما يكون متعمدا بسبب الكلفة العالية او المخاطرة او بسبب عدم توافر الامكانيات ومنها ما يكون غير متعمد مثل تعطل اجهزة التسجيل او بسبب عدم توافر المستلزمات الضرورية لعملية الانتاج او بسبب الكوارث الطبيعية والحروب وغيرها ومهما اختلفت وتعددت اسباب الفقدان فاننا سنواجه مشكلة معقدة وغير بسيطة وهي ان المشاهدات والبيانات تكون غير تامة (incomplete) .[1].

+ تاريخ استلام البحث ٢٠٠٤/٣/٣١ تاريخ قبول النشر ٢٠٠٥/٢/١٢

[*] استاذ مساعد / قسم الاحصاء / كلية الادارة والاقتصاد/ لجامعة المستنصرية

ان مجرد اطلاق كلمة مشكلة على تقدير معلمات الانموذج في حالة البيانات غير التامة تعني ان المقدرات التي سيتم الحصول عليها من البيانات غير التامة ستكون مقدرات غير كفوءة ويطلق مصطلح بيانات غير تامة لنماذج الانحدار والتي فيها يفقد متغير او اكثر من متغير ما لبعض المشاهدات لاي سبب كان. وفي هذه الحالة يجب معالجة هذه المشكلة وايجاد الحلول الناجمة لها عن طريق استخدام التحليل الاحصائي الذي يهتم بالبيانات غير التامة والتي عن طريق يتم الحصول على نتائج دقيقة.

ان البيانات غير التامة تاخذ انماطا (patterns) مختلفة فمنها ما يكون مرتباً او بشكل يمكن ترتيبه ويدعى هذا النمط الخاص (special pattern) ومنها ما يكون غير مرتب وياخذ النمط العام (general pattern) وان اختلاف نمط البيانات يتطلب اساليب مختلفة تتناسب مع نوع النمط.[2]

ان فقدان البيانات نتيجة الالية التي فقدت بسببها هذه البيانات يؤدي الى اختلاف مواصفات العينات المتاحة للدراسة. فمن البيانات ما يفقد بشكل عشوائي تام (missing completely at random) (MCAR) او ان الفقدان لقيم متغير معين تام المشاهدة يتم بسبب فقدان مشاهدات قيم متغير اخر غير تام المشاهدة أي ان البيانات تفقد بشكل عشوائي (missing at random) (MAR) وفي احيان اخرى تفقد البيانات بشكل متعمد أي بشكل غير عشوائي (not missing at random) (OT MAR) وان التعرف على الية الفقدان يعد المفتاح لتشخيص اسلوب التحليل الاحصائي المعتمد لدراسة الظاهرة. [3]

<u>هدف البحث</u>

ان الهدف من هذا البحث هو عرض اهم الطرائق الاحصائية التي تستخدم لتقدير معلمات انموذج الانحدار عندما يفقد متغيراً توضيحياً واحداً او اكثر لبعض المشاهدات وبفرض ان متغير الاستجابة يكون تام المشاهدات. ومن ثم عرض طريقتين من الطرائق المستخدمة لتقدير معلمات انموذج الانحدار الخطي البسيط وهما طريقة المشاهدات التامة وطريقة المشاهدات المتاحة واللتان تعالجان حالة البيانات غير التامة ومن ثم دمج وربط هاتين الطريقتين معاً وافترض ان طريقة المشاهدات المتاحة هي اسلوب لتقدير معلمات انموذج الانحدار الخطي البسيط للحصول على المقدرات الاعتيادية او التقليدية وافتراض طريقة المشاهدات التامة اسلوب لتقدير معلمات الانموذج للحصول على القيمة الاولية وبعد ذلك تكوين تركيبة او توليفة خطية (linear combination) بين طريقة المشاهدات التامة وطريقة المشاهدات المتاحة وبالصيغة الاتية:

$$\tilde{q} = \Psi(\hat{q})\hat{q} + [1 - \Psi(\hat{q})]q_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

حيث ان $\hat{q}$ تمثل تقدير لمعلمات انموذج الانحدار بطريقة المشاهدات المتاحة و q_0 تمثل تقدير للمعلمة انموذج الانحدار بطريقة المشاهدات التامة وعدها قيمة اولية و $\Psi(\hat{q})$ تمثل دالة للربط بين القيمة التقديرية والقيمة الاولية وتكون قيمتها محصورة بين الصفر والواحد و $\tilde{q}$ تمثل مقدر التقلص ذات المرحلة الواحدة.

كما تم في هذا البحث اشتقاق وايجاد مقدار التحيز ومتوسط مربعات الخطأ والكفاءة النسبية لمقدر التقلص المقترح والذي يستخدم لاول مرة في البيانات غير التامة لتقدير معالم انموذج الانحدار الخطي البسيط.

تحليل المشاهدات التامة[2]

استنبط هذا التحليل في عام (١٩٣٢) من قبل (YATES) الذي يعد اول من استخدم هذا الاسلوب لمعالجة البيانات غير التامة لانموذج الانحدار الخطي البسيط فان الانموذج يعرف بالصيغة الاتية:

$$Y_i = a + b\,x_{ij} + e_i$$

حيث ان $j = 1,2,.....n_x$, $i = 1,2,.....n_c$ ، ان n_x تمثل عدد مشاهدات المتغير التوضيحي و n_c تمثل عدد المشاهدات المشتركة بين كل من المتغير التوضيحي ومتغير الاستجابة ومن اجل ايجاد تقدير معالم انموذج الانحدار الخطي البسيط في حالة وجود بيانات غير تامة فاننا نطبق مبدا طريقة المربعات الصغرى الاعتيادية وان صيغة المقدارات تكون كما يلي:

$$\hat{a} = \overline{Y}_c - \hat{b}\,\overline{\mathrm{X}}_c \qquad , \qquad \hat{b} = \frac{\sum_{j=1}^{n_c}(Y_j - \overline{Y}_c)(\mathrm{X}_j - \overline{\mathrm{X}}_c)}{\sum_{j=1}^{n_c}(\mathrm{X}_j - \overline{\mathrm{X}}_c)^2}$$

$$\overline{\mathrm{X}} = \frac{\sum_{j=1}^{n_c}\mathrm{X}_j}{n_c} \qquad , \qquad \overline{Y}_c = \frac{\sum_{j=1}^{n_c}Y_j}{n_c}$$

ان مقدرات المعالم في تحليل المشاهدات التامة تمتلك نفس خاصية مقدرات طريقة المربعات الصغرى الاعتيادية بكونها مقدرات غير متحيزة.

ان تباين مقدرات المعالم بطريقة تحليل المشاهدات التامة فانها تكون كالاتي:

$$\mathrm{var}(\hat{a}) = s^2\left(\frac{1}{n} + \frac{\overline{\mathrm{X}}_c^2}{\sum_{j=1}^{n_c}(x_j - \bar{x}_c)^2}\right) \qquad , \qquad \mathrm{var}(\hat{b}) \frac{s^2}{\sum_{j=1}^{n_c}(x_j - \bar{x}_c)^2}$$

تحليل المشاهدات المتاحة

في عام(١٩٦٤) استنبط هذه الطريقة لتقدير معلمات انموذج الانحدار الخطي البسيط وذلك بالاعتماد على مصفوفة التباين والتباين المشترك $\mathrm{cov}(x_j, x_k)$ للمتغيرات X,S والمتجه $\mathrm{cov}(x_j, y)$ وان مقدر معلمات الانحدار تكون وفق الصيغة الاتية:[2]

$$\hat{b} = (\mathrm{cov}(x_k, x_j))^{-1}(\mathrm{cov}(x_j, Y))$$

ان استخدام تحليل المشاهدات التامة يؤدي الى هدر كبير للبيانات مما يؤثر في التقـديرات الخاصـة بالمتغير الذي حذفت بعض مفرداته. وتتلخص خطوات هذه الطريقة بما يلي:

أ-حساب مصفوفة التباينـات والتباينـات المـشتركة للمتغيـرات التوضـيحية (variance convariance matrix).

ب-حساب متجه التباينات المشتركة بين متغير الاستجابة وكل متغير توضيحي.

ولتقدير المتوسطات ومصفوفة التباينات والتباينات المشتركة يجب توافر شرط اساس هـو ان تكون الية فقدان البيانات من نوع فقدان البيانات بشكل عشوائي تـام وان حـساب التبـاين المـشترك للمتغيـرات X_j, X_k باستخدام البيانات المتاحة كما يلي:

$$S_{jk} = \frac{\sum_{j=1}^{n_{jk}} (X_{ij} - \overline{X}_j)(X_{ik} - \overline{X}_k)}{(n_{jk} - 1)}$$

حيث ان $J, K = 1,2,.....,P$ $i = 1,2,.....,n_{jk}$ وان n_{jk} تمثل عدد المـشاهدات التي تظهر فيها X_j, X_k معا وان

$$\overline{X}_j = \frac{\sum_{j=1}^{n_{jk}} X_{ij}}{n_{jk}} \quad , \quad \overline{X}_k = \frac{\sum_{j=1}^{n_{jk}} X_{ik}}{n_{jk}}$$

ومن الجدير بالذكر ان $\overline{X}_j, \overline{X}_k$ تحسب من المشاهدات المشتركة نفـسها للمتغيـرين X_j, X_k امـا تباينات X_j, X_k فيتم حسابها كما يلي:

$$\tilde{S}_{jj} = \frac{\sum_{i=1}^{n_j} (X_{ij} - \tilde{X}_j)}{n_j} \quad , \quad \tilde{S}_{kk} = \frac{\sum_{i=1}^{n_k} (X_{ik} - \tilde{X}_k)}{n_k}$$

حيث ان

$$\tilde{X}_j = \frac{\sum X_{ij}}{n_j} \quad , \quad \tilde{X}_k = \frac{\sum X_{ik}}{n_k}$$

وان n_j تمثل عدد المشاهدات التامة للمتغير X_j و n_k تمثل عدد المشاهدات التامة للمتغير X_k وان $\tilde{X}_k, \tilde{X}_j, \tilde{S}_{kk}, \tilde{S}_{jj}$ تحسب من استخدام المشاهدات المتاحة لكل متغير في العينة.

يعد تقدير معالم انموذج الانحدار الخطي البسيط باستخدام تحليل المشاهدات المتاحة احدى الطرائـق السهلة المباشرة، حيث ان صيغة تقدير المعلمة b يكون كما يلي:

$$\hat{b} = (\tilde{S}_{xx})^{-1}\tilde{S}_{xy}$$

حيث ان

$$\tilde{S}_{xx} = \frac{\sum_{j=1}^{n_c}(\mathrm{X}_j - \tilde{\mathrm{X}})^2}{n_x - 1} \quad , \quad \tilde{S}_{xy} = \frac{\sum_{j=1}^{n_x}(\mathrm{X}_j - \tilde{\mathrm{X}})(Y_j - \bar{Y})}{n_x - 1}$$

وان مقدر الحد الثابت a فانه يكون كما يلي:

$$\hat{a} = \bar{Y} - \hat{b}\,\bar{\mathrm{X}}$$

ان مقدرات تحليل المشاهدات المتاحة تكون غير متحيزة عن مقدرات طريقة المربعات الصغرى الاعتياديــة وان صيغة تباين $\hat{b}$ في هذه الحالة تكون بالشكل الاتي:

$$\mathrm{var}(\hat{b}) = \frac{s^2}{\sum_{j=1}^{n_j}(\mathrm{X}_j - \tilde{\mathrm{X}})^2}$$

وان مقدر تباين المجتمع $\hat{s}^2$ يكون بالصيغة الاتية:

$$\hat{s}^2 = \frac{(n-1)[\tilde{S}_{yy} - \hat{b}\tilde{S}_{xy}]}{(n-k-1)}$$

وان تباين الحد الثابت يكون بالشكل التالي:

$$\begin{aligned}\mathrm{var}(\hat{a}) &= \mathrm{var}(\bar{Y}) + \bar{\mathrm{X}}_j^2\,\mathrm{var}(\hat{b}) \\ &= \mathrm{var}(s^2/n) + \bar{\mathrm{X}}_j^2\,\mathrm{var}(\hat{b})\end{aligned}$$

مقدر التقلص[4]

تعتمد فكرة هذه الطريقة في الاستفادة من المعلومات المسبقة (القيمة الاولية) التي توفرها طرائق بيز وتقدير المعلمات التي توفرها الطرائق الاعتيادية في التقدير وعمل ربط ودمج بين هذه الطرائــق للحــصول على مقدرات التقلص التي يمكن عدها مقدرات اكثر كفاءة من المقدرات التي تستخدم بيانات العينة فقط.

سنعد دالة التقلص الموزونة $\Psi(\hat{q})$ دالة ثابتة مقدارها K حيث ان $0 \le K \le 1$ والتي تمثــل مقدار الثقة بـــ $\hat{q}$ والتي تقدر باستخدام المشاهدات المتاحة وبمقدار ثقة (1-K) للمعلمــة $\hat{q}$ والتــي تقــدر باستخدام تحليل المشاهدات التامة وبعد تعويض $\Psi(\hat{q}) = k$ في العلاقة (١) فــان مقــدار الــتقلص ذات المرحلة الواحدة يكون بالشكل الاتي:-

$$\begin{aligned}\tilde{q} &= K\hat{q} + (1-K)q_0 \\ &= K(\hat{q} - q_0) + q_0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)\end{aligned}$$

ومقدر التقلص للمعلمة a يكون بالشكل التالي:-

$$\begin{aligned}\tilde{a} &= K\hat{a} + (1-K)a_0 \\ &= K(\hat{a} - a_0) + a_0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)\end{aligned}$$

فيما يلي سنقوم بحساب التحيز ومتوسط مربعات الاخطاء والكفاءة النسبية للمقدرين $\tilde{a}, \tilde{b}$

نفترض ان

$$\hat{a} \sim N(a, V(\hat{a})) \quad \ldots\ldots(4)$$

ولغرض تسهيل حساب التكاملات نعمل التحويل التالي

$$Z = \frac{\hat{a} - a}{\sqrt{\mathrm{var}(\hat{a})}}, \qquad Z \sim N(0,1) \quad \ldots\ldots(5)$$

من العلاقة (٥) فان

$$Z\sqrt{\mathrm{var}(\hat{a})} + a = \hat{a} \qquad , \sqrt{\mathrm{var}(\hat{a})}\,dz = d\hat{a} \quad \ldots\ldots(6)$$

الان سنجد مقدار التحيز لمقدر التقلص $\tilde{a}$ وذلك باستخدام العلاقة التالية

$$\begin{aligned} Bias(\tilde{a}) &= E(\tilde{a} - a) \\ &= E[K(\hat{a} - a_0) + a_0 - a] \\ &= KE(\hat{a} - a_0) - E(a - a_0) \quad \ldots\ldots(7) \end{aligned}$$

وباستخدام العلاقات (٥) و (٦) فان

$$\begin{aligned} E(\hat{a} - a_0) &= \int_{-\infty}^{\infty} (\hat{a} - a_0) \frac{1}{\sqrt{2p\ \mathrm{var}(\hat{a})}} e^{\frac{-(\hat{a}-a)^2}{2\,\mathrm{var}(\hat{a})}} d\hat{a} \\ &= \int_{-\infty}^{\infty} (Z\sqrt{\mathrm{var}(\hat{a})} + a - a_0) \frac{1}{\sqrt{2p}} e^{-z^2/2}\, dz \\ &= \sqrt{\mathrm{var}(\hat{a})} \int_{-\infty}^{\infty} Z \frac{1}{\sqrt{2p}} e^{-z^2/2}\, dz + (a - a_0) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2p}} e^{-z^2/2}\, dz \\ &= l\sqrt{\mathrm{var}(\hat{a})} \quad \ldots\ldots(8) \end{aligned}$$

حيث ان

$$l = \frac{a - a_0}{\sqrt{\mathrm{var}(\hat{a})}} \quad \ldots\ldots(9)$$

وان

$$\begin{aligned} E(a - a_0) &= a - a_0 \\ &= l\sqrt{\mathrm{var}(\hat{a})} \quad \ldots\ldots(10) \end{aligned}$$

بتعويض العلاقتين (٨) و (١٠) في العلاقة (٧) نحصل على

$$\begin{aligned} Bias(\tilde{a}) &= Kl\sqrt{\mathrm{var}(\hat{a})} - l\sqrt{\mathrm{var}(\hat{a})} \\ &= l\sqrt{\mathrm{var}(\hat{a})}(K - 1) \quad \ldots\ldots(11) \end{aligned}$$

وبالطريقة نفسها يمكننا ايجاد القيمة النهائية لمقدار التحيز للمعلمة b والذي نتيجته هي:

$$Bias(\tilde{b}) = l\sqrt{\text{var}(\hat{b})}(K-1).....(12)$$

لغرض ايجاد متوسط مربع الاخطاء للمقدر $\tilde{a}$ فاننا نطبق العلاقة التالية:

$$MSE(\tilde{a}) = E(\tilde{a}-a)^2$$

$$= E[K(\hat{a}-a_0)-(a-a_0)]^2$$

$$= E[K^2(\hat{a}-a_0)^2 - 2K(\hat{a}-a_0)(a-a_0)+(a-a_0)^2]$$

$$= K^2A_1 - 2KA_2 + A_3.....(13)$$

حيث ان

$$A_1 = E(\hat{a}-a_0)^2,\quad A_2 = E[(\hat{a}-a_0)(a-a_0)],\quad A_3 = E(a-a_0)^2.....(14)$$

وباستخدام العلاقات (٥) و (٦) و (٩) فان حساب A_1 , A_2 , A_3 يكون كما يلي:

$$A_1 = E(\hat{a}-a_0)^2$$

$$A_1 = \int_{-\infty}^{\infty}(\hat{a}-a_0)^2\frac{1}{\sqrt{2p\ \text{var}(\hat{a})}}e^{\frac{-(\hat{a}-a)^2}{2\,\text{var}(\hat{a})}}d\hat{a}$$

$$A_1 = \int_{-\infty}^{\infty}(Z\sqrt{\text{var}(\hat{a})}+(a-a_0))^2\frac{1}{\sqrt{2p}}e^{-z^2/2}dz$$

$$A_1 = \int_{-\infty}^{\infty}(Z^2\,\text{var}(\hat{a})+2Z\sqrt{\text{var}(\hat{a})}(a-a_0)+(a-a_0)^2)\frac{1}{\sqrt{2p}}e^{-z^2/2}dz$$

$$A_1 = [\text{var}(\hat{a})\int_{-\infty}^{\infty}Z^2\frac{1}{\sqrt{2p}}e^{-z^2/2}dz + 2l\,\text{var}(\hat{a})\int_{-\infty}^{\infty}Z\frac{1}{\sqrt{2p}}e^{-z^2/2}dz + l^2\,\text{var}(\hat{a})\int_{-\infty}^{\infty}\frac{1}{\sqrt{2p}}e^{-z^2/2}dz]$$

$$A_1 = \text{var}(\hat{a}) + l^2\,\text{var}(\hat{a}) = \text{var}(\hat{a})(1+l^2).....(15)$$

حيث ان $E(Z) = 0$, $E(Z^2) = Var(Z) = 1$.

وان

$$A_2 = E[(a-a_0)(a-a_0)]$$

$$A_2 = \int_{-\infty}^{\infty}(\hat{a}-a_0)(a-a_0)\frac{1}{\sqrt{2p\ \text{var}(\hat{a})}}e^{\frac{-(\hat{a}-a)^2}{2\,\text{var}(\hat{a})}}d\hat{a}$$

$$A_2 = \int_{-\infty}^{\infty}[Z\sqrt{\text{var}(\hat{a})}+a-a_0](a-a_0)\frac{1}{\sqrt{2p}}e^{\frac{z^2}{2}}dz$$

$$A_2 = l^2 \operatorname{var}(\hat{a}) \int_{-\infty}^{\infty} Z e^{\frac{-z^2}{2}} \frac{1}{\sqrt{2p}} dz + l^2 \operatorname{var}(\hat{a}) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2p}} e^{\frac{-z^2}{2}} dz$$

$$A_2 = l^2 \operatorname{var}(\hat{a}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (16)$$

واخيرا فان

$$A_3 = E(a - a_0)^2 = (a - a_0)^2 = l^2 \operatorname{var}(a) \ldots\ldots\ldots\ldots\ldots\ldots (17)$$

وبتعويض العلاقــات (١٥،١٦،١٧) في (١٣) يكون $\hat{a}$ فان متوسط مربعات الخطأ لمقدر التقلص بالصيغة النهائية الاتية:

$$MSE(\tilde{a}) = K^2 \operatorname{var}(\hat{a})(1 + l^2) - 2kl^2 \operatorname{var}(\hat{a}) + l^2 \operatorname{var}(\hat{a})$$

$$= \operatorname{var}(\hat{a})[K^2 + l^2(k-1)^2] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (18)$$

وبالطريقة نفسها يمكننا ايجاد متوسط مربعات الخطأ لمقدر التقلص $\tilde{b}$ والذي صيغته النهائية بالشكل الاتي:

$$MSE(\tilde{b}) = \operatorname{var}(\hat{b})[K^2 + l^2(K-1)^2] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (19)$$

اما الكفاءة النسبية لمقدر التقلص $\tilde{a}$ فهي:

$$R.E(\tilde{a}) = MSE(\hat{a}) / MSE(\tilde{a}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (20)$$

ولكون المقدر $\tilde{a}$ غير متحيز فان

$$MSE(\hat{a}) = Var(\hat{a}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (21)$$

$$R.E(\tilde{a}) = [K^2 + l^2(K-1)^2]^{-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (22)$$

وبالطريقة نفسها نجد الكفاءة النسبية لمقدر التقلص $\tilde{b}$ والذي تكون صيغته هي:

$$R.E(\tilde{b}) = [K^2 + l^2(K-1)^2]^{-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (23)$$

وان قيمة K التي تجعل $MSE(\tilde{a})$ اقل ما يمكن تحسب بواسطة اخذ المشتقة بالنسبة الى K للعلاقــة (١) ومساواة هذه المشتقة بالصفر وكما يلي:

$$\frac{\partial MSE(\hat{a})}{\partial K} = 2K \operatorname{var}(\hat{a})(1 + l^2) - 2l^2 \operatorname{var}(\hat{a})$$

$$0 = 2K \operatorname{var}(\hat{a})(1 + l^2) - 2l^2 \operatorname{var}(\hat{a})$$

فنحصل على قيمة K والتي تكون

$$K = \frac{l^2}{1 + l^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (24)$$

وبالطريقة نفسها تحسب قيمة K التي تجعل $MSE(\tilde{b})$ اقل ما يمكن والتي تكون:

$$K = \frac{l^2}{1+l^2} \quad \text{.....(25)}$$

الجانب التجريبي

تم اعداد برنامج باستخدام لغة (Qbasic) لمحاكاة التجارب قيد الدراسة وقد تولى البرنامج مهمة توليد البيانات الخاصة بالمتغير التوضيحي والذي يتوزع توزيعا منتظما ضمن الفترة (٠,١) وتوليد البيانات الخاصة بمتغير الخطأ العشوائي الذي يتوزع طبيعيا تم افتراض قيمة الى كل من المعالم b, a وذلك لاجراء المقارنة بين الطرائق المستخدمة في انموذج الانحدار الخطي البسيط ودراسة تاثير كل من حجم ونسبة المشاهدات المفقودة، فقد تم في هذه الدراسة توليد عينات بحجم (١٠٠، ٥٠، ٣٠، ١٠ =n)، ولجميع هذه الحجوم فقد تم اختيار المشاهدات بشكل عشوائي تام وان نسب البيانات المفقودة هي (١٠%، ٢٠%، ٣٠%، ٤٠%) وتم حساب مقدار التحيز ومتوسط مربعات الخطأ لمقدرات المعالم باستخدام الطرائق الاتية:

أ-تحليل المشاهدات التامة CC.

ب-تحليل المشاهدات المتاحة AC.

ج-التقدير باستخدام مقدرات التقلص SH.

ومن ثم مقارنة متوسط مربعات الخطأ للمقدرات مع متوسط مربعات الخطأ لمقدر طريقة المربعات الصغرى الاعتيادية باستخدام الكفاءة النسبية.

$$R.E(\tilde{a}) = \frac{MSE(\hat{a})\ in OLS}{MSE(\hat{a})\ in another method} \quad \text{.....(26)}$$

ان الصيغة الرياضية للانموذج كانت كالاتي:

$$Y_i = 0.5 + 1.5X_i + e_i \quad \text{.....(27)}$$

جدول رقم (١): يوضح القيمة المطلقة للتحيز لمقدرات معلمات الانموذج الخطي البسيط

نسبة البيانات المفقودة		١٠%		٢٠%		٣٠%		٤٠%	
		a	b	a	b	a	b	a	b
N=10	SHI	٠,١٥٩٨	٠,٣٤١١٢	٠,١٥٠٨	٠,٣٥٧٦	٠,٠٨٩٥	٠,٢٧٦٩	٠,٠٤١٣	٠,١١١٢
N=30	SHI	٠,٠٢٩٥	٠,٠٥٥٠	٠,٠٤٩٤	٠,١٢٢١	٠,٠٢٧٢	٠,٠٩٩٣	٠,٠٢٧٨	٠,٠٨٣٨
N=50	SHI	٠,٠٢٥٠	٠,١٧٢٣	٠,٠٢٨٨	٠,٠٧٥٠	٠,٠٠٣١	٠,٠١٣٠	٠,٠٠٦٢	٠,٠٢٢٧
N=100	SHI	٠,٢٤٨١	٠,٢١٣٩	٠,٠٩٢٣	٠,٢٥٤٥	٠,٠٣٩٢	٠,٢٩١٥	٠,٠٣٧٨	٠,١٢٩٥

وعن طريق ملاحظة الجدول (١) والذي يمثل القيمة المطلقة لمقدار التحيز لمقدرات المعالم وفق طريقة SH حيث ان كلا من طريقتي CC و AC غير متحيزة استنتج آلاتي:

١-للعينة ذات حجم (n=10,100) لوحظ ان مقدار التحيز للمعلمة a يتناسب تناسبا عكسيا مع نسبة فقدان البيانات (١٠%، ٢٠%، ٣٠%، ٤٠%) حيث ان مقدار التحيز يقل بازدياد نسب فقدان البيانات.

٢-للعينة ذات حجم (n=30) لوحظ تذبذب بمقدار التحيز للمعلمة a ولنسب فقدان البيانات (١٠%، ٢٠%، ٣٠%، ٤٠%).

٣-للعينة ذات حجم (n=50) لوحظ ان مقدار التحيز للمعلمة a يتناسب تناسبا عكسيا مع نسبة فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ زيادة بمقدار التحيز عن ما هو عليه لنسبة فقدان بيانات (٣٠%).

٤-للعينة ذات حجم (n=10,30,50) لوحظ ان مقدار التحيز للمعلمة a يتناسب تناسبا عكسيا مع حجم العينة حيث يقل مقدار التحيز بازدياد حجم العينة، اما لحجم العينة (n=100) فلوحظ زيادة بمقدار التحيز للمعلمة a عن ما هو عليه حجم العينة (n=50)، أي ان مقدار التحيز للمعلمة a يتناسب تناسبا طرديا مع حجم العينات الكبيرة.

٥-لوحظ ان مقدار التحيز للمعلمة b يتصف بصفات مقدار التحيز للمعلمة a نفسها.

جدول رقم (٢): يوضح متوسط مربعات الخطأ لمقدرات معلمات الانموذج الخطي البسيط

نسبة البيانات المفقودة		%١٠		%٢٠		%٣٠		%٤٠	
		a	b	a	b	a	b	a	b
N=10	CC	٠,٤٨٩٥	٢,٤٩١٤	٠,٣٤٨١	١,٣٥٦٦	٠,٠٦٨٨	٠,٤٠٠٥	٠,٣٢٧١	١,٤١٣٧
	AC	١,٠٠٦٢	٥,٤٤٤٧	٠,٩٧٧٢	٤,٧١٣٦	٠,٣٥٩٠	٢,٥٢٩٨	١,٠٧٤٢	٥,٨١٩٩
	SHI	٠,٠٢٦٥	٠,١٧٣٣	٠,٠٢٣٦	٠,١٣٣٧	٠,٠٠٨٧	٠,٠٧٨٥	٠,٠٠١٥	٠,٠١٢٣
N=30	CC	٠,١٣٥٤	٠,٤٩٤٦	٠,١٢٣٧	٠,٤٩٣٤	٠,٠٣٢٨	٠,٢٠٤٠	٠,٠٩٠٢	٠,٣٩٦٥
	AC	٠,٢٢٤١	٠,٨٦٣٤	٠,٢٢٧٦	١,٠٤١٦	٠,٠٨٤٢	٠,٦٥٩١	٠,١٨٨٠	١,٠٥٨٨
	SHI	٠,٠٠٠٢	٠,٠٠٣١	٠,٠٠٢٧	٠,٠١٥٥	٠,٠٠٠٦	٠,٠٠٩٥	٠,٠٠٠٥	٠,٠٠٦٢
N=50	CC	٠,٠٧١٤	٠,٢٥٣١	٠,٠٥٤٠	٠,٢٢٨٧	٠,٠٣٢٣	٠,١٩٩٦	٠,٠٢٩١	٠,١٥٣٢
	AC	٠,١٠٦١	٠,٤٤٧٩	٠,٠٩٤٣	٠,٤٤٧٧	٠,٠٥٦٨	٠,٤٢١٥	٠,٠٦٧٤	٠,٤٣٠١
	SHI	٠,٠٠٦٩	٠,٠٣١٤	٠,٠٠٠٣	٠,٠٠٥٠	٠,٠٠٠٠١	0.0001	٠,٠٠٠٠٧	٠,٠٠٠٢
N=100	CC	٠,٠٢٤٣	٠,٠٩٠٠	٠,٠٢٥١	٠,١٠١٨	٠,٠١٩١	٠,٠٩٥٤	٠,٠١٦٢	٠,٠٨٨٨
	AC	٠,٠٣٩١	٠,١٥٥٩	٠,٠٤٥٣	٠,٢٠٩٥	٠,٠٣٤١	٠,٢٠٣٢	٠,٠٣٥٧	٠,٢٣٧١
	SHI	٠,٠٠١٧	٠,٠٠٩٧	٠,٠٠٠٢	٠,٠٠٠٩	٠,٠٠٠٤	٠,٠٠٣٨	٠,٠٠١٥	٠,٠١٦٥

وعن طريق ملاحظة الجدول (٢) والذي يمثل متوسط الخطأ لمقدرات معالم الانموذج الخطي البسيط استنتج الاتي:

١-تم حساب متوسط مربعات الخطأ للمعالم بطريقة OLS للبيانات التامة ولحجوم العينة المختارة كافة.

٢-للعينة ذات حجم (n=10,30) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة CC يتناسب تناسبا عكسيا مع نسبة فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة CC عن ما هو عليه لنسبة فقدان بيانات (٣٠%).

٣-للعينة ذات حجم (n=50) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة CC يتناسب تناسب عكسيا مع نسبة فقدان البيانات ولنسب فقدان البيانات المختارة.

٤-للعينة ذات حجم (n=100) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة CC يتناسب تناسبا عكسيا مع نسبة فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة CC عن ما هو عليه لنسبة فقدان بيانات (١٠%).

٥-لحجوم العينة المختارة كافة لوحظ استجابة طريقة CC لحجم العينة حيث ان متوسط مربع الخطأ للمعلمة a يقل بازدياد حجم العينة.

٦-لوحظ ان متوسط مربع الخطأ للمعلمة b بطريقة CC لها خصائص المعلمة a نفسها.

٧-للعينة ذات حجم (n=10,50) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة AC يتناسب تناسب عكسيا مع نسبة فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما نسبة فقدان بيانات (٤٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة AC عن ما هو عليه لنسبةفقدان بيانات (٣٠%).

٨-للعينة ذات حجم (n=30,100) لوحظ تذبذب بمتوسط مربع الخطأ للمعلمة a بطريقة AC .

٩-لحجوم العينة المختارة كافة لوحظ استجابة طريقة AC لحجم العينة حيث ان متوسط مربع الخطأ للمعلمة a يقل بازدياد حجم العينة.

١٠-لوحظ ان متوسط مربع الخطأ للمعلمة b بطريقة AC لها خصائص المعلمة a نفسها.

١١-للعينة ذات حجم (n=10) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH يتناسب تناسبا عكسيا مع نسبة فقدان البيانات حيث يقل متوسط مربع الخطأ بزيادة نسبة فقدان البيانات ولجميع نسب الفقدان.

١٢-للعينة ذات حجم (n=30) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH يتناسب تناسبا عكسيا مع نسبة فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة SH عن ما هو عليه لنسبة فقدان بيانات (١٠%).

١٣-للعينة ذات حجم (n=50) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH يقل بزيادة نسب فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة SH .

١٤-للعينة ذات حجم (n=100) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH يتناسب تناسبا طرديا مع نسب فقدان البيانات حيث يزداد متوسط مربع الخطأ للمعلمة a بزيادة نسب فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH اقل من متوسط مربع الخطأ للمعلمة a لنسبة فقدان بيانات (١٠%).

١٥-لحجوم العينة (n=10, 30, 50, 100) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH يقل بزيادة حجم العينة ما عدا حجم العينة (n=50) ولنسبة فقدان بيانات (١٠%) فلوحظ زيادة بمتوسط مربع الخطأ للمعلمة a بطريقة SH عن ما هو عليه لحجم العينة (n=30)، وكذلك لحجم العينة (n=100) ولنسبة فقدان بيانات (٣٠%، ٤٠%) لوحظ ان متوسط مربع الخطأ للمعلمة a بطريقة SH اكثر مما هو عليه للنسبة نفسها ولحجم العينة (n=50).

١٦-لوحظ ان متوسط مربع الخطأ للمعلمة b بطريقة SH لها خصائص المعلمة a نفسها.

١٧-لحجوم العينة (n=10, 30, 50, 100) لوحظ ان متوسط مربع الخطأ للمعالم بطريقة SH هي الاقل تليها بطريقة CC ثم بطريقة AC ولجميع نسب فقدان البيانات.

جدول رقم (٣): يوضح الكفاءة النسبية لمقدرات معلمات الانموذج الخطي البسيط

نسبة البيانات المفقودة		١٠%		٢٠%		٣٠%		٤٠%	
		a	*b*	*a*	*b*	*a*	*b*	*a*	*b*
N=10	CC	2.3346	١,٤٨٦١	٣,٢٨٢٩	٢,٧٢٩٣	١٦,٦١٠٤	٩,٢٤٤٩	٠,٣٢٧١	٢,٦١٩٠
	AC	١,١٣٥٧	٠,٦٨٠٠	١,١٦٩٤	٠,٧٨٥٥	٣,٠١٥٣	١,٤٦٣٥	١,٠٦٣٨	٠,٦٤٦١
	SHI	٤٣,١٢٤٥	٢١,٣٦٥٢	48.4237	٢٧,٦٩٣٣	١٣١,٣٥٦٣	٤٧,١٦٦٨	٧٦١,٨٦٦٦	٣٠١,٠٢٤٣
N=30	CC	١,٢٠٦٧	١,١٦١٥	١,٣٢٠٩	١,١٦٤٣	٤,٩٨١٧	٢,٨١٦١	١,٨١١٥	١,٤٤٨٩
	AC	٠,٦٠٨٣	٠,٦٦٥٣	٠,٧١٧٩	٠,٥٥١٥	١,٩٤٠٦	٠,٨٧١٦	٠,٨٦٩١	٠,٥٤٢٥
	SHI	٨١٧,٠	١٨٥,٣٢٢٥	٦٠,٥١٨٥	٣٧,٠٦٤٥	٢٧٢,٣٣٣٣	٦٠,٤٧٣٦	٣٢٦,٨	٩٢,٦٦١٢
N=50	CC	١,٠٧٤٢	١,٠١٦٥	١,٤٢٠٣	١,١٢٥٠	٢,٣٧٤٦	١,٢٨٩٠	٢,٦٣٥٧	١,٦٧٩٥
	AC	٠,٧٢٢٩	٠,٥٧٤٤	٠,٨١٣٣	٠,٥٧٤٧	١,٣٥٠٣	٠,٦١٠٤	١,١٣٧٩	٠,٥٩٨٢
	SHI	١١,١١٥٩	٨,١٩٤٢	٢٥٥,٦٦٦٦	٥١,٤٦	٧٦٧٠,٠	٢٥٧٣,٠	١٠٩٥,٧١٤٢	1286.5
N=100	CC	١,٤٥٢٦	٠,٨٢٨٧	١,٤٠٦٣	١,٠٦٦٧	١,٨٤٨١	١,١٣٨٣	٢,١٧٩٠	١,٢٢٢٩
	AC	٠,٩٠٢٨	٠,٦٩٦٦	٠,٧٧٩٢	٠,٥١٨٣	١,٠٣٥١	٠,٥٣٤٤	٠,٩٨٨٧	٠,٤٥٨٠
	SHI	٢٠,٧٦٤٧	١١,١٩٥٨	١٧٦,٥	١٢٠,٦٦٦٦	٨٨,٢٥	٢٨,٥٧٨٩	٢٣,٥٣٣٣	٦,٥٨١٨

وعن طريق ملاحظة الجدول (٣) والذي يمثل الكفاءة النسبية لمقدرات معالم الانموذج الخطي البسيط استنتج الاتي:

١-للعينة ذات حجم (n=10, 30) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة CC تتناسب تناسبا طرديا مع نسب فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة CC عن ما هو عليه لنسبة فقدان بيانات (٣٠%).

٢-للعينة ذات حجم (n=50) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة CC تتناسب تناسبا طرديا مع نسب فقدان البيانات.

٣-للعينة ذات حجم (n=100) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة CC تتناسب تناسبا طرديا مع نسب فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة CC عن ما هو عليه لنسبة فقدان بيانات (١٠%).

٤-لحجوم العينة المختارة كافة لوحظ استجابة طريقة CC لحجم العينة حيث ان الكفاءة النسبية للمعلمة a تزداد بازدياد حجم العينة.

٥-لوحظ ان الكفاءة النسبية للمعلمة b بطريقة CC لها خصائص المعلمة a نفسها.

٦-للعينة ذات حجم (n=10, 50) لوحظ ان الكفاءة النسبية للمعلمة a بطريقةAC تتناسب تناسبا طرديا مع نسب فقدان البيانات (١٠%، ٢٠%، ٣٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة AC عن ما هو عليه لنسبة فقدان بيانات (٣٠%).

٧-للعينة ذات حجم (n=30, 100) تذبذبت الكفاءة النسبية للمعلمة a بطريقة AC .

٨-لحجوم العينة المختارة كافة يتضح استجابة طريقة AC لحجم العينة حيث ان الكفاءة النسبية للمعلمة a تزداد بازدياد حجم العينة.

٩-لوحظ ان الكفاءة النسبية للمعلمة b بطريقة AC لها خصائص المعلمة a نفسها.

١٠-للعينة ذات حجم (n=10) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH تتناسب تناسبا طرديا مع نسبة فقدان البيانات حيث تزداد الكفاءة النسبية بزيادة نسبة فقدان البيانات ولجميع نسب الفقدان.

١١-للعينة ذات حجم (n=30) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH تتناسب تناسبا طرديا مع نسب فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة SH عن ما هو عليه لنسبة فقدان بيانات (١٠%).

١٢-للعينة ذات حجم (n=50) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH تزداد بزيادة نسب فقدان البيانات (٣٠%، ٢٠%، ١٠%) اما لنسبة فقدان بيانات (٤٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة SH عن ما هو عليه لنسبة فقدان بيانات (٣٠%).

١٣-للعينة ذات حجم (n=100) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH تتناسب تناسبا عكسيا مع نسب فقدان البيانات حيث تقل الكفاءة النسبية للمعلمة a بزيادة نسب فقدان البيانات (١٠%، ٣٠%، ٤٠%) اما لنسبة فقدان بيانات (٢٠%) فلوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH اكثر من الكفاءة النسبية للمعلمة a لنسبة فقدان بيانات (١٠%).

١٤-لحجوم العينة (n=10, 30, 50, 100) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH تزداد بزيادة حجم العينة ما عدا حجم العينة (n=50) ولنسبة فقدان بيانات (١٠%) فلوحظ نقصان بالكفاءة النسبية للمعلمة a بطريقة SH عن ما هو عليه لحجم العينة (n=30)، وكذلك لحجم العينة (n=100) ولنسبة فقدان بيانات (٣٠%، ٤٠%) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH اقل مما هو عليه للنسبة نفسها ولحجم العينة (n=50).

١٥-لوحظ ان الكفاءة النسبية للمعلمة b بطريقة SH تمتاز بصفات وخصائص المعلمة a نفسها.

١٦-احجوم العينة (n=10, 30, 50, 100) لوحظ ان الكفاءة النسبية للمعلمة a بطريقة SH هي الاكثر تليها بطريقة CC ثم بطريقة AC ولجميع نسب الفقدان.

جدول رقم (٤): يوضح قيم K لمقدرات معلمات الانموذج الخطي البسيط

نسبة البيانات المفقودة		١٠%		٢٠%		٣٠%		٤٠%	
		a	b	a	b	a	b	a	b
N=10	K	0.0336	٠,٠٢٧٨	٠,٠٢٠٤	٠,٠٣٥١	٠,٠٢٧٩	٠,٠٢٧١	٠,٠٠١٥	٠,٠٠٢٢
N=30	K	0.0025	٠,٠٠٣٧	٠,٠٠٩٩	٠,٠١٦٨	٠,٠٠٨٠	٠,٠١٦٥	٠,٠٠٤٤	٠,٠٠٦١
N=50	K	0.0846	٠,٠٦٤٥	٠,٠٠٩٩	٠,٠١١٥	٠,٠٠٠٢	٠,٠٠٠٤	٠,٠٠٠٦	٠,٠٠١٢
N=100	K	٠,٠٥٨٩	٠,٠٥٣٥	٠,٠٠٣١	٠,٠٠٤٢	٠,٠١٣٠	٠,٠١٧١	٠,٠٥٧٣	٠,٠٥٩٨

ومن خلال ملاحظة الجدول (٤) والذي يمثل قيمة K لمعالم الانموذج الخطي البسيط استنتج الاتي:

١-ان قيمة K تتراوح بين الصفر والواحد أي ان $0 \le K \le 1$.

٢-لوحظ عدم تأثر قيمة K بحجم العينة.

٣-لوحظ عدم تأثر قيمة K بنسبة فقدان البيانات.

الاستنتاجات والتوصيات

١-ان مقدر المربعات الصغرى (OLS) ومقدر المشاهدات التامة (CC) ومقدر المشاهدات المتاحة (AC) تمتاز جميعها بخاصية عدم التحيز اما مقدر التقلص فانه مقدر متحيز لاحجام العينات كافة ولنسب فقدان المشاهدات المعتمدة.

٢-ان مقدر التقلص المقترح لتقدير معلمات الانموذج الخطي البسيط كان افضل واحسن من مقدر طريقة OLS ومقدر طريقة CC ومقدر طريقة AC لانه يمتلك اقل متوسط مربعات خطأ قياسا الى مقدرات الطرق اعلاه.

٣-ان الكفاءة النسبية لمقدر التقلص المقترح كانت كبيرة وعالية جدا لاحجام العينات كافة ولنسب الفقدان المعتمدة بالنسبة لجميع الطرق.

٤-ان قيم K في مقدر التقلص الثابت ولنسب الفقدان جميعها كانت محصورة بين الصفر والواحد وهو يلائم الشرط الذي وضع في الجانب النظري على المقدر المقترح.

٥-نوصي بتطوير مقدرات التقلص وذلك عن طريق اقتراح دوال تقلص اخرى ويفضل ان تكون متغيرة.

المصادر

1-Afifi,A.A., and Elashoff, R.M,"Missing Observation in Multivariate Statistics, Anote on Simple Linear Regression" , *JASA*, Vol 64, pp 358-365, 1969

2-Little,R.J.A., "Regression with Missing X'S " ,*JASA*, Vol 87, pp 1227-1237, 1992.

4-Mehta,J.S., and Srinvasan R., "Estimation of the Mean by Shrinkage to a Point" ,*JASA*,Vol 66, pp86-90, 1971.

3-Rubin,D.B "Comparing Regression when Some Predictor Variables are Missing", *Technometrics*,Vol 18,pp 201-205, 1976.