طريقة (James – Stein) لتقدير المشاهدات المفقودة في نماذج الانحدار الخطية

باسم منصور عبد الأحد*

المستخلص

يتضمن هذا البحث استخدام طريقة (James - Stein) لتقدير متجه المشاهدات المفقودة (من المتغير التابع) في نموذج الانحدار الخطي، وتبين بأنه إذا كان هناك أكثر من قيمتين مفقودتين فأن التقديرات المستحصل عليها تمتلك خطورة اقل من تلك المقدرة بالطرق الاعتيادية بالإضافة إلى ذلك عند اعتماد تقدير (James - Stein) لمتجه المشاهدات المفقودة لن يتطلب الأمر أية قيود على الشكل الدالي للنموذج . وقد تم إعداد برنامج على الحاسوب مع إعطاء مثال توضيحي لتطبيق الطريقة أعلاه حيث تم إلغاء المشاهدات الخمس الأخيرة للمتغير التابع واعتبرت مفقودة وبتطبيق الطريقة أعلاه وفقاً للبرنامج المعد لها تم الحصول على التقديرات المطلوبة.

ABSTRACT

This research includes using (James - Stein) method to estimate the missing observation vector (from the depent veriable) in liner regression model and we found that , if there were more than two missing values , then the estimata that we got had less dangerous form those that had been estimated in normal ways , Besides the (James - Stein) vector for missing observation vector won't demand any substruct on the formal function and it had been done a program on the computer with an example explein the method that mentioned above and the last fifth observation for the depend veriable had been deleted and considered missing , so by applying the method mentioned above , we had got the demanded estimates .

المقدمة:

من الملاحظ ان نماذج الانحدار الخطية واسعة الاستخدام في كثير من العلوم الطبيعية والانسانية وما لهذه النماذج من اهمية كبيرة في بناء القرارات المستقبلية من خلال تقدير معالم النموذج الخطي والذي يعتمد بصورة اساسية على المشاهدات المأخوذة من مجتمع قيد الدراسة. ان المشكلة الاساسية التي تواجه الباحثين هي حالة وجود نقص او فقدان بعض المشاهدات والتي قد تكون ضرورية ومؤثرة في القرارات المتخذة . من هنا تراء لنا ان نستعرض في هذا البحث تقدير جيمس ستين (James Stein) لمتجه المشاهدات المفقودة والبرهنة على انه افضل من التقدير

* مدرس /قسم الاحصاء/ كلية الإدارة والاقتصاد /الجامعة المستنصرية

بطريقة المربعات الصغرى (Least Squres Method) باستخدام مفهوم دوال المخاطرة وكذلك استنباط تعديل لتقدير جيمس ستين (Modified James Stein) بحيث يعطي خطورة اقل من التقدير الاصلي .

في بعض الحالات يمكن النظر الى المشاهدات المفقودة على انها معالم اضافية Additional (Parameteres) يتوجب علينا تقديرها ، فقد قدر (Selove , 1972) متجه المشاهدات المفقودة في نموذج الانحدار الخطي بواسطة تقليل مجموع مربعات البواقي .

(Residual Sum of Squres) كما وحصل (Radhakrishnan 1982) على فئة من المقدرات لمتوسط التـــوزيع الحدي لمتغير عشوائي طبيعي في حالـــة وجود مشاهدات مفقودة ، وهذه المقدرات تمتلك دائماً اقل خطورة من تقديرات الامكان الاعظم .

في هذا البحث تم الحصول على تقدير (James- Stein) لمتجه المشاهدات المفقودة (Y_m) ، عناصر هذا التقدير يمكن استخدامها بصورة مبدئية كتنبؤات لمشاهدات مستقبلية . وسوف نوضح في فقرة لاحقة انه عندما ($N > 2$) فأن تقدير (James – Stein) يمتلك وفي كافة الاحوال خطورة اقل من تلك التي يمتلكها تقدير المربعات الصغرى الاعتيادية . كذلك تم أدخال على هذا التقدير تحسينات اخرى باستخدام مفهوم المخاطرة ، من خلال الاخذ بنظر الاعتبار الجزء الموجب المناظر للتقدير ، كذلك لابد من الا شارة هنا انه لغرض تطبيق طريقة (James-Stein) يتطلب الامر وجود ثلاث متغيرات توضيحية او اكثر وان يكون عدد المشاهدات المفقودة اكثر من مشاهدتين .

هدف البحث :

يهدف هذا البحث الى استعراض طريقة جيمس ستين (James Stein Method)لتقدير متجه المشاهدات المفقودة والبرهنة على انه افضل من التقدير بطريقة المربعات الصغرى (Least Squres Method) وذلك من خلال استخدام مفهوم دوال المخاطرة وكذلك استنباط تعديل لتقدير جيمس ستين (Modified James Stein) بحيث يعطي خطورة اقل من التقدير الاصلي .

استعراض البحث :

لقد اعتمدت السياقات اللاحقة في متن البحث على النموذج الخطي التالي:

$$Y = X\beta + e \quad ; \quad e \sim N(O, \sigma^2 Ir)$$

حيث ان :

(Y) و (e) : متجهين ذات مرتبة $(r * 1)$ الاول (Y) يمثل متجه المشاهدات عن المتغير المعتمد والثاني (e) يمثل متجه الاخطاء العشوائية .

(X) : مصفوفة ذات مرتبة $(r * P)$ تمثل مصفوفة المشاهدات عن المتغيرات التوضيحية البالغ عددها $(P - 1)$ وان (r) تمثل عدد المشاهدات عن كل متغير توضيح وان رتبة (X) هي (P) .

اي ان $(r(X) = P)$.

(β) : يمثل متجه معالم النموذج المطلوب تقديره

$(I r)$: يمثل المصفوفة الاحادية ذات المرتبة $(r * r)$.

الان لنفرض ان (N) تمثل عـدد المشاهدات المفقـودة عـن المتغير التابع حيث ان (N> 2)، بعد اعادة ترتيب عناصر المتجه (Y) بحيث يؤدي ذلك الى امكانية تجزئة واعادة صيغة النموذج اعلاه بالشكل التالي :

$$\begin{bmatrix} Y_a \\ Y_m \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \qquad \ldots(1)$$

حيث ان الدليل " a " يشير الى المشاهدات المتاحة وان الدليل " m " يشير الى المشاهدات المفقودة .

كذلك نفرض ان (n = (r- N) > P) اي ان عدد المشاهدات المتاحة اكبر من عدد المعالم المطلوب تقديرها والبالغ عددها (P) (بضمنها المقطع) .

كذلك سوف نفترض انه لا يوجد اي عنصر مفقود من المصفوفة

$$X' = [X'_1, X'_2]$$

نلاحظ من الصيغة (1) ان الجزئين (e_2 , e_1) يتوزعان بشكل مستقل وكما يلي :

$$e_1 \sim N(0, \sigma^2 {}_{In}) \quad ; \quad e_2 \sim N(0, \sigma^2 {}_{IN})$$

ولغرض عرض نتائج البحث لابد من التطرق لبعض التعاريف والرموز والمصطلحات التي سوف نستخدمها لاحقاًوهي :

$t_r(A)$: يمثل اثر المصفوفة المربعة (A) ويساوي مجموع عناصر القطر الرئيسي للمصفوفة (A) .

$U \sim N(0, I_m)$: المتجه العشوائي (U) يتوزع طبيعياً بمتجه الاوساط (o) وبمصفوفة تباين وتباين مشترك (I_m) .

$V \sim \chi^2_r$: المتغيـر العشوائي (V) يتوزع وفق توزيع مربع كاي بـ (r) درجة حرية .

$K \sim P_{ois}(M)$: المتغير العشوائي (K) يتوزع وفق توزيع بواسون بالمعلمة (M) .

وكذلك يجب ملاحظة الرموز التالية وما يكافئها:

لاي متجه $\|\eta\|^2 = \eta'\eta$

$$M = X_2(X_1'X)^{-1} \quad ; \quad \Sigma = M'M = X_2(X_1'X)^{-1}X'_2$$

$$\theta = \Sigma^{-1}X_2\beta/\sigma \quad ; \quad W = M'\Sigma^{-1}M$$

على افتراض ان (Σ^{-1}) موجودة ومعرفة .

الا ن ليكن :

$$\hat{Y}_m = X_2\hat{\beta} \qquad ...(2)$$

يمثل تقدير المربعات الصغرى لـ (Y_m) حيث ان :

$$\hat{\beta} = (X_1'X)^{-1}X_1'Y_a \sim N\left(\beta, \sigma^2(X_1'X)^{-1}\right)$$

عندئذ فأن :

$$D(\hat{Y}_m) = E[Y_m - E(Y_m)]\left[\hat{Y}_m - E(\hat{Y}_m)\right]' = \sigma^2\Sigma \qquad ...(3)$$

مفترضين ان $D(\hat{Y}_m)$ هي مصفوفة موجبة محددة وان :

$$\tilde{Y}_m = C\hat{Y}_m \quad ; \quad 0 < C < 1 \qquad ...(4)$$

والتي تجعل الصيغة التالية اقل ما يمكن :

...(5)

$$R\left(\tilde{Y}_m ; \beta , \sigma^2\right)= E\left[\left(\tilde{Y}_m - Y_m\right)'\left(D^{-1}\tilde{Y}_m\right)\left(\tilde{Y}_m - Y_m\right)\right]$$

علماً ان التوقع يؤخذ على التوزيع المشترك لـ (Y_m) و $(\tilde{Y}_m)$ الان يمكن البيان ان :

...(6)

$$R\left(\tilde{Y}_m ; \beta, \sigma^2\right)=\left(C-1\right)^2 \beta' X_2' \Sigma^{-1} X_2 \beta\sigma^2 + C^2 t_r(w) + t_r\left(\Sigma^{-1}\right)$$

علماً ان $R\left(\tilde{Y}_m ; \beta , \sigma^2\right)$ تكون اصغر ما يمكن عندما تكون :

...(7)

$$C = \left[1 - \sigma^2 t_r(w) \Big/ \left(\beta' X_2' \Sigma^{-1} X_2 \beta + \sigma^2 t_r(w)\right)\right]$$

القيمة المثلى المعطاة لـ (C) في الصيغة (7) ليست عملية بسبب كونها دالة بدلالة المعلمتين المجهولتين (β) و (σ^2) وان

...(8)

$$\hat{\sigma}^2 = \left(Y_a - \hat{Y}_a\right)'\left(Y_a - \hat{Y}_a\right)\Big/(n-p) = Y_a' Ya - \beta' X_1' Y_a/(n-p)$$

هو تقدير غير متحيز للمعلمة (σ^2) كماوانه من خلال الصيغة (2) فأن :

... (9)

$$E\left[\hat{\beta}' X_2' \Sigma^{-1} X_2 \beta\right] = \beta' X_2' \Sigma^{-1} X_2 \beta + \sigma^2 t_r(w)$$

ومن خلال الاستعاضة عن التقدير غير المتحيز لـ (σ^2) وعن المقدار $\left(\beta' X_2' \Sigma^{-1} X_2 \beta + \sigma^2 t_r(w)\right)$ في الصيغة (7) فأن التقدير الى (C) يمكن الحصول عليه بالصيغة :

$$\hat{C} = \left[1 - \hat{\sigma}^2 t_r(w) \Big/ \hat{\beta}' X_2' \Sigma^{-1} X_2 \beta\right]$$

الان يمكن تعريف متجه تقديرات (James – Stein) الى (Y_m) من خلال الصيغة :

...(10)

$$\hat{\hat{Y}}_m = \left[1 - \lambda\sigma^2 t_r(w) / \left(\hat{\beta}' X_2' \Sigma^{-1} X_2 \hat{\beta}\right)\right] X_2 \hat{\beta}$$

حيث ان $(\lambda>0)$ وهي كمية ثابتة احتباطية وسوف نحدد قيمتها في فقرة لاحقة .

ان التقديرات التي حصلنا عليها سوف تقارن مع بعضها من خلال مفهوم دوال المخاطرة .

$$R\left(\hat{Y}_m;\beta,\sigma^2\right) = E\left[\left(\hat{Y}_m - Y_m\right)'\left(D^{-1}\hat{Y}_m\right)\left(\hat{Y}_m - Y_m\right)\right] \quad ...(11)$$

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right) = E\left[\left(\hat{\hat{Y}}_m - Y_m\right)'\left(D^{-1}\hat{Y}_m\right)\left(\hat{\hat{Y}}_m - Y_m\right)\right] \quad ...(12)$$

حيث ان التوقعات تؤخذ على التوزيع المشترك ولكل المتغيرات العشوائية ذات العلاقة .

يقال ان تقديرها مهمين بشكل كامل على تقدير اخر اذا كانت مخاطرته هي اقل وبشكل كامل من مخاطرة التقدير الاخير لجميع قيم معلمة التقدير .

لو جعلنا (C=1) في الصيغة (6) نحصل على :

$$R\left(\hat{Y}_m;\beta,\sigma^2\right) = t_r(w) + t_r\left(\Sigma^{-1}\right) \quad ...(13)$$

❖ **اشتقاق صيغة تقدير جيمس ستين لـ Y_m**

James – stein Estimator for Y_m

لاشتقاق صيغة تقدير جيمس ستين لـ Y_m سوف نحتاج الى المساعدات التالية لحل دالة المخاطرة

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right)$$

المساعدة Lemma (1) :

ليكن (u) متجه عشوائي ذي بعد (m) يتوزع طبيعياً بمتجه الاوساط (o) وبمصفوفة تباين وتباين مشترك (Im) وان (K) يتوزع توزيع بواسون بالمعلمة $\left(\frac{1}{2}\|\theta\|^2\right)$

عندئذ :

$$(i)\ E\left(1/\|U\|^2\right) = E\left[1/(2K + m - 2)\right]$$

$$(ii)\ E\left(1 - \theta' U/\|V\|^2\right) = (m - 2)E\left[1/(2K + m - 2)\right]$$

المساعدة Lemma (2) :

ليكن K) متغير عشوائي يتوزع توزيع بواسون بالمعلمة $\left(\frac{1}{2}\|\theta\|^2\right)$ وان :

$$\theta = \left(\sum^{-\frac{1}{2}} X_2 \beta / \sigma\right)$$

عندئذ:

$$R\left(\hat{Y}_m ; \beta, \sigma^2\right) = R\left(\hat{Y}_m ; \beta, \sigma^2\right) - \left[2\lambda(N - 2)t_r(w) - \lambda^2 \frac{n - p + 2}{n - p}\right.$$
$$\left.\{t_r(w)\}^2\right] E\left[1/(2K + N - 2)\right]$$

حيث ان $R\left(\hat{Y}_m ; \beta, \sigma^2\right)$ معطاة في الصيغة (13)

البرهان:

باستخدام الصيغتين (3) و (10) في الصيغة (12) مع اعتماد تعريف Y_m في الصيغة (2)واخذ التوقع على توزيع (Y_m) نحصل على :

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right)=R\left(\hat{Y}_m;\beta,\sigma^2\right)-2\left(\lambda/\sigma^2\right)t_r(w)*\frac{E\left[\hat{\sigma}^2\left(X_2\hat{\beta}-X_2\beta\right)'\right]\Sigma^{-1}X_2\hat{\beta}}{\left(\hat{\beta}'X_2'\Sigma^{-1}X_2\hat{\beta}\right)}+\left(\lambda^2/\sigma^2\right)\left\{t_r(w)\right\}^2*\frac{E\left[\hat{\sigma}^4/\left(\hat{\beta}'X_2'\Sigma^{-1}X_2\hat{\beta}\right)\right]}{}$$

...(14)

حيث ان $R\left(\hat{Y}_m;\beta,\sigma^2\right)$ معطاة في الصيغة (13) ، انه لمعلوم وبشكل واضح ان $(\hat{\beta})$ و$(\hat{\sigma}^2)$ يتوزعان بشكل مستقل وان :

$$\left[(n-p)\hat{\sigma}^2/\sigma^2\right]\sim\chi^2(n-p)$$

وعليه فأن :

...(15)

$$E\left(\hat{\sigma}^2\right)=\sigma^2\ ;\ E\left(\hat{\sigma}^4\right)=(n-p+2)\sigma^4/(n-p)$$

كذلك وباخذ التوقع الرياضي على توزيع $(\hat{\sigma}^2)$ نجد ان :

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right)=R\left(\hat{Y}_m;\beta,\sigma^2\right)-2\lambda t_r(w)E\left[\frac{1-\beta'X_2'\Sigma^{-1}X_2\hat{\beta}}{\left(\beta'X_2'\Sigma^{-1}X_2\hat{\beta}\right)}\right]-\lambda^2\frac{(n-p+2)}{(n-p)}\left\{t_r(w)\right\}^2E\left[\sigma^2/\left(\hat{\beta}'X_2'\Sigma^{-1}X_2\hat{\beta}\right)\right]$$

...(16)

ليكن :

$$Z=\frac{1}{\sigma}\sum^{-\frac{1}{2}}X_2\hat{\beta}\sim N(\theta,I_N);=\frac{1}{\sigma}\sum^{-\frac{1}{2}}X_2\beta \qquad ...(17)$$

الان يمكننا اعادة كتابة الصيغة (16) بالشكل التالي

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right)=R\left(\hat{Y}_m;\beta,\sigma^2\right)-2\lambda t_r(w)E\left[1-\theta' Z/\|Z\|^2\right]+$$

$$\lambda^2\left(\frac{n-p+2}{n-p}\right)\{t_r(w)\}^2E\left[1/\|Z\|^2\right]$$

حيث ان (θ) و (Z) معرفتان في الصيغة (17) وبتطبيق المساعدة (1) نحصل على :

$$R\left(\hat{\hat{Y}}_m;\beta,\sigma^2\right)=R\left(\hat{Y}_m;\beta,\sigma^2\right)-\left[2\lambda(N-2)tr(w)-\right.$$

$$\left.\lambda^2\left(\frac{n-p+2}{n-p}\right)\{t_r(w)\}^2E\{1/(2k+N-2)\}\right]$$

حيث ان :

$$K\sim pois\left(\frac{1}{2}\|\theta\|^2\right)$$

وهكذا يكمل البرهان.

وبما ان $(N>2)$ و $(K\geq\theta)$ فأن :

$$E[1/(2K+N-2)]>\theta$$

وبشكل واضح فأن $R\left(\hat{Y}_m;\beta,\sigma^2\right)$ تكون اقل ما يمكن عندما :

$$\lambda=(n-p)(N-2)/[(n-p+2)t_r(w)] \qquad ...(18)$$

عندئذ وبأفتراض ان :

$$\hat{Y}^{*}m = \left[1 - \propto (n-p)\hat{\sigma}^{2} / \left(\hat{\beta}' X_2' \Sigma^{-1} X_2 \hat{\beta}\right)\right] X_2 \hat{\beta} \qquad ...(19)$$

حيث ان :

$$\propto = (N-2)/(n-p+2)$$

عندئذ فأن $(\hat{Y}^{*}m)$ يمثل تقدير غير متحيز لـ (Y_m) ويطلق عليه اسم تقدير (James –Stein) لمتجه المشاهدات المفقودة (Y_m) وتعرف دالة المخاطرة لهذا التقدير بالصيغة التالية :

$$R\left(\hat{Y}^{*}m; \beta, \sigma^{2}\right) = E\left[\left(\hat{Y}^{*}m - Y_m\right)' \left(D^{-1}\hat{Y}_m\right)\left(\hat{Y}^{*}m - \hat{Y}_m\right)\right]$$

حيث ان التوقع يؤخذ على التوزيع المشترك لكل المتغيرات العشوائية ذات العلاقة.

نظرية (1) :

اذا كان $(N>2)$ عندئذ ولتقدير (Y_m) فأن التقدير $(\hat{Y}^{*}m)$ يهيمن وبشكل كامل على تقدير المربعات الصغرى $(\hat{Y}_m)$ وتكون دالة المخاطرة لـ $(\hat{Y}^{*}m)$ معطاة بالصيغة :

$$R\left(\hat{Y}^{*}m; \beta, \sigma^{2}\right) = R\left(\hat{Y}_m; \beta, \sigma^{2}\right) - \propto (n-p)(N-2)E\left[(2K+N-2)^{-1}\right]$$

حيث ان :

$$\propto = (N-2)/(n-p+2), K \sim pois\left(\left(\beta' X_2' \Sigma^{-1} X_2 \beta\right)/2\sigma^{2}\right)$$

البرهان:

نعوض عن قيمة (λ) المعطاة في الصيغة (18) في المساعدة الثانية علماً بأنه ولجميع قيم المعلمة فأن :

$$R\left(\hat{Y}_m; \beta, \sigma^{2}\right) - R\left(\hat{Y}^{*}m; \beta, \sigma^{2}\right) = \propto (n-p)(N-p)E\left[1/(2K+N-2)\right]$$

ولما كانت :

$$n = (r-N) > p$$

وان $(N>2)$ لذلك فأن :

$$R\left(\hat{Y}_m;\beta,\sigma^2\right)-R\left(\hat{Y}^*{}_m;\beta,\sigma^2\right)>Zero$$

وهذا يؤدي الى :

$$R\left(\hat{Y}_m;\beta,\sigma^2\right)>R\left(\hat{Y}^*{}_m;\beta,\sigma^2\right)$$

اي ان المخاطرة باستخدام التقدير $\left(\hat{Y}^*_m\right)$ تكون اقل من نظيرتها في حالة استخدام تقدير المربعات الصغرى لـ $\left(\hat{Y}_m\right)$ وهذا يعني ان $\left(\hat{Y}^*_m\right)$ مهيمن بشكل كامل على $\left(\hat{Y}_m\right)$ ، وحيث ان $(\theta\le K)$ فأن $E\left[1/(2K+N-2)\right]$ موجود عندما تكون $(N>2)$ وعندئذ اذا كان عدد المشاهدات المفقودة في نموذج انحدار طبيعي اكثر من مشاهدتين عندئذ فأن طريقة (James – Stein) يمكن استخدامها لتقدير تلك المشاهدات المفقودة وان التقدير المستحصل عليه باستخدام هذه الطريقة يمتلك اقل مخاطرة من تلك التي يمتلكها تقدير المربعات الصغرى .

❖ **تقدير (James –Stein) المعدل لـ (Y_m)**

The Modified James –Stein Estimator for (Y_m)

انه لمن المعلوم بأن العامل المنكمش

...(20)

$$h\left(\hat{\beta},\hat{\sigma}^2\right)=\left[1-(n-p)\propto\sigma^2/\left(\hat{\beta}'X_2'\Sigma^{-1}X_2\hat{\beta}\right)\right]$$

كما هو معطى في الصيغة (19) ليس من الضروري ان يكون غير سالب وعليه يمكن الاستعاضة عنه بـ

...(21)

$$h^+\left(\hat{\beta},\hat{\sigma}^2\right)=\left[1-(n-p)\propto\sigma^2/\left(\hat{\beta}'X_2'\Sigma^{-1}X_2\hat{\beta}\right)\right]^+$$

حيث ان $h^+\left(\hat{\beta},\hat{\sigma}^2\right)$ تمثل الجزء الموجب من الدالة $h\left(\hat{\beta},\hat{\sigma}^2\right)$ اي ان :

$$h^+\left(\hat{\beta},\hat{\sigma}^2\right)=Max\left[o;h\left(\hat{\beta},\hat{\sigma}^2\right)\right]$$

الآن لنفرض تقدير أخر الى (Y_m) مثل $(\tilde{Y}_m)$ المعطى بالصيغة :

$$\tilde{Y}_m = h^+(\hat{\beta}, \hat{\sigma}^2) X_2 \hat{\beta} \qquad \ldots(22)$$

ان التقدير $(\tilde{Y}_m)$ يمكن ان يسمى بتقدير (James –Stein) المعدل لـ (Y_m) وسوف نختبر سلوك $(\tilde{Y}_m)$ من خلال حساب دالة المخاطرة له :

$$R(\tilde{Y}_m; \hat{\beta}, \sigma^2) = E\left[(\tilde{Y}_m - Y_m)'(D^{-1}Y_m)(\tilde{Y}_m - Y_m)\right] \qquad \ldots(23)$$

حيث ان التوقع يؤخذ على التوزيع المشترك لـ $(\tilde{Y}_m)$ و $(\hat{\beta})$ $(\hat{\sigma}^2)$.

المساعدة (3) Lemma :

لتكن (Z) تتوزع توزيعاً طبيعياً متعدد المتغيرات بمتجه الاوساط (θ) ومصفوفة التباين والتباين المشترك (I_N) وان (S) يتوزع كتوزيع مربع كاي بـ (m) درجة حرية مستقلاً عن (Z) . لتكن (h) دالة بحيث ان التوقعات التالية تكون معرفة عندئذ :

$$E\left[h\left(\|Z\|^2/S\right)\theta' Z\right] = \|\theta\|^2 E\left[h\left(\chi^2_{(2K+N+2)}\right)/\chi^2_m\right]$$

نظرية 2 :

لتقدير (Y_m) فأن التقدير

$$\tilde{Y}_m = h^+(\hat{\beta}, \hat{\sigma}^2) X_2 \hat{\beta}$$

يهيمن وبشكل كامل على التقدير $(\hat{Y}^*_m)$ المعطى في الصيغة (19) حيث ان $h^+(\hat{\beta}, \hat{\sigma}^2)$ معرفة في الصيغة (21) .

البرهان:

من الصيغ (19) و (20) و (22) و (23) لدينا :

$$R\left(\hat{Y}^{*}_{m};\beta,\sigma^{2}\right)-R\left(\tilde{Y}_{m};\beta,\sigma^{2}\right)=F_{1}-2F_{2} \qquad ...(24)$$

حيث ان :

$$F_{1}=E\left[\left\{\left(h\left(\hat{\beta},\hat{\sigma}^{2}\right)\right)^{2}-\left(h^{+}\left(\hat{\beta},\hat{\sigma}^{2}\right)\right)^{2}\right\}\left(\hat{\beta}'X_{2}'\left(D^{-1}\hat{Y}_{m}\right)X_{2}\hat{\beta}\right)\right] \qquad ...(25)$$

$$F_{2}=E\left[\left\{h\left(\hat{\beta},\hat{\sigma}^{2}\right)-h^{+}\left(\hat{\beta},\hat{\sigma}^{2}\right)\right\}\left\{Y'_{m}\left(D^{-1}\hat{Y}_{m}\right)X_{2}\hat{\beta}\right\}\right] \qquad ...(26)$$

ومن الصيغة (3) نرى ان :

$$D\left(\hat{Y}_{m}\right)=\sigma^{2}\Sigma \Rightarrow D^{-1}\hat{Y}_{m}=\sigma^{2}\Sigma^{-1}$$

وبأخذ التوقع لـ $\left(Y_{m}\right)$ والاستعاضة بالصيغتين (20) و (21) يمكن اعادة كتابة (F_2) بالصيغة:

$$F_{2}=E\left[\left\{1-\frac{\propto\left(n-p\right)\hat{\sigma}^{2}}{\hat{\beta}'X_{2}'\Sigma^{-1}X_{2}\hat{\beta}}\right\}\left\{\frac{\hat{\beta}'X_{2}'\Sigma^{-1}X_{2}\hat{\beta}}{\sigma^{2}}\right\}-\left\{1-\frac{\propto\left(n-p\right)\hat{\sigma}^{2}}{\hat{\beta}'X_{2}'\Sigma^{-1}X_{2}\hat{\beta}}\right\}^{+}\left\{\frac{\hat{\beta}'X_{2}'\Sigma^{-1}X_{2}\hat{\beta}}{\sigma^{2}}\right\}\right] \qquad ...(27)$$

حيث ان :

$$\propto=\left(N-2\right)/\left(n-p+2\right)$$

وبأستخدام الصيغة (17) في الصيغة (27) نحصل على :

$$F_2 = E\left[\left(1-\frac{\propto (n-p)\hat{\sigma}^2/\sigma^2}{\|Z\|^2}\right)\theta' Z-\left(1-\frac{\propto (n-p)\hat{\sigma}^2/\sigma^2}{\|Z\|^2}\right)^{+}\theta' Z\right]$$

ونلاحظ من الصيغتين (17) و (8) ان :

$$Z \sim N(\theta, I_N)$$

مستقلة عن :

$$\left((n-p)\hat{\sigma}^2/\sigma^2\right) \sim \chi^2_{(n-p)}$$

وبتطبيق المساعدة (3) نلاحظ بأن :

$$F_2 = \|\theta\|^2 E\left[\left\{1-\propto\left(\chi^2_{(n-p)}/\chi^2_{(2K+N+2)}\right)\right\}-\left\{1-\propto\left(\chi^2_{(n-p)}/\chi^2_{(2K+N+2)}\right)\right\}^{+}\right]$$

...(28)

حيث ان $(\propto)$ معرفة بالصيغة (19) وان $K \sim pois\left(\frac{1}{2}\|\theta\|^2\right)$ وبتعويض الصيغتين (25) و (28) في الصيغة (24) نحصل على :

$$\begin{aligned} R(\hat{Y}^*_m;\beta,\sigma^2)-R(\tilde{Y}_m;\beta,\sigma^2) = {} & E[\{(h(\hat{\beta},\hat{\sigma}^2))^2-(h^+(\hat{\beta},\hat{\sigma}^2))^2\} \\ & \{(\hat{\beta}' X_2'(D^{-1}\hat{Y}_m)X_2\hat{\beta})\}]-2\|\theta\|^2 \\ & E[\{1-\propto(\chi^2_{(n-p)}/\chi^2_{(2K+N+2)})\}- \\ & \{1-\propto(\chi^2_{(n-p)}/\chi^2_{(2K+N+2)})\}^{+}] \end{aligned}$$

...(29)

ومن تعريف الجزء الموجب للدالة نستنتج ان الطرف الايمن من الصيغة (29) يكون موجباً لكل قيم المعلمة اي ان :

$$R\left(\hat{Y}^{*}{}_{m};\beta,\hat{\sigma}^{2}\right)-R\left(\tilde{Y}_{m};\beta,\hat{\sigma}^{2}\right)>zero$$

وهذا يؤدي الى :

$$R\left(\hat{Y}^{*}{}_{m};\beta,\hat{\sigma}^{2}\right)>R\left(\tilde{Y}_{m};\beta,\hat{\sigma}^{2}\right)$$

وبعبارة اخرى ان المخاطرة في استخدام تقديـــر (James – Stein) المعدل $\left(\tilde{Y}_{m}\right)$ اقل من المخاطرة في حالة استخدام تقدير (James – Stein) $\left(\hat{Y}^{*}_{m}\right)$ وبالتالي فالنتيجة النهائية ستكون :

$$R\left(\hat{Y}_{m};\beta,\hat{\sigma}^{2}\right)>R\left(\hat{Y}^{*}{}_{m};\beta,\hat{\sigma}^{2}\right)>R\left(\tilde{Y}_{m};\beta,\hat{\sigma}^{2}\right)$$

ولغرض توضيح الفكرة الواردة متن البحث تم الاستعانة بمثال توضيحي يدرس تاثير اربعة كميـــات لمركبات كيميائية داخلة فى صناعة السمنت علىعدد السعرات الحرارية المنبعثة من كل غرام لحظة صنعة والبيانات مقاسة كنسب من وزن المادة (clinkers) لحظة صنع السمنت وقد تعمـــدنا اســـقاط المشاهدات الخمسة الاخيرة والمشار اليهـــا لتوضح دقة الطريقة فى التقدير مع العلم ان :

Y : كمية الحرارة المنبعثة من كل غرام من السمنت مقاسة بالسعرات

X_1 : يمثل المركب ($3CaO.AL_2O_3$)

X_2 : يمثل المركب ($3\ CaO.SiO_2$)

X_3 : يمثل المركب ($4CaO.AL_2O_3\ Fe_2O_3$)

X_4 : يمثل المركب ($2CaO.SiO_2$)

والقياسات الخاصة بالمتغيرات التوضيحية (X_4 ,X_3 , X_2 , X_1) والمتغير المعتمد (Y) الواردة اعلاه مبينة في الجدول التالي .

N	$Y=X_0$	X_1	X_2	X_3	X_4
1	78.5	7	26	6	60
2	74.3	1	29	15	52
3	104.3	11	56	8	20
4	87.6	11	31	8	47
5	95.9	7	52	6	33
6	109.2	11	55	9	22
7	102.7	3	71	17	6
8	72.5	1	31	22	44
9	93.1*	2	54	18	22
10	115.9*	21	47	4	26
11	83.8*	1	40	23	34
12	113.3*	11	66	9	12
13	109.4*	10	68	8	12

وبعد تطبيق طرق التقدير التي تم التطرق لها وفقاً للبرنامج المعد لها تم الحصول على التقديرات المطلوبة وكما هي موضحة بالاتي : -

Ym1	Ym2	Ym3	Real-Y_m
90.60547	90.61092	90.61092	93.1
115.6587	115.6657	115.6657	115.9
79.78321	79.78801	79.78801	83.8
112.7481	112.7548	112.7548	113.3
112.458	112.4648	112.4648	109.4

حيث ان :

Y_{m1} : التقديربأستخدام طريقة المربعات الصغرى .

Y_{m2} : التقدير بأستخدام طريقة James –Stein .

Y_{m3} : التقدير بأستخدام طريقة (James – Stein) المعدل .

المصادر :

1. دراسة مقارنة لبعض طرق التقدير المتحيزة ، رسالة ماجستير ، احمد ثامر احمد ، كلية الادارة والأقتصاد ، الجامعة المستنصرية ، 1987.

2. تقدير معاملات تأثير التصميم في االمسوح الأحصائيـــة،رسالـــة ماجستير ، اسماعيـــل هادي جلوب ، كلية االأدارة والأقتصاد ، الجامعة المستنصرية ، 2002 .

3. Belesly , D.A , & Kuh , E . & Welsch , R . E .(1980) "Reagression Diagonstics" John wiley & Son ; New York .

4. Hanagal , D .D . (1997) . " Estimation of Reliability when stress is consored at strength", communications in statistics – theory and methods , 26(4) b, 911-919 .

5.Gibbons . D .G . (1981) , "Asimulation study of some Ridge Estimators". Journal of the American statistical (Association , 76, p . 131-139 .

6. Gunst , R. F . & Mason , R . L . (1977) , "Biased Estimation in Regression An Evaluation using Mean Squared Error " Journal of the American statistical Association ,72 , P. 616 –628 .

7. M . A . Ali , M . S . Abu – Salih (1988) " , on Estimation of Missing observation in Linear Regression Models , volume 50 Series B , Pt .3 , PP . 404 –411 .

.....

.....

تقدير العلاقة الفنية بين إنتاج الجوز و العوامل المؤثرة عليه في إقليم كوردستان العراق باستخدام تحليل الانحدار اللاخطي المتعدد و تحليل المسار

بروين محمد حمه خان*

ABSTRACT

The walnut fruit has an important role among agricultural product Kurdistan region has a ood level of advantages of walnut production in Iraq.

Because of wrong policies of the former government in Iraq, that destroyed about 4500 own and villages in this area, this product has witnessed a down trend.

The production characteristics and technical relation among variables were related to the alnut production has been studded, using regression and path analysis approaches.

The results show that the Sulaimani province comes at first in production quantity and ultivated area with this crop, and the cultivated area is a dominant variable of walnut production omparing with weak effect of capital variable on

this product. That refers to the back ward using of modern technologies in production rocesses.

The natural logarithm function was a best model was selected depending on statistical and conomic theory criteria.

Therefore, increasing and developing of production of this crop become today the necessary

المقدمة

تحليل الانحدار هو وسيلة إحصائية تستخدم لتحليل العلاقة بين متغير توضيحي واحد أو أكثر وبين متغير معتمد، و يعد تحليل الانحدار من أكثر الطرائق الإحصائية استخداما في مختلف العلوم لأنه يصف العلاقة بين المتغيرات وفقا لنموذج رياضي نستطيع من خلاله ان نقدر أو نتنبأ بقيمة المتغير المعتمد مقابل قيمة معطاة للمتغير (المتغيرات) التوضيحية في النموذج.

الاسم العلمي للجوز هو (Juglans Regia) حيث تعتبر منطقة حوض قزوين في إيران الموطن الأصلي للجوز وخاصة الصنف العجمي ،ولكن وجد بان موطنه يمتد من قوقاز الى ايران ومنه الى تركستان إضافة إلى كوردستان العراق ، ومن الدول المشتهرة بزراعته هي الولايات المتحدة الأمريكية ،إذ يقدر إنتاجها السنوي بأكثر من 130000 آلف طن ثم تليها إيطاليا وفرنسا والهند وإيران وتركيا

* مدرس مساعد / كلية الإدارة و الاقتصاد / جامعة السليمانية

ويوغسلافيا وسورية [1]إن زراعة الجوز في العراق متأخرة نوعا ما ،حيث تشير بعض الإحصائيات السابقة إلى ان عدد أشجار الجوز العجمي في القطر يبلغ حوالي 82 آلاف شجرة ،واحتلت محافظة السليمانية المرتبة الأولى من حيث عدد الأشجار وكان حوالي 36 آلف شجرة ثم تليها محافظة أربيل وبحدود 14 آلاف ومحافظة دهوك بحدود 8 آلاف شجرة و نينوى بحدود 3 آلاف شجرة أما معدل إنتاج الشجرة الواحدة في القطر بلغ 27 كغم [2] ألا ان هذه الأعداد وكذلك إنتاجية الشجرة قد تأثرت كثيرا بالسياسات القمعية الخاطئة من قبل النظام العراقي السابق ونتيجة لحرق وتدمير المناطق الرئيسية والمنتجة لهذا المحصول الحيوي ،مما أدى إلى انخفاض كبير في إعداد هذه الأشجار إلى أدني مستوياتها،مما اضطر العراق إلى استيراد كميات كبيرة من الجوز سنويا لسد حاجة الاستهلاك المحلي ،إذ يذكر ان الجوز له أهمية اقتصادية لا يستهان بها سواء لاستعمال ثمارها في الكثير من الصناعات الغذائية والمعجنات واستخراج الزيوت ،وكذلك استخدام أخشابه كإحدى أجود أنواع الأخشاب و الاستفادة من قشرته في صناعة بعض المواد التجميلية وصناعة الورق وغيرها .

لذا أوجب العمل على تطوير وتحسين البساتين القائمة وخاصة بساتين المناطق التي تتميز بميزة نسبية في إنتاجها من إنشاء بساتين حديثة للجوز وتطوي الموجود منها واستخدام مستويات أعلى من التكنولوجيا لسد حاجة البلد منها مع العلم بان مستلزمات الطبيعة لنجاح زراعة الجوز متوفرة في الكثير من المواقع في إقليم كوردستان العراق وحتى ان أحد الأصناف المعروفة بصنف هورامان وذلك نظرا لنجاحها وتأقلمها في هذه المنطقة [3] وانطلاقا من هذه المفاهيم ونظرا لأهمية هذا المحصول من الناحية الاقتصادية والغذائية التي تحتلها ،فقد تم دراسة العوامل المؤثرة على إنتاج الجوز وبشكل خاص في إقليم كوردستان من اجل حصر هذه العوامل وجعلها في شكل علاقة نظامية يسهل التعامل معها خاصة من الناحية الإحصائية وقد استخدم الباحث تحليل الانحدار اللاخطي المتعدد وكذلك تحليل المسار وحيث انهما أسلوبين ملائمين في دراسة هكذا مشاكل.

مشكلة البحث

تتمثل مشكلة البحث في:

1-عدم وجود دراسات واضحة تحدد مناطق تخصص إنتاج هذا المحصول سوءا على مستوى العراق او على مستوى إقليم كوردستان

2-انخفاض إنتاجية الأراضي المزروعة والشجرة الواحدة من .هذا المحصول خلال فترة العقود الماضية

3-ضعف استخدام التكنولوجيا المتطورة في إنتاج هذا المحصول والذي في رأينا يرجع إلى عدم معرفة العلاقة الفنية بين هذه التقنيات من السماد والمبيدات و الأصناف المقاومة للأمراض وما إلى ذلك والكميات المنتجة من هذا المحصول

فرضية البحث

تستند البحث إلى الفرضيات الآتية:--

1- تجانس في مستوى التكنولوجيا المستخدمة (المكننة،البذور،الأسمدة،المكافحة،طريقة الزراعة،الصنف المزروع.....الخ)من قبل مزارعي محصول الجوز في إقليم كوردستان العراق.

2- تجانس الظروف البيئية في هذه المنطقة المزروعة بهذا المحصول .

3-متغيري راس المال والمساحة المزروعة هما العاملين الأكثر أهمية من الناحية الإحصائية و لهما تأثير على إنتاج هذا المحصول مع بقاء العوامل الأخرى على حالها دون تغير بحيث يمكن إهمالها من الناحية الإحصائية .

هدف البحث

يهدف البحث الى:

1-تحديد المناطق التي تتميز بميزة نسبية في إنتاج محصول الجوز في إقليم كوردستان وتحديد خصائص إنتاجه.

1-تقدير العلاقة الفنية بين الإنتاج كمتغير تابع والمتغيرين التوضيحين (راس المال والمساحة المزروعة) باستخدام تحليل الانحدار اللاخطي المتعدد.

2-التحقق من صحة النتائج وتعزيزها باستخدام تحليل المسار.

مصادر البيانات

اعتمد البحث على بيانات المقطع العرضي (Cross Sectional Data) للموسم الزراعي (1997/1998) لمنظمة الأغذية والزراعة التابعة للأمم المتحدة (F.A.O) حول محصول الجوز في إقليم كوردستان لعراق بالإضافة إلى الأرقام والبيانات الأخرى المنشورة في الكتب والمجلات والمصادر المشار أليها في متن البحث .

خصائص إنتاج الجوز في إقليم كوردستان

بغية التعرف على بعض المؤشرات المتعلقة بإنتاج الجوز في إقليم كوردستان ،إذ تم حساب عدد من المؤشرات كما موضح في الجدول الآتي:

جدول (1)

الأهمية النسبية للمساحة وعدد الأشجار و الإنتاج لمحافظات الإقليم الثلاث

المحافظة	الأهمية النسبية%	الأهمية النسبية لعدد الأشجار %			الأهمية النسبية للإنتاج الكلي %
		المثمرة	غير المثمرة	الكلي	
اربيل	44.27	39.420	61.55	55.57	36.70
دهوك	4.36	9.47	2.00	4.00	8.00
السليمانية	51.37	51.11	36.45	40.43	55.30

- احتسب بالاعتماد على جدول في الملحق الأول .

إذ نلاحظ في الجدول أعلاه بان محافظة السليمانية تحتل المركز الأول من حيث المساحة المزروعة بمساهمتها بحوالي 51 % من إجمالي المساحات المزروعة وفيما يتعلق بعدد الأشجار المثمرة نلاحظ أيضا تفوق محافظة السليمانية من حيث عدد هذه الأشجار بحوالي 51 % . في حين نلاحظ ان محافظة اربيل تأتي في المرتبة الأولى من حيث عدد الأشجار غير المثمرة و الأشجار الكلية.بحوالي 61 % و 56 % على التوالي.مما يدل على ان لهذه المحافظة مستقبلا جيدا في مساهمة إنتاج هذا المحصول ولكن يجب ان نتذكر ان ليست كل الأشجار المزروعة قابلة للإثمار ماعدا التي تقع ضمن المناطق الباردة أو معتدلة الجو لأنها قد تنجح زراعتها ولكن لا تشترط ان تزهر هذه الأشجار و تؤتي أكلها [4] ، وكذلك من حيث المساهمة في الإنتاج ساهمت السليمانية بحوالي 55 % من الإنتاج الكلي ،والجدير بالذكر هنا ان هذه المحافظة حققت أدنى كلفة الإنتاج للشجرة الواحدة والبالغ حوالي (5.14) دينار [5]

مما يدل على ان هذه المحافظة لها ميزة نسبية لانتاج هذا المحصول وتحقيق وفورات السعة له ،مما تطلب التركيز من قبل دوائر الزراعة والهيئات الحكومية في هذه المحافظة بغية الحصول على درجة ملائمة من الاكتفاء الذاتي والتي تتلاءم مع الموارد البشرية والطبيعية والاجتماعية المتاحة من هذا المحصول ولاسيما في المناطق المعروفة بالخبرة و الإنتاج مثل هورامان والتي سميت أحد الأصناف

باسم هذه المنطقة .مع عدم إهمال المناطق الأخرى .وبهذا نستطيع ان نضيف سلعة إضافية و منفعة جديدة إلى سلة المستهلك المحلي و تحقيق دخل إضافي للمزارع المحلي ،و كذلك على مستوى الاقتصاد الكلي نستطيع من خلال هذه الأجزاء على تحقيق معدلات نمو ملائمة في جميع القطاعات الاقتصادية .

ومن جانب آخر نلاحظ انخفاض إنتاجية الدونم لهذا المحصول والتي لم يتجاوز 460 كغم/دونم في احسن أحوالها.مما يدل على ان مزارعي هذا المحصول ما يزالون في وضع غير لائق و الإمكانيات المتاحة ،وذلك بسبب ضعف المهارة الفنية والخبرة الكافية لانتاج هذا المحصول هذا من ناحية واعتمادهم على تكنولوجيا غير ملائمة ومتخلفة من ناحية أخرى مثل استخدام السماد، خدمة الأشجار ومكافحة الأمراض ، خدمة وصيانة التربة ، التعشيب و غيرها من العمليات الزراعية وضعف الجهاز الإرشادي لإيصال الخبرات والمهارات والمعلومات الحديثة إلى المزارعين لتحسين أدائهم لانتاج هذا المحصول بكفاءة اقتصادية وفنية مقبولة .

الجانب النظري

إن العديد من الظواهر الطبيعية تحتوي على متغيرات مرتبطة فيما بينها بعبارة أخرى اعتماد متغير على عدد من المتغيرات السببية والمؤثرة عليه حيث ان هذه العلاقة غير مسيطر عليها بواسطة قانون طبيعي أو صيغة رياضية محددة ودقيقة ولكن عندما نمتلك معلومات عن اثنين أو اكثر من المتغيرات المتلازمة أو المرتبطة فانه من الطبيعي ان نبحث عن الطريقة التي يمكننا التعبير عن هذه العلاقة الدالية الموجودة بالإضافة إلى ذلك فانه من المرغوب معرفة قوة هذه العلاقة أي بمعنى آخر ليس فقط البحث عن دالة رياضية تفيدنا في معرفة كيفية و مدى دقة التنبؤ بالاعتماد على أحد المتغيرات وخصوصا إذا عرفنا قيم التغيرات الأخرى ذات العلاقة.

ان الأساليب المختلفة المستخدمة لتحقيق هذين الهدفين تعرف بطرق الانحدار (Regression Methods) و طرق الارتباط (Correlation Methods) ، إذ يعتبر الأسلوب الأول اكثرا اعتمادا لتحديد افضل علاقة دالية بين المتغيرات بينما طرق الارتباط تستخدم لقياس درجة التوافق بين متغيرات مختلفة في هذه الحالة يمكن القول بان العلاقات الدالية بصورة عامة يمكن التعبير عنها رياضيا كما يلي

$$Y=f(X1,X2,X3,.....Xk)$$

حيث ان y هو رمز إلى متغير الاستجابة (Response Variable) أو المتغير المعتمد (Dependent Variable)

X 1,X2, X3...Xk تمثل المتغيرات المؤثرة أو المتغيرات التوضيحية (Independent Variables)وتجدر الإشارة هنا ان كلمة الانحدار استخدمت لاول مرة بصيغتها الحاضرة من قبل العالم الوراثي البريطاني كالتون (Francis galton) حيث كان واحدا من أوائل الباحثين الذين تعاملوا مع

موضوع دراسة أو وصف متغير واحد بالاعتماد على واحد أو اكثر من المتغيرات عمل كالتون تركز على دراسة أطوال الأبناء مقارنة بأطوال آبائهم .فلاحظ وجود علاقة واضحة وسببية وهي ميل أطوال الأبناء نحو المتوسط لاطوال آبائهم فالآباء قصار القامة يميلون لإنجاب أبناء معدل أطوالهم أعلى (أطول من آبائهم)بينما العكس صحيح في حالة الآباء طوال القامة بشكل غير اعتيادي لذلك فان العالم كالتون ذكر بان أطوال الأبناء ترتد أو تنحدر (Regress) نحو المتوسط ولذلك ظهرت كلمة الانحدار (Regression) .[6]

ولذلك يمكن ان ينظر إلى تحليل الانحدار بشكل رئيسي كتحليل للعلاقة بين المتغيرات و يعتبر احد الأدوات الإحصائية الأكثر استعمالا لانه يعطينا طريقة سهلة لتحديد العلاقة بين المتغيرات ، هذه العلاقة يمكن التعبير عنها بشكل معادلة تحتوي على متغير الاستجابة y مع واحد أو اكثر من المتغيرات التوضيحية X 1,X2, X3...Xk .

وان تحليل الانحدار تخدم أربعة أغراض رئيسية وهي[7]

1-وصف البيانات data description

2-تقدير المعلمات parameter estimation

3- التنبؤ prediction

4-السيطرة control

أنواع الانحدار

يمكن تصنيف الانحدار لغرض تبسيط دراسته إلى:

1- الانحدار الخطي

أ- الانحدار الخطي البسيط (simple linear regression)ويضم متغيرا توضيحيا واحدا

ومعادلة الانحدار البسيط هي .....(1) Yi = B0 + B1Xi1 +ei

ب- الانحدار الخطي المتعدد multiple linear regression

ويضم عدة متغيرات توضيحية

ومعادلة الانحدار الخطي المتعدد بصورة عامة هي (2) ..... Yi = B0 BkXik + ei + B1Xi1 + B2Xi2 +

2- انحدار غير خطي Non-linear regression

ويقسم إلى قسمين

أ- الانحدار غير خطي البسيط Simple Curvilinear

ويضم متغيرا توضيحيا واحدا

ومعادلته بصورة عامة هي $Yi = B0 + B1Xi1 + B2Xi1^2 + \ldots\ldots + ei \ldots\ldots\ldots\ldots\ldots\ldots (3)$

ولكن هناك معادلات أخرى غير خطية قد تأخذ الشكل التالي: .(4) ..... $Yi = B0e^{B1^{Xi1}} Ui$

وهذه يمكن تحويلها إلى معادلة خطية باستخدام التحويل اللوغاريتمي

ب- الانحدار غير الخطي المتعدد Multiple Curvilinear

ويضم عدة متغيرات توضيحية مثل

$Yi = B00 + B10 + B01Xi2 + B11Xi1Xi2 + B20Xi1^2 + B02Xi2^2 + \ldots + ei \ldots\ldots (5)$

وهناك معادلات أخرى غير الخطية مثل $Y = B0X1^{B1}X2^{B2}e$

يمكن تحويلها إلى معادلات خطية باستخدام التحويل اللوغاريتمي[8]

و من خلال دراسة الصيغ المختلفة لتحديد العلاقة المناسبة للمتغيرات المؤثرة على إنتاج محصول الجوز ،فقد تم اعتماد الصيغة الاتية و هي خاضعة للاختبارات $B0 + B3\ Xi2^2 + ei$

$e^{Y_i} = Xi1^{B1}Xi2^{B2}e \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$

حيث أن Yi يمثل قيم المتغير التابع أو مقدار الاستجابة.

Xi1 ، Xi2 يمثل قيم المتغيرات التوضيحية .

B0، B1،B2 تمثل ثوابت معلمات معادلة الانحدار.

ei الخطأ العشوائي أو المتبقي .(ويتوزع توزيعا طبيعيا بوسط حسابي قدره صفر أي $E(e)=0$ و تباين قدره $\sigma^2 e$.

واستخدم طريقة المربعات الصغرى في تقدير معلمات النموذج لانها تجعل مجموع مربعات الخطأ اقل ما يمكن .

فروض تحليل الانحدار Analysis Assumptions of Regression

أن تحليل الانحدار يعتمد على فروض عديدة أهمها هي :--

1- أن المتغير المعتمد y هو متغير عشوائي وقيمه مستقلة إحصائيا الواحدة عن الأخرى وتتوزع توزيعا طبيعيا بوسط حسابي قدره μ وتباين σ^2.

2- ان حد الخطأ ei هو خطأ عشوائي يتوزع توزيعا طبيعيا بوسط حسابي صفر وتباين قدره $\acute{o}^2$ ان التباين المشترك بين $cov\ (ei,ej)=0$ حيث ان $j \neq i$.

3- عدم وجود علاقة خطية محددة أو تامة بين المتغيرات التوضيحية[9]

ثانيا: تحليل المسار Path Analysis

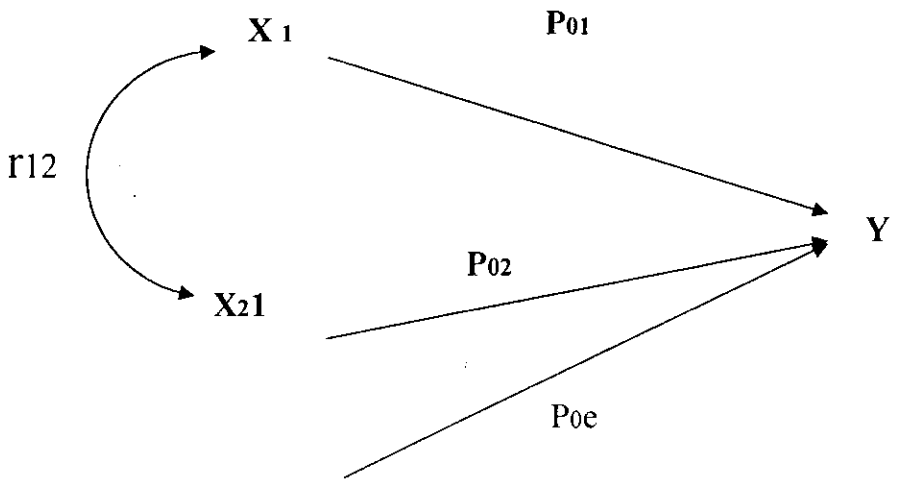
ان تحليل المسار كأسلوب إحصائي يعتمد على وجود نماذج سببية (Causal Models) فهو يفترض وجود علاقة سببية بين السبب (Cause) والأثر (Effect) وان إحدى فوائد تحليل المسار هي تجزئة معامل الارتباط R بين متغيرين إلى مكوناته التالية:

ا-التأثير المباشر Direct Effect للسبب على الأثر.

ب-التأثير غير المباشر Indirect Effect للسبب على الأثر من خلال مسالك Paths عبر مسببات أخرى [10] .

الشكل (1)

مخطط تحليل المسار



- المصدر : د.خاشع محمود الراوي(1987)،المدخل الى تحليل الانحدار ،المكتبة الوطنية،بغداد،ص

النتائج و التحليل

تعتمد نتائج وتحليل البحث كما ذكرنا على أسلوبين إحصائيين هما :

أولا: تقدير نموذج الانحدار و تحليل نتائجه

أ – تقدير النموذج الانحدار اللاخطي المتعدد

بعد إدخال البيانات ومحاولة تقدير العلاقة الفنية بين إنتاج محصول الجوز كمتغير تابع وكل من راس المال والمساحة المزروعة كمتغيرين توضيحين مع ثبات العوامل الأخرى على حالها كالعمل والعوامل البيئية وما إلى ذلك أي :

$$Yi = f(X1, X2 | X3, X4, \ldots Xk)$$

وباستخدام برنامج الإحصائي (Statigraph) واختبار نماذج عدة وبالاستناد إلى الاختبارات الإحصائية تم اختيار افضل نموذج لتمثيل هذه العلاقة وكانت النتائج كما موضح في جدول (ا)

جدول (1)

المعلمات القدرة لدالة إنتاج الجوز (y) كمتغير تابع *

المتغيرات التوضيحية	المعلمات القدرة Estimated Coefficients
Xo	– 80471.9 (-2.84664)
اللوغاريتم الطبيعي لراس المال (ln X1)	16521.9 (2.18908)
اللوغاريتم الطبيعي للمساحة المزروعة (ln X2)	22542 (4.28021)
مربع المساحة المزروعة ($X2^2$)	0.0970187 (5.15192)
الاختبارات الإحصائية	
% 89.5985	R^2
% 87.5182	R^{-2}
2.07257	D.W
43.07	F

- المصدر : تم تقديره بالاعتماد على بيانات (F.A.O)
- تشير الارقام داخل الأقواس إلى قيمة اختبار (t)
- للتفصيل انظر الملحق الثاني

إذ نلاحظ من الجدول أعلاه بان :

1- ان جميع المعلمات الفردية تتمتع بدرجة عالية من المعنوية بناء على اختبار (t)

2- ان النموذج ذو معنوية عالية أيضا اعتمادا على اختبار F والبالغ (43.07)

3- ان قيمة R^2 بلغت (89.5985 %) مما يدل على ان حوالي 89 % من التقلبات الكلية التي تحدث في المتغير التابع تكون مسؤولة عنها المتغيرين التوضيحيين(Explanatory Variables) الداخلة في النموذج والباقي تعزى إلى عوامل أخرى لم يتضمنها النموذج أما لعدم توفر البيانات عنها أو عدم ثبوت أهميتها الإحصائية أو صعوبة تقديرها كالعمل والظروف البيئية و هذا يعني ان هذه الدراسة قد أسهمت في تحديد أهم العوامل المؤثرة على الإنتاج بنسبة 89 % [11] و ان قيمة اختبار(Durban - watson) البالغة 2.07257 **تدل** على عدم وجود مشكلة الارتباط التسلسلي (Autocorrelation) بصورة خطيرة على مستوى 1 % عند مقارنتها بالقيمة الجدولية ، حيث ظهر ان قيمة(d) تقع في منطقة قبـــول عدم وجـــــــود المشكلـــة و ان (du <D.W < 4-du) وهذا يدل على عدم وجود المشكلة[12]

ونظرا لاعتماد التقدير على بيانات المقطع العرضي (Cross-section Data) فانه يتوقع وجود مشكلة عدم ثبوت تجانس التباين (Heteroscedasticity) فقد تم استخدام اختبار (park-Glejser Test) للكشف عن هذه المشكلة [13] :-

1-اختبار المتغير العشوائي مع عنصر راس المال ((X1) (للتفصيل انظر الملحق الثاني)

$$Lnei^2 = -2.09076 + 0.321396LnX1$$

(t) = -2.03833 0.708943

F=0.50

2-اختبار المتغير العشوائي مع المساحة المزروعة (X2)

$$Lnei^2 = -2.10702 + 0.165922LnX2$$

(t) = -1.91507 0.666861

F= 0.44

نستنتج من الاختبارات السابقة (اختبار F , t) وان النموذج المقدر لا يعاني من مشكلة عدم ثبات تجانس الخطا،وهذا يمكن توقعه في الصيغة اللوغاريتمية إذ أنها تقلل من شدة هذه المشكلة[14].

وفيما يتعلق بمشكلة الارتباط الخطي فقد تم استخدام اختبار فراير وكلوبر(Farrar-Glauber) إذ يستند هذا الاختبار على احصاءة ($\chi 2$) حيث ان صيغة الاختبار $\chi 2 = -[n-1-1/6(2k+5)].\ln| D |$ | في حين ان n يمثل حجم العينة و k يمثل عدد المتغيرات التوضيحية و D مصفوفة معاملات الارتباط ، وتبين بان القيمة $\chi 2$ المحسوبة هي (9.30459466 و كانت اصغر من القيمة الجدولية

و البالغة (11.3449) و بمستوى معنوي 0.01 و بـــــــدرجة حريـــــــــة (d. f=3) و التي تشير إلى عدم وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية [15] .

يلاحظ ان الحد الثابت سالب وهذا لا يؤثر على صلاحية المعادلة ،حيث ان الحد الثابت يتم تفسيره إذا كان موجبا أو سالبا كبيرا او صغيرا انه قيمة رياضية ليس لها مدلول اقتصادي في معظم الحالات [16] .

ب- تحليل النتائج

إذ نلاحظ من هذا النموذج ان المتغيرين راس المال و المساحة يؤثران بشكل واضح على إنتاج هذا المحصول ، و هذا يدل على انه مازال هناك إمكانية التوسع في إنتاج هذا المحصول سواء عن طريق التوسع الأفقي و الذي يعني زراعة مساحات أخرى و زيادة حجم المزارع القائمة ، و لكن هذا الأسلوب لا يعول عليه كثيرا لما يتميز به عنصر الأرض من المحددات و المشاكل منها محدودية الأراضي الصالحة للزراعة و ما ترتب عليها من استصلاحها و توفير المستلزمات الأساسية لها [17] .

أما الأسلوب الأخر للتوسع في الإنتاج هو التوسع العمودي و الذي يعني تكثيف راس المال و التكنولوجيا الحديث و المتقدمة في زراعة هذا المحصول ، و هذا في رأينا ملائم مع إنتاج محصول الجوز ، و خصوصا إذا علمنا ان اغلب مناطق نجاح زراعة الجوز تكون إما جبلية او صخرية مما يصعب اعتماد الأسلوب الأول ، لذا فالمطلوب الاعتماد على هذا الأسلوب و توفير المستلزمات الإنتاجية الحديثة من السماد و المبيدات و محاولة مكننة إنتاج هذا المحصول و استخدام أصناف مبكرة الإثمار و المقاومة للأمراض و الظروف الطبيعية و محاولة تطبيق الخبرات و المهارات و المعارف و النتائج العلمية الحديثة و إيصالها إلى مزارعي هذا المحصول عن طريق جهاز إرشادي كفوء ، و كذلك العمل على تحسين الظروف المعيشية لسكان الريف و تطوير أحوالهم الثقافية و الاجتماعية ، مما ينعكس إيجابا من حيث الكمية و النوعية على إنتاج هذا المحصول في نهاية المطاف .

ثانيا: نتائج تحليل المسار

أ – تقدير علاقات المسار

ان تحليل المسار كما نوهنا عنه سابقا يعتمد على دراسة العلاقات السببية بين المتغيرات ، حيث ان:

$$\begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} P_{o1} \\ P_{o2} \end{pmatrix} = \begin{pmatrix} r_{10} \\ r_{20} \end{pmatrix}$$

اذ ان:

r12 يشير إلى معامل الارتباط بين X1 , X2 .

r10 يشير إلى معامل الارتباط بين X1 , Y.

P01 يشير إلى معامل المسار (Path Coefficient) أو قيمة المسار من X1 إلى Y.حيث (0) يشير إلى المتغير Y و(1) يشير الى المتغير X1 .

P02 يشير إلى معامل المسار (Path Coefficient) او قيمة المسار من X2 الى Y.حيث (0) يشير الى المتغير Y و(2) يشير الى المتغير X2 .

و من مصفوفة الارتباطات بين المتغيرات تبين ان rX1X2 = -0.0352 , rX1Y = 0.0429, rX2Y = 0.9232(كما هو موضح في الملحق الثاني)

وان RP=r

حيث ان R مصفوفة معاملات الارتباط بين المتغيرات التوضيحية (السبب)

P معاملات المسار .

r =معاملات الارتباط بين كل من المتغيرات التوضيحية (السبب) والمتغير التابع (الاثر)

أي ان:

$$\begin{pmatrix} 1 & -0.0352 \\ -0.0352 & 1 \end{pmatrix} \begin{pmatrix} P_{o1} \\ P_{o2} \end{pmatrix} = \begin{pmatrix} 0.0429 \\ 0.9232 \end{pmatrix}$$

$$\begin{pmatrix} P_{o1} \\ P_{o2} \end{pmatrix} = \begin{pmatrix} 0.07549974 \\ 0.9232 \\ 0.925857253 \end{pmatrix}$$

و على هذا الأساس بإمكاننا إيجاد درجة التحديد للمتغير Y من قبل المسببين X1 , X2 ويرمز له ب $Ro^2(12)$

$$Ro^2(12) = P01^2 + P02^2 + 2\ P01\ r12\ P02$$

$$= (0.07549974)^2 + (0.925857253)^2 + 2$$

$$(0.07549974)(0.925857253)(-0.0352)$$

$$0.00570021+0.857211652-0.004921099$$

$$= 0.857990763$$

$$Ro(e)^2 = 1 - Ro(12)^2 = P0e^2 = 0.142009237$$

$$P0e = 0.3768411$$

أي ان X1 يحدد $P01^2 = 0.00570021$ من تباين Y وان X2 يحدد $P02^2 = 0.857211652$

من تباين Y وان التحديد المشترك ل X1، X2هو $2\ P01\ r12\ P02 = 0.004921099$ –

وعليه فان مجموع التحديد ل Y من قبل X1،X2= =0.857990763

آما درجة التحديد ل Y من قبل e فهو 0.142009237 (أي ان 0.142009237 من تباين Y يحددها الباقي e والتي تشمل عوامل أخرى سببية لم تدخل في هذه الدراسة)

وبناء على هذه التقديرات بإمكاننا تحديد تأثيرات المباشرة وغير المباشرة والمشتركة للمتغيرين راس المال والمساحة على إنتاج الجوز كما موضح في الجدول (2).

جدول (2)

تحليل المسار المباشر وغير المباشر X1 (راس المال) ،X2 (المساحة المزروعة)على Y(إنتاج الجوز)

نوع التأثير	قيمة المعامل
1- تأثير X1 على Y أ- التأثير المباشر =P01 ب- التأثير غير المباشر عن طريق X2 (r12P02)	0.07549974 0.032590175-
مجموع التأثير الكلي =r10	0.04290799
2- تأثير X2 على Y أ-التأثير المباشر =P02 ب-التأثير غير المباشر عن طريق X1 (r12P01)	0.925857253 0.00265759-
مجموع التأثير الكلي =r20	0.923199662

-المصدر :احتسب بالاعتماد على بيانات ممثلية (F..A.O) في شمال العراق ،مكتب تنسيق (F.A.O) شمال العراق (1998) ،نشرة إحصائية، وحدة الإحصاء الزراعي ،العراق -اربيل.

ب-تحليل النتائج

من خلال تحليل المسار نلاحظ الآتي:

1- عند تغير راس المال (X1) بمقدار انحراف قياسي واحد فان ذلك سيؤدي الى تغير مباشر في معدل إنتاج الجوز (كغم) بمقدار (0.0754) =P01 والى تغير مباشر عبر المتغير (X2) المساحة بمقدار r12P02 - 0.03259 و عليه فالتأثير الكلي لراس المال (X1) = التأثير المباشر + التأثير غير المباشر =0.042907

والقيم هي نفسها لمعامل الارتباط بين X1 ،Y اي ان 0.042907 = r10

2- عند تغير المساحة X2بوحدة قياسية واحدة فان ذلك سيؤدي إلى تغير مباشر في معدل إنتاج الجوز بمقدار P02=0.9258572 والى تغير غير مباشر عبر المتغير X1 بمقدار r12P01 = 0.00265759- و كما نرى بان التأثير الكلي للمساحة X2 =التاثير المباشر + التاثير غير المباشر= 0.923199662 والقيمة هي نفسها لمعامل الارتباط بين X2 على Y أي ان r20 .

3- من خلال هذا التحليل ظهرت هيمنة متغير المساحة لتأثيرها على الإنتاج وبمساهمة كبيرة بالمقارنة مع متغير راس المال الذي تكون نسبة مساهمته ضئيلة ومنخفضة مما يدل على اعتماد مزارعي بساتين الجوز على أساليب تكنولوجية متخلفة وتقليدية في جميع عمليات إنتاج هذا المحصول من الأصناف المزروعة وخدمة الأشجار والمكافحة والري وغيرها ...الخ والتي هي سمة مميزة للقطاع الزراعي بشكل عام في العراق وفي إقليم كوردستان على وجه الخصوص ،واكتفاء هؤلاء المزارعين بتبني استراتيجية التوسع الأفقي لزيادة الإنتاج بزراعة مساحات اكبر بدلا من اعتمادهم على الأساليب التكنولوجية الحديثة. و هذه النتيجة متطابقة ومتجانسة مع النتائج التي تم التوصل إليها بواسطة تحليل الانحدار .

إذ نرى ان استجابة هذا المحصول للمتغيرين بد أ واضحا ونستطيع التوسع في إنتاج هذا المحصول توسعا عموديا وذلك بتكثيف راس المال واستخدام التكنولوجيا المتقدمة من استخدام الأصناف المحسنة واستثمار العلوم الزراعية في هذا المجال وذلك نرى ان راس المال اقل تأثيرا من متغير المساحة وهذا ما يفسر نمط الإنتاج المتخلف والذي يعكس الأساليب الإنتاجية المتخلفة والمعتمدة في القطاع الزراعي بشكل عام وهذا المحصول بشكل خاص لانتاج هذا المحصول [18].

وفيما يتعلق بمتغير المساحة نرى استجابة إنتاج استجابة كبيرة بالمقارنة مع نظيرتها راس المال .مما يعني ان هناك إمكانية كبيرة لزيادة إنتاج هذا المحصول وخصوصا ان إقليم كوردستان العراق يتمتع بمساحات مروية كبيرة وغير مستعملة إذ بلغ مجموع المساحات المروية حوالي (400) ألف هكتار في هذه المنطقة [19] والتي تكون نسبة لا باس بها من هذه المساحات واقعة داخل خط نجاح زراعة الجوز .

الاستنتاجات

من خلال استعراض هذه الدراسة تم التوصل إلى الاستنتاجات الآتية:

1- هناك شحة كبيرة في البيانات و المعلومات المتعلقة بالقطاع الزراعي بشكل عام و إنتاج الجوز بشكل الخاص ، و ذلك بسبب عدم وجود هيئات و دوائر إحصائية كفوءة و مخولة لجمع و إعداد مثل هذه البيانات و التي يكون هذا الإقليم أحوج ما يكون لمثل هذه البيانات ، لانه لا يتوفر فيه قاعدة بيانات زراعية تفصيلية و موثوقة ، بدرجة ان بعض البيانات التي يتم الاعتماد عليها في بض الاحيان قد تعود إلى اكثر من ثلاثة عقود ، ما يجعل الاعتماد عليها أمرا غير منطقي و غير علمي لان اغلب البيانات ان لم تكن جميعها قابلة للتغير و خاصة في فترة الأجل الطويل .

2- ان إنتاج القطر من الجوز باعتباره محصولا ستراتيجيا لازال منخفضا ويشكل نسبة ضئيلة من الاكتفاء الذاتي من هذا المحصول.

3-أظهرت نتائج بان محافظة السليمانية تحتل المركز الأول من المساحة المزروعة و عدد الأشجار المثمرة بحوالي 51 % في كلتا الحالتين ، في حين تأتي محافظة أربيل في المقدمة من حيث عدد الأشجار الكلية بنسبة 56 % و كذلك حققت السليمانية أدنى التكلفة للشجرة الواحدة و البالغ 5.14 دينارا .

4- أظهرت النتائج التقديرات ان افضل صيغة لتمثيل العلاقة بين الإنتاج كمتغير تابع و كل من راس المال والمساحة المزروعة كمتغيرين توضيحين هي صيغة اللوغاريتم الطبيعي و ذلك بناءا على المعايير و الاختبارات الإحصائية ، و ظهرت أيضا ان النموذج المقدر لا يعاني من المشاكل القياسية ويتمتع بدرجة عالية من المعنوية الإحصائية ، و كذلك و جود علاقة طردية بين المتغيرين التوضيحين و المتغير المعتمد.

5- من خلال تحليل الانحدار و التحليل المسار نلاحظ أهمية متغير المساحة بتأثيرها على الإنتاج والتأثير المتواضع لعنصر راس المال، و هذا ما يفسر نمط الإنتاج المتخلف السائد سوءا بالنسبة للقطاع الزراعي بشكل عام و إنتاج هذا المحصول بشكل خاص و اعتماد وسائل تكنولوجية بدائية و عدم الاستفادة من التقنيات الحديثة و التي بدوره انعكست سلبا على إنتاج و إنتاجية هذا المحصول.

6- هناك إمكانية التوسع في إنتاج هذا المحصول سوءا كان التوسع الأفقي او العمودي و يفضل التوسع الثاني لملاءمته مع طبيعة إنتاج هذا المحصول بغية الوصول إلى مستوى من الاكتفاء الذاتي اللائق من محصول الجوز.

التوصيات

من خلال هذه الدراسة و انطلاقا من الاستنتاجات سابقة الذكر نوصي بالآتي:

1-العمل على إجراء تعدادا زراعيا شامل و متكامل من كل النواحي و حسب التوصيات الدولية ؟، و توفير كافة المستلزمات و المتطلبات الضرورية لها من كوادر الفنية الناجحة و الموارد المالية و التقنية الحديثة لمثل هذه العملية .

2- ينبغي بحث السبل الكفيلة لرفع إنتاج و إنتاجية هذا المحصول إلى مستوى من الاكتفاء الذاتي اللائق ،وذلك من خلال التوسع العمودي او التوسع الأفقي أو معا ،حيث يتم التوسع العمودي عن طريق إدخال التحسينات التكنولوجية الحديثة ،و بالأخص من خلال زراعة الأصناف المحسنة التي تم الحصول عليها من مراكز البحوث العلمية في القطر ،اما التوسع الأفقي باستصلاح وزراعة مساحات إضافية جديدة او من خلال التوسعين معا،والعمل على خفض كلفة وحدة الناتج.

3- التأكيد على الجانب الإرشادي لمزارعي الجوز من خلال فتح المراكز التعليمية والتثقيفية في أنحاء الإقليم ،بغية رفع الجانب العملي والثقافي لمزارعي الإقليم.

4- القيام بأجراء المزيد من مثل هذه الدراسات و لمختلف المحاصيل من اجل تحقيق الوفورات الاقتصادية المناسبة في زراعة هذا المحصول .

هوامش البحث

1-د.يوسف حنا يوسف(بدون سنة النشر)،إنتاجية الفاكهة النفضية،ص357

2- اذ نشك في صحة هذه البيانات ونرى إنها بعيدة عن الواقع لان عدد الأشجار سواء في العراق أو إقليم كوردستان تفوق كثيرا هذه الأعداد بحيث لايترك مجال للشك بان هذه الارقام خاطئة وغير مقصودة ومنشورة بدون تركيز ،اما مقصودة وكانت عبارة عن امتداد لسيلسة تهميش هذه المنطقة وتجريدها من كل امكانات المادية والبشرية والتي تعرضت لها العراق عموما واقليم كوردستان خصوصا ،وقد اوردنا هذه الارقام لغرض التصحيح والمقارنة فقط للاسهاب انظر:

- د.يوسف حنا يوسف،المصدر السابق،ص358

3-عبد الهادي اسماعيل غني(1978)،زراعة الفاكهة في العراق،اعدادية الزراعة 1 حزيران

4-هناك مشكلة التي تكتنف المساحات المزروعة باشجار الجوز ،حيث ان اشجار الجوز ليست لها مساحات مخصصة في الغالب ان لم يكن في كل المناطق وانما يزرع مع اشجار اخرى مثل الرمان والتفاح وغيرها لذا فان من الصعب من الناحية العملية عزل هذه المساحات.

5-للمزيد راجع الملحق الأول

6- Gujarati, D., (1985), Basic Econometrics, McGraw-Hill Book Company, 5th printing, PP: 11-12.

7-د.خاشع محمود الراوي(1987)،المدخل الى تحليل الانحدار ،المكتبةالوطنية،بغداد،ص66

8- Johnston, J (1984), Econometric Methods, 3rd Edition, McGraw-Hill -Book Co, and PP: 12-74.

9- للمزيد انظر :

- N.R.DRAPER .& Smith, H(1981) ,Applied Regression Analysis (2nd .ed)New York .John Wiley&Sons,PP:22-23

10-د.خاشع محمودالراوي،مصدر سابق،ص523

11-ا.د.اموري هادى كاظم،باسم شليبه مسلم (2002)،القياس الاقتصادي المتقدم،بغداد ،مطبعة الطيف،ص69 .

12 -William. E. Griffits & others (1992), Learning and Practicing Econometrics John Wiley & Sons, PP: 530-534.

13- Pindyck, R. S. & Rubinfeld, D.L (1985), Econometrics Model and Economic Forecasts, 2nd Edition, McGraw-Hill, Inc.4th printing, PP: 150-152.

14 -Gujarati, D (1985), Basic Econometrics, McGraw-Hill Book Company, 5th printing. PP: 210.

15 - Koutsoyiannis, A. (1981), Theory of Econometrics, 2nd Edition, Macmillan Inc, and PP: 244.

16- عادل عبد الغني المحبوب (1992) ، تفسيرات تحليلية لمفهوم الحد الثابت في نماذج الاقتصاد القياسي،مجلة تنمية الرافدين،مجلد 14،العدد 35،سنة ، ص 109.

17- عبالوهاب المطر الداهري (1987) ، الاقتصاد الزراعي ، جامعة بغداد الطبعة الثانية ، صص: 159- 162.

18- رسلي جميل القدو (1997)، الانتاجية والحجم الامثل لمزرعة محصول الشلب في النجف، مجلة العلوم الزراعية العراقية، المجلد 28، العدد الاول ، ص ص 107-116.

19- الجهاز المركزي للاحصاء (1988)، المجموعة الاحصائية السنوية ، جمهورية العراق ، وزارة التخطيط ، ص: 98 .

المصادر

المصادر العربية

1- ا.د.اموري هادى كاظم،باسم شليبه مسلم (2002)،القياس الاقتصادي المتقدم،بغداد ،مطبعة الطيف .

2- الجهاز المركزي للاحصاء (1988)، المجموعة الاحصائية السنوية ، جمهورية العراق ، وزارة التخطيط .

3- د.خاشع محمود الراوي(1987)،المدخل الى تحليل الانحدار ،المكتبةالوطنية،بغداد .

4- عادل عبد الغني المحبوب.تفسيرات تحليلية لمفهوم الحد الثابت في نماذج الاقتصاد القياسي،مجلة تنمية الرافدين،مجلد 14،العدد 35،سنة 1992،

5- عبالوهاب المطر الداهري (19878) ، الاقتصاد الزراعي ، جامعة بغداد الطبعة الثانية .

6- ممثلية (F. . A.O) في شمال العراق (1998) ،مكتب تنسيق (F.A.O) شمال العراق ،نشرة احصائية،وحدة الاحصاء الزراعي ،العراق –اربيل.

7- د.يوسف حنا يوسف(بدون سنة النشر)،انتاجية الفاكهة النفضية .

المصادر الاجنبية

1- Gujarati. D.. (1985). Basic Econometrics. McGraw-Hill Book Company. 5th printing.

2- Johnston, J (1984), Econometric Methods, 3rd Edition, McGraw-Hill-Book Co.

3- Koutsoyiannis, A. (1981), Theory of Econometrics, 2nd Edition, Macmillan In.

4- N.R.DRAPER. & Smith. H.(1981) .Applied Regression Analysis (2nd .ed)New York. John Wiley & Sons.

5- William. E. Griffits & others (1992), Learning and Practicing Econometrics John Wiley & Sons.

الملاحق

- الملحق الاول: جدول المساحة المزروعة ، عدد الاشجار ، غلة الدونم و كلفة الشجرة الواحدة.
- الملحق الثاني : نتائج تقدير الارتباط ونماذج الانحدار المقدرة و تقديرات (Park- Gljser Test) باستخدام برنامج (StatiGraph)

الملحق الاول

المساحة المزروعة وعدد الأشجار وإنتاج وإنتاجية الدونم وتكلفة الشجرة الواحدة لمحافظات الإقليم الثلاث

المحافظة	المساحة المزروعة	عدد الأشجار			معدل عدد الاشجار في الدونم	انتاج الشجرة المثمرة الواحدة	الانتاج الكلي*	غلة الدونم(كغم)	تكلفة الشجرة الواحدة (دينار)
		المثمرة	غير المثمرة	الكلى					
اربيل	2111	40635	175989	216624	30	11	446985	211.75	10
دهو ك	208	9766	5630	15396	77	9.8	95707	460.13	12.35
السليمانية	2449	52681	104827	157508	74	12.75	674386	275.37	5.14
المجموع	4768	103082	286446	389528	60.33	11.18	1217078	315.75	9.16

-المصدر:احتسب بالاعتماد على بيانات ممثلية (F. .A.O) في شمال العراق ،مكتب تنسيق (F.A.O) شمال العراق (1998) ،نشرة إحصائية،وحدة الإحصاء الزراعي ،العراق -اربيل.

- تم استخراجه بضرب كمية الإنتاج للشجرة الواحدة في عدد الأشجار المثمرة

الملحق الثاني

1 -مصفوفة الارتباطات

Correlations

	y	x1
y		0.0429 (19) 0.8617
x1	0.0429 (19) 0.8617	
x2	0.9232 (19) 0.0000	-0.0352 (19) 0.8862

Correlation
(Sample Size)
P-Value

2- نتائج تقدير نموذج الانحدار

Multiple Regression Analysis

Dependent variable: y

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-80471.9	28269.1	-2.84664	0.0122
LN21	16521.9	7547.42	2.18908	0.0448
lnx2	22542.0	5266.56	4.28021	0.0007
x22	0.0970187	0.0188316	5.15192	0.0001

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	1.31493E11	3	4.38309E10	43.07	0.0000
Residual	1.5265E10	15	1.01767E9		
Total (Corr.)	1.46756E11	18			

R-squared = 89.5985 percent
R-squared (adjusted for d.f.) = 87.5182 percent
Standard Error of Est. = 31900.9
Mean absolute error = 20659.5
Durbin-Watson statistic = 2.07257

Stepwise regression

Method: backward selection

F-to-remove: 4.6

Step 0:

3 variables in the model. 15 d.f. for error.

R-squared = 89.60% Adjusted R-squared = 87.52% MSE = 1.01767E9

Final model selected.

3- تقديرات (Park- Gljser Test) باستخدام برنامج (StatiGraph)

أ- تقدير العلاقة بين رأ المال(X1) و المتغير العشوائي (ei)

Regression Analysis - Linear model: Y = a + b*X

Dependent variable: lne2
Independent variable: LNX1

Parameter	Estimate	Standard Error	T Statistic	P-Value
Intercept	-2.09076	1.02573	-2.03833	0.0574
Slope	0.321396	0.453345	0.708943	0.4880

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	2.01169	1	2.01169	0.50	0.4880
Residual	68.0435	17	4.00256		
Total (Corr.)	70.0552	18			

Correlation Coefficient = 0.169457
R-squared = 2.87158 percent
Standard Error of Est. = 2.00064

ب – تقدير العلاقة بين المساحة(X2) و المتغير العشوائي (ei)

Regression Analysis - Linear model: Y = a + b*X

Dependent variable: lne2
Independent variable: lnx2

Parameter	Estimate	Standard Error	T Statistic	P-Value
Intercept	-2.10702	1.10024	-1.91507	0.0725
Slope	0.165922	0.24881	0.666861	0.5138

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	1.78586	1	1.78586	0.44	0.5138
Residual	68.2693	17	4.01584		
Total (Corr.)	70.0552	18			

Correlation Coefficient = 0.159663
R-squared = 2.54922 percent
Standard Error of Est. = 2.00396

قائمة المختصرات

lnY = اللوغاريتم الطبيعي للرأسمال .

ln X1 = اللوغاريتم الطبيعي للرأسمال .

ln X2 = اللوغاريتم الطبيعي للرأسمال .

X22 = تربيع متغير راس المال .

ei = المتغير العشوائي .

.....

.....