الفصل الحادي عشر

الأساليب الكمية في العلوم الإدارية

Quantitative Techniques

الفصل الحادي عشر

الأساليب الكمية في العلوم الإدارية

Quantitative Techniques

1- 1. مفاهيم عامة (General Concepts)

(1) المجتمع (Population) أو (Universe)

ويقصد به جميع مفردات أو وحدات (Units) الظاهرة تحت البحث. فقد يكون المجتمع مكونا من سكان مدينة أو مجموعة من المزارع في منطقة معينة أو مجموعة من الحيوانات أو وحدات سلعة معينة ينتجها معمل معين.

وعليه يمكن القول:

أن المجتمع الإحصائي (Statistical Population): هو مجموعة من الوحدات الإحصائية (Statistical Units) معرفة بصورة واضحة بحيث تميز الوحدات الإحصائية التي تدخل ضمن هذا المجتمع عن غيرها.

(2) العينة (Sample)

هي جزء من المجتمع (مجموع وحدات إحصائية) يجري إختيارها من المجتمع الإحصائي وفق قواعد خاصة لكي تمثل المجتمع تمثيلا صحيحا. ونظرا لأهمية العينة في لبحث العلمي سيتم تخصيص فصلا كاملا لها لاحقا.

(3) المعاينة (Sampling)

هي الطريقة التي يختار بها هذا الجزء بحيث تكون خواص المجتمع بما فيها الاختلاف بين الوحدات منعكسة في العينة بأحسن ما يسمح به حجم العينة. حيث تستند هذه الطريقة على قواعد مستمدة من النظرية الإحصائية والتي تعتمد على نظرية الاحتمالات وقواعد رياضية أخرى.

فوائد المعاينة:

يمكن للمعاينة أن تحقق الفوائد التالية

أ- إختصار الوقت والجهد والتكاليف.

ب- إمكانية الحصول على بيانات أكثر بواسطة العينة مما نستطيع الحصول عليه من أفراد المجتمع كله وعليه يمكن توسيع مجال البحث.

جـ- في طريقة العينات هناك طرائق لتحديد مدى الدقة للنتائج ونسبة تمثيلها للمجتمع.

وعند البدء بالمعاينة يجب تحديد:

1- وحدة المعاينة بصورة واضحة.

2- الإطار (Frame) والذي يشمل كل الوحدات المحتملة في المجتمع تحت الدراسة.

3- إجراء إختبار مسبق (Pre- Test) على مجموعة من وحدات المجتمع كعينة إختبارية وعلى ضوء هذا الاختبار التجريبي يمكن القيام

بمجموعة من التقديرات كتقدير حجم العينة والكلفة النسبية والوقت اللازم للعينة المطلوب سحبها.

المعاينة العشوائية البسيطة (Simple Random Sampling)

هي طريقة إختيار عينة بطريقة عشوائية إعتباطية بحيث يكون لجميع وحدات المعاينة في المجتمع نفس النصيب أو الاحتمال في الاختيار. فإذا كان عدد مفردات المجتمع هو (N) فإن إحتمالإختيار أي مفردة منه هو (1/N).

التقديرات (المقدرات) (Estimators)

بالرغم من أن المعاينة ترمي إلى أهداف إلا أن الاهتمام ينصب غالبا إلى دراسة الصفات الآتية والتي تدعى بمعالم (Parameters) المجتمع:

أ - الوسط الحسابي للمجتمع (Population Mean) ويرمز له بالرمز M أو ($\bar{Y}$).

ب - المجموع الكلي (Total) ويرمز له بالرمز (Y)

جـ - نسبة مجموعتين (Ratio) كنسبة أجور الكهرباء إلى الدخل الشهري للعائلة ويرمز لها

$$R=\frac{\bar{Y}}{\bar{X}}$$

د – نسبة القيم للمفردات التي تمتاز بميزة خاصة (Proportion) كنسبة الأفراد العاطلين أو نسبة المواطنين بالولادة أو نسبة الأشخاص الذين لا يزيد دخلهم عن حد معين ويرمز لها P= A / N.

ويكون التقدير لهذه المعالم من معطيات العينة وعلى النحو التالي:

العينة المسحوبة من المجتمع Sample عدد المفردات (n)Estimator	المجتمع Population عدد المفردات NParameters
1- $\hat{\bar{Y}} = \bar{y} = \sum \frac{yi}{n}$ 2- $\hat{Y} = N\bar{y}$ 3- $\hat{R} = \frac{\bar{y}}{\bar{x}}$ 4- $\hat{P} = a/n$	1-Mean = M or $\bar{Y} = \sum \frac{yi}{N}$ 2- Total = $Y = N\bar{Y}$ 3- Ration = R = $\bar{Y} / \bar{X}$ 4- Proportion = P = A/N

ولمعرفة دقة هذه التقديرات نحسب عادة التباين (Variance) والذي يرمز لـه بـالرمز S^2 أو الخطـأ المعيـار (Standard Error) للتقدير S وعلـى النحـو التالي:

1) $S^2(\bar{y}) = S^2y/n\ (1-f)$ لمعرفة الدقة (الخطأ المعياري للوسط الحسابي)

2) $S^2(\hat{Y}) = N^2 * S^2y/n\ (1-f)$ الخطأ المعياري للمجموع الكلي

3) $S^2(\hat{R}) = \left[\frac{1-f}{n\,\bar{x}^2}\right] * \left[\frac{\Sigma\left((y)^2 i - 2R\Sigma yixi + R^2 \Sigma^n x^2 i\right)}{n-1}\right]$

4) $S^2(Pn) = P8/n-1\ (1-f)$

حيث تشير (F) إلى كسر المعاينة Sampling Faction

ويعطى بالصيغة التالية $f=h/N$ أما $\xi=1-P$

وإن التباين:

$$S^2y=[\sum^{n} y^2 - (\sum^{n} yi)^2 /n]/ (n-1)$$

تقدير حجم العينة (Estimation of Sample size)

يعتمد حجم العينة التي تسحب من مجتمع ما على الخطأ المسموح به (d) والاحتمال المطلوب للحصول على الدقة المطلوبة، حيث يكون حجم العينة

فـــي حالــــة تقـــدير متوســـط المجتمــــع ← $n_0=(Z\sigma /d)^2$ ، $n= n_0/(1+n_0/N)$

في حالة تقدير المجموع الكلي ← $n_0=(ZN\sigma /d)^2$

في حالة تقدير النسبة ← $n_0=Z^2P ξ/d^2$

إن الصيغة أعلاه تعطي حجم العينة (n) التي تسمح بخطأ مساوي إلى (d)بإحتمال(w) إما (z) فهي القيمة الحدودية للتوزيع الطبيعي القياسي التي تقابل كفة (1-w).

التقديرات (Estimators)

من أهداف المعاينة دراسة صفات المجتمع والتي تدعى عادة معالم المجتمع (parameters). منها

أ ـالوسط الحسابي للمجتمع (Population Mean) = $(\bar{Y})$

ب ـ المجموع الكلي أو القيمة الكلية (Total): $Y=N\bar{Y}$

جـ - نسبة الو□دات التي تملك خاصية معينة (Proportions): $P = A / N$

والتي تقدر على النـ□و التالي من عينة عشوائية □جمها (n)

a) $\hat{\bar{Y}} = \bar{y} = \sum_{i=1}^{n} yi \,/\, n$

ولمعرفة دقة هذا التقدير نحسب تباينه، حيث:

$S^2(\bar{y}) = (\sigma^2 y/n)\ (1-f)$

حيث $(\sigma^2 y)$ هو تباين المجتمع الذي سحبت منه العينة، وإن (f) هو كسر المعاينة حيث $f = n/N$

وإن حجم العينة البدائي اللازم لتقدير متوسط المجتمع هو:

$n_0 = (Z\sigma\ / d)^2$

b) $\hat{Y} = N\bar{Y}$

وتباين

$S^2(\hat{Y}) = N^2 S^2(\bar{y})$

إما حجم العينة البدائي اللازم لتقدير المجموع الكلي:

$n_0 = (ZN\sigma\ /d)^2$

$C.\hat{P} = Y = a / n$

وأن تباين هذه النسبة هو:

$S^2(\psi) = \frac{\psi\xi}{n-1}(1-f)$، $\xi = 1-\psi$

إما□جم العينة البدائي المطلوب لتقدير تلك النسبة:

$$N_0 = Z^2 \psi \xi / d^2$$

المعاينة العشوائية الطبقية: Stratified Random Sampling

في هذا الأسلوب من المعاينة يقسم المجتمع الذي يحتوي علـى (N) مـن الوحدات إلى (L) من الطبقات تسمى "Strata" أحجامها N1،N2،NL..... على التوالي ويتم التقسيم والتجزئة وفقا لخصائص أو شروط معينة بحيث تكون كل طبقة متجانسة في داخلها وتختلف عن الطبقات الأخرى وتسحب من كل طبقة عينة عشوائية ذات حجم معين علـى إعتبـار أن كل طبقة تمثل مجتمعا مستقلا وبهذه الطريقة تكون العينـة الكبرى التـي هـي مجموع عينـات الطبقـات تمثـل المجتمع المذكور تمثيلا أدق.

تقدير حجم العينة:

لحساب الحجوم الجزئية للعينة الكلية التي تنتخب من كل طبقة:

أ . في حالة التوزيع المتناسب (Proportional Allocation) يكون

$$nh = n\,(Nh / N), \quad h = 1, 2, \ldots L$$

ب . طريقة التوزيع الأمثل "Optimum Allocation": فـي هذه الطريقة نفـرض أن كلفـة وحجـم البيانـات لأي وحـدة مـن وحـدات المعاينـة للطبقـة (h) هــي Ch وإن مجمـوع الكلفـة الأوليـة لتهيئـة الوسـائل المختلفة لإجراء المعاينة هي C0 فالكلفة الكلية إذاً تكون:

$$C = C0 + \sum_{h=1}^{n} C_h n_h$$

وإن قيمة الأجزاء (nh) والتي تجعل تباين $\bar{y}st$ أصغر ما يمكن

$nh= n [(wh\sigma h/\sqrt{Ch}) / (\sum wh\sigma h/\sqrt{C\sigma})]$ ، $Wh = Nh / N$

$nh = [(C – C0)(Nh\sigma n / \sqrt{C\sigma})] / [\sum Nh\sigma n \sqrt{C\sigma}]$

جـ - طريقـة توزيـع نيمـان (Neyman Allocation): إذا كانـت تكـاليف المعاينة متساوية لكل الطبقات فإن أجزاء العينة للطبقات تكون:

$nh = n (Wh\sigma h / \sum Wh\sigma h)$

المقدرات (التقديرات):

أ) إن متوسط المعاينة الطبقية هو:

$$\bar{y}st. =(\sum_{h=1}^{l} n_h \bar{y}_h) / n$$

وبتباين للوسط الحسابي

$$S^2(\bar{y}st) = \sum_{h=1}^{l} W^2{}_h * \sigma^2 h / nh (1-fh)$$

ب) إن تقدير الإجمالي (الكلي) total هو كالتالي:

$Y^\wedge = N \bar{y}st.$

وبتباين للمجموع الكلي:

$$S^2(Y\hat{}) = \sum_{h=1}^{l} \square^2 h \; (\square^2 h/nh)(1-fh)$$

إختبار الفرضيات (Testing of Hypothesis)

يعتبر موضوع إختبار الفرضيات من أهم المواضيع في مجال إتخاذ القرارات حيث تعرف الفرضية الإحصائية (Statistical Hypothesis) بإنها إدعاء أو تصرف قد يكون صائبا أو خاطئا حول معلمة (Parameter) أو أكثر لمجتمع أو لمجموعة مجتمعات وعادة تؤخذ عينة من المجتمع ذات العلاقة. وتستخدم جميع المعلومات منها للوصول إلى قرار بقبول أو رفض الفرضية الإحصائية، فترفض الفرضية عندما تكون بيانات العينة لا تساند النظرية وهذا يعني بأن الفرضية خاطئة لذلك فأن الباحث يحاول دائما أن يضع الفرضية بشكل يأمل أن يرفضها. وأن الفرضية التي يضعها البا□ث على أمل أن يرفضها تدعى بفرضية العدم (Nil hypothesis) ويرمز لها بالرمز H0. ورفضنا لفرضية العدم يقودنا إلى قبول فرضية بديلة عنها، هذه الفرضية تدعى بالفرضية البديلة(Alternative hypothesis) ويرمز لها بالرمز H1.

إن طريقة إتخاذ القرارات قد يقودنا إلى الوقوع في نوعين من الخطأ هما:

أ- **الخطأ من النوع الأول(Type I Error):** يقع البا□ث فيه إذا رفض فرضية العدم عندما تكون هي الفرضية الص□ي□ة.

ب- **الخطأ من النوع الثاني (Type II Error):** يقع البا□ث فيه إذا قبل فرضية العدم عندما تكون هي الفرضية الخاطئة.

حيث يمكن تلخيص ذلك في الجدول التالي:

الحالة الحقيقة / القرار	H_0 صحيحة	H_1 خاطئة
قبول H_0	قرار صائب	الخطأ من النوع الأول
رفض H_0	الخطأ من النوع الثاني	قرار صائب

إن رفض H0 أو قبولها يتم على أساس قياسات العينة لحساب المختبر الإحصائي أو معيار الاختبار أو إحصاء الاختبار (Test Statistic)وإعتمادا على الفرضية البديلة يتحدد نوع الاختبار فيما إذا كان من جانب واحد (one – tailed test)أو من جانبين (Two – tailed test).

أولا: الاختبارات المستندة إلى التوزيع الطبيعي

أ -**إختبارات تتعلق بالمتوسطات:** في هذا الاختبار يكون هدف الباحث منه هل أن عينة ما تنتمي لمجتمع معين أو لا؟ حيث يضع الباحث فرضيته على النحو التالي:

H0: $\mu : \mu_0$

Against H1: $\mu \neq \mu_0$

Or H1: $\mu > \mu_0$

Or H1: $\mu < M_0$

وإن إحصاءة الاختبار هي:

$$Z = (\bar{y} - \mu_0) / \left(\frac{\sigma\,(s.d)}{\sqrt{\square}} \right)$$

حيث $\bar{y}$ هو متوسط العينة، $\square$ هو الانحراف المعياري للمجتمع، n عدد مفردات العينة (يزيد على 30)

ب – إختبارات تتعلق بمتوسطين، يهدف الباحث مقارنة متوسطي مجتمعين معروف بتباينها، حيث يضع الباحث فرضيته $\sigma^2 1, \sigma^2 2$

H0 = $\mu 1 = \mathcal{M}_2$

Against H1 = $\mu 1 \neq \mathcal{M}_2$

وإن إحصاءة الاختبار هي

$$Z = (\bar{y}1 - \bar{y}2) / \sqrt{\left(\frac{\sigma^2 1}{n1}\right) + \left(\frac{\sigma^2 2}{n2}\right)}$$

جـ - إختبارات تتعلق بالنسب: الاختبار الفرضية:

H0: P=P0

Against H1: P $\neq$ P$\square$

ويكون معيار الاختبار:

$$Z = \frac{\Psi - P0}{\sqrt{\frac{P0\,\xi\square}{n}}}$$

، $\xi 0=1-P0$

د – لاختبار نسبتين

H0: P1=P2 against H1= P1 $\neq$ P$\square$

$\psi_1 \neq \square_2$

توزيع مربع كاي (Chi – Square Distribution)

يستخدم مربع كاي(X^2) في الفرضيات وعلى النحو التالي:

أولا – اختبار يتعلق بتباين المجتمع (إختبار التجانس - TestHomogeneity)

وتشمل الفرضية هنا مقارنة تباين المجتمع التطبيقي بقيمة معينة أي:

$H0 = \square^2 = \square^2 0$

$H1 = \square^2 > \square^2 0$

وتتلخص طريقة الاختبار بإختيار عينة عشوائية ذات حجم n من المجتمع وحساب تباين هذه العينة

$$S^2 = \sigma^2 = \frac{\Sigma yi^2 - \frac{(\Sigma yi)^2}{\square}}{n-1}$$

ومن ثم نحسب إحصاءة الاختبار

$$X^2 = \frac{(n-1)S^2}{\sigma 0^2}$$

إختبار الاستقلال (بين متغيرين) (Test For Independence)

يمكن استخدام مربع كاي عندما يراد اختبار فرضية حول استقلال متغيرين(أو ظاهرتين سواءا أكانت نوعية أو وصفية). وعادة تنظم البيانات في جدول مزدوج يسمى جدول التوافق (Contingency table) وكالآتي:

الصفوف الأفقية Rows	الأعمدة الرأسية Columns C1 C2 Ci			المجموع
r1	O11	O12	O1c	R1
r2	O21	O22	O2c	R2
rr	Or1	Or2	Orc	Rr
المجموع	C1	C2	Cc	T

فالأعمدة تمثل متغير (ظاهرة معينة) والصفوف الأفقية تمثل متغير آخر بينما التكرار المشاهد فهو المشاهد في الخلايا (Cells) داخل الجدول ويرمز له Oij ولاختبار الفرضية القائلة بأنه الأعمدة والصفوف تمثل تصنيفات مستقلة تحتاج إلى حساب التكرارات المتوقعة لكل خلية وعلى النحو التالي:

$$Eij = \frac{(Ri)(Cj)}{T}$$

وبعد إيجاد التكرار المتوقع لجميع الخلايا في الجدول نطبق:

$$X^2=\sum \frac{(Oij - Eij)^2}{Eij}$$

والذي له درجات حرية تساوي:

(r-1)(C-1)

أما الفرضية التي ستختبر فهي أن الظاهرتان مستقلتان ؟ أي لا توجد علاقة بين الظاهرتين. ولقياس درجة العلاقة (Relationship) أو الاعتماد (Dependence) بين ظاهرتين نحسب قيمة معامل التوافق حيث:

$$C = \sqrt{\frac{X^2}{\square^2 + T}}$$

وكلما زادت قيمة C كانت درجة العلاقة قوية بين الظاهرتين

إختبار فرضيات حول تساوي عدة أوساط حسابية

عندما يراد مقارنة أكثر من مجموعتين تجريبيتين نلجأ إلى طريقة تحليل التباين (Analysis of Variance) والذي يرمز إليها إختصارا(ANOVa).

إن ت□ليل التباين عبارة عن عملية رياضية يقسم فيها التباين الكلي إلى مكوناته أو مصادره المختلفة ويوضع في جدول يسمى جدول تحليل التباين. لنفرض أن أحد الباحثين لديه نتائج K من العينات العشوائية كل ذات حجم nأختيرت من K من المجتمعات المختلفة وعادة يطلق على هذه العينات اسم المعاملات (Treatments) وإن رغبة الباحث هي في إختبار أن الأوساط الحسابية متساوية. أي:

H0: $\mu 1 = \mu 2 = \cdots = \mu \square$

H1: على الأقل اثنان منهم غير متساويان

فإنه يمكن تنظيم المشاهدات أو نتائج العينات كالآتي:

	المعاملات (المجاميع، العينات) 1 2 3 .... K	
Total	y11y21 y31 yk1 y21y22 y32 yk2 ⋮ ⋮ ⋮ ⋮ y1n y2n y3nykn	المجموع الكلي
المجموع	y1. y2. y3. yk.	y....

Grand Total

ونقوم بحساب مجموع المربعات الكلي المصحح (Corrected total sum of squares) ويرمز له بالرمز SSTحيث:

$$SST = \sum_{i,j} y^2 ij - \frac{(y..)^2}{hk}$$

ومجموع مربعات مابين المجاميع (المعاملات) (Between treatments sum of squares) ويرمز له SSt

$$SSt = \frac{\sum_i y^2 i.}{n} - \frac{(y..)^2}{hk}$$

إما مجموع مربعات الخطأ ويدعى أحيانا (داخل المعاملات) (Within treatments sum of squares) ويرمز له بالرمز SSE

ويصبح جدول تحليل التباين على النحو التالي: $SSE = SST - SSt$

طرق المقارنات المتعددة بين المعاملات

(Multiple Comparisons among treatments)

1. لمقارنة متوسطين حسابيين لعينتين عشوائيتين أو مجموعتين تجريبيتين نستخدم عادة الاختبار الطبيعي (Z) عندما يكون حجم العينة كبير (أكثر من 30) حيث تكون الفرضية كالآتي:

$H0: \mathcal{M}1 = \mathcal{M}2$

$H1: \mathcal{M}1 \neq \mathcal{M}2$

Or $H1: \mathcal{M}1 < \mathcal{M}2$

Or $H1: \mathcal{M}1 > \mathcal{M}2$

إما معيار الاختبار فهو:

$$Z = \frac{(y1 - y2) - (\mathcal{M}1 - \mathcal{M}2)}{\sqrt{\frac{\sigma^2 1}{h1} + \frac{\sigma^2 2}{h2}}}$$

إما إذا كان حجم العينة صغير (30 فأقل) نستخدم إختبار(t) ويكون معيار الاختبار

$$t = \frac{(y1 - y2) - (\mathcal{M}1 - \mathcal{M}2)}{\sqrt{\frac{S^2 1}{h1} + \frac{S^2 2}{h2}}}$$

2. إما عندما يراد مقارنة أكثر من مجموعتين تجريبيتين فإن إختبار(t) يكون غير عملي ونلجأ إلى طريقة لاختبار المتوسطات تدعى طريقة تحليل التباين Analysis of Variance حيث تكون الفرضية:

H0: $\mu_1 = \mu_2 = \cdots = \mu_k$

H1: على الأقل اثنان منهم غير متساويان

مثال / الجدول الآتي هو نتائج في تجربة صممت لمقارنة الأوساط الحسابية لست مجاميع

1	2	3	4	5	6	Total
19.4 32.6 27.0 32.1 33.0	17.7 24.8 27.9 25.2 24.3	17.0 19.4 9.1 11.9 15.8	20.7 21.0 20.5 18.8 18.6	14.3 14.4 11.8 11.6 14.2	17.3 19.4 19.1 16.9 20.8	X..(y..)
144.1	119.7	37.2	99.6	66.3	93.5	596.6
28.8	24.0	14.6	19.9	13.3	18.7	

$\bar{y}\bar{x}$ $\bar{x}i$

فهل تعتقد إن هناك فروقا أساسية ما بين المتوسطات وأي من المتوسطات هو الأفضل؟ في مستوى دلالة 0.05

الحل:

H0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$

H1: هناك فروقات معنوية بين متوسطين على الأقل

$$SST = (19.4)^2 + \ldots + (93.5)^2 - \frac{(596.6)^2}{6(5)}$$ (مجمــوع المربعــات

الكلي)....

$$= 12999.36 - 11864.38 = 1129.98$$ Total SS

$$SSt = \frac{(144.1)^2 + \ldots\ldots + (93.5)^2}{5} - 11864.38$$

$$= 847.05$$ Treatment SS (مجموع مربعات معاملات مابين المجاميع)

$$SSE = SST - SSt = 1129.98 - 847.05$$

$$= 282.93$$ Error SS (مجموع مربعات الخطأ)

وعليه فأن جدول تحليل التباين هو:

F	Mean squares متوسط المربعات	sum of squares مجموع المربعات	d.f درجات الحرية	Source of variation مصادر التباين
14.37	169.41	847.05	5	Among groups بين المجاميع
	11.79	282.93	24	Within groups داخل المجاميع
		1129.98	29	المجموع الكلي

وبما أن قيمة F المحسوبة تقع في منطقة الرفض Ho أي ـن الفروق المشاهدة لها دلالة معنوية (Significant)

3) عندما ترفض H0 فيجب التقدم مرحلة أخرى في التحليل لمعرفة أي من المتوسطات هو الأفضل أما إذا قبلت H0 فنقف في التحليل الإحصائي لهذا الحد.

طرق المقارنة بين المتوسطات

وهناك عدة طرق للمقارنة بين المتوسطات منها:

1. The least significant difference (Lsd) أصغر فرق معنوي
2. The honestly significant difference (Hsd) الأدق فرق معنوي

3. Duncan's new multiple – range test طريقة دنكان الجديدة

4. Student–Newman–Keuls' test (SNK) طريقة ستيودنت نيومان كويل

5. Dunnett's method طريقة دونت

6. Sheffe's method طريقة شيفي

7. Parzen method طريقة بارزن

وبالنظر لانتشار طريقة أصغر فرق معنوي (Lsd) بالبحوث الزراعية على الرغم من وجود عيب فيها كبير. وهي أن المتوسط لأية مجموعة يظهر مرة واحدة في المقارنة، لذلك فسنعطي فكرة عنها معتمدين على نتائج المقال السابق ذكره.

4) طريقة أصغر فرق معنوي:

أن قيمة أصغر فرق معنوي تحت مستوى دلالة 0.05 هو:

$$Lsd(0.05) = t = \sqrt{\frac{2S^2}{\square}} \text{ ، } S^2 = MSE$$

حيث أن S^2 هو متوسط المربعات للخطأ في جدول تحليل التباين

وعليه فأن $Lsd(0.05) = 2.064\sqrt{\frac{2\,(11.79)}{5}} = 4.5$

ويعتبر الفرق الظاهري ما بين المتوسطين معنويا إذا كان الفرق المشاهد يزيد على Lsd وعليه فأن = الفروق مابين المتوسطات حسب هذه الطريقة هي:

$\bar{X}1 - \bar{X}2 = 28.8 - 24.0 = 4.8$

$\bar{X}3 - \bar{X}4 = 14.6 - 19.9 = -5.3$

$\bar{X}_5 - \bar{X}_6 = 13.3 - 18.7 = -5.4$

وأن الفروق المشاهدة كلها معنوية تحت مستوى 0.05

5) طريقة دانكن الجديدة:

وفي هذه الطريقة يمكن مقارنة كل وسط حسابي مع كل من الأوساط الأخرى.

وتتم المقارنة على النحو التالي:

أ – ترتب الأوساط الحسابية بصورة تصاعدية أو تنازلية ⇄.

ب – تستخرج قيم ستيودنت المعنوية والذي يرمز له بالرمز (SSR).

جـ - تتم المقارنة مع أصغر فرق في المدى (LSR) حيث يكون

$$LSR = SSR\sqrt{\frac{S^2}{\square}} \quad ، \quad S^2 = MSE$$

من نتائج المثال $\sqrt{\frac{4.79}{5}} = \sqrt{\frac{S^2}{\square}}\,1.54 =$

Value P	2	3	4	5	6
SSR	2.92	3.07	3.15	3.22	3.28
LSR	4.5	4.7	4.9	5.0	5.1

	6	5	4	3	2
	28.8	24.0	19.9	18.7	14.6
13.3	15.5*	10.7*	6.6*	5.4*	13.3 Ns
14.6	14.2*	9.4*	5.3*	4.1 Ns	
18.7	10.1*	5.3*	4.1 Ns		
19.9	8.9*	4.1 Ns			
24.0	9.8*				

تحليل الانحدار التطبيقي "Applied Regression Analysis"

يمكن أن يعرف تحليل الانحدار بشكل رئيسي بأنه تحليل العلاقات بين المتغيرات كما أنه لإقامة العلاقة الدالية بين المتغيرات، مثل هذه العلاقة الدالية يعبر عنها بصيغة معادلة تتعلق بالاستجابة (Response)، (y) مع واحد أو أكثر من المتغيرات التوضيحية {X1،X2،.......،Xk}

إن معادلة الانحدار العامة تأخذ الصيغة التالية:

$$Yi = B0+B1X1i+B2X2i+......BkXki+uI, \quad i = 1,2,3 .... n$$

حيث أن B0، هي المقطع على المحور العمودي (y- intercept) أما B1،B2،Bk... فهي معاملات الانحدار(regression coefficients) وتحسب عادة من البيانات. أما ui فهو الخطأ العشوائي (random error).

إن معادلـة الانحـدار المحتويـة علـى متغيـر توضـيحي واحـد فقـط تسـمى معادلة الانحدار البسيط (Simple regression)، أما المعادلة المحتوية على أكثر مـن متغيـر توضـيحي واحـد فتـدعى معادلـة الانحـدار المعقـد (Multiple regression). وأن معادلة الانحدار يمكن أستخدامها لعدة أغراض. فقد تستخدم لتقييم أهميـة المتغيـرات التوضـيحية (X' s). أو للتنبـؤ بقـيم الاسـتجابة(y) عنـد مجموعة معينة من قيم X's.

الانحدار البسيط Simple regression

إن البيانـات (Data) تحتـوي علـى (n) مـن المشـاهدات حـول متغيـر الاستجابة (y) مع المتغير التوضيحي X1، حيث يتم تسجيل المشاهدات كالآتـي:

رقـم المشاهدة	y	X1
1	y1	x11
2	y2	x12
3	y3	x13
.	.	.
.	.	.
n	yn	xn

ويمكن أن تصاغ العلاقة الخطية بين X1،y لنموذج خطي بالشكل التالي:

$$yi = B0 + B1X1i + Ui \quad , \quad i = 1,2,3 \ldots\ldots, n$$

وتقدر المعالم B0،(parameters)B1 بالاعتماد على قيم المشاهدات (معطيات العينة) وبواسطة طريقة معروفة لا تعتمد على توزيع البيانات تدعى بطريقة المربعات الصغرى (Least squares) ويرمز لها بالرمز (OLS) والتي تتضمن تصغير مجموع مربعات الخطأ $\sum_{i=1}^{n} \square^2_i$، وواضح أن (B0) هو الحد معدل قيمة yi عندما تكون $X2 = 0$. أما B1 فهو معامل انحدار y على x ويعرف بأنه معدل المتغير في (y) عندما تتغير قيمة x وحدة واحدة. وأن Ui الخطأ العشوائي، وهي كميات مستقلة بوسط حسابي قدره صفر وتباين قدره $\square^2$. ومن معطيات عينة عشوائية حجمها n من المشاهدات يمكن تقدير المعالم وعلى النحو التالي:

$$B\hat{}1 = \frac{\sum_{i=1}^{n} yix1i - \frac{\left(\sum_{i=1}^{n} yi\right)\left(\sum_{i=1}^{n} x1i\right)}{n}}{\frac{\sum_{i=1}^{n} \square 1i^2 - \left(\sum_{i=1}^{n} x1i\right)^2}{n}}$$ ،

$$B\hat{}0 = \bar{y} - B\hat{}1\bar{x}1$$

أما قيمة الاستجابة المتنبأ بها بواسطة النموذج والمناظرة للمشاهدة ن، فهي:

$$y\hat{}i = B\hat{}0 + B\hat{}1\, x1i$$

ولأجل اختبار فرضية العدم: $H0: B1=0$

$$\text{Versus } H1: B1 \neq 0$$

فأن معيار الاختبار هو :

$$T = B\hat{}1 - 0 / s.e(B\hat{}1)$$

ولتقييم جودة المطابقة للنموذج فأن المؤشر الإحصائي الأكثر انتشارا فهو معامل التحديد والذي يرمز له بالرمز R^2 والذي يمكن تفسيره بنسبة التغير الكلي في y المشروحة بواسطة x1، حيث:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^{\square\square}\right)^2}{\sum_{i=1}^{n}\overline{y_i - (y)^2}}$$

وكما يلي الشرح الكامل للقانون:

إذا كان R^2 قريبا من الواحد فأن x1 تفسر أو تشرح الجزء الأكبر من الاختلاف في y.

وأن الخلاصة الإحصائية الضرورية ستولد من برامج الانحدار الجاهزة بإستخدام مكتبة ال□اسوب "The statistical package for the social science" والتي يرمز لها بالرمز Spss

مثال تطبيقي /

شركة لتسوية وتصليح ال□اسبات الصغيرة ترغب بالتنبؤ بعدد مهندسي الخدمة طلب الخدمة. □يث أن مدة الطلب (y) تعتمد على عدد الأجزاء الألكترونيةx1 في الحاسوب التي يجب إصلاحها أو إستبدالها، ولتعيين هذه العلاقة أخذت عينة من سجلات طلب الخدمة، حيث أن البيانات تحتوي على مدة طلبات الخدمة (بالدقائق) وعدد الأجزاء التي تم تصليحها.

(y) = 23 29 49 64 74 87 96 97 109 119 149 145 154 166

والآتي نتائج(x1) = 1 2 3 4 4 5 6 6 7 8 9 9 10 10

الحاسوب:

Coefficients	SE	t
B1^ = 15.509 B0^ = 4.162	0.505 3.355	30.71 * 1.24
h = 14، R^2 = 0.987	S = 5.392	

ولتفسير نتائج الحاسوب "Print out"

أ) نبدأ التحليل بفحص ملائمة المتغير التوضيحي (x1)، حيث أن معامل التحديد يشير إلى أن 99% من الاختلاف في مدة طلب الخدمة (y) يمكن تفسيره أو توضيحه بواسطة عدد الوحدات (x1) التي تم تصليحها.

ب) يمكننا تقييم القابلية التفسيرية للوحدات (x1) باختبار

H0: B1 = 0

Against H1: B1 ≠ 0

حيث B1 تعرف بمعامل الانحدار لمدة طلب الخدمة (دقائق) على الوحدات. إما معيار الاختبار فهو

$$t = \frac{15.509 - 0}{0.505} = 30.71$$

* وبما أن قيمة (t) المحسوبة تقع في منطقة الرفض وعليه ترفض H0 وهذا يؤكد أهمية المتغير (x1).

جـ) إما المعادلة التقديرية فهي:

$$y\hat{}i = 4.162 + 15.509\,x1i$$

حيث يعبر الحد الثابت عن الوقت اللازم للبدء بالتصليح وهو تقريبا (4) دقائق. أما معامل الوحدات فيعبر عن الزيادة في مدة طلب الخدمة لكل زيادة وحدة إضافية من الأجزاء التي يجب تصليحها. حيث قدرت بحوالي 16 دقيقة لكل جزء إضافي يتم تصليحه.

وللتنبؤ بمدة طلب الخدمة التي فيها يتم تصليح أربعة أجزاء

$$y\hat{} = 4.162 + 15.509(y) = 66.198 \text{ minutes}$$

أما الخطأ المعياري لهذا التقدير فهو

$$S.e(y\hat{}) = S\{1+1/n + \frac{x - (x)^2}{\sum(x1 - x1)^2}\}^{\frac{1}{2}}$$

$$= 5.392\ \{\ 1 + 1/14 + \frac{(4-6)^2}{114}\}^{\frac{1}{2}} = 5.672$$

الانحدار المتعدد (Multiple Regression)

إن البيانات المتكونة من (n) من المشاهدات للمتغير المعتمد أو متغير الاستجابة (y) مع (k) من المتغيرات المستقلة (التفسيرية): {x1، x2، ...، xk}

يعبر عنها: المتغيرات التفسيرية

رقم المشاهدة	متغير الاستجابة	x1	x2	……..	xk
1	y1	x11	x21		xk1
2	y2	x12	x22		xk2
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
n	yn	x1n	x2n		xkn

حيث إن العلاقة بين Y ، {x1، x2، …، xk} يمكن أن تصاغ كنموذج خطي على النحو التالي:

$$y_i = B_0 + B_1x_{1i} + B_2x_{2i} + \ldots + B_kx_{ki} + U_i \quad ، \quad i = 1، 2، \ldots، n$$

حيث أن (B0، B1، ….، Bk) هي ثوابت مجهولة تشير إلى معاملات الانحدار (Ui) كميات عشوائية مستقلة تتوزع بمتوسطات صفرية وتباين ثابت مقداره σ^2 وواضح أن:

أ) معامل الانحدار Bi هي الزيادة في متغير الاستجابة y المناظرة إلى زيادة وحدة واحدة من xi عندما تكون جميع المتغيرات الأخرى ثابتة.

ب) قيم المعالم B's تقدر عن طريق تصغير مجموع مربعات البواقي والتي تعرف بطريقة المربعات الصغرى.

جـ) أن برامج الحاسبات متاحة وتعطي حلول عددية دقيقة.

د) أن القيمة التنبؤية(Predicated Value) تعرف كما يلي:

y^i = B^0+ B^1x1i + ….. + B^kxki

وإن المتبقي المشاهد لكل مشاهدة هو: ei = yi - y^i

هـ) بعد توفيق أو مطابقة النموذج الخطي للبيانات المعطاة فأن دقة المطابقة تقاس بواسطة معامل التحديد R^2 حيث أن:

$$R^2 = 1 - \frac{\Sigma(yi - y^i)^2}{\sum[(yi - \bar{y}])^2}$$

فعندما يكون النموذج ملائما للبيانات فمن الواضح إن قيمة R^2 تقترب من الواحد ونظرا لأن إضافة أي متغير توضيحي إلى معادلة الانحدار يؤدي إلى تخفيض درجات الحرية وبالتالي يرفع من قيمة معامل التحديد لذلك يلجأ الباحثون إلى حساب معامل التحديد المعدل (Adjusted R^2) ويرمز له بالرمز $\bar{R}_2$ حيث:

$$\bar{R}_2 = 1 - \left[\frac{(1 - R^2)(n - 1)}{\square - k - 1}\right]$$

و) إن تقييم الأهمية النسبية للمتغيرات التوضيحية يتم باختبار جوهرية معاملات الانحدار ومقارنة t المحسوبة مع t النظرية (الجدولية) تحت مستوى معنوية ودرجات معنوية معلومتين أو باختبار الانخفاض في معامل التحديد من النموذج التام إلى النموذج المقيد.

مثال / أجريت الدراسة التالية في مؤسسة مالية كبيرة، حيث تضمنت الدراسة أسئلة تتعلق بقناعة الموظف بالمشرفين عليه، وأحد الأسئلة وضع لقياس الأداء العام للمشرف بالإضافة إلى أسئلة تتعلق بالنشاطات المعينة للتفاعل بين المشرف والموظف لمحاولة تفسير العلاقة بين صفات المشرف

والرضا عن المشرفين كما يدركها الموظفون، والجدول التالي يعطي وصفا للمتغيرات في الدراسة:

المتغير	الوصف
y	التقييم العام للعمل المنجز من قبل المشرف.
x1	معالجة مشاكل المستخدم.
x2	لا يسمح بإمتيازات خاصة.
x3	الفرصة لتعلم أشياء جديدة.
x4	الزيادات على أساس الانجاز.
x5	إنتقادي أكثر مما ينبغي للانجازات الرديئة.
x6	نسبة التقدم إلى أعمال أفضل.

وواضح أن المتغيرات x1،x2،x5 تتعلق بالعلاقات الخاصة المباشرة بين الموظفين والمشرف، بينما المتغيرات x3،x4، فهي ذات طبيعة تتعلق بالعمل ككل.

أما المتغير x6 فهو لا يعبر عن التقييم المباشر للمشرف لكنه يفيد كثيرا كمقياس عام لكيفية فهم الموظف لتقدمه في الشركة وقد قام الباحث في توثيق النموذج.

$$Yi = B0+B1x1i+B2x2i+\ldots..+B6x6i+ui$$

وعند إدخال البيانات في الحاسوب تم التوصل إلى النتائج التالية:

والمطلوب تفسير النتائج.

Variable	Coefficient	SE	t
x1	0.6130	0.1610	3.81 *
x2	- 0.0730	0.1357	- 0.54 Ns
x3	0.3200	0.1685	1.90 *
x4	0.0810	0.2215	0.37 Ns
x5	0.0380	0.1470	0.26 Ns
x6	- 0.2170	0.1782	- 1.22 Ns
constant	10.7870	11.5890	0.93 Ns
n= 30	$R^2 = 0.7326$	S = 7.068	

الحل:

(أ) تقييم الأهمية النسبية للمتغيرات ظهرت في "print out " قيم "t" والتي هي قيم معيار الاختبار لاختبار الفرضية

$H0: Bi = 0$ ، i=1،2،....، n

$H1: Bi \neq 0$

ومن النتائج فأن المتغيرين x1،x3 فقط لهما معاملات انحدار تقترب من الاختلاف المعنوي عن الصفر.

(ب) أن قيمة (R^2) المساوية (0.7326) تعني أن حوالي 37% من الاختلافات الكلية في التقييم العام للعمل المنجز من قبل المشرف يمكن تفسيرها.

(جـ) أن من أهداف تحليل الانحدار المهمة هو الوصول إلى وصف ملائم للظاهرة المشاهدة، بدلالة أقل عدد ممكن من المتغيرات المعنوية قدر الإمكان وتسمى هذه العملية (parsimony) وعليه فان اختبارات (t) تستخدم لاقتراح

المتغيرات الأكثر أهمية والتي يجب أن تضمن في المعادلة ومن وجهة النظر هذه فأن x1،x3 قد أقترحت وأن النموذج المختزل (Reduced Model) في هذه الحالة سيكون:

$$Y_i = \delta 0 + \delta 1x1i + \delta x3i + ui$$

وعند إدخال البيانات الخاصة بـ x1، x3،y في الحاسوب فأن معادلة المربعات الصغرى التقديرية هي:

$$\hat{y} = 9.871 + 0.643x1 + 0.211x3$$
$$(7.062) \quad (0.118) \quad (0.134)$$

حيث أن الكميات داخل الأقواس تحت المعاملات هي الأخطاء المعيارية لهذه المعاملات التقديرية على التوالي.

*ملاحظات المعاملات: تستخدم هذه المعاملات لتقييم الأهمية النسبية للمتغيرات أي هل المتغير ضروري أن يبقى في النموذج أم لا؟

عدد المعالم بما فيها الثابت. = p ، حيث n-p-1 = درجات الحرية

❖ إذن لاختبار الفرضية:

H0: B1 = 0

Versus H1: B1 ≠ 0 حيث $t = \frac{Coefficient}{SE}$

$$t = \frac{0.6130 - 0}{0.1610} = 3.81$$

0.05