

الفصل الثالث عشر

التحليل التمييزي

يمكن استخدام التحليل التمييزي لتصنيف الأفراد في جماعات على مقياس أو أكثر، أو التمييز بين الجماعات على أساس التجميع الخطي لعدة مقاييس بعد الحصول على قيمة 'ف' دالة في اختبار تحليل التباين المتعدد (انظر الفصل السابق). ويجب أن يكون لكل حالة في تحليل التمييز درجة أو درجات على متغير كمي أو أكثر وقيمة في متغير تصنيفي يحدد عضوية الجماعة.

وكثيرا ما يطلق على المتغيرات الكمية في التحليل التمييزي متغيرات مستقلة أو منبئة، كما يشار لمتغير عضوية الجماعة بالمتغير التابع أو المتغير المحكي. وقد يبدو استخدام هذه المصطلحات محيرا وبخاصة عند استخدام التحليل التمييزي كاختبار تتبعي عقب الحصول على نتائج دالة في اختبار تحليل التباين المتعدد. فالمتغيرات التابعة في التحليل التمييزي يطلق عليها متغيرات مستقلة أو عوامل في تحليل التباين المتعدد، كما أن المتغيرات التابعة في تحليل التباين المتعدد هي المتغيرات المستقلة في التحليل التمييزي.

تطبيق التحليل التمييزي:

هناك نوعان من التطبيقات في التحليل التمييزي:

- الدراسات التي تهدف إلى تصنيف الأفراد في مجموعات على أساس متغيرات كمية منبئة.
- الدراسات التي تستخدم تحليل التباين المتعدد لبعد واحد كتحليل محوري والتحليل التمييزي كإجراء تتبعي.

أسس استخدام التحليل التمييزي:

يستخدم التحليل التمييزي تجمعا من متغيرات كمية منبئة يطلق عليها دوال التمييز. وعدد الدوال الممكنة لتحليل به 'ن' مجموعة و'ك' من المتغيرات الكمية يبلغ إما (ن - 1) أو (ك)، أيهما أصغر. مثال ذلك إذا كان لدينا ثلاث مجموعات وأربعة

متغيرات كمية فإن عدد الدوال يبلغ ٢ لأن ٢ هي القيمة الأصغر، لأن
(نم - ١) = (٣ - ١) = ٢ و 'ك' = ٤.

وتستخرج الدالة الأولى بحيث تزداد الفروق على هذه الدالة بين المجموعات. ويمكن بعد ذلك استخراج دالة ثانية بحيث تزيد الفروق بين المجموعات بالنسبة لهذه الدالة، مع إضافة قيد بأن الدالة الثانية ليست مرتبطة بالدالة الأولى. ويمكن بعد ذلك إضافة دوال أخرى ولكن دائما مع وجود شرط عدم ارتباطها بالدوال السابقة المستخرجة.

وتشير الجذور الكامنة المرتبطة بدوال التمييز على قدرة الدوال على التمييز بين المجموعات، وكلما زادت قيمة الجذر الكامن كان التمييز بين المجموعات أفضل. والجذر الكامن لدالة التمييز هو النسبة بين مجموع المربعات لـ بين المجموعات إلى مجموع المربعات لداخل المجموعات لتحليل التباين الذي يكون المتغير التابع فيه هو دالة التمييز، والمجموعات كمستويات العامل. ونظرا لأن الجذور الكامنة تعكس قدرة الدوال على التمييز بين المجموعات فإن أكبر جذر كامن يرتبط بدالة التمييز الأولى، ويرتبط الجذر الكامن التالي في الترتيب بدالة التمييز الثانية، وهكذا.

والإجراء الأكثر استخداما في التحليل التمييزي في برنامج SPSS هو تصنيف الحالات في مجموعات. ففي عملية التصنيف تجمع المنبئات مع بعضها البعض تجميعا خطيا للتنبؤ بعضوية الجماعة كما حددها المتغير التصنيفي. ويطلق SPSS على هذه التجمعات الخطية دوال التصنيف أو (Fisher's linear discriminant function). ويشار إلى معاملاتها بمعاملات دوال فيشر Fisher's function coefficients. ويتم تقويم دقة التصنيف بحساب النسبة المئوية للحالات التي صنفت تصنيفا صحيحا بناء على دالة التصنيف. وهناك عملية إحصائية بديلة هي كبا kappa التي تُقوّم أيضا النسبة المئوية التي صنفت تصنيفا صحيحا إلا أنها تصحح عامل الاتفاق بالصدفة.

مسلمات التحليل التمييزي:

المسلم رقم ١: المتغيرات الكمية موزعة توزيعا اعتداليا متعددا لكل مجتمع ويحدد هذه المجتمعات مستويات المتغير التصنيفي.

إذا كانت المتغيرات التابعة موزعة معا توزيعا اعتداليا يكون كل متغير منها موزعا توزيعا اعتداليا بغض النظر عن المتغيرات الأخرى ويكون كل متغير منها موزعا توزيعا اعتداليا مع أي تجميع من درجات المتغيرات الأخرى. ومن الصعب تصور تحقق

هذا المسلم، ولذلك يمكن القول أن التحليل التمييزي يعطي نتائج صادقة نسبيا في ضوء الخطأ من النوع الأول وذلك باستخدام عينات ذات حجم متوسط أو كبير.

المسلم رقم ٢: تباينات وتغايرات المتغيرات التابعة في المجتمع واحدة في جميع مستويات العامل.

إذا اختلف حجم العينات وكانت تباينات وتغايرات المتغيرات التابعة غير متساوية فإن التحليل التمييزي لن يعطي نتائج سليمة. ويسمح SPSS باختبار مسلم تجانس التباينات والتغايرات باستخدام إحصاءة Box's M. ويجب الحذر عند تفسير اختبار 'ف' من هذه الإحصاءة، لأن النتيجة الدالة قد تكون راجعة إلى انتهاك مسلم اعتدال التوزيع الذي يتطلبه تحليل التباين المتعدد لبعد واحد، كما أن النتيجة غير الدالة قد تكون راجعة إلى نقص في القوة.

المسلم رقم ٣: اختيار العينة اختيارا عشوائيا، كما أن درجة أي فرد في العينة في أي متغير مستقلة عن جميع درجات أفراد العينة الآخرين.

لا يجب الثقة في اختبار الدلالة للتحليل التمييزي إذا انتهك مسلم الاستقلالية.

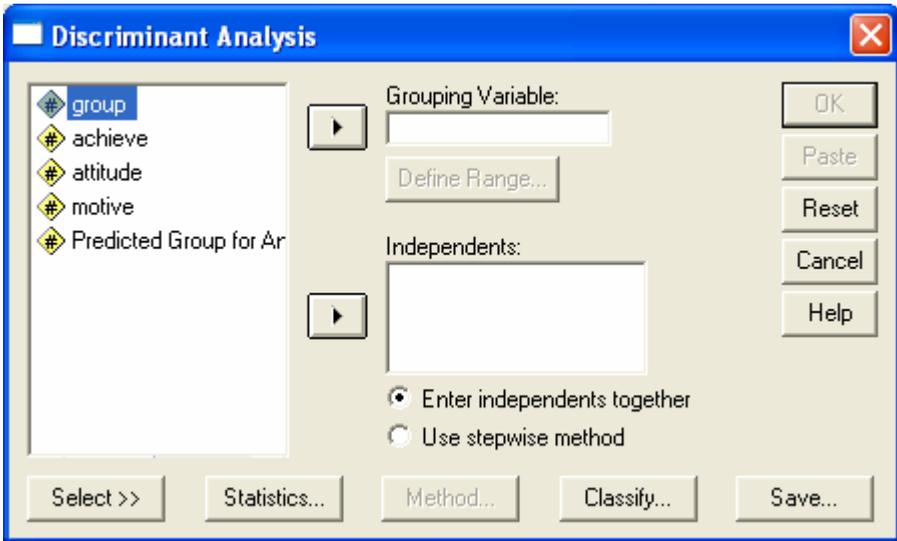
تنفيذ التحليل التمييزي:

قامت الباحثة مایسة فاضل في البحث الذي حصلت فيه على درجة الماجستير من جامعة القاهرة بدراسة العلاقة بين فاعلية البيئة المدرسية وبعض المتغيرات المعرفية (التحصيل الدراسي) وغير المعرفية (الاتجاه نحو المدرسة ودافعية الإنجاز). وقد تساءلت الباحثة عن مدى الاختلاف بين المدرسة ذات الفاعلية الأكبر والمدرسة ذات الفاعلية الأقل من حيث مستويات تحصيلهم الدراسي، ودافعية الإنجاز لديهم وطبيعة اتجاههم نحو المدرسة. أي هل يتفوق تلاميذ المدرسة ذات الفاعلية الأعلى على تلاميذ المدرسة ذات الفاعلية الأدنى في مستوى التحصيل الدراسي ودافعية الإنجاز والاتجاه نحو المدرسة. وقد قامت الباحثة بدراسة على عينة عشوائية من أعضاء هيئة التدريس وإدارة المدرسة في مدارس الحلقة الثانية بنات من التعليم الأساسي (المرحلة الإعدادية)، وقد طبقت الباحثة على أعضاء هيئة التدريس مقياس "فاعلية المدرسة". واستخدمت درجات هذا المقياس في تصنيف المدارس إلى نوعين: مدارس ذات فاعلية عالية، ومدارس ذات فاعلية أقل.

وقد اختارت الباحثة عينة عشوائية من تلميذات أربع مدارس (اثنتين ذات فاعلية عالية، واثنتين ذات فاعلية أقل) وبلغ حجمها ٢٤٠ تلميذة بواقع ١٢٠ تلميذة من كل نوع من المدارس. وقد طبقت الباحثة على تلميذات مدارس المجموعتين مقياس "دافعية الإنجاز" ومقياس "الاتجاه نحو المدرسة" كما حصلت على درجات تلميذات العينة في اختبار نهاية الفصل الدراسي الأول للعام ٢٠٠٠/١٩٩٩ وهو اختبار موحد على مستوى المديرية التعليمية. وبيّن ملف Effect.sav (على الأسطوانة المرنة) البيانات التي حصلت عليها الباحثة من تطبيق الاختبارات الثلاث. أدخل هذه البيانات في محرر البيانات في SPSS. لاحظ أن الأسماء التالية أعطيت للمتغيرات: المجموعة **group** - التحصيل الدراسي **achieve** - الاتجاه نحو المدرسة **attitude** - دافعية الإنجاز **.motive**

طريقة التأشير والضغط:

- ١- اضغط على **Statistics** (الإصدار الثامن) أو **Analyze** (الإصدار التاسع وما بعده).
- ٢- اضغط على **Classify** ثم على **Discriminant** وسوف يظهر مربع الحوار الموضح بشكل ١٣-١.
- ٣- اضغط على **group** ثم اضغط على السهم الأوسط لنقل المتغير إلى مربع **.Grouping Variable**.



شكل ١٣-١ مربع حوار التحليل التمييزي

٤- اضغط على Define Range التي تنشيط بعد نقل المتغير **group**.

٥- اكتب ١ أمام Minimum و ٢ أمام Maximum.

٦- اضغط على **Continue**.

٧- اضغط على مفتاح Ctrl واستمر في الضغط أثناء الضغط على المتغيرات **achieve** و **attitude** و **motive** ثم اضغط على السهم الأوسط لنقل هذه المتغيرات إلى مربع Independents.

٨- اضغط على **Statistics** وسوف يظهر مربع حوار الإحصاء المبين في شكل ١٣-٢.

٩- اضغط على **Box's M - Univariate ANOVAs - Means** في مربع **Descriptives**.

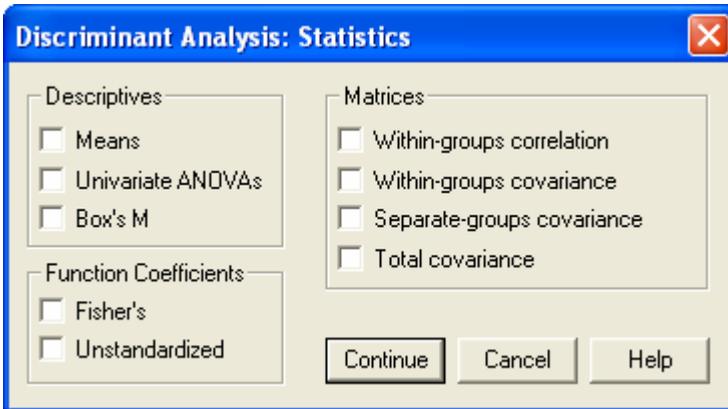
١٠- اضغط على **Unstandardized - Fisher's** في مربع **Function Coefficients**.

١١- اضغط على الاختيارات الأربعة جميعا في مربع **Matrices**.

١٢- اضغط على **Continue**.

١٣- اضغط على **Classify** وسوف يظهر مربع حوار **Discriminant Analysis Classification** (شكل ١٣-٣).

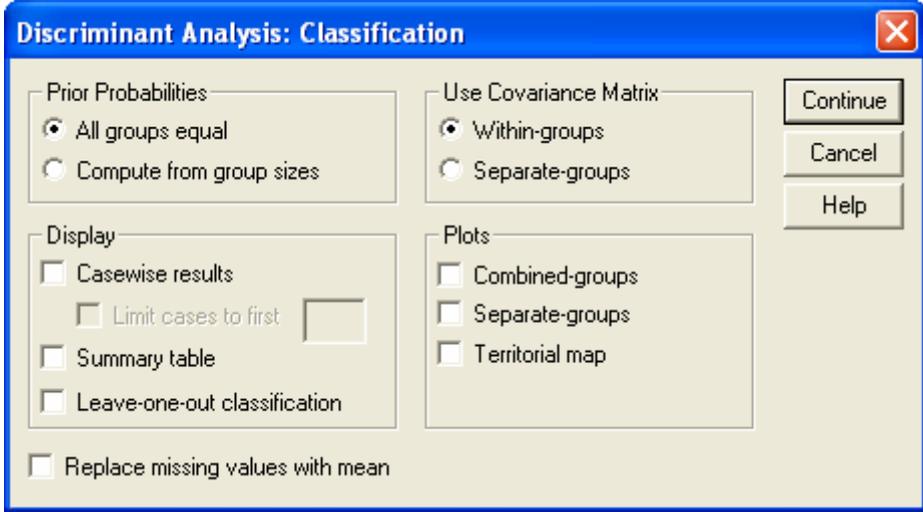
١٤- اضغط على **Compute from Group Sizes** في مربع **Prior Probabilities** إذا كان الحجم النسبي للعينة في المجموعات تقديرات لنسب المجتمع. أما إذا كان نسب مجموعات المجتمع متساوية اختر **All groups equal**.



شكل ١٣-٢ مربع حوار الإحصاء

١٤- اضغط على **Combined Groups** و **Separate Groups** في مربع الرسوم Plots.

١٥- اضغط على **Within Groups** في مربع **Use Covariance Matrix**، إلا أنه إذا كنت لا تستطيع التسليم بتساوي تغايرات جميع المجموعات اضغط على **Separate groups**.



شكل ١٣-٣ مربع حوار التصنيف في التحليل التمييزي

١٦- اضغط على **Summary table** و **Leave-one-out classification** في مربع Display.

١٧- اضغط على **Continue**.

١٨- اضغط على **Save**. وسوف يظهر مربع حوار **Discriminant Analysis: Save New Variables** (شكل ١٣-٤).

١٩- اضغط على **Predicted Group Membership**.

٢٠- اضغط على **Continue**، ثم اضغط على **OK**.

الطريقة اللغوية:

افتح المحرر اللغوي **Syntax Editor** واكتب الأمر التالي (ولا تنسى النقطة في نهاية الأمر) ويمكنك استرجاع ملف **Effect.sps**. اضغط على زر **Run** لتنفيذ التحليل.

DISCRIMINANT

/GROUPS=group(1 2)

/VARIABLES=achiev attitude motive

/ANALYSIS ALL

/SAVE=CLASS

/PRIORS EQUAL

/STATISTICS=MEAN STDDEV UNIVF BOXM COEFF RAW CORR COV GCOV

TCOV TABLE CROSSVALID

/PLOT=COMBINED SEPARATE MAP

/CLASSIFY=NONMISSING POOLED .



شكل ١٣-٤ مربع حوار الحفظ في التحليل التمييزي

نتائج التحليل التمييزي:

كما يتبين من شكل (١٣-٥) تعطي النتائج عددا من الإحصاءات الأولية. إذ نجد المتوسطات والانحرافات المعيارية للمنبتات داخل المجموعتين. كما نجد اختبارا لتحليل التباين لتقويم الفروق بين متوسطات المنبتات الثلاثة. وهناك أيضا مصفوفة التغايرات للمجموعتين، ومصفوفة لاختبار تساوي التغاير داخل المجموعات.

وتشير هذه الإحصاءات إلى فروق دالة إحصائية بين متوسطات المنبتات في المجموعتين (حيث تتراوح قيم 'ل' بين ٠,٠١ و ٠,٢٣).

Group Statistics

GROUP		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
1	ACHIEVE	102.49	16.976	120	120.000
	ATTITUDE	272.42	37.963	120	120.000
	MOTIVE	215.46	21.889	120	120.000
2	ACHIEVE	84.55	20.310	120	120.000
	ATTITUDE	248.07	37.967	120	120.000
	MOTIVE	209.40	19.096	120	120.000
Total	ACHIEVE	93.52	20.729	240	240.000
	ATTITUDE	260.24	39.801	240	240.000
	MOTIVE	212.43	20.721	240	240.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
ACHIEVE	.812	55.130	1	238	.000
ATTITUDE	.906	24.683	1	238	.000
MOTIVE	.979	5.220	1	238	.023

شكل ١٣-٥ الإحصاءات الأولية للتحليل التمييزي

ويشير شكل ١٣-٦ أنه لا توجد فروق دالة إحصائية في مصفوفة التباين للمجموعتين ($\alpha = 0.05$) في اختبار Box's M.

اختبار الدلالة وقوة العلاقة:

يبين شكل (١٣-٧) اختبارات الدلالة وقوة العلاقة للتحليل التمييزي. وفي المربع المعنون Wilk's Lambda اختبار مربع كاي ويحدد هذا الاختبار ما إذا كانت هناك فروق دالة إحصائية بين المجموعتين في المتغيرات المنبئة، بعد استبعاد أي أثر لدوال التمييز السابقة. وفي مثالنا الحالي نجد أن اختبار ويلكس لا يبلغ ٧٨١، وتبلغ قيمة

مربع كاي ٦٤,٦٨٩ ول = ٠,٠٠١. وهذا الاختبار دال عند مستوى ٠,٠٥, ويشير إلى وجود فروق بين المجموعتين في المتغيرات المنبئة الثلاثة في المجتمع.

Covariance Matrices^a

GROUP		ACHIEVE	ATTITUDE	MOTIVE
1	ACHIEVE	288.193	82.772	10.063
	ATTITUDE	82.772	1441.153	321.522
	MOTIVE	10.063	321.522	479.124
2	ACHIEVE	412.489	22.392	54.967
	ATTITUDE	22.392	1441.458	373.780
	MOTIVE	54.967	373.780	364.662
Total	ACHIEVE	429.688	162.039	59.667
	ATTITUDE	162.039	1584.125	383.231
	MOTIVE	59.667	383.231	429.342

- a. The total covariance matrix has 239 degrees of freedom.

Test Results

Box's M		12.027
F	Approx.	1.977
	df1	6
	df2	410401.8
	Sig.	.065

Tests null hypothesis of equal population covariance matrices.

شكل ١٣-٦ اختبار تساوي التباير في المجتمع

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.315 ^a	100.0	100.0	.489

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.761	64.689	3	.000

شكل ١٣-٧ اختبارات الدلالة وإحصاءات قوة العلاقة للتحليل التمييزي

وهناك اختبار آخر مرتبط بدالة التمييز في الجدول المعنون الجذور الكامنة Eigenvalues. وإذا كان لدينا أكثر من دالة للتمييز فإن هذا الاختبار يدلنا على أي دوال التمييز يجب تفسيرها. وفي اختبارنا الحالي ليس لدينا سوى دالة واحدة للتمييز حيث أن لدينا مجموعتين (نم = ١ - ١). ودالة التمييز للجذر الكامن تبلغ ٣١٥، كما أن الارتباط القانوني يبلغ ٤٨٩، ومربع هذا الارتباط يبلغ حوالي ٢٣٩، وهذه هي قيمة مربع أينتا التي نحصل عليها عندما نجري تحليل التباين الأحادي. وبمعنى آخر فإن حوالي ٢٤٪ من تباين الدرجات يرجع إلى الفروق بين المجموعتين.

معاملات دوال التمييز:

يبين شكل ١٣-٨ معاملات دوال التمييز. ويمكن إعطاء مسمى لدالة التمييز وفقا للنتائج وذلك بتحديد أكثرها ارتباطا بالدالة. ولذلك فإننا نفحص حجم المعاملات المعيارية للمتغيرات المنبئة في الدالة، ومعامل الارتباط بين المتغيرات المنبئة والدالة داخل المجموعة (معاملات مصفوفة البنية). وبالنسبة لدالة التمييز في مثالنا الحالي نجد أن لمتغير التحصيل أكبر قيمة إذ بلغت ٨٢٢، في الدالة المعيارية، و٨٥٨، في مصفوفة بنية الدالة. وعلى أساس هذه النتيجة فإننا سوف نطلق على دالة التمييز في هذه الدراسة "التحصيل الدراسي".

Standardized Canonical Discriminant Function Coefficients

	Function
	1
ACHIEVE	.822
ATTITUDE	.532
MOTIVE	-.043

Structure Matrix

	Function
	1
ACHIEVE	.858
ATTITUDE	.574
MOTIVE	.264

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

شكل ٨-١٣ المعاملات المعيارية والارتباطات بين المجموعات للتحليل التمييزي

دوال تمرکز المجموعة:

يبين شكل (٩-١٣) دوال تمرکز المجموعة. ويقصد بها متوسط قيم دوال التمييز في التحليل التمييزي للمجموعتين. ومنها يتبين أن المجموعتين تقعان موقعا معاكسا من بعضهما البعض مما يعزز أن المجموعة الأولى هي الأعلى فاعلية.

Functions at Group Centroids

	Function
GROUP	1
1	.559
2	-.559

Unstandardized canonical discriminant functions evaluated at group means

شكل ٩-١٣ دوال تمرکز المجموعة

Classification Results^{b,c}

		GROUP	Predicted Group Membership		Total
			1	2	
Original	Count	1	88	32	120
		2	43	77	120
	%	1	73.3	26.7	100.0
		2	35.8	64.2	100.0
Cross-validated ^a	Count	1	84	36	120
		2	43	77	120
	%	1	70.0	30.0	100.0
		2	35.8	64.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
- b. 68.8% of original grouped cases correctly classified.
- c. 67.1% of cross-validated grouped cases correctly classified.

شكل ١٣-١٠ تصنيف المجموعتين في التحليل التمييزي

نتائج التصنيف:

يبين شكل (١٣-١٠) نتائج التصنيف. وتدلنا هذه النتائج على جودة التنبؤ بعضوية الجماعة باستخدام التحليل التمييزي. والجزء العلوي من الجدول والمعنون "Original" يحدد مدى جودة دالة التمييز وتظهر الحالات المصنفة تصنيفا صحيحا في قطر الجدول. مثال ذلك نجد أن ٨٨ (٧٣,٣٪) من المجموعة ذات الفاعلية الأعلى (المجموعة الأولى) قد صنفت تصنيفا صحيحا. وبالنسبة للمجموعة الثانية (المجموعة ذات الفاعلية الأقل) نجد أن هناك ٧٧ (٦٤,٢٪) تلميذة قد صنفت تصنيفا صحيحا. أي أن المجموع الكلي المصنف تصنيفا صحيحا هو ١٧٦ (٦٨,٨٪) من العينة الكلية (٢٤٠ تلميذة).

أما الجزء السفلي من الجدول والمعنون "Cross-validated" فقد بني باستخدام أحد الاختيارات في مربع حوار التصنيف وهو "leave-one-out"، ويترتب على هذا الاختيار تصنيف جميع الحالات باستثناء حالة تترك دون تصنيف، ثم تصنف الحالة

المتروكة. وتكرر هذه العملية حتى يتم ترك جميع الحالات مرة واحدة، ويكون التصنيف على أساس ن - ١ من الحالات. ويبين النصف السفلي من الجدول مدى حسن التصنيف القائم على ترك حالة في كل مرة. ويمكن استخدام هذه النتائج لتقدير مدى جودة التصنيف باستخدام جميع الحالات إذا اخترنا عينة جديدة.

حساب كابا Kappa:

تبلغ نسبة الحالات المصنفة تصنيفا صحيحا في العينة حوالي ٦٨,٨٪. وتتأثر هذه النسبة بعامل الصدفة. وتعتبر $Kappa (k)$ مؤشرا لتصحيح عامل الصدفة. ويمكن الحصول على نتائج مؤشر التصحيح هذا مع نسبة الأفراد المصنفين تصنيفا صحيحا.

وسوف نقوم بحساب كابا لتقدير دقة التنبؤ بعضوية الجماعة. ولحساب كابا سوف نستخدم المتغير **dis_1** الذي يمكن إنشاؤه باختيار التنبؤ بعضوية المجموعة Predicted Group Membership في مربع حوار Discriminant Analysis: Save. ولحساب كابا نقوم بالخطوات التالية:

طريقة التأشير والضغط:

- ١- اضغط على **Statistics** (الإصدار الثامن) أو **Analyze** (الإصدار التاسع والإصدارات التالية).
- ٢- اضغط على **Descriptive Statistics**، ثم اضغط على **Crosstabs**.
- ٣- اضغط على **group** ثم اضغط على السهم الأوسط لنقل هذا المتغير إلى المربع المعنون **Row(s)**.
- ٤- اضغط على **dis_1** ثم على السهم الأوسط لنقله إلى المربع المعنون **Column(s)**.
- ٥- اضغط على **Statistics**.
- ٦- اضغط على **Kappa**.
- ٧- اضغط على **Continue**.
- ٨- اضغط على **cells**.
- ٩- اضغط على **Expected** في مربع **Counts**. وتأكد من اختيار **Observed**.
- ١٠- اضغط على **Row** في مربع **Percentage**.
- ١١- اضغط على **Continue**.
- ١٢- اضغط على **OK**.

الطريقة اللغوية:

افتح المحرر اللغوي Syntax Editor واكتب الأمر التالي (ولا تنسى النقطة في نهاية الأمر) ويمكنك استرجاع ملف Effect2. اضغط على زر Run لتنفيذ التحليل.

CROSSTABS

```
/TABLES=group BY dis_1  
/FORMAT= AVALUE TABLES  
/STATISTIC=KAPPA  
/CELLS= COUNT EXPECTED ROW .
```

نتائج تحليل كبا:

Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement Kappa	.375	.060	5.834	.000
N of Valid Cases	240			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

شكل ١٣-١١ نتائج تحليل كبا

كما يتبين من شكل ١٣-١١ تبلغ قيمة كبا ٣٧٥. وتشير هذه القيمة إلى تنبؤ معتدل الدقة. وتتراوح قيمة كبا بين -١ و +١. وتدل القيمة ١ على تنبؤ تام بينما تدل القيمة صفر على تنبؤ على مستوى الصدفة. كما تدل القيمة السالبة على تنبؤ أسوأ من التنبؤ الذي يرجع إلى الصدفة. وكلما بعدت القيمة عن الصفر كان هذا مؤشرا بتنبؤ أفضل من الصدفة.