

تحليل التباين (ANOVA)

تمهيد:

يستخدم تحليل التباين (ANOVA) كأحد الأساليب الإحصائية الاستدلالية في البحوث السوسولوجية والديموغرافية، بغرض اختبار دلالة الفروق بين متوسطات عدة مجموعات، عند دراسة تأثير متغيرات اجتماعية أو ديموغرافية مختلفة على ظاهرة معينة. ويعتمد هذا الأسلوب عندما يكون المتغير التابع كمي وموز توزيعاً طبيعياً، في حين يصنف المتغير المستقل ضمن المتغيرات النوعية متعددة المستويات.

ويستند تحليل التباين إلى مبدأ تفكيك التباين الكلي للبيانات إلى تباين يعزى إلى الاختلاف بين المجموعات المدروسة، وتباين آخر ناتج عن الفروق داخل كل مجموعة. ويتم تقييم دلالة هذه الفروق من خلال مقارنة هذين المصدرين للتباين باستخدام إحصاء (F)، الذي يسمح بالحكم على مدى تأثير المتغير المستقل في المتغير التابع عند مستوى معنوية محدد.

وتصاغ الفرضية الصفرية (H_0) في تحليل التباين على أساس عدم وجود فروق ذات دلالة إحصائية بين متوسطات المجموعات، بينما تفترض الفرضية البديلة (H_1) وجود فروق معنوية بين هذه المتوسطات، بما يعكس تأثير العوامل السوسيو-ديموغرافية محل الدراسة. ويشترط لتطبيق هذا الاختبار توفر مجموعة من الافتراضات الإحصائية، من أهمها: استقلالية الملاحظات، والتقارب من التوزيع الطبيعي للمتغير التابع داخل كل فئة، إضافة إلى تجانس التباين بين المجموعات.

وتكمن أهمية تحليل التباين في كونه يسمح بدراسة الفروق بين عدة فئات اجتماعية أو ديموغرافية في إطار تحليلي موحد، دون اللجوء إلى تعدد الاختبارات الثنائية، وهو ما يساهم في رفع دقة النتائج وتقليل احتمال الخطأ الإحصائي. لذلك يعد تحليل التباين من الأدوات المنهجية الأساسية في الدراسات التي تتناول التفاوتات الاجتماعية، المجالية أو السكانية، وتحليل محددات الظواهر الديموغرافية المختلفة.

1. تعريف تحليل التباين:

تحليل التباين هو أسلوب إحصائي استدلالي يهدف إلى اختبار دلالة الفروق بين متوسطات أكثر من مجموعتين، من خلال مقارنة التباين الناتج عن اختلاف المجموعات بالتباين داخل كل مجموعة، وذلك للتحقق مما إذا كانت هذه الفروق تعكس أثراً حقيقياً للمتغير المستقل أم أنها تعود إلى التباين العشوائي.

2. أنواع تحليل التباين:

تختلف أنواع تحليل التباين تبعا لطبيعة التصميم البحثي وعدد المتغيرات المستقلة والمتغيرات التابعة، ويتيح كل نوع امكانية معالجة وضعيات تحليلية متنوعة في الدراسات الاجتماعية، ومن أهم أنواعه نجد:

2.1. تحليل التباين الأحادي:

يستخدم تحليل التباين الأحادي عندما يكون هناك متغير مستقل واحد ذو عدة مستويات، ويهدف الى اختبار دلالة الفروق بين متوسطات المجموعات الناتجة عنه. ويعد هذا النوع مناسباً لدراسة أثر عامل تنظيمي واحد على متغير كمي.

مثال:

دراسة الفروق في متوسط عدد ساعات التكوين المهني السنوي بين العمال حسب نوع المؤسسة (عمومية، خاصة، مختلطة).

2.2. تحليل التباين الثنائي:

يطبق تحليل التباين الثنائي عندما يشمل التصميم البحثي متغيرين مستقلين، مما يسمح بدراسة اثر كل متغير على حدة، اضافة الى اثر التفاعل بينهما على المتغير التابع. ويساعد هذا النوع على فهم التداخل بين العوامل التنظيمية والاجتماعية.

مثال تطبيقي:

دراسة اثر كل من المستوى التعليمي (ثانوي، جامعي، دراسات عليا) ونوع المنصب (تنفيذي، اشرافي، اداري) والتفاعل بينهما على متوسط الاجر الشهري.

2.3. تحليل التباين المتعدد:

يستخدم تحليل التباين المتعدد عندما يكون هناك أكثر من متغير تابع كمي، ويتم تحليلها في وقت واحد لدراسة اثر المتغيرات المستقلة عليها مجتمعة. ويعد هذا النوع ملائماً للدراسات التي تتناول عدة ابعاد للعمل والتنظيم.

مثال:

دراسة اثر نمط التنظيم الداخلي للمؤسسة (مركزي، لا مركزي) على كل من عدد ايام الغياب عن العمل، ومتوسط ساعات العمل الاضافي، ومعدل الحوادث المهنية.

2.4. تحليل التباين للقياسات المتكررة

يستعمل هذا النوع عندما يتم قياس نفس المتغير التابع عدة مرات على نفس الافراد او الوحدات، سواء عبر الزمن او في ظروف تنظيمية مختلفة. ويستخدم بكثرة في الدراسات الطولية في علم الاجتماع التنظيمي.

مثال:

تحليل تطور متوسط انتاجية العمال قبل تطبيق نظام تقييم اداء جديد، وبعد ستة اشهر، وبعد سنة من تطبيقه داخل نفس المؤسسة.

2.5. تحليل التباين المشترك

يطبق تحليل التباين المشترك عندما يراد دراسة أثر متغير مستقل نوعي على متغير تابع كمي مع التحكم في أثر متغير كمي اخر قد يؤثر في النتائج. ويسمح هذا النوع بعزل اثر بعض العوامل الدخيلة في التحليل التنظيمي.

مثال:

دراسة اثر نوع عقد العمل (دائم، مؤقت) على متوسط عدد ايام الغياب عن العمل، مع التحكم في اثر سنوات الاقدمية المهنية

ملاحظة: يجب التنويه إلى أن دلالة اختبار تحليل التباين لا تحدد موضع الفروق بين المجموعات، مما يستدعي في بعض الحالات اللجوء إلى اختبارات المقارنات البعدية لتحديد الفئات التي تظهر بينها الفروق المعنوية.

3. افتراضات وفروض اختبار تحليل التباين:

3.1. الافتراضات:

- أن المجتمع الذي تسحب منه المجموعات يكون موزعا توزيعا طبيعيا.
- أن التباينات بين المجتمعات التي تسحب منها العينات تكون متجانسة (أي لا توجد فروق ذات دلالة

$$H_0: s_1^2 = s_2^2 = s_3^2 = s_4^2 \quad \text{احصائية:}$$

3.2. الفروض:

يقوم اختبار تحليل التباين على صياغة فرضيتين إحصائيتين أساسيتين هما الفرضية الصفرية والفرضية البديلة، وذلك

بهدف الحكم على دلالة الفروق بين متوسطات المجموعات المدروسة

- الفرضية الصفرية (H_0): تنص الفرضية الصفرية على عدم وجود فروق ذات دلالة احصائية بين متوسطات المجموعات، أي أن جميع المتوسطات متساوية احصائيا $H_0: \mu_1 = \mu_2 = \mu_3 \dots \dots \mu_k$
- الفرضية البديلة (H_1): تنص الفرضية البديلة على وجود فروق ذات دلالة احصائية بين متوسطات المجموعات، أي أنه على الأقل متوسطا واحدا يختلف عن باقي المتوسطات: $H_1: \mu_1 \neq \mu_2 \neq \mu_3 \dots \dots \mu_k$

4. خطوات تحليل التباين:

4.1. الخطوات المنهجية:

يتم تطبيق تحليل التباين وفق مجموعة من الخطوات المنهجية المتسلسلة، تسمح باختبار دلالة الفروق بين متوسطات المجموعات المدروسة بطريقة علمية دقيقة.

- **الخطوة الاولى: تحديد مشكلة الدراسة والمتغيرات**
يتم في هذه المرحلة تحديد المتغير التابع الكمي المراد دراسته، والمتغير المستقل النوعي ذي المستويات المتعددة، مع توضيح طبيعة العلاقة المراد اختبارها بينهما.
- **الخطوة الثانية: صياغة الفروض الاحصائية**
يقوم الباحث بصياغة الفرضية الصفرية التي تنص على تساوي متوسطات المجموعات، والفرضية البديلة التي تشير إلى وجود فروق ذات دلالة احصائية بين هذه المتوسطات.
- **الخطوة الثالثة: التحقق من شروط التطبيق**
قبل إجراء تحليل التباين، يجب التأكد من توفر الشروط الاحصائية اللازمة، والمتمثلة في استقلالية المشاهدات، والتقارب من التوزيع الطبيعي للمتغير التابع داخل كل مجموعة، وتجانس التباين بين المجموعات.
- **الخطوة الرابعة: حساب مكونات التباين**
يتم حساب التباين الكلي للبيانات، ثم تفكيكه إلى تباين بين المجموعات وتباين داخل المجموعات، مع تحديد درجات الحرية الخاصة بكل مصدر من مصادر التباين.
- **الخطوة الخامسة: حساب إحصاء F**
يتم حساب قيمة إحصاء F من خلال قسمة متوسط التباين بين المجموعات على متوسط التباين داخل المجموعات، وذلك لتقييم مدى دلالة الفروق بين المتوسطات.
- **الخطوة السادسة: اتخاذ القرار الاحصائي**
تتم مقارنة القيمة المحسوبة لإحصاء F بالقيمة الجدولية عند مستوى معنوية معين، أو الاعتماد على قيمة الدلالة الاحصائية (Sig). وبناء على ذلك يتم قبول أو رفض الفرضية الصفرية.

• الخطوة السابعة: اجراء المقارنات البعدية عند الضرورة

في حال رفض الفرضية الصفرية وإثبات وجود فروق ذات دلالة احصائية، يتم اللجوء إلى اختبارات المقارنات البعدية لتحديد المجموعات التي توجد بينها هذه الفروق.

• الخطوة الثامنة: تفسير النتائج

يتم تفسير النتائج إحصائيا ومنهجيا في ضوء إشكالية الدراسة وفرضياتها، وربطها بالإطار النظري والدراسات السابقة في مجال علم الاجتماع والتنظيم والعمل.

4.2. الخطوات الحسابية:

لأجراء التحليل نتبع الخطوات التالية:

1- يتم استخراج مجموع مربعات التباين الكلي

$$SST = \sum_i \sum_j x_{ij}^2 - \frac{(\sum_i \sum_j x_{ij})^2}{N}$$

حيث:

SST: مجموع المربعات الكلي

$\sum_i$: مجموع عبر المجموعات (الفئات أو المعالجات)

$\sum_j$: مجموع عبر المفردات داخل كل مجموعة

x_{ij} : قيمة المتغير المدروس للمفردة رقم j في المجموعة رقم i

x_{ij}^2 : مربع قيمة المشاهدات

$\sum_i \sum_j x_{ij}$: المجموع الكلي لجميع القيم

$\sum_i \sum_j x_{ij}^2$: مجموع مربعات جميع القيم المشاهدة

$(\sum_i \sum_j x_{ij})^2$: مربع المجموع الكلي

ويمثل هذا القانون إجمالي التشتت الكلي للبيانات حول المتوسط العام في تحليل التباين.

2- يتم استخراج مجموع مربعات بين المجموعات

$$SSB = \sum_i \frac{(\sum_j x_{ij})^2}{n_i} - \frac{(\sum_i \sum_j x_{ij})^2}{N}$$

3- مجموع مربعات داخل المجموعات

$$SSW = \sum_i \sum_j x_{ij}^2 - \sum_i \frac{(\sum_j x_{ij})^2}{n_i}$$

مصدر التباين S.V	مجموع المربعات S.S	درجة الحرية d.f	متوسط المربعات S.M	القيمة الفائية المحسوبة F
B بين المجموعات	$\sum_i \frac{(\sum_j x_{ij})^2}{n_i} - \frac{(\sum_i \sum_j x_{ij})^2}{N}$	K-1 عدد المجموعات	$\frac{\sum S.SB}{K-1}$	$\frac{S.M.B}{S.M.W}$
W داخل المجموعات	$\sum_i \sum_j x_{ij}^2 - \sum_i \frac{(\sum_j x_{ij})^2}{n_i}$	N-K عدد المجموعات ناقص عدد الأفراد الكلية	$\frac{\sum S.SW}{N-K}$	
T مجموع المجموعات الكلية	$\sum_i \sum_j x_{ij}^2 - \frac{(\sum_i \sum_j x_{ij})^2}{N}$	N-1		

مثال عددي:

أراد باحث معرفة ما إذا كان متوسط عدد أيام الغياب يختلف حسب نوع المؤسسة. جمع الباحث البيانات التالية (أيام غياب خلال السنة)

- مؤسسة صناعية: 6، 8، 5، 7
 - مؤسسة إدارية: 3، 4، 5، 2
 - مؤسسة خدمائية: 9، 7، 8، 10.
- عدد المجموعات (k) = 3

عدد الأفراد الكلي (N) = 12

1. حساب مجموع مربعات التباين الكلي:

$$SST = \sum_i \sum_j x_{ij}^2 - \frac{(\sum_i \sum_j x_{ij})^2}{N}$$

أ- حساب مجموع مربعات القيم:

- مؤسسة صناعية: $6^2 + 8^2 + 5^2 + 7^2 = 36 + 64 + 25 + 49 = 174$
- مؤسسة إدارية: $3^2 + 4^2 + 5^2 + 2^2 = 9 + 16 + 25 + 4 = 54$
- مؤسسة خدماتية: $9^2 + 7^2 + 8^2 + 10^2 = 81 + 49 + 64 + 100 = 294$

المجموع:

$$\sum_i \sum_j x_{ij}^2 = 174 + 54 + 294 = 522$$

ب- حساب مربع مجموع القيم على N:

$$6 + 8 + 5 + 7 + 3 + 4 + 5 + 2 + 9 + 7 + 8 + 10 = 74$$

$$\frac{(\sum x_{ij})^2}{N} = \frac{74^2}{12} = \frac{5476}{12} = 456.33$$

ت- حساب SST:

$$SST = 522 - 456.67 = 65.67$$

2. حساب التباين بين المجموعات (SSB):

$$SSB = \sum_i \frac{(\sum_j x_{ij})^2}{n_i} - \frac{(\sum_i \sum_j x_{ij})^2}{N}$$

أ) حساب مجموع اليم لكل مجموعة:

- صناعية: $26 = 7+5+8+6$
- إدارية: $14 = 2+5+4+3$
- خدماتية: $34 = 10+8+7+9$

ب) قسمة مربع مجموع كل مجموعة على حجم المجموعة: ($n_i=4$)

• صناعية: $\frac{(26)^2}{4} = \frac{676}{4} = 169$

• إدارية: $\frac{(14)^2}{4} = \frac{196}{4} = 49$

• خدماتية: $\frac{(34)^2}{4} = \frac{1156}{4} = 289$

مجموعها: $(\sum_i \frac{(\sum_j x_{ij})^2}{n_i})$

$.507 = 289+49+169$

لدينا أيضا $\frac{(\sum_i \sum_j x_{ij})^2}{N} = 456.33$

إذن:

$$SSB = \sum_i \frac{(\sum_j x_{ij})^2}{n_i} - \frac{(\sum_i \sum_j x_{ij})^2}{N} = 507 - 456.33 = 50.67$$

3. حساب التباين داخل المجموعات:

$$SSW = \sum_i \sum_j x_{ij}^2 - \sum_i \frac{(\sum_j x_{ij})^2}{n_i}$$

لدينا:

$$\sum_i \sum_j x_{ij}^2 = 522$$

$$\sum_i \frac{(\sum_j x_{ij})^2}{n_i} = 507$$

إذن:

$$SSW = 522 - 507 = 15$$

4. حساب متوسط المربعات (SMB و SMW):

لدينا:

- درجة الحرية بين المجموعات (درجات الحرية الأفقية) $2 = 1 - 3 = k - 1$
 - درجة الحرية داخل المجموعات (درجات الحرية العمودية) $9 = 3 - 12 = N - k$
- أ) حساب متوسط المربعات بين المجموعات:

$$SMB = \frac{SSB}{k - 1} = \frac{50.67}{2} = 25.335$$

ب) حساب متوسط المربعات داخل المجموعات:

$$SMW = \frac{SSW}{N - k} = \frac{15}{9} = 1.67$$

ت) حساب قيمة F:

$$F = \frac{SMB}{SMW} = \frac{25.335}{1.67} = 15.17$$

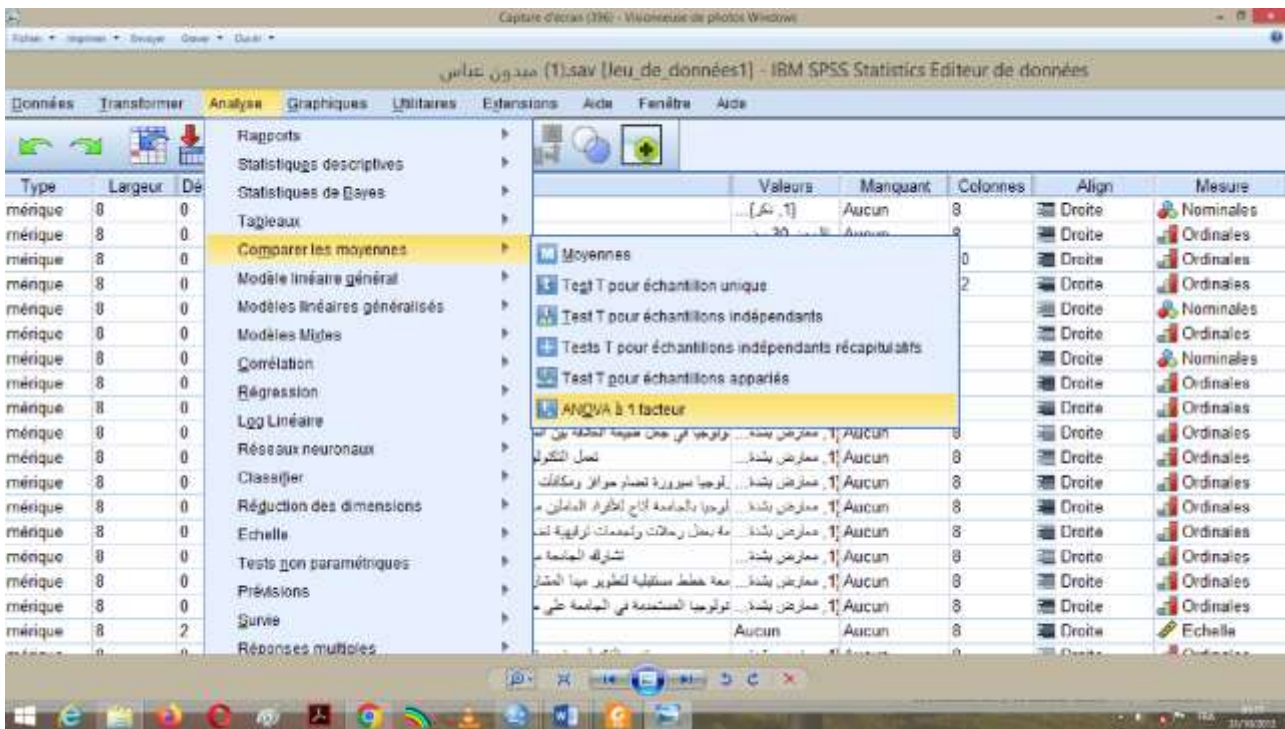
نقوم باستخراج القيمة الفائية من الجدول عند مستوى الدلالة = 0.05، درجة حرية أفقية = 2، ودرجة حرية عمودية = 9، والتي تساوي **4.26**

القيمة الفائية الجدولية = 4.26. بما أن القيمة الفائية المحسوبة أكبر من القيمة الفائية الجدولية

$(F_c(15.17) > F_t(4.26))$ نرفض الفرض الصفري ونقبل الفرض البديل الذي مفاده بأنه توجد فروق ذات دلالة إحصائية بين المؤسسات الثلاثة. وبالتالي يجب القيام باختبار المقارنات البعدية (post-Hoc) باستخدام شيفيه أو توكي.

4.3. حساب تحليل التباين الأحادي باستخدام SPSS:

من القائمة Analyse-- comparer les moyennes—ANOVA à 1 facteur



ضع المتغير التابع في liste variable dépendantes et le facteur en facteur



تظهر لنا النتائج كالتالي:

Test d'homogénéité des variances

		Statistique de Levene	ddl1	ddl2	Sig.
المشاركة_تحقيق	Basé sur la moyenne	1,054	2	187	,351
	Basé sur la médiane	,990	2	187	,373
	Basé sur la médiane avec ddl ajusté	,990	2	185,992	,373
	Basé sur la moyenne tronquée	1,043	2	187	,354

ANOVA

المشاركة_تحقيق

	Somme des carrés	ddl	Carré moyen	F	Sig.
Intergruppes	,866	3	,289	,621	,602
Intragruppes	86,881	187	,465		
Total	87,747	190			

- ظهر اختبار تجانس التباين (Levene) أن الفروق في التباينات بين المجموعات ليست دالة ($\text{Sig} = 0.351 > 0.05$)، وهو ما يسمح باستخدام تحليل التباين الأحادي.
- كما بينت نتائج ANOVA عدم وجود فروق ذات دلالة إحصائية في متوسط تحقيق المشاركة بين المجموعات

$$\text{Sig} = 0.602, (F(3,187) = 0.621)$$

مثال آخر:

Test d'homogénéité des variances

		Statistique de Levene	ddl1	ddl2	Sig.
المرور احداث التعرض محور متوسط	Basé sur la moyenne	3,019	2	83	,054
	Basé sur la médiane	2,845	2	83	,064
	Basé sur la médiane avec ddl ajusté	2,845	2	71,323	,065
	Basé sur la moyenne tronquée	3,030	2	83	,054

ANOVA

المرور احداث التعرض محور متوسط

	Somme des carrés	ddl	Carré moyen	F	Sig.
Intergruppes	1,632	2	,816	3,998	,022
Intragruppes	16,946	83	,204		
Total	18,579	85			

- اظهر اختبار Levene أن الفروق في التباينات بين المجموعات ليست ذات دلالة (Sig = 0.054 > 0.05)، مما يسمح باعتماد تحليل التباين الأحادي.
- كما بينت نتائج ANOVA وجود فروق ذات دلالة إحصائية في متوسط التعرض لحوادث المرور بين المجموعات
Sig = 0.022). (F(2,83) = 3.998
ولتحديد اتجاه هذه الفروق تم اللجوء إلى اختبارات المقارنات البعدية

الاختبار البعدي (post-hoc)

Comparaisons multiples :

Variable dépendante: المرور احداث التعرض محور متوسط

Différence significative de Tukey

المؤسسة في الوظيفة (I)	المؤسسة في الوظيفة (J)	Différence moyenne (I-J)	Erreur standard	Sig.	Intervalle de confiance à 95 %	
					Borne inférieure	Borne supérieure
عامل	(فريق أو مصلحة) رئيس	,22844	,10978	,100	-,0335	,4904
	تقني	,33681*	,13044	,031	,0255	,6481
(فريق أو مصلحة) رئيس	عامل	-,22844	,10978	,100	-,4904	,0335
	تقني	,10836	,13313	,695	-,2093	,4261
تقني	عامل	-,33681*	,13044	,031	-,6481	-,0255
	(فريق أو مصلحة) رئيس	-,10836	,13313	,695	-,4261	,2093

*. La différence moyenne est significative au niveau 0.05.

- ظهرت نتائج اختبار Tukey وجود فرق دال احصائيا بين فئة العمال وفئة التقنيين (Sig = 0.031)، في حين لم تظهر فروق ذات دلالة بين باقي الفئات الوظيفية.
- ويشير ذلك إلى أن مستوى التعرض لحوادث العمل يختلف أساسا بين العمال والتقنيين، بينما يبقى متقاربا بين رؤساء المصالح/الفرق وبقيّة الفئات.

ويمكنك إضافة تفسير سوسيولوجي حسب سياق الدراسة: طبيعة المهام، ارتباط العمل الميداني بالمخاطر... إلخ