كلية العلوم الاجتماعية- قسم الديموغرافيا

**مادة الإحصاء الرياضي وتطبيق الاختبارات الإحصائية**

**ماستر 1 ديموغرافيا اجتماعية**

د. راشدي خضرة

## المحاضرة 12: تحليل التباين Analyse de la variance

هو مقارنة عدة متوسطات (أكثر من متوسطين) حيث يكون المتغير التابع عدديا و المتغيرات المستقلة اسمية او رتبية . ويشترط ان تتبع البيانات التوزيع الطبيعي و وتستند إجراءات تحليل التباين إلى توزيع احتمالي متصل هو توزيع ف (F-distribution). و يوجد تحليل التباين الاحادي ANOVA و تحليل التباين المتعدد

1- **تحليل التباين في اتجاه واحد :One-way Analysis of Variance** (بافتراض تساوي أحجام المجموعات): هو اختبار الفروق بين متوسطات ثلاثة مجتمعات فأكثر. ويهدف التحليل إلى بيان ما إذا كانت متوسطات المجتمعات متساوية تقريباً وأن أية اختلافات بينهما تعزى للصدفة ويمكن توقعها أم أن المتوسطات مختلفة (غير متساوية) والفروق بينها جوهرية أو ذات دلالة. ويقتصر على متغير تابع و متغير مستقل.

**مثال :** اثر المستوى التعليمي على الخصوبة حيث الخصوبة هي المتغير التابع و المستوى التعليمي متغير مستقل

**شروطه :**

- عينات مستقلة، ويعني أن تحديد أي عنصر في مجموعة من المجموعات لن يؤثر على كيفية اختيار العناصر في المجموعات الأخرى.
- أن يكون توزيع المتغير التابع طبيعياً في المجتمعات التي سحبت منها العينات.

- التجانس في التباين: ويعني ذلك أن يكون الانحراف المعياري للمتغير التابع متساو للمجموعات.
- أن يكون المتغير التابع كمي (نسبي أو فئوي).

**خطواته :**

- يتم حساب الوسط الحسابي والتباين لكل مجموعة .
- بعدها يتم حساب التباين بين المجموعات والتباين داخل المجموعات.
- إذا كان التباين بين المجموعات يقارب التباين داخلها، فهذا يدل على أن العينات تمثل مجتمعاً واحدا. وإذا كان التباين بين المجموعات أكبر من التباين داخلها بدرجة كبيرة فهذا يدل على اختلاف المتوسطات.
- و يكون الاختبار ب:

1. الفرض الصفري : جميع المتوسطات متساوية $H_0 : \mu_1 = \mu_2 = \mu_3 = ... = \mu_k$
2. الفرض البديل : ليس كل المتوسطات متساوية، أو على الأقل واحد من المتوسطات يختلف عن المتوسطات الباقية
3. تحديد مستوى الدلالة ( $\alpha$ ): وتحدد مستويات المعنوية سلفاً وهي عادة 0.05 أو 5%
4. حساب إحصائية الاختبار (F) حيث يتم حسابها بإعداد جدول تحليل التباين التالي:

| مصدر التباين | درجات الحرية | مربع الانحرافات | التباين | قيمةF |
|---|---|---|---|---|
| بين المجموعات | K−1 | SSTR | $MSTR = \frac{SSTR}{K-1}$ | $F = \frac{MSTR}{MSE}$ |
| داخل المجموعات | n−K | SSE | $MSE = \frac{SSE}{N-K}$ | |
| المجموع | n−1 | TSS | | |

5. اتخاذ قرار حول بيانات العينة من خلال مقارنة قيمة إحصائية الاختبار بالقيمة النظرية لتوزيع ف عند مستوى المعنوية المحددة. ويشمل اختبار ف تطبيق معيار التفريق بين ما هو فرق جوهري وما هو ناتج عن الصدفة. فإذا كانت قيمة الاحصاءة F أكبر من القيمة النظرية لتوزيع ف عند درجتي حرية (k−1 وn−k) يتم رفض فرض العدم ويعني هذا أن هناك فروقا ذات دلالة احصائية بين المتوسطات.

حيث K−1 درجات الحرية بين المجموعات و n−K درجات الحرية داخل المجموعات و n عدد المشاهدات و K عدد المجموعات .

**مثال** : البيانات التالية تخص 12 سيدة حسب عمر الزواج و المستوى التعليمي . هل هناك فروق ذات دلالة احصائية في عمر الزواج بين مختلف المستويات التعليمية

| بدون مستوى | ابتدائي | متوسط | ثانوي فاكثر |
|---|---|---|---|
| 19 | 25 | 25 | 28 |
| 20 | 21 | 22 | 25 |
| 19 | 19 | 28 | 29 |

فرض العدم: جميع المتوسطات متساوية $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

الفرض البديل: ليس كل المتوسطات متساوية، أو على الأقل واحد من المتوسطات يختلف عن المتوسطات الباقية.

مستوى الدلالة ( 0.05 أو 5%).

**أولاً حساب مجموع قيم المشاهدات لكل مجموعة**

المجموعة الأولى: 58

المجموعة الثانية: 65

المجموعة الثالثة: 75

المجموعة الرابعة: 82

المجموع الكلي: 280

**ثانياً حساب مجموع مربعات الانحرافات الكلي(TSS)** :حيث

$$TSS = \sum (x - \bar{x})^2 \text{ or } \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$TSS = 19^2+20^2+.....+25^2+30^2 - \frac{280^2}{12} = 6751-6533.33=217.67$$

ثالثاً حساب مجموع مربعات الانحرافات بين المجموعات (SSTR):

$$\sum n_j(\bar{x}_j-\bar{x})^2 \text{ or } \sum\left(\frac{T_j^2}{nj}\right)-\frac{\left(\sum X\right)^2}{n}$$

$$= [\frac{58^2}{3} +\frac{65^2}{3} + \frac{75^2}{3} +\frac{82^2}{3} ] - \frac{280^2}{12}= 6646- 6533.33=112.67$$

**رابعاً حساب مجموع مربعات الانحرافات داخل المجموعات (SSE):**

$$SSE = TSS – SSTR = 217.67 – 112.67 = 105$$

**خامساً جدول تحليل التباين**

| مصدر التباين | درجات الحرية | مربع الانحرافات | التباين | قيمة F |
|---|---|---|---|---|
| بين المجموعات | 3=1-4 | 112.67 | 37.56=3/112.67 | 2.86=13.125/37.56 |
| داخل المجموعات | 8=4-12 | 105 | 13.125=8/105 | |
| المجموع | 11=1-12 | 217.67 | | |

**سادسا: اتخاذ القرار** من خلال مقارنة قيمة إحصائية الاختبار (2.86) بالقيمة النظرية لتوزيع ف بدرجتي حرية (3 ،11) عند مستوى معنوية (0.05). ومن الجدول نجد قيمة توزيع ف عند بدرجتي حرية (3 ،11) يساوي (3.10). وبما أن قيمة F المساوية (2.86) أصغر من قيمة توزيع ف والمساوية (3.58)، فإننا نقبل فرض العدم ونقول أنه لا يوجد فرق ذو دلالة إحصائية عند مستوى معنوية (5%) بين اعمار الزواج حسب المستوى التعليمي .

**تمرين1 : من خلال البيانات التالية نريد اختبار ما اذا كان هناك اختلاف في توزيع عدد الاطفال حسب المستوى المعيشي للأسر .**

| ضعيف | متوسط | جيد |
|---|---|---|
| 5 | 5 | 2 |
| 7 | 4 | 3 |
| 5 | 3 | 2 |
| | 3 | 1 |
| | | 3 |