المحاضرة الخامسة في الاحصاء الاستدلالي: معامل كاف تربيعي (Chi-Sequared)

تعريف معامل كاف تربيعي:

هو مقياس احصائي يرمز له بالحرف اليوناني، اول من اوجده هو كارل بيرسون سنة 1900، لذلك يعرف كذلك ب كاي سكوير بيرسون. يستخدم هذا المعامل للبيانات الفئوية المعبر عنها بتكرارات في جدول التوزيع التكراري.

استخداماته: يعتبر كاي سكوير أحد الاساليب المعروفة والمعتمدة في التحليل الاحصائي، ويستخدم الاختبار لقياس المدى الذي تقترب او تبتعد فيه التكرارات المشاهدة عن التكرارات المتوقعة. فكلما كان هذا المدى كبيرا نرفض فرضية العدم، وتقبل هذه الاخيرة في حالة التطابق. ويستخدم كذلك لاختبار دلالة الفرق بين مجموعين. ويمكن تلخيص استخدامات معامل X^2 في مجالين هما:

1- لاختبار مدى تطابق (goodness-of-fit) التوزيع المتوقع مع التوزيع الحقيقي؛
2- لتحديد وجود علاقة (ارتباط) بين مجموعتين او أكثر استنادا الى صفات معينة بينهما وهو ما يعرف بالتصنيف الثنائي.

يتم حساب قيمة X^2 من خلال المعادلة التالية:

$$X^2 = \sum \frac{(fo - fe)^2}{fe}$$

حيث: fo = التكرارات المشاهدة

fe = التكرارات المتوقعة

إذا كانت قيمة كاي سكوير المحسوبة أكبر من القيمة الجدولية عند مستوى الدلالة ودرجات الحرية المحددة، نرفض الفرض الصفري والذي يقر بعدم وجود اختلاف هام احصائيا بين التوزيع الحقيقي المشاهد وبين التوزيع النظري المتوقع. أي ان التكرارات المشاهدة في هذه الحالة تختلف معنويا عن التكرارات المتوقعة (الفرضية البديلة).

تتحدد درجة الحرية كما يلي: **df= N-1** حيث ان N تمثل عدد المجموعات

وفي حالة التصنيفات الثنائية تحسب درجة الحرية كالتالي: $df= (r-1)(c-1)$ حيث: r عدد الصفوف في الجدول

C عدد الاعمدة في الجدول

ملاحظة:

تكون قيمة كاي سكوير موجبة دائما وتكون منطقة الرفض فيها على جهة واحد فقط، فاختبار كاي تربيع يفترض ان التكرارات المتوقعة تكون كبيرة نسبيا اي لا يصلح للاستخدام إذا كانت احد التكرارات اقل من (5) ويؤدي الى نتائج غير موثوقة.

مثال:

نفترض انه لدينا البيانات التالية تخص معهد علوم وتقنيات النشاطات البدنية لجامعة العربي بن مهيدي، ويريد المعهد ان يعرف ما اذا كان توزيع الطلاب المسجلين في التخصصات المختلفة يتوافق مع التوزيع المتوقع الذي حدده المعهد مسبقا.

المساق	التكرار المتوقع (عدد الطلاب)	التكرار الفعلي (عدد الطلاب المسجلين)
تدريب رياضي	50	45
ادارة وتسيير رياضي	30	35
تربية بدنية	20	22

الحل:

- **تحديد المشكلة:** هل يوجد فرق ذو دلالة احصائية بين التوزيع المتوقع والتوزيع الفعلي لعدد الافراد المسجلين في التخصصات
- **صياغة الفرضيات:** H_0 = لا يوجد فرق ذو دلالة احصائية

  H_1 = يوجد فرق ذو دلالة احصائية

- **تحديد نوع الاختبار:** بما ان مستوى القياس اسمي فإننا نختار معامل كاي سكوير لاختبار مدى التطابق بين التوزيع الفعلي والتوزيع المتوقع
- **حساب معامل كاي سكوير** لتقييم مدى التوافق بين التوزيعين:

$$X^2 = \sum \frac{(fo - fe)^2}{fe}$$

$$\mathbf{X^2 = \ \ 1.53}$$

- **تحديد درجة الحرية:** $df= (3-1)*(2-1)=2$

مستوى الدلالة: $\alpha= 0.05$

- **تحديد القيمة الجدولية لكاي سكوير:** $X^{2\,(t)} = \ \ 5.99$
- **اتخاذ القرار الاحصائي:**

نجد ان قيمة كاي سكوير المحسوبة اصغر من القيمة الحرجة (أي قيمة كاي سكوير الجدولية) التي تساوي 5.99 عند مستوى الدلالة ودرجة الحرية المحددة . وبالتالي نقبل الفرضية الصفرية وهذا يعني عدم وجود فروق ذات دلالة احصائية بين التوزيع المتوقع والتوزيع الفعلي لعدد الافراد المسجلين.