المستوى: طلبة السنة الثانية ماستر تخصص تسويق خدمات أستاذ المقياس: د. لمياء مكرسي

المحاضرة الرابعة عشرة: المعالجة الإحصائية للبيانات

ثالثا: مقاييس التشتت

تعكس هذه المعايير درجة تقارب هذه البيانات أو تباعدها فيما بينها، فالتشتت يكون صغيرا إذا كانت البيانات متقاربة فيما بينها والعكس بالعكس. ومقاييس التشتت كثيرة ويمكن ذكر الأكثر استخداما منها فيما يلي:

أ **المدى**: وهو الفرق بين أكبر وأصغر قيمة في البيانات الإحصائية موضع الدراسة أي أن:

المدى= أكبر قيمة–أصغر قيمة

ب **الانحراف المعياري (التباين)**: وهو أكثر المقاييس التي تستخدم لقياس مدى تفرق أو تناغم البيانات عن متوسطها الحسابي، حيث يتم إدخال كل القيم وحسابها وليس من خلال قيمتين أو ثلاثة فقط، ومن هنا تكمن دقته عن باقي مقاييس التشتت. ويأخذ الصيغة التالية:

$$s = \sqrt{\frac{\sum \left(x_i - \bar{x}\right)}{n-1}}$$

ويتم حسابه وفقا للخطوات التالية:

- **الخطوة الأولى**: يحتسب المتوسط الحسابي للقيم وذلك بجمع القيم وتقسيمها على عددها.
- **الخطوة الثانية**: يحسب مقدار انحراف كل قيمة عن المتوسط الحسابي، وذلك بطرح الوسط الحسابي من كل قيمة.
- **الخطوة الثالثة**: يربع انحراف كل قيمة على حدة، ومن ثم تجمع مربعات انحرافات القيم.
- **الخطوة الرابعة**: يطبق قانون الانحراف المعياري وهو مجموع مربعات انحراف القيم عن المتوسط تقسيم عدد القيم.

مثال: أحسب الانحراف المعياري للقيم التالية:7، 9، 10، 3، 6.

الحل: الوسط الحسابي= 5/35= 7

$\sum \left(x_i - \bar{x}\right)^2$ =(1)+ (16)+ (9)+ (4)+ (0

وباستخدام الصيغة السابقة للانحراف المعياري نجد أن: $S=\sqrt{\frac{30}{4}} = 2.74$

رابعا: مقاييس الارتباط الخطي

الارتباط هو دراسة العلاقة بين متغيرين (ظاهرتين) أو أكثر. إذا كانت العلاقة بين متغيرين يسمى ارتباط بسيط Simple Correlation، أما إذا كانت العلاقة بين أكثر من متغيرين فيسمى ارتباط متعدد Multiple Correlation. وقد عرف هذا المعامل من قبل كارل بيرسون وتكمن أهميته في مساعدة الباحث على التنبؤ بالعلاقة بين المتغير المستقل والتابع.

ويقيس معامل الارتباط قوة الارتباط الخطي بين القيم المرصودة X و Y، وتتراوح قيمة المعامل بين 1 و −1. وتدل القيمة المحسوبة على قوة العلاقة بينما تدل الإشارة على اتجاهها، حيث تدل الإشارة الموجبة على العلاقة الطردية في حين تدل الإشارة السالبة على العلاقة العكسية. فإذا كانت قيمته مساوية ل 1 فهذا يعني وجود علاقة خطية طردية تماما بين المتغيرين. أما إذا كانت قيمته −1 فهذا يدل على وجود علاقة خطية عكسية تماما. وإذا كانت قيمته 0 فهذا يعني لا توجد علاقة بين متغيرات الدراسة.

ومن أهم المقاييس المستخدمة لحساب معامل الارتباط ما يلي:

معامل ارتباط بيرسون: يمتاز هذا المعامل بأنه أكثر مقاييس الارتباط وأخطاؤه المعيارية صغيرة وهو الأكثر استخداما.

معامل ارتباط سيبرمان: يتطلب هذا المعامل إعطاء قيم كل متغير رتبا حسب تسلسلها، فيتم ترتيب قيم المتغير الأول ترتيبا تصاعديا أو تنازليا، ثم ترتيب قيم المتغير الثاني وفقا لذلك. فمثلا تعطى أكبر قيمة في المتغير الأول رتبة (1) ثم تليها رتبة (2) وهكذا. ثم يتم التعامل مع قيم المتغير الثاني بنفس الطريقة.

خامسا الاختبارات الإحصائية

لتحضير المعطيات الكمية بهدف تحليلها، فإن الجداول والرسومات البيانية لها فائدة استعمالية كبيرة، إلا أنها غير كافية أحيانا في تحديد ما نبحث عنه، لا سيما في حالة إذا كنا نريد تحديد قوة العلاقة بين متغيرين أو اختبار مدى وجود ارتباط بين متغيرات الدراسة؛ لهذا الغرض يمكن استعمال عدة اختبارات في العلوم الإنسانية.

المقصود بالفرضيات الإحصائي(Statistical Hypotheses) بمعنى الفرضيات التي تتعلق بالمجتمع الإحصائي المسحوبة منه العينة، أو توزيع هذا المجتمع أو معالمه كالوسط الحسابي أو النسبة في المجتمع. وسيتم التركيز هنا على نوعين مهمين جدا من الاختبارات هما:

1.اختبارات الفرضية:

يشير مصطلح الفرضية في اختبارات الفرضية إلى الفرضية الإحصائية، وليس الفرضية التي يكون الباحث قد أعدها خلال المرحلة الأولى من البحث. على العكس من ذلك فإن الفرضية الإحصائية تساعد على التحقق من فرضية البحث وذلك بالتأكد إن كان المتغير (X) يؤثر في المتغير (Y) أم لا. فمثلا: قد يفترض الباحث أن للتسويق الإلكتروني تأثير إيجابي على سلوك المستهلك الشرائي (بناء على ما يراه من مستوى المعيشة في هذا البلد وسلوك المستهلك العام)، ويحتاج إلى اختبار إحصائي لمعرفة مدى صحة هذه الفرضية.

وتسمى الفرضية التي يمكنها أن تنبئ بعدم وجود علاقة بين المتغيرين في الإحصاء بالفرضية الصفرية ورمزها H_0، أما الفرضية التي تنبئ بوجود علاقة بين المتغيرين في الإحصاء فتسمى بالفرضية البديلة H_1.

مستوى المعنوية Level of Significance:

يعتبر مصطلح "مستوى المعنوية" واحدا من أهم المصطلحات المستخدمة في دراسة نظرية اختبارات الفرضيات. والمقصود بمستوى المعنوية هو" احتمال حدوث الخطأ من النوع الأول "، أي احتمال رفض الفرضية العدمية بينما هي صحيحة. وعادة ما يرمز إلى مستو المعنوية بالرمز اللاتيني ألفا α وأشهر قيمتين لمستوى المعنوية هما 5% و1%.

إن كل اختبارات الفرضية لن يكون لها معنى إلا أذا كانت معطيات البحث قد جمعت عن عينة من نوع احتمالي والتي توفر ضمانات عن درجة تمثيليتها بالنسبة لمجتمع البحث الذي أخذت منه. لكن نظرا إلى كون العينة ليست دائما انعكاسا صادقا وصحيحا لمجتمع البحث فإن هناك هامشا للخطأ لا مفر منه ضمن الفروق والاختلافات المبينة بواسطة الاختبارات الإحصائية، لهذا نتحدث فيما يخص هذه الاختبارات عن عتبة الدلالة أي مستوى من الثقة يكون مقبولا للتمكن فيما بعد من القول إن الفرق الملاحظ دال أو غير دال. ومن الملاحظات المهمة هو أن" مستوى المعنوية" والذي يسمى أحيانا "مستوى الدلالة" هو المكمل لدرجة الثقة، بمعنى أن مجموعها يساوي 100% أو واحد صحيح. فإذا كانت درجة الثقة 95% فإن مستوى المعنوية يساوي 5% والعكس صحيح.

ويتم اختيار الاختبارات المناسبة وفقا لطبيعة الفرضية إن كانت فرضية ارتباطية مثل: هل توجد علاقة ذات دلالة إحصائية بين التسوق الالكتروني ومستوى الدخل. أو فرضية دلالة الفروق مثل: توجد فروق ذات دلالة إحصائية بين التسوق الالكتروني ومستوى الدخل، كما هو موضح فيما يلي:

اختيار الأسلوب الإحصائي المناسب–الفرضية الارتباطية

الفرضية	ارتباطية	ارتباطية	تنبؤية	عاملية
المتغير / مستوى القياس	متغيران 2 Variables	Multivariable متعدد المتغيرات		
Nominal اسمية	معامل التنبؤ لجتمان معامل فاي ϕ PHI معامل الاقتران Association	معامل التوافق Contingency معامل لامدا Lambda معامل كرامر Cramer تشيبرو Tachuprou		
Ordinal رتبية	معامل سبيرمان Spearman معامل جاما Gammaمعامل كندال Kendall			
Interval or Ratio فترية+ نسبية	معامل بيرسون Pearson معامل ايتا ETA الانحدار الخطي Linear regression	معامل الارتباط المتعدد Multiple correlation	تحليل الانحدار المتعدد Multiple regression التحليل التمييزي Discriminant Analysis تحليل المسار Bath Analysis السلاسل الزمنية Time Series	التحليل العاملي Factor analysis

اختيار الأسلوب الإحصائي المناسب–دلالة الفروق

الفرضية	التحقق من المطابقة	الفروق بين المجموعات	الفروق بين القياسات	الفروق بين المجموعات	الفروق بين القياسات
عينة الدراسة / مستوى القياس	1-sample عينة واحدة	2 Related Samples عينتان مترابطتان	2 Independent Samples عينتان مستقلتان	k Related Samples عدة عينات مترابطة	k Independent Samples عدة عينات مستقلة
Nominal اسمية	Chi-square χ^2 مربع كاي(كا٢)	McNemar ماكنمار	Chi-square χ^2 Median Test اختبار الوسيط Fisher Test اختبار فيشر	Cochran Test Q اختبار كوجران	Chi-square χ^2 مربع كاي (كا٢)
Ordinal رتبية	Kolmogorov Smirnow كولموجروف-سمير نوف-KS	Wilcoxon- Z ولكوكسن Sign Test اختبار الاشارة	Kolmogorov Smirnow KS Mann-Whitney U مان وتني	Friedman فريدمان	Median Test اختبار الوسيط Kruskal-Wallis –H كروسال-واليز
Interval or Ratio فترية+ نسبية	Z-test Z اختبار t-test اختبار ت	Paired t-test اختبار ت	Independent t-test اختبار ت	ANOVA with repeated measures تحليل التباين للقياسات المتكررة	ANOVA (F) تحليل التباين (ف) Covariance تحليل التغاير

2.اختبارات التجميع:

تهدف اختبارات التجميع أو الارتباط إلى قياس شدة العلاقة الموجودة بين متغيرين. إنها تسمح بإتمام دراسة العلاقة بين هذه المتغيرات. لكل قياس من قياسات التجميع مداه في التغير الذي لا بد من معرفته أولا إذا أردنا تقييم المعنى فيما بعد. من بين أهم الاختبارات هو **معامل التكرار المزدوج**.

معامل التكرار المزدوج: ويستخدم هذا النوع من الجداول في التبويب في حالة وجود ظاهرتين (متغيرين) يعتمد أحدهما على الآخر وهي الجداول التي غالبا ما تستخدم في اختبار الفروض، واختبار التباين.

فهو يعبر عن درجة التجميع بين المتغيرات، ويكون متأسسا انطلاقا من **كاي التربيعي** ومن حجم العينة، وتتراوح قيمته بين صفر (0) وواحد (1) أو أقل من ذلك بقليل، وذلك بحسب الخطوط والأعمدة الموجودة داخل الجدول.