الاختبارات الإحصائية غير المعلمية

اختبار مان ويتني
واختبار كروسكال واليس

دليل شامل لفهم الخطوات والشروط والتطبيقات العملية
للاختبارات البديلة غير المعلمية

استخلاص محتوى من فيديوهات يوتيوب التعليمية

رابط الفيديو الأول: youtu.be/rI5RiHlIydc

رابط الفيديو الثاني: youtu.be/C1oioGHlnI8

مقدمة في الاختبارات غير المعلمية

تُعدّ الاختبارات الإحصائية غير المعلمية (Non-parametric Tests) من الأدوات الأساسية في التحليل الإحصائي، خاصةً عندما لا تستوفي البيانات شروط التوزيع الطبيعي. وتتميز هذه الاختبارات بمرونتها العالية في التعامل مع البيانات الرتبية والفئوية والبيانات التي تحتوي على قيم متطرفة، مما يجعلها الخيار الأمثل في العديد من مجالات البحث العلمي والاجتماعي والطبي والتربوي.

يهدف هذا الدليل إلى تقديم شرح مفصّل لاختبارين من أهم الاختبارات غير المعلمية، وهما: اختبار مان ويتني (Mann-Whitney U Test) الذي يُستخدم لمقارنة عينتين مستقلتين، واختبار كروسكال واليس (Kruskal-Wallis Test) الذي يُستخدم لمقارنة ثلاث مجموعات فأكثر. وقد تم استخلاص المحتوى من فيديوهين تعليميين متخصصين يشرحان هذين الاختبارين بالتفصيل.

> **لماذا الاختبارات غير المعلمية؟**
>
> عندما لا تكون البيانات موزعة توزيعاً طبيعياً، أو عندما تكون البيانات من النوع الرتبي (الترتيبي)، فإن الاختبارات المعلمية التقليدية مثل اختبار t واختبار تحليل التباين (ANOVA) تصبح غير مناسبة. وهنا تبرز الحاجة إلى استخدام الاختبارات غير المعلمية كبديل موثوق وقوي إحصائياً.

القسم الأول: اختبار مان ويتني (Mann-Whitney U Test)

الفيديو المرجعي: youtu.be/C1oioGHlnI8

تعريف الاختبار

اختبار مان ويتني يو (Mann-Whitney U Test) هو اختبار إحصائي غير معلمي يُستخدم لتحديد ما إذا كان هناك فرق ذو دلالة إحصائية بين مجموعتين مستقلتين من البيانات غير الموزعة بشكل طبيعي. ويُعرف أيضاً باختبار مجموع رتب ويلكوكسون (Wilcoxon Rank-Sum Test)، وهو بديل لاختبار t للعينات المستقلة عندما لا تتحقق شروط التوزيع الطبيعي.

يقوم هذا الاختبار على مبدأ ترتيب الملاحظات من كلا المجموعتين معاً بترتيب تصاعدي، ثم حساب إحصائية U لمقارنة الرتب بين المجموعتين. وبفضل اعتماده على الرتب بدلاً من القيم الفعلية، يتميز الاختبار بمقاومة عالية للقيم المتطرفة

والبيانات المنحرفة، مما يجعله أداة قوية وموثوقة في مختلف مجالات البحث العلمي.

متى نستخدم اختبار مان ويتني؟

- **عند مقارنة عينتين مستقلتين:** مثل مقارنة أداء الذكور والإناث في اختبار معين، أو مقارنة نتائج مجموعتين علاجيتين مختلفتين.
- **عندما لا تتبع البيانات التوزيع الطبيعي:** إذا فشلت البيانات في اختبار الشهودية (Shapiro-Wilk) أو اختبار كولموغوروف-سميرنوف للتعرف على التوزيع الطبيعي.
- **مع البيانات الرتبية أو الترتيبية:** مثل تصنيف المستويات (ممتاز، جيد جداً، جيد، مقبول) أو ترتيب التفضيلات.
- **عند وجود قيم متطرفة:** التي تؤثر بشكل كبير على المتوسطات الحسابية وتجعل اختبار t غير موثوق.

تطبيقات عملية

في البحوث الصيدلانية، يمكن استخدامه لمقارنة فعالية دواءين مختلفين. وفي التعليم، يمكن استخدامه لتحليل ما إذا كانت طريقة التدريس "أ" تحقق درجات اختبار أعلى من الطريقة "ب". والمفتاح أنه يسمح بمقارنة مجموعتين على نتيجة مستمرة أو ترتيبية دون الحاجة لافتراض التوزيع الطبيعي.

فرضيات الاختبار

يعتمد اختبار مان ويتني يو على عدة افتراضات أساسية يجب التحقق منها قبل تطبيقه:

- **استقلال الملاحظات:** يجب أن تكون كل ملاحظة مستقلة عن الملاحظات الأخرى، بحيث لا يوجد أي ارتباط أو تبعية بين الملاحظات الفردية داخل كل مجموعة أو بين المجموعتين.
- **أخذ عينات عشوائية:** يجب أن يتم سحب البيانات بشكل عشوائي من المجتمع الإحصائي، بحيث يتم استخلاص كل ملاحظة بشكل مستقل.
- **البيانات ترتيبية أو عددية مستمرة:** يكون الاختبار مناسباً بشكل خاص للبيانات الترتيبية (المرتبة) أو العددية المستمرة التي لا تتبع التوزيع الطبيعي. وإذا كانت البيانات مستمرة وتتبع التوزيع الطبيعي، فإن اختبار t البارامتري يكون أكثر ملاءمة لتمتعه بقدرة إحصائية أكبر.

تحذير مهم

يمكن أن تؤدي انتهاكات هذه الافتراضات إلى نتائج متحيزة أو غير صحيحة. ولذلك، فإن فهم هذه الافتراضات والتحقق من صحتها أمر بالغ الأهمية قبل تنفيذ اختبار مان ويتني يو. لا يمكن استخدام هذا الاختبار مع البيانات المقترنة أو المرتبطة (كالقياسات قبل وبعد على نفس الأفراد)؛ في هذه الحالة يُستخدم اختبار ويلكوكسون للرتب الموقعة.

خطوات إجراء الاختبار

1. **فرز البيانات:** ابدأ بدمج مجموعتي البيانات معاً وفرز جميع القيم بترتيب تصاعدي. قم بتعيين أرقام رتبية لكل ملاحظة، بحيث تحصل أصغر نقطة بيانات على الرتبة 1. وإذا كانت نقطتا بيانات أو أكثر متطابقتين (أي مرتبطتين)، فستحصلان على رتبة متوسطة.

2. **حساب مجموع الرتب:** قم بتلخيص الرتب لكل مجموعة بشكل منفصل. يمنحك هذا إجماليين، واحد لكل مجموعة من المجموعتين اللتين تقارنهما (R1 للمجموعة الأولى و R2 للمجموعة الثانية).

3. **حساب إحصائية U:** تُحسب إحصائية U لكل مجموعة باستخدام الصيغ التالية

$$U_1 = n_1 \times n_2 + n_1(n_1 + 1) / 2 - R_1$$

$$U_2 = n_1 \times n_2 + n_2(n_2 + 1) / 2 - R_2$$

حيث n_1 و n_2 هما حجم العينتين، و R_1 و R_2 هما مجموع الرتب في المجموعة الأولى والثانية على التوالي.

4. **تحديد قيمة U الأصغر:** يتم استخدام قيمة U الأصغر بين الإحصائيتين المحسوبتين للاختبار، وهي تمثل إحصاءة الاختبار الفعلية.

5. **تحديد الدلالة الإحصائية:** قارن إحصائية U المحسوبة بالقيمة الحرجة من جداول توزيع Mann-Whitney U، أو استخدم القيمة الاحتمالية (p-value) من البرامج الإحصائية.

6. **اتخاذ القرار:** إذا كانت القيمة الاحتمالية (p-value) أقل من مستوى الدلالة المختار (عادة 0.05)، نرفض الفرضية العدمية ونستنتج وجود فرق ذو دلالة إحصائية بين المجموعتين.

فرضيات الاختبار الإحصائية

الفرضية	الوصف
الفرضية العدمية (H_0)	لا يوجد فرق ذو دلالة إحصائية بين المجموعتين (تتساوى توزيعات المجموعتين)
الفرضية البديلة (H_1)	يوجد فرق ذو دلالة إحصائية بين المجموعتين (تختلف توزيعات المجموعتين)

تفسير النتائج

- **إحصائية U:** توفر إحصائية U معلومات عن مجموع ترتيب البيانات من المجموعتين. إذا كانت إحصائية U صغيرة، فإنها تشير إلى فرق كبير بين المجموعتين، أي أن إحداهما تميل للحصول على رتب أعلى من الأخرى.
- **القيمة الاحتمالية (p-value):** تشير القيمة p الأقل من مستوى الدلالة المختار (عادة 0.05) إلى أن الفرق بين المجموعتين ذو دلالة إحصائية، مما يعني أن الفرق المرصود ليس ناتجاً عن الصدفة وحدها.
- **حجم التأثير:** غالباً ما يتم قياس حجم التأثير باستخدام ارتباط الرتب الثنائية (Rank-Biserial Correlation)، ويمكن أن تتراوح قيمته من -1 إلى +1، حيث تشير القيم القريبة من الصفر إلى تأثير ضعيف والقيم القريبة من 1 أو -1 إلى تأثير قوي.

ملخص اختبار مان ويتني

- اختبار غير معلمي لمقارنة مجموعتين مستقلتين
- بديل لاختبار t عندما لا تتحقق شروط التوزيع الطبيعي
- يعتمد على ترتيب البيانات (الرتب) بدلاً من المتوسطات
- إحصائية U الصغيرة تشير إلى فرق أكبر بين المجموعتين
- p-value أقل من 0.05 تعني فرق ذو دلالة إحصائية
- لا يُستخدم للبيانات المقترنة (يُستبدل باختبار ويلكوكسون)

القسم الثاني: اختبار كروسكال واليس (Kruskal-Wallis Test)

الفيديو المرجعي: youtu.be/rI5RiHlIydc

تعريف الاختبار

اختبار كروسكال واليس (Kruskal-Wallis Test) هو اختبار إحصائي غير معلمي يُستخدم لمقارنة ثلاث مجموعات أو أكثر من البيانات المستقلة عن بعضها البعض. يتم تطبيق هذا الاختبار عندما تكون البيانات غير موزعة توزيعاً طبيعياً،

أي عندما لا يمكن استخدام الاختبارات المعلمية التقليدية مثل اختبار تحليل التباين الأحادي (One-Way ANOVA). ويعتمد الاختبار على ترتيب البيانات ضمن المجموعات بدلاً من مقارنة المتوسطات الحسابية.

يُعتبر اختبار كروسكال واليس امتداداً لاختبار مان ويتني من مجموعتين إلى ثلاث مجموعات فأكثر، ويُعرف أحياناً باختبار H أو اختبار كروسكال واليس H. وهو يماثل في وظيفته اختبار تحليل التباين الأحادي (ANOVA) ولكن في إطار غير معلمي، مما يجعله ملائماً عندما يتعذر تحقيق شروط ANOVA المتعلقة بالتوزيع الطبيعي وتجانس التباينات.

متى نحتاج اختبار كروسكال واليس؟

- **عندما لا تكون البيانات موزعة توزيعاً طبيعياً:** في العديد من الدراسات، لا تكون البيانات التي يتم جمعها موزعة توزيعاً طبيعياً (مثل البيانات التي تحتوي على قيم شاذة أو غير متجانسة). في هذه الحالات، يصعب استخدام اختبارات معلمية مثل ANOVA، وهنا يأتي دور اختبار كروسكال واليس ليكون الخيار الأمثل.
- **عندما تكون البيانات غير متجانسة:** يُستخدم هذا الاختبار عندما تكون هناك اختلافات كبيرة في التباين بين المجموعات التي نقارنها، مما يجعل اختبار ANOVA غير موثوق بنتائجه.
- **البيانات الرتبية أو الفئوية:** يُعد هذا الاختبار مثالياً عندما تكون البيانات نوعية أو رتبية، مثل تصنيف الأشخاص أو الأشياء ضمن فئات معينة (مثل التقييمات أو التفضيلات) بدلاً من قياسات رقمية دقيقة.

خطوات إجراء اختبار كروسكال واليس

1. **تحضير البيانات:** تأكد من وجود ثلاث مجموعات أو أكثر من البيانات المستقلة. وتأكد أن البيانات غير موزعة توزيعاً طبيعياً (يمكن التحقق من ذلك باستخدام اختبار الشهودية). كما يجب أن تكون البيانات رتبية أو نوعية أو عددية مستمرة لا تتبع التوزيع الطبيعي.

2. **تحديد الفرضيات:** الفرضية العدمية (H_0): لا توجد فروق ذات دلالة إحصائية بين المجموعات. والفرضية البديلة (H_1): هناك فرق دال إحصائياً بين المجموعات (على الأقل بين مجموعتين).

3. **اختيار مستوى الدلالة:** اختر مستوى دلالة (عادة 0.05 أو 0.01 حسب الحاجة) للمقارنة مع القيمة الاحتمالية.

4. **دمج وترتيب البيانات:** قم بدمج جميع بيانات المجموعات معاً في إطار تحليلي واحد، ثم رتّب جميع القيم ترتيباً تصاعدياً وأسند الرتب المناسبة لكل قيمة، مع استخدام الرتب المتوسطة للقيم المتساوية.

5. **حساب مجموع رتب كل مجموعة:** قم بحساب مجموع الرتب لكل مجموعة على حدة (R_i للمجموعة i)، ثم استخدم الصيغة التالية لحساب إحصائية H.

$$H = [12 / N(N+1)] \times \sum(R_i^2 / n_i) - 3(N+1)$$

حيث N هو إجمالي حجم جميع العينات، و n_i هو حجم العينة في المجموعة i، و R_i هو مجموع رتب المجموعة i.

6. **تحليل النتائج واتخاذ القرار:** إذا كانت القيمة الاحتمالية (p-value) أقل من مستوى الدلالة (0.05)، يمكنك رفض الفرضية العدمية والاستنتاج بوجود فرق دال إحصائياً. أما إذا كانت أكبر من 0.05، فلا يمكنك رفض الفرضية العدمية.

7. **الاختبارات التالية (Post-hoc):** إذا تم رفض الفرضية العدمية وتبين وجود فرق دال إحصائياً، يمكن إجراء اختبارات إضافية مثل اختبار مان ويتني للمقارنات الزوجية لمعرفة أي المجموعات تختلف عن الأخرى تحديداً.

فرضيات الاختبار الإحصائية

الفرضية	الوصف
الفرضية العدمية (H_0)	لا توجد فروق ذات دلالة إحصائية بين جميع المجموعات (تتساوى التوزيعات)
الفرضية البديلة (H_1)	يوجد فرق دال إحصائياً بين المجموعات (على الأقل مجموعتان تختلفان)

مزايا اختبار كروسكال واليس

- **مرونة في التعامل مع البيانات غير الطبيعية:** بما أن الاختبار لا يعتمد على فرضية التوزيع الطبيعي، فإنه يمكن استخدامه مع العديد من أنواع البيانات المختلفة دون قيود صارمة.
- **سهولة التطبيق:** مقارنةً ببعض الاختبارات المعقدة، يُعد اختبار كروسكال واليس سهل التطبيق سواء يدوياً أو باستخدام البرامج الإحصائية، ويعطي نتائج دقيقة وموثوقة.
- **إجراء تحليل متعدد المجموعات:** يمكن مقارنة أكثر من مجموعتين في نفس الوقت، مما يجعله أكثر كفاءة من إجراء مقارنات زوجية متعددة باستخدام اختبار مان ويتني.
- **مقاومة للقيم المتطرفة:** نظراً لاعتماده على الرتب بدلاً من القيم الفعلية، فإن القيم المتطرفة لا تؤثر بشكل كبير على نتائجه.

متى لا يُفضل استخدام اختبار كروسكال واليس؟

في حال كانت البيانات تتبع توزيعاً طبيعياً، أو إذا كان هناك حاجة لتحليل المتوسطات الحسابية بدلاً من ترتيب البيانات، فقد يكون من الأفضل استخدام اختبارات أخرى مثل ANOVA لأن الاختبارات المعلمية تتمتع بقدرة إحصائية

أكبر عندما تتحقق شروطها. كما لا يمكن استخدام هذا الاختبار إذا كانت المجموعات ليست مستقلة عن بعضها البعض (كالقياسات المتكررة على نفس الأفراد)، أو إذا كانت البيانات لا تتبع مقياساً رتبياً أو نوعياً على الأقل.

ملخص اختبار كروسكال واليس

- اختبار غير معلمي لمقارنة ثلاث مجموعات فأكثر
- بديل لاختبار ANOVA عندما لا تتحقق شروط التوزيع الطبيعي
- امتداد لاختبار مان ويتني من مجموعتين إلى عدة مجموعات
- يعتمد على الرتب ولا يتطلب تجانس التباينات
- إحصائية H تُقارن بتوزيع كاي-تربيع
- عند رفض الفرضية العدمية، تُستخدم اختبارات Post-hoc للمقارنات الزوجية

مقارنة بين الاختبارين

يُعدّ اختبارا مان ويتني وكروسكال واليس من أهم الاختبارات غير المعلمية في الإحصاء، وهما يتشابهان في المبدأ الأساسي (الاعتماد على الرتب) لكنهما يختلفان في نطاق التطبيق. يقدم الجدول التالي مقارنة تفصيلية بينهما من حيث الشروط والاستخدامات والخصائص:

المعيار	اختبار مان ويتني (Mann-Whitney U)	اختبار كروسكال واليس (Kruskal-Wallis H)
عدد المجموعات	مجموعتان مستقلتان	ثلاث مجموعات فأكثر
البديل المعلمي	اختبار t للعينات المستقلة	اختبار تحليل التباين الأحادي (ANOVA)
إحصائية الاختبار	U	H (تتبع توزيع كاي-تربيع)
نوع البيانات	رتبية أو عددية مستمرة	رتبية أو فئوية أو عددية مستمرة
شرط التوزيع	لا يتطلب التوزيع الطبيعي	لا يتطلب التوزيع الطبيعي
الفرضية العدمية	تساوي توزيعي المجموعتين	تساوي توزيعات جميع المجموعات
العلاقة بينهما	الحالة الخاصة (مجموعتان)	الامتداد العام (3+ مجموعات)
الاختبارات اللاحقة	غير مطلوبة (مجموعتان فقط)	مطلوبة (اختبار مان ويتني زوجي)
حجم التأثير	ارتباط الرتب الثنائية	إيتا-تربيع (Eta-squared)

مجالات التطبيق المشتركة

يجد هذان الاختباران تطبيقات واسعة في العديد من المجالات البحثية والمهنية، ومن أبرزها:

- **البحوث الطبية والبيولوجية:** مقارنة فعالية العلاجات المختلفة أو مقارنة المؤشرات الحيوية بين مجموعات المرضى عندما لا تتوزع البيانات طبيعياً.
- **البحوث الاجتماعية:** تحليل الفروق بين فئات اجتماعية مختلفة في متغيرات مثل مستوى الرضا أو الاتجاهات.
- **دراسات التعليم:** مقارنة أداء الطلاب بين فصول دراسية أو طرق تدريس مختلفة.
- **أبحاث السوق والتحليل التجاري:** مقارنة تفضيلات المستهلكين بين منتجات أو علامات تجارية متعددة.
- **البحوث النفسية:** تحليل الفروق في درجات المقاييس النفسية بين مجموعات مختلفة.
- **دراسات البيئة:** مقارنة مستويات التلوث أو التنوع البيولوجي بين مواقع مختلفة.

الخطوات العامة للاختبارات غير المعلمية

بصرف النظر عن نوع الاختبار غير المعلمي المستخدم، توجد خطوات عامة مشتركة ينبغي اتباعها لضمان صحة النتائج وموثوقيتها:

1. **فحص التوزيع الطبيعي:** قبل اختيار الاختبار، تحقق من توزيع البيانات باستخدام اختبارات مثل Shapiro-Wilk أو Kolmogorov-Smirnov. إذا لم تكن البيانات طبيعية، انتقل للاختبارات غير المعلمية.

2. **اختيار الاختبار المناسب:** مجموعتان مستقلتان (مان ويتني)، ثلاث فأكثر (كروسكال واليس)، بيانات مقترنة (ويلكوكسون).

3. **صياغة الفرضيات:** حدد الفرضية العدمية والبديلة بوضوح قبل جمع البيانات وتحليلها.

4. **تنفيذ الاختبار:** استخدم برنامج إحصائي مثل SPSS أو R أو Python لحساب الإحصائية والقيمة الاحتمالية.

5. **تفسير النتائج:** قارن القيمة الاحتمالية بمستوى الدلالة، واحسب حجم التأثير لتقييم الأهمية العملية وليس فقط الإحصائية.

6. **الإبلاغ عن النتائج:** تضمين أحجام العينات وإحصائية الاختبار والقيمة الاحتمالية وحجم التأثير في التقرير النهائي.

نصيحة عملية مهمة

حتى عندما تكون بياناتك تتوزع طبيعياً، يمكن استخدام الاختبارات غير المعلمية كأداة تحقق إضافية (Triangulation). وإذا توافقت نتائج الاختبارات المعلمية وغير المعلمية، فإن ذلك يزيد من ثقتك في النتائج. أما إذا اختلفت، فينبغي فحص البيانات بشكل أعمق لمعرفة السبب.

تم إعداد هذا الدليل بناءً على محتوى فيديوهات يوتيوب التعليمية:
الفيديو الأول (كروسكال واليس): https://youtu.be/rI5RiHlIydc?t=119
الفيديو الثاني (مان ويتني): https://youtu.be/C1oioGHlnI8?t=6
مع الإثراء من مصادر إحصائية أكاديمية متخصصة.