جامعة البعث

كلية التمريض

المعهد التقاني لخدمات الطوارئ

الإحصاء الحيوي

المحاضرة الخامسة

الدكتور ياسر العمر

مدخل إلى اختبار الفرضيات

AN INTRODUCTION TO HYPOTHESIS TESTING

مقدمة: Introduction

يمكن أن نقسم عموماً الفرضية أو النظرية الإحصائية إلى جزئيين :

١- هناك ما يسمى بالإحصاء الوصفي descriptive statistics حيث تستخدم فيه أدوات العرض المناسبة كالجداول النموذجية والأشكال والقياسات الرقمية لوصف قاعدة البيانات وتقديم عرض مختصر لتوزيعها .

٢- إضافة إلى ذلك هناك ما يدعى بالإحصاء الاستنتاجي أو الاستدلالي inferential statistics الذي يركز على إعطاء الاستنتاجات لمجتمع ما مستخدماً النتائج التي حصلنا عليها من العينة الممثلة للمجتمع المدروس والمأخوذة من هذا الأخير .

a- وإن إحدى النقاط في الاستدلال الإحصائي هي قضايا التقدير estimation لحدود المجتمع المدروس باستخدام العينة الإحصائية المناسبة (على سبيل المثال ، وسط المجتمع المدروس يمثل بوسط العينة المأخوذة منه) . كما أن قضايا التقدير تكتمل فقط عندما تكون هناك دقة في عملية التقدير التي تحدد بالخطأ المعياري أو يعطى مؤشر لهذه العملية حد الثقة المدروس .

b- النقطة الثانية في الاستدلال الإحصائي هي اختبار الفرضية Hypothesis testing . وفي هذه الحالة نفحص الفرضية ضمن عبارة واحد أو أكثر من المجتمعات المدروسة . كما أننا نريد أن نعرف ما إذا كانت الفرضية حول هذا المجتمع أو المجتمعات الأخرى يمكن أن تدعم أو تمثل ببيانات العينة .

إن قضايا التقدير للقيم تركز على عملية وصف القيم بينما يركز اختبار الفرضية بلا تحديد على اتخاذ القرار المناسب .

المفاهيم الأساسية لاختبار الفرضية :

The Main Concept of the Hypothesis Test

إن اختبار الفرضية عملية تركز على إجراء الاستنتاجات حول المجتمع المدروس مستخدمين المعلومات المحصول عليها من العينة المأخوذة . ويجب أن نعي أنه من المستحيل أن نكون قادرين على التأكيد المطلق لصحة استنتاجاتنا حول المجتمع . إذ أن عينة واحدة مختارة من المجتمع من غير المحتمل أن تعطي نفس النتائج تماماً كما هي في العينة الثانية .

ولذلك يجب أن نحدد نتائجنا حول المجتمع المدروس باستخدام مفهوم الاحتمالية وهذا يعطي مؤشراً لفرصة الحصول على النتائج المشاهدة إذا كانت الفرضية حقيقية .

وهنا لا بد من استبدال عبارة الحالة المطلقة بعبارة الاحتمالية Probabilistic وهذه الأشكال هي المضمون أو المفهوم لاختبار الفرضية وهكذا يمكن أن يتلخص مفهوم الفرضية الإحصائية بدراسة صحة ادعاء معين عن قيمة معينة ما للمجتمع الإحصائي كالوسط الحسابي أو النسبة أو التباين . كأن نقول مثلاً إن هذا الدواء ناجع بنجاح بنسبة 95% لعلاج التهابات الكبد إلى غيرها من الافتراضات التي نود التأكد من صحتها.

نظرية العدم أو فرضية العدم

The Null Hypothesis H0

ترتكز الدراسات للتقصي عن فرضية معينة أو نظرية علمية حول المجتمع المدروس عادة على عملية مقارنة بالحالة الطبيعية مشمولة ببعض التأثيرات الرقمية لموضوع الدراسة وغالباً ما تكون هذه المقارنة لمعالجات مختلفة وقياس استجابة هذه المعالجة يدعى بالتأثير العلاجي . على سبيل المثال الفروقات بين مختلف العلاجات من حيث الوسط الحسابي .

لنفترض أننا نريد التقصي عن نقص مغنزيوم الدم عند أشخاص يتناولون في وجباتهم الجرعات اللازمة من عنصر المغنزيوم وأشخاص لا يحصلون على الجرعات الضرورية من هذا العنصر في غذائهم فعندئذ نريد أن نفحص قيم معدل بلازما الدم. وهذا يقترح أن هناك خطورة في لعدم وجود الاحتياجات الطبيعية للمنغنزيوم . وبهذه الحالة يتشكل ما يدعى بنظرية العدم حول معدل القيم الحقيقية لمنغنزيوم البلازما والتي يفترض أن لا تختلف بين المجموعتين .

هذا ما يدعى بنظرية العدم null hypothesis ويرمز لها بالرمز Ho وقد تكون فرضية أو نظرية العدم غير حقيقية وعندئذ تدعى بالفرضية أو النظرية البديلة alternative hypothesis ويرمز لها بالرمز H1 .

نتيجة :

هناك عادة فرضيتان في دراسة أي اختبار ، الأولى منها تدعى بنظرية أو فرضية العدم وهي الفرضية الأولية والتي غالباً ما تؤيد صحة الادعاء ويرمز لها بالرمز Ho والثانية هي الفرضية البديلة والتي هي أساساً لا تؤيد صحة الادعاء ويرمز لها بالرمز H1 .

عادة تذكر الفرضية البديلة أنه يوجد اختلاف بين قيم الحد لكن الاتجاه لهذا الاختلاف غير معروف . وهذا يقودنا إلى فرضية ذات جانبين أو اختبار الذيلين (two- sided or a two tailed test) .

- الاختبار الإحصائي وقيمة p

The Test Statistic and the p-Value

من البيانات التي حسبنا منها قيمة الاختبار الإحصائي - (وهو تعبير جبري معين للفرضية التي نختبرها) ، وعادة يستخدم الحاسوب لعملية الحساب هذه وأحياناً تحسب يدوياً - ويرفق بكل قيمة الاختبار الإحصائي ما يدعى بالاحتمالية probability وهو ما يدعى بالقيمة p–

(p-value) . وهي وهذه القيمة تصنف فرضية الحصول على تأثير القيم المشاهدة (وبعبارة أخرى أكثر توضيحاً) إذا كانت فرضية العدم حقيقية .

اتخاذ القرار باستخدام قيمة p

Making a Decision Using the p-Value

١- إذا كانت النتائج المشاهدة غير مطابقة لما نتوقع وعندما تكون نظرية العدم حقيقية ، فإننا نستنتج أنه يوجد دلالة كافية لرفض reject نظرية أو فرضية العدم . ونقول أن نتيجة الاختبار إحصائياً هي معنوية Statistically significant .

٢- وإذا كانت النتائج المشاهدة ثابتة مع ما نتوقع إذا كانت نظرية العدم حقيقية ونستنتج أنه لا يوجد دلالة كافية لتكون حقيقية وعندها نرفض فرضية (نظرية) العدم . ونقول أن نتيجة الاختبار غير معنوية non – significant . وتسمح لنا قيمة (p) بأن نحدد فيما إذا كنا نملك دليلاً كافياً لرفض فرضية العدم بالمقارنة مع الفرضية البديلة .

- إذا كانت قيمة (p) صغيرة جداً فعندئذٍ من غير المحتمل أن نستطيع الحصول على النتائج المشاهدة إذا كانت فرضية العدم حقيقية ولذلك ترفض فرضية العدم Ho .

- أما إذا كانت قيمة p كبيرة جداً فعندئذ ثمة فرصة كبيرة للحصول على النتائج المشاهدة إذا كانت فرضية العدم حقيقية وعندئذ ترفض فرضية العدم لا ترفض . وبوضوح إن التمييز بين قيم (P) الصغيرة والكبيرة هو موضوع متروك لاتخاذ القرار المناسب . فيجب أن نحدد قبل البدء بعملية جمع البيانات المكونة لقيم (p) سواء بقيمة صغيرة أم كبيرة وهنا لا بد من استخدام مصطلح مستوى المعنوية Significance level للاختبار . كما أن اختبار مستوى المعنوية يعتمد على طبيعة البيانات والظروف المتعلقة بالتقصي عن الحالة إذ هناك إمكانية تجعلنا مثلاً نختار قيمة منخفضة جداً لمستوى المعنوية كأن نقول المستوى 0.01 ، إذا كنا نريد أن ننظر إلى الجانب

السببي لرفض نظرية العدم وهذا يعني أنه في مثالنا نحصل على قيمة (p) والتي هي أقل من 0.01 فإننا نرفض نظرية العدم ونقول إن النتيجة معنوية عند المستوى 1% .

على سبيل المثال نريد التقصي عما إذا كانت المعالجة بالصادات الحيوية مكلفة فإننا نريد أن نكون واثقين جداً من فوائدها خلال فترة المعالجة وكذا تكلفتها المادية .

إن اختيار الفاصل القاطع 0.05 Cut off point هو غالباً اختيار لمستوى المعنوية مثلاً إذا كانت قيمة p أصغر من 0.05 فعندئذ ترفض نظرية العدم ، أما إذا كانت قيمة p أكبر أو تساوي 0.05 فعندها نقبل فرضية العدم . وهناك تمييزات إضافية تستخدم أحياناً بواسطة الإشارة النجمية :

(***) تمثل p ($p< 0.0001$) وهي معنوية واضحة جداً .

(**) تمثل قيمة ($P<0.001$) وهي معنوية واضحة .

(*) تمثل قيمة ($0.01 < p<0.05$) وهي معنوية وعندما لا توجد معنوية يرمز لها بالرمز (NS) وعندها تكون قيمة p أكبر من 0.05 .

وبشكل آخر المعنوية الواضحة جداً تكون ضمن المجال [0.0001-0.0000]

المعنوية الواضحة تكون ضمن المجال [0.001-0.0001]

معنوية فقط [0.01-0.001]

ونذكر كلما صغرت قيمة (p) كان هناك دليل أوضح ومؤكد ضد نظرية العدم . ومن المهم الإشارة هنا إلى أن قرار الأخذ بصحة الفرضية أو رفضها لا يعني القطيعة في الحكم لأننا نقبل بالفرضية أو نرفضها باحتمال ثقة معين يفترض أن يكون كبيراً ومقنعاً مما يعني أننا قد نرفض فرضية وهي صحيحة أو نقبل بفرضية وهي خاطئة إلا أن احتمال وقوعنا بالخطأ سيصغر مع ارتفاع احتمال الثقة .

اشتقاق قيمة(p): Deriving the p-Value

نموذجياً نحصل على قيمة p من التوزيعات المعروفة للاختبار الإحصائي وهي التوزيعات الطبيعية ، وتوزيع t ، F ، α (والتي سوف تدرس في فصول لاحقة) . وإذا أنجز التحليل بالحاسوب فإن قيمة p التي نطلبها نحصل عليها مباشرة بواسطة نتائج الحاسوب وإذا كان من الضروري يمكن أن نحصل على قيمة p بالإشارة إلى الاختبار الإحصائي من الجدول المدرج بالتوزيعات المذكورة أعلاه .

درجات الحرية للاختبار الإحصائي

The Degrees of Freedom of the Test Statistic

سوف نجد ن عبارة درجات الحرية (degrees of freedom (df مستخدمة بإستمرار في التحليل الإحصائي . فإذا ما قمنا باستخدام جداول لربط قيمة الاختبار الإحصائي إلى قيمة p فإننا عادة يجب أن نعرف درجات الحرية للتوزيع المناسب لاختبارنا الإحصائي المستخدم . إن درجات الحرية للاختبار الإحصائي هي عبارة عن عدد المشاهدات المستقلة المساهمة بتشكيل الاختبار الإحصائي فعلى سبيل المثال أعداد المشاهدات لتقييم الاختبار الإحصائي ناقص عدد المشاهدات المحددة بقيمها بحدود معينة . إذ إن أسهل طريقة لحساب درجات الحرية لأي اختبار إحصائي هو أن نأخذها من حساب الفرق بين عدد المشاهدات في العينة وعدد الحدود التي يجب أن نقدرها لكي نقيم هذه المشاهدات إحصائياً .

ولذلك لنفرض مثلاً أن تقديرنا لفرق المجتمع بكامله σ^2 لمتغير x في حجم عينة n بواسطة عينة إحصائية S^2 يعطى بالقانون : $S^2 = \frac{\sum(X-\overline{X})^2}{(n-1)}$ (الفرق) يجب أن نقدر المعدل لكي نقيم قيمة الصورة ولذلك تكون درجات الحرية لهذا الفرقS^2 هي (n-1) بإعتبار أن n هي عدد قيم المشاهدات .

ملخص لإجراء اختبار الفرضية / النظرية / :

إن إجراء اختبار الفرضية :يتلخص في ست مراحل :

١- حدد فرضية العدم Ho أو الفرضية البديلة المطلوبة في الاختبار .

٢- اجمع البيانات وانظر نظرة شاملة إذا كان هناك إمكانية للتقصي عن توزيعات هذه البيانات فهناك العديد من الاختبارات تحدد الافتراضات التوزيعية حول البيانات : افحص التوزيعات المحددة لهذا الاختبار .

٣- على الحاسب ، اختر الاختبار المناسب ، أو باستخدام الحساب اليدوي المناسب لاختيار الاختبار الإحصائي المناسب مستخدمين بيانات العينة .

٤- اربط القيمة المحسوبة بالاختبار الإحصائي بقيمة p .

٥- اعتبر قيمة p لمحاكمات أو تحكيم فيما إذا كانت البيانات قابلة للتحكيم باستخدام فرضية العدم ومن ثم قرر ما إذا كانت البيانات ترفض أو تقبل فرضية العدم .

٦- إلا أنه من الأجدى والأنسب أن نحسب حد الثقة لتأثير موضوع الدراسة .

حيث أن نحدد المعنوية لنتيجة الدراسة في ضوء الإجراء السابق ولذلك يدعى هذا أحياناً باختبار المعنوية Significance test .

وإن اختبار الاختبار – كما سنرى لاحقاً – ليس سهلاً دائماً بل إنه يعتمد على طبيعة البيانات .

الخطأ نمط I والخطأ نمط II Type I and Type II Errors

عندما نجد نتيجة الاختبار معنوية نرفض فرضية العدم في مستوى المعنوية المدرجة . فإذا كان الاستنتاج غير صحيح و حقيقة المعدلين المدروسين متساوية فعندئذ نرفض فرضية العدم وعندما لا ترفض (عندما تكون حقيقية ..) نكون قد حددنا نمط الخطأ I .

وبشكل آخر أو بديل ، عندما نجد نتيجة الاختبار غير معنوية لا ترفض فرضية العدم عند مستوى المعنوية المدرج ، ولذلك لا يمكن أن نستدل إحصائياً على معدلات مجتمعين مختلفين .

وهنا تكون الحالة غير صحيحة والحقيقة لمعدلين مختلفين وعندئذ لا ترفض فرضية العدم وعندما ترفض الفرضية (عندما تكون خاطئة) نحدد نمط الخطأ II كما هو موضح في الجدول التالي.

الأخطاء في فرضية العدم

لا ترفض فرضية العدم Ho	رفض فرضية العدم Ho	الوصف
القرار صحيح	نمط الخطأ I	فرضية العدم حقيقية Ho
نمط الخطأ II	القرار صحيح	فرضية العدم خاطئة Ho