الفصل 9. اختبارات الفروض

اختبار المتوسط - اختبار النسبة - اختبار التباين - اختبارات المقارنة بين مجتمعين - اختبار التجانس والتعديل

توطئة

غالبا ما يحتاج الباحث في مرحلة ما من بحثه إلى اختبار فرضية أو أكثر بخصوص المجتمع المدروس. من أمثلة ذلك: اختبار فرضية بخصوص معدل الدخل في منطقة معينة، اختبار فرضية نسبة شفاء لدواء معين، ... يتم ذلك بصياغة فرضية عن المجتمع المدروس (أو المجتمعات المدروسة) ومن ثم محاولة الحصول على دليل إحصائي ينفي أو يثبت هذه الفرضية وذلك من خلال بيانات عينة (أو أكثر) عشوائية بسيطة. في هذا الفصل[1] سنتناول كيفية اختبار فرضيات موضوعة حول معالم مجتمع أو أكثر. رياضيا تخص الفرضية أحد معالم المجتمع الإحصائي كالمتوسط، النسبة أو التباين، ونعتمد في إثباتها أو رفضها على خصائص إحصائية معاينة مختارة، عادة ما تكون مناظرة لمعلمة المجتمع. من أجل ذلك يعتمد هذا الدرس، كما هو الحال بالنسبة للتقدير، على درس المعاينة.

9-1-1. اختبار المتوسط

الاختبار ثنائي الاتجاه للمتوسط
الاختبار أحادي الاتجاه للمتوسط
استخدام S كمقدر لتباين المجتمع في اختبار المتزسط
اختبار المتزسط باستخدام توزيع t

الغرض من اختبار المتوسط هو اختبار فرضية أن متوسط المجتمع μ (متوسط الدخل في بلد ما أو لفئة ما، متوسط وزن منتج معين، ...) يساوي قيمة ما μ_0. للقيام بالاختبار نستخرج عينة عشوائية بسيطة ونحسب فيها المتوسط m[2] . إذا كانت الفرضية صحيحة يفترض أن تكون m وμ_0 متقاربتين. اختبار الفرضية يكون بتقييم الفرق بين هاتين القيمتين، بحيث نرفضها إذا كان الفرق كبيرا. لتقييم الفرق بين القيمتين نحتاج إلى تحديد طبيعة التوزيع الاحتمالي ل M (و هو ما يعيدنا إلى نظريات توزيع المعاينة)، لحساب احتمال الفرق بين m وμ_0 .

9-1-2. الاختبار ثنائي الاتجاه للمتوسط.

لنتناول هذا المثال: نريد اختبار، بمستوى معنوية 5%، فرضية أن متوسط دخل الطالب في السنة الأولى من تخرجه يساوي القيمة 15000دج التي هي متوسط الدخل في عينة من 100 خريج علما أن الانحراف المعياري للدخل هو 1500 وأن الدخل يتبع التوزيع الطبيعي. نحتاج إلى الخطوات التالية:

– صياغة الفرضيات،

---

[1] تضمن البرنامج فصلين حول موضوع اختبار الفروض، الأول مفاهيم أساسية والثاني تطبيقات اختبار الفروض؛ غير أننا نرى أن الفصل بين هذين الجزئين سوف يؤدي إلى تكرار التطرق للمفاهيم الأساسية. من جهة أخرى، يصعب شرح المفاهيم الأساسية بمعزل عن تطبيقاتها أي بمعزل عن بنود المبحث الثاني. لذلك فسوف نخوض مباشرة في تطبيقات اختبارات الفروض، وسوف نشرح المفاهيم الأساسية بالتدريج عندما تظهر الحاجة إليه.

2 نرمز ب M للمتغيرة ع متوسط العينة، و نرمز ب m لقيمة المتغيرة.

- تحديد قاعدة القرار،
- حساب القيمة الجدولية للمتغيرة،
- حساب القيمة الفعلية للمتغيرة،
- اتخاذ القرار.

9-1-2-1. تحديد الفرضيات (الصفرية والبديلة):

الخطوة الأولى في الاختبار هو تحديد الفرضية المراد اختبارها، وتسمى **الفرضية الصفرية أو فرضية العدم** H_0 ، والفرضية التي ستقبل في حالتي **عدم قبول**[1] **أو رفض** الفرضية الصفرية وتسمى **الفرضية البديلة** H_1. في مثالنا تكتب الفرضيتان كما يلي:

$$H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu \neq \mu_0$$

يؤدي الاختبار إما إلى رفض H_0 ونكتب RH_0 وفي هذه الحالة نقبل الفرضية البديلة $H_1 : \mu \neq \mu_0$،

أو إلى عدم رفض الفرضية الصفرية ونكتب $R'H_0$.

μ_0 القيمة الافتراضية ل μ هي في هذه الحالة 15000 لذلك نكتب الفرضيات كما يلي:

$$H_0 : \mu = 15000 \leftrightarrow H_1 : \mu \neq 15000$$

عادة ما تكون μ_0 محددة بناءا على بيانات عينة عشوائية بسيطة $(\mu_0 = m)$، وفي هذه الحالة يمكن استخدام الخاصية $M \sim N(\mu, \sigma_m)$ لإجراء اختبار صدقية $H_0 : \mu = \mu_0$ بتقييم الفرق بين μ_0 وm أي المسافة $(\mu_0 - m)$، حيث نعلم أنه إذا كانت $H_0 : \mu = \mu_0$ صحيحة فإن :

$$P(\mu_0 - 1.64(\sigma_m) \leq m \leq \mu_0 + 1.64(\sigma_m)) = 0.90$$
$$P(\mu_0 - 1.96(\sigma_m) \leq m \leq \mu_0 + 1.96(\sigma_m)) = 0.95$$
$$P(\mu_0 - 2.58(\sigma_m) \leq m \leq \mu_0 + 2.58(\sigma_m)) = 0.99$$

وبصفة عامة نكتب:

$$P[\mu_0 - z_{1-\alpha/2}(\sigma_m) \leq m \leq \mu_0 + z_{1-\alpha/2} (\sigma_m)] = 1- \alpha$$

أو حسب الكتابة الأكثر شيوعا:

$$P(-z_{1-\alpha/2} \leq \frac{m-\mu_0}{\sigma_m} \leq z_{1-\alpha/2}) = 1-\alpha$$

حيث:

- $z_{cal} = (m - \mu_0)/\sigma_m$: تسمى **متغيرة القرار** أو **z المحسوبة**، حيث $z_{cal} \sim N(0, 1)$.
- σ_m هو خطأ المعاينة حيث: $\sigma_m = \sigma/\sqrt{n}$ في حالة المعاينة بالإرجاع (أو $n \leq 0.05N$) و $\sigma_m = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ إذا كان المجتمع محدود و المعاينة نفادية.

---

[1] عدم القبول لا يعني بالضرورة الرفض، لذلك ينبغي التمييز بينهما.

– $1-\alpha/2$: المساحة على يسار z .

– n : حجم العينة.

إذا كان m خارج المجال الذي احتماله $1-\alpha$ نرفض الفرضية الصفرية $H_0 : \mu = \mu_0$ التي حدد على أساسها هذا المجال ونقبل بالتالي الفرضية البديلة $H_1 : \mu \neq \mu_0$. تسمى هذه (الخطة) **قاعدة القرار**.

9-1-2-2. تحديد قاعدة القرار

تكتب قاعدة القرار في المثال الذي بين أيدينا، وهي قاعدة **اختبار ثنائي الاتجاه للمتوسط** (أنظر الرسم الموالي)، كما يلي:

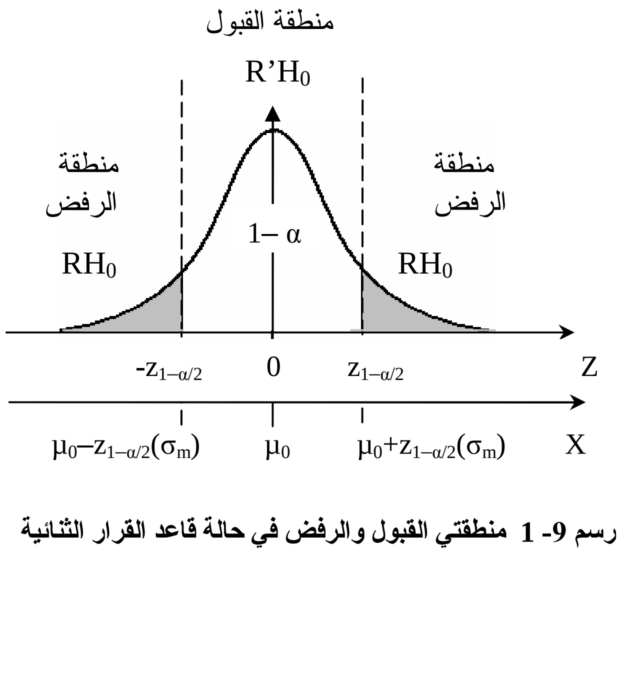
$$\begin{cases} z_c = \dfrac{m-\mu_0}{\sigma_m} \notin \left[-z_{1-\frac{\alpha}{2}} \;\; ; \;\; z_{1-\frac{\alpha}{2}} \right] \rightarrow RH_0 \\ z_c = \dfrac{m-\mu_0}{\sigma_m} \in \left[-z_{1-\frac{\alpha}{2}} \;\; ; \;\; z_{1-\frac{\alpha}{2}} \right] \rightarrow \overline{R}H_0 \end{cases}$$

بمعنى: نرفض الفرضية الصفرية إذا كان $z_c \notin [-z_{1-\alpha/2} \; ; \; z_{1-\alpha/2}]$ ولا نرفضها إذا كانت z_c تنتمي إلى المجال المذكور. تسمى z_c القيمة المحسوبة.

ويمكن أن تكتب قاعد القرار بطرق أخرى لها نفس الدلالة منها:

$$\begin{cases} \left| \dfrac{m-\mu_0}{\sigma_m} \right| > Z_{1-\frac{\alpha}{2}} \rightarrow RH_0 \\ \left| \dfrac{m-\mu_0}{\sigma_m} \right| \leq Z_{1-\frac{\alpha}{2}} \rightarrow \overline{R}H_0 \end{cases}$$

الرسم الموالي يبين في أي حالة يمكن رفض الفرضية الصفرية وفي أي حالة لا يمكن ذلك.



رسم 9- 1 منطقتي القبول والرفض في حالة قاعد القرار الثنائية

لاحظ:

- الوصول إلى قرار خاطيء يكون إما بقبول الفرضية H_0 بينما هي خاطئة أو رفضها بينما هي صحيحة: فقد تكون الفرضية H_0 صحيحة بينما نتوصل بسبب قيمة m بعيدة عن μ_0 إلى رفضها، ويسمى هذا **الخطأ من النوع الأول**، ويكتب: $P(RH_0 / H_0) = \alpha$، أو قد تكون الفرضية H_0 خاطئة بينما نتوصل بسبب قيمة m قريبة من μ_0 إلى قبولها، ويسمى هذا **الخطأ من النوع الثاني** (أنظر الملحق) ويكتب :
$$P(R'H_0 / H_1) = \beta$$
- β غير $1 - \alpha$ على عكس ما قد يبدر إلى الأذهان، فالحدثان RH_0 / H_0 و$R'H_0 / H_1$ ليسا حدثان متعاكسان وإنما الحدث المعاكس ل RH_0 / H_0 هو $R'H_0 / H_0$.
- يمكن تقليص احتمال أحد الخطأين على حساب الثاني، ولكن لا يمكن تقليص احتمال كلا الخطأين معا إلا بزيادة حجم العينة.

9-1-2-3. حساب z_{tab} أي z الجدولية:

وهي في حالتنا (اختبار ثنائي بمستوى معنوية 5%) :

$$z_{tab} = z_{1-\alpha/2} = z_{1-0.05/2} = z_{1-0.025} = z_{0.975}$$

ومن الجدول نجد أن $z_{0.975} = 1.96$.

9-1-2-4. حساب z_{cal} أي z الفعلية أو المحسوبة أي متغيرة القرار:

وهي المتغيرة المعيارية ل M:

$$z_{cal} = \frac{m-\mu_0}{\sigma_m} = \frac{15800-15000}{1500/\sqrt{100}} = 5{,}33$$

9-1-2-5. القرار:

نقرر قبول أو رفض H_0 حسب قاعدة القرار. وفي حالتنا نرفض H_0 لأن $z_{cal} > z_{tab}$ ونقبل H_1 أي أن متوسط دخل الخريج حديث التوظيف ليس 15000دج.

9-1-3. الاختبار أحادي الاتجاه للمتوسط.

يتميز الاختبار الثنائي عن الأحادي في الفرضية البديلة. هذه الأخيرة هي عدم مساواة في الاختبار الثنائي وأكبر تماما أو أصغر تماما (حسب كون الاختبار من اليمين أم من اليسار) في الاختبار الأحادي، وهذا يترتب عليه تغيير في قاعدة القرار. يوجد نوعان من الاختبار أحادي الاتجاه:

9-1-3-1. الاختبار أحادي الاتجاه من اليمين.

لنرجع إلى المثال السابق مع تغيير محدد هو أننا نريد اختبار ما إذا كان متوسط الدخل للخريج 15000دج أم أكثر (اختبار من اليمين).

أ- **الفرضيات** : $H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu > \mu_0$

في هذه الحالة $\mu_0 = 15000$ لذلك نكتب : $H_0 : \mu = 15000 \leftrightarrow H_1 : \mu > 15000$

ب- **قاعدة القرار**:

$$\begin{cases} z_{ca} = \dfrac{m - \mu_0}{\sigma_m} > z_{1-\alpha} \rightarrow RH_0 \\ z_c = \dfrac{m - \mu_0}{\sigma_m} \leq z_{1-\alpha} \rightarrow \overline{R}H_0 \end{cases}$$

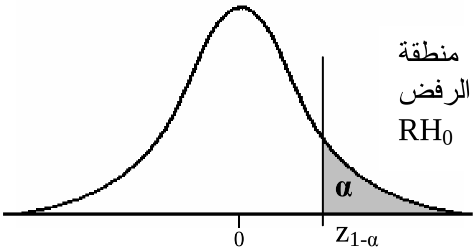
ج- **حساب z_{tab} الجدولية**: (اختبار على اليمين بمستوى معنوية 5%) :

$$z_{tab} = z_{1-\alpha} = z_{1-0.05} = z_{0.95} = 1.645$$

د- **حساب z_{cal} أو اختصارا z_c الفعلية**:

$$z_{cal} = \frac{m-\mu_0}{\sigma_m} = \frac{m-\mu_0}{\sigma/\sqrt{n}} = \frac{15800-1500}{1500/\sqrt{100}} = 5.33$$

ه- **القرار**: نرفض H_0 لأن $z_{cal} > z_{tab}$ ونقبل H_1 أي أن متوسط دخل الخريج حديث التوظيف ليس 15000دج وإنما هو أكبر.



رسم 9-2. منطقة الرفض للاختبار أحادي الاتجاه من اليمين

9-1-3-2. الاختبار أحادي الاتجاه من اليسار

نعود إلى مثالنا ونفترض أن متوسط العينة كان 14200دج ونريد أن نختبر ما إذا كان متوسط الدخل مساويا أم أقل من 15000دج.

أ- **الفرضيات** : $H_0 : \mu = 15000 \leftrightarrow H_1 : \mu < 15000$

ب- **قاعدة القرار**:

$$\begin{cases} z_{cal} = \dfrac{m-\mu_0}{\sigma_m} < -z_{1-\alpha} \rightarrow RH_0 \\ z_{cal} = \dfrac{m-\mu_0}{\sigma_m} \geq -z_{1-\alpha} \rightarrow \overline{R}H_0 \end{cases}$$

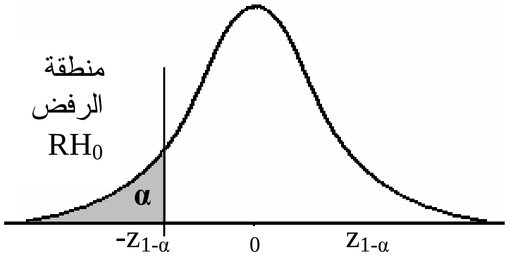
ج- **حساب z_{tab}**: (اختبار على اليسار بمستوى معنوية 5 %) :

$$z_{tab} = -z_{1-\alpha} = -z_{1-0.05} = -z_{0.95} = -1.645$$

د- **حساب z_{cal}** :

$$z_{cal} = \frac{m-\mu_0}{\sigma_m} = \frac{m-\mu_0}{\sigma/\sqrt{n}} = \frac{14200-1500}{1500/\sqrt{100}} = -5.33$$

ه- **القرار**: نرفض H_0 لأن $z_{cal} < z_{tab}$ ونقبل H_1 أي أن متوسط دخل الخريج حديث التوظيف أقل من 15000دج .



رسم 2 منطقة الرفض للاختبار أحادي الاتجاه من اليسار

9-1-4. استخدام S كمقدر ل σ في اختبار المتوسط.

في الواقع غالبا ما يكون الانحراف المعياري للمجتمع مجهولا ونحتاج بالتالي إلى استخدام الانحراف المعياري للعينة (S) عند حساب σ_m (أنظر درس التقدير)، حيث نعوض العبارة

$$\sigma_m = \frac{\sigma}{\sqrt{n}} \quad \text{ب} \quad \sigma_m = \frac{S}{\sqrt{n-1}} \quad \text{أو} \quad \sigma_m = \frac{\widehat{S}}{\sqrt{n}}$$

في هذه الحالة نستخدم التوزيع ستيودنت إلا إذا كانت العينة كبيرة فنستخدم التوزيع الطبيعي.

مثال: في المثال السابق نفترض أن الانحراف المعياري للدخل الشهري للطالب مجهول، لكن الانحراف المعياري للعينة $S = 1600$. كيف يمكن اختبار ما إذا كان الدخل الشهري أقل من 15000دج؟

الخطوات أ، ب وج تبقى بدون تغيير.

د- **حساب z الفعلية:**

$$z_{cal} = \frac{m-\mu_0}{\sigma_m} = \frac{m-\mu_0}{S/\sqrt{n-1}} = \frac{14200-1500}{1600/\sqrt{100-1}} = -4.97$$

ه- **القرار:** نرفض H_0 لأن $z_{cal} < z_{tab}$ ونقبل H_1 أي أن متوسط دخل الخريج حديث التوظيف ليس 15000دج وإنما هو أقل.

9-1-5. استخدام التوزيع t في اختبار المتوسط.

في حالة $n < 30$ وσ مجهولا، لا يمكن استخدام التوزيع الطبيعي، ولكن لدينا:

$$\frac{m-\mu}{S/\sqrt{n-1}} \sim t_{n-1} \quad \text{و تحت } H_0 \ (\mu = \mu_0): \quad \frac{m-\mu_0}{S/\sqrt{n-1}} \sim t_{n-1}$$

يمكن إذا استخدام توزيع ستيودنت (بشرط أن يكون توزيع المجتمع طبيعيا أو على الأقل جرسي الشكل).

و تتغير قاعدة القرار تبعا لهذا التغيير فتكتب في حالة الاختبار الثنائي كما يلي:

$$\begin{cases} \left|\dfrac{m-\mu_0}{S/\sqrt{n-1}}\right| > t_{n-1;1-\alpha/2} \rightarrow RH_0 \\ \left|\dfrac{m-\mu_0}{S/\sqrt{n-1}}\right| \leq t_{n-1;1-\alpha/2} \rightarrow \overline{R}H_0 \end{cases}$$

في حالة اختبار من اليمين: $\begin{cases} \dfrac{m-\mu_0}{S/\sqrt{n-1}} > t_{n-1;1-\alpha} \rightarrow RH_0 \\ \dfrac{m-\mu_0}{S/\sqrt{n-1}} \leq t_{n-1;1-\alpha} \rightarrow \bar{R}H_0 \end{cases}$

في حالة اختبار من اليسار: $\begin{cases} \dfrac{m-\mu_0}{S/\sqrt{n-1}} < -t_{n-1;1-\alpha} \rightarrow RH_0 \\ \dfrac{m-\mu_0}{S/\sqrt{n-1}} \geq -t_{n-1;1-\alpha} \rightarrow \bar{R}H_0 \end{cases}$

9-1-6. خلاصة

يتم اختبار الفرضيات من خلال 5 خطوات متتالية هي: صياغة الفرضيات (الصفرية والبديلة)، تحديد قاعدة القرار (متى نرفض أو لا نرفض H_0)، حساب القيمة الجدولية للمتغيرة، حساب القيمة الفعلية للمتغيرة، واتخاذ القرار.

تتحدد كيفية إتمام كل خطوة حسب طبيعة الاختبار (ثنائي أو أحادي الاتجاه)، حسب طبيعة المجتمع وطبيعة وحجم العينة، ... وتستخدم في ذلك نظريات توزيع المعاينة.

9-1-7. ملحق

احتمال الخطأ في القرار.

ذكرنا أنه عند اتخاذ القرار يمكن أن نقع في أحد خطأين، الخطأ الأول هو رفض الفرضية H_0 بينما هي صحيحة، واحتماله هو مساحة منطقة الرفض أي α، أما الخطأ من النوع الثاني فهو قبول الفرضية H_0 بينما هي خاطئة، ورمزنا لاحتماله ب β ، فكيف يمكن حساب β ؟.

لتوضيح ذلك نعود إلى المثال المعطى في بداية الدرس، ونفترض أن دخل الخريج الجامعي في سنة عمله الأولى هو 15000 دج الذي يمثل متوسط دخل العينة المدروسة المكونة من 100 خريج، مع العلم أن $\sigma = 1500$.

احتمال الخطأ من النوع الأول: ترفض الفرضية H_0 إذا وجد أن m لا ينتمي إلا مجال القبول:

$$m \notin [15000 - 1.96(1500/100)\ ;\ 15000 + 1.96(1500/100)]$$

$$P(RH_0/H_0) = P(RH_0/\ X \sim N(15000, 1500))$$
$$= P(m \notin [15000 - 1.96(1500/100)\ ;\ 15000 + 1.96(1500/100)]/X \sim N(15000, 1500))$$
$$= P(m \notin [14706\ ;\ 15294]/X \sim N(15000, 1500))$$
$$\Rightarrow P(RH_0/H_0) = \alpha = 0.05$$

بصفة عامة الاحتمال (غير الشرطي) $P(RH_0)$ يتغير حسب القيمة الحقيقة ل μ كما سنرى عند بحث قوة وفعالية الاختبار.

احتمال الخطأ من النوع الثاني (قبول فرضية خاطئة): تقبل الفرضية H_0 إذا وجد أن m تنتمي إلى مجال القبول،

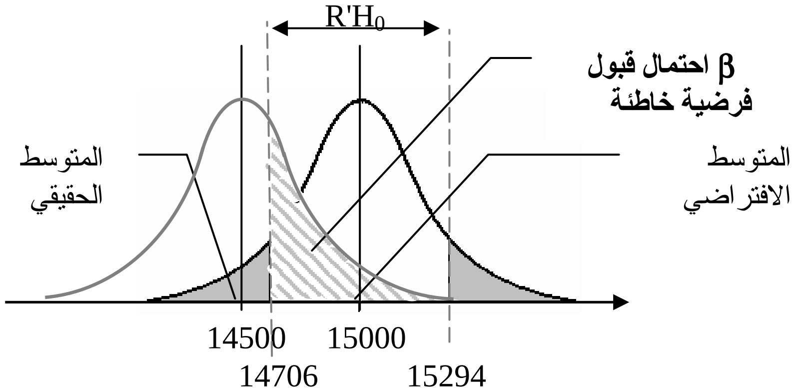
$$m \in [15000 - 1.96(1500/100)\ ;\ 15000 + 1.96(1500/100)]$$
$$m \in [14706\ ;\ 15294]$$

ويكون هذا القبول خطأ إذا كان المتوسط الحقيقي للمجتمع μ مختلفا عن $m = 15000$. ففي حالـــة $\mu = 14500$ مثلا يكتب احتمال الخطأ من النوع الثاني ويحسب كما يلي:

$$P(R'H_0/H_1) = \beta = P(m \in [14706\ ;\ 15294]/X \sim N(14500, 1500))$$

$$\beta = P\left(\frac{14706-14500}{1500/\sqrt{100}} \leq z \leq \frac{15294-14500}{1500/\sqrt{100}}\right) = P(1.37 \leq z \leq 5.29)$$

$$= F(5.29) - F(1.37) = 0.9999 - 0.9152 \Rightarrow \beta = 0.0848$$



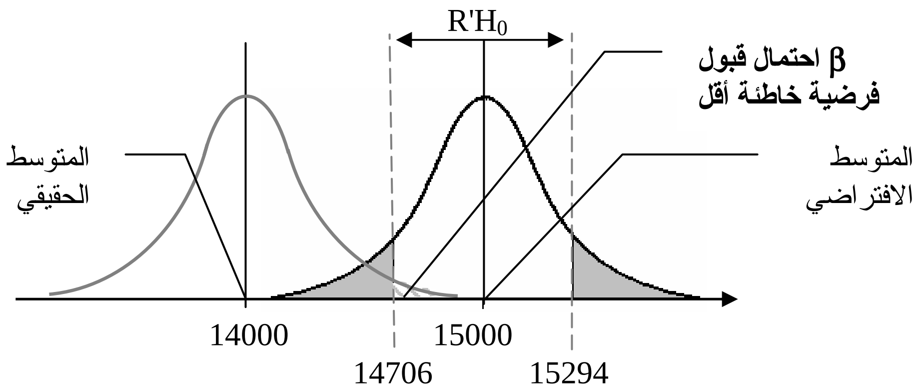
رسم 9- 2 احتمال الخطأ من النوع الثاني : قبول فرضية خاطئة.

في حالة $\mu = 14000$ ، يكون احتمال الخطأ من النوع الثاني أقل:

$$P(R'H_0/H_1) = \beta = P(m \in [14706\ ;\ 15294]/X \sim N(14000, 1500))$$

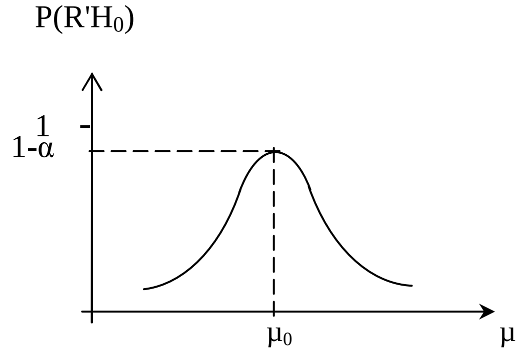
$$\beta = P\left(\frac{14706-14000}{1500/\sqrt{100}} \leq z \leq \frac{15294-14000}{1500/\sqrt{100}}\right) = P(4.71 \leq z \leq 8.63)$$

$$= F(8.63) - F(4.71) = 0.00000126 < 0.0848$$

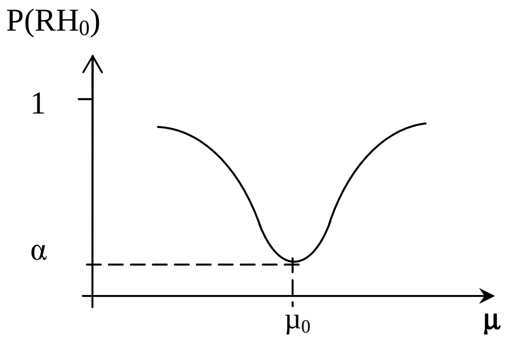


قوة الاختبار وفعاليته.

يقيس احتمال رفض الفرضية الصفرية $P(RH_0)$ **قوة الاختبار** أي القدرة على رفض الفرضية، فيما يقيس احتمال قبولها $P(R'H_0)$ **فعالية الاختبار** (أنظر الرسم أدناه). وتتفاوت الاختبارات من حيث القوة والفعالية بحسب تصميم الاختبار وبحسب نوعه[1]. ويتوقف احتمال قبول الفرضية الصفرية واحتمال رفضها على القيمة الحقيقية ل μ لكن ليس بنفس الطريقة؛ وهو ما يبينه الرسم التالي الخاص باختبار المتوسط في حالة المجتمع الطبيعي.



رسم 9- 3 (1)
منحنى فعالية الاختبار



(2) منحنى قوة الاختبار

9-2. اختبار النسبة واختبار التباين

اختبار النسبة
اختبار التباين

9-2-1. اختبار النسبة

يتعلق هذا الاختبار بنسبة مفردات المجتمع التي تتصف بخاصية ما (p)، حيث يؤكد الاختبار أو ينفي صحة فرضية معينة بخصوص قيمة p . يرمز للقيمة الافتراضية ب p_0 وتكتب الفرضية كما يلي: $H_0 : p = p_0$

للقيام بالاختبار نستخدم خصائص p' النسبة في العينة (أنظر توزيع المعاينة للنسبة : نظرية 6).

$$E(p') = \mu_{p'} = p \qquad ; \qquad \sigma^2_{p'} = \frac{pq}{n}$$

عند $n \geq 30$: $p' \approx N(p, \sigma_{p'})$ (قانون موافر– لابلاس)

استنادا إلى هذه الخصائص وتحت H_0 : $p' \approx N\left(p_0 ; \sqrt{\frac{p_0 q_0}{n}}\right)$

و من ثم يمكن تحديد قاعدة القرار بحسب طبيعة الاختبار كما يلي:

في حالة الاختبار الثنائي:
$$\begin{cases} RH_0 \ si \left| \dfrac{p' - p_0}{\sqrt{p_0 q_0 / n}} \right| > z_{1-\alpha/2} \ . \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

[1] الاختبارات **بدون معالم** (Les tests non paramétriques)هي عادة أكثر قوة من الاختبارات **بمعلم** (tests paramétriques)، الاختبارات بمعالم هي التي تتضمن فرضية بخصوص متوسط المجتمع أو طبيعة توزيعه أما الاختبارات بدون معالم فهي التي لا نضع فيها فرضيات بخصوص معالم المجتمع. سوف نقتصر في دراستنا على النوع الأول من الاختبارات دون الثاني، هذا الأخير يمكن مطالعته مثلا في: رونالد سيهيسات (Ronald Céhéssat) 1976 ص 253.

في حالة اختبار من اليمين:
$$\begin{cases} RH_0 \ si \dfrac{p'-p_0}{\sqrt{p_0q_0/n}} > z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليسار:
$$\begin{cases} RH_0 \ si \dfrac{p'-p_0}{\sqrt{p_0q_0/n}} < -z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

مثال: تقدر الدوائر الرسمية نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم ب 70% وجدت دراسة أجريت على عينة من 900 طالب أن نسبة الحصول على عمل 67% . كيف يمكن اختبار بمستوى معنوية 5 % ما إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها ؟

$$H_0 : p = 0.70 \leftrightarrow H_1 : p < 0.70$$

$$\frac{p'-p_0}{\sqrt{p_0q_0/n}} = \frac{0.67-0.7}{\sqrt{0.7(0.3)/900}} \cong -196.34 < -z_{1-0.05} = -1.64$$
=> RH_0

9-2-2. اختبار التباين

لاختبار صدقية فرضية بخصوص قيمة تباين مجتمع ما،

$$H_0 : \sigma^2 = \sigma_0^2 \ \leftrightarrow H_1 : \sigma^2 = \sigma_0^2$$

نستعمل المقدر غير المنحاز $\hat{S}^2 = \dfrac{\sum_i (X_i - m)^2}{n-1}$ حيث في حالة العينة الكبيرة ($n \geq 50$ في أحسن الأحوال)،

وتحت H_0 فإن:

$$T = \frac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \approx N(0.1). \qquad \mu_4 = E(X-\mu)^4$$

حيث μ_4 هو العزم المركزي من الدرجة الرابعة. وبهذا الشكل تكتب قاعدة القرار للاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \ si \left| \dfrac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \right| > z_{1-\alpha/2} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

وفي حالة μ_4 مجهول يمكن استخدام كمقدر : $m_4 = E(x_i - m)^4$.

وإذا كان المجتمع طبيعيا، حيث $\mu_4 = 3\sigma^4$ ، فإن **متغيرة القرار** يمكن أن تكتب كما يلي:

$$T = \frac{\hat{S}^2 - \sigma_0^2}{\sigma_0^2 \sqrt{2/n}} \approx N(0.1).$$

3-9. اختبار المقارنة بين مجتمعين

اختبار تساوي متوسطي مجتمعين
اختبار تساوي تبايني مجتمعين

يتناول هذا الاختبار مقارنة بين مجتمعين من خلال المتوسط أو التباين لكل منهما ... وسنركز هنا على متغيرة القرار، إذ من السهل على الطالب استنتاج كيفية إتمام الخطوات الأخرى على ضوء ما سبق.

1-3-9. اختبار تساوي متوسطي مجتمعين

الغرض من الاختبار هو تأكيد أو نفي تساوي متوسطي مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين.

تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي: $H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2$

لتحديد متغيرة القرار نميز بين حالة كون تباينا المجتمعين معلومين وحالة كون تباينا المجتمعين مجهولين، فنعتمد على متغيرة القرار T أو 'T بحسب الحالة.

1-1-3-9. تباينا المجتمعين معلومين

1– المجتمعين طبيعيين:

$$T = \frac{m_1 - m_2}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}} \sim N(0,1)$$

2– مجتمعين ما $(n_1 , n_2 \geq 30)$:

$$T = \frac{m_1 - m_2}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}} \approx N(0,1)$$

2-1-3-9. تباينا المجتمعين مجهولين

1– المجتمعان طبيعيان:

إذا كان تباينا المجتمعين متساويين $T' = \frac{m_1 - m_2}{\sqrt{\frac{n_1 S^2{}_1 + n_2 S^2{}_2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$

2– مجتمعين ما $(n_1 , n_2 \geq 30)$: $T = \frac{m_1 - m_2}{\sqrt{\frac{S^2{}_1}{n_1 - 1} + \frac{S^2{}_2}{n_2 - 1}}} \approx N(0,\ 1)$

مثال: نسحب من مجتمعين طبيعيين متساويي التباين عينتين حجم الأولى 18 وحجم الثانية 21. وجدنا النتائج التالية:

$m_1 = 81,\ m_2 = 76,\ S^2{}_1 = 9,\ S^2{}_2 = 8.$

كيف يمكن إجراء اختبار تساوي متوسطي المجتمعين بمستوى معنوية 5% ؟

$$H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2$$

$$T' = \frac{m_1 - m_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{81 - 76}{\sqrt{\frac{18(9) + 21(8)}{18 + 21 - 2}\left(\frac{1}{18} + \frac{1}{21}\right)}} \cong 5.43$$

$$t_{0.975;\,37} \cong 2.336 < 5.43 \qquad \Rightarrow RH_0$$

9-3-2. اختبار تساوي تبايني مجتمعين

الغرض من الاختبار هو تأكيد أو نفي تساوي تباينا مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين.

تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي: $H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$

نعتمد في الاختبار على متغيرة القرار (T أو T') بحسب كون المجتمعين طبيعيين أم غير ذلك.

9-3-2-1. مجتمعين طبيعيين

1– الحالة العامة:

$$T' = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_1 - 1;\, n_2 - 1}$$

2- في حالة $n_1 , n_2 \geq 30$

$$T = \frac{1}{2}\ln\left(\frac{\hat{S}_1^2}{\hat{S}_2^2}\right) / \sqrt{\frac{1}{2}\left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right)} \approx N(0;1)$$

9-3-2-2. مجتمعين ما ($n_1, n_2 \geq 50$)

1– $\mu_4^{(1)} ; \mu_4^{(2)}$ معروفين :

$$T = (\hat{S}_1^2 - \hat{S}_2^2) \;/\; \sqrt{\frac{\mu_4^{(1)} - \hat{S}_1^4}{n_1} - \frac{\mu_4^{(2)} - \hat{S}_2^4}{n_2}} \approx N(0;1)$$

2– في حالة $\mu_4^{(1)} ; \mu_4^{(2)}$ غير معروفين : نعوض μ_4 ب m_4 .

مثال : نسحب من مجتمعين طبيعيين عينتين حجم الأولى 18 وحجم الثانية 21. وجدنا النتائج التالية:

$m_1 = 81,\ m_2 = 76,\ S^2_1 = 9,\ S^2_2 = 8.$

كيف يمكن اجراء اختبار تساوي متوسطي المجتمعين بمستوى معنوية 5 % ؟

$$H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$$

$\hat{S}^2_1 = S^2_1 \,.\, n_1 / (n_1 - 1) = 9\,(18)/17 \approx 9.53$;

$\hat{S}^2_2 = S^2_2 \,.\, n_2 / (n_2 - 1) = 8\,(21)/20 = 8.4$

$\hat{S}^2_1 / \hat{S}^2_2 \approx 1.135$; $F_{0.05\,;\,17\,;\,20} \approx 2.17$

$$T < F_{\alpha;\, n_1 - 1;\, n_2 - 1} \Rightarrow \overline{R}H_0$$

4-9. اختبار ك2 أو اختبار التجانس و اختبار التعديل.

9-4-1. اختبار التجانس

لنعد إلى اختبار النسبة، ونفترض أن لدينا عددا k من الخصائص المتنافية، نسبة تحقق كل منها في المجتمع هي p_i حيث $\sum p_i = 1$. نريد اختبار فرضية حول قيم هذه النسب:

$$H_0 : p_i = p_{i0} \ \ (i = 1, 2, . . . k) \leftrightarrow H_1 : p_i \neq p_{i0}$$

(الفرضية البديلة هي أن إحدى النسب الافتراضية p_{i0} على الأقل غير مساوية للقيمة الحقيقية.)

متغيرة القرار : لإنجاز الاختبار نستخرج عينة (نرمز لحجم العينة ب n) ونحسب فيها ni عدد مرات تحقق كل خاصية. إذا تحققت الشروط الثلاثة التالية:

$n \geq 30$ ، $np_{i0} \geq 1$ وعلى الأقل في 80 % من الحالات $np_{i0} \geq 5$ نبرهن أن :

$$T = \sum_i \frac{(n_i - np_{i0})^2}{np_{i0}} \approx \chi^2_{k-1}$$

مثال: يتقدم إلى انتخابات معينة 3 مرشحين: أ، ب وج. نريد اختبار فرضية بمستوى معنوية 5 % حول شعبيتهم كما يلي:

$$H_0 : p_1 = 0.4, p_2 = 0.35, p_3 = 25$$

أجري استجواب ل $n = 400$ ناخب فكان توزعي فئات المساندين على التوالي : $n_i = 170, 135, 95$

لدينا $n = 400 \geq 30$، و الأعداد الافتراضية:

$$np_{i0} = 400(0.4) = 160, 400(0.35) = 140, 400(0.25) = 100 \geq 1$$

و الشرط $np_{i0} \geq 5$ في أكثر من 80 % من الحالات محقق.

$$T = \sum_i \frac{(n_i - np_{i0})^2}{np_{i0}} = \frac{(170-160)^2}{160} + \frac{(135-140)^2}{140} + \frac{(95-100)^2}{100} = 1.05$$

$X^2_{2\,;\,0.95} = 5.99 > 1.05 \Rightarrow H_0$.

بما أن T أقل بكثير من X^2 يمكن قبول الفرضية الصفرية.

9-4-2. اختبار التعديل

تستخدم هذه الطريقة أيضا لاختبار تعديل توزيع معين بتوزيع آخر، وفي هذه الحالة نقارن بين تكرارات العينة (التكرارات الحقيقية) n_i وتكرارات افتراضية n_{i0} ، حيث تصاغ الفرضيات كما يلي:

على الأقل إحدى التكرارات النظرية n_{i0} غير مساوية للتكرار الحقيقي n_i : $H_0: n_i = n_{i0} \ \ (i = 1, 2, \dots k) \leftrightarrow H_1$

لإجراء الاختبار نستخدم القانون التالي:

$$T = \sum_i \frac{(n_i - n_{i0})^2}{n_{i0}} \approx \chi^2_{k-m-1}$$

حيث m عدد معالم المجتمع المقدرة انطلاقا من بيانات العينة لتحديد التكرارات النظرية.

مثال. طلبنا من 10 أفراد أن يعطونا كل منهم بطريقة عشوائية 10 أرقام تتراوح بين 0 و 9. نريد أن نختبر، بمستوى معنوية 5%، ما إذا كانت العينات المحصلة فعلا عشوائية علما أن المشاهدات الفعلية n_i من كل رقم كانت كالتالي:

x_i	0	1	2	3	4	5	6	7	8	9
n_i	10	8	9	14	8	9	11	9	12	10

الحل.

أن تكون العينات عشوائية يعني أن يكون لكل رقم من الأرقام 0 إلى 9 نفس الاحتمال في أن يتم اختياره في كل مرة. معنى هذا أنه إذا اعتبرنا الرقم المختار من قبل الفرد في كل مرة كمتغيرة عشوائية X فإن X تتبع توزيع احتمالي متعادل (uniforme) أي أن لكل القيم نفس الاحتمال 0.1. لدينا إذا توزيعان فعلي ونظري نريد اختبار المطابقة أو التقارب بينهما، و هو ما يسمى باختبار التعديل (Teste d'ajustement).

x_i	0	1	2	3	4	5	6	7	8	9
n_i	10	8	9	14	8	9	11	9	12	10
n_{io}	10	10	10	10	10	10	10	10	10	10

$$T = \sum_{i=1}^{k} \frac{\left(n_i - n_{i_0}\right)^2}{n_{i_0}} = \frac{1}{10}\left(0 + 2^2 + 1^2 + 4^2 + 2^2 + 1^2 + 1^2 + 1^2 + 2^2 + 0\right) = 3.2$$

$$\chi^2_{10-1,\,0.95} = 16.9 > 3.2$$

بما أن T أقل بكثير من X^2 يمكن قبول الفرضية الصفرية بأن X تتبع التوزيع المتعادل ما يعني أن الأرقام اختيرت بطريقة عشوائية.

مثال 2. لمدة ساعتين تم عد X : عدد الزبائن الذين يصلون في الدقيقة إلى شباك معين. نريد اختبار بمستوى معنوية 0.05 فرضية أن X تتبع توزيع بواسون. المشاهدات كانت كالتالي:

xi	0	1	2	3	4	5	6	7	8	9	10	11
ni	4	9	24	25	22	18	6	5	3	2	1	1

الحل.

لتحديد القيم النظرية حسب التوزيع النظري بواسون نحتاج إلى تحديد معلمة هذا الأخير. نعلم أن معلمة توزيع بواسون تساوي التوقع وتساوي في الوقت ذاته التباين. بحساب هذين المؤشرين للمتغيرة X نجدهما على التوالي 3.7 و 4.41 لذلك نأخذ كقيمة تقديرية ل λ القيمة 4. نحسب إذا باستخدام توزيع بواسون بمعلمة 4 احتمالات القيم من 0 إلى 11 ثم نضرب هذه الاحتمالات في حجم العينة n = 120 للحصول على القيم n_{i0} مع التقريب لحذف الأرقام بعد الفاصلة و الحرص على أن يبقى المجموع 120.

xi	0	1	2	3	4	5	6	7	8	9	10	11	∑
ni	4	9	24	25	22	18	6	5	3	2	1	1	120
nio	**2**	9	18	23	23	19	13	7	3	2	1	0	120

نعيد تجميع القيمتين الأولين والأربع الأخيرة لتلبية الشرط $np_{i0} \geq 5$ ما يجعل عدد القيم k = 8 بدلا من 12. بهذا تكون عدد درجا ت الحرية 6 = 1-1-8 .بما أن هنا معلم واحد للتوزيع النظري تم تقديره انطلاقا من العينة.

$$T = \sum_{i=1}^{k} \frac{\left(n_i - n_{i_0}\right)^2}{n_{i_0}} = 7.14 \qquad \chi^2_{6,\,0.95} = 12.6 > 7.14$$

بما أن T أقل بكثير من X^2 يمكن قبول الفرضية الصفرية بأن عدد الزبائن X الذين يصلون إلى الشباك في الدقيقة يتبع التوزيع بواسون.