الفصل الثالث:

اختبار الفرضيات

1. مفاهيم أساسية

قبل التطرق لاختبارات الفرضيات نذكر ببعض المفاهيم الخاصة التي تمكن من فهمها وتطبيقها، والمتمثلة فيما يلي:

- **الفرضية الإحصائية:** هي أية مقولة تتعلق بمعلمة المجتمع الإحصائي أو بشكل توزيعه وتحتمل الصحة أو الخطأ.
- **الفرضية الصفرية:** وتسمى فرضية العدم وتصاغ عادة بحيث تنفي وجود فروق جوهرية بين معالم المجتمع وإحصاءات العينة، وهي الفرضية التي يتم اختبار إمكانية رفضها على اعتبار أنها صحيحة ويرمز لها بالرمز H_0.
- **الفرضية البديلة:** هي فرضية مكملة لفرضية العدم يتم قبولها عند رفض H_0 أو رفضها عند قبول H_0 ويرمز لها بالرمز H_1.
- **دالة الاختبار:** هي إحصاءة للعينة تساعد على اتخاذ القرار.
- **الخطأ من النوع الأول:** هو رفض فرضية العدم H_0 وهي صحيحة، ويرمز لاحتمال الوقوع في هذا النوع من الخطأ بـ α ويسمى بمستوى أهمية الاختبار أو مستوى المعنوية.
- **الخطأ من النوع الثاني:** وهو قبول H_0 عندما تكون خاطئة، ويرمز لاحتمال الوقوع في هذا النوع من الخطأ بـ β.
- **قوة الاختبار:** هي مقياس لكفاءة الاختبار وبالتالي لدقة الاستدلال الإحصائي، وهي مكملة لاحتمال الخطأ من النوع الثاني، أي:

$$P\left(\mathrm{H}_0 \text{ رفض} / \mathrm{H}_0 \text{ خاطئة}\right) = 1 - \beta$$

- **المنطقة الحرجة:** هي المساحة التي تقع أسفل منحنى المستخدم في عملية التحليل الإحصائي وتمثل احتمال رفض H_0 وهي صحيحة، وتسمى هذه المنطقة بمنطقة الرفض وتحدد حسب نوع الفرضية البديلة ويحدد قيمتها مستوى المعنوية α، وتسمى المساحة المتبقية أسفل المنحنى منطقة القبول.
- **القيمة الحرجة المعيارية:** هي قيم التوزيع الاحتمالي التي تفصل بين منطقة قبول H_0 ومنطقة رفضها.

2. الاختبارات المعلمية

هي اختبار الفرضيات الإحصائية المتعلقة بمعلمات المجتمعات الإحصائية علما أن توزيعاتها معروفة، فإذا كانت θ معلمة لمجتمع إحصائي وكانت الفرضية موضع الاختبار $\mathrm{H}_0: \theta = \theta_0$ فإن الفرضية البديلة H_1 تكون في إحدى الحالات التالية:

الفرضية البديلة ذات الطرفين	الفرضية البديلة ذات الطرف الأيسر	الفرضية البديلة ذات الطرف الأيمن
وتصاغ على الشكل $H_1: \theta \neq \theta_0$ فإذا كانت α هي مستوى المعنوية وكانت دالة الاختبار φ_0 فإننا نضع $\frac{\alpha}{2}$ في كل طرف توزيع إحصاءة الاختبار φ_0 : $$P(\varphi_0 < k_0) = P(\varphi_0 > k_1) = \frac{\alpha}{2}$$ ونقبل الفرضية H_0 إذا وقعت قيمة احصاء اختبار H_0 المحسوبة من العينة بين العددين k_0 و k_1 اللذان يمثلان الحدين الأعلى والأدنى لمجال الثقة للمعلمة θ.	وتصاغ على الشكل $H_1: \theta > \theta_0$ وفي هذه الحالة نضع α في الطرف الأيمن من توزيع إحصاءة الاختبار φ_0 وتحدد k بحيث: $P(\varphi_0 > k) = \alpha$ ونقبل H_0 اذا كانت $\theta_0 < k$ ونرفضها إذا كانت $\theta_0 > k$.	وتصاغ على الشكل $H_1: \theta < \theta_0$ وفي هذه الحالة نضع α في الطرف الأيسر من توزيع إحصاءة الاختبار φ_0 وتحدد k بحيث: $P(\varphi_0 < k) = \alpha$ ، ونقبل H_0 اذا كانت $\theta_0 > k$ ونرفضها إذا كانت $\theta_0 < k$.

3. تطبيقات لاختبار الفرضيات

من بين الاختبارات المهمة في الإحصاء الاستدلالي التي تستعمل بشكل كبير في التطبيقات العملية الاختبارات التالية:

1.3. اختبارات الفرضيات حول متوسط مجتمع طبيعي

يمكن التمييز بين حالتين:

مجتمع طبيعي تباينه معلوم

لدينا:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \backsim N(0,1)$$

أي أن مجال الثقة للمتوسط μ هو:

$$\left[\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

بمستوى $100(1-\alpha)\ \%$ من الثقة.

ملاحظة:

إذا كانت العينات العشوائية المسحوبة من مجتمعات غير طبيعية ذات تباين σ^2 معلوم وحجمها $n \geq 30$ يمكن تطبيق نظرية النهاية المركزية والحصول على نفس المجال السابق الذي يكون تقريبيا.

مجتمع طبيعي تباينه مجهول

حجم العينة $n \geq 30$

لدينا إحصاءة الاختبار هي:

$$Z = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \backsim N(0,1)$$

وعندما نعتبر صحيحة H_0 تكون قيمة الاحصاءة هي:

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

وحسب مستوى المعنوية α نجري الاختبار حسب الحالات التالية:

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

حجم العينة $n < 30$

لدينا إحصاءة الاختبار هي:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \backsim t_{n-1}$$

وعندما نعتبر صحيحة H_0 تكون قيمة الإحصاءة هي:

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

وحسب مستوى المعنوية α نجري الاختبار حسب الحالات التالية:

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

لدينا: وتكون قاعدة اتخاذ القرار إذا كانت الفرضية البديلة ذات الطرفين على الشكل التالي:

$$\begin{cases} Z_0 \epsilon \left[-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}\right] \Longrightarrow H_0 \text{ الفرضية} \\ Z_0 \notin \left[-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}\right] \Longrightarrow H_0 \text{الفرضية} \end{cases}$$

أما إذا كانت الفرضية البديلة ذات الطرف فتكون قاعدة اتخاذ القرار على الشكل التالي:

$$\begin{cases} |\bar{X}-\mu_0| \leq Z_{\,1-\alpha} \dfrac{\sigma}{\sqrt{n}} \Longrightarrow H_0 \text{ رضية} \\ |\bar{X}-\mu_0| > Z_{\,1-\alpha} \dfrac{\sigma}{\sqrt{n}} \Longrightarrow H_0 \text{الفرضية} \end{cases}$$

تكون قاعدة اتخاذ القرار إذا كانت الفرضية البديلة ذات الطرفين على الشكل التالي:

$$\begin{cases} Z_0 \epsilon \left[-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}\right] \Longrightarrow H_0 \text{ فرضية} \\ Z_0 \notin \left[-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}\right] \Longrightarrow H_0 \text{لفرضية} \end{cases}$$

أما إذا كانت الفرضية البديلة ذات الطرف فتكون قاعدة اتخاذ القرار على الشكل التالي:

$$\begin{cases} |\bar{X}-\mu_0| \leq Z_{\,1-\alpha} \dfrac{S}{\sqrt{n}} \Longrightarrow H_0 \text{ ضية} \\ |\bar{X}-\mu_0| > Z_{\,1-\alpha} \dfrac{S}{\sqrt{n}} \Longrightarrow H_0 \text{رضية} \end{cases}$$

أو $H_1: \mu > \mu_0$

أو $H_1: \mu < \mu_0$

تكون قاعدة اتخاذ القرار إذا كانت الفرضية البديلة ذات الطرفين على الشكل التالي:

$$\begin{cases} T_0 \epsilon \left[-t_{1-\frac{\alpha}{2},n-1}, t_{1-\frac{\alpha}{2},n-1}\right] \Longrightarrow \\ T_0 \notin \left[-t_{1-\frac{\alpha}{2},n-1}, t_{1-\frac{\alpha}{2},n-1}\right] \Longrightarrow \end{cases}$$

أما إذا كانت الفرضية البديلة ذات الطرف فتكون قاعدة اتخاذ القرار على الشكل التالي:

$$\begin{cases} |\bar{X}-\mu_0| \leq t_{\,1-\alpha,n-1} \dfrac{S}{\sqrt{n}} \Longrightarrow H \\ |\bar{X}-\mu_0| > t_{\,1-\alpha,n-1} \dfrac{S}{\sqrt{n}} \Longrightarrow H_0 \end{cases}$$

أمثلة:

المثال الأول:

أخذت عينة من 64 تلميذا من إحدى المدارس فوجد أن متوسط طولهم هو 155 سم، فإذا كان الانحراف المعياري للمجتمع هو 5 سم اختبر عند مستوى المعنوية 0.05 الفرضية التالية:

$$H_0: \mu = 160$$
$$H_1: \mu \neq 160$$

لدينا إحصاءة الاختبار هي:

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \backsim N(0,1)$$

وعندما نعتبر صحيحة H_0 تكون قيمة الإحصاءة هي:

$$Z_0 = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{155-160}{\frac{5}{\sqrt{64}}} = -8$$

عند مستوى المعنوية 0.05 وبما أن فرضية العدم ذات الطرفين فإن القيمة الجدولية تحسب كما يلي:

$$Z_{1-\frac{\alpha}{2}} = 1.96 \Longrightarrow -Z_{1-\frac{\alpha}{2}} = -1.96$$

نلاحظ أن القيمة المحسوبة Z_0 تقع في منطقة رفض H_0 ومنه نرفضها ونقبل الفرضية البديلة H_1.

المثال الثاني:

إذا كان متوسط ربح سهم إحدى الشركات هو 15 وحدة نقدية في العام الماضي تم أخذ عينة من 7 مساهمين عن توقعاتهم عن متوسط ربح السهم في العام الحالي فوجد أنه 17 وحدة نقدية بانحراف معياري 2 وحدة نقدية. هل توافق المساهمين في تقديرهم لارتفاع متوسط ربح السهم هذا العام عند مستوى المعنوية 0.05؟

نريد إجراء اختبار الفرضيات التالية:

$$H_0: \mu = 15$$
$$H_1: \mu > 15$$

لدينا σ مجهول و $n < 30$ ومنه احصاءة الاختبار هي:

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \backsim t_{n-1}$$

وعندما نعتبر صحيحة H_0 تكون قيمة الإحصاءة هي:

$$T_0 = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} = \frac{17-15}{\frac{2}{\sqrt{7}}} = 2.65$$

عند مستوى المعنوية 0.05 وبما أن فرضية العدم ذات الطرف فإن القيمة الجدولية تحسب كما يلي: $\alpha = 0.05 \Longrightarrow T_{1-\alpha,n-1} = T_{0.95,6} = 1.943$

نلاحظ أن القيمة المحسوبة Z_0 تقع في منطقة رفض H_0 ومنه نقبل الفرضية البديلة H_1 وأوافق المساهمين في تقديرهم لارتفاع متوسط ربح السهم هذا العام عند مستوى المعنوية 0.05.