الفصل الثاني
اختبار الفرضيات

1.2 المقدمة: Introduction

يرتبط مفهوم اختبار الفرضيات بمجال الاستدلال الاحصائي اذ يعتبر من المواضيع المهمة في التطبيقات العملية. ان الغرض من اختبار الفرضية هو مساعدة المختصين في المجالات المختلفة في اتخاذ قرارات بشان الظواهر المدروسة سواءا كانت ظواهر اقتصادية او ادارية او حيوية او غيرها من الظواهر، وذلك من خلال تقدير خصائص المجتمع المدروس عن طريق اختيار عينة مناسبة من ذلك المجتمع. وسيتضمن الفصل استعراض لبعض بعض المفاهيم الاساسية الضرورية لفهم موضوع اختبار الفرضيات فضلا عن دراسة انواع الاختبارات واستخداماتها.

2.2 مفاهيم اساسية: Basic concepts

من الضروري التطرق مبدئيا الى بعض المفاهيم الاساسية التي تساعد في الخوض والتوسع في تفاصيل اختبار الفرضيات.

3.2 الفرضية (Hypothesis):

تعرف الفرضية على انها تعبير او ادعاء او راي حول خاصية (معلمة parameter) من خصائص المجتمع المدروس او من عدة مجتمعات يراد دراستها والتاكد من صحتها، وذلك من خلال اسلوب او خطوات منطقي ومتسلسل تعرف باختبار الفرضيات.

لتحديد نوع الفرضية واسلوب اختبار البيانات يجب الاخذ بنظر الاعتبار مایاتي:

1. نوع وطبيعة البيانات: ان فهم نوع وطبيعة البيانات المدروسة هو امر مهم في تحديد نوع الفرضية واسلوب الاختبار، فقد تكون البيانات كمية او وصفية او قد تكون البيانات كاملة او غير كاملة او قد تكون بيانات مراقبة او غير مراقبة.

2. الافتراضات: يشترط التقدير في اغلب الاحيان شروطا معينة تؤدي الى اختيار اختبارات مناسبة تتلائم مع تلك الشروط، على سبيل المثال التوزيع الاحتمالي لبيانات المجتمع المدروس، او طريقة اختيار اسلوب المعاينة او حجم العينة الملائم لاجراء التقدير او الاختبار.

في مفهوم اختبار الفرضيات هنالك نوعين من الفرضيات هما:

1. **فرضية العدم Null Hypothesis:**

وهي الفرضية التي يضعها الباحث على امل ان يرفضها، او بمعنى اخر هي الفرضية التي تعبر عن راي معين يرفض تصديقه، وبالتالي ففرضية العدم تكون اما مقبولة او مرفوضة فاذا لم يتم رفضها من خلال الاختبار فيمكن ان نستنتج ان البيانات المدروسة او عملية التقدير لاتحقق دليلا كافيا للرفض، اما اذا رفضت فرضية العدم من خلال الاختبار المستخدم فيمكن القول ان عملية التقدير او البيانات المدروسة قد لا تنسجم مع فرضية العدم. يرمز عادة لفرضية العدم بالرمز H_0.

2. **الفرضية البديلة Alternative Hypothesis:**

هي الفرضية المكملة لفرضية العدم او هي الفرضية المعاكسة لفرضية العدم، ويرمز لها بالرمز H_1. تبين الفرضسة البديلة ان التقدير او البيانات المدروسة غير مطابقة لخصائص المجتمع المدروس.

على سبيل المثال في تجربة لفحص مدى تاثير علاج يخص تقليل الوزن عندها يمكن قياس متوسط اوزان مجموعة الاشخاص بعد تلقيهم علاج تخيف الوزن، هنا قد نحصل على مجموعتين احداهما تشمل الاشخاص الذين لم تتناقص اوزانهم والاخرى تتناقصت اوزانهم. عليه يمكن ان نتحقق من الادعاء الذي ينص على ان الاوزان لم تتغير (او ان الاوزان في حالتها الاصلية)، وهذا الادعاء يمثل فرضية العدم، عندها يتم اختيار عينة بحجم مناسب من المجتمع المدروس وحساب متوسط الاوزان لها، فاذا تبين ان انها اقل من متوسط الاوزان الاصلي عندها فان العلاج قد حقق النتائج ويتم رفض فرضية العدم، اما اذا كان متوسط الاوزان للعينة المسحوبة اكبر او يساوي الاوزان الاصلية فهذا يشير الى ان الادعاء بان العلاج قد حقق تقدما هو ادعاء مرفوض وغير صحيح (رفض فرضية العدم). مما تقدم فانه يمكن صياغة فرضية العدم على اساس ان متوسط الاوزان للمجموعة المدروسة يساوي متوسط الاوزان الاصلية للمجموعة وليكن 85 كغم، هنا يتم وضع فرضية العدم على امل رفضها

(بمعنى ان يكون العلاج قد ساهم في تخفيض الوزن). عندها فان الفرضية البديلة ستكون ان العلاج قد ساهم بتخفيض الوزن، اي ان:

$$H_0: \mu = 85$$
$$H_1: \mu \neq 85 \quad \dots (1)$$

وبصورة اكثر ملائمة فان:

$$H_0: \mu > 85$$
$$H_1: \mu \leq 85 \quad \dots (2)$$

$$H_0: \mu \geq 85$$

$$H_1: \mu < 85$$

ان الاختبار في الصيغة (1) يسمى بالاختبار من جانبين (Two tailed test)، اما الاختبار في الصيغة (2) فيسمى بالاختبار من جانب واحد (One tailed test).

4.2 الاختبار من جانب واحد ومن جانبين:

بافتراض ان الاختبار يخص متوسط مجتمع ما، عندها فانه يمكن صياغة فرضيات الاختبار من جانب واحد على النحو الاتي:

$$H_0: \mu = \mu_0 \quad \text{or} \quad H_0: \mu \geq \mu_0 \quad \text{or} \quad H_0: \mu \leq \mu_0$$

$$H_1: \mu < \mu_0 \quad \text{or} \quad H_1: \mu > \mu_0 \quad \text{or} \quad H_1: \mu > \mu_0$$

$$H_1: \mu > \mu_0$$

كما ويمكن صياغة فرضيات الاختبار من جانبين على النحو الاتي:

$H_0: \mu = \mu_0 \quad or \quad H_0: \mu > \mu_0 \quad or \quad H_0: \mu < \mu_0$

$H_1: \mu \neq \mu_0 \quad or \quad H_1: \mu \leq \mu_0 \quad or \quad H_1: \mu \geq \mu_0$

ان مسالة تحديد اتجاه الاختبار فيما اذا كان من جانب واحد او من جانبين مهمة جدا كونه سيؤثر بالتيجة على مايعرف بمستوى المعنوية الملائم للاختبار. ويعتمد تحديد اتجاه الاختبار اساسا على الفرضية البديلة، فاذا كانت الفرضية البديلة مركبة (تتضمن الاشارات $\neq, \geq, \leq$) هذا يعني ان الاختبار من جانبين، اما اذا كانت الفرضية البديلة بسيطة (تتضمن الاشارات $<, >$) فهذا يعني ان الاختبار من جانب واحد.

ويمكن تصنيف اتجاهات الاختبار على النحو الاتي:

$H_1: \mu < \mu_0$: اختبار من الجانب الايسر Left tailed test

$H_1: \mu > \mu_0$: اختبار من الجانب الايمن Right tailed test

5.2 مستوى المعنوية: Level of Significance

عند اتخاذ قرار بشان قبول او رفض فرضية العدم فان هذا القرار يستند الى معطيات الظاهرة او التجربة المدروسة اي الى المعلومات والبيانات التي تم الحصول عليها من العينة الماخوذة من تلك الظاهرة، وعليه فان نتائج التجربة قد لاتكون صحيحة دائما. هنا قد يقع الباحث في نوعين من الخطأ هما:

1. الخطا من النوع الاول Type I error: هو الخطا الذي يقع فيه الباحث عندما يرفض فرضية العدم وهي صحيحة (قبول الفرضية البديلة وهي خاطئة).
2. الخطا من النوع الثاني Type II error: هو الخطا الذي يقع فيه الباحث عندما يقبل فرضية العدم وهي غير صحيحة (رفض الفرضية البديلة وهي صحيحة).

يمكن من خلال الجدول (1) ادناه توضيح انواع الخطا.

جدول (1) انواع الاخطاء في اختبار الفرضيات

Decision / Hypothesis	Reject H_0	Accept H_0
H_0 is true	False Decision Type I error	True Decision
H_0 is not true	True Decision	False Decision Type II error

من خلال ماتقدم فان مستوى المعنوية يعرف على انه احتمال الوقوع بالخطا من النوع الاول، اي احتمال رفض فرضية العدم وهي صحيحة، وعادة مايرمز لمستوى المعنوية بالرمز (α)، وعليه فان:

$$\alpha = \Pr\{\text{rejecting } H_0 \text{ when } H_0 \text{ is true}\}$$

ان تحديد مستوى المعنوية يجب ان قبل يكون دقيقا بحيث يتلائم مع طبيعة الظاهرة المدروسة ويتم قبل البدء بالاختبار، فعند اختيار مستوى معنوية وليكن (0.1) هذا يعني ان مستوى الدقة المراد الاختبار ضمنه هو $(1 - 0.1 = 0.9)$، ويسمى الحد $(1 - \alpha)$ بمستوى الثقة اي كلما زاد مستوى الثقة قل احتمال الوقوع في الخطا من النوع الاول. بصورة عامة فان تحديد مستوى المعنوية يجب ان يكون متلائما مع طبيعة الظاهرة المراد اختبارها، فمن المنطق ان يتم اختيار متسوى معنوية صغير جدا عندما يكون الاختبار حساسا في الدراسات الطبية مثل اجراء العمليات الجراحية اي ان يكون مستوى الثقة المطلوب عال، ويمكن ان يكون مستوى المعنوية اكبر اذا كان الاختبار غير حساس.

اما احتمال الوقوع في الخطا من النوع الثاني اي احتمال رقبول فرضية العدم وهي غير صحيحة، ويرمز له بالرمز (β)، اي ان:

$$\beta = \Pr\{\text{accepting } H_0 \text{when } H_0 \text{ is false}\}$$

عندها فان الحد $(1 - \beta)$ يسمى بقوة الاختبار (Power of test) ويعرف على انه احتمال رفض فرضية العدم وهي غير صحيحة، اي ان:

$$\text{Power of test} = 1 - \beta = \Pr\{\text{rejecting } H_0 \text{when } H_0 \text{ is false}\}$$

مما تقدم يمكن ملاحظة انه كلما كان احتمال حدوث الخطا من النوع الثاني قليل فذلك مؤشر على زيادة قوة الاختبار، كما هو موضح في الجدول (1) ادناه:

جدول (2) قيم قوة الاختبار

β	0.1	0.05	0.01	0.001
$1-\beta$	0.9	0.95	0.99	0.999

6.2 العينات الكبيرة والعينات الصغيرة:

ان المقاييس والمؤشرات الاحصائية المحسوبة من العينة يمكن اعتبارها تقديرات جيدة لمعلمات المجنمع، وان النظرية الاحصائية تنص على ان زيادة حجم يعني اقتراب قيم هذه المقدرات من قيم معلمات المجنمع. لذلك فقد تصنيف احجام العينات الى عينات كبيرة $n \geq 30$ وعينات صغيرة $n < 30$، وهناك توزيعات تقترن بمفهوم العينات الكبيرة منها التوزيع الطبيعي، وتوزيعات اخرى تقترن بمفهوم العينات الصغيرة منها توزيع t.

6.2 احصاءة الاختبار: Test Statistics

بافتراض ان (w) مقدر (مؤشر احصائي) وان $E(w)$ تمثل القيمة المتوقعة لهذا المقدر عند فرضية العدم H_0، وان الخطا المعياري (Standard Error) للمقدر هو $S.E.(w)$، عندها فان احصاءة الاختبار (ايا كان التوزيع الاحتمالي للمتغير العشوائي) ولتكن (V) تسلك سلوك د\الة احتمالية تتبع توزيعا معينا وتحسب وفق الصيغة الاتية:

$$V = \frac{w - E(w)}{S.E.(w)} \sim f(V) \ , \qquad \text{where } S.E.(w) = \sqrt{\frac{s^2(w)}{n}} = \frac{S(w)}{\sqrt{n}} \qquad \dots (3)$$

اذ ان $s^2(w)$ هو تباين (Variance) المقدر المحسوب من العينة.

وان $S(w)$ هو الانحراف المعياري (Standard deviation) للمقدر.

ان قيمة (V) الناتجة تمثل قيمة احصاءة الاختبار المحسوبة.

7.2 درجات الحرية: Degrees of freedom

ان حساب قيمة احصاءة الاختبار (V) فان ذلك يعتمد على قياسات العينة العشوائية بحجم (n)، توظف تلك العينة لتقدير معلمتين مجهولتين هما (w) و $S(w)$، وهذا يعني وضع قيد او قيود على هذه العينة تتمثل بعدد المعلمات المقدرة. لذلك فان هي عدد مفردات العينة التي يتم اختيارها للتقدير مطروحا منه عدد القيود المفروضة على تلك العينة.

8.2 خطوات اختبار الفرضيات:

يمكن ايجاز الخطوات المتبعة في اختبار فرضية العدم (V) كما ياتي:

1. تحديد فرضية العدم H_0.
2. تحديد الفرضية البديلة (H_1) ونوعها فيما اذا كانت من جانب واحد او من جانبين.
3. اختيار مستوى معنوية ملائم وذلك يعتمد على الدقة المطلوبة لاجراء للتجربة قيد الاختبار.
4. حساب قيمة احصاءة الاختبار (V) عند فرضية العدم (H_0). تسمى بقيمة احصاءة الاختبار المحسوبة ويرمز لها عادة V_{cal}.
5. تحديد القيمة الحرجة (القيمة الجدولية) استنادا الى جداول التوزيع الاحتمالي لاحصاءة الاختبار (V) واعتمادا على مستوى المعنوية ودرجات الحرية المحددين. ويطلق عليها عادة بقيمة احصاءة الاختبار الجدولية ويرمز لها عادة بالرمز V_{tab}.
6. قاعدة القرار الاحصائي اذ يتم رفض فرضية العدم اذا كانت $(|V_{cal}| > V_{tab})$ وتقبل الفرضية البديلة (H_1) وهنا يقال ان (V) معنوية، اي ان الفرق مابين w و $E(w)$ هو فرق جوهري. اما اذا كانت $(|V_{cal}| < V_{tab})$ عندها يتم قبول فرضية العدم ورفض الفرضية البديلة، وهذا يشير الى ان الفرق غير جوهري مابين w و $E(w)$.
7. القرار الحيوي او الاستنتاج ان الاختبار الاحصائي ليس مجرد رقم بحد ذاته بل ان الاهمية الرئيسية للاختبار الاحصائي تكمن في مساعدة الباحثين في المجالات الحيوية الطبية والصحية والسكانية على اتخاذ قرارات تخص الظاهرة المراد اختبارها، لذلك يجب ان تفضي نتائج الاختبارات الاحصائية الى استنتاجات حيوية.

9.2 اختبارات متوسط عينة عشوائية من مجتمع يتبع التوزيع الطبيعي:

لتكن $(x_1, x_2, \dots, x_n)$ تمثل قياسات عينة عشوائية ماخوذة من مجتمع يتبع التوزيع الطبيعي بمتوسط (μ) وتباين (σ^2)، ولتكن $(\bar{x})$ و (S^2) تمثلان وعلى التوالي الوسط الحسابي وتباين العينة. عندها فنحن بصدد اختبار فرضية العدم الاتية ضد اي فرضية بديلة:

$$H_0: \mu = \mu_0$$

اذ ان μ_0 قيمة محددة.

عندها تبرز عدة حالات يمكن اجراء الاختبار حولها وهي:

1. اذا كان تباين المجتمع (σ^2) معلوم فان احصاءة الاختبار تعطى وفق الصيغة الاتية:

$$Z = \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \qquad \text{where } Z \sim N(0,1) \qquad \dots (4)$$

2. اذا كان تباين المجتمع (σ^2) غير معلوم وحجم العينة كبير $(n \geq 30)$ فان احصاءة الاختبار تعطى وفق الصيغة الاتية:

$$Z = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \qquad \text{where } Z \sim N(0,1) \qquad \dots (5)$$

3. اذا كان تباين المجتمع (σ^2) غير معلوم وحجم العينة صغير $(n < 30)$ فان احصاءة الاختبار تعطى وفق الصيغة الاتية:

$$T = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \qquad \text{where } T \sim t(1-\alpha, n-1) \qquad \dots (6)$$

مثال 1.2: طورت احدى شركات صناعة الادوية دواءا يستخدم لعلاج ارتفاع نسبة الكوليسترول في الدم، اذ تامل الشركة في تخفيض نسبة الكوليسترول الى نسبة اقل او تساوي 5 مل مول / لتر. فاذا علمت ان عينة المرضى التي تم تجريب الدواء عليها بحجم 100 مريض ماخوذة من مجتمع له متوسط

نسبة كوليسترول تبلغ 5.2 مل مول / لتر وبتباين يبلغ 9 مل مول / لتر. فهل تعتقد الشركة ستحقق ماخططت له عند مستوى معنوية ملائم.

الحل:

- الفرضيات:

$H_0: \mu \leq 5$

$H_1: \mu > 5$

- احصاءة الاختبار:

$$Z_{cal} = \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \quad , \quad Z \sim N(0,1)$$

$$Z_{cal} = \frac{5.2 - 5}{3/\sqrt{100}} = \frac{0.2}{0.3} = 0.667$$

- مستوى المعنوية: يمكن اجراء هذا الاختبار عند مستوى معنوية 0.05.
- القيمة الحرجة (الجدولية): الاختبار من جانب واحد لذلك فان القيمة الجدولية للاختبار تستخرج من جداول التوزيع الطبيعي القياسي وكما ياتي:

$$Z_{tab} = Z_{\alpha} = Z_{0.05} = 1.645$$

- القرار الاحصائي:

لما كانت $Z_{cal} < Z_{tab}$ لذلك تقبل فرضية العدم وترفض الفرضية البديلة، وهذا يعني ان متوسط نسبة الكوليسترول في الدم للعينة التجريبية بعد تلقي العلاج الجديد هو اقل او يساوي 5 مل مول / لتر.

- القرار الحيوي:

ان العلاج الجديد قد نجح في تخفيض ارتفاع نسبة الكوليسترول.

مثال 2.2: يدعي احد الاطباء ان متوسط تقييم العمليات الجراحية التي يجريها يبلغ 18 درجة من اصل 20 درجة، وقد تبين ان هذا الطبيب قد اجرى 49 عملية جراحية سابقا حصل من خلالها متوسط تقييم بلغ 16 درجة وبانحراف معياري بلغ 3 درجات. هل تعتقد ان ادعاء الطبيب صحيح. اختر مستوى المعنوية الذي تجده ملائما للاختبار.

الحل:

- الفرضيات:

$$H_0: \mu = 18$$

$$H_1: \mu \neq 18$$

- احصاءة الاختبار:

$$Z_{cal} = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \quad , \quad Z \sim N(0,1)$$

$$Z_{cal} = \frac{16 - 18}{4/\sqrt{49}} = -3.5$$

- مستوى المعنوية: يمكن اجراء هذا الاختبار عند مستوى معنوية 0.01.
- القيمة الحرجة (الجدولية): الاختبار من جانبين لذلك فان القيمة الجدولية للاختبار تستخرج من جداول التوزيع الطبيعي القياسي وكما ياتي:

$$Z_{tab} = Z_{\alpha/2} = Z_{0.01/2} = Z_{0.005} = 2.575$$

- القرار الاحصائي:

لما كانت $|Z_{cal}| > Z_{tab}$ لذلك ترفض فرضية العدم وتقبل الفرضية البديلة، وهذا يعني ان متوسط درجات تقييم العمليات الجراحية التي اجراها الطبيب لـ 49 مريض لم يحصل من خلالها على متوسط يساوي 18 درجة.

- القرار الحيوي:

ان الطبيب لم ينجح في رفع مستوى تقييم عملياته الجراحية الى 18 درجة من اصل 20 درجة.

مثال 3.2: تم اجراء فحص نسبة السكر في الدم لعينة مؤلفة من 16 شخص، فتبين ان متوسط نسبة السكر في الدم لهذه العينة بلغ 96 وحدة لكل 100 مللتر وبانحراف معياري بلغ 36 مللتر لكل 100 وحدة. فهل يمكن القول ان متوسط نسبة السكر في الدم للعينة المفحوصة يقع ضمن المستوى الطبيعي لنسبة السكر في الدم البالغة 120 وحدة لكل 100 مللتر عند مستوى معنوية ملائم.

الحل:

- الفرضيات:

$H_0: \mu > 120$

$H_1: \mu \leq 120$

- احصاءة الاختبار:

$$T_{cal} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \quad , \quad T \sim t_{\left(1-\frac{\alpha}{2}, n-1\right)}$$

$$T_{cal} = \frac{96 - 120}{36/\sqrt{16}} = \frac{-24}{9} = -2.667$$

- مستوى المعنوية: يمكن اجراء هذا الاختبار عند مستوى معنوية 0.05.
- القيمة الحرجة (الجدولية): الاختبار من جانبين لذلك فان القيمة الجدولية للاختبار تستخرج من جداول توزيع t وكما ياتي:

$$T_{tab} = t_{(1-\alpha/2, n-1)} = t_{(0.975,15)} = 2.132$$

- القرار الاحصائي:

لما كانت $|T_{cal}| > T_{tab}$ لذلك ترفض فرضية العدم وتقبل الفرضية البديلة، وهذا يعني ان متوسط نسبة السكر في الدم للعينة اقل او يساوي 120 وحدة لكل 100 مللتر.

- القرار الحيوي:

ان المفحوصين لايعانون من ارتفاع نسبة السكر في الدم.