اعداد: نادية علي عايد
قسم الاحصاء

الاستدلال الإحصائي: اختبار الفروض

اختبار الفروض

اختبار الفروض عن خصائص المجتمع (مثل μ و σ) هو جانب أساسي أخر من جوانب الاستدلال والتحليل الإحصائي. وفي اختبار الفروض نبدأ بعمل فرض ما عن خاصية المجتمع غير المعلومة. ثم نأخذ عينة عشوائية من المجتمع، وعلى أساس الخاصية المناظرة في العينة، أما أن نقبل وإما أن نرفض الفرض بدرجة ثقة محددة.

مثال (1):نفرض أن باحثا اجتماعيا ادعى ان متوسط اعمار طلاب الجامعة لايختلف عن متوسط أعمار الطالبات. للتأكد من ذلك فإن الشئ الطبيعي أن نقوم بحصر اعمار الطلاب والطالبات ومنها نحسب المتوسط لكل منهما ثم نقرر من منهما اكبر..ولكن عملية الحصر صعبة ومجهدة لذلك نضطر الى اختيار عينة عشوائية من بين الطلاب وعينة عشوائية من بين الطالبات ونحسب متوسط العمر في كل عينة منهما,,,,فإذا كان متوسط عمر الطالب هو 24 وكان متوسط عمر الطالبة هو 22 فهل يعني ذلك ان متوسط عمر الطالب اكبر من متوسط عمر الطالبة ؟؟ هل الفرق راجع لمجرد الصدفة ؟؟ متى يكون الفرق نتيجة للصدفة؟؟ متى يكون الفرق دالا على وجود اختلاف حقيقي أو جوهري بين متوسطي المجتمعين الأصليين ..

مفاهيم مهمة :

هناك بعض المفاهيم المتعلقة باختبارات الفروض لابد من معرفتها:

- الفرض الاحصائي statistical hypothesis

هو عبارة عن إدعاء او تخمين معين حول معلمة من معالم المجتمع ويكون المطلوب اختبار صحة هذا الادعاء أو التخمين ...هناك نوعين من الفروض :

- فرض العدم (null hypothesis) ويرمز له بالرمز H_0 ويصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو عدم وجود تغير – مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن فرض العدم هو

H_0 : نفترض عدم وجود اختلاف بين متوسطي اعمار الطلاب والطالبات

- الفرض البديل (alternative hypothesis) ويرمز له بالرمز H_1 وهو الفرض الذي يجب أن يكون صحيحا اذا كان فرض العدم غير صحيح – مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن الفرض البديل هو

H_1: يوجد اختلاف حقيقي وليس ظاهري بين متوسط اعمار الطلاب والطالبات .

مستوى المعنوية α ودرجة الثقة $(1-\alpha)$:

- إن القرار الذي سوف نتخذه بناء على الاختبار الإحصائي لا يمكن اعتباره صحيح % 100 فهناك مقدار من الخطأ لأن المعلومات التي نتخذ قرارنا بناء عليها بيانات مأخوذة من عينة وليس من المجتمع الأصلي
- في اختبار فرض معين، فإن مقدار ثقتنا في القرار المتخذ بالرفض أو القبول يسمى بدرجة الثقة ويرمز له بالرمز $(1-\alpha)$ كما وأن مقدار عدم الثقة أو مقدار الخطأ يسمى بمستوى المعنوية ويرمز له بالرمز α

وعادة يحدد الباحث مستوى المعنوية أو درجة الثقة قبل البدء في عملية الاختبار.

عند اختبار فرض العدم H_o ضد الفرض البديل H_1 نجد أننا امام احدى الحالات الاربع الاتية :

H0 خطأ	H0 صحيح	
خطأ من النوع الثاني	قرار سليم	قبول H0
قرار سليم	خطأ من النوع الاول	رفض H0

1) أن يكون فرض العدم صحيحا ويكون القرار بقبوله ....وهذا قرار سليم
2) أن يكون فرض العدم صحيحا ويكون القرار برفضه .... وهذا قرار خاطئ (الخطأ من النوع الأول : رفض H_o عندما يكون H_o صحيحا ويرمز لحجم هذا الخطأ بالرمز α)
3) أن يكون فرض العدم خطأ ويكون القرار برفضه وهذا قرار سليم
4) أن يكون فرض العدم خطأ ويكون القرار بقبوله ..وهذا قرار خاطئ (الخطأ من النوع الثاني : قبول H_o عندما يكون H_o خاطئ ويرمز لحجم هذا الخطأ بالرمز β)

- احتمال الوقوع في الخطأ من النوع الأول يسمى مستوى المعنوية ويرمزله بالرمز α أي ان α = احتمال رفض فرض العدم H_o عندما يكون صحيح = مستوى المعنوية

- احتمال الوقوع في خطأ من النوع الثاني يرمز له بالرمز β أي أن β = احتمال قبول فرض العدم H_o عندما يكون خطأ

ويمكن تلخيص ما سبق فيما يلي:

- الخطأ من النوع الأول هو رفض فرضية عدم صحيحة ويرمز لاحتمال وقوعه بالرمز ويطلق علية مصطلح مستوى المعنوية.
- الخطأ من النوع الثاني هو قبول فرضية عدم خاطئة ويرمز لاحتمال وقوعه بالرمز .
- مستوى الثقة هو احتمال قبول فرضية عدم صحيحة.
- قوة الاختبار هو احتمال رفض فرضية عدم خاطئة.

خطوات اختبار الفرض الإحصائي حول متوسط المجتمع لعينة كبيرة

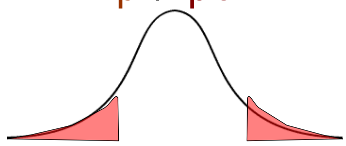
لإجراء الاختبار الإحصائي فإننا نتبع الخطوات التالية :

1- صياغة فرض العدم Ho

$Ho : \mu = \mu0$

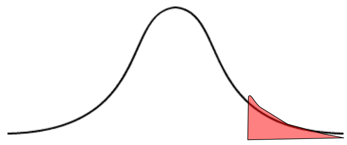
والفرض البديل هو احد الحالات التالية :

1- $H1: \mu \neq \mu0$



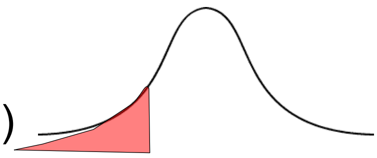
(اختبار من طرفين)

2- $H1: \mu > \mu0$



(اختبار من طرف واحد، الجهة اليمنى)

3- $H1: \mu < \mu0$



(اختبار من طرف واحد، الجهة اليسرى)

2- تحديد قيمة احصاءة الاختبار(قيمة Z المحسوبة):

حيث أن هذا الاحصاءة يتبع تقريبا توزيعا طبيعيا قياسياً

$$Z_C = \frac{\overline{X} - \mu_O}{\frac{\sigma}{\sqrt{n}}}$$

3- تحديد القيمة الجدولية و تحدد على حسب نوع الاختبار وقيمة α :

الدرجة المعيارية	درجه الثقة $(1-\alpha)$	مستوى المعنوية α	نوع الاختبار
$\pm 1.96 = Z_{\alpha/2}$	$95\% = 0.95$	$5\% = 0.05$	اختبار من طرفين
$\pm 2.58 = Z_{\alpha/2}$	$99\% = 0.99$	$1\% = 0.01$	

القيمة الجدولية (القيمة الحرجة)	درجه الثقة $(1-\alpha)$	مستوى المعنوية α	نوع الاختبار
$Z\alpha = 1.64$	$95\% = 0.95$	$5\% = 0.05$	اختبار من طرف واحد (الجهة اليمنى)
$Z\alpha = 2.33$	$99\% = 0.99$	$1\% = 0.01$	
$Z\alpha = -1.64$	$95\% = 0.95$	$5\% = 0.05$	اختبار من طرف واحد(الجهةاليسرى)
$Z\alpha = -2.33$	$99\% = 0.99$	$1\% = 0.01$	

4- اتخاذ القرار:

تتخذ القرار بناءاً على قيمة احصاءة الاختبار

نرفض H0 إذا وقعت قيمة احصاءة الاختبار في منطقة الرفض

لا نرفض H0 إذا وقعت قيمة احصاءة الاختبار في منطقة القبول

إذا كان الاختبار من طرفين : نقبل فرض العدم إذا تحققت المعادلة التالية : $-Z_{\alpha/2} < Z_C < Z_{\alpha/2}$

نرفض فرض العدم إذا تحققت إحدى المعادلتين : $Z_C > Z_{\alpha/2}$ ، $Z_C < -Z_{\alpha/2}$

إذا كان الاختبار من طرف واحد الجهة اليمنى :

نقبل فرض العدم إذا تحققت المعادلة : $Z_C < Z_\alpha$

نرفض فرض العدم إذا تحققت المعادلة : $Z_C > Z_\alpha$

إذا كان الاختبار من طرف واحد الجهة اليسرى :

نقبل فرض العدم إذا تحققت المعادلة : $Z_C > -Z_\alpha$

نرفض فرض العدم إذا تحققت المعادلة : $Z_C < -Z_\alpha$

-

مثال (1): أفترض أن الشركة ترغب في اختبار ما إذا كان يمكنها الادعاء بأن متوسط عمر المصباح من إنتاجها هو 1000 ساعة احتراق. وأنها قامت بأخذ عينة عشوائية حجمها $n = 100$ من إنتاجها فوجدت أن متوسط العينة $\overline{X} =$ 980 ساعة والانحراف المعياري للعينة $s = 80$ ساعة. فإذا أرادت الشركة القيام بالاختبار عند مستوى معنوية 5%، فعليهما أن تمضي كالآتي. حيث أن μ يمكن أن تساوي، تزيد عن، أو تقل عن 1000

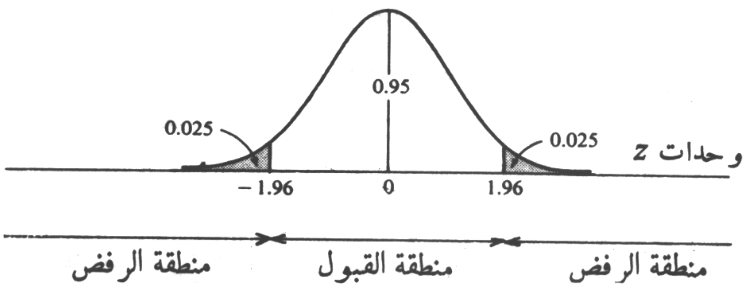
$$H_0 : \mu = 1000$$

$$H_1 : \mu \neq 1000$$

وحيث أن $n > 30$ ، فأن توزيع المعاينة للوسط يكون تقريباً طبيعياً (ويمكن استخدام s كتقدير بدلاً من σ) . وتكون منطقة القبول للاختبار عند مستوى المعنوية 5% بين 1.96 تحت التوزيع الطبيعي القياسي وتكون منطق الربط خارجها (أنظر شكل 5-1). وحيث أن منطقة الربط تقع عند ذيل التوزيع، فأن الاختبار يسمى اختبار ذو ذيلين. وتكون

الخطوة الثالثة إيجاد قيمة المناظرة لقيمة $\overline{X}$:

$$z = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{\overline{X} - \mu_0}{80 / \sqrt{100}} = \frac{-20}{8} = -2.5$$



شكل (٥ – ١)

وحيث أن قيمة z المحسوبة تقع داخل منطقة الرفض، فأن على الشركة أن ترفض H_0 أي $\mu = 1000$ وتقبل H_1 أي $\mu \neq 1,000$ عند مستوى معنوية 5%.

مثال (2):
شركة متخصصة في صناعة لعب الأطفال تعاقدت لشراء نوع جديد من الخيوط الصناعية، يدعي صانع هذه الخيوط أن متوسط قوة تحمل الخيط 15 كجم بانحراف معياري نصف كغم .
ولاختبار صحة ادعاء الصانع أخذت عينة عشوائية من 50 خيطاً وتم اختبارها فوجد أن متوسط قوة التحمل في العينة 14.8كغم . فهل يمكننا تأييد ادعاء المدير؟ استخدمي مستوى معنوية 5 % .

الحل:

$$n=50 \qquad \mu_O=15\ kg$$
$$\overline{X}=14.8\ kg \qquad \sigma=0.5\ kg$$

1- *صياغة الفرض الإحصائي:*

الفرض العدم : عدم وجود فرق معنوي بين متوسط الخيوط القديمة والحديثة $H_O:\mu=15$

ضـــــــــــــد

الفرض البديل : وجود فرق معنوي بين متوسط الخيوط القديمة والحديثة $H_1:\mu\neq15$

حيث μ هي متوسط قوة تحمل الخيط .

2- *إيجاد قيمة إحصاء الاختبار:*

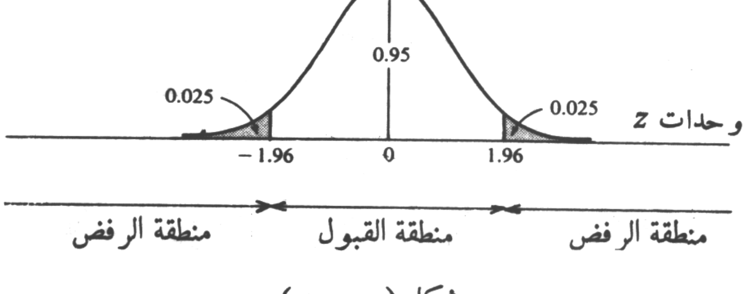
لان σ مجهولة والعينة كبيرة فانه يمكن استخدام S بدلاً من σ :

$$Z_C=\frac{\overline{X}-\mu_O}{\frac{\sigma}{\sqrt{n}}}=\frac{14.8-15}{\frac{0.5}{\sqrt{50}}}=-2.83$$

3- تحديد القيمة الجدولية :

$$Z_{\alpha/2}=\pm1.96$$

4- اتخاذ القرار:
بما أن القيمة المحسوبة وقعت في منطقة الرفض، فإن القرار هو: رفض فرض العدم



شكل (٥ – ١)

.

مثال(3) :

إذا كان من المعروف أن جسم الإنسان البالغ يحتاج يوميا في المتوسط 800 ميللكغرام من الكالسيوم لكي يقوم بوظائفه خير قام.ويعتقد احد علماء التغذية أن الأفراد ذوي الدخل المنخفض لا يستطيعون تحقيق هذا المتوسط , ولاختبار ذلك تم اختيار عينة من 50 شخصا بالغا من بين ذوي الدخل المنخفض فكان متوسط ما يتناوله من كالسيوم يوميا هو 755.3 ميللغرام والانحراف المعياري هو 239.3 ميللغرام .فهل تدل هذه النتائج على أن متوسط ما يتناوله الأشخاص البالغون من ذوي الدخل المنخفض من كالسيوم يقل عن 800 ميللجرام؟ استخدمي مستوى معنوية 0.05

الحل :

1- صياغة الفرض الإحصائي:

فرض العدم هو

$$H_0 : \mu = 800$$

والفرض البديل

$$H_1 : \mu < 800$$

حيث μ هي متوسط ما يتناوله الإنسان البالغ ذوي الدخل المنخفض من الكالسيوم .

2- إيجاد قيمة إحصاء الاختبار:

لان σ مجهولة والعينة كبيرة فانه يمكن استخدام $s=239.3$ بدلا منها .وبالتعويض نجد ان قيمة احصاءة الاختبار هي

$$Z= \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{755.3-800}{239.3/\sqrt{50}} = -1.32$$

3- تحديد القيمة الجدولية :

ونلاحظ هنا أن الاختبار ذو جانب (طرف) واحد هو الجانب (الطرف) الأيسر وحيث أن مستوى المعنوية $\alpha=0.05$ فانه من جدول التوزيع الطبيعي القياسي نجد أن القيمة الحرجة هي :

$Z_\alpha = -1.64$

4- اتخاذ القرار:

بما أن قيمة الاحصاءة -1.32 اكبر من القيمة الحرجة -1.64 وهي تقع في منطقة القبول وبالتالي فإننا لانرفض فرض العدم H_0 وهو أن متوسط ما يتناوله الإنسان البالغ ذو الدخل المنخفض من الكالسيوم يساوي 800 ميللجرام .

مثال (3) :

في عينة عشوائية مكونة من تسجيل 100 حالة وفاة في قرية معينة تبين أن متوسط العمر في العينة 67.5 عاما والانحراف المعياري 8 أعوام. فهل هذا يوضح أن متوسط العمر في هذه القرية اكبر من 65 عاما ؟

استخدمي مستوى معنوية % 5.

الحل:

نفرض أن μ متوسط العمر في هذه القرية .

1- صياغة الفرض الإحصائي:

$$H_0 : \mu = 65$$

$$H_1 : \mu > 65$$

2- إيجاد قيمة إحصاء الاختبار:

$$Z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{67.5-65}{8/\sqrt{100}} = 3.125$$

3- تحديد القيمة الجدولية :

ونلاحظ هنا أن الاختبار ذو جانب (طرف) واحد هو الجانب (الطرف) الأيمن وحيث أن مستوى المعنوية $\alpha=0.05$ فانه من جدول التوزيع الطبيعي القياسي نجد أن القيمة الحرجة هي :

$Z_\alpha = 1.64$

4- اتخاذ القرار:

نجد أن قيمة Z المحسوبة 3.125 اكبر من القيمة الجدولية 1.64 لهذا فان Z المحسوبة تقع في منطقة الرفض لهذا فان القرار هو رفض H_0

ونستنتج من ذلك أن متوسط العمر في هذه القرية اكبر من 65 عاما .

حاله العينة صغير والتباين معلوم للعينة

مثال (4) : ترغب شركة أن تعرف بدرجة ثقة 95% ما إذا كان يمكنها الادعاء بأن صناديق الصابون المسحوق الذي تبيع تحتوي على اكثر من 500 غرام (حوالي 1.1 رطل) من الصابون. وتعرف الشركة من الخبرة الماضية أن أوزان الصابون بالصناديق تتبع التوزيع الطبيعي. وقد أخذت الشركة عينة عشوائية حجمها $n = 25$ ووجدت أن $\overline{X} = 520$ غرام و $s = 75$ جرام. وحيث أن الشركة ترغب في اختبار ما إذا كانت $\mu > 500$ ، فإن :

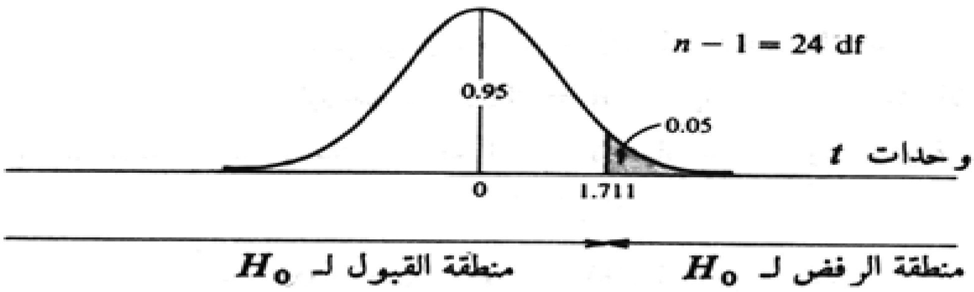
$$H_0 : \mu = 500$$

$$H_1 : \mu > 500$$

وحيث أن التوزيع طبيعي ، $n < 30$ ، وكذلك σ غير معلومة، فعلينا أن نستخدم توزيع t (بدرجة حرارية $n - 1 = 24$) لتحديد المنطقة الحرجة، أي منطقة الرفض، للاختبار بمستوى معنوية 5%. ويسمى هذا اختبار الذيل الأيمن . وأخيراً ، حيث أن

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{520 - 500}{75/\sqrt{25}} = \frac{20}{15} = 1.33$$

وهي تقع داخل منطقة القبول، وتقبل H_0 أى $\mu = 500$ ، عند مستوى معنوية 5% (أو بدرجة ثقة 95%) .



شكل (٥ – ٢)