اختبارات الفروض

حول الفرق بين متوسطي مجتمعين

أولا: إذا كانت العينتان مستقلتان

بافتراض أن كل مجتمع من المجتمعين له توزيع طبيعي، خصائص كل منهما كمايلي:

المجتمع الثاني	المجتمع الأول	خصائص المجتمع
μ_2	μ_1	المتوسط
σ_2^2	σ_1^2	التباين
$x_2 \sim N(\mu_2, \sigma_2^2)$	$x_1 \sim N(\mu_1, \sigma_1^2)$	الخصائص

للحصول على التقدير بنقطة للفرق $(\mu_1 - \mu_2)$، يجب إتباع الآتي:

١- **سحب عينة عشوائية من المجتمع الأول، حجمها n_1 ، وبياناتها هي: $x_{11}, x_{12}, ..., x_{1n_1}$ ، ويحسب لها الوسط الحسابي $\bar{x}_1$ ، وكذلك التباين S_1^2 ، كما يلي:**

$$S_1^2 = \frac{\sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2}{n_1 - 1} \ , \ \bar{x}_1 = \frac{\sum_{j=1}^{n_1} x_{1j}}{n_1}$$

٢- **سحب عينة عشوائية من المجتمع الثاني، حجمها n_2 ، وبياناتها هي: $x_{21}, x_{22}, ..., x_{2n_2}$ ، ويحسب لها الوسط الحسابي $\bar{x}_2$ ، وكذلك التباين S_2^2 ، كما يلي:**

$$S_2^2 = \frac{\sum_{j=1}^{n_1} (x_{2j} - \bar{x}_2)^2}{n_2 - 1} \ , \ \bar{x}_2 = \frac{\sum_{j=1}^{n_1} x_{2j}}{n_2}$$

٣- التقدير بنقطة للفرق $(\mu_1 - \mu_2)$ هو:

$$\widehat{(\mu_1 - \mu_2)} = (\bar{x}_1 - \bar{x}_2) \qquad \textbf{(3-1)}$$

٤- وتوزيع المعاينة لهذا التقدير هو التوزيع الطبيعي، وله الخصائص التالية:

$$(\bar{x}_1 - \bar{x}_2) \sim N\left((\mu_1 - \mu_2) \ , \ \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right) \right) \qquad \textbf{(3-2)}$$

- ***اختبار فرض حول الفرق بين متوسطي مجتمعين***

الغرض من هذا الاختبار هو مقارنة متوسطي مجتمعين من خلال اختبار الفرق بينهما، وبإتباع نفس الخطوات السابقة في اختبار فرض حول متوسط مجتمع، يمكن اختبار فرض حول الفرق بين المتوسطين $(\mu_1 - \mu_2)$ كما يلي:

١- صياغة الفروض.

الفرض البديل H_a	*الفرض العدم* H_o
$H_a : (\mu_1 - \mu_2) \neq (\mu_1 - \mu_2)_O$	$H_o : (\mu_1 - \mu_2) = (\mu_1 - \mu_2)_O$
$H_a : (\mu_1 - \mu_2) \succ (\mu_1 - \mu_2)_O$	$H_o : (\mu_1 - \mu_2) = (\mu_1 - \mu_2)_O$
$H_a : (\mu_1 - \mu_2) \prec (\mu_1 - \mu_2)_O$	$H_o : (\mu_1 - \mu_2) = (\mu_1 - \mu_2)_O$

٢- **تحديد مستوى المعنوية α ، وتوزيع المعاينة، ومن ثم القيم الحرجة Critical Values**

٣- **توزيع المعاينة الذي يتبعه إحصائية الاختبار، ويتحدد بناءً على ما إذا كان تبايني المجتمع $(\sigma_1^2 \text{ ، } \sigma_2^2)$ معلوم، أو غير معلوم، وفيما يلي بيان ذلك:**

توزيع المعاينة	إحصائية الاختبار	تبايني المجتمع $(\sigma_1^2 \text{ ، } \sigma_2^2)$	
$\sim N(0, 1)$	$Z^* = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_o}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$		معلوم
$\sim t_{(n_1 + n_2 - 2)}$	$t^* = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_o}{\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$	$\sigma_1^2 = \sigma_2^2$	غير معلوم
$\sim t_{(\nu)}$	$t^* = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_o}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	$\sigma_1^2 \neq \sigma_2^2$	

حيث أن S_p^2 هو التباين المرجح أو (المشترك) Pooled Variance، ويستخدم كتقدير غير متحيز لتباين المجتمع في حالة $\sigma_1^2 = \sigma_2^2$ ، ويحسب بإيجاد المتوسط المرجح لتبايني العينتين كما يلي:

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1+n_2-2)} = \frac{\sum_{j=1}^{n_1}(x_{1j}-\bar{x}_1)^2+\sum_{j=1}^{n_2}(x_{2j}-\bar{x}_2)^2}{(n_1+n_2-2)} \qquad (3\text{-}3)$$

كما أن ν هي درجات الحرية المحسوبة لتوزيع t، في حالة $\sigma_1^2 \neq \sigma_2^2$، ويعبر عنها بالمعادلة التالية:

$$\nu = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{(n_1-1)}+\frac{(S_2^2/n_2)^2}{(n_2-1)}} \sim (\text{integer number}) \qquad (3\text{-}4)$$

الخلاصة

إذا كانت العينتان مستقلتان، فإن إحصائية الاختبار، وكذلك فترات الثقة في حالة معلومية، وعدم معلومية التباينين هي:

تبايني المجتمعين		إحصائية الاختبار	فترة الثقة
معلـــــــــــــوم		$Z^* = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)_o}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	$(\bar{x}_1-\bar{x}_2)\mp\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ Z_{(1-\alpha/2)}$
غير معلوم	$\sigma_1^2=\sigma_2^2$	$t^* = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)_o}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$	$(\bar{x}_1-\bar{x}_2)\mp\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\ t_{(1-\alpha/2\,,\ n_1+n_2-2)}$

$(\bar{x}_1-\bar{x}_2)\mp\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\; t_{(1-\alpha/2\,,\,v)}$	$t^*=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)_o}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$	$\sigma_1^2 \neq \sigma_2^2$	

تطبيقات

مثال(٢-١) صـ٣٣

بيانات العينتين

عينة البرتقال(٢)	عينة التفاح(١)	
4.72	4.86	
4.81	5.11	
5.22	5.23	
5.67	5.19	
5.52	5.61	
4.96	5.32	
5.35	5.20	
5.34	4.95	
	4.98	
8	9	n_i

$\sum x_{ij}$	**46.45**	**41.59**
$\sum x_{ij}^2$	**240.1381**	**217.0219**

ومن ثم يمكن حساب المتوسط والتباين لكل عينة، كما يلي:

العينة الأولى	*العينة الثانية*
$\bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{46.45}{9} = 5.16$	$\bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{41.59}{8} = 5.2$
$S_1^2 = \frac{\sum x_1^2 - \frac{\left(\sum x_1\right)^2}{n_1}}{(n_1 - 1)} = \frac{240.1381 - \frac{(46.45)^2}{9}}{8} = 0.05056$	$S_2^2 = \frac{\sum x_2^2 - \frac{\left(\sum x_2\right)^2}{n_2}}{(n_2 - 1)} = \frac{217.0219 - \frac{(41.59)^2}{8}}{7} = 0.11513$
$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)} = \frac{8(0.05056) + 7(0.11513)}{(9 + 8 - 2)} = 0.08069$	

بفرض أن :

μ_1 : **متوسط كمية الصوديوم (ml/100gr) في النوع الأول (التفاح).**

μ_2 : **متوسط كمية الصوديوم (ml/100gr) في النوع الثاني(البرتقال).**

وفيما يلي خطوات اختبار الفرض:

- **صياغة الفروض**

$$H_o : (\mu_1 = \mu_2) \ \ or \ \ H_o : (\mu_1 - \mu_2) = 0$$

$$H_a : (\mu_1 - \mu_2) \neq 0$$

- **توزيع المعاينة ومناطق الرفض والقبول**

تبايني المجتمع غير معلوم، وهما متساويان $(\sigma_1^2 = \sigma_2^2)$ **، لذا يكون توزيع المعاينة هو توزيع t بدرجات حرية** $(n_1 + n_2 - 2) = 9 + 8 - 2 = 15$ **، وحيث أن مستوى المعنوية هو:** $\alpha = 0.05$ **، تكون القيمة الحرجة هي :** $t_{(0.975,15)} = 2.131$

- **إحصائية الاختبار**

$$t^* = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_o}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(5.16 - 5.2) - 0}{\sqrt{0.8069\left(\frac{1}{9} + \frac{1}{8}\right)}} = -0.29$$

- **القرار**

بما أن : $\left|t^*\right| = 0.29 < t = 2.131$ **إذا لا يمكن رفض الفرض العدم**

واجب منزلي: أنشاء فترة ثقة 95% للفرق بين المتوسطين.

اختبار تساوي تبايني مجتمعين

Test of Equal Two Variances

خطوات الاختبار

- **صياغة الفروض**

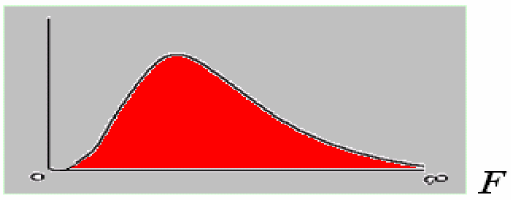
الفرض البديل H_a	الفرض العدم H_o
$H_a : \sigma_1^2 \neq \sigma_2^2$	$H_o : \sigma_1^2 = \sigma_2^2$

- **تحديد التوزيع الاحتمالي لإحصائية الاختبار**

إذا كان S_1^2 ، S_2^2 ، هما تبايني العينتين المسحوبتين من المجتمعين بحيث أن ($S_1^2 \succ S_2^2$)، فإن إحصائية الاختبار هي:

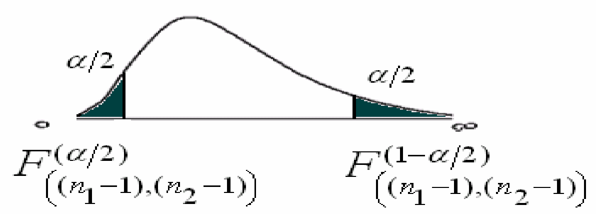
$$F^* = \frac{S_1^2}{S_2^2} \sim F_{\left((n_1-1),(n_2-1)\right)} \qquad \textbf{(3-5)}$$

حيث أن F^* تحت صحة الفرض العدم لها توزيع احتمالي يسمى توزيع F بدرجات حرية بسط = n_1-1 ، ودرجات حرية مقام n_2-1 ، وهما معالم هذا التوزيع ، كما يعتبر توزيع F توزيع موجب الالتواء ، ويعبر عنه بمنحنى يأخذ الصورة التالية:



- **تحديد مستوى المعنوية، ومناطق الرفض والقبول من خلال استخراج القيمة الجدولية، وهي:**

$$F^{(\alpha/2)}_{((n_1-1),(n_2-1))} \quad , \quad F^{(1-\alpha/2)}_{((n_1-1),(n_2-1))}$$



وحيث أننا وضعنا تباين العينة الأكبر في البسط، وتباين العينة الأصغر في المقام، سوف نستخدم منطقة الرفض اليمنى فقط.

- **القرار:**

إذا كان قيمة F^* المحسوبة أكبر من $F^{(1-\alpha/2)}_{((n_1-1),(n_2-1))}$ الجدولية، يرفض فرض العدم $H_o : \sigma_1^2 = \sigma_2^2$ ، ويستدل من ذلك على أن تبايني المجتمعين غير متساويين.

اختبار التجانس في مثال(٢-١) صـ ٣٣

- **البيانات المتاحة:**

	العينة الأولى	*العينة الثانية*
n	$n_1 = 9$	$n_2 = 8$
S_i^2	$S_1^2 = 0.05056$	$S_2^2 = 0.11513$

- **خطوات الاختبار**

صياغة الفروض:

$$H_o : \sigma_1^2 = \sigma_2^2 \qquad H_a : \sigma_1^2 \neq \sigma_2^2$$

إحصائية الاختبار هي:

$$F^* = \frac{Max(S^2)}{Min(S^2)} = \frac{S_2^2}{S_1^2} = \frac{0.115}{0.051} = 2.255$$

مناطق الرفض والقبول.

$$\alpha/2 = 0.025 \ , \ \alpha = 0.05$$

درجات حرية البسط: $n_2 - 1 = 8 - 1 = 7$ **(التباين الأكبر)**

درجات حرية المقام : $n_1 - 1 = 9 - 1 = 8$ **(التباين الأصغر)**

إذا القيمة الجدولية هي:

$$F^{(1-\alpha/2)}_{\left((n_2-1),(n_1-1)\right)} = F^{(0.975)}_{(7,8)} = 4.53$$

القرار:

بما أن $F^* = 2.255 \prec F^{(0.975)}_{(7,8)} = 4.53$ **إذا لا يمكننا رفض الفرض العدم ، ويستدل من ذلك على أن التباينين متساويين.**

ثانيا:إذا كانت العينتان غير مستقلتين

يقصد بالعينتين غير المستقلتين ، أن مفردات العينة لم يتغير ، وإنما كل ما يتم ، هو أخذ القياس على كل مفردة مرتين ، مرة قبل تطبيق المعالجة أو الطريقة ، ومرة بعد تطبيق المعالجة أو الطريقة ، وفي هذه الحالة إذا وجد اختلاف بين المتوسطين إنما يمكن إرجاع ذلك لأثر الطريقة.

- ***طريقة إجراء الاختبار***

بفرض أن :

i	العينة قبل x_1	العينة بعد x_2	عينة الفروق $d = x_2 - x_1$
1	x_{11}	x_{21}	$d_1 = (x_{21} - x_{11})$
2	x_{12}	x_{22}	$d_2 = (x_{22} - x_{12})$
3	x_{13}	x_{23}	$d_3 = (x_{23} - x_{13})$
i	x_{1i}	x_{2i}	$d_i = (x_{2i} - x_{1i})$
n	x_{1n}	x_{2n}	$d_n = (x_{2n} - x_{1n})$

يتم حساب المتوسط والتباين للفروق:

$$\bar{d} = \sum d_i \Big/ n \qquad\qquad S_d^2 = \sum (d_i - \bar{d})^2 \Big/ (n-1)$$

في هذه الحالة يتم اختبار فرض حول متوسط الفرق $\mu_d = \mu_2 - \mu_1$ **، باعتباره اختبار خاص بمتوسط واحد.**

ومن ثم يمكن تلخيص خطوات الاختبار في النقاط التالية:

١- **حساب عينة الفروق** $d_i = (x_{2i} - x_{1i})$ **، ومن ثم حساب الوسط الحسابي لها** $\bar{d} = \sum d_i / n$ **، وكذلك الانحراف المعياري** $S_d = \sqrt{\sum (d_i - \bar{d})^2 / (n-1)}$.

٢- **صياغة الفروض:**

الفرض البديل H_a	الفرض العدم H_o
$H_a : \mu_d \neq \mu_{do}$	$H_o : \mu_d = \mu_{do}$
$H_a : \mu_d \succ \mu_{do}$	$H_o : \mu_d = \mu_{do}$
$H_a : \mu_d \prec \mu_{do}$	$H_o : \mu_d = \mu_{do}$

٣- **حساب إحصائية الاختبار.**

$$t^* = \frac{\bar{d} - \mu_{do}}{\frac{S_d}{\sqrt{n}}}$$

٤- **مناطق الرفض والقبول.**

بتحديد مستوى المعنوية، ونوع اتجاه الفرض البديل يمكن استخراج القيمة الجدولية من توزيع t عند درجات حرية *(n-1)* ، ومن ثم يمكن تحديد مناطق الرفض والقبول.

٥- **القرار.**

إذا كانت $|t|^* \succ t_{(A,n-1)}$ **نرفض فرض العدم، ونقبل الفرض البديل.**

تطبيق {مثـــال (٢-٢) صـ ٣٧} .

بإتباع نفس الخطوات المذكورة أعلاه، يمكننا إجراء الاختبار.

pair	عليقة (1)	عليقة (2)	d=(2-1)	d^2
1	30.51	36.32	5.81	33.7561
2	29.37	37.51	8.14	66.2596
3	28.72	35.47	6.75	45.5625
4	31.33	38.20	6.87	47.1969
5	31.56	36.52	4.96	24.6016
6	29.80	37.22	7.42	55.0564
7	30.50	38.95	8.45	71.4025
Sum			48.4	343.8356

$$\bar{d} = \frac{\sum d}{n} = \frac{48.4}{7} = 6.91$$

$$S_d = \sqrt{\frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n-1}} = \sqrt{\frac{343.8356 - \frac{48.4^2}{7}}{6}} = 1.237$$

١- صياغة الفروض.

يود الباحث أن يوصي باستخدام العليقة الثانية لأنها تؤدي إلى زيادة في الوزن عن العليقة الأولى ، أي أن الباحث يريد إثبات أن :

$$H_a : \mu_d = (\mu_2 - \mu_1)_o \succ 0$$

ومن ثم يصاغ الفرض العدم والبديل كالتالي:

$$H_o : \mu_d = 0$$
$$H_a : \mu_d \succ 0$$

٢- **إحصائية الاختبار.**

$$t^* = \frac{\bar{d} - \mu_{do}}{\frac{S_d}{\sqrt{n}}} = \frac{6.91 - 0}{\left(\frac{1.237}{\sqrt{7}}\right)} = 14.8$$

٣- **قيمة t الجدولية.** $(\alpha = 0.05)$

$$t_{((1-\alpha),(n-1))} = t_{(0.95,6)} = 1.943$$

٤- **القرار.**

بما أن $(t^* = 14.8) \succ (t_{(0.95,6)} = 1.943)$ **، إذا نرفض فرض العدم ، ونقبل الفرض البديل، ونوصي الباحث باستخدام العليقة الثانية، وذلك عند مستوى معنوية *5%*.**