

## رابعاً: توزيع المعاينة للفرق بين المتوسطين الحسابيين لعينتين

### 1. توزيع الفرق بين متوسطي عينتين مستقلتين

التوزيعات الطبيعية المستقلة المجهولة التباين

التوزيعات الطبيعية المستقلة المعروفة التباين

$n_1$  و  $n_2$  صغيرين أو أحدهما صغير

$n_1$  و  $n_2$  كبيرين

إذا كانت  $X_1, X_2, \dots, X_{n_1}$  عينة من توزيع وسطه  $\mu_1$  وتباينه  $\sigma_1^2$ ، و  $Y_1, Y_2, \dots, Y_{n_2}$  عينة أخرى من توزيع آخر مستقل عن الأول متوسطه  $\mu_2$  وتباينه  $\sigma_2^2$ ، وكان  $\sigma_1^2$  و  $\sigma_2^2$  مجهولين وحجم العينتين  $n_1, n_2 < 30$  فإن توزيع  $(\bar{X} - \bar{Y})$  هو:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{n_1+n_2-2}$$

إذا كان  $\sigma_1^2 = \sigma_2^2 = \sigma^2$  فإن:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

حيث:  $S_p^2$  هو تقدير لـ  $\sigma^2$ .

إذا كانت  $X_1, X_2, \dots, X_{n_1}$  عينة من توزيع وسطه  $\mu_1$  وتباينه  $\sigma_1^2$  و  $Y_1, Y_2, \dots, Y_{n_2}$  عينة أخرى من توزيع آخر مستقل عن الأول متوسطه  $\mu_2$  وتباينه  $\sigma_2^2$ ، وكان  $\sigma_1^2$  و  $\sigma_2^2$  مجهولين يمكن استبدالهما بتقديريهما غير المتحيزين  $S_1^2$  و  $S_2^2$ ، فإذا كان حجم العينتين يفوق أو يساوي 30 فإن توزيع  $(\bar{X} - \bar{Y})$  هو:

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)$$

ومنه:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

إذا كانت  $X_1, X_2, \dots, X_{n_1}$  عينة من توزيع وسطه  $\mu_1$  وتباينه  $\sigma_1^2$ ، و  $Y_1, Y_2, \dots, Y_{n_2}$  عينة أخرى من توزيع آخر مستقل عن الأول متوسطه  $\mu_2$  وتباينه  $\sigma_2^2$  حيث:

$$V(\bar{X}) = \frac{\sigma_1^2}{n_1}$$

$$V(\bar{Y}) = \frac{\sigma_2^2}{n_2}$$

من الاستقلالية:

$$V(\bar{X} - \bar{Y}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

ومنه:

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

ونكتب:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

## أمثلة:

المثال الأول:	المثال الثاني:
<p>سحبت عينة من كلية الاقتصاد حجمها 36 طالبا تباين معدلات توجيههم هو 4 وعينة أخرى من كلية الآداب حجمها 42 طالبا تباين معدلات توجيههم هو 3.5، فإذا كان المتوسطان الحقيقيان للمعدلات في كل من الكليتين متساويين أحسب احتمال أن يكون الفرق بين الوسطيين الحسابيين للعينتين مساويا لـ 1.5.</p>	<p>مصنع ينتج 700 كغ من العجائن الغذائية كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 20 يوما فبلغ انحرافها المعياري 40 كغ، في حين ينتج مصنع آخر 500 كغ كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 15 يوماً فبلغ انحرافها المعياري 20 كغ. المطلوب حساب الاحتمال التالي <math>P(\bar{X} - \bar{Y} \leq 210)</math> إذا علمت أن <math>\sigma_1^2 = \sigma_2^2</math></p>
<p><b>الحل:</b>  <math>\sigma_1^2</math> و <math>\sigma_2^2</math> مجهولان و <math>n_1, n_2 \geq 30</math> ومنه يمكن استبدالهما بتقديرهما غير المتحيزين <math>s_1^2</math> و <math>s_2^2</math>، فينتج لدينا:</p> $Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$ <p>والاحتمال المطلوب هو:</p> $P(\bar{X} - \bar{Y} = 1.5) = P\left(\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{1.5 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}\right)$ $= P(Z = 3.41) = 0.9997$	<p><math>\sigma_1^2</math> و <math>\sigma_2^2</math> مجهولان ومتساويان و <math>n_1, n_2 &lt; 30</math> ومنه:</p> $T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$ <p>والاحتمال المطلوب هو:</p> $P(\bar{X} - \bar{Y} \leq 210) = P\left(\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \leq \frac{210 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right)$ $= P\left(T \leq \frac{210 - (700 - 500)}{27.41 \sqrt{\frac{1}{20} + \frac{1}{15}}}\right) = P(T \leq 1.07)$ $= 0.850$

## 2. توزيع الفرق بين متوسطي عينتين غير مستقلتين

أحيانا تكون العينتان اللتان يتم سحبهما مرتبطتين، فمثلا لقياس فاعلية برنامج تدريبي على عينة من الطلبة يتم قياس مستواهم قبل التعرض للبرنامج وبعده، فتشكل القراءات قبل وبعد التعرض للبرنامج عينتين مرتبطتين، ولكي نجد توزيع الفرق بين متوسطي المجتمعين الذين سحبنا منهما العينتين نلجأ لتحويل المسألة لإيجاد توزيع متوسط واحد.

فإذا كانت  $X_1, X_2, \dots, X_{n_1}$  عينة من توزيع وسطه  $\mu_1$  وتباينه  $\sigma_1^2$ ، و  $Y_1, Y_2, \dots, Y_{n_2}$  عينة أخرى من توزيع آخر مستقل عن الأول متوسطه  $\mu_2$  وتباينه  $\sigma_2^2$ ، و  $X_i$  و  $Y_i$  القيمتين المتناظرتين في العينتين، نضع  $d_i = x_i - y_i$  وعليه تشكل  $D_1, D_2, \dots, D_n$  عينة من مجتمع طبيعي وسطه  $\mu_D = \mu_1 - \mu_2$  وتباينه  $\sigma_D^2$  ونفرض أن متوسط العينة  $\bar{D}$  وتباينها  $S_D^2$ .

وفي هذه الحالة يمكن التمييز بين حالتين:

- حجم العينة  $n \geq 30$ :

$$\bar{D} \sim N\left(\mu_D, \frac{\sigma_D}{\sqrt{n}}\right)$$

وبمعلومية تباين عينة الفروق ووسطها كمقدرين لوسط مجتمع الفروق وتباينه يكون:

$$Z = \frac{\bar{D} - \mu_D}{\frac{\sigma_D}{\sqrt{n}}} \sim N(0,1)$$

- حجم العينة  $n < 30$ :

$$T = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim t_{n-1}$$

مثال:

في تجربة لبيان تحسن أداء العمال، تم سحب عينة عشوائية بحجم 16 عاملا في المصنع فكان قياس الكفاءة قبل وبعد دخولهم دورة تحسين الأداء كما هو موضح

في الجدول التالي:

8	8	8	7	7	7	8	9	9	9	8	8	7	7	8	7	$x_i$
5	7	7	5	6	5	8	7	6	5	8	7	4	5	5	5	$y_i$

المطلوب: أحسب احتمال أن الفرق في الأداء قبل وبعد الدورة لا يقل عن 2.5.

الحل:

نحسب الفروق  $d_i = x_i - y_i$  كما هو موضح في الجدول الموالي:

3	1	1	2	1	2	0	2	3	4	0	1	3	2	3	2	$d_i$
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-------

لدينا :

حجم العينة  $n < 30$  ومنه:

$$T = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim t_{n-1}$$

ويكون الاحتمال المطلوب حسابه كما يلي:

$$\begin{aligned} P(\bar{D} > 2.5) &= P\left(\frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} > \frac{2.5 - \mu_D}{\frac{S_D}{\sqrt{n}}}\right) \\ &= P(T > 2.71) = 1 - P(T \leq 2.71) \\ &= 1 - 0.99 = 0.01 \end{aligned}$$

### خامسا: توزيع المعاينة للفرق بين نسبي عينتين

إذا اختيرت عينتان عشوائيتان من مجتمعين مستقلين يخضع كل منهما لتوزيع الثنائي فإن الفرق بين نسبي العينتين يخضع تقريبا حسب نظرية النهاية المركزية

عندما يكون حجم كل من العينتين كبيرا للتوزيع التالي:

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} \sim N(0,1)$$

مثال:

الحل:	المثال:
<p>لدينا:</p> $\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} \sim N(0,1)$ <p>والاحتمال المطلوب هو:</p> $P(\hat{p}_1 - \hat{p}_2 \leq 0.1) = P\left(\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} \leq \frac{0.1 - (p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}}\right)$ $= P(Z \leq 0.51) = 0.6950$	<p>إذا كانت نسبة النجاح في الامتحان التوجيهي في إحدى الجامعات هي 0.7 وكانت نسبة النجاح في الامتحان التوجيهي في جامعة أخرى هي 0.65، تم اختيار عينة عشوائية حجمها 70 طالبا من الجامعة الأولى وعينة عشوائية أخرى من الجامعة الثانية حجمها 35 طالبا، فما هو احتمال أن تزيد نسبة النجاح في الجامعة الأولى على نسبة النجاح في الجامعة الثانية بمقدار 0.1 على الأكثر؟</p>

## سادسا: توزيع المعاينة لتباين عينة

نحتاج في كثير من الأحيان في تطبيقات الإحصاء الاستقرائي لمعرفة توزيع تباين العينة الذي يختلف حسب الحالتين التاليتين:

### متوسط المجتمع مجهول

إذا أخذت عينة عشوائية حجمها  $n$  من توزيع طبيعي متوسطه  $\mu$  وتباينه  $\sigma^2$  مجهول و  $S^2$  تقديره غير المتحيز فإن:

$$\frac{(n-1)S^2}{\sigma^2} \sim \lambda_{n-1}^2$$

حيث يستعمل الوسط الحسابي للعينة  $\bar{X}$  كمقدر لمتوسط المجتمع  $\mu$ :

$$S^2 = \frac{\sum(x_i - \bar{X})^2}{n-1}$$

### متوسط المجتمع معلوم

إذا أخذت عينة عشوائية حجمها  $n$  من توزيع طبيعي متوسطه  $\mu$  وتباينه  $\sigma^2$  فإن:

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \lambda_n^2$$

حيث:

$$\hat{\sigma}^2 = \frac{\sum(x_i - \mu)^2}{n}$$

مثال:

المثال:	الحل:
سحبت عينة عشوائية حجمها 11 من توزيع طبيعي تباينه 70، أوجد احتمال أن يقل تباين العينة عن 87.5.	لدينا: $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ <p>ومنه يحسب الاحتمال المطلوب كما يلي:</p> $P(S^2 \leq 87.5) = P\left(\frac{(n-1)S^2}{\sigma^2} \leq \frac{87.5(n-1)}{\sigma^2}\right) = P(\chi^2 \leq 12.5)$ $= 0.75$

### سابعاً: توزيع المعاينة للنسبة بين تبايني عينتين

للمقارنة بين تبايني مجتمعين نعتمد على النسبة بين تبايني عينتين مسحوبتين منهما، والتي يحدد توزيعها في حالة المعاينة من مجتمعين طبيعيين مستقلين استناداً إلى النظرية التالية:

إذا كان  $S_1^2$  تباين عينة عشوائية من مجتمع طبيعي متوسطه  $\mu_1$  وتباينه  $\sigma_1^2$  و  $S_2^2$  تباين عينة عشوائية من مجتمع طبيعي مستقل عن الأول متوسطه  $\mu_2$  وتباينه  $\sigma_2^2$  فإن:

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1-1, n_2-1}$$

مثال:

المثال:

سحبت عينة حجمها 11 من مجتمع طبيعي تباينه 9 وسحبت عينة أخرى حجمها 21 من مجتمع طبيعي مستقل عن الأول تباينه 25، أحسب احتمال أن تكون النسبة بين تبايني العينتين أقل من 0.8.

نعلم أن:

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1-1, n_2-1}$$

وبالتالي يكون الاحتمال المطلوب:

$$\begin{aligned} P\left(\frac{S_1^2}{S_2^2} < 0.8\right) &= P\left(\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} < 0.8 \left(\frac{25}{9}\right)\right) \\ &= P(F < 2.22) = 0.95 \end{aligned}$$