

توزيعات المعاينة الفرق بين متوسطين $(\bar{X}_1 - \bar{X}_2)$

من المقارنت بين مجموعتين من حيث المتوسطات ،
تفكير ، اختبار فرضية

من مجموعتين N_1 مع المتوسط الأول ، و N_2 مع المتوسط الثاني .
الفرق بين المتوسط $(\bar{X}_1 - \bar{X}_2)$ هو إحصاء - لها توزيع
إحصائي ، توزيع طبيعي ، ومعلمة :

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_{\bar{X}_1} - \mu_{\bar{X}_2}$$

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2$$

المسألة الأولى : اختبار ت

المسألة الأولى : اختبار ت : μ_1 و μ_2 : متوسطات المجتمعين ، σ_1^2 و σ_2^2 : تباينات المجتمعين .

n_1 و n_2 : أحجام العينات

المسألة الثانية : اختبار ت : $(\bar{X}_1 - \bar{X}_2)$: الفرق بين المتوسطات

$\mu_1 - \mu_2$: الفرق بين المتوسطات : $\mu_1 - \mu_2$

$$\left\{ \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right. : \text{تباين}$$

المسألة الثالثة :

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

$\sigma_1^2 = 10$ Cakes $\mu_1 = 50$ Cakes $\bar{X}_1$ $n_1 = 5$ Alt

$\sigma_2^2 = 8$ Cakes $\mu_2 = 40$ Cakes $\bar{X}_2$ $n_2 = 4$

Find the $\bar{X}_1$: $n_1 = 5$ Find the $\bar{X}_2$: $n_2 = 4$

Find the $\bar{X}_2$: $n_2 = 4$ Find the $\bar{X}_1$: $n_1 = 5$

15 Cakes $(\bar{X}_1 - \bar{X}_2)$ Find the $\bar{X}_1$: $n_1 = 5$

Find the $(\bar{X}_1 - \bar{X}_2)$: 50 Cakes

$$\left\{ \begin{aligned} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} &= 50 - 40 = 10 \end{aligned} \right.$$

$$\left\{ \begin{aligned} \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 &= \frac{10}{5} + \frac{8}{4} = 4 \end{aligned} \right.$$

$$Pr((\bar{X}_1 - \bar{X}_2) > 15) = Pr\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > \frac{15 - 10}{\sqrt{4}}\right)$$

$$= Pr(Z > 2.5) = 1 - Pr(Z \leq 2.5) = 1 - 0.9938 = 0.0062$$

الافتراضات في اختبار ز: الافتراضات
 $n_1 > 30, n_2 > 30$ افتراضات

افتراضات في اختبار ز، افتراضات في اختبار ز

افتراضات في اختبار ز، افتراضات في اختبار ز

افتراضات في اختبار ز، افتراضات في اختبار ز

افتراضات في اختبار ز، افتراضات في اختبار ز

افتراضات في اختبار ز، افتراضات في اختبار ز

$$\begin{cases} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2 \\ \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \end{cases}$$

$$(\bar{X}_1 - \bar{X}_2) \sim N\left((\mu_1 - \mu_2), \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) : \text{افتراضات}$$

$\sigma_1^2 = 120$ Cubes, $\mu_1 = 500$ Cubes, $\bar{X}_1$ Sample Mean, $n_1 = 40$ Sample Size

$\sigma_2^2 = 180$ Cubes, $\mu_2 = 400$ Cubes, $\bar{X}_2$ Sample Mean, $n_2 = 30$ Sample Size

Find the probability that the difference between the sample means is greater than 105.

Find the probability that the difference between the sample means is greater than 105.

105 Cubes, $(\bar{X}_1 - \bar{X}_2)$ Difference between sample means

Find the probability that the difference between the sample means is greater than 105.

$$\begin{cases} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = 500 - 400 = 100 \\ \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{120}{40} + \frac{180}{30} = 9 \end{cases} \quad n_1 > 30, n_2 > 30$$

$$Pr((\bar{X}_1 - \bar{X}_2) > 105) = Pr\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > \frac{105 - 100}{\sqrt{9}}\right)$$

$$= Pr(Z > 1.66) = 1 - Pr(Z \leq 1.66) = 1 - 0.9515 = 0.0485$$

الحالة الثالثة: حالة مستقلة $n_1 > 30, n_2 > 30$

$$n_1 > 30, n_2 > 30$$

إذا كانت العينات لها توزيع برنولي (نجاح 1، فشل 0).

المعجم الأول: معدل النجاح p_1 ، المعجم الثاني: معدل النجاح p_2

وإذا كانت $n_1 > 30, n_2 > 30$ كبيرة، يمكن كتابة

المتباينة: $(\bar{X}_1 - \bar{X}_2)$ التي هي نسبة التوزيع

بالجبر توزيع طبيعي:

$$\begin{cases} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2 = p_1 - p_2 \\ \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} \end{cases}$$

وذلك:

$$(\bar{X}_1 - \bar{X}_2) \sim N\left((p_1 - p_2), \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}\right)$$

مثال 1

نفس الشيء في كل مرة ، نأخذ عينات عشوائية من المجتمع الأول

، $n_1 = 100$ ، $p_1 = 0.8$ ، $n_2 = 125$ ، $p_2 = 0.5$ ،

نأخذ عينات عشوائية من المجتمع الثاني ، $n_2 = 100$ ، $p_2 = 0.5$ ،

نأخذ عينات عشوائية من المجتمع الثالث ، $n_3 = 125$ ، $p_3 = 0.5$ ،

نأخذ عينات عشوائية من المجتمع الرابع ، $n_4 = 100$ ، $p_4 = 0.8$ ،

نأخذ عينات عشوائية من المجتمع الخامس ، $n_5 = 125$ ، $p_5 = 0.5$ ،

$$\mu_{\bar{x}_1} - \mu_{\bar{x}_2} = p_1 - p_2 = 0.8 - 0.5 = 0.3$$

$$\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} = \frac{0.8 \cdot 0.2}{100} + \frac{0.5 \cdot 0.5}{125} = 0.0036$$

$$Pr(\bar{x}_1 - \bar{x}_2 \leq 0.45) = Pr\left(\frac{\bar{x}_1 - \bar{x}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \leq \frac{0.45 - 0.3}{\sqrt{0.0036}}\right)$$

$$= Pr(Z \leq 2.5) = 0.9938$$

الحالة الرابعة: مجموعتين متجانستين

إذا كانت المجموعتان متجانستين، لهما نفس التباين σ^2 ، $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

في هذه الحالة، يكون التباين المشترك σ^2 هو المتوسط الحسابي للتباينات S_1^2 و S_2^2 .

لنأخذ S_1^2 من المجموعة الأولى n_1 و S_2^2 من المجموعة الثانية n_2 .

فإن S^2 هي المتوسط الحسابي لـ S_1^2 و S_2^2 .

أي: $(\bar{X}_1 - \bar{X}_2)$: الفرق

$$\begin{cases} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2 \\ \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \end{cases}$$

يجب أن نلاحظ أن S^2 هي المتوسط الحسابي لـ S_1^2 و S_2^2 .

$$S^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

درجات الحرية لـ S^2 هي $V = n_1 + n_2 - 2$.

$$S_1^2 = \frac{\sum (X_1 - \bar{X})^2}{n_1 - 1} \quad ; \quad \text{المجموع}$$

$$S_2^2 = \frac{\sum (X_2 - \bar{X})^2}{n_2 - 1}$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

: سبوق

$V = n_1 + n_2 - 2$ درجه آزادی student $\bar{X}_1, \bar{X}_2$

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{(n_1 + n_2 - 2)}$$

: توزیع

مثال

James Cakes $\mu_1 = 80$ كوب زينة $S_1^2 = 61$ Small Cakes $\mu_2 = 50$ كوب زينة $S_2^2 = 12$ Small Cakes $n_1 = 5$ since List

والجواب James Cakes $\mu_2 = 50$ كوب زينة $S_2^2 = 12$ Small Cakes $n_2 = 4$ since List. $(\bar{X}_1 - \bar{X}_2)$ 3.6 $\bar{X}_1 - \bar{X}_2$ 3.6 $\bar{X}_1 - \bar{X}_2$

$$\begin{cases} \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2 = 80 - 50 = 30 & : \text{الفرق} \\ S_{\bar{X}_1}^2 + S_{\bar{X}_2}^2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = 6^2 \left(\frac{1}{5} + \frac{1}{4} \right) = 6^2 \cdot \frac{9}{20} & : \text{الخطأ} \end{cases}$$

الخطأ S_1^2 ، S_2^2 : $S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$= \frac{4 \cdot 61 + 3 \cdot 12}{7} = \frac{280}{7} = 40$$

$S^2 \cdot \frac{9}{20} = 40 \cdot \frac{9}{20} = 18$: $(\bar{X}_1 - \bar{X}_2)$ 3.6 $\bar{X}_1 - \bar{X}_2$

$$\begin{aligned} \Pr((\bar{X}_1 - \bar{X}_2) \leq 3.6) &= \Pr\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \leq \frac{3.6 - 30}{\sqrt{18}}\right) \\ &= \Pr(t \leq 1.4142) = 0.90 \end{aligned}$$

1.414 C.C. $n = 5 + 4 - 2 = 7$ $t_{(7, 0.90)} = 1.415$

(٧-٦-١) توزيع المعاينة لنسبة S_1^2 / σ_1^2 إلى S_2^2 / σ_2^2 : توزيع F:The Sampling Distribution of the Ratio of S_1^2 / σ_1^2 to S_2^2 / σ_2^2 : The F Distribution

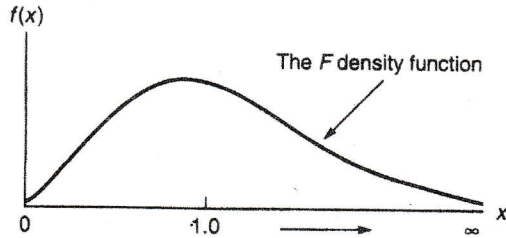
بفرض أننا سحبنا عينات عشوائية من مجتمعين، كل مجتمع يتبع التوزيع الطبيعي. لكل عينة نتناول نسبة تباين العينة إلى تباين المجتمع الذي سحبت منه، أي: S_1^2 / σ_1^2 ، S_2^2 / σ_2^2 . بفرض أننا كونا إحصاءاً عبارة عن نسبة من هذه القيم على الصورة:

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \quad (7.32)$$

توزيع المعاينة لهذا الإحصاء يعرف على أنه توزيع: F ، F distribution .

مثلاً كان توزيع T وتوزيع كاي تربيع دالة في درجات الحرية، فإن توزيع F أيضاً هو دالة في درجات الحرية، ولكن بخلاف توزيع T وتوزيع كاي تربيع، فإن توزيع F له نوعين من درجات الحرية: درجات حرية مقترنة بتباين العينة S_1^2 في البسط، ودرجات حرية مقترنة بتباين العينة S_2^2 في المقام. ويرمز لدرجات الحرية بـ γ_1, γ_2 على التوالي، وعلى ذلك فدرجات الحرية في البسط والمقام هي على التوالي: $\gamma_1 = n_1 - 1$ ، $\gamma_2 = n_2 - 1$. ويتحدد توزيع F تماماً بدلالة درجات الحرية ولا يتوقف على أي معالم أخرى.

أي متغير عشوائي يتبع توزيع F لا يمكن أن يأخذ قيمة سالبة، وهذا واضحاً من الصيغة (7.32)، حيث يلاحظ أن مكوناتها لا يمكن أن تكون سالبة. يتركز توزيع F حول القيمة واحد، ويرجع ذلك إلى أن تباينات المجتمعين يتم تقديرهما بتباينات العينتين، وبالتالي فمن المتوقع أن يكون كل من S_1^2 / σ_1^2 ، S_2^2 / σ_2^2 قريباً من القيمة واحد، لذلك نجد أن النسبة $(S_1^2 / \sigma_1^2) / (S_2^2 / \sigma_2^2)$ تقترب أيضاً من الواحد الصحيح. وتوزيع F هو توزيع ملتوي إلى اليمين ومدها نظرياً من الصفر إلى ما لا نهاية وشكل (٧-٧) يوضح توزيع F.



شكل (٧-٧): شكل توزيع F

وتوزيع F مثل توزيع T وتوزيع كاي تربيع، نجده متوجداً في معظم البرامج الإحصائية الجاهزة، بالإضافة إلى كونه مجدولاً بتوسع. وحيث أن توزيع F يعتمد على قيمتين من درجات الحرية، فإن هذا يقتضي إعداد مجلد كامل لجداول F. في هذا الكتاب نجد جزءاً من جداول F عند توليفات مختلفة من درجات الحرية تناظر الاحتمالات: 0.01، 0.025، 0.05، 0.10، 0.90، 0.95، 0.975، 0.99. وذلك في جدول E بالملحق. لاحظ أن هناك جدول منفصل لكل احتمال من هذه الاحتمالات.

لاستخدام هذا الجدول، نحدد أولاً الاحتمال المرغوب فيه (محدد بالمساحة A وهي المساحة التي تقع على يسار القيمة الجزئية المطلوبة)، بعد ذلك نبحث عن درجات حرية البسط $\gamma_1 = n_1 - 1$ من بين قمم الأعمدة، ونبحث عن درجات حرية المقام $\gamma_2 = n_2 - 1$ من بين قمم الصفوف. قيمة F التي تنتج من تقاطع درجتَي حرية البسط والمقام هي القيمة الجزئية المطلوبة، حيث أن المساحة التي تقع على يسار تلك القيمة تمثل الاحتمال المرغوب فيه.

الفصل السابع، الاستنتاج الإحصائي المتعلق بمجتمعين

كمثال، نفرض أن $n_2=20, n_1=16$ ونرغب في إيجاد القيم الجزئية (أي الجدولية) التي تناظر الاحتمالات 0.95, 0.025. هنا درجات حرية البسط $\gamma_1 = n_1 - 1 = 15$ ودرجات حرية المقام $\gamma_2 = n_2 - 1 = 19$. عند $A=0.025$ ، نجد أن القيمة الجزئية هي 3.6. وعند $A=0.95$ نجد أن القيمة الجزئية هي 2.23. هذا يعني أن احتمال أن متغير عشوائي يتبع توزيع F بدرجات حرية 15, 19 أن يأخذ قيمة لا تزيد عن 3.6 هو 0.025. وأن احتمال أن هذا المتغير يأخذ قيمة لا تزيد عن 2.23 هو 0.95. وقيم F الجزئية هذه يرمز لها بالرمز: $f_{0.025, 15, 19} = 3.6$ ، $f_{0.95, 15, 19} = 2.23$ وهي موضحة بيانياً في شكل (٧-٨). أو يمكن أن تكتب بصورة رمزية أخرى كما يلي:

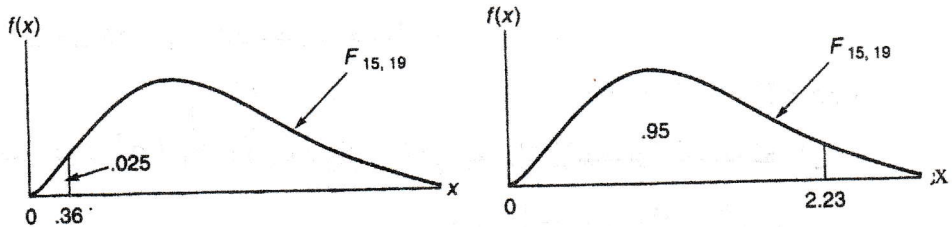
$$P(F_{15, 19} \leq 3.6) = 0.025$$

$$P(F_{15, 19} \leq 2.23) = 0.95$$

لاحظ أنه يمكن استخدام قاعدة الاحتمال للحوادث المكملة على الصورة:

$$P(F_{15, 19} > 3.6) = 0.975$$

$$P(F_{15, 19} > 2.23) = 0.05$$



شكل (٧-٨) : توضيح قيم F عند درجات الحرية 15, 19

استخدام الحاسب الآلي:

مثلاً حدث مع توزيع T ومع توزيع كاي تربيع، يمكن أيضاً استخدام برنامج Minitab لتوليد القيم الجزئية المطلوبة لتوزيع F. وكما رأينا من قبل، يستخدم الأمر INVCDF مصحوباً بالمساحة الاحتمالية التي تقع على يسار القيمة الجزئية المطلوبة، ثم الأمر الفرعي F مصحوباً بدرجات حرية البسط والمقام بهذا الترتيب. القيم الجزئية المبينة في شكل (٧-٨) تم توليدها ببرنامج Minitab كما يلي:

```
MTB > INVCDF .025 ;
SUBC > f 15 19
0.0250 0.3606
MTB > INVCDF .95 ;
SUBC > f 15 19
0.9500 2.2341
```

(٧-٦-٢) فترة الثقة واختبارات الفروض حول σ_1^2 / σ_2^2 :

Confidence Intervals and Hypothesis Testing for σ_1^2 / σ_2^2

كما أشرنا من قبل فإن أفضل إحصاء لـ σ_1^2 / σ_2^2 هو S_1^2 / S_2^2 وأن الاستنتاج الإحصائي حول σ_1^2 / σ_2^2 يعتمد على توزيع F.

مجموعة طمس ، الموزع الأول $\mu_1 = 800$ و $\sigma_1^2 = 300$ و $n_1 = 10$ عينات
 و S_1^2 عينات المجموعة الثانية ، $\mu_2 = 1000$ و $\sigma_2^2 = 690$ و $n_2 = 25$ عينات
 و S_2^2 عينات

(١) ما احتمال أن S_1^2 أكبر من S_2^2

(٢) ما احتمال أن S_1^2 أكبر من $\frac{3}{2} S_2^2$

الجزء الأول : S_1^2 المعلوم : $S_1^2 = 300$, $n_1 = 10$ درجة ، $V_1 = n_1 - 1$ $\rightarrow$ $\frac{(n_1 - 1) S_1^2}{S_1^2}$: المعلوم

الجزء الثاني : S_2^2 المعلوم : $S_2^2 = 690$, $n_2 = 25$ درجة ، $V_2 = n_2 - 1$ $\rightarrow$ $\frac{(n_2 - 1) S_2^2}{S_2^2}$: المعلوم

$F = \frac{S_1^2 / S_1^2}{S_2^2 / S_2^2}$: Fisher المعلوم

$Pr(S_1^2 \leq S_2^2) = Pr\left(\frac{S_1^2}{S_2^2} \leq 1\right)$: Jhao! (α_1)
 $= Pr\left(\frac{S_1^2 / S_1^2}{S_2^2 / S_2^2} \leq \frac{S_2^2}{S_1^2}\right) = Pr\left(F \leq \frac{690}{300}\right) = Pr(F \leq 2.3) = 0.95$

$F_{9 \ 24 \ 0.95} = 2.3$ | $\alpha = 0.05$ | $V_1 = 10 - 1 = 9$: درجة الحرية الأولى
 $1 - \alpha = 0.95$ | $V_2 = 25 - 1 = 24$: درجة الحرية الثانية
 0.95 المعلوم

$Pr\left(S_1^2 \leq \frac{3}{2} S_2^2\right) = Pr\left(\frac{S_1^2}{S_2^2} \leq \frac{3}{2}\right)$: المعلوم (α_2)
 $= Pr\left(\frac{S_1^2 / S_1^2}{S_2^2 / S_2^2} \leq \frac{3}{2} \cdot \frac{S_2^2}{S_1^2}\right) = Pr\left(F \leq \frac{3}{2} \cdot \frac{690}{300}\right)$
 $= Pr(F \leq 3.45) = 0.99$

$3.45 \leq F_{9 \ 24 \ 0.99} = 3.25$

$\alpha = 0.01$
 $1 - \alpha = 0.99$: المعلوم