**Republic of Iraq**

**Ministry of Higher Education & Scientific Research**

**University of Anbar**

**College of Education for Pure Sciences**

**Department of Mathematics**

# محاضرات الاحصاء ٢

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

# References:

1) Introduction to Mathematical Statistics, R. V. Hogg and A.T. Craig, (4, 5, 6) edition.
2) Mathematical Statistics with Applications, K. M. Ramachandran and C. P. Tsokos, 2009.
3) Probability and Statistical Inference, Robert V. Hogg ,Elliot A. Tanis and ale L. Zimmerman, Ninth Edition, Pearson Education, USA,2015.
4) Probability and mathematical Statistics, Prasanna Sahoo, University of Louisville,, USA, 2008.

٥)امير حنا هرمز، الاحصاء الرياضي، مطبوعات جامعة الموصل، ١٩٩٠

٦) صفاء يونس الصفاوي، الاحصاء، مطبوعات جامعة الموصل، ٢٠٠٨

الفصل الاول

# نظرية التقدير

في الفصل السابق درسنا من خلال مجموعة من النظريات العلاقة الرياضية بين معالم العينة والمعالم المناظرة لها في المجتمع مثل المتوسط، التباين، النسبة...كما درسنا العلاقة بين شكل توزيع المجتمع وشكل التوزيع الاحتمالي لمعالم العينة. تظهر هذه العلاقات كتوصيف لخصائص العينة ومعالمها ولكنها تستخدم أكثر لتقدير خصائص ومعالم المجتمع محل الدراسة، وهذا ما سنتعرف عليه في هذا الفصل.

## بعض خصائص المقدر

لتقدير معلمة من معالم مجتمع محل دراسة، نحتاج إلى اختيار الإحصائية المناسبة في العينة لتقدير هذه المعلمة. غالبا ما تكون المعلمة المناظرة في العينة هي أحسن مقدر، كأن نقدر متوسط المجتمع $\mu$ من خلال متوسط العينة $\mu_m$. تسمى الإحصائية المستخدمة في التقدير المقدر.

### المقدر غير المتحيز

نقول عن إحصائية ما بأنها مقدر غير متحيز unbiased لمعلمة المجتمع إذا كان متوسطها أو توقعها الرياضي مساويا لمعلمة المجتمع.

**مثال**: نقول عن متوسط العينة m أنه مقدر غير متحيز لمتوسط المجتمع $\mu$ لأن $E(m) = \mu$ . في المقابل نسمي الإحصائية $S^2$ في معاينة بالإرجاع أنها مقدر متحيز ل $\sigma^2$ لأن $E(S^2) = \sigma^2 (n-1)/n \neq \sigma^2$ ، بينما تعتبر الاحصائية $S'^2 = S^2n/(n-1)$ مقدرا غير متحيز في معاينة بالإرجاع.

### الكفاءة

تتعلق كفاءة ( efficient) مقدر ما بمقدار التباين لتوزيع المعاينة للإحصائية، فإذا كان لمقدرين (إحصائيتين) نفس المتوسط نقول عن المقدر ذو توزيع المعاينة الأقل تباينا أنه الأكثر كفاءة.

**مثال**: لكل من توزيعي المعاينة للمتوسط والوسيط نفس المتوسط هو متوسط المجتمع $\mu$ ، لكن يعتبر المتوسط m مقدرا أكثر كفاءة لمتوسط المجتمع $\mu$ من الوسيط لأن تباين توزيع المعاينة للمتوسطات $V(m) = \sigma^2/n$ أقل من تباين توزيع المعاينة للوسيط :

$$V(méd) = \sigma^2\pi/2n = (\sigma^2/n)\,(3.14159/2) \quad > \quad \sigma^2/n.$$

من البديهي أن استخدام مقدرات فعالة وغير متحيزة هو الأفضل، إلا أنه قد يلجأ لمقدرات أخرى لسهولة الحصول عليها.

### التقارب convergence

نقول عن مقدر أنه متقارب إذا كان يؤول إلى قيمة المعلمة المقدرة عندما يؤول حجم العينة إلى ما لا نهاية.

**مثال**: يعتبر متوسط العينة مقدرا متقاربا لمتوسط المجتمع لأن:

$$E(m) = \mu \quad , \qquad V(m) = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0.$$

## التقدير النقطي والتقدير بفترة.

قد نحتاج إلى تقدير لمعلمة مجتمع بقيمة واحدة ونقول عن هذا التقدير أنه **تقدير نقطي**، و أحيانا أخرى نحتاج إلى تقدير معلمة المجتمع بنقطتين يحددان فترة لقيمة المعلمة ونقول عن هذا النوع من التقدير أنه **تقدير بفترة**.

**مثال** : إذا قدرنا دخل الأسرة في منطقة ما ب ١٨٠٠٠ دع، نكون قد قدرنا دخل الأسرة تقديرا نقطيا. يكون تقديرنا بفترة إذا قلنا مثلا أن الدخل يساوي ١٨٠٠٠ ± ٢٠٠٠ أي أنه يتراوح بين ١٦٠٠٠ و٢٠٠٠٠دع.

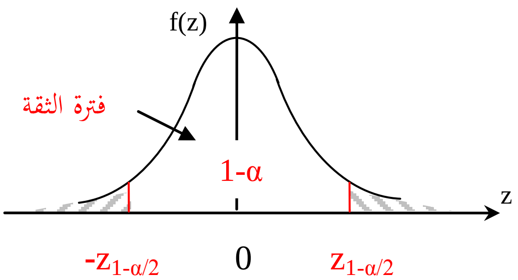
### درجة التأكد

لكي يكون التقدير علميا ينبغي تقييم احتمال أن تكون المعلمة تنتمي فعلا إلى الفترة المحدد، لذلك نلحق بالفترة ما يسمى بدرجة أو مستوى الثقة، ويرمز له ب p. الاحتمال المعاكس يسمى احتمال الخطأ ويرمز له ب $\alpha$ ، ويسمى أيضا ""مستوى المعنوية".

**مثال**: دخل الأسرة في المنطقة (أ) ينتمي إلى الفترة [١٦٠٠٠٠، ٢٠٠٠٠٠] بمستوى معنوية ٥ % أي بمستوى ثقة ٩٥ % . وتسمي الحدود ١٦٠٠٠٠ و٢٠٠٠٠٠ **حدود الثقة**.

### تعيين حدود فترة الثقة

تحدد حدود الثقة من خلال معاملات الثقة التي بدورها تحدد من خلال مستوى المعنوية (مستوى الثقة). ففي حالة استخدام التوزيع الطبيعي للتقدير تكون القيمتين ± ١.٩٦ معاملات الثقة من أجل مستوى ثقة ٩٥% بينما القيمتين ±2.58 تمثلان معاملات الثقة من أجل مستوى ثقة ٩٩ % .




**رسم 1 فترة الثقة للتوزيع الطبيعي**

**مثال**: ليكن $\mu_s$ و$\sigma_s$ متوسط وانحراف معياري توزيع المعاينة لإحصائية ما s حيث $\mu_s = \mu$ . إذا كان توزيع المعاينة ل s توزيعا طبيعيا (كما هو الحال بالنسبة لأغلب الإحصائيات عندما $(n \geq 30)$ ) فإننا نقدر مثلا وبالنظر إلى توزيع s أن:

القيمتين $\mu_s \pm 1.96\sigma_s$ تمثلان **حدود الثقة** ب ٩٥ %، و $\mu_s \pm 2.58\sigma_s$ حدود الثقة ب ٩٩ %.

في حالة التوزيع الطبيعي يرمز لحدود الثقة ب $Z_c$ أو $Z_{1-\alpha/2}$ (أنظر الرسم ١).

## التقدير بفترة

## فترة الثقة للمتوسط

يقدر متوسط المجتمع $\mu$ من خلال الإحصائية m.

### تقدير μ باستخدام التوزيع الطبيعي

نستخدم التوزيع الطبيعي لتحديد فترة الثقة إذا علمنا أن المجتمع الذي سحبت منه العينة يتبع التوزيع الطبيعي.
وفي حالة العينة الممتدة $(n \geq 30)$ يمكن كذلك الاستفادة من نظرية النهاية المركزية أن m تتبع التوزيع الطبيعي.
تكتب حدود فترة الثقة كما يلي:

$$m \pm z_c \cdot \frac{S}{\sqrt{n-1}} \quad \text{أو} \quad m \pm z_c \frac{S'}{\sqrt{n}}$$

وفي حالة σ مجهول: $$m \pm z_c \cdot \frac{\sigma}{\sqrt{n}}$$

و نستخدم هذه الصيغة إلا إذا كان المجتمع محدود (ذا حجم N) والمعاينة نفادية حيث تصبح الصيغة كالآتي:

$$m \pm z_c \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

إلا أنه غالبا ما يكون الانحراف المعياري للمجتمع σ مجهولا، ولذلك نعوض σ في الصيغ السابقة بالمقدر S’ أو S.
الجدول الآتي يبين قيم $z_c$ التي تمثل حدود فترة الثقة بحسب مستوى الثقة :

| مستوى الثقة 1-α | 0.99 | 0.98 | 0.95 | 0.90 | 0.8 | 0.5 |
|---|---|---|---|---|---|---|
| مستوى المعنوية α | 0.01 | 0.02 | 0.05 | 0.10 | 0.2 | 0.5 |
| 1- α/2 | 0.995 | 0.99 | 0.975 | 0.95 | 0.9 | 0.75 |
| $Z_{1-\alpha/2}$ | 82.5 | 2.326 | 1.96 | 1.645 | 1.282 | 0.674 |

**مثال** : نقدر أن μ يوجد داخل الفترة $m \pm 1.96\sigma_m$ بمستوى ثقة 95% (٠.٩٥) أي بمستوى معنوية ٥ % (٠.٠٥)، وداخل الفترة $m \pm 2.58\sigma_m$ بمستوى ثقة 99% أي بمستوى معنوية ٠.٠١ .

### تقدير μ باستخدام التوزيع t :

في حالة العينة الصغيرة $(n < 30)$ و σ مجهول نستخدم توزيع ستيودنت لتحديد فترات الثقة ل μ. مثلا القيم $-t_{0.975}$ ؛ $t_{0.975}$ تحد ٩٥ % من المساحة تحت المنحنى ونقول أن $-t_{0.975}$ ; $t_{0.975}$ تمثل القيم الحرجة أو معاملات الثقة عند مستوى ثقة ٩٥ % ونكتب:

$$-t_{0.975} < \frac{m-\mu}{\hat{S}/\sqrt{n}} < t_{0.975}$$

ومنه نستخلص فترة الثقة ل μ كما يلي:

$$m - t_{0.975} \cdot \frac{\hat{S}}{\sqrt{n}} < \mu < m + t_{0.975} \cdot \frac{\hat{S}}{\sqrt{n}}$$

## فترة الثقة للنسبة

### ١) *حالة المجتمع غير محدود أو المعاينة غير نفادية و العينة الممتدة $(n \geq 30)$ :*

لتكن s إحصائية تمثل نسبة "نجاحات" في عينة ذات حجم $n \geq 30$ مستخرجة من مجتمع ثنائي حيث p هي نسبة النجاحات. تستعمل التوزيع الطبيعي لتقدير p فنعين حدود الثقة ل p كما يلي: $p' \pm z_c\sigma_p$ أين p’ نسبة النجاحات في العينة،

نعلم من الفصل السابق أن $\sigma_p = \sqrt{\frac{pq}{n}}$ ومنه يحدد فترة الثقة ل p كما يلي:

$$p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}}$$

٢) ***حالة كون المجتمع محدود ذا حجم N والمعاينة نفادية:***

$$p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

## فترة الثقة للتباين

لتقدير التباين والانحراف المعياري لمجتمع بفترة ثقة نستعمل الخاصية : $\frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$.

**مثال:** فترة الثقة ب٩٥ % يحدد كما يلي:

$$\chi^2_{0.025} \le \frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \le \chi^2_{0.975}$$

ومنه نستنتج فترة الثقة ل σ كما يلي:

$$\frac{\sqrt{n}S}{\chi_{0.975}} \le \sigma \le \frac{\sqrt{n}S}{\chi_{0.025}} \qquad ou \qquad \frac{\sqrt{n-1}\hat{S}}{\chi_{0.975}} \le \sigma \le \frac{\sqrt{n-1}\hat{S}}{\chi_{0.025}}$$

نظرا لأن توزيع مربع كاي غير متماثل فإن الفترة أعلاه ليس الأمثل، إذ توجد طريقة لتضييق فترة الثقة أكثر إذا لم نشأ أن تكون أطراف المنحنى متساوية، وهذا بخلاف التوزيعات المتماثلة كالطبيعي وستيودنت.

## فترات الثقة لنسبة تباينين

رأينا أنه إذا كان لدينا مجتمعان طبيعيان تبايناهما $\sigma^2_1$ , $\sigma^2_2$ وسحبنا منهما عينتين عشوائيتين حجمهما على التوالي $n_1$ , $n_2$

فإن :
$$F = \frac{\left[\frac{S_1^2 n_1}{n_1 - 1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{S_2^2 n_2}{n_2 - 1}\right]\frac{1}{\sigma_2^2}} = \frac{\hat{S}_1^2 / \sigma_1^2}{\hat{S}_2^2 / \sigma_2^2} \to F_{n_1-1;\, n_2-1}$$

إذا يمكن تكوين تقدير بفترة ل F عند مستوى ثقة ٠.٩٨ كما يلي:

$$F_{0.01} \le \frac{\hat{S}_1^2 / \sigma_1^2}{\hat{S}_2^2 / \sigma_2^2} \le F_{0.99}$$

و من ثم يمكن تقدير النسبة بين تبايني المجتمعين كما يلي:

$$\frac{1}{F_{0.99}} \frac{\hat{S}_1^2}{\hat{S}_2^2} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{1}{F_{0.01}} \frac{\hat{S}_1^2}{\hat{S}_2^2}$$

## خلاصة

لتقدير إحصائية مجتمع نستخدم نظريات توزيع المعاينة. هذه النظريات تتناول خصائص إحصائيات العينة من متوسط العينة، النسبة في العينة، ... و علاقتها بالإحصائيات المناظرة لها في المجتمع.

**جدول1** توزيع المعاينة للمتوسطات حسب طبيعة توزيع المجتمع، معلومية التباين و حجم العينة.

| **قانون المجتمع** | **تباين المجتمع ($\sigma^2$)** | **N** | $\sigma_{\bar{x}}$ | **قانون** $\bar{x}$ |
|---|---|---|---|---|
| طبيعي | معلوم | $n \geq 30$ أو $n < 30$ | $\sigma/\sqrt{n}$ | $N(\mu ; \sigma/\sqrt{n})$ |
| | غير معلوم | $n \geq 30$ | $S'/\sqrt{n}$ | $N(\mu ; S'/\sqrt{n})$ |
| | | $n < 30$ | $S'/\sqrt{n}$ | $t_{\alpha;\, n-1}$ |
| غير معلوم | معلوم | $n \geq 30$ | $\sigma/\sqrt{n}$ | $N(\mu ; \sigma/\sqrt{n})$ |
| | غير معلوم | $n \geq 100$ | $S'/\sqrt{n}$ | $N(\mu ; S'/\sqrt{n})$ |

**جدول ٢** تحديد فترة الثقة للنسبة، للتباين وللنسبة بين تباينين

| **المجتمع** | **التوزيع الاحتمالي للإحصائية** | **فترة الثقة** |
|---|---|---|
| مجتمع غير محدود أو معاينة غير نفادية و عينة ممتدة ($n \geq 30$) | التوزيع الطبيعي | $p' \pm z_c \cdot \sqrt{\dfrac{p(1-p)}{n}}$ |
| مجتمع محدود ذا حجم **N** والمعاينة نفادية | التوزيع الطبيعي | $p' \pm z_c \cdot \sqrt{\dfrac{p(1-p)}{n}} \cdot \sqrt{\dfrac{N-n}{N-1}}$ |
| غير معلوم | $\dfrac{nS^2}{\sigma^2} = \dfrac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$ | $\dfrac{\sqrt{n}S}{\chi_{0.975}} \leq \sigma \leq \dfrac{\sqrt{n}S}{\chi_{0.025}}$ أو $\dfrac{\sqrt{n-1}\hat{S}}{\chi_{0.975}} \leq \sigma \leq \dfrac{\sqrt{n-1}\hat{S}}{\chi_{0.025}}$ |
| مجتمعين طبيعيين، أو عينتين مسحوبتين من مجتمع طبيعي واحد. | $F = \dfrac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}} = \dfrac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \rightarrow F_{n_1-1;\, n_2-1}$ | مثلا عند مستوى ثقة ٠.٩٨: $\chi^2_{0.025} \leq \dfrac{nS^2}{\sigma^2} = \dfrac{(n-1)\hat{S}^2}{\sigma^2} \leq \chi^2_{0.975}$ $\dfrac{1}{F_{0.99}}\dfrac{\hat{S}_1^2}{\hat{S}_2^2} \leq \dfrac{\sigma_1^2}{\sigma_2^2} \leq \dfrac{1}{F_{0.01}}\dfrac{\hat{S}_1^2}{\hat{S}_2^2}$ |

## ملحق. فترات الثقة للفروق والمجاميع

إذا كانت $s_1$ و$s_2$ إحصائيتا معاينة لها توزيع يقترب من التوزيع الطبيعي، والعينتان مستقلان، تكتب حدود الثقة للفروق بين المعالم التي تمثلها الإحصائيتين كما يلي:

$$S_1 - S_2 \pm z_c \cdot \sigma_{S_1-S_2} = S_1 - S_2 \pm z_c \cdot \sqrt{\sigma^2_{S_1} + \sigma^2_{S_2}}$$

في حالة المجموع :

$$S_1 + S_2 \ \pm \ z_c \cdot \sigma_{S_1 - S_2} = S_1 + S_2 \ \pm \ z_c \cdot \sqrt{\sigma^2{}_{S_1} + \sigma^2{}_{S_2}}$$

**مثال**: إذا كانت الإحصائيتان هما متوسطا عينتين مستقلتين، مسحوبتين من مجتمعين غير محدودين، نحدد فترة الثقة للفرق (و للمجموع) بين متوسطي المجتمعين $\mu_1 - \mu_2$ كما يلي :

$$m_1 - m_2 \ \pm \ z_c \cdot \sigma_{m_1 - m_2} = m_1 - m_2 \ \pm \ z_c \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

**مثال ٢** : إذا كانت الإحصائيتان هما نسبتان في عينتين مستقلتين، مسحوبتان من مجتمعين غير محدودين :

$$p'_1 - p'_2 \ \pm \ z_c \cdot \sigma_{p'_1 - p'_2} = p'_1 - p'_2 \ \pm \ z_c \cdot \sqrt{\frac{pq_1}{n_1} + \frac{pq_2}{n_2}}$$

## طرق تأسيس المقدر

أحد الطرق لاختيار مقدر معلمة ما للمجتمع أن نأخذ مباشرة نظيرتها في العينة، وإذا كان هذا المقدر لا يتصف بالخصائص المطلوبة نجري عليه تعديلا (استخدام $S'^2$ بدلا من $S^2$ لتقدير $\sigma^2$). توجد طرق أخرى لتحديد المقدر الأنسب منها طريقة المعقولية العظمى والتي تدعى أيضا طريقة الاحتمال الأكبر والتي تنسب إلى العالم فيشر وكذا طريقة العزوم.

## طريقة العزوم

ليكن المطلوب تقدير عدد K من معالم المجتمع : $\theta_1, \theta_2, \ldots, \theta_k$. نكون جملة معادلات عددها K. تتضمن كل معادلة مساواة العزم الصفي من الدرجة k لمتغير المجتمع X : $\mu'_k = E(X^k)$، بنظيره لمتغير المعاينة x :

$$m'_k = (1/n)\sum_i x_i^k \qquad k = 1, 2, \ \ , K$$

**مثال**: ليكن $X \sim B(20; p)$ . تقدير p بطريقة العزوم انطلاقا من عينة يتم كما يلي:

لدينا عدد المعالم المراد تقديرها $K = 1$ إذا نحتاج إلى معادلة واحدة : $\mu = 20p$ .ومنه $p = 20/\mu$، نأخذ إذا كمقدر لp القيمة: $p'$ ونحسبها كما يلي : $p' = m/20$.

في حالة تقدير معلمتين للمجتمع نحتاج أن نستعمل جملة المعادلتين: $\mu = m \ , \ \mu'_2 = m'_2$

**مثال٢** : لتكن $X \sim N(\mu; \sigma^2)$. نسحب عينة ذات متوسط m، وتباين $S^2$. لتقدير $\mu$ و$\sigma^2$ نحتاج إلى حل جملة المعادلتين:

$$\begin{cases} \mu'_2 = m'_2 \\ \mu = m \end{cases} \text{ or } \begin{cases} \mu'_2 = \mu^2 + \sigma^2 \\ m'_2 = m^2 + S^2 \end{cases} \quad \Longrightarrow \quad \begin{cases} \hat{\mu} = m \\ \hat{\sigma}^2 = S^2 \end{cases}$$

هذه الطريقة قد تعطي مقدرات متحيزة كما في هذه الحالة.

## طريقة المعقولية العظمى (طريقة الاحتمال الأكبر)

حالة كون المتغير المجتمع متقطعة : نريد تقدير معلمة $\theta$ واحدة للمجتمع، ولدينا عينة غير نفاديه (المتغيرات التي تمثل قيم المحصل عليها في العينة مستقلة) لها نفس التوزيع للمجتمع. من البديهي أن احتمال تحقق عينة بذاتها مرتبط ب قيمة المعلمة المجهولة : $P(x_1, x_2, \ldots, x_n) = L(\theta)$. هناك قيمة ل $\theta$ تعظم احتمال الحصول على العينة المحصل عليها، ونفترض أن تلك

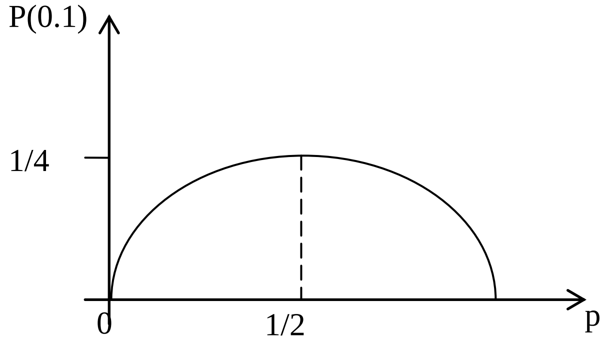
القيمة هي الصحيحة بما أن العينة حصلت بالفعل. تتمثل طريقة المعقولية العظمى في البحث عن هذه القيمة. أي البحث عن $\theta$ التي تعظم $L(\theta)$ ، حيث :

$$L(\theta) = f(x_1, \ldots, x_n ; \theta) = f(x_1) . f(x_2) \ldots f(x_n).$$

تعتمد طريقة المعقولية العظمى على تعظيم دالة الاحتمال المشتركة $L(\theta)$ .

**مثال:** ليكن $X \sim B(p)$، حيث النجاح هو وجود الخاصية " أ " لدى فرد مسحوب عشوائيا من المجتمع. نرد تقدير p من خلال عينة حجمها ٢. ما هي القيمة ’p ل p التي تجعل النتيجة ١، ٠ هي الأكثر احتمالا؟ أي ما هي ’p التي تجعل $p(0.1) = pq$ أكبر ما يمكن؟

من الواضح أن أكبر قيمة ل$p(0.1)$ هي ¼ والقيمة التي تحققها هي $p' = 1/2$ ، وبهذا نجيب على التساؤل.




**رسم 2 أقصى قيمة ل P(0,1)**

# الفصل الثاني

## اختبار الفرضيات

## مفاهيم اختبارات الفروض وتطبيقاتها

في الفصل السابق تناولنا كيفية تقدير معالم المجتمع من خلال بيانات العينة وبعض خصائص المقدر الجيد. في هذا الفصل سنتناول كيفية اختبار فرضيات موضوعة حول معالم مجتمع أو أكثر. يحتاج الدارس أحيانا في مرحلة ما من بحثه إلى اختبار فرضية أو أكثر بخصوص المجتمع المدروس. من أمثلة ذلك: اختبار فرضية بخصوص معدل الدخل في منطقة معينة، اختبار فرضية نسبة شفاء لدواء معين، ... يتم ذلك بصياغة فرضية عن المجتمع المدروس (أو المجتمعات المدروسة) ومن ثم محاولة الحصول على دليل إحصائي ينفي أو يثبت هذه الفرضية وذلك من خلال بيانات عينة (أو أكثر) عشوائية بسيطة. تخص الفرضية أحد معالم المجتمع كالمتوسط، النسبة أو التباين، ونعتمد في إثباتها أو رفضها على خصائص إحصائية معاينة مختارة. من أجل ذلك يعتمد هذا الدرس، كما هو الحال بالنسبة لدرس التقدير، على درس المعاينة.

### اختبار المتوسط

يتناول هذا الاختبار متوسط المجتمع ($\mu$)، مثل متوسط الدخل، متوسط وزن منتج معين، .. ويؤكد اختبار المتوسط فرضية مساواته لقيمة ما $\mu_0$. و للقيام بالاختبار نستخرج عينة عشوائية نحسب فيها المتوسط m ثم نستخدم التوزيع الاحتمالي ل m لقياس قرب أو بعد هذه القيمة من $\mu_0$.

#### اختبار ثنائي الاتجاه للمتوسط.

لنتناول هذا المثال: نريد اختبار فرضية حول متوسط دخل الطالب في السنة الأولى من تخرجه، ولتكن القيمة الافتراضية هي ١٥٠٠٠٠دع كمتوسط للدخل الشهري. نحتاج إلى الخطوات التالية: تحديد الفرضيات، تحديد قاعدة القرار، حساب القيمة الجدولية للمتغير، حساب القيمة الفعلية للمتغير، اتخاذ القرار.

**تحديد الفرضيات** (**الصفرية والبديلة**):

$$H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu \neq \mu_0$$

تسمى الفرضية $H_0$ **الفرضية الصفرية** أو فرضية العدم، ويؤدي الاختبار إما إلى رفضها ونكتب $RH_0$ وفي هذه الحالة نقبل **الفرضية البديلة** أو المعاكسة أو عدم رفضها ونكتب $R'H_0$. $\mu_0$ هي القيمة الافتراضية ل $\mu$ وهي في هذه الحالة ١٥٠٠٠٠ لذلك نكتب الفرضيات كما يلي:

$$H_0 : \mu = 150000 \leftrightarrow H_1 : \mu \neq 150000$$

عادة ما تكون $\mu_0$ محددة بناءا على بيانات عينة عشوائية بسيطة ($\mu_0 = m$)، وفي هذه الحالة يمكن استخدام الخاصية $m \sim N(\mu, \sigma^2/n)$ لأجراء الاختبار، حيث أنه تحت $H_0$ فإن : $m \sim N(\mu_0 , \sigma^2/n)$

مما يعني معلومية احتمال أن يكون m قريب إلى درجة ما من $\mu_0$ فمثلا :

$$P(\mu_0 - 1.64(\sigma_m) \leq m \leq \mu_0 + 1.64(\sigma_m)) = 0.90$$
$$P(\mu_0 - 1.96(\sigma_m) \leq m \leq \mu_0 + 1.96(\sigma_m)) = 0.95$$
$$P(\mu_0 - 2.58(\sigma_m) \leq m \leq \mu_0 + 2.58(\sigma_m)) = 0.99$$

وبصفة عامة نكتب:

$$P[\mu_0 - z_{1-\alpha/2}(\sigma_m) \leq m \leq \mu_0 + z_{1-\alpha/2}(\sigma_m)] = 1-\alpha$$

أو حسب الكتابة الأكثر شيوعا:

$$P(-z_{1-\alpha/2} \leq \frac{m-\mu_0}{\sigma_m} \leq z_{1-\alpha/2}) = 1-\alpha$$

حيث:

- $(m - \mu_0)/\sigma_m$ : (**متغير القرار**) هو المتغير المعياري ل m ونرمز لها ب $z_c$، حيث $z \sim N(0, 1)$ .
- $\sigma_m$ تحدد كما يلي: $\sigma_m = \sigma/\sqrt{n}$ في حالة المعاينة بالإرجاع (أو $n \leq 0.05N$ ) و $\sigma_m = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ في الحالة المعاكسة.
- $1 - \alpha/2$ : المساحة على يسار z .
- n : حجم العينة.

يمكن إذا كان m خارج الفترة $1-\alpha$، أن نرفض الفرضية الصفرية التي حدد على أساسها هذه الفترة ونقبل بالتالي الفرضية البديلة. تسمى هذه (الخطة) قاعدة القرار.

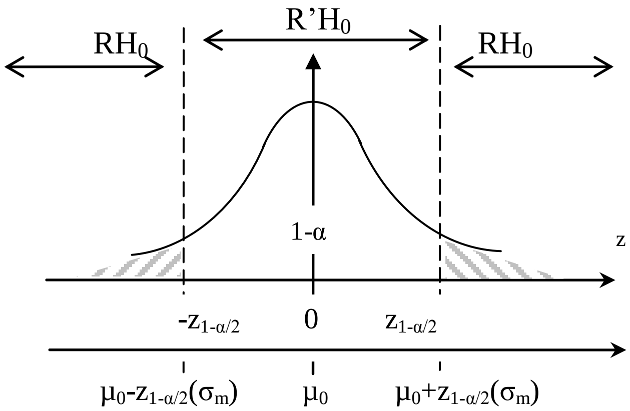
### تحديد قاعدة القرار

تكتب قاعدة القرار في المثال الذي بين أيدينا، وهي قاعدة **اختبار ثنائي الاتجاه** (أنظر الشكل ١) ، كما يلي:

$$\begin{cases} RH_0 \text{ for } z_c = \dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} \notin [-z_{1-\alpha/2};\ z_{1-\alpha/2}] \\ \bar{R}H_0 \qquad \text{otherwise.} \end{cases}$$

أو

$$\begin{cases} RH_0 \text{ for } \left|\dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}}\right| > z_{1-\alpha/2} \\ \bar{R}H_0 \qquad \text{otherwise.} \end{cases}$$




**رسم 3** منطقتي القبول و الرفض في حالة قاعد القرار الثنائية

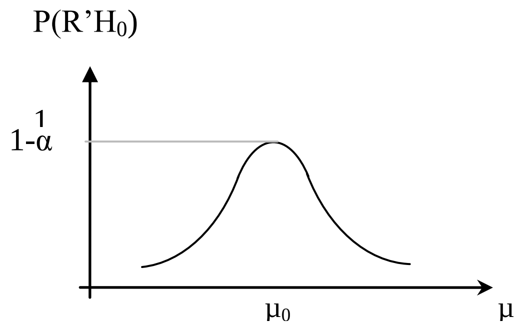
تتضمن هذه الخطة مخاطرة تتمثل في الوصول إلى قرار خاطئ: فقد تكون الفرضية $H_0$ صحيحة بينما تقودنا قيمة m المحصلة إلى رفضها، ويسمى هذا **الخطأ من النوع الأول**، واحتماله α ، ويكتب : $P(RH_0 / H_0) = \alpha$،

و قد تقودنا قيمة m إلى قبول $H_0$ فيما هي خاطئة، ويسمى هذا **الخطأ من النوع الثاني** واحتماله α -1 ويكتب :

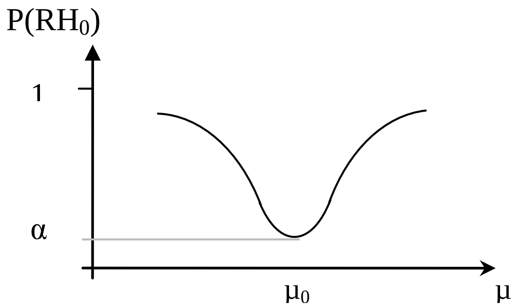
$$P(R'H_0 / H_1) = 1-\alpha$$

و يمكن تقليص احتمال أحد الخطأين على حساب الثاني، ولكن لا يمكن تقليص احتمال كلا الخطأين معا إلا بزيادة حجم العينة.

و يقيس احتمال رفض الفرضية الصفرية $P(RH_0)$ قوة الاختبار (أنظر الشكل ٤) فيما يقيس احتمال قبولها $P(R'H_0)$ فعالية الاختبار (أنظر الشكل ٤). ويتوقف كلا الاحتمالين على القيمة الحقيقية ل μ .




رسم 4 (1) منحنى الفعالية




(2) منحنى القوة

**(ب) حساب z الجدولية:**

ويرمز لها ب $z_t$ حيث، وهي المشار إليها في قاعدة القرار (الشكل الثاني)، وفي حالتنا (اختبار ثنائي بمستوى معنوية ٥%):

$$z_t = z_{1-\alpha/2} = z_{1-0.05/2} = z_{1-0.025} = z_{0.975}$$

ومن الجدول نجد أن $z_{0.975} = 1.96$ .

**(ج) حساب z الفعلية:**

ويرمز لها ب $z_c$ وهي المتغير المعيارية ل m (أنظر قاعدة القرار الشكل الأول) :

$$z_c = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \frac{158000 - 150000}{1500 / \sqrt{100}} = 5.33$$

**(د) القرار:**

نقرر قبول أو رفض $H_0$ حسب قاعدة القرار. وفي حالتنا نرفض $H_0$ لأن $z_c > z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف ليس ١٥٠٠٠٠دع.

## الاختبار أحادي الاتجاه للمتوسط.

يتميز الاختبار الثنائي عن الأحادي في الفرضية البديلة التي هي عدم مساواة في الاختبار الثنائي وأكبر تماما أو أصغر تماما (حسب الحالة) في الاختبار الأحادي، وهذا يترتب عليه تغيير في قاعدة القرار.

### ١- الاختبار أحادي الاتجاه من اليمين.

لنرجع إلى المثال السابق مع تغيير محدد هو أننا نريد اختبار ما إذا كان متوسط الدخل للخريج ١٥٠٠٠٠دع أم أكثر (اختبار من اليمين).

أ- **الفرضيات** : $H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu > \mu_0$

في هذه الحالة $\mu_0 = 150000$ لذلك نكتب : $H_0 : \mu = 150000 \leftrightarrow H_1 : \mu > 150000$

ب- **قاعدة القرار**:

$$\begin{cases} RH_0 \text{ for} \dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} > z_{1-\alpha} & . \\ \bar{R}H_0 & \text{otherwise.} \end{cases}$$

ج- **حساب z الجدولية**: (اختبار على اليمين بمستوى معنوية ٥%) : $z_t = z_{1-\alpha} = z_{1-0.05} = z_{0.95}$

ومن الجدول نجد أن $z_{0.95} = 1.645$

د- **حساب z الفعلية**:

$$z_c = \frac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{158000-150000}{1500/\sqrt{100}} = 5.33$$

هـ- **القرار**: نرفض $H_0$ لأن $z_c > z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف ليس ١٥٠٠٠٠دع وإنما هو أكبر.

### ٢- الاختبار أحادي الاتجاه من اليسار

نعود إلى مثالنا ونفترض أن متوسط العينة كان ١٤٢٠٠٠دع ونريد أن نختبر ما إذا كان متوسط الدخل مساوي أم أقل من ١٥٠٠٠٠دع.

أ- **الفرضيات** : $H_0 : \mu = 150000 \leftrightarrow H_1 : \mu < 150000$

ب- **قاعدة القرار**:

$$\begin{cases} RH_0 \text{ for} \dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} < -z_{1-\alpha} & . \\ \bar{R}H_0 & \text{otherwise.} \end{cases}$$

ج- **حساب z الجدولية**: (اختبار على اليسار بمستوى معنوية ٥ %) :

$z_t = -z_{1-\alpha} = -z_{1-0.05} = -z_{0.95} = -1.645$

د- حساب z الفعلية:

$$z_c = \frac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{142000 - 150000}{1500/\sqrt{100}} = -5.33$$

هـ- **القرار**: نرفض $H_0$ لأن $z_c < z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف أقل من ١٥٠٠٠٠دع .

## استخدام S كمقدر ل σ في اختبار المتوسط.

في الأمثلة السابقة افترضنا أن σ معلوم، في الواقع غالبا ما يكون الانحراف المعياري مجهولا ونحتاج بالتالي إلى استخدام الانحراف المعياري للعينة (S) عند حساب $\sigma_m$ (أنظر درس التقدير)، حيث نعوض العبارة

$$\sigma m = \sigma/\sqrt{n} \quad \text{ب} \quad \sigma_m = \frac{S}{\sqrt{n-1}} \quad \text{أو} \quad \sigma_m = \frac{S'}{\sqrt{n}}$$

**مثال**: في المثال السابق نفترض أن الانحراف المعياري للدخل الشهري للطالب مجهول، لكن الانحراف المعياري للعينة S = 1600. كيف يمكن اختبار ما إذا كان الدخل الشهري أقل من ١٥٠٠٠دع؟

الخطوات أ، ب ، ج تبقى بدون تغيير.

د- حساب z الفعلية:

$$z_c = \frac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}} = \frac{14200-1500}{1600/\sqrt{99}} = -4.97$$

هـ- **القرار**: نرفض $H_0$ لأن $z_c < z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف ليس ١٥٠٠٠دع وإنما هو أقل.

## استخدام التوزيع t في اختبار المتوسط.

في حالة $n < 30$ و σ (الانحراف المعياري للمجتمع) مجهولا، لا يمكن استخدام التوزيع الطبيعي، ولكن لدينا :

$$\frac{\bar{X}-\mu}{S/\sqrt{n-1}} \sim t_{n-1} \quad \text{و تحت } H_0 \ (\mu = \mu_0): \quad \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}} \sim t_{n-1}$$

يمكن إذا استخدام التوزيع ستيودنت (بشرط أن يكون توزيع المجتمع طبيعيا أو على الأقل جرسي الشكل) .

و تتغير قاعدة القرار تبعا لهذا التغيير فتكتب في حالة الاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \text{ for } \left|\dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}}\right| > t_{n-1;\,1-\alpha/2} \\ \bar{R}H_0 \text{ otherwise} \end{cases}$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \text{ for } \dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}} > t_{n-1,\,1-\alpha} \\ \bar{R}H_0 \text{ otherwise} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \text{ for} \dfrac{\bar{X} - \mu_0}{S / \sqrt{n-1}} < -t_{n-1;\,1-\alpha} \\ \bar{R}H_0 \text{ otherwise} \end{cases}$$

**خلاصة**

يتم اختبار الفرضيات من خلال ٥ خطوات متتالية وهي:

- تحديد الفرضيات (الصفرية والبديلة)
- تحديد قاعدة القرار
- حساب القيمة الجدولية للمتغير
- حساب القيمة الفعلية للمتغير
- اتخاذ القرار.

تتحدد كيفية إتمام كل خطوة حسب طبيعة الاختبار (ثنائي أو أحادي الاتجاه)، حسب طبيعة المجتمع و طبيعة و حجم العينة، ... و تستخدم في ذلك نظريات توزيع المعاينة.

## اختبار النسبة واختبار التباين

اختبار النسبة
اختبار التباين

### اختبار النسبة

يتعلق هذا الاختبار بنسبة مفردات المجتمع التي تتصف بخاصية ما (p)، حيث يؤكد الاختبار أو ينفي صحة فرضية معينة بخصوص قيمة p . يرمز للقيمة الافتراضية ب $p_0$ وتكتب الفرضية كما يلي: $H_0 : p = p_0$

للقيام بالاختبار نستخدم خصائص ’p النسبة في العينة (أنظر توزيع المعاينة للنسبة : نظرية ٦).

$$E(p') = \mu_{p'} = p \qquad ; \qquad \sigma^2_{p'} = \frac{pq}{n}$$

عند $n \geq 30$ : $p' \approx N(p, \sigma_{p'})$ (نظرية موافر- لابلاس)

استنادا إلى هذه الخصائص وتحت H0 : $p' \approx N\left(p_0 ; \sqrt{\dfrac{p_0 q_0}{n}}\right)$

و من ثم يمكن تحديد قاعدة القرار بحسب طبيعة الاختبار كما يلي:

في حالة الاختبار الثنائي:

$$\begin{cases} RH_0 \text{ for} \left| \dfrac{p' - p_0}{\sqrt{p_0 q_0 / n}} \right| > z_{1-\alpha/2} \;. \\ \bar{R}H_0 \qquad \text{otherwise.} \end{cases}$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \text{ for } \dfrac{p'-p_0}{\sqrt{p_0q_0/n}} > z_{1-\alpha} \\ \overline{R}H_0 \qquad \text{otherwise.} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \text{ for } \dfrac{p'-p_0}{\sqrt{p_0q_0/n}} < -z_{1-\alpha} \\ \overline{R}H_0 \qquad \text{otherwise.} \end{cases}$$

**مثال:** تقدر الدوائر الرسمية نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم ب ٧٠ % . وجدت دراسة أجريت على عينة من ٩٠٠ طالب أن نسبة الحصول على عمل ٦٧ % . كيف يمكن اختبار ما إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها، بمستوى معنوية ٥ %.

$$H_0 : p = 0.70 \leftrightarrow H_1 : p < 0.70$$

$$\frac{p'-p_0}{\sqrt{p_0q_0/n}} = \frac{0.67-0.7}{\sqrt{0.7(0.3)/900}} \cong -196.34 < -z_{1-0.05} = -1.64$$

ومنه نرفض الفرضية $H_0$ .

## اختبار التباين

لاختبار صدق الفرضية بخصوص قيمة تباين مجتمع ما،

$$H_0 : \sigma^2 = \sigma_0^2 \quad \leftrightarrow H_1 : \sigma^2 = \sigma_0^2$$

نستعمل المقدر غير المنحاز $\hat{S}^2 = \dfrac{\sum_i (X_i - \overline{X})^2}{n-1}$ . حيث في حالة العينة الكبيرة ($n \geq 50$ في أحسن الأحوال) ، وتحت $H_0$ فإن

$$T = \frac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \approx N(0.1). \qquad \mu_4 = E(X-\mu)^4$$

حيث $\mu_4$ هو العزم المركزي من الدرجة الرابعة. وبهذا الشكل تكتب قاعدة القرار للاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \; for \left| \dfrac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \right| > z_{1-\alpha/2} \\ \overline{R}H_0 \qquad \text{otherwise} \end{cases}$$

وفي حالة $\mu_4$ مجهول يمكن استخدام كمقدر : $m_4 = E(xi - m)^4$ .

وإذا كان المجتمع طبيعيا، حيث $\mu_4 = 3\sigma^4$ ، فإن **متغير القرار** يمكن أن تكتب كما يلي:

$$\begin{cases} RH_0 \; for \left| \dfrac{\hat{S}^2 - \sigma_0^2}{\sqrt{(3\sigma^4 - \sigma_0^4)/n}} \right| > z_{1-\alpha/2} \\ \overline{R}H_0 \qquad \text{otherwise} \end{cases}$$

## اختبار المقارنة بين مجتمعين

يتناول هذا الاختبار مقارنة بين مجتمعين من خلال المتوسط أو التباين لكل منهما ... وسنركز هنا على المتغير القرار، إذ من السهل على الطالب استنتاج كيفية إتمام الخطوات الأخرى على ضوء ما سبق.

### اختبار تساوي متوسطي مجتمعين

الغرض من الاختبار هو تأكيد أو نفي تساوي متوسطي مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين. تكتب الفرضيات (في حالة الاختبار الثنائي) كما يلي: $H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2$

لتحديد المتغير القرار نعتمد في الاختبار على المتغير القرار T أو 'T بحسب الحالة (نترك للطالب استنتاج قاعدة القرار)، حيث نميز بين حالة كون تباينا المجتمعين معلومين وحالة كون تباينا المجتمعين مجهولين.

#### تباينا المجتمعين معلومين

١- المجتمعين طبيعيين:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0.1)$$

٢- مجتمعين ما ($n_1 , n_2 \geq 30$):

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \approx N(0,1)$$

#### تباينا المجتمعين مجهولين

١- المجتمعان طبيعيان: إذا كان تباينا المجتمعين متساويين

$$T' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$$

٢- مجتمعين ما $(n_1 , n_2 \geq 30)$:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} \approx N(0,1)$$

**مثال**: نسحب من مجتمعين طبيعيين متساويي التباين عينتين حجم الأولى ١٨ وحجم الثانية ٢١. وجدنا النتائج التالية: $m_1 = 81, m_2 = 76, S^2_1 = 9, S^2_2 = 8$. كيف يمكن اجراء اختبار تساوي متوسطي المجتمعين بمستوى معنوية ٥ %.

$$H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2$$

$$T' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{81 - 76}{\sqrt{\frac{18(9) + 21(8)}{18 + 21 - 2}\left(\frac{1}{18} + \frac{1}{21}\right)}} \cong 5.43$$

$$t_{0.975;37} \cong 2.336 < 5.43 \qquad \Rightarrow RH_0$$

### اختبار تساوي تبايني مجتمعين

الغرض من الاختبار هو تأكيد أو نفي تساوي تباينا مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين.

تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي: $H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$

نعتمد في الاختبار على المتغير القرار T أو T’ بحسب الحالة، حيث نميز بين حالة كون المجتمعين طبيعيين أم غير ذلك.

#### مجتمعين طبيعيين

١- الحالة العامة:

$$T' = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_1-1;\, n_2-1}$$

٢- في حالة $n_1 , n_2 \geq 30$

$$T = \frac{1}{2}\ln\left(\frac{\hat{S}_1^2}{\hat{S}_2^2}\right) / \sqrt{\frac{1}{2}\left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}\right)} \approx N(0;1)$$

#### مجتمعين ما $(n_1, n_2 \geq 50)$

**١-** $\mu_4^{(1)} ; \mu_4^{(2)}$ معروفين :

$$T = (\hat{S}_1^2 - \hat{S}_2^2) / \sqrt{\frac{\mu_4^{(1)} - \hat{S}_1^4}{n_1} - \frac{\mu_4^{(2)} - \hat{S}_2^4}{n_2}} \approx N(0;1)$$

٢- في حالة $\mu_4^{(2)}$ ; $\mu_4^{(1)}$ غير معروفين : نعوض $\mu_4$ ب $m_4$ .

**مثال** : نسحب من مجتمعين طبيعيين عينتين حجم الأولى ١٨ وحجم الثانية ٢١. وجدنا النتائج التالية:

$m_1 = 81, m_2 = 76, S^2_1 = 9, S^2_2 = 8.$ كيف يمكن اجراء اختبار تساوي متوسطي المجتمعين بمستوى معنوية ٥ % ؟

$$H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$$

$S'^2_1 = S^2_1 * n_1 / (n_1-1) = 9 (18)/17 \approx 9.53$ ; $S'^2_2 = S^2_2 * n_2 / (n_2-1) = 8 (21)/20 = 8.4$

$S'^2_1 / S'^2_2 \approx 1.135$ ; $F_{0.05 ;17 ;20} \approx 2.17$

$$T < F_{\alpha\ ; n1-1 ;n2-1} \Rightarrow R'H_0$$

## اختبار الاستقلال والتجانس

### اختبار التجانس

لنعد إلى اختبار النسبة، ونفترض أن لدينا عددا $k$ من الخصائص المتنافية، نسبة تحقق كل منها في المجتمع $p_i$ حيث $\sum p_i = 1$. نريد اختبار فرضية تساوي هذه النسب:

$$H_0 : p_i = p_{i0} \ \ (i = 1, 2, . . . k ) \leftrightarrow H_1 : p_i \neq p_{i0}$$

(الفرضية البديلة هي أن إحدى النسب النظرية $p_{i0}$ على الأقل غير مساوية للقيمة الحقيقية.)

**المتغير القرار** : لإنجاز الاختبار نستخرج عينة نحسب فيها عدد مرات تحقق الخصائص $(n_i)$ . إذا تحققت الشروط

$n \geq 30$ ، $p_{i0} \geq 1$ وعلى الأقل في ٨٠ % من الحالات $np_{i0} \geq 5$ نبرهن أن :

$$T = \sum_i \frac{(n_i - np_{i0})^2}{np_{i0}} \approx \chi^2_{k-1}$$

### اختبار التعديل

تستخدم هذه الطريقة أيضا لاختبار تعديل توزيع معين بتوزيع آخر، وفي هذه الحالة نقارن بين تكرارات العينة (التكرارات الحقيقية) $n_i$ وتكرارات افتراضية $n_{i0}$ ، حيث تصاغ الفرضيات كما يلي : **على الأقل إحدى التكرارات النظرية** $n_{i0}$ غير **مساوية للتكرار الحقيقي** $n_i$

$$H_0: n_i = n_{i0} \ \ (i = 1, 2, \dots k) \leftrightarrow H_1: n_i$$

$$T = \sum_i \frac{(n_i - n_{i0})^2}{n_{i0}} \approx \chi^2_{k-m-1}$$

m عدد من معالم من المجتمع المقدرة انطلاقا من بيانات العينة لتحديد التكرارات النظرية.

**مثال**: يتقدم إلى انتخابات معينة ٣ مرشحين: أ، ب ، ج. نريد اختبار فرضية بمستوى معنوية ٥ % حول شعبيتهم كما يلي:

$$H_0 : p_1 = 0.4, p_2 = 0.35, p_3 = 25$$

أجري استجواب ل ٤٠٠ ناخب فكان توزع فئات المساندين على التوالي : ١٧٠، ١٣٥، ٩٥ . لدينا $n = 400 \geq 30$، الأعداد الافتراضية $np_{i0} = 160, 140, 100 \geq 1$ ، وأكثر من ٨٠ % من $np_{i0} \geq 5$ .

$$T = \sum_i \frac{(n_i - np_{i0})^2}{np_{i0}} = \frac{(170-160)^2}{160} + \frac{(135-140)^2}{140} + \frac{(95-100)^2}{100} = 1.05$$

$X^2_{2\ ; 0.95} = 5.99 > 1.05 \Rightarrow R'H_0$ .