جـامعة الجزائر 3

كلية العلوم الاقتصادية، العلوم التجارية و علوم التسيير

محاضرات في مقياس الإحصاء3

(الإحصاء الاستدلالي)

مع تطبيقات للأعمال الموجهة

مطبوعة موجهة لطلبة السنة الثانية (السداسي الثالث)

شعبة علوم اقتصادية، علوم تجارية و علوم التسيـير

إعداد : د. توات عثمان

أستاذ محاضر أ ، قسم العلوم الاقتصادية

2018

هذه المطبوعة

هذه المطبوعة هي عبارة عن محاضرات مقياس الإحصاء 3 و التي تتضمن دروس الإحصاء الاستدلالي، حسب البرنامج الوزاري الجديد و الموحد لمقياس "إحصاء 3" للسداسي الثالث ميدان علوم اقتصادية، علوم تجارية و علوم تسيير. برمج هذا المقياس لطلبة السنة الثانية، لكي يستفيد الطلبة من القاعدة التي اكتسبوها عند دراسة الإحصاء الوصفي و الرياضي في السنة الأولى، لكن هدفه الأساسي هو التمهيد لدراسة التطبيقيات الإحصائية في السداسيات اللاحقة. هدف هذا المقياس هو تقديم علم الإحصاء الاستدلالي، أي الأساس الرياضي للاقتصاد. و باعتباره فرعا من الإحصاء، يدرس هذا المقياس في كليات العلوم و الهندسة، بالإضافة إلى كليات العلوم الاقتصادية و الاجتماعية.

تعد هذه المطبوعة ثمرة تجربة الأستاذ في تدريس المادة بكلية العلوم الاقتصادية، العلوم التجارية و علوم التسيير بجامعة الجزائر 3، سواء كأعمال موجهة أو كمحاضرة. ولقد حاولنا أن نستفيد من هذه التجربة لصياغة محتوى المقياس بطريقة تلائم مستوى طلبة هذه الكلية و طبيعة التخصص، من جهة. و من جهة أخرى تم تصميم المحاضرات لتتناسب مع التوزيع الوزاري الجديد الذي جعل هذا المقياس يقدم خلال سداسي واحد (السداسي الثالث) و بعدد 14 محاضرة موزعة على 14 أسبوع بحجم ساعي يقدر بـ 45 ساعة موزعة بين محاضرة و أعمال موجهة. لتحقيق هذا الغرض حرصنا على عدم التعمق خاصة في المبرهنات على الخواص الرياضية كما تم ربط المفاهيم والقواعد النظرية باستخداماتها التطبيقية؛ فعملنا على إعطاء أمثلة محلولة عن كل مفهوم جديد، تم إلحاق كل فصل بمجموعة من التمارين التطبيقية للأعمال الموجهة و هذا حتى يتمكن الطالب من التمرن عليها.

هذا وننبه الطلبة الأعزاء إلى أنه يفترض بهم عند دراسة الإحصاء الاستدلالي أن يكونوا قادرين على استيعاب المفاهيم الرياضية بعموميتها و لا يبقوا خيالهم حبيس الأمثلة والمسائل المعطاة، فالإحصاء الاستدلالي غير الإحصاء الوصفي، إذ يعنى بتطبيقات هذه المفاهيم التي تم اكتسباها من قبل.

من أجل الاستجابة لحدود التوزيع الزمني للمقياس تم الموازنة بين الفصول كما رأينا أن نتبع البرنامج الوزاري، فتم توزيعه على ثلاثة فصول. ولقد قسمنا الفصول إلى عدة عناصر حسب متطلبات الدرس. بحيث يوافق الفصل أربع أو خمس محاضرات في أغلب الأحيان، و التزمنا في الغالب الأعم بالمنهج المقرر، لكن سوف يجد القارئ أننا لم نتوسع في بعض الجوانب، فله أن يتوسع من خلال الاطلاع على المراجع الموضوعة في آخر المطبوعة، إن رأى أنه قد تمكنه من فهم النقاط الرئيسية المقررة، و إلا فإننا ننصحه بأن يمر عليها مرور الكرام. و غني عن الذكر أن محتوى هذه المطبوعة من نظريات وقواعد ليس من إبداع مؤلفها، و إنما هي قواعد موجودة في المراجع جمعناها وعرضناها بأسلوب رأينا أنه الأنسب لمستوى طالب كلية العلوم الاقتصادية. وإذ نقدم لطلبتنا و زملائنا هذا العمل المتواضع، نهيب بهم أن لا يبخلوا علينا بملاحظاتهم وتعليقاتهم حتى نستفيد منها لطبعات مقبلة بحول الله.

فهرس المحتويات

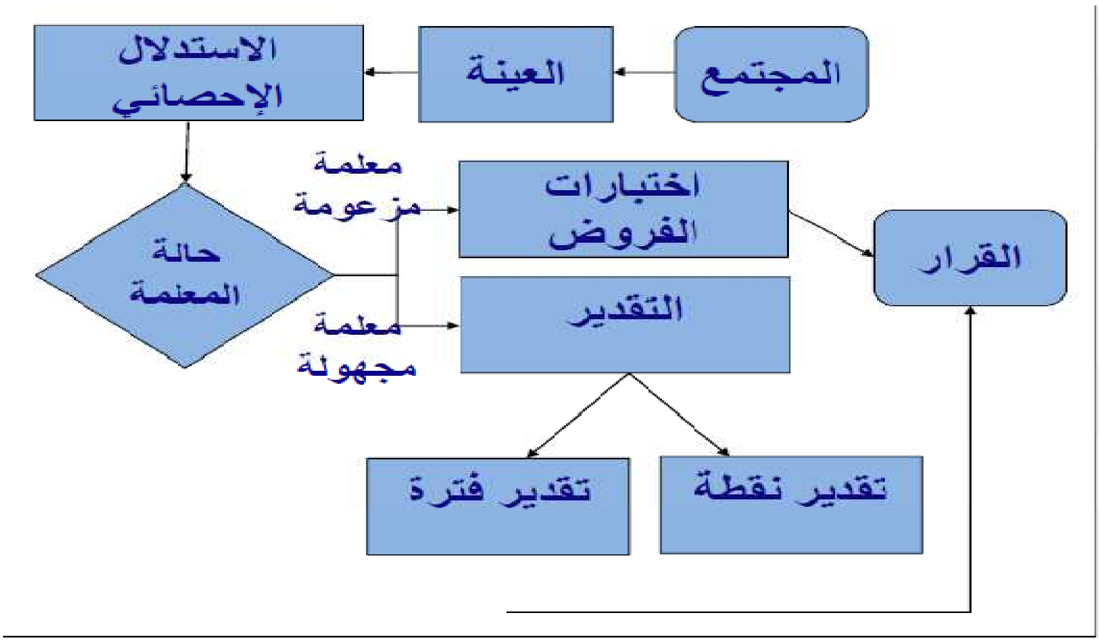
الفصل الأول: توزيع المعاينة

تمهيد:

الهدف من الإحصاء الاستدلالي استنتاج خصائص المجتمع من خصائص عينة سحبت منه, فعند استخدام بيانات عينة (Statistiques) للاستدلال عن المجتمع ولكوننا لا نملك كل حقائق المجتمع فنبحث عن طريقة عملية نستطيع من خلالها تقدير معلمات(Paramètres) المجتمع المطلوب محاولين الوصول للقيم (العددية) لمعالم المجتمع من خلال بيانات العينة المسحوبة منه.

و ينقسم الاستدلال الإحصائي لقسمين، الأول التقدير الإحصائي (estimation statistique) والثاني اختبارات الفروض الإحصائية (tests des hypothèses) فالأول يشير للطرق المختلفة لتقدير معالم المجتمع المجهولة في حين يشير الثاني إلى الاستدلال على معالم المجتمع المجهولة بالنسبة للسؤال محل البحث ومن الممكن استخدام كلاهما معاً حال تحليل البيانات.

يمكن تمثيل ملخص محتوى المقياس و أهدافه من خلال المخطط التالي:



1.مفاهيم أساسية:

لابد في البداية من التطرق إلى بعض المفاهيم الأساسية و التي تتضمن:

أ.مفهوم المجتمع الإحصائي Population

يعرف المجتمع على أنه مجموعة من الأفراد أو الوحدات محل الدراسة و التي لها خصائص مشتركة، أي مجموعات من المفردات تشترك في صفة أو صفات معينة وتكون موضوع دراسة أو بحث.

ويقسم المجتمع الإحصائي إلى قسمين:

-مجتمع محدود :و الذي يكون فيه عدد محدود من الأفراد مثل عدد أجهزة الكمبيوتر في المعمل،

عدد طلاب الفرقة الأولى في كلية ما...الخ.

-مجتمع غير محدود :هو المجتمع الذي يكون فيه عدد الأفراد غير منته مثل عدد النجوم في السماء، عدد حبات القمح المحصود في مزرعة معينة...الخ.

ب.معالم المجتمع Paramètres d'une population

المعلمة Paramètre عبارة عن خاصية أو مقياس يتم حسابه من مجتمع الدراسة، حيث يتكون المجتمع الإحصائي من مجموعة من المفردات التي يهمنا دراستها و هذا المجتمع له بعض معالم أو الخصائص مثل متوسط المجتمع μ و انحرافه المعياري σ و نسبة صفة معينة p في المجتمع.

مثال: - نسبة البطالة في المجتمع.

- متوسط العمر المتوقع عند الولادة.

ج.العينة échantillon

بالنظر للصعوبات التي تواجه الباحثين في الحصول علي بيانات المجتمع ككل يتم اللجوء إلى أسلوب العينة في جمع البيانات ، وفيه يتم جمع البيانات عن جزء من مفردات المجتمع يختار بطريقة أو بأخرى ويطلق عليه عينة (Sample, échantillon) ثم بعد ذلك يتم تعميم نتائج الدراسة على المجتمع بأكمله. أي أن أسلوب المعينة يقصد به دراسة خصائص المجتمع من خلال دراسة عينه مسحوبة منه،

ويمكن تعريف العينة بأنها جزء من المجتمع يتم اختيارها بطرق مختلفة بغرض دراسة هذا المجتمع، بحيث تمثل خصائصه و تمثله أحسن تمثيل.

تنقسم العينات عادة إلى قسمين رئيسين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لكل قسم منها:

- **العينة العشوائية échantillon aléatoire**

هي تلك العينات التي يتم اختيار مفرداتها حسب طريقة إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي وحدة فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة.

- **العينة غير العشوائية:**

هي تلك العينات التي لا تكفل لجميع وحدات المجتمع احتمال ثابت ومحدد للاختيار، وغالباً يتدخل الباحث في عملية الاختيار بصورة أو بأخرى.

د.إحصائية العينة Statistique de l’échantillon

المقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله يطلق عليها معالم المجتمع (Parameters of population, Paramètres d’une population)، أما المقاييس الإحصائية التي تحسب من بيانات عينة مسحوبة من مجتمع الدراسة فيطلق عليها إحصائية (Statistic, Statistique)، و تعتبر كل إحصائية منها بمثابة تقدير أو قيمه تقديرية لمعلمة المجتمع المناظر، فيكون المتوسط الحسابي المحسوب من بيانات العينة تقدير لمعلمة المجتمع المناظرة وهي المتوسط الحسابي المسحوب منه هذه العينة وهكذا.. و يجب ألا يغيب عن الأذهان بأن حساب قيمة المتوسط الحسابي من بيانات العينة ليس هدفاً في حد ذاته ولكن وسيلة للتعرف على المتوسط الحسابي للمجتمع موضوع الدراسة.. وهكذا بالحال بالنسبة لباقي المقاييس الإحصائية التي تحسب من العينة.

للتفرقة بين المعالم والإحصاءات يجب أن نرمز لكل منها برموز تختلف عن رموز الأخرى، على سبيل المثال يرمز للمتوسط الحسابي للمجتمع بالرمز μ بينما يرمز للمتوسط الحسابي للعينة بالرمز $\bar{x}$ ، أيضاً للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز s وهكذا.

2. أهم توزيعات المعاينة:

نفرض أن لدينا مجتمع حجمه(N) ، اخترنا منه عينة عشوائية حجمها (n) مفردة و حسبنا وسطها وليكن x_1 ثم عينة ثانية لها نفس الحجم (n) و حسبنا وسطها الحسابي وليكنx_2 ، ثم عينة ثالثة لها نفس الحجم(n) وحسبنا وسطها الحسابي وليكن x_3 ، وهكذا بالنسبة لجميع العينات التي يمكن سحبها من هذا المجتمع.

سيتوفر لدينا عدد كبير من القيم للوسط الحسابي، لا نتوقع أن تكون كلها متساوية وعلى ذلك يمكن النظر إلى هذا المقياس (الوسط الحسابي)على أنه متغير عشوائي له توزيع احتمالي و يسمى هذا المجتمع الجديد :مجتمع المتوسطات الحسابية أو توزيع المعاينة.

نفرض أننا أخذنا عينة حجمها n من مجتمع ما، ثم حسبنا منها بعض المقاييس الإحصائية مثل المتوسط الحسابي، نسبة صفة، معينة التباين،... فإن كل مقياس من هذه المقاييس يعتبر متغير عشوائي في ذاته يختلف من عينه إلى أخرى – هذا المتغير العشوائي يخضع لتوزيع معين – هذا التوزيع يسمى بتوزيع المعاينة. فمثلاً نقول أن توزيع المعاينة للمتوسط الحسابي وهو عبارة عن توزيع جميع المتوسطات الحسابية للعينات المأخوذة من نفس هذا المجتمع ذات الحجم n ، وكذلك فإن توزيع المعاينة للتباين هو توزيع جميع التباينات المحسوبة من عينات لها نفس الحجم n ومأخوذة من نفس المجتمع ، وهكذا ..الخ.

أ.توزيع المعاينة للمتوسط

Sampling Distributions of Means, Distribution D'échantillonnage de Moyenne

نفرض أننا سحبنا عينة حجمها n من مجتمع غير محدود، القيمة المتوقعة (المتوسط الحسابي) له تساوي μ والانحراف المعياري هو σ فإن المتوسط الحسابي $\overline{X}$ للعينة يخضع لتوزيع ما، متوسط هذا التوزيع وانحرافه المعياري هما:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \quad ، \quad \mu_{\overline{X}} = \mu$$

ملاحظة: في حالة سحب عينة عشوائية حجمها n من مجتمع محدود حجمه N و كان السحب بدون إرجاع فإن: $\mu_{\overline{X}} = \mu$ و $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

مثال:

افترض أن المجتمع يتكون من 900 عنصر بوسط حسابى 20 وحدة وانحراف معيارى 12 وحدة . الوسط الحسابى والانحراف المعيارى لتوزيع المعاينة لوسط عينة حجمها 36 هما :

$$\mu_{\overline{X}} = \mu = 20 \text{ units}$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = 2$$

لو كانت n تساوى 64 وقمنا بسحب بالإرجاع لاختيار العينة ، فإن

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{12}{\sqrt{64}}\sqrt{\frac{900-64}{900-1}} = \frac{12}{8}\sqrt{\frac{836}{899}} = (1.5)(0.96) = 1.44$$

بدلا من $\sigma X = 1.5$ بدون معامل التصحيح للمجتمعات المحدودة .

نظرية 1:

في الحالة التي يكون فيها المجتمع الأصلي المسحوبة منه العينة مجتمع طبيعي أي يتبع توزيع طبيعي $N(\mu, \sigma^2)$ فإن توزيع المعاينة للمتوسط الحسابي $\overline{X}$ يكون في هذه الحالة توزيع طبيعي أيضاً له نفس المتوسط الأصلي μ ولكن انحرافه المعياري يساوي $\sigma/\sqrt{n}$، أي بمعنى أن :

$$X \sim N(\mu,\sigma^2) \quad \Rightarrow \quad \overline{X} \sim N(\mu,\frac{\sigma^2}{n})$$

ومن ثم يكون بعد التحول إلى قانون التوزيع الطبيعي المعياري:

$$z = \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \sim N(0,1)$$

مثال1:

إذا كان ضغط الدم لمجموعة من الأفراد يتبع توزيعا طبيعيا متوسطه 90 و انحرافه المعياري 9، سحبنا عينة حجمها 16 فردا من هذه المجموعة فما هو احتمال أن يكون متوسط ضغط الدم في العينة أكبر من86 ؟.

الحل:

لدينا $X \sim N(90, 81)$ حيث $\mu = 90$ و $\sigma = 9$ و منه $\mu_{\overline{X}} = 90$ و

$\sigma_{\overline{X}} = \frac{9}{\sqrt{16}} = 2.25$ و منه :

$P(\overline{x} > 86) = P\left(\frac{x-\mu}{\sigma} > \frac{86-90}{2.25}\right) = P(z > -1.77) = P(z \leq 1.77) = 0.9625$

مثال2:

إذا كان بدل السكن المعطى لموظف بإحدى الشركات الكبرى يتبع توزيع طبيعي ، متوسطه :

دينار $\sigma = 170$ و انحرافه المعياري $\partial = 8$ دينار .

- اخترنا موظف عشوائيا ، فما هو احتمال أن يقل بدل سكنه عن 166 دينار .
- سحبت عينة من 64 موظف ، فما هو احتمال أن يكون متوسط بدل سكنهم أكبر من 172 دينار.

الحل:

لدينا $X \sim N(170, 64)$ حيث $\mu = 170$ و $\sigma = 8$ و منه $\mu_{\overline{X}} = 170$ و

$\sigma_{\overline{X}} = \frac{8}{\sqrt{64}} = 1$ و منه:

.1

$$P(X<166)=P(Z<-0.5)=0.5-0.1915=\boxed{0.3085}$$

.2

$$\bar{X}>172 \Rightarrow Z>\frac{172-\mu(\bar{X})}{\sigma(\bar{X})} \Rightarrow Z>\frac{172-\mu}{\sigma/\sqrt{n}} \Rightarrow Z>\frac{172-170}{8/\sqrt{64}} \Rightarrow Z>2$$

$$P(\bar{X}>172)-P(Z>2)-0.5-0.4772-\boxed{0.0228}$$

نظرية 2:

إذا كان المجتمع غير طبيعي فإن $\bar{X}$ لا يخضع للتوزيع الطبيعي ولكنه يتوزع توزيع يكون قريباً من التوزيع الطبيعي لقيم n الكبيرة $(n \geq 30)$ حيث أن:

$$z=\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \xrightarrow[n\to\infty]{as} N(0,1)$$

وتعتبر النتيجة السابقة الهامة جداً في الإحصاء وخاصة في التطبيقات العلمية وتسمى نظرية النهاية المركزية Central Limit Theorem والتي تنص على أنه في حالة العينات الكبيرة الحجم فإن المتوسط الحسابي $\bar{X}$ يخضع للتوزيع الطبيعي بالمعاملات μ و $\frac{\sigma^2}{n}$ ، حيث أن μ, σ^2 هما متوسط وتباين المجتمع الأصلي بغض النظر عن شكل توزيع المجتمع الأصلي. ومن ثم فإن قيم n الكبيرة تحقق العلاقة (3-4) بصرف النظر عن توزيع المجتمع الأصلي.

مثال:

إذا كانت الحوادث الأسبوعية على إحدى الطرق تتبع توزيع بواسون بمتوسط (4) حوادث،أخذت عينة من (64) أسبوعاء، فما هو احتمال أن يكون متوسط الحوادث فيها يزيد عن 4.2 حادثة؟

الحل:

من خصائص توزيع بواسون $\mu = \lambda = 4$

$\sigma^2 = \lambda = 4$

x يتبع توزيع طبيعي $x \sim N(\mu_{\bar{x}} = \mu \,,\ \sigma^2_{\bar{x}} = \frac{\sigma^2}{n})$

$\mu_{\bar{x}} = \mu = 4$

و انحرافه المعياري

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{64}} = 0.25$$

$$p(\bar{x} > 4.2) = P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} > \frac{4.2-4}{0.25}\right)$$

$$= p\,(z > 0.8)$$

$$= 1 - p\,(z < 0.8) = 0.2119$$

نظرية 3:

كذلك فإنه إذا كان $\bar{X}_1$ هو المتوسط الحسابي لعينه عشوائية مسحوبة من مجتمع غير محدود متوسطه هو μ_1 وانحرافه المعياري هو σ_1، وكان $\bar{X}_2$ هو المتوسط الحسابي لعينة عشوائية مسحوبة من مجتمع لانهائي آخر متوسط μ_2 وانحرافه المعياري σ_2 وكانت العينتين مستقلتين فإن المجموع /الفرق الجبري لمتوسط العينتين يخضع لتوزيع المعاينة بالمعاملات:

$$\mu_{(\bar{x}\pm\bar{x}_2)} = \mu_1 \pm \mu_2 \quad \text{and} \quad \sigma^2_{(\bar{x}_1\pm\bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

حيث n_1, n_2 هما حجم العينة الأولى والثانية.

وإذا كان المجتمعين الأصليين طبيعيين فإن $(\bar{x}_1 \pm \bar{x}_2)$ يخضع لتوزيع طبيعي أيضاً بالمعلمات المعطاة سابقا و عليه فإنه في هذه الحالة:

$$z = \frac{(\bar{x}_1 \pm \bar{x}_2) - (\mu_1 \pm \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

أما إذا كان أحد المجتمعين أو كليهما لا يتوزع توزيعاً طبيعياً فإن $(\bar{x}_1 \pm \bar{x}_2)$ لا يتوزع توزيعاً طبيعياً كذلك ، ولكن لقيم n_1, n_2 الكبيرة $(n \geq 30)$ طبقاً لنظرية النهاية المركزية السابقة فإن $(\bar{x}_1 \pm \bar{x}_2)$ تتوزع توزيعاً قريباً من التوزيع الطبيعي وبذلك يمكننا استخدام نفس العلاقة (4-6)في حالة العينات الكبيرة.

مثال:

ليكن المجتمع U_1 : 3، 7، 8. والمجتمع U_2 : 2، 4. تحقق من أن :

$\mu_{U1-U2} = \mu_{U1} - \mu_{U2}$; $\sigma^2_{U1-U2} = \sigma^2_{U1} + \sigma^2_{U2}$.

الحل:

U1			U1 – U2	
8	7	3		
6	5	1	2	U2
4	3	-1	4	

$\mu U1 = (3 + 7 + 8)/3 = 6$; $\mu U2 = (2 + 4)/2 = 3 \Rightarrow$

$\mu U1 - \mu U2 = 6 - 3 = 3$

$\mu U1 - U2 = (1 + 5 + 6 - 1 + 3 + 4)/6 = 3$

$\sigma^2 U1 = (3^2 + 7^2 + 8^2)/3 - 6^2 = 14/3$;

$\sigma^2 U2 = (2^2 + 4^2)/2 - 3^2 = 1 \Rightarrow \sigma^2 U1 + \sigma^2 U2 = 17/3$

$\sigma^2 U1 - U2 = (1^2 + 5^2 + 6^2 + 1^2 + 3^2 + 4^2) / 6 - 3^2 = (1 + 25 + 36 + 1 + 9 + 16) / 6 - 9 = 17/3$

ب. توزيع المعاينة للنسبة

نفرض أن لدينا مجتمعا ما و أن بعض مفردات هذا المجتمع تتوفر فيها صفة معينة هي p ويمكن أن تكون p مثلا هي نسبة الأفراد الأميين في إحدى المدن أو نسبة الإنتاج التالف من مجموع إنتاج شركة ما، فإذا اخترنا عينة عشوائية حجمها n من هذا المجتمع ووجدنا أن نسبة الصفة في العينة هي r .

و إذا اخترنا عينات أخرى حجم كل منها n فإن هذه النسبة تتغير من عينة لأخرى وعلى ذلك فإنها تكون متغير عشوائيا له توزيع احتمالي تحدده النظرية الآتية:

نظرية4 :

إذا كانت P هي نسبة وجود ظاهرة معينة في مجتمع ما واختيرت من هذا المجتمع عينات كبيرة حجم كل منها n وكانت r تمثل نسبة وجود هذه الظاهرة في العينات فإنr تتبع توزيعا طبيعيا وسطه P و تباينه: $\frac{p(1-p)}{n}$

و نكتب :

$$r \approx N(\mu_r = P, \sigma_r^2 = \frac{P(1-P)}{n})$$

حيث :

$$\mu(r) = \mathrm{P}, \quad \sigma(r) = \sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}}$$

ملاحظة: في حالة عدم إعطاء النسبة r يتم حسابها كما يلي: $r = x/n$ حيث x عدد المفردات التي بها الصفة في العينة.

مثال 1:

إذا علمت أن نسبة البيض التالف الذي ينتجه أحد مراكز إنتاج الدواجن هي 0.03 اشترى شخص 400 بيضة من إنتاج هذا المركز ، ما هو احتمال أن يجد من بينها 20 بيضة على الأقل تالفة؟

الحل:

n =400

$$r = \frac{x}{n} = \frac{20}{400} = .05$$

P= 0.03 نسبة البيض التالف الذي ينتجه أحد مراكز انتاج الدواجن.

وحيث أن r تتبع توزيع طبيعي فإن

$$r \approx N(\mu_r = P, \sigma_r^{\ 2} = \frac{P(1-P)}{n})$$

$$\sigma_r = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{0.03 \times 0.97}{400}} = 0.0085$$

$$P(r \geq 0.05) = P(\frac{r-\mu_r}{\sigma_r} \geq \frac{0.05-0.03}{0.0085}) = P(Z \geq 2.35) = 1 - P(Z \prec 2.35) =$$
$$= 1 - 0.9906 = 0.0094$$

مثال 2:

إذا كان نسبة المصابيح الكهربائية التالفة التي ينتجها أحد المصانع هي 3%، اشترى شخص **400** مصباح كهربائي من هذا المصنع، ما احتمال أن يجد من بينها **20** مصباح على الأقل تالفة؟

الحل:

$$\mathrm{P} = 0.03, \quad n = 400$$
$$\Rightarrow \mu(r) = \mathrm{P} = 0.03$$
$$\Rightarrow \sigma(r) = \sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}} = \sqrt{\frac{0.03 \times 0.97}{400}} \approx 0.01$$
$$X > 20 \overset{r=\frac{X}{n}}{\Rightarrow} r > \frac{20}{400} \Rightarrow r > 0.05 \Rightarrow Z > \frac{0.05-0.03}{0.01} \Rightarrow Z > 2$$
$$\Rightarrow P(r > 0.05) = P(Z > 2) = 0.5 - 0.4772 = \boxed{0.0228}$$

مثال 3:

إذا كان نسبة الموظفين الذين حصلوا على زيادة في الراتب بإحدى الشركات هي **91%**،

اخترنا **1000** موظف عشوائياً ، ما احتمال أن يجد من بينه م **70** على الأكثر لم يحصلوا على زيادة في الراتب؟

الحل:

$$\mathrm{P} = 1 - 0.91 = 0.09, \quad n = 1000$$
$$\Rightarrow \mu(r) = \mathrm{P} = 0.09$$
$$\Rightarrow \sigma(r) = \sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}} = \sqrt{\frac{0.09 \times 0.91}{400}} \approx 0.01$$
$$X < 70 \overset{r=\frac{X}{n}}{\Rightarrow} r < \frac{70}{1000} \Rightarrow r < 0.07 \Rightarrow Z < \frac{0.07-0.09}{0.01} \Rightarrow Z < -2$$
$$P(r < 0.07) = P(Z < -2) = 0.5 - 0.4772 = \boxed{0.0228}$$

ج.توزيع المعاينة للتباين

إذا كان $s^2 = \frac{\sum (x_i - \overline{x})^2}{n}$ هو تباين عينه عشوائية حجمها n مأخوذة من مجتمع متوسطه μ وتباينه σ^2 فإنه يمكن تمييز الحالات التالية:

- **متوسط تباين العينة:**

في مجتمع ما غير محدود و عينة حجمها n فإن متوسط تباين العينة يعطى بالعلاقة التالية:

$$\mu_{s^2} = \left(\frac{n-1}{n}\right)\sigma^2$$

$$\mu_{S^2} = \sigma^2\left(\frac{n-1}{n}\right)\left(\frac{N}{N-1}\right)$$

ملاحظة: في حالة سحب عينة عشوائية حجمها n من مجتمع محدود حجمه N و كان السحب بدون إرجاع فإن:

في حالة ما إذا كان حجم العينة $(n \geq 30)$ فإن متوسط تباين العينة:

$E(S^2) \approx \sigma^2$

الانحراف المعياري لتباين العينة:

في مجتمع ما غير محدود و عينة حجمها n بحيث قيم n كبيرة $(n \geq 30)$. فإن متوسط تباين العينة و الانحراف المعياري لتباين العينة يعطيان بالعلاقة التالية:

$$\mu_{S^2} = \sigma^2 \quad \text{and} \quad \sigma_{S^2} = \sqrt{\frac{\mu_4 - \sigma^4}{n}}$$

حيث μ_4 العزم الرابع لتباين العينة حول المتوسط.

و بالتالي يكون تباين توزيع المعاينة للتباين:

$$\sigma^2{}_{S^2} = \frac{\mu_4 - \sigma^4}{n}$$

وإذا كان المجتمع طبيعي فإن $\mu_4 = 3\sigma^4$ وبالتالي فإن : $\sigma^2_{s^2} = \left(\frac{2}{n}\right)\sigma^4$ أي:

$$\sigma_{S^2} = \sigma^2\sqrt{2/n}$$

ملاحظة:

نلاحظ هنا أن S^2 لا تتوزع طبيعيا حتى ولو كان المجتمع طبيعي، ولكنه يتوزع توزيع قريب من التوزيع الطبيعي وذلك لقيم n الكبيرة $(n \geq 100)$.

أما إن كان المجتمع الأصلي يخضع للتوزيع الطبيعي فإن المتغير $\frac{(n-1)s^2}{\sigma^2}$ يخضع لتوزيع يسمى توزيع كي مربع χ^2 بعدد درجات حرية يساوي $n-1$. أي أن:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

ويعتبر توزيع كي مربع من التوزيعات الهامة في الإحصاء التطبيقي ودالة كثافته هي:

$$\boxed{f(x) = \begin{cases} \dfrac{x^{(v/2)-1} e^{-x/2}}{2^{v/2}\Gamma(v/2)} & x > 0 \\ 0 & x \leq 0 \end{cases}}$$

وتكون القيمة المتوقعة لهذا التوزيع هي v و تباينه هو $2v$ أي بمعنى أن:

$E(y)=\mu_y=v$ و $V(y)=\sigma^2=2v$

حيث $\Gamma(\alpha)$ هي الدالة قاما و هي تتميز بجملة من الخصائص الرياضية أهمها:

- $\forall \alpha \in N : \Gamma(\alpha) = (\alpha-1)!$ و $\Gamma(1/2) = \sqrt{\pi}$

-

و بالعودة إلى توزيع المعاينة و وضع: $W = \frac{(n-1)S^2}{\sigma^2}$

نقول أن: $W \sim \chi^2_{n-1}$ حيث:

$$\mathrm{V}(W) = 2(n-1) \quad \text{و} \quad \mathrm{E}(W) = n-1$$

و يكون لدينا:

- $P(S^2 \leq b) = P\left(\frac{(n-1)S^2}{\sigma^2} \leq \frac{(n-1)b}{\sigma^2}\right) = P\left(W \leq \frac{(n-1)b}{\sigma^2}\right)$
- $E(S^2) = E\left(\frac{\sigma^2}{n-1}W\right) = \frac{\sigma^2}{n-1}E(W) = \sigma^2$
- $V(S^2) = V\left(\frac{\sigma^2}{n-1}W\right) = \frac{\sigma^4}{(n-1)^2}V(W) = \frac{2\sigma^4}{n-1}$

سلسلة تمارين حول الفصل الأول

التمرين الأول:

إذا كان لدينا مجتمع يتكون من المفردات التالية: $x_1 = 1$, $x_2 = 2$, $x_3 = 3$

1. أوجد متوسط وتباين المجتمع.
2. اوجد توزيع المعاينة للوسط الحسابي للعينات ذات الحجم n=2 التي يمكن سحبها من هذا المجتمع في كل من الحالات التالية:
   أ-إذا كان السحب بإرجاع. ب-إذا كان السحب بدون إرجاع.
3. أوجد المتوسط الحسابي و الانحراف المعياري لتوزيع المعاينة في الحالات التالية:
   أ-إذا كان السحب بإرجاع. ب-إذا كان السحب بدون إرجاع.

التمرين الثاني:

من مجتمع مكون من الأعداد الخمسة الآتية: 1، 3، 5، 7، 9. أوجد:

(أ) متوسط المجتمع μ و انحرافه المعياريσ. (ب) توزيع المعاينة للمتوسط لعينة حجمها 2 مع السحب بإرجاع. (ج) متوسط توزيع المعاينة $\mu_{\bar{x}}$ و تباينه $\sigma^2_{\bar{x}}$.

التمرين الثالث:

إذا كان لدينا متغير عشوائي X يتبع التوزيع الطبيعي بمتوسط $\mu = 80$ و تباين $\sigma^2 = 49$

أوجد توزيع المعاينة لمتوسط العينة من الحجم 25 مسحوبة من المجتمع. ثم أوجد $P(\bar{X} \leq 83)$, $P(\bar{X} \geq 78)$.

التمرين الرابع:

إذا كان علامات طلبة السنة الثانية لمقياس الإحصاء3 تتبع توزيع طبيعي بمتوسط $\mu = 100$ وأنحرف معياري $\sigma = 75$. تم اختيار25 طالباً عشوائيا من بين طلبة الجامعة. أوجد:

1. احتمال أن يكون متوسط العلامات المحسوب من العينة أكبر من 125.

2. احتمال ان يكون متوسط العلامات المحسوب من العينة اصغر من 80
3. احتمال أن يكون متوسط العلامات المحسوب من العينة محصورا بين 70 و 130.

التمرين الخامس:

لدى بنك محلى صغير 1450 حساب ادخار شخصي يتبع توزيع بواسون برصيد متوسط قدرة 1600$

إذا أخذ البنك عينة عشوائية من 100 حساب، أوجد توزيع المعاينة للمتوسط و ما هو احتمال أن متوسط المدخرات لهذه الحسابات المائة سيكون أقل من 1600$ ؟.

التمرين السادس:

ليكن لدينا متغيران مستقلان X_1 و X_2 يمثلان مجتمعين 1 و 2 حيث:

$X_1 \sim N(15, 32)$ و $X_2 \sim N(20, 50)$

-ما هو توزيع المعاينة للفرق بين متوسطي عينتين الأولى من المجتمع1 حجمها 16 و الثانية من مجتمع 2 حجمها25 ؟.

-ما هو احتمال أن لا يزيد الفرق بين المتوسطين عن 2؟.

التمرين السابع:

إذا كان لدينا البيانات التالية الخاصة بمجتمعين 1 و 2 ، حيث :

المجتمع 1: متوسطه $\mu_1 = 40.7$ و تباينه $\sigma_1^2 = 3.2$. المجتمع 2: متوسطه $\mu_2 = 39.2$ و تباينه $\sigma_2^2 = 2.9$

وتم سحب عينتين مستقلتين من المجتمعيين حيث $n_1 = 17$ و $n_2 = 15$ ، أوجد الاحتمال التالي بفرض أن التوزيعات طبيعية: $P(0.289 \le \bar{X}_1 - \bar{X}_2 \le 2.711)$

التمرين الثامن:

إذا كان لدينا البيانات التالية الخاصة بمجتمعين 1 و 2 ، حيث :

المجتمع 1: متوسطه $\mu_1 = 8.34$ و تباينه $\sigma_1^2 = 2.8$. المجتمع 2: متوسطه $\mu_2 = 6.45$ و تباينه $\sigma_2^2 = 2.25$.

إذا وتم سحب عينتين مستقلتين من المجتمعين ، حيث $n_1 = 50$ و $n_2 = 40$ ، فإذا كانت المتغيرات قيد الدراسة لا تتبع التوزيع الطبيعي ، أوجد قيمة تقريبية للاحتمال التالي:

$P\left(1.19 \leq \overline{X}_1 - \overline{X}_2 \leq 2.59\right)$

<u>التمرين التاسع:</u>

مجتمع مكون من 5 أطفال فقيس إذا كان الطفل يحب عصير الليمون أم لا فكانت النتائج التالية:

x_1= yes, x_2= no, x_3= yes, x_4= no, x_5= no

- أوجد نسبة الأطفال في المجتمع الذين لا يحبون عصير الليمون؟
- أوجد متوسط وتباين توزيع المعاينة للنسبة r في عينة حجمها 2 ؟

<u>التمرين العاشر</u>:

مجتمع مكون من 70 شخص من بينهم 21 شخص مدخن، أوجد:

1. توزيع المعاينة لنسبة المدخنين في عينة عشوائية حجمها 10 مسحوبة من هذا المجتمع.
2. إذا سحبنا عينة عشوائية حجمها 10 من هذا المجتمع، ما هو احتمال أن تزيد نسبة المدخنين في العينة عن.0.44 .

<u>التمرين الحادي عشر</u>:

مجتمع مكون من 7 أرقام متوسطها 40 و انحرافها المعياري 3، سحبت عينة من المجتمع حجمها 5.

أوجد متوسط توزيع المعاينة للتباين إذا كان السحب بإرجاع و بدون إرجاع.

<u>التمرين الثاني عشر</u>:

مزرعة حجمها 8000 شجرة. إذا كان طول الأشجار يتبع توزيع طبيعي بمتوسط ارتفاع الشجرة 5.6 متر والانحراف المعياري 0.9 متر. أخذت عينة عشوائية حجمها 50 شجرة من المجتمع .

أوجد متوسط توزيع المعاينة للتباين و الانحراف المعياري.

الفصل الثاني: التقدير

تمهيد

الهدف الأساسي من دراسة أي مجتمع هو إيجاد أو تقدير كاستدلال لبعض خصائصه أو معامله ، مثلما رأينا في الفصل السابق مثل متوسط المجتمع و انحرافه المعياري أو نسبة صفة معينة في المجتمع، و هذه تعتبر أهم معالم المجتمع. و هذه المعالم غالبا ما تكون مجهولة و نرغب في تقديرها. و حيث أن العينة تعتبر صورة مصغرة عن المجتمع، فإننا نلجأ إلى حساب ما يقابل معالم المجتمع في العينة من مقاييس من بيانات العينة و التي نسميها الإحصائيات ، و يمكن استخدام قيمة الإحصائية كتقدير للمعلمة المناظرة له في المجتمع.

1. التقدير Estimate و المقدر Estimator

المقصود بالتقدير هو تقدير معالم المجتمع الإحصائي (أو التوزيع الاحتمالي) والتي غالباً ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها -عادة– من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة، أو تقدير متوسط عمر الناخب... وغيرها. و منه نقول أن:

التقدير Estimate ، هو القيمة العددية للمقدر .

المقدرEstimator، هو إحصاء أو معادلة تستخدم للحصول على تقدير لمعلمة المجتمع.

وهناك نوعان (أو أسلوبان) للتقدير يسمى الأول تقدير النقطة (أو القيمة الواحدة)، ويسمى الثاني تقدير الفترة (أو فترة التقدير أو الثقة).

2.التقدير النقطي (التقدير بالنقطة)

في حالة تقدير النقطة نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. وكمثال آخر لو أخذنا نسبة الناخبين في العينة الذين يؤيدون مرشحا معيناً كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين.

3. التقدير بمجال

في التقدير بمجال نحصل على مجال intervalle يتحدد بحدين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن مجال التقدير يحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

مثال: إذا قدرنا متوسط دخل الفرد سنويا في دولة ما ب 10000 دولار، نكون قد قدرنا دخل الفرد تقديرا نقطيا. يكون تقديرنا بمجال إذا قلنا مثلا أن الدخل يساوي 10000 ± 2000 أي أنه يتراوح بين 8000 و 12000 دولار.

4.خصائص المقدر الجيد :

يوجد ثلاث خصائص يجب أن تتوفر في المقدر حتى يوصف بالمقدر الجيد وهي:

خاصية عدم التحيز: يقال عن التابع الإحصائي انه مقدر غير متحيز للثابت الإحصائي (لمعلمة المجتمع)

إذا كانت القيمة المتوقعة للتابع الإحصائي تساوي الثابت الإحصائي.

خاصية الكفاءة :يقال عن المقدر انه أكثر كفاءة إذا كان له تباين أقل بمعنى آخر أن المقدر الأقل

تباينا هو الأكثر كفاءة لأنه كلما قل التباين زادت فعالية التقدير.

خاصية التقارب :يقال عن الإحصائية أنها مقدر متقارب للمعلمة كلما زاد حجم العينة واقترب من

الملا نهاية.

5.مجال الثقة لمتوسط المجتمع:

أ. حالة مجتمع طبيعي معلوم التباين و/أو عينة كبيرة:

إذا كان المجتمع المسحوب منه العينة ذا توزيعاً طبيعياً وتباينه معروفاً أو كانت العينة كبيرة (أي حجمها ثلاثون مفردة أو أكثر، حتى و إن لم يكن التوزيع طبيعي) فإن توزيع المعاينة للوسط الحسابي يكون توزيعاً طبيعياً، وطالما نحن نتكلم عن تقدير متوسط المجتمع فإن أول ما نفكر فيه هو الوسط الحسابي للعينة.

ومجال التقدير (أو الثقة) للمتوسط الحسابي للمجتمع يأخذ الشكل التالي:

$$\left(\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

تحديد مجال الثقة للمعلمة يرفق بتحديد احتمال تحقق، أي باحتمال أن تنتمي المعلمة إلى المجال المذكور ، يرمز لهذا الاحتمال بـ ($1-\alpha$) و يسمى درجة التأكد أو **مستوى الثقة** ، الاحتمال المعاكس α يسمى **مستوى أو درجة المعنوية .**

مثال : فإذا كان متوسط أعمار الناخبين يتراوح ما بين 46 و 34 سنة، ودرجة الثقة هي 95 % فإن هذا معناه أنه لو تكررت التجربة مائة مرة، فإن التقدير سيكون محصوراً بين هذين الرقمين في 95 من الحالات (أي احتمال أن يكون صحيحاً هو 95% بمعنى أن احتمال أن تكون فترة التقدير صحيحة هو 0.95 وهكذا...).

يمكن كتابة أشهر وأهم درجات ومعاملات الثقة (للتوزيع الطبيعي) في الجدول التالي (مع ملاحظة أن 95%، 99% هي أشهرها على الإطلاق):

معامل الثقة Z	درجة الثقة
1	68.26%
1.65	90%
1.96	95 %
2	95.44%
2.58	99%
3	99.72%

مثال:

لو سحبت عينة عشوائية من مجموع مجتمع الناخبين في دولة ما حجمها 100 ناخب فإذا كان الوسط الحسابي للدخل اليومي للناخبين بالعينة هو 90 دولار والانحراف المعياري لدخل الناخبين في المجتمع 25 دولار، أوجد فترة تقدير للمتوسط الحسابي للدخل اليومي لمجموع الناخبين في هذه الدولة بدرجة ثقة 95% ؟

الحل : بما أن فترة تقدير الوسط الحسابي للمجتمع هي :

$$\hat{\mu} = \overline{X} \pm Z \frac{\delta}{\sqrt{n}}$$

والمعلومات المعطاة هي :

حجم العينة n = 100 الوسط الحسابي للعينة $\overline{X} = 90$ والانحراف المعياري للعينة δ = 25

وحيث أن درجة الثقة هي % 95 فإن : Z = 1.96 حسب ما هو موضح في الجدول السابق.

وبالتالي فإن فترة تقدير الوسط الحسابي للدخل اليومي لمجتمع الناخبين بدرجة ثقة % 95 هي :

$$\hat{\mu} = 90 \pm 1.96 \frac{25}{\sqrt{100}} = 90 \pm 1.96(2.5) \quad = 90 \pm 4.9$$

$$\mu = \begin{cases} 85.1 \\ 94.9 \end{cases}$$

أي أن الوسط الحسابي للدخل اليومي لمجتمع الناخبين يتراوح بين 85.1 دولاراً كحد أدنى، 94.9 كحد أعلى، وذلك بدرجة ثقة % 95.

ب. حالة مجتمع طبيعي أو مجتمع ما مجهول التباين و عينة كبيرة:

إذا كان المجتمع طبيعي أو كانت العينة كبيرة حتى و إن لم يكن المجتمع طبيعي و كان تباين المجتمع مجهول فإننا نستعمل الانحراف المعياري للعينة بدلا من الانحراف المعياري للمجتمع لتقدير مجال اليقة للمتوسط كما يلي:

مثال:

مصنع لإنتاج المصابيح الكهربائية، اختير من إنتاجه عينة حجمها 100 مصباح، فإذا كان المتوسط الحسابي لعمر المصباح في العينة 1200 ساعة و انحرافه المعياري 250 ساعة .

قدر عند مستوى ثقة 95 % متوسط عمر المصباح من إنتاج المصنع ككل.

الحل: حيث أن الانحراف المعياري للمجتمع σ مجهول سنستخدم الانحراف المعياري للعينة.

$$n=100,\quad \bar{x}=1200,\quad s_x=250,\quad 1-\alpha=0.95 \Rightarrow Z_{\alpha/2}=1.96$$

$$P\left[\bar{x}-Z_{\alpha/2}\frac{s_x}{\sqrt{n}} \le \mu \le \bar{x}+Z_{\alpha/2}\frac{s_x}{\sqrt{n}}\right]=1-\alpha$$

$$P\left[1200-(1.96)\frac{250}{\sqrt{100}} \le \mu \le 1200+(1.96)\frac{250}{\sqrt{100}}\right]=0.95$$

$$P\left[1200-49 \le \mu \le 1200+49\right]=0.95$$

$$P\left[1151 \le \mu \le 1249\right]=0.95$$

$\therefore$ متوسط عمر المصباح من إنتاج المصنع كله يتراوح بين 1151 و1249 ساعة بدرجة ثقة 95%.

ج.حالة مجتمع طبيعي مجهول التباين و عينة صغيرة:

إذا كان المجتمع المسحوب منه العينة ذا توزيعاً طبيعياً وتباينه مجهولا و كانت العينة صغيرة (أي حجمها أقل من ثلاثون مفردة) فإن التوزيع الإحصائي المتبع في مثل هذه الحالات هو ما يطلق عليه " توزيع t " فعند تقدير متوسط عمر الناخب في مدينة ما عن طريق سحب عينة صغيرة (حجمها أقل من 30 ناخب) التوزيع الطبيعي يكون في مثل هذه الحالات غير مناسب لصغر حجم العينة أولاً، ثم عدم معرفة الانحراف المعياري لعمر الناخب ثانياً. لذا فإن الأسلوب الإحصائي المتبع في حالات كهذه هو استخدام " توزيع t " والذي يسميه البعض "توزيع العينات الصغيرة".

و لعل الاختلاف الأساسي بين توزيع t والتوزيع الطبيعي هو أن الانحراف المعياري للعينة هو المستخدم في الأول بدلا من الانحراف المعياري للمجتمع في الثاني.

ويمكن تحديد الشروط الثلاثة لاستخدام توزيع t كما يلي:

1- أن يكون المجتمع المسحوبة منه العينة له توزيع طبيعي.

2- والانحراف المعياري للمجتمع σ غير معروف (أو مجهول).

3- والعينة صغيرة (حجمها أقل من 30 مفردة).

و منه يمكن استخدام توزيع ستيودنت لتقدير μ بمجال فنكتب حدود مجال الثقة كما يلي:

$$\mu \in \left[m - t_{1-\alpha/2;n-1} \frac{S}{\sqrt{n-1}} \; ; \; m + t_{1-\alpha/2;n-1} \frac{S}{\sqrt{n-1}} \right]$$

أو:

$$\hat{\mu} = \overline{X} \pm t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}$$

ولعل من أهم الملاحظات على المعادلة الإحصائية السابقة احتواؤها على مفهومين مهمين هما:

1) مستوى المعنوية أو الدلالة Level of Significance والذي رمزنا له بالرمز اللاتيني ألفا α والذي يعني أنه المكمل لدرجة الثقة أي نسبة الخطأ. وبالتالي فإذا كانت درجة الثقة 99 % أي احتمال أن يكون التقدير صحيحاً بنسبة 99 %، فإن مستوى المعنوية، والذي يعني هنا درجة احتمال الخطأ يساوي 1%. وعند الكشف في جدول(t)، ولأنه توزيع متماثل، فإنه يتم قسمة مستوى المعنوية على 2.

2) درجات الحرية، ويساوي في هذه الحالة n - 1. حيث n هو حجم العينة وطرحنا منه 1 لأنه تم تقدير الانحراف المعياري للمجتمع المجهول باستخدام الانحراف المعياري للعينة S.

ملاحظة:

تعرف درجات الحرية بأنها عدد المشاهدات المستقلة في العينة والتي تساوي حجم العينة مطروحاً منه عدد القيود أو معالم المجتمع التي يتم تقديرها من بيانات العينة. وكمثال مبسط لشرح فكرة درجات الحرية نفترض أن لدينا 3 قيم واشترطنا أن مجموع القيم يساوي 10 فإن لدى الباحث في هذه الحالة حرية في اختيار الرقم الأول (وليكن 2) والثاني (وليكن 3) لذلك فإن قيمة الثالثة لابد وأن تكون (5) بالتالي نستطيع القول بأن درجة الحرية المتاحة لدى الباحث هي (2) أي 1 - 3 = 2 أي أن درجات الحرية في هذه الحالة هي : n - 1

حيث n تساوي حجم العينة (والتي تساوي في المثال السابق 3) والرقم (1) والذي طرحناه يعني الشرط الذي يحتم أن مجموع القيم =10 وبصفة عامة إذا كان عدد القيود k فإن درجات الحرية تساوي n - k

مثال1:

إذا كانت دخول مجموعة من الأفراد في دولة ما تتبع التوزيع الطبيعي، وسحبت منهم عينة عشوائية حجمها 10 أفراد بوسط حسابي دولاراً $\overline{X} = 72$ وانحراف معياري بلغ دولاراً =S 6.4 أنشئ فترة تقدير للوسط الحسابي للدخل اليومي لجميع الأفراد بدرجة ثقة % 95

الحل :

نلاحظ أولاً: أن العينة صغيرة (حجمها 10 أفراد فقط) وأن المجتمع طبيعي وانحرافه المعياري غير معروف لذلك نستخدم فترة تقدير الوسط للعينات الصغيرة التي تعتمد على توزيع t

وحيث أن n = 10 فإن درجات الحرية لها هي:

$$n - 1 = 10 - 1 = 9$$

وحيث أن درجة الثقة المطلوبة هي $1-\alpha = 0.95$ فإن مستوى المعنوية $\alpha = 0.05$ وبالتالي فإن نصف مستوى المعنوية هو :

$$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$$

أي يتم الكشف في جدول (II) عند درجات حرية تساوي 9 تحت احتمال (نصف مستوى المعنوية) 0.025 أي أن :

$$t_{0.025,9} = 2.262$$

وبالتعويض عن فترة تقدير الوسط نحصل على المعادلة التالية :

$$\hat{\mu} = \overline{X} \pm t\frac{S}{\sqrt{n}} \qquad \hat{\mu} = 72 \pm 2.262\frac{6.4}{\sqrt{10}} \qquad \hat{\mu} = 72 \pm \frac{14.48}{3.16}$$

$$\hat{\mu} = 72 \pm 4.6$$

$$\therefore \hat{\mu} = \begin{cases} 67.4 \\ 76.6 \end{cases}$$

أي أن الوسط الحسابي للدخول اليومية يتراوح بين 67.4 دولاراً كحد أدنى. 76.6 دولاراً كحد أعلى وذلك بدرجة ثقة % 95.

مثال2. نريد تقدير متوسط مجتمع طبيعي، بمستوى ثقة 0.95، انطلاقا من عينة حجمها 10 متوسطها 15 و انحرافها المعياري 27.

لدينا $\frac{(m-\mu)}{S/\sqrt{n-1}} \sim t_{n-1}$ إذن:

$$\mu \in \left[m - t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n-1}} \; ; \; m - t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n-1}} \right] \Rightarrow$$

$$\mu \in \left[15 - t_{0.975.10-1} \frac{27}{\sqrt{10-1}} \; ; \; 15 + t_{0.975.10-1} \frac{27}{\sqrt{10-1}} \right] \rightarrow$$

$$\mu \in [15 - 2{,}262(3) \; ; \; 15 + 2{,}262(3)] \Rightarrow \mu \in [8.214 \; ; \; 21.687]$$

6.مجال الثقة للنسبة في المجتمع:

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر السياسية، وبالذات الوصفية منها كقياس اتجاهات الرأي العام، وقياس نسبة قتلى الحروب، ونسبة الدول التي أوفت بالتزاماتها في المنظمات الدولية أو الإقليمية... وغيرها ونظراً لأنه من الصعوبة بمكان في كثير من الأحيان حساب هذه النسبة مباشرة من المجتمع، فإننا غالبا ما نلجأ لتقدير هذه النسبة من عينة عشوائية مسحوبة من هذا المجتمع.

فلو افترضنا أن نسبة المؤيدين للسياسة الاقتصادية التي تنتهجها دولة ما هي P وأن العينة العشوائية كبيرة بدرجة كافية وأن نسبة مؤيدي هذه السياسة في العينة هي r فإن خطوات تقدير النسبة في المجتمع تكون كما يلي :

$$P\left[r - Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}} \le \mathrm{P} \le r + Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right] = 1-\alpha$$

مثال **(1)**: أخذت عينة من مصنع لإنتاج قطع غيار السيارات، اختير من إنتاجه عينة حجمها **(500)** قطعة، وجد من بينها **100** قطعة معيبة، أوجد بدرجة ثقة **95%** نسبة القطع المعيبة في إنتاج المصنع كله.

الحل:

$$r = \frac{100}{500} = 0.2$$

$$P\left[r - Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}} \le \mathrm{P} \le r + Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right] = 1-\alpha$$

$$P\left[0.2-(1.96)\sqrt{\frac{0.2\times 0.8}{500}} \le P \le 0.2+(1.96)\sqrt{\frac{0.2\times 0.8}{500}}\right] = 0.95$$

$$P\left[0.2-0.04 \le \mathrm{P} \le 0.2+0.04\right] = 0.95$$

$$P\left[0.16 \le \mathrm{P} \le 0.24\right] = 0.95$$

أي أن نسبة قطع الغيار المعيبة في إنتاج المصنع كله تتراوح بين 16% و24% وذلك بدرجة ثقة 95%.

مثال (2) :

عينة عشوائية حجمها 144 ناخباً سحبت من إحدى المدن فوجد أن عدد المؤيدين في العينة لمرشح معين هو 60 ناخباً، أنشئ فترة تقدير لنسبة المؤيدين لهذا المرشح في المدينة كلها بدرجة ثقة % 95.

الحل :

1- نحسب أولاً نسبة المؤيدين للمرشح في العينة r التي نحصل عليها بقسمة عدد المؤيدين له على العدد الكلي للعينة (حجم العينة) أي أن :

$$r = \frac{60}{144} = 0.42$$

وحيث أن درجة الثقة المطلوبة هي % 95 فإن معامل الثقة المناسب هو: Z = 196 وفترة تقدير نسبة المؤيدين لهذا المرشح في المدينة تأخذ الشكل التالي:

وبالتعويض عن حجم العينة 144 = n والنسبة في العينة

$1 - r = 1 - 0.42 = 0.58$، $r = 0.42$

ومعامل الثقة Z = 1.96 ، نحصل على:

$$P = 0.42 \pm 1.96\sqrt{\frac{0.42 \times 0.58}{144}}$$

$$= 0.42 \pm 0.08$$

$$\therefore \mathrm{P}\begin{cases} 0.34 \\ 0.50 \end{cases}$$

اي ان نسبة المؤيدين للمرشح في المدينة تتراوح بين 0.34 , 0.50 وذلك بدرجة ثقة 95 % بمعنى آخر أن نسبة مؤيدي هذا المرشح في هذه المدينة لا تتجاوز % 50 كحد أعلى، وبالتالي ففرصته في الفوز كمرشح قد لا تكون كبيرة وذلك بدرجة ثقة % 95 بمعنى أن هذا الحكم لا تتجاوز نسبة الخطأ فيه % 5.

و فيما يلي جدول يلخص أهم توزيعات المعاينة و طرق التقدير التي تم تناولها خلال هذا الفصل:

فترة الثقة للمعالم	حجم العينة	نوع التوزيع الذي تتبعه فترة الثقة	معامل الثقة	فترة الثقة الحقيقة
متوسط مجتمع معلوم التباين أو مجتمع طبيعي معلوم التباين	$\mathbf{n \geq 30}$	يتبع توزيع الطبيعي $Z_{\frac{\alpha}{2}}$	$Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$	$(\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$
	$\mathbf{n} < 30$	يتبع توزيع $t_{\gamma,\frac{\alpha}{2}}$ حيث $\gamma = n - 1$	$t_{\gamma,\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$	$(\bar{x} - t_{\gamma,\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\gamma,\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$
متوسط مجتمع مجهول التباين	$\mathbf{n \geq 30}$	يتبع توزيع الطبيعي $Z_{\frac{\alpha}{2}}$	$Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$	$(\bar{x} - Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}})$
	$\mathbf{n} < 30$	يتبع توزيع $t_{\gamma,\frac{\alpha}{2}}$ حيث $\gamma = n - 1$	$t_{\gamma,\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$	$(\bar{x} - t_{\gamma,\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\gamma,\frac{\alpha}{2}}\frac{s}{\sqrt{n}})$
النسبة في مجتمع	$\mathbf{n \geq 30}$	يتبع توزيع الطبيعي $Z_{\frac{\alpha}{2}}$ حيث $\square = \frac{\square}{\square}$ حيث N تمثل عدد	$\square_{\frac{\square}{\square}}\sqrt{\frac{\square(\square - \square)}{\square}}$	$\square - \square_{\frac{\square}{\square}}\sqrt{\frac{\square(\square - \square)}{\square}} \leq \square \leq \square + \square_{\frac{\square}{\square}}\sqrt{\frac{\square(\square - \square)}{\square}})$

		n مفردات المجتمع و تمثل مفردات العينة		
الفرق بين مجتمعين طبيعيين	$(\square_{\square}, \square_{\square}) \geq \square\square$	**يتبع توزيع الطبيعي** $\square_{\frac{\square}{\square}}$	$\square_{\frac{\square}{\square}}\sqrt{\frac{\square^{\square}{}_{\square}}{\square_{\square}}+\frac{\square^{\square}{}_{\square}}{\square_{\square}}}$	$((\overline{\square_{\square}} - \overline{\square_{\square}}) - \square_{\frac{\square}{\square}}\sqrt{\frac{\square^{\square}{}_{\square}}{\square_{\square}}+\frac{\square^{\square}{}_{\square}}{\square_{\square}}} \leq \square_{\square} - \square_{\square} \leq (\overline{\square_{\square}} - \overline{\square_{\square}}) - \square_{\frac{\square}{\square}}\sqrt{\frac{\square^{\square}{}_{\square}}{\square_{\square}}+\frac{\square^{\square}{}_{\square}}{\square_{\square}}}$
	$(\square_{\square}, \square_{\square}) < \square\square$	**يتبع توزيع** $\square_{\square,\frac{\square}{\square}}$ **حيث** $\gamma = n_1 + n_2 - 2$	$t_{\gamma,\frac{\alpha}{2}} sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ **حيث** $sp = \sqrt{\frac{(n_1-1){s_1}^2 +}{n_1 + n}}$	$((\overline{x_1} - \overline{x_2}) - t_{\gamma,\frac{\alpha}{2}} sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) - t_{\gamma,\frac{\alpha}{2}} sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

الفرق بين نسبتين	$(n_1, n_2) \geq 30$	يتبع توزيع الطبيعي $z_{\frac{\alpha}{2}}$	$z_{\frac{\alpha}{2}}\sqrt{\frac{r_1(1-r_1)}{n_1} + \frac{r}{}}$	$((r_1 - r_2) - z_{\frac{\alpha}{2}}\sqrt{\frac{r_1(1-r_1)}{n_1} + \frac{r_2(1-r_2)}{n_2}} \leq p_1 - p_2 \leq (r_1 - r_2) + z_{\frac{\alpha}{2}}\sqrt{\frac{r_1(1-r_1)}{n_1} + \frac{r_2(1-r_2)}{n_2}})$

سلسلة تمارين حول الفصل الثاني

التمرين الأول:

أخذت عينة عشوائية من 36 طالباً من بين طلبة السنة الثانية، متقدمين لامتحان الإحصاء. وجد أن متوسط درجات العينة هو 380، والانحراف المعيار للمجتمع كله هو 40. إذا علمت أن متوسط علامات الطلبة تتبع توزيع طبيعي:

أوجد تقدير لمتوسط علامات الطلبة في المجتمع كله بمستوى الثقة,90% و 95%.

التمرين الثاني:

يرغب باحث في تقدير متوسط الأجر الأسبوعي لعدة آلاف من العاملين بأحد المصانع وبدرجة ثقة 99%، أخذ عينة من 49 عامل بمتوسط 20 دولار. ويعرف الباحث من خبرته الماضية أن توزيع الأجر الأسبوعي للعاملين يتبع التوزيع الطبيعي بانحراف معياري قدره 40 دولار. أوجد التقدير المناسب ؟.

التمرين الثالث:

في دراسة لمعرفة عدد الساعات التي ينامها الطفل يومياً في عمر سنه واحدة اختيرت عينة من 36 طفل وكان متوسط عدد ساعات النوم يومياً 11.35 ساعة وبانحراف معياري 0.4 ساعة. بافتراض أن التوزيع طبيعي.

أوجد تقدير لمتوسط عدد ساعات النوم يومياً للطفل في عمر سنه بمستوى ثقة 95% ؟.

التمرين الرابع:

تسبب المقاعد الخالية لشركات الطيران في خسارة لمصدر الدخل، بفَرض أن إحدى شركات الطيران الكبرى أرادت تقدير عدد المقاعد الخالية لكل رحلة خلال العالم الماضي . ولهذا الغرض، تم اختيار عشوائي لعدد 225 رحلة طيران وتسجيل عدد المقاعد الخالية في كل رحلة من هذه العينة. وكان المتوسط الحسابي و الانحراف المعياري لعدد المقاعد الخالية في هذه هما: $\bar{x}=11.6$ و $S=4.1$

قدر متوسط عدد المقاعد الخالية للرحلات خلال العام الماضي بمستوى ثقة 90 % و 99%.

التمرين الخامس:

تنتج إحدى الشركات الغذائية نوع من العصير المشكل زنة العبوة 125 غرام، قام مدير مراقبة الإنتاج بسحب عينة عشوائية حجمها 36 عبوة و قياس كمية الكربوهيدرات بالغرام ، و وجد أن متوسط كمية الكربوهيدرات 12 غرام، و الانحراف المعياري هو 2.4 غرام. فإذا أراد قسم مراقبة الإنتاج تقدير فترة ثقة 95% لمتوسط كمية الكربوهيدرات في العبوات ، علمان أن وزن الكربوهيدرات يتبع التوزيع الطبيعي.

التمرين السادس:

سحبت عينة عشوائية حجمها n=10 بطارية بمتوسط عمر $\bar{x}=5$ ساعة ، وانحراف معياري للعينة s=1 ساعة من خط إنتاج مصنع ينتج بطاريات عمرها موزع طبقاً للتوزيع الطبيعي .

أوجد تقدير لمتوسط عمر بطاريات المصنع عند مستوى ثقة 90 % و 99%.

التمرين السابع:

إذا كان وزن الدجاج بالغرام في أحد المزارع بعد 45 يوم يتبع توزيع طبيعي ، سحبت عينة عشوائية من هذه المزرعة المنتجة للدجاج حجمها 25 دجاجة ، و وجد أن متوسط وزن الدجاج في العينة 890 غرام، و الانحراف المعياري لها 200 غرام.

قدر مجال الثقة لمتوسط وزن الدجاج في المزرعة باستخدام مستوى ثقة 95 %.

التمرين الثامن:

أخذت عينة عشوائية حجمها 700 طالب بالجامعة فوجد أن 25% من الطلاب يجيدون اللغة الإنجليزية. قدر نسبة الطلبة بالجامعة الذين يجيدون اللغة الإنجليزية بمستوى ثقة: 90%,95% ؟

التمرين التاسع:

في عينة عشوائية مكونة من 150 طائر من نوع معين وجد أن 50 طائر في العينة مصاب بمرض ما. قدر نسبة الطيور المصابة بالمرض بمستوى الثقة 99% ؟

التمرين العاشر:

في دراسة لتقدير مستوى مادة كيميائية ما في نوعين من النباتات A,B أخذت عينة عشوائية من كل نوع وكانت النتائج كما يلي:

	حجم العينة	الوسط الحسابي	الانحراف المعياري للمجتمع
A	40	7.208	0.727
B	50	6.64	1.57

إذا علمت أن مستوى المادة يتبع توزيع طبيعي قدر فرق بين متوسطي النوعين A,B عند مستوى ثقة 90% ؟

التمرين الحادي عشر:

تم أخذ عينتين مستقلتين من الرجال والنساء لقياس الزمن المستغرق لانجاز مهمة معينة فكانت البيانات التالية:

الرجال	10.4	11.6	15.3	14.4	9.7	14.3	13.6	12.9	19.8	16.9
النساء	8.3	12.6	9.6	13.3	10.1	12.7	11.8	14.2	12.9	14.7

قدر الفرق بين متوسطي الزمن للرجال والنساء عند مستوى ثقة 95% ؟

الفصل الثالث: الاختبارات الإحصائية

تمهيد:

يحاول الباحث اتخاذ قرار ما لمشكلة محددة بشأن خواص توزيع ما (المتوسط – النسبة) لعينة عشوائية تم سحبها من المجتمع ، ولكي نصل إلى قرار إحصائي لابد من وضع فروض عن معالم المجتمع، ومن هنا نختبر مدى صحة هذا الفرض من عدمه وذلك عن طريق العينة العشوائية التي تم سحبها من المجتمع. وهذه الفروض هي ما نطلق عليه الفروض الإحصائية.

تتمحور خطوات إجراء أي اختبار للفروض الإحصائية بشكل عام كما يلي:

- صياغة فرضان يسميان فرض العدم H_0 و الفرض البديل H_1 حول معلمة (أوخاصية) في مجتمع الدراسة.
- حساب بعض الإحصائيات كالمتوسط، و الانحراف المعياري ..الخ.
- نحسب من قيم الإحصائيات إحصائية الاختبار.
- نتخذ القرار برفض أو قبول فرض العدم.

وكما ذكرنا من قبل تعتمد اختبارات الفروض على بيانات العينة وفرض قيمة معينة لمعلمة من معالم المجتمع حيث يكون الاختبار :هل هناك فرق بين قيمة معلمة المجتمع المفروضة والقيمة المقدرة لها من خلال بيانات العينة؟ فإذا كان هناك فرق فهل يرجع هذا الفرق إلى خطأ المعاينة؟ أم هو فرق" حقيقي" معنوي فإذا كان الفرق معنويا فيكون القرار هو عدم قبول فرض العدم وعليه فإننا نقبل بالفرض البديل .أما إذا كان الفرق غير معنوي فإننا نقبل فرض العدم.

1. اختبار الفروض عن المتوسط

يتناول هذا الاختبار متوسط المجتمع μ ، مثل متوسط الدخل ، متوسط وزن منتج معين ،... و يؤكد اختبار المتوسط فرضية مساواته لقيمة ما μ_0 . و للقيام بالاختبار نستخرج عينة عشوائية نحسب فيها المتوسط $\overline{X}$ ثم نستخدم التوزيع الاحتمالي لـ $\overline{X}$ لقياس قرب أو بعد هذه القيمة من μ_0 .

الخطوات الرسمية لاختبار فروض عن متوسط المجتمع هي كالآتي:

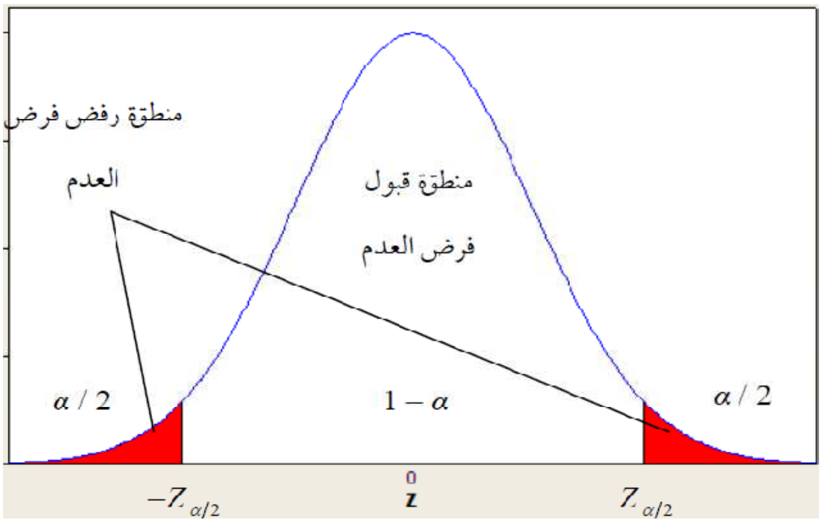
- افترض أن μ متوسط المجتمع تساوي قيمة افتراضية μ_0. يمكن تمثيل ذلك بالعبارة

$$H_0 : \mu = \mu_0$$ ويسمى بفرض العدم

وتكون الفروض البديلة إما:

$H_1: \mu < \mu_0$ أو $H_1 : \mu > \mu_0$ أو $H_1 : \mu \neq \mu_0$

- حدد مستوى معنوية للاختبار (وتمسى α مستوى المعنوية، و $1 - \alpha$ مستوى الثقة للاختبار) وعرف منطقة القبول ومنطقة الرفض للاختبار لاستخدام التوزيع الملائم



- نحسب متغيرة القرار Z حيث:

إحصاء الاختبار (Z):

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

و نقارنها بـ z المجدولة (المستخدمة في نظرية التقدير حسب مستوى المعنوية).

إذا وقعت قيمة (Z) في منطقة القبول $(1 - \alpha)$ نقبل فرض العدم H_0. وإذا وقعت قيمة (Z) في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1.

ملاحظة: عند حساب متغيرة القرار Z و كان تباين أو الانحراف المعياري σ للمجتمع مجهول، نستخدم الانحراف المعياري للعينة s كتقدير لحساب Z .

مثال1: أظهرت إحدى الدراسات أن متوسط طول طلبة أحد المدارس 160 سم يتبع توزيع طبيعي، أخذت عينة من 64 طالب من هذه المدارس فوجد أن متوسط الطول هو 155 سم . فإذا كان الانحراف المعياري للمجتمع يساوى 5 سم .اختبر الفرض القائل و ذلك عند مستوى معنوية 5%.

$o\ H : \mu = 160,\ 1\ H : \mu \neq 160$

الحل:

$$Z = \frac{\bar{x} \quad \mu_o}{\sigma / \sqrt{n}}$$

حيث:

$\bar{x} = 155, \qquad \mu_o = 160, \qquad \sigma = 5, \qquad n = 64$

$\alpha = 0.05$

$$\Rightarrow Z = \frac{155 - 160}{\frac{5}{\sqrt{64}}},$$

$Z - -8 \qquad (1)$

و هذه تمثل قيمة Z المحسوبة.

عند مستوى معنوية ($\alpha = 0.05$) وبما أن الاختبار من طرفين إذا قيم Z الجدولية سوف تكون كالتالي:

$$Z_{\alpha/2} = Z_{0.05/2} = Z_{0.025}$$
$$Z_{0.025} = 1.96$$

وعليه فإن قيمة $-Z_{\alpha/2}$ ستكون:

$$-Z_{\alpha/2} = -1.96$$

ومن هنا نجد أن قيمة Z المحسوبة أصغر من قيمة Z الجدولية، أى أن قيمة Z تقع فى منطقة الرفض، ومن هنا يتم رفض الفرض العدمى بأن $\mu = 160$.

أي أن متوسط طول الطلاب في المدارس لا يساوي 160 سم.

ملاحظة:

قيم $Z_{\alpha/2}$ نفسها المستخدمة في التقدير.

مثال2:

إذا كان متوسط الزيادة في أجور العاملين في إحدى المؤسسات سنة 2012 هو 36 وحدة نقدية ، و في عام 2015 أخذت عينة من 64 عامل في هذه المؤسسة ، فوجد أن المتوسط الحسابي للزيادة في أجورهم يبلغ 40 وحدة نقدية و الانحراف المعياري 8 وحدات نقدية.

هل يدل ذلك على أن متوسط الزيادة في أجور العاملين سنة 2015 قد اختلف عن متوسط الزيادة عام 2012 ؟ و ذلك عند مستوى معنوية $\alpha = 0.05$

الحل:

فرض العدم والفرض البديل:

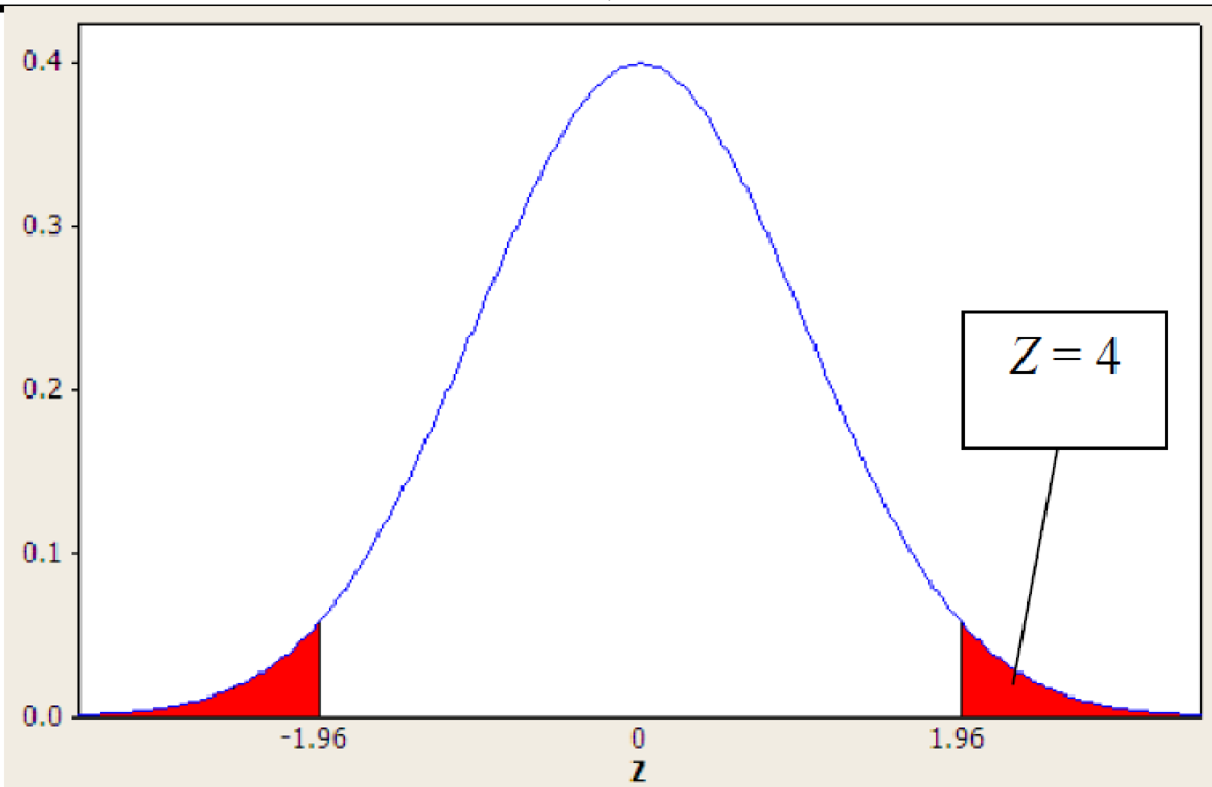
$$H_0 : \mu = 36 \quad (\text{vs}) \quad H_1 : \mu \neq 36$$

معطيات:

$$n = 64 \qquad \bar{x} = 40 \qquad s_x = 8$$

إحصاء الاختبار (Z):

$$Z = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}} = \frac{40 - 36}{8/\sqrt{64}} = 4$$



نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 ؛ أي أن متوسط الزيادة في الأجور عام 2001م قد اختلف عن متوسط الزيادة في الأجور عام 1998م. وذلك عند مستوى معنوية $\alpha = 0.05$.

مثال3:

إذا كان أحد مصانع المواد الغذائية ينتج نوعا من الألبان حيث يصل متوسط وزن العبوة 240 جرام، وذلك بانحراف معياري 18 جرام حيث كانت أوزان العبوات تتبع التوزيع الطبيعي، تم أخذ عينة من 9 عبوات وذلك عند إجراء اختبار الرقابة على الجودة فوجد أن متوسط وزن العبوة **235** جرام .هل ترى أن هناك عيبا بالإنتاج مما أدى إلى انخفاض متوسط وزن العبوة؟ مستوى المعنوية 10%.

الحل:

نضع فروض الاختبار وهى كالتالي:

$_o H : \mu = 240$

$_1 H : \mu < 240$

نقوم بحساب متغيرة القرار Z.

حيث:

$\bar{x} = 235, \quad \mu_o = 240, \quad \sigma = 18, \quad n = 9$

$\alpha = 0.01$

$$\Rightarrow Z = \frac{235 - 240}{\frac{18}{\sqrt{9}}},$$

$Z = -0.83 \qquad (1)$

نقوم الأن بحساب قيمة Z الجدولية

عند مستوى معنوية $(\alpha = 0.1)$ وبما أن الاختبار من طرف واحد إذا قيمة Z الجدولية سوف تكون كالتالي:

$-Z_\alpha = -Z_{0.1}$

$-Z_{0.1} = -1.28 \qquad (2)$

من (١) و (٢) نجد أن قيمة Z المحسوبة أصغر من قيمة Z الجدولية، أى أن قيمة Z تقع فى منطقة القبول، ومن هنا يتم قبول الفرض العدمى بأن $\mu = 240$.

ملاحظات:

في حالة مجتمع غير طبيعي لكن حجم العينة كبير (أكبر من 30) حسب نظرية النهاية المركزية يتم تقريب توزيع المعاينة للتوزيع الطبيعي و بالتالي فإن متغيرة القرارz تتبع توزيع طبيعي.

في حالة عينة صغيرة $n \leq 30$ و ∂ الانحراف المعياري للمجتمع مجهول ، لا يمكن استخدام التوزيع الطبيعي ، حيث نستخدم في هذه الحالة توزيع سيتودنت (بشرط أن يكون توزيع المجتمع طبيعي)

مثال4:

إذا كان متوسط ربح سهم إحدى الشركات هو 15 دولار فى العام الماضي، تم أخذ عينة من 70 مساهم عن توقعاتهم عن متوسط ربح السهم العام الحالي فوجد أنه 17 جنية بانحراف معياري 2% هل توافق المساهمين في تقديرهم لارتفاع متوسط ربح السهم هذا العام؟ وذلك بمستوى معنوية 5%.

الحل: 1-نضع فروض الاختبار وهى كالتالي:

$H_0 : \mu = 15$

$H_1 : \mu > 15$

٢- نقوم بحساب قيمة t المحسوبة

$$t = \frac{\bar{x} - \mu_o}{S/\sqrt{n}}$$

حيث:

$\bar{x} = 17, \qquad \mu_o = 15, \qquad S = 2, \qquad n = 7$

$\alpha = 0.05$

$$\rightarrow t = \frac{17 - 15}{\frac{2}{\sqrt{7}}},$$

$t = 2.65 \qquad (1)$

وهذه تمثل قيمة t المحسوبة.

٣- نقوم الآن بحساب قيمة t الجدولية.

عند مستوى معنوية $(\alpha = 0.05)$ وبما أن الاختبار من طرف واحد إذا قيمة t الجدولية سوف تكون كالتالي:

$t_{(n-1,\alpha)} = t_{(7-1,0.05)}$

$t_{(6,0.05)} = 1.943 \qquad (2)$

من (١) و (٢) نجد أن قيمة t المحسوبة أكبر من قيمة t الجدولية، أى أن قيمة t تقع في منطقة الرفض، ومن هنا يمكن القول بأننا نوافق المساهمين على توقعهم بارتفاع متوسط ربح السهم هذا العام.

2. اختبار الفروض عن 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:

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

فرض العدم والفرض البديل:

$$H_0 : \mathrm{P} = \mathrm{P}_0 \quad (\mathrm{vs}) \quad H_1 : \mathrm{P} \neq \mathrm{P}_0$$

إحصاء الاختبار (Z):

$$Z = \frac{r - \mathrm{P}_0}{\sqrt{\frac{\mathrm{P}_0(1-\mathrm{P}_0)}{n}}}$$

نحدد مكان وقوع قيمة إحصاء الاختبار (Z) في الرسم التوضيحي. حيث إذا وقعت قيمة (Z) في منطقة القبول نقبل فرض العدم H_0. وإذا وقعت قيمة (Z) في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1.

مثال1:

تقدر الدوائر الرسمية نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم ب 70 % . وجدت دراسة أجريت على عينة من 900 طالب أن نسبة الحصول على عمل 67 % . كيف يمكن اختبار ما إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها، بمستوى معنوية 5 %.

$$H0 : p = 0.70 \leftrightarrow H1 : p < 0.70$$

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

ومنه نرفض الفرضية H0 .

مثال 2 من بين (900) شخص وجد أن عدد المؤيدين منهم لرأي معين هو (738) شخص. اختبر الفرض أن نسبة المؤيدين لذلك الرأي في المجتمع هو (0.8)، وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل:

فرض العدم والفرض البديل:

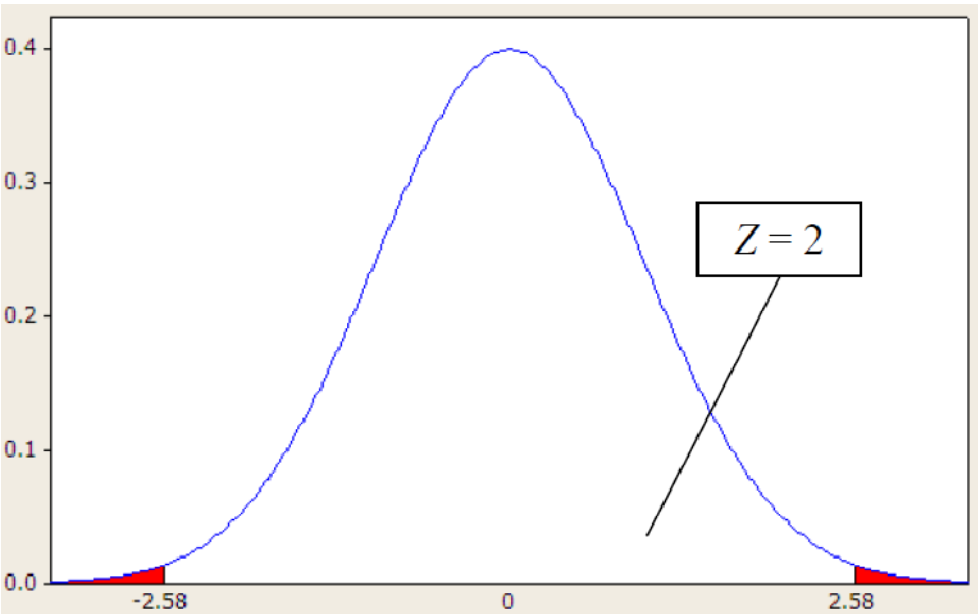
$$H_0 : P = 0.8 \quad (vs) \quad H_1 : P \neq 0.8$$

معطيات:

$$n = 900 \qquad\qquad r = \frac{738}{900} = 0.82$$

إحصاء الاختبار (Z):

$$Z = \frac{r - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.82-0.8}{\sqrt{\frac{0.8\times0.2}{900}}} = \frac{0.02}{0.01} = 2$$



تقبل فرض العدم H_0 ؛ نسبة المؤيدين لذلك الرأي في المجتمع هو (0.8)، وذلك عند مستوى معنوية $\alpha = 0.01$.

3. اختبار الفروض للفرق بين متوسطي مجتمعين

هناك العديد من الأبحاث التي يكون المطلوب فيها المقارنة بين مجتمعين مختلفين أو منطقتين مختلفتين أو أسلوبين لتدريس مقرر ما...الخ .من هنا جاء اختبار الفرق بين متوسطي المجتمعين عن طريق سحب عينة من كل مجتمع ويكون الاختبار لنرى الفرق بين متوسطي العينتين، وهل هو فرق حقيقي أم يرجع هذا الفرق إلى الصدفة البحتة.

إذا كان لدينا عينتين، متوسط العينة الأولى هو $\overline{X}_1$ و متوسط العينة الثانية $\overline{X}_2$ ، بانحراف معياري S_1 و S_2 على الترتيب و كان حجم العينتين هو n_1 و n_2 .

تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي: $H_0 : \mu_1 = \mu_2 \leftrightarrow H_1 : \mu_1 \neq \mu_2$

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

(أ) الحالة الأولى: تباينا ا تمعين معلومين

1- المجتمعين طبيعيين:

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

2- مجتمعين ما $(n1 , n2 \geq 30)$:

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

ب) الحالة الثانية: تباينا ا تمعين مجهولين

- المجتمعان طبيعيان: إذا كان تباينا المجتمعين متساويين

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

2- مجتمعين ما $(n1 , n2 \geq 30)$:

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

مثال 1: إذا اختيرت عينة عشوائية من 60 طالب من جامعة خاصة، فوجد أن متوسط ذكائهم 69 درجة وتباين قدرة 230 درجة، كذلك تم اختيار عينة عشوائية أخرى من ٨٥ طالب من جامعة حلوان فوجد أن متوسط ذكائهم ٧٤ درجة وتباين قدرة ٢١٥ درجة. اختبر الفرض القائل بأن متوسط ذكاء طالب الجامعة الخاصة أقل من متوسط ذكاء طالب جامعة **عامة**.

وذلك بمستوى معنوية $(\alpha = 0.05)$.

الحل:

نضع فروض الاختبار وهي كالتالي:

$$H_o: \mu_1 = \mu_2$$
$$H_1: \mu_1 < \mu_2$$

نقوم بحساب قيمة Z المحسوبة 30 و 30).

$$Z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)_o}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z = \frac{(69 - 74) - 0}{\sqrt{\frac{230}{60} + \frac{215}{85}}}$$

$$Z = \frac{-5}{\sqrt{6.36}}$$

$$Z = -1.98, \qquad (1)$$

عند مستوى معنوية $(\alpha - 0.05)$ وبما أن الاختبار من طرف واحد إذا قيمة Z الجدولية سوف تكون كالتالي:

$$-Z_\alpha = -Z_{0.05}$$

$$-Z_{0.05} = -1.65 \qquad (2)$$

من (١) و (٢) نجد أن قيمة Z المحسوبة أصغر من قيمة Z الجدولية، أى أن قيمة Z تقع فى منطقة الرفض، ومن هنا يمكن القول بأن متوسط ذكاء طلبة الجامعة الخاصة أقل من متوسط ذكاء طلبة جامعة حلوان.

سلسة تمارين حول الفصل الثالث

التمرين الأول:

يرغب منتج صفائح من الزجاج اختبار ما إذا كانت الصفائح التي ينتجها لديها قوة مقاومة للكسر قدرها 5,000 وحدة مقاومة. فقوة مقاومة للكسر أقل من 5,000 وحدة لن تكون ملائمة، وقوة مقاومة للكسر أكبر من 5,000 وحدة ترفع التكاليف بدون مبرر. يأخذ المنتج عينة عشوائية من 64 قطعة ويجد أن متوسط قوة المقاومة للكسر هو 5,100 وحدة، إذا علمت أن التوزيع طبيعي والانحراف المعياري هو 480. هل يجب أن يقبل المنتج الفرض أن صفائح الزجاج لها قوة مقاومة للكسر5,000 وحدة عند مستوى معنوية 5%؟

التمرين الثاني:

يعرف مركز تجنيد بالجيش من الخبرة الماضية أن وزن المجند يتبع التوزيع الطبيعي بوسط μ يساوي 80 كيلو جراماً وانحراف معياري σ يساوي 10 كيلو جراما. ويرغب مركز التجنيد أن يختبر، عند مستوى معنوية 1%، ما إذا كان متوسط وزن مجند هذا العام أكبر من 80 كيلو جراماً. ولعمل هذا، فقد أخذ عينة عشوائية من 25 مجنداً حيث وجد أن متوسط الوزن في العينة 85 كيلو جراماً. كيف يمكن إجراء هذا الاختبار؟

التمرين الثالث:

في دراسة على متوسط أعمار الطلاب المتقدمون لشغل إحدى الوظائف أخذت عينة عشوائية من 30 متقدم للوظيفة، فوجد أن متوسط أعمارهم 25 سنة وذلك بانحراف معياري 5 سنوات . هل يمكن القول بأن متوسط أعمار جميع المتقدمين يساوى 28 سنة وذلك في 95 %من الحالات؟

التمرين الرابع:

في دراسة لمعرفة عدد الساعات التي ينامها الطفل يومياً في عمر سنه واحدة اختيرت عينة من 36 طفل وكان متوسط عدد ساعات النوم يومياً 11.35 ساعة بافتراض أن التوزيع طبيعي،وبانحراف معياري 0.4 ساعة.

هل يمكن القول بأن متوسط عدد ساعات النوم يومياً للطفل في عمر سنه أقل من 11.5 ساعة عند $\alpha=0.10$ ؟

التمرين الخامس:

يريد مستشفى أن يختبر أن 90% من جرعات عقار يشتريه يحتوي على 100mg من العقار. لعمل هذا، يأخذ المستشفى عينة من $n = 100$ جرعة، ويجد أن 95% منها فقط يحتوي على الكمية المناسبة. كيف يمكن للمستشفى أن يختبر هذا عند:

(أ) $\alpha = 1\%$ ؟ (ب) $\alpha = 5\%$ ؟ (ج) $\alpha = 10\%$ ؟

التمرين السادس:

يدعى متحدث حكومي لمكافحة التلوث أن أكثر من 80% من المصانع في المنطقة تستوفى معايير مكافحة التلوث. ولكن أحد جمعيات أنصار مكافحة التلوث لا تصدق إدعاء الحكومة. فهي تأخذ عينة عشوائية من البيانات المنشورة عن مكافحة التلوث في 64 مصنعاً في منطقة وتجد أن منها 56 مصنعاً مستوفى معايير المكافحة. هل تؤيد بيانات العينة إدعاء الحكومة عند مستوى معنوية 5%؟

التمرين السابع:

أخذت عينة عشوائية مكونة من 125 طالب في جامعة الجزائر 3 في السنة الثانية فوجد 95 طالب منهم قد نجحوا في جميع المقاييس. اختبر الفرض القائل بأن نسبة الطلبة الناجحين في جميع المقاييس أقل من60% عند $\alpha = 0.1$ ؟

التمرين الثامن:

يرغب مشتر كبير للمصابيح الكهربائية أن يقرر، أي صنف يشترى من بين صنفين لهما نفس السعر. لعمل هذا، فأنه يأخذ عينة عشوائية من 100 مصباح من كل صنف فيجد أن الصنف الأول يعيش في المتوسط $X = 980$ ساعة، مع انحراف معياري، s1 قدره 80 ساعة وبالنسبة للصنف الثاني، $X = 1,010$ ساعة و $s2 = 120$ ساعة. أي الصنفين يجب شراءه إذا كان المشتري يرغب في أن يصل إلى قرار عند مستوى معنوية: (أ) 5% ؟ (ب) 1% ؟

التمرين التاسع:

متوسط الدرجات في امتحان القبول لمواصلة الماستر لعام 2015 لـ 64 طالباً مترشحا هو 640 درجة بانحراف معياري 20 درجة. وفي عام 2016 تقدم 81 طالباً للالتحاق بالماستر فكان متوسط درجاتهم في امتحان القبول 650 درجة بانحراف معياري 40.

هل مستوى المتقدمين عام 2015 أقل من مستوى المتقدمين عام 2016 عند مستوى معنوية 1%؟

التمرين العاشر:

يرغب اتحاد أطباء الأسنان في اختبار أي معجون من بين معجوني أسنان أفضل في محاربة التسوس. أخذت عينة عشوائية من 21 شخصاً من مستعمل كل من المعجونين موضع الاختبار. ووجد أن متوسط عدد الفجوات للمجموعة الأولى على مدى 10 سنوات هو 25 بانحراف معياري 5 وبالنسبة للمجموعة الثانية، متوسط الفجوات 23 بانحراف معياري 5 بافتراض أن توزيع الفجوات طبيعي لمستعمل المعجون الأول والمعجون الثاني، وأن $\sigma^2_1 = \sigma^2_2$، حدد إذا كانت $\mu_1 = \mu_2$ عند مستوى معنوية 5% .

قائمة المراجع:

1-حسن ياسين طعمه، أساليب الإحصاء التطبيقي، دار صفاء،2010

2-جامعة الملك عبد العزيز ،مقدمة في الإحصاء لطلاب الدراسات الاقتصادية و الإدارية.

3-المحمدي شاكر مصلح، " الإحصاء وتصميم الاختبارات" ، دار أسامة للنشر، عمان، 2011

4- ANDERSON, SWEENEY et WILLIAMS : Statistiques pour l’économie
Et la gestion, Editions De Boeck, 2010

5- KHALDI KHALEDE, méthodes statistiques, rappels de cours et exercices corrigés, OPU, 2005.