محاضرات مقياس الإحصاء 3

لفائدة طلبة السنة الثانية علوم اقتصادية، علوم تجارية وعلوم التسيير

قسم علوم التسيير

للأستاذة: طالب دليلة

الموسم الجامعي: 2021 / 2022

المحاضرة الاولى:

التوزيع الطبيعي

التوزيع الطبيعى

NORMAL DISTRIBUTION

يعتبر التوزيع الطبيعى من أهم التوزيعات المتصلة حيث يلعب دوراً أساسياً فى عملية المعاينة Sampling ، كما يستخدم لوصف الأنماط التكرارية للعديــد مــن الظــواهر الإحصائية مثل التغيرات الطبيعية التى تحدث للإنسان والحيوان والعوامل البيئية بشكل عام . والتوزيع الطبيعى يدرس سلوك المتغيرات العشوائية المتصلة مثل درجة الحرارة والطــول والوزن والدخل والأخطاء العشوائية الناتجة عند تحليل الانحدار .

وبدراسة شكل منحنى التوزيع الطبيعى نجد أنه منحنى متماثل حول الوسط الحسابى للتوزيع ، ويأخذ الشكل الجرسى وله قمة واحدة ويمتد طرفاه إلى ما لانهاية مقتربين مــن المحور الأفقى شيئا فشيئا دون أن يتماسا مع هذا المحور . وإذا أسقطنا عموداً من قمــة المنحنى على المحور الأفقى فإن هذا العمود يعتبر محوراً للتماثل لأنه يقسم المساحة تحتى المنحنى إلى قسمين متساويين تماماً وينطبق كل منهما على الآخر تمام الانطباق ومساحة كل قسم تساوى 50% من المساحة الكلية تحت المنحنى وينتج عن هذا التماثل أن قيم الوسط الحسابى والوسيط والمنوال للتوزيع الطبيعى تكون متساوية .

وحيث أن التوزيع الطبيعى يعبر عن متغيرات عشوائية متصلة فإننا لا نستطيع استخدام دالة التوزيع له لإيجاد احتمال أن يأخذ المتغير العشوائى الذى يتبعه لقيمة معينة حيــث أن هذا الاحتمال يساوى صفراً . ولكن يمكن استخدام دالة التوزيع الطبيعى لإيجاد احتمال أن يأخذ المتغير العشوائى أقل من قيمة معينة أو أكثر من قيمة معينة أو تقع قيمته بين قيمتــين معلومتين.

وترجع أهمية التوزيع الطبيعى إلى الأسباب الآتية :

1. يستخدم التوزيع الطبيعى فى وصف العديد من المتغيرات العشوائية فى الواقع العملــى مثل الأطوال والأوزان لمجموعة من الأشخاص ، درجات الطلاب فى امتحــان مقــرر معين ، أخطاء القياس الناتجة من تجربة ما ، حيث تتــوزع القياســات المشــاهدة (

التكرارات) بشكل متماثل حول قيمة مركزية هى قيمة متوسط التوزيع وتقـــل هـــذه القياسات تدريجيا كلما ابتعدنا عن هذه القيمة إلى جهة اليمين أو إلى جهة اليســـار ، حيث يأخذ منحنى التوزيع الشكل الجرسى .

2. يعتبر التوزيع الطبيعى تقريبا مفيدا للعديد من التوزيعات المتقطعة وغير المتقطعة مثـــل توزيع ذى الحدين وتوزيع بواسون وتوزيع هيبرجيومترك وغيرها .

3. يعتبر التوزيع الطبيعى حجر الزاوية فى الاستدلال الإحصائى حيث تتوزع معالم المجتمع المقدرة من العينة مثل المتوسط مثلا كالتوزيع الطبيعى وهذا ما سوف تناوله فى معرض حديثنا عن نظرية النهاية المركزية .

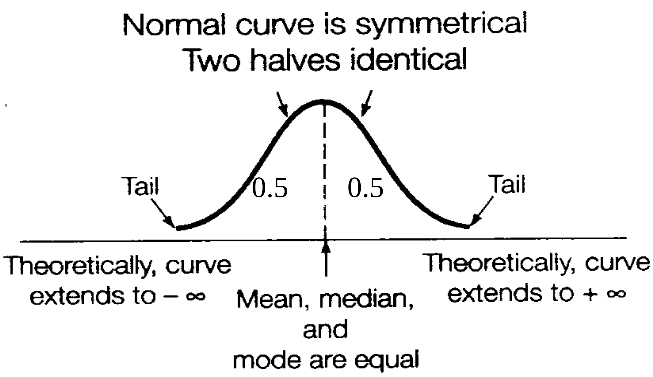
– خصائص التوزيع الطبيعى

يتميز التوزيع الطبيعى بالخصائص الآتية إجمالا :

1. منحنى التوزيع الطبيعى منحنى متماثل حول متوسط التوزيع والذى يرمز له بالرمز μ .

2. المتوسط والوسيط والمنوال للتوزيع الطبيعى متساوية القيمة .

3. المساحة تحت منحنى التوزيع الطبيعى تساوى الوحدة المربعة وهذه المساحة تعبر عن مجموع الاحتمالات . وكما ذكرنا أن الخط الرأسى الساقط من قمة المنحنـــى علـــى المحور الأفقى عند المتوسط يقسم المساحة إلى نصفين متساويين 50% من المساحة الكلية على يمين الخط العمودى ، 50 % من المساحة الكلية على يساره .

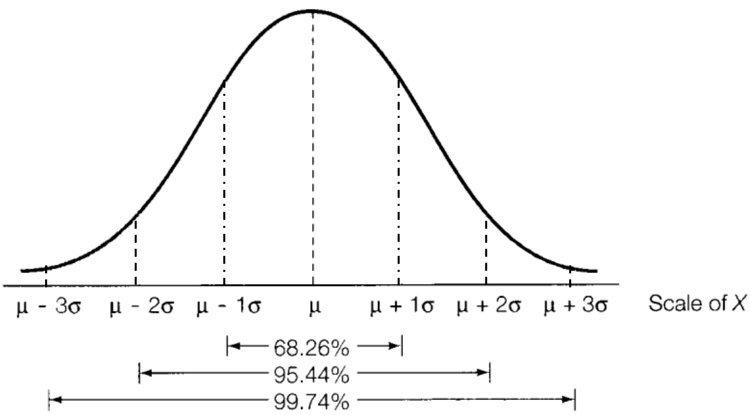


4. إذا أسقطنا عمودين على بعد انحراف معيارى واحد إلى يمين ويسار متوسط التوزيـــع فإن المساحة التى يحصرها هذين العمودين تمثل 68.26% تقريبا من المساحة الكلية تحت المنحنى . كذلك نجد أن المساحة التى يحصرها العمودان المقامـــان عنـــد 2 انحراف معيارى على جانبى متوسط التوزيع تمثل 95.44% تقريبا من المساحة الكلية تحـــت المنحنـــى . أمـــا المســـاحة التـــى يحصـــرها العمـــودان المقامـــان عند 3 انحرافات معيارية على جانبى متوسط التوزيع فتمثل 99.74% تقريبـــا مـــن المساحة الكلية تحت المنحنى . هذا ويمكن التعبير عما سبق رمزيا كالآتى :

$$p(\mu - 1\sigma < x < \mu + 1\sigma) = 0.6826$$

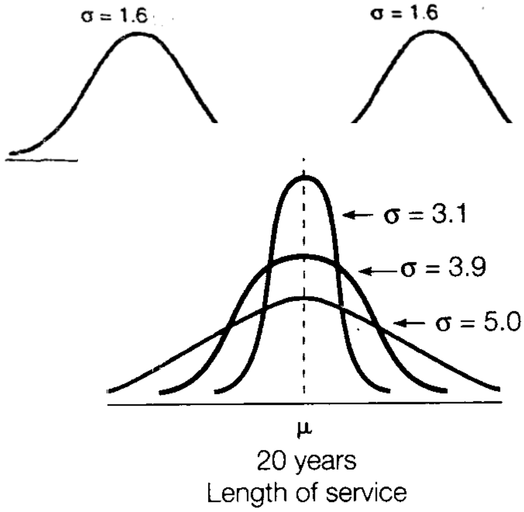
$$p(\mu - 2\sigma < x < \mu + 2\sigma) = 0.9544$$

$$p(\mu - 3\sigma < x < \mu + 3\sigma) = 0.99.74$$



5. التوزيع الطبيعى ليس توزيعاً وحيدا ولكنه يمثل عائلة من التوزيعات الطبيعية ، وكـــل توزيع من هذه العائلة يمكن تحديده وتعريفه بدقة متى عرفت معالمه وهى المتوســـط ويرمز له بالرمز μ ، والانحراف المعيارى ويرمز له بالرمز σ ، حيث تحـــدد قيمتـــى هاتين المعلمتين شكل منحنى التوزيع ، حيث تحدد قيمة المتوسط μ موقع التوزيـــع

على المحور الأفقى من نقطة الأصل ، كما تحدد قيمة الانحراف المعيارى مدى تشتت التوزيع فكلما كانت σ كبيرة كلما زاد تشتت القيم وبالتالى اتساع المنحنى .



- دالة كثافة الاحتمال Probability Density Function

إذا كان X متغيرا عشوائيا متصلا يتبع التوزيع الطبيعى بمتوسط μ وانحراف معيارى σ فإن دالة كثافة الاحتمال للمتغير العشوائى X تعطى بالعلاقة :

$$\mathbf{f(X)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}} \qquad , -\infty < X < \infty$$

$$, -\infty < \mu < \infty \qquad , \ \sigma > 0$$

وحيث أن استخدام هذه الدالة لحساب الاحتمالات المختلفة تكتنفه صعوبات رياضية كثيرة تعتمد بالدرجة الأولى على معرفة تامة بعلم التكامل ، علاوة على أنه كما ذكرنا يوجد عدد لا نهائى من التوزيعات الطبيعية والتى تحدد كل منها قيم المعلمتين μ ، σ فإنه لحسن الحظ يمكن تحويل أى توزيع طبيعى إلى توزيع طبيعى عيارى له جداول خاصة تعرف باسم جداول المساحات تحت منحنى التوزيع الطبيعى العيارى متى علم متوسطه وانحرافه المعيارى .

التوزيع الطبيعى العيارى Standard Normal Distribution

إذا كان X متغير عشوائى يتبع التوزيع الطبيعى بمتوسط μ وانحراف معيـــارى σ فإن Z تتبع التوزيع الطبيعى المعيارى بمتوسط صفر، وانحراف معيارى يساوى الواحـــد الصحيح حيث:

$$Z = \frac{X - \mu}{\sigma}$$

وتسمى Z أيضا بالقيمة المعيارية أو الدرجة المعيارية وهى تســـاعد علـــى إيجـــاد المساحة أو الاحتمال المطلوب أسفل أى منحنى توزيع طبيعى وذلك باستخدام جـــدول المساحات تحت منحنى التوزيع الطبيعى العيارى . هذا وتأخذ دالـــة كثافـــة الاحتمـــال للمتغير Z الشكل الآتى :

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}Z^2} \qquad , -\infty < Z < \infty$$

ومن الجدير بالذكر هنا أن الدالة لا تعتمد على معالم مجهولة القيمة وبالتـــالى فقـــد استخدمت فى حساب جدول المساحات تحت منحنى التوزيع الطبيعى العيارى المشـــار إليه بعاليه .

خصائص المنحنى الطبيعى العيارى

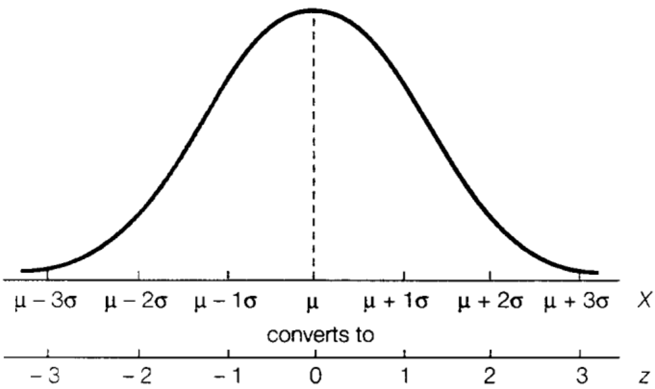
1. المساحة الكلية تحت منحنى التوزيع الطبيعى العيارى تساوى الواحد الصحيح بحيث يقسم الخط العمودى الواصل من قمة المنحنى إلى الصـــفر (متوســـط التوزيـــع) المساحة إلى قسمين متساويين مساحة كل منهما تساوى 0.05
2. منحنى التوزيع الطبيعى العيارى متماثل حول متوسطه ، وبالتالى فإن التواءه يســـاوى صفر وتفرطحه يساوى 3 .

3. المســـاحة المحصـــورة بـــين ± 1 درجـــة معياريـــة تســـاوى $68.26\ \%$ تقريبـــا مـــن المســـاحة الكليـــة تحـــت المنحنـــى . والمســـاحة المحصـــورة بين ± 2 درجة معيارية تساوى $95.44\ \%$ تقريبا من المساحة الكليـــة تحـــت المنحنى . أما المساحة المحصورة بين ± 3 درجات معيارية تساوى $99.74\ \%$ تقريبا من المساحة الكلية تحت المنحنى . هذا ويمكن التعبير عن ذلك رمزيا كالآتى :

$$P\,(\ -1 < Z < 1\ ) = 0.6826$$

$$P\,(\ -2 < Z < 2\ ) = 0.9544$$

$$P\,(\,(-3 < Z < 3\ ) = 0.9974$$



مثـــــــال (1)

إذا كان العمر الافتراضى لمشغل الأقراص المرنة والمنتج بمعرفة إحـــدى شـــركات الكمبيوتر يتبع التوزيع الطبيعى بمتوسط 1000 ساعة وانحراف معيارى 100 ساعة . فإذا تم اختيار وحدة مشغل من إنتاج هذه الشركة بطريقة عشوائية ، فأوجد احتمال أن يكون عمرها الافتراضى :

1. بين 1000 ، 1150 ساعة .
2. أقل من 930 ساعة .
3. أكبر من 780 ساعة .

4. بين 700 ، 1200 ساعة .

5. بين 750 ، 850 ساعة .

الحـــــل :

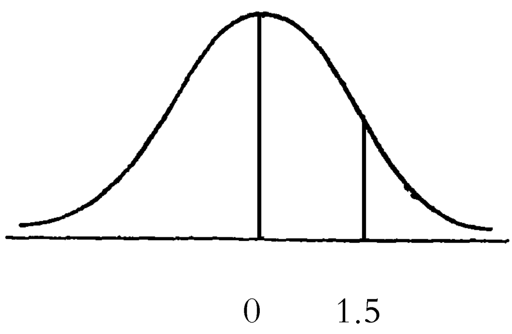
نفرض أن المتغير العشوائى χ يعبر عن العمر الافتراضى لمشغل الأقراص المرنة ، أى أن :

المتوسط μ = 1000 ساعة

الانحراف المعيارى σ = 100 ساعة

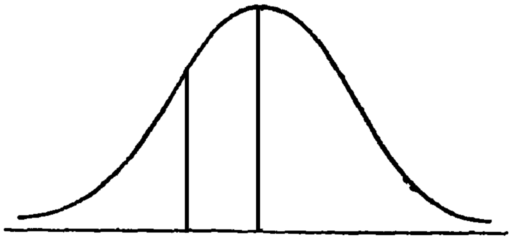
1. احتمال أن يكون العمر الافتراضى لمشغل الأقراص المرنة بين 1000 ، 1150 ساعة

$$P(1000 \le X \le 1150) = P\left(\frac{1000-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{1150-1000}{100}\right)$$
$$= P(0 \le Z \le 1.5) = 0.4332$$



02 احتمال أن يكون العمر الافتراضى لمشغل الأقراص المرنة أقل من 930 ساعة

$$P(X < 930) = P\left(\frac{X-\mu}{\sigma} < \frac{930-1000}{100}\right) = p\left(Z < \frac{-70}{100}\right)$$
$$= p(Z < -0.7) = 0.5 - 0.2580 = 0.2420$$



-.7 0

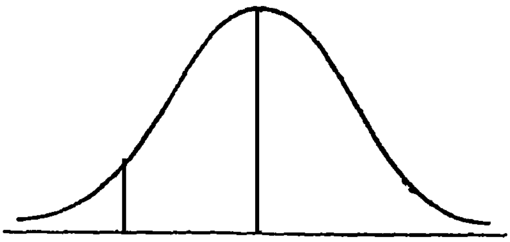
3. احتمال أن يكون العمر الافتراضى لمشغل الأقراص المرنة أكبر من 780 ساعة

$$P(X>780)=P\left(\frac{X-\mu}{\sigma}>\frac{780-1000}{100}\right)=P\left(Z>\frac{-220}{100}\right)$$

$$=P(Z>-2.2)=0.5+P(-2.2<Z<0)$$

$$=0.5+P(0<Z<2.2)$$

$$=0.5+0.4861=0.9861$$



-2.2 0

4. احتمـــال أن يكـــون العمـــر الافتراضـــى لمشـــغل الـــأقراص المرنـــة بين 700 ، 1200 ساعة

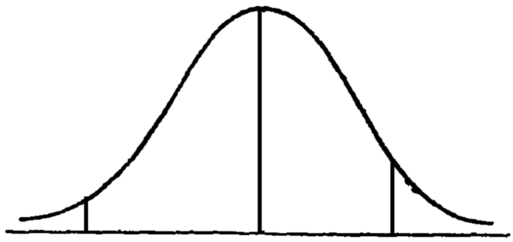
$$p(700 \le X \le 1200) = P\left(\frac{700-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{1200-1000}{100}\right)$$

$$= P\left(\frac{-300}{100} \le Z \le \frac{200}{100}\right) = P(-3 \le Z \le 2)$$

$$= P(-3 \le Z \le 0) + P(0 \le Z \le 2)$$

$$= P(0 \le Z \le 3) + P(0 \le Z \le 2)$$

$$= 0.4987 + 0.4772 = 0.9759$$



5. احتمال أن يكون العمر الافتراضى لمشغل الأقراص المرنة بين 750 ، 850 ساعة

$$P(750 \le X \le 850) = P\left(\frac{750-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{850-1000}{100}\right)$$

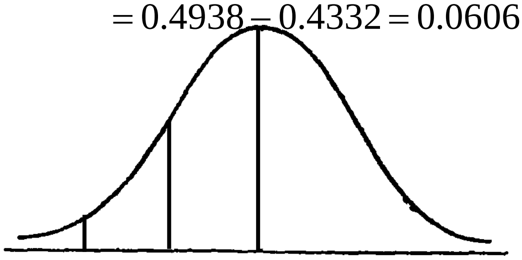
$$= P\left(\frac{-250}{100} \le Z \le \frac{-150}{100}\right)$$

$$= P(-2.50 \le Z \le -1.50)$$

$$= P(-2.50 \le Z \le 0) - P(-1.50 \le Z \le 0)$$

$$= P(0 \le Z \le 2.50) - P(0 \le Z \le 1.50)$$

$$= 0.4938 - 0.4332 = 0.0606$$



$-2.5 \quad -1.5 \quad 0$

التوزيع الطبيعى كتقريب لتوزيع ذى الحدين

يمكن استخدام التوزيع الطبيعى كتقريب لتوزيع ذى الحدين إذا تحققت الشـــروط الآتية :

1. حجم العينة يكون كبيرا ($n \geq 30$)
2. المتوسط يكون أكبر من أو يساوى 5 ($\mu = np \geq 5$)
3. التباين يكون أكبر من أو يساوى 5 ($\sigma^2 = np(1-p) \geq 5$)

مثـــــــال (2)

إذا كانـــت نســـبة الوحـــدات المعيبـــة المنتجـــة بواســـطة إحـــدى الآلـــات هى 15% ، فإذا تم اختيار عينة حجمها 150 وحدة بطريقة عشوائية من إنتاج هذه الآلة . فاحسب احتمال أن تحتوى هذه العينة على :

1. أقل من 20 وحدة معيبة .
2. 20 وحدة معيبة على الأقل .
3. ما بين 15 ، 20 وحدة معيبة .
4. أكثر من 18 وحدة معيبة .
5. أكثر من 28 وحدة معيبة .

الحـــــــل :

نلاحظ أن شروط استخدام التوزيع الطبيعة كتقريب لتوزيع ذى الحدين متوفرة حيث أن :

1. $n = 150 > 30$

2. $\mu = np = 150 \times 0.15 = 22.5 > 5$

3. $\sigma^2 = np(1-p) = 150 \times 0.15 \times 0.85 = 19.125 > 5$

وعلى فرض أن المتغير العشوائى X يعبر عن عدد الوحدات المعيبة فإن الاحتمالـــات المطلوبة يمكن حسابها كالآتى :

1. احتمال أن تحتوى العينة على أقل من 20 وحدة معيبة

$$P(X \prec 20) = P\left(\frac{X-\mu}{\sigma} < \frac{20-22.5}{4.4}\right)$$

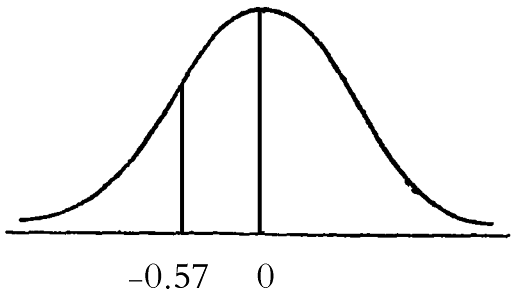
$$= p\left(Z < \frac{-2.5}{4.4}\right)$$

$$= p(Z < -0.57)$$

$$= 0.5 - p(-0.57 < Z < 0)$$

$$= 0.5 - P(0 < Z < 0.57)$$

$$= 0.5 - 0.2157 = 0.2843$$



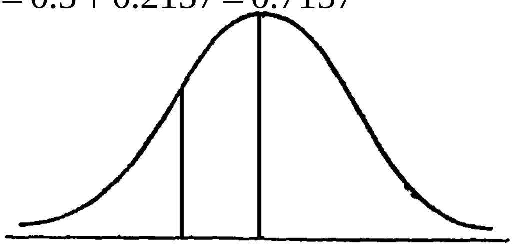
2. احتمال أن تحتوى العينة على 20 وحدة معيبة على الأقل

$$P(X \geq 20) = P\left(\frac{X-\mu}{\sigma} \geq \frac{20-22.5}{4.4}\right)$$

$$= p(Z \geq -0.57) = 0.5 + p(-0.57 \leq Z \leq 0)$$

$$= 0.5 + P(0 \leq Z \leq 0.57)$$

$$= 0.5 + 0.2157 = 0.7157$$



−0.57 0

3. احتمال أن تحتوى العينة على ما بين 15 ، 20 وحدة معيبة

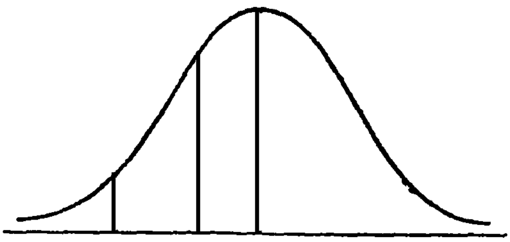
$$P(15 \le X \le 20) = P\left(\frac{15-22.5}{4.4} \le \frac{X-\mu}{\sigma} \le \frac{20-22.5}{4.4}\right)$$

$$= P(-1.71 \le Z \le -0.57)$$

$$= P(-1.71 \le Z \le 0) - p(-0.57 \le Z \le 0)$$

$$= P(0 \le Z \le 1.71) - P(0 \le Z \le 0.57)$$

$$= 0.4564 - 0.2157 = 0.2407$$



−1.71 −0.57 0

4. احتمال أن تحتوى العينة على أكثر من 18 وحدة معيبة

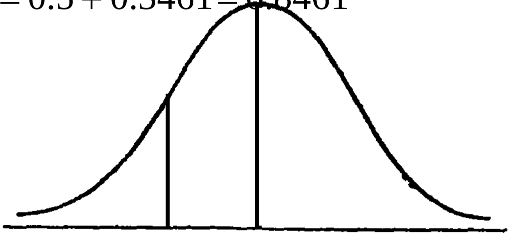
$$P(X > 18) = P\left(\frac{X-\mu}{\sigma} > \frac{18-22.5}{4.4}\right)$$

$$= P\left(Z > \frac{-4.5}{4.4}\right) = P(Z > -1.02)$$

$$= 0.5 + P(-1.02 < Z < 0)$$

$$= 0.5 + P(0 < Z < 1.02)$$

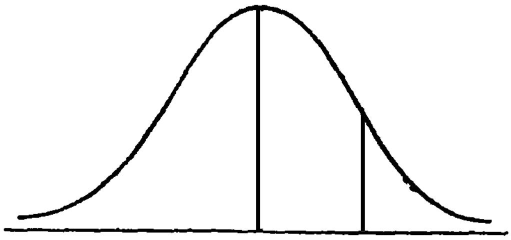
$$= 0.5 + 0.3461 = 0.8461$$



$-1.02 \quad 0$

5. احتمال أن تحتوى العينة على أكثر من 28 وحدة معيبة

$$P(X > 28) = P\left(\frac{X-\mu}{\sigma} > \frac{28-22.5}{4.4}\right)$$
$$= P\left(Z > \frac{5.5}{4.4}\right) = P(Z > 1.25)$$
$$= 0.5 - P(0 < Z < 1.25)$$
$$= 0.5 - 0.3944 = 0.1056$$



نظرية النهاية المركزية **Central Limit Theorem**

إذا سحبنا عدة عينات من الحجم n بطريقة عشوائية من مجتمع متوسطه μ وانحرافه المعيارى σ فإن توزيع المعاينة للمتوسطات يتوزع تقريبا كالتوزيع الطبيعى بمتوسط μ وانحراف معيارى $\frac{\sigma}{\sqrt{n}}$، وغالبا ما يسمى الانحراف المعيارى للمعاينة باسم الخطأ المعيارى Standard Error وبالتالى فإن المقدار $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ يتوزع كالتوزيع الطبيعى المعيارى بمتوسط صفر وانحراف معيارى يساوى الواحد الصحيح .

مثـــــال (3)

إذا كان الزمن اللازم لأداء خدمة معينة للعملاء بأحد البنوك يتــوزع كــالتوزيع الطبيعــى بمتوسط 25 ثانية وانحراف معيارى 4 ثوانى . فإذا تم سحب عينة عشوائية من 25 شــخص يقومون بأداء هذه المهمة ، فاحسب احتمال أن يكون متوسط الزمن اللازم لأداء هذه الخدمــة البنكية :

1. 26 ثانية أو أكثر .
2. بين 24 ، 27 ثانية .
3. 26 ثانية أو أقل .
4. 23 ثانية على الأقل .

الحـــــل : $n = 25 \qquad \sigma = 5 \qquad \mu = 25$

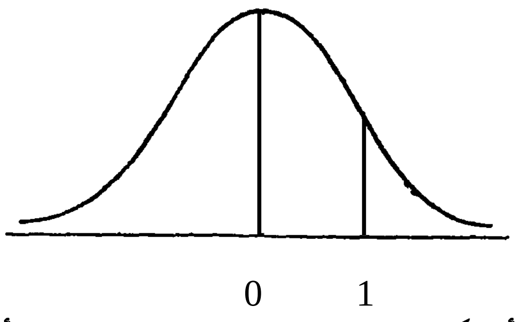
1. احتمــــال أن يكـــون متوســـط الـــزمن اللـــازم لـــأداء الخدمـــة 26 ثانيـــة أو أكثر

$$P(\overline{X} \geq 26) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{26-25}{5/\sqrt{25}}\right)$$

$$= P(Z \geq 1)$$

$$= 0.5 - P(0 \leq Z \leq 1)$$

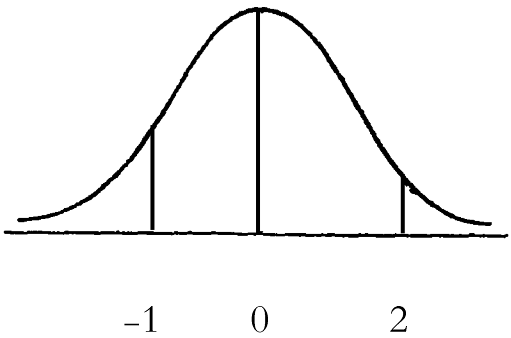
$$= 0.5 - 0.3413 = 0.1587$$



2. احتمــــال أن يكـــون متوســـط الـــزمن اللـــازم لـــأداء الخدمـــة مـــا بين 24 , 27 ثانية

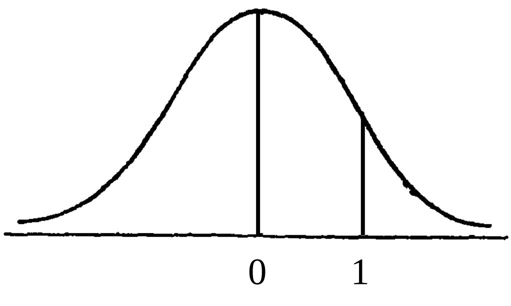
$$P(24 < \overline{X} < 27) = P\left(\frac{24-25}{5/\sqrt{25}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{27-25}{5/\sqrt{25}}\right)$$

$$= P(-1 < Z < 2)$$
$$= P(-1 < Z < 0) + P(0 < Z < 2)$$
$$= P(0 < Z < 1) + P(0 < Z < 2)$$
$$= P(0 < Z < 1) + P(0 < Z < 2)$$
$$= 0.3413 + 0.4772 = 0.8180$$



3. احتمــــال أن يكـــون متوســـط الـــزمن اللـــازم لـــأداء الخدمــــة 26 ثانيـــة أو أقل

$$P(\overline{X} \le 26) = P\left(\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \le \frac{26 - 25}{5 / \sqrt{25}}\right)$$
$$= P(Z \le 1)$$
$$= 0.5 + p(0 \le Z \le 1)$$
$$= 0.5 + 0.3413 = 0.8413$$



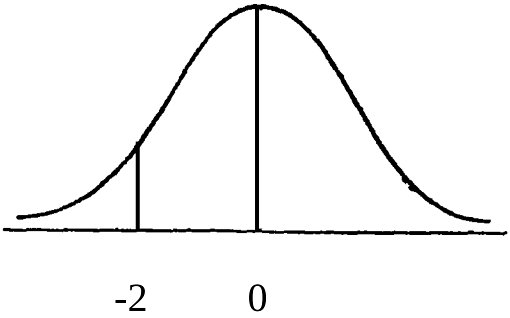
4. احتمــــال أن يكـــون متوســـط الـــزمن اللـــازم لـــأداء الخدمــــة 23 ثانيـــة على الأقل

$$P(\overline{X} \geq 23) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{23-25}{5/\sqrt{25}}\right)$$

$$= P(Z \geq -2) = 0.5 + P(-2 \leq Z \leq 0)$$

$$= 0.5 + P(0 \leq Z \leq 2)$$

$$= 0.5 + 0.4772 = 0.9772$$



المحاضرة الثانية:

نظرية المعاينة

1. مفاهيم أساسية

من أهم المفاهيم المرتبطة بالمعاينة الإحصائية ما يلي:

- **المجتمع الإحصائي:** هو مجموعة المفردات أو العناصر التي تشترك في خصائص أو خاصية محددة تميزها عن غيرها من المجتمعات.

- **الوحدة الإحصائية**: هي المفردة الأساسية التي تشكل المجتمع الإحصائي المدروس.

- **العينة:** هي جزء من المجتمع الإحصائي يتم اختيارها بطرق مختلفة بهدف تعميم النتائج المتوصل إليها من خلال دراستها على المجتمع، والذي يتحدد مدى تمثيل العينة له بعدة عوامل منها:

- حجم العينة.
- تباين خصائص المجتمع.
- طريقة اختيار العينة.

- **الإحصاءة:** هي دالة في مفردات العينة كالوسط الحسابي للعينة، انحرافها المعياري ومعامل الارتباط...الخ.

- **المعلمة:** هي مقياس يحدد خصائص المجتمع الإحصائي كالوسط الحسابي، الوسيط، معامل الارتباط...الخ.

2. المعاينة

يواجه الباحث في مختلف المجالات العلمية صعوبة في دراسة كل مشاهدة من مشاهدات المجتمع لعدة أسباب، مما يضطره لإجراء الدراسة على مجموعة جزئية من مجتمع الدراسة، وذلك بالاعتماد على أسلوب المعاينة.

1.2 تعريف المعاينة

المعاينة هي عملية اختيار جزء من مفردات المجتمع الإحصائي المدروس بطريقة علمية للاستدلال على خواصه.

2.2 مراحل تصميم خطة المعاينة

خطة المعاينة هي وصف للإجراءات المنهجية الخاصة بكيفية تنفيذ كل مرحلة من مراحل المعاينة، والمتمثلة فيما يلي:

- تحديد مشكلة وهدف الدراسة: تعد هذه المرحلة من أهم مراحل تصميم خطة المعاينة، فهي تسمح بتحديد الإطار العام للدراسة المتضمن طبيعة وحجم البيانات المطلوبة، متغيرات الدراسة، المجتمع الإحصائي المدروس وغيرها من المتغيرات.

- تحديد المجتمع الإحصائي عن طريق تحديد المتغيرات والمفردات المراد دراستها والحدود الزمنية والمكانية.

- تحديد البيانات المطلوبة، مصادرها وطريقة جمعها وفقا لموضوع الدراسة، هدفها، فرضياتها وتكاليف الحصول عليها وغيرها.

- تحديد حجم العينة الذي يحقق التوافق بين دقة المعلومات وتكاليف الدراسة، والذي يعتمد على تطبيق القواعد الإحصائية الخاصة بنظرية المعاينة وخبرة الباحث.

- جمع البيانات عن مفردات العينة المختارة، وهي مرحلة هامة لأنها عرضة لكثير من أخطاء القياس.

- مراجعة وترميز البيانات باستخدام طرق إحصائية معينة للوصول إلى نتائج ودراستها.

- تحليل البيانات وتقدير معالم المجتمع: بعد الحصول على إحصاءات العينة يتم دراستها وتحليلها للاستدلال على خصائص المجتمع.

3. توزيعات المعاينة

1.3 توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$

1.1.3 شكل توزيع المعاينة للمتوسط الحسابي $\bar{X}$

يمكن التمييز بين حالتين:

1.1.1.3 توزيع المعاينة للوسط الحسابي $\bar{X}$ عند المعاينة من مجتمع طبيعي

يمكن التمييز بين حالتين:

- **المعاينة من مجتمع طبيعي تباينه معلوم:**

نظرية:

إذا كانت $X_1, X_2, \ldots, X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ وتباينه σ^2 فإن المتغير العشوائي $\bar{X}$ يخضع للتوزيع الطبيعي متوسطه μ وتباينه $\sigma_{\bar{X}}$.

$$\bar{X} \sim N(\mu, \sigma_{\bar{X}})$$

ونكتب:

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \sim N(0,1)$$

ملاحظة: إذا كانت قيمة $n \geq 30$ فإن التقريب يكون جيدا.

مثال:

إذا كان الدخل الأسبوعي لعمال إحدى الشركات يخضع للتوزيع الطبيعي بمتوسط قدره 1200 وحدة نقدية وانحراف معياري قدره 100 وحدة نقدية، ما هو احتمال أن يكون متوسط دخل عينة حجمها $n = 64$ أكبر من 1180 وحدة نقدية في الأسبوع؟

الحل:

ليكن المتغير العشوائي X_i يمثل الدخل الأسبوعي للعمال.

$$X \sim N(\mu, \sigma)$$
$$X \sim N(1200,100)$$

لدينا:

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

ومنه:

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

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{X} > 1180) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{1180-\mu}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{1180-1200}{100/\sqrt{64}}\right)$$
$$= P(Z > -1.6) = P(Z \leq 1.6) = 0.9452$$

- **المعاينة من مجتمع طبيعي تباينه مجهول:**

من الناحية العملية غالبا ما يكون تباين مجتمع الدراسة مجهولا فيستبدل بتقديره غير المتحيز S^2 ، ويمكن التمييز بين حالتين لتحديد توزيع المتغير:

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$$

- **الحالة الأولى:** إذا كان حجم العينة $n \geq 30$ فإن المتغير $Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ يقترب توزيعه من التوزيع الطبيعي المعياري استنادا إلى نظرية النهاية المركزية.[1]

مثال:

إذا علمت أن نقاط الطلبة في مقياس الإحصاء تخضع للتوزيع الطبيعي بمتوسط قدره 11 نقطة، تم اختيار عينة حجمها 36 طالبا من هذا المجتمع فوجد أن انحرافها المعياري هو 3 نقاط .

المطلوب: حساب $P(\bar{X} > 13)$

الحل:

ليكن المتغير العشوائي X_i يمثل نقاط الطلبة في مقياس الاحصاء.

$$X \sim N(\mu, \sigma)$$
$$X \sim N(11, \sigma)$$

لدينا σ مجهول و $n \geq 30$ ومنه:

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \sim N(0,1)$$

وبالتالي يكون الاحتمال المطلوب:

---

[1] نظرية النهاية المركزية هي: مجموعة نتائج لنظرية الاحتمالات تنص أن مجموع عدة متغيرات عشوائية مستقلة ومتطابقة التوزيع يميل إلى التوزع حسب توزيع احتمالي معين.

$$P(\bar{X} > 13) = P\left(\frac{\bar{X}-\mu}{s/\sqrt{n}} > \frac{13-\mu}{s/\sqrt{n}}\right) = P\left(Z > \frac{13-11}{3/\sqrt{36}}\right) = P(Z > 4)$$

$$= 1 - P(Z \leq 4) = 1 - 0.99995 = 0.00005$$

- **الحالة الثانية:** إذا كان حجم العينة $n < 30$ فإن المتغير $T = \frac{\bar{X}-\mu}{s/\sqrt{n}}$ يتبع توزيع ستودنت

استنادا إلى النظرية التالية:

نظرية:

إذا كانت$X_1, X_2, \dots, X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ وتباينه σ^2 حيث:

- $\bar{X}$ وS^2 مستقلان.
- $\frac{(n-1)S^2}{\sigma^2} \sim \lambda^2_{n-1}$

فإن:

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$$

مثال:

تخضع أوزان أكياس الطحين الذي تنتجه إحدى المؤسسات للتوزيع الطبيعي بمتوسط قدره 50 كلغ. تم سحب عينة حجمها 10 أكياس من إنتاج هذه المؤسسة فوجد أن انحرافها المعياري هو 1 كلغ.

المطلوب: إيجاد احتمال أن يزيد متوسط العينة عن 51 كلغ.

الحل:

ليكن المتغير العشوائي X_i يمثل أوزان أكياس الطحين.

$$X \sim N(\mu, \sigma)$$
$$X \sim N(50, \sigma)$$

لدينا σ مجهول و $n < 30$ ومنه:

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{X} > 51) = P\left(\frac{\bar{X} - \mu}{s/\sqrt{n}} > \frac{51 - \mu}{s/\sqrt{n}}\right) = P\left(T > \frac{51 - 50}{1/\sqrt{10}}\right) = P(T > 3.16)$$

$$= 1 - P(T \leq 3.16) = 1 - 0.995 = 0.005$$

2.1.1.3 توزيع المعاينة للوسط الحسابي $\bar{X}$ عند المعاينة من مجتمع غير طبيعي

إذا اختيرت عينة كبيرة الحجم من مجتمع غير طبيعي فإن توزيع المعاينة للوسط الحسابي $\bar{X}$ يحدد استنادا إلى نظرية النهاية المركزية.

نظرية:

إذا كانت المتغيرات $X_1, X_2, \ldots, X_n$ متتالية من المتغيرات العشوائية المستقلة المتطابقة التوزيع بمتوسط μ وتباينه σ^2 فإن توزيع المتغير العشوائي $Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$ يقترب من التوزيع الطبيعي كلما اقتربت قيمة n من ∞.

مثال:

إذا سحبت عينة عشوائية حجمها 36 وتباينها 9 من مجتمع متوسطه 7 . ما هو احتمال أن يفوق الوسط الحسابي للعينة 8 ؟

الحل:

لدينا σ مجهول و $n \geq 30$ ومنه:

$$Z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim N(0,1)$$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{X} > 8) = P\left(\frac{\bar{X} - \mu}{s/\sqrt{n}} > \frac{8 - \mu}{s/\sqrt{n}}\right)$$

$$= P\left(Z > \frac{8-7}{3/\sqrt{36}}\right) = P(Z > 2)$$

$$= 1 - P(Z \leq 2) = 1 - 0.9772 = 0.0228$$

المحاضرة الثالثة:

نظرية التقدير

تمهيد

يلعب الإحصاء الإستدلالي دورا مهما في عملية اتخاذ القرارات في عدة ميادين، بحيث يهتم بإيجاد أو تقدير بعض الخصائص أو معلمات المجتمع محل الدراسة (المتوسط m، الإنحراف المعياري **σ**، نسبة صفة معينة في المجتمع P، ...)، لذلك غالبا ما تكون هذه المعلمات مجهولة و نرغب في تقديرها انطلاقاً من العينة. إن صعوبة الوصول إلى كل المجتمع بسبب ضيق الوقت و التكاليف الباهظة للبحث، يحتم على المؤسسة اللجوء إلى حساب ما يقابل معلمات المجتمع في العينة التي يشترط أن تكون ممثلة تمثيلاً صادقاً للمجتمع. تعتبر العينة صورة مصغرة من المجتمع، بحيث يمكن استخدام قيمة معلماتها كتقدير للمعلمة المناظرة لها في المجتمع، كأن يتم استخدام $\bar{x}$ لتقدير m و σ' لتقدير σ و هكذا.

1. تعريف نظرية التقدير

يعرف بأنه الطريقة العلمية التي تنطلق من الخاص إلى العام، أي الانطلاق من الدليل لإظهار نتيجته، بحيث يكون فيها استنباط القوانين من استنطاق الواقع. كما يعرف بأنه العلم الذي يدخل ضمن الإحصاء الاستدلالي الذي يهتم بمعالجة الإشارات التي تتعامل مع تقدير القيم من المعلَمات بناء على بيانات مقاسة أو تجريبية متعلقة بالعينة التي تم اختيارها عشوائياً من المجتمع المراد دراسته. المعلَمات تصف خصائص المجتمع وفقاً لتوزيع البيانات المقاسة، لذلك المقدّر يحاول أن يضع قيم تقريبية للمَعلَمات المجهولة للمجتمع، باستخدام قياسات (معلمات) مستنبطة من العينة المدروسة. ضمن نظرية التقدير، يفترض أن البيانات المقاسة تكون عشوائية مع توزيع احتمالي غالباً ما يكون طبيعي، مرتكزاً على مَعلَمات العينة التي اختيرت بالصدفة. على سبيل المثال، في الدراسات التسويقية، غالبا ما ترتبط القياسات التي تحتوي على معلومات حول محفزات الشراء بالآراء التي يدلي بها زبائن العينة المدروسة (التي تم اختيارها بالصدفة). فبدون عشوائية المشكلة تكون حتمية و لن تكون هناك حاجة تقدير.

2. معالم جودة التقدير

عندما نكون بصدد تقدير معلمة من معالم مجتمع محل دراسة، غالباً ما نحتاج إلى اختيار المعلمة ($\bar{x}$،σ'، ..) المناسبة في العينة لتقدير معلمة المجتمع (m،σ، ..). تسمى المعلمة المستخدمة في عملية التقدير بـ"**المقدر**". يتميز المقدر بمجموعة من الخصائص نذكر منها:

2.1 عدم التحيز: متوسط العينة $\bar{x}$ بأنه متحيز لمتوسط المجتمع m لأن $E(\bar{x})\neq m$

2.2 فعالية: تشير إلى مقدار التباين الفعال لتوزيع المعاينة للإحصائية، التقدير الذي يجب اختياره يكون تباينه هو الأصغر.

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

للقيام بالتقدير يمكن للباحث اللجوء إلى طريقتين، الأولى التقدير بقيمة المتوسط، النسبة أو بالانحراف المعياري، الذي يسمى بالتقدير النقطي المعروف ببساطته و الذي لا يعطي نتائج دقيقة، و الثانية التقدير بمجال الثقة الذي يعتبر رائج الاستخدام.

3. التقدير النقطي

يمكن لقيمة وحيدة من العينة ($\bar{X}$) أن تستخدم لتقدير المعلمة المقابلة لها في المجتمع (m) و هذا ما يعرف بالتقدير النقطي. كأن نقول: بما أن متوسط رقم الأعمال اليومية لمحل معين هو $\bar{X}$=100000 فإن متوسط رقم أعمال كل المحلات المتواجدة بنفس المنطقة هو m=100000.

4. التقدير بفترة (مجال الثقة)

التقدير بنقطة نادراً ما يساوي المعلمة التي المراد تقديرها، لذلك يجب تحديد فترة أو مجال يحتوي على مجموعة من القيم تتضمن قيمة معلمة المجتمع، تسمى هذه الفترة بفترة الثقة. احتمال وقوع المعلمة في هذا المجال يسمى بدرجة الثقة و يرمز لها بـ [1-α]، و مكمل هذه القيمة يسمى بمستوى **المعنوية** و يرمز له بـ [α]. مثلاً إذا كانت درجة الثقة 95% يكون مستوى المعنوية 0.05. كما يشار لهذا الأخير بأنه مستوى خطأ التقدير.

فإذا كان أمامنا تقدير المعلمة m بفترة ثقة، و إيجاد مجال من الشكل $[m_1 - m_2]$ بحيث تنتمي هذه المعلمة إلى هذا المجال باحتمال معين، يسمى هذا الاحتمال **بقياس الثقة** (1 - α) يعبر عنه بالمعادلة التالية:

$$\mathbf{P\,[m \in (m_1 - m_2)] = P\,(m_1 \leq m \leq m_2) = 1 - \alpha}$$

بحيث أن m_1, m_2: تعبر عن حدود مجال الثقة أي الحد الأدنى و الحد الأقصى.

5. درجة و معامل الثقة

كما رأينا في الدرس الأول بأنه عندما تتوزع x توزيعاً طبيعياً، فإنه من بين خصائص المنحنى الطبيعي أن: **68.26%** من البيانات محصورة بين $[\bar{x} \mp 1.\sigma_x]$، و **95.44%** بين $[\bar{x} \mp 2.\sigma_x]$، و أن **99.72%** من البيانات محصورة بين $[\bar{x} \mp 3.\sigma_x]$، بحيث يمثل x المتغير و الانحراف (الخطأ) المعياري. بنفس المبدأ إذا كان لدينا مجتمعاً إحصائياً يتوزع توزيعاً طبيعياً، و سحبنا منه عينات كبيراً نسبياً لها نفس الحجم $(n_i \geq 30)$، فإن توزيع المعاينة للوساط للوساط الحسابي للمجتمع **(m)** يكون له هو الآخر توزيعاً طبيعياً، بحيث يرتكز تقديره على الوسط الحسابي للعينة $(\bar{X}_i)$ ، بمعنى أن تقدير المتوسط الحسابي للمجتمع يساوي المتوسط الحسابي للعينة ∓ الانحراف المعياري لمتوسط للعينة.

لو قمنا بالتعبير عن هذا التقدير بالرموز لوجدنا ما يلي:

$$P[m = \bar{x} \mp 1.\sigma_{\bar{x}}] = 0.6826$$

$$P[m = \bar{x} \mp 2.\sigma_{\bar{x}}] = 0.9544$$

$$P[m = \bar{x} \mp 3.\sigma_{\bar{x}}] = 0.9972$$

بحيث أن:

- **P:** يعبر عن الاحتمال؛
- **m:** الوسط الحسابي المقدر للمجتمع؛
- **1، 2، و 3:** معامل التقدير أو $t\alpha$ **الجدولية؛**
- **m:** الخطأ المعياري للمتوسط الحسابي
- **درجة الثقة:** 0.6826، 0.9544، 0.9972؛

6. تقدير الوسط الحسابي للمجتمع بفترة ثقة:

ليكن لدينا مجتمع إحصائي يتوزع توزيع طبيعي [أي $N(m, \sigma)$]، وسطه الحسابي m مجهول و انحرافه المعياري σ معلوم، سحبنا منه عينة عشوائية ذات حجم n، عناصرها على التوالي: x_1، x_2،...، x_n. ، وسطها الحسابي $\bar{X}$ وانحرافها المعياري σ'. لا يمكننا إيجاد قيمة محددة لمتوسط المجتمع m بل نحاول حصر قيمتها داخل مجال معين بين m_1 و m_2 يسمى هذا المجال المحصور بفترة الثقة، وذلك يكون تحت احتمال P، بحيث أن هذا الأخير يساوي $1-\alpha$، وتكتب على النحو الآتي:

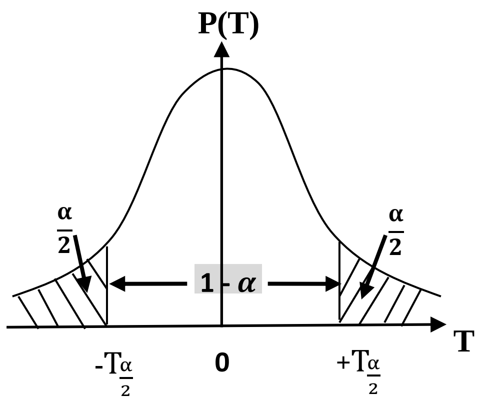
إذا كان لدينا: $(n \leq 0.05N)$ $\quad N \rightarrow +\infty$

فإن: $P / m \in [m_1 - m_2] = 1 - \alpha$

ليكن $\bar{X}$ متغير عشوائي لأي عينة غير حصرية (شاملة) ذات حجم n أكبر من 30 مرفقة لمتوسط هذه العينة، $\bar{X}$ يأخذ بصفة متتالية القيم التالية: $\bar{X}_1$، $\bar{X}_2$، ....

نفرض أن كل الشروط متوفرة لاستخدام نتائج نظرية النهاية المركزية والقيام بالتقريبات التالية:

المتغير العشوائي يتبع القانون الطبيعي $N(m, \frac{\sigma}{\sqrt{n}})$ بمعنى المتغير العشوائي $T = \frac{\bar{X}-m}{\frac{\sigma}{\sqrt{n}}}$ يتبع القانون الطبيعي الموسط المختصر $N(0, 1)$.



$$\bar{X} \sim N(m, \frac{\sigma}{\sqrt{n}}), \quad T \sim N(0, 1), \quad T = \frac{\bar{X}-m}{\frac{\sigma}{\sqrt{n}}}$$

$$\Leftrightarrow P(-T_{\frac{\alpha}{2}} \leq T \leq +t_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$\Leftrightarrow P(-T_{\frac{\alpha}{2}} \leq \frac{\bar{X}-m}{\frac{\sigma}{\sqrt{n}}} \leq +T_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$\Leftrightarrow P(-\frac{\sigma}{\sqrt{n}} T \leq \bar{X} - m \leq +\frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$\Leftrightarrow P(\bar{X} - \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}}) = 1 - \alpha$$

إذاً مجال تقدير متوسط المجتمع **m** هو: $P(\bar{X} - \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}}) = 1 - \alpha$

نقول أن $[\bar{X} - \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}} ، \bar{X} + \frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}}]$ هو مجال الثقة لـm مع معامل الثقة $1 - \alpha$، مركزه هو المتوسط $\bar{X}$ للعينة التي تم اختيارها بالصدفة. و نلاحظ أننا طرحنا القيمة $\frac{\sigma}{\sqrt{n}} T_{\frac{\alpha}{2}}$ على يسار المتوسط $\bar{X}$ كما أننا أضفنا تلك القيمة على اليمين.

σ	الحالة الأولى: $n \geq 30$	الحالة الثانية: $n < 30$ نفرض أن المجتمع يخضع إلى التوزيع الطبيعي
معلومة	نقرأ $T_{\frac{\alpha}{2}}$ من الجدول(**02**) و نعوض في معادلة التقدير	نقرأ $T_{\frac{\alpha}{2}}$ من الجدول (02) و نعوض في معادلة التقدير
مجهولة	نقرأ $T_{\frac{\alpha}{2}}$ من الجدول (02) ؛ نقرأ أو نقدر σ بـ S بحيث أن: **تباين المجتمع**: $S^2 = \frac{\sum(x_i-\bar{X})^2}{n-1}$ **وتباين العينة**: $S'^2 = \frac{\sum(x_i-\bar{X})^2}{n}$	نقرأ **σ** بـ **S** بحيث أن: **المجتمع: التباين**: $\sigma^2=S^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$ **الإنحرف المعياري**: σ=S **العينة:التباين**: $\sigma'^2=S'^2=\frac{\sum(x_i-\bar{x})^2}{n}$ **الإنحرف المعياري**: σ'=S' نقرأ $T_{\frac{\alpha}{2}}$ من الجدول (03) و نعوض في معادلة التقدير
ملاحظة: **الجدول(2) هو** جدول القانون الطبيعي الموسط المختصر الذي يدعى بالفرنسية بـ la Loi Normale Centrée Réduite		

المثال1: تقوم مؤسسة معينة بإنتاج نوع من البطاريات، بحيث أن متوسط مدة اشتغال 50 بطارية من تلك التي أنتجتها هو 1000 سا. إذا علمت أن الإنحراف المعياري لمدة إشتغال البطاريات التي تنتجها المؤسسة هو 40 سا. قدر مستوى إشتغال هذه البطاريات بفترة ثقة (قياس الثقة 95%).

الحل: $\bar{X}=1000$ h, $n = 50$, $\sigma = 40$ h, $1-\alpha = 95\%$

$$\bar{X} \rightarrow N\left(m, \frac{\sigma}{\sqrt{n}}\right), \quad T \rightarrow N(0, 1), \quad T = \frac{\bar{X}-m}{\frac{\sigma}{\sqrt{n}}}$$

$$\mathbf{P\left(-t_{\frac{\alpha}{2}} \leq T \leq +t_{\frac{\alpha}{2}}\right) = 1-\alpha}$$

$$\mathbf{P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} t_{\frac{\alpha}{2}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{n}} t_{\frac{\alpha}{2}}\right) = 1-\alpha}$$

من الجدول (02): $t_{\frac{\alpha}{2}} = \mathbf{1.96}$ لأن $\alpha = 5\%$ و $1-\alpha = 95\%$

$$P\left(1000 - \frac{40}{\sqrt{50}}(1.96) \leq m \leq 1000 + \frac{40}{\sqrt{50}}(1.96)\right) = 0.95$$

$$\mathbf{P(988,90 \leq m \leq 1071,08) = 95\%}$$

المثال2: متوسط الدخل الفردي الشهري لعينة مكونة من 80 مواطن يقطنون بحي معين هو 32000 ون، بانحراف معياري (للعينة) قدره 600 ون. قدر متوسط الدخل الفردي الشهري لسكان هذا الحي بفترة ثقة. (قياس الثقة بـ 98%)

الحل: $n = 80$, $\bar{X}=12000$, $1-\alpha = 98\%$, $\sigma' = 600$, $\sigma = ?$

$$\bar{X} \rightarrow N\left(m, \frac{\sigma}{\sqrt{n}}\right), \quad T \rightarrow N(0, 1), \quad T = \frac{\bar{X}-m}{\frac{\sigma}{\sqrt{n}}}$$

$$\mathbf{P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} t_{\frac{\alpha}{2}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{n}} t_{\frac{\alpha}{2}}\right) = 1-\alpha}$$

σ **للمجتمع مجهولة لذلك تقدرها بـ S**

نعلم أن: $\sum(x_i-\bar{x})^2 = \sigma^2.n$ إذن $\mathbf{S^2} = \frac{\sum(x_i-\bar{x})^2}{n-1}$

$$\sigma^2 = \frac{\sum(x_i-\bar{X})^2}{n}, \quad S^2 = \frac{n.\sigma^2}{n-1} \Rightarrow S^2 = \frac{80.(600)^2}{80-1} = 603.78 \Rightarrow$$

من الجدول.2: $t_\alpha = \mathbf{2.32}$ لأن $\alpha = 2\%$ و $1-\alpha = 98\%$

$$P\left(12000 - \frac{603.78}{\sqrt{80}}(2.32) \leq m \leq 12000 + \frac{603.78}{\sqrt{80}}(2.32)\right) = 0.98$$

$$P(988,90 \leq m \leq 1071,08) = 95\%$$

المثال.4: **الدخل الشهري للعائلات**

متوسط الدخل الشهري العائلي لعينة مكونة من 80عائلة (سحبت بدون إرجاع) متواجدة بحي معين هو 32000دج، بانحراف معياري قدره 600دج. إذا علمت أنه يسكن بهذا الحي 945 عائلة، قدر متوسط الدخل الشهري لعائلات هذا الحي بفترة ثقة. (قياس الثقة بـ 98%).

الحل: $\bar{X}$ =32000 DA, **n** = 80, σ = ?**σ'** = 600 DA, **1 – α** = 98%

بما أن $n=80 > [0.05(945)=47.25]$ إذا نقوم بتعديل الانحراف المعياري بضربه بـ $\sqrt{\frac{N-n}{N-1}}$.

$$\mathbf{P(-T_{\frac{\alpha}{2}} \leq T \leq +T_{\frac{\alpha}{2}}) = 1 - \alpha}$$

$$\mathbf{P(\bar{X} - \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.T_{\frac{\alpha}{2}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.T_{\frac{\alpha}{2}}) = 1 - \alpha}$$

σ للمجتمع مجهولة لذلك نقدرها بـS، **كذلك** نعلم أن تباين العينة: $\sigma'^2 = S'^2 = \frac{\sum(x_i-\bar{x})^2}{n}$

ومنه $\sum(x_i - \bar{x})^2 = \sigma'^2.n = (600)^2.80 = 28800000$

– نعلم كذلك أن تباين المجتمع: $\sigma^2 = \frac{28800000}{80-1} = 364556.96 = \frac{\sum(x_i-\bar{x})^2}{n-1}$

ومنه الانحراف المعياري للمجتمع يساوي: **σ=603.78**

من جدول (02) التوزيع الطبيعي المركز المختصر: $T_{\frac{\alpha}{2}} = 2.326$ لأن α = 2% و 1- α = 98%

$$P(32000 - \frac{603.78}{\sqrt{80}}\sqrt{\frac{945-80}{945-1}}(2.326) \leq m \leq 32000 + \frac{603.78}{\sqrt{80}}\sqrt{\frac{945-80}{945-1}}(2.326)) = 0.98$$

$$\mathbf{P(31849.78 \leq m \leq 32150.296) = 98\%}$$

الاستنتاج: متوسط الدخل الشهري لعائلات هذا الحي محصور بين 31849.78ون و 32150.296ون و هذا بنسبة 98%.

7. تقدير النسبة في المجتمع بفترة ثقة

نسبة عناصر المجتمع التي تتميز بخاصية معينة هي **P**، نسحب منه عينة ذات حجم **n**، نسبة عناصرها التي تتميز بالخاصية هي $\hat{p}$، تقدير نسبة المجتمع **P** بمجال الثقة باستعمال نسبة العينة $\hat{p}$ يتم كالآتي:

$$\mathbf{T \sim N(0, 1),\ \hat{P} \sim N(P, \sqrt{\frac{p.q}{n}})} \qquad T = \frac{\hat{P}-P}{\sqrt{\frac{p.q}{n}}}$$

$$P(-t_{\frac{\alpha}{2}} \leq T \leq +t_{\frac{\alpha}{2}}) = 1-\alpha \iff P(-t_{\frac{\alpha}{2}} \leq \frac{\hat{P}-P}{\sqrt{\frac{p.q}{n}}} \leq +t_{\frac{\alpha}{2}}) = 1-\alpha$$

$$\iff P(-\sqrt{\frac{p.q}{n}}\, t_{\frac{\alpha}{2}} \leq \hat{P} - P \leq +\sqrt{\frac{p.q}{n}}\, t_{\frac{\alpha}{2}}) = 1-\alpha$$

ملاحظة هامة: نعوض **P** (الموجود تحت الجذر التربيعي) بـ $\hat{P}$ في جدول المتراجحة نجد:

$$P(\hat{P} - \sqrt{\frac{p.q}{n}}\, t_\alpha \leq P \leq \hat{P} + \sqrt{\frac{p.q}{n}}\, t_\alpha) = 1-\alpha \qquad \mathbf{P(\hat{P} - \sqrt{\frac{\hat{p}.\hat{q}}{n}}\, t_{\frac{\alpha}{2}} \leq P \leq \hat{P} + \sqrt{\frac{\hat{p}.\hat{q}}{n}}\, t_{\frac{\alpha}{2}}) = 1-\alpha}$$

حيث أن t_α تقرأ من جدول.2 للتوزيع الطبيعي و n ≥30، و [q=1-p]

مثال.5: ضمن عينة مكونة من 80 عائلة في مدينة معينة، وجدنا 45 عائلة تملك سيارة. قدر نسبة العائلات في هذه المدينة التي تملك سيارات بفترة ثقة (قياس الثقة 98%).

الحل: $n=80$، $P=\frac{45}{80}=0.56$، $1-\alpha=0.98$

$$P\left(\hat{P}-\sqrt{\frac{\hat{p}.\hat{q}}{n}}\,t_{\frac{\alpha}{2}}\le P\le \hat{P}+\sqrt{\frac{\hat{p}.\hat{q}}{n}}\,t_{\frac{\alpha}{2}}\right)=1-\alpha$$

حيث أن $t_{\frac{\alpha}{2}}$ تقرأ من جدول.2 للتوزيع الطبيعي و $n\ge 30$: $t_{\frac{\alpha}{2}}=2.32$ $\alpha=0.02$ $1-\alpha=0.98$

$$P\left(0.56-\sqrt{\frac{0,56.0.44}{80}}(2.32)\le P\le 0.56+\sqrt{\frac{(0.56)(0.44)}{80}}(2.32)\right)=98\%$$

$$P(0.43\le P\le 0.68)=98\%$$

8. تقدير الفرق للوسطين الحسابين للمجتمعين بفترة ثقة

نعتبر مجتمعين الأول وسطه الحسابي m_1 و انحرافه المعياري σ_1، و الثاني وسطه الحسابي m_2 و انحرافه المعياري σ_2. نسحب من الأولى عينة ذات حجم n_1 و وسطها الحسابي $\bar{X}_1$، و نسحب من الثانية عينة ذات حجم n_2 و وسطها الحسابي $\bar{X}_2$. عندما يطلب منا إيجاد فترة الثقة للفرق بين متوسطي المجتمعين أي فترة الثقة لـ m_1-m_2، نفرض أن حجم المجتمع **الأول** يؤول إلى ∞ و نفس الشيء بالنسبة للمجتمع **الثاني** ثم نكتب ما يلي:

$$P\left(\bar{X}_1-\bar{X}_2-\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\,t_{\frac{\alpha}{2}}\le (m1-m2)\le \bar{X}_1-\bar{X}_2+\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\,t_{\frac{\alpha}{2}}\right)=1-\alpha$$

σ_1 و σ_2	الحالة الأولى: $n\ge 30$	الحالة الثانية: $n<30$ نفرض أن المجتمع يخضع إلى التوزيع الطبيعي
معلومتين	نقرأ من الجدول$t_{\frac{\alpha}{2}}$ (**02**) و نعوض في معادلة التقدير	نقرأ $t_{\frac{\alpha}{2}}$ من الجدول (02) و نعوض في معادلة التقدير
مجهولتين	نقرأ $t_{\frac{\alpha}{2}}$ من الجدول (02) ؛ نقدر σ_1 بـ S_1 و σ_2 بـ S_2 بحيث أن: تباين المجتمع: $S^2=\frac{\sum(x_i-\bar{X})^2}{n-1}$ وتباين العينة: $S'^2=\frac{\sum(x_i-\bar{X})^2}{n}$	نفرض أن للمجتمعين نفس التباين بحيث أننا **نقرأ** (أو **نقدر**) σ_1^2 و σ_2^2 **بالتباين مشترك** S^{*2} $S^{*2}=\frac{\sum(x_{1i}-\bar{X}_1)^2+\sum(x_{2i}-\bar{X}_2)^2}{n_1+n_2-2}$ **نقرأ من الجدول.3** (درجة الحرية $dl=n_1+n_2-2$)
ملاحظة: الجدول(2) هو جدول القانون الطبيعي الموسط المختصر الذي يدعى بالفرنسية بـ la Loi Normale Centrée Réduite		

مثال.4: متوسط أجور عينة من 24 عامل يشتغلون في مؤسسة أ هو 32000 دج بانحراف معياري (للعينة) قدره 900 دج و متوسط أجور عينة من 22 عامل يشتغلون في المؤسسة ب هو 28000 دج بانحراف معياري (للعينة) قدره 600 دج.

– قدر الفرق بين متوسطي أجور العمال في المؤسسة بفترة ثقة (قياس الثقة **95%**).

الحل: $n_1=24<30$، $\bar{X}_1=32000$، $\sigma'_1=900$ $n_2=22<30$، $\bar{X}_2=28000$، $\sigma'_2=600$

تخضع أجور العمال إلى توزيعين طبيعيين لهما نفس التباين. **σ_1 و σ_2 مجهولين**: نفرض أن للمجتمعين نفس التباين و **نقدر** σ_1^2 و σ_2^2 بـ S^{*2}.

$$S^{*2}=\frac{\sum(x_{1i}-\bar{X}_1)^2+\sum(x_{2i}-\bar{X}_2)^2}{n_1+n_2-2}$$

$$\sigma_1^2=\frac{\sum(x_{1i}-\bar{X}_1)^2}{n_1} \quad \sum(x_{1i}-\bar{X}_1)^2=\sigma_1^2.n_1=900^2.24=\mathbf{19440000}$$

$$\sigma_2^2=\frac{\sum(x_{2i}-\bar{X}_2)^2}{n_2} \quad \sum(x_{2i}-\bar{X}_2)^2=\sigma_2^2.n_2=600^2.22=\mathbf{7920000}$$

$$S^{*2}=\frac{19440000+7920000}{24+22-2}=\mathbf{621818.18}$$

$$\left.\begin{array}{l}\mathbf{1-\alpha}=0.95 \quad \boldsymbol{\alpha}=\mathbf{0.05}\\ \mathbf{dl}=25+20-2=\mathbf{43}\end{array}\right\}$$

عندما ننظر في الجدول.3 سنجد $\Longrightarrow$ $\mathbf{t_\alpha=1.95}$

$$P\left(\bar{X}_1-\bar{X}_2-\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ t_{\frac{\alpha}{2}}\le m_1-m_2\le\bar{X}_1-\bar{X}_2+\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ t_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$P\left(4000-\sqrt{\frac{621818.18}{24}+\frac{621818.18}{22}}.(1.95)\le m_1-m_2\le 4000+\sqrt{\frac{621818.18}{24}+\frac{621818.18}{22}}.(1.95)\right)=95\%$$

$$\mathbf{P(3546.14\le m_1-m_2\le 4453.86)=0.95}$$

9. تقدير الفرق بين نسبتين للمجتمعين بفترة ثقة:

لنفرض أنه لدينا مجتمعين يتميز كل واحد منهما على التوالي بالخاصية P_1 و P_2. نسحب من المجتمع الأول عينة ذات حجم n_1 بحيث أن نسبة الخاصية فيها هي $\widehat{P}_1$، و نسحب من المجتمع الثاني عينة ذات حجم n_2 و نسبة الخاصية فيها هي $\widehat{P}_2$. عندما يطلب منا إيجاد فترة الثقة للفرق بين متوسطي المجتمعين أي فترة الثقة لـ m_1 – m_2، نفرض أن حجم المجتمع **الأول** يؤول إلى ∞ و نفس الشيء بالنسبة للمجتمع **الثاني** ثم نكتب ما يلي

$$P\left(\widehat{P}_1-\widehat{P}_2-\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n_1}+\frac{\hat{p}_2.\hat{q}_2}{n_2}}\ t_{\frac{\alpha}{2}}\le P_1-P_2\le\widehat{P}_1-\widehat{P}_2+\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n_1}+\frac{\hat{p}_2.\hat{q}_2}{n_2}}\ t_{\frac{\alpha}{2}}\right)=1-\alpha$$

مثال.6: اقترحت صيغة جديدة لبيع السيارات، فكانت نسبة الموافقين على هذه الصيغة في عينة مكون من 70 شخص يسكنون بمدينة "أ" هي 68%، و كانت نسبة الموافقين على نفس الصيغة في عينة مكونة من 95 شخص في المدينة "**ب**" هي 62%.

– قدر الفرق بين نسبتي الموافقين على تلك الصيغة بين المدينتين "أ" و "**ب**" بفترة ثقة. (قياس الثقة 95%)

الحل $\mathbf{n_1=70,\quad P_1=0.68\ (q_1=0.32),\quad n_2=95,\quad P_2=0.62\ (q_2=0.38)}$

حيث أن $\mathbf{t_{\frac{\alpha}{2}}}$ تقرأ من جدول.**2**: $t_{\frac{\alpha}{2}}=1.96 \quad \alpha=0.05 \quad 1-\alpha=0.95$

$$P\left(\widehat{P}_1-\widehat{P}_2-\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n_1}+\frac{\hat{p}_2.\hat{q}_2}{n_2}}\ t_{\frac{\alpha}{2}}\le P_1-P_2\le\widehat{P}_1-\widehat{P}_2+\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n_1}+\frac{\hat{p}_2.\hat{q}_2}{n_2}}\ t_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$P\left(0.68-0.62-\sqrt{\frac{(0.68).(0.32)}{70}+\frac{(0.62).(0.38)}{95}}(1.96)\le P_1-P_2\le 0.68-0.62+\sqrt{\frac{(0.68).(0.32)}{70}+\frac{(0.62).(0.38)}{95}}(1.96)\right)=95\%$$

$$\mathbf{P(-0.086\le P_1-P_2\le 0.2062)=95\%}$$

المحاضرة الرابعة:

اختبار الفرضيات

اختبار الفرضيات هي طريقة إحصائية لاتخاذ القرارات باستخدام البيانات، هذه الأخيرة قد تكون تجربتها خاضعة للسيطرة أو من دراسة رصدية. في هذا المجال غالباً ما نستخدم مصطلح "الدلالة الإحصائية"، و قد قام العالم الأسطوري "Ronald Fisher" و لأول مرة بصياغة مصطلح "اختبار الدلالة". هذه الأخيرة تعني أن نتيجة التجربة من الغير المتوقع أن تكون قد حدثت بالصدفة، و أن لها عتبة احتمالية تم تحديدها مسبقاً تسمى بـ"درجة الدلالة"، أو بعبارة أخرى احتمال خطأ التقدير (**α أو β**) الذي يعبر عنه باحتمال معين يشار إليه بـ: "P-value"، يتواجد بموقع يسمى بالمنطقة الحرجة.

عندما نكون أمام اختبار فرضية إحصائية و تكون نتيجة تجربتها داخل المنطقة الحرجة فإنه يتم رفض الفرضية العدمية و قبول الفرضية البديلة. فالفرضية الإحصائية هي في حد ذاتها ادعاء أو تصريح أو إقتراح يتعلق بمعلمة واحدة أو مجموعة من المعلمات بحيث أن هذا الإدعاء يمكن أن يكون صحيح أو خاطئ. يضع الباحثين فرضية تسمى بالفرضية **العدمية** $[H_0]$ مقابل فرضية أخرى تسمى بالفرضية **البديلة** $[H_1]$. الفرضية العدمية يضعها الباحث دائما على أمل رفضها و بالتالي إلغائها. إن قبول فرضية لا يعني أنها صحيحة و إنما عدم وجود أدلة كافية من العينة تدفع لرفضها أدى إلى قبولها، بالمقابل فإن رفض الفرضية لا يعني أنها خاطئة، لكن عدم وجود أدلة كافية من العينة تدفع إلى قبولها أدى إلى رفضها.

مثال: ينتج مصنع معين نوع من القطع الميكانيكية، أراد زبون إمضاء صفقة مع هذا المصنع من خلال شراءه لعدد كبير من القطع التي يقدر طولها بـ **1** سم. يقبل الزبون منتج هذا المصنع إذا كانت نسبة القطع الفاسدة التي يقل طولها عن **1** سم هي P=%5. عملياً يستحيل مراقبة قياسات كل القطع، لذلك يجب سحب عينة و مراقبة أطوال القطع المكونة لها ثم القيام باختبار:

$$H_0: \ P = 5\%$$
$$H_1: \ P < 5\%$$

1. الخطأ من النوع (1) و الخطأ من النوع (2)

- قد يقع الباحث في خطأ بسبب قبوله الفرضية $[H_1]$ بينما تكون الفرضية الصحيحة هي $[H_0]$. احتمال وقوع هذا نسميه بالخطأ من النوع (**1**) و نرمز له بـ **α**، يشار إليه بالعبارة التالية: $[\alpha = P(H_1/H_0)]$
- و قد يقع الباحث في خطأ لكونه اختار الفرضية $[H_0]$ بينما تكون الفرضية الصحيحة هي $[H_1]$. احتمال وقوع هذا نسميه بالخطأ من النوع (**2**) و نرمز له بـ **β**، يعبر عليه كما يلي: $[\beta = P(H_0/H_1)]$

الفرضية المختارة		الفرضية الصحيحة H_1	الفرضية الصحيحة H_0
	H_0	الخطأ من النوع الثاني (β)	مستوى الثقة لا يوجد خطأ (1-α)
	H_1	لا يوجد خطأ (1-β)	الخطأ من النوع الأول (α)

2. المُخْتَبَرْ الإحصائي

هو متغير عشوائي نحسب قيمه من العينات و على أساس قيمته يتم قبول فرضية و رفض الأخرى.

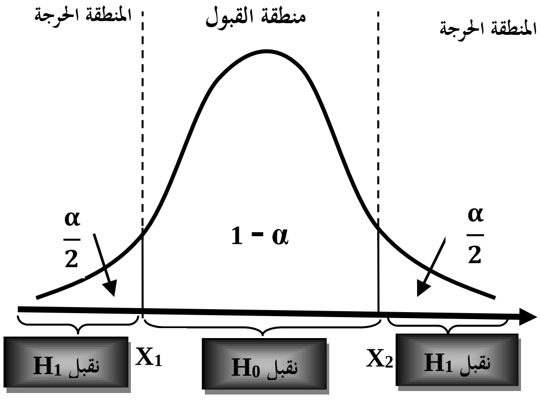
يرمز له بـ **T'** أو **Z'**

في حالة الأوساط الحسابية $t' = \dfrac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}}$

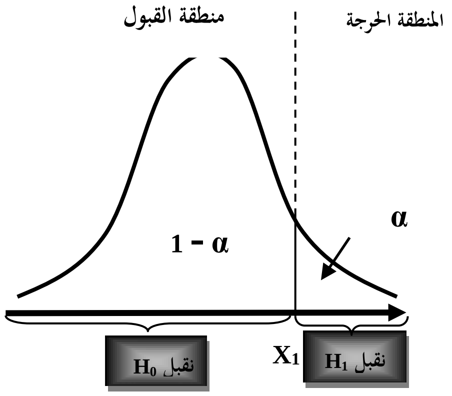
في حالة النسب $t' = \dfrac{\hat{P} - P}{\sqrt{\frac{p.q}{n}}}$

3. المنطقة الحرجة

هي مجال أو اتحاد مجالات بحيث إذا كانت قيمة المختبر الإحصائي تنتمي إلى هذه المنطقة يقبل الباحث الفرضية البديلة $[H_1]$ و يرفض الفرضية العدمية $[H_0]$. لذلك المنطقة الحرجة هي منطقة رفض الفرضية العدمية.



شكل.1: منطقة القبول محصورة بين X1 و X2



شكل.2: منطقة القبول أقل من X1

منطقة رفض أو عدم رفض الفرضية العدمية $[H_0]$ (Région critique).

من خلال تحديدنا لمستوى الدلالة α و اختيار $[H_0]$ كفرضية لا نريد رفضها، و اعتمادا على المعلومات المحصل عليها من العينة نعرف المنطقتين:

– منطقة **رفض** الفرضية $[H_0]$ و هي مجموعة قيم **الإختبار إحصائي** التي من أجلها نرفض الفرضية $[H_0]$.

– منطقة **عدم الرفض** تتكون من مجموعة قيم الاختبار احصائي التي من أجلها لا نرفض الفرضية $[H_0]$ و يتعلق حجم المنطقة الأولى و الثانية بمستوى الدلالة التي نختارها للإختبار.

المنطقة (1): لا يوجد دليل على أن الشخص مخطئ $[H_0$ لا يمكن رفضها$]$؛

المنطقة (2): الدليل قاطع على أن الشخص مخطئ $[H_0$ يمكن رفضها$]$؛

ملاحظة: كلما ابتعدنا عن النقطة (0) نحو اليمين، كلما زادت قناعتنا أن الشخص مخطئ

نقطة فاصلة (C) النقطة الحرجة أو القيمة الحرجة:

– إذا كانت النقطة على يسار (C) نعلن قبول $[H_0]$ و رفض $[H_1]$. و إذا كانت النقطة على اليمين نعلن قبول $[H_1]$ و رفض $[H_0]$.

تعرف النقطة الحرجة للمختبر الاحصائي و يمكن تعيينها اعتمادا على حجم العينة، فإذا كان حجم العينة كبيراً، يمكن أن تقترب من التوزيع الطبيعي، أما إذا كانت صغيرة فتتبع توزيع T de student.

معنى مستوى الدلالة α: إلى أي مدى يقبل المخاطرة في رفض الفرضية H_0 بالرغم من صحتها (**مبدأ الخطورة الخطأ** 1)

- **الفرضية العدمية** (Hypothèse Nulle): سميت بهذا الإسم لأنها تمثل حالة عدم التغير أي $[m = m_0]$
- **الفرضية البديلة** (Hypothèse Alternative): سميت بهذا الإسم لأنها تمثل حالة التغير أي $[m \neq m_0$ إذا كانت من الجانبين و $m > m_0$ أو $m < m_0$ إذا كانت من جانب واحد$]$

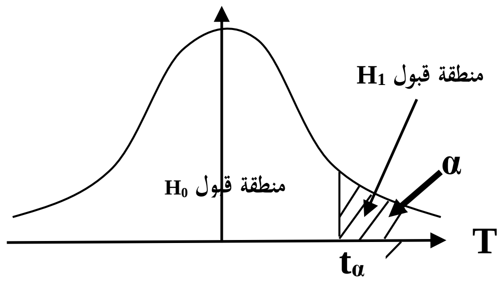
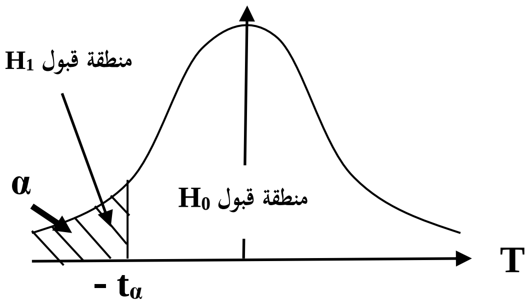
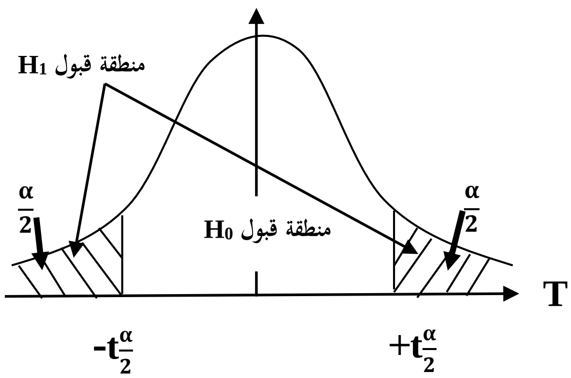
4. الخطوات اللازمة لإجراء الاختبار:

1. تحديد الفرضية العدمية $[H_0]$ و البديلة $[H_1]$؛
2. تحديد مستوى الدلالة و اختيار المختبر الإحصائي ($t_{\frac{\alpha}{2}}$ أو t_α)؛
3. تحديد مناطق الرفض و القبول في رسم بياني لمنحنى التوزيع الطبيعي
4. حساب قيمة المختبر الاحصائي الناجم عن العينة (t').
5. الإستنتاج و اتخاذ القرار.

ملاحظة:

–$(n\geq30)$: عندما تكون σ معلومة، نقرأ من جدول التوزيع الطبيعي (ج.2: Table de l'écart-réduit) و نستعمل المتغير المعيا

- $(n<30)$: عندما تكون σ غير معلومة و توزيع يكون طبيعي نستعمل توزيع Student T (ج.3: Table de T de Student)

<table>
<tr><th></th><th>اختبار الأوساط الحسابية</th><th>اختبار النسب</th></tr>
<tr><td rowspan="2">الحالة الأولى : الذيل على اليمين</td><td>$H_0 : m=m_0$
$H_1: m=m_1 \Rightarrow [m= m_1>m_0]$
-نحسب t' : $t' = \frac{\bar{X}- m_0}{\frac{\sigma}{\sqrt{n}}}$</td><td>$H_0 : P=P_0$
$H_1: P=P_1 \Rightarrow [H_1: P_1>P_0]$
نحسب t': $t' =\frac{\hat{P}- P}{\sqrt{\frac{p.q}{n}}}$</td></tr>
<tr><td>- نضرب α في 2 و من خلال نتيجتها نقرأ قيمة $+t_\alpha$ من الجدول.(2)
(على اليمين إشارة t_α تكون موجبة)
- نعتمد على قاعدة إتخاذ القرار:
نقبل H_1 و نرفض H_0 إذاكان * $t' \geq t_\alpha$
نقبل H_0 و نرفض H_1 إذاكان *$t' < t_\alpha$</td><td>
</td></tr>
<tr><td rowspan="2">الحالة الثانية: الذيل على اليسار</td><td>$H_0 : m=m_0$
$H_1: m=m_1 \Rightarrow [H_1: m_1<m_0]$
نحسب t': $t' = \frac{\bar{X}- m_0}{\frac{\sigma}{\sqrt{n}}}$</td><td>$H_0 : P=P_0$
$H_1: P=P_1 \Rightarrow [H_1: P_1<P_0]$
نحسب t': $t' =\frac{\hat{P}- P}{\sqrt{\frac{p.q}{n}}}$</td></tr>
<tr><td>* - نضرب α في 2 و من خلال نتيجتها نقرأ قيمة $-t_\alpha$ من الجدول.(2)
(على اليسار إشارة t_α تكون سالبة)
* نعتمد على قاعدة إتخاذ القرار:
نقبل H_1 و نرفض H_0 إذاكان * $t' \leq -t_\alpha$
نقبل H_0 و نرفض H_1 إذاكان *$t' > -t_\alpha$</td><td>
</td></tr>
<tr><td rowspan="2">الحالة الثالثة: الذيلين</td><td>$H_0 : m=m_0$
$H_1: m\neq m_1$
نحسب t' $t' = \frac{\bar{X}- m_0}{\frac{\sigma}{\sqrt{n}}}$</td><td>$H_0 : P=P_0$
$H_1: P\neq P_1$
نحسب t' : $t' =\frac{\hat{P}- P}{\sqrt{\frac{p.q}{n}}}$</td></tr>
<tr><td>* نقرأ قيمة $t_{\frac{\alpha}{2}}$ مباشرة من الجدول (2) من خلال قيمة α.
* نعتمد على قاعدة إتخاذ القرار:
نقبل H_1 و نرفض H_0 إذاكان * $t' \in]-\infty, -t_{\frac{\alpha}{2}}] \cup [t_{\frac{\alpha}{2}}, +\infty[$
نقبل H_0 و نرفض H_1 إذاكان * $t' \in]-t_{\frac{\alpha}{2}}, +t_{\frac{\alpha}{2}}[$</td><td>
</td></tr>
</table>

مثال.1: يعلم صاحب شركة لصناعة السيارات من خبراته السابقة أن المدة الزمنية المتوسطة التي تستغرقها صناعة كل سيارة على حدا هي **8** سا بانحراف معياري قدره **2** سا، و للتأكد من ذلك قام بسحب (بالإرجاع) عينة مكونة من **40** سيارة فوجد أن متوسط الفترة الزمنية لصناعتها هي **10** ساعات، ثم قام بوضع الفرضيتين التاليتين تحت مستوى معنوية **5%:** * أي من الفرضيتين تختار الشركة الفرضية العدمية (**H_0**) أو البديلة (**H_1**)؟

الحل:

$H_0: m = m_0 = \mathbf{8\ h}$

$H_1: m = m_1 = \mathbf{10\ h} \quad m_1 > m_0$

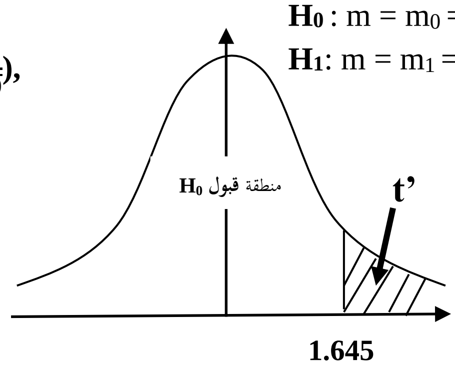
$\bar{X}=10$ ، $m_0 = 8$ ، $\sigma=40$

$\bar{X} \sim N(m_0 , \frac{\sigma}{\sqrt{n}})$, $\quad \bar{X} \sim N(8 , \frac{2}{\sqrt{40}})$,

$$t' = \frac{\bar{X}-8}{\frac{2}{\sqrt{40}}} = \frac{10-8}{\frac{2}{\sqrt{40}}} = 6.324$$

$2\alpha=0.1$, و منه من الجدول.2: $t_\alpha=1.645$

$t' > t_\alpha$: نقبل H_1 و نرفض H_0



مثال.2: ترغب شركة إنتاج المشروبات الغازية أن تعرف بدرجة ثقة **98%** ما إذاكان بإمكانها الادعاء أن القارورات التي تنتجها تحتوي في المتوسط على **2** لتر من المشروب. يعلم صاحب الشركة أن سعت القارورات تخضع إلى توزيع طبيعي لذلك أخذ عينة عشوائية حجمها **35** قارورة فوجد أن متوسط عدد اللترات التي تحتويها هو **2.06** لتر بانحراف معياري قدره **0.2** لتر. لذلك أراد اختبار ما إذاكان المتوسط الحسابي للمجتمع أكبر من **2** لتر.

الحل:

$H_0: m = m_0 = \mathbf{2L}$

$H_1: m = m_1 > m_0$

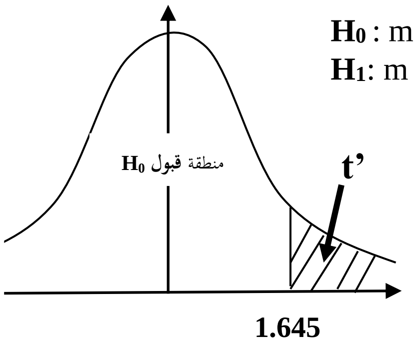
$\bar{X}=2.06$ ، $m_0 = 2$ ، $\sigma=0.2$, $n=35$

$\bar{X} \sim N(m_0 , \frac{\sigma}{\sqrt{n}})$, $\quad \bar{X} \sim N(2 , \frac{0.2}{\sqrt{35}})$,

$$t' = \frac{\bar{X}-2}{\frac{0.2}{\sqrt{35}}} = \frac{2.06-2}{\frac{0.2}{\sqrt{35}}} = 1.775$$

$2\alpha=0.04$, و منهمن الجدول.2: $t_\alpha=2.054$

$t' < t_\alpha$: نقبل H_0 و نرفض H_1



مثال.3: أراد مدير شركة إنتاج الدرجات النارية إمضاء صفقة مع شركة إنتاج الإطارات لكي تموله بـ 100000 إطار. تدعي هذه الأخيرة بأن نسبة الإطارات الفاسدة هو 4%. مدير الشركة لم يصدق هذا الادعاء فقام بمراقبة جودة 200 إطار فوجد 6% فاسدة. على هذا الأساس أراد المدير اختبار فرضية أن نسبة الإطارات الفاسدة هي أكبر من 4%. و اختار مستوى الثقة 99%.

$H_0: P = P_0 = \mathbf{4\%}$

$H_1: P = P_1 > P_0$

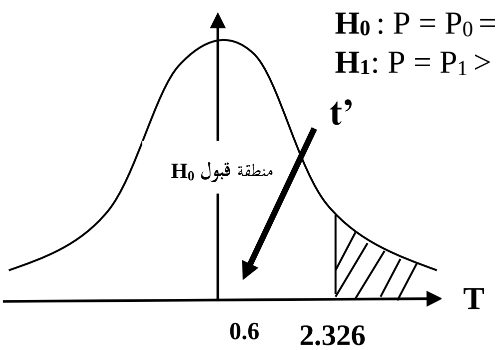
$\hat{P}=0.06$ ، $P = 0.04$ ، $n=200$

$\hat{P} \sim N(P , \sqrt{\frac{p.q}{n}})$, $\quad \hat{P} \sim N(0.04 , \sqrt{\frac{(0.04).(0.96)}{200}})$,

$$t' = \frac{\hat{P}-0.04}{\sqrt{\frac{(0.04).(0.96)}{200}}} = \frac{0.06-0.04}{0.0138} = 1.44$$

$2\alpha=0.02$, و منه من الجدول.2: $t_\alpha=2.326$

$t' < t_\alpha$: نقبل H_0 و نرفض H_1



مثال.4: يتلقى مكتب مصلحة المستهلك لمؤسسة معينة شكاوى كثيرة من زبائنه، فحواها أن وزن علب مسحوق التنظيف الذي تبيعها إياهم يزن أقل من 450غ. للتحقق من هذه المعلومة قام مكتب البحوث و الدراسات لهذه الشركة، بوزن عينة من 45 علبة (سحبتها بالإرجاع) فوجدت أوزانها 435غ. تشير الخبرات السابقة لهذه المؤسسة إلى أن الانحراف المعياري لوزن العلب هو 25غ. هل يمكن لهذه المؤسسة أن تؤيد شكاوى زبائنها عند مستوى معنوية **5%**.

الحل

$H_0: m = m_0 = 450$

$H_1: m = m_1 < m_0$

X: وزن علب مسحوق التنظيف

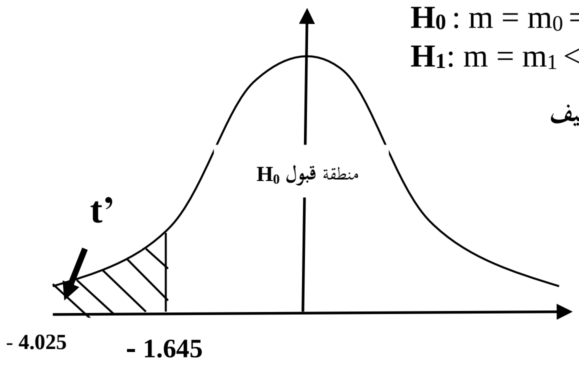
$\bar{X}=435$ ، $m_0 = 450$ ، $\sigma=25$, $n=45$

$\bar{X} \sim N(m_0 , \frac{\sigma}{\sqrt{n}})$, $\quad \bar{X} \sim N(450 , \frac{25}{\sqrt{45}})$,

$$t'=\frac{\bar{X}-450}{\frac{25}{\sqrt{45}}} = \frac{435-450}{3.726} = -4.025$$

$-t_\alpha = -1.645$:2.و منه من الجدول $2\alpha=0.1$,

$t'<-t_\alpha$: نقبل H_1 و نرفض H_0



مثال.5 : تخضع مدة الاستجابة إلى طلبية الزبائن في إحدى المطاعم إلى توزيع طبيعي وسطه الحسابي 15د و انحراف معياري 2د. تم إدخال تحسينات على هذه الخدمة من خلال دمج برنامج إدارة العلاقات مع الزبائن، و للتأكد من نجاعة هذا النظام قام رجل التسويق باختيار 30 زبون ممن طبق عليهم هذا النظام فوجد أن متوسط مدة تلبية طلبيتهم هي 11د. حسب رأيك هل النظام الجديد أحسن من القديم عند مستوى معنوية **5%**.

الحل: X: مدة الخدمة

$\bar{X}=11$ ، $m_0 = 15$ ، $\sigma=2$, $n=30$

$\bar{X} \sim N(m_0 , \frac{\sigma}{\sqrt{n}})$, $\quad \bar{X} \sim N(15 , \frac{2}{\sqrt{30}})$,

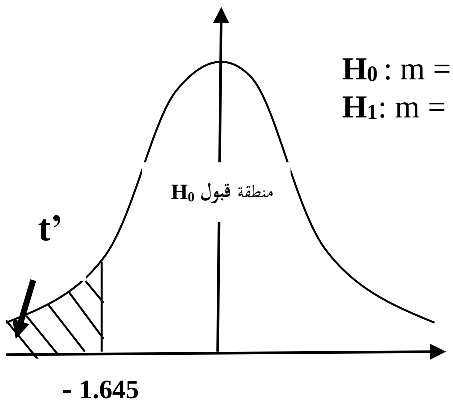
$$t'=\frac{\bar{X}-15}{\frac{2}{\sqrt{30}}} = \frac{11-15}{0.365} = -10.95$$

$t_\alpha = 1.645$:2.و منه من الجدول $2\alpha=0.1$,

$t'<-t_\alpha$: نقبل H_1 و نرفض H_0

$H_0: m = m_0 = 15$

$H_1: m = m_1 < m_0$



مثال.6: صرحت مديرية الضرائب في مدينة معينة أن أكثر من **80%** من التجار يدفعون الضرائب، لكن أحد الموظفين بالمديرية لم يصدق هذا التصريح، فأخذ عينة من **64** تاجر فوجد **48** منهم يدفعون الضرائب. هل يمكن تصديق بيانات العينة تحت مستوى معنوية **5%**.

الحل: P: نسبة الذين يدفعون الضرائب

$\hat{P}=\frac{48}{64}=0.75$ ، $P = 0.8$ ، $n=64$

$\hat{P} \sim N(P , \sqrt{\frac{p.q}{n}})$, $\quad \hat{P} \sim N(0.8 , \sqrt{\frac{(0.75).(0.25)}{64}})$,

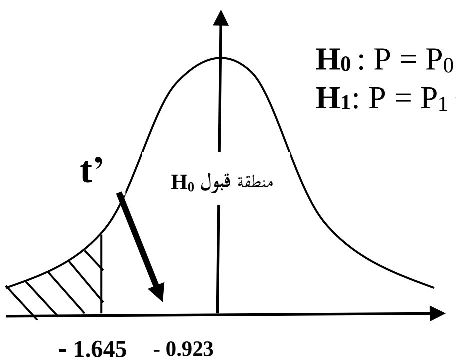
$$t'=\frac{\hat{P}-0.04}{\sqrt{\frac{(0.75).(0.25)}{64}}} = \frac{0.75-0.80}{0.05412} = -0.9238$$

$t_\alpha=1.645$:2.و منه من الجدول $2\alpha=0.1$,

$t'>-t_\alpha$: نقبل H_0 و نرفض H_1

$H_0: P = P_0 = 80\%$

$H_1: P = P_1 < P_0$

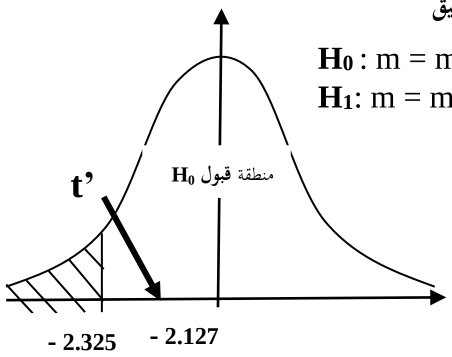


مثال.7: تلقت مؤسسة إنتاج المواد الغذائية شكاوى من مستهلكين فهواها أن أكياس الدقيق التي تنتجها تزن أقل من 10كغ، للتحقق من ذلك قامت بوزن عينة من 40 كيس فوجدت أن متوسط أوزانها 9.8كغ. تشير الخبرة السابقة إلى أن الانحراف المعياري لوزن الأكياس هو 0.3كغ. هل هذه المعطيات تؤيد شكاوى المستهلكين عند مستوى معنوية 1%.

الحل: X: **وزن الأكياس الدقيق**

$H_0: m = m_0 = 15$

$H_1: m = m_1 < m_0$



$\bar{X}=9.9$ ، $m_0 = 10$ ، $\sigma=0.3$, $n=40$

$\bar{X} \sim N(m_0 , \frac{\sigma}{\sqrt{n}})$, $\bar{X} \sim N(10 , \frac{0.3}{\sqrt{40}})$,

$$t' = \frac{\bar{X} - 10}{\frac{0.3}{\sqrt{40}}} = \frac{9.9 - 10}{0.047} = -2.127$$

$2\alpha=0.1$, و منه من الجدول.2: $t_\alpha = 2.325$

$t' > -t_\alpha$: نقبل H_0 و نرفض H_1

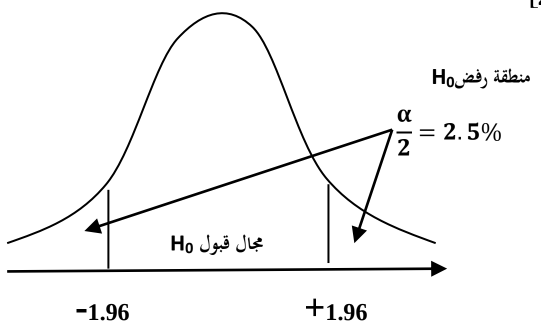
مثال.8: متوسط أعمار خريجي دفعة معينة هو 22 سنة. بعد اختبار عينة مكونة من 150 طالب من هذه الدفعة، تبين أن متوسط أعمارهم هو 23 سنة و الانحراف المعياري هو 5.65 سنة. هل يمكن استنتاج أن متوسط أعمار خريجي تلك الدفعة يختلف عن 22 سنة عند مستوى معنوية 5%.

الحل: $\bar{X} = 23$، $n= 150$ ، $m_0= 22$، $\sigma = 5.65$

X: **متوسط أعمار خريجي الجامعة**

نحن أمام الحالة الثالثة [$n>30$، σ معلومة، نقرأ $t_{\frac{\alpha}{2}}$ من جدول التوزيع الطبيعي ج.2]

$\alpha=5\%=0.05$، **إذن** من جدول.2 نجد أن $t_{\frac{\alpha}{2}}=1.96$



$$t' = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} = \frac{23-22}{\frac{5.65}{\sqrt{150}}} = 2.17$$

بما أن: $t' = 2.17 > 1.96$

إذن نرفض H_0 و نقبل H_1 و نستنتج أن $m_0 \neq 22$

مثال.9: لصنع منتج معين تحتاج مؤسسة A قطع ميكانيكية متوسط طولها 12.5 سم بانحراف معياري 0.8 سم. اختيرت عينة من 18 قطعة، بعد قياسها تبين لها أن متوسط أطوالها هو 12.9 سم. هل يمكن استنتاج أن متوسط طول القطع الميكانيكية يختلف عن 12.5 سم عند مستوى معنوية **2%**.

الحل: (n<30): [عندما تكون **σ** غير معلومة و توزيع يكون طبيعي نستعمل توزيع student T]

$\alpha = 2\%$ ، $\sigma' = 0.8$ ، $\bar{X} = 12.5$ ، $n=18$

1. نضع الفرضيتين التاليتين: $\begin{cases} H_0: m_0 = 12.5 \\ H_1: m_0 \neq 12.5 \end{cases}$

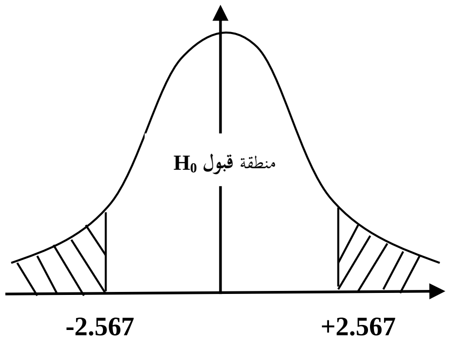
2. حساب t':

$$\sigma * = \frac{0.8}{\sqrt{18}} = \frac{\sigma}{\sqrt{n}} = 1.88,$$

$$t' = \frac{\bar{X} - m_0}{\sigma} = \frac{12.9 - 12.5}{1.88} = 2.121,$$

3. **تعيين القيمة الجدولية لـ t** من جدول Student (ج.3) **لأن** (n<30)

بما أن $\frac{\alpha}{2} = 0.02$ نجد $t_{\frac{\alpha}{2}} = 2.567$



4. **اتخاذ القرار:**

نرفض الفرضية العدمية H_0 إذا كانت **t** في منطقة الرفض.

إذن بما أن $t = 2.121 < 2.567$ عدم رفض H_0

مثال.5: صرح مدير إنتاج الاسمنت بمؤسسة معينة بأن متوسط أوزان أكياس الإسمنت هو **48**كغ و انحراف معياري **3**كغ. للتأكد من وزن الأكياس أخذ عينة من **25** كيس فوجد متوسط وزنها **50** كغ. هل يمكننا القول بأن المدير غير مخطأ في تصريحه عند مستوى معنوية **2%**.

الحل: $n=25$ ، $m_0=48$ ، $\sigma = 3$ ، $\bar{X} = 50$

الحالة الثالثة: $[m = m_0 \neq m_1]$

1. نضع الفرضيتين التاليتين: $\begin{cases} H_0: m_0 = 48 \\ H_1: m_0 \neq 48 \end{cases}$

2. نحسب t': بحيث $t' = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} = \frac{50 - 48}{\frac{3}{\sqrt{25}}} = 3.33$

3. نقرأ $t_{\frac{\alpha}{2}}$ الموافقة لـ $\left(\frac{\alpha}{2}\right)$ **من الجدول (2)**

$\frac{\alpha}{2} = 0.02$ **لذلك من الجدول.2 نجد :** $t_{\frac{\alpha}{2}} = 2.326$

4. **نعتمد على قاعدة إتخاذ القرار التالية:**

* $t' \in]-\infty, -t_\alpha[\cup [t_\alpha + \infty[$: نقبل H_1 و نرفض H_0

* $t' \in]-t_\alpha, +t_\alpha[$: نقبل H_0 و نرفض H_1

إذا نقبل الفرضية البديلة H_1 $t' > +t_\alpha$

لأن $3.33 > 2.326$ و تنتمي إلى $[2.326, +\infty[$

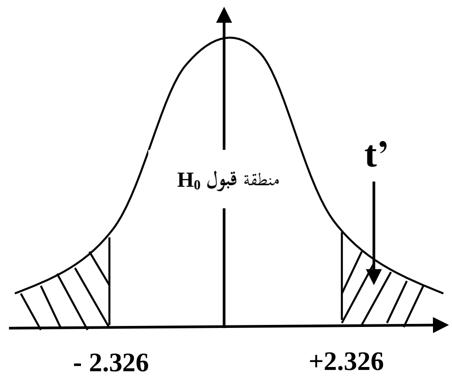
ملاحظة: لو قيل في المعطيات بأن المدير يريد التأكد من أن

* وزن الأكياس أكثر من 48كغ الفرضيات تصبح:

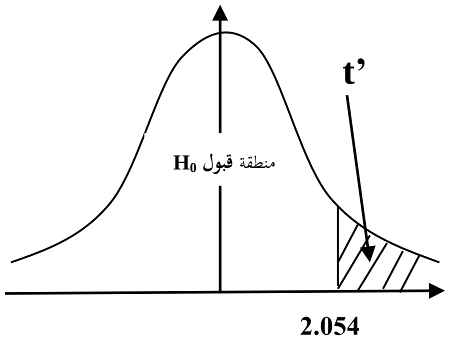
$$\begin{cases} H_0: m_0 = 48 \\ H_1: m_1 > 48 \end{cases}$$

* وزن الأكياس أقل من 48كغ الفرضيات تصبح:

$$\begin{cases} H_0: m_0 = 48 \\ H_1: m_1 < 48 \end{cases}$$



مثال.6: مدة الصلاحية المتوسطة لإحدى القطع الميكانيكية التي ينتجها المصنع هو 900سا بانحراف معياري 95سا. يدعي مدير الانتاج أنه أدخل بعض التعديلات على عملية الانتاج أدت إلى تحسين جودة القطع الميكانيكية (أي مدة صلاحيتها). لاختبار صحة ادعائه أخذت عينة من **49** قطعة و تم استخدامها كلها فكانت مدة صلاحيتها المتوسطة **960**سا. هل يمكن تأييد ادعاء مدير الانتاج عند مستوى معنوية **2%**.



الحل: $n=49$ ، $m_0=900h$ ، $\sigma = 95h$ ، $\bar{X} = 960h$

1.نضع الفرضيتين التاليتين:

$$\begin{cases} H_0: m_0 = 900 \\ H_1: m_1 > 900 \end{cases}$$

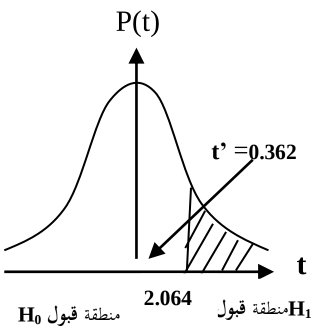
2. نحسب t': بحيث $t' = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} = \frac{960 - 900}{\frac{95}{\sqrt{49}}} = 4.42$

3. نقرأ t_α الموافقة لـ $(2\alpha=0.04)$ **من الجدول (2) لذلك من الجدول.2:** $t_\alpha=2.054$

4. نعتمد على قاعدة إتخاذ القرار التالية:

* $t' \geq t_\alpha$: نقبل H_1 و نرفض H_0 ، * $t' < t_\alpha$: نقبل H_0 و نرفض H_1 **إذاً نقبل الفرضية البديلة H_1**

التمرين.7: قامت شركة لإنتاج الدرجات النارية بإمضاء صفقة مع شركة إنتاج الإطارات لكي تزودها بـ 100000 إطار، ادعت هذه الأخيرة بأن نسبة الإطارات التي بها عيب هو 4%. مدير شركة الدرجات النارية لم يصدق بهذا الادعاء فقام بصورة عشوائية بمراقبة جودة 200 إطار فوجد أن 9 منها بها عيب، ثم قام باختبار فرضية أن تكون نسبة الإطارات الفاسدة أكبر من 4%، و ذلك عند مستوى معنوية 99%.



حل: $\hat{P}=\frac{9}{200}=0.045$ ، $P_0 = 0.04$ ، $n=200$ ، $\alpha=1\%$

- نضع الفرضيتين التاليتين: $H_0: P = P_0 = 4\%$ ، $H_1: P = P_1 > 4\%$
- نحسب t': بحيث $t' = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} = \frac{0.045 - 0.04}{\sqrt{\frac{0.04.(1-0.04)}{200}}} = \mathbf{0.362}$
- **$2\alpha=0.04$,** و منه من الجدول.2: $t_\alpha=2.064$

وجدنا أن $t'=0.362$ و $t_\alpha=2.064$ إذاً

- بما أن $t' < t_\alpha$: نقبل H_0 و نرفض H_1

إستنتاج: نستنتج أن نسبة الإطارات الفاسدة4%.

مثال.8: يعتبر مدير إحدى الإدارات أن القيمة المتوسطة لمجتمع كبير من عروض الأسعار هي 10 وحدات، لكن المحاسب يختلف معه في هذا الاعتقاد و يظن أنها أقل من ذلك الرقم. للتأكد من صحة وجهة نظره أخذ المحاسب عينة من 36 عرضا، و بعد دراسته وجد أن وسطها الحسابي هو $\overline{X}=9.5$ و تباينها $S^2=3.88$. هل ما يعتقده المدير هو الصح تحت مستوى معنوية [5%].

الحل: $\alpha=0.05$، $n=36>30$، $S^2=3.88$، $X=9.5$، $m_0=10$

1. نضع الفرضيتين التاليتين:

$$\begin{cases} H_0: m_0 = 10 \\ H_1: m_0 < m_1 \end{cases}$$

2. حساب t' :

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

$$\sum(x_i - \overline{X})^2 = 3.88.(36) = 139.68$$

$$\sigma = \sqrt{\frac{\sum(x_i - X)^2}{n-1}} = \sqrt{\frac{139.68}{35}} = 1.96$$

$$\sigma' = \frac{0.8}{\sqrt{18}} = \frac{\sigma}{\sqrt{n}} = 1.88,$$

$$t' = \frac{\overline{X} - m_0}{\sigma} = \frac{9.5-10}{\frac{1.96}{\sqrt{36}}} = -1.5,$$

3. تعيين القيمة الجدولية لـt_α

من جدول Student (ج.2) **(n>30)**

$\frac{\alpha}{2} = 0.05$ إذاً $\alpha = 0.1$ و منه $t_\alpha = 1.645$

4. **اتخاذ القرار** نرفض الفرضية العدمية H_0 إذا كانت t' في منطقة الرفض.

$$t' = -1.5 > -t_\alpha = -1.645$$

إذن نختار H_0 و نرفض H_1 بذلك نستخلص أنه لا وجود لدليل كافٍ لرفض ما يعتقده المدير.(المدير هو الأقرب إلى الصح)