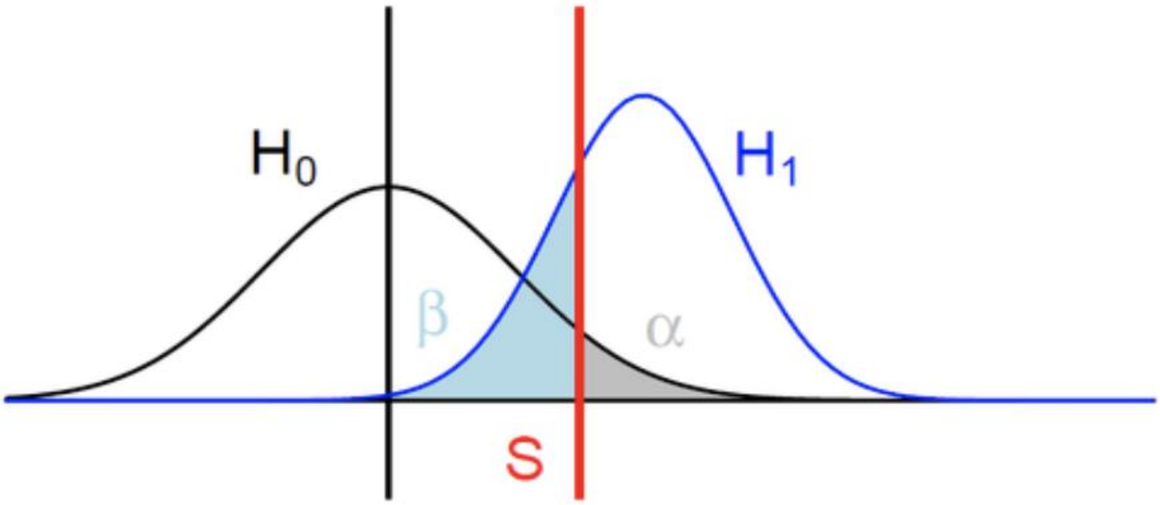
Université Abou Bekr Belkaid
Tlemcen Algérie
جامعة أبي بكر بلقايد
تلمسان الجزائر

كلية العلوم الاقتصادية، التجارية وعلوم التسيير

مطبوعة بيداغوجية

سلسلة محاضرات و تمارين في الاحصاء الرياضي

موجهة إلى طلبة السنة الثانية كلية العلوم الاقتصادية، التجارية وعلوم التسيير



من اعداد : د.طالب دليلة

السنة الجامعية 2022/2021

بسم الله الرحمن الرحيم

تقديم

أصبح لدى الإنسان في وقتنا الحاضر ميلا متعاظما باتجاه تكميم سمات وخصائص الأشياء والافراد، بحيث أصبحت أنشطة الإنسان تتلخص في صورة أرقام، وأضحى التعامل والتفاهم مبنيا على لغة الرقم والكم أكثر من لغة الوصف والكيف لدقة الأولى، وما التطور العلمي والتكنولوجي الذي يشهده العالم اليوم إلا حصيلة لما اكتسبه الإنسان من مهارة التعامل مع الأرقام جمعا وتلخيصا وتحليلا لتسهيل الوصول على الاستنتاجات وزيادة دقة القرارات في مجالات الحياة المختلفة.

ويعتبر علم الإحصاء ركنا أساسيا في حياة الأفراد والمؤسسات باعتباره رياضيات جمع البيانات تلخيصها وتحليلها وصولا إلى قرارات مبينة على جزئيات لتعمم بصورة إجمالية، كما استفاد الباحثون من طرق علم الإحصاء في إنجاز أبحاثهم في مختلف مجالات الحياة، وذلك عبر توظيف تقنيات الإحصاء في جمع البيانات وتحليلها.

هذه المطبوعة هي ثمرة تجربة سنوات عديدة في تدريس الإحصاء بكلية العلوم الاقتصادية لجامعة ابو بكر بلقايد بتلمسان، ولقد حاولنا أن نستفيد من هذه التجربة لصياغة محتوى المقياس بطريقة تلائم مستوى طلبة هذه الكلية وطبيعة التخصص، لتحقيق هذا الغرض حرصنا على ربط المفاهيم والقواعد النظرية باستخداماتها التطبيقية؛ فعملنا على إعطاء أمثلة محلولة عن كل مفهوم جديد. ولأن فهم القواعد الرياضية يكون أسهل إذا كان للمتلقي خلفية عن المشكلة التي يحتاج حلها إلى استخدام هذه القواعد، عملنا في كثير من الأحيان إلى التقديم لبعض الدروس أو النظريات بمسألة تكون بمثابة التمهيد، وأحيانا بمثابة مشكلة ننطلق منها لنتوصل إلى النظرية.

يتضمن البرنامج المقرر حسب البرنامج الوزاري لمقياس "إحصاء 3" للسنة الثانية علوم اقتصادية، التجارية وعلوم التسيير، برمج هذا المقياس لطلبة السنة الثانية، لكي يستفيد الطلبة من القاعدة التي اكتسبوها عند دراسة الإحصاء الوصفي في السنة الأولى، لكن هدفه الأساسي هو التمهيد لدراسة الإحصاء التطبيقي في السنة الثالثة، هدف هذا المقياس هو تقديم علم الإحصاء الرياضي، أي الأساس الرياضي للإحصاء التطبيقي.

يتضمن هذه المطبوعة على اربعة فصول حيث يعبر كل فصل على محاضرة ، يتناول الفصل الأول على التوزيع الطبيعي من خلال سرد مختلف خصائصه و تطبيقاته ، ليتناول الفصل الثاني نظرية المعاينة أو ما يسمى بتوزيع المعاينة يقصد به التوزيع لمختلف القيم التي يمكن أنت يأخذها إحصاء أو مقدر العينة لعينات كثيرة من نفس الحجم، وهكذا وبالرغم أننا عمليا سنحصل على عينة عشوائية أو مجموعة قياسية جزئية واحدة إلا أننا نعرف أن الإحصاء المعين للعينة الذي تحدد مثل الوسط الحسابي أو الوسيط للعينة لا يساوي تماما معلمة المجتمع المقابل له، في حين يتضمن الفصل الثالث نظرية التقدير و المقصود بالتقدير هو تقدير معالم المجتمع الإحصائي(أو التوزيع الاحتمالي) والتي غالباً ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة، أو تقدير متوسط عمر الناخب ...وغيرها، أما عن الفصل الرابع فيتضمن اختبار الفرضيات حيث يعتبر إختبار الفروض أحد المواضيع الرئيسية في الإحصاء الرياضي،

ويستهدف الوصول إلى القرار بشأن معلمة المجتمع من خلال قبول او رفض تقديرها المعتمد على معطيات العينة المسحوبة من ذلك المجتمع.

كلمة إلى الطلبة

كثيرا ما نلاحظ أن الطلبة يستخدمون التمارين المقدمة في السلاسل كنماذج أو شبه قوانين في حد ذاتها يحاولون حفظها بينما هي في الحقيقة مجرد وسيلة لفهم الدرس. هذا التشبث بالشكل دون المضمون في محاولة يائسة لمواجهة الامتحان دون فهم حقيق لمضمون المادة هو نتيجة حتمية بالنسبة لمن لا يتابع المحاضرات والتطبيقات بالمراجعة المستمرة و الفورية. وحسب رأينا فإن الصعوبة التي يواجهها الطلبة في هذا المقياس سببها أنه مقياس يعتمد أساسا على الفهم أكثر مما يعتمد على التذكر. وهذا الفهم لا يتأتى عن طريق التلقي من الأستاذ، مهما بذل هذا الأخير ومهما كانت مهارته، وإنما يحتاج إلى جهد مستقل يبذله الطالب بمفرده مع قدر من التركيز و المثابرة. "الوصفة السحرية" لفهم هذه المادة، هي المراجعة بجرعات منتظمة و فورية (بعد كل محاضرة قبل النوم[1]!) مع شيء من التركيز على القواعد والمفاهيم حتى يتم فهمها فهما جيدا، وليعمل الطالب على تعميق فهمه من خلال تمارين السلاسل و لكن لا يتخذها "نماذج" جامدة أو قواعد إضافية.

إن هدف الأستاذ والجامعة ككل هو إعداد الطالب لمواجهة المشكلات المعقدة للتسيير، وهذا الهدف لا يتحقق إلا بتنمية الذكاء والتزود بعدد من التقنيات المساعدة. إن الجائزة الحقيقية التي يجب أن يتوقعها الطالب من دراسة بالجامعة هي تكوين قدرة على التعلم الذاتي أكثر من تجميع كم من المعارف التي قد لا يحتاجها أبدا، و هي من جهة أخرى، تكوين ذهنية مستقلة قادرة على تحليل المشكلات والوضعيات المعقدة وصياغتها في شكل واضح ودقيق ومن ثم إبداع حلول لها من خلال تفكيره الخاص. هذه القدرة لا تتأتى إذا عود الطالب نفسه على إعمال فكره مطولا في المسائل التي تطرحها التمارين مما يعطي الطالب القدرة على التحليل والتركيب والاستنباط والاستدلال كأسس التفكير المنتج والمبدع. إن الوصول إلى هذه القدرة على مواجهة مشكلات وحلها هي غاية أساسية للتعليم الجامعي وهي أحسن رأسمال يجمعه الطالب ليستثمره حياته العامة والخاصة معا.

و غني عن الذكر أن محتوى هذه المطبوعة البيداغوجية من نظريات وقواعد ليس من إبداع مؤلفها، و إنما هي قواعد مبسوطة في المراجع جمعناها وعرضناها وفق البرنامج الوزاري بأسلوب رأينا أنه الأنسب لمستوى طالب سنة ثانية كلية العلوم الاقتصادية التسيير و العلوم التجارية.

[1] يعلمنا علم نفس التربية أن أكثر من 60 بالمئة من المعلومات التي نتعلمها ننساها في التسع ساعات الأولى. فأنقذ معلوماتك في الساعات الأولى قبل أن تتبخر!

الفهرس

الفصل الرابع : اختبار الفرضيات

مقدمة

يعتبر التوزيع الطبيعى من أهم التوزيعات المتصلة حيث يلعب دوراً أساسياً فى عملية المعاينة Sampling ،كما يستخدم لوصف الأنماط التكرارية للعديد من الظواهر الإحصائية مثل التغيرات الطبيعية التى تحدث للإنسان والحيوان والعوامل البيئية بشكل عام . والتوزيع الطبيعى يدرس سلوك المتغيرات العشوائية المتصلة مثل درجة الحرارة والطول 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 من المحور الأفقى شيئا فشيئا دون أن يتماسا مع هذا المحور . وإذا أسقطنا عموداً من قمة 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 حيث 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 قيمتين معلومتين.

وترجع أهمية التوزيع الطبيعى إلى الأسباب الآتية :

1. يستخدم التوزيع الطبيعى فى وصف العديد من المتغيرات العشوائية فى الواقع العملى مثل الأطوال والأوزان لمجموعة من الأشخاص ، درجات الطلاب فى امتحان مقرر معين ، أخطاء القياس الناتجة من تجربة ما ، حيث تتوزع القياسات المشاهدة (التكرارات) بشكل متماثل حول قيمة مركزية هى قيمة متوسط التوزيع وتقل هذه القياسات تدريجيا كلما ابتعدنا عن هذه القيمة إلى جهة اليمين أو إلى جهة اليسار ، حيث يأخذ منحنى التوزيع الشكل الجرسى .

2. يعتبر التوزيع الطبيعى تقريبا مفيدا للعديد من التوزيعات المتقطعة وغير المتقطعة مثل 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 .

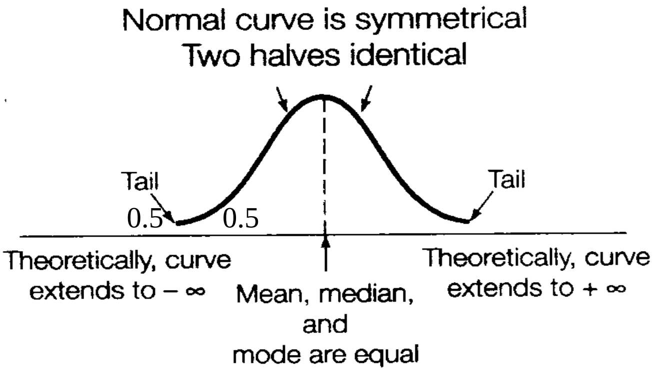
1- خصائص التوزيع الطبيعى

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 المنحنى على المحور الأفقى عند المتوسط يقسم المساحة إلى نصفين متساويين 50% من المساحة الكلية على يمين الخط العمودى ، 50 % من المساحة الكلية على يساره .

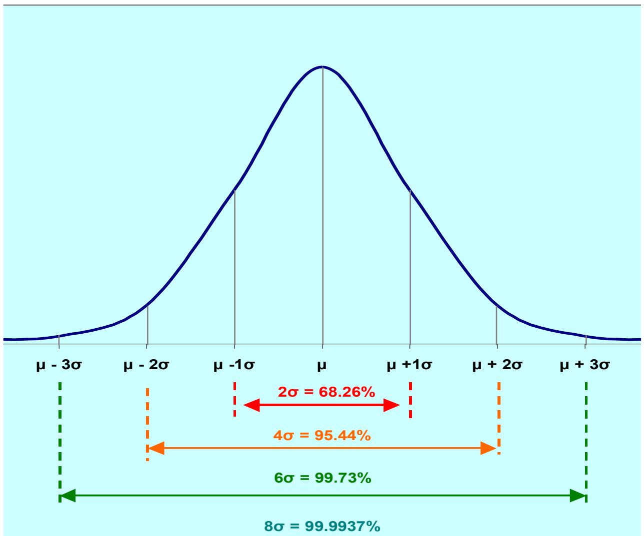


4. إذا أسقطنا عمودين على بعد انحراف معيارى واحد إلى يمين ويسار متوسط التوزيع فإن المساحة التى يحصرها هذين العمودين تمثل 68.26% تقريبا من المساحة الكلية تحت المنحنى . كذلك نجد أن المساحة التى يحصرها العمودان المقامان عند 2 انحراف معيارى على جانبى متوسط التوزيع تمثل 95.44% تقريبا من المساحة الكلية تحت المنحنى . أما المساحة التى يحصرها العمودان المقامان عند 3 انحرافات معيارية على جانبى متوسط التوزيع فتمثل 99.74% تقريبا من المساحة الكلية تحت المنحنى . هذا ويمكن التعبير عما سبق رمزيا كالآتى :

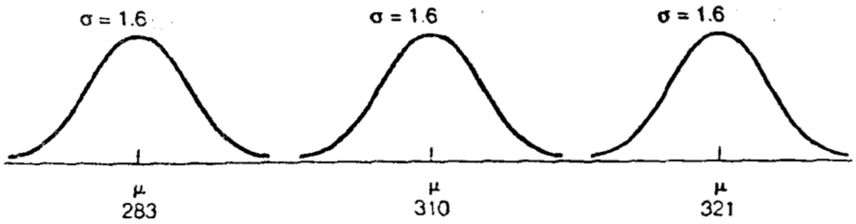
$$p(\mu - 1\sigma < x < \mu + 1\sigma) = 0.6826$$

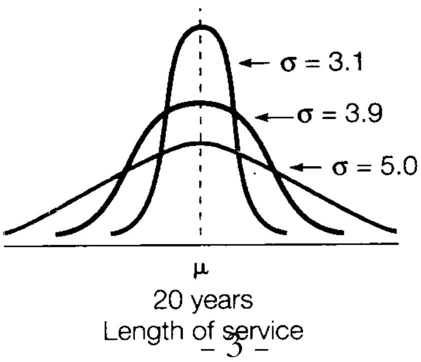
$$p(\mu - 2\sigma < x < \mu + 2\sigma) = 0.9544$$

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



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





2- 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 :

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

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

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 نهائي من التوزيعات الطبيعية والتي 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 المعيارى متى علم متوسطه وانحرافه المعيارى .

3- التوزيع الطبيعى المعيارى Standard Normal Distribution

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

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

وتسمى Z أيضا بالقيمة المعيارية أو الدرجة المعيارية وهى تساعد على إيجاد المساحة أو الاحتمال المطلوب أسفل أى منحنى توزيع طبيعى وذلك باستخدام جدول المساحات تحت منحنى التوزيع الطبيعى العيارى . هذا وتأخذ دالة كثافة الاحتمال 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 وبالتالى فقد استخدمت فى حساب جدول المساحات تحت منحنى التوزيع الطبيعى المعيارى المشار إليه سلفا .

3-1خصائص المنحنى الطبيعى المعيارى

1. المساحة الكلية تحت منحنى التوزيع الطبيعى المعياري تساوى الواحد الصحيح بحيث يقسم الخط العمودي الواصل من قمة المنحنى إلى الصفر (متوسط التوزيع) المساحة إلى قسمين متساويين مساحة كل منهما تساوى 0.05

2. منحنى التوزيع الطبيعى المعيارى متماثل حول متوسطه ، وبالتالى فإن التواءه يساوى صفر وتفرطحه يساوى 3 .

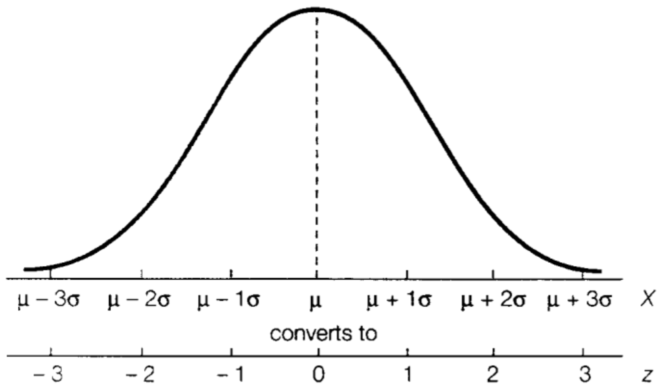
3. المساحة المحصورة بين ± 1 درجة معيارية تساوى % 68.26 تقريبا من المساحة الكلية تحت المنحنى . والمساحة المحصورة

بين ± 2 درجة معيارية تساوى % 95.44 تقريبا من المساحة الكلية تحت المنحنى . أما المساحة المحصورة بين 3 $\pm$ درجات معيارية تساوى % 99.74 تقريبا من المساحة الكلية تحت المنحنى . هذا ويمكن التعبير عن ذلك رمزيا كالآتى :

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

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

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



مثال (1)

إذا كان العمر الافتراضى لمشغل الأقراص المرنة والمنتج بمعرفة إحدى شركات 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 .

الحل :

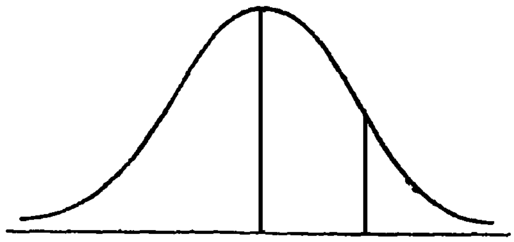
نفرض أن المتغير العشوائى χ يعبر عن العمر الافتراضى لمشغل الأقراص المرنة ، أى أن :

المتوسط μ = 1000 ساعة

الانحراف المعيارى σ = 100 ساعة

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

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

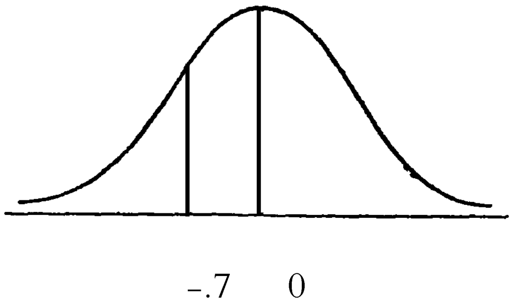


0 1.5

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

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

$$= p(Z < -0.7) = 0.5 - 0.2580 = 0.2420$$



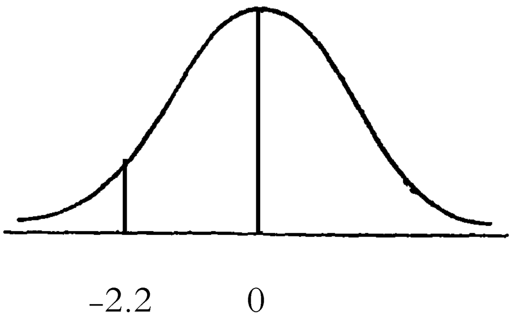
3. احتمال أن يكون العمر الافتراضي لمشغل الأقراص المرنة أكبر من 780 ساعة

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

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

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

$$= 0.5 + 0.4861 = 0.9861$$



4. احتمال أن يكون العمر الافتراضي لمشغل الأقراص المرنة بين 700 ، 1200 ساعة

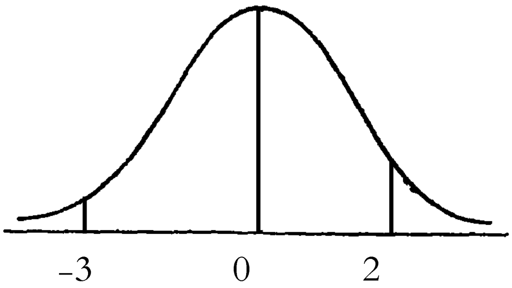
$$p(700 \le X \le 1200) = P\left(\frac{700-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{1200-1000}{100}\right)$$

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

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

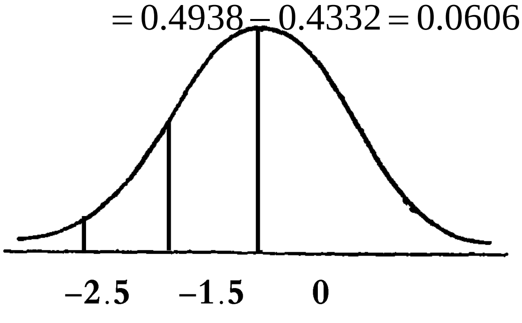
$$= P(0 \le Z \le 3) + P(0 \le Z \le 2)$$

$$= 0.4987 + 0.4772 = 0.9759$$



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

$$P(750 \le X \le 850) = P\left(\frac{750-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{850-1000}{100}\right)$$
$$= P\left(\frac{-250}{100} \le Z \le \frac{-150}{100}\right)$$
$$= P(-2.50 \le Z \le -1.50)$$
$$= P(-2.50 \le Z \le 0) - P(-1.50 \le Z \le 0)$$
$$= P(0 \le Z \le 2.50) - P(0 \le Z \le 1.50)$$
$$= 0.4938 - 0.4332 = 0.0606$$



التوزيع الطبيعى كتقريب لتوزيع ذى الحدين

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

1. حجم العينة يكون كبيرا ($n \ge 30$)

2. المتوسط يكون أكبر من أو يساوى 5 ($\mu = np \ge 5$)

3. التباين يكون أكبر من أو يساوى 5 ($\sigma^2 = np\,(1-p) \geq 5$)

مثال (2)

إذا كانت نسبة الوحدات المعيبة المنتجة بواسطة إحدى الآلات 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 .

الحل :

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

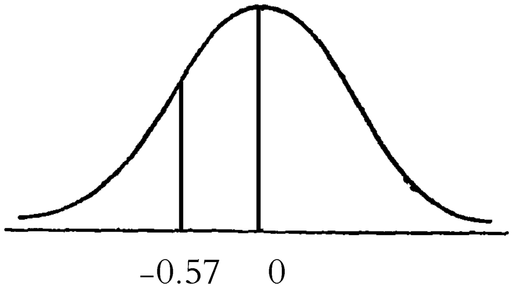
وعلى فرض أن المتغير العشوائى X يعبر عن عدد الوحدات المعيبة فإن الاحتمالات المطلوبة يمكن حسابها كالآتى :

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

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

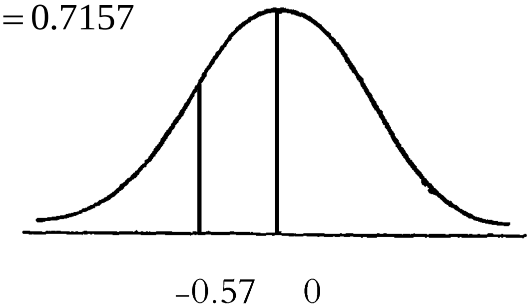
$$= p\left(Z < \frac{-2.5}{4.4}\right)$$

$$= p(Z < -0.57)$$
$$= 0.5 - p(-0.57 < Z < 0)$$
$$= 0.5 - P(0 < Z < 0.57)$$
$$= 0.5 - 0.2157 = 0.2843$$



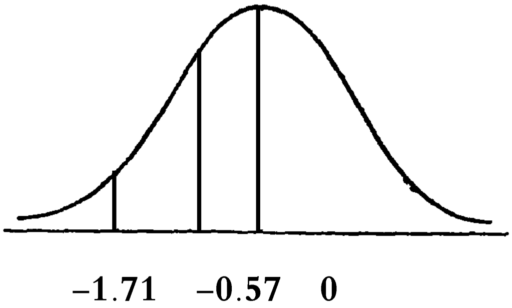
2. احتمال أن تحتوى العينة على 20 وحدة معيبة على الأقل

$$P(X \geq 20) = P(\frac{X - \mu}{\sigma} \geq \frac{20 - 22.5}{4.4})$$
$$= p(Z \geq -0.57) = 0.5 + p(-0.57 \leq Z \leq 0)$$
$$= 0.5 + P(0 \leq Z \leq 0.57)$$
$$= 0.5 + 0.2157 = 0.7157$$



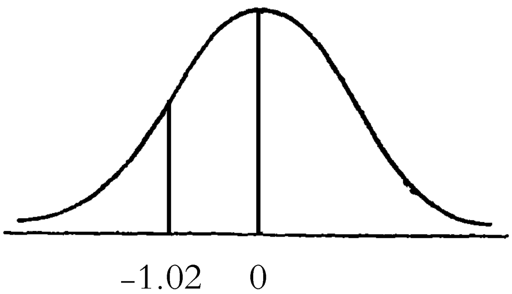
3. احتمال أن تحتوى العينة على ما بين 15 ، 20 وحدة معيبة

$$P(15 \leq X \leq 20) = P\left(\frac{15 - 22.5}{4.4} \leq \frac{X - \mu}{\sigma} \leq \frac{20 - 22.5}{4.4}\right)$$
$$= P(-1.71 \leq Z \leq -0.57)$$
$$= P(-1.71 \leq Z \leq 0) - p(-0.57 \leq Z \leq 0)$$
$$= P(0 \leq Z \leq 1.71) - P(0 \leq Z \leq 0.57)$$
$$= 0.4564 - 0.2157 = 0.2407$$



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

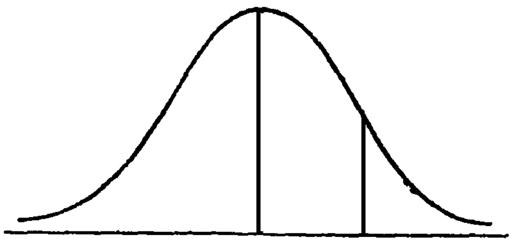
$$P(X>18)=P\left(\frac{X-\mu}{\sigma}>\frac{18-22.5}{4.4}\right)$$
$$=P\left(Z>\frac{-4.5}{4.4}\right)=P(Z>-1.02)$$
$$=0.5+P(-1.02<Z<0)$$
$$=0.5+P(0<Z<1.02)$$
$$=0.5+0.3461=0.8461$$



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

$$P(X>28)=P\left(\frac{X-\mu}{\sigma}>\frac{28-22.5}{4.4}\right)$$

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



4–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 .

مثال (3)

إذا كان الزمن اللازم لأداء خدمة معينة للعملاء بأحد البنوك يتوزع كالتوزيع الطبيعى بمتوسط 25 ثانية وانحراف معيارى 4 ثوانى . فإذا تم سحب عينة عشوائية من 25 شخص يقومون بأداء هذه المهمة ، فاحسب احتمال أن يكون متوسط الزمن اللازم لأداء هذه الخدمة البنكية :

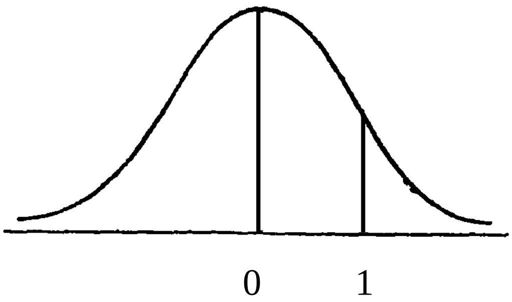
1. 26 ثانية أو أكثر .
2. بين 24 ، 27 ثانية .
3. 26 ثانية أو أقل .

4. 23 ثانية على الأقل .

الحل : $\mu = 25 \quad \sigma = 5 n = 25$

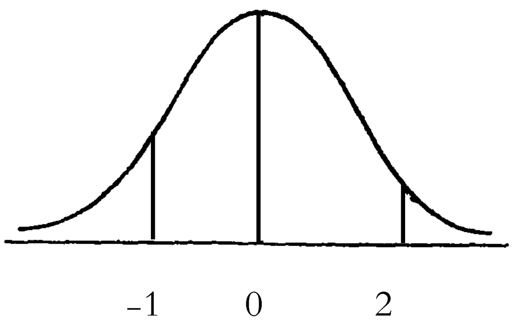
1. احتمال أن يكون متوسط الزمن اللازم لأداء الخدمة 26 ثانية أو أكثر

$$P(\overline{X} \geq 26) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{26-25}{5/\sqrt{25}}\right)$$
$$= P(Z \geq 1)$$
$$= 0.5 - P(0 \leq Z \leq 1)$$
$$= 0.5 - 0.3413 = 0.1587$$



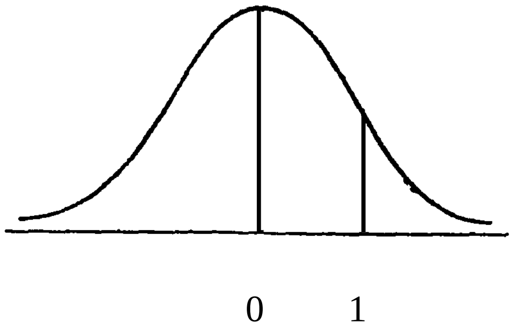
2. احتمال أن يكون متوسط الزمن اللازم لأداء الخدمة ما بين 24 , 27 ثانية

$$P(24 < \overline{X} < 27) = P\left(\frac{24-25}{5/\sqrt{25}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{27-25}{5/\sqrt{25}}\right)$$
$$= P(-1 < Z < 2)$$
$$= P(-1 < Z < 0) + P(0 < Z < 2)$$
$$= P(0 < Z < 1) + P(0 < Z < 2)$$
$$= P(0 < Z < 1) + P(0 < Z < 2)$$
$$= 0.3413 + 0.4772 = 0.8180$$



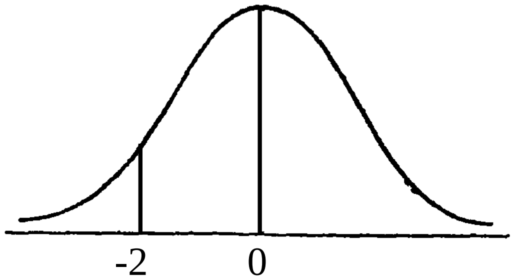
3. احتمال أن يكون متوسط الزمن اللازم لأداء الخدمة 26 ثانية أو أقل

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



4. احتمال أن يكون متوسط الزمن اللازم لأداء الخدمة 23 ثانية على الأقل

$$P(\overline{X} \ge 23) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \ge \frac{23-25}{5/\sqrt{25}}\right)$$
$$= P(Z \ge -2) = 0.5 + P(-2 \le Z \le 0)$$
$$= 0.5 + P(0 \le Z \le 2)$$
$$= 0.5 + 0.4772 = 0.9772$$



تمارين مختارة

التمرين الأول:

إذا كانت سرعة السيارات في طريق رئيسي بإحدى المدن عبارة عن متغير عشوائي يتبع التوزيع الطبيعي وسطه الحسابي 50 كلم /سا وانحرافه معياري 10 كلم / سا، أوجد ما يلي:

1 -احتمال أن تكون سرعة السيارة أقل من 40 كلم / سا

2- احتمال أن تتراوح سرعة السيارة بين 45 كلم / سا و 55 كلم / سا

3 - بافتراض وجود" مراقبة رادار "على هذا الطريق، ما هي أقصى سرعة مسموح السير بها مع العلم أن نسبة السائقين الذين لم يلتقطهم الرادار هي 85% ؟

التمرين الثاني:

أعطيت احدى الشعب امتحانا في الإحصاء من عشر علامات، وكانت النتائج تتدرج من الصفر حتى (10) ، وكان متوسط علامات الطلاب في هذا الامتحان 6.5 ، بانحراف المعياري 1.5 ، إذا افترضنا أن العلامات تتوزع طبيعيا، فأوجد ما يلي:

-النسبة المئوية لعدد الطلاب الذين تحصلوا على العلامة 7 فما فوق.

-أكبر علامة سجلها ال 20% من الطلاب ذوي العلامات المتدنية في الفصل.

التمرين الثالث:

تقوم مستثمرة فلاحية بإنتاج البرتقال بهدف التصدير، حيث يتبع وزن البرتقالة توزيعا طبيعيا وسطه الحسابي 200 غ وانحراف معياري 20 غ، إذا علمت أن شروط تصدير هذا البرتقال تتطلب أن يكون وزن البرتقالة الواحدة يتراوح بين 180 غ و 210غ

-أحسب نسبة البرتقال التي يمكن تصديرها وفق هذه المعايير.

-أحسب الكمية المصدرة الموافقة لهذه النسبة إذا كان حجم الإنتاج الكلي يساوي 10000 طن.

-إذا علمت أنه تم امضاء اتفاقية لتصدير 40% من البرتقال الأكبر وزنا نحو دول الاتحاد الأوروبي، فما هو أقل وزن محتمل للبرتقالة في الشحنة التي سيتم تصديرها ؟

1-مقدمة:

تستند معظم البحوث والدراسات العلمية على مبدأ دراسة مجموعة مؤشرات وقياسات مستخرجة من مشاهدات العينة ومحاولة تعميمها لتكون مؤشرات وقياسات المجتمع وهذا هو مفهوم الاستدلال الإحصائي حيث أن استخدام العينة في البحث العلمي بدلا م المجتمع الإحصائي له مبررات كثيرة منها ما يتعلق باختصار الوقت والجهد والتكاليف الذي يحققها استخدام العينة بدلا من دراسة المجتمع، كما أن بعض المجتمعات الإحصائية تكون غير محددة وغير مغلقة ولا يمكن الوصول إلى جميع مشاهداتها، وهنا أصبح استخدام أسلوب المعاينة ضرورة قصوى لإتمام الدراسة وجعل استخدام العينة أمرا ضروريا وحتميا، لذا لا بد أن نؤكد على اختيار الأسلوب العلمي لاختيار المشاهدات لهذه العينة وتحديد نوعها دون غيرها سوف يمكننا من اختيار عينة تحمل كل سمات المجتمع وخواصه أي أننا اخترنا عينة تكون بمثابة مجتمع مصغر عن المجتمع الأصلي وبذلك فإن التعميم لنتائج العينة يصبح صحيحا وممكناً.

2-أنواع العينة:

للوصول إلى الدقة المطلوبة في النتائج ينبغي اختيار العينة الناسبة لخصائص وطبيعة مجتمع الدراسة بحيث تحتوي على معلومات كافية عن المجتمع ولكي يكون من الممكن تعميم نتائجها على المجتمع.

وتنقسم العينات إلى نوعين رئيسيين:

أولا: العينات الاحتمالية أو العشوائية

وهي العينات التي لا يكون للباحث أي دخل في اختيار مفرداتها، أي أن العينات التي يتحقق فيها مبدأ تساوي الفرصة لظهور أي مفردة من مفردات المجتمع ضمن مفردات العينة، وهي من أهم أنواع العينات الاحتمالية (العشوائية):

1.2. العينة العشوائية البسيطة Simple Randon Sample:

إذا ا اتصف المجتمع الإحصائي بأنه مجتمع تما التجانس أي أن كل مشاهدة فيه لها خصائص وصفات لا تختلف عن المشاهدات الأخرى استنادا للدراسة أو البحث المطلوب أمكن سحب عينة عشوائية بسيطة من هذا المجتمع لأن كل

مفردة أو مشاهدة في المجتمع سيكون لها نفس الفرصة بالظهور في العينة، فعندما نريد اختيار عينة من الطلبة المدخنين في الجامعة فإن كل طلبة الجامعة المدخنين هم مشاهدات مجتمع البحث، وأصبح هنا المجتمع تام التجانس من حيث الظاهرة المدروسة (بعض النظر عن جميع المتغيرات الأخرى كالجنس والعمر وغيرها من المتغيرات) ولهذا النوع من العينات بعض الخصائص نذكر منها:

1) عدد العينات ممكن الاختيار بحجم (n) من مجتمع حجمه (N) هو:

$$C_n^N=\frac{N!}{(N-n)!n!} \qquad (1-1)$$

مثال (1):

في إحدى الكليات ستة أقسام علمية أرددنا اختيار قسمين لإجراء دراسة ما هي عدد العينات التي يمكن اختيارها بواقع قسمين فقط.

الحل:

نفترض أننا رمزنا لأقسام هذه الكلية بالرموز :

(F ,E, D,C,B,A)

ستكون العينات ممكنة الاختبار هي:

$$C_2^6=\frac{6!}{(6-2)!2!}$$

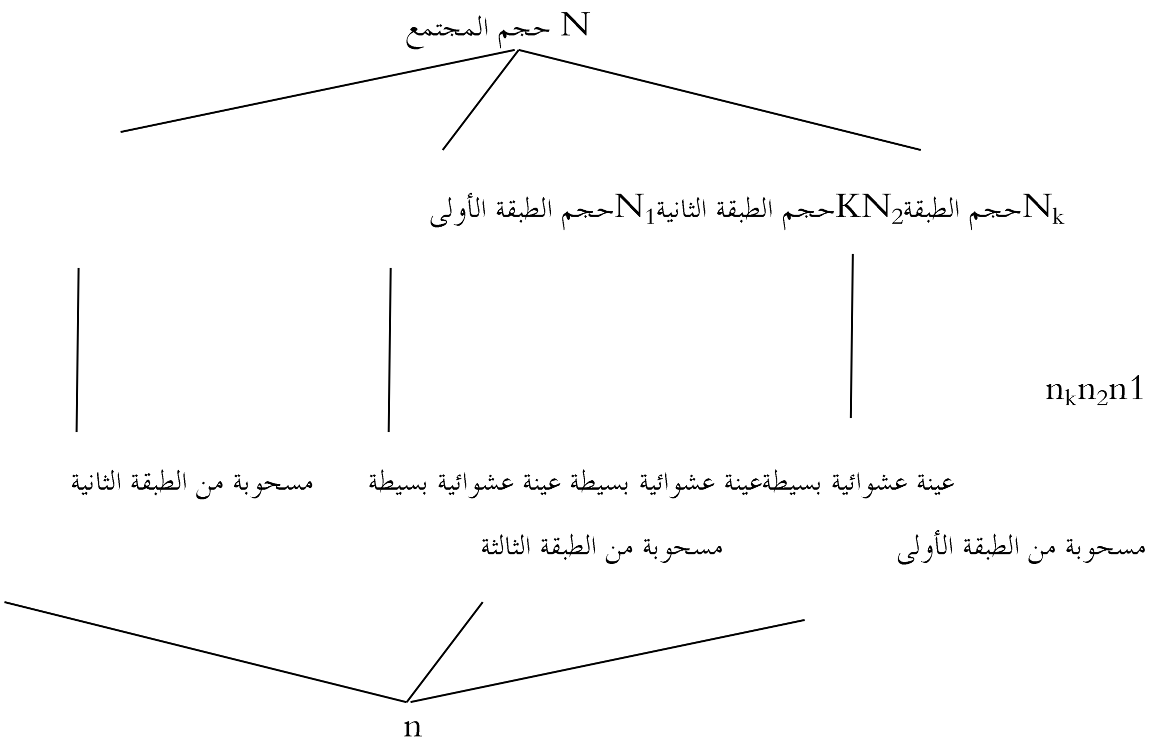
ويمكن توضيح هذه العينات الخمسة عشر كالتالي:

العينة	الأقسام	العينة	الأقسام
1	AB	9	BF
2	AC	10	CD

3	AD	11	CE
4	AE	12	CF
5	AF	13	DE
6	BC	14	DF
7	BD	15	EF
8	BE		

2.2. العينة الطبقية العشوائية:

المجتمع في هذه الحالة مقسم إلى مجموعات كل مجموعة تسمى طبقة وكل طبقة تامة التجانس في مفرداتها أو مشاهداتها، لذا يمكن سحب عينة عشوائية بسيطة من كل طبقة ليكون مجموع العينات المسحوبة من الطبقات هو العينة الطبقية العشوائية المطلوبة، لاحظ الشكل (1):



شكل رقم (1) العينة الطبقية العشوائية

لاحظ أن:

$$N=N_1+N_2+\ldots\ldots\ldots\ldots +N_k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1-2)$$

$$n=n_1+n_2+\ldots\ldots\ldots\ldots +n_k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1-3)$$

إن حجم العينة العشوائية المسحوب من الطبقة (i) يستخرج وفق العلاقة التالية:

$$n_i = N_i \times \frac{n}{N}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1-4)$$

حيث أن:

$i = 1,2,3\ldots\ldots,k$

$N =$ حجم المجتمع

$N_i =$ حجم الطبقة

$n =$ حجم العينة

مثال(2):

في أحد المعامل الإنتاجية (50) مهندس و(20) موظف إداري و(100) عامل ماهر و(30) عامل غير ماهر، وأردنا سحبة عينة طبقية عشوائية بحجم (20) من المنتسبين للمعمل لدراسة واقع الرضا الوظيفي لديهم.

وضح كيفية سحب العينة

الحل:

حجم المجتمع هو:

$N= 50+20+100+30=200$

حجم العينة المسحوبة من الطبقة الأولى (المهندسين):

$$n_1= 50x\{\frac{20}{200}\}=5$$

حجم العينة المسحوبة من الطبقة الثانية (الموظفين الإداريين):

$$n_1= 20x\{\frac{20}{200}\}=5$$

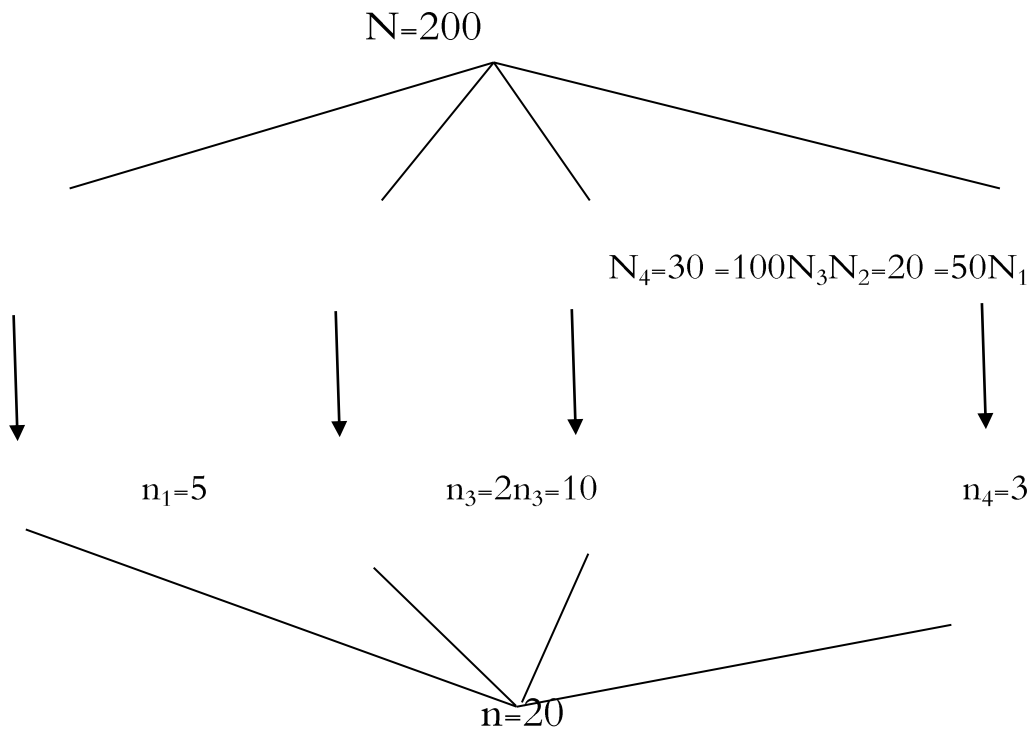
حجم العينة المسحوبة من الطبقة الثالثة (العمال الماهرين):

$$n_1= 100x\{\frac{20}{200}\}=5$$

حجم العينة المسحوبة من الطبقة الثالثة (العمال الغير الماهرين):

$$n_1= 30x\{\frac{20}{200}\}=3$$

ويمكن توضيح هذه العملية بالشكل التالي:



لاحظ أنه تم تحقيق التناسب الطبقي المطلوب في العينة حيث أن الطبقة الكبيرة (العمال الماهرين) كان لها أكبر إسهام والطبقة الصغيرة (الموظفين الإداريين) كان لها أقل إسهام.

3-2 العينة المنتظمة Systematic Sampling :

إذا كان مجتمع الدراسة يتصف إنه يمكن ترتيبه بشكل تصاعدي أو تنازلي تكون العينة المناسبة للسحب هي العينة المنتظمة، ومثال على ذلك إذا أردنا سحب عينة من دفتر الشيكات فإن الشيكات مرتبة ترتيبا تصاعديا يمكن الاعتماد عليه في سحب عينة منتظمة.

لذا فإن هذه العينة تتطلب ترقيم مشاهدات المجتمع ثم اختبار مشاهدة من أول K من المشاهدات، وإذا أردنا اختيار المشاهدة الثانية فإننا نفتقر إلى:

$$K+\frac{N}{n} \qquad \text{.........(1-5)}$$

أي أن مقدار القفزة أو الفترة بين المشاهدة الأولى والمشاهدة الثانية هو $\frac{N}{n}$ و هكذا نضيف $\frac{N}{n}$ للمشاهدة الثانية لنجد المشاهدة الثالثة وهكذا لبقية المشاهدات المختارة.

مثال (3):

لدراسة تتعلق بواقع الإنفاق العائلي أردنا سحب عينة بحجم عشرة دور في حي سكني رقمت دوره من 1 على 200 ما هيي المشاهدات (الدور) التي سيتم اختيارها من العينة.

الحل:

عدد المجموعات التي يمكن تقسيم المجتمع لها هي:

$$\frac{N}{n}=\frac{200}{10}=20$$

عدد المجاميع هي عشرون مجموعة وطول الفترة أيضا هو 20.

من المجموعة الأولى التي تحتوي على الدور المرقمة من 1 إلى 20 نختار أحد الدور عشوائيا وليكن الدار رقم 2 ومن المجموع الثانية نختار الدار رقم 20+2=22 وهكذا من المجموعة الثالثة نختار الدار رقم 22+20=42 ونستمر لبقية المجموعات بإضافة 20 للدار الذي اخترناه من المجموعة السابقة فيكون لدينا عينة عشوائية منتظمة مفرداتها:

2، 22، 42، 62، 82، 102، 122، 142، 162، 182.

2-4العينة متعددة المراحل:

تسحب هذه العينة عندما يكون المجتمع كبيرا جدا إذ يتم تقسيمه إلى وحدات تسمى بالوحدات الأولية يتم سحب عينة عشوائية منها. كمرحلة أولى يتم تقسيم المجتمع إلى وحدات تسحب منها عينة للمرحلة الأولى ثم تقسم مفردات العينة المختارة من المرحلة الأولى إلى وحدات أصغر تسمى بالوحدات الثانوية وتأخذ منها عينة للمرحلة الثانية وهكذا نستمر بالتقسيم والاختيار حتى نصل إلى العينة الأخيرة التي تحتوي على مشاهدات تستنبط منها المعلومات المطلوبة.

مثال (1):

قررت إحدى المؤسسات منح اثنين من موظفيها جائزة تميز وكان عدد المؤهلين لهذه الجائزة ستة أشخاص.

1)ما نوع السحب في هذه الحالة.

2)كم عينة (اختيار) يمكن اختيارها في هذه الحالة.

الحل :

1. العينة المسحوبة هي عينة عشوائية بسيطة لأن المجتمع تام التجانس ولا فرق بين الموظفين الستة من حيث تأهلهم للجائزة.
2. يمكن حساب عدد العينات العشوائية التي يمكن اختيارها باستخدام الصيغة (1.1) أي:

$$C_2^6=\frac{6!}{(6-2)!2!}=15$$

ويمكن توضيح العينات الخمسة عشر إذا افترضنا أن الموظفين الستة هم (F ,E, D, C, B, A):

العينة	الموظفين	العينة	الموظفين	العينة	الموظفين
1	AB	6	BC	11	CE
2	AC	7	BD	12	CF
3	AD	8	BE	13	DE
4	AE	9	BF	14	DF
5	AF	10	CD	15	EF

مثال (2):

في إحدى المعامل يوجد 40 عاملا ماهرا و 60 عاملا غير ماهر نريد سحب عينة عشوائية من 20 عاملا

1.ما نوع العينة المسحوبة.

2.كيف يتم سحبها.

الحل:

1.نوع العينة المسحوبة عينة طبقية عشوائية لأن المجتمع مقسم إلى طبقتين وكل طبقة هي تامة التجانس بالنسبة لمفرداتها.

2.وزن طبقة العمال الماهرين:

$$W_1=\frac{40}{60+40}=0.40$$

العمال الماهرين في العينة المطلوبة هو:

$$n_1=n\times w_1=20\times 0.40=8$$

وزن طبقة العمال غير الماهرين:

$$W_2=\frac{60}{60+40}=0.60$$

عدد العمال غير الماهرين في العينة المطلوبة:

$$n_2=n.w_2=20(0.60)=12$$

مثال (3):

أريد سحب عينة بحجم خمسة شيكات من دفتر أحد العملاء في البنك لدراسة إنفاقه الفردي وكانت الشيكات مرقمة من 1 إلى 30.

1.ما نوع العينة الواجب سحبها.

2.كيف سيتم السحب.

الحل:

1.بما أن المفردات (الشيكات) مرتبة تصاعديا ستكون العينة المختارة عينة منتظمة.

2. $K=\frac{N}{n}=\frac{30}{5}=6$

وعليه تكون مفردات العينة إذا اخترنا المفردة رقم (1) ستكوم المفردة رقم (2) هي:

$$1+6=7$$

أما المفردة الثالثة فهي:

$$7+6=13$$

والمفردة الرابعة:

$13+6=19$

والمفردة الخامسة:

$19+7=26$

أي أن جميع الشيكات المختارة هي: (1، 7، 13، 19، 26).

مثال (4):

حدد نوع العينة للحالات التالية:

1. اختيار طلبة من الفصل للمشاركة في وضع البرنامج الإمتحاني.

الجواب:عينة عشوائية بسيطة.

2.لدراسة واقع الإنفاق العائلي في مناطق مختلفة من محتفظة عمان.

الجواب: عينة طبقية عشوائية حيث يتم اختبار عينة من عوائل مرتفعة الدخل وعينة من عوائل متوسطة الدخل وعينة من عوائل واطئة الدخل.

3.دراسة الإنفاق الصحي والمطلوب سحب عينة من (100) دار مرقمة من (1-100).

الجواب: عينة غير منتظمة.

4.سحب عينة لدراسة إمكانية تصدير محصول زراعي.

الجواب: عينة متعددة المراحل.

3-توزيعات المعاينة للعينة العشوائية:

إن الهدف من أية عملية معاينة هي تقدير المعالم الأساسية للمجتمع الإحصائي الكلي(المتوسط الحسابي، الانحراف المعياري) عن طريق معالم العينة، ومن حساب المعالم المقابلة لها في المجتمع.

3-1 توزيعات المعاينة للأوساط الحسابية:

كانت لدينا مجموعة من العينات مأخوذة من مجتمع ما وسطه الحسابيμ وانحرافه المعياري σ، فإن معظم الأوساط الحسابية لهذه العينات تختلف عن بعضها. نسمي التوزيع الاحتمالي لمتوسطات هذه العينات بتوزيع المعاينة للوسط الحسابي، ولسوف نجد أن لهذه الأوساط (التوزيع) متوسط حسابي يرمز له ب$\mu_{\bar{x}}$، حيث أن:

$$\mu_{\bar{x}}=\frac{\bar{x}1+\bar{x}2+\cdots+\bar{x}n}{C_N^n}=\mu$$

انحرافه المعياري:

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{N}}\sqrt{\frac{N-n}{n-1}}$$ (سحب بدون إرجاع) في حالة وجود مجتمع محدود

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{N}}$$ في حالة وجود مجتمع غير محدود (سحب مع الإرجاع)

ولقيم n الكبيرة ($n>30$) فإن توزيع المعاينة للأوساط الحسابية يقترب من التوزيع الطبيعي بمتوسط$\mu_{\bar{x}}$ وانحراف معياري$\sigma_{\bar{x}}$، إن هذه الخاصية في الحقيقة تنطبق على المجتمعات الغير المحدودة والتي تثبت دقة التقريب من التوزيع الطبيعي، وهي حالة خاصة من نظرية النهاية المركزية، التي سنراها لاحقا.

➢ **مثال:**

مجتمع إحصائي حجمه $N=05$يمثل أعمار 05 أطفال هي:

$X1=6, x2=8, x3=10, x4=12, x5=14$

لنقسم المجتمع إلى مجموعة من العينات، حجم كل عينة 2 وذلك بالسحب بدون إرجاع. إن غدد العينات اللازم تشكيلها في هذه الحالة هي:

$$C_5^2=\frac{5!}{2!(5-2)!}=10$$

وأن : $\mu=\frac{6+8+10+12+14}{5}=10$

كما أن : $\sigma=2.82$

العينة	الوسط الحسابي للعينة
(6,8)	7
(6,10)	8
(6,12)	9
(6,14)	10
(8,10)	9
(8,12)	10
(8,14)	11
(10,12)	11
(10,14)	12
(14,12)	13

نجد أن:

$$\mu_{\bar{x}} = \frac{7+8+9+\cdots+12+13}{10} = 10 = \mu$$

أما الانحراف المعياري في حالة مجتمع محدود:

$$\sigma\bar{x} = \frac{1.82}{\sqrt{5}}\sqrt{\frac{5-2}{5-2}} = 1.73 \neq \sigma$$

- **حالة عينتين من مجتمعين مستقلين:**

إذا كانت $x_1.,.......x_2.......x_n$ عينة مأخوذة من توزيع وسطه الحسابي μ_1 وانحرافه المعياري σ_1، وكانت $y_1,y_2....y_n$ عينة أخرى من توزيع أخر وسطه الحسابي μ_2 وانحرافه المعياري σ_2، فإنه يمكن الحصول على توزيع الفروق للأوساط الحسابية وفق العلاقة التالية:

$$\mu\bar{x}-\mu\bar{y}=\mu_1-\mu_2$$

أما الإنحراف المعياري لهذا الفرق فيعرف بالعلاقة التالية:

- في حالة مجتمع غير محدود:

$$\sigma\bar{x}-\bar{y}=\sqrt{\sigma^2\bar{x}+\sigma^2\bar{y}}=\sqrt{\frac{\sigma^2 1}{n1}+\frac{\sigma^2 2}{n2}}$$

- في حالة مجتمع محدود:

$$\sigma\bar{x}-\bar{y}=\sqrt{\sigma^2\bar{x}+\sigma^2\bar{y}}=\sqrt{\frac{\sigma^2 1}{n1}.\frac{N1-n1}{\mathrm{N1}-1}}+\sqrt{\frac{\sigma^2 2}{n2}.\frac{N2-n2}{\mathrm{N2}-1}}$$

مثال:

أخذت عينة حجمها 30 وحدة من توزيع وسطه الحسابي 75 وتباينه 25، وأخذت عينة ثانية حجمها 60 من توزيع أخر مستقل وسطه الحسابي 60 وتباينه 15. أوجد الفروق للوسط الحسابي والانحراف المعياري لتوزيع المعينة للتوزيعين.

➢ **الحل:**

بما أن المجتمعين غير محدودين فإن:

$$\mu\bar{x}-\mu\bar{y}=75-60=15$$

أما الفروق بين التباينين:

$$\sigma\bar{x}-\bar{y}=\sqrt{\frac{25}{30}+\frac{15}{60}}=1.04$$

- **نظرية النهاية المركزية:**

إذا كان حجم العينة كبيرا ($n>30$) فإن توزيع المعينة للمتوسط الحسابي يقترب من التوزيع الطبيعي، بغض النظر عن شكل المجتمع الأصلي، لذلك يمكن حساب احتمال أن يكون $\bar{x}$ لعينة عشوائية داخل فترة معينة من خلال حساب القيمة الإحصائية ل Z وفق العلاقة الشهيرة التالية:

$$Z=\frac{\bar{x}-\mu\bar{x}}{\sigma\bar{x}}$$

- **مثال:**

ضمت إحدى الشركات سيارة بحيث أ أكبر حمولة لها هي 3000 كغ وتتسع إلى 30 راكبا. إذا علمت أن الأوزان تخضع للتوزيع الطبيعي الذي وسطه 70 كغ وانحرافه المعياري 54.77. أحسب احتمال أن تحمل هذه السيارة أكثر من حمولتها.

- **الحل:**

نلاحظ أن: $100= \frac{30000}{30} = \bar{x}$ والمطلوب إيجاد $p(\bar{x}>100)$

الحل:

$$p(\bar{x}>100)=P\left(\frac{\bar{x}-\mu\bar{x}}{\sigma\bar{x}}>\frac{100-70}{\frac{54.77}{\sqrt{30}}}\right)$$

$$p(z>3)= 0.5-0.4987=0.00013$$

- **في حالة مجتمعين مستقلين:**

إذا كانت $x_1, x_2, \ldots\ldots x_n$ عينة مأخوذة من توزيع وسطه الحسابي μ_A وانحرافه المعياري σ_A ، وكانت $y_1, y_2, \ldots. y_n$ عينة أخرى من توزيع أخر مستقل ووسطه الحسابي μ_B وانحرافه المعياري σ_B ، فإننا يمكن إيجاد القيمة الإحصائية Z وفق العلاقة التالية:

$$Z=\frac{(\bar{x}_A-\bar{x}_B)-\mu_{\bar{X}A-\bar{y}B}}{\sigma_{\bar{X}A-\bar{y}B}}$$

- **مثال:**

تنتج الشركة A مصابيح كهربائية متوسطة مدة حياتها 1400 ساعة، بانحراف معياري 200 ساعة، وتنتج الشركة B مصابيح مماثلة متوسط مدة حياتها 1200 ساعة بانحراف معياري 100 ساعة. قمنا باختيار عينة عشوائية حجمها 125 وحدة من كل شركة. أوجد احتمال أن الشركة A تنتج مصابيح كهربائية متوسطة مدة حياتها على أكبر من الشركة B ب:

-160 ساعة، 250 ساعة.

- **الحل**

$\mu\bar{x}_{A}-\bar{x}_{B}=\mu\bar{x}_{A}-\mu\bar{x}_{B}=1400-1200=200.$

$$\sigma\bar{x}_{A}-\bar{x}_{B}=\sqrt{\sigma_{\bar{x}A}{}^{2}+\sigma_{\bar{x}B}{}^{2}}=\sqrt{\frac{\sigma_A{}^2}{n_A}+\frac{\sigma_B{}^2}{n_B}}=\sqrt{\frac{200^2}{125}+\frac{100^2}{125}}=20$$

احتمال أن الشركة A تنتج مصابيح كهربائية متوسطة مدة حياتها على أبر من الشركة B ب 160 ساعة هو:

$$P((\bar{x}_A-\bar{x}_B)>160)=p(\frac{(\bar{x}_A-\bar{x}_B)-\mu_{\bar{X}A-\bar{y}B}}{\sigma_{\bar{X}A-\bar{y}B}}>\frac{160-200}{20})$$

$=p(Z>-2)=0.5+0.4772=0.9772$

واحتمال أن الشركة A تنتج مصابيح كهربائية متوسط مدة حياتها على أكبر من الشركة B ب 250 ساعة

هو:

$$P((\bar{x}_A-\bar{x}_B)>250)=p(\frac{(\bar{x}_A-\bar{x}_B)-\mu_{\bar{X}A-\bar{y}B}}{\sigma_{\bar{X}A-\bar{y}B}}>\frac{250-200}{20})$$

$=p(Z>2.50)=0.5+0.4938=0.0062$

2.3- المعاينة باستعمال توزيع T :

إذا كان حجم العينة صغيرا ($n<30$) فإنه يمكن استخدام توزيع كبديل للتوزيع الطبيعي، ويمكن حساب احتمال $\bar{x}$ لعينة عشوائية، داخل فترة معينة، من خلال حساب القيمة الإحصائية ل t وفق العلاقة التالية:

$T=\frac{\bar{x}-\mu\bar{x}}{S\bar{x}}$ بدرجات حرية n-1

مثال:

تخضع علامات طلبة معهد ما للتوزيع الطبيعي بوسط حسابي 70 أخذت عينة عشوائية حجمها 16 طالبا بانحراف معياري 8.

أحسب احتمال أن تزيد علامات الطلبة عن 78 درجة.

- **الحل:**

المطلوب هو حساب:

$$p(\bar{x}>74)=\frac{\bar{x}-\mu\bar{x}}{S\bar{x}}=\frac{74-70}{\frac{8}{\sqrt{16}}}= \quad p(t>2)=0.025$$

عدد درجات حرية 15 نجد المساحة 0.975

3.3-توزيعات المعاينة للتباينات:

إذا كانت $x_1,x_2......x_n$عينة مأخوذة من توزيع وسطه الحسابي μ وتباينه σ^2 معلومة، فإن تباين العينة يخضع لتوزيع مربع كاي وفق العلاقة التالية:

$$\chi^2=\frac{(n-1)S^2}{\sigma^2} \text{ بدرجات حرية } n-1$$

- **مثال:**

إذا كانت $x_1,x_2......x_n$عينة عشوائية من توزيع طبيعي $N(\mu,9)$ تباينها S^2، أوجد النقطة C بحيث تكون $P(S^2\leq c)$

- **الحل:**

$$P(S^2\leq c)= p\left(\frac{n-1}{\sigma^2}.S^2\leq\frac{c(n-1)}{\sigma^2}\right)$$

$$P\left(\chi^2_9\leq\frac{c(10-1)}{9}\right)= 0.95\Rightarrow P(\chi^2_9\leq c)=0.95$$

وباستعمال جدول كاي مربع نجد أن $c=16.91$

- **حالة σ مجهولة:**

إذا كانت $x_1,x_2......x_n$عينة عشوائية من توزيع طبيعي وسطه الحسابي μ وتباينه σ^2، وكان S^2هو تبيان العينة بحيث σ^2 مجهولة فإن تباين العينة يخضع لتوزيع t وفق العلاقة التالية:

$t=\frac{X-\mu}{\frac{s}{\sqrt{n-1}}}$ بدرجات حرية n-1

4.3- توزيعات المعاينة بالنسب:

لنفترض أنه لدينا مجتمع غير محدود به جميع العينات الممكنة ذات الحجم n، ولكل عينة نسبة النجاح p حيث

أن p+q=1، فإن توزيع المعاينة للنسب يكون بمتوسط μ_p وانحراف معياري قدره σ_p، حيث:

$$\mu_p = p$$

$$\sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

أما إذا كان المجتمع محدود (المعاينة بدون إرجاع) فإن العلاقة تصبح كما يلي:

$$\mu_p = p$$

$$\sigma_p = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$$

- **مثال:**

نرمي قطعة نقود متزنة 120 مرة، أوجد احتمال الحصول على 40% و60% صورة.

- **الحل:**

إن احتمالات النجاح (الحصول على صورة) p=1/2 واحتمالات الفشل هي q=1/2 وحيث أن $n \geq 30$ فإن احتمالات الحدين تقترب من التوزيع الطبيعي ويكون:

40 %x120=48 و   60%x120=72

فإن المطلوب حساب $p(48 \leq x \leq 72)$؛ وبما أننا سنقوم بعملية تقريب توزيع منفصل إلى توزيع متصل فإن الاحتمال السابق يأخذ الشكل التالي: $p(45.5 \leq x \leq 72.5)$

$$\mu_p = np = 120.1/2 = 60$$

$$\sigma = \sqrt{npq} = \sqrt{120 \times 0.5 \times 0.5}$$

$$= 5.48$$

$$P(47.5\leq x\leq 72.5)=p\left(\frac{47.5-60}{5.48}\leq z\leq\frac{72.5-60}{5.48}\right)$$

$$P(-2.28\leq Z\leq 2.28)=0.9774$$

تمارين مختارة

التمرين الأول:

مجتمع يتكون من 12000 عنصرا بوسط حسابي 100 وانحراف معياري 60. أوجد الوسط الحسابي والانحراف المعياري لتوزيع المعاينة للمتوسط الحسابي، عندما يكون حجم العينة 100.

التمرين الثاني:

مجتمع مكون من 3000 امرأة متوسط أعمارها 25 سنة. فإذا كانت أوزان هؤلاء النسوة تتبع التوزيع الطبيعي بوسط 58 كلغ وانحراف معياري 4. أخذت عينة عشوائية حجمها 25 امرأة لكل عينة. أوجد المتوسط الحسابي والانحراف المعياري لتوزيع المعاينة، وذلك في حالة:

-السحب بدون إرجاع -السحب بالإرجاع.

التمرين الثالث:

إذا كان متوسط وزن 500 كرة هو 5.02 كلغ، بانحراف معياري قدره 0.3 كلغ. أخذت عينة عشوائية حجمها 100؛ أوجد احتمال أن يكون وزن الكريات: -محصورا بين 496 و500 كلغ – أكبر من 510 كلغ.

التمرين الرابع:

بينت الإحصائيات في بلد ما أن 64% من السكان يملكون سيارة سياحية، أخذت عينة عشوائية مكونة من 225 شخص، أوجد احتمال أن نسبة الذين يملكون سيارة سياحية: -محصورا بين 40% و 70%

-أكبر من 60% -أقل من 25%.

التمرين الخامس:

ماكنة صناعية مختصة في إنتاج قطع حديد دائرية متوسطة قطرها 25 مم. بعد مدة أصاب الماكنة عطل وتم تصليحه؛ أخذت عينة عشوائية من منتوج الماكنة بعد تصليحها وتم قياس الأقطار المنتجة فكانت النتائج كما يلي (بالمم):

22، 23، 21، 25، 24، 22، 26، 21.

هل يمكن الجزم بأن الماكنة مازالت تعمل بشكل جيد، عند احتمال 95%، بفرض أن أقطار القطع تتبع التوزيع الطبيعي؟

التمرين السادس:

أظهرت دراسة طبية وجود نوعين من مرض الروماتيزم، االتهابي وغير االتهابي، من أصل 220 مصاب بمرض الروماتيزم، وجد أن 167 شخص مصابون بالنوع اإللتهابي.

-أعطي تقديرا نقطيا وتقديرا بمجال لنسبة الأشخاص المصابون بالنوع اإللتهابي.

قامت هذه الدراسة بتحليل المناعة X في دم المصابين بالنوعين من المرض، فكانت النتائج كالتالي:

	النوع اإللتهابي	النوع الغير إلتهابي
$\sum X$	420	104
$\sum X^2$	100	292

- أوجد مجال الثقة لمتوسط عامل المناعة $X\alpha$ = (10%) لدى النوعين من المرض.ماذا تقترح؟

التمرين السابع:

ينتج مصنع للأدوية نوعا من الدواء يحتوي على مادة فعالة ويجب أن تكون هذه المادة محددة بشكل دقيق. ولدراسة هذه الدقة قمنا بتحليل عينة حجمها 25 حبة. أحسب احتمال أن لا يزيد الانحراف المعياري لكمية هذه المادة في الحبوب عن 1.30 غ، علما أن الانحراف المعياري لوزن المادة الفعالة في إنتاج المصنع كله هو 1.20 غ.

مقدمة:

يهتم الإستدلال الإحصائي بالتقدير واختبار الفروض، والتقدير هو عملية استنتاج أو تقدير أحد معالم المجتمع (مثل الوسط الحسابي أو الانحراف المعياري) من الإحصاء المناظر والخاص بعينة مسحوبة م المجتمع.

ولكي يكون التقدير (واختبار الفروض) سليما، ينبغي أن يبنى على عينة من المجتمع. ويمكن تحقيق ذلك بالمعاين العشوائية حيث يكون لكل مفردة في المجامع فرصة متكافئة للدخول في العينة.

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

ففي حالة تقدير النقطة نحصل على قيمة واجدة من العينة، واستخدم هذه القيمة للوحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة فمثلا لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. ومثال آخر لو أخذنا نسبة الناخبين في العينة الذين يرصدون مرشحا معينا كتقدير لهذه السبة في المجتمع تكون حصلا على تقدير لنسبة في مجتمع الناخبين. وسوف نناقش في هذا الفصل التقدير بنقطة الذي يسمى أحيانا التقدير النقطي، وهناك عدة طرق لهذا التقدير، وجميعها تؤدي إلى نتيجة واحدة من حيث إيجاد المقدر المطلوب ومن هذه الطرق طريقة العزوم، طريقة مربع كاي، طريقة المسافة الصغرى، وسوف نناقش هنا طريقتين مهمتي هما طريقة العزوم وطريقة الترجيح واللتين تنسبان إلى العالم فيشركونهما من الطرق التي يكثر استخدامها وتعطي نتائج جيدة في حساب وإيجاد المقدرات الثابتة.

I. التقدير بنقطة Point Estimation

1-1 طريقة العزوم Methode of Moments:

إحدى الطرق لاختيار مقدر معلمة ما للمجتمع أن نأخذ مباشرة نظيرتها في العينة، وإذا كان هذا المقدر لا يتصف بالخصائص المطلوبة نجري عليه تعديلا مناسبا ليصبح مقدارا جيدا (سنطرق في الفصل الثالث إلى خواص المقدرات).

ليكن المطلوب تقدير عدد K من معلم المجتم$Ø1$, $Ø2,......,Øk$, نكون جملة معادلات عددها K، تتضمن كل معادلة مساواة العزم المرتبط بالأصل من الدرجة K لمتغير المجتمع X:

$$\mu'_K = E(x^k)$$

بنظيره لمتغير العينة x:

$$M'=(\frac{1}{n})\sum_{i=1}^{n} x_i^{k} \quad , k=1,2.....,k$$

مثال (1-1)

ليكن $x\approx B(20\ ,p)$ (توزيع برنولي)، أوجد تقدير p بطريقة العزوم للعينة x من المتغير العشوائي X.

الحل:

لدينا عدد المعالم المراد تقديرها K=1، لذا نحتاج إلى معادلة واحدةμ =20p ومنه μ /20=p، وعليه يكون $\hat{p}$ مقدراً للـ p ويحسب كما تأتي:

$\hat{p}=\frac{m}{20}$، (حيث $m=\mu$ يمثل العزم الأول).

في حالة تقدير معلمتين للمجتمع نحتاج أن نستعمل جملة المعادلتين:

$$\mu'_2=m'_2,\ \mu=m$$

وكما في المثال التالي:

مثال (1-2)

ليكن$x\sim N(\mu;\sigma^2)$ أوجد تقديرا للمعالم μ وσ^2؟

الحل:

هنا$K=2$ وهو يمثل عدد المعالم للتوزيع، سنحتاج إلى حل جملة معدلتين التاليتين:

في حالة المجتمع يكون:

$$M'_1=E(x)=\mu$$

$$M'_2=E(x^2)=\mu^2+\sigma^2$$

أما في حالة العينة فيكون:

$$m'_1\frac{1}{n}\sum_{i=1}^{n} xi=\bar{x} \Rightarrow \bar{x}=\mu$$

$$m'_2=\frac{\sum_{i=1}^{n} xi^2}{n}$$

وعليه نستنتج أن:

$\mu = \bar{x}$ (تقديرا لمعلمة المتوسط) ....(1)

$\mu^2+\sigma^2=\frac{1}{n}\sum_{i=1}^{n} xi^2$..................(2)

وبحل المعادلتين نحصل من المعادلة (3) على:

$\sigma^2=\frac{1}{n}\sum_{i=1}^{n} x_i^2-\mu^2$ وبتعويض المعادلة (1) في (3) نحصل على:

$\widehat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n} x_i^2-(\bar{x})^2=S^2$

$\widehat{\sigma}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{xi}-(\bar{x})^2)}= S$

وهو يمثل تقديرا للمعلمة σ.

وهذه المقدرات هي مقدرات متحيزة كما سيأتي شرحه لاحقا.

مثال (1-3)

ليكن $X\approx p(\lambda)$، حيث إن X متغير عشوائي (r.v)، يتوزع توزيع بواسون، أوجد تقدير المعلمة λ ؟

الحل:

من العزم الأول نحصل على:

$m'_1=E(X)= \lambda$

كذلك فإن: $m'_1=\frac{1}{n}\sum_{i=1}^{n} xi=\bar{x}$

وبمساواة العزمين $M'_1=m'_1$ نحصل على $\bar{x}\lambda=$ أي أن المقدار لـ λ هو:

$\widehat{\lambda}=\bar{x}= \lambda$

1-2 طريقة الاحتمالية العظمة(طريقة الترجيح): Methode Of Maximum Likelihood

تعتبر طريقة الاحتمالية العظمى من أقوى الطرق الإحصائية في إيجاد تقدير لعالم المجتمع سواءا أكان ذلك للتوزيعات المتقطعة أو المستمرة، وقد وضعها العالم الإنجليزي فيشر (R.A.Fisher)، والمسمى أبو الإحصاء، وهي تؤدي إلى ننفس

النتائج التي أدت إليها طريقة العزوم وإذا لم تتطابق النتائج فتكون هذه الطريقة هي الأفضل. وتسمى الدالة التي يسحب منها التقدير بدالة الترجيح (Likelihood function) ويعرف بالشكل التالي:

هي دالة الكثافة(p.d.f) أو دالة الكتلة الإحتمالية المشتركة (p.m.f) لمفردات عينة عشوائية $x_1,x_2,\ldots\ldots x_n$، وحجمها n بمعلمة θ وتكتب الدالة كالتي:

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

$$=\prod_{i=1}^{n} f(xi,\theta)$$

حيث $f(x\theta)$ هي دالة الاحتمال بالمعلمة θ

1-2-1 طريقة الحصول على التقدير $\hat{\theta}$:

تشتق دالة الترجيح بالنسبة للمعلمة θ ونساويها بالصفر ومن ثم نحل المعادلة، أما في حالة وجود عدة معالم مجهولة مثل $\theta^1, \theta_2,\ldots\ldots \theta_n$ فإننا نشتق لكل معلمة من المعالم فنكون عدة معادلات، ومن ثم نقوم بحل هذه المعادلات للحصول على تقدير مقابل لكل معلمة، أي أن:

$$\frac{\partial L(\theta)}{\partial\theta_1}=0, \frac{\partial L(\theta)}{\partial\theta_2}=0, \ldots\ldots\ldots, \frac{\partial L(\theta)}{\partial\theta_n}=0$$

$$\frac{\partial^2 L(\theta)}{\partial{\theta_1}^2}<0, \frac{\partial^2 L(\theta)}{\partial{\theta_2}^2}<0, \ldots\ldots\ldots, \frac{\partial^2 L(\theta)}{\partial{\theta_{n1}}^2}<0$$

ولتسهيل الحسابات نضع $In(L()) =K(\theta)$

مثال (1-4)

ليكن x متغيرا عشوائيا له التوزيع الطبيعي الذي متوسطه μ وتباينه σ^2 أوجد تقدير للمعلمة μ و σ^2 باستخدام طريقة الإحتمالية العظمى؟

$$L(\mu,\sigma^2)=\prod_{i=1}^{n} f(x_i,\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$=\left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}$$

وبالتالي فإن:

$\ln L(\mu,\sigma^2) = -n\ln\sigma^2 - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

$= -\frac{n}{2}\ln\sigma^2 - n\ln\sqrt{2\pi}\,\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

ولإيجاد قيم μ,σ^2 التي تجعل هذه الدالة نهاية عظمى فإننا نفاضل هذه الدالة جزئيا مرة بالنسبة إلى μ وأخرى بالنسبة إلى σ^2 ومن ثم مساواتهم بالصفر أي أن:

$\frac{\partial}{\partial\mu}[lnL(\mu,\sigma^2)]=0, \qquad \frac{\partial}{\partial\sigma^2}[lnL(\mu,\sigma^2)]=0,$

وبالتالي فإن:

$\frac{\partial}{\partial\mu}[lnL(\mu,\sigma^2)] = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)\ (-1)=0$

$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0, \sigma^2 \neq 0$

$\sum_{i=1}^{n}(x_i-\mu) = 0$

$\sum_{i=1}^{n}x_i = n\mu$

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i = \bar{x}$

وهذا يمثل تقديرا لمعلمة متوسط التوزيع.

وبإجراء الاشتقاق بالنسبة للمعلمة λ نحصل على:

$\frac{\partial K(\lambda)}{\partial\lambda} = \frac{\sum_{i=1}^{n}x_i}{\lambda} - n = \frac{\sum_{i=1}^{n}x_i}{\lambda} = n$

$n\,\lambda = n\sum x_i \qquad \Rightarrow \bar{x} = \frac{\sum_{i=1}^{n}x_i}{n} = \lambda \Rightarrow \bar{x} = \lambda$

$\lambda = \bar{x}$

وهذا يمثل تقدير للمعلمةλ ، ويعني بأن متوسط العينة $\bar{x}$ هو تقدير مرجح للمعلمة λ في توزيع بواسون، وهو نفس التقدير الذي توصلنا إليه في طريقة العزوم في مثال (1-5)

مثال (5-1)

إذا كانت $x_1, x_2, \ldots\ldots x_n$ عينة عشوائية حجمها n من المتغير العشوائي x الذي يتوزع توزيع برنولي،

فأوجد تقدير للوسط p مستخدما الطريقة الاحتمالية العظمى.

الحل:

من دالة توزيع برنولي $f(x)=p^x.q^{1-x}$ نحصل على دالة الترجيح التالية:

$$L(p)=\prod_{i=1}^{n} p^{xi}.q^{1-xi} \Rightarrow L(p)=p^{\sum_{i=1}^{n} xi}(1-p)^{\sum_{i=1}^{n} xi}$$

$$K(p)= In(L(p))=In(p^{\sum_{i=1}^{n} xi}(1-p)^{\sum_{i=1}^{n} xi})$$

$$K(p)=\sum_{i=1}^{n} xi \ln p+(n-\sum_{i=1}^{n} xi)\ln(1-p)$$

ثم تشتق بالنسبة لـ p فنحصل على:

$$\frac{\partial K(\lambda)}{\partial p}=\frac{\sum_{i=1}^{n} x_i}{p}-=\frac{-n\sum_{i=1}^{n} x_i}{1-n}=0$$

$$\frac{\sum_{i=1}^{n} x_i}{p}-=\frac{-n\sum_{i=1}^{n} x_i}{1-n}=0 \qquad \Rightarrow \frac{(1-p)\sum_{i=1}^{n} x_{i-p(n-\sum_{i=1}^{n} x_i)}}{p(1-p)}$$

$$\frac{(1-p)\sum_{i=1}^{n} x_{i-p(n-\sum_{i=1}^{n} x_i)}}{p(1-p)}=0 \qquad \Rightarrow (1-p)\sum_{i=1}^{n} x_i-p(n-\sum_{i=1}^{n} x_i)=0$$

$$\sum_{i=1}^{n} x_i-p\sum_{i=1}^{n} x_i - np + p\sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} x_i=np \quad \Rightarrow \quad p=\frac{\sum_{i=1}^{n} x_i}{n}=\bar{x}$$

$$\hat{p}=\bar{x}$$

وهذا يعني بأن الوسط الحسابي للعينة هو مقدر للوسيط p في توزيع برنولي.

تمارين مختارة

1.إذا كانت إذا كانت $x_1, x_2, \ldots\ldots x_n$عينة عشوائية حجمها n من المتغير العشوائي المستمر x الذي يتوزع توزيع كاما، فأوجد تقديرا للوسط α مستخدما طريقة الإحتمالية العظمى؟

2. إذا كانت $x_1, x_2, \ldots\ldots x_n$عينة عشوائية حجمها n من المتغير العشوائي المتقطع x الذي يتوزع توزيعا هندسيا سالب، فأوجد تقدير للوسط p مستخدما طريقة العزوم؟

3.استنتج دالة الترجيح العظمى في حالة للتوزيع معلمتين؟

4.قارن بين طريقة العزوم وطريقة الاحتمالية العظمى من حيث دقة النتائج وبين أيهما أفضل؟

5. إذا كانت $x_1, x_2, \ldots\ldots x_n$عينة عشوائية حجمها n من المتغير العشوائي المستمر x الذي يتوزع توزيع أي سالب، فأوجد تقدير لمعلمة التوزيع مستخدما طريقة الاحتمالية العظمى ومرة وطريقة العزوم مرة أخرى؟

6.قارن بين التقدير النقطي لعينة مسحوبة من مجتمع والتقدير المجالي، موضحا الأفضل من وجهة نظرك؟

7. إذا كانت $x_1, x_2, \ldots\ldots x_n$عينة عشوائية حجمها n من المتغير العشوائي المستمر x الذي يتوزع توزيع ويبل، فأوجد تقديرا لمعالم التوزيع مستخدما طريقة العزوم أو طريقة الاحتمالية العظمى؟

II. خصائص المقدرات Properties of Estimation

لتقدير معلمة من معالم المجتمع محل الدراسة، نحتاج على اختيار الإحصائية المناسبة في العينة لتقدير هذه المعلمة، غالبا ما تكون المعلمة المناظرة في العينة هي أحسن مقدر، كأن نقدر المجتمعμ من خلال متوسط العينة μm. تسمى الإحصائية المستخدمة في التقدير بالمقدر، وتكون قيمة المقدر الذي نرمز له بالرمز$\hat{\theta}$ دائما قريبة من الوسيط المراد تقديره θ، وبعبارة أخرى نقول بأن التوقع لمربع الفرق بين قيمة المقدر والوسيط يكون أقل ما يمكن أي أن $E(\hat{\theta},\theta)^2$ وتسمى هذه الخاصية بخاصية القرب(closeness)، وأن من أهم مقاييس قرب المقدرات عن الدوال المقدرة هو مقياس متوسط مربع الخطأ (Mean Square error)، وسنتناول في هذا الفصل أهم خصائص هذه المقدرات لكي نصبح قادرين على تمييز القدر الجيد من أجل استخدامه في دراسة العينات.

2-1- متوسط مربع الخطأ:

إذا كان $\hat{\theta}$ مقدرا للوسيط θ فإن $E(\hat{\theta},\theta)^2$ يسمى متوسط مربع الخطأ للوسيط θ ويرمز له بالرمز MSE (θ).

ويمكن إثبات العلاقة أعلاه:

$$(\hat{\theta}-\theta)^2=[\,(E\hat{\theta})+E\hat{\theta})-\theta]^2$$

$$(E\hat{\theta}-\theta)^{2+}=E[(\hat{\theta}-E\hat{\theta}))+(E\hat{\theta}-\theta)]^2$$

$$=E(\hat{\theta}-\theta)^2=E[(\hat{\theta}-E\hat{\theta})^2]+2\,[(\hat{\theta}-E(\hat{\theta})][\,E(\hat{\theta}-\theta)]+E[E(\hat{\theta}-\theta]^2$$

وبما أنه لدينا$[2\,[(\hat{\theta}-E(\hat{\theta})][\,E(\hat{\theta}-\theta)]=0$

وعليه يصبح$(E\hat{\theta}-\theta)^2[E(\hat{\theta})-E(\hat{\theta})^2]+E[(\hat{\theta}-E\hat{\theta})^2]=$

ولكي يكون المقدر في نهايته الدنيا أي يقترب من الصفر فإن:

$$E[E(\hat{\theta})-E(\hat{\theta})^2]0=\Rightarrow \quad E(\hat{\theta})-\theta=0 \;\Rightarrow E(\hat{\theta})=\theta$$

وهذه هي خاصية عدم التحيز وعليه يكون: $E[E(\hat{\theta})-E(\hat{\theta})^2=v(\hat{\theta})=min$

وهو يمثل أقل تباين للمقدر $\hat{\theta}$ في حالة عدم التحيز، ويسمى أحيانا مقدرا غير متحيز متشتت تشتتا أصغر.

2-2- المقدر الغير متحيز:

نقول عن إحصائية ما بأنها غير متحيزة **Unbiased** لمعلمة المجتمع إذا كان متوسطها أو توقعها الرياضي مساويا لمعلمة المجتمع.

بعبارة أخرى نقول عن متوسط العينة $\bar{x}$ أنه مقدر غير متحيز لوسط المجتمع μ إذا كان μ =$E(\bar{x})$.

بالمقابل نسمي الإحصائية S^2 في معاينة بالإرجاع أنها مقدر متحيز لـ σ^2

لأن: $E(S^2) = \frac{(n-1)\sigma^2}{n} \neq \sigma^2$

وبصياغة أكثر دقة نستطيع أن نعبر عن مفهوم عدم التحيز بالشكل التالي:

ليكن لدينا مجتمع معالمه θ هذا المجتمع أخذا عينة عشوائية واحدة حجمها n ومن هذه العينة أمكن تعيين التقدير الإحصائي $\hat{\theta}$ كتقدير للمعلمة (البارامتر) θ، وبالحقيقة فإن $\hat{\theta}$ ما هو إلا متغير عشوائي يتغير من عينة إلى أخرى علما أن حجم العينة n ثابت.

يقال إن التقدير الإحصائي $\hat{\theta}$ هو تقدير غير متحيز للمعلمة $\hat{\theta}$ إذا كان

$$E(\hat{\theta}) = \theta$$

مثال (2-1)

أثبت أن المتوسط الحسابي $\bar{x}$ للعينة العشوائية $x1, x2, \ldots\ldots xn$ هو تقدير غير متحيز لمتوسط المجتمع μ

الحل:

$$\mathbf{E}(\bar{x})=E\left(\frac{1}{n}\sum_{i=1}^{n} xi\right)=\frac{1}{n}\sum_{i=1}^{n}(xi) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}\ \mu = \mu$$

ومن ذلك نستنتج بأن $\bar{x}$ تقديرا غير متحيز لمتوسط المجتمع μ.

مثال (2-2)

بين أن تباين المجتمع $\widehat{\sigma^2}=\sum_{i=1}^{n}(xi\text{-}x_i)^2$ (الذي سبق حسابه بطريقة العزوم وطريقة الاحتمالية العظمى) مقدرا غير متحيز لتباين المجتمع σ^2 ولكن بعد تعديله؟

الحل:

$$E(\hat{\sigma}^2)= E(\frac{1}{n}\sum_{i=1}^{n}(xi\text{-}x_i)^2)=\frac{1}{n}E[\![\sum_{i=1}^{n}(xi-\mu)-(\bar{x}-\mu]\!]^2$$

$$E(\hat{\sigma}^2)=\frac{1}{n}E[\![\sum_{i=1}^{n}(xi-\mu)^2-2(xi-\ \mu)(\bar{x}-\ \mu)]\!]$$

$$E(\hat{\sigma}^2)=\frac{1}{n}[\![\sum_{i=1}^{n}E(xi-\mu)^2-n(xi-\ \mu)^2 2nE(\bar{x}-\ \mu)(\bar{x}-\ \mu)]\!]$$

$$E(\hat{\sigma}^2)=\frac{1}{n}[\![\sum s^2\ +nV(\bar{x})-22nV\bar{x}]\!]$$

$$E(\hat{\sigma}^2)=\sigma^2-V(\bar{x})=S^2-\frac{s^2}{n}=S^2(1-\frac{1}{n})=S^2(\frac{n-1}{n})\neq\sigma^2$$

ومنه نستنتج أن $E(\hat{\sigma}^2)\neq\sigma^2$ ، وهذا يعني بأن $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(xi\text{-}x_i)^2$ هو مقدار غير متحيز وعليه يجب تعديله لكي يصبح مقدرا غير متحيز وبالتالي يكون مقدر جيدا لتباين للمجتمع وكالآتي:

$$E(\hat{\sigma}^2)= S^2(\frac{n-1}{n})$$

$$(\frac{n-1}{n})\ E(\hat{\sigma}^2)=S^2\Rightarrow E(\frac{n-1}{n})\hat{\sigma}^2=S^2$$

$$S^2=\frac{\sum_{i=1}^{n}(xi-\mathrm{xi})2}{n-1}$$

وهذا هو مقدار تباين العينة، وهو مقدر غير متحيز لتباين المجتمع σ^2.

لهذا السبب نجد في كثير من التجارب الإحصائية بأن $S^2=\frac{\sum_{i=1}^{n}(xi-\mathrm{xi})2}{n-1}$ يستخدم بدلا من $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(xi-\mathrm{xi})^2$ ، وخاصة في التجارب ذات العينات الصغيرة.

وعليه يمكن تعريف المقدر المتحيز (biasad) بأنه المقدر الذي يحقق الخاصية الآتية: $E(\hat{\theta})$، وبعبارة أخرى نقول بأن المقدر إذا لم يكن متحيز فإنه يكون متحيزا.

2-3- الكفاءة:

تتعلق كفاءة مقدر ما بمقدار التباين لتوزيع المعاينة الإحصائية، فإذا كان للمقدرين (إحصائيتين) نفس المتوسط نقول عن المقدر ذوي توزيع المعاينة الأقل تباينا أنه الأكثر كفاءة، بمعنى أخر إذا كان $\hat{\theta}_1$و$\hat{\theta}_2$ تقديران أن إحصائيتين للوسط $\hat{\theta}$، وكان كل منهما مقدرا غير متحيزا أي أن: $E(\hat{\theta}_1)= E(\hat{\theta}_2)=\theta$ فإن التقدير الأكفأ هو التقدير الذي يكون تباينه أصغر.

مثال (3-2)

إذا كان x_1,x_2,x_3 عينة عشوائية(r.s) من x تتوزع توزيع بواسون بالوسط λ، أثبت أن كلا من $\bar{x}$ و $\bar{y} = \frac{1}{5}(x_1,x_2,x_3)$ مقدرات غير متحيزة للوسط λ، ثم عين أيهما أفضل؟

الحل:

1.في حالة $\bar{x}$ فإن:

$E(\bar{x})=\mu$، وبما أن $\mu = \lambda$ في توزيع بواسون فإن $E(\bar{x})$ وهو تقدير غير متحيز للوسط λ.

2.في حالة $\bar{y}$ فإن:

$$E(\bar{y}) = E\left[\!\left[\frac{1}{5}(x_1, 3x_2, x_3\right]\!\right]$$

$$E(\bar{y})=\frac{1}{5}E(x_1)+\frac{3}{5}E(x_2)+\frac{3}{5}E(x_3)=\frac{\lambda}{5}+\frac{3\lambda}{5}+\frac{\lambda}{5}=\frac{5\lambda}{5}=\lambda$$

$$E(\bar{y})= \lambda$$

إذا كانت المقدرات متحيزة للوسط λ

أحسب تباين كل من $\bar{x}$ و $\bar{y}$:

$$V(\bar{x})=V\frac{\sum_{i=1}^{n} xi}{n}=\frac{1}{n^2}\sum_{i=1}^{n} vx_i=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{1}{n^2}\sum_{i=1}^{n} n\,\sigma^2$$

وبما أن تباين توزيع بواسون λ (التباين= المتوسط) فإن$\sigma^2=\lambda$ وبما أن n=3 فإن: $V(\bar{x}) = \frac{\lambda}{n} = \frac{\lambda}{3}$

وبنفس الأسلوب نحسب تباين $\bar{y}$ كالآتي:

$$v(\bar{y}) = E\left[\!\left[\frac{1}{5}(x_1, 3x_2, x_3\right]\!\right]$$

$$v(\bar{y}) = V\frac{1}{25}\,V(x_1)+9V(x_2)+V(x_3)=\frac{1}{25}(\lambda+9\,\lambda+\lambda)=\frac{11}{25}(\lambda)$$

$$v(\bar{y}) = \frac{11}{25}(\lambda)$$

وبعد المقارنة بين التباينين $V(\bar{x}) = \frac{\lambda}{3}$ و $v(\bar{y}) = \frac{11}{25}(\lambda)$ يتضح أن المقدر $(\bar{x})$ أفضل من المقدر $\bar{y}$ لأن $v(\bar{y}) = \frac{11}{25}(\lambda) > V(\bar{x}) = \frac{\lambda}{3}$.

ملاحظة: لكل من توزيع المعاينة للمتوسط والوسيط نفس المتوسط وهو متوسط المجتمع μ، لكن يعتبر المتوسط $\bar{x}$ أكثر كفاءة لمتوسط المجتمع μ من الوسيط لأن تباين توزيع المعاينة للمتوسطات $V(\bar{x}) = \frac{\sigma^2}{n}$ أقل من تباين توزيع المعاينة للوسيط، ومن البديهي أن استخدام مقدرات فعالة وغير متحيزة هو الأفضل، إلا أنه قد يلجأ لمقدرات أخرى لسهولة الحصول عليها.

- متباينة كرامز-رو Kramer-Rano Inequality

إذا كان $\hat{\theta}$ إحصائيا غير متحيزا للوسط θ لدالة التوزيع الاحتمالي $f(x)$، فإن المتباينة الآتية تحقق:

$V(\hat{\theta}) \geq \frac{1}{nE\left(\frac{\partial Inf(x)^2}{\partial\theta}\right)}$، وتسمى متباينة كراماز-رو.

وتلعب هذه المتباينة أهمية كبيرة في إثبات نوع آخر من المقدرات غير المتحيزة والذي يسمى المقدر الفعال.

4-2 المقدرات الفعالة (الفعالة) Efficient Estimators:

يسمى المقدر غير المتحيز $\hat{\theta}$ للوسط θ، مقدرا فعالا أو (أفضل المقدرات الغير متحيزة) إذا كان يحقق الأدنى من متباينة كرامز-رو أي أن:

$$V(\hat{\theta}) = \frac{1}{nE\left(\frac{\partial Inf(x)^2}{\partial\theta}\right)}$$

ويسمى أحيانا مقدر غير متحيز بأقل تباين.

مثال(2-4)

في التوزيع الطبيعي الذي متوسطة μ وتباينه σ^2، فإنه متوسط العينة العشوائية $\bar{x}$ التي حجمها n، يكون مقدرا غير متحيز بأقل تباين (فعال)؟

الحل:

بما أن $\bar{x}$ غير متحيز لـμ أي $E(\bar{x})=\mu$ (سبق إثباته).

الآن نريد أن نثبت أن $\bar{x}$ مقدرا فعالا بأقل تباين؟

من دالة التوزيع الطبيعي:

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)}{\sigma}} \quad , \qquad -\infty<x<+\infty$$

$$Inf(x)=in(\frac{1}{\sigma\sqrt{2\pi}})-\frac{1}{2\sigma^2}(x-\mu)^2$$

$$Inf(x)=in\ 1-in\ (\sigma\sqrt{2\pi})-\frac{1}{2\sigma^2}(x-\mu)^2$$

$$\frac{\partial Inf(x)}{\partial\mu}=-\frac{1}{2\sigma^2}(2)(x-\mu)(-1)=\frac{x-\mu}{\sigma^2}$$

$$E\left[\!\left[\frac{\partial Inf(x)}{\partial\mu}\right]\!\right]^2=E\left[\!\left[\frac{x-\mu}{\sigma^2}\right]\!\right]$$

$$E\left[\!\left[\frac{\partial Inf(x)}{\partial\mu}\right]\!\right]^2=\frac{1}{\sigma^2}E(x-\mu)^2=\frac{1}{\sigma^2}V(x)=\frac{\sigma^2}{\sigma^4}=\frac{1}{\sigma^2}$$

$$\Rightarrow \frac{1}{nE\left[\!\left[\frac{\partial In\, f(x)}{\partial\mu}\right]\!\right]^2}=\frac{1}{n\frac{1}{\sigma^2}}=\frac{1}{\frac{1}{\sigma^2}}=\frac{\sigma^2}{n}$$

وبما أن:

$$V(\bar{x})=V\frac{\sum_{i=1}^{n}x_i}{n}=\frac{1}{n^2}\sum_{i=1}^{n}V(x_i)=\frac{1}{n^2}n\sigma^2$$

وعليه نستنتج أن: $V(\bar{x})=\frac{1}{nE\left[\!\left[\frac{\partial Inf(x)}{\partial\mu}\right]\!\right]^2}$ أي أن $\bar{x}$ مقدارا فعالا بأقل تباين لأنه يساوي الحد الأدنى من متباينة كرامر-رو.

5-2 التقارب Convergence:

نقول عن مقدار إنه متقارب إذا كان يؤول إلى قيمة المعلمة عندما يؤول حجم العينة إلى ما لا نهاية.

مثال(2-5):

يعتبر متوسط العينة مقدارا متقاربا لمتوسط المجتمع لأن:

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

6-2 الاتساق:

الخاصية الأخرى التي يجب أن تتوافر حتى نقول إن التقدير جيد هو الاتساق، نفترض أن لدينا مجتمعا حجمه N، وسحبنا منه عينة عشوائية حجمها n، ونفترض أن معلمة المجتمع المجهولة هي θ، وأن التقدير من العينة هو $\hat{\theta}$ تقدير متسق للمعلمة θ إذا تقارب هذا التقدير للمعلمةθ عن طريق الإحتمال، أي أن:

$$\lim_{n\to n} \mathrm{p}\left\{\left|\theta, \hat{\theta}\right| \geq \mathrm{c}\right\} \to 1$$

حيث c مقدار اختياري موجب c>0.

كما يمكن التعبير عن التعريف السابق بالشكل الآتي:

يسمى التقدير $\hat{\theta}$ تقديرا متسقا للمعلمة θ إذا كان:

$\hat{\theta}$تقديرا غير متحيزا للمعلمة θ .

$$\lim_{x\to\infty} \left[\underbrace{A}_{\theta}\right] \to 0$$

مثال (2-6)

في التوزيع الطبيعي الذي متوسطه μ وتباينه σ^2 فإن تباين العينة العشوائية S^2 الذي حجمها n يكون مقدرا غير متحيز ومتسق للتباين σ^2؟

الحل:

سبق أن برهنا بأن العينة هو مقدر غير متحيز لتباين المجتمع أي أن:

$$E(S^2)=E\left[\!\left[\frac{\sum_{i=1}^{n}(xi-\mathrm{xi})2}{n-1}\right]\!\right]=\sigma^2$$

ولبرهان أن S^2 مقدرا متسقا للتباين فإن:

$$E(S^2)=E\left[\!\left[\frac{\sum_{i=1}^{n}(xi-\mathrm{xi})2}{n-1}\right]\!\right]$$

$$=\frac{1}{n-1}[\![\sum_{i=1}^{n}(xi-\text{xi})^2]\!]\frac{1}{n-1}=\sum_{i=1}^{n}V(\bar{x})=\frac{1}{n-1}n\frac{\sigma^2}{n}$$

وبأخذ النهاية لها نحصل على :

$$\lim_{n\to\infty}V(S^2)=\lim_{n\to\infty}\frac{\sigma^2}{n-1}=0$$

وعليه فإن تباين العينة يكون مقدرا غير متحيز ومتسق لتباين المجتمع.

7-2 الكفاية Sufficiency:

التقدير $\hat{\theta}$ تقدير كاف للمعلمة θ لأي مجتمع إذا أمكن التعبير عن دالة الكثافة الإحتمالية المشتركة لمفردات العينة كحاصل ضرب دالتين إحداهما تعتمد فقط على المعلمة θ والتقدير $\hat{\theta}$ والأخرى لا تعتمد على المعلمة θ.

أي أن:

$$F(x_1, x_2, \ldots\ldots x_n;\theta)= g[\![\theta,\hat{\theta}]\!]\ h\ [\![(x1, x2, \ldots\ldots xn;\hat{\theta})]\!]$$

مثال (2-7)

$$\bar{x}=\frac{1}{3}(x_1+x_2+x_3)\ ,\qquad \bar{Y}=\frac{1}{6}(x_1+2x_2+3x_3)$$

$$E(\bar{x})=\frac{1}{3}E(x_1+x_2+x_3)=\frac{1}{3}(\mu+\mu+\mu)=\mu$$

$$E(\bar{Y})=\frac{1}{6}E(x_1+2x_2+3x_3)=\frac{1}{6}(\mu+2\mu+3\mu)=\mu$$

أي أن كلي التقديرين هو تقدير غير متحيز للمعلمة μ

للمقارنة بين التباينين لكل من التقديرين ومعرفة أيهما أفضل نحسب تباين كل منهما:

$$V(\bar{x})=\frac{1}{9}V(x_1+x_2+x_3)=\frac{1}{9}(\sigma^2+\sigma^2+\sigma^2)=\frac{1}{3}\sigma^2=\frac{6}{18}\sigma^2$$

$$V(\bar{Y})=\frac{1}{36}V(x_1+2x_2+3x_3)$$

$$=\frac{1}{36}(\sigma^2+4\sigma^2+9\sigma^2)=\frac{14}{36}\sigma^2=\frac{7}{18}\sigma^2$$

أي أن تشتت $\bar{x}$ أقل من $\bar{Y}$ وبالتالي $\bar{x}$ أفضل.

مثال (2–8)

لتكن $x_1, x_2, \ldots\ldots x_n$ مفردات عينة عشوائية من x حجمها n وتخضع لتوزيع برنولي، فأوجد مقدارا كافيا للوسط θ .

الحل:

من توزيع برنولي فإن:

$$f(x, \theta)=\theta^{x}(1-\theta)^{1-x} \qquad ,x=0,1$$

$$F(x_1, x_2, \ldots\ldots x_n;\theta)=\prod_{i=1}^{n}\theta^{x_i}(1-\theta)^{1-x}, x=0,1$$

$$F(x_1, x_2, \ldots\ldots x_n;\theta)=\theta^{\sum_{i=1}^{n}x_i} \qquad , x=0,1$$

$$F(x_1, x_2, \ldots\ldots x_n;\theta)=\theta^{n\bar{x}}(1-0)^{n-\bar{x}}$$

$$=h(x_1, x_2, \ldots\ldots x).g(x_1, x_2, \ldots\ldots x_n;\bar{x})$$

وعليه فإن $\bar{x}$ مقدرا كافيا للوسط θ في توزيع بارنولي.

تمارين مختارة

1. عدد أهم خصائص المقدرات الإحصائية مع إعطاء مثال عن كل خاصية؟

2. لتكن لتكن $x_1, x_2, \ldots\ldots x_n$ مفردات عينة عشوائية من x حجمها n و تخضع لتوزيع الطبيعي بالوسطين μ و σ^2

أوجد مقدرا كافيا للوسط μ؟

3. في التوزيع الآسي السالب الذي متوسطه λ، فإن متوسط العينة العشوائية $\bar{x}$ التي حجمها n، يكون مقدرا غير متحيز بأقل تباين(فعال) المتوسط التوزيع؟

4. أعط نص متباينة كرامر-رو ووضح أهميتها في المقدرات الفعالة؟

5. إذا كان $x_1, x_2, \ldots\ldots x_n$ عينة عشوائية من x تتوزع توزيع ثنائي الحدين بالوسط p، أثبت أن كلا من $\bar{x}$ و

$(\bar{Y})=\frac{1}{5}(x_1, x_2, \ldots\ldots x_n)$ مقدرات غير متحيزة للوسط p ثم عين أيهما أفضل (أكفأ)؟

6.عينة عشوائية حجمها 10 ومتوسطها 3 وتباينها 9 تتوزع توزيعا طبيعيا، أثبت أن متوسط العينة مقدرا غير متحيزا فعال؟

7. أثبت أن مقدر تباين المجتمع المحسوب بطريقة العزوم والاحتمالية العظمى مقدرا متحيزا لـ σ^2 ثم عدله ليصبح متحيزاً إن أمكن ذلك؟

III. التقدير المجالي Interaval Estimation

يسمى هذا التقدير بالتقدير المجالي أو تقدير الفترة، فنحصل على مدى (Range) أو فترة تحدد بحدين (حد أدنى وحد أعلى) وذلك من العينة.

ويلاحظ أن تقدير الفترة يحتوي على أكثر قيمة، بل قد يكون عدد القيم غير محدود أو لا نهائي في كثير من الحالات. فمثلا إذا قدرنا، أن الوسط الحسابي لأعمار الناخبين يتراوح بين 40$\pm$ 6سنة أي يتراوح بين 34 سنة كحد أدنى و 46 سنة كحد أعلى، قد نكون قد حصلنا على تقدير فترة للوسط الحسابي لأعمار الناخبين في المجتمع. ونلاحظ أن هذه الفترة (34.46) تحتوي على عدد لا نهائي من الأعمار، بمعنى أن العدد لا يقتصر فقط على الأعداد الصحيحة والتي تشمل السنوات، ولكنها تشمل أيضا كسور السنوات، والأيام والشهور، والساعات...إلخ، وسوف نرى كيف نحدد التقدير للفترة في بعض الحالات.

وتتميز تقديرات الفترة بالإضافة إلى أنها تحتوي على عدد كبير جداً من القيم، بأنه يمكن حساب احتمال أن يكون التقدير صحيحا، وبالتالي فإنه يمكن معرفة مدى دقة التقديرات. لذا فإن تقدير الفترات يسمى أيضا فترات الثقة Confidence Intervable)) لأن هذه الفترات تعتمد في تكوينها الإحصائي على مستويات ثقة معينة (Confidence Levels) مثل 95 % أو 99% وغيرها، بمعنى أن احتمال أن تكون فترة التقدير صحيحة هو 0.95 أو 0.99 وهكذا....

فإذا كان متوسط أعمار الناجحين يتراوح بين 34 و 46 سنة، ودرجة الثقة هي 95%، فإن معناه أنه لو تكررت التجربة مائة مرة، فإن التقدير سيكون محصورا بين هذين الرقمين في 95 من الحالات (أي احتمال أن يكون صحيحا هو 95%).

3-1 المعاينة في حالة المتوسطات

إذا أخذننا عينات متكررة من مجتمع ما، وقمنا بقياس متوسط لكل عينة، فإننا نجد أن معظم هذه المتوسطات x5، تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات توزيع المعاينة للوسط. ولكن توزيع المعاينة للوسط له أيضا وسط، يعبر عنه بالرمز $\mu_{\bar{x}}$، وانحراف معياري أو خطأ معياري $\sigma_{\bar{x}}$.