الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

المركز الجامعي سي الحواس – بريكة

معهد الحقوق والعلوم الاقتصادية

رقم: ....24............. م.ح.ع.إ/ م.ج.ب/ 2021   بريكة في: 2021/05/30

مستخرج من محضر اجتماع المجلس العلمي للمعهد

بتاريخ: 27 ماي 2021 - على الساعة العاشرة ($10^{H}00$) صباحا

بتاريخ: 27 ماي 2021، اجتمع أعضاء المجلس العلمي لمعهد الحقوق والعلوم الاقتصادية، وذلك بعد توجيه الدعاوى لهم بصفة رسمية، وذلك بحضور الأساتذة الآتية أسماؤهم:

- أ.د/ **بن سعيد عمر**: رئيسا.
- د/ **بولحية شهيرة**: مدير المعهد.
- د/ **مرادسي حمزة**: عضو منتخب.
- د/ **دعاس عز الدين**: عضو منتخب.
- دأ/ **سلالي بوبكر**: رئيس قسم العلوم الاقتصادية.
- د/ **ذبيح هشام**: عضو منتخب.
- د/ **بن سعيد صبرينة**: عضو منتخب.
- د/ **ونوغي نبيل**: عضو منتخب.
- د/ **مرجال عائشة**: عضو منتخب.
- د/ **محمودي سماح**: مدر المخبر.
- د/ **عباسي سهام**: نائب مدير المعهد لما بعد التدرج والبحث العلمي
- د/ **نويس نبيل**: نائب مدير المعهد للبيداغوجيا
- د/ **سايب رامي**: عضو منتخب
- د/ **بوهنتالة ياسين**: رئيس قسم الحقوق

وذلك لاعتماد المطبوعات، وبناء على التقارير الإيجابية الواردة من الخبراء إلى المعهد، وبعد المناقشة والمداولة تقرر ما يلي:

- اعتماد المطبوعة المعنونة بـ: **محاضرات في مقياس الإحصاء 03** موجهة لطلبة السنة 02 L.M.D بالعلوم الاقتصادية والتجارية وعلوم التسيير، للدكتور: **مرادسي حمزة**.

نسخة طبق الأصل من سجل المداولات للمجلس العلمي

مدير المعهد

مديرة معهد الحقوق و العلوم الإقتصادية

د.شهيرة بولحية

وزارة التعليم العالي والبحث العلمي - معهد الحقوق والعلوم الإقتصادية - المركز الجامعي - بريكة

المركز الجامعي سي الحواس - بريكة -

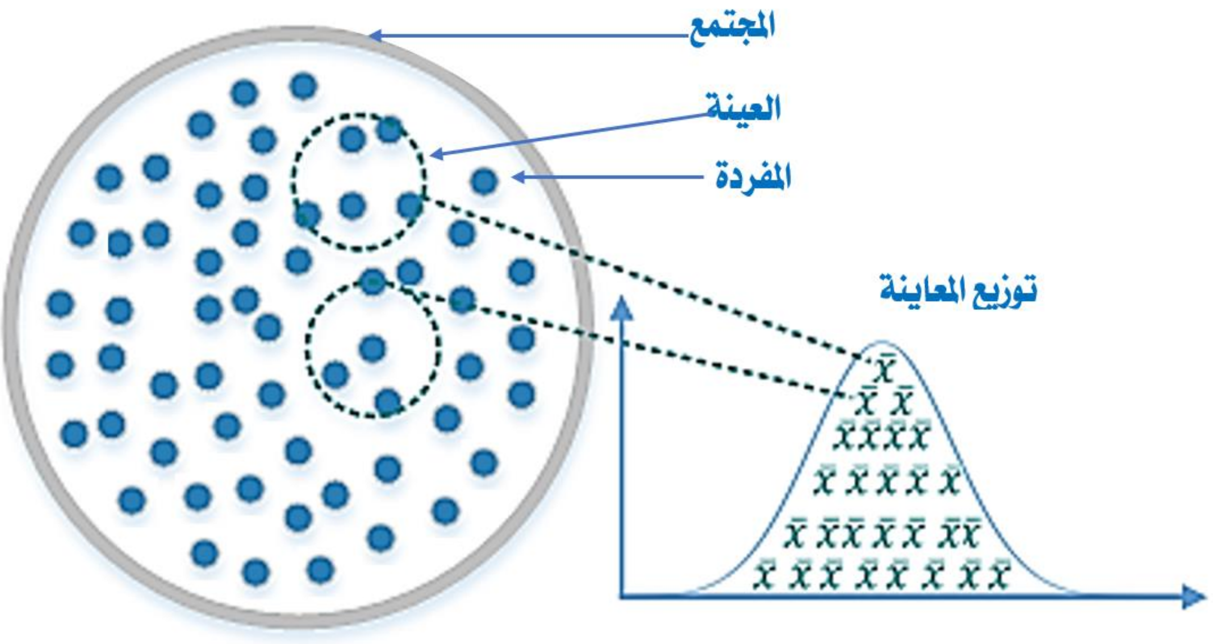
معهد الحقوق والعلوم الاقتصادية

قسم الاقتصاد

مطبوعة موجهة لطلبة السنة الثانية L.M.D بالعلوم الاقتصادية والتجارية وعلوم التسيير بعنوان:

محاضرات في مقياس الإحصاء 3

من إعداد: د/ حمزة مرادسي



السنة الجامعية 2019 – 2020

□

بسم الله الرحمن الرحيم

كلمة المؤلف

تتناول هذه المطبوعة محاضرات في علم الإحصاء الاستدلالي حسب البرنامج الوزاري لمقياس "إحصاء (3) للسنة الثانية ليسانس ميدان العلوم الاقتصادية والتجارية وعلوم التسيير. إذ برمج هذا المقياس لطلبة السنة الثانية، حتى يستكمل الطالب معارفه في علم الإحصاء التطبيقي ويستفيد من القواعد الأساسية التي اكتسبها عند دراسة مقياسي الإحصاء (1) والإحصاء (2) في السنة الأولى من التعليم الأساسي.

وتعتبر هذه المطبوعة ثمرة تجربة سنوات عدية في تدريس مقياس الإحصاء (3) لطلبة السنة ثانية ميدان العلوم الاقتصادية والتجارية وعلوم التسيير في المركز الجامعي سي الحواس بريكة، ولقد حاولت أن أضع خبرتي التي اكتسبتها من هذه التجربة بصياغة محتوى هذا المقياس في شكل مطبوعة توافق البرنامج المحدد من قبل الوزارة وطبيعة التخصص. ولتحقيق هذا الغرض وحرصنا مني على ربط المفاهيم والقواعد النظرية باستخداماتها التطبيقية؛ قمت بتلخيص بعض المفاهيم الأساسية في صورة تذكير لأهم المصطلحات والدروس حول ماهية علم الإحصاء وأهم المصطلحات المستخدمة فيه ومراجعة لموضوع المتغيرات العشوائية التي درسها الطالب سابقا كونها تعتبر القاعدة النظرية التي يجب على الطالب الإحاطة بها حتى يستوعب محتوى المقياس.

ويتضمن البرنامج المقرر على ثلاثة فصول، أولها فصل يتناول "توزيع المعاينة". أما الفصل الثاني فقد خصص لنظرية التقدير الاحصائي. في حين يدرس الفصل الثالث موضوع اختبار الفرضيات وتحليل الانحدار وقد حاولت الالتزام بالمنهج المواضيع المقررة في عرض التكوين، وتجدر الإشارة في هذا المقام أن محتوى هذه المطبوعة من نظريات وقواعد ليس من إبداع مؤلفها، وإنما هي قواعد مبسوطة في المراجع جمعناها وعرضناها بأسلوب ارتأينا أنه الأنسب لمستوى طالب سنة ثانية ليسانس. وإذ نقدم لطلبتنا وزملائنا هذا العمل المتواضع، نهيب بهم أن لا يبخلوا علينا بملاحظاتهم وتعليقاتهم حتى نستفيد منها في الطبعات المقبلة.

فيما يتعلق بما يحتاجه الطالب لمتابعة وفهم هذا المقياس، من المهم الإلمام بمبادئ الإحصاء الوصفي والاحتمالات وبعض القواعد الأساسية للرياضيات. لذلك على الطالب أن يقوم بمراجعة عدد من المفاهيم التي تمت دراستها في مقياسي الإحصاء (1) و (2).

فهرس المحتويات

□

المحور الأول:
مدخل لعلم الإحصاء
والمفاهيم الإحصائية

I. تعريف علم الاحصاء statistics:

كلمة (الاحصاء) في الماضي كانت تهدف الى العد والحصر حتى سمي الاحصاء بعلم العد. The) (Science of canting أما الآن فقد تطور مفهوم الاحصاء كثيرا وخاصة في القرن العشرين واصبح علما مستقلا له أهميته البالغة باعتباره وسيلة واداة في البحث العلمي في تطور العلوم الانسانية الاجتماعية والطبيعية.

ويعرف علم الاحصاء بأنه ذلك الفرع من العلوم الذي يهتم بتجميع وتنظيم البيانات عن طريق تبويبها ثم عرضها في صور ملخصة في شكل جداول أشكال بيانية أو مقاييس ومؤشرات تعطي معلومات قيمة من خلال إبراز معالمها وخصائصها وتحديد العلاقات بينها واستكشاف وتفسير الحقائق غير الظاهرة والتنبؤ بالمستقبل بطرق علمية تحليلية منظمة.

كما يعرف علم الإحصاء بأنه فرع من فروع الرياضيات يشمل النظريات والطرق التي تهتم بجمع البيانات وتنظيمها وتلخيصها ثم عرضها في شكل يسمح بإظهار خصائصها لتحليلها واستخلاص النتائج وتفسيرها بغية الوصول إلى اتخاذ القرارات السليمة على ضوء تلك النتائج، كما يهتم بتحديد العلاقات بين الظواهر المختلفة، وتفسير الحقائق والروابط غير الظاهرة لاستخدامها في التنبؤ ومحاولة فهم سلوك الظواهر مستقبلا، لذلك يعتبر علم الإحصاء أداة هامة من أدوات البحث العلمي وضرورية لتطور كافة العلوم الاجتماعية والانسانية والطبيعية.

من خلال التعريفين السابقين يتضح أن علم الإحصاء هو مجموعة أساليب وطرق منهجية تنظم التفكير، تعتمد على جمع البيانات والحقائق عن الظواهر محل الدراسة، وتنظيمها وترتيبها ثم تلخيصها وعرضها في شكل جداول او بيانات او مؤشرات، لاستخلاص وتحليل المعلومات ومعرفة اتجاهات الظواهر والعلاقات فيما بينها، ثم تفسيرها لاستخدامها في اتخاذ القرار أو في التنبؤ ومحاولة فهم سلوك الظواهر مستقبلا.

ويعد الاهتمام بتطور وتقدم الأساليب الإحصائية أساس هام لتطور البحوث العلمية وأداة لإثبات مصداقيتها، كما يمنح فرص ثمينة لنجاح المنظمات والمشروعات المختلفة لمعرفة ما يخفيه المستقبل إذا ما تم استخدامه في التنبؤ.

II. مراحل البحث الإحصائي:

من خلال تعريف علم الاحصاء يتضح انه علم منظم يعتمد على منهجية واضحة للوصول إلى نتائج ذات مصداقية، ويمر أي تحليل إحصائي على عدة مراحل ترتب كما هو موضح في النقاط التالية:

1. تحديد مشكلة او فرضية البحث او الدراسة.
2. جمع البيانات من السجلات والدفاتر والقوائم والمراجع المختلفة أو من الميدان (الواقع) عن طريق قائمة استقصاء. سواء باستخدام أسلوب الحصر الشامل أو بأسلوب العينات*.
3. تصنيف وتبويب البيانات في جداول إحصائية.
4. عرض البيانات في أشكال ورسوم بيانية لتفسير الظواهر والتعبير عنها.
5. تلخيص البيانات وتفسيرها وتحديد العلاقات بينها باستخدام النظريات والمقاييس والأدوات الإحصائية (المؤشرات) المختلفة.
6. التحليل والاستنباط واستقراء النتائج والتنبؤ بالظواهر مستقبلاً.
7. تفسير نتائج البحث وعملية اتخاذ القرار بشأن فرضيات البحث.

III. فروع علم الإحصاء

لدراسة المراحل السابقة تم تقسيم علم الإحصاء إلى فرعين يكمل أحدهما الأخر،

1. **الاحصاء الوصفي Descriptive Statistic**: وهو علم يهتم باستخدام البيانات الخام بهدف إظهار سماتها وخصائصها من خلال مراحل البحث الإحصائي الخمسة الأولى السابقة وهي جمع وتبويب وحساب بعض المؤشرات الإحصائية وعرضها بشكل بسيط وسهل في شكل مؤشرات أو جدول إحصائي يسهل قراءته أو في رسوم بيانية، والغرض من كل ذلك هو إعطاء وصفا أوليا للظاهرة المدروسة ذو مدلول لكن دون تحليل معمق ودون الوصول إلى استنتاجات أو استدلالات خاصة بالبيانات، فهو يعطينا نتائج عن البيانات في حد ذاتها دون تعميم لتلك النتائج.

من خلال تعريف الاحصاء الوصفي يتضح انه علم يهتم بدراسة خصائص المجتمعات الاحصائية ووصفها، وهو يسبق في التحليل وظيفة الاحصاء الاستدلالي الذي يهتم بعملية التعميم والاستنتاج.

2. **الاحصاء الاستدلالي (الاستنتاجي) Inferential statistics**: وهو أحد فروع علم الإحصاء اذ يعتبر الفرع الثاني بعد الاحصاء الوصفي، ويستند على فكرة جمع البيانات على جزء فقط من المجتمع يسمى العينة، يتم اختيارها بطريقة علمية مناسبة، بغرض استخدام بيانات هذه العينة في التوصل إلى

* المسح الشامل يتم فيه جمع البيانات من جميع مفردات المجتمع دون استبعاد أي مفردة، فمثلا إذا أردنا التعرف على مستوى طلاب الجامعة في مادة الإحصاء نقوم برصد درجات جميع طلاب الجامعة في مادة الإحصاء. لكن هذه الطريقة عادة تكون طويلة ومكلفة وتحتاج إلى الكثير من الوقت ناهيك عن عدم إمكانية تطبيقاتها في الحالات التي تؤدي فيها جمع البيانات عن مفردات البحث إلى فناء هذه المفردات. لذلك يتم اللجوء إلى أسلوب العينات أين يتم اختيار عينة تمثل المجتمع وتجرى عليها الدراسة وتعمم النتائج على المجتمع، وكلما كانت العينة مختارة بطريقة صحيحة وممثلة تمثيلا صادقا المجتمع كلما كانت النتائج صادقة ودقيقة.

نتائج يمكن تعميمها على مجتمع الدراسة، وعليه نقول قمنا بالاستدلال على معلمات المجتمع على أساس إحصاءات العينة، وهذا عكس الاستنباط الذي يعني استخراج خواص الجزء انطلاقا من خواص الكل.

ويعرف الاحصاء الاستدلالي بأنه عبارة عن مجموعة الطرق العلمية التي تتعلق بطرق تحليل وتفسير وتقدير واستخلاص الاستنتاجات بالاعتماد على جزء من المجتمع الأصلي يسمى العينة، بسبب صعوبة أو استحالة* أسلوب الحصر الشامل للبيانات الخاصة بكل المفردات الخاصة بظاهرة معينة. ويستهدف الوصول إلى تعميمات تخص المجتمع الاحصائي، حيث يعمل على الاستدلال على معلمات المجتمع بناءً على البيانات الإحصائية التي جمعت من عينة من هذا المجتمع وفق طرق إحصائية محددة، وتعتبر نظرية الاحتمالات والتوزيعات الاحتمالية المختلفة ونظرية العينات والتقدير واختبارات الفروض الإحصائية هي الإطار الأساسي لهذا الفرع من فروع الإحصاء.

وقد بدأ الإحصاء كعلم بجانبه الوصفي البحت، بعد ذلك تطور ليصبح أداة قوية لاتخاذ القرارات من خلال ظهور فرع الاستدلال الاحصائي وانتشاره في جميع المجالات، حتى أصبح التحليل الإحصائي في مختلف المجالات ينصب أساسا على الاستدلال الذي يعطي تعميمات واستنتاجات ذات قيمة. وحتى يكون الاستدلال الإحصائي صحيحا يجب أن تحقق العينة مجموعة من الشروط تحقق تماما صفات المجتمع الذي سحبت منه وخصائصه، وتسمى حينها بانها عينة ممثلة.

IV. بعض المصطلحات والمفاهيم الإحصائية العامة:

1. البيانات Data: هي سلسلة غير مترابطة من المعطيات والقياسات والمشاهدات والحقائق الموضوعية، التي يمكن الحصول عليها عن طريق الملاحظة، أو عن طريق البحث والتسجيل. وهي المادة الخام لإنتاج المعلومات حيث تعبر عن مختلف الحقائق خلال فترة زمنية محددة، وقد تكون البيانات رقمية (كمية) مثل أطوال وأوزان مجموعة من الطلاب أو دخول مجموعة من الأسر أو بيانات غير رقمية (وصفية) في شكل كلمات، معاني لغوية، أرقام، رموز...، مثل لون البشرة والجنس ... إلخ. وتعتبر البيانات من أهم الاشياء التي يعتمد عليها المحققون في ربط الاحداث للتعرف على ظاهرة معينة.

* إن تحليل خصائص كل مفردات المجتمع قد يعرض تلك المفردات للتلف أو قد يكون بالغ التكلفة أو يتطلب مدة زمنية طويلة للحصول على البيانات اللازمة لذلك يمكن التغلب على هذه المشكلة بأخذ عينة ممثلة للمجتمع.

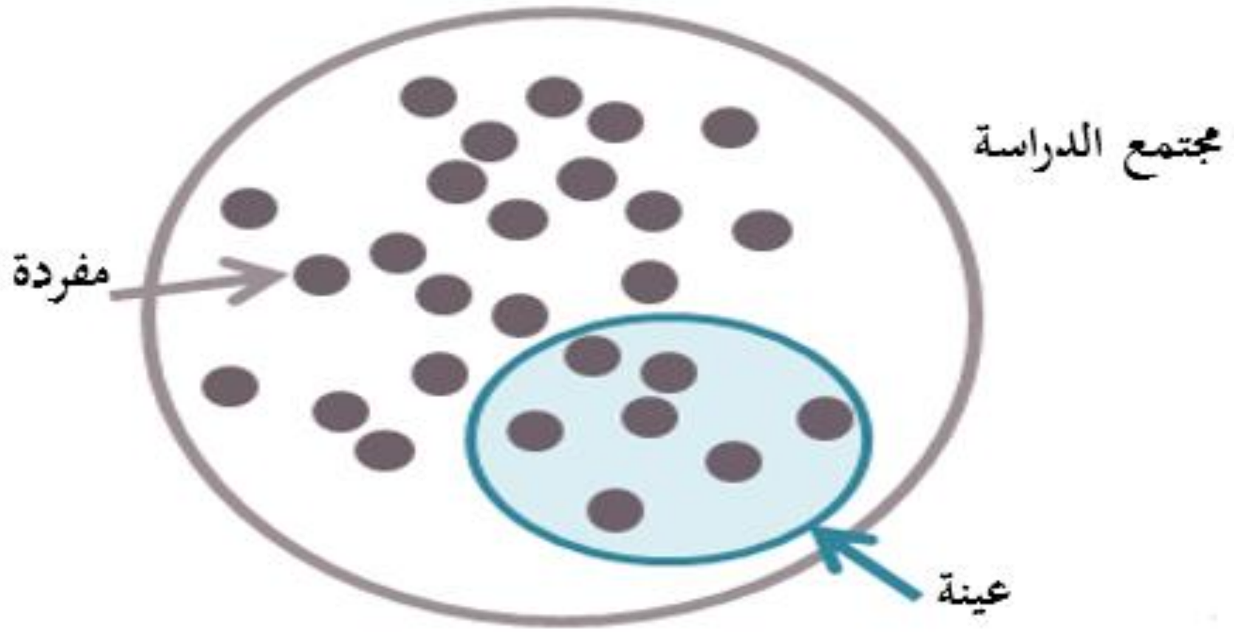
والبيانات لا يكون لها معنى إذا تركت بشكل خام، ولهذا يتم تجميعها حتى يتم استخدامها في الحصول على المعلومات، من خلال معالجتها وتصنيفها وتنظيمها وتحليلها، ليصبح لها معنى وتحقق هدف معين حتى توفّر ما يسمى بالمعرفة.

2. **الوحدة (المفردة) الإحصائية:** هي أصغر جزء في المجتمع وهو المكون الأساسي الذي تجر عليه الدراسة الإحصائية وهي التي يتم جمع البيانات لأجلها وعنها. فالمفردة هي ما نريد معرفة خصائصه وتحديد الحقائق عنه سواء كانت هذه المفردة شخص، سلعة، وحدة زمن، مكان أو غير ذلك.

3. **العينة Sample**: هي جزء من المجتمع أو هي كل مجموعة جزئية داخل المجتمع الإحصائي تكون ممثلة للمجتمع الإحصائي أحسن تمثيل، بحيث تمثل جميع صفات المجتمع، ويتم اختيارها بطريقة مناسبة، لإجراء الدراسة عليها ومن ثم استخدام تلك النتائج، وتعميمها على كامل المجتمع الأصلي تحت الدراسة، ويلجأ الباحث إلى الاعتماد على العينة في الدراسة بدل المجتمع بسبب كبر حجم المجتمع وربما للوقت والجهد والمال؛ أو أحيانا قد يكون الفحص مؤذيا أو متلفا للوحدات المدروسة لذلك لا يمكن إتلاف جميع وحدات المجتمع. بالإضافة إلى ذلك قد يكون أسلوب الحصر الشامل مستحيل في حالة حجم المجتمعات غير المحدودة. وهذا ما استدعى بروز الحاجة لدراسة العينة كبديل من خلال جمع معلومات عن جزء من المجتمع للإجابة عن إشكالية تخص مجتمع الدراسة، إذ تعمم نتائج العينة على المجتمع كله، مثل أخذ عينة من دم مريض لفحصها حيث إننا لا نستطيع فحص كل دم المريض لأن ذلك يؤدى إلى الوفاة. وهناك نوعين من العينات حسب طبيعة سحب العينات هي العينة النفادية والعينة غير النفادية.

4. **المجتمع Population**: المجتمع الإحصائي هو مجموعة من المفردات التي تشترك في صفات وخصائص محددة تخص ظاهرة معينة تهم الباحث في دراسته، كما يعرف بانه عبارة عن جميع المفردات التي تشترك فيما بينها في الصفة الأساسية محل اهتمام الباحث، ويمكن أن يكون المجتمع الإحصائي بهذا التصور محدوداً مثل عدد أجهزة الكمبيوتر في معمل معين، عدد الطلبة في قسم معين، أو قد يكون غير محدود مثل عدد النجوم في السماء، عدد الطيور في مدينة معينة. ويوضح الشكل ادناه الفرق بين المفردة العينة والمجتمع:

الشكل رقم 1-1: المجتمع العينة والمفردة



وتسمى دراسة خصائص كل مفردات المجتمع بطريقة الحصر الشامل، أين يتم جمع البيانات من جميع المفردات المشمولة بالبحث وهذه الطريقة تعطي بيانات متكاملة ونتائج دقيقة، وتستخدم هذه الطريقة في التعدادات السكانية والزراعية والصناعية...الخ التي تجريها الحكومات وتدعمها بإمكانيات ضخمة، لأنها تحتاج الى جهد ووقت ومال كثير، كما قد يؤدي هذا الأسلوب إلى تلف الوحدات محل الدراسة، لذلك يلجأ الباحثون إلى استخدام أسلوب المعاينة الإحصائية، أين يتم دراسة خصائص العينة المدروسة ثم يتم تعميم نتائج تلك الدراسة على كل مفردات المجتمع قيد الدراسة ويعتمد نجاح أسلوب المعاينة على مدى دقة العينة المسحوبة ومدى تمثيلها للمجتمع المسحوبة منه.

5. معلمة المجتمع parameter: نقصد بمعالم المجتمع مجموعة المقاييس التي تستخدم لعرض خصائص المجتمع أو استخلاص نتائج مجموعة من خصائصه مثل المتوسط ويرمز له بالرمز μ، التباين ويرمز له بالرمز σ^2، النسبة p أو θ ...الخ، ومن خصائص المجتمع أيضا طبيعة توزيعه الاحتمالي $f(x)$ كأن تكون مفردات المجتمع تتبع توزيعا طبيعيا أو توزيع ستودنت أو توزيع ثنائي الحد ...الخ. وعليه يمكن القول بأن المعلمة هي شيء يميز المجتمع ككل، مثل متوسط الدخل الشهري للأسر في دولة معينة، أو متوسط طول الطلاب في مدرسة ما، أو نسبة المدخنين في مجتمع معين، أو نسبة المعيب في إنتاج إحدى السلع، في مصنع معين.

ويمكن حساب أهم معلمات المجتمع من خلال استخدام العلاقات التالية:

المتوسط الحسابي (التوقع الرياضي) للمجتمع:
$$\mu = \bar{X} = \frac{\sum X_i}{N}$$

التباين للمجتمع: $\sigma_X{}^2 = var_x = \frac{\sum(X_i - \bar{X})^2}{N} = E(X^2) - [E(X)]^2$

النسبة للمجتمع: $p = \frac{X}{N}$

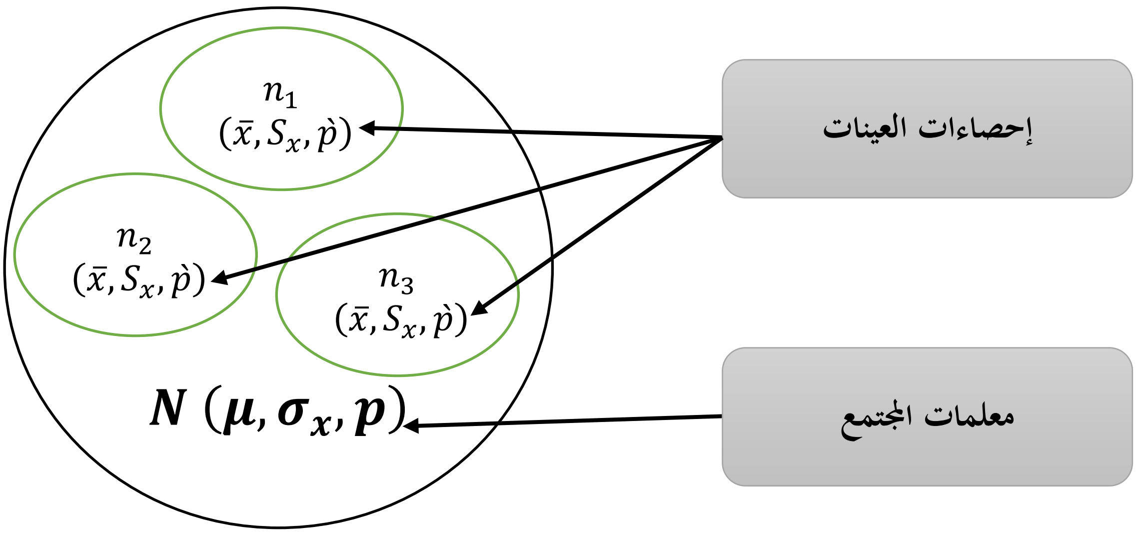
6. **إحصاءة العينة statistic**: الإحصاءات هي مجموعة من المقاييس تستخدم لعرض خصائص العينة أو استخلاص نتائج عنها. ولتقدير معالم المجتمع (متوسط المجتمع μ، تباين المجتمع $\boldsymbol{\sigma^2}$، النسبة $\boldsymbol{p}$) ننطلق من خصائص العينة، لذلك نحتاج إلى حساب إحصاءات مثل متوسط العينة $\bar{x}$ ، تباين العينة S^2، النسبة في العينة $\grave{p}$. وبصفة عامة نسمي كل قيمة تحسب انطلاقا من بيانات العينة من أجل تقدير قيمة معالم المجتمع **بالإحصاءة**. ونظريا (رياضيا) إحصاءة المعاينة هي كل دالة في المتغيرات العشوائية التي تمثل القيم المحصل عليها في العينة. ومن أمثلة إحصاءات العينة نعطي مثال بمتوسط الدخل الشهري لعينة مكونة من 100 أسرة في دولة ما، أو متوسط الطول لعينة مكونة من 50 طالب في مدرسة ما.

ويمكن حساب أهم معلمات المجتمع من خلال استخدام العلاقات التالية:

المتوسط الحسابي (التوقع الرياضي) للعينة: $\bar{x} = \frac{\sum x_i}{n}$

التباين العينة **غير المتحيز**: $S_x{}^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$

النسبة للعينة: $\grave{p} = \frac{x}{n}$



وتجدر الإشارة هنا أن معالم المجتمع دائما ثابتة في حين تشكل أن إحصاءات العينة متغير عشوائي يتغير بتغير العينة المسحوبة. فقد نجد أن إحصاءة العينة تساوي معلمة المجتمع، أو قد تكون أصغر أو أكبر في عينة أخرى من نفس المجتمع، ويسمى الفرق بين إحصاءة العينة ومعلمة المجتمع بخطأ المعاينة، أو الخطأ العشوائي.

7. **المتغير Variable**: هو الخاصية أو الصفة (النوعية أو الكمية) المشتركة لكل الوحدات (المفردات) الإحصائية التي تشكل المجتمع الإحصائي، وهي قابلة للتغير من مفردة إلى أخرى في المجتمع الإحصائي أو العينة الاحصائية مثل: الدخل، الجنس، الحالة الاجتماعية، الطول، السن، المستوى التعليمي، الإنتاج، ...إلخ.

المحور الثاني:
تذكير بالمتغيرات العشوائية والتوزيعات الاحتمالية

سنحاول في هذا الفصل أن نستعرض بعض الأساسيات في موضوع المتغيرات العشوائية ودوالها الاحتمالية، وذلك بالاعتماد على مفهوم أساسي يقضي بأن فضاء العينة ما هو إلا مجموعة النتائج المتوقعة من إجراء تجربة ما، وأن الحادث ما هو غلا مجموعة جزئية من فضاء العينة Ω.

I. المتغيرات العشوائية The Random Variable

إن نتائج أي تجربة تكون غير معروفة بصورة مسبقة إذ لا يمكن التنبؤ بنتيجة تجربة بشكل قطعي ومؤكد قبل حدوثها، ولكن من الممكن معرفة كل النواتج المتوقعة مسبقا، إذ أن نتائج كل تجربة تكون محصورة في مجال معين من القيم أو الصفات يطلق عليها مصطلح فراغ العينة، حيث أن إجراء نفس تجربة ما بصورة متكررة سيؤدي إلى ظهور نتائج مختلفة، ومثل هذه المحاولات تسمى التجربة العشوائية، ويلاحظ أن النتائج التي نحصل عليها تعتمد فقط على الصدفة أو العشوائية Randomness، وقد تكون نتائج التجربة في صورة عددية أو وصفية. ومن ثم يمكن القول أنه يرافق نتائج كل تجربة عشوائية مقدار يطلق عليه المتغير العشوائي، وهذا المقدار يأخذ قيما مختلفة حسب النتائج المحتملة للتجربة العشوائية، ويمكننا القول إن المتغير العشوائي هو النتائج الممكنة للتجربة العشوائية سواءً كانت قيماً عددية حقيقية أو وصفية متغيرة حسب عدد المحاولات المختلفة لتجربة عشوائية خاضعة للصدفة. كما يمكن معرفة أو حساب فرص ظهور (احتمال) كل نتيجة من النتائج المحتملة للتجربة قبل حدوثها.

1. تعريف المتغيرات العشوائية: يعرف المتغير العشوائي بانه خاصية تتميز بها نواتج تجربة عشوائية معينة، وتختلف قيمته من محاولة إلى أخرى حسب عدد المرات التي تتكون منها التجربة، وذلك حسب طبيعة النواتج في فراغ العينة للتجربة العشوائية تحت الدراسة.

كما يعرف المتغير العشوائي رياضيا بأنه دالة ذات قيم عددية معرفة على فضاء العينة* Ω، ومجاله المقابل هو مجموعة الأعداد الحقيقية. ذلك نقول إن المتغير العشوائي X هو تطبيق مجاله فضاء العينة S ومجاله المقابل هو مجموعة الأعداد الحقيقية R. ونكتب: $X: S \rightarrow R$

المتغير العشوائي هو دالة في فضاء تجربة ما إلى مجموعة الأعداد الحقيقية R، حيث يقابل كل عنصر من فضاء التجربة (فضاء العينة) عددا حقيقيا هو المتغير العشوائي عن هذا العنصر، ويرمز للمتغيرات العشوائية

* فضاء العينة للتجربة العشوائية هي المجموعة المكونة من جميع النتائج الممكنة للتجربة، ونرمز لفضاء العينة بالرمز Ω. ونسمي كل نتيجة من نتائج التجربة العشوائية عنصرا في فضاء العينة أو نقطة العينة.

عادة بأحرف لاتينية كبيرة مثل: ... X, Y, Z ويرمز لقيم هذه المتغيرات العشوائية بأحرف لاتينية صغيرة مثل: ...x, y, z.

مثال 1: في تجربة إلقاء حجر نرد عدة مرات يمكن أن نسمي نتيجة التجربة المتمثلة في عدد النقاط التي تظهر على السطح العلوي للحجر في كل مرة بأنها متغيرة عشوائية X. حيث يمكن أن يأخذ إحدى القيم الممكنة هي: 1، 2، 3، 4، 5، 6. بكل قيمة يمكن أن نلحق احتمال تحققها، وهو هنا 6/1. ونكتب:

$P(X = 1) = f(1) = 1/6, \ P(X = 2) = f(2) = 1/6$

نلاحظ أن القيم الممكنة لـ: X : (1، 2، 3، 4، 5، 6) هي متنافية، ولذلك فإن مجموع احتمالاتها يساوي 1 .

مثال 2. في تجربة إلقاء قطعة نقدية مرتين يمكن أن نعين المتغيرة العشوائية X التي تمثل عدد مرات الحصول على كتابة. في هذه الحالة القيم الممكنة ل X هي 0، 1، 2.

2. أنواع المتغيرات العشوائية

حسب التعريفات السابقة فإن المتغير العشوائي يعتبر دالة مجالها هو فراغ عينة تجربة عشوائية معينة ومجالها المقابل هو مجموعة الأعداد الحقيقية ومدى هذه الدالة هو مجموع جزئية من مجموعة الأعداد الحقيقية. فإذا كان فراغ العينة يتكون من نواتج عبارة عن قيم يفصل بينهما فجوات فإن المتغير العشوائي يعرف على مثل هذا الفراغ يسمى متغيراً عشوائيا متقطعاً أو منفصلا Discrete Random Variable أما إذا كان فراغ العينة يتكون من نواتج عبارة عن قيم متصلة أو مستمرة ولا يمكن الفصل بينها بفجوات مثل الفترات التي تعبر عن الأعداد الحقيقية فإن أي متغير يعرف على مثل هذه الفراغ يسمى متغيراً عشوائياً متصلاً أو مستمراً Continuous Random Variable.

أ- **المتغير العشوائي المتقطع (المنفصل)**: هو المتغير الذي يمكن أن يأخذ إما عددًا محدودًا من القيم، أو مجموعة لا نهائية من القيم القابلة للعد ، مثل 0 ، 1 ، 2 ، ...الخ، كما يمكن تعريف المتغير العشوائي المنفصل بأنه المتغير الذي يأخذ قيما **محددة قابلة للعد** في مجال تغيره، وعليه نقول عن المتغير العشوائي X أنه متغير عشوائي متقطع إذا كانت مجموعة القيم الممكنة لـ فضاء العينة **Ω منتهية أو غير منتهية لكن قابلة للعد** (مُحَدَّدَة)، بعبارة أخرى إذا احتوى فراغ العينة على عدد منته أو غير منته ولكنه معدود فإننا نقول إن هذا الفراغ متقطع ونسمي المتغير العشوائي المعرف على فراغ هذه العينة بالمتغير العشوائي المتقطع. ونكتب:

$X(\Omega)= \{x_1, x_2, x_3, ...x_n\}$ أو $X(\Omega)= \{x_1, x_2, ...\}$

ومن بين الأمثلة على المتغيرات العشوائية المتقطعة:

إذا كان X يعبر على عدد الأولاد الذكور الممكن تواجدهم في أسرة مكونة من أربع أبناء نكتب:

$X(\Omega) : \{x=0,1,2,3,4\}.$

إذا كان Y يعبر على عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق، نكتب:

$Y(\Omega) : \{y=0,1,2, 3, ...\}.$

ب-المتغير العشوائي المستمر (المتصل): ليكن X متغير عشوائي حقيقي، فإذا كان مجا تعريفه فترة (مجال) أو مجموعة فترات فإنه يعتبر متغير عشوائي مستمر، وبناءً على ذلك يقال أن المتغير العشوائي (X) متغير مستمر إذا كان يأخذ جميع القيم الصحيحة والكسرية في مجال تعريفه.

ويكون مجال تعريف المتغير العشوائي المستمر مجال واحد ونكتب: $X=\{a<x_i<b\}$، كما يمكن ان يكون مجال تعريفه عدة مجالات ونكتب حينها: $X=\{a<x_i<b \ , \ a<x_i<b \}$.

وعليه نقول أن المتغير العشوائي $X: \Omega \rightarrow R$ هو متغير عشوائي مستمر إذا كانت مجموعة القيم الممكنة لـ $X(\Omega)$ هي مجموعة غير منتهية وغير قابلة للعد، أي أنها عبارة عن مجال من المحور الحقيقي أو اتحاد عدة مجالات.

ويعرف المتغير العشوائي المستمر بأنه متغير عشوائي يمكن أن يأخذ عددا لا متناهيا من القيم في مجال محدود محصور بين قيمتين، أي أنه يأخذ أي قيمة داخل مجال تعريفه (فضاء العينة). كما يمكن اعتبار المتغير العشوائي المتصل ذلك المتغير الذي تتكون واتجه الممكنة من كل القيم التي تقع في مجال واحد أو في مجالات عدة على خط مستقيم، ويكون عدد القيم الممكنة لأي متغير عشوائي مستمر غير منتهية وغير قابلة للعد. من أجل هذا فإن وحدات قياس المتغيرة العشوائية المستمرة تكون متصلة مثل الزمن، الوزن، المسافة، الحجم. فإذا كان X متغير عشوائي مستمر، ويقع في المدى (a,b)، أي:

$X(\Omega) : \{X=x_i : a<x_i<b\}.$

إن للمتغير X عدد لانهائي من القيم تقع بين الحدين الأدنى والأعلى (a,b).

وكنتيجة يمكن القول بأنه إذا كان المتغير العشوائي معرفا على فراغ عينة متقطع كان المتغير العشوائي متقطعاً أما إذا كان معرفا على فراغ عينة مستمر كان المتغير العشوائي مستمرا. ومن الظواهر التي يمكن التعبير عنها بمتغيرات عشوائية منفصلة عدد السيارات التي تصل إلى محطة خدمة السيارات في الساعة، عدد المرضى

الذين يترددون عل عيادة معينة في اليوم الواحد، عدد الحوادث التي تقع في الأسبوع الواحد على إحدى الطرق السريعة وغيرها. بينما الظواهر التي تأخذ قيما في فترة متصلة مثل الأطوال والأوزان ودرجات الحرارة والأعمار فإنه يتم التعبير عنها بمتغيرات عشوائية متصلة.

وترتبط بكل المتغيرات العشوائية دالة تسمى دالة الاحتمال أو تسمى بدالة التوزيع الاحتمالي، وهناك نوعان من الدوال، دالة الاحتمال للمتغير العشوائي المنفصل، ودالة الاحتمال للمتغير العشوائي المستمر.

3. دالة التوزيع الاحتمالي للمتغيرات العشوائية

التوزيع الاحتمالي هو توضيح العلاقة التي تربط بين القيم المختلفة للمتغير العشوائي (X) واحتمالات تحقق هذه القيم $f(x)$، وهذه العلاقة يمكن أن تكون في صورة جدول أو في صورة دالة رياضية. ونرمز للمتغيرات العشوائي بأحرف لاتينية كبيرة، وللقيم التي تأخذها المتغير بأحرف لاتينية صغيرة. وعليه يمكن اعتبار التوزيع الاحتمالي كدالة أو جدول يوضح مجموعة القيم الممكنة مع احتمالات الحدوث المرتبطة بالنتائج الممكنة التي يمكن ان يأخذها المتغير في فراغ العينة، وبمعنى آخر هو التكراري النسبي للقيم التي يمكن أن يأخذها المتغير.

أ- **دالة التوزيع الاحتمالي للمتغير العشوائي المنفصل**: إذا كان X متغير عشوائي منفصل يأخذ القيم $X=\{x_1, x_2, x_3, \ldots x_n\}$، فإن احتمال أن يأخذ المتغير العشوائي X القيمة x_i يسمى بدالة التوزيع الاحتمالي $f(x)$ وتكتب بالشكل التالي:

$$p(X=x_i)=f(x_i)$$

ويمكن تمثيل التوزيع الاحتمالي للمتغيرات العشوائية المتقطعة من خلال تكوين جدول التوزيع الاحتمالي للمتغير العشوائي X كما هو مبين أدناه:

الجدول رقم 1: جدول التوزيع الاحتمالي للمتغير العشوائي المتقطع X.

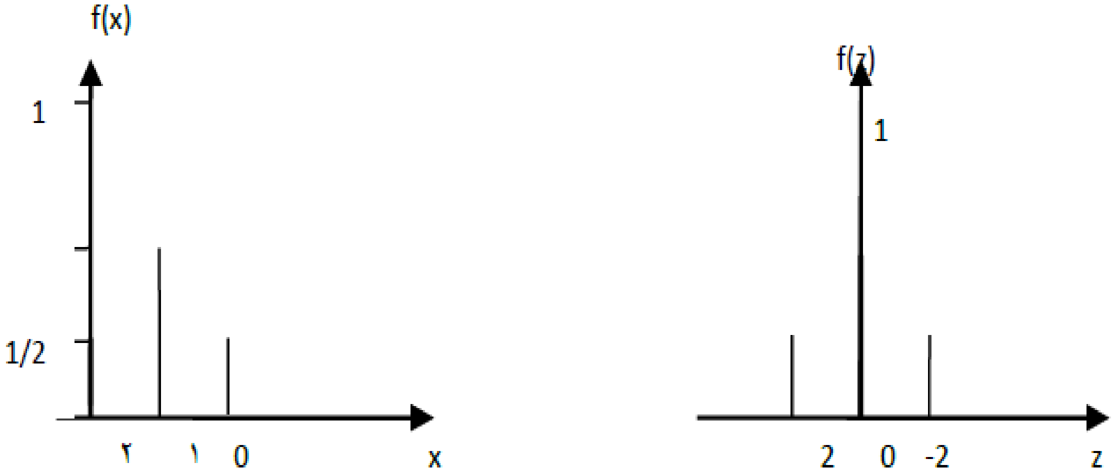
$\sum P(X=x_i)$	x_n	....	x_2	x_1	X
1	$p_n=f(x_n)$	....	$P_2=f(x_2)$	$P_1=f(x_1)$	$P(X=x_i)$

ولكي نعتبر دالة ما $f(x)$، دالة توزيع احتمالي لمتغير عشوائي منفصل يجب أن يتحقق شرطان اثنان هما:

- **الشرط الأول :** $0\leq f(x_i)\leq 1$ أو $0\leq P_i\leq 1$
- **الشرط الثاني :** $\sum f(x_i)=1$ أو $\sum P(X=x_i)=1$

يتم تمثيل التوزيع الاحتمالي للمتغيرات العشوائية المنفصلة في الأشكال البيانية عن طريق العمدة كما يوضحه الشكل أدناه:

الشكل رقم 2–1: التمثيل البياني لدالة التوزيع الاحتمالي للمتغيرات العشوائية المتقطعة



ولإيجاد احتمال أن يأخذ المتغير العشوائي قيمة أصغر من أو يساوى قيمة معينة في فراغ العينة Ω، سنقوم حينها بجمع الاحتمالات المناظرة للقيم التي يأخذها المتغير العشوائي في المجال المحدد، وهو ما يطلق عليه الاحتمال التجميعي أو التراكمي، والدالة التي تستخدم لهذا الغرض تسمى دالة الاحتمال التجميعية أو التراكمية، وهي دالة تراكمية أو تجميعية لدالة التوزيع الاحتمالي، وتعطى بالصيغة التالية:

$$f(x)=P(X\leq x_i)=\Sigma P(X=x_i)$$

ب- **التوقع الرياضي والتباين للمتغير العشوائي المنفصل**: التوقع الرياضي أو الأمل الرياضي هو المتوسط الحسابي للتوزيع الاحتمالي للمتغير العشوائي X، ويرمز للتوقع الرياضي بالرمز $E(X)$ أو μ وهو العزم الأول حول الصفر، ويعرف بالصيغة التالية:

$$\mu= E(X)= \Sigma[x_i*f(x_i)] = \Sigma[x_i*P(x_i)]$$

بالنسبة للتباين يرمز له بالرمز σ^2 وهو العزم الثاني حول التوقع الرياضي، ويعطى بالصيغة التالية:

$$\sigma^2= E(xi-\mu)^2 \Rightarrow \sigma^2= \Sigma (x_i-\mu)^2*f(x_i) \Rightarrow \sigma^2= \Sigma (x_i-\mu)^2*P(x_i)$$

$$\sigma^2= \Sigma[x_i^2*P(xi)] - E(X)^2 \Rightarrow \sigma^2= E(X^2) - E(X)^2$$

مثال1: إذا ألقيت قطعة نقود متزنة مرتين متتاليتين، وكان المتغير العشوائي X عبارة عن عدد الصور التي تظهر.

المطلوب: أوجد القيم التي يمكن أن يأخذها المتغير العشوائي X؟

- عين التوزيع الاحتمالي لهذا المتغير، ثم تأكد أنه فعلا توزيع احتمالي؟
- أحسب التوقع الرياضي، التباين والانحراف المعياري؟

الحل: إذا كان X يمثل عدد مرات ظهور الصورة فإن النتائج المحتملة لقيم المتغير العشوائي يمكن حصرها في فضاء العينة Ω حيث:

$X(\Omega)$: {(P.P), (P.F), (F.P), (F.F)}

$X(\Omega) : \{0,1,2\}$

التوزيع الاحتمالي لقيم المتغير العشوائي X يمكن تمثيلها في الجدول التالي:

X_i	0	1	2	Σ
$P(x_i)$	1/4	2/4	1/4	4/4=1
$X_i p(x_i)$	0	2/4	2/4	1
$(X_i)^{2*}p(x_i)$	0	2/4	4/4	6/4=1.5

من خلال جدول التوزيع الاحتمالي نلاحظ أن جميع قيم p(x) أكبر من الصفر وأقل من الواحد $0<p(X)<1$، كما أن مجموع قيم p(x) تساوي الواحد $\Sigma p(X)=1$ وهذا يبين أش الشرطين محققين وبالتالي فإن الجدول أعلاه هو جدول توزيع احتمالي.

حساب التوقع الرياضي E(X):

$$E(X)= \Sigma x_i{*}p(x_i)=1$$

حساب التباين σ^2:

$$\sigma^2= E(X^2) - E(X)^2=1.5-(1)^2= 0.5$$

حساب الانحراف المعياري σ:

$$\sigma = \sqrt{\sigma^2} = \sqrt{0.5} = 0.7071$$

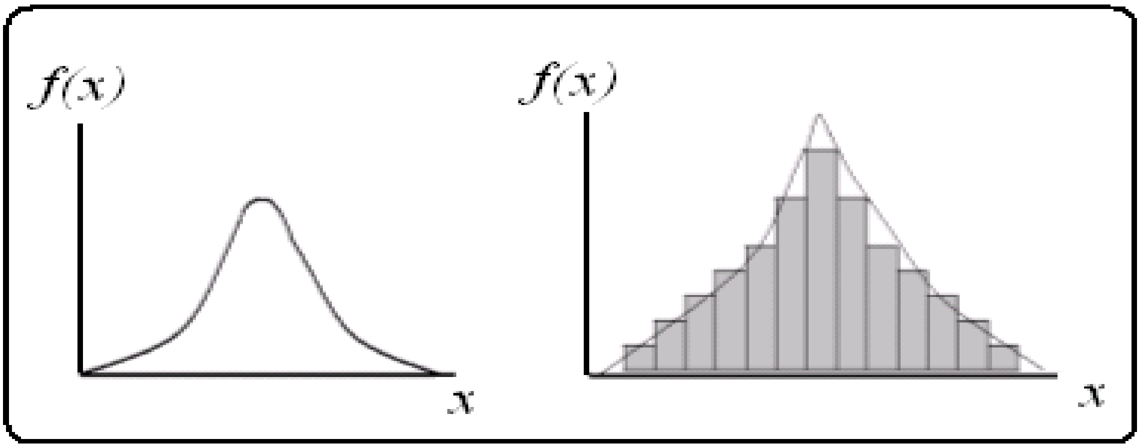
ج- **دالة الكثافة الاحتمالية للمتغير العشوائي المتصل:** سبق وأن ذكرنا أنه إذا كان المتغير العشوائي المعرف على فراغ عينة ما، عبارة عن فترة متصلة من الأعداد الحقيقية كان هذا المتغير عشوائيا متصلا.

فإذا كان X متغير عشوائي متصل، فإن الدالة الاحتمالية التي تصف سلوك المتغير العشوائي لا تعطى احتمال أن يأخذ هذا المتغير قيمة معينة داخل مجال تعريفه، بل تعطى احتمال أن يقع هذا المتغير بين قيمتين من قيم في مجال تعريفه، ولذلك عادة لا يتم تمثيل التوزيع الاحتمالي للمتغير العشوائي بجدول كما سبق أن فعلنا مع المتغيرات العشوائية المتقطعة، بل نعبر عن احتمال المتغير العشوائي المتصل بدالة احتمال، ويقدر الاحتمال بالمساحة المحصورة بين القيمتين تحت المنحنى الخاص بتلك الدالة ومحور السينات، وتسمى هذه الدالة **دالة كثافة الاحتمال أو دالة الكثافة الاحتمالية** *Probability Distribution Function*، ونرمز لها بالرمز *f(x)*، وتعطى بالصيغة الرياضية التالية:

$$f(x) = \begin{cases} f(x) & si: x \in \Omega(x) \text{ أي } a < x < b \\ 0 & si: x \notin \Omega(x) \text{ أي } a > x > b \end{cases}$$

ويتم تمثيل التوزيع الاحتمالي للمتغيرات العشوائية المستمرة في شكل مدرج تكراري النسبي أو في شكل مضلع. ونجد أن شكل المدرج أو المضلع هو أقرب وصف لمنحنى التوزيع الاحتمالي للمتغير المستمر، وكلما ضاقت الفترات بين مراكز الفئات، يمكن الحصول على رسم دقيق للمضلع الخاص بدالة احتمال المتغير المستمر، كما هو مبين بالشكل التالي:

الشكل رقم 2-2: التمثيل البياني لدالة الكثافة الاحتمالية للمتغيرات العشوائية المستمرة



والمساحة أسفل المنحنى أو في المضلع تعبر عن مجموع الاحتمالات الكلية، ولذا تساوي هذه المساحة الواحد الصحيح، وبفرض أن المتغير العشوائي المستمر معرف على المجال]a-b[نكتب:

$X(\Omega): \{X=x_i : a<x_i<b\}$

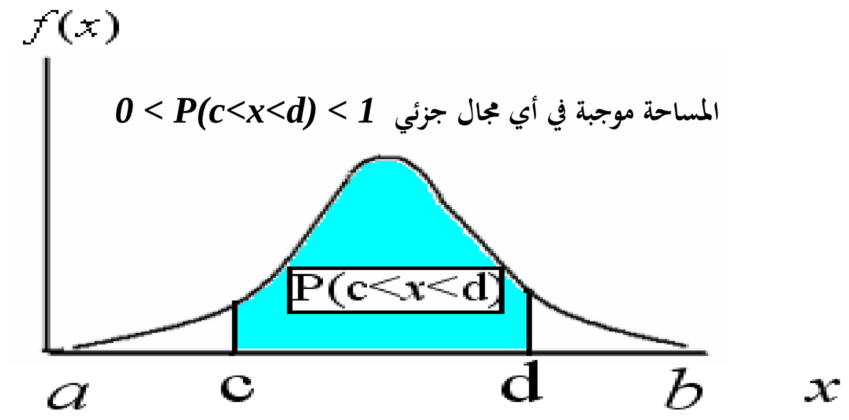
تستخدم دالة الكثافة لحساب احتمال وقوع المتغير العشوائي داخل مجال معين ضمن مجال تعريفه. ويكون ذلك بحساب المساحة تحت منحنى *f(x)*، بين حدود المجال المحدد من خلال حساب تكامل الدالة في ذلك

المجال، فإذا كانت X متغيرة عشوائية مجال تعريفها *[a-b]* وأردنا حساب احتمال وقوع x ضمن مجال جزئي *[c-d]* فإننا نقوم بحساب الدالة التالية:

$$p(c < x < d) = \int_{c}^{d} f(x)dx$$

حيث أن الشكل الرياضي أعلاه يسمى بالتكامل المحدد من x=c حتى x=d، وهذا يعني إيجاد المساحة أسفل المنحني بين (a.b). وهذا ما سنوضحه من خلال الشكل أدناه:

الشكل رقم 2–2: التمثيل البياني لكيفية حساب المساحة تحت المنحنى دالة الكثافة



وبصفة عامة فإن احتمال وقوع المتغير العشوائي X في أي فترة تحت منحنى دالة كثافة الاحتمال يساوي نسبة المساحة فوق ذلك المجال الجزئي إلى مجموع المساحة تحت المنحنى على مجال تعريف الدالة العام.

ولكي تعتبر الدالة *f(x)* دالة كثافة احتمال يجب تحقق شرطين أساسيين في التوزيع الاحتمالي هما:

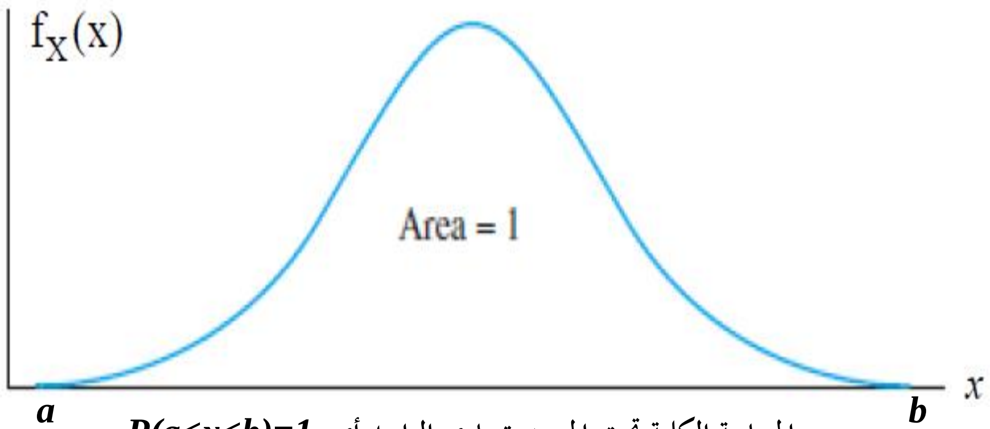
الشرط الأول: الدالة *f(x)* موجبة داخل مجال تعريفها ضمن أي مجال جزئي من]a–b[أي:

- $\forall\, x \in\ ]a-b[\Longrightarrow 0 < f(x) < 1$

الشرط الثاني: التكامل على حدود مجال تعريف المتغير العشوائي المستمر من الحد الأدنى a إلى الحد الأعلى b تساوي مجموع الاحتمالات الكلية وتساوي الواحد الصحيح أي:

- $p(a < x < b) = \int_{a}^{b} f(x)dx = 1$

الشكل رقم 2–3: التمثيل البياني يوضح كل المساحة تحت دالة الكثافة الاحتمالية



المساحة الكلية تحت المنحنى تساوي الواحد أي $P(a<x<b)=1$

من خلال ما تم التطرق إليه نستنتج أن خصائص دالة الكثافة الاحتمالية للمتغير العشوائي المستمر هي:

- دالة كثافة الاحتمال للمتغير العشوائي المتصل س تعطى الاحتمال بين أي قيمتين من قيم المجال المعرف عليه المتغير العشوائي وهذا الاحتمال يكافئ المساحة تحت المنحنى المحصورة بين هاتين القيمتين.
- المساحة تحت دالة كثافة الاحتمال والمحصورة بين الحد الأدنى والحد الأعلى للفترة المعرف عليها المتغير المتصل تساوى الواحد الصحيح.
- احتمال أن يأخذ المتغير العشوائي نقطة معينة في الفترة المعرف عليها يساوى صفرا لأنها الدالة تعطى الاحتمال كمساحة بين نقطتين تحت المنحنى.

د- **التوقع الرياضي والتباين للمتغير العشوائي المنفصل**: إذا كانت $f(x)$ دالة كثافة الاحتمال للمتغير العشوائي X، حيث أنه معرف على المجال]a–b[أي أن a<x<b، فإن التوقع الرياضي للدالة $f(x)$ يعرف بالصيغة الرياضية التالية:

$$\mu = E(x) = \int_a^b x * f(x)dx$$

أما تباين المتغير العشوائي المستمر يعطى بالصيغة الرياضية التالية:

$$\sigma^2 = E(X^2) - E(X)^2 \Rightarrow \sigma^2 = E(X^2) - \mu^2$$

مع العلم أن $E(X^2)$ يحسب بالطريقة التالية:

$$E(x^2) = \int_a^b x^2 * f(x)dx$$

مثال 2: إذا كان إنفاق الأسر في مجتمع معين له دالة كثافة احتمالية تأخذ الشكل الرياضي التالي:

$$f(x) = \begin{cases} 0.006x * (10 - x) \ , & 0 < x < 10 \\ 0 & sinon \end{cases}$$

المطلوب: أوجد المتوسط الحسابي، التباين والانحراف المعياري؟

الحل:

- **إيجاد المتوسط الحسابي** *E(X)*:

$$\mu = E(x) = xf(x)dx = \int_0^{10} x(0.006x * (10 - x)dx$$

$$\mu = 0.006\int_0^{10}(10x^2 - x^3)dx = 0.006\left[10\frac{x^3}{3} - \frac{x^4}{4}\right]_0^{10}$$

$$\mu = 0.006\left(10\frac{10^3}{3} - \frac{10^4}{4}\right) - 0.006\left(10\frac{0^3}{3} - \frac{0^4}{4}\right)$$

$$\mu = 0.006\left(\frac{10000}{3} - \frac{10000}{4}\right) - 0 \Rightarrow \mu = E(x) = 5$$

- **إيجاد التباين** σ^2:

$$\sigma^2 = E(x^2) - \mu^2$$

$$E(x^2) = \int_0^{10} x^2 f(x)dx = \int_0^{10} x^2(0.006x * (10 - x)dx$$

$$E(x^2) = 0.006\int_0^{10}(10x^3 - x^4)dx = 0.006\left[10\frac{x^4}{4} - \frac{x^5}{5}\right]_0^{10}$$

$$E(x^2) = 0.006\left(10\frac{10^4}{4} - \frac{10^5}{5}\right) - 0.006\left(10\frac{0^4}{4} - \frac{0^5}{5}\right)$$

$$E(x^2) = 0.006\left(\frac{100000}{4} - \frac{100000}{5}\right) - 0 \Rightarrow E(x^2) = 30$$

$$\sigma^2 = 30 - 5^2 = 30 - 25 = 5$$

- **إيجاد الانحراف** σ:

$$\sigma = \sqrt{\sigma^2} = \sqrt{5}$$

II- التوزيعات الاحتمالي الأكثر استخداما للمتغيرات العشوائية المتقطعة والمستمرة

عند إجراء أي تجربة عشوائية أو دراسة أي ظاهرة إحصائية، فإن أول هدف يسعى إليه الباحث هو تعريف وتحديد المتغير العشوائي الذي يصف هذه الظاهرة، بغية اشتقاق التوزيع الاحتمالي لهذا المتغير العشوائي. ويساعد فيهم سلوك المتغير العشوائي في دراسة سلوك الظاهرة التي نريد دراستها، كما يسمح معرفة التوزيع الاحتمالي للمتغيرات العشوائية المستخدمة بالدراسة في حساب الكثير من المقاييس الإحصائية العامة التي تساعد على اتخاذ القرار المناسب للمشكلة تحت الدراسة مثل الوسط الحسابي والتباين والانحراف المعياري.

وذلك بهدف تيسير دراسة وفهم سلوك الظواهر باستخدام هذه التوزيعات واتخاذ القرار بشأنها أو التنبؤ بسلوكها المستقبلي. توصل علماء الإحصاء إلى اشتقاق العديد من التوزيعات الاحتمالية التي تصلح للتعبير عن الكثير من الظواهر وذلك تحت شـروط معينة. وحســب طبيعة المتغيرات العشــوائية فإن هذه التوزيعات بدورها منها ما هو متقطع مثل توزيع برنولي، توزيع ذي الحدين، وتوزيع بواســون، ومنها ما هو متصــل مثل التوزيع الأسي والتوزيع الطبيعي وغيرها.

1. التوزيعات الاحتمالية للمتغيرات العشوائية المتقطعة

ترتبط أهم التوزيعات الاحتمالية ارتباطا وثيق بنتائج التجارب العشـــوائية الثنائية، وحياتنا اليومية زاخرة بهذا النوع من التجارب نظرا لأهميتها التطبيقية في مختلف العلوم. وهناك عدة توزيعات احتمالية متقطعة منها توزيع برنولي، توزيع ثنائي الحدين، توزيع ثنائي الحدين الســـالب، توزيع كثير الحدود، توزيع الهيبروجيومتريك وتوزيع بواســون وتوزيع كثير الحدود، ولكننا ســتقتصــر دراســتنا فقط على توزيع برنولي، توزيع ثنائي الحدين وتوزيع بواسون نظرا لما لها من تطبيقات واسعة في ميادين الإدارة والاقتصاد.

أ- **توزيع برنولي *Bernoulli Distribution***: عندما نقوم برمي قطعة نقود فإن أمامنا احتمالين فقط إما ظهور صورة أو الرقم، أثناء الضغط على زر قاطعة المصباح هناك احتمالين إما أن يشتغل المصباح او لا يشتغل، إذا اخترنا عشوائيا قطعة من انتاج مصنع معين قد تكون تلك القطعة جيدة وسليمة أو قد تكون معيبة، وعلى الرغم من تنوع هذا النوع من التجارب إلا أن جميعها لها خصائص تجربة الثنائية التي لها نتيجتين متنافيتين وتسمى هذه التجربة بتجربة برنولي.

وكتعريف نقول إن توزيع برنولي ينشأ من تجربة عشوائية لها نتيجتين فقط، نسمي النتيجة الأولى اصطلاحا بالنجاح ونرمز لها بالرمز S، ونسمي النتيجة الثانية بالفشل ونرمز لها بالرمز F. لذلك فإن فراغ العينة في توزيع

برنولي هو Z حيث $Z: \{S.F\}$ ونرمز لاحتمال النجاح بالرمز P ونرمز لاحتمال الفشل بالرمز q مع العلم أن $q=1-p$

ونكتب أن Z متغير عشوائي يتبع توزيع برنولي، فإن:

$$Z(S)=1$$
$$Z(P)=0$$

وعليه فإن القيم الممكنة للمتغير العشوائي هي: $Z(S): \{0.1\}$ واحتمالاته هي:

$$P(Z=1)=P(S)=P$$
$$P(Z=0)=P(F)=1-P=q$$

- **التوقع الرياضي والتباين في توزيع برنولي:** يمكن حساب قيمة التوقع الرياضي والتباين للمتغير العشوائي الذي يتبع توزيع برنولي وفق العلاقة الرياضية التالية:

$$\mu_x = p$$
$$\sigma_x{}^2 = p * q$$

ب- **توزيع ثنائي الحد *The Binomiale Distribution*:** يعتبر توزيع ثنائي الحد او ذو الحدين من تكثر التوزيعات الاحتمالية شيوعا، ومجالات تطبيقه متعددة منها فحص الجودة، المبيعات، التسويق، بحوث الرأي والاستطلاع ...الخ، ويستخدم توزيع ذي الحدين في الحالات التي يتم فيها الاعتماد على العينة لتقدير نسبة معينة في المجتمع، ويكون للظاهرة المدروسة نتيجتان متنافيتان فقط كما في تجربة برنولي، لكن مع تكرار التجربة أكثر من مرة، يعني أن عدد المحاولات هي $n>1$، وبافتراض أن هذه المحاولات مستقلة عن بعضها البعض أي أن نتيجة كل محاولة مستقلة عن نتيجة المحاولات الأخرى، كما يكون احتمال النجاح ثابت في جميع المحاولات $P=P(S)$.

ويعرف المتغير العشوائي X على أنه عدد مرات النجاح عند تكرار إجراء تجربة برنولي n مرة وفق الشروط المبينة أعلاه، وبالتالي فإن هذا التوزيع معرف بالمعلمتين n وP، ومجموعة القيم الممكنة (فضاء العينة) لهذا المتغير العشوائي هي:

$$X(S): \{0.1.2.3\ldots\ldots n\}$$

وفيما يتعلق لدالة التوزيع الاحتمالي للمتغير العشوائي X الذي يتبع توزيع ثنائي الحد فهي تعطى بالصيغة التالي:

$$f(x) = p(X = x_i) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & ; \quad x = 0,1,2,3 \dots n \\ 0 & \quad x \neq 0,1,2,3 \dots n \end{cases}$$

حيث:
$$\binom{n}{x} = \frac{n\,!}{x\,!(n-x)\,!}$$

وهناك حالة خاصة لتوزيع ذي الحدين التي يكون فيها $n=1$ ويصبح حينها توزيع ذي الحدين نفسه توزيع برنولي بمعلمة واحدة p وتأخذ الدالة في هذه الحالي الشكل:

$$f(x) = p^x (1-p)^{1-x} \quad ; \quad x = 0\,,1$$

- **التوقع الرياضي والتباين في توزيع ذي الحدين**: إذا كان لدينا X متغير عشوائي يتبع توزيع ذي الحدين بالمعلمتين n عدد المحاولات وp احتمال النجاح، فإن التوقع الرياضي والتباين للمتغير العشوائي X يعطى وفق العلاقة الرياضية التالية:

$$E(x) = \mu(x) = n * p$$
$$\sigma^2 = n * p * q$$

مثال: إذا كانت نسبة القطع التالفة لأحد مصانع الأجهزة الخلوية هي 10%. وتم أخذ عينة من 5 أجهزة بشكل عشوائي من أنتاج هذا المصنع، ما هو احتمال الحصول على:

- جميع الأجهزة غير عاطلة (تعمل)؟
- جهاز خليوي واحد تالف؟
- جهاز واحد على الأقل تالف؟
- جميع الأجهزة عاطلة؟
- أحسب التوقع الرياضي الانحراف المعياري

الحل:

ليكن المتغير X الذي يدل على عدد الأجهزة الخلوية التالفة. نتيجة النجاح تعني الحصول على جهاز خليوي عاطل، بالتالي احتمال النجاح $(p=0.1)$ أما نتيجة الفشل فهي الحصول على جهاز خليوي سليم، بالتالي احتمال الفشل $(q=0.9)$.

يمكن القول أن المتغير العشوائي X يتبع إذن توزيع ثنائي الحد بوسيطين $p=1.0$ و$n=5$ ونكتب:

$$X \backsim b(x: 5, 0.1)$$

نكتل بشكل عام صيغة دالة التوزيع الاحتمالي في توزيع ذي الحدين كما يلي:

$$f(x) = P(X = x) = \binom{5}{x} (0{,}1)^x (0{,}9)^{5-x} \quad ; \quad x = 0,1,2,3,4,5$$

انطلاقا من هذه العلاقة العامة يمكن حساب:

- $f(1) = P(X = 1) = \binom{5}{1}(0{,}1)^1(0{,}9)^4 = 0{,}328055$
- $f(1) = P(X = 1) = \binom{5}{1}(0{,}1)^1(0{,}9)^4 = 0{,}328055$
- $f(1) = P(X = 1) = \binom{5}{1}(0{,}1)^1(0{,}9)^4 = 0{,}328055$
- $f(5) = P(X = 5) = \binom{5}{5}(0{,}1)^5(0{,}9)^0 = 0{,}00001$
- $f(0) = P(X = 0) = \binom{5}{0}(0{,}1)^0(0{,}9)^5 = 0{,}59049$
- $\int_1^5 f(x) = P(X \geq 1) = 1 - P(x = 0) = 1 - 0{,}59049 = 0{,}40951$

التوقع الرياضي: $E(X) = \mu_X = n * p = 0{,}5 * 0{,}1 = 0{,}05$

التباين: ${\sigma_X}^2 = n * p * q = 0{,}5 * 0{,}1 * 0{,}9 = 0{,}045$

ج- **توزيع بواسون *Poisson Distribtion***: يسمى أيضا بقانون الاحتمالات الصغيرة وكذا قانون الأحداث النادرة، وهو توزيع من التوزيعات الاحتمالية المنفصلة، حيث يستخدم للتعبير عن عدد مرات حدوث ظاهرة معينة في فترة زمنية محددة أو مساحة محددة أو طول محدد وبشرط أن تتصف هذه الأحداث بالندرة والاستقلال.

كما يستخدم هذا لتوزيع للتعبير عن الظواهر والأحداث نادرة الوقوع، ومن أمثلة الأحداث: عدد الأخطاء المطبعية في صفحة مختارة عشوائيا من كتاب، عدد الحوادث المرورية في فترة زمنية محدودة، عدد المرات التي يذهب فيها أفراد أسرة معينة إلى الطبيب خلال شهر معين، عدد الآلات الإنتاجية التي تتعطل في مصنع خلال فترة زمنية معينة، عدد الحوادث التي تقع على إحدى الطرق السريعة في يوم ما حيث تعتبر هذه الحوادث نادرة الوقوع إذا ما قورنت بعدد السيارات التي تسلك هذه الطريق في ذلك اليوم...الخ. ويمثل توزيع بواسون النهاية التي يؤول إليها توزيع ذي لحدين عندما $n \longrightarrow \infty$ و $p \longrightarrow 0$.

ومن خصائص توزيع بواسون:

- إن احتمال الحدوث يكون نفسه في فترتين إذا كان لهما نفس الطول.
- إن حدوث أو عدم حدوث حدث في فترة زمنية واحدة مستقل عن حدوث أو عدم حدوث هذا الحدث في فترة زمنية أخرى.

فإذا كان X متغير عشوائي يعبر عن عدد حالات النجاح لحدث نادر الوقوع، باحتمال P حيث P صغيرة جدا $(p \longrightarrow 0)$ وفى عدد كبير جدا من المحاولات المستقلة يساوى n حيث $(n \longrightarrow \infty)$ فإننا نقول

عن المتغير العشوائي X أنه يتبع توزيع بواسون بمعلمة λ ونكتب باختصار $X \sim p(x;\lambda)$، ودالة التوزيع الاحتمالي للمتغير العشوائى X تعطى بالعلاقة الآتية:

$$P(X;\lambda) = f(x) = \frac{\lambda^x * e^{-\lambda}}{x!} \quad ; \quad x = 1,2,3, \dots \dots . \infty$$

حيث أن: $f(x)$ هو احتمال حدوث x في مجال معين، $\lambda = (p * n) > 0$: تسمى معلمة التوزيع معلمة التوزيع وهي عبارة عن مقدار ثابت يعبر على التوقع الرياضي (عدد مرات الحدوث في الفترة) والتباين في نفس الوقت، e: أساس اللوغاريتم الطبيعي $(e=2.71828)$.

نلاحظ أنه في توزيع بواسون، X هو متغير عشوائي منفصل يشير إلى عدد التكرارات في مجال معين. ونظرًا لعدم وجود حد أعلى لعدد تكرارات X، فإن دالة الاحتمال $f(x)$ قابلة للتغير والزيادة بلا حدود لهذا تم في دالة التوزيع وضع الافتراض $x = 1,2,3, \dots \dots . \infty$

مثال: إذا كان متوسط وصول الطائرات إلى مطار مدينة باتنة طائرتين كل ساعة، فما هو احتمال وصول 3 طائرات إلى هذا المطار في ساعة واحدة.

الحل:

من خلال المعطيات نستنتج أن قيمة $\lambda = 2$ وعليه سنقوم بحساب $P(X = 3)$ حيث:

$$P(X = 3) = \frac{2^3 * e^{-2}}{3!} = 0.18045$$

مثال: إذا كان عدد الأخطاء المطبعية في كتاب يتكون من 600 صفحة هو 50 خطأ، فإذا كانت الأخطاء تتوزع توزيعا عشوائيا، فما احتمال إذا اختيرت 10 صفحات عشوائيا ألا تحتوي على أي خطأ؟

الحل:

بفرض أن X يمثل عدد الأخطاء في كل صفحة وأن عدد المحاولات (الصفحات) تمثل سلسلة من محاولات برنولي عددها n=10 فنسبة الخطأ (النجاح) هي p=50/600=0.083، وعليه فإن:

$$\lambda = n * p = 10(0.083) = 0.83$$

وبالتالي فإن X يتبع توزيع بواسون حيث $X \sim p(x; 0.83)$ وأن:

$$P(X = 0) = \frac{0.83^0 * e^{-0.83}}{0!} = 0.436$$

- **التوقع الرياضي والتباين في توزيع بواسون:** إذا كان X متغير عشوائي يتبع توزيع بواسون $X \sim p(x;\lambda)$ فإن توقع وتباين متغير بواسون العشوائي X يساوي معلمة التوزيع أي يساوي λ:
- $E(x) = {\sigma_x}^2 = \lambda$

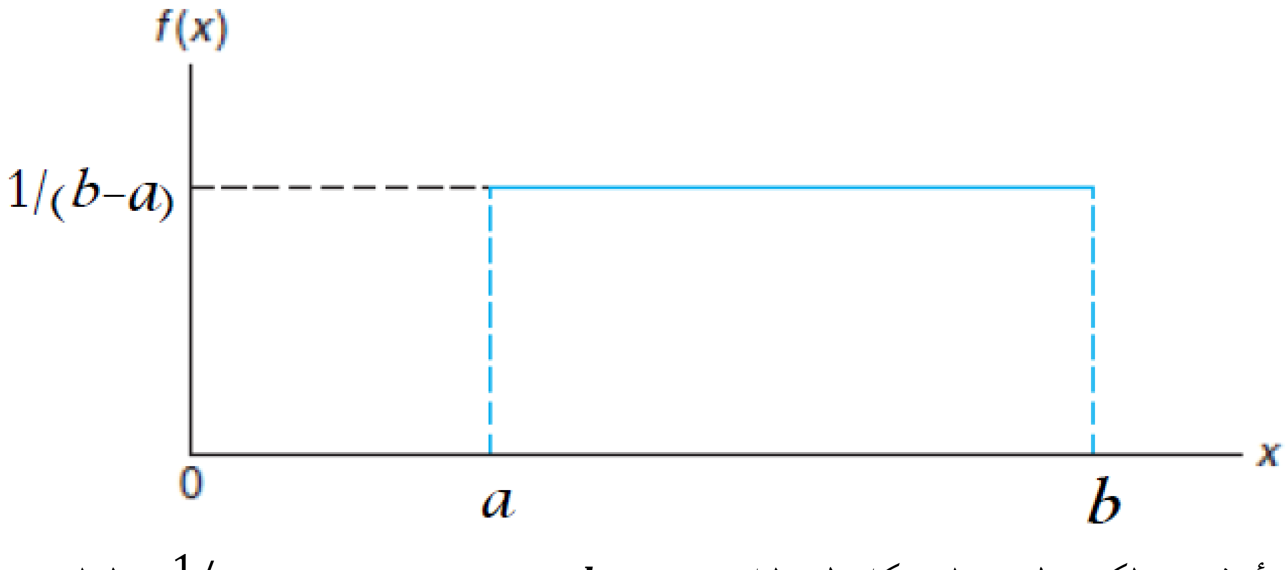
2. التوزيعات الاحتمالية للمتغيرات العشوائية المستمرة

هناك العديد من التوزيعات الاحتمالية المستمرة، لكل منها دالة كثافة احتمال محددة، ومن بين أهم التوزيعات الاحتمالية المستمرة نجد (التوزيع المنتظم، التوزيع الطبيعي، التوزيع الأسي، توزيع ستودنت، توزيع قاما، توزيع كاي مربع، التوزيع الطبيعي اللوغاريتمي، توزيع ويب، توزيع رايلي وتوزيع رايس ...الخ) وسنحاول في هذ العنصر التطرق لبعض هذه التوزيعات بعض هذه التوزيعات:

أ- **التوزيع الاحتمالي المنتظم *Uniform distribution*:** يعتبر التوزيع المنتظم من أبسط التوزيعات المستمرة في الحصاء، حيث أن الاحتمال فيه منتظم (متساو) على مجال مغلق نسميه $[a-b]$، أي أن له دالة احتمال ثابتة، ويستخدم في الظواهر التي يمكن أن تحدث بشكل منتظم، ونقول عن متغير عشوائي X أنه يتبع التوزيع المنتظم مداه هو المجال $[a-b]$ فإن تابع الاحتمال (دالة كثافة الاحتمال) $f(x)$ الخاصة به تعطى بالصيغة التالية:

$$f(x;a,b) = \begin{cases} \dfrac{1}{b-a} & ; \quad a \le x \le b \\ 0 & \quad sinon \end{cases}$$

الشكل رقم 2–4: يوضح التمثيل البياني لدالة التوزيع الاحتمالي المنتظم



ويأخذ تابع الكثافة الاحتمالي شكل المستطيل قاعدته ***b-a*** وارتفاع ثابت يساوي $1/_{b-a}$، وبالتالي فإنه من الواضح رؤية أن مساحة هذا المستطيل تساوي الواحد.

مثال: بفرض أن طول فترة صيانة السيارات في محطة الصيانة بالساعات تعتبر متغير عشوائي X يخضع لتوزيع منتظم على المجال $[0,4]$، المطلوب إيجاد تابع الكثافة الاحتمالية، وما هو احتمال أن تستغرق صيانة سيارة معينة أكثر من ثلاث ساعات.

الحل:

- تابع الكثافة الاحتمالية الخاصة بالمتغير العشوائي X هي:

$$f(x;a,b)=f(x;0,4)=\frac{1}{4-0}=\frac{1}{4} \qquad ; \qquad 0\le x\le 4$$

- احتمال أن تستغرق صيانة سيارة معينة أكثر من ثلاث ساعات$(p(x\ge 3)$:

$$p(x\ge 3)=\int_3^4 \frac{1}{b-a}dx=\int_3^4 \frac{1}{4-0}dx=\int_3^4 \frac{1}{4}dx=\frac{1}{4}$$

- **التوقع الرياضي والتباين للتوزيع الاحتمالي المنتظم**: إذا كان X متغير عشوائي يتبع توزيع منتظم على المجال $[a-b]$ فإن:
  - التوقع الرياضي: $E(X)=\mu_X=\frac{a+b}{2}$
  - التباين: $\sigma_X{}^2=\frac{\boldsymbol{(a+b^2)}}{\mathbf{12}}$

- **دالة التوزيع التراكمي (التجميعي) للتوزيع الاحتمالي المنتظم**: إذا كان X متغير عشوائي يتبع توزيع منتظم على المجال $[a-b]$ فإن تابع التوزيع التراكمي يعطى بالصيغة الرياضية التالية:

$$f(x)=p(X\le x)=\begin{cases} 0 & si \quad x<a \\ \frac{1}{b-a}(x-a) & si\ x\epsilon[a,b] \\ 0 & si \quad x>b \end{cases}$$

ب- **التوزيع الاحتمالي الأسي *Exponential Distribution***: يعد التوزيع الأسي من التوزيعات المستمرة المهمة التي تستخدم على نطاق واسع في الحياة، وهو يعالج حالات مثل:

- الزمن العشوائي لمدة بقاء المريض في المشفى لتلقي العلاج؛
- الزمن العشوائي لصلاحية الدواء؛
- الوقت اللازم لتحميل شاحنة.

ونقول عن متغير عشوائي أنه يتبع التوزيع الأسي بالمعلمة λ حيث $\lambda > 0$ إذا كان له دالة كثافة احتمالية بالصيغة الموالية:

$$f(x) = \begin{cases} \frac{1}{\mu}.e^{\frac{-x}{\mu}} & ; \quad si\ x > 0 \\ 0 & ; \quad si\ x \leq 0 \end{cases}$$

حيث أن μ هو التوقع الرياضي أو المتوسط.

وإذا كانت $\lambda = \frac{1}{\mu}$ تصبح دالة كثافة الاحتمال للتوزع الأسي معطاة بالصيغة التالية:

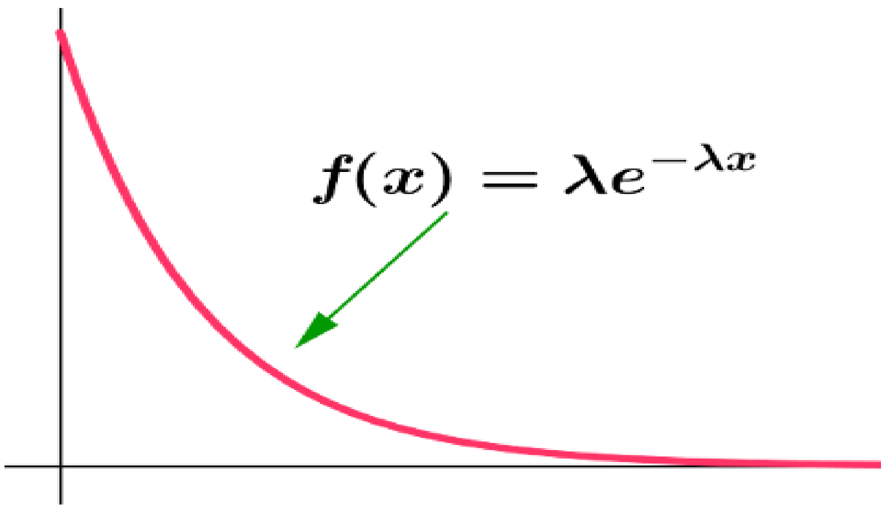
$$f(x) = \begin{cases} \lambda.e^{-\lambda.x} & ; \quad si\ x > 0 \\ 0 & ; \quad si\ x \leq 0 \end{cases}$$

ومن خصائص القانون التوزيع الأسي أن التوقع الرياضي والانحراف المعياري للتوزيع متساويان. وبالتالي يكون التباين هو مربع التوقع الرياضي:

- $\sigma_x = E(x) = \sqrt{{\sigma_x}^2}$

ويأخذ منحنى الدالة *f(X)* الشكل البياني التالي:

الشكل رقم 2–5: يوضح التمثيل البياني لدالة التوزيع الاحتمالي الأسي



كما هو الحال مع أي توزيع مستمر، تعبر المساحة الواقعة أسفل المنحنى في فترة زمنية معينة على احتمال أن يأخذ المتغير العشوائي قيمة تنتمي إلى هذه الفترة. وسنحتاج في هذه الحالة إلى دالة الكثافة الاحتمالية التراكمية (التجميعية) والتي تعطى حسب هذا القانون بالصيغة التالية:

$$f(x) = p(X \leq x_0) = \int_0^{x_0} f(x)dx = 1 - e^{\frac{-x_0}{\mu}}$$

- **التوقع الرياضي والتباين في التوزيع الأسي:**

مثال: بفرض أن وقت تحميل الشاحنة بالبضائع في مخزن معين يتبع التوزيع الأسي. وكان متوسط وقت تحميل الشاحنة 15 دقيقة ($\mu = 15$).

المطلوب: ماهي دالة الكثافة الاحتمالية المناسبة لهذا المتغير؟

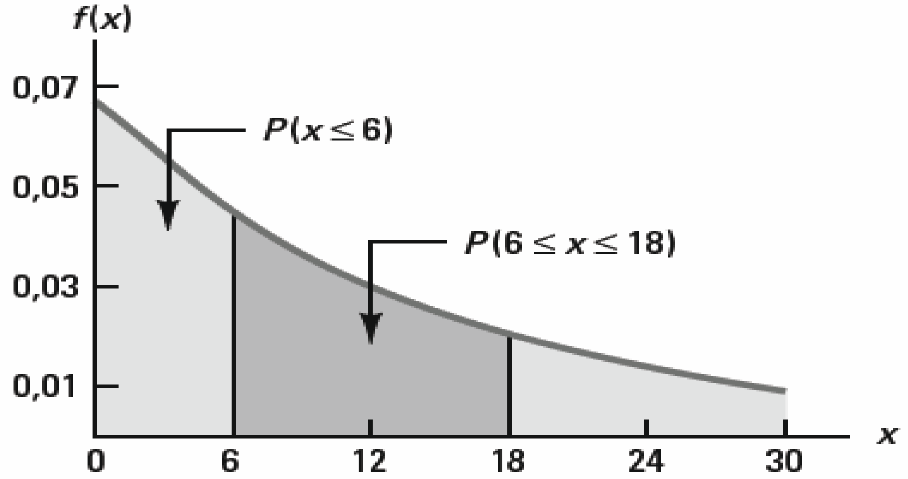
- مثل بيانيا شكل الدالة $f(x)$؟
- ما هو احتمال أن يتم تحميل الشاحنة في ست دقائق على الأكثر؟
- ما هو احتمال أن يتم تحميل الشاحنة في ثمانية عشرة دقيقة على الأكثر؟
- ما هو احتمال أن يتم تحميل الشاحنة في ثمانية عشرة دقيقة على الأكثر وست دقائق على الأقل؟
- أحسب تباين التوزيع في هذا المثال؟

الحل:

- **دالة الكثافة الاحتمالية:** إذا كان X متغير عشوائي ذو توزيع احتمالي أسي يعبر على وقت تحميل الشاحنة بالبضائع في المخزن، وكان متوسط وقت تحميل الشاحنة 15 دقيقة فإننا نقوم بكتابة دالة الكثافة المناسبة بالصيغة التالية:

$$f(x) = \frac{1}{15}.e^{\frac{-x}{15}}$$

- **التمثيل البياني للدالة $f(x)$:**



- لحساب الاحتمالات الأسية المطلوبة أعلاه، نستخدم الصيغة التراكمية.

$$p(x \leq 6) = 1 - e^{\frac{-6}{15}} = 0.3297$$

$$p(x \leq 18) = 1 - e^{\frac{-18}{15}} = 0.6988$$

احتمال أن يكون وقت تحميل الشاحنة بين 6 و18 دقيقة يساوي:

$$p(x \leq 18) - p(x \leq 6) = 0.6988 - 0.3297 = 0.3691$$

يمكن حساب احتمالات أي فترة زمنية أخرى بنفس الطريقة.

إن احتمال تحميل الشاحنة خلال 6 دقائق على الأكثر، $P(x \leq 6)$، يتوافق مع المنطقة الواقعة أسفل المنحنى، التي يمثلها الشكل أعلاه، بين $x=0$ و$x=6$. وبالمثل، فإن احتمال تحميل الشاحنة خلال 18 دقيقة على الأكثر، $P(x \leq 18)$ يتوافق مع المنطقة الواقعة أسفل المنحنى بين $x=0$ و $x=18$. ونلاحظ أيضًا أن احتمال أن يكون وقت تحميل الشاحنة بين 6 و18 دقيقة $P(6 \leq x \leq 18)$ تقابل المنطقة الواقعة تحت المنحنى بين $x=6$ و$x=18$.

إذا كان متوسط وقت تحميل الشاحنة هو 15 دقيقة. وحسب خاصية القانون الأسي فإن المتوسط والانحراف المعياري للتوزيع متساويان. وبالتالي فإن الانحراف المعياري لوقت تحميل الشاحنة هو 15 دقيقة $\sigma_x = E(x) = 15$ وعليه فإن قيمة التباين هي مربع قيمة الانحراف المعياري:

- $\sigma_x{}^2 = (\sigma_x)^2 = (15)^2 = 225$

ج- **التوزيع الاحتمالي الطبيعي (توزيع لابلاس-قوس) *Normal Distribution***: يعد التوزيع الطبيعي من أهم التوزيعات الاحتمالية المستمرة شيوعا واستخداما في النواحي التطبيقية، لما له من خصائص تنطبق على نسبة كبيرة من الظواهر الطبيعية والاجتماعية والاقتصادية، ويطلق على التوزيع الطبيعي الرمز $\boldsymbol{N}$. فلو اخترنا بالصدفة مئة أو ألفا من المارين في شارع ما، وقمنا بقياس أطوالهم سنجد نسبة كبيرة منهم تقترب أطوالهم من قيمة المتوسط، ونسبة قليلة من المارة أطوالهم أكثر بكثير من المتوسط أو أقل بكثير من قيمة المتوسط. ولو مثلنا هذه البيانات في معلم متعامد متجانس لكان المنحنى الذي يمثل النسبة، أو ما يمكن أن نسميه الاحتمال، ذا شكل جرسي متماثل حول المتوسط.

ويقال أن المتغير العشوائي المستمر $\boldsymbol{X}$ أن له توزيعا طبيعيا مداه هو $-\infty < \boldsymbol{x} < +\infty$ بمتوسط μ وتباين $\sigma_x{}^2$ إذا كانت دالة كثافته الاحتمالية $f(x)$ تأخذ الصيغة التالية:

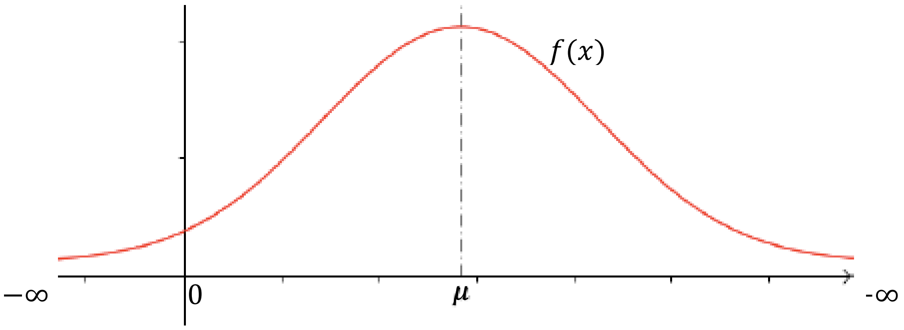
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad -\infty < x < +\infty$$

حيث: $\sigma > 0$، π: ثابت، $-\infty < \mu < +\infty$

وتوجد معلمتين يعرف بها التوزيع الطبيعي N هكا التوقع الرياضي $E(x) = \mu$، والتباين $\boldsymbol{Var(x) = \sigma^2}$. لهذا نكتب: $x \sim N(\mu, \sigma^2)$

وإذا أردنا تمثيل دالة الكثافة الاحتمالية للتوزيع الطبيعي سنجد انها تأخذ الشكل الجرسي المتماثل بالنسبة للمتوسط كما يوضحه الشكل أدناه:

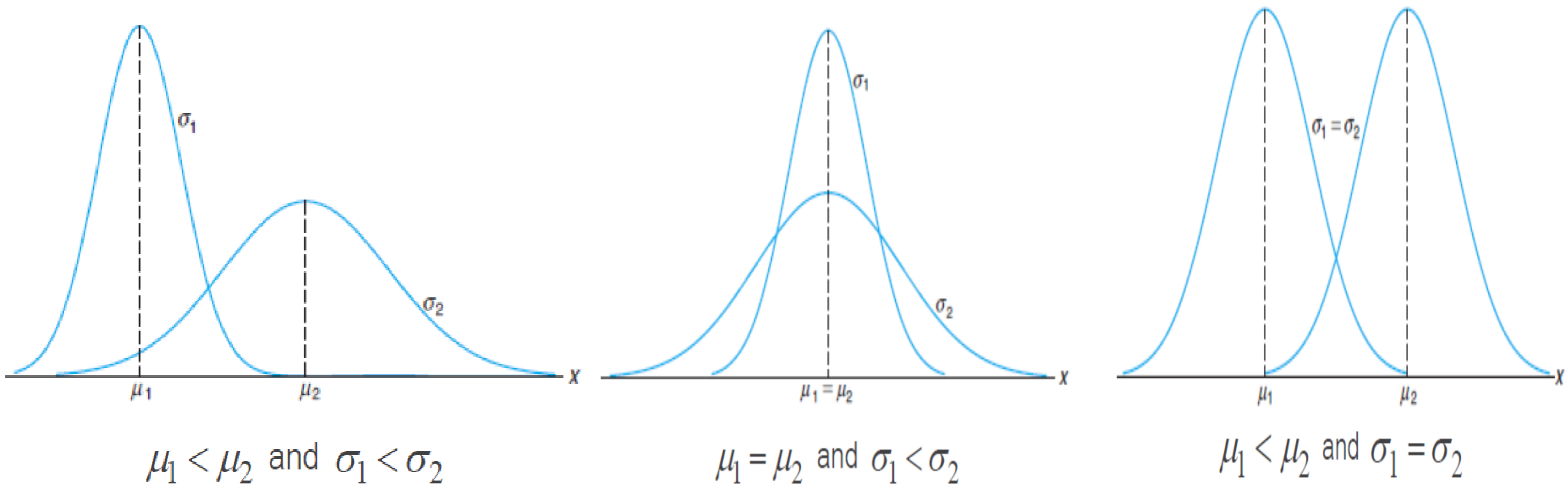
الشكل رقم 2–6: يوضح التمثيل البياني لدالة الكثافة الاحتمالية للتوزيع الطبيعي



كما في التوزيعات المستمرة السابقة تساوي المساحة تحت منحنى دالة الكثافة الاحتمالية للتوزيع الطبيعي الواحد الصحيح.

وهناك عدد ا نهائي من منحنيات التوزيع الطبيعي وتختلف عن بعضها البعض تبعا لاختلاف معلمتي كل توزيع التوقع الرياضي والانحراف المعياري (μ_1, σ_1)، وقد تتفق منحنيات المختلف الظواهر في أحد المعلمتين كأن تتفق ظاهرتين في الانحراف المعياري ولكن تختلف في قيمة التوقع الرياضي (المتوسط)، أو يمكن أن تتفق ظاهرتين في قيمة التوقع الرياضي بينما تختلف قيمتي الانحراف المعياري للظاهرتين، أو يمكن أن تختلف قيمة كلا المعلمتين وهذه الحالات جميعها تجعل من منحنى دالة الكثافة الاحتمالية للظواهر يختلف باختلاف معلمات التوزيع، فبافتراض أن لدينا توزيعين لظاهرتين يتبعان التوزيع الطبيعي $f(x_1;\, \mu_1, \sigma_1)$ و $f(x_2;\, \mu_2, \sigma_2)$ سنحاول أن نبين من خلال الأشكال التالية تأثير تغير قيم كل من معلمتي التوزيع الطبيعي على شكل دالة كثافة الاحتمال $f(x)$.

الشكل رقم 2–7: يوضح تغير شكل دالة كثافة احتمال التوزيع الطبيعي بتغير قيم (μ, σ)



نلاحظ من خلال الشكل أعلاه أنه عند تمثيل دالة كثافة الاحتمال للتوزيع الطبيعي يجب أن يكون المتوسط μ يناظر القمة الوحيدة لمنحنى التوزيع، وهذا ما يجعل منحنى الدالة متماثل حول المستقيم الرأسي x.)

السلسلة رقم 1:

تمرين رقم 1: لدينا تجربة رمي حجر نرد مرتين، وX متغير عشوائي يمثل الفرق المطلق بين العددين الظاهرين على الوجه العلوي في الرميتين.

المطلوب:

- أوجد القيم التي يمكن أن يأخذها المتغير العشوائي X؟
- عين التوزيع الاحتمالي لهذا المتغير، ثم تأكد أنه فعلا توزيع احتمالي؟
- أحسب التوقع الرياضي، التباين والانحراف المعياري؟

تمرين رقم 2: لاعب يرمي قطعة نرد مرة واحدة، ولكن ضمن الشروط التالية:

- إذا ظهر الرقم 1 على الوجه العلوي يحص اللاعب على نقطتين.
- إذا ظهر الرقم 6 على الوجه العلوي يخسر اللاعب نقطتين.
- فيما عدا ذلك يحصل اللاعب على نقطة واحدة.

فإذا كان x هو هعدد النقاط التي يمكن أن يحصل عليها اللاعب.

المطلوب:

- عين التوزيع الاحتمالي لهذا المتغير ثم تأكدا فعلا أنه توزيعا احتماليا؟
- أحسب التوقع الرياضي، التباين والانحراف المعياري؟

تمرين رقم 3: إذا كان X متغير عشوائي مستمر معرف بدالة الكثافة الاحتمالية التالية:

$$f(x)\begin{cases}\frac{1}{8}x & 0<x<4\\ 0 & sinon\end{cases}$$

المطلوب:

- أثبت أن الدالة $f(x)$ تمثل دالة كثافة احتمالية؟
- أحسب قيمة الاحتمال $p(0<x<\frac{1}{2})$؟
- أحسب التوقع الرياضي، التباين والانحراف المعياري؟

تمرين رقم 4: إذا كان X متغير عشوائي مستمر معرف بدالة الكثافة الاحتمالية التالية:

$$f(x)\begin{cases}kx & 0<x<2\\ 0 & sinon\end{cases}$$

المطلوب:

- أحسب قيمة k؟
- أحسب قيمة الاحتمال $p(0<x<3)$، $p(0.5<x<2)$؟

المحور الثالث:
التوزيع الطبيعي

I- التوزيع الطبيعي N:

تمهيد:

يعتبر التوزيع الطبيعي حجر الزاوية في تطبيقات الاستنتاج الاحصائي في الدراسات التي تعتمد على العينات العشوائية، ويسمى التوزيع الطبيعي من قبل البعض بالتوزيع المعتدل، ويعود اصل اكتشاف التوزيع الطبيعي إلى عالم الرياضيات الفرنسي **دي موافر** *Abraham de Moivre* سنة *1733*، وشارك في تطويره عدد من العلماء من أشهرهم الفرنسي *Pierre-Simon de Laplace*، وعالم الرياضيات الألماني عام *1809* *Carl Friedrich Gaus* الذي يسمى التوزيع أحيانا باسمه.

وقد ألصقت صفة الطبيعي بهذا التوزيع كون اغلب الظواهر الطبيعية تخضع لهذا التوزيع، حيث لو تم تمثيل قيم كل ظاهرة بيانيا سنجد أن أغلب القيم توم حول المتوسط الحسابي، وكلما ابتعدنا عن هذا الأخير فإن عدد المشاهدات(البيانات) ستنخفض باستمرار، لذلك نجد أن هناك عدد قليل من البيانات التي قيمتها أكبر بكثير من المتوسط او أقل بكثير من قيمة المتوسط، ليتخذ بذلك توزيع البيانات في المدرج التكراري أو المنحنى الشكل الجرسي المتماثل حول المتوسط أو قريب من ذلك الشكل، وهذا ما يجعل التوزيع الاحتمالي الطبيعي يلعب دورا أساسيا في عملية الاستقراء الاحصائي.

ويستخدم التوزيع الطبيعي غالبا لتقريب التوزيعات الاحتمالية الأخرى، كون معظم التوزيعات يمكن تقريبها إلى هذا التوزيع إذا كان حجم العينة المدروسة كبيرة $n>30$. بالإضافة إلى ذلك فإن توزيع الكثير من الإحصاءات مثل وسط العينة غالبا ما تتبع توزيعا طبيعيا بغض النظر عن شكل المجتمع الأصلي. وإن عدد كبير من الدراسات التجريبية أشارت إلى أن التوزيع الطبيعي غالبا ما يعطي تمثيلا مناسبا لمجتمعات أو عمليات في بيئات إدارية متنوعة وفي مجالات مختلفة، فمثلا توزيع الأطول، الأوزان، حجم المبيعات الاسبوعية في مطعم معين أو نسبة المعيب في منتج معين، حيث وجدت الدراسات أنه في الغالب تكون التوزيعات الاحتمالية لهذه الظواهر قريبة من التوزيع الطبيعي.

1. دالة الكثافة الاحتمالية للتوزيع الطبيعي N:

كما تم التطرق إليه في المحور السابق فإنه يقال أن المتغير العشوائي **x** يتبع التوزيع الطبيعي بالمعلمتين $x \sim N(\mu, \sigma)$ إذا كانت دالة كثافته الاحتمالية تأخذ الصيغة الرياضية التالية:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{(x-\mu)^2}{\sigma^2}\right]} \quad ; \quad -\infty < x < +\infty$$

حيث: X: متغير عشوائي مستمر معرف على المجال $]-\infty , +\infty[$

μ: هو المتوسط الحسابي للتوزيع ونكتب $-\infty < \mu < +\infty$

σ^2: تباين التوزيع حيث $\sigma^2 > 0$

σ: الانحراف المعياري وهو جذر التباين حيث $\sigma > 0$

2. **خصائص التوزيع الطبيعي** N: نذكر أهم خصائص التوزيع الطبيعي N في النقاط التالية:

- مجموع المساحة تحت منحنى التوزيع تساوي الواحد، أي:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{(x-\mu)^2}{\sigma^2}\right]} dx = 1$$

ومنه فإن:

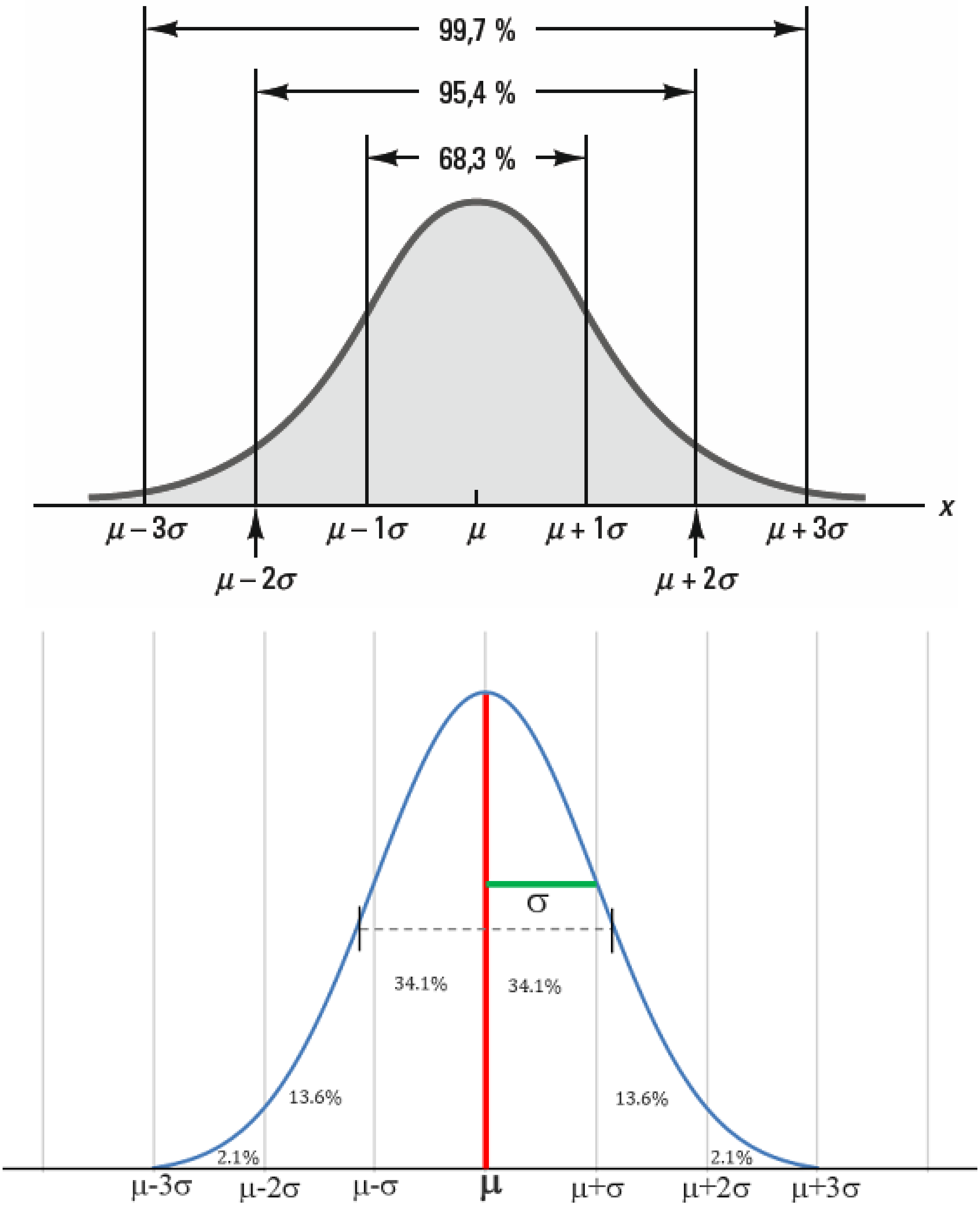
$$= \int_{a}^{b} f(x)dx = \int_{a}^{b} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{(x-\mu)^2}{\sigma^2}\right]} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{a}^{b} e^{\frac{-1}{2}\left[\frac{(x-\mu)^2}{\sigma^2}\right]} dx$$

- لمنحنى التوزيع الطبيعي قمة واحدة، وهذا ما يجعل من المتوسط الحسابي يساوي الوسيط ويساوي المنوال $\overline{x} = Med = Mod$.
- منحنى التوزيع الطبيعي ناقوسي الشكل أي أنه متماثل حول المحور الرأسي الذي يمر بالنقطتين $(\boldsymbol{x.\mu})$، وعندها تقسم المساحة تح المنحنى إلى نصفين متماثلين ومتساويين.
- هناك نسب معينة من المساحة الواقعة ضمن أي عدد من الانحرافات العيارية عن الوسط الحسابي والوسط الحسابي نفسه معرفة كما يلي:

$$p(\mu - \sigma < x < \mu + \sigma) = 68.26\%$$
$$p(\mu - 2\sigma < x < \mu + 2\sigma) = 95.44\%$$
$$p(\mu - 3\sigma < x < \mu + 3\sigma) = 99.74\%$$

ويمكن تمثيل هذه النسب على منحنى التوزيع كما يبينه الشكل التالي:

الشكل رقم 3−1: يوضح تغير شكل دالة كثافة احتمال التوزيع الطبيعي بتغير قيم (μ, σ)



II-التوزيع الطبيعي المعياري Z:

1. تعريف التوزيع الطبيعي المعياري Z: يطلق مصطلح التوزيع الطبيعي المعياري على المتغير العشوائي المستمر الذي يتبع توزيع طبيعي N بوسط مقداره $\mu = 0$ وتباين $\sigma^2 =$

1، ونكتب $x \sim N(0,1)$، أو $Z(0,1)$، وعليه تصبح دالة كثافة الاحتمال للتوزع الطبيعي في هذه الحالة تأخذ الصيغة الرياضية التالية:

$$\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}[Z^2]}$$

وتستند معالم التوزيعات الطبيعية على جداول التوزيع الطبيعي المعياري (القياسي) $\boldsymbol{Z}$.

2.تحويل المتغير العشوائي الطبيعي $N(\mu,\sigma)$ **إلى متغير عشوائي طبيعي معياري** $Z(0,1)$: إذا كان لدينا متغير عشوائي $\boldsymbol{X}$ يتبع التوزيع الطبيعي أي $X \sim N(\mu,\sigma)$ فإنه بالإمكان تحويله إلى متغير عشوائي Z إذا ما عوضت قيم X بقيم Z حيث:

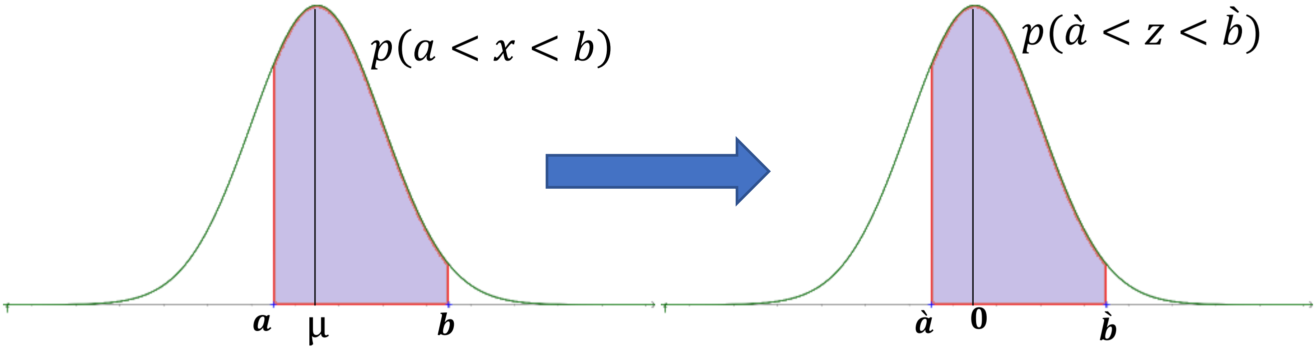
$$Z = \frac{X-\mu}{\sigma}$$

ومن خلال تطبيق العلاقة السابقة يمكن تحويل أي متغير عشوائي يتبع التوزع الطبيعي $\boldsymbol{N}$ $x \sim N(\mu,\sigma)$ إلى متغير عشوائي يتبع التوزيع الطبيعي المعياري $\boldsymbol{Z}$، وبتعويض قيم $\boldsymbol{X}$ بقيم $\boldsymbol{Z}$ فإن معلمات التوزيع الطبيعي تصبح تحقق شروط التوزيع الطبيعي المعياري أي $\mu = 0$ وتباين $\sigma^2 = 1$، **وعندئذ نكتب**: $Z \sim N(0,1)$،

وهناك جداول إحصائية تعطينا قيم $\phi(Z)$ لمختلف قيم X، لذلك عند حساب أي احتمال للمتغير X يجب تحويله أولا إلى الصيغة المعيارية Z ثم بعد ذلك نستعمل الجداول المعيارية لقياس قيمة الاحتمال المطلوب.

فإذا أردنا حساب احتمال أن يقع المتغير X بين قيمتين أي $p(a < x < b)$ فإننا نقوم بتحويل قيم a و b إلى قيم معيارية باستخدام نفس المعادلة السابقة:

$$p(a < x < b) = p\left(\frac{a-\mu}{\sigma} < z < \frac{b-\mu}{\sigma}\right) = p(\grave{a} < z < \grave{b})$$



وهناك عدة أنواع من الجداول للتوزيع الطبيعي المعياري Z تختلف من حيث كيفية الحساب، إلا أنها متماثلة في النتائج، وسوف نعتمد في هذه الدروس على الجداول التي تقيس المساحة بين الصفر وقيمة z المعيارية الموجبة، أي $p(0 < Z < z)$.

مثال 1: إذا كان x متغير عشوائي يتبع التوزيع الطبيعي بمتوسط 3 وتباين 4، فما هو احتمال أن يقع المتغير x بين القيمتين 3 و 5؟

الحل:

إذا كان تباين المتغير العشوائي X هو 4 أي $\sigma^2 = 4$ فإن الانحراف هو جذر التباين وبالتالي $\sigma = 2$، وعليه يمكن أن نكتب $x \sim N(3,2)$ ولحساب قيمة الاحتمال $p(3 < x < 5)$ يجب أن نحول التوزيع الطبيعي على توزيع طبيعي معياري، وعليه نكتب:

$$Z = \frac{X - \mu}{\sigma} \Longrightarrow Z_1 = \frac{3 - 3}{2} = 0$$

$$Z_2 = \frac{5 - 3}{2} = 1$$

$$p(3 < x < 5) = p(0 < z < 1)$$

وبالاعتماد على جداول التوزيع الطبيعي لمعياري Z نجد ان قيمة الاحتمال تساوي **0.3413** أي أن:

$$p(3 < x < 5) = p(0 < z < 1) = 0.3413$$

مثال 2: إذا علمنا أن دخل مجموعة من العائلات يتبع التوزيع الطبيعي بوسط حسابي مقداره 16000دج وانحراف معياري بقيمة 2000دج، فإذا اختيرت عائلة بطريقة عشوائية فما هو احتمال أن يكون دخلها محصور بين 15000دج و18000دج، وأقل من 15000دج، وأكبر من 18000دج.

الحل:

- **حساب احتمال ان يكون دخل العائلة بين 15000دج و18000دج**

$$p(a < x < b) = p(15000 < x < 18000)$$

تحويل قيم كل من a و b إلى القيم المعيارية

$$p(a < x < b) = p\left(\frac{15000 - \mu}{\sigma} < z < \frac{18000 - \mu}{\sigma}\right)$$

$$= p\left(\frac{15000 - 16000}{2000} < z < \frac{18000 - 16000}{2000}\right) = p(-0.5 < Z < 1)$$

$$p(-0.5 < Z < 1) = p(-0.5 < Z < 0) + p(0 < Z < 1)$$

بما أن توزيع الطبيعي المعياري هو توزيع متناظر فإن:

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

$$p(-0.5 < Z < 1) = p(0 < Z < 0.5) + p(0 < Z < 1)$$

من خلال الجداول الإحصائية لتوزيع Z فإن قيم الاحتمال تساوي:

$$p(15000 < x < 18000) = p(-0.5 < Z < 1) = 0.3413 + 0.1915$$

$$p(15000 < x < 18000) = p(-0.5 < Z < 1) = 0.5328$$

- **حساب احتمال ان يكون دخل العائلة أقل من 15000دج**

$$p(x < a) = p(x < 15000) = p\left(z < \frac{15000 - \mu}{\sigma}\right)$$

$$p(x < 15000) = p\left(z < \frac{15000 - 16000}{2000}\right) = p(z < -0.05)$$

بما أن التوزيع الطبيعي المعياري توزيع متناظر فإن:

$$p(z < -0.05) = p(z > 0.05)$$

$$p(z > 0.05) = p(0 < Z < +\infty) - p(0 < Z < 0.5)$$

$$p(z > 0.05) = 0.5 - p(0 < Z < 0.5)$$

من خلال جداول التوزيع الطبيعي المعياري Z فإن قيمة:

$$p(0 < Z < 0.5) = 0.1915$$

$$p(z > 0.05) = 0.5 - 0.1915$$

$$p(x < 15000) = p(z > 0.05) = 0.3085$$

- **حساب احتمال ان يكون دخل العائلة أكبر من 18000دج**

$$p(x > a) = p(x > 18000) = p\left(z > \frac{18000 - 16000}{2000}\right)$$

$$p(x > 18000) = p(z > 1) = p(0 < Z < +\infty) - p(0 < Z < 1)$$

$$p(z > 1) = 0.5 - 0.3413$$

$$p(x > 18000) = p(z > 1) = 0.1587$$

السلسلة رقم (2)

تمرين رقم1

أ- إذا كان المتغير العشوائي المستمر X له توزيع طبيعي $N\ (u, \sigma^2)$، أوجد ما يلي:

1. $P(u - \sigma < X < u + \sigma)$
2. $P(u - 2\sigma < X < u + 2\sigma)$
3. $P(u - 3\sigma < X < u + 3\sigma)$

ب- نفرض أن Z متغير عشوائي مستمر يتبع توزيع طبيعي قياسي $N(0,1)$ احسب ما يلي:

1.$p(0< Z < 1.42)$ / 2.$p(-0.73< Z < 0)$ / 3.$p(-1.79<Z< -0.54)$/ 4.$p(Z> 1.13)$/ 5.$p(-1.37< Z < 2.01)$ / 6.$p(0.65< Z < 1.26)$ / 7.$p(-0.5<Z< 0.5)$

تمرين رقم 2: إذا كان Z متغير عشوائي مستمر يتبع توزيع طبيعي قياسي $N(0.1)$، أوجد قيمة a و b في الحالات التالية:

1.$p(0< Z < a)=0.1915$/ 2.$p(Z < a) = 0.9$/ 3.$p(Z < a) = 0.05$/
4.$p(Z > a) = 0.95$/ 5.$p(a< Z < b) = 0.9974$

علما أن a و b في الحالة رقم 5 نقطتين متماثلتين على منحنى التوزيع الطبيعي القياسي.

تمرين رقم 3: إذا كان سن السكان في مدينة ما يتبع توزيع طبيعي بوسط مقداره 32 سنة وانحراف معياري 18 سنة. فإذا اختير شخص بطريقة عشوائية من هذه المدينة، فما هو احتمال أن يكون هذا الشخص:

1. عمره بين 32 و 59 سنة.
2. عمره أكبر من 50 سنة.
3. عمره أقل من 18 سنة.
4. عمره بين 20 و 40 سنة.
5. إذا أراد صاحب قاعة السينما منح تخفيض مقداره 5% على سعر التذكرة للمسنين في هذه المدينة، فما هو السن الذي سيمنح له هذا التخفيض؟
6. إذا كان هذا التخفيض سيمس 15% من المسنين، فما هو السن الموافق لذلك؟

تمرين رقم 4: نفرض أن أوزان 800 طالب تتبع توزيع طبيعي بوسط مقداره 66 كلغ وانحراف معياري 5 كلغ.أوجد عدد الطلبة الذين يكون وزنهم:

1. بين 65 و 70 كلغ.
2. أكبر من 72 كلغ.

المحور الرابع:
نظرية العينات
وتوزيعات المعاينة

تمهيد:

يقصد عادة بنظرية العينة استنباط العلاقات الموجودة بين المجتمع الاحصائي وبين العينات المسحوبة من هذا المجتمع، وهي علاقات ذات أهمية أساسية في تقدير القيم التي اميز المجتمع مثل المتوسط الحسابي، الوسيط، النسبة والتباين ...الخ، كما تهتم نظرية العينة بدراسة الفوارق الموجودة بين العينات وتحاول معرفة هل هذه الفوارق تعود إلى مجرد صدفة أو هي فوراق ذات دلالة إحصائية؟ وبعبارة موجزة عادة ما يقصد بنظرية العينة مفهوم الاستدلال الاحصائي الذي يهتم بكيفية استخلاص نتائج العينة ومحاولة تعميمها على المجتمع بعد التأكد من صحة هذه النتائج وإمكانية تعميمها. كما سنتطرق في هذا السياق إلى كيفيات سحب العينات العشوائية، وحجم العينة، وكيفية حساب مختلف خصائص توزيع المعاينة.

I. المعاينة (نظرية العينات)

المعاينة هي عملية اختيار العينة، أي اختيار جزء من المجتمع الاحصائي محل الدراسة، للاستدلال على خواص المجتمع بأكمله عن طريق تعميم نتائج العينة بشرط ان تكون هذه العينة ممثلة للمجتمع أحسن تمثيل، وهناك العديد من الطرق لسحب عينة ما من مجتمع معين، هذه الطرق يمكن تصنيفها على أساس عامل العشوائية في السحب إلى صنفين رئيسيين:

الصنف الأول: يتضمن العينات غير العشوائية والتي يكون فيها اختيار قصدي لوحدات العينات (عدم تكافؤ الفرص) في سحب مفردات العينة وتسمى هذه الطريقة بالمعاينة القصدية (العمدية).

الصنف الثاني: يشمل العينات العشوائية التي يعتمد الباحث في اختيار وحداتها على نظرية الاحتمالات وتسمى بطريقة المعاينة العشوائية.

1. المعاينة غير العشوائية: هي تلك العينات التي لا تكفل لجميع المفردات في المجتمع نفس احتمال الاختيار، وغالبا ما يتدخل الباحث في عملية الاخيار، ما يجعل عملية اختيار مفردات العينات في هذه الحالة تتم بطريقة غير احتمالية، إذ يلجأ الباحث في هذه الحالة إلى اختيار وحدات محددة لإدراجها في العينة على اعتبار أنها -**حسب رأي الباحث**- الاختيار الأمثل والعقلاني الذي يمثل المجتمع المدروس تمثيلا جيدا، ما يجعل العينة المختارة تتمتع بالقدرة على عكس خصائص المجتمع بصورة جيدة. ومن أهم أنواع اختيار العينات غير العشوائية: المعاينة العمدية او القصدية، العينة الحصصية، العينة المختارة بواسطة الخبراء، العينة الطبقية العمدية، العينة الحكمية أو التقديرية، ...الخ. ويكثر استخدام هذه العينات من قبل مراكز البحوث التي تقوم بإجراء دراسات استطلاع

الرأي، أو أغراض الدعاية. كما يتطلب هذا النوع من العينات أن يكون حجم المجتمع المدروس محدود وبالتالي حجم العينة صغير. ويعيب هذه العينات عدم القدرة على التحكم مسبقا في تحديد وحساب الأخطاء التي يمكن أن تنتج عنها وهذا راجع إلى أن النظريات والمعادلات الرياضية المبنية عليها هذه العينات لم تتبلور بعد وعليه يندر استخدامها في الدراسات التي تتسم بالطابع العلمي الدقيق.

2. المعاينة العشوائية: هي العينات التي يتم اختيار مفرداتها باستخدام الاختيار العشوائي، بحيث تكون لجميع المفردات في المجتمع نفس فرصة الظهور في العينة المختارة، وهذا بهدف أن تكون العينة المسحوبة تتمتع بالقدرة على عكس خصائص المجتمع بصورة جيدة، وتتعدد طرق وأساليب السحب في المعاينة العشوائية نذكرها فيما يلي:

أ- **المعاينة العشوائية البسيطة**: تعتبر من أبسط الطرق في اختيار العينات، وتستخدم عندما يكون المجتمع متجانسا من حيث الغرض أو الصفة التي تتعلق بها الدراسة، حيث يتم اختيار المفردات بطريقة فردية مباشرة من خلال عملية عشوائية (بإرجاع أو بدون إرجاع) عن طريق الخلط أو القرعة إذا كان حجم المجتمع صغيرا نسبيا، كما يمكن استخدام جداول الأرقام العشوائية أو حتى الاستعانة ببرامج الحاسوب إذا كان حجم المجتمع كبير.

ب- **المعاينة العشوائية المنتظمة**: يتم في هذه الحالة سحب العينة بطريقة من منتظمة من قائمة مرتبة، عن طريق تحديد فترات منتظمة للاختيار، وتنطلق عملية اختيار العينة بتقسيم المجتمع إلى فئات متساوية العدد من خلال حساب نسبة المجتمع إلى العينة $K = \frac{N}{n}$، ثم يتم اختيار رقم بشكل عشوائي بين القيمتين 1 وK، ليكون رقم الوحدة الأولى ثم نقوم بإضافة القيمة K ومضاعفتها إلى ان نصل إلى حجم العينة المرغوب.

مثال: إذا كان حجم المجتمع $N=1000$، ونريد سحب عينة حجمها $n=50$، فإن نسبة المعاينة هي $K=1000/50=20$، ثم نقوم باختيار رقم من المجال $[1\text{-}20]$ بشكل عشوائي، وليكن مثلا 15، ثم نقوم بعد ذلك بتقسيم حجم المجتمع إلى مجموعات عددها يساوي حجم العينة المطلوبة n، وداخل كل مجموعة عدد من المفردات يساوي $K=20$، لنختار من كل مجموعة المفردة رقم 15، فإذا كانت المجموعات مرتبة ترتيبا تصاعديا فإن أو مفردة تكون رقم 15 ثم 35 ثم 55، 75، 95، 105، 125، وهكذا حتى نصل إلى حجم العينة المطلوب وهو 50 مفردة.

ج- **العينة العشوائية الطبقية**: عندما يكون المجتمع الاحصائي غير متجانس، تصبح العينة العشوائية البسيطة غير مناسبة للاستخدام باعتبارها لا تكون ممثلة للمجتمع الذي تسحب منه، لذا يتطلب الامر اللجوء الى

العينة العشوائية الطبقية التي تعتبر إحدى بدائل العينة البسيطة والمنتظمة، إذ يلجأ إليها لزيادة دقة النتائج في حالة المجتمعات التي تتميز بتباين نوعيات المفردات وعدم تجانسها، حيث نقوم بتقسيم المجتمع إلى طبقات لكل طبقة خصائص ومميزات معينة مثل (العمر، الجنس، اللون، المنطقة الجغرافية...الخ)، ثم نقوم بتحديد عدد أفراد كل طبقة، بعد ذلك نحدد نسبة المفردات لكل طبقة (مجموعة) بالنسبة للعدد الكلي لمفردات مجتمع الدراسة، وعند اختيار المفردات من كل طبقة نستخدم ألوب المعاينة العشوائية البسيطة أو المنتظمة. تتلخص اختيار وحداتها بما يلي:

د- **العينة العشوائية العنقودية**: تستخدم هذه الطريقة عندما تكون وحدات المجتمع متباعدة في مناطق جغرافية واسعة، فنقوم بتقسيم المجتمع الأم إلى وحدات أولية ثم يقسم كل منها إلى عدد من الوحدات الفرعية، بعد ذلك تقسم هذه الفروع إلى وحدات ثانوية أصغر وهكذا، إلى أن نصل إلى المجموعات الفرعية النهائية لتختار من كل مجموعة نهائية عينة جزئية أو مفردة باستخدام المعاينة العشوائية البسيطة، وبذلك يضمن في أغلب الأحوال سحب عينة ممثلة لكل فئات المجتمع وتحمل جميع خصائصه وتكون ممثلة له أحسن تمثيل.

II- عدد العينات المسحوبة بإرجاع وبدون إرجاع:

إذا كان لدينا مجتمع إحصائي مؤلف من N وحدةّ، ونريد أن نسحب منه مجموعة عينات، حجم كل عينة يساوي n، فإن عدد العينات العشوائية التي يمكن سحبها نرمز له بالرمز m، ويتغير مقدار m بحسب حجم المجتمع هل هو محدود او غير محدود، إذ أن حالة المجتمعات غير محدودة* تكون المعاينة فيها دائما معاينة غير نفادية، أما المجتمعات المحدودة فتختلف عدد العينات الممكن سحبها حسب طريقة السحب المستخدمة في المعاينة، إذا كانت بإرجاع أو دون إرجاع كما هو مبين أدناه:

1. المعاينة باستخدام طريقة السحب بإرجاع: في هذه الحالة يمكن أن تظهر المفردة أكثر من مرة في عملية اختيار مفردات العينة الواحدة، وتسمى المعاينة في هذه الحالة معاينة غير نفادية، ويعتبر المجتمع في هذه الحالة حتى وإن كان محدودا مجتمعا غير محدود، لأن تكرار عملية السحب لا يؤدي إلى تقليص عدد المفردات في المجتمع، ويتم حساب عدد العينات الممكن سحبها في هذه الحالة باستخدام العلاقة التالية وتسمى علاقة القائمة:

$$m = N^n$$

---

* **يعد المجتمع الاحصائي غير محدود إذا كان عدد مفرداته لا نهائي، أو كان حجم المجتمع كبير جدا مقارنة بالنسبة لحجم العينة المسحوبة أي $\frac{n}{N} \leq 5\%$، كما يعتبر المجتمع الاحصائي غير محدود في حالة العينة المسحوبة بإرجاع حتى وإن كان $\frac{n}{N} > 5\%$.**

2. **المعاينة باستخدام طريقة السحب بدون إرجاع:** في هذه الحالة لا يمكن أن تظهر المفردة الواحدة أكثر من مرة في العينة المسحوبة، لذلك تسمى المعاينة هذه الحالة بالمعاينة النفادية، لأن تكرار عملية السحب يؤدي إلى تقليص عدد المفردات في المجتمع، ولمعرفة عدد العينات الممكن سحبها في هذه الحالة يمكن استخدام علاقة التوفيقة C_N^n ذا كان السحب ودون إرجاع وفي آن واحد، كما هو موضح فيما يلي:

- **التوفيقة:** هي عدد التشكيلات الممكنه لانتقاء مجموعة جزئية من مجموعة كلية من العناصر عندما يكون ليس هناك أهمية للترتيب. ونسمي توفيقة n حيث $(0 \leq n \leq N)$ هي عبارة عن عدد الطرق التي يمكن فيها انتقاء n عنصر من مجموعة ذات N عنصر دون مراعاة لترتيب تسلسل العناصر المنتقاة ضمن التشكيلات الممكنة للمجموعة الجزئية، ويمكن كتابة عدد التوفيقات C_N^n بالعلاقة التالية:

$$m = C_N^n = \frac{N\,!}{n\,!\,(N-n)\,!}$$

III- توزيع المعاينة:

إذا كان لدينا مجتمع إحصائي محدود حجمه N، وتم سحب مجموعة من العينات حجم كل عينة هو n، ستكون لكل عينة من هذه العينات وسطها الحسابي وتباينها أو انحرافها المعياري قيمة النسبة فيها...الخ، وبذلك نحصل على مجموعة من المتوسطات الحسابية والتباينات والنسب، المحسوبة للعينات العشوائية المسحوبة. وطالما تم سحب العينات بشكل عشوائي ستشكل الإحصاءات المحسوبة للعينات بدورها متغيرات عشوائية لها توزيعات احتمالية تعرف بتوزيعات المعاينة.

إذا فإذا كان لدينا مجتمع إحصائي محدود حجمه N، وتم سحب مجموعة من العينات عددها m حجم كل عينة هو n، وقمنا بحساب المتوسط لكل عينة من العينات الممكنة فإنه سيكون لدينا مجموعة من الأوساط الحسابية عددها m مطابق لعدد العينات الموجودة، ليتشكل بذلك توزيعا تكراريا يطلق عليه توزيع المعاينة للأوساط الحسابية، وإن دراسة التوزيع الاحتمالي لمتوسطات مختلف العينات المسحوبة يطلق عليه توزيع المعاينة للمتوسط الذي هو عبارة عن البحث عن القيمة المتوقعة لهذه المتوسطات وانحرافها المعياري وما هي علاقة هذه المعلمات بالمتوسط الحسابي للمجتمع وبتباينه وانحرافه المعياري، وما هو طبيعة توزيعه الاحتمالي هل هو يتبع التوزيع الطبيعي أم يتبع توزيع آخر، وبشكل متشابه نحصل على توزيعات المعاينة للانحرافات المعيارية وللنسب المئوية ...الخ.

لذلك تعرف توزيعات المعاينة لأي إحصاءة من إحصاءات العينة بأنها توزيع احتمالي نظري لجميع القيم الممكنة والمحتملة لإحصاءة العينات المسحوبة، والتي نحصل عليها إذا ما تصورنا سحب كل العينات الممكنة من نفس المجتمع بنفس الحجم ونفس الكيفية، كما تعرف توزيعات المعاينة بأنها التوزيع النظري للتكرارات النسبية لأحد المقاييس الإحصائية الوصفية (الإحصاءة) المحسوبة لعدد كبير من العينات العشوائية المتساوية الحجم والمسحوبة بنفس الطريقة من نفس المجتمع.

1. **توزيع المعاينة للوسط الحسابي $\bar{x}$**: رأينا في الفصل التمهيدي أن الوسط الحسابي للمجتمع μ هو مقدار ثابت، أما الوسط الحسابي للعينة $\bar{x}$ فتختلف قيمته من عينة لأخرى مسحوبة من نفس المجتمع، ومن ثم يمكن اعتبار أن الوسط الحسابي متغيرا عشوائيا له توزيع احتمال يسمى بتوزيع المعاينة للوسط الحسابي. ونحصل على توزيع المعاينة بأخذ جميع العينات الممكن سحبها نفس المجتمع N والتي لها نفس الحجم. لذلك يعرف توزيع المعاينة للوسط بأنه التوزيع الاحتمالي لجميع القيم التي يمكن أن يأخذها المتوسط الحسابي $\bar{x}$ للعينات المسحوبة ذات الحجم n،

فإذا تم سحب عينة عشوائية حجمها n وحسبنا متوسطها الحسابي سوف نجده قيمة ثابتة، أما إذا سحبنا عدة عينات بنفس الحجم ومن نفس المجتمع فمن المتوقع أن يأخذ المتوسط الحسابي قيما مختلفة في هذه العينات، أما إذا قمنا بسحب كل العينات الممكنة m والتي لها نفس الحجم n وحسبنا المتوسطات الحسابية لهذه العينات، فإنه يمكن حساب المتوسط الحسابي لهذه المتوسطات (التوقع الرياضي لهذه المتوسطات) نرمز له بالرمز $\mu_{\bar{x}}$، لكما يمكن حساب الانحراف المعياري لهذه المتوسطات ونرمز له بالرمز $\sigma_{\bar{x}}$ حيث:

أ- **متوسط متوسطات العينات (التوقع الرياضي لمتوسطات العينات) $\mu_{\bar{x}}$** :

$$\mu_{\bar{x}} = E(\bar{x}) = E\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \frac{1}{n}\sum E(x_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n*\mu = \mu$$

وهذا يعني أن التوقع الرياضي لمتوسط العينة (الوسط الحسابي للأوساط الحسابية للعينات) مساوي لمتوسط المجتمع سواء كانت المعاينة نفادية أو غير نفادية.

ب-التباين والانحراف المعياري لمتوسطات العينات $\sigma_{\bar{x}}$:

- **الحالة الأولى: المعاينة غير النفادية**: إذا كان المجتمع غير محدود أو مجتمع محدود والسحب بإرجاع أو مجتمع محدود والسحب بدون إرجاع لكن $\frac{n}{N} \leq 5\%$ وهو ما يحقق بالاستقلال الاحصائي، فإن التباين والانحراف المعياري للمتوسطات الحسابية للعينات تعطى بالعلاقة:

$$\sigma_{\bar{x}}{}^2 = VAR(\bar{x}) = VAR\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \frac{1}{n^2}VAR(\sum_{i=1}^{n} x_i)$$

$$\sigma_{\bar{x}}{}^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} VAR(x_i)\right) = \frac{1}{n^2}nVAR(x_i) = \frac{1}{n^2}n = \frac{1}{n} * \sigma_x{}^2 = \frac{\sigma_x{}^2}{n}$$

$$\sigma_{\bar{x}}{}^2 = \frac{\sigma_x{}^2}{n} \Longrightarrow \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

مثال 1: ليكن لدينا المجتمع الإحصائي التالي $\Omega\,\{8,6,5,3,1\}$ أحسب التوقع الرياضي والتباين والانحراف المعياري لتوزيع المعاينة للمتوسط، إذا كانت العينة المسحوبة $n = 2$ وكان السحب بالإرجاع، ثم قارن بين التوقع الرياضي والتباين والانحراف المعياري لتوزيع المعاينة للمتوسط مع متوسط المجتمع وانحرافه المعياري.

الحل: لدينا المجتمع الاحصائي التالي: $\Omega\,\{8,6,5,3,1\}$، أي أن $N = 5$، $n = 2$، السحب بإرجاع.

بما ان السحب بإرجاع فإن عدد العينات m الممكن سحبها من هذا المجتمع هو:

$$m = N^n = 5^2 = 25$$

حساب المتوسط الحسابي (التوقع الرياضي) للمجتمع:

$$\mu = E(X) = \frac{\sum X_i}{N} = \frac{1+3+5+6+8}{5} = \frac{23}{5} = 4.6$$

حساب التباين والانحراف المعياري للمجتمع:

$$\sigma_X{}^2 = VAR_X = \frac{\sum(X_i - \bar{X})^2}{N}$$

$$\sigma_X{}^2 = \frac{(1-4.6)^2 + (3-4.6)^2 + (5-4.6)^2 + (6-4.6)^2 + (8-4.6)^2}{5}$$

$$\sigma_X{}^2 = 5.84 \Longrightarrow \sigma_X = \sqrt{\sigma_X{}^2} = \sqrt{5.84} = 2.41$$

إيجاد العينات الممكن سحبها ومتوسطاتها:

8	6	5	3	1	
(8,1) **4.5**	(6,1) **3.5**	(5,1) **3**	(3,1) **2**	(1,1) **1**	1
(8,3)	(6,3)	(5,3)	(3,3)	(1,3)	3

	2	3	4	4.5	5.5
5	(1,5) **3**	(3,5) **4**	(5,5) **5**	(6,5) **5.5**	(8,5) **6.5**
6	(1,6) **3.5**	(3,6) **4.5**	(5,6) **5.5**	(6,6) **6**	(8,6) **7**
8	(1,8) **4.5**	(3,8) **5.5**	(5,8) **6.5**	(6,8) **7**	(8,8) **8**

حساب المتوسط الحسابي (التوقع الرياضي) لمتوسطات العينة:

$$\mu_{\bar{x}} = E(\bar{x}) = \frac{\sum \bar{x}_i}{m}$$

$$= \frac{1+2+3+3.5+4.5+2+3+4+4.5+5.5+3+4+5+5.5+6.5+3.5+4.5+5.5+6+7+4.5+5.5+6.5+7+8}{25}$$

$$\mu_{\bar{x}} = \frac{115}{25} = 4.6$$

ومنه يمكن القول أن:

$$\mu_{\bar{x}} = \mu = 4.6$$

تؤكد هذه النتيجة المساواة بين التوقع الرياضي للمتوسطات الحسابية للعينات ومتوسط المجتمع في حالة المعاينة غير النفادية.

حساب التباين والانحراف المعياري لمتوسطات العينة:

$$\sigma_{\bar{x}}{}^2 = \frac{\sum(\bar{x}_i - \mu_{\bar{x}})^2}{m}$$

$$\sigma_{\bar{x}}{}^2 = \frac{(1-4.6)^2 + (2-4.6)^2 + (3-4.6)^2 + \cdots \dots + (7-4.6)^2 + (8-4.6)^2}{25}$$

$$\sigma_{\bar{x}}{}^2 = 2.92 \Longrightarrow \sigma_X = \sqrt{\sigma_X{}^2} = \sqrt{2.92} = 1.7$$

إذا ما طبقنا العلاقة الخاصة بتوزيع المعاينة سنجد أن:

$$\sigma_{\bar{x}}{}^2 = \frac{\sigma_x{}^2}{n} = \frac{5.84}{2} = 2.92 \Longrightarrow \sigma_X = \sqrt{\sigma_X{}^2} = \sqrt{2.92} = 1.7$$

تؤكد هذه النتيجة صحة المعادلة التي توضح العلاقة بين تباين المجتمع وتباين المتوسطات الحسابية للعينات المسحوبة في حالة المعاينة غير النفادية.

- **الحالة الثانية**: **المعاينة النفادية**: إذا كان المجتمع محدود والسحب بدون إرجاع و$\frac{n}{N} > 5\%$ (**أي في حالة عدم تحقق الاستقلال الاحصائي**) من الواجب في هذه الحالة ندخل معامل التصحيح $\left(\frac{N-n}{N-1}\right)$، ويعطى التباين والانحراف المعياري للمتوسطات الحسابية للعينات المسحوبة بالعلاقة التالية:

$$\sigma_{\bar{x}}{}^2 = \frac{\sigma^2}{n}\frac{N-n}{N-1} \Longrightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

مثال 2: نفس معطيات ومطلوب المثال السابق لكن السحب في هذه الحالة دون إرجاع وفي آن واحد.

الحل: لدينا المجتمع الاحصائي التالي: $\Omega\ \{8,6,5,3,1\}$، أي أن $N = 5$، $n = 2$، السحب دون إرجاع،

بما ان السحب بإرجاع فإن عدد العينات m الممكن سحبها من هذا المجتمع يعطى بالعلاقة التالية:

$$m = \frac{N!}{n!\,(N-n)!} = \frac{5!}{2!\,(5-2)!} = 10$$

إيجاد العينات الممكن سحبها ومتوسطاتها:

	1	3	5	6	8
1					
3	(1,3) **2**				
5	(1,5) **3**	(3,5) **4**			
6	(1,6) **3.5**	(3,6) **4.5**	(5,6) **5.5**		
8	(1,8) **4.5**	(3,8) **5.5**	(5,8) **6.5**	(6,8) **7**	

حساب المتوسط الحسابي (التوقع الرياضي) لمتوسطات العينة:

$$\mu_{\bar{x}} = E(\bar{x}) = \frac{\sum \bar{x}_i}{m} = \frac{2+3+4+3.5+4.5+5.5+4.5+5.5+6.5+7}{10}$$

$$\mu_{\bar{x}} = \frac{46}{10} = 4.6$$

ومنه يمكن القول في حالة السحب دون إرجاع (معاينة نفادية) تعتبر المساواة بين التوقع الرياضي للمتوسطات الحسابية للعينات ومتوسط المجتمع صحيحة: $\mu_{\bar{x}} = \mu = 4.6$

حساب التباين والانحراف المعياري لمتوسطات العينة:

$$\sigma_{\bar{x}}{}^2 = \frac{\sum(\bar{x}_i - \mu_{\bar{x}})^2}{m}$$

$$\sigma_{\bar{x}}{}^2 = \frac{(2-4.6)^2 + (3-4.6)^2 + (4-4.6)^2 + \cdots \ldots + (6.5-4.6)^2 + (7-4.6)^2}{10}$$

$$\sigma_{\bar{x}}{}^2 = 2.19 \Rightarrow \sigma_X = \sqrt{\sigma_X{}^2} = \sqrt{2.19} = 1.48$$

إذا ما طبقنا العلاقة الخاصة بتوزيع المعاينة في حالة السحب دون إرجاع سنجد أن:

$$\frac{n}{N} = \frac{2}{5} = 0.4 > 0.05$$

وبالتالي نطبق معامل التصحيح في حساب قيمة $\sigma_{\bar{x}}{}^2$

$$\sigma_{\bar{x}}{}^2 = \frac{\sigma_x{}^2}{n} * \frac{N-n}{N-1} = \frac{5.84}{2} * \frac{5-2}{5-1} = 2.19 \Rightarrow \sigma_{\bar{X}} = \sqrt{\sigma_{\bar{X}}{}^2} = \sqrt{2.19} = 1.48$$

تؤكد هذه النتيجة صحة المعادلة التي توضح العلاقة بين تباين المجتمع وتباين المتوسطات الحسابية للعينات المسحوبة في حالة المعاينة النفادية $\frac{n}{N} > 0.05$.

مثال 3: إذا كان لدينا مجتمع حجمه $N = 900$. متوسطه الحسابي $\mu = 20$ وانحرافه المعياري $\sigma = 12$.

فإذا سحبت عينة عشوائية دون إرجاع حجمها $n = 36$.

المطلوب: أحسب المتوسط الحسابي والانحراف المعياري للمتوسط الحسابي لهذه العينة.

أحسب المتوسط الحسابي والانحراف المعياري للمتوسط الحسابي إذا كان حجم العينة $n = 64$.

الحل:

الحالة الأولى: لدينا: $N = 900$، و $n = 36$ والسحب دون إرجاع، وعليه يجب حساب النسبة $\frac{n}{N}$.

$$\frac{n}{N} = \frac{36}{900} = 0.04 < 0.05$$

وبالتالي فإن المعاينة تعتبر معاينة غير نفادية ولا نطبق معامل التصحيح في حساب الانحراف المعياري لتوزيع المعاينة للمتوسط.

- **حساب التوقع الرياضي لمتوسط العينة:**

$$\mu_{\bar{x}} = \mu = 20$$

- **حساب الانحراف المعياري لتوزيع المعاينة لمتوسط العينة:**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = \frac{12}{6} = 2$$

الحالة الثانية: لدينا: $N = 900$، و $n = 64$ والسحب دون إرجاع، وعليه يجب حساب النسبة $\frac{n}{N}$.

$$\frac{n}{N} = \frac{64}{900} = 0.07 > 0.05$$

وبالتالي فإن المعاينة تعتبر معاينة نفادية ونطبق في هذه الحالة معامل التصحيح في حساب الانحراف المعياري لتوزيع المعاينة للمتوسط.

- **حساب التوقع الرياضي لمتوسط العينة:**

$$\mu_{\bar{x}} = \mu = 20$$

حساب الانحراف المعياري لتوزيع المعاينة لمتوسط العينة:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}} = \frac{12}{\sqrt{64}} * \sqrt{\frac{900-64}{900-1}} = 1.45$$

2. طبيعة وشكل توزيع المعاينة لـ $\bar{x}$: إن تحديد شكل توزيع المتوسطات المحسوبة من العينات المسحوبة لب اهتمام موضوع توزيع المعاينة، وسنحاول معرفة أي من التوزيعات الاحتمالية يتبع توزيع المعاينة للمتوسط الحسابي للعينات.

أ- إذا كان المجتمع يتبع توزيع بطبيعي وبتباين σ^2 معلوم:

نظرية: إذا كان لدينا متغير عشوائي X يتبع التوزيع الطبيعي بوسط μ، وانحراف σ **معلوم**، واخترنا منه عينة عشوائية حجمها n، فإن الوسط الحسابي للعينات المسحوبة $\bar{x}$ يخضع كذلك إلى التوزيع الطبيعي، بمتوسط $\mu_{\bar{x}}$ وتباين ${\sigma_{\bar{X}}}^2$ ونكتب:

$$\bar{x} \sim N(\mu_{\bar{x}}, {\sigma_{\bar{x}}}^2) \Longrightarrow \bar{x} \sim N((\mu_{\bar{x}} = \mu, {\sigma_{\bar{x}}}^2 = \frac{\sigma^2}{n}) \Longrightarrow \bar{x} \sim N((\mu, \frac{\sigma^2}{n})$$

وفي حالة المجتمع المحدود والمعاينة النفادية $\left(\frac{n}{N} > 0.05\right)$ نستبدل العبارة ${\sigma_{\bar{x}}}^2 = \frac{\sigma^2}{n}$ بالصيغة:

$${\sigma_{\bar{x}}}^2 = \frac{\sigma^2}{n} * \frac{N-n}{N-1}$$

مثال 1: إذا كانت لدينا عينة عشوائية حجمها 25 تم سحبها من مجتمع كبير يتبع توزيعا طبيعيا بمتوسط 50 وانحراف 15. أحسب توزيع المعاينة للمتوسط الحسابي في هذه الحالة؟

أحسب احتمال أن يكون متوسط العينة أكبر من 55.

الحل: بما أن المجتمع يتبع التوزيع الطبيعي فإن توزيع المعاينة لمتوسط الحسابي $\bar{x}$ يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{x}}$، وتباين ${\sigma_{\bar{x}}}^2$، حيث:

$$\mu_{\bar{x}} = \mu = 50$$

$${\sigma_{\bar{x}}}^2 = \frac{\sigma^2}{n} \Longrightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{25}} = \frac{15}{5} = 3$$

ومنه نقول أن الوسط الحسابي لتوزيع المعاينة يتبع التوزيع الطبيعي بمتوسط 50 وانحراف معياري 3، ونكتب:

$$\bar{x} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \Longrightarrow \bar{x} \sim N((\mu_{\bar{x}} = 50, \sigma_{\bar{x}} = 3) \Longrightarrow \bar{x} \sim N((50,3)$$

- **حساب احتمال أن يكون متوسط العينة أكبر من 55.** $P(\bar{x} > 55)$

لحساب الاحتمال لابد من التحويل من التوزيع الطبيعي N إلى التوزيع الطبيعي المعياري Z، ونكتب:

$$P(\bar{x} > 55) = p\left(Z > \frac{55 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) = p\left(Z > \frac{55 - 50}{3}\right)$$

$$P(\bar{x} > 55) = p(Z > 1.66) = p(1.66 < Z < +\infty)$$

$$P(\bar{x} > 55) = 0.5 - p(0 < Z < 1.66) = 0.5 - 0.4515$$

$$P(\bar{x} > 55) = 0.0485$$

ب- إذا كان المجتمع لا يتبع التوزيع الطبيعي وتباينه σ^2 معلوم (نظرية النهايات المركزية):

في كثير من الحلالات يكون التوزيع الاحتمالي للمجمع الأصلي للدراسة لا يتبع التوزيع الطبيعي، إذا كان لدينا المتغير العشوائي X يتبع توزيع ما (لا يتبع التوزيع الطبيعي) وسطه μ، وانحرافه σ **معلوم**، واخترنا عينة عشوائية حجمها n حيث $(n \geq 30)$، فإن القيمة المعيارية $\mu_{\bar{x}}$ للعينات المسحوبة حسب نظرية لنهايات المركزية تؤول إلى التوزيع الطبيعي بمتوسط $\mu_{\bar{x}}$ وتباين ${\sigma_{\bar{x}}}^2$ لأنه كلما زاد حجم العينة اقترب توزيع متوسطها من التوزيع الطبيعي، بصرف النظر عن طبيعة توزيع المجتمع الأصلي، ونكتب:

$$\bar{x} \sim N(\mu_{\bar{x}}, {\sigma_{\bar{x}}}^2)$$

مثال 1: في أحد اختبارات الذكاء كان متوسط الدرجات هو 1000، والانحراف المعياري هو 125، فإذا أعطى الاختبار لعينة عشوائية من 100 شخص، فما هو احتمال أن قيمة المتوسط الحسابي في هذه العينة أكبر من 970 وأقل من 1030؟

الحل: بما أن المجتمع لا يتبع بالضرورة التوزيع الطبيعي (توزيع غير معلوم) ولدينا أن حجم العينة $(n = 100 > 30)$ فإنه حسب نظرية النهايات المركزية يمكن اعتبار توزيع المعاينة لمتوسط الحسابي $\bar{x}$ يتبع توزيع قريب من التوزيع الطبيعي بمتوسط $\mu_{\bar{x}}$، وتباين ${\sigma_{\bar{x}}}^2$، حيث:

$$\mu_{\bar{x}} = \mu = 1000$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \Longrightarrow \sigma_{\bar{x}} = \frac{125}{\sqrt{100}} = 12.5$$

ومنه نقول أن الوسط الحسابي لتوزيع المعاينة يتبع التوزيع الطبيعي بمتوسط 1000 وانحراف معياري 12.5 أي:

$$\bar{x} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \Longrightarrow \bar{x} \sim N((1000,12.5)$$

حساب احتمال أن يكون متوسط العينة أكبر من 970 وأقل من 1030.

$$P(970 < \bar{x} < 1030) = p\left(\frac{970 - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < Z < \frac{1030 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$
$$P(970 < \bar{x} < 1030) = p\left(\frac{970 - 1000}{12.5} < Z < \frac{1030 - 1000}{12.5}\right)$$
$$P(970 < \bar{x} < 1030) = p(-2.4 < Z < 2.4)$$
$$P(970 < \bar{x} < 1030) = p(-2.4 < Z < 0) + p(0 < Z < 2.4)$$

بما أن التوزيع متناظر بالنسبة للصفر فإن: $p(-2.4 < Z < 0) = p(0 < Z < 2.4)$

$$P(970 < \bar{x} < 1030) = 2p(0 < Z < 2.4)$$
$$P(970 < \bar{x} < 1030) = 2 * 0.4918 = 0.9836$$

مثال 2: يقدر متوسط وصول الطائرات إلى مطار مدينة باتنة أربع طائرات في الساعة، فإذا أخذت عينة من 64 ساعة، فما هو احتمال أن يكون معدل وصول الطائرات إلى هذا المطار يزيد عن **4.2** طائرة في الساعة الواحدة.

الحل: لدينا المتغير العشوائي X يتبع توزيع بواسون (توزيع غير طبيعي)، والذي يعرف بالمعلمة $\lambda = 4$، ونعلم أنه في توزيع بواسون، X هو متغير عشوائي منفصل يشير إلى عدد التكرارات في مجال معين. ومن خصائص توزيع بواسون أن معلمة التوزيع λ تساوي التوقع الرياضي μ والتباين σ^2 أي نكتب:

$$\lambda = \mu = \sigma^2 = 4$$

وبما أن حجم العينة $n = 64 > 30$ فإنه يمكن تقريب توزيع بواسون إلى التوزيع الطبيعي حسب نظرية النهايات المركزية، حيث يصبح الوسط الحسابي $\bar{x}$ في العينة n يقترب من التوزيع الطبيعي بالمعلمات:

$$\bar{x} \sim N\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right) \Longrightarrow \bar{x} \sim N\left(4, \frac{2}{\sqrt{64}}\right) \Longrightarrow \bar{x} \sim N\left(4, \frac{2}{8}\right)$$
$$\Longrightarrow \bar{x} \sim N(4,0.25)$$

حساب احتمال أن يكون متوسط وصول الطائرات المطار يزيد عن 4.2 طائرة في الساعة الواحدة.

$$P(\bar{x} > 4.2) = p\left(Z > \frac{4.2 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) = p\left(Z > \frac{4.2 - 4}{0.25}\right) = p(Z > 0.8)$$
$$P(\bar{x} > 4.2) = p(Z > 0.8) = p(0.8 < Z < +\infty) = 0.5 - p(0 < Z < 0.8)$$
$$P(\bar{x} > 4.2) = p(Z > 0.8) = 0.5 - 0.2881 = 0.2119$$

ج- إذا كان تباين المجتمع σ^2 غير معلوم:

يحدث في كثير من الحالات العملية أن يكون تباين المجتمع غير معلوم، ولحساب تباين توزيع متوسطات العينة نستخدم تباين العينة S^2 بدلا من تباين المجتمع σ^2، وبالتالي فإن الانحراف المعياري للأوساط الحسابية يرمز له في هذه الحالة $S_{\bar{x}} = \frac{S}{\sqrt{n}}$ بدلا من $\sigma_{\bar{x}}$، ويتبع توزيع المعاينة للمتوسط الحسابي للعينات في هذه الحالة التوزيع الطبيعي إذا كان حجم العينة أكبر من 30، بينما يتبع توزيع ستودنت إذا كان حجم العينة أقل من 30.

- **الحالة الأولى تباين مجهول و$n \geq 30$**: في هذه الحالة يتبع توزيع المعاينة للمتوسط الحسابي التوزيع الطبيعي مع استبدال تباين انحراف المجتمع σ بانحراف العينة S، ونكتب:

$$\bar{x} \sim N\left(\mu_{\bar{x}}, \frac{S}{\sqrt{n}}\right) \Rightarrow Z = \left(\frac{\bar{x} - \mu_{\bar{x}}}{\frac{S}{\sqrt{n}}}\right) \Rightarrow Z \sim N(0,1)$$

- **الحالة الثانية تباين مجهول و$n < 30$**: إذا أخذت عينة عشوائية من مجتمع يتبع التوزيع الطبيعي متوسطه μ وتباينه σ^2 غير معلوم، وكان $\bar{x}$ المتوسط الحسابي للعينة لعينة حجمها $n < 30$ وانحرافها المعياري S فإن القيمة $T = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$ تخضع لتوزيع **ستودنت** $Student\ \ t$ بدرجة حرية $df = v = (n-1)$، حيث:

$$T = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

حيث: T متغير عشوائي لا تخضع للتوزيع الطبيعي بل تخضع لتوزيع ستودنت $\boldsymbol{Student}$ بدرجة حرية $df = v = (n-1)$، وما يميز توزيع t أنه يشبه التوزيع الطبيعي المعياري من حيث أنه متماثل حول مركزه الصفر، لكنه أكثر تشتتا وتباينا من التوزيع الطبيعي المعياري Z.

ولاستخدام توزيع ستودنت هناك ثلاث شروط يجب تحققها:

- أن يكون المجتمع الأصلي الذي سحبت منه العينة يتوزع توزيعا طبيعيا.
- أن يكون تباين المجتمع غير معلوم.
- أن يكون حجم العينة صغير $n < 30$.

مثال: إذا كان الزمن اللازم لشحن عربة في مصنع معين هو أربعين دقيقة، واقترح صاحب المصنع إجراءات جديدة لعملية الشحن، فإذا طبقة هذه الإجراءات على 25 عربة اختيرت بطريقة عشوائية، وكانت النتيجة أن متوسط زمن الشحن في العينة هو 37.5 دقيقة، بانحراف 4.5 دقيقة، فغذا علمت أن توزيع أزمنة الشحن قريبة من التوزيع الطبيعي. أوجد احتمال أن يكون متوسط زمن الشحن ل 25 عربة هو اقل من 37.5 دقيقة.

الحل:

حساب احتمال أن يكون متوسط زمن الشحن في العينة أقل من 37.5 دقيقة.

$$p(\bar{x} < 37.5)$$

بما أن توزيع المجتمع الاحصائي يتبع التوزيع الطبيعي، وانحراف المجتمع غير معلوم، وحجم العينة $n = 25 < 30$ فإن توزيع المعاينة للمتوسط الحسابي للعينة لا يتبع التوزيع الطبيعي بل يتبع توزيع **ستودنت $\boldsymbol{t}$** بدرجة حرية:

$$dl = v = (n-1) = (25-1) = 24$$

ونكتب:

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

$$p(\bar{x} < 37.5) = p\left(t < \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}\right) = p\left(t < \frac{37.5 - 40}{\frac{4.5}{\sqrt{25}}}\right) = p(t < -2.78)$$

من خلال جداول t ستودنت ومن السطر الخاص بدرجة حرية تساوي 24 نبحث عن القيمة 2.78 فنجد:

$$p(\bar{x} < 37.5) = p(t < -2.78) \cong 0.005$$

3. توزيع المعاينة للنسبة p: عادة ما يكون من الاهمية معرفة نسبة مفردات المجتمع الاحصائي التي تتصف بخاصية معينة، مثل نسب النجاح، نسبة المعيب في منتج معين، نسبة الإناث في صف معين، ...الخ. فإذا كان لدينا مجتمع إحصائي N يتبع توزيع ذي الحدين بالمعلمتين $\mu = n * p$، والتباين $\sigma^2 = n * p * q$.

حيث $q = 1 - p$ ، و $p = \frac{X}{N}$ هي نسبة النجاح (نسبة المفردات التي تتوفر فيها خاصية أو صفة معينة من مجموع مفردات المجتمع)، وتحسب قيمة النسبة p بقسمة مجموع المفردات ذات الخاصية X على حجم المجتمع N.

فإذا سحبت عينة بطريقة عشوائية من هذا المجتمع فإن نسبة هذه الخاصية في العينة هي $\grave{p} = \frac{x}{n}$. وسنهتم في هذا العنصر بطبيعة توزيع النسبة $\grave{p}$ التي تمثل نسبة خاصية ما في العينات المسحوبة من المجتمع، كما سنتعرف على كيفية حساب التوقع الرياضي للنسبة والتباين لتوزيع المعاينة لهذه الإحصاءة $\grave{p}$ في العينة.

أ- **التوقع الرياضي والانحراف المعياري لتوزيع المعاينة للنسبة $\acute{p}$**: إذا أخذت عينة عشوائية حجمها n من مجتمع إحصائي يتبع توزيع ذي الحدين ذو النسبة p فإن التوقع الرياضي لتوزيع المعاينة للنسبة في كل العينات الممكن سحبها يرمز له بالرمز $\mu_{\acute{p}}$، أما الانحراف المعياري فيرمز له بالرمز $\sigma_{\acute{p}}$ حيث:

$$\boldsymbol{\mu_{\acute{p}} = E(\acute{p}) = E\left(\frac{x}{n}\right) = \frac{1}{n}E(x) = \frac{1}{n}n * p = p}$$

وهذا يعني أن التوقع الرياضي للنسبة في العينة تساوي إلى قيمة النسبة في المجتمع

$$\sigma_{\acute{p}} = \sqrt{\sigma_{\acute{p}}{}^2}$$

$$\sigma_{\acute{p}}{}^2 = \left(\sigma_{\frac{x}{n}}\right)^2 = Var\left(\frac{x}{n}\right) = \frac{1}{n^2}Var(x) = \frac{1}{n^2}n * p * q = \frac{1}{n}p * q$$

$$\sigma_{\acute{p}}{}^2 = \frac{p * q}{n} \Longrightarrow \sigma_{\acute{p}} = \sqrt{\frac{p * q}{n}} = \sqrt{\frac{p * (1-p)}{n}}$$

ملاحظة: في حالة المعاينة النفادية (المجتمع محدود والسحب دون إرجاع مع $\left(\frac{n}{N} > 0.05\right)$ فإنه لابد من الأخذ بعين الاعتبار معامل التصحيح في حساب تباين أو انحراف توزيع المعاينة للنسبة، حيث تصبح العلاقة السابقة بالشكل الموالي:

$$\sigma_{\acute{p}}{}^2 = \frac{p * q}{n} * \frac{N-n}{N-1} \Longrightarrow \sigma_{\acute{p}} = \sqrt{\frac{p * q}{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$\boldsymbol{\sigma_{\acute{p}} = \sqrt{\frac{p * (1-p)}{n} * \frac{N-n}{N-1}}}$$

ب- طبيعة توزيع المعاينة للنسبة $\acute{p}$: إذا كان لدينا X متغير عشوائي يتبع توزيع ذي الحدين، حيث p هي نسبة النجاح لخاصية معينة في المجتمع، و$q = 1 - p$، فإذا سحب عينة عشوائية من هذا المجتمع ذات الحجم

n فإن توزيع المعاينة للخاصية $\acute{p}$ في العينة يعتبر متغير عشوائي بمتوسط $\boldsymbol{\mu_{\acute{p}}}$ وانحراف $\boldsymbol{\sigma_{\acute{p}}}$، ويتبع التوزيع الطبيعي إذا توفرت الشروط الثلاثة التالية:

$$n \geq 30 \qquad n * p \geq 5 \qquad n * (1-p) \geq 5$$

ونكتب:

$$si: \{n > 30\,, \quad np \geq 5, \quad n(1-p) \geq 5\} \Longrightarrow \acute{p} \rightsquigarrow N(\mu_{\acute{p}}, \sigma_{\acute{p}})$$

مثال: إذا كانت نسبة الديمقراطيين في مجتمع الناخبين تساوي 60%، وتم سحب عينة عشوائية حجمها 150 ناخب من هذا المجتمع. فما هو احتمال أن تكون نسبة الديمقراطيين في هذه العينة أكبر من 65%؟

الحل: لدينا X متغير عشوائي يتبع توزيع ذي الحدين بالمعلمة $p = 0.6$، وحجم العينة $n = 150$.

حساب القيمة المتوقعة $\boldsymbol{\mu_{\acute{p}}}$ والانحراف المعياري $\boldsymbol{\sigma_{\acute{p}}}$ للنسبة في العينة:

$$\boldsymbol{\mu_{\acute{p}} = p = 0.6}$$

$$\boldsymbol{\sigma_{\acute{p}} = \sqrt{\frac{p*(1-p)}{n}} = \sqrt{\frac{0.6*(1-0.6)}{150}} = 0.04}$$

طبيعة توزيع الإحصاءة $\boldsymbol{\acute{p}}$: التأكد من تحقق الشروط الثلاثة

$$n = 150 \leq 30$$

$$n * p = 150 * 0.6 = 90 \geq 5$$

$$n * (1-p) = 150 * 0.4 = 60 \geq 5$$

بما أن الشروط الثلاثة محققة فإن توزيع المعاينة للإحصاءة $\acute{p}$ يتبع التوزيع الطبيعي ونكتب:

$$\acute{p} \rightsquigarrow N(\mu_{\acute{p}} = 0.6, \sigma_{\acute{p}} = 0.04) \Longrightarrow \acute{p} \rightsquigarrow N(0.6, 0.04)$$

حساب احتمال أن تكون نسبة الديمقراطيين في هذه العينة أكبر من 65%؟

$$p(\acute{p} > 0.65) = p\left(z > \frac{0.64 - \mu_{\acute{p}}}{\sigma_{\acute{p}}}\right) = p\left(z > \frac{0.64 - 0.6}{0.04}\right)$$

$$p(\acute{p} > 0.65) = p(z > 1.25) = p(1.25 < z < +\infty)$$

$$= p(0 < z < +\infty) - p(0 < z < 1.25) = 0.5 - p(0 < z < 1.25)$$

$$p(\acute{p} > 0.65) = 0.5 - 0.3944 = 0.1056$$

4. توزيع المعاينة لمجموع أو فرق متوسطين $(\bar{x}_1 \pm \bar{x}_2)$: ليكن لدينا مجتمعين إحصائيين N_1 و N_2 بمتوسطين μ_1 و μ_2 لكل منهما تباينه ${\sigma_1}^2$ و ${\sigma_2}^2$، فإذا سحبنا مجموعة من العينات العشوائية من كل مجتمع بصورة متكررة n_1 و n_2، سينتج لنا متغيرين عشوائيين، $\bar{x}_1$ كمتغير عشوائي يمثل متوسطات العينات n_1 المسحوبة من المجتمع N_1، والإحصاءة الثانية هي $\bar{x}_2$ كمتغير عشوائي يمثل متوسطات العينات n_2 المسحوبة من المجتمع N_2، كما أن المجموع أو الفرق بين متوسطي العينات المسحوبة من المجتمعين سيكون بدوره متغير عشوائي هو $(\bar{x}_1 + \bar{x}_2)$ أو $(\bar{x}_1 - \bar{x}_2)$، إذ سنحصل على توزيع لهذا المقدار يسمى توزيع المعاينة لمجموع او فرق متوسطي عينتين، وهو متغير عشوائي له متوسط حسابي هو $\mu_{(\bar{x}_1 \pm \bar{x}_2)}$، وتباين ${\sigma^2}_{(\bar{x}_1 \pm \bar{x}_2)}$ حيث:

أ- **التوقع الرياضي أو متوسط توزيع المعاينة لمجموع أو فرق متوسطين**

$$\mu_{(\bar{x}_1+\bar{x}_2)} = \mu_{\bar{x}_1} + \mu_{\bar{x}_2} = \mu_1 + \mu_2$$

$$\mu_{(\bar{x}_1-\bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$$

ب- التباين والانحراف المعياري لمجموع أو فرق متوسطي عينتين

$$\sigma^2{}_{(\bar{x}_1+\bar{x}_2)} = \sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2} = \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}$$

$$\sigma^2{}_{(\bar{x}_1-\bar{x}_2)} = \sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2} = \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}$$

$$\sigma_{(\bar{x}_1\pm\bar{x}_2)} = \sqrt{\sigma^2{}_{(\bar{x}_1\pm\bar{x}_2)}} = \sqrt{\sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2}} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$$

ملاحظة: في حالة المعاينة النفادية (مجتمعين محدودين والسحب دون إرجاع مع $\left(\frac{n_i}{N_i} > 0.05\right)$ فإنه لابد من الأخذ بعين الاعتبار معامل التصحيح في حساب التباين والانحراف المعياري لمجموع أو فرق متوسطي عينتين، حيث تصبح العلاقة السابقة بالشكل الموالي:

$$\sigma_{(\bar{x}_1\pm\bar{x}_2)} = \sqrt{\sigma^2{}_{(\bar{x}_1\pm\bar{x}_2)}} = \sqrt{\sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2}} = \sqrt{\frac{{\sigma_1}^2}{n_1} * \frac{N-n}{N-1} + \frac{{\sigma_2}^2}{n_2}\frac{N-n}{N-1}}$$

ج- طبيعة توزيع المعاينة لمجموع أو فرق متوسطين

- **الحالة الأولى: إذا كان المجتمعين يتوزعان توزيعا طبيعيا وتباين كل منهما معلوم**

نظرية 1: إذا كان لدينا العينة العشوائية $(x_1,\ x_2,\ x_3, \dots \dots\ ,\ x_n)$ مسحوبة من مجتمع N_1 له توزيع طبيعي بمتوسط μ_1 وتباين معلوم $\sigma_1{}^2$، وكانت لدينا عينة عشوائية ثانية $(x_1,\ x_2,\ x_3, \dots \dots\ ,\ x_n)$ مستقلة عن العينة الأولى ومسحوبة من مجتمع آخر N_2 له توزيع طبيعي بمتوسط μ_2 وتباين معلوم $\sigma_2{}^2$، ورمزنا لمتوسط العينة الأولى بالرمز $\bar{x}_1$، ولمتوسط العينة الثانية بالرمز $\bar{x}_2$، فإن توزيع المعاينة لمجموع أو فرق متوسطي العينات المسحوبة من المجتمعين $(\bar{x}_1 \pm \bar{x}_2)$ يعتبر متغير عشوائي يخضع للتوزيع الطبيعي ونكتب:

$$(\bar{x}_1 \pm \bar{x}_2) \backsim N(\mu_{(\bar{x}_1 \mp \bar{x}_2)}, \sigma^2{}_{(\bar{x}_1 \mp \bar{x}_2)})$$

مثال: ينتج مصنع A محركات سيارات لها متوسط عمر 3.5 سنة بانحراف معياري 0.45 سنة، بينما ينتج مصنع B محركات سيارات مماثلة لكن لها متوسط عمر 3.3 سنة وبانحراف معياري 0.3 سنة، فإذا سحب عينة عشوائية حجمها 30 محرك من إنتاج المصنع A، وعينة عشوائية حجمها 36 من إنتاج المصنع B.

المطلوب: فإذا علمت الإنتاج في كل مصنع يتبع التوزيع الطبيعي، فما هو احتمال أن يكون متوسط عمر محركات المصنع A يزيد بقيمة 0.4 سنة عن عمر محركات المصنع B في العينتين المسحوبتين؟

الحل: من خلال معطيات المثال نستخلص ما يلي:

N_1

$$\mu_1 = 3.5 \Longrightarrow \mu_{\bar{x}_1} = \mu_1 = 3.5$$
$$n_1 = 30$$
$$\sigma_1 = 0.45 \Longrightarrow \sigma_1{}^2 = (0.45)^2 = 0.2025$$
$$\sigma_{\bar{x}_1}{}^2 = \frac{\sigma_1{}^2}{n_1} = \frac{0.2025}{30} = 0.00675$$
$$\bar{x}_1 \backsim N(3.5\ , 0.08215)$$

N_2

$$\mu_2 = 3.3 \Longrightarrow \mu_{\bar{x}_2} = \mu_2 = 3.3$$
$$n_2 = 36$$
$$\sigma_2 = 0.3 \Longrightarrow \sigma_2{}^2 = (0.3)^2 = 0.09$$
$$\sigma_{\bar{x}_2}{}^2 = \frac{\sigma_2{}^2}{n_2} = \frac{0.09}{36} = 0.0025$$
$$\bar{x}_2 \backsim N(3.3\ , 0.09)$$

- **حساب توزيع المعاينة للفرق بين متوسطي العينتين** $(\bar{x}_1 - \bar{x}_2)$:

$$\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 3.5 - 3.3 = 0.2$$
$$\sigma^2{}_{(\bar{x}_1 - \bar{x}_2)} = \sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2} = 0.00675 + 0.0025 = 0.00925$$
$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\sigma^2{}_{\bar{x}_1} + \sigma^2{}_{\bar{x}_2}} = \sqrt{0.00925} = 0.09617$$

بما أن العينتين سحبتا من مجتمعين هما المصنع A، والمصنع B واللذان يتبعان التوزيع الطبيعي فإن المتغير العشوائي $(\bar{x}_1 - \bar{x}_2)$ يتبع أيضا التوزيع الطبيعي، ونكتب:

$$(\bar{x}_1 - \bar{x}_2) \sim N(\mu_{(\bar{x}_1-\bar{x}_2)}, \sigma_{(\bar{x}_1-\bar{x}_2)}) \Longrightarrow (\bar{x}_1 - \bar{x}_2) \sim N(0.2\,, 0.09617)$$

- **احتمال أن يكون متوسط عمر المحركات في عينة المصنع A يزيد بقيمة 0.4 سنة عن عمر المحركات في عينة المصنع B:**

$$p[(\bar{x}_1 - \bar{x}_2) > 0.4] = p\left(Z > \frac{0.4 - \mu_{\bar{x}_1-\bar{x}_2}}{\sigma_{\bar{x}_1-\bar{x}_2}}\right) = p\left(Z > \frac{0.4 - 0.2}{0.09617}\right)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 0.4] = p(Z > 2.08) = 0.5 - p(0 < Z < 2.08)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 0.4] = 0.5 - 0.4812 = 0.0188$$

إن احتمال أن يكون متوسط عمر المحركات في عينة المصنع A يزيد بقيمة 0.4 سنة عن عمر المحركات في عينة المصنع B هو $\%1.88$.

- **الحالة الثانية: إذا كان المجتمعين لا يتوزعان توزيعا طبيعيا وتباين كل منهما معلوم**

نظرية 2: إذا كان لدينا العينة العشوائية $(x_1,\ x_2,\ x_3, \dots\dots,\ x_n)$ مسحوبة من مجتمع N_1 **لا يتبع التوزيع الطبيعي** بمتوسط μ_1 وتباين معلوم ${\sigma_1}^2$، وكانت لدينا عينة عشوائية ثانية $(x_1,\ x_2,\ x_3, \dots\dots,\ x_n)$ مستقلة عن العينة الأولى ومسحوبة من مجتمع آخر N_2 **لا يتبع التوزيع الطبيعي** بمتوسط μ_2 وتباين معلوم ${\sigma_2}^2$، ورمزنا لمتوسط العينة الأولى بالرمز $\bar{x}_1$، ولمتوسط العينة الثانية بالرمز $\bar{x}_2$، فإن توزيع المعاينة لمجموع أو فرق متوسطي العينتين $(\bar{x}_1 \pm \bar{x}_2)$ حسب نظرية النهايات المركزية يعتبر متغير عشوائي يخضع لتوزيع يقترب من التوزيع الطبيعي **إذا كان حجم العينتين اكبر أو يساوي ثلاثين** أي $n_1 \geq 30\,, n_2 \geq 30$ ونكتب:

$$(\bar{x}_1 \pm \bar{x}_2) \sim N(\mu_{(\bar{x}_1 \mp \bar{x}_2)}, {\sigma^2}_{(\bar{x}_1 \mp \bar{x}_2)})$$

مثال 2: إذا كان متوسط الدخل الشهري للعائلات في المدينة A هو 16500دج بانحراف معياري 2500 دج، بينما متوسط الدخل الشهري للعائلات في المدينة B هو 15700 دج وبانحراف معياري 2000 دج، فإذا سحب عينة عشوائها حجمها 125 عائلة من المدينة A، وعينة عشوائها أخرى مستقلة عن الأولى حجمها 100 عائلة من المدينة B.

- ما هو احتمال أن يكون متوسط الدخل في العينة A أكبر من متوسط الدخل في العينة B بأكثر من 330 دج؟

الحل: من خلال معطيات المثال نستخلص ما يلي:

N_1

$\mu_1 = 16500 \Longrightarrow \mu_{\bar{x}_1} 16500$

$n_1 = 125$

$\sigma_1 = 2500 \Longrightarrow {\sigma_1}^2 = 6250000$

$${\sigma_{\bar{x}_1}}^2 = \frac{{\sigma_1}^2}{n_1} = \frac{6250000}{125} = 50000$$

$$\bar{x}_1 \sim N(16500\,, 50000)$$

N_2

$\mu_2 = 15700 \Longrightarrow \mu_{\bar{x}_2} = 15700$

$n_2 = 100$

$\sigma_2 = 2000 \Longrightarrow {\sigma_2}^2 = 4000000$

$${\sigma_{\bar{x}_2}}^2 = \frac{{\sigma_2}^2}{n_2} = \frac{4000000}{100} = 40000$$

$$\bar{x}_2 \sim N(15700\,, 40000)$$

- **حساب توزيع المعاينة للفرق بين متوسطي العينتين** $(\bar{x}_1 - \bar{x}_2)$:

$$\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 16500 - 15700 = 800$$

$$\sigma^2_{(\bar{x}_1 - \bar{x}_2)} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} = 50000 + 40000 = 90000$$

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}} = \sqrt{90000} = 300$$

لا نعلم طبيعة توزيع المجتمعين الأصليين ولكن بما أن حجم العينتين المسحوبتين من مجتمعين كبيرتين:

$$n_1 = 125 > 30\,, n_2 = 100 > 30$$

فإنه حسب نظرية النهايات المركزية يمكن اعتبار أن المتغير العشوائي $(\bar{x}_1 - \bar{x}_2)$ يتبع التوزيع الطبيعي:

$$(\bar{x}_1 - \bar{x}_2) \sim N(\mu_{(\bar{x}_1 - \bar{x}_2)}, \sigma_{(\bar{x}_1 - \bar{x}_2)}) \Longrightarrow (\bar{x}_1 - \bar{x}_2) \sim N(800\,, 300)$$

- **احتمال أن يكون متوسط الدخل في العينة A أكبر من متوسط الدخل في العينة B بأكثر من 330:**

$$p[(\bar{x}_1 - \bar{x}_2) > 330] = p\left(Z > \frac{330 - \mu_{\bar{x}_1 - \bar{x}_2}}{\sigma_{\bar{x}_1 - \bar{x}_2}}\right) = p\left(Z > \frac{330 - 800}{300}\right)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 330] = p(Z > -1.56) = 0.5 + p(0 < Z < 1.56)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 330] = 0.5 + 0.4406 = 0.9406$$

- **الحالة الثالثة: إذا كان تبايني المجتمعين غير معلومين**

لحساب تباين توزيع مجموع او فرق متوسطي عينتين في حالة تبيان المجتمعين غير معلوم، نقوم باستخدام تباين العينة S^2 بدلا من تباين المجتمع σ^2، وبالتالي فإن تباين الأوساط الحسابية للمتغير العشوائي الذي يعبر على مجموع او فرق متوسطي العينتين يرمز له في هذه الحالة $S^2_{(\bar{x}_1 \pm \bar{x}_2)} = \frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}$ بدلا من $\sigma^2_{(\bar{x}_1 \pm \bar{x}_2)}$، ويتبع توزيع المعاينة للمتوسط الحسابي للعينات في هذه الحالة التوزيع الطبيعي إذا كان حجم العينتين كلا على حدى $n_1 \geq 30\,, n_2 \geq 30$، بينما يتبع توزيع ستودنت إذا كان حجم كل عينة أقل من 30.

- **إذا كان تبايني المجتمع مجهولين و$n \geq 30$**: في هذه الحالة يتبع توزيع المعاينة للمتغير العشوائي الذي يعبر على مجموع او فرق متوسطي العينتين $(\bar{x}_1 \mp \bar{x}_2)$ التوزيع الطبيعي مع استبدال تباين انحراف المجتمعين σ_1 , σ_2 بانحراف العينتين S_1 , S_2، ونكتب:

$$(\bar{x}_1 \pm \bar{x}_2) \backsim N(\mu_{(\bar{x}_1 \mp \bar{x}_2)}, S^2{}_{(\bar{x}_1 \mp \bar{x}_2)})$$

مثال 3: إذا كان لدينا مجتمعين إحصائيين مستقلين يتبعان التوزيع الطبيعي، حجم المجتمع الأول N_1 بمتوسط 20 وتباين مجهول $X_1 \backsim N(20, \sigma^2)$، وحجم المجتمع الثاني N_2 بمتوسط 16 وتباين مجهول $X_2 \backsim N(16, \sigma^2)$، وتم سحب عينتين من هذين المجتمعين $n_1 = 100$ و$n_2 = 80$ على التوالي، وكان تباين العينتين هو على التوالي ${S_1}^2 = 49$ و${S_2}^2 = 36$.

المطلوب: أوجد توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$ ثم أحسب $p[(\bar{x}_1 - \bar{x}_2) \leq 2]$

الحل: بما ان تبايني المجتمعين الأصليين مجهولين، إلا ان حجم العينتين كبير $n_1 = 100 > 30$، و$n_2 = 80 > 30$، وبالتالي فإن توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$ يتبع التوزيع الطبيعي، ونكتب:

$$(\bar{x}_1 - \bar{x}_2) \backsim N\left[\left(\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}\right), \left(S^2{}_{(\bar{x}_1 - \bar{x}_2)} = S^2{}_{\bar{x}_1} + S^2{}_{\bar{x}_2}\right)\right]$$

- **حساب توزيع المعاينة لفرق متوسطي عينتين مستقلتين**

$$\mu_{\bar{x}_1} = \mu_1 = 20$$

$$\mu_{\bar{x}_2} = \mu_2 = 16$$

$$\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 20 - 16 = 4$$

$$S^2{}_{\bar{x}_1} = \frac{{S_1}^2}{n_1} = \frac{49}{100} = 0.49$$

$$S^2{}_{\bar{x}_2} = \frac{{S_2}^2}{n_2} = \frac{36}{80} = 0.45$$

$$S^2{}_{(\bar{x}_1 - \bar{x}_2)} = S^2{}_{\bar{x}_1} + S^2{}_{\bar{x}_2} = 0.49 + 0.45 = 0.94$$

$$S_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{0.94} = 0.969$$

$$(\bar{x}_1 - \bar{x}_2) \backsim N(4 , 0.94)$$

- **حساب احتمال** $p[(\bar{x}_1 - \bar{x}_2) \leq 2]$

$$p[(\bar{x}_1 - \bar{x}_2) \leq 2] = p\left(Z \leq \frac{2 - \mu_{(\bar{x}_1 - \bar{x}_2)}}{S_{(\bar{x}_1 - \bar{x}_2)}}\right) = p\left(Z \leq \frac{2-4}{0.969}\right)$$

$$p[(\bar{x}_1 - \bar{x}_2) \leq 2] = p(Z \leq -2.06) = 0.5 - p(0 < Z < 2.06)$$

$$[(\bar{x}_1 - \bar{x}_2) \leq 2] = 0.5 - 0.4803 = 0.0197$$

- **إذا كان تبايني المجتمع مجهولين و$n < 30$**: إذا سحبت عينتين عشوائيتين من مجتمعين مستقلين يتبعان التوزيع الطبيعي بمتوسطين μ_1 وμ_2 وتباين${\sigma_1}^2$ و${\sigma_2}^2$ غير معلومين، وكان المقدار $(\bar{x}_1 \pm \bar{x}_2)$هو المتغير العشوائي الذي يعبر على مجموع أو فرق متوسطي العينتين، وكان حجم العينتين صغيرتين $n_i < 30$، فإن توزيع المعاينة للمتغير العشوائي $(\bar{x}_1 \pm \bar{x}_2)$ يتبع توزيع t ستودنت حيث:

$$T = \frac{(\bar{x}_1 \pm \bar{x}_2) - (\mu_1 \pm \mu_2)}{\sqrt{\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}}}$$

ويخضع $Student\ \ t$ لدرجة حرية مركبة تعطى بالصيغة التالية: $df = v = n_1 + n_2 - 2$

مثال 4: إذا كان لدينا مجتمعين إحصائيين مستقلين يتبعان التوزيع الطبيعي، حجم المجتمع الأول N_1 بمتوسط 40 وتباين مجهول $X_1 \sim N(40, \sigma^2)$، وحجم المجتمع الثاني N_2 بمتوسط 50 وتباين مجهول $X_2 \sim N(50, \sigma^2)$، وتم سحب عينتين من هذين المجتمعين $n_1 = 22$ و$n_2 = 25$ على التوالي، وكان تباين العينتين هو على التوالي ${S_1}^2 = 9$ و${S_2}^2 = 16$.

المطلوب: أوجد توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$ ثم أحسب $p[(\bar{x}_1 - \bar{x}_2) > 12]$

الحل: بما ان تبايني المجتمعين الأصليين مجهولين، وحجم العينتين صغير $n_1 = 22 < 30$، و$n_2 = 25 <$ 30، وبالتالي فإن توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$ يتبع توزيع t ستودنت، بدرجة حرية:

$$df = n_1 + n_2 - 2 = 22 + 25 - 2 = 45$$

$$(\bar{x}_1 - \bar{x}_2) \sim T(n_1 + n_2 - 2) \Longrightarrow (\bar{x}_1 - \bar{x}_2) \sim T(45)$$

- **حساب توزيع المعاينة لفرق متوسطي عينتين مستقلتين**

$$\mu_{\bar{x}_1} = \mu_1 = 50$$

$$\mu_{\bar{x}_2} = \mu_2 = 40$$

$$\mu_{(\bar{x}_1-\bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 50 - 40 = 10$$

$$S^2{}_{\bar{x}_1} = \frac{{S_1}^2}{n_1} = \frac{9}{22} = 0.409$$

$$S^2{}_{\bar{x}_2} = \frac{{S_2}^2}{n_2} = \frac{16}{25} = 0.64$$

$$S^2{}_{(\bar{x}_1-\bar{x}_2)} = S^2{}_{\bar{x}_1} + S^2{}_{\bar{x}_2} = 0.409 + 0.64 = 1.049$$

$$S_{(\bar{x}_1-\bar{x}_2)} = \sqrt{1.049} = 1.024$$

- **حساب احتمال** $p[(\bar{x}_1 - \bar{x}_2) > 12]$

$$p[(\bar{x}_1 - \bar{x}_2) > 12] = p\left(T > \frac{12 - \left(\mu_{(\bar{x}_1-\bar{x}_2)}\right)}{S_{(\bar{x}_1-\bar{x}_2)}}\right) = p\left(T > \frac{12 - 10}{1.024}\right)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 12] = p(T > 1.953)$$

من خلال القراءة في جداول توزيع t ستودنت نبحث عن القيمة المساوية لـ 1.953 أو الأكبر منها مباشرة، عند درجة حرية $df = 45$ ثم نبحث على درجة الاحتمال المقابلة لها في السطر العلوي الأول للجدول، وبعد الاطلاع على الدول يتضح أن:

$$p[(\bar{x}_1 - \bar{x}_2) > 12] = p(T > 1.953) = 0.025$$

- **حالة خاصة إذا كان تبايني المجتمع متجانسين (متساويين) ومجهولين و$n < 30$** : إذا سحبت عينتين عشوائيتين من مجتمعين مستقلين يتبعان التوزيع الطبيعي بمتوسط μ_1 وμ_2 وتباين${\sigma_1}^2$ و${\sigma_2}^2$ ومتجانسين أي ${\sigma_1}^2 = {\sigma_2}^2 = \sigma^2$، وكان المقدار $(\bar{x}_1 \pm \bar{x}_2)$ متغير العشوائي يمثل مجموع أو فرق متوسطي العينتين، وكان حجم العينتين صغيرتين حيث $n_i < 30$، فإن توزيع المعاينة للمتغير العشوائي $(\bar{x}_1 \pm \bar{x}_2)$ يتبع توزيع t ستودنت، حيث:

$$T = \frac{(\bar{x}_1 \pm \bar{x}_2) - (\mu_1 \pm \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

مع العلم أن S_p^2 يحسب يقدر من خلال العلاقة التالية:

$$S_p^2 = \frac{(n_1 - 1)S^2{}_{\bar{x}_1} + (n_2 - 1)S^2{}_{\bar{x}_2}}{n_1 + n_2 - 2}$$

وتعطى القيمة t بدرجة حرية التالية: $df = v = n_1 + n_2 - 2$

مثال 5: إذا كان لدينا مجتمعين إحصائيين مستقلين يتبعان التوزيع الطبيعي، حيث $X_1 \sim N(15, \sigma^2)$، و$X_2 \sim N(10, \sigma^2)$، وتم سحب عينتين من هذين المجتمعين حيث $n_1 = 12$ بتباين ${S_1}^2 = 16$ أما العينة الثانية فكان حجمها $n_2 = 10$ وبتباين ${S_2}^2 = 25$. فإذا علمت أن تباين المجتمعين متجانسين (متساويين) ${\sigma_1}^2 = {\sigma_2}^2 = \sigma^2$ أوجد توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$، ثم أحسب $p[(\bar{x}_1 - \bar{x}_2) > 8]$

الحل: بما ان تبايني المجتمعين الأصليين مجهولين، وحجم العينتين صغير $n_1 = 22 < 30$، و$n_2 = 25 < 30$، وبالتالي فإن توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{x}_1 - \bar{x}_2)$ يتبع توزيع t ستودنت، بدرجة حرية:

$$df = n_1 + n_2 - 2 = 12 + 10 - 2 = 20$$

$$(\bar{x}_1 - \bar{x}_2) \sim T(n_1 + n_2 - 2) \Longrightarrow (\bar{x}_1 - \bar{x}_2) \sim T(20)$$

- **حساب توزيع المعاينة لفرق متوسطي عينتين مستقلتين**

$$\mu_{\bar{x}_1} = \mu_1 = 15$$

$$\mu_{\bar{x}_2} = \mu_2 = 10$$

$$\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 15 - 10 = 5$$

$$S^2{}_{\bar{x}_1} = \frac{{S_1}^2}{n_1} = \frac{16}{12} = 1.33$$

$$S^2{}_{\bar{x}_2} = \frac{{S_2}^2}{n_2} = \frac{25}{10} = 2.5$$

$$S_p^2 = \frac{(n_1 - 1)S^2{}_{\bar{x}_1} + (n_2 - 1)S^2{}_{\bar{x}_2}}{n_1 + n_2 - 2} = \frac{(12-1)16 + (10-1)25}{12 + 10 - 2} = 20.05$$

- **حساب احتمال** $p[(\bar{x}_1 - \bar{x}_2) > 8]$

$$p[(\bar{x}_1 - \bar{x}_2) > 8] = \left(T > \frac{8 - (\mu_1 \pm \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right)$$

$$p[(\bar{x}_1 - \bar{x}_2) > 8] = p\left(T > \frac{8 - 5}{\sqrt{20.05\left(\frac{1}{12} + \frac{1}{10}\right)}} \right) = p(T > 1.56)$$

من خلال القراءة في جداول توزيع t ستودنت نبحث عن القيمة المساوية لـ 1.56 أو الأكبر منها مباشرة، عند درجة حرية $df = 20$ ثم نبحث على درجة الاحتمال المقابلة لها في السطر العلوي الأول للجدول، وبعد الاطلاع على الدول يتضح أن:

$$p[(\bar{x}_1 - \bar{x}_2) > 8] = p(T > 1.56) = 0.05$$

5. توزيع المعاينة لمجموع أو فرق نسبتي عينتين $(\acute{p}_1 \pm \acute{p}_2)$: إذا سحبنا عينتان مستقلتان حجمهما n_1 وn_2، بصورة متكررة من مجتمعين إحصائيين N_1 وN_2 يخضعان إلى توزيع ذي الحدين، بمعلمتين p_1 وp_2 على التوالي، فإن المجموع أو الفروق بين نسبتي العينتين $(\acute{p}_1 \pm \acute{p}_2)$ هو متغير عشوائي يخضع للتوزيع الطبيعي وسطه $\mu_{(\acute{p}_1 \pm \acute{p}_2)}$ وتباينه $\sigma^2_{(\acute{p}_1 \pm \acute{p}_2)}$ حيث:

أ- التوقع الرياضي أو متوسط توزيع المعاينة لمجموع أو فرق نسبتي عينتين

$$\mu_{(\acute{p}_1 + \acute{p}_2)} = \mu_{\acute{p}_1} + \mu_{\acute{p}_2} = p_1 + p_2$$

$$\mu_{(\acute{p}_1 - \acute{p}_2)} = \mu_{\acute{p}_1} - \mu_{\acute{p}_2} = p_1 - p_2$$

ب- التباين والانحراف المعياري لمجموع أو فرق نسبتي عينتين

$$\sigma^2{}_{(\acute{p}_1 + \acute{p}_2)} = \sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2} = \frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}$$

$$\sigma^2{}_{(\acute{p}_1 - \acute{p}_2)} = \sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2} = \frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}$$

$$\sigma_{(\acute{p}_1 - \acute{p}_2)} = \sqrt{\sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2}} = \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}}$$

أ- **طبيعة توزيع مجموع أو فرق نسبتي عينتين:** يتبع توزيع المعاينة للفرق أو مجموع نسبتي عينتين التوزيع الطبيعي إذا تحققت الشرط التالية:

$n_1 \geq 30$، و $n_1 * p_1 \geq 5$، و $n_1 * q_1 \geq 5$،

$n_2 \geq 30$، و $n_2 * p_2 \geq 5$، و $n_2 * q_2 \geq 5$

ويمكن كتابة: $$(\acute{p}_1 - \acute{p}_2) \sim N\left(\mu_{(\acute{p}_1 - \acute{p}_2)}, \sigma^2{}_{(\acute{p}_1 - \acute{p}_2)}\right)$$

ملاحظة: في حالة المعاينة النفادية (مجتمعين محدودين والسحب دون إرجاع مع $\left(\frac{n_i}{N_i} > 0.05\right)$ فإنه لابد من الأخذ بعين الاعتبار معامل التصحيح في حساب التباين والانحراف المعياري لمجموع أو فرق نسبتي عينتين، حيث تصبح العلاقة السابقة بالشكل الموالي:

$$\sigma_{(\acute{p}_1-\acute{p}_2)} = \sqrt{\sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2}} = \sqrt{\frac{p_1 * q_1}{n_1} * \frac{N-n}{N-1} + \frac{p_2 * q_2}{n_2} * \frac{N-n}{N-1}}$$

مثال: إذا كانت نسبة النجاح في طلبة السنة الثانية تخصص تسيير هو 70%، بينما كانت نسبة النجاح في طلبة السنة الثانية تخصص مالية ومحاسبة هو 75%، فإذا تم سحب عينة عشوائية حجمها 45 طالب من كل تخصص. فما هو احتمال أن تكون نسبة النجاح في عينة طلبة التسيير أكبر من نسبة النجاح في عينة طلبة المالية والمحاسبة؟

الحل: من خلال معطيات المثال نستخلص ما يلي:

N_1

$$p_1 = 0.7 \Longrightarrow q_1 = 0.3$$
$$n_1 = 45$$
$$\mu_{\acute{p}_1} = p_1 = 0.7$$
$$\sigma^2{}_{\acute{p}_1} = \frac{p_1 * q_1}{n_1} = \frac{0.7 * 0.3}{45}$$
$$\sigma^2{}_{\acute{p}_1} = 0.00466$$

N_2

$$p_2 = 0.75 \Longrightarrow q_2 = 0.25$$
$$n_2 = 45$$
$$\mu_{\acute{p}_2} = p_2 = 0.75$$
$$\sigma^2{}_{\acute{p}_2} = \frac{p_2 * q_2}{n_2} = \frac{0.75 * 0.25}{45}$$
$$\sigma^2{}_{\acute{p}_2} = 0.00416$$

- **حساب توزيع المعاينة للفرق بين نسبتي العينتين** $(\acute{p}_1 - \acute{p}_2)$:

$$\mu_{(\acute{p}_1-\acute{p}_2)} = \mu_{\acute{p}_1} - \mu_{\acute{p}_2} = 0.7 - 0.75 = -0.05$$

$$\sigma^2{}_{(\acute{p}_1-\acute{p}_2)} = \sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2} = 0.00466 + 0.00416 = 0.00883$$

$$\sigma_{(\acute{p}_1-\acute{p}_2)} = \sqrt{\sigma^2{}_{\acute{p}_1} + \sigma^2{}_{\acute{p}_2}} = \sqrt{0.00883} = 0.0939$$

- **طبيعة توزيع المعاينة للفرق بين نسبتي العينتين**: لكي يتبع المتغير العشوائي $(\acute{p}_1 - \acute{p}_2)$ التوزيع الطبيعي لابد من توفر ثلاث شروط سنتحقق منها فيما يلي:

$$n_1 = 45 \geq 30$$
$$n_1 * p_1 = 45 * 0.7 = 31.5 \geq 5$$
$$n_1 * q_1 = 45 * 0.3 = 13.5 \geq 5$$

$$n_2 = 45 \geq 30$$
$$n_2 * p_2 = 45 * 0.75 = 33.75 \geq 5$$
$$n_2 * q_2 = 45 * 0.25 = 11.25 \geq 5$$

بما ان الشروط محققة فإن المتغيرة العشوائية $(\acute{p}_1 - \acute{p}_2)$ تتبع توزيع قريب من التوزيع الطبيعي ونكتب:

$$(\acute{p}_1 - \acute{p}_2) \sim N\left(\mu_{(\acute{p}_1-\acute{p}_2)} = -0.05, \sigma^2{}_{(\acute{p}_1-\acute{p}_2)} = 0.00883\right)$$

$$(\acute{p}_1 - \acute{p}_2) \sim N(-0.05\ \ , 0.00883)$$

- حساب احتمال أن تكون نسبة الناجحين في عينة طلبة التسيير أكبر من نسبة الناجحين في عينة طلبة المالية والمحاسبة $p[(\acute{p}_1 - \acute{p}_2) > 0]$:

$$p[(\acute{p}_1 - \acute{p}_2) > 0] = p(Z > \frac{0 - \mu_{(\acute{p}_1-\acute{p}_2)}}{\sigma_{(\acute{p}_1-\acute{p}_2)}} = p(Z > \frac{0 - (-0.05)}{0.0939}$$

$$p[(\acute{p}_1 - \acute{p}_2) > 0] = p\left(Z > \frac{0.05}{0.0939}\right) = p(Z > 0.53)$$

$$p[(\acute{p}_1 - \acute{p}_2) > 0] = p(Z > 0.53) = 0.5 - p(0 < Z < 0.53)$$

$$p[(\acute{p}_1 - \acute{p}_2) > 0] = 0.5 - 0.2019 = 0.2981$$

احتمال أن تكون نسبة الناجحين في عينة طلبة التسيير أكبر من نسبة الناجحين في عينة طلبة المالية والمحاسبة هو 29.81%.

6. توزيع المعاينة لتباين العينة S^2: يعتبر تباين العينة S^2 أفضل إحصاءة يمكن استخدامها للاستدلال على تباين المجتمع σ^2، حيث أن الإحصاءة S^2 تقيس درجة تشتت قيم مفردات داخل العينة العشوائية الواحدة، ويستخدم تباين العينة بكثرة في دراسات مراقبة جودة المنتجات ومدى مطابقتها للمعايير والمواصفات، فمثلا في عملية تصنيع علب الطماطم، إذا كان وزن العلب المنتجة تختلف كثير عن الوزن النمطي (500 غرام مثلا) فإنها تعتبر منتجات معيبة لا تطابق مواصفات الجودة المطلوبة، لتشتت أوزانها عن المتوسط بشكل كبير، وسنحاول في هذا العنصر معرفة كيفية تحديد الوسط الحسابي والانحراف المعياري لتباين العينة ثم توزيع المعاينة لتباين العينات الممكن سحبها من مجتمع معين.

إذا تم أخذ عينة عشوائية حجمها n من مجتمع يتبع التوزيع الطبيعي متوسطه μ وتباينة σ^2 وكان S^2 هو مقدار التباين للعينة المسحوبة حيث: $S^2 = \frac{\sum(x_i-\bar{x})^2}{n}$، وتتغير قيمة S^2 من عينة لأخرى، وإن اختلاف تباين العينات فيما بينها يشكل بدوره توزيع للمعاينة للتباين له توقع رياضي وتباين خاص به. لذلك سنقوم بتوضيح خواص توزيع المعاينة للتباين بحساب قيمة التوقع الرياضي والانحراف المعياري لهذه الإحصاءة.

أ- التوقع الرياضي والانحراف المعياري لتوزيع المعاينة لتباين العينة S^2:

- **التوقع الرياضي لتباين العينة في حالة في حالة مجتمع لا يتبع التوزيع طبيعي والمعاينة غير النفادية:**

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

ويمكن البرهنة على هذه العلاقة فيما يلي:

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{1}{n}\sum(x_i - \bar{x})^2 = \frac{1}{n}\sum(x_i - \mu + \mu - \bar{x})^2$$

$$S^2 = \frac{1}{n}\sum[(x_i - \mu) + (\mu - \bar{x})]^2$$

$$S^2 = \frac{1}{n}\sum(x_i - \mu)^2 + \frac{2}{n}\sum(x_i - \mu)(\mu - \bar{x}) + \frac{1}{n}\sum(\mu - \bar{x})^2$$

$$S^2 = \frac{1}{n}\sum(x_i - \mu)^2 + \frac{2}{n}(\mu - \bar{x})\sum(x_i - \mu) + \frac{1}{n}\sum(\mu - \bar{x})^2$$

من المعلوم أن مجموع الانحرافات عن الوسط الحسابي تساوي الصفر $\sum(x_i - \mu) = 0$ وعليه:

$$S^2 = \frac{1}{n}\sum(x_i - \mu)^2 + (\mu - \bar{x})^2$$

$$E(S^2) = E\left(\frac{1}{n}\sum(x_i - \mu)^2\right) - E(\bar{x} - \mu)^2$$

$$E(S^2) = \frac{1}{n}\sum E(x_i - \mu)^2 - Var(\bar{x})$$

$$E(S^2) = \sigma^2 - \sigma_{\bar{x}}^2$$

لقدا توصلنا في توزيع المعاينة للوسط الحسابي أن $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ وعليه يمكن كتابة:

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

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

وهذا يعني أنه إذا استخدمنا التباين $S^2 = \frac{\sum(x_i - \bar{x})^2}{n}$ سيكون مقدار متحيز لأن متوسط تباينات العينات المسحوبة (التوقع الرياضي لتباين العينة) ليس هو نفسه قيمة تباين المجتمع الأصلي الذي سحبت منه العينة.

ملاحظة: في حالة المعاينة النفادية (المجتمع محدود والسحب دون إرجاع مع $\left(\frac{n}{N} > 0.05\right)$ فإنه لابد من الأخذ بعين الاعتبار معامل التصحيح في حساب التوقع الرياضي لتباين العينة، فإذا كانت X متغير عشوائي تمثل مجتما ما محدود وS^2 هي متغيرة عشوائية تمثل تباين العينات النفادية الممكن سحبها من ذات المجتمع فإن القيمة المتوقعة للتباين تأخذ في هذه الحالة معامل التصحيح بعين الاعتبار وتعطى بالصيغة التالية:

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

التوقع الرياضي في حالة مجتمع يتبع التوزيع الطبيعي: في هذه لحالة نستخدم المقدار $\dot{S}^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$ ويطلق عليه مصطلح التباين غير المتحيز للعينة حيث:

$$\dot{S}^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} \Rightarrow \dot{S}^2 = \frac{n}{n}\frac{\sum(x_i-\bar{x})^2}{n-1} \Rightarrow \dot{S}^2 = \frac{n}{n-1}\frac{\sum(x_i-\bar{x})^2}{n}$$

إذا كان المقدار $\frac{\sum(x_i-\bar{x})^2}{n}$ هو مقدار تباين العينة المتحيز S^2 وعليه فإن قيمة $\dot{S}^2$ تعطى بالعلاقة التالية:

$$\dot{S}^2 = \frac{n}{n-1}S^2$$

وهذا يعني أن تباين العينة غير المتحيز يساوي إلى تباين العينة مضروبا في المقدار $\frac{n}{n-1}$

$$E(\dot{S}^2) = E\left(\frac{n}{n-1}S^2\right) \Rightarrow E(\dot{S}^2) = \frac{n}{n-1}E(S^2)$$

ولقد توصلنا أعلاه إلى أن $E(S^2) = \mu_{S^2} = \frac{n-1}{n}\sigma^2$ وعليه يمكن أن نكتب:

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

وهذا يعني أنه إذا استخدمنا التباين غير المتحيز $\dot{S}^2$ هو مقدار غير متحيز لأن متوسط تباينات العينات المسحوبة (التوقع الرياضي لتباين العينة) هو نفسه قيمة تباين المجتمع الأصلي الذي سحبت منه العينة.

- **تباين تباين العينة**: تعطى علاقة تباين تباين العينة $Var(S^2)$ بالصيغة الرياضية التالية:

$$Var(S^2) = \frac{(n-1)^2}{n^3}\left[\mu_4 + \frac{(n-3)}{n-1}\mu_2^2\right]$$

ويعبر المقدار μ_4 على العزم الرابع حول المتوسط الحسابي حيث: $\mu_4 = E(X_i-\mu)^4 = \frac{\sum(X_i-\mu)^4}{N}$، بينما يعبر المقدار μ_2^2 على مربع العزم الثاني حول المتوسط الحسابي: $\mu_2^2 = [E(X_i-\mu)^2]^2 = \left(\frac{\sum(X_i-\mu)^2}{N}\right)^2$.

وإذا حاولنا استخدام التباين غير المتحيز $Var(\dot{S}^2)$ بدلا التباين من $Var(S^2)$ فإن علاقة تباين التباين تصبح كما يلي:

نعلم أن $\dot{S}^2 = \frac{n}{n-1}S^2$ وعليه فإن $Var(\dot{S}^2)$ تصبح بالعلاقة التالية:

$$Var(\dot{S}^2) = Var\left(\frac{n}{n-1}S^2\right) \Rightarrow Var(\dot{S}^2) = \left(\frac{n}{n-1}\right)^2 Var(S^2)$$

$$Var(\dot{S}^2) = \frac{n^2}{(n-1)^2}\frac{(n-1)^2}{n^3}\left[\mu_4 + \frac{(n-3)}{n-1}\mu_2^2\right]$$

$$Var(\dot{S}^2) = \frac{1}{n}\left[\mu_4 + \frac{(n-3)}{n-1}\mu_2^2\right]$$

طبيعة توزيع المعاينة لتباين العينة $\dot{S}^2$: إذا كان لدينا X متغير عشوائي يتبع التوزيع الطبيعي، وتم سحب كل العينات الممكنة ذات الحجم n وحسبنا لكل عينة تباينها سينتج لنا متغير عشوائي $\dot{S}^2$ يعبر على تباينات العينات المسحوبة وعددها m (عدد العينات المسحوبة)، ويتبع هذا المتغير العشوائي توزيع احتمالي مستمر يسمى كاي مربع ***khi-deux*** بدرجة حرية $dl = v = n-1$، ويعتبر توزيع كاي مربع ***khi-deux*** من التوزيعات المستمرة الهامة في الإحصاء التطبيقي، وتقع قيم هذا التوزيع بأكملها في النصف الموجب، كما أن منحنى دالة الكثافة الاحتمالية لهذا التوزيع غير متماثل، فهو يعتبر من المنحنيات موجبة الالتواء، إلا انه يقترب من التماثل كلما زادت درجات الحرية $v = n-1$. وتعطى قيمة χ^2 بالصيغة التالية:

$$\dot{S}^2 \sim \chi^2_{(n-1)}$$

حيث:

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

وبتعويض تباين العينة بقيمته $\dot{S}^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$ تصبح العلاقة كما يلي:

$$\chi^2_{n-1} = \frac{(n-1)\frac{\sum(x_i-\bar{x})^2}{(n-1)}}{\sigma^2} = \frac{\sum(x_i-\bar{x})^2}{\sigma^2}$$

$$\chi^2_{n-1} = \frac{1}{\sigma^2}\sum(x_i - \bar{x} + \mu - \mu)^2 = \frac{1}{\sigma^2}\left[\sum(x_i - \mu)^2 - \sum(\bar{x} - \mu)^2\right]$$

$$\chi^2_{n-1} = \frac{1}{\sigma^2}\left[\sum(x_i - \mu)^2 - n(\bar{x} - \mu)^2\right]$$

$$\chi^2_{n-1} = \left[\frac{\sum(x_i - \mu)^2}{\sigma^2} - \frac{n(\bar{x} - \mu)^2}{\sigma^2}\right] \Rightarrow \chi^2_{n-1} = \frac{\sum(x_i - \mu)^2}{\sigma} - \frac{(\bar{x} - \mu)^2}{\frac{\sigma^2}{n}}$$

$$\chi^2_{n-1} = \left(\frac{\sum(x_i - \mu)}{\sigma}\right)^2 - \left(\frac{(\bar{x} - \mu)}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

تبين المعادلة أعلاه أن المعادلة $\chi^2_{n-1} = \frac{(n-1)\dot{S}^2}{\sigma^2}$ تتكون من مقدارين، المقدار الأول هو مجموع مربعات الدرجة المعيارية للتوزيع الطبيعي للمجتمع للمتغير العشوائي x_i، أما المقدار الثاني فيعبر على مربع الدرجة المعيارية للتوزيع الطبيعي للمجتمع ذو المتغير العشوائي x_i.

مثال 1: إذا تم سحب عينة حجمها 25 من مجتمع موزع بشكل طبيعي حجمه 100 وتباينه 36، فما هو احتمال أن يكون التباين في العينة المسحوبة أكبر أو يساوي 20.

الحل: من خلال معطيات المثال نستخلص:

$$n = 25 \qquad N = 100 \qquad \sigma^2 = 36$$

$$p(S^2 \geq 20) = p\left(\chi^2_{n-1} \geq \frac{(n-1)\dot{S}^2}{\sigma^2}\right) = p\left(\chi^2_{24} \geq \frac{(24)20}{36}\right)$$

$$(\chi^2_{24} \geq 13.33) = 0.95$$

من خلال جداول توزيع توزيع كاي مربع ***khi-deux*** χ^2 نبحث في عمود درجة الحرية على **24** ثم نبحث في السطر المقابل لها عن القيمة 13.33 أو الأكبر منها مباشرة لنجد قيمة الاحتمال مقابلة لها في السطر العلوي تساوي **0.95**.

مثال 2: إذا تم سحب عينة حجمها 6 من مجتمع موزع بشكل طبيعي بمتوسط 10 وتباينه 5، فما هو احتمال أن يكون التباين في العينة المسحوبة أقل من 3. ثم أجد القيمة المتوقعة للتباين في العينة.

الحل: من خلال معطيات المثال نستخلص:

$n = 6 \qquad \mu = 10 \qquad \sigma^2 = 5$

$$p(S^2 < 3) = p\left(\chi^2_{n-1} < \frac{(n-1)\dot{S}^2}{\sigma^2}\right) = p\left(\chi^2_5 < \frac{(5)3}{5}\right)$$

$$(\chi^2_5 < 3) = 1 - (\chi^2_5 \geq 3) = 1 - 0.7 = 0.3$$

$$E(\dot{S}^2) = \sigma^2 = 5$$

من خلال جداول توزيع كاي مربع $\boldsymbol{khi\text{-}deux}\ \chi^2$ نبحث في عمود درجة الحرية على 5 ثم نبحث في السطر المقابل لها عن القيمة 3 أو الأكبر منها مباشرة لنجد قيمة الاحتمال مقابلة لها في السطر العلوي تساوي 0.3.

7. توزيع المعاينة لنسبة بين تباين عينتين: للمقارنة بين تبايني مجتمعين فإننا نحتاج إلى النسبة بين تبايني عينتين مسحوبتين من هذين المجتمعين، ونلجأ لحساب النسب والتباينات وليس الفرق بينهما لسهولة دراسة النسب وتفسيرها، وسنعطي توزيع هذه النسبة في حالة المعاينة من مجتمعين طبيعيين ومستقلين.

نظرية: إذا كانت n_1 عينة عشوائية مسحوبة من مجتمع N_1 يتبع التوزيع الطبيعي بمتوسط μ_1 وتباين σ_1^2، وكانت n_2 عينة عشوائية مسحوبة من المجتمع N_2 مستقل عن N_1 ويتبع التوزيع الطبيعي بمتوسط μ_2 وتباين σ_2^2، فإذا كان $\dot{S}_1^2$ و $\dot{S}_2^2$ هما تباين العينتين الأولى والثانية على التوالي، فإننا يمكن كتابة:

$$\frac{(n_1-1)\dot{S}_1^2}{\sigma_1^2} \backsim \chi^2_{(n_1-1)} \qquad , \qquad \frac{(n_2-1)\dot{S}_2^2}{\sigma_2^2} \backsim \chi^2_{(n_2-1)}$$

وبما أن العينتين مستقلين فإنه حسب توزيع فيشر $\boldsymbol{Fisher\ F}$ فإن النسبة بين تبايني العينتين يمكن التعبير عليها بالمتغير العشوائي $\boldsymbol{F(n_1-1\,,n_2-1)}$ الذي له درجة حرية البسط $\boldsymbol{v_1 = n_1 - 1}$ ودرجة حرية المقام $\boldsymbol{v_2 = n_2 - 1}$ ويعطى بالصيغة: التالية:

$$F = \frac{\left.\dfrac{(n_1-1)\dot{S}_1^2}{\sigma_1^2}\right/(n_1-1)}{\left.\dfrac{(n_2-1)\dot{S}_2^2}{\sigma_2^2}\right/(n_2-1)} \backsim F(n_1-1\,,n_2-1)$$

$$F = \left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{\sigma_2^2}{\sigma_1^2}\right) \backsim F(n_1-1\,,n_2-1)$$

وتجر الإشارة إلى أن توزيع فيشر من أهم التوزيعات المستمرة المستخدمة في الإحصاء، وهو توزيع ملتوي إلى اليمين ومداه نظريا يكون من الصفر إلى زائد مالانهاية، كما يمكن تعريفه بانه نسبة بين توزيعين مستقلين يتبعان توزيع χ^2 وكل منهما له درجة حرية مختلفة، وبالتالي فإن منحنى الدالة سيقع بالكامل في النصف الموجب للمعلم المتعامد المتجانس كما في حالة توزيع كاي مربع، ويعتبر توزيع فيشر أيضا من التوزيعات غير المتماثلة ولكنه يقترب من التماثل (يقترب من التوزيع الطبيعي) كلما زادت درجات الحرية $\boldsymbol{v_1}$ و $\boldsymbol{v_2}$.

مثال 1: سحب عينة عشوائية حجمها 21 من مجتمع موزع بشكل طبيعي تباينه 36، كما سحب عينة عشوائية أخرى حجمها 25 من مجتمع طبيعي آخر مستقل عن المجتمع الأول تباينه 25، فما هو احتمال أن تكون النسبة بين تبايني العينتين المسحوبتين أكبر من 3.95.

الحل: من خلال معطيات المثال نستخلص:

$$n_1 = 21 \quad , \quad \sigma_1^2 = 36 \quad , \quad n_2 = 25 \quad , \quad \sigma_2^2 = 25$$

يكون توزيع النسبة بين تبايني متغير عشوائي F الذي يأخذ الصيغة التالية:

$$F = \left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{\sigma_2^2}{\sigma_1^2}\right) \backsim F(n_1 - 1\,, n_2 - 1) \Longrightarrow F = \left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{25}{36}\right) \backsim F(20\,, 24)$$

$$p\left(\frac{\dot{S}_1^2}{\dot{S}_2^2} > 3.95\right) = p\left(\left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{25}{36}\right) > 3.95 * \frac{25}{36}\right)$$

$$p(F > 2.74) = 0.01$$

من خلال جداول توزيع فيشر وباستخدام درجة حرية البسط $v_1 = 20$ ودرجة حرية المقام $v_2 = 20$، نستنتج قيمة الاحتمال للمتغير العشوائي من خلال الاحتمال الذي يعطي أقرب قيمة لـ 2.74، وفي هذه الحالة هو 0.01.

مثال 2: عينتين حجمهما 8 و10 مسحوبتين من مجتمعين طبيعيين تبايناهما على التوالي 20 و36. ما احتمال أن يكون تباين الأولى أكبر من ضعف تباين الثانية؟

الحل: من خلال معطيات المثال نستخلص:

$$n_1 = 8 \quad , \quad \sigma_1^2 = 20 \quad , \quad n_2 = 10 \quad , \quad \sigma_2^2 = 36$$

يكون توزيع النسبة بين تبايني متغير عشوائي F الذي يأخذ الصيغة التالية:

$$p\left(\dot{S}_1^2 > 2\left(\dot{S}_2^2\right)\right) = p\left(\frac{\dot{S}_1^2}{\dot{S}_2^2} > 2\right)$$

$$F = p\left[\left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{\sigma_2^2}{\sigma_1^2}\right) > \left(2 * \frac{\sigma_2^2}{\sigma_1^2}\right)\right] \backsim F(n_1 - 1\,, n_2 - 1)$$

$$p\left[\left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{36}{20}\right) > \left(2 * \frac{36}{20}\right)\right] \Longrightarrow F = \left(\frac{\dot{S}_1^2}{\dot{S}_2^2} * \frac{36}{20}\right) \backsim F(7\,, 9)$$

$$p[F > 3..6] = 0.05$$

السلسلة رقم (3)

تمرين رقم1

ب- إذا كان المتغير العشوائي المستمر X له توزيع طبيعي $N\ (u,\sigma^2)$، أوجد ما يلي:

4. $P(u-\partial<X<u+\partial)$
5. $P(u-2\,\partial<X<u+2\,\partial)$
6. $P(u-3\,\partial<X<u+3\,\partial)$

ب- نفرض أن Z متغير عشوائي مستمر يتبع توزيع طبيعي قياسي N(0,1) احسب ما يلي:

1.p($0<Z<1.42$) / 2.p($-0.73<Z<0$) / 3.p($-1.79<Z<-0.54$)/ 4.p($Z>1.13$)/ 5.p($-1.37<Z<2.01$) / 6.p($0.65<Z<1.26$) / 7.p($-0.5<Z<0.5$)

تمرين رقم 2: إذا كان Z متغير عشوائي مستمر يتبع توزيع طبيعي قياسي N(0.1)، أوجد قيمة a و b في الحالات التالية:

1.p($0<Z<a$)=0.1915/ 2.p($Z<a)=0.9$/ 3.p($Z<a)=0.05$/
4.p($Z>a)=0.95$/ 5.p($a<Z<b)=0.9974$

علما أن a و b في الحالة رقم 5 نقطتين متماثلتين على منحنى التوزيع الطبيعي القياسي.

تمرين رقم 3: إذا كان سن السكان في مدينة ما يتبع توزيع طبيعي بوسط مقداره 32 سنة وانحراف معياري 18 سنة. فإذا اختير شخص بطريقة عشوائية من هذه المدينة، فما هو احتمال أن يكون هذا الشخص:

7. عمره بين 32 و 59 سنة.
8. عمره أكبر من 50 سنة.
9. عمره أقل من 18 سنة.
10. عمره بين 20 و 40 سنة.
11. إذا أراد صاحب قاعة السينما منح تخفيض مقداره 5% على سعر التذكرة للمسنين في هذه المدينة، فما هو السن الذي سيمنح له هذا التخفيض؟
12. إذا كان هذا التخفيض سيمس 15% من المسنين، فما هو السن الموافق لذلك؟

تمرين رقم 4: نفرض أن أوزان 800 طالب تتبع توزيع طبيعي بوسط مقداره 66 كلغ وانحراف معياري 5 كلغ.أوجد عدد الطلبة الذين يكون وزنهم:

3. بين 65 و 70 كلغ.
4. أكبر من 72 كلغ.