المحاضرة 01: الإحصاء الوصفي والاستدلالي

يمكن تعريف علم الإحصاء بشكل عام بأنه العلم الذي يهتم بتوفير الحقائق الرقمية للظواهر المختلفة، ومن ثم عرضها وتحليلها للوصول إلى نتائج محددة نستطيع من خلالها فهم الظاهرة وتفسيرها، ومن ثم عرضها وتحليلها للوصول إلى نتائج محددة نستطيع من خلالها فهم الظاهرة وتفسيرها. والتنبؤ بسيرها مستقبلا يتضمن إجراءها اتباع الخطوات التالية.

1. **جمع البيانات**: وهي تمثل جميع المعلومات العددية (الرقمية) أو الوصفية عن الظاهرة محل الدراسة، يتم جمعها بالاعتماد على أمرين:

- **مصادر جمع البيانات**: منها ما هو مباشر، ويتم من خلال النزول للميدان، وجمع المعلومات بشكل مباشر، ويتم الاعتماد في ذلك على الاستبيان، والاختبار، والمقابلات الشخصية، ومنها ما هو غير مباشر أو مصادر ثانوية لجمع المعلومات كالاعتماد على وسيط مثلا.

أساليب، أو طرق جمع البيانات: ويتم بطريقتين.

المسح الشامل: ونقصد به جمع البيانات من جميع عناصر المجتمع الإحصائي، ويعتمد إذا كان الغرض من البحث هو حصر جميع مفردات المجتمع، وفي هذه الحالة يتم تجميع بيانات عن كل مفردة من مفردات المجتمع دون استثناء.

أسلوب المعاينة: يقوم هذا الأسلوب باختيار جزء من المجتمع محل الدراسة. ومن ثم تعميم النشأ اجعليه، مع شرط أن تكون العينة ممثلة تمثيلا جيدا للمجتمع الإحصائي، كما تعرف العينة على أنها جزء من المجتمع يتم اختياره لتمثيل المجتمع بأكمله، وعليه فإن المعاينة تعرف على أنها عملية اختيار جزء من المجتمع الإحصائي للاستدلال على خواص المجتمع بأكمله عن طريق تعميم نتائج العينة على المجتمع، وبالتالي نقصد بحجم المجتمع جميع وحدات المعاينة التي يتكون منها المجتمع، أما حجم العينة فهي عدد وحدات المعاينة ويعتبر حجم العينة صغيرة إذا كانت أن أصغر من 30.

2. **تنظيم وتبويب البيانات وعرضها**: وذلك بالاعتماد على الجداول الإحصائية أو الرسوم البيانية حتى يسهل قراءتها، والاطلاع على مؤشراتها وخصائصها، وتسمى ببيانات مبوبة أو مجدولة، بحيث يتم تنظيمها في جداول خاصة تسمى بجداول التوزيع التكراري والفكرة الأساسية، التي. والفكرة الأساسية، التي. والفكرة الأساسية

في التوزيع التكراري هي تحديد عدد مرات ظهور أي قيمة، ثم تسجيل تلك القيم في جدول، وأمام كل منها عدد مرات ظهورها.

من أمثلة ذلك عرض وتنظيم البيانات في جداول إحصائية أو جداول توزيع تكرارية وتكون إما بسيطة، أو مركبة حسب عدد المتغيرات أو تنظيمها في جداول توزيع تكراري على شكل فئات كما هو موضح في الجداول التالية: مثال : لنفترض لدينا بيانات عن عدد العمال في مصنع ما، موزعين على عدة أقسام أو أفرع، فيتم تنظيم البيانات هنا حسب المتغيرات المعتمدة في الدراسة، فسيتم الاعتماد على متغيري الجنس والفرع أو القسم كمثال:

جدول (01): توزيع أفراد عينة الدراسة حسب هام، وتغير هي الجنس وأفرع العمل.

الفرع / الجنس	ذكور		إناث		المجموع
	التكرار	النسبة المئوية	التكرار	النسبة المئوية	
التسويق					
المحاسبة					
الموارد البشرية					
المجموع					

- كما يمكن الاعتماد على مقاييس النزعة المركزية ومقاييس التشتت في وصف البيانات

تكوين جدول توزيع تكراري على شكل فئات:

نتبع الخطوات التالية:

- نوجد المدى المطلق للبيانات حسب القانون التالي: المدى= أكبر قيمة –أصغر قيمة
- نحدد عدد فئات مناسب لعدد البيانات بحيث لا يقل عن 5 ولا يزيد عن 15
- نحدد طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$
- نحدد حدود الفئات: الحد الأعلى= الحد الأدنى +طول الفئة –1

3. **المعالجة الرياضية:** إذ تتم من خلالها معالجة البيانات رياضيا بالاعتماد على مجموعة من الطرق والأساليب الإحصائية على حسب نوع الفرض ونوع البيانات في شكل مؤشرات إحصائية، مثل مقاييس النزعة المركزية أو التشتت، أو معاملات الارتباط وغيرها.

4. **تحليل النتائج واتخاذ القرارات:** وتعدّ من أهم مراحل العملية الإحصائية وبدونها فإن النتائج تبقى صماء لا معنى لها إذ يتم أولا إجراء القراءة الإحصائية للنتائج بالاعتماد على مجموعة من المؤشرات والمعطيات الرقمية وبناء عليها يتم الخروج باستنتاجات تناسب واقع الظاهرة المدروسة، وغالبا ما تكون على شكل تقديرات أو تنبؤات أو تعميمات أو قرارات رفض أو قبول للفرضيات الإحصائية. ومنها يمكن الإشارة إلى الإحصاء التطبيقي على أنه العلم الذي يهتم أو ويبحث في جمع البيانات وتنظيمها وعرضها وتحليلها واستقراء النتائج واتخاذ القرارات بناء عليها.

- **الإحصاء الوصفي:** يتضمن جمع وعرض ووصف للبيانات وتنظيمها وعرضها في جدال ورسوم بيانية وحساب مقاييس التشتت ومقاييس النزعة المركزية مثال ذلك معرفة آراء المجتمع حول مشكلة ما، أو دراسة أداء العمال في مهارة معينة، أو معرفة مقدار استهلاك المجتمع الجزائري لمنتج ما. ومنه فإن الإحصاء الوصفي يهدف إلى إعطاء صورة عامة عن الظاهرة محل الدراسة وإتجاهها ويصف خصائصها وعلاقاتها بغيرها من الظواهر.
- **الإحصاء الاستدلالي:** يستخدم لاستخلاص استنتاجات حول مجتمع الدراسة أو للاستدلال على خواص المجتمع بناء على عينة منه بمعنى أنه يساعد في التنبؤ واتخاذ قرارات بناء على البيانات المتاحة (المستخرجة من عينة الدراسة) ويستخدم في ذلك معاملات الارتباط والانحدار والأساليب الفرقية (يتعامل مع التنبؤات والاستدلالات وتقدير المعلمات) بمعنى أن الباحث يستخدم الإحصاء الاستدلالي لغرضين، أحدهما يتعلق بتقدير قيم بارامترات المجتمع الإحصائي، والثاني يتعلق باختبار صحة الفروض المتعلقة بهذه القيم.

المحاضرة 02: الأساليب الإحصائية البارامترية واللابارمترية

الإحصاء البارامتري هو أحد أنواع الأساليب الإحصائية الإستدلالية التي تهتم بالكشف والإستدلال عن المجتمع اعتمادا على ما توافر من بيانات لدى الباحث والمأخوذة من عينة منه كما تتناول أساليب اتخاذ القرارات الإحصائية واختبار الفرضيات أي أن الإحصاء الاستدلالي يهتم بالاستدلال على خصائص خصائص المجتمع استنادا على معلومات نحصل عليها من العينات، وبذلك فهو يختلف عن الإحصاء الوصفي الذي يهتم بدراسة الظواهر ويصف خصائصها وعلاقاتها بغيرها من الظواهر كما نجده يهتم بتنظيم البيانات وعرضها في جداول ورسوم بيانية وأشكال هندسية وحساب مقاييس النزعة المركزية ومقاييس التشتت.

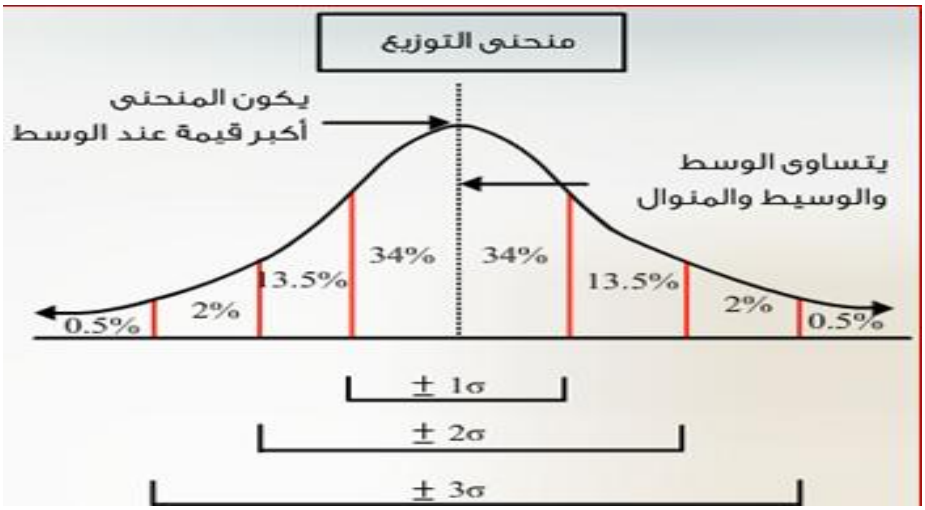
يستخدم الإحصاء البارامتري في حالة العينات الكبيرة التي تشترط فيها توفر معلومات عن مجتمعاتها (معلمات الأصل) بحيث أن n تساوي او أكبر من30، توزيع البيانات توزيعا اعتداليا، تجانس التباين، المعاينة العشوائية، خطية العلاقة (يكشف عليها بشكل الانتشار)، المتغير التابع يجب أن يكون كمي ويستخدم فقط مع البيانات من نوع المسافة أو النسبة، ويعدّ بذلك أدق وأكثر كفاءة من الإحصاء اللابارامتري والذي يندرج أيضا ضمن

الأساليب الإحصائية الاستدلالية التي لا تتقيد بالشروط الواجب توافرها لاستخدام الإحصاء البارامتري .فهو يتحرر من التوزيع الاعتدالي للمجتمع الذي سحبت منه العينة، كما يتحرر من حجم العينة بالإضافة .إلى ذلك، فهو يصلح لمعالجة البيانات التصنيفية أو الرتبية ويؤخذ عليها بأنها أقل كفاءة ودقة من البارامترية كما أن نتائجها يمكن تعميمها بحذر .

1. التوزيع الطبيعي للبيانات: Normal Distribution

تنقسم التوزيعات التكرارية إلى توزيعات متماثلة وأخرى غير متماثلة أو ملتوية.

التوزيعات الاعتدالية أو المتماثلة تعتمد على العلاقة بين المتغير والتكرارات بحيث تتزايد التكرار ات تدريجيا كلما زاد المتغير حتى تصل التكرارات إلى قمة معينة ثم تبدأ بعدها في النزول، بمعنى أن التوزيع متماثل في صعوده وهبوطه حول قيمة معينة للمتغير، ويرجع الفضل في اكتشاف الأساس النظري لهذا المنحنى إلى Gausse & Laplace ولهذا يسمى بمنحنى جاوس ويكون التوزيع متماثلا عندما تتطابق فيه قيم مقاييس النزعة المركزية (**المتوسط =الوسيط= المنوال**)، كما هو موضح في الشكل أدناه:



- بين ±1 تقع 68% من البيانات أي أن الظاهرة تتوزع توزيع طبيعيا إذا كان 68% من بياناتها تنحصر بين ±1+
- بين ±2 تقع 95% من البيانات أي أن الظاهرة تتوزع توزيعا طبيعيا إذا كان 95% من بياناتها تنحصر بين ±2.
- بين ±3 تقع 99% من البيانات أي أن الظاهرة تتوزع توزيعا طبيعيا إذا كان 99% من بياناتها تنحصر بين ±3

من خواصه أيضا أن يكون **معامل التوائه يساوي الصفر وتفرطحه يساوي 03**.

ويمكن تحديد شكل التوزيع من خلال حساب معامل الالتواء

بحيث أن **معامل الالتواء=** $\frac{\text{المتوسط}-\text{المنوال}}{\text{الانحراف المعياري}}$ كما يمكن صياغة المعادلة بدلالة المتوسط والوسيط كالتالي: **المنوال = 3 الوسيط– 2 المتوسط**، وبالتعويض عن قيمة المنوال في المعادلة يمكن أيضا حساب معامل الالتواء من المعادلة التالية:

معامل الالتواء = $\frac{\mathbf{3}(\text{المتوسط}-\text{الوسيط})}{\text{الانحراف المعياري}}$

فإذا كانت قيمة معامل الالتواء تساوي الصفر أو تقترب منه فيمكن القول بأن منحنى التوزيع اعتدالي.

- **التوزيعات الملتوية:** يعرف الالتواء بأنه درجة التماثل أو البعد عن التماثل لتوزيع ما ويتضمن التوزيعات التالية:

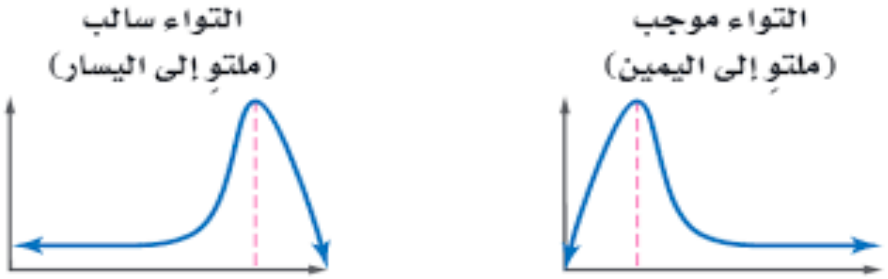
التواء موجب (توزيع ملتو نحو اليمين)

التواء سالب (توزيع ملتو نحو اليسار)

التواء يساوي الصفر (توزيع متماثل)

والالتواء يعطي فكرة أيضا عن تمركز قيم البيانات، فإذا كانت هذه البيانات تتمركز حول قيم البيانات الصغيرة أكثر من تمركزها حول قيم البيانات الكبيرة فإن هذا التوزيع يعتبر ملتويا نحو اليمين ويسمى موجب الالتواء، وعليه عندما يكون التوزيع ملتويا نحو اليمين فإن القيم المتطرفة تؤثر على الوسط الحسابي بسحبه نحو اليمين ولتوضيح ذلك أكثر نقول بأن الالتواء الموجب هو الذي يتزايد التكرار فيه بسرعة كلما تزايدت قيم المتغير حتى تصل إلى قمتها ثم تنخفض التكرارات ببطء كلما تزايدت القيم بعد ذلك وبذلك يكون المتوسط> الوسيط>المنوال

أما المنحنى الملتو التواء سالب هو التوزيع الذي تتركز قيمه حول قيم البيانات الكبيرة وبهذا فإن التوزيع يكون ملتويا نحو اليسار وهو الذي يتميز ببطء صعود تكراراتها كلما زادت قيم المتغير حتى تصل التكرارات إلى قمتها عند نقطة معينة ثم تبدأ بالنزول وفيه يكون المتوسط< الوسيط< المنوال كما هو موضح في الشكلين التاليين:



2. تجانس التباين:

يعتبر التباين من مقاييس التشتت التي تعتمد في تحديد صفات مجتمع الدراسة فلا يمكن أن نتصور تساوي الإنتاج مثلا في جميع المؤسسات الصناعية أو تساوي دخل جميع العمال وبهذا فإن التباين يقيس مدى بعد القيم أو قربها من المتوسط، فإذا كانت قيمة التباين صغيرة فهذا يعني أن القيم قريبة من المتوسط والعينة أكثر تجانسا أما إذا كانت قيمة التباين كبيرة فهذا يعني أن القيم متباعدة عن المتوسط مما يدل ذلك على تشتت أكبر وعدم تجانس في بيانات العينة ولتقديره نقوم بحساب الانحراف المعياري وهو الجذر التربيعي للتباين لغرض توحيد البيانات الأصلية وحتى يسهل تفسير البيانات وهو يعبر عن التشتت بنفس وحدة القياس الأصلية ومنه فإن تجانس التباين هو أسلوب احصائي يشير إلى أنّ تباين المجموعات المختلفة في العينة يجب ان يكون متقاربا أو متساويا ويحسب وفق القانون التالي:

$$S^2 = \frac{\varepsilon(x_i - \tilde{x})^2}{n-1}$$

مثال:

$(x_i - \tilde{x})^2$	Xi -X	(القيم)Xi
9	3	17
1	1	15
81	9	23
49	7-	7
25	5-	9
1	1-	13
166		المجموع

$$\bar{x} = \frac{84}{6} = 14$$

$$s^2 = \frac{166}{6-1} = 33{,}2$$

الانحراف المعياري هو جذر التباين $s = \sqrt{33{,}2} = 5{,}76$