

تمهيد

لعلم الإحصاء دور كبير في جميع المجالات والعلوم، مثل العلوم الطبية، علم الاقتصاد، علم الوراثة، الدراسات الفلكية، تكنولوجيا المعلومات والاتصالات وغيرها من العلوم. حيث يزودهم الإحصاء بالأدوات التي تساعد على تحليل المعطيات بشكل علمي دقيق، ومن ثم استخراج النتائج التي يتم على أساسها اتخاذ القرارات السليمة. سنتناول من خلال هذا الفصل، المفاهيم الأساسية لعلم الإحصاء.

1- علم الإحصاء: نشأته وتطوره

نشأ علم الإحصاء في العصور الوسطى لاهتمام الدول بتعداد أفراد المجتمع، حتى تتمكن كل دولة، من تكوين جيش قوي يستطيع الدفاع عنها، في حال وقوع اعتداء من جانب إحدى الدول، وذلك طمعا في التوسع والثروة. وكذلك اهتمت الدول بخصر ثروات الأفراد، حتى تتمكن من فرض الضرائب وتجميع الأموال اللازمة لتمويل الجيش وإدارة شؤون البلاد.

ثمّ توسعت عمليات التعداد والحصر، لتشمل بيانات عن المواليد، والوفيات، والإنتاج والاستهلاك، وبذلك نشأت الحاجة إلى تنظيم وتلخيص هذه البيانات، ووضعها في صورة جداول أو رسم بياني، حتى يسهل الرجوع إليها والاستفادة منها بأسرع وقت ممكن، وقد أطلق على هذه الطرق "علم الدولة" أو علم الملوك ثم "علم الإحصاء".

فالإحصاء كان مقتصرًا على الحقائق الخاصة بالدولة، وهذا ما يدل عليه مصطلح الإحصاء باللغة الأجنبية "statistics" فهو مشتق من كلمة "status" أي الدولة. هذا كل ما يعرف عن علم الإحصاء في ذلك الوقت، حيث كان التحليل الإحصائي يستخدم للوصول إلى نتائج، يستفاد بها في اتخاذ القرارات الخاصة بالدولة.

إلا أن البوادر العلمية للإحصاء كنظرية لم تظهر إلا بداية من القرن الثامن عشر مع تطور نظرية الاحتمالات، التي كان لها دور كبير في تحليل البيانات واتخاذ القرارات المناسبة. وقد امتد التطبيق الإحصائي إلى مجالات العلوم المختلفة (الطب، الزراعة، الأدب، ... إلخ).

في النصف الثاني من القرن العشرين، تطورت الحاسبات الالكترونية وتنوعت أحجامها وقدرتها ودقتها، الأمر الذي ساعد على تقدم علم الإحصاء بشكل كبير.

2- مفهوم علم الإحصاء وفروعه:

لكلمة الإحصاء مفهومان رئيسيان:

المفهوم الأول: مفهوم عام اعتيادي يتصل بفكرة الأرقام، وهو المفهوم الذي يخطر في أذهان الناس مباشرة، عند ذكر كلمة إحصاء.

المفهوم الثاني: هو ذو مدلول علمي، يعبر عن الأصول والمبادئ والطرق العلمية، التي تتبع في الحصول على المعلومات الكمية ودراسة خصائصها، من أجل الوصول إلى نتائج صحيحة لقطع الشك أو مواجهة المستقبل المجهول.

ويمكن أن نعرف الإحصاء، حسب المفهوم الثاني بأنه: ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات الخاصة بمختلف الظواهر، عرضها وتحليلها، وذلك للوصول إلى نتائج مقبولة وقرارات سليمة.

أما الإحصائيات، فهي البيانات العددية المتعلقة بموضوع ما، وبالتالي فهي المادة الخام لعلم الإحصاء.

ينقسم علم الإحصاء إلى فرعين (قسمين) أساسيين هما:

الإحصاء الوصفي: وهو ذلك الفرع من الإحصاء الذي يختص، بجمع المعلومات وتلخيصها وعرضها في صورة مبسطة. والغرض من كل ذلك، هو إعطاء وصفا أوليا للظاهرة المدروسة بدون تحليل معمق.

الإحصاء الاستدلالي: وهو ذلك الفرع من الإحصاء الذي يهتم بطرق الوصول إلى نتائج معينة أو توقعات ما عن المجتمع، من خلال دراسة عينة من ذلك المجتمع، يتم اختيارها بطريقة علمية مناسبة. وعليه فالإحصاء الاستدلالي يهتم بموضوعين هما، التقدير واختبارات الفروض.

3- مصادر جمع البيانات:

المعلومات الكمية هي الموارد الأولية التي يستعملها الإحصائي، لذلك فإن أول ما يهتم به هو جمع هذه المعلومات، ويمكن أن تجمع هذه المعلومات، مباشرة من مصادرها الأساسية أو أن تأخذ من مصادر ثانوية.

3-1- المصادر الأولية: وهي المصادر التي نحصل منها على البيانات بشكل مباشر، حيث يقوم الباحث نفسه بجمع المعلومات الإحصائية من المفردة محل البحث مباشرة، هذه الطريقة لجمع المعلومات تكلف غالبا من الجهد والوقت والمال. إلا أنّ المعلومات أو الإحصاءات التي تجمع من المصادر الأصلية (الأساسية)، أحسن وأفضل من التي تجمع من المصادر الثانوية.

3-2- المصادر الثانوية: وهي المصادر التي نحصل منها على البيانات بشكل غير مباشر، وهي تشمل جميع البيانات والمعلومات الإحصائية المتوفرة، من وثائق ومطبوعات ونشرات إحصائية التي تصدرها الهيئات والدواوين المختلفة، وكذلك الهيئات الدولية ومنظماتها المختلفة، ولهذه الطريقة فوائد متعددة أهمها أنها تؤدي إلى اقتصاد كبير في وقت الباحث ونفقاته، إلا أنها تعاني أيضا من عدد من العيوب أهمها، عدم وجود الدقة المطلوبة.

ويتحدد الأسلوب المتبع في جمع البيانات، حسب الهدف من البحث وحجم المجتمع محل البحث.

❖ أساليب جمع البيانات: هناك أسلوبين لجمع البيانات هما:

أسلوب الحصر الشامل: في هذه الحالة، تتم دراسة كل الوحدات الإحصائية المكونة للمجتمع الإحصائي، أي أخذ المعلومات المراد الحصول عليها مباشرة من الوحدة الإحصائية، ومن مميزات هذا الأسلوب، أنه يوفر حظوظ الحصول على معلومات دقيقة إذا ما توفرت شروط البحث الإحصائي، الشيء الذي يجعل نتائج الدراسة الإحصائية غير مشكوك فيها. غير أن هذه الطريقة، صعبة التنفيذ وتحتاج إلى تكاليف باهظة وجهاز إحصائي كبير ومتخصص.

أسلوب المعاينة الإحصائية: حيث يتم دراسة جزء من المجتمع الإحصائي وذلك بأخذ عينة عشوائية من ذلك المجتمع، ودراسة خواصها واستخلاص المعلومات اللازمة منها، ثم تعميم نتائجها على المجتمع الذي سحبت منه. يتميز هذا الأسلوب، بتقليل الوقت والجهد والتكلفة، لكن ما يعاب عليه هو أنه أقل دقة من الأسلوب السابق، خاصة إذا كانت العينة المختارة لا تمثل المجتمع تمثيلا جيدا.

ومهما تكن طرق جمع البيانات الإحصائية، سواء بالحصر الشامل أو المعاينة، فإنه توجد وسيلتان لجمع البيانات، الاستمارة الإحصائية بالإضافة إلى المقابلة الشخصية.

4- المصطلحات والمفاهيم الإحصائية العامة

4-1- المجتمع الإحصائي:

هو جميع العناصر المشتركة في الصفة، التي يمكن أن يشملها موضوع الدراسة، والتي تهم الباحث في دراسته، ويقسم المجتمع الإحصائي إلى:

- **مجتمع منته (محدود):** وهو الذي يمكن حصر مفرداته. مثلا: عدد سكان مدينة أو طلبة جيجل، أو المساحات الزراعية في الجزائر، أو إنتاج آلة معينة... إلخ.
- **مجتمع غير منته (غير محدود):** وهو الذي لا يمكن حصر مفرداته. مثل: عدد النجوم في السماء، عدد حبات القمح المحصود في مزرعة معينة... إلخ.

4-2- العينة الإحصائية:

هي جزء من المجتمع الإحصائي محل الدراسة، يتم اختيارها بطريقة ما، بغرض دراسة هذا المجتمع، تمثله أحسن تمثيل. مثل: مجموعة من سكان مدينة، مجموعة من طلبة جامعة جيجل، بعض المساحات الزراعية في الجزائر... إلخ.

4-3- الوحدة الإحصائية:

هي كل فرد أو كائن أو عنصر واحد وواحد فقط من المجتمع الإحصائي، يشترط في الوحدة أن تكون خاضعة لتعريف دقيق وواضح، قد تكون الوحدة شيئا حيويا مثل: طالب، موظف، ... وقد تكون شيئا ماديا، مثل: مؤسسة، سيارة... كما قد تكون شيئا معنويا.

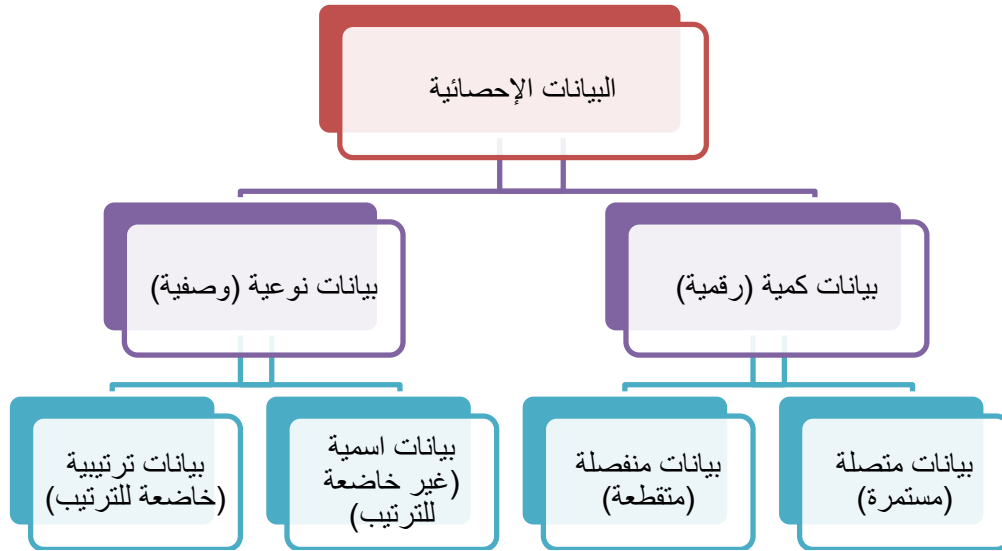
4-4- المتغير الإحصائي:

هو مقدار له خصائص رقمية وغير رقمية (وصفية)، يتغير من عنصر لآخر من عناصر المجتمع أو العينة، بواسطته يمكن للباحث أن يفرق بين الوحدات الإحصائية، يرمز له بالرمز: $x, y, z, \dots$ إلخ.

5- أنواع البيانات (المتغيرات) الإحصائية

يمكن تقسيم البيانات (المتغيرات) الإحصائية إلى مجموعتين هما:

- بيانات كمية (رقمية)
- بيانات نوعية (وصفية)

الشكل رقم (1-4): أنواع البيانات (المتغيرات) الإحصائية**5-1- بيانات كمية (رقمية) variables quantitatives**

وهي البيانات التي تأخذ قيما عددية وتنقسم إلى:

بيانات كمية متقطعة (منفصلة): وهي البيانات التي تأخذ قيما عددية صحيحة، ولا يمكن تجزئة وحدة قياسها إلى وحدات أصغر. مثل: عدد أفراد الأسرة، عدد موظفي الشركة... إلخ.

بيانات كمية مستمرة (متصلة): وهي البيانات التي تعتمد على وحدات القياس التي تأخذ قيم في مجال تغيراتها، يمكن تجزئة وحدة قياسها إلى وحدات أصغر. مثل: الطول، الوزن، الزمن، السرعة... إلخ.

5-2- بيانات نوعية (كيفية) variables qualitatives

وهي البيانات التي تصف ظاهرة من الظواهر، دون أن تأخذ قيما عددية وتنقسم البيانات النوعية إلى:

بيانات نوعية غير خاضعة للترتيب (اسمية): تعتمد على التصنيف النوعي بغض النظر عن أهمية الترتيب. مثل: تصنيف موظفي إحدى الشركات حسب الجنسية أو حسب التخصص، نوع الجنس ذكر أو أنثى. الحالة الاجتماعية: متزوج، أعزب، أرمل، مطلق. المهنة: أستاذ، تاجر، طبيب، إلخ.

بيانات نوعية ترتيبية: وهي البيانات النوعية التي يمكن ترتيبها مثل: ترتيب موظفي إحدى الشركات حسب المؤهل (المستوى العلمي) إلخ.

6- أنواع العينات

يمكن تقسيم العينات وفقاً لأسلوب اختيارها إلى نوعين هما:

1-6- العينات الإحصائية (العشوائية): وهي العينات التي يتم اختيار مفرداتها وفقاً لقواعد الاحتمالات، بمعنى آخر هي التي يتم اختيار مفرداتها من مجتمع الدراسة بطريقة عشوائية. بهدف تجنب التحيز الناتج عن اختيار المفردات، وتنقسم إلى:

- **العينة العشوائية البسيطة:** هي العينة التي نعطي فيها لجميع مفردات المجتمع المراد بحثه نفس الفرصة في الاختيار، وهذا يعني عدم الإهتمام ببعض المفردات أكثر من البعض الآخر وإتاحة الفرص المتكافئة أمام كل مفردة للظهور في العينة. وهناك العديد من الطرق يمكن استخدامها في تحديد مفردات العينة وفقاً لهذه الطريقة منها:
 - إعطاء أرقام متسلسلة لعناصر المجتمع الأصلي، ووضع كل رقم في ورقة منفصلة في وعاء أو كيس، ثم سحب العدد المطلوب من الوعاء بشكل عشوائي.
 - استخدام جداول الأرقام العشوائية، وهي جداول معدة خصيصاً لذلك، حيث يتم من خلالها تحديد المفردات التي ستدخل في عينة الدراسة.
 - استخدام الدواليب، كالتى تستخدم في السحب على جوائز اليانصيب.
- بالرغم من سهولة اختيار العينة العشوائية البسيطة، إلا أنها لا تضمن تمثيل المجتمع في حالة عدم تجانس النسبة للظاهرة موضوع الدراسة، وكذلك يصعب تطبيقها في بعض أنواع البحوث التي لا يمكن حصر جميع عناصر مجتمع الدراسة الأصلي فيها، وكذا ارتفاع تكلفة استخدامها لانتشار أفرادها في مناطق بعيدة.

- **العينة العشوائية المنتظمة:** تستخدم في حالة تجانس المجتمع الأصلي، ويتم تحديدها بإتباع الخطوات التالية:
 - تحديد المجتمع: بحيث نعطي لكل عنصر من عناصر المجتمع الأصلي رقم متسلسل $\{1, 2, \dots, N\}$.
 - تحديد مجال اختيار المفردة الأولى في العينة: وهو من $\{1, \dots, K\}$ حيث: K (كسر المعينة)، $K = \text{حجم المجتمع} / \text{حجم العينة}$.
 - إيجاد باقي المفردات بحيث تكون بصورة منتظمة بإضافة القيمة K .
 - المفردة الأخيرة في العينة = المفردة الأولى + $[(\text{حجم العينة} - 1) \times K]$.

لحساب قيمة أي مفردة في العينة يكفي فقط معرفة رتبها ثم نطبق القاعدة:

$$\text{قيمة المفردة المراد حسابها} = \text{قيمة المفردة الأولى} + (\text{رتبة المفردة المراد حسابها} - 1) \times K$$

فمثلاً: إذا كان لدينا مجتمع يتكون من 1200 علة كبريت أنتجها أحد المصانع، وأردنا اختيار عينة مكونة من 40 علة كبريت من هذا المجتمع، فإنه يمكن إجراء هذا الاختيار بطريقة منظمة وسريعة، وذلك كما يلي:

1- تحديد المجتمع: وذلك بتقييم عناصر المجتمع من {1، 2،، 1200}.

2- تحديد مجال اختيار المفردة الأولى في العينة:

$$K = 1200 \div 40 = 30 \text{ ومنه مجال اختيار المفردة الأولى هو من } \{1، 2، \text{ إلى } 30\}.$$

نفرض المفردة الأولى في العينة هي: 23.

3- إيجاد باقي المفردات في العينة، والتي تكون بصورة منتظمة بإضافة القيمة $K = 30$.

$$\{23، 53، 83، 113،، 1193\}$$

4- المفردة الأخيرة في العينة = المفردة الأولى + [(حجم العينة - 1) × K].

$$1193 = [30 \times (40 - 1)] + 23 =$$

■ **العينة العشوائية الطبقية:** إذا كان المجتمع يتكون من مجموعات من المفردات، تتصف بالتجانس داخل كل مجموعة وبالتباين بين المجموعات المختلفة، ويراد أخذ عينة تكون ممثلة بقدر الإمكان لهذا المجتمع، فلا بد من تمثيل هذه المجموعات في العينة، وذلك بتقسيم المجتمع إلى أقسام تعرف بالطبقات، ثم تؤخذ عينة عشوائية من كل طبقة وبذلك نضمن تمثيل العينة لكل طبقات المجتمع.

مثال (1-5): بين كيفية اختيار عينة مكونة من 200 طالب من مجتمع طلبة كلية العلوم الاقتصادية والتجارية وعلوم التسيير البالغ عددهم 5000 طالب. بحيث تمثل جميع الطلبة بصورة مماثلة، إذا علمت أن الكلية تتكون من أربعة أقسام وهي: قسم التجارة 1800 طالب، قسم التسيير 1600 طالب، قسم الاقتصاد 1200 طالب، قسم المالية والمحاسبة 400 طالب.

الطريقة الأولى:

نقوم أولاً بتحديد نسبة كل طبقة في هذا المجتمع كما يلي:

$$\text{نسبة طلبة قسم التجارة في المجتمع} = (5000/1800) \times 100 = 36\%$$

$$\text{نسبة طلبة قسم التسيير في المجتمع} = (5000/1600) \times 100 = 32\%$$

$$\text{نسبة طلبة قسم الاقتصاد في المجتمع} = (5000/1200) \times 100 = 24\%$$

$$\text{نسبة طلبة قسم المالية والمحاسبة في المجتمع} = (5000/400) \times 100 = 8\%$$

ثم نقوم بحساب عدد مفردات كل عينة جزئية في العينة حيث تحافظ كل طبقة على نفس النسبة في هذا المجتمع كما يلي:

$$\begin{aligned} \text{عدد طلبة قسم التجارة في العينة} &= (36\% / 100\%) \times 200 = 72 \\ \text{عدد طلبة قسم التسيير في العينة} &= (32\% / 100\%) \times 200 = 64 \\ \text{عدد طلبة قسم الاقتصاد في العينة} &= (24\% / 100\%) \times 200 = 48 \\ \text{عدد طلبة قسم المالية والمحاسبة في العينة} &= (8\% / 100\%) \times 200 = 16 \end{aligned}$$

بعد ذلك نقوم باختيار مفردات كل عينة جزئية من طبقات المجتمع المقابلة لها بطريقة عشوائية، حيث نقوم باختيار 72 طالبا بصورة عشوائية من طلبة قسم التجارة، وذلك بترقيم الطلبة من 1 إلى 1800، ثم نختار 72 رقما عشوائيا. وبنفس الطريقة نختار 64 طالبا من قسم التسيير، 48 طالبا من قسم الاقتصاد و16 طالبا من قسم المالية والمحاسبة، وبهذا نحصل على العينة المراد تشكيلها والمكونة من 200 طالب.

الطريقة الثانية:

- نقوم بتحديد حجم كل عينة جزئية في العينة كما يلي:

$$\begin{aligned} \text{نصيب طلبة قسم التجارة في العينة} &= (5000/1800) \times 200 = 72 \\ \text{نصيب طلبة قسم التسيير في العينة} &= (5000/1600) \times 200 = 64 \\ \text{نصيب طلبة قسم الاقتصاد في العينة} &= (5000/1200) \times 200 = 48 \\ \text{نصيب طلبة قسم المالية والمحاسبة في العينة} &= (5000/400) \times 200 = 16 \end{aligned}$$

- بعد ذلك نقوم باختيار مفردات كل عينة جزئية من طبقات المجتمع المقابلة لها بطريقة عشوائية، حيث نقوم باختيار 72 طالبا بصورة عشوائية من طلبة قسم التجارة، وذلك بترقيم الطلبة من 1 إلى 1800، ثم نختار 72 رقما عشوائيا. وبنفس الطريقة نختار 64 طالبا من قسم التسيير، 48 طالبا من قسم الاقتصاد و16 طالبا من قسم المالية والمحاسبة، وبهذا نحصل على العينة المراد تشكيلها والمكونة من 200 طالب.

■ **العينة متعددة المراحل (العنقودية):** إذا كان المجتمع يتكون من أقسام متجانسة، نبدأ باختيار بعض هذه الأقسام عشوائيا (كمرحلة أولى)، ثم نختار عينة عشوائية بسيطة من كل قسم من هذه الأقسام (كمرحلة ثانية)، وقد يحتاج الأمر إلى اختيار عينة عشوائية بسيطة، من كل قسم من الأقسام التي تم اختيارها، في المرحلة الثانية وهكذا، إلخ.

فالعينة متعددة المراحل، عبارة عن مجموعة من العينات العشوائية البسيطة، لسحب مفردات مجتمع دراسة واحد، هذه المجموعة من العينات، لا تقل عن مرحلتين، وتزيد حسب طبيعة الدراسة.

مثلاً: إذا أردنا التعرف على النمط الإستهلاكي لسكان مدينة ما، فيما إذا كانوا يفضلون الأسماك أم اللحوم. فإنه في المرحلة الأولى، يتم الإختيار العشوائي للمديريات والمرحلة الثانية، الإختيار العشوائي للمراكز، ثم المرحلة الثالثة، الإختيار العشوائي للأسر، حسب مساكنهم والمرحلة الرابعة، الإختيار العشوائي للأفراد من هذه الأسر، وبهذا نكون قد شكلنا العينة المراد دراستها والمثلة لسكان هذه المدينة باستخدام العينة العنقودية.

2-6- العينات غير الإحتمالية (غير العشوائية): هي التي يتم اختيار مفرداتها بطريقة غير عشوائية، حيث يقوم الباحث باختيار مفردات العينة بالصورة التي تحقق الهدف من المعاينة. وتنقسم إلى:

■ العينة العمدية.

■ العينة بالمصادفة.

■ العينة الحصصية.

ملاحظة: هناك وجه شبه كبير بين العينة الحصصية والعينة العشوائية الطبقية، إذ أن الفكرة في كليهما قائمة على تقسيم المجتمع إلى مجموعات غير متداخلة ومن ثم سحب عينة تمثل المجتمع، على أن تأخذ عدم تجانسه بعين الإعتبار.

أما الفرق الجوهرى بين هذين النوعين من العينات، يتمثل في عملية اختيار المفردات، إذ يخضع اختيار مفردات العينة الطبقية لمبادئ الإحتمالات (العشوائية)، بينما لا يتحقق ذلك في المعاينة الحصصية.