الفصل الرابع عشر

العينات وطرق انتقائها
Samples and Sampling

14.1. مقدمة Introduction

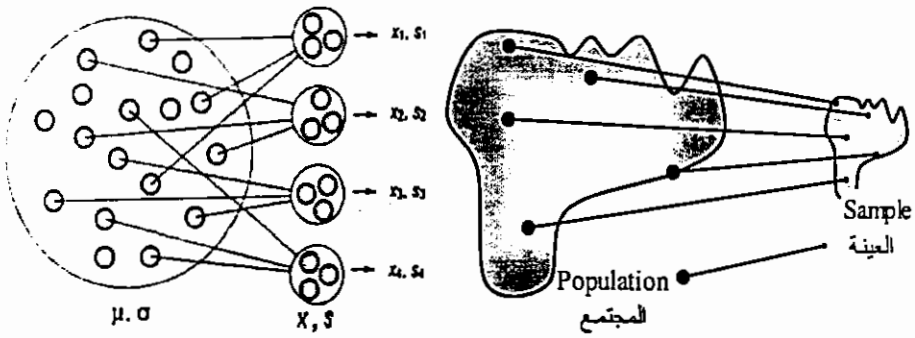
كما ذكرنا فى الأجزاء السابقة من هذا الكتاب وأكدنا على أن الإحصاء الاستدلالي Inferential Statistic يمكننا من التوصل إلى نتائج بشأن المجتمع من خلال دراسة العينات Samples ، ويرتكز بصورة كبيرة على دقة هذه العينات وكيفية أخذها وكذلك على حجمها.

وقبل أن نسترسل فى الموضوع نود أن نورد بعض التعريفات المهمة والتي ستساعدنا على فهم هذا الجزء الهام.

14.2. المجتمع و العينة The Population and The Sample

المجتمع الإحصائى Population: هو جميع العناصر التي تندرج تحت لواء الدراسة الإحصائية، و هذه العناصر قد تكون أفراد أو عائلات أو شركات أو مؤسسات أو غير ذلك.

أما العينة Sample: فهي مجموعة جزئية من المجتمع Population، يفترض أنها تحمل نفس خصائصه وتعبر عنه، وكلما كان حجم العينة كبيراً كانت أكثر تعبيراً عن المجتمع، والشكل 14-1 يعرض هذا المفهوم.



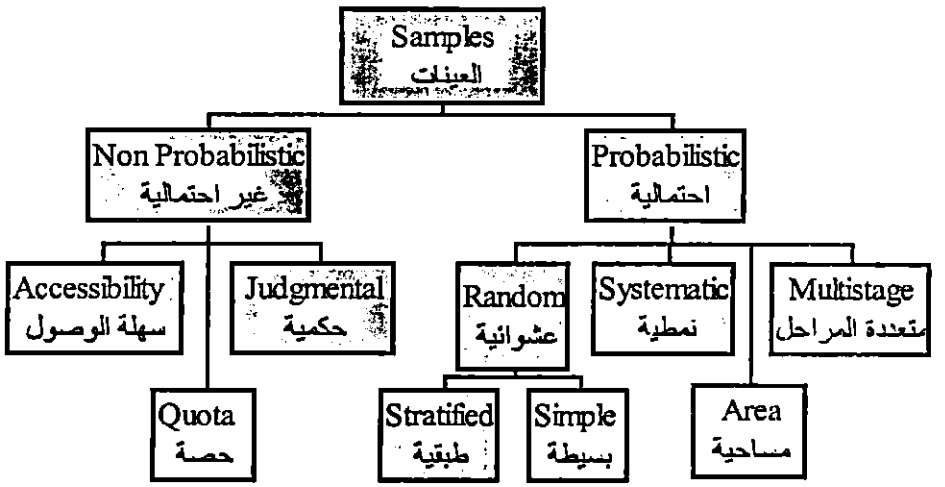
شكل رقم 14- 1 التناظر بين المجتمع والعينة

اما الجدول 14-1 فيعرض مقارنة بسيطة بين كل من العينة والمجتمع:

جدول رقم 14- 1 مقارنة بين العينة والمجتمع

المجتمع Population	العينة Sample	
Parameters المعلمة	Statistic الإحصاءة	المقياس Measures
(μ) mu	$\bar{X}$ (X Bar)	المتوسط Average (Mean)
$\mu = \frac{X_1 + X_2 + X_3 + \ldots + X_N}{N}$	$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$	
$\tilde{\mu}$ Mu til-dah	$\tilde{x}$ X til-dah	الوسيط Median
σ(sigma)	S	الانحراف المعياري Standard deviation
$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$	$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$	
منحنى ناعم Smooth curve	هستوجرام Histogram	التوزيع التكراري Frequency distribution
غالية Expensive	رخيص Cheep	التكاليف Observation Cost
طويل Long	قصير Short	الوقت المستهلك Time consumed
غير دقيق Not Accurate	دقيق Accurarte	دقة القياس Measurement Accurcy
To check if the sample is from the population or not (variation cause identification) لبحث ما اذا كانت العينة تتبع المجتمع ام لا	To infer the population characteristics لاستنباط خصائص المجتمع	التحليل Analysis

وكما أشرنا في الأجزاء السابقة إلى أنه توجد طريقتان لجمع البيانات الإحصائية بهدف دراستها (يرجى مراجعة الفصل العاشر تحت عنوان تقنيات وطرق جمع البيانات) ، وهما الحصر الكامل، وتتم إذا كان من الممكن قياس كافة عناصر المجتمع محل الدراسة، أو أخذ العينات وهو الشائع، وفيه نلجأ إلى دراسة جزء من كل، ويتم ذلك في الحالات التي يتعذر فيها قياس كل عناصر المجتمع، إما لكثرتها أو لصعوبتها أو لاستحالتها، و الشكل 14-2 يجمل الأنواع الرئيسية للعينات ويلية عرض لهذه الأنواع بشيء من التفصيل كما يلي:



شكل رقم 14- 2 الانواع المختلفة للعينات

14.3. العينات غير الاحتمالية Non Probabilistic Sampling

وفى هذا النوع كما يتضح من اسمه فإن العينات المنتقاة منه تتم عمداً، ولذا فهى لا تعبر بصدق وكفاءة عن المجتمع ولا تمثله تماما، ونلجأ له عندما تكون الدراسة المطلوبة بسيطة وسريعة ولا يترتب عليها قرارات هامة، بل تعطى مؤشرات مفيدة عن المجتمع المأخوذة منه، وينقسم هذا النوع إلى عدة أقسام منها:

أ- العينات سهلة الاختيار Accessibility Sampling

وهى اختيار العينة باعتبار الأولوية لسهولة الوصول إليها، فعلى سبيل المثال إذا أردنا دراسة حالات الموظفين المدخنين، فإننا نطلب أى مدخنين متطوعين للإجابة عن الأسئلة التي ستطرح عليهم.

كذلك إذا أردنا تقريرا سريعاً عن جودة منتج ما أثناء الإنتاج والتعبئة فى صناديق، فإننا نأخذ الصندوق الأول من كل مناوبة (وردية).

ب- العينات الحُكمية Judgmental Sampling

وفيها يتم اختيار العينة عن قصد بناءً على حكم شخصي ورأى ذاتي، مثل أخذ عينات من خط إنتاج بواسطة خبير الجودة، والذي يعتمد على خبرته فى انتقاء العينات المعيبة.

ويعتمد مدى تمثيل العينة للمجتمع فى هذه الحالة على خبرة وقدرة الشخص الذي يأخذ العينات، وهى طريقة غير دقيقة إذ تختلف من شخص لآخر، ولكن لها مبرراتها، فمثلاً إذا

كانت شركتك تتعامل مع خمسين مورداً، ونريد دراسة دوافع بعض الموردين لإحجامهم عن التوريد، فمن المنطقي أن يتم اختصار عدد هذه الشركات إلى خمس أو ست شركات، خاصة إذا كان هذا العدد يمثل 90%من حجم التعامل مع هؤلاء الموردون، كذلك تلك العينات المختارة للكشف على مدمني المخدرات فى الأندية الرياضية فيتم بخبرة السلوك والمنظر العام، حينئذ يقال أن هذه العينة هي العينة الحكمية Judgmental Sample أى عينه محددة حكماً.

ج- العينات بالحصة Quota Sampling

وهذا النوع تستخدمه الكثير من المؤسسات الصحفية، ومعاهد استطلاع الرأي، ومن اشهرها معهد جالوب Gallup بأمريكا، والذي يستشف نتائج الانتخابات العامة قبل إجرائها بسرعة وبتكاليف بسيطة، فيُطلب من عدد من الباحثين استطلاع رأى مجموعة من الأشخاص (حصة) فى أحد الأحياء أو المناطق، فيقوم كل باحث بسؤال من يصادفه من الناس فى المكان المحدد له، حتى يتم الحصة المنوطة به أو العدد المحدد له.

14.4. العينات الاحتمالية Probabilistic Sampling

وهذا النوع من العينات يحقق الموضوعية فى الاختيار والبعد عن الذاتية والتحيز إلى حدٍ كبير، فهو يقدم عينة تصلح لتعميم النتائج على المجتمع ، وهي تمكن الباحث من التحكم فى دقتها و قياس هذه الدقة، و كما يتضح من اسمها، فإن العينات المنتقاة يُحتمل أن تمثل المجتمع وتعبر عنه بصدق وكفاءة، إذ أن كل عينة فى هذا النوع من العينات يحظى بنفس فرصة انتقائه كعينة من ذلك المجتمع، ونلجأ لهذا النوع عندما تكون الدراسة المطلوبة هامة وعميقة ويترتب عليها قرارات حاسمة، وينقسم هذا النوع إلى عدة أقسام ومنها:

أ- العينة المنتظمة Systematic Sample

ويتم اختيار العينات فى هذه الطريقة باستخدام المتواليات العددية، فمثلاً لو كان لدينا مجمع يتألف من 100 مفردة، وأردنا أن نأخذ العناصر التي أرقامها الخامس والعاشر والخامس عشر والعشرين والخامس والعشرين وهكذا، أى متوالية حدها الأول خمسة وتزداد بقيمة خمسة، فإن هذه العناصر تمثل عينة منتظمة Systematic Sample، كذلك لو أخذنا العناصر التى أرقامها الثالث والتاسع والخامس عشر والواحد والعشرون وهكذا، أى متوالية حدها الأول ثلاثة وتزداد بقيمة ستة، فإن هذه العناصر تمثل عينة منتظمة Systematic Sample، ويتميز هذا النوع بسهولة وسرعة انتقائه وقلة تعرضه للخطأ والتحيز حتى على الشخص غير المدرب، إذ يكفى تحديد عدد عشوائي واحد لانتقاء كافة العناصر، ويجب أن تغطى العناصر المنتقاة مدى البيانات بالكامل.

ب- العينة العشوائية البسيطة Simple Random Sample

العشوائية لا تعنى أن نأخذ جزءاً من المجتمع بشكل جزافي، بل هي إجراء يُصَمم بدقة بحيث يضمن عدم وجود تحيز من أى نوع قد يؤثر فى عملية اختيار العينة، وبحيث يكون لكل عنصر من عناصر المجتمع نفس فرصة اختياره كعينة، أى أنها عشوائية Random، وعادة ما تستخدم عندما يكون المجتمع محتويا على نوع واحد من البيانات، وبالتالي تكون العينة نوع واحد من البيانات أى أنها بسيطة Simple، وهذه الطريقة تؤدى فى كثير من الأحيان إلى نتائج جيدة ومقبولة لسمات المجتمع وبتكاليف منخفضة نسبيا. (يحيى سعد زغلول و مصطفى عبد المنعم الخواجة، 1993)

جدول رقم 14- 2 جزء من جدول الأرقام العشوائية

0695
0437
6242
7090
0683
7013
8808
9876
1873
2581
3785
8626
0113
4646
7873
3755
2673
0187
7976

ولاختيار عينة عشوائية بسيطة حجمها عشر أرقام وذلك من مجموعة بيانات مكونة من 300 رقم فإننا نقوم بالآتي:

1-نقوم باعطاء أعداد تسلسلية لمجموعة البيانات تبدأ برقم 001 ثم 002 ثم 003 حتى تنتهي برقم 300، مع ملاحظة أن كل عدد من أعداد المجموعة يتألف من ثلاثة أرقام.

2-نستعمل جدول الأرقام العشوائية،(أحد الجداول الإحصائية الشهيرة والهامة وقد تم ادراجه ضمن ملاحق هذا الكتاب)، ونقوم باختيار صفحة منه، ثم نبدأ بالعمود الأول من اليسار ونقرأ الأعداد واحداً تلو الآخر، بحيث نقرأ ثلاثة أرقام، (تم اختيار ثلاثة ارقام لأن أعداد المجتمع الذي بين أيدينا مكونة من ثلاثة أرقام)، فإذا كان العدد الذي نقرؤه من الجدول واقعاً ضمن الأرقام التي قمنا بتجهيزها فإننا نعتبره عنصراً من عناصر العينة، وإلا فنتركه وننتقل إلى الرقم الذي يليه، وهكذا.

3-لا نكرر الأعداد التي أخذناها كعناصر فى العينة أى نرفض أى رقم أخذناه فى قراءة سابقة.

4-نستمر فى القراءة حتى نحصل على حجم العينة المطلوب.

وبتطبيق ما سبق على جزء من جدول الأعداد العشوائية كالمبين بجدول 14-2، فتكون العينة هي القيم التي أرقامها: الرقمين الأولين 695 و 437 نتركهما لأنهما ليسا من ضمن المجتمع لأنهما أكبر من 300، الرقمين 242 و 90 نأخذهما كأفراد للعينة ، الرقم 683 نتركه لأنه أكبر من 300، الرقم 13 نأخذه، وهكذا حتى نحصل على حجم العينة المطلوب.

ج- العينة العشوائية الطبقية Stratified Random Sample

عندما يكون المجتمع محتوياً على أكثر من نوع من البيانات، وبالتالي لكي تكون العينة معبرة عن هذا المجتمع فيجب أن يكون هناك عينة لكل نوع من البيانات، أى أنها طبقية Stratified.

وأول خطوات اختيار العينة العشوائية الطبقية هي تقسيم طبقات البيانات بوضوح، ثم تحديد حجم كل طبقة كنسبة من الحجم الكلى للمجتمع ، ثم تحديد حجم كل عينة من كل طبقة، ومن ثم اختيار عينة عشوائية بسيطة من كل طبقة باستخدام الطريقة التي تم شرحها سابقا، أى أننا تعاملنا مع كل طبقة من البيانات على أنها مجموعة من البيانات المستقلة.

مثال رقم 14- 1:

بلغ عدد المشاريع بإحدى الشركات كما يلي مشروعات كبيرة 120، ومشروعات متوسطة 100، ومشروعات صغيرة 80 والمطلوب اختيار عينة حجمها 60 عينة.

الحل:

حيث إنه توجد أنواع مختلفة من المشروعات فسنستخدم طريقة العينات الطبقية العشوائية Stratified Random Sample.

حجم المشاريع الإجمالى = 120+100+80=300 مشروع

حجم عينة المشروعات الكبيرة =60 *(300/120) =24 عينة

حجم عينة المشروعات المتوسطة =60 *(300/100) =20 عينة

حجم عينة المشروعات الصغيرة =60 *(300/80) =16 عينة

وبالتالي تحولت عملية الانتقاء إلى عينة بسيطة عشوائية Simple Random Sample، أى نختار 24 عينة من 120 مشروعا كبيرا، و نختار 20 عينة من 100 مشروعا متوسطا، ونختار 16 عينة من 80 مشروعا صغيرا. وذلك باستخدام الجداول العشوائية كما فى المثال السابق.

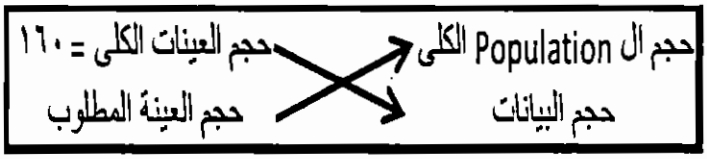
مثال رقم 14- 2: إذا كان لديك مجتمع يحتوى على عدة عناصر كما بالجدول 14-2

جدول رقم 14- 3 بيانات المثال14-2

Data 1	Data 2	Data 3	Data 4	Data 5
500	400	280	200	220

فما هو حجم العينة من كل نوع من البيانات إذا كان الحجم الكلى للعينات هو 160 عينة؟

الحل: حجم المجتمع الكلى = 500+400+280+200+220=1600



حجم عينة 1 = 160*1600/500=50 عينة

حجم عينة 2 = 160*1600/400=40 عينة

حجم عينة 3 = 160*1600/280=28 عينة

حجم عينة 4 = 160*1600/200=20 عينة

حجم عينة 5 = 160*1600/220=22 عينة

د- العينة متعددة المراحل Multistage Sample

عندما يكون المجتمع كبيراً نضطر أحياناً إلى اختيار العينة عن طريق سلسلة من المراحل، وكمثال لذلك نفرض أننا نريد اختيار عينة لتقدير عدد المرضى الذين فحصوا بالأشعة فى أسبوع فى المستشفيات الحكومية فى دولة ما، فى هذه الحالة يصعب بل يستحيل تجميع هذه البيانات من المرضى أنفسهم مباشرة، لذلك نلجأ للمعاينة على مراحل كأن نجرى حصراً بالمناطق الجغرافية التى بها مستشفيات حكومية، ثم نختار عينات عشوائية بسيطة للمناطق حجمها n_1 وهذه هى المرحلة الأولى، ثم نأخذ من كل منطقة من المناطق السابق اختيارها عينة عشوائية أخرى حجمها n_2 من المستشفيات وهذه هى المرحلة الثانية، وبذلك يكون لدينا عدد من عناصر العينات $n_1 * n_2$، ثم نأخذ من كل المستشفيات عينة عشوائية بسيطة من المرضى الذين دخلوها أو كانوا مقيمين فى الأسبوع المحدد وليكن حجمها n_3 مريضاً، وهذه هى المرحلة الثالثة والأخيرة، وبالتالى نكون قد حصلنا على عينة إجمالية حجمها $n_1 * n_2 * n_3$ من المرضى، ونستطيع حينئذ أن نفحص ملفاتهم لاستنتاج عدد الذين فحصوا بالأشعة.

وقد يكون من المفيد أحياناً استخدام التقسيم الطبقي فى واحدة أو أكثر من مراحل أخذ العينات إذا استدعى الأمر ذلك، فتقسم المناطق الجغرافية مثلا إلى مدن صغيرة وكبيرة وقرى، أو تقسم المستشفيات بحسب التخصص، أو يقسم المرضى بحسب الجنس، أى **ننتقى من كل بستان زهرة .**

ه- العينة المساحية Area Sample

العينة المساحية هي تلك التي يتم اختيارها من مساحة، وتستخدم بكثرة فى الدراسات السكانية، والزراعة والجيولوجيا، وقد تعمدت ذكرها للعلم فقط، وفى الواقع فإننى اعتبرها ليست نوعا منفصلا بل هي تطبيق لما تم ذكره، إذ تؤخذ العينات فيها بنفس الطرق سابقة الذكر، فقد تكون منتظمة Systematic، أو تكون عشوائية بسيطةSimple Random ، أو تكون عشوائية طبقية Stratified Random ، ولا يحتاج الأمر إلا إلى تعديل بسيط يساعدنا فى تحديد العناصر التي تدخل فى العينة.

ومن هذه التعديلات رسم خريطة مصغرة للمسطح المعطى، يحدد بمحورين متعامدين أحدهما يعين الشمال والجنوب والأخر يعين الشرق والغرب مع تحديد نقطة أصل مناسبة، و يتم اختيار القطع أو المناطق أو المربعات التي تدخل فى العينة عن طريق الاختيار العشوائي لأزواج من الأعداد تتخذ كإحداثيات للنقط التي تحدد على الخريطة، ومن ثم على المسطح الأصلى، والمعتاد هو اختيار هذه الأزواج من الأعداد عن طريق جداول الأرقام العشوائية الملحقة بنهاية هذا الكتاب.

مثال رقم 14- 3:

لدينا مساحة من الأرض مستطيلة الشكل عرضها 100 متر وطولها 200 متراً، ونريد دراسة بعض الخواص الكيميائية لتربة هذه الأرض فى الحالات الآتية:

- أولا:اختيار عينة منتظمة من 16 نقطة.
- ثانيا:اختيار عينة عشوائية بسيطة من 15 نقطة.
- ثالثا:اختيار عينة طبقية من 15 نقطة إذا علم أن الأرض مقسمة إلى طبقتين بنسبة 2 إلى 3 على وجه التقريب.
- رابعا:اختيار عينة عشوائية من 20 نقطة تؤخذ على ثلاث مراحل.

الحل :

- **أولا:اختيار عينة منتظمة من 16 نقطة.**

نقسم كلا من الطول والعرض إلى أربعة أقسام متساوية البعد، فنحصل على 16 قسماً كل منها مستطيل طوله 50 متراً وعرضه 25 متراً، ثم نحدد عدداً عشوائيا بين 0، 25 وليكن 22، و نحدد عددا عشوائيا بين 0، 50 وليكن 18، فنحصل على إحداثيات النقطة الأولى وهى (22، 18)، ثم نحدد النقاط الأخرى بانتظام بحيث تبعد كل نقطة عن سابقتها بمسافة 25 على المحور الأفقى، و علي المحور الرأسي.

- **ثانيا:اختيار عينة عشوائية بسيطة من 15 نقطة.**

نستخدم جداول الأرقام العشوائية لاختيار 15 عددا عشوائيا يقع بين الصفر، 100 مثلا 74، 10، 11، 44، 16..... ثم نختار عددا عشوائيا يقع الصفر، صفر، 200 مثلا 168، 150، 174، 91، 136..... وبهذا يتحدد لنا 15 زوجا من الأعداد هي (74، 168)، (10، 150)، (11، 174)، (44، 91)، (16، 136).... كل منها يحدد نقطة فى المستوى، فتكون النقط الناتجة هي عناصر العينات المطلوبة.

- **ثالثا:اختيار عينة طبقية من 15 نقطة إذا علم أن الأرض مقسمة إلى طبقتين بنسبة 2 إلى 3 على وجه التقريب.**

نحدد عدد العناصر التى ينبغى سحبها كعينات، بحيث يتناسب حجمهما مع حجمى الطبقتين ، أى نسحب $15*\frac{2}{5}=6$ وحدات من الطبقة الأولى ، $15*\frac{3}{5}=9$ من الطبقة الثانية ، وعلى ذلك نختار 6 أزواج من الأعداد العشوائية (كما سبق فى أولا وثانيا) للطبقة الأولى ، وكذلك 9 أزواج من الأعداد العشوائية للطبقة الثانية.

- **رابعا:اختيار عينة عشوائية من 20 نقطة تؤخذ على ثلاث مراحل.**

نقسم الأرض إلى مربعات متساوية المساحة، وليكن 25 مترا لطول المربع، وعليه فإن الأرض تنقسم إلى اثنين وثلاثين مربعا، ثم نرقم هذه المربعات من 1 إلى 32، ونبدأ بأخذ عينة عشوائية من خمسة مربعات، مثلا المربعات ذوات الأرقام 17، 27، 6، 29، 12 وهذه هي المرحلة الأولى (العناصر الابتدائية)، ثم نقسم كلا من هذه المربعات الخمسة إلى أربعة أجزاء متساوية المساحة ونأخذ من كل مربع جزئين عشوائيا، وهذه هي المرحلة الثانية، وتحتوى على أجزاء كل منها ربع مربع، وأخيرا نختار من كل ربع مربع نقطتين عشوائيا فنحصل على 20 نقطة تمثل العينة المطلوبة.

14.5. إحصاءات العينة Sample Statistics

كما سيتضح لنا عند دراسة التوزيع الاحتمالي لمتغير عشوائي Probability Distributions For Variables (يرجى مراجعة الفصل الثالث عشر تحت عنوان العينات وطرق انتقائها)، فإن أى مجتمع يخضع لتوزيع معين، وهو التوزيع الاحتمالي لمتغير عشوائي يمثل كل عناصر ذلك المجتمع ، ويوجد لكل توزيع احتمالي عدة ثوابت تعين وتحدد هذا التوزيع تحديدا دقيقا وتسمى المعلمات Parameters .

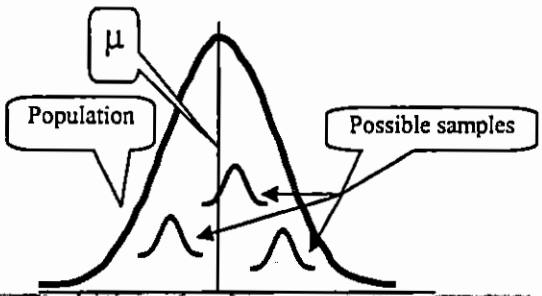
فمثلا إذا كان المجتمع يخضع للتوزيع الهندسي Geometric Distribution فإن المعلمات هي احتمال نجاح المحاولة أو Probability of Success أو (P)، فإذا علمت P فإن

التوزيع الاحتمالي للمجتمع يكون معلوماً تماماً، وحينئذ فإننا نستطيع الإجابة عن أى أسئلة احتمالية حول المتغير X ، والذي يمثل أى عنصر من عناصر المجتمع .

وكذلك إذا كان المجتمع يخضع للتوزيع "تى" T Distribution فإن المعلمات هى درجات الحرية Degree of Freedom فإذا علمت قيمة الدرجات، فإن التوزيع الاحتمالى للمجتمع يكون معلوماً تماماً ، وحينئذ فإننا نستطيع الإجابة عن أى أسئلة احتمالية حول المتغير T والذى يمثل أى عنصر من عناصر المجتمع .

كذلك لو كان المجتمع يخضع لتوزيع طبيعى فإن المعلمات هى الوسط (μ) Mean و التباين Variance σ^2 . فإذا علمت قيمة كل من (μ) و σ^2 فاننا نستطيع الاجابة عن الأسئلة الاحتمالية حول هذا المجتمع ، ويصبح بامكاننا الاجابة عن الأسئلة المتعلقة بالاحتمالات حول ذلك المتغير X.

ولابد أن نفهم جيدا أننا حين نأخذ عينة من مجتمع ما لدراسة بعض المقاييس الرقمية الخاصة بهذه العينة، مثل المتوسط والتباين و الوسيط فإن هذه المقاييس تسمى إحصاءات، كما يجب أن ندرك أنه بتعدد العينات المأخوذة من مجتمع ما فإن قيمة إحصاءات هذه العينات ستتغير، ويتضح ذلك من الشكل 14-3 والذي يوضح اختلاف الوسط الحسابي للعينات المختلفة، وتتوقف قيمة هذا الاختلاف على مدى دقة أخذ العينات .



شكل رقم 14- 3 تعدد العينات الممكنة من المجتمع

وإذا كان حجم المجتمع كبيرا جدا فإننا نلجأ لأخذ عينات للتعبير عن المجتمع وعند حساب المتوسط $\bar{X}$ والانحراف المعيارى S للعينات نجد أنها تختلف عن μ & σ للمجتمع وتعتمد قيمة هذا الاختلاف علي حجم العينات، فكلما زاد حجم العينة قلت نسبة الخطأ Sample Error والعكس صحيح، وكمثال فإن 360 عينة ستكون بدقة Precisions في حدود $\pm 5\%$ ، أي أن المتوسط للعينات سيختلف بمقدار +& - & 5% عن المتوسط للمجتمع .

14.6. الأخطاء التي تتعرض لها العينات Sampling Errors

تتعرض العينات أثناء سحبها إلى العديد من الأخطاء منها ما يلي:

أ- أخطاء التحيز Bias Error

وتتعرض لها البيانات سواء جمعت على أساس الحصر الشامل Total Population Measuring أو على أساس العينات Samples، ويكون لها أثرها الواضح على درجة دقة وشمولية البيانات، وتنشأ عن واحد أو أكثر من المصادر التالية:

- عدم تحديد البيانات التي سيتم تجميعها تحديداً واضحاً.
- القصور فى تحديد المجتمع محل الدراسة، فاستخدام دليل التليفونات كإطار لدراسة خاصة بالرأى العام فى مشكلة تهم المجتمع ككل يؤدى إلى إغفال غيرالحائزين على أجهزة التليفون، وهو ما يؤدى إلى تقديرات معيبة دون شك.
- الإجابات الخاطئة Response Error ومنها الأخطاء الغير متعمدة نتيجة الفهم الخاطئ لأسئلة الاستبيان مثلاً، أو أخطاء السهو والنسيان عند الإجابة عن أحداث وقعت من مدة طويلة، أو أخطاء عمدية عندما تسبب الأسئلة حرجاً اجتماعياً أو تسبب مشكلات قانونية، أو أخطاء ترتبط بالأوقات والتواريخ والظروف المناسبة لجمع البيانات، أو أخطاء منشؤها تبعية الدراسة لجهاز أو جهة محددة، الأمر الذي قد يدفع مقدم البيانات إلى تغيير البيانات بما يتفق مع مصالحه.
- بيانات مبيعات الشركة خلال الخمس سنوات المنقضية قد لا يكون مناسبا لدراسة اتجاه الشركة لفتح أسواق جديدة.

ب- الأخطاء العشوائية Random Error.

وهى الأخطاء الناتجة عن استخدام مبدأ جمع العينات Sampling بدلاً من الحصر الشامل Total Population Measuring، وتعرف بخطأ المصادفة أو خطأ المعاينة وهو يعنى ببساطة الشك فى إمكانية تمثيل العينة لكل خصائص المجتمع، ويتناقص حجم هذا الخطأ بازدياد حجم العينة، كما يمكن تقدير حجم العينة التي تعطى حداً مقبولاً من الخطأ يحدده الباحث.

14.7. معالجة أخطاء العينات Sampling Errors Correction

يمكن التغلب على الأخطاء التي تحدث عادة عند الرغبة فى تجميع البيانات عن طريق العينات ومنها:

- وضع خطة واضحة المعالم، لجمع البيانات، أى من سيجمع ماذا ومتى وأين وكيف؟

- الحصول على البيانات التي تحتمل الخطأ بأكثر من سؤال واحد فى أكثر من موقع واحد فى الاستبيان، ولا نعنى بالطبع تكرار نفس السؤال وبنفس الصياغة.
- إعادة الدراسة على أساس عينة أخرى من نفس المجتمع بعد مرور فترة من الوقت تسمح بتلاشي الأخطاء الناشئة عن تذكر مقدمي البيانات لما أدلوا به من معلومات فى الدراسة الأولى.
- استخدام جامعي بيانات أكثر دقة ومراناً وخبرة.

14.8. تحديد حجم العينة Sample Size Determination

إن تحديد حجم العينة يعد خطوة هامة وأساسية فى مرحلة التجهيز لتجميع البيانات، وكما قلنا مراراً، كلما كبُر حجم العينة كلما زادت ثقتنا فيما نستخلصه من نتائج، ولذلك ينبغي أن نحرص على ألا يكون حجم العينة صغيراً بدرجة تكون معها تقديراتنا أقل مما يجب، وألا يكون حجمها كبيراً بدرجة تثقل كاهلنا بالجهد والتكاليف، وبالتالي فإن الخطوة الأولى فى عملية التجريب هي تحديد الحجم المناسب للعينة، ويتحدد حجم العينة للحصول على قيم الإحصاءة Statistics بخطأ محدد لا نريد أن نتجاوزه بالعلاقة التالية إذا كان توزيع المجتمع يتبع التوزيع الطبيعي:

$$n \geq \left(\frac{2\sigma Z_{1-\frac{\alpha}{2}}}{d} \right)^2$$

حيث d هي قيمة الخطأ الذي نريد ألا نتجاوزه، و حيث $Z_{1-\frac{\alpha}{2}}$ هي المساحة تحت المنحنى والمناظرة لفترة ثقة محددة، وحيث σ هي قيمة الانحراف المعياري وإذا كانت مجهولة يمكن استخدام الانحراف المعياري للعينة S.

14.9. قيود حجم العينة Sample Size Constraints

من أكثر الأسئلة التي تراود الدارس إذا ما قرر الاعتماد على أسلوب جمع العينات فى جمع البيانات المتعلقة بدراسته، ما هو حجم العينة؟ هل يجب أن يكون 5% أم 10% أم 50% من إجمالى المجتمع ؟

والواقع أن هناك احتياطات واعتبارات يَفرض الدارس بعضها ، وبعضها يُفرض عليه حتى يكون متأكدا من أن العينة تمثل المجتمع ، وهذه الاعتبارات تنقسم إلى اعتبارات إحصائية، واعتبارات غير إحصائية، بعدها يتمكن الدارس من تحديد حجم العينة فى ضوء هذه الاعتبارات كما يلي:

أ- الاعتبارات الإحصائية **Statistical Guidelines**

- يدخل مستوى المعنوية Level of Significance الذى يرغبه الباحث فى معادلة رياضية تحدد حجم العينة، كما أنه فى ضوء هذا المستوى يستطيع الدارس قبول أو رفض الفرضية الصفرية Null Hypothesis (يرجى مراجعة الفصل السابع عشر من هذه الكتاب تحت عنوان اختبار الفرضيات) .
- عدد المتغيرات المقاسة ودرجة الدقة التي يبغيها الدارس.

ب- الاعتبارات العامة General Guidelines

- كلما زادت درجة تجانس المجتمع وقل تباينه، كان حجم العينة صغيرا والعكس صحيح.
- تتناسب الميزانية المخصصة للبحث طرديا مع حجم العينة.
- يتناسب الوقت المخصص للدراسة طرديا مع حجم العينة.
- الإمكانيات المتاحة، ومدى توافر جامعي البيانات ذوى الخبرة من عدمه.
- الدواعي الأخلاقية التي تتطلب تخفيض حجم العينة لتقليل الأضرار التي تتعرض لها الوحدات محل البحث، كما فى التجارب التي تجرى على الإنسان، وعلى الحيوان، حيث تقضى المواثيق الدولية بتخفيض حجم العينة إلى أقل حد ممكن يسمح بالتوصل إلى نتائج دقيقة.