العينة العشوائية

1 - المقدمـة :

تعتمد العينات لجمع معلومات حديثة عن منطقة الدراسة واستكمالا للنقص الحاصل في المتوفر من بيانات ، ولكثرة استخدامها فقد تنوعت طرق جمعها . **ويعتمد اختيار الطريقة المناسبة على طبيعة مجتمع الدراسة ، وعلى المتوفر لدى الباحث من مصادر مادية وبشرية و الزمن المتاح ، وعلى درجة الدقة المطلوبة في البيانات و النتائج .** وبالتأكيد فان الطريقة المناسبة لأخذ العينة من مجتمع مكون من (100) أو (1000) وحدة سكنية تختلف عن تلك المعتمدة على مستوى المحافظة ، أو القطـر .

وقبل وصف بعض الطرق من الضروري التأكيد على أن اخذ العينة الممثلة لمجتمعها بصورة علمية يتطلب اعتماد إجراءات مرسومة و محددة مسبقا بدقة و وضوح . فمن السهل على المبتدئين اختيار عينات قد لا تكون ممثلة لمجتمعها ، ولكن ذوي الخبرة والدراية يؤكدون على أنها مهمة صعبة والتدريب عليها ضروري لضمان صحة النتائج وموضوعيتها ، فمعظم الأشخاص معرضون للانحياز بشكل واعي أو غير واعي مما يؤدي إلى اخذ عينات معبرة عن نفسها وليس مجتمعها ، ولهذا تكون الفائدة منها محدودة جدا .

لقد أكدت تجارب عديدة صعوبة اختيار العينة بموضوعية ، ففي كتاب كونوي نجد مثالين عمليين : في التجربة الأولى نشرت (1200) حصاة على طاولة وطلب من (12) طالب اختيار ثلاث عينات تضم كل منها عشرة حصى . وقد تم وزن الحصى فجاءت النتائج أعلى من المعدل العام للمجموع وذلك لأن الطلبة قد مالوا إلى اختيار الحصى الكبيرة . أشرت هذه التجربة صعوبة اخذ عينات دون انحياز سواء أكانت العينة حجرا أم بشرا ما لم يلتزم بقواعد اخذ العينة . وفي التجربة الثانية اختير عدد من تلاميذ المدارس وتم توزيعهم إلى فئتين وفق إجراءات العينة العشوائية. أخبرت المعلمات بان المجموعة التي تضم اطفالاً تعاني سوء التغذية ستعطى كمية إضافية من الحليب ، لذا قامت المعلمات باستبدال بعض التلاميذ لاعتقادهم بحاجتهم إلى التغذية . وكان وزن تلاميذ المجموعتين وأطوالهم متساويا . وبعد انتهاء التجربة لوحظ فرق مساو لنمو ثلاثة أشهر فقط وليس ستة أشهر (Conway 1967) . بعبارة أدق ، مغريات الانحياز كثيرة ، الجاذب منها والطارد، **فعلى الباحث أن يعي بعمق أن العينة التي يأخذها هي ليست له ذاتيا أو لذاتها، بل لتمثل مجتمعها وللخروج بعمومية تخدم المجتمع البشري والعلم .والأمانة العلمية أسمى صفة يجب أن يتحلى بها الباحث وأغلى من جميع الأهداف الطارئة ، لأنها العلم وما سواها حالة تطفل وتمسح بأذيال العلم.**

2 - اختيار العينة عشوائيا :

العينة العشوائية البسيطة Simple Random Samples (SRS) ، وهي الأكثر شيوعا في الاستخدام ، تختار من قائمة أسماء أو هيكل العينة بطريقة عشوائية ، أو باعتماد جداول الأرقام العشوائية (الموجودة في معظم كتب الإحصاء) . حسب هذه الطريقة ، لكل

العشوائية . فعلى سبيل المثال ، إذا كان الرقم الأول المختار من جداول الأرقام العشوائية هو (76) فيعني هذا اختيار الاسم الذي يحمل هذا الرقم في هيكل العينة . وجداول الأرقام العشوائية هي قائمة بأرقام لها فرص متساوية للحدوث في أي موقع . وعند احتواء هيكل العينة الأرقام من (1) إلى (1000) فيجب أن تكون قائمة الأرقام العشوائية مكونة من ثلاثة أرقام (من 001 إلى 999) أما إذا انتهت القائمة بالرقم (1001) عندها من الضروري أن تكون قائمة الأرقام العشوائية مكونة من أربع أرقام ، وهكذا . وإذا وردت أرقام عشوائية ليس لها ما يقابلها في هيكل العينة فإنها تهمل . ومن الضروري البدء عشوائيا برقم من جداول الأرقام العشوائية وليس بأول رقم في الجدول .

تفترض نظرية العينات انه في حالة اخذ عينة (فرد) مرتين فإنها تدخل في الحساب مرتان (Dixon & Leach 1978) ، وفي الواقع يؤثر هذا على تمثيل التباين في خصائص مجتمع الدراسة ، لذا يفضل عدم إعادة العينة المختارة (الرقم العشوائي) إلى القائمة (سلة الاختيار) حتى لا تختار ثانية . وفي الجغرافيا تؤخذ العينة لمرة واحدة دون إعادة اختيارها (Conway 1967) .

تسمح إجراءات اختيار العينة العشوائية البسيطة باختيار الحجم المطلوب للعينة، وقد تبرز مشكلة عدم تمييز الأرقام بين الأفراد مما قد يسبب تركزا في جزء من المجتمع دون غيره . مع هذا فهذه الطريقة هي الأسهل والأسرع وهي معيارية تقاس فاعلية الطرق الأخرى على أساسها . ولهذا أيضا جرت بعض التعديلات عليها ، وتم مزجها مع الطرق الأخرى لجمع العينات .

3 - عوامل تؤثر على تقدير خصائص المجتمع :

تتأثر تقديرات خصائص مجتمع الدراسة بثلاثة عوامل، هي: حجم مجتمع الدراسة، حجم العينة، ودرجة التباين في الخصائص (قيمتي الانحراف والخطأ المعياريين) .

(أ) مجتمع الدراسة: عندما يكون مجتمع الدراسة محددا Finite، وتؤخذ العينة دون إعادتها إلى (سلة) الاختيار ثانية without replacement حينها تصاغ معادلة تقدير خصائص مجتمع الدراسة بطريقة مختلفة عن تلك المعتمدة في مجتمع غير محدد، أو عند منح فرصة إضافية للعينة لتختار ثانية . يتمثل الفرق بين المعادلتين بوجود نسبة العينة إلى مجتمعها (ن) (f) ، والهدف هو تعديل قيمة التقديرات لتقترب من الواقع. (Kalton 1984)

التعديل = (حجم المجتمع - حجم العينة)\(حجم المجتمع - 1)

$$\text{Finite pop. correction} = (N - n) / (N-1)$$

هذا عندما يعتمد معدل القيم ، أما عند التركيز على نسبة الخاصية قيد الدرس فيكون التعديل :

(1 - ن) (1-f) ، وتشتق (ن) (f) بقسمة تكرار الخاصية على حجم العينة . بعبارة أخرى، تكون معادلة حساب قيمة الخطأ المعياري للمجتمع و/أو عندما تعطى العينة فرصة إضافية :

الخطأ المعياري = جذر (التعديل X مربع الانحراف المعياري \ حجم العينة)

$$SE=\sqrt{(N-n)/(N-1)}*(SD^2/n)$$

أو

$$=\sqrt{(1-f)}*(SD^2/n)$$

أما في حالة المجتمع غير المحدد، وعندما لا تعطى العينة فرصة أخرى للاختيار حينها تكون المعادلة بالصيغة الآتية: الخطأ المعياري = جذر (مربع الانحراف المعياري\حجم العينة)

$$SE=\sqrt{SD^2/n}$$

وفي العديد من الحالات العملية يكون مجتمع الدراسة كبيرا، كذلك حجم العينة، لذا تكون نسبة العينة صغيرة . في مثل هذه الحالات فان إعطاء فرص إضافية للعينة للاختيار ثانية من عدمه لا يشكل فرقا لصغر فرصة إعادة الاختيار. ومن المعتاد إهمال نسب العينة من مجتمعها عندما تكون اقل من (1 من 10) أو حتى (1 من 20) (Kalton 1984) .

(ب) حجم العينة: بديهي أن تقل التباينات بين القيم بزيادة حجم العينة، ولكن ما هو مهم أيضا انه في المجتمعات الكبيرة تفوق أهمية حجم العينة نسبتها . فعينة بحجم (2000) مأخوذة من مجتمع يضم (200) مليون تعطي نتائج بالدقة ذاتها (تقريبا) عند أخذها من مجتمع يضم (40000) بافتراض تشابه التباين بين المجتمعين. لذا يمكن القول بـ**أن الفائدة من اخذ العينات تزداد بزيادة حجم مجتمعها** . وللفائدة نذكر بما قيل في المبحث الأول من هذا الفصل ، عندما يكون حجم العينة اقل من (30) يعتمد تعديل بيسل لحساب قيمتي الانحراف و الخطأ المعياريين (ع - 1) (n - 1) .

(ج) التباين في الخصائص : تتباين مجتمعات الدراسة في الكثير من الخصائص حسب طبيعتها ، دقة تحديدها وهدف دراستها . ويقاس التباين بقيمتي الانحراف و الخطأ المعياريين . وكلما كبر مجتمع الدراسة اقترب توزيع القيم فيه من التوزيع الطبيعي، وهو الأكثر شيوعا في الاستخدام . وقد أشير في المبحث الأول إلى توزيعات (ت) (T) والثنائي وعلاقتهما بحجم العينة والتباين في تقدير معدل أو نسبة (ن) و طبيعة توزيع قيم العينة وبالتالي مجتمعها . وللتوضيح نذكر :

أخذت عينة بحجم (250) من مؤسسات صناعية عددها (1872) وكان معدل عدد العاملين في المؤسسة الواحدة (192ر2) عامل وبانحراف معياري قدره (008ر1) ، فما هو تقدير معدل عدد العاملين في المؤسسة الواحدة في مجتمع الدراسة ؟ وبمعرفة أن (165) مؤسسة تصنع الملابس الجاهزة ، فما هي نسبة صناعة الملابس في منطقة الدراسة ؟

التقدير = معدل العينة ± درجة معيارية × جذر ((1 - (حجم العينة \ حجم المجتمع) ×

(الانحراف المعياري \ حجم العينة))

= 192ر2 ± 96ر1 × جذر ((1 - (250 \ 1872) × (008ر1 \ 250))

= 2ر192 ± 0ر116 = 2ر308 و 2ر076

إذن باحتمالية قدرها (95%) يقع معدل عدد العاملين في المؤسسات الصناعية في منطقة الدراسة بين (2ر076) و (2ر308) عامل .

وفي حالة عدم معرفة حجم مجتمع الدراسة ، وفي حالة عدم إعطاء العينة فرصة ثانية للاختيار تكون النتيجة كما مبين في أدناه :

التقدير = معدل العينة ± درجة معيارية × جذر (الانحراف المعياري \ حجم العينة)

= 2ر192 ± 1ر96 × جذر (1ر008 \ 250)

= 2ر192 ± 0ر1244 = 2ر3164 و 2ر0676

أما إذا اهتم الباحث بنسبة وجود خاصية معينة في مجتمع الدراسة ، تحسب هذه النسبة من قسمة تكرار الخاصية على مجموع العينة و يرمز لها (ن) (f) . وفي المثال أعلاه تكون النسبة المئوية لصناعة الملابس في العينة (165 × 100) \ 250 = 66% ، ويقدر التباين في نسب مجتمع الدراسة باعتماد المعادلة الآتية :-

التباين = التعديل × (حجم المجتمع × نسبة الخاصية × نسبة عدم وجود الخاصية \ (حجم المجتمع - 1) × حجم العينة

$$Var = (1 - f) (NPQ / (N - 1) n)$$

= ((250 × (1 - 1872) \ (34 × 66 × 1872)) × ((1872 \ 250) - 1)

= (1 - 0ر133) × (467750 \ 4200768) = 0ر867 × 8ر98

التباين = 7ر785

الخطأ المعياري = جذر التباين = 2ر890

التقدير = النسبة ± درجة معيارية × الخطأ المعياري

= 66 ± 1ر96 × 2ر79

= 71ر4689% و 60ر5311%

ويحسب التباين في نسب العينة باعتماد المعادلة :-

التباين = التعديل × (نسبة الخاصية × نسبة عدم وجود الخاصية) \ (حجم العينة - 1)

$$Var = (1 - f) * (PQ / (n - 1))$$

= ((1872 \ 250)- 1) × ((34 × 66) \ (1- 250))

= 0ر867 × 9ر012

التباين = 7ر813 الخطأ المعياري = 2ر795

التقدير = 66 ± 1ر96 × 2ر795

= 71ر4782% و = 60ر5218%

وعندما تكون قيمة التعديل ضئيلة حينها تهمل ، وتحسب قيمة (ن) (f) بقسمة حجم العينة على حجم مجتمعها ، وقيمة (Q) تحسب من (1 - ن) (1 - f) .

التباين = (نسبة الخاصية × نسبة عدم وجود الخاصية) \ حجم العينة Var = PQ / n

= (66 × 34) \ 250 = 8ر976 الخطأ المعياري = 2ر995

التقدير = 66 ± 1ر96 × 2ر995

= 71ر8702% و = 60ر1298%

ولنعد الآن إلى صيغة المعادلة الأولى ونرى الفرق في النتيجة :

التقدير = 66 ±1ر96 × جذر (1 - (250 \ 1872)) × ((66 × 34) × 249)

= 71ر5% و = 60ر5%

إذن باحتمالية قدرها (95%) تكون نسبة صناعة الملابس في منطقة الدراسة بين (60ر5%)

و (71ر5% . قارن بين قيمة الخطأ المعياري في المعادلات أعلاه : 5ر46 ، 5ر47 ، 5ر87 و 5ر5 . إلى أية درجة من الدقة ترغب في ان تكون النتائج في بحثك ؟

4 - اختيار العينة نظاميا :

الطريقة الأخرى البسيطة المعتمدة لجمع العينات والتي تضمن تغطية شاملة لهيكل العينة هي الطريقة النظامية Systematic Sampling ، فبدلا من اخذ الأرقام عشوائيا تختار وبينها مجال فاصل محدد مسبقا . ويتم تحديد المجال الفاصل بنسبة حجم مجتمع الدراسة إلى حجم العينة K=N/n فإذا كان حجم مجتمع الدراسة (5000) و حجم العينة المطلوبة (50) حينها يكون المجال الفاصل بين عينة وأخرى (5000 \ 50 = 100) ويختار الرقم الأول عشوائيا بين الرقم (1) و (100) وليكن فرضا الرقم (73) وبهذا تختار العينات التي تحمل الأرقام (73 ، 173 ، 273 ، 373 ، ..... ، الخ) . وتعامل قائمة الأرقام بصيغة مدورة Circular حيث يأتي الرقم الأول بعد الأخير لإكمال العدد المطلوب من العينة .

هناك أسباب عديدة تجعل العينات النظامية لا تعامل بالصيغة التي تتبع مع الطريقة العشوائية البسيطة ، ومنها :-

(1) لا تعطي الطريقة النظامية فرصا متساوية لأفراد المجتمع لتحدد الفرص بالمسافة الفاصلة ،

(2) اختيار العينة غير مستقل ، فاخذ العينة الثانية مرتبط بالعينة الأولى ، واختيار الثالثة متعلق بالثانية ، وهكذا (Dixon & Leach 1978) .

(3) لتنظيم هيكل العينة اثر كبير في لا عشوائية الاختيار وانحيازه ، فقد تكون قائمة الأسماء مرتبة على أساس مكاني أو مهني أو هجائي أو غيره . فإذا نظمت أسماء الساكنين على أساس مواقعهم في شارع معين وأخذت الأرقام الفردية فقط حينها يعنى ذلك إنها ستكون على جانب واحد من الشارع . وقد تكون منازل هذا الجانب من الشارع متجانسة في المستوى العمراني وتختلف عن منازل الجانب الآخر من الشارع . وقد تنظم بعض القوائم على أساس تدرجي على ضوء خاصية معينة : العمر ، تاريخ التعيين ، اللقب

العلمي ، الرتبة ، الأبجدية فإذا كانت الأسماء مرتبة هجائيا يمكن تصنيفها إلى فئات مباشرة وفي هذه الحالة تؤخذ كعينات طبقية وليس نظامية (Kalton 1984) .

5 - المزج بين العشوائية و النظامية :

بالإمكان تجاوز معظم مشاكل طريقة جمع العينات نظاميا مع الاحتفاظ بايجابيات الطريقة العشوائية Systematic Random Sampling وذلك بدمج الطريقتين مع بعض . قدم دكسن و ليج أكثر من طريقة للمزج بين الاثنين . (Dixon & Leach (1978

(أ) **الطريقة الأولى :** وتتم بتوزيع بدايات عشوائية على هيكل العينة ومن ثم تختار العينات الأخرى على أساس المجال الفاصل (نظاميا) . فإذا كان المطلوب (100) عينة من هيكل يضم (1000) حينها يكون المجال الفاصل (10) . تختار خمس بدايات عشوائيا موزعة على هيكل العينة ، كأن تكون بأرقام : 4 ، 26 ، 64 ، 505 ، 787 ، ثم تؤخذ العينات الأخرى على أساس المجال الفاصل بين هذه الأرقام وكأنها أربعة قوائم مستقلة ، القائمة الأولى تضم الأرقام (4 - 25) ، الثانية (26 - 63) وهكذا . في هذه الحالة ستكون العينات المختارة تحمل الأرقام : 4 ، 14 ، 24 ، 26 ، 36 ، 46 ، 56 ، 64 ، 74 ، ... ، وهكذا .

(ب) **الطريقة الثانية :** وهي مشابهة للأولى إلا أن نقاط البداية تكون ذات قيم واطئة ، كأن تكون بين (1 - 50) ، ولنفترض إنها كانت : 3 ، 17 ، 21 ، 35 ، 44 ثم تختار الأرقام الأخرى على أساس الفاصلة والتي هي هنا (50) . بهذه الطريقة تكون الفاصلة وكأن قيمتها تساوي (10) ويتم توزيع العينة على الهيكل بصورة متوازنة تقريبا .

(ج) **الطريقة الثالثة :** إن توزيع العينة على مجال هيكلها له ميزته الايجابية المفضلة على الطريقة العشوائية البسيطة إلا انه يبطئ عملية الاختيار ، وهو أكثر فائدة عند اختيار العينة مكانيا . تحدد نقاط البداية اولاً ، ثم أما إن تضاف الفاصلة أو تطرح من الرقم المختار نظاميا وذلك باعتماد القرعة (وجهي عملة معدنية ، الأول إضافة والثاني طرح) .

(د) **الطريقة الرابعة :** لما كانت الفاصلة بين عينة وأخرى تحرم الأفراد بين العينتين من فرص الاختيار ، لذا توجه الباحثون إلى تقليل هذا الإجحاف وزيادة فرص الاختيار وذلك بقسمة الفاصلة على (2) يضاف ناتج عملية القسمة إلى الرقم المختار نظاميا و يطرح في الوقت نفسه ، أي تختار بدايات عشوائيا ومن ثم تطرح الفاصلة من كل رقم من الأرقام المختارة ، بعد ذلك تضاف الفاصلة إليه. فإذا كان الرقم المختار (135) والفاصلة قبل القسمة تساوي (20) فيعني هذا اختيار الأرقام (125، 135 ، 145).

(هـ) **الطريقة الخامسة :** وقد تكون بعض الهياكل على شكل مجاميع مقسمة نظاميا ، لذا يمكن أن تؤخذ العينة عشوائيا من كل مجموعة . أما عندما لا تكون المجاميع متساوية في العدد (الحجم) فالمجموعة الأصغر تضاف إليها بطاقات خالية لإكمال العدد . وإذا حدث أن

اختيرت إحدى هذه البطاقات الخالية فلا يجوز احتيار بديل عنها . وبهذه الطريقة تتساوى فرص الاختيار بين جميع أفراد المجتمع .

اختيرت إحدى هذه البطاقات الخالية فلا يجوز احتيار بديل عنها . وبهذه الطريقة تتساوى فرص الاختيار بين جميع أفراد المجتمع .

اختيرت إحدى هذه البطاقات الخالية فلا يجوز احتيار بديل عنها . وبهذه الطريقة تتساوى فرص الاختيار بين جميع أفراد المجتمع .