مـــركـــز الإحـــصـــاء
STATISTICS CENTRE

دليل المعاينـــــة الإحصـــائيـــة

أدلة المنهجية والجودة- دليل رقم (1)

قائمة المحتويات

المقدمة.....3

الفصل الأول: مفاهيم أساسية في العينات.....4

الفصل الثاني: أساليب المعاينة الإحصائية.....9

1. أسلوب المعاينة العشوائية البسيطة.....9
2. أسلوب المعاينة العشوائية المنتظمة.....10
3. أسلوب المعاينة المتناسبة مع الحجم.....10
4. أسلوب المعاينة العشوائية الطبقية.....11
5. أسلوب المعاينة العنقودية.....17
6. أسلوب المعاينة متعددة المراحل.....18

الفصل الثالث: تقدير حجم العينة.....19

1. تحديد المؤشرات اللازمة لحساب حجم العينة.....19
2. التقدير المسبق لتباين المجتمع.....19
3. اختيار المتغير المناسب لتقدير حجم العينة.....21
4. تقدير حجم العينة لتقدير مؤشر النسبة (p).....21
5. تقدير حجم العينة لتقدير المتوسط (μ).....23
6. تقديرات حجم العينة في المعاينة الطبقية.....24
7. دور أثر التصميم في تقدير حجم العينة بالمعاينة العنقودية.....26

الفصل الرابع: أساليب المعاينة في مركز الإحصاء –أبو ظبي.....28

1. أطر المعاينة في مركز الإحصاء – أبو ظبي.....28
2. تصميم عينات المسوح الإحصائية.....31

المراجع.....34

المقدمة

يأتي إعداد دليل المعاينة الإحصـــائية في مركز الإحصـــاء – أبو ظبي، ضـــمن إطار عمل إدارة البحث الإحصـــائي والمنهجيات ومعايير الجودة على توثيق الأدلة الخاصة بالعمليات الإحصائية. ويهدف هذا الدليل الى اطلاع الاحصائيين الفنيين داخل المركز، ومســـتخدمي البيانات خارج المركز على تفاصـــيل إجراءات تصـــميم وســـحب عينات المســـوح الإحصائية التي ينفذها المركز سواء كانت هذه المسوح اقتصادية أو أسرية أو غيرها.

يتضمن هذا الدليل أربعة فصول رئيسية، الأول يستعرض المفاهيم والمصطلحات الإحصائية الخاصة بأساليب وطرق المعاينة الإحصـــائية، تأتي هذه المفاهيم من واقع الاســـتعراض المرجعي لنظرية العينات ومن واقع أدلة ومعاجم المصطلحات الإحصائية المعتمدة لدى المنظمات الدولية والإقليمية.

اما الفصل الثاني، فيستعرض أساليب وطرق المعاينة الإحصائية المختلفة، هذا إضافة الى محددات ومزايا استخدام كل منها ومعايير اختيار الأســـلوب الأمثل للمعاينة الاحصـــائية. في حين يتناول الفصـــل الثالث تقديرات حجم العينة بحســـب أســـاليب تصـــميم المعاينة المســـتخدمة كما ويعرض المتطلبات الرئيســـية التي يجب توفرها لتحديد حجم العينة.

أخيرا يتناول الفصل الرابع وصفا إحصائيا لأطر المعاينة الإحصائية المعتمدة في المركز سواء كانت هذه الأطر اقتصادية او أسرية او غيرها. إضافة الى تصاميم المعاينة الإحصائية المستخدمة في المركز وآلية تحديد أحجام العينات، إضافة الى الطرق والأساليب العشوائية المستخدمة لسحب العينات.

الفصل الأول: مفاهيم أساسية في العينات

يتناول هذا الفصل المفاهيم الأساسية والتعاريف التي تتعلق بالجوانب النظرية والتطبيقية الخاصة بتصميم وسحب العينات، وهي تنسجم مع المفاهيم والتعاريف الدولية في هذا المجال:

المجتمع الإحصائي

جميع الوحدات الإحصائية التي يراد إجراء البحث الإحصائي عليها، ومن الضروري تعريف هذه الوحدات بشكل واضح بحيث تجمعها صفة واحدة أو صفات مشتركة. ومعظم المجتمعات الإحصائية مؤلفة من وحدات إحصائية تتغير حسب الزمن (مجتمعات متجددة)، وبعضها الآخر مجتمعات ثابتة لا تتغير حسب الزمن.

المسح الإحصائي

هو عمل إحصائي منظم مبني على أسس علمية ويقوم على مبدأ شمول جزء من المجتمع الإحصائي وتختار المفردات في الغالب باعتماد أحد أساليب المعاينة الاحتمالية، أو شمول جميع وحدات المجتمع واخضاعها للمشاهدة من خلال المسح الشامل.

المسح الشامل

هو العمل الإحصائي المنظم الذي يقوم على مبدأ الشمول لجميع مفردات المجتمع الإحصائي بعملية جمع البيانات واخضاعها للمشاهدة الإحصائية، وعادة يجري العد الشامل في التعدادات كالتعداد السكاني والتعداد الزراعي والتعداد الصناعي، وأحيانا يتم العد الشامل إذا كان المجتمع المنوي دراسته مجتمع صغير حيث يكون أسلوب المعاينة غير فعال. أيضا إذا كان ليس لدى الإحصائي خلفية واضحة حول طبيعة المجتمع من الممكن أن يجري عد شامل له بدلا من اسلوب المعاينة.

أسلوب المعاينة

هو أسلوب يستخدم لاختيار مفردات من المجتمع واخضاعها للعمل الاحصائي، بحيث تكون النتائج التي يتم التوصل اليها بناء على معطيات العينة تمثل مؤشرات المجتمع المراد تقديرها.

الاختيار العشوائي

هي عملية اختيار مفردات من المجتمع الإحصائي بطريقة تبعد أي تحكم شخصي للتدخل في اختيار او استبعاد أي مفردة من مفردات المجتمع، مع ضمان إعطاء فرصة متساوية للمفردات كافة لأن تظهر في العينة المنتقاة.

إطار المعاينة

قائمة أو سجل يشمل جميع وحدات المجتمع الإحصائي، ويتضمن عادة أسماء وعناوين الوحدات الإحصائية وبعض المعلومات المتعلقة بها، يمكن أن يكون الإطار مجموعة قوائم او خرائط تساعد في الوصول إلى الوحدات الإحصائية لجمع البيانات منها.

تصميم العينة

تصميم العينة هو خطة محددة للحصول على عينة من مجتمع معين، يشير تصميم العينة إلى الأسلوب أو الإجراء الذي سيتبعه الإحصائي في اختيار عناصر العينة.

العينـة

هي مجموعة جزئية من المجتمع الإحصـائي يتم اختيارها وفق أحد أســاليب المعاينة الإحصـائية ويشــترط أن تكون العينة ممثلة لمجتمع الدراسة، ولكي تكون العينة ممثلة للمجتمع يجب أن تتضمن خصائص المجتمع بشكل يمكن من تعميم نتائجها لتقدير أهم معالم المجتمع الإحصائي.

أنواع العينات

تقسم العينات الاحصائية الى قسمين رئيسيين:

أولاً _ عينات احتمالية:

يتم ســحب العينات الاحتمالية وفقا لقوانين الاحتمالات، بحيث تســحب مفرداتها بشــكل متتالي وباحتمال معروف، ومن أنواعها العينة العشــوائية البســيطة Simple Random Sample، العينة الطبقية Stratified Sample، العينة المنتظمة Systematic Sample، والعينة العنقودية Cluster Sample.

إن اهم ما يميز العينات الاحتمالية انها تمكن من تعميم نتائجها على كافة وحدات المجتمع من خلال حســاب معاملات رفع او اوزان، بحيث يعتمد مقدار الوزن لوحدة العينة على احتمال ســحب تلك الوحدة من المجتمع. أيضا تمكن العينات الاحتمالية من تحليل نتائج العينة وحساب الأخطاء المعيارية، ومعاملات التغير إضــافة الى أثر التصــميم، وبالتالي فإن العينة الاحتمالية تمكن من تقدير نســبة الخطأ ومقدار الثقة في التقديرات الناتجة عنها.

من هنا إن مؤشرات الإحصاءات الرسمية تعتمد بشكل أساسي على تصاميم العينات الاحتمالية بحيث ترفع نتائجها على مستوى المجتمع ككل وتقدر قيم ونسب الأخطاء فيها.

ثانياً _ عينات غير احتمالية:

يتم اختيار وحدات العينات غير الاحتمالية وفقا لمعايير يضــعها الباحث دون التقيد بقوانين الاحتمالات ومن أنواع هذه العينات، العينة العمدية، العينة الحصــصــية، العينة الملائمة، عينة كرة الثلج، الخ. غالبا ما يطبق هذا النوع من العينات في مســوح اســتطلاعات الرأي، وفي دراســات بعض الظواهر المحدودة ضــمن المجتمعات، وهي تعطي نتائج تأشيريه تعتمد على بيانات تمثل وحدات العينة وليس المجتمع ككل.

العينة المتناسبة مع الحجم

يتميز هذا النوع من العينات بأن احتمال ظهور كل وحدة معاينة فيها يتناسب مع حجم تلك الوحدة للصفة المدروسة، فمثلا المنشــأة الاقتصــادية يقاس حجمها بعدد العاملين، وعند اختيار عينة المنشــآت الاقتصــادية بأســلوب المعاينة المتناسبة مع الحجم تعطى احتمالية او فرصة أكبر لظهور المنشآت ذات الحجم الكبير أي ذات عدد العمال الأكبر.

العينة متعددة الأغراض

هي العينة التي يتم من خلالها جمع بيانات لأكثر من موضوع في مسح احصائي واحد مثل الدخل والانفاق والصحة والتغذية.

العينة المتوالية

هو أحد أســاليب المعاينة الذي يغطي المجتمع على عدة ســنوات حيث يقســم المجتمع الى عدة أجزاء، بحيث يغطى كل جزء في سنة، وعادة يستخدم هذا الأسلوب في المجتمعات الصغيرة كبديل عن التعدادات.

العينة الاستطلاعية (البحث التجريبي)

اختيار عدد من الوحدات الإحصائية وجمع البيانات عنها، وتدوينها في استمارات متخصصة لهدف اختبار دقة تصميم الاستمارة الإحصائية، والوقوف على الصعوبات التي قد يوجهها الباحث عند تنفيذ المسح.

العينة المتطابقة

هي العينات التي تتضمن وحدات عينة على شكل ازواج متشابهة او متطابقة بحيث تقاس الصفة المدروسة لهذه العينة مرتين وتحت ظروف مختلفة وبصورة عامة تدعى بالقياسات المكررة.

العينة الموزونة ذاتيا

هو تصميم المعاينة الذي يتضمن أوزان (معاملات رفع) متساوية لجميع الوحدات. أي بمعنى اخر تكون احتمالات اختيار جميع وحدات العينة متساوية.

العينة الملائمة

هي أسلوب معاينة غير احتمالي، يتم من خلاله سحب عينة من مجتمع الدراسة فقط بسبب سهولة توفرها وتوظيفها للدراسة من قبل الاحصائي، وأن الإحصائي لم يأخذ في الاعتبار اختيارها بأن تمثل كامل مجتمع الدراسة.

وحدة المعاينة

هي المفردة او الوحدة التي يتم اختيارها في العينة، وتشكل عنصرا في المجتمع الإحصائي الذي يخضع لعملية العد او عملية المعاينة، أي هي الوحدة التي تجمع عنها البيانات أو المعلومات الإحصائية المطلوبة.

وحدة المعاينة الأولية

هي وحدات المعاينة التي تسحب في المرحلة الأولى من تصميم عينة متعددة المراحل، وغالبا ما تمثل وحدة المعاينة الأولية عنقودا، وهي تشكل مجموعة من وحدات المعاينة الثانوية.

وحدة المعاينة الثانوية

هي وحدات المعاينة التي يتم سحبها في المرحلة الثانية من تصميم عينة متعددة المراحل وتعتبر كل وحدة معاينة ثانوية جزءا من وحدات المعاينة الأولية.

وحدة التحليل

هي الوحدة التي تستخدم في تحليل البيانات الإحصائية التي يتم جمعها لتحقيق اهداف المسح الاحصائي. ومن الممكن ان تكون وحدة التحليل هي ذاتها وحدة العد المستخدمة في المعاينة او قد تكون غير ذلك.

أخطاء عدم التغطية

أخطاء تحصل نتيجة عوامل عديدة اثناء إعداد الإطار من نقص في الشمول او زيادة فيه، او ادخال مفردات غريبة عن المجتمع او عدم توصيف المفردات في الإطار بشكل صحيح. تنقسم أخطاء عدم التغطية الى نوعين: الأول النقص في التغطية أي عدم شمول مفردات ينبغي شمولها. والثاني أخطاء الزيادة في التغطية أي شمول مفردات ينبغي عدم شمولها.

أخطاء عدم الاستجابة

هي أخطاء ناتجة عن عدم اســـتجابة بعض المســـتجوبين في العينة، وتكون عدم الاســـتجابة نتيجة للرفض أو أية أســـباب أخرى، ويندرج هذا النوع من عدم الاســـتجابة تحت اســـم عدم اســـتجابة كلية لوحدة المعاينة، اما عدم الاستجابة الجزئية فهي تكون عندما يرفض المستجوب الإجابة على أسئلة أو متغيرات معينة دون غيرها.

الأخطاء العشوائية

هي انحرافات القيم عن القيمة الحقيقية والتي تعود الى الصـــدفة، ويمكن التعرف عليها من مشـــاهدة انتشـــار البيانات. وهذه الأخطاء لا تختفي عند اســـتخدام أســـلوب الحصـــر الشـــامل وذلك لأنها تنتج عن اختلاف العدادين أو اختلاف الدافع الشـــخصـــي للإجابة على أســـئلة البحث وفي معظم الأحيان يكون مقدارها ضـــئيلا ويمكن قياســـها ومعرفة حدودها. ويتوقف مقدارها على عاملين أساسيين هما مدى الاختلاف أو التباين بين وحدات المجتمع وحجم العينة بالنســـبة للمجتمع الذي ســـحبت منه فكلما ازداد التباين بين وحدات المجتمع ازداد احتمال الوقوع في الخطأ العشوائي، أما بالنسبة لحجم العينة فكلما كبر حجم العينة انخفض احتمال الوقوع في هذا الخطأ.

خطأ التحيز

(أ) خطأ التحيز في التقدير:

وهو انحراف متوسط جميع التقديرات الممكنة لمعلمة المجتمع عن قيمتها الحقيقية، ومن الصعب اكتشاف هذا الخطأ والتخلص منه إلا بأجراء تعديلات جذرية على تصـــميم الدراســـة أو طريقة جمع البيانات أو تعديل النتائج.

(ب) خطأ التحيز في المعاينة:

وهو التحيز الذي يكون مقصودا، وينشأ بسبب الإدلاء بمعلومات لا تطابق الواقع من قبل المزود بالبيانات، أو بســـبب خلل من قبل مصـــممي المســـح وفقا لميول أو أغراض مقصـــودة، وأما أن يكون التحيز غير مقصـــود نتيجة لعدم فهم المستجيب للبيانات المطلوب تقديمها أو لعدم إتاحة الفرصة لتحضير إجابات صحيحة.

الخطأ المعياري

هو الانحراف المعياري لتوزيع مجتمع متوســـطات العينات، بمعنى اخر هو انحراف متوســـطات العينات عن متوســـط المجتمع الذي تنتمي اليه تلك العينات. ويدعى أيضا بخطأ العينة.

الخطأ المعياري النسبي

هو قيمة الخطأ المعياري المطلق لقيم المشاهدات مقسوما على قيمة الاحصاء المحسوبة من تلك المشاهدات، ويدعى بمعامل التغير النسبي.

دقة المسح بالعينة

هو الفرق بين تقدير الصفة المدروسة عن طريق العينة والقيمة الحقيقية للصفة المدروسة في المجتمع، وكلما كان الفرق قليلا كلما كانت الدقة اعلى.

نقطة البداية العشوائية

هو رقم عشوائي يتم اختياره عشوائيا عند استخدام أسلوب العينة المنتظمة، بحيث يقع في الفترة (من 1 الى K) حيث kهو مقدار فترة الانتظام.

التوزيع الأمثل

هو أحد أساليب توزيع وحدات المعاينة الطبقية على مختلف الطبقات بحيث تكون حصة كل طبقة تتناسب طرديا مع حجم الطبقة والتباين داخل الطبقة وعكسيا مع كلفة جمع بيانات وحدة المعاينة في تلك الطبقة.

توزيع نيمان

هو أحد أساليب توزيع المعاينة الطبقية على مختلف الطبقات بحيث تكون حصة كل طبقة تتناسب طرديا مع كل من حجم الطبقة ومقدار التباين ضمن الطبقة الواحدة.

التوزيع المتناسب

هو أحد أساليب توزيع وحدات المعاينة الطبقية على مختلف الطبقات بحيث تكون حصة كل طبقة تتناسب طرديا مع حجم تلك الطبقة من حيث عدد وحدات المعاينة.

الثقة في تقدير العينة

تقيس مدى الاعتماد على النتائج المقدرة عن طريق العينة، وتزداد الثقة في النتائج كلما ازداد حجم العينة، وتقترب من المتوسط الحقيقي أو معلمات المجتمع الأخرى.

حد الخطأ

هو قيمة الخطأ المعياري مضروبة بقيمة Z أو قيمة t الجدولية عند حدود ثقة معينة.

تأثير التصميم

هو نسبة تباين أحد التقديرات في أسلوب تصميم معاينة محدد الى تباين العينة العشوائية البسيطة من نفس الحجم.

الترجيح أو التوزين

هو اجراء حساب الأوزان أو معاملات الرفع لوحدات العينة (وزن وحدة المعاينة يساوي معكوس احتمال سحبها من المجتمع) لاستخدامها في حساب التقديرات الإحصائية بناء على نتائج العينة.

الفصل الثاني: أساليب المعاينة الإحصائية

تتضمن نظرية العينات عدة طرق وأساليب للمعاينة الإحصائية، وجميع هذه الأساليب تركز على الحصول على عينة إحصائية تفرز مؤشرات وتقديرات بأقل خطأ معاينة ممكن، مع الاخذ بالاعتبار أيضا الى حجم العينة المقدر. يتناول هذا الفصل طرق وأساليب المعاينة الإحصائية.

1. أسلوب المعاينة العشوائية البسيطة

يعتبر اسلوب المعاينة العشوائية البسيطة من أبسط الطرق وأكثرها انتشارا في أساليب المعاينة، يمتاز هذا الأسلوب بانه يعطي كل وحده من وحدات المعاينة الموجودة في المجتمع فرص أو احتمالات متساوية للاختيار أو الظهور بالعينة.

هذا ويندرج ضمن المعاينة العشوائية البسيطة اسلوبين لسحب العينة، الأول في حالة سحب العينة وإرجاعها إلى المجتمع بعد عملية السحب لإعطائها فرصة الظهور مرة أخرى ويسمى أسلوب سحب العينة مع الإرجاع، اما الأسلوب الثاني فهو في حال سحب العينة واستثناء كل عينة اخرى مسحوبة من المجتمع يسمى سحب عينة دون إرجاع حيث أن هذا الأسلوب لا يعطي فرصة لتكرار ظهور العينة الواحدة أكثر من مره.

1.1. شروط استخدام العينة العشوائية البسيطة

نظرا لسهولة وبساطة تطبيق هذا الأسلوب من العينات فهو منتشر بشكل كبير، ولكن قبل استخدام هذا الأسلوب يجب ملاحظة أن هناك شروط يجب أن تتوفر في وحدات المعاينة للمجتمع المستهدف في الدراسة، إذ يجب أن تكون وحدات المعاينة في المجتمع متجانسة بالنسبة للصفة المدروسة، أي أن التباين بين وحدات المعاينة في المجتمع للصفة المدروسة قليل نسبيا، هذا بالإضافة إلى المحددات الأخرى من قبل مستخدمي البيانات مثل مستويات النشر، فمثلا إذا كان مطلوب استخراج النتائج على مستوى المنطقة فلا يجوز سحب عينة عشوائية بسيطة على مستوى الإمارة ككل، كذلك وجود محدد التكاليف والإمكانيات الأخرى مثل توفر الكوادر الميدانية المدربة وتكاليف التنقل داخل المجتمع، مما يحدد من استخدام هذا الأسلوب. إضافة الى ما سبق إن تأثر هذا الأسلوب بوحدة تكلفة المسح قد يكون أكثر من معظم أنواع العينات الاخرى، لذا يجب أخذ جانب الحيطة عند استخدام أسلوب المعاينة العشوائية البسيطة.

1.2. طريقة سحب العينة العشوائية البسيطة

عند اختيار أسلوب العينة العشوائية البسيطة، لابد من اختيار طريقة تضمن العشوائية في اختيار وحدات المعاينة، بحيث تعطى كل وحدة من وحدات المجتمع فرصة متساوية للظهور. في الجانب التطبيقي هناك عدة طرق مستخدمة لسحب عينات عشوائية بسيطة، أحد هذه الطرق إعطاء رقم متسلسل لكافة وحدات المجتمع الخاضعة للمعاينة، ومن ثم اختيار رقم عشوائي من جدول الأرقام العشوائية. ومن ثم يتم سحب وحدة المعاينة التي تتطابق مع الرقم العشوائي، اما إذا كانت قيمة الرقم العشوائي خارج إطار تسلسل وحدات المعاينة فيتم إعادة سحب رقم اخر ... وهكذا.

مثال (1):

إذا كان لدينا مجتمع مكون من 60 منشأة واردنا سحب عينة عشوائية بسيطة مقدارها 4 منشات. في هذه الحالة نعطي المنشآت في المجتمع ارقام متسلسلة بدءا من 01، 02، 03، .....، 60. ومن ثم يتم اختيار رقم عشوائي ذو منزلتين عشريتين من جدول الأرقام العشوائية أو باستخدام الحاسوب، وليكن الرقم 45، في هذه الحالة تختار

المنشاة ذات التسلسل 45 ضمن العينة، وفي عملية السحب التالية إذا كان الرقم العشوائي 73 مثلا يتم تجاهل هذا الرقم ويعاد سحب رقم اخر ... وهكذا، حتى نحصل على 4 منشآت تمثل عشوائية بسيطة.

2. أسلوب المعاينة العشوائية المنتظمة

أسلوب العينة العشوائية المنتظمة هو أحد أساليب المعاينة العشوائية التي تمتاز بالسهولة والبساطة في التطبيق، إضافة الى انه يضمن انتشار العينة على أكبر مساحة من المجتمع بسبب أن أسلوب السحب يتم وفق انتظام متسلسل. تعتبر المعاينة المنتظمة الخطية هي الأسلوب الأكثر شيوعا في العينات المنتظمة، ويتلخص أسلوب تطبيقها بما يلي:

أفرض أن المجتمع يتكون من N من وحدات المجتمع وأن حجم العينة المطلوب سحبها هو n فاذا ما قسمنا حجم المجتمع Nعلى حجم العينة المطلوب n نحصل على المقدار k حيث nk=N ويعرف احصائيا مقدارk بفترة الانتظام، بعد ذلك يجري اختيار رقم عشوائي يقع بين 1 وk يسمى هذا الرقم برقم البداية العشوائية ويرمز له بالرمز(I)، يكون الرقم المتسلسل للعينة الأولى هو (I) والعينة الثانية هي K+I والثالثة K+I2.. الخ.

مثال (2):

لتقدير عدد عمال المنشآت ضمن إقليم معين يحتوي على 400 منشأة سحبت عينة منتظمة من تلك المنشآت حجمها 16. يكون أسلوب سحب العينة كما يلي:

$$N = 400 \ , \ n=16$$
$$k = N/n = 400/16 = 25$$

يتم سحب رقم عشوائي يقع بين 1 و25 وليكن الرقم 14، بذلك تكون الأرقام المتسلسلة لوحدات المعاينة المسحوبة في العينة هي:

$$14, 14 + 25, 14 + 2 \times (25), 14 + 3 \times (25) \ \ldots$$

وتساوي: 14، 39، 64، 89، ...

2.1. مزايا ومحددات استخدام العينة العشوائية المنتظمة

من أهم مزايا العينة المنتظمة هو سهولة سحب العينات، وتوزيع العينة على المجتمع بشكل جيد، وتعتبر العينة المنتظمة فعالة مقارنة بالعينة العشوائية البسيطة لكثير من المجتمعات وخاصة إذا كان ميل الصفة المدروسة في المجتمع خطي.

أما من محددات هذه العينة هي صعوبة الحصول على تقدير متحيز للتباين، كذلك في حالة وجود صفة دورية في المجتمع قد يؤدي إلى وجود التحيز في العينات المسحوبة وفي التقدير فمثلا ضمن مجموعة من الأسر تم سحب عينة فرد واحد من كل أسرة وكان رقم البداية هو 1 وفترة العينة هي 2 وكان الأفراد الذي تسلسلهم 3 هم المختارين في العينة ، سنلاحظ أن جميع أفراد العينة هم ذوو الترتيب الثالث في الأسرة ، واذا كان ترتيب الأفراد في الأسرة الأب ، الأم ، الابن ، الابنة ، سنلاحظ أن العينة جميعها أبناء مما يؤدي الى تحيز في التقديرات المطلوبة.

3. أسلوب المعاينة المتناسبة مع الحجم

كما ذكر سابقا، إن أهم ما يميز العينة العشوائية البسيطة هو أن احتمالية ظهور كل وحدة من وحدات المعاينة للصفة المدروسة تكون متساوية لجميع وحدات المجتمع، فعلى سبيل المثال عند سحب عينة من المنشآت الاقتصادية بأسلوب المعاينة العشوائية البسيطة، تكون جميع المنشآت وبأحجامها المختلفة سواء كانت متناهية الصغر أو كبيرة لها نفس الفرصة في الظهور بالعينة. عند دراسة كثير من الصفات بمختلف الظواهر قد تتطلب طبيعة الصفة المدروسة إعطاء احتمال أكبر لظهور وحدات معينة من المجتمع في العينة المسحوبة. فمثلا قد تكون طبيعة

الصفة المدروسة في المنشآت الاقتصادية تتطلب إعطاء فرصة أكبر لاختيار المنشآت ذات الحجم الأكبر (عدد العمال الأكثر)، في هذه الحالة يتم اتباع أسلوب المعاينة المتناسبة مع الحجم.
كذلك يعرف أسلوب سحب العينة المتناسبة مع الحجم بأسلوب التجميع التراكمي، ويتلخص هذا الأسلوب بإعطاء كل وحدة معاينة رقم يوازي الصفة التي تحملها.

مثال (3):
ضمن إطار المنشآت الاقتصادية، المنشأة التي تتضمن 1000 عامل يعتبر وزنها 1000 أي أنها تحتوي على 1000 وحدة فرضية والمنشأة التي تتضمن 100عامل يكون وزنها 100 وهكذا.
وللتوضيح من خلال الجدول التالي:

المنشأة	عدد العمال	عدد العمال التجميعي	الأرقام التي ترافقها
1	1000	1000	1000- 1
2	700	1700	1700-1001
3	1200	2900	2900-1701
4	500	3400	3400-2901
5	300	3700	3700-3401
6	800	4500	4500-3701

لاختيار ثلاث منشآت يتم سحب ثلاثة أرقام عشوائية تقع بين 1و4500 باستخدام الجدول العشوائي او الحاسوب، ومن ثم يتم حصر الرقم العشوائي على عمود عدد العمال التجميعي في الجدول أعلاه ويتم اختيار المنشأة التي يكون فيها عدد العمال التجميعي أكبر من او يساوي الرقم العشوائي. ففي هذا المثال إذا اختيرت الأرقام العشوائية 75،2000،4000 فإن منشآت العينة تكون المنشآت ذات الارقام 1،3،6. على التوالي.
عندما يكون المجتمع كبير نسبيا فإن عملية سحب العينة بالطريقة أعلاه تستغرق وقتا طويلا، لذا هناك أسلوب أو طريقة أخرى لسحب العينة المتناسبة مع الحجم ، وتعرف بطريقة لاهير (Lahiri) ، حيث يتضمن هذا الأسلوب سحب أزواج من الأرقام العشوائية، يمثل الرقم الأول من كل زوج رقم وحده المعاينة بحيث يسحب الرقم العشوائي بين 1 وNg ، حيثN تمثل عدد وحدات المعاينة الإجمالية في المجتمع ويمثل الرقم الثاني في الزوج حجم وحدة المعاينة بحيث يسحب الرقم العشوائي الثاني بين 1وM حيث M تمثل حجم اكبر وحدة معاينة موجودة في المجتمع للصفة التي يتم الوزن على أساسها.

4. أسلوب المعاينة العشوائية الطبقية

كما بينا سابقا، يعتمد تطبيق أسلوب المعاينة العشوائية البسيطة على شرط التجانس بين جميع وحدات المجتمع وغير ذلك فإن العينة ستؤدي الى نتائج متحيزة وغير دقيقة، أضف الى ذلك أن العينة ستنتشر بطريقة عشوائية على منطقة واسعة من المجتمع مما يؤدي الى زيادة حجم العبء والكلفة في عملية جمع البيانات. ونظرا لصعوبة تحقق التجانس في كثير من المجتمعات، يلجأ الى تطبيق أسلوب المعاينة العشوائية الطبقية، حيث يقسم المجتمع إلى عدد من المجموعات غير المتداخلة، كل مجموعة تكون متجانسة للصفة المدروسة وتسمى طبقة، وذلك بهدف الحصول على نتائج أكثر دقة. فمثلا عند دراسة متوسط دخل الأسرة أو المستوى التعليمي يمكن تقسيم المجتمع إلى ريف وحضر، وللحصول على نتائج جيده واستخدام أسلوب المعاينة الطبقية بفاعلية عالية يجب أن يراعى الدقة وخاصة عند إجراء الأمور التالية:

- تكوين الطبقات
- عدد الطبقات المراد عملها

- حجم العينة في كل طبقة
- تحليل البيانات لتصميم العينة الطبقية

مثال (4):

إذا كان لدينا ثماني منشآت وكان عدد العمال في هذه المنشآت كما يلي:

رقم المنشأة	عدد العمال	رقم المنشأة	عدد العمال
1	3000	5	6000
2	1500	6	1200
3	7000	7	4500
4	2500	8	5000

للحصول على تقدير متوسط عدد العمال في المنشأة الواحدة. يمكن من خلال أسلوب المعاينة البسيطة سحب ثلاث عينات، فاذا كانت العينة المسحوبة هي المنشآت التي أرقامها 1،2،6 فإن متوسط عدد العمال هو 1900 عامل بينما المتوسط الحقيقي للمجتمع هو (3838) أي ما يزيد على ضعف المتوسط المقدر بالعينة، من الواضح وجود خلل نتيجة استخدام العينة العشوائية البسيطة لمجتمع غير متجانس.

عند تطبيق اسلوب المعاينة الطبقية، من خلال تقسيم هذا المجتمع إلى طبقات حسب فئات عدد العمال كما يلي:

الطبقة الأولى عدد العمال (أقل من 3000)		الطبقة الثانية عدد العمال (5000 - 3000)		الطبقة الثالثة عدد العمال (أكثر من 5000)	
رقم المنشأة	عدد العمال	رقم المنشأة	عدد العمال	رقم المنشأة	عدد العمال
2	1500	1	3000	3	7000
4	2500	7	4500	5	6000
6	1200			8	5000

وتم اختيار عينة مقدارها منشأة واحدة من كل طبقة عشوائيا، وكانت العينة هي المنشآت ذات الأرقام (2، 7، 5).

فإن متوسط عدد العمال المقدر = $(\frac{3}{8}\times 1500)+(\frac{2}{8}\times 4500)+(\frac{3}{8}\times 6000)=3938$

يلاحظ أن هذه النتيجة قريبة جدا من المتوسط الحقيقي للمجتمع مقارنة بما افرزته العينة العشوائية البسيطة من متوسط، لذا فإن استخدام العينة الطبقية في مثل هذه الحالات يعتبر ضرورة لابد منها.

4.1. مبادئ تقسيم المجتمع إلى طبقات

عند تقسيم المجتمع إلى طبقات يجب مراعاة النقاط التالية:

- أن يكون حاصل جمع وحدات المجتمع في جميع الطبقات مساوي لمجموع الوحدات في المجتمع، ويجب ألا يكون تداخل بين الطبقات.
- أن تكون وحدات المجتمع داخل كل طبقة من الطبقات متجانسة بالنسبة للصفة المدروسة.
- الحصول على نتائج حسب تقسيمات جغرافية أو أداريه (إقليم، منطقة،) يترتب عليه تقسيم المجتمع الى طبقات جغرافية بحسب مستويات التمثيل المطلوبة.
- مراعاة الطرق العلمية عند تقسيم المجتمع إلى فئات حيث تعتبر كل فئة طبقة

تقسيم المجتمع إلى طبقات:

تعتبر عملية تقسيم قيم وحدات المجتمع إلى فئات كنوع من التقسيم الطبقي للمجتمع، حيث يمكن اعتبار كل فئة من الفئات طبقة، ويهدف هذا الأسلوب إلى تقليل التباين بين مفردات المجتمع للصفة المدروسة، فمثلا إذا كانت

الدراسة عن معدل دخل الأسرة فيمكن تقسيم الأسر إلى فئات حسب مستويات الدخل، وإذا كانت الدراسة عن عدد العاملين في المصانع يمكن تقسيم المصانع إلى فئات حسب عدد العاملين وهكذا.

أما تحديد عدد الفئات المقترح فهناك عدة طرق نذكر منها استخدام القانون التالي:

عدد الفئات = 1+3 Log(n)

مثال (5):

إذا كان عدد المنشآت في قطاع معين 1850 منشأة فإن عدد الفئات التي يمكن أن يقسم المجتمع إليها هي:

$$1+3(\text{Log } 1850) = 11$$

لتقسيم المجتمع الى طبقات يمكن استخدام أسلوب Cochran والذي يأخذ عدد وحدات المجتمع ووزن كل وحدة بالاعتبار، حيث أن عدد الطبقات المطلوب محدد مسبقا. ويمكن توضيح ذلك بالمثال التالي:

مثال (6):

إذا كان عدد المنشآت في المجتمع هو 1850 مقسمة إلى فئات مبدئية حسب عدد العمال إلى 12 فئة والمطلوب إعادة تقسيم هذه المنشآت إلى أربع طبقات فإذا كان لدينا الفئات التالية:

الفئة	عدد المنشآت	عدد العمال (C)	C التراكمي	$\sqrt{C}$
1 - 5	200	720	720	26.8
6 - 10	250	2000	2720	52.2
11-15	300	3900	6620	81.4
16-20	300	5250	11870	108.9
21-25	250	5750	17620	132.7
26-30	200	5400	23020	151.7
31-35	100	3650	26670	163.3
36-40	80	3000	29670	172.2
41-45	80	3360	33030	181.7
46-50	50	2400	35430	188.2
51-55	30	1590	37020	192.4
56 فأكثر	10	950	37970	194.9

نقسم الجذر التربيعي لعدد العمال التجميعي ($\sqrt{C}$) في الفئة الأخيرة على عدد الفئات المقترحة

$$\frac{194.9}{4} = 48.7$$

الطبقة الأولى هي التي تقابل الجذر التربيعي لعدد العمال التجميعي (48.7) وبالتقريب يمكن اعتبار الطبقة الأولى هي 1-10 عمال، وباستخدام نفس الأسلوب نحصل على الطبقات التالية:

- الطبقة الأول 1× 48.7 1-10
- الطبقة الثانية 2× 48.7 11-20
- الطبقة الثالثة 3× 48.7 21-30
- الطبقة الرابعة 4× 48.7 أكثر من 30

4.2. فوائد تقسيم المجتمع إلى طبقات

1. في العينة الطبقية يكون المجتمع متجانسا في كل طبقة ويتم تمثيل المجتمع بشكل جيد حيث تؤخذ عينات من مختلف الطبقات، وخاصة التي قد يكون لبعضها أهمية خاصة.

2. اســـتخدام العينة الطبقية فعال أكثر من العينات الأخرى خاصـــة في حالة وجود مجتمع غير متجانس وفي حالة وجود قيم متطرفة لبعض وحدات المعاينة.
3. تؤدي المعاينة الطبقية إلى تخفيض الكلفة لأنها تخفض من حجم العينة المطلوب تغطيتها عند مســـتوى معين الدقة.
4. يمكن اســـتخدام العينة الطبقية للحصـــول على نتائج على مســـتويات جغرافية او إدارية معينة (أقاليم، قطاع، منطقة ...).
5. تكون السيطرة على العمل الميداني والأشراف عليه وتنظيمه وتحديد منطقة عمل كل مجموعة بشكل أفضل وذلك عند تقسيم المجتمع إلى طبقات حسب المناطق الجغرافية والإدارية.

4.3. توزيع العينة على مختلف الطبقات

هناك عدة طرق لتوزيع العينة الإجمالية (n) على مختلف الطبقات بحيث يكون حجم العينة في الطبقة h هو n_h ومن أهمها:

أسلوب التوزيع المتساوي:

تستخدم هذه الطريقة عادة عند الحاجة للحصول على نتائج على مستوى المناطق الإدارية (في حالة كون الطبقة تمثل منطقة إدارية) أو في حالة توزيع العمل بشـــكل متســـاوي بين جميع الطبقات (حســـب توفر الإمكانيات للعمل الميداني) كذلك عندما يكون عدد وحدات المجتمع متقارب في جميع الطبقات.

في هذه الحالة يكون حجم العينة في الطبقة الواحدة حسب المعادلة التالية:

$$n_h = \frac{n}{L} \quad \ldots\ldots (1)$$

حيث L هو عدد الطبقات.

مثال (7):

إذا كان حجم العينة 2000 وحدة معاينة وعدد الطبقات يســـاوي (8) فإن حجم العينة في الطبقة h يســـاوي:

$$\frac{2000}{8} = 250$$

أسلوب التوزيع المتناسب مع حجم الطبقة:

إن هذا الأسـلوب من أكثر الأسـاليب شـيوعا نظرا لسـهولته، إذ عندما لا توجد معلومات سـوى عدد وحدات المجتمع في كل طبقة فيمكن استخدام المعادلة التالية لتقدير عدد العينات في الطبقة (h)

$$n_h = n(\frac{N_h}{N}) \quad \ldots\ldots (2)$$

حيث N_h هو عدد وحدات المجتمع في الطبقة (h).

مما سبق إن (نسبة العينة في الطبقة الواحدة) متساوية لجميع الطبقات مما يؤدي إلى عينة موزونة ذاتيا، لا حاجة عندها الى حســـاب أوزان المعاينة عند بناء المؤشـــرات والتقديرات. مما يعطي قدرة على إجراء التقديرات بســـرعة ودقة عالية.

مثال (8):

عند تقدير عدد العمال في مجتمع ما سحبت عينة منشآت من كل إقليم بأسلوب العينة الطبقية وباستخدام التوزيع المتناسب وبحجم عينة إجمالي قدرة 35 منشأة. يكون حجم المنشآت في كل طبقة حسب المعادلة (2) أعلاه، على الشكل التالي:

الطبقة (الإقليم)	عدد المنشآت الكلي	حجم العينة في الطبقة
1	200	20
2	100	10
3	50	5
المجموع	**350**	**35**

أسلوب توزيع نيمان (Nyman):

إن أهم ما يميز أسلوب توزيع نيمان هو انه يستخدم للتقليل من حجم التباين وزيادة دقة وكفاءة البيانات، حيث يؤخذ في الاعتبار بالإضافة الى حجم الطبقة عند توزيع العينة الكلية على الطبقات تباين كل طبقة. حيث يكون حجم العينة في الطبقة يتناسب طرديا مع الانحراف المعياري لتلك الطبقة وذلك من أجل جعل تصميم المعاينة أكثر فعالية من أسلوب التوزيع المتناسب، وعادة يستخدم هذا الأسلوب عندما يكون الانحراف المعياري مختلف من طبقة لأخرى، وعندما يكون حجم العينة الكلي ثابت وكلفة العينة ثابتة لمختلف الطبقات. يقدر حجم العينة للطبقة (h) حسب المعادلة التالية:

$$n_h = \frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h} n \quad \ldots\ldots\ldots (3)$$

أما الانحراف المعياري على مستوى كل طبقة فيمكن الحصول عليه من خلال تعداد سابق أو يمكن تقديره من المسوح السابقة او مسوح مشابهة أخرى.

مثال (9):

عند تقدير متوسط دخل الأسرة في أحد الأقاليم، قسم المجتمع إلى ثلاث طبقات حسب فئات الدخل وسحبت عينة من كل طبقة بطريقة توزيع نيمان وبحجم عينة إجمالي قدرة 15 أسرة فإذا كانت لدينا البيانات التالية:

الطبقة (فئة الدخل)	عدد الأسر في الطبقة (N_h)	الانحراف المعياري لكل طبقة (S_h)	$N_h S_h$
أقل من 1000 دينار	30	20	600
1000 - 3000 دينار	50	30	1500
أكثر من 3000 دينار	20	20	1000
المجموع	**100**	**-**	**3100**

إن عدد العينات للطبقة الأولى=

$$n_h = \frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h} = 15 \times \frac{600}{3100} = 3$$

عدد العينات للطبقة الثانية=

$$= 15 \times \frac{1500}{3100} = 7$$

عدد العينات للطبقة الثالثة=

$$= \frac{15 \times 1000}{3100} = 5$$

أسلوب التوزيع الأمثل:

يهدف هذا التوزيع إلى تخفيض التباين لأقل قدر ممكن بكلفة محددة، أو لتقليل الكلفة اقل ما يمكن بمستوى دقة معين؛ حيث يدخل عامل الكلفة في توزيع العينات على الطبقات، وعادة يستخدم هذا الأسلوب عندما يكون تفاوت في كلفة جمع البيانات بين الطبقات فمثلا كلفة جمع البيانات من مناطق معينة أعلى بكثير من كلفة جمع البيانات من مناطق أخرى، وهناك عدة معادلات تستخدم لهذا الغرض نعرض منها المعادلة التالية:

$$n_h = \frac{N_h S_h / \sqrt{C_h}}{\sum_{h=1}^{L} N_h S_h / \sqrt{C_h}} \; n \quad \ldots\ldots\ldots (4)$$

حيث (C_h) تشكل كلفة وحدة المعاينة الواحدة في الطبقة h.

مثال (10):

في دراسة لتقدير متوسط إنتاج الحقل في مناطق مختلفة أخذت عينة طبقية حيث شكلت كل منطقة طبقة وكان عدد العينات الإجمالي هو 200 عينة، وفق المعادلة (4) أعلاه يكون حجم العينة في الطبقة الواحدة كما في العمود الأخير من الجدول أدناه:

الطبقة (المنطقة)	عدد الحقول N_h	الانحراف المعياري للطبقة S_h	$N_h S_h$	كلفة الوحدة C_h	$\frac{N_h S_h}{\sqrt{C}}$	حجم العينة (n)
1	200	10	2000	4	1000	58
2	100	20	2000	6	816	47
3	150	10	1500	6	612	36
4	100	15	1500	9	500	29
5	50	20	1000	12	287	17
6	50	20	1000	20	224	13
المجموع	**650**				**3439**	**200**

أسلوب التوزيع المتناسب مع صفة معينة:

هناك بعض أنواع العينات يتطلب تصميمها وسحبها حسب أهمية متغيرات الصفة المدروسة في العينة، في بعض الحالات تكون الأهمية لصفة معينة والتوزيع على أساسها أهم من عدد مفردات المجتمع ولتوضيح ذلك يمكن أخذ المثال التالي:

إذا كان الهدف تقدير إيرادات المنشآت، في مختلف المناطق (الطبقات) وكان لدينا البيانات التالية:

الطبقة	عدد المنشآت	كمية الانتاج
1	1000	20000
2	500	30000
3	1500	10000

يشكل حجم الطبقة الثالثة بحسب عدد المنشآت فيها 50% من وحدات المجتمع على الرغم من انها تحتوي على 0.1667 فقط من كمية الانتاج، لذلك في حالة ا ستخدام عدد المن شآت لوزن الطبقة فأنه يعطي نتائج م ضللة، حيث يكون نصـــيب الطبقة الثالثة 50% من حجم العينة المســـحوبة، بينما يكون نصـــيب الطبقة الثانية 17% من العينة المســـحوبة على الرغم من أنها تحتوي على 50% من إجمالي كمية الإنتاج والتي لها علاقة بالإيرادات، مما يقلل من دقة وفعالية العينة ويزيد في التكاليف.

لذلك فإن الأسلوب الأفضل هو توزيع العينة على الطبقات حسب كمية الانتاج كونها الصفة المراد دراستها ويتم توزيع العينة بين الطبقات بناء عليها بدلا من عدد المنشآت. وباستخدام المعادلة التالية لتوزيع العينة بالأسلوب المتناسب مع صفة معينة.

$$n_h = \frac{x_h}{\sum_{n=1}^{L} x_h} n \quad \ldots\ldots (5)$$

حيث X_h: الصفة التي تحملها الطبقة (h) وهي كمية الانتاج في المثال اعلاه.

فاذا كان حجم العينة المطلوب هو 100 منشـأة مثلا، بالأسـلوب المتناسـب وحسـب صـفة كمية الإنتاج يكون توزيع العينة على الطبقات كما يلي:

الطبقة	حجم العينة
1	33
2	50
3	17
المجموع	**50**

5. أسلوب المعاينة العنقودية

يقوم أســـلوب المعاينة العنقودية على مبدأ تقســـيم المجتمع إلى مجموعات بشـــكل مناســـب بحيث تكون هذه المجموعات متقاربة بالحجم ومتجانسة بالنسبة للصفة المدروسة، حيث كل مجموعة من هذه المجموعات تسمى عنقود، وتشكل العناقيد المجتمع كاملا دون حذف أو تكرار.

إن اهم ما يميز العينة العنقودية إنها فعالة بالنســـبة لوحدة التكاليف حيث تعطي دقة أكثر لوحدة الكلفة، كذلك يلجأ الى هذا الاســـلوب في كثير من الأحيان خاصـــة في المجتمعات التي لا يتوفر لديها اطر معاينة او يصـــعب توفير إطار حديث بكل مفردة من مفردات المجتمع، ولكن من الممكن توفير إطار بالعناقيد مما يوفر بالجهد والوقت. كذلك يمتاز هذا الأسلوب بإمكانية توفير التكاليف المادية والوقت والجهد اثناء العمل الميداني بين وحدات المعاينة. لكن يجب ألا ننســـى أن من محددات العينة العنقودية إنها اقل فاعلية من العينة العشـــوائية البســـيطة، كونها اقل انتشـــارا. وعند استخدام العينة العنقودية يجب مراعاة ما يلي:

- الموائمة بين عدد العناقيد واحجامها، بحيث يكون حجم العنقود صغير وعدد العناقيد كبير نسبيا.
- عند تكوين العناقيد تؤخذ مفردات المجتمع المتقاربة مكانيا أو ضـــمن منطقة معينة حيث تكون غالبا متشابهة للصفة المدروسة.
- الاتساق في احجام العناقيد، بحيث تكون متقاربة في الحجم ما أمكن.
- يجب أن يكون كل عنقود واضح المعالم والحدود بحيث يمكن تمييزه عن غيره.

6. أسلوب المعاينة متعددة المراحل

إن التحدي الرئيسي في كثير من المسوح هو عدم توفر إطار حديث لوحدات المعاينة الرئيسية مثل المنشآت أو المساكن وغيرها، ويكون من الصعب إعداد إطار حديث لها، وفي نفس الوقت يتوفر قائمة او إطار بمتغير على مستوى تجميعي وليس تفصيلي مثل تجمعات سكانية او مناطق رئيسية، في هذه الحالة يمكن استخدام أسلوب المعاينة متعددة المراحل.

6.1. ميزات العينة متعددة المراحل

- توفير في الوقت والمال حيث يكتفى بإعداد إطار بوحدات المعاينة الرئيسية.
- تتسم العينة متعددة المراحل بالمرونة حيث من الممكن استخدام أسلوب سحب عينات في كل مرحلة مختلف عن المراحل الأخرى.

هذا ويفضل ضمن أسلوب المعاينة المتعددة المراحل تقسيم المجتمع إلى وحدات معاينة رئيسية متساوية وذلك:

- عندما يكون حجم وحدات المعاينة الرئيسية كبير نسبيا يحتاج وقت كبير في إعداد إطار بوحدات المعاينة الثانوية.
- عندما يكون حجم وحدات المعاينة الرئيسية صغير يحتاج إلى وقت في عملية التنقل ميدانيا بين وحدات العينة.

6.2. أسلوب اختيار وحدات المعاينة الرئيسية

- إذا كانت وحدات المعاينة الرئيسية متجانسة يمكن استخدام العينة العشوائية البسيطة.
- إذا وجد تفاوت يمكن تقسيم وحدات المعاينة الرئيسية إلى طبقات وسحب عينة من كل طبقة.
- إذا كان التفاوت كبير يمكن استخدام العينة المتناسبة مع الحجم.
- يمكن استخدام العينة المنتظمة ولكن قد يكون من الصعوبة بمكان الحصول على تقدير غير متحيز لخطأ المعاينة.

أما وحدات المعاينة الثانوية فيمكن اختيارها أيضا بأي من أساليب المعاينة: العشوائية البسيطة، المنتظمة، الطبقية، العنقودية، أو المتناسبة مع الحجم.

الفصل الثالث: تقدير حجم العينة

يتناول هذا الفصل موضوع تقدير حجم العينة اللازم عند تنفيذ مسح احصائي معين، وما هي المتطلبات الأساسية التي يجب توفرها للوصول الى تقدير أمثل لحجم العينة. إضافة الى استعراض الية تقدير حجم العينة وفق طرق وأساليب المعاينة الإحصائية المختلفة.

1. تحديد المؤشرات اللازمة لحساب حجم العينة

إن عملية تقدير حجم العينة تقوم على أساس صيغ ومعادلات رياضية تعتمد على عدد من المتغيرات التي يجب توافرها عند اجراء عملية حساب حجم العينة المناسب. إن هذه المتغيرات هي كما يلي:

1. معرفة مستوى الثقة بالتقديرات التي سوف تبنى بالاعتماد على هذا الحجم من العينة، كأن يكون (90%، 95%، 99% ،..) وهي تمثل احصائيا مساحات التوزيع الطبيعي تحت المنحنى الطبيعي القياسي عندما تكون قيم Z (1.64، 1.96، 2.58 ....) على التوالي. وترتبط درجة الثقة بقيمة التقدير ارتباطا موجبا مع حجم العينة، أي كلما زاد حجم العينة يزاد مستوى الثقة بالتقدير.
2. هامش الخطأ في التقدير، وهو عبارة عن قيمة الخطأ المسموح به أي الفرق بين القيمة الحقيقية والقيمة التقديرية للمعلمة المطلوب إيجاد تقدير لها باستخدام بيانات العينة، هذا ويتناسب حجم العينة طرديا مع مستوى دقة التقدير أي يزداد التقدير دقة (أي يقل الخطأ) كلما زاد حجم العينة.
3. مقدار التباين في وحدات المجتمع الخاص بالمعلمة التي سيتم تقديرها، فاذا كانت قيمة تباين المجتمع غير معروفة للمؤشر المطلوب تقديره فيجب ايجاد تقدير مناسب لها. أحيانا قد يكون الهدف من المسح تقدير أكثر من مؤشر واحد. فذا كان الهدف من المسح تقدير مؤشرات عديدة. لابد من اختيار مؤشر مناسب (indicator key) لتقدير حجم عينة يعتمد عليها في تقدير كافة المؤشرات المطلوبة بمستوى كاف من الدقة.

لابد للباحث هنا من دراسة لاختيار كل من (مستوى الثقة وهامش الخطأ) في حساب حجم العينة، فكلما كانت قيمة هامش الخطأ (أي حد الخطأ المطلوب عند التقدير d) صغيرا، وكلما كان مستوى الثقة بتجاوز الخطأ المسموح به القيمة (d) عالي، كلما احتاج حجم عينة أكبر. إن دراسة هذا الموضوع تكون من خلال دراسة عدة بدائل لكل من مستوى الثقة (z) وهامش الخطأ (d) وعلى أساس ذلك يتم وضع عدة سيناريوهات مختلفة لحجم العينة، لتقوم إدارة المسح بالموازنة بينها وفقا للتكاليف والمستلزمات المادية والبشرية المتاحة.

أما موضوع اختيار مقدار التباين المناسب لتقدير حجم العينة، فيكون وفق أحد الطرق التالية:

2. التقدير المسبق لتباين المجتمع

إن بناء وتركيب معادلات تقدير احجام العينة تقوم على أساس توفر قيمة لتباين المجتمع (σ^2) للمؤشرات الرقمية، من هنا لابد من اجراء عملية تقدير لهذه التباينات. احصائيا هنالك عدة أساليب يمكن من خلالها تقدير تباين المجتمع لأغراض حساب حجم العينة:

2.1. الأسلوب الأول: تجزئة العينة على مرحلتين

ضمن هذه الطريقة يتم تقسيم العينة الى جزئين يتم تنفيذهما على مرحلتين، في المرحلة الأولى يتم سحب عينة عشوائية بسيطة بحجم (n_1) تقدر على أساسها قيم التباين بالمقدارين s_1^2 أو $\widehat{P_1}$. هذا ويستخدم هذ التقدير لحساب حجم العينة النهائي (n).

وفي المرحلة الثانية بعد تقدير قيمة التباين يحســـب حجم العينة اعتمادا على هذا التقدير. يتم ســـحب وحدات المعاينة المتبقية من حجم العينة الكلي.

مثال (1):

إذا كان حجم العينة المتوقع لدراســـة متوســـط إنفاق الأســـرة هو 2000 أســـرة. وكان هذا الحجم ليس مبنيا على أســـاس وجود قيمة معروفة لتباين الأســـر في الانفاق على الســـلع. يتم هنا الاكتفاء باختيار (500) أســـرة اختيارا عشوائيا من مجتمع الدراسة ويتم على أساسها تقدير التباين، ثم يستخدم هذا التقدير لحساب حجم العينة، فقد يظهر أن حجم العينة الكافي (بمســـتوى الثقة والدقة المحددتين مســـبقا) هو (2300) أســـرة. عندها يتم اســـتكمال المسح بوحدات العينة المتبقية وهي مقدار الفرق بين 2300 و500 أي 1700 منشأة.

إن أهم ما يميز هذا الأسلوب هو توفير تقديرات موثوق بها لمعلمة S^2 أو p، أما محددات استخدام هذ الأسلوب فهي أنه يحتاج إلى جهد كبير ووقت طويل. إضافة الى ذلك إن أســـلوب ســـحب العينة انقســـم الى جزئين كل جزء تم من خلاله ســـحب وحدات عينة بكســـر معاينة أي باحتمالية مختلفة عن الآخر، أي لم توظف نظرية الاحتمالات في اختيار العينة بشكلها الأمثل.

2.2. الأسلوب الثاني: اجراء مسح تجريبي

يقوم هذا الأسلوب على الا ستفادة من بيانات المسوح التجريبية التي يتم تنفيذها قبيل المسوح الرئيسية لتخدم اهداف أخرى كاختبار اســـتمارات المســـح وقواعد التدقيق الضـــبط إضـــافة الى تقديرات اعدا د الكوادر اللازمة لتنفيذ المســـح. إن واحداً من الأهداف الأخرى التي يمكن وضـــعها للمســـوح التجريبية هو الاســـتفادة من بياناتها في تقدير تباين المجتمع وحساب الأخطاء المعيارية اللازمة لتقدير حجم عينة المسح الرئيسي.

2.3. الأسلوب الثالث: الاعتماد على نتائج مسوح سابقة

يعتبر هذا الأسلوب الأكثر شيوعا في الجانب التطبيقي، اذ عند تقدير حجم العينة المناسب يتم الرجوع الى مسوح ســـابقة تم تنفيذها على نفس المجتمع او حتى على مجتمع اخر مشـــابه وذلك لتقدير قيمة الخطأ المعياري (Standard error) وعلى الرغم من أن التباين يعتبر أكثر اســـتقرار من التغيرات التي تحصـــل في مؤشـــرات النزعة المركزية للظاهرة قيد الدرا سة الا أن مقدار الانحراف المعياري المعتمد على هذا التباين قد يحتاج أحيانا اجراء تعديل مناسب عليه لاحتمال حصول تغيرات في سلوك الظاهرة مع الزمن.

2.4. الأسلوب الرابع: تقدير النسبة للخصائص الوصفية

إن مؤشرات النسبة (عدد الحالات التي تتصف بها الظاهرة المدروسة إلى عدد الحالات الكلي) هي من المؤشرات الأساسية في الدراسات والمسوح. فمن الممكن إعطاء تقدير للنسبة p. ومن المعلوم أن كلما كانت النسبة قريبة من الواقع كلما كان تقدير التباين دقيقا. فإذا كانت مثلا النســـبة الحقيقية في المجتمع (p=0.3) فإن التباين هو

$$(S^2 = p(1-p) = 0.3 \times 0.7 = 0.21)$$

يلاحظ أن مقدار التباين يكون في حده الأعلى عندما تكون قيمة النسبة تساوي 0.5، أي عندها تكون:

$$(S^2 = p(1-p) = 0.5 \times 0.5 = 0.25)$$

وهذا الافتراض يتطلب اختيار أكبر حجم عينة في ظل مستوى الثقة والدقة المعتمدين مسبقا.

من هنا إذا كان مقدار التباين غير معلوم لظاهرة معينة، وأردنا تقدير حجم عينة مناســـب لها يتم تحديد مقدار التباين بالمقدار 0.25 فهو يعطي أكبر حجم عينة ممكن كأجراء احترازي.

3. اختيار المتغير المناسب لتقدير حجم العينة

من المعلوم في الجانب التطبيقي إن الهدف من تنفيذ مسح احصائي لا يقتصر على عملية جمع بيانات خاصة بمؤشر واحد فقط، اذ لابد وأن يكون هناك عدة مؤشرات تجمع عنها بيانات المسح وذلك لتحليل الظاهرة المدروسة من مختلف الجوانب. من جانب اخر أصبح هناك تركيزا من قبل معظم المؤسسات الإحصائية الوطنية والدولية نحو تنفيذ مسوح إحصائية تكون ذات مؤشرات متعددة، ففي ظل تطور أساليب جمع البيانات ومعالجتها آليا أصبحت المسوح متعددة المؤشرات (Multi-indicators Surveys) شائعة الاستخدام. حيث توفر نتائج مثل هذه المسوح مؤشرات عديدة ومختلفة. من أمثلة ذلك المسح العنقودي متعدد المؤشرات (Multi-indicators cluster Surveys) الذي يهدف إلى جمع مؤشرات تفصيلية عن وضع الطفل والأم وما يحيط بهما من ظروف وعوامل.

إن التحدي الرئيسي الذي يمكن أن نواجه عند تحديد حجم عينة هذا النوع من المسوح هو الية وأسلوب اختيار المؤشر المناسب الذي يمكن أن يفرز حجم عينة كاف لإعطاء تقدير لقيمة هذا المؤشر، وتقديرات للمؤشرات الأخرى في المسح بحيث تتسم جميعها بالدقة والكفاءة.

إن هذا التحدي يقودنا الى نتيجة وهي أن الأسلوب الأمثل في اختيار المؤشر المناسب لتحديد حجم العينة يكون وفقا لما يلي:

- تحديد جميع المؤشرات المهمة التي يتضمنها المسح، واختيار مؤشر مهم منها، بحيث يتطلب أكبر حجم عينة
- الحرص على ألا تقل نسبة الخطأ (عن حد معين) للمؤشرات المختلفة التي يجري العمل على جمع بيانات عنها.
- تحديد الفئات المستهدفة، واختيار الفئة الأقل نسبة في المجتمع مع مراعاة أهمية تلك الفئة في مجمل أهداف المسح.

مثال (2):

إذا كان مؤشر نسبة الأسر التي لديها إنفاق على السلع المعمرة في مسح دخل وأنفاق الأسرة هو (82%) وكان مؤشر نسبة الأسر التي لديها طلاب في المدارس (60%) على التوالي. فإن حجم العينة المسحوبة من المؤشر الأول ينبغي أن يكون هو المعتمد على الرغم من أن المؤشر الثاني يتطلب حجم عينة أكبر، كما يتضح من العمليات الحسابية الآتية بافتراض أن Z=1.96 و d=0.05.

$$n_1 = \frac{(1.96)^2\,(0.82)(0.18)}{(0.05)^2} = 202$$

$$n_2 = \frac{(1.96)^2\,(0.6)(0.4)}{(0.05)^2} = 369$$

يعتبر المؤشر الأول في هذا المسح من المؤشرات الرئيسية واهميته تتفوق كثيرا على المؤشر الثاني، وعليه فإن زيادة حجم العينة من 202 إلى 369 ليست مبررة للوصول الى مؤشر لا يعد أساسيا في مسح دخل وأنفاق الأسرة.

4. تقدير حجم العينة لتقدير مؤشر النسبة (p)

إن تقدير حجم العينة المناسب يرتبط بطبيعة ونوع المؤشر الرئيسي الذي سوف تستخدم بيانات العينة في تقديره، فقد تكون المؤشرات رقمية كالمتوسط الحسابي، كمتوسط الطول (بالسنتمتر)، أو متوسط وزن العبوة (بالغرام)، أو متوسط درجات الطلبة (بالدرجة)، أو قد تكون المؤشرات تهتم بنسبة حدوث الظاهرة (أي نسبة تكراراها)، مثل نسبة المدخنين البالغين.

تقوم القواعد الإحصائية في تقدير حجم العينات سواء كانت تهتم في تقدير مؤشرات رقمية او مؤشرات نسبة، تقوم على أساس خصائص التوزيع الطبيعي.

4.1. حجم العينة لتقدير نسبة المجتمع بدلالة مقدار نقاط الخطأ المطلق (d) المحدد مسبقا

إذا كان $\hat{p}$ هو تقدير لنسبة المجتمع p ، وذلك بالاعتماد على مسح سابق أو بناء على تجربة قبلية أو بافتراضه قيمة معينة، وفق أسلوب المعاينة العشوائية البسيطة يكون توزيع المعاينة للتقدير هو التوزيع الطبيعي التقريبي حيث

$E(\hat{p}) = p$ وتباين $var(\hat{P}) = \frac{p(1-p)}{n}$

تمثل القيمة d هامش الخطأ وهو عبارة عن الفرق بين نسبة المجتمع الحقيقية والنسبة المقدرة من العينة، ويتم التعبير عنها بالصيغة الآتية:

$$D = Z_{\frac{1-\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \quad \ldots\ldots (1)$$

حيث أن: $\hat{q} = (1 - \hat{p})$

تعبر قيمة $Z_{\frac{1-\alpha}{2}}$ في العلاقة (1) عن عدد الأخطاء المعيارية التي تبتعد فيها النسبة المقدرة من متوسط العينة (p).

(أي الفرق بينهما) أي أن قيمة (d) تعطي دلالة لمستوى الدقة، للحصول على قيمة صغيرة ل (d) يتطلب زيادة حجم العينة، فإذا اختيرت القيمة Z=1.96 فإن تفسير ذلك هو 95% من نسب العينات الممكنة ستقع ضمن ما مقداره 1.96من الأخطاء المعيارية standard errors.

إن ما سبق يقود إلى القول بأن الحرص على أن يكون الفرق بين النسبة الحقيقية والنسبة المقدرة من العينة صغيرا جدا يعني اختيار حجم عينة كبير نسبياً. وللحصول على حجم العينة يمكن حل المعادلة (1) أعلاه بدلالة قيمة (n) لنحصل على التالي:

$$n = \frac{\left(Z_{\frac{1-\alpha}{2}}\right)^2}{d^2} \hat{p}\hat{q} \quad \ldots\ldots (2)$$

يتضح من المعادلة أعلاه أن حجم العينة يزداد بزيادة قيمة البسط في العلاقة (2)، عند قيمة معينة ل Z فإن n تتغير تبعا لتغير قيمة $\hat{p}$.

إن أكبر حجم عينة ممكن بافتراض ثبات هامش الخطأ ومستوى الثقة، هو عندما تكون قيمة النسبة $\hat{p}$ مساوية ل (0.5) عندها يكون مقدار $\hat{p}\hat{q}$ =0.25.

4.2. حجم العينة لتقدير نسبة المجتمع بدلالة الفرق النسبي ε

يمكن تقدير حجم العينة ليس بالاعتماد على مقدار الخطأ المطلق وانما على مقدار الخطأ كنسبة من قيمة التقدير p، أي عندما يكون مقدار الخطأ d= εp . وفق المعادلة (1) أعلاه يكون تقدير حجم العينة كالتالي:

$$n = \frac{\hat{q}\left(Z_{\frac{1-\alpha}{2}}\right)^2}{(\hat{p})(\epsilon)^2} \ldots\ldots (3)$$

على سبيل المثال إذا كانت قيمة $\hat{p}$ =0.2 ، وكانت قيمة التقدير المطلوبة ضمن ε=0.10 فإن قيمة الخطأ المطلق بالنقاط هي d=0.2X0.10=0.02، في هذه الحالة يكون حجم العينة المقدر هو:

$$n = \frac{0.8(1.96)^2}{(0.2)(0.10)^2} = 1536$$

5. تقدير حجم العينة لتقدير المتوسط (μ)

إذا كانت المؤشـرات المطلوب تقديرها هي ليسـت نسـب وانما متوسـطات او مجاميع، في هذه الحالة إن المبدأ الإحصائي لحساب حجم العينة اللازم اختياره لا يختلف عن حساب حجم العينة في حال تقدير المؤشرات التي تكون على شكل نسب.

5.1. حجم العينة لتقدير متوسط المجتمع بدلالة الخطأ بالنقاط (d)

إن حجم العينة اللازم للو صول إلى تقدير منا سب للمتو سط (μ) لمتغير معين، يح سب حجم العينة باعتماد ال صيغة التالية:

$$n=\frac{\left(Z_{1-\frac{\alpha}{2}}\right)^2 \sigma^2}{d^2} \quad \ldots\ldots (4)$$

حيث أن:

σ^2: هو تباين المجتمع، ويمكن تقديره با ستخدام مقدار s^2 لعينة من م سح سابق م شابه او بالاعتماد على نتائج مسح قبلي.

d: تمثل حد الخطأ المسـموح به في التقدير وهي: القيمة المطلقة للفرق المرغوب عدم تجاوزه بين متوسـط المجتمع الحقيقي (μ) والمتوسط المقدر من بيانات العينة ($\bar{x}$).

مثال (3):

تشـير المعلومات إلى أن التباين في انتاج عدد من المزارع هو (140) كغم ... عند تقدير متوسـط الإنتاج في الحيازة الزراعية الواحدة بحيث لا يزيد الفرق بين المتوسـط المقدر من معطيات العينة وبين المتوسـط الحقيقي عن (كغم واحد)، بمستوى ثقة 95%. فإن حجم العينة اللازم (عدد الحيازات) لتقدير متوسط عدد انتاج الحيازة كالتالي:

أن الخطأ المطلق المسموح به هو d=1

قيمة Z التي تحقق ثقة قدرها 95% هي z=1.96، وتباين المجتمع هو σ^2 =140

وفق المعادلة (4) أعلاه يكون تقدير حجم العينة كالتالي:

$$n= \frac{(140)(1.96)^2}{1^2} =538$$

إن حجم العينة البالغ (538) هو حجم كبيرا نسـبيا , والسـبب هو اختيار حجم خطأ مطلق صـغير، فإذا قبلنا بفرق لا يتجاوز 2 كغم (أي d=2) فإن حجم العينة سـينخفض بشكل كبير، حيث:

$$n= \frac{(140)(1.96)^2}{2^2} =134$$

5.2. حجم العينة لتقدير متوسط المجتمع بدلالة الخطأ مقاسا بنسبة من المعلمة (ε)

إذا ما تم اعتماد فرق نسـبي للخطأ المسـموح به بدلا من الفرق بالنقاط، عند ذلك تسـتبدل الصـيغة (4) بالصـيغة الآتية:

$$n= \frac{Z^2\sigma^2}{\varepsilon^2(\mu)^2} \quad \ldots\ldots (5)$$

حيث أن:

ε : نسبة الخطأ المسموح به

μ : معلمة المجتمع (الوسط الحسابي)

لكن القيم الحقيقية μ σ^2 قد لا تكون متوفرة، عند ذلك يستعاض عنها بالقيم $\bar{x}$ و s^2 على الترتيب من مسوح أو دراسات سابقة، أو الاستعانة ببعض المسوح التجريبية المحدودة. بحيث يصبح حجم العينة كالتالي:

$$n= \frac{Z^2 s^2}{\varepsilon^2 (\bar{x})^2} \quad \ldots\ldots (6)$$

مثال (4):

في المثال السابق إذا كان الخطأ المسموح به هو أن يقع متوسط العينة ضمن 5% من القيم الحقيقية للمتوسط، التي تشير دراسة سابقة إلى أنها تبلغ (80) كغم (أي80= μ), وأن مقدار التباين 250، فإن حجم العينة اللازمة سيكون:

$$n=\frac{(250)(1.96)^2}{(0.05)^2(80)^2} = 60$$

إذا ما نظرنا الى المعادلة (5) نلاحظ أن نسبة σ^2 إلى μ^2 ما هي إلا قيمة مربع معامل الاختلاف (Coefficient of Variation) حيث أن:

$$C.V= \frac{\sigma}{\mu}$$

إن هذا المعامل في الغالب غير متوفر يتم تقديره من بيانات العينة لمسح سابق.

فإذا كان المرغوب فيه اعتماد قيمة معامل الاختلاف لتقدير حجم العينة للمسح الحالي، فيمكن التعبير عن ذلك بالصيغة التالية:

$$n= \frac{\left(Z_{\frac{1-\alpha}{2}}\right)^2 (C.V)^2}{\varepsilon^2} \quad \ldots\ldots (7)$$

6. تقديرات حجم العينة في المعاينة الطبقية

إن تقديرات حجم العينة السابقة تنطبق بشكل أساسي عند أسلوب تطبيق المعاينة العشوائية البسيطة، ولكن وكما بينا سابقا إن أسلوب المعاينة الطبقية يختلف عنه في أسلوب العشوائية البسيطة، فهو يقوم على أساس تقسيم المجتمع الى طبقات مستقلة يتم تقدير حجم العينة ومعاينة كل طبقة بشكل مستقل تبعا لعدد من المتغيرات وهي مقدار التباين في الطبقة الواحدة وكلفة معاينة الوحدة الواحدة وحجم الطبقة.

إن صيغة حجم العينة تنطلق من العلاقة بين الخطأ المسموح به d والخطأ المعياري للتقدير حيث أن:

$$d = Z.S(\bar{x})$$

فاذا حددنا قيمة الخطأ d ومستوى الثقة Z، وبالاعتماد على الصيغة التالية للخطأ المعياري لتقدير العينة الطبقية:

$S^2(\bar{x}_{st}) = \frac{d^2}{Z^2} = B^2$. يكون حجم العينة الطبقية الكلي كالتالي:

- في حالة التوزيع المتساوي:

$$n_{eq} = \frac{L\sum_{h=1}^{L} N_h^2 S_h^2}{N^2 B^2 + \sum_{h=1}^{L} N_h S_h^2} \quad \ldots\ldots (8)$$

- في حالة التوزيع المتناسب:

$$n_{prop} = \frac{N\sum_{h=1}^{L} N_h^2 S_h^2}{N^2B^2+\sum_{h=1}^{L} N_h S_h^2} \quad \ldots\ldots (9)$$

- في حالة توزيع نيمان:

$$n_{Ney} = \frac{\left(\sum_{h=1}^{L} N_h S_h\right)^2}{N^2B^2+\sum_{h=1}^{L} N_h S_h^2} \quad \ldots\ldots (10)$$

- في حالة التوزيع الأمثل:

$$n_{opt} = \frac{\sum(N_h S_h \sqrt{C_h})\sum\frac{N_h S_h}{\sqrt{C_h}}}{N^2B^2+\sum_{h=1}^{L} N_h S_h^2} \quad \ldots\ldots (11)$$

6.1. موائمة حجم العينة الطبقية وفق الكلف المحددة للمسح

إذا كانت الميزانية المتاحة للمسح (بالعينة) هي المحدد الوحيد لحجم العينة، وكانت كلفة العمل الميداني المرتبطة بحجم العينة محددة الكمية (C) وهي مبينة على أســـاس كلفة الوحدة الواحدة لجميع البيانات (C_h) (أي أن كلفة الوحدة الواحدة ثابتة لكل الطبقات)، يحسب حجم العينة بموجب الصيغة التالية:

$$n = \frac{C}{c_h} \quad \ldots\ldots (12)$$

مثال (5):

إذا كانت التكاليف الكلية المتاحة للعمل الميداني 160000 درهم وكانت كلفة جمع البيانات من وحدة المعاينة الواحدة 100 درهم، فإن حجم العينة الكلي يكون:

$$\text{n= } \frac{c}{c_h} = \frac{160000}{100} = 1600$$

- الكلفة الكلية محددة مسبقا وتباينات الطبقات ثابتة مع اختلاف الكلف بين الطبقات

بافتراض أن الكلفة الكلية للمســـح معلومة مســـبقا وتســـاوي c ، وأن كلفة وحدة المعاينة الواحدة (c_h) تختلف من طبقة لأخرى، فإن حجم العينة الكلي يحسب كالتالي:

$$n = C.\frac{\sum\frac{N_h}{\sqrt{C_h}}}{\sum N_h\sqrt{c_h}} \quad \ldots\ldots (13)$$

مثال (6):

إذا حددت الكلفة الكلية ب 160,000 درهم، وكانت تفاصيل الحجم والكلفة، كما في الجدول:

	الطبقات			$N_h\sqrt{C_h}$	$N_h/\sqrt{C_h}$
Stratum No.	N_h	C_h	$\sqrt{C_h}$		
1	4000	36	6	24000	667
2	2700	81	9	24300	300
3	1600	100	10	16000	160
المجموع	**8300**			**64300**	**1127**

إن حجم العينة المقدر هو:

$$n = C.\frac{\sum\frac{N_h}{\sqrt{C_h}}}{\sum N_h\sqrt{C_h}} = \frac{(160000)(1127)}{64300} = 2805$$

- الكلفة الكلية محددة مسبقا مع اختلاف كل من احجام الطبقات والتباينات والكلف بين الطبقات حال أن الكلفة الكلية للمسح معلومة مسبقا، بينما كان كل من احجام الطبقات، ومقدار التباين لكل منها قيم مختلفة فأن تقدير حجم العينة يكون وفق المعادلة التالية:

$$n = C.\frac{\sum_{h=1}^{L} N_h S_h/\sqrt{c_h}}{\sum_{h=1}^{L} N_h S_h\sqrt{c_h}} \quad \ldots\ldots (14)$$

- الكلف والتباين محددة مسبقا على مستوى الطبقة، عند مستوى من الدقة معلوم مسبقا

إذا ما تم ربط كل من حجم العينة وكلفة المسح المعلومة مسبقا بمستوى الدقة المطلوب، فيتم تقدير حجم العينة على النحو التالي:

$$n = \frac{\sum(N_h S_h\sqrt{C_h})\left(\sum\frac{N_h S_h}{\sqrt{C_h}}\right)}{N^2B^2 + \sum_{h=1}^{L} N_h S_h} \quad \ldots\ldots (15)$$

حيث أن B هو مقدار مستوى دقة معينة يساوي $B = \frac{d}{z}$ ، أما المقدار d فهو هامش الخطأ المحدد مسبقا. بينما قيمة Z هي تتعلق بمستوى الثقة المطلوب في التقديرات.

7. دور أثر التصميم في تقدير حجم العينة بالمعاينة العنقودية

يعتبر اسلوب المعاينة العشوائية البسيطة حجر الاساس في تصاميم المعاينة الاخرى، ونتيجة لمحددات مختلفة مثل عدم توفر التجانس في المجتمع، الكلفة العالية، عدم توفر إطار معاينة دقيق أحيانا، يلجأ الى اساليب معاينة أكثر تعقيد ولكنها بنيت على اساس العينة العشوائية البسيطة. ففي المعاينة العنقودية هناك قيم احتمالية مختلفة ترتبط بخصوصية كل عنقود وهذه تؤدي الى خصائص قد لا تنطبق مع افتراض الاستقلالية والتطابق في التوزيع الاحتمالي للمجتمع وانما الى توزيعات أكثر تعقيد الامر الذي يؤدي الى ظهور مشاكل وتعقيدات في التحليلات الاحصائية.

إذا ما رغبنا بتطبيق المعاينة العنقودية في مسح معين. لتحديد حجم العينة المطلوب ضمن مستوى ثقة وحد خطأ معلومين يلجأ في هذه الحالة الى افتراض الاستقلالية والتطابق في البيانات وأنها تتوزع توزيع طبيعي وبالاعتماد على المعادلات المشتقة وفق هذا الفرض نحصل على حجم العينة المقدر.

إلا أن حجم العينة المسحوب هو مقرر لعينة عشوائية بسيطة وليست عنقودية، ولكن من المعلوم أن تباين المعاينة العنقودية أعلى من تباين المعاينة العشوائية البسيطة مما يعني أننا بحاجة الى التغلب على ذلك من خلال زيادة حجم العينة في المعاينة العنقودية إلى حد معين يكون فيه التباين أقرب ما يكون إلى تباين العشوائية البسيطة.

بناء على ما سبق، يستخدم معامل نسبي يبين التأثير المتوقع لاستخدام تصميم معاينة معين بدلا من استخدام العينة العشوائية البسيط يسمى هذا المعامل بتأثير التصميم.

تأثير التصميم: هو نسبة تباين التصميم الفعلي الى تباين التصميم العشوائي البسيط، وهو مقياس للكفاءة النسبية ويعبر عنه رياضياً:

$$deff(\hat{\theta}) = \frac{Var_{cluster\ Des.}(\hat{\theta})}{Var_{SRS}(\hat{\theta})} \quad \ldots\ldots (16)$$

يتضح مما سبق إن التضخم في التباين في اسلوب المعاينة العنقودية يعود الى التصميم العنقودي، وعليه يمكن تفسير إثر التصميم على انه مقدار التضخم في تباين التقدير نتيجة لاستخدام التصميم العنقودي بدلا من العينة العشوائية البسيطة.

من جانب آخر إن عدد وحدات المعاينة الذي يجب اختياره من منطقة العد الواحدة يرتبط ارتباطا مباشرة بقيمة أثر التصميم وفق المعادلة التالية:

$$deff = 1 + \delta_x \times (\bar{n} - 1)$$

حيث أن:

$deff$ = أثر التصميم

δ_x = مقدار الارتباط الداخلي بين وحدات المعاينة الثانوية الواقعة ضمن منطقة العد الرئيسية الواحدة.

$\bar{n}$ = هو متوسط عدد وحدات العينة الثانوية المسحوب من منطقة العد الرئيسية الواحدة.

يتضح من المعادلة أعلاه أن مقدار أثر التصميم يتناسب طرديا مع كل من عدد وحدات العينة الثانوية في منطقة العد الرئيسية، ومع مقدار الارتباط الداخلي بين وحدات المعاينة في منطقة العد الرئيسية الواحدة.

مثال (7):

لدراسة متوسط دخل الأسرة مثلا في مجتمع ما، وبالاعتماد على مسوح مشابهة سابقا لنفس المجتمع، بحيث اختيرت آنذاك عينتين عشوائيا الاولى عنقودية والثانية عشوائية بسيطة، وكانت نسبة التباين الناتج بين العينتين (أثر التصميم) هي 1.4 ، فإن حجم عينة المسح الحالي اللازم للوصول الى تقدير لمتوسط دخل الأسرة وفق مستوى ثقة مقداره 95% أي z=1.96 وهامش خطأ مقداره d=1.5 ، وفق العينة العشوائية البسيطة:

$$n_{SRS} = \frac{z^2 S_{SRS}}{d^2} = \frac{1.96^2 370}{1.5^2} = 632$$

أي إن حجم العينة العشوائية البسيطة المطلوب ضمن مستوى ثقة 95% هو 632 منشأة، إذا استخدمنا العينة العنقودية بالتأكيد أن مقدار التباين سيكون اعلى منه في العينة البسيطة وعليه لابد من تعديل حجم العينة للإحاطة بمقدار التباين والمحافظة على التقدير ضمن حد الخطأ المحدد. هنا لابد من تعديل حجم العينة بتأثير التصميم أي أن:

$$n_{cluster} = deff \times n_{SRS}$$

وبما أن قيمة تأثير التصميم تساوي 1.4 ، فإن حجم العينة العنقودية يكون 632 مضروبا بقيمة تأثير التصميم 1.4، أي 885 وحدة معاينة ثانوية.

الفصل الرابع: أساليب المعاينة في مركز الإحصاء –أبو ظبي

ضـمن إطار المهام الموكل بتنفيذها مركز الإحصـاء أبو ظبي، والتي تتناول تنفيذ المسـوح الإحصـائية بكافة أنواعها الاقتصادية والأسرية والزراعية والبيئية التي تهدف الى توفير إحصاءات رسمية خاصة بإمارة أبو ظبي. هذا إضافة الى دوره في تصميم وسحب العينات التي قد تطلب من مختلف مؤسسات ودوائر حكومة أبو ظبي.

يقوم المركز ببناء اطر المعاينة بأنواعها المختلفة ويتابع بشـكل دوري ومسـتمر عمليات تحديث هذه الأطر، بحيث تستخدم في تصميم وسحب عينات إحصائية قادرة على تمثيل المجتمع الإحصائي بأعلى درجة من الكفاءة والدقة. وبناء على هذه الأطر تتم إجراءات تقدير احجام العينات المطلوبة، وسحب العينات الإحصائية بمختلف أنواعها.

1. أطر المعاينة في مركز الإحصاء – أبو ظبي

إن إطار المعاينة هو عبارة عن قائمة بجميع وحدات المجتمع قيد البحث والدراسـة، أو قد تكون على شـكل خرائط جغرافية توضح كافة وحدات المجتمع. هذا وتتضمن قائمة الإطار متغيرات تعريفية وجغرافية يستدل من خلالها على وحدات المجتمع، هذا إضافة الى متغيرات فنية أخرى تساعد في دراسة طبيعة وخصائص وحدات المجتمع ضمن الإطار، كعدد العمال في المنشـأة الاقتصـادية مثلا في إطار المنشـآت الاقتصـادية، او عدد الأسـر في منطقة العد الواحدة ضمن إطار الوحدات السكنية والأسر.

يعتبر كل من إطار المنشآت الاقتصادية وأطار الوحدات السكنية والاسر، إطاران يعتمد عليهما بشكل أساسي في تصميم وسحب عينات المسوح الاقتصادية والأسرية في إمارة أبو ظبي.

1.1. إطار المنشآت الاقتصادية

المحتوى: يتضـمن إطار المنشـآت الاقتصـادية، قائمة بجميع المنشـآت الاقتصـادية العاملة التي تمارس نشـاط اقتصادي واحد او أكثر في إمارة أبو ظبي، وذلك بحسب كيانها القانوني، اذ قد تكون مركزا رئيسيا او منشأة مفردة او فرع يمسـك حسـابات لمنشـأة مركزها الرئيسـي إمارة أبو ظبي، أو فرع المنشـأة لا تمسـك حسـابات ومركزها الرئيسي خارج إمارة أبو ظبي.

الطبقية: تصنف ضمن هذا الإطار المنشآت الاقتصادية الى طبقات بحسب النشاط الاقتصادي ISIC -4 على أربعة حدود، وعلى مستوى حجم المنشأة وفقا لعدد العمال فيها، وفق التصنيف الخاص بالمنشآت الاقتصادية لحكومة أبو ظبي، وكما يلي:

قطاع الصناعة التحويلية:
المنشـآت متناهية الصـغر (1-9) عمال، المنشـآت الصـغيرة (10 -100) عامل، المنشـآت المتوسـطة (101-250) عامل. المنشآت الكبيرة (250 عامل فأكثر).

باقي القطاعات (التجارة، الخدمات،. الخ):
المنشآت متناهية الصغر (1-5) عمال، المنشآت الصغيرة (6 -50) عامل، المنشآت المتوسطة (51-200) عامل. المنشآت الكبيرة (200 عامل فأكثر).

أما أهم المتغيرات التي يتضمنها الإطار فيمكن تصنيفها على النحو التالي:

أنواع المتغيرات	
المتغيرات الجغرافية	الإقليم، المنطقة، رقم الحوض، موقع المنشأة بالمبنى، عنوان المنشأة، وما الى ذلك.
المتغيرات التعريفية	اسم المنشأة، ورقم الرخصة، اسم صاحب المنشأة والمدير المسؤول ورقم هاتف المنشأة.
المتغيرات التحليلية	خصائص حالة المنشأة، والكيان القانوني، والنشاط الاقتصادي، وصفة المنشأة،

مصدر البيانات: تم بناء إطار المنشآت الاقتصادية بناء على حصر شامل للمنشآت الاقتصادية في إمارة أبو ظبي.

تحديث إطار المنشآت الاقتصادية: بهدف مواكبة الإطار لعملية التغطية والشمول لكافة المنشآت الاقتصادية التي تستحدث، او تغلق، او تغير نشاطها، تتم عمليات تحديث مستمرة تعتمد على نتائج المسوح الاقتصادية السنوية التي ينفذها المركز، اذ يتم سنويا اجراء التحديث على حالة المنشآت فيما إذا توقف نشاطها او غيرت من نوع النشاط، او إذا تغير عدد العمال في المنشاة، هذا إضافة الى تحديث المتغيرات التعريفية والجغرافية الدالة على المنشأة.

إضافة لما سبق من عمليات تحديث، يقوم المركز بتنفيذ مشروع تحديث إطار المنشآت الاقتصادية بالاعتماد على بيانات السجلات الإدارية المتوفرة من مؤسسات حكومة أبو ظبي لحصر قوائم المنشآت التي دخلت مؤخرا الى السوق ومن ثم زيارتها ميدانيا لجمع البيانات اللازمة واضافتها الى طار المنشآت وذلك بهدف الوصول الى أكبر شمول ممكن في الإطار الحالي.

1.2. إطار الاسر والوحدات السكنية

المحتوى: يتضمن قائمة بجميع الاسر المقيمة ضمن الوحدات السكنية المأهولة القائمة على أرض إمارة أبو ظبي.

مصدر البيانات: يتم بناء الإطار بالاعتماد على سجلات حصر المباني والمساكن والأسر من واقع تعداد السكان والاسر والمساكن، بحيث تدرج جميع الوحدات السكنية المأهولة بأسر ضمن الإطار.

التنظيم الجغرافي للإطار: يأتي تنظيم الإطار جغرافيا منسجما مع التقسيمات الإدارية المعتمدة لدى حكومة أبو ظبي، إضافة الى تقسيمات إحصائية تفصيلية تستخدم لأغراض المعاينة:

اما مستويات التقسيم الإدارية المعتمدة فهي: المنطقة الرئيسية Region، المنطقة الفرعية district، الحوض Sector. في حين أن مستويات التقسيم الإحصائي هي مناطق العد الإحصائية Enumeration Area.

تعريف: منطقة العد هي عبارة عن منطقة جغرافية ذات حدود طبيعية أو صناعية تتضمن مباني ووحدات سكنية وأسر، بحيث يكون متوسط عدد الأسر فيها يتراوح ما بين 100 الى 200 أسرة بالمتوسط، مع بعض الاستثناءات الخاصة بالمناطق الممتدة على مساحات كبيرة وذات كثافة سكانية قليلة.

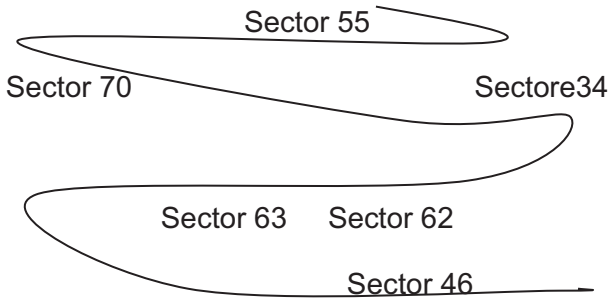
بناء على ما سبق، يأتي بناء وتنظيم إطار الوحدات السكنية والأسر على مرحلتين، بحيث ينسجم مع تصاميم المعاينة الإحصائية التي يمكن تطبيقها مستقبلا:

المرحلة الأولى: بناء (مناطق العد Enumeration areas)، بالاعتماد على الخرائط الجغرافية والبيانات الخاصة بالمباني والمساكن والأسر، تم تقسيم إمارة أبو ظبي الى مناطق عد مستقلة بحيث تحتوي المنطقة الواحدة 100 -200 أسرة.

بناء على ذلك، من الممكن أن تشكل منطقة العد الواحدة حوض او جزء من حوض، أو أحيانا قد يتم دمج مجموعة من الاحواض لتشكيل منطقة عد واحدة، وذلك بحسب حجم وتوزيع الأسر والوحدات السكنية في المنطقة. ومن

الجدير بالذكر إن هذه المناطق بعد عملية بناء الإطار يتم اعتمادها كوحدات معاينة أولية Primary Sampling units (PSUs).

بعد إجراء عمليات تشكيل وبناء مناطق العد تم ترتيبها بحسب التقسيمات الإدارية التي وردت أعلاه بدءا من الإقليم ومن ثم القطاع. وضمن القطاع الواحد تم ترتيب الاحواض ترتيبا حلزونيا بدءا من اتجاه الشمال وانتهاء بالجنوب وذلك لضمان انتشار العينة جغرافيا.



أما ترتيب مناطق العد ضمن الحوض الواحد فكان أيضا ترتيبا حلزونيا بحيث يضمن أكبر انتشار داخل الحوض خاصة في الأحواض التي تتضمن عدد كبير من مناطق العد.

بالإضافة الى القائمة التي تتضمن جميع الوحدات السكنية والأسر ضمن منطقة العد الواحدة هناك خرائط جغرافية توضح المواقع التفصيلية للمباني والوحدات السكنية ضمن منطقة العد الواحدة.

المرحلة الثانية: إعداد قوائم الوحدات السكنية المأهولة بأسر وقت إعداد الإطار، يتضمن إطار الوحدات والأسر ضمن منطقة العد الواحدة قائمة بأسماء وعناوين جغرافية تفصيلية يمكن من خلالها الوصول الى الوحدة السكنية والأسر المقيمة فيها. هذا إضافة الى أسماء أرباب الأسر المقيمين في الوحدات السكنية ونوع الأسرة المقيمة في الوحدة بحسب جنسية رب الأسرة (أسرة مواطنة، أسرة غير مواطنة، أسرة جماعية). من هنا تصنف الأسر ضمن الإطار الى نوعين رئيسيين:

الأسرة الخاصة: هي عبارة عن فرد او كثر يعيشون ضمن وحدة سكنية واحدة ويتشاركون في وجبات الطعام، وهناك شخص متعارف عليه من بينهم على انه رب للأسرة وهو مسؤول عن الترتيبات المعيشية للأسرة، وغالبا ما ينفق أفراد الأسرة من دخل رب الأسرة. وإذا كان جنسية رب الأسرة مواطن تصنف الأسرة على أنها أسرة خاصة مواطنة، اما إذا كانت جنسية رب الأسرة غير مواطنة فتعتبر الأسرة خاصة غير مواطنة.

الأسرة الجماعية: هي مجموعة من الأفراد يقيمون معا في وحدة سكنية واحدة، لا تربطهم علاقة قرابة ولا يوجد رب للأسرة، وغالبا لا يشتركون في تناول وجبات الطعام والانفاق عليها معا.

ويجب التمييز هنا بين الأسرة الجماعية وفق هذا المفهوم ومعسكرات العمال، فمعسكرات العمال تكون وحدات سكنية كبيرة وتكون اداراتها من قبل مؤسسة أو شركة معينة تشغل الأفراد المقيمين في المعسكر، بينما الأسرة الجماعية يكون الأفراد المقيمين في الوحدة السكنية هم المسؤولين عن الترتيبات المعيشية وإدارة المسكن الذي يقيمون فيه.

تحديث إطار الوحدات السكنية والمساكن

إن عملية المحافظة على تحديث إطار المعاينة تعتبر واحدة من الأوليات الرئيسية التي تبنى عليها دقة الإطار وبالتالي كفاءة وجودة المسوح التي تبنى على هذا الإطار.

إن عملية تحديث الإطار تكون ضمن اسلوبين مستقلين يمكن الاعتماد على أي منهما بحسب الإمكانيات والموارد المتوفرة والوقت المتاح.

الاسلوب الأول: يعتمد على استخدام بيانات السجلات الإدارية التي توفر معلومات تفصيلية عن تطورات الوحدات السكنية والأسر المقيمة فيها، وغالبا ما توفر سجلات الخدمات كالماء والكهرباء مثلا بيانات يمكن الاستفادة منها في تحديث مناطق العد الجديدة التي ظهرت بعد فترة من انشاء الإطار، فقد تكون هناك مدن جديدة واسكانات أنشئت بعد التعداد او كانت تحت الانشاء والآن أصبحت مأهولة، فبالاعتماد على بيانات السجلات الإدارية يمكن إضافة هذه الوحدات الى الإطار بهدف زيادة مستوى الشمول والتغطية.

الاسلوب الثاني: يكون من خلال اجراء عمليات حصر شاملة لمناطق العد يتم خلالها تحديث المباني والوحدات السكنية والأسر المقيمة ضمنها، وفي الغالب لا تتم عملية التحديث وفق هذا الأسلوب لكافة مناطق العد وذلك نظرا للوقت والتكلفة الباهظة التي سوف تترتب على ذلك. من هنا تتجه بعض الدول الى اجراء عمليات تحديث جزئية للمناطق التي يتوقع أن يكون قد حصل فيها تغيرات معنوية بإعداد المباني والوحدات والأسر.

من جانب اخر هناك اسلوب اخر يتناول تحديث جزئي فقط لمناطق العد الواقعة ضمن العينة الرئيسية Master Sample، والعينة الرئيسية هي عبارة عن عينة كبيرة نسبية يتم سحبها بهدف خدمة عدة مسوح تكون مقسمة الى مجموعات تسمى مكررات Replicates بحيث تسحب المكررة الواحدة بشكل مستقل وتكون ممثلة للمجتمع بأكمله. عند اجراء أي مسح يتم اختيار مكررة واحدة أو أكثر بحسب حجم عينة المسح المقررة، ويتم تنفيذ المسح بالاعتماد عليها.

يقوم مركز الإحصاء – أبو ظبي بتحديث أطار المساكن والمباني الحالي اعتمادا على الاسلوب الثاني وذلك من خلال اجراء عمليات تحديث ميداني سنويا بشكل جزئي لنسبة معينة من مناطق العد في الإمارة، وإضافة مناطق جديدة لم تكن موجودة على الواقع أثناء فترة تجهيز الإطار.

2. تصميم عينات المسوح الإحصائية

2.1. تصميم عينات المسوح الاقتصادية

ضمن أهداف المسوح الاقتصادية السنوية المتمثلة في توفير مؤشرات اقتصادية على مستوى الأنشطة الاقتصادية على حدين، إضافة الى التمثيل ضمن النشاط الواحدة لحجم المنشأة (كبير، متوسط، صغير، صغير جدا). فإن أسلوب المعاينة الأمثل المعتمد في تصميم هذه العينة هو اسلوب المعاينة الطبقية العشوائية، اذ يتم تقسيم المنشآت الى طبقات مستقلة:

- بحسب النشاط الاقتصادي على حدين، وفي بعض الأنشطة تكون على مستوى الأربعة حدود.
- بحسب حجم المنشأة الاقتصادية الواقع ضمن 4 مجموعات. وهي كبيرة، متوسطة، صغيرة، متناهية الصغر.

مما سبق يتم التعامل مع كل طبقة على انها مجتمع احصائي مستقل، وتتحدد منشآت الطبقة الواحدة بالنشاط الاقتصادي، وبفئة عدد العمال (متناهية الصغر، او صغيرة، او متوسطة، او كبيرة).

تقدير حجم العينة

يتم تقدير حجم عينة المسح الاقتصادي السنوي ضمن كل طبقة او مجموعة من الطبقات وفق استراتيجيات تحددها تلك الطبقات:

- يتم تحديد حجم العينة على مستوى الطبقة الواحدة. أي على مستوى النشاط الاقتصادي وحجم المنشآت ضمن النشاط. مع مراعاة أن جميع الأنشطة ضمن أحجام الطبقات المختلفة التي تتضمن في الغالب 5 منشآت فأقل يتم مسحها بشكل شامل.
- أما في الأنشطة والطبقات الاخرى يتم تحديد حجم العينة وفق معادلات تقدير حجم العينة، وبالاعتماد على نتائج المسوح الاقتصادية للسنوات السابقة، بحيث لا يتجاوز مقدار هامش الخطأ لمتغير العينة الرئيسي وهو عدد العمال بين 15% و20%، باستثناء بعض الطبقات التي تتصف المنشآت فيها بالتباين العالي في حجم العمالة قد يصل هامش الخطأ فيها الى 25%. بعد تحديد حجم العينة ضمن الطبقة الواحدة يتم اضافة ما نسبته بالمتوسط 15% تعويضا عن حالات عدم الاستجابة وذلك بناء على نتائج المسوح الاقتصادية السابقة.
- إن تقدير حجم العينة ضمن المعايير أعلاه، يفرز مؤشرات لعدد العمال، تعويضات العاملين، مجموع الإيرادات، مجموع القيمة المضافة، على مستوى النشاط الاقتصادي على حدين بهامش خطأ لا يتجاوز 20% لمعظم المتغيرات التي سبق ذكرها.

أسلوب سحب العينة

يتم سحب عينة المنشآت الاقتصادية ضمن كل طبقة من الطبقات بأسلوب المعاينة المتناسبة مع الحجم الذي سبق شرحه، حيث أن حجم المنشأة الواحدة يعبر عنه بعدد العاملين في تلك المنشاة.

إن عملية سحب العينة ضمن طبقة المنشآت الكبيرة تختلف بشكل طفيف عن باقي الطبقات. فمن المعلوم ان هذه الطبقة تحتوي على المنشآت الاقتصادية الكبيرة ذات الوزن الاقتصادي المهم في الحسابات الاقتصادية، وبنفس الوقت لا يمكن مسحها بشكل شامل لأن عددها في الإطار كبير، وإن حجم العينة يكون محدود. في هذه الحالة تقسم الطبقة الواحدة الى طبقتين جزئيتين تتضمن الاولى المنشآت الاكبر حجما في النشاط وتسحب في العينة بشكل شامل، في حين باقي المنشآت ضمن الطبقة الجزئية الثانية يتم اختيار عينة عشوائية منها وفق اسلوب المعاينة المتناسبة مع الحجم.

اما الية تقسيم طبقة المنشآت الكبيرة الى طبقتين جزئيتين فتعتمد على توزيع تركز العمال في الطبقة، بحيث تتكون الطبقة الجزئية الاولى من 40% على الاقل من المنشآت ذات الحجم الاكبر في الطبقة. وهذه النسبة تتفاوت في بعض الانشطة بحسب عدد العمال في الطبقة الكبيرة وتوزيعهم.

حساب أوزان العينة

من المعلوم أن وزن وحدة المعاينة هو عبارة عن معكوس احتمال اختيار تلك الوحدة من العينة. وبما أن اسلوب المعاينة طبقي، فيتم حساب أوزان المعاينة لمنشآت العينة ضمن كل طبقة بشكل منفصل. إن وزن عينة المنشآت الاقتصادية ضمن طبقة معينة يساوي حاصل قسمة عدد المنشآت الاقتصادية في الإطار على عدد المنشآت الاقتصادية المستجيبة لبيانات المسح.

أما في طبقة المنشآت الكبيرة وتحديدا في الطبقة الجزئية الاولى فيكون وزن المنشأة الاقتصادية يساوي واحد اي تمثل نفسها فقط، باعتبار انها سحبت بشكل مؤكد وليست بالأسلوب الاحتمالي. في حين أن باقي المنشآت في الطبقة الجزئية الثانية يتم حساب أوزان كل منها بنفس الاسلوب السابق.

2.2. تصميم عينات المسوح الأسرية:

لقد تم إعداد وتنظيم إطار المسوح الأسرية بما يتناسب مع أهداف تلك المسوح بحيث يضمن تمثيل كفؤ للنتائج. وكأحد الإجراءات الأساسية التي اعتمدت عند بناء الإطار تم تقسيم مجتمع الأسر في إمارة أبو ظبي

الى أربعة طبقات مستقلة وذلك لضمان الحد الأدنى من التباين في وحدات المجتمع بما يكفل التقليل ما أمكن من حجم العينة مع المحافظة على مستوى عال من الدقة، أما الطبقات فقد كانت بالشكل التالي:

الطبقة الاولى: تتضمن جميع مناطق العد التي فيها - أسر مواطنة بنسبة تقل عن 25%

الطبقة الثانية: تحتوي على جميع مناطق العد التي فيها أسر مواطنة بنسبة 25% الى 50%.

الطبقة الثالثة: تحتوي على جميع مناطق العد التي فيها أسر مواطنة بنسبة 50% الى 75%.

الطبقة الرابعة: تحتوي على جميع مناطق العد التي فيها أسر مواطنة بنسبة 75% الى 100%.

بناء على ذلك فقد تم ترتيب مناطق العد ضمن الإطار وفق التسلسل التالي: المنطقة الرئيسية، المنطقة الفرعية، الحوض، الطبقة. بحيث تصنف ضمن المنطقة الواحدة (أبو ظبي، العين، الظفرة) الى 4 طبقات بحسب التصنيف أعلاه. بناء على ذلك يكون عدد الطبقات 12 طبقة ضمنية

أما أسلوب المعاينة المستخدم فهو المعاينة الطبقية العنقودية المسحوبة على مرحلتين Stratified Two Stage cluster Sample Design. ضمن المرحلة الأول يتم سحب عينة مناطق العد من كل طبقة من الطبقات المذكورة أعلاه، وفي هذه المرحلة تسمى مناطق العد بوحدات المعاينة الأولية PSUs، وفي المرحلة الثانية يتم سحب عينة الوحدات السكنية المأهولة بأسر، من كل وحدة معاينة أولية تم اختيارها في المرحلة الأولى.

أسلوب سحب عينات المسوح الأسرية

بعد تحديد حجم العينة وتوزيعها بحسب الإقليم والطبقة، تتم إجراءات السحب ضمن مرحلتي المعاينة بالطرق التالية:

- يتم سحب مناطق العد من قائمة المناطق في الإطار بأسلوب المعاينة العشوائية المتناسبة مع الحجم Proportional to the size، إن هذا الأسلوب يعطي فرصة أكبر لمناطق العد ذات عدد الأسر الكبير نسبيا بالظهور في العينة مما يرفع من الكفاءة والجودة.
- يتم سحب عينة الأسر من كل منطقة عد، بأسلوب المعاينة العشوائية المنتظمة Systematic Sampling، إن هذا الأسلوب يضمن انتشار توزيع عينة الأسر المختارة أن تكون منتشرة جغرافيا في منطقة العد الواحدة. مما يعطي تباين عالي نسبيا بين الأسر داخل منطقة العد الواحدة. وبنفس الوقت يعطي تباين قليل نسبيا ما بين مناطق العد. مما يؤدي الى تفعيل دور هذا النوع من تصاميم المعاينة في إعطاء نتائج ذات دقة عالية.

المراجع:

- Theory and Analysis of Sample Survey Designs, Caroga Singh, 1986.
- Sampling Techniques, William. Cochran, 1953.
- معجم المصطلحات الإحصائية، المعهد العربي للتدريب والبحوث الإحصائية 2005.
- الأساليب الإحصائية في ميدان التطبيق، د. مهدي العلاق، أ ذ. عدنان شهاب حمد، 2001.

مـركـز الإحـصـاء
STATISTICS CENTRE

الرؤية: ببياناتنا نمضي نحو غدٍ أفضل
Vision: Driven by data for a better tomorrow

www.scad.gov.ae
adstatistics