المركز الاتحادي
للتنافسية والإحصاء
FEDERAL COMPETITIVENESS
AND STATISTICS CENTRE

United Arab Emirates

# الدليل الوطني
# لطرق المعاينة الإحصائية

Ministry of Cabinet Affairs | وزارة شؤون مجلس الوزراء

# دليل طرق المعاينة

إعــداد خميـــــس رداد
مراجعة د. زينب الـدبـاغ
د. عادل الزيتوني

# تقديـــم

في إطار حرص المركز الاتحادي للتنافسية والإحصاء بنشر الوعي والمعرفة في مختلف المجالات الإحصائية وتوحيد المصطلحات والتعاريف المستخدمة في النظام الإحصائي في الدولة، يسرني بأن أقدم هذا الدليل باسم المركز الاتحادي للتنافسية والاحصاء، والذي يهدف إلى تسليط الضوء على أهمية موضوع العينات في العمل الاحصائي، وتأثيره المباشر على دقة البيانات الاحصائية التي يتم جمعها باسلوب المعاينة.

يشتمل هذا الدليل على أساسيات نظرية وتطبيقات عملية في مجال الإحصاء، ويتناول أهم المصطلحات والتعاريف الخاصة بعمل العينات، بالإضافة إلى استعراض أنواع العينات الإحتمالية وبعض تصاميم العينات المستخدمة.

نأمل أن يكون هذا الدليل ذات اضافة علمية قيمة لكل من يعمل في مجال الإحصاء سواء في الأجهزة الإحصائية أو في الوزارات والمؤسسات الحكومية، وأن يسهم في رفع مستوى الكفاءة والفعالية في العمل الإحصائي.

أتقدم بجزيل الشكر لكل من ساهم في إعداد هذا الدليل، كما أرحب بأي اقتراحات بناءة تهدف إلى تحسين الدليل أو أية أدلة أخرى نخطط لإعدادها في المستقبل.

# فهــرس المحتويات

3 تقديم

5 مقدمـة

12 الفصل الأول
مصطلحات وتعاريف

31 الفصل الثاني
العينة العشوائية البسيطة

47 الفصل الثالث
العينة العشوائية الطبقية

69 الفصل الرابع
العينة المنتظمة

83 الفصل الخامس
التقدير النسبي

90 الفصل السادس
العينة المتناسبة مع الحجم

102 الفصل السابع
العينة العنقودية

110 الفصل الثامن
العينة متعددة المراحل

118 الفصل التاسع
تطبيقات أسلوب المعاينة

146 المراجع

## مقدمـــة

يعتبر أسلوب المعاينة أسلوباً علمياً يهدف إلى قياس المعلومات الإحصائية بشكل موثوق وليس بديلاً للتغطية الشاملة أو عوضا عنها. ويعتمد هذا الأسلوب على نظريات الاحتمالات، ويستخدم لجمع بيانات من جزء محدود من المجتمع يُعرف بأسم " **العينة**".

ويحظى موضوع العينات بأهمية خاصة نظراً لكونه المقياس الأول الذي يحدد دقة نتائج أي مسح أو دراسة تُجرى بالعينة، لذا عند تقييم أي دراسة، يبدأ بتقييم العينة المستخدمة، فإذا كانت العينة غير صحيحة، تُرفض النتائج المستخلصة منها، أما إذا كانت العينة صحيحة يُجرى بعد ذلك تقييم الدراسة من الجوانب الأخرى.

يُعد تصميم العينة من العوامل الرئيسية التي تحدد كُلفة المسح. حيث يحدد التصميم حجم العينة وتوزيعها، ويؤثر بذلك على اتخاذ القرار بشأن قبول أو رفض تصميم معين. يتطلب اختيار البديل الأنسب من بين تصاميم العينات المتاحة مراعاة عدة عوامل منها تكلفة التنفيذ، الوقت المتاح لإجراء المسح واستخراج النتائج، مستوى الدقة المطلوب، وإمكانات الجهاز الإحصائي أو الجهة المنفذة.

نظراً لندرة المختصين في هذا المجال، يلجأ العديد من مصممي العينات إلى استخدام أساليب معاينة بسيطة يمكن التعامل معها بسهولة، أو الاستعانة بتصاميم عينة جاهزة كنماذج. وفي بعض الأحيان، يقوم الخبراء بتطبيق نفس التصميم على مسوح متعددة رغم التباين في خصوصيات كل مسح، أو حتى نقل نفس التصميم من دولة إلى أخرى تختلف من حيث الخصائص المتعلقة بالدراسة التي ستنفذ.

لذا، من الضروري أن تتوفر خبرات متخصصة في مجال تصميم العينات لدى الجهاز المركزي للإحصاء.كما ينبغي أيضاً أن يكون لدى الفنيين في الإدارات الفنية في الأجهزة الإحصائية المختلفة معلومات كافية حول موضوع العينات لضمان اتخاذ القرارات المناسبة واختيار البديل الأمثل في تصميم العينة.

## ميزات ومجالات استخدام أسلوب المعاينة

إن استخدام أسلوب المعاينة له ميزات عديدة، وهناك مجال لاستخدام العينة لا يمكن أن يستخدم فيه المسح الشامل، وبالمقابل هناك العديد من المجالات التي تستدعي إجراء مسح شامل، أما من أهم ميزات ومجالات استخدام أسلوب المعاينة:

1. التوفير في الوقت والموارد .
2. دقة أعلى من التعداد الشامل وذلك لأن أخطاء غير المعاينة تكون قليلة حيث أن الذين يعملون في المسح لديهم الخبرة الجيدة ودقة أعلى عند جمع البيانات مقارنة بمن يعملون في التعدادات.
3. نتائج دقيقة بتكاليف محددة أو الحصول على مستويات دقة معينة بأقل عدد ممكن من مفردات المجتمع (بأقل تكاليف ممكنة).
4. عندما يكون هناك تباين بين مفردات المجتمع المدروس.
5. عندما يكون عدد أفراد المجتمع غير معروف بالكامل أو يكون مجال العد واسع جدا مثل احصاءات الأسماك في البحر أو عدد الأشجار في غابة.
6. عندما يؤدي المسح الشامل الى اتلاف جميع مفردات المجتمع مثل فحص الدم، اذا أخذ بالمسح الشامل يؤدي الى وفاة المريض. في كثير من المسوح يصعب جمع بيانات عن جميع مفردات المجتمع ففي فحص الدم لعدد من الخراف مثلا فإن عملية الفحص الشامل تعني سحب جميع الدم الموجود في كل خروف وإجراء الفحوص عليها، وهذا يؤدي إلى وفاة جميع الخراف، كذلك فإن عملية إجراء فحص دم لجميع القطعان الموجودة في منطقة معينه قد يحتاج إلى وقت طويل وجهاز ضخم لإجراء هذه الفحوص.
7. عندما يكون عدد أفراد المجتمع كبير جدا ويكون إجراء العد الشامل مستحيلا.
8. عندما يكون الوقت والأموال والمصادر الأخرى محددة.

## محددات استخدام أسلوب المعاينة

1. إن استخدام أسلوب المعاينة لا يغطي المجتمع بالكامل ولا يعطي نفس النتائج التي يعطيها المسح الشامل.
2. تحتاج نظريات المعاينة في تطبيقها إلى أشخاص مدربين تدريبا جيدا ومؤهلين.
3. يجب أن يتم تخطيط وتنفيذ المسح بالعينة بعناية فائقة حيث أن الخطأ في العينة يتضاعف عند وزن النتائج.

## متطلبات إجراء مسح بالعينة

كما أشير سابقا فإن إجراء المسح بالعينة يحتاج إلى أشخاص مؤهلين، كما أن إجراء المسح يحتاج إلى عناية فائقة أثناء التنفيذ والتخطيط للمسح بالإضافة لما تقدم فإن تصميم العينة يحتاج إلى:

1. توفر إطار حديث وشامل لجميع وحدات المعاينة في المجتمع دون حذف او تكرار ويتضمن البيانات الأساسية التي يحتاجها تصميم العينة.
2. وضوح أهداف المسح ومستويات النشر وغيرها من الأمور المذكورة في البند الثاني.
3. ان يكون مجتمع الدراسة والمجتمع المستهدف واضح ومحدد ويمكن الوصول إلى كل مفردة من مفرداته.
4. قابلية تطبيق تصميم العينة في تنفيذ المسح.
5. توفر المستلزمات الأخرى مثل الأجهزة والبرمجيات.

## تعريف المجتمع المراد دراسته

لكل دراسة إحصائية مجتمع خاص بها، فعند تقدير إنتاج حليب الأبقار مثلا فإن المجتمع يكون مزارع الأبقار، أو الأبقار الموجودة في منطقة معينة، وإذا كان الموضوع دراسة معدل البطالة في منطقة معينة فإن المجتمع يكون هو سكان تلك المنطقة الذين هم في سن العمل، ولا يجوز أن نأخذ منطقة أخرى مثلا أو نأخذ طلاب المدارس والجامعات بدلاً من الناشطين اقتصاديا، إذ يجب أن يكون المجتمع المستهدف معروف بوضوح ودون أي لبس أو غموض، وتحدد معالمه الجغرافية والديموغرافية والتقسيمات الإدارية والحدود الأخرى التي تحدد المجتمع بحيث يكون المجتمع يحقق أهداف الدراسة التي حددت في البند السابق.

## وحدات المعاينة

هي مفردات المجتمع ومن الممكن أن تكون وحدات طبيعية مثل السكان أو تجمعات طبيعية مثل الأسر، أو وحدات صناعية مثل المزارع. قبل إجراء المعاينة يجب أن يكون المجتمع مقسم إلى أجزاء معرفة غير متداخلة، وغير غامضة، كل عنصر يعتبر وحدة معاينة واحدة فقط، ويجب أن تكون وحدات المعاينة معرفة بحيث يكون لكل عنصر في المجتمع احتمال الظهور في وحدة معاينة واحدة وواحدة فقط، مثال ذلك إذا كانت وحدة المعاينة هي الأسرة فيجب أن يكون كل فرد منتميا لأسره واحدة (لا يوجد فرد ينتمي إلى أسرتين في نفس الوقت أو لا ينتمي لأي أسره).

## تحديد إطار المعاينة ووحدات المعاينة

يعرف إطار المعاينة بأنه القائمة التي تشمل جميع مفردات المجتمع المنوي دراسته أو ما يسمى بوحدات المعاينة، ويمكن أن تكون وحدات المعاينة قائمة معرفة أو خريطة توضح وحدات المعاينة أو الإثنين معا، وهو القاعدة التي تسحب منها العينات .

إن إعداد إطار المعاينة من التحديات العملية وذلك للصعوبات التي تواجه معدي الإطار، ويعتبر الإطار فعّالا إذا كان:

1. شاملا مستوعبا جميع وحدات المعاينة الموجودة في مجتمع الدراسة.
2. حديث ضمن مدة صلاحيته، بالنسبة لوحدات المعاينة والصفة المدروسة.
3. يجب أن تكون وحدات المعاينة من نفس النوع.
4. يجب أن يكون خالي من الأخطاء مثل الحذف والتكرار.

## المعاينة الاحتمالية والمعاينة غير الاحتمالية

**1. المعاينة الاحتمالية:**

عندما نستخدم قوانين الاحتمالات عند اختيار وحدات المعاينة بشرط أن يكون لكل وحدة معاينة في المجتمع احتمال ظهور في العينة، وعادة يبحث علم العينات في هذا النوع فقط.

**2. المعاينة غير الاحتمالية:**

في هذا النوع من العينات تعتمد عملية الاختيار على الحكم الشخصي، أو الحدس للشخص الذي يختار العينات ويسمى أحيانا العينة العمدية ، حيث يتم اختيار وحدات المعاينة التي يتوقع أن تكون من متوسط المجتمع، وهذا النوع من العينات يستخدم لبناء موقف معين (أخذ فكره) ولا يمكن تعميم النتائج، إذا كان الشخص الذي يختار العينات له خبرة جيدة فإنه قد يخرج بنتائج جيده، ولكن لا يمكن الاعتماد على تلك النتائج في أية تقديرات.

## ميزات التصميم الفعّال

إن اختيار تصميم المعاينة المناسب يعتبر من أهم وأصعب الأعمال التي يقوم بها مصمم العينة، حيث ان أفضل التصاميم هو التصميم الذي يعطي أعلى دقة بأقل كلفة ممكنة، هذا بالإضافة إلى ميزة التصميم بالسهولة والبساطة والقابلية للتطبيق، بالإضافة إلى المرونة وإمكانية إجراء تعديلات على التصميم، أو سحب عينات جزئية من العينة الرئيسية، كما ان التصميم يتميز بإمكانية التحليل لسلسلة زمنية.

## اختيار أسلوب معاينة مناسب

يجري اختيار الأسلوب المناسب من قبل المختصين بعد دراسة جميع المتغيرات بحيث يعطي أعلى دقة، ويكون الاختيار وفق أحد الخيارين التاليين:

1. تحديد مدى الدقة من قبل صاحب القرار وتقليل الكلفة لأقل حد ممكن من قبل مصمم العينة.
2. تحديد الكلفة من قبل صاحب القرار ورفع مستوى الدقة لأعلى حد ممكن من قبل مصمم العينة.

## بعض الأمور التي تؤخذ بالاعتبار عند تصميم العينة

1. **الإمكانيات الفنية المتاحة:** وتشمل الخبرات المحلية المتوفرة والقدرة على استخدام البرمجيات التي تعالج قضايا العينات وتختصر في الوقت والجهد، ففي حالة توفر خبرات جيدة يمكن عمل تصاميم معقدة تعطي دقة عالية، ولكن في حالة عدم وجود خبرات كافية فإنه يوصى باللجوء الى التصاميم البسيطة التي يستطيع العاملون في الأجهزة الإحصائية التعامل معها، كذلك الحال بالنسبة لتوفير البرمجيات والتعامل معها فإذا كانت متوفرة ويعمل المستخدمون بها بكفاءة فيمكن استخدام أساليب معاينة اكثر فعالية وتعقيد، أما في حالة عدم وجود هذه الإمكانيات فإنه يستخدم التصاميم التي تحتاج أقل ما يمكن من العمليات الحسابية لاستخراج النتائج وحساب خطأ المعاينة.
2. **الإمكانيات المادية المتوفرة:** تعتبر عاملا محددا لحجم العينة وأسلوب العمل، حيث أنه لا يجوز تصميم عينة يحتاج تنفيذها أكثر مما هو متوفر من أموال، أو يحتاج وسائل نقل أكثر مما هو متاح، لذلك يجب الموازنة بين التصميم المقترح والإمكانيات المتوفرة.
3. **مستويات النشر:** تُؤخذ بالاعتبار مستويات النشر فمثلا إذا كان مطلوب استخراج النتائج على مستوى الدولة يختلف عن تصميم العينة إذا كان مطلوب استخراج النتائج على مستوى الإمارة أو البلدية.
4. **الوقت المتاح لاستخراج النتائج:** يجب أن يكون تصميم العينة يحقق هدف الدراسة بحيث تُستخرج النتائج حسب الوقت المحدد، وذلك لأن المعلومة لها أهمية وقيمة في وقت معين، وتفقد قيمتها وأهميتها في حالة تأخر صدورها عن الوقت المحدد لذلك.
5. **مستوى الدقة المطلوب:** إن أكثر التصاميم فاعلية هو الذي يعطي أعلى دقة بأقل كلفة ممكنة، ويعتمد حجم العينة على مستوى الدقة المطلوب إضافة إلى تصميم العينة المستخدم، وكلما زاد حجم العينة تزيد الدقة لنفس التصميم.
6. **أهداف المسح أو الدراسة:** من الملاحظ أن تصميم عينة لمسح له هدف واحد يختلف عن تصميم عينة لمسح متعدد الأهداف.
7. **عدم الاستجابة المتوقعة:** ففي حالة توقع حدوث نسبة عالية من عدم الاستجابة فإن ذلك يتطلب استخدام تصميم فيه مرونة كافيه، وبحجم يسمح بالتعويض عن عدم الاستجابة، وفي حالة وجود استجابة مرتفعة يمكن البحث عن التصميم الذي يعطي أعلى دقة بأقل كلفة.
8. **دورية المسح:** ففي حالة تنفيذ المسح سنويا أو في كل فصل وكان الأمر يتطلب إجراء مقارنة بين السنوات او الفصول فإن تصميم العينة يختلف عن الدراسة التي تُجرى لمرة واحدة فقط.

# الفصـل الأول

# مصطلحات وتعـاريف

**العد الشامل**

**COMPLETE CENSUS**

هو الأسلوب الذي ندرس فيه حالة جميع أفراد المجتمع المبحوث، وعادة يجري العد الشامل لجمع بيانات الإحصاءات الأساسية (Basic Statistics) مثل التعداد السكاني والتعداد الزراعي والتعداد الصناعي، كذلك يجري العد الشامل إذا كان المجتمع المنوي دراسته صغير نسبيا، حيث يكون أسلوب المعاينة غير فعّال، كذلك إذا كان الباحث يجهل طبيعة المجتمع المنوي دراسته.

**أسلوب المعاينة**

**SAMPLING TECHNIQUES**

هو أسلوب علمي يقيس المعلومات الإحصائية الموثوقة وليس بديلاً عن التغطية الشاملة أو عوضا عنها، ويقوم هذا العلم على نظريات الاحتمالات، ويمكن القول إنه أسلوب علمي تجمع بواسطته بيانات عن جزء محدد من المجتمع المنوي دراسته وهذا الجزء يسمى عينه (Sample).

**إطار المعاينة**

**SAMPLING FRAME**

هو القائمة التي تشمل جميع وحدات المعاينة ويمكن أن تكون وحدات معرفة أو خريطة توضح وحدات المعاينة، وهو القاعدة التي تسحب منها العينات.

**تأثير التصميم**

**DESIGN EFFECT**

هو نسبة تباين أحد التقديرات في تصميم عينة معينة إلى تباين عينة عشوائية بسيطة من نفس الحجم.

| | |
|---|---|
| **تحليل التباين**<br>**ANALYSIS OF VARIANCE** | هو عملية حساب التباين الناتج عن كل مصدر من مصادر التباين بشكل مستقل مثل حساب التباين بين العناقيد والتباين داخل كل عنقود في العينة العنقودية من مرحلتين. |
| **التحيز**<br>**BIAS** | هو انحراف متوسط جميع التقديرات الممكنة لدليل المجتمع عن قيمته الحقيقية. |
| **الترجيح أو التوزين**<br>**WEIGHTING** | هو إجراء العمليات الحسابية للوصول إلى المعلمات المقدرة للمجتمع بإستخدام نتائج المسح بالعينة. |
| **تصميم العينة**<br>**SAMPLE DESIGN** | هي عملية اختيار التوليفة المناسبة من عدة أأنواع من العينات للوصول الى العينة التي تحقق النتائج المرجوة منها |
| **تقدير**<br>**ESTIMATION** | هو عبارة عن قيمة رقمية يجري حسابها من البيانات بإستخدام العينة بدلا من معلمة المجتمع المراد تقديرها والتي تكون غير معروفة. |
| **التقدير بفترة**<br>**INTERVAL ESTIMATE** | هو مدى من القيم المستخدمة لتقدير معلمة من معلمات المجتمع. |

**التقدير بنقطة**

**POINT ESTIMATE**

هو تقدير معلمة من معلمات المجتمع بقيمة واحدة.

**التقديرات المتحيزة**

**BIAS ESTIMATION**

هي التقديرات التي تعطي قيم متوقعة لمعلمات المجتمع المقدرة بإستخدام أسلوب معاينة مختلفة عن القيم الحقيقية للمجتمع مثل استخدام أسلوب التقدير النسبي.

**التقدير النسبي**

**RATIO ESTIMATION**

هو احد أساليب التقدير المستخدمة بحيث يستخدم معلومات سابقة للعينة والإطار وقد تكون متوفرة من خلال تعداد او مصدر آخر للمعلومات وذلك بهدف تحسين النتائج ويعتبر هذا الأسلوب من الأساليب التي تعطي تقديرات متحيزة.

**الثقة في تقديرات العينة**

**CONFIDENCE IN SAMPLE ESTIMATE**

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

**حدود الثقة**

**CONFIDENCE INTERVAL**

هي المدى الذي تقع فيه القيمة الحقيقية لمعلمة ما من معلمات المجتمع باحتمال معين.

**حد الخطأ**

**BOUND OF ERROR**

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

**خطأ المعاينة**

**SAMPLING ERROR**

هو الخطأ الناشيء عن اختلاف وحدات المعاينة المسحوبة في العينة فيما بينها وهو موجود في الدراسات التي تجرى بالعينة، ولا يمكن التخلص منه ولكن يمكن تقليله عن طريق زيادة حجم العينة وتحسين تصميم العينة.

**الخطأ المعياري**

**STANDARD ERROR**

هو الجذر التربيعي لتباين العينة المقدر.

**خطأ المعاينة**

**SAMPLE ERROR**

هو الفرق بين تقديرات العينة ومعلمات المجتمع الحقيقية نتيجة الحقيقة القائلة إن العينة لا يمكن أن تعطي نفس نتائج المسح الشامل.

**الخطأ المعياري النسبي أو ما يسمى معامل التباين**

**Relative Standard Error**

هو الخطأ المعياري مقسوما على المعلمة المحسوب لها الخطأ المعياري مضروبا بـ 100% وهو مساوٍ لمعامل التباين

**دقة المسح بالعينة**

**Accuracy of the Sample Survey**

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

**طبقية بعدية**

**POST STRATIFICATION**

هو أحد الأساليب الإحصائية التي تستخدم في حالة اكتشاف ظاهرة معينة لها علاقة بالصفة المدروسة ولم تكن قد أخذت بالإعتبار عند تقسيم المجتمع إلى طبقات، ويؤدي هذا الأسلوب إلى الحصول على تقديرات متحيزة.

**كسر المعاينة**

**SAMPLING FRACTION**

هو حاصل قسمة عدد وحدات المعاينة المسحوبة في العينة على إجمالي وحدات المعاينة الموجودة في المجتمع.

**المجتمع المستهدف**

**TARGET POPULATION**

هو جزء من المجتمع الكلي وهو المجتمع المعني بالدراسة مباشرة مثل النساء اللواتي سبق لهن الزواج في دراسة الخصوبة ويكون الوصول اليهن من خلال مجتمع الأسر.

**المسح الإحصائي**

**STATISTICAL SURVEY**

هو أسلوب الحصول على معلومات عن مجموعة كبيرة من المجتمع عن طريق أخذ مجموعة قليلة من مفردات المجتمع في حالة استخدام العينة، أو أخذ جميع مفردات المجتمع للمسح الشامل.

**المسح الاستطلاعي**

**PILOT SURVEY**

هو المسح الذي يجرى على عينة صغيرة بهدف اختبار الاستمارات وخطة العمل ويمكن استخدامه لتقدير تباين المجتمع بهدف تقدير حجم العينة إذا لم يكن ذلك معروفا للصفة المراد دراستها.

**معامل التكبير**

**RAISING FACTOR**

هو الوزن الذي يضرب في كل صفة لكل متغير في العينة بهدف الوصول إلى تقديرات على مستوى المجتمع.

**المُقدر (بكسر الدال)**

**ESTIMATOR**

هو الصيغة الرياضية أو القاعدة التي تستخدم نتائج العينة للحصول على تقدير معلمة المجتمع قيد البحث.

**اتساق التقديرات**

**CONSISTENT ESTIMATE**

هي العينة التي تحقق نتائج للصفة المراد دراستها ضمن حدود ثقة محددة مسبقا بغض النظر عن حجم وتصميم العينة وتتسم التقديرات المتسقة بأنها كلما زاد حجم العينة فإن تقديرات العينة تقترب من المعلمات الحقيقية في المجتمع.

**اتساق المقدر**

**CONSISTENT ESTIMATOR**

هو المقدر الذي يعطي قيم قريبة لمعلمات المجتمع كلما زاد حجم العينة.

**الاختيار الضابط**

**CONTROLLED SELECTION**

هي عملية التحكم في توزيع العينة بين الطبقات بحيث تعطي نتائج متسقة، وقد يكون ذلك بإستخدام كسر معاينة متغير من طبقة إلى أخرى.

**أخطاء عدم الإستجابة**

**NON RESPONSE ERROR**

هي عدم استجابة بعض المستجيبين في العينة للإجابة على الإستمارة ككل نتيجة الرفض أو أية أسباب أخرى، وتعتبر هذه الحالة عدم استجابة كلية لوحدة المعاينة، وهناك عدم استجابة جزئية عندما يرفض المستجيب الإجابة على أسئلة معينة دون غيرها مثل الأسئلة المتعلقة بالدخل.

**الارتباط داخل الفئة**

**INTRACLASS CORRELATION**

وهو مقياس لدرجة تجانس الوحدات الثانوية داخل الوحدات الأولية للصفة المدروسة.

**أسلوب المعاينة**

**SAMPLING TECHNIQUES**

هو أسلوب علمي يقيس المعلومات الإحصائية الموثوقة وليس بديلا عن التغطية الشاملة أو عوضا عنها، ويقوم هذا العلم على نظريات الاحتمالات، ويمكن القول أنه أسلوب علمي تجمع بواسطته بيانات عن جزء محدد من المجتمع المنوي دراسته وهذا الجزء يسمى عينه (Sample).

**إطار المعاينة**

**SAMPLING FRAME**

هو القائمة التي تشمل جميع وحدات المعاينة ويمكن أن تكون وحدات معرفة أو خريطة توضح وحدات المعاينة، وهو القاعدة التي تسحب منها العينات .

**تأثير التصميم**

**DESIGN EFFECT**

هو نسبة تباين أحد التقديرات في تصميم عينة معينة إلى تباين عينة عشوائية بسيطة من نفس الحجم.

**تحليل التباين**

**ANALYSIS OF VARIANCE**

هو عملية حساب التباين الناتج عن كل مصدر من مصادر التباين بشكل مستقل مثل حساب التباين بين العناقيد والتباين داخل كل عنقود في العينة العنقودية من مرحلتين.

**التحيز**

**Bias**

هو انحراف متوسط جميع التقديرات الممكنة لدليل المجتمع عن قيمته الحقيقية.

**الترجيح أو التوزين**

**Weighting**

هو إجراء العمليات الحسابية للوصول إلى المعلمات المقدرة للمجتمع بإستخدام نتائج المسح بالعينة.

**تصميم العينة**

**Sample Design**

هي عملية اختيار التوليفة المناسبة من عدة أنواع من العينات للوصول إلى العينة التي تحقق النتائج المرجوة منها

**تقدير**

**ESTIMATION**

هو عبارة عن قيمة رقمية يجري حسابها من البيانات بإستخدام العينة بدلاً من معلمة المجتمع المراد تقديرها والتي تكون غير معروفة.

**التقدير بفترة**

**Interval Estimate**

هو مدى من القيم المستخدمة لتقدير معلمة من معلمات المجتمع.

**التقدير بنقطة**

**POINT ESTIMATE**

هو تقدير معلمة من معلمات المجتمع بقيمة واحدة.

**التقديرات المتحيزة**

**BIAS ESTIMATION**

هي التقديرات التي تعطي قيم متوقعة لمعلمات المجتمع المقدرة بإستخدام أسلوب المعاينة مختلفة عن القيم الحقيقية للمجتمع مثل استخدام أسلوب التقدير النسبي.

**التقدير النسبي**

**RATIO ESTIMATION**

هو أحد أساليب التقدير المستخدمة بحيث يستخدم معلومات سابقة للعينة والإطار وقد تكون متوفرة من خلال تعداد أو مصدر آخر للمعلومات وذلك بهدف تحسين النتائج ويعتبر هذا الأسلوب من التقديرات متحيزاً.

**التوزيع الأمثل**

**OPTIMUM ALLOCATION**

هو أحد أساليب توزيع وحدات المعاينة في العينة الطبقية على مختلف الطبقات بحيث يكون نصيب كل طبقة متناسبا تناسباً طردياً مع حجم الطبقة والتباين داخل كل طبقة ويتناسب عكسيا مع كلفة جمع بيانات لوحدة المعاينة في تلك الطبقة.

**توزيع نيمان**

**NYMAN ALLOCATION**

أحد أساليب توزيع وحدات المعاينة في العينة الطبقية على مختلف الطبقات بحيث يكون نصيب كل طبقة متناسباً تناسباً طردياً مع حجم الطبقة والتباين داخل الطبقة.

**التوزيع المتناسب**

**PROPORTIONAL ALLOCATION**

هو أحد أساليب توزيع وحدات المعاينة في العينة الطبقية على مختلف الطبقات بحيث يكون نصيب كل طبقة متناسبا مع حجم تلك الطبقة من حيث عدد وحدات المعاينة.

**الثقة في تقديرات العينة**

**Confidence in Sample Estimate**

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

**حدود الثقة**

**CONFIDENCE INTERVAL**

هي المدى الذي تقع فيه القيمة الحقيقية لمعلمة ما من معلمات المجتمع باحتمال معين.

**حد الخطأ**

**BOUND OF ERROR**

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

**خطأ المعاينة**

**SAMPLING ERROR**

هو الخطأ الناشيء عن اختلاف وحدات المعاينة المسحوبة في العينة فيما بينها وهو موجود في الدراسات التي تجرى بالعينة، ولا يمكن التخلص منه ولكن يمكن تقليله عن طريق زيادة حجم العينة وتحسين تصميم العينة.

**الخطأ المعياري**

**STANDARD ERROR**

هو الجذر التربيعي لتباين العينة المقدر.

**خطأ المعاينة**

**SAMPLE ERROR**

هو الفرق بين تقديرات العينة ومعلمات المجتمع الحقيقية نتيجة الحقيقة القائلة إن العينة لايمكن أن تعطي نفس نتائج المسح الشامل.

**الخطأ المعياري النسبي**

**RELATIVE STANDARD ERROR**

هو الخطأ المعياري مقسوما على المعلمة المحسوب لها الخطأ المعياري وهو مساوٍ لمعامل التباين.

**دقة المسح بالعينة**

**ACCURACY OF THE SAMPLE SURVEY**

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

**طبقية بعدية**

**POST STRATIFICATION**

هو أحد الأساليب الإحصائية التي تستخدم في حالة اكتشاف ظاهرة معينة لها علاقة بالصفة المدروسة ولم تكن قد أخذت بالإعتبار عند تقسيم المجتمع الى طبقات، ويؤدي هذا الأسلوب إلى الحصول على تقديرات متحيزة.

**ظروف الإستجابة**

**CONDITIONING OF RESPONDENTS**

هي أحد أنواع الخلل الذي يحدث نتيجة استجابة شخص معين في الأسرة غير الشخص الذي سحب في العينة، وهذا يحدث عادة في حالة أخذ بديل عن الشخص العشوائي الذي سحب من الأسرة نتيجة عدم تواجد الشخص المسحوب في العينة وقت المقابلة، وهذا يؤدي إلى تحيز في العينة لصالح الأشخاص الذين يتواجدون في البيت خلال النهار أو وقت الزيارة على حساب الأشخاص الذين يكونون خارج البيت مثل العمل أو الدراسة، وهذا يؤدي الى تحيز في النتائج. واحيانا تحدث هذه المشكلة نتيجة أخذ أسرة بديلة عن أسرة غير متواجدة وقت الزيارة.

**العد الشامل**

**COMPLETE CENSUS**

هو الأسلوب الذي يتم من خلاله دراسة جميع أفراد المجتمع المبحوث، وعادة يجري العد الشامل في الإحصاءات الأساسية (Basic Statistics) مثل التعداد السكاني والتعداد الزراعي والتعداد الصناعي، كذلك يجري العد الشامل إذا كان المجتمع المنوي دراسته صغير نسبيا حيث يكون أسلوب المعاينة غير فعال، كذلك إذا كان الباحث يجهل طبيعة المجتمع المنوي دراسته.

**عدم الإستجابة الجزئي**

**PARTIAL NON RESPONSE**

هو عدم الحصول على إجابة لسؤال واحد أو أكثر من الأسئلة الواردة في الإستمارة مع وجود اجابة أسئلة أخرى في الإستمارة.

**عدم الإستجابة الكلي**

**NON RESPONSE**

هو عدم القدرة على إجابة أي جزء في الإستمارة قيد الدراسة نتيجة الرفض أو عدم تواجد المستجيب أو عدم وجود شخص مؤهل يدلي بالبيانات وتكون الاستمارة فارغة.

**عينة متعددة الأغراض**

**MULTI-PURPOSE SAMPLE**

هي العينة التي يتم من خلالها جمع بيانات لعدة مواضيع في مسح واحد مثل الدخل والإنفاق والصحة والتغذية وغيرها، وعادة ما يكون هناك ترابط في المواضيع قيد البحث.

**العينة المتوالية**

**SUCCESSIVE SAMPLE**

هو أحد تصاميم العينات الذي يغطي المجتمع على عدة سنوات حيث يقسم المجتمع الى عدة أجزاء تكون متقاربة في الحجم، ويغطى كل جزء في سنة ويستخدم هذا التصميم عادة في المجتمعات الصغيرة وكبديل عن التعدادت في بعض الأحيان وضمن شروط خاصة.

**العينة الموزونة ذاتيا**

**Self -Weighting Designs**

هو أي تصميم من تصاميم العينات الذي يكون فيه الوزن (معامل التكبير) لكل وحدة من وحدات المعاينة المسحوبة في العينة مساويا لوزن وحدات المعاينة الأخرى.

**كسر المعاينة**

**SAMPLING FRACTION**

هو حاصل قسمة عدد وحدات المعاينة المسحوبة في العينة على إجمالي وحدات المعاينة الموجودة في المجتمع.

**متغير**

**VARIABLE**

صفة أو ميزة يمكن أن تحمل قيم مختلفة.

| | |
|---|---|
| **متغير عشوائي**<br>**RANDOM VARIABLE** | هو المتغير الذي تظهر قيمه عن طريق الصدفة (بعشوائية). |
| **مجتمع الدراسة**<br>**POPULATION** | هو المجال المراد دراسته وقد يكون مجتمع من أفراد أو أسر أو منشآت اقتصادية أو وحدات زراعية وغيرها. |
| **المجتمع المستهدف**<br>**TARGET POPULATION** | هو جزء من المجتمع الكلي وهو المجتمع المعني بالدراسة مباشرة مثل النساء اللواتي سبق لهن الزواج في دراسة الخصوبة ويكون الوصول اليهن من خلال مجتمع الأسر. |
| **مجتمع دوري**<br>**PERIODIC POPULATION** | هو مجتمع تأخذ عناصره الشكل الدوري لصفة معينة. |
| **المجتمع المرتب**<br>**ORDERED POPULATION** | هو المجتمع الذي تكون عناصره مرتبة بناء على مقادير لصفة معينة تحملها تلك العناصر مثل الدخل. |
| **المجتمع العشوائي**<br>**RANDOM POPULATION** | هو المجتمع الذي لا تتبع مفرداته ترتيبا معينا. |

**المحدد المكاني**

**ADMINISTRATIVE RESTRICTION**

هي أحد العيوب التي قد تحدث عند سحب عينة مكانية حيث يتم سحب عينة من مناطق معينة واستثناء مناطق أخرى لعدم القدرة على سحب عينة منها، مثل المناطق المكتظة بالسيارات. أو عدم القدرة على سحب عينة سيارات لمقابلة السائقين عند الإشارة الضوئية، أو المناطق الحدية في الخرائط.

**المستجيب**

**RESPONDENT**

هو الشخص أو الجهة التي تدلي بالبيانات أو المعلومات المطلوبة لملئ الاستمارة.

**المسح الإحصائي**

**STATISTICAL SURVEY**

هو أسلوب الحصول على معلومات عن مجموعة كبيرة من المجتمع عن طريق أخذ مجموعة قليلة من مفردات المجتمع في حالة استخدام العينة، أو أخذ جميع مفردات المجتمع للمسح الشامل.

**المسح الاستطلاعي**

**PILOT SURVEY**

هو المسح الذي يجرى على عينة صغيرة بهدف اختبار الاستمارات وخطة العمل ويمكن استخدامه لتقدير تباين المجتمع بهدف تقدير حجم العينة إذا لم يكن ذلك معروفاً للصفة المراد دراستها.

**المسوحات متعددة الأطر**

**MULTI FRAME SURVEY**

هي تصميمات العينات الاحتمالية التي تجمع بين أكثر من تصميم للعينة للحصول على تقديرات مسح تجمع بين تصميمات إطار المساحة، وتصميمات إطار القائمة. ويتألف المسح متعدد الأطر عادة من عنصر عينة المساحة وعنصر عينة القائمة.

**مصداقية نتائج المسح بالعينة**

**Reliability of the Sample Result**

هو الفرق بين نتائج المسح بالعينة للصفة المدروسة، ونتائج المسح الشامل لنفس الصفة المدروسة وتحت نفس الظروف، وكلما كان الفرق أقل كلما كانت المصداقية أكبر

**المعاينة الاحتمالية**

**PROBABILITY SAMPLE**

هو أحد أساليب المعاينة ويستخدم قوانين الاحتمالات عند اختيار العينات، بحيث يكون لكل وحدة معاينة في المجتمع احتمال ظهور في العينة، وتتميز العينة الإحتمالية بأنها تحد من خطورة التحيز وتسمح بأن يستدل من خلال العينة على المجتمع الذي سحبت منه، وعادة يستخدم هذا الأسلوب في الأجهزة الإحصائية.

**المعاينة غير الاحتمالية**

**NON PROBABILITY SAMPLE**

هو أحد أساليب المعاينة والذي يتم اختيار وحدات المعاينة بطريقة متعمدة بناءاً على الحكم الشخصي أو الحدس للشخص الذي يختار العينات، ويسمى أحيانا العينة العمدية، حيث يتم اختيار وحدات المعاينة التي يتوقع أن تكون من متوسط المجتمع، وهذا النوع من العينات يستخدم لبناء موقف معين (أخذ فكره) ولا يمكن تعميم النتائج، وإذا كان الشخص الذي يختار العينات له خبرة جيدة فإنه قد يخرج بنتائج جيده، ولكن لا يمكن الاعتماد على تلك النتائج في أية تقديرات.

**المقِدر (بكسر الدال)**

**ESTIMATOR**

هو الصيغة الرياضية أو القاعدة التي تستخدم نتائج العينة للحصول على تقدير معلمة المجتمع قيد البحث.

**المعلمة**

**PARAMETER**

هي مؤشر من المؤشرات الدالة على صفة أو مقياس في المجتمع أو هي كمية تحسب من جميع قيم مفردات المجتمع

**المعاينة المزدوجة**

**DOUBLE SAMPLING**

هو أحد الأساليب المستخدمة في العينات بحيث يسحب جزء من العينة ويستخدم كعينة استطلاعية، وبعد جمع البيانات وتقدير حجم العينة يجري سحب بقية العينة بحيث يشكل كلا الجزئين العينة.

**المعاينة متعددة الأطوار**

**MULTI-PHASE SAMPLE**

هي الطريقة التي يمكن أن تستخدم بإقترانها ببعض الأنواع من المعاينة الاحتمالية والمعاينة غير الاحتمالية، وقد يكون من المناسب جمع بعض البيانات من جميع وحدات المعاينة وكذلك جمع بعض المعلومات لعينة فرعية فقط من العينة الكلية، وتعرف هذه الطريقة باسم المعاينة ذات الطورين ويمكن التوسع في هذا المبدأ الى ثلاثة أطوار أو أكثر.

**وحدات المعاينة**

**SAMPLING UNIT**

هي الأجزاء التي يتكون منها المجتمع ومن الممكن أن تكون وحدات طبيعية مثل السكان أو تجمعات طبيعية مثل الأسر، أو وحدات صناعية مثل المزارع. ويشترط أن تكون معرفة غير متداخلة، ويعتبر كل عنصر وحدة معاينة واحدة فقط، بحيث يكون لكل عنصر في المجتمع احتمال الظهور في وحدة معاينة واحدة، مثال ذلك إذا كانت وحدة المعاينة هي الأسرة فيجب أن يكون كل فرد منتمياً لأسره واحدة (لا يوجد فرد ينتمي إلى أسرتين في نفس الوقت أو لا ينتمي لأي أسره).

**وحدات المعاينة الأولية**

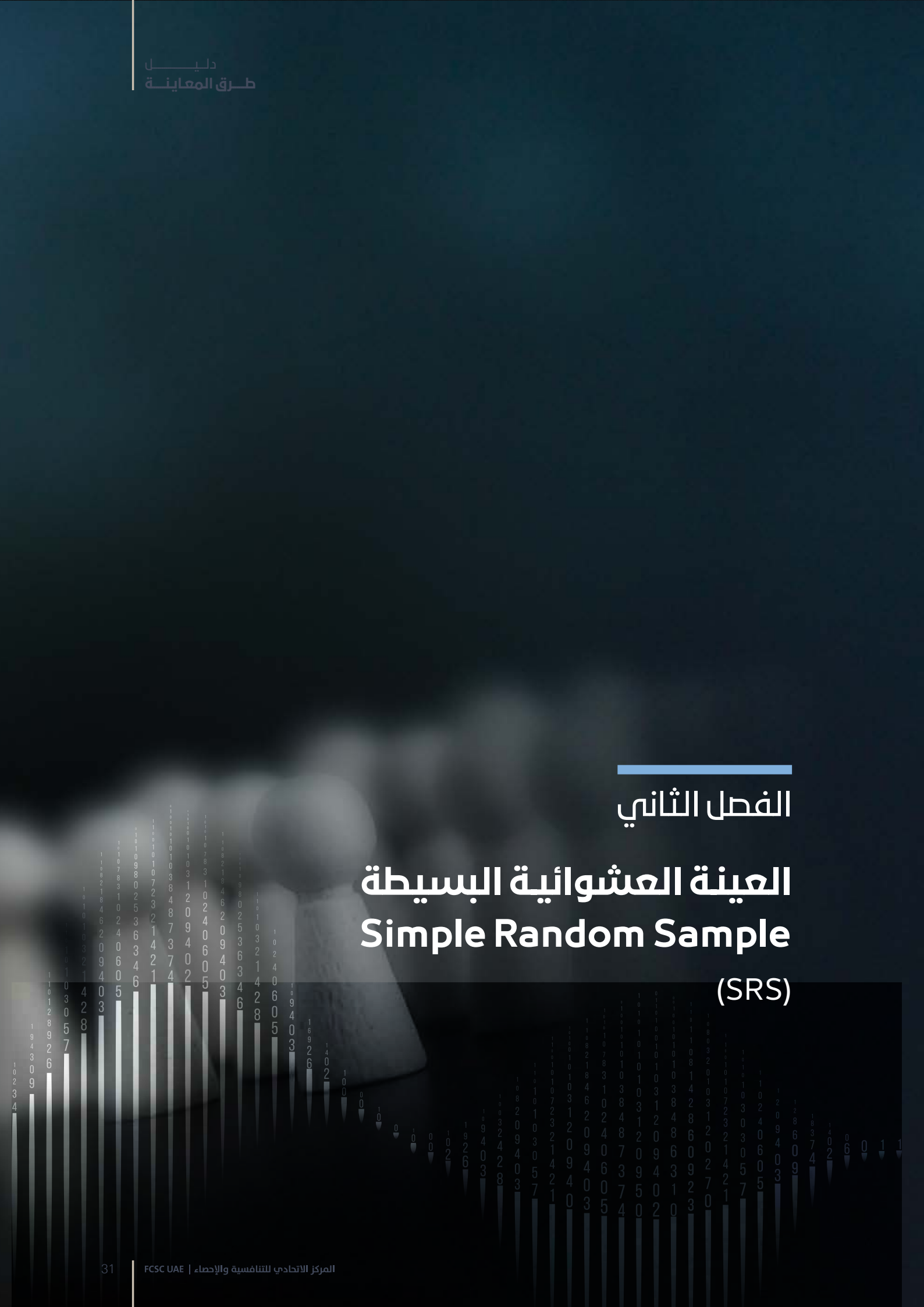
**PRIMARY SAMPLING UNIT**

حي وحدات المعاينة التي تسحب في المرحلة الاولى لعينة متعددة المراحل وغالبا ما تمثل كل وحدة معاينة عنقود وتتكون كل وحدة معاينة اولية من عدد من وحدات المعاينة الثانوية.

**وحدات المعاينة الثانوية**

**SECONDARY SAMPLING UNIT**

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



الفصل الثاني

# العينة العشوائية البسيطة
# Simple Random Sample
# (SRS)

## مقدمة

يعتبر أسلوب المعاينة العشوائية البسيطة من أبسط طرق المعاينة وأكثرها انتشارا، وهو أسلوب المعاينة الذي يعطي احتمالات متساوية لاختيار كل وحدة من وحدات المعاينة الموجودة في المجتمع في السحبة الأولى، وكذلك احتمالات متساوية لاختيار أي وحدة من وحدات المعاينة الممكن سحبها في السحبات التالية.

في حالة سحب العينة وإرجاعها إلى المجتمع بعد عملية السحب لإعطائها فرصة الظهور مرة أخرى يسمى هذا الأسلوب أسلوب سحب العينة مع الإرجاع، وفي حالة سحب العينة واستثناء كل وحدة معاينة مسحوبة من المجتمع يسمى سحب عينة دون إرجاع، حيث إن هذا الأسلوب لا يعطي فرصة لتكرار ظهور وحدة المعاينة الواحدة أكثر من مرة.

## شروط استخدام العينة العشوائية البسيطة

نظرا لسهولة هذا الأسلوب وبساطته فهو منتشر بشكل كبير، ولكن قبل استخدام هذا الأسلوب يجب ملاحظة أن هناك شروطاً يجب أن تتوفر في وحدات المعاينة للمجتمع المستهدف في الدراسة، فمثلا 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 الامارة او الاقليم فلا يجوز سحب عينة عشوائية بسيطة على مستوى الدولة، كذلك وجود محدد التكاليف والإمكانيات الأخرى مثل توفر السيارات والقوى العاملة المدربة يحد من استخدام هذا الأسلوب، بالإضافة إلى فعّالية هذا الأسلوب بالنسبة لوحدة التكاليف قد يكون اقل من معظم أنواع العينات الأخرى، لذا يجب أخذ جانب الحيطة عند استخدام أسلوب المعاينة العشوائية البسيطة.

## عدد العينات التي يمكن سحبها

لو كان لدينا صف في إحدى المدارس يتكون من 6 طلاب وأردنا سحب عينة تتكون من طالبين لأجراء فحص المناعة لدى الطلاب فان عدد العينات التي يمكن سحبها دون إرجاع هي (الأول، الثاني)، (الأول، الثالث)، (الأول،الرابع)، (الأول،الخامس)، (الأول، السادس)، (الثاني،الثالث)، (الثاني، الرابع)، (الثاني،الخامس)، ( الثاني، السادس)، (الثالث، الرابع)،(الثالث، الخامس)، (الثالث، السادس)، (الرابع، الخامس)، (الرابع، السادس)، (الخامس، السادس) وبذلك يكون عدد العينات التي يمكن سحبها هو 15 عينة، ويمكن حساب عدد العينات بإستخدام المعادلة التالية:

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!} = \frac{6!}{2!\,(6-2)!} = 15 \qquad (1)$$

# أسلوب اختيار العينة العشوائية البسيطة

### 1. الكيس المثالي:

يتلخص هذا الأسلوب بكتابة أسماء أو أرقام لجميع وحدات المعاينة في المجتمع على بطاقات متشابهة تماما وتوضع جميع البطاقات في كيس مثالي (ليس له حواف) وتخلط بشكل جيد ثم يتم اختيار عدد من البطاقات بعدد وحدات المعاينة المنوي سحبها من المجتمع بشرط أن لا يدخل فيها أي نوع من أنواع التحيز أثناء الاختيار، وهذا الأسلوب بدائي وغير قابل للتطبيق خاصة في حالة وجود مجتمع كبير نسبيا مما يعني استحالة تطبيق هذه الطريقة، لذا فقد ابتكرت طرق أخرى.

### 2. الجداول العشوائية:

يتكون الجدول العشوائي من عدد من الأرقام، كل رقم يتكون من نفس العدد من الخانات، أما أن تكون من خانة واحدة أو من خانتين أو من ثلاث خانات...الخ .

ويكون كل رقم موجود في الجدول بشكل مستقل وغير مرتبط بأي رقم آخر، تكون فرصة الظهور لأي رقم في الجدول مساوية لفرصة الظهور لأي رقم آخر، (من الأمثلة على ذلك جداول كيش العشوائية).

### 3. استخدام برنامج اكسل لسحب العينة العشوائية البسيطة

يجري سحب العينة عن طريق كتابة المعادلة التالية :

**rand ()*N =** حيث **N**: عدد وحدات المعاينة في المجتمع

مثال إذا كان لدينا مجتمع يتكون من **180** وحدة معاينة فان المعادلة تكتب بالطريقة التالية: **=rand *180**

الاجابة **103.5097**

لذا فإن وحدة المعاينة 104 هي التي تختار في العينة

# تقدير مَعلَمات المجتمع

### تقدير مجموع المجتمع ومتوسط المجتمع لصفة معينة

بعد ان تُجرى عملية سحب العينة فإن الخطوة التالية تكون هي تقدير مَعلمات المجتمع بالاعتماد على المشاهدات لوحدات المعاينة المسحوبة، إن المَعلمات الشائعة التي يجري تقديرها هي متوسط المجتمع، ومجموع المجتمع، ونسبة الذين يحملون صفة معينة في المجتمع.

### أولاً :التقدير بقيمة واحدة

لتقدير متوسط صفة معينة (متغير ) للمجتمع أو مجموع تلك الصفة للمجتمع.
لنفرض أن:

| | |
|---|---|
| N = | عدد وحدات المعاينة في المجتمع (عدد مفردات المجتمع). |
| n = | عدد وحدات المعاينة في العينة المسحوبة (حجم العينة). |
| Y = | إجمالي الصفة المدروسة في المجتمع (مثل إجمالي الدخل، المساحة المزروعة...الخ). |
| $Y_i$ = | قيمة الصفة المدروسة لوحدة المعاينة i في المجتمع. |
| $\bar{Y}$ = | متوسط المجتمع للصفة المدروسة. |
| $\hat{\bar{Y}}$ = | المتوسط المقدر للمجتمع. |
| $V(\bar{Y})$ = | تباين المجتمع للصفة المدروسة. |
| $\hat{v}(\bar{y}_n)$ = | التباين المقدر للمجتمع. |
| S. E = | الخطأ المعياري (خطأ المعاينة). |
| S.E% = | الخطأ المعياري النسبي (خطأ المعاينة النسبي) أو معامل التباين (C.V %). |

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \quad (2)$$

$$\hat{\bar{Y}} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (3)$$

$$v(\bar{y}) = \frac{N-n}{N-1} \times \frac{s^2}{n} \quad (4)$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 \quad (5)$$

**لتقدير تباين المتوسط والخطأ المعياري بإستخدام مشاهدات وحدات المعاينة في العينة**

$$\hat{v}(\bar{y}_n) = \left(\frac{N-n}{N}\right) \times \frac{s^2}{n} = \left(\frac{1}{n} - \frac{1}{N}\right) s^2_{n-1} = \frac{1-f}{n} s^2 \quad (6)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}_n)^2 \quad (7)$$

$$S.E.(\bar{y}) = \sqrt{\left(\frac{1}{n} - \frac{1}{N}\right) s^2} \quad (8)$$

$$S.E.\% = \frac{\sqrt{\hat{v}(\bar{y})}}{\bar{y}} \times 100\% \quad (9)$$

**يعتبر تقدير المتوسط والتباين بأسلوب المعاينة العشوائية البسيطة تقديرات غير متحيزة**

**مثال 1:**

في دراسة لتقدير متوسط إنتاج الدونم من محصول القمح في إحدى القرى سحبت عينه حجمها 27 حقل من أصل 431 حقل موجود في تلك القرية ، وقد كانت نتائج تلك المشاهدات للعينة كما يلي:

| الرقم المتسلسل للحقل | متوسط إنتاج الحقل كغم \دونم ( yi) | الرقم المتسلسل للحقل | متوسط إنتاج الحقل كغم \دونم ( yi) |
|---|---|---|---|
| 1 | 100.0 | 15 | 270.0 |
| 2 | 320.0 | 16 | 276.0 |
| 3 | 320.0 | 17 | 216.0 |
| 4 | 332.0 | 18 | 216.0 |

| الرقم المتسلسل للحقل | متوسط إنتاج الحقل كغم \دونم ( yi) | الرقم المتسلسل للحقل | متوسط إنتاج الحقل كغم \دونم ( yi) |
|---|---|---|---|
| 5 | 160.0 | 19 | 360.0 |
| 6 | 328.0 | 20 | 120.0 |
| 7 | 296.0 | 21 | 216.0 |
| 8 | 301.2 | 22 | 180.0 |
| 9 | 224.0 | 23 | 180.0 |
| 10 | 208.0 | 24 | 100.0 |
| 11 | 180.0 | 25 | 180.0 |
| 12 | 408.0 | 26 | 20.0 |
| 13 | 192.0 | 27 | 100.0 |
| 14 | 325.2 | | |
| المجموع | | 6128.4 | |

**الإجابة:**

عدد الحقول في المنطقة (N)= 431

عدد الحقول في العينة (n)= 27

المتوسط المقدر للمجتمع = متوسط العينة $\bar{y}$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{6128.4}{27} = 227.0$$

ولتقدير التباين

**1. تقدير متوسط مربعات الفروق في العينة** ( $s^2$ )

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{5513.8}{26} = 212$$

التباين المقدر للمتوسط $\hat{v}(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)s^2 = \left(\frac{1}{27} - \frac{1}{431}\right) \times 212.07 = 6.65$

الخطأ المعياري $S.E. = \sqrt{\hat{v}(\bar{y})} = 2.7$

الخطأ المعياري النسبي (خطأ المعاينة النسبي) $S.E.\% = \frac{\sqrt{\hat{v}(\bar{y})}}{\bar{y}} \times 100\% = 1.2\%$

**ثانياً: التقدير بحدود الثقة**

- معرفة حجم العينة هل هو كبير أو صغير نسبيا.
- إذا كان حجم العينة كبير نسبيا فإن

$$p(\bar{Y} - 1.96\ S_{\bar{y}} \leq \bar{y} \leq \bar{Y} + 1.96\ S_{\bar{y}}) = 0.95$$

وعليه فإن متوسط الصفة المدروسة باحتمال وقوع المتوسط الحسابي الحقيقي للمجتمع بنسبة 95%، وهذا يعني أننا واثقون بنسبة 95% من أن المتوسط الحسابي الحقيقي للمجتمع يقع ضمن حدود الثقة التالية:

$$\bar{y} \mp z \times S_{\bar{y}} \qquad (10)$$

إذا كان حجم العينة صغير نسبيا فإن حدود الثقة تحدد على النحو التالي:

$$\bar{y} \mp t \times s_{\bar{y}} \quad (11)$$

حيث أن **t** الجدولية المناظرة لدرجات الحرية **(n-1)** ومستوى معنوية $\alpha$ = (1-confidence level)
ويمكن كتابتها بالصيغة التالية:

$$\bar{y} \mp t \times s \sqrt{\left(\frac{1}{n} - \frac{1}{N}\right)} \quad (12)$$

**مثال 2:**

في المثال 1 إذا طلب منا حساب حدود الثقة التي يقع فيها متوسط المجتمع الحقيقي باحتمال 95% هو:

$$P\left(\bar{Y} - 1.96S_{\bar{y}} \leq \bar{y} \leq \bar{Y} + 1.96S_{\bar{y}}\right) = 0.95$$

$$S_{\bar{y}} = s$$

$$\bar{y} \mp 1.96 \times s \sqrt{\left(\frac{1}{n} - \frac{1}{N}\right)}$$

$$= 227 \mp 28.54 \sqrt{\left(\frac{1}{27} - \frac{1}{431}\right)}$$

$$= 227 \mp 5.32$$

وعليه فإننا 95% واثقون أن المتوسط الحقيقي لإنتاج الدونم من القمح في منطقة زراعية يتراوح بين 221.68 و 232.32 كغم لكل دونم.

### ب. تقدير النسبة الحقيقية للمجتمع:

عندما يقسم المجتمع إلى مجموعتين بحيث تكون المجموعة الأولى تحمل صفة مختلفة عن المجموعة الثانية او احدى المجموعتين تحمل الصفة والأخرى لا تحمل تلك الصفة، وتكون احتمالية ظهور الصفة على كل مفرده من مفردات المجتمع إما (1) أو صفر ، مثل الإصابة بمرض معين أو نسبة الأمية في المجتمع أو نسبة الذكور في المجتمع ...الخ .

#### 1. التقدير بقيمة واحدة فقط

نفرض **Ni** عدد الوحدات التي تحمل الصفة من المجتمع الذي يتكون من **N** وحدة

**P:** نسبة وحدات المعاينة التي تحمل الصفة في المجتمع

$$P = \frac{N_i}{N} \qquad (13)$$

**Q:** نسبة وحدات المعاينة التي لا تحمل الصفة في المجتمع

$$Q = I - P \qquad (14)$$

$n_i$: عدد وحدات المعاينة التي تحمل الصفة في العينة

$$p = \frac{n_i}{n}$$

$$q = 1 - p$$

وحيث أن العينة تقدر المجتمع تقديراً غير متحيز فإن:

$$\hat{p} = p = \bar{y}_n$$

$$v(p) = \frac{PQ}{n} \quad \frac{\text{N-n}}{\text{N-1}}$$

$$\hat{v}(p) = \frac{pq}{n-1} \quad \frac{\text{N-n}}{N}$$

**مثال 3:**

لتقدير نسبة الأميين من حائزي الثروة الحيوانية أخذت عينة عشوائية بسيطة حجمها 25 حائزاً ووجد أن 10 حائزين أميين من تلك العينة، وعلى افتراض أن المجتمع كبير نسبيا.

**الإجابة:**

**أ. تقدير النسبة الحقيقية للأميين في المجتمع** $(\hat{p})$

$$\hat{p} = p = \bar{y}_n$$
$$= \frac{10}{25} = .4$$

وهذا يعني أن النسبة الحقيقية للأمية لمجتمع حائزي الثروة الحيوانية تقدر بـ 40%.

**ب. التقدير بحدود ثقة 95%**

قيمة t الجدولية عند حدود ثقة 95% ودرجة حريـة 24 هـي

$$t_{0.05,24} = 2.064$$

$$p \mp t\sqrt{\hat{v}(p)} = 0.4 \mp 2.064\sqrt{\frac{0.4(1-0.4)}{25-1}} = 0.4 \mp 0.2$$

وهذا يعني أننا واثقون بنسبة 95% أن النسبة الحقيقية للحائزين الأميين تتراوح بين 20% و 60%.

### 2. تقدير مجموع الوحدات التي تحمل الصفة

لتقدير مجموع الوحدات التي تحمل الصفة (A)

- العدد الحقيقي لمفردات المجتمع التي تحمل الصفة

$$A = \sum_{i=1}^{N} N_i \quad (15)$$

- العدد المقدر لمفردات المجتمع التي تحمل الصفة

$$\hat{A} = NP \quad (16)$$

- تباين العدد المقدر لمفردات المجتمع التي تحمل الصفة ($\hat{A}$) هي

$$V(\hat{A}) = \frac{PQ}{n}\frac{N-n}{N-1}N^2 \quad (17)$$

- التباين المقدر

$$\hat{V}(\hat{A}) = \frac{N(N-n)}{n-1}pq \quad (18)$$

تقدير حجم العينة (n)

$$n = \frac{\frac{t^2 s^2}{p^2 \bar{Y}^2}}{1 + \frac{1}{N} \times \frac{t^2 s^2}{p^2 \bar{Y}^2}} \quad (19)$$

**مثال 4:**

لتقدير حجم العينة للمثال رقم 1 إذا كان الخطأ المسموح به يساوي 5% بإستخدام المعادلة أعلاه.

$$n = \frac{\frac{1.96^2 \times 212.068}{(227)^2 \times .05^2}}{\left[1 + \left(\frac{1}{431} \times \frac{1.96^2 \times 212.068}{(227^2) \times (.05)^2}\right)\right]}$$

فإن عدد الحقول التي يقترح سحبها بخطأ مسموح به 5% هو 7 حقول.
نستطيع الحصول على عدة أحجام للعينات حسب حجم المجتمع ومستوى الدقة وهامش الخطأ بهدف تقدير النسبة الحقيقية للمجتمع على افتراض أن نصف المجمتع يحمل الصفة المراد دراستها والنصف الآخر لا يحمل الصفة.

حجم العينة حسب حجم المجتمع ومستوى الدقة وهامش الخطأ

| حجم المجتمع | حجم العينة | | |
|---|---|---|---|
| | Alpha =.1<br>t = 1.65 | Alpha =.05<br>t = 1.96 | Alpha =.01<br>t = 2.58 |
| 100 | 41 | 80 | 100 |
| 200 | 51 | 132 | 198 |
| 300 | 56 | 169 | 295 |
| 400 | 59 | 196 | 391 |
| 500 | 60 | 218 | 486 |
| 600 | 62 | 235 | 580 |
| 700 | 63 | 249 | 672 |
| 800 | 63 | 260 | 764 |
| 900 | 64 | 270 | 854 |
| 1000 | 64 | 278 | 944 |
| 1500 | 66 | 306 | 1376 |
| 2000 | 66 | 323 | 1786 |
| 4000 | 67 | 351 | 3225 |
| 6000 | 68 | 362 | 4410 |
| 8000 | 68 | 367 | 5403 |
| 10000 | 68 | 370 | 6247 |

# تمارين

## السؤال الأول

مزرعة أبقار تتكون من 100 رأس من الأبقار الحلوب، كان إنتاج هذه الأبقار من الحليب خلال سنة كما يلي:

| الرقم المتسلسل | الإنتاج بالطن | الرقم المتسلسل | الإنتاج بالطن | الرقم المتسلسل | الإنتاج بالطن | الرقم المتسلسل | الإنتاج بالطن |
|---|---|---|---|---|---|---|---|
| 1 | 3.0 | 26 | 9.7 | 51 | 6.3 | 76 | 7.1 |
| 2 | 6.0 | 27 | 8.4 | 52 | 4.8 | 77 | 3.8 |
| 3 | 4.5 | 28 | 5.5 | 53 | 2.7 | 78 | 7.2 |
| 4 | 5.2 | 29 | 8.6 | 54 | 9.0 | 79 | 8.5 |
| 5 | 6.8 | 30 | 9.1 | 55 | 2.9 | 80 | 2.0 |
| 6 | 2.2 | 31 | 6.5 | 56 | 7.1 | 81 | 9.5 |
| 7 | 9.0 | 32 | 9.7 | 57 | 7.7 | 82 | 9.0 |
| 8 | 7.1 | 33 | 8.8 | 58 | 8.4 | 83 | 2.7 |
| 9 | 8.3 | 34 | 7.5 | 59 | 7.2 | 84 | 5.9 |
| 10 | 5.6 | 35 | 6.1 | 60 | 3.3 | 85 | 6.5 |
| 11 | 5.3 | 36 | 4.1 | 61 | 4.1 | 86 | 3.1 |
| 12 | 4.8 | 37 | 2.8 | 62 | 7.5 | 87 | 4.6 |
| 13 | 4.2 | 38 | 7.9 | 63 | 8.4 | 88 | 3.9 |
| 14 | 6.4 | 39 | 4.5 | 64 | 5.6 | 89 | 4.6 |
| 15 | 2.8 | 40 | 8.9 | 65 | 4,7 | 90 | 2.9 |
| 16 | 8.6 | 41 | 5.1 | 66 | 6.2 | 91 | 5.8 |
| 17 | 7.4 | 42 | 4.8 | 67 | 7.8 | 92 | 8.3 |
| 18 | 5.8 | 43 | 9.3 | 68 | 8.4 | 93 | 2.7 |
| 19 | 4.2 | 44 | 6.7 | 69 | 6.2 | 94 | 7.6 |
| 20 | 4.6 | 45 | 3.1 | 70 | 5.9 | 95 | 6.7 |
| 21 | 4.8 | 46 | 8.7 | 71 | 4.4 | 96 | 9.1 |
| 22 | 5.0 | 47 | 7.6 | 72 | 5.8 | 97 | 7.0 |
| 23 | 6.2 | 48 | 7.8 | 73 | 7.7 | 98 | 8.5 |
| 24 | 4.2 | 49 | 5.8 | 74 | 8.0 | 99 | 5.0 |
| 25 | 6.0 | 50 | 4.3 | 75 | 4.8 | 100 | 3.6 |

**احسب ما يلي:**

1. احسب متوسط إنتاج الرأس الواحد من الحليب سنويا.
2. احسب التباين لمتوسط إنتاج الرأس الواحد من الحليب سنويا.
3. اسحب عينة عشوائية بسيطة حجمها 25 رأس من الأبقار وقدر من خلالها متوسط إنتاج الرأس الواحد بحدود ثقة 95 %.
4. من نفس العينة المسحوبة في (3) قدر نسبة الأبقار التي يزيد إنتاجها عن 5 طن سنويا بحدود ثقة 95%.
5. قدر حجم العينة التي يمكن سحبها لتقدير متوسط إنتاج الرأس بحدود ثقة 95%.

**السؤال الثاني**

في دراسة لتقدير عدد العاملين في الفنادق في إحدى المدن السياحية والتي يوجد فيها 50 فندقا كانت النتائج كما يلي:

| رقم الفندق في العينة | عدد العاملين |
|---|---|
| 1 | 20 |
| 2 | 12 |
| 3 | 8 |
| 4 | 50 |
| 5 | 10 |
| 6 | 25 |
| 7 | 13 |
| 8 | 40 |
| 9 | 32 |
| 10 | 18 |

**قّدر ما يلي:**

1. قّدر إجمالي عدد العاملين في الفنادق في المدينة بحدود ثقة 95%.
2. قّدر نسبة الفنادق التي يعمل فيها أكثر من 20 عاملا بحدود ثقة 95%.
3. قّدر حجم العينة التي يمكن سحبها بحدود ثقة 95 %.

### السؤال الثالث

لاحظ المدير الإداري وجود تأخر عدد من الموظفين عن الدوام في الوقت المحدد، فقام برصد عينة تتكون من 10 موظفين و5 موظفات من أصل 500 موظف وموظفة يعملون في المؤسسة فوجد أن

| الموظف | عدد الساعات الضائعة نتيجة التأخر |
|---|---|
| 1 | 3 |
| 2 | 10 |
| 3 | 2 |
| 4 | 5 |
| 5 | 20 |
| 6 | 0 |
| 7 | 2 |
| 8 | 1 |
| 9 | 0 |
| 10 | 2 |

| الموظفة | عدد الساعات الضائعة نتيجة التأخر |
|---|---|
| 1 | 5 |
| 2 | 6 |
| 3 | 0 |
| 4 | 10 |
| 5 | 4 |

إذا علمت أن عدد الموظفين الذكور في المؤسسة 300 شخص.

**احسب ما يلي:**

1. إجمالي عدد الساعات الضائعة شهرياً للموظفين خلال شهر بحدود ثقة 95%.
2. متوسط عدد الساعات الضائعة للموظفة الواحدة شهريا بحدود ثقة 95%.
3. إجمالي عدد الساعات الضائعة للموظفين والموظفات العاملين في المؤسسة.

الفصل الثالث

# العينة العشوائية الطبقية
# Stratified Random Sample

(STr)

## مقدمـــــة

أن من شروط استخدام العينة العشوائية البسيطة وجود تجانس بين وحدات المعاينة (مفردات المجتمع) للصفة المدروسة، ونظرا لصعوبة تحقق هذا الشرط في كثير من المسوح بالعينة، فأنه يُلجأ إلى طرق أخرى بهدف ايجاد هذا التجانس، وغالبا يستخدم أسلوب العينة الطبقية، حيث يقسم المجتمع إلى عدة مجموعات، كل مجموعة تكون متجانسة بالنسبة 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:

1. **تكوين الطبقات.**
2. **عدد الطبقات المراد تكوينها.**
3. **حجم العينة في كل طبقة.**
4. **تحليل البيانات لتصميم العينة الطبقية.**

**مثال 1:**

لو كان لدينا ثمانية محلات تجارية في إحدى التجمعات السكنية وكان صافي إيراد هذه المحلات كما يلي:

| رقم المحل | الإيراد ( درهم ) | رقم المحل | الإيراد ( درهم ) |
|---|---|---|---|
| 1 | 3000 | 5 | 6000 |
| 2 | 1500 | 6 | 1200 |
| 3 | 7000 | 7 | 4500 |
| 4 | 2500 | 8 | 5000 |

فلو طلب منا تقدير متوسط إيراد المحل الواحد بإستخدام أسلوب المعاينة بحيث يسحب ثلاث وحدات معاينة بإستخدام العينة العشوائية البسيطة وظهر في العينة المحال التي أرقامها 1،2،6 فإن متوسط إيراد المحل الواحد هو 1900 درهم، بينما المتوسط الحقيقي للمجتمع هو (3837.5) أي ما يزيد على ضعف المتوسط المقدر بالعينة، من الواضح وجود خلل نتيجة استخدام العينة العشوائية البسيطة لمجتمع غير متجانس، فلو قسم هذا المجتمع إلى طبقات حسب فئات الدخل كما يلي:

| الطبقة الأولى (اقل من 3000) | |
|---|---|
| رقم المحل | الإيراد |
| 2 | 1500 |
| 4 | 2500 |
| 6 | 1200 |

| الطبقة الثانية (5000 - 3000) | |
|---|---|
| رقم المحل | الإيراد |
| 1 | 3000 |
| 7 | 4500 |

| الطبقة الثالثة 5000 فأكثر | |
|---|---|
| رقم المحل | الإيراد |
| 3 | 7000 |
| 5 | 6000 |
| 8 | 5000 |

فلو سحبت وحدة معاينة واحدة من كل طبقة كما يلي:

| رقم الطبقة | رقم المحل | الإيراد ( درهم ) |
|---|---|---|
| 1 | 2 | 1500 |
| 2 | 7 | 4500 |
| 3 | 5 | 6000 |

المتوسط المقدر للمجتمع = متوسط العينة= $(\frac{3}{8} \times 1500) + (\frac{2}{8} \times 4500) + (\frac{3}{8} \times 6000) = 3937.5$

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

**مبادئ تقسيم المجتمع إلى طبقات:**

**عند تقسيم المجتمع إلى طبقات يجب مراعاة النقاط التالية:**

1. أن يكون مجموع وحدات المعاينة لجميع الطبقات مساوٍ لمجموع وحدات المعاينة للمجتمع، ويجب أن لا يوجد تداخل بين الطبقات.
2. أن تكون وحدات المعاينة في داخل كل طبقة متجانسة بالنسبة للصفة المدروسة.
3. هناك حالات تؤدي إلى استخدام أسلوب العينة الطبقية فمثلا إذا طلب الحصول على نتائج حسب الوحدات الإدارية (امارة، بلدية ...).
4. مراعاة الطرق العلمية عند تقسيم المجتمع إلى فئات حيث تعتبر كل فئة طبقة مستقلة.

**أسلوب تقسيم المجتمع إلى فئات:**

تعتبر عملية تقسيم 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 متوسط 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 الأسلوب التالي:بإستخدام القانون التالي:
عدد الفئات = 1+3Log n

**مثال 2:**

إذا كان عدد المصانع في أحد الأقاليم هو 1850 مصنعا فإن عدد الفئات التي يمكن أن يقسم المجتمع إليها هي:

1+3( Log 1850)
= 10.8 ⇐ 11 فئة

وهناك أسلوباً مقترحاً يأخذ عدد وحدات المعاينة ووزن كل وحدة بالاعتبار ويمكن توضيح ذلك بالمثال التالي:

**مثال 3:**

إذا كان عدد المصانع في أحد الأقاليم هو 1850 مصنعا مقسمة إلى فئات مبدئية حسب عدد العمال إلى 12 فئة والمطلوب إعادة تقسيم هذه المصانع إلى أربع فئات فإذا كان لدينا البيانات التالية:

| الفئة المبدأية | عدد المصانع | عدد العمال (c) | C التراكمي | √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 (√C) على عدد الفئات المقترحة
48.7=194.9/4

الفئة الأولى هي التي تقابل الجذر التربيعي لعدد العمال التجميعي (48.7) وبالتقريب يمكن اعتبار الفئة الأولى هي 1 - 10 عمال، وبإستخدام نفس الأسلوب لبقية الفئات نحصل على الفئات التالية:

| | | |
|---|---|---|
| الفئة الأولى | 1×48.7 ⇐ | 10-1 |
| الفئة الثانية | 2×48.7 ⇐ | 20-11 |
| الفئة الثالثة | ×48.73 ⇐ | 30-21 |
| الفئة الرابعة | 4×48.7 ⇐ | 31- فأكثر |

**ميزات وفوائد العينة الطبقية:**

1. في العينة الطبقية يكون المجتمع متجانسا في كل طبقة، ويكون المجتمع مُمثل بشكل جيد حيث تؤخذ عينات من مختلف الطبقات، والتي قد يكون لبعضها أهمية خاصة ، وقد تسقط هذه الطبقات عند استخدام أساليب معاينة أخرى.
2. استخدام العينة الطبقية اكثر فعّالية من العينات الأخرى خاصة في حالة وجود مجتمع غير متجانس، وفي حالة وجود قيم متطرفة لبعض وحدات المعاينة.
3. يُؤدي إلى تخفيض الكلفة لأنه يخفض عدد وحدات المعاينة المطلوب تغطيتها عند ثبات مستوى الدقة.
4. يمكن استخدام العينة الطبقية للحصول على نتائج على مستويات إدارية معينة (إمارة، بلدية...).
5. تكون السيطرة على العمل والأشراف علية وتنظيمه وتحديد منطقة عمل كل مجموعة بشكل افضل وذلك عند تقسيم المجتمع إلى طبقات حسب المناطق الإدارية .

**تقدير المعدل والمجموع للعينة العشوائية الطبقية:**

لنفرض أن المجتمع مقسم إلى (L) طبقة.
**Nh =** عدد وحدات المعاينة في الطبقة (n).

$$N = \sum_{h=1}^{L} N_h \quad (20)$$

**N=** عدد وحدات المعاينة في المجتمع.
**h=** عدد وحدات المعاينة لجميع الطبقات.

$$N = \sum_{h=1}^{L} n_h \quad (21)$$

$Y_{hi}$ **=** قيمة الصفة المدروسة للوحدة(i) في الطبقة(h)

معدل المجتمع $(\overline{Y})$

$$\bar{Y} = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} Y_{hi} \quad (22)$$

معدل الطبقة $(\overline{Y}_h)$

$$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} Y_{hi} \quad (23)$$

معدل المجتمع $(\bar{Y})$ يمكن حسابه بالطريقة التالية:

$$\bar{Y} = \frac{1}{N} \sum_{h=1}^{L} \frac{N_h}{N} \bar{Y}_h = \sum_{h=1}^{L} W_h \bar{Y}_h \quad (24)$$

حيث $W_h$ هو وزن الطبقة **(h)**

$$\sum_{h=1}^{L} W_h = 1\,, \qquad W_h = \frac{N_h}{N} \tag{25}$$

يمكن تقدير معدل المجتمع تقديرا غير متحيزا
لنفرض أن $\hat{\bar{Y}}_h$ هو تقدير غير متحيز لمعدل الطبقة $\bar{Y}_h$

$$\hat{\bar{Y}}_{st} = \sum_{h=1}^{L} W_h \; \hat{\bar{Y}}_h \tag{26}$$

حيث

$$\hat{\bar{Y}}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi} \tag{27}$$

**أما حساب التباين لمعدل المجتمع $V(\bar{y}_{st})$ فيمكن حسابه:**

$$V(\bar{Y}_{st}) = \sum_{i=1}^{L} w_h{}^2 \, V(\bar{y}_h) + \sum\sum W_h W_g \left(\bar{y}_h \bar{y}_g\right) = \sum w_h{}^2 V(\bar{y}_h) \tag{28}$$

حيث أن التباين المشترك (covariance) يساوي صفر نظراً لكون وحدات المعاينة تسحب بشكل مستقل في كل طبقة عن الطبقات الأخرى.

**مثال 4:**

في دراسة لتقدير إنتاجية الدونم من محصول البطاطا في أحد المقاطعات سحبت عينة من كل منطقة زراعية (بحيث اعتبرت كل منطقة طبقة مستقلة) بأسلوب العينة العشوائية البسيطة فكانت النتائج كما يلي:

| الطبقة | عدد الحقول الكلية ($N_h$) | عدد حقول العينة ($n_h$) | معدل إنتاج الدونم للعينات طن/دونم ($\bar{y}_i$) |
|---|---|---|---|
| 1 | 60 | 8 | 2, 3, 5, 4, 4, 3.5, 2, 5 |
| 2 | 100 | 9 | 3, 4, 2, 3, 5, 4, 2, 2.5, 3 |
| 3 | 20 | 5 | 3, 2, 1.5, 2.5, 2 |
| 4 | 20 | 5 | 3, 2.5, 2, 2, 1.5 |
| المجموع | 200 | 27 | |

لحساب معدل إنتاج الدونم من البطاطا في المقاطعة يمكن حسابه كما يلي:

- معدل إنتاج الطبقة الأولى= $\hat{\bar{Y}}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi} = \frac{28.5}{8} = 3.56$
- معدل إنتاج الطبقة الثانية= $\frac{1}{9} \times 27 = 3$
- معدل إنتاج الطبقة الثالثة= $\frac{1}{5} \times 11 = 2.2$
- معدل إنتاج الطبقة الرابعة= $\frac{1}{5} \times 11 = 2.2$
- معدل إنتاج جميع الطبقات= $\hat{\bar{Y}}_{st} = \sum_{h=1}^{L} W_h \ \hat{\bar{Y}}_h$

- وزن الطبقة الأولى= $\frac{60}{200} = 0.3$
- وزن الطبقة الثانية= $\frac{100}{200} = 0.5$
- وزن الطبقة الثالثة= $\frac{20}{200} = 0.1$
- وزن الطبقة الرابعة= $\frac{20}{200} = 0.1$

**معدل جميع الطبقات:** $(0.3 \times 3.56) + (0.5 \times 3) + (0.1 \times 2.2) + (0.1 \times 2.2) = 3.01$ دونم /طن

لحساب التباين لجميع الطبقات

$$S_n^2 = \frac{1}{n_N - 1}\sum (y_{ni} - \bar{y}_n)^2$$
$$S_1^2 = \frac{1}{7}(9.3) = 1.39$$
$$S_2^2 = \frac{1}{8}(4.5) = 0.56$$
$$S_3^2 = \frac{1}{4}(5.2) = 1.3$$
$$S_4^2 = \frac{1}{4}(1.3) = 0.33$$

التبـــايـــن الـمـقـــدر لجميع الطبقات $v(\hat{\bar{Y}}_{st})$

$$v(\hat{\bar{Y}}_{st}) = \sum_{h=1}^{L} w_h^2\, v(\hat{\bar{Y}}_h) = \sum_{h=1}^{L} w_h^2 \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\sigma_h^2$$
$$= 0.3^2\left(\frac{1}{8} - \frac{1}{60}\right)1.39 + 0.5^2\left(\frac{1}{9} - \frac{1}{100}\right)0.56$$
$$+ 0.1^2\left(\frac{1}{5} - \frac{1}{20}\right)1.39 + 0.1^2\left(\frac{1}{5} - \frac{1}{20}\right)0.33$$
$$= 0.07905$$
$$S.E = \sqrt{v(\bar{y}_{st})} = 0.28, \qquad S.E\% = \frac{S.E}{\bar{y}_{st}} \times 100 = 9.34\%$$

### تقدير الزيادة في الدقة الناتجة عن استخدام الطبقية:

يمكن تقدير الزيادة في الدقة وذلك عن طريق حساب تباين العينة الطبقية ومقارنته بتباين العينة العشوائية البسيطة، وهذا يدل على مدى الاستفادة من استخدام العينة الطبقية، ويتم التقدير عن طريق الحصول على تقدير لتباين العينة فيما لو استخدم أسلوب العينة العشوائية البسيطة من خلال مشاهدات العينة الطبقية،

ويمكن استخدام المعادلة التالية للتقدير: $v(\bar{Y}_{sr}) = \left(\frac{1}{n} - \frac{1}{N}\right)\sum w_i {s_i}^2$

$$+\left(\frac{N-n}{N-1}\right)\left[\sum w_i(\bar{y}_i - y_{st})^2 - \sum w_i(1-w_i)\frac{{s_i}^2}{n_i}\right] \quad (29)$$

أما الدقة المكتسبة

$$\frac{v(\bar{Y}_{sr}) - v(\bar{Y}_{st})}{v(\bar{Y}_{st})} \quad (30)$$

### تأثير التصميم

يمكن حساب تأثير التصميم وهو يساوي التباين المقدر عن طريق استخدام العينة الطبقية مقسوما على التباين للعينة العشوائية البسيطة، واذا كان تأثير التصميم أقل من 1 فهذا يعني أن العينة الطبقية أفضل من العينة العشوائية البسيطة وإذا كان تأثير التصميم أكبر من 1 يعني أن العينة العشوائية البسيطة أفضل من العينة الطبقية وإذا كان تأثير التصميم يساوي 1 فيكون كل من العينة العشوائية البسيطة والطبقية لها نفس مستوى الدقة.

**مثال 5:**

يمكن حساب الدقة المكتسبة من استخدام العينة الطبقية في المثال السابق

| الطبقة | $W_i$ | $s_i^2$ | $W_i s_i^2$ | $W_h(\bar{y}_h - \bar{y}_{st}^2)$ | $W_i(1-W_i)s_i^2$ |
|---|---|---|---|---|---|
| 1 | 0.3 | 1.39 | 0.417 | 0.09075 | 0.036 |
| 2 | 0.5 | 0.56 | 0.28 | 0.00 | 0.016 |
| 3 | 0.1 | 1.30 | 0.13 | 0.085 | 0.002 |
| 4 | 0.1 | 0.33 | 0.033 | 0.022 | 0.000 |
| المجموع | | | 0.86 | 0.19775 | 0.054 |

$$v(\bar{Y}_{sr}) = \left(\frac{1}{27} - \frac{1}{100}\right) \times 0.86 + \left(\frac{200-27}{199}\right) \times (0.19775 - 0.054) = 0.1525$$

$$\text{الدقة المكتسبة} = \frac{v(\bar{Y}_{sr}) - v(\bar{Y}_{st})}{v(\bar{Y}_{st})} = \frac{0.1525 - 0.07905}{0.07905} = 0.929$$

=93%

## 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:

### 1. العدد المتساوي من العينات لكل طبقة:

تستخدم هذه الطريقة عادة عند الحاجة للحصول على نتائج على مستوى المناطق الإدارية (في حال كون الطبقة تمثل منطقة إدارية) أو في حال توزيع العمل بشكل متساوٍ بين جميع الطبقات (حسب توفر الإمكانيات للعمل الميداني) كذلك عندما يكون عدد وحدات المعاينة متقارب في جميع الطبقات.

في هذه الحالة يكون عدد وحدات المعاينة في الطبقة الواحدة حسب المعادلة التالية:

$$n_h = \frac{n}{L} \quad (31)$$

**مثال 6:**

إذا كان حجم العينة 2000 وحدة معاينة وعدد الطبقات يساوي (8) فإن عدد وحدات المعاينة في الطبقة h يساوي

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

**التوزيع المتناسب:**

هذا أسلوب من أكثر الأساليب شيوعا نظراً لسهولته، عندما لا توجد معلومات عدا عدد وحدات المعاينة في كل طبقة فيمكن استخدام المعادلة التالية لتقدير عدد وحدات المعاينة في الطبقة (h)

$$n_h = \frac{N_h}{N} n \quad (32)$$

**مثال 7:**

لتقدير عدد أشجار الزيتون في احدى المناطق الزراعية والتي تتكون من اقليمين حيث اعتبر كل اقليم طبقة مستقلة، سحبت عينة قرى من كل طبقة بأسلوب العينة الطبقية وبإستخدام التوزيع المتناسب، وبحجم عينة إجمالي قدره 35 قرية.

| الطبقة (المحافظة) | عدد القرى الكلي |
|---|---|
| 1 | 200 |
| 2 | 150 |

عدد وحدات المعاينة في الطبقة الأولى يساوي = $n(\frac{N_h}{N}) = 35(\frac{200}{350}) = 20$

عدد وحدات المعاينة في الطبقة الثانية يساوي= $n(\frac{N_h}{N}) = 35(\frac{150}{350}) = 15$

## توزيع نيمان (Neyman):

يستخدم هذا التوزيع لتخفيض حجم التباين حيث يؤخذ في الاعتبار تباين كل طبقة، فيكون حجم العينة في الطبقة يتناسب طرديا مع الانحراف المعياري لتلك الطبقة وذلك من أجل جعل تصميم المعاينة أكثر فعّالية من أسلوب التوزيع المتناسب، وعادة يستخدم هذا الأسلوب عندما يكون الانحراف المعياري مختلفاً من طبقة لأخرى، وعندما يكون حجم العينة ثابتاً وكلفة العينة ثابتة لمختلف الطبقات، ويقدر حجم العينة للطبقة (h) حسب المعادلة التالية:

$$n_h = \frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h} n \qquad (33)$$

أما الانحراف المعياري على مستوى كل طبقة فيمكن الحصول علية من خلال تعداد سابق أو يمكن تقديره من مسح سابق.

**مثال 8:**

في دراسة لتقدير معدل دخل الأسرة في إحدى المدن قُسم المجتمع إلى ثلاث طبقات حسب فئات الدخل، وسحبت عينة من كل فئة بطريقة توزيع نيمان وبحجم عينة إجمالي قدره 15 وحدة معاينة، فإذا كانت لدينا البيانات التالية:

| الطبقة (فئة الدخل) | عدد الأسر في المجتمع $N_h$ | 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 | 50 | 1000 |
| | | | 3100 |

فإن عدد وحدات المعاينة للطبقة الأولى= $n_h = \frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h} = 15 \times \frac{600}{3100} = 3$

فإن عدد وحدات المعاينة للطبقة الثانية= $= 15 \times \frac{1500}{3100} = 7$

فإن عدد وحدات المعاينة للطبقة الثالثة= $= \frac{15 \times 1000}{3100} = 5$

### التوزيع الأمثل:

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

$$n_h = \frac{\frac{N_h S_h}{\sqrt{C_h}}}{\frac{\sum_{h=1}^{L} N_h S_h}{\sqrt{C_h}}} n \qquad (34)$$

حيث ($C_h$) تشكل كلفة إحصاء وحدة المعاينة الواحدة للطبقة h.

### مثال 9:

في دراسة لتقدير معدل إنتاج الدونم من محصول الحمص في احدى المقاطعات أستخدمت العينة الطبقية حيث شكلت كل منطقة زراعية طبقة مستقلة، وكان عدد وحدات المعاينة الإجمالي هو 200 وحدة معاينة وكانت البيانات التالية عن الطبقات:

| الطبقة (المحافظة) | عدد الحقول $N_h$ | الانحراف المعياري للطبقة $S_h$ | $N_h S_h$ | كلفة الوحدة $C_h$ | $\frac{N_h S_h}{\sqrt{C_h}}$ |
|---|---|---|---|---|---|
| 1 | 200 | 10 | 2000 | 4 | 1000 |
| 2 | 100 | 20 | 2000 | 6 | 816 |
| 3 | 150 | 10 | 1500 | 6 | 612 |
| 4 | 100 | 15 | 1500 | 9 | 500 |
| 5 | 50 | 20 | 1000 | 12 | 287 |
| 6 | 50 | 20 | 1000 | 20 | 224 |
| | | | | | 3439 |

$$n_h = \frac{N_h S_h / \sqrt{C_h}}{\sum_{h=1}^{L} N_h S_h / \sqrt{C_h}} \quad n = 200 \times 1000/3439 = 58$$ عدد وحدات المعاينة في الطبقـة الأولـى=

عدد وحدات المعاينة في الطبقـة الثـانيــة= $47 = (816/3439)200$

عدد وحدات المعاينة في الطبقــة الثالثـة= $36 = (612/3439)200$

عدد وحدات المعاينة في الطبقـة الرابعــة= $29 = (500/3439)200$

عدد وحدات المعاينة في الطبقة الخامسـة= $17 = (287/3439)200$

عدد وحدات المعاينة في الطبقة السادسة= $13 = 0(224/3439)20$

**التوزيع المتناسب مع صفة معينة:**

في بعض الحالات تكون الأهمية لصفة معينة، ويكون توزيع العينة على أساسها أهم من عدد مفردات المجتمع، ولتوضيح هذه الفكر يمكن أخذ المثال التالي:

إذا كان الهدف تقدير المساحة المزروعة بالمحاصيل الحقلية وكان لدينا البيانات التالية:

| الطبقة | عدد الحيازات الزراعية | مساحة المحاصيل الحقلية (دونم) |
|---|---|---|
| 1 | 1000 | 20000 |
| 2 | 500 | 30000 |
| 3 | 1500 | 10000 |

في حالة استخدام أسلوب التوزيع المتناسب فإن وزن الطبقة الثالثة يساوي 0.5 على الرغم من أن الطبقة الثالثة تحتوي على 0.1667 من مساحة المحاصيل الحقلية، لذلك في حالة استخدام عدد الحيازات لوزن الطبقة فإنه يعطي نتائج مضللة ويكون نصيب الطبقة الثالثة 50% من حجم العينة المسحوبة، بينما يكون نصيب الطبقة الثانية 17% من العينة المسحوبة على الرغم من أنها تحتوي على 50% من إجمالي المساحة مما يقلل من فعّالية العينة ويزيد التكاليف.

لذلك فإن الأسلوب الأفضل هو توزين الطبقات حسب مساحة المحاصيل الحقلية لكل طبقة كونها الصفة المراد دراستها، ويتم توزيع العينة بين الطبقات بمختلف الأساليب التي مرت معنا، ولكن بإستخدام المساحة لوزن الطبقة بدلا من عدد الحيازات.

وبإستخدام المعادلة التالية لتوزيع العينة بالأسلوب المتناسب مع صفة معينة.

$$n_h = \frac{x_h}{\sum_{n=1}^{L} x_h} n$$

حيث $x_h$: الصفة التي تحملها الطبقة (h) وهي مساحة المحاصيل الحقلية في المثال الوارد لدينا.

فلو طلب سحب عينة حجمها 100 حيازة بالأسلوب المتناسب وحسب صفة مساحة المحاصيل الحقلية فيكون توزيع العينة كما يلي:

| الطبقة | حجم العينة |
|---|---|
| 1 | 33 |
| 2 | 50 |
| 3 | 17 |

## تمارين

1. ما هي أهم مزايا تقسيم المجتمع إلى طبقات؟
2. ما هي أهم أسس تقسيم المجتمع إلى طبقات؟
3. ما هي أهم طرق توزيع حجم العينة بين الطبقات؟
4. في دراسة لتقدير المساحة المزروعة ومعدل إنتاج الدونم لمحصول القمح اختير 5 حقول من أصل 100 حقل موجودة إحدى القرى فكانت النتائج كما يلي:

| الطبقة | حجم العينة |
|---|---|
| 1 | 120 |
| 2 | 100 |
| 3 | 140 |
| 4 | 80 |
| 5 | 110 |

**احسب ما يلي:**

- المعدل المقدر لإنتاج الدونم في القرية.
- تباين العينة.
- معامل التكبير (وزن العينة).
- كسر المعاينة.
- خطأ المعاينة.

5. في دراسة لتقدير معدل دخل الأسرة في إحدى التجمعات السكانية قسمت المجتمع إلى ثلاث فئات حسب معدلات الدخل المتوقع وسحبت عينة من كل فئة فكانت النتائج كما يلي:

| فئة الدخل | عدد الأسر في المجتمع | عدد العينات | معدل دخل الأسرة | تباين العينة لكل طبقة |
|---|---|---|---|---|
| أقل من 1000 دينار | 30 | 5 | 600 | 500 |
| 3000-1000 | 50 | 5 | 1800 | 1200 |
| أكثر من 3000 | 20 | 5 | 4200 | 2800 |

**احسب ما يلي:**

- المعدل المقدر لدخل الأسرة في التجمع السكاني.
- عدد الطبقات التي قسم إليها هذا المجتمع.
- عدد وحدات المعاينة لجميع الطبقات.
- لتباين المقدر للطبقة الأولى.
- ما هو أسلوب توزيع العينات على الطبقات.

6. في دراسة لتقدير متوسط مبيعات قطاع التجارة الداخلية في إحدى المدن قسم المجتمع إلى ثلاث طبقات حسب نوع التجارة وسحبت عينة من كل طبقة بأسلوب التوزيع فكانت النتائج كما يلي:

| الطبقة | $N_i$ | $n_i$ | $\bar{y}_i$ | تباين العينة لكل طبقة |
|---|---|---|---|---|
| 1 | 20 | 2 | 2000 | 400 |
| 2 | 60 | 6 | 5000 | 600 |
| 3 | 40 | 4 | 6000 | 1200 |

**احسب ما يلي:**

- احسب متوسط المبيعات لجميع الطبقات بحدود ثقة 95%.
- اعد توزيع العينة على الطبقات بأسلوب التوزيع الأمثل.

7. في دراسة لتقدير نسبة الأشغال في الفنادق لإحدى المدن السياحية قسمت الفنادق إلى ثلاث فئات حسب حجمها وسحبت عينة من كل فئة فكانت النتائج كما يلي:

| الفئة | نسبة الأشغال | $n_i$ | $N_i$ |
|---|---|---|---|
| 1 | 0.6 | 6 | 30 |
| 2 | 0.7 | 4 | 20 |
| 3 | 0.5 | 4 | 20 |

**احسب ما يلي:**

- ما هو أسلوب توزيع العينة على الطبقات.
- قدر نسبة الأشغال لجميع الطبقات بحدود 95% ثقة.

8. في دراسة لتقدير متوسط دخل الأسرة في إحدى المدن قسمت المدينة إلى أربع طبقات حسب الخصائص الاجتماعية والاقتصادية وسحبت عينة من كل طبقة فكانت النتائج كما يلي:

| الطبقة | $N_i$ | $n_i$ | $\bar{y}_i$ | تباين العينة لكل طبقة |
|---|---|---|---|---|
| 1 | 400 | 20 | 1200 | 2000 |
| 2 | 200 | 10 | 2000 | 3000 |
| 3 | 200 | 10 | 2500 | 5000 |
| 4 | 100 | 5 | 4000 | 6000 |

**أجب عما يلي:**

- ما هو أسلوب توزيع العينة بين الطبقات.
- احسب متوسط دخل الأسرة الشهري في المدينة.
- اعد توزيع العينة بين الطبقات بإستخدام توزيع نيمان.

9. في دراسة لتقدير متوسط إنتاج الشجرة الواحدة من الزيتون قسمت المنطقة إلى ثلاث طبقات وسحبت عينة من كل طبقة فكانت نتائج الدراسة كما يلي:

| الطبقة | $N_i$ | $n_i$ | $\bar{y}_i$ | تباين العينة لكل طبقة |
|---|---|---|---|---|
| 1 | 200 | 20 | 8 | 4 |
| 2 | 150 | 15 | 4 | 2 |
| 3 | 50 | 5 | 20 | 8 |

**احسب ما يلي:**

- احسب متوسط إنتاج الشجرة لجميع الطبقات و أوجد 90% فترة ثقة للوسط الحسابي.
- احسب معامل التغير c.v للمتوسط.
- احسب مجموع الإنتاج لجميع الطبقات وتباينه.
- قارن بين النتائج في حالة استخدام التوزيع الأمثل مقارنة مع التوزيع المناسب.

10. في مسح بالعينة لتقدير عدد حقول الزيتون في إحدى المحافظات تم تقسيم قرى المحافظة إلى أربع طبقات حسب المساحات في المزرعة في تلك القرى، تم اختيار عينة عشوائية بسيطة من قرى كل طبقة وعدد حقول كل قرية كما هو موضح في الجدول المرفق.

| الطبقة | عدد قرى الطبقة | عدد قرى العينة في الطبقة | عدد الحقول في قرى العينة لكل طبقة |
|---|---|---|---|
| 1 | 275 | 15 | 7،0،7،4،0،5،0،0،3،0 ،6 ،9 ،1 ، 5 ،2 |
| 2 | 146 | 10 | 21،11،7،5،6،19،5،24،30،24 |
| 3 | 93 | 12 | 3،10،4،11،38،11،4،46،4،18،1،19 |
| 4 | 62 | 11 | 30،42،20،38،29،22،31،28،66،41،15 |

**احسب ما يلي:**

- قدر عدد الحقول الزيتون في المحافظة.
- احسب خطأ المعاينة بحدود 95% ثقة.
- احسب فاعلية العينة الطبقية نسبة إلى العينة العشوائية البسيطة.

الفصل الرابـع

# العينة المنتظمة
# Systematic sample
(Sy)

# مقدمة

في الفصل الثاني ورد أسلوب سحب العينات العشوائية، كذلك سيرد في فصول لاحقة أساليب أخرى لسحب العينات، ولكن جميع تلك الأساليب تحتاج إلى استخدام الجداول العشوائية عند سحب كل وحدة معاينة باستثناء العينة المنتظمة التي تحتاج استخدام الجداول العشوائية في السحبة الأولى فقط، ومن المعروف أن استخدام الجداول العشوائية تحتاج إلى وقت وجهد، لذا فقد وُجد أسلوب المعاينة المنتظمة للتقليل من هذا الجهد والوقت، ويعتبر استخدام العينة المنتظمة مفيد لعدد من الدراسات مثل احصاءات الغابات وإحصاءات السياحة وإنتاج حليب الأبقار وغيرها.

# أسلوب اختيار وحدات المعاينة العينات

## أ. المعاينة المنتظمة الخطية

هذا الأسلوب هو الأكثر شيوعا في العينات المنتظمة، ويتلخص في ما يلي:

نفترض أن المجتمع يتكون من N من وحدات المعاينة وأن عدد وحدات المعاينة المطلوب سحبها هو n وأن المتمم ل n هوk حيث nk=N ويجري اختيار رقم عشوائي يقع بين 1 وK يسمى هذا الرقم برقم البداية ويرمز له بالرمز(i) وتسمى K فترة الانتظام ويكون الرقم المتسلسل للعينة الأولى هو (i) والعينة الثانية هي K+i والثالثة 2K+i..الخ.

**مثال 1:**

لتقدير معدل إنتاج الرأس الواحد من الأبقار الحلوب في احدى المناطق والتي تحتوي على 380 مزرعة تجارية لتربية الأبقار الحلوب سحبت عينة منتظمة من تلك المزارع حجمها 19 مزرعة.

**أسلوب سحب العينة كما يلي:**

$$k = \frac{N}{n} = \frac{380}{19} = 20 \qquad N = 380, \qquad n = 19$$

يجري سحب رقم عشوائي من الجدول العشوائي يقع بين 1و20 ولنفترض أننا حصلنا على الرقم 12 لذا فإن الأرقام المتسلسلة لوحدات المعاينة المسحوبة في العينة هي:

12, 12 + 20, 12 + 2(20), 12 + 30(20) ...

وتساوي

12،32،52،72،92،112،152،172،192،212،232،252،272،292،312،332،352،372.

### ب. المعاينة المنتظمة الدائرية

في حالة كون وحدات المعاينة في المجتمع N لا تقبل القسمة على n بدون باقي قسمة تكونK غير مكملة لـ n وإنما تكون اقرب ما يمكن من N/n، فعلى سبيل المثال لو كان حجم المجتمع (N)=11 وعدد وحدات المعاينة n=4 فان k=3 والعينات المحتملة هي (1،4،7،10) ،(2،5،8،11) و (3،6،9) ويلاحظ عدم تساوي العينات.

للتغلب على التفاوت في حجم العينة عندما لا تكون K رقماً صحيحاً يتم استخدام أسلوب آخر في سحب وحدات المعاينة للحصول على عدد ثابت من وحدات المعاينة في جميع الأحوال، وتتلخص هذه الطريقة بسحب رقم عشوائي يقع بين 1 و N ثم تسحب عينة من كل فترة انتظام (K) من وحدات المعاينة ،حيث K متممة n وقريبه من N/n.

**مثال 2:**

لو أخذنا المثال السابق فإن العينات المحتملة هي:
(1،4،7،10)،(2،5،8،11)،(3،6،9،1)،(4،7،10،2)،(5،8،11،3)،(6،9،1،4)،(7،10،2،5)،(8،11،3،6)،
(1،4،7،9)،(10،2،5،8)،(11،3،6،9).

**ميزات وعيوب العينة المنتظمة**

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

أما من عيوب هذه العينة هي صعوبة الحصول على تقدير غير متحيز للتباين، كذلك في حالة وجود صفة دورية في المجتمع قد يؤدي إلى وجود التحيز أثناء سحب وحدات المعاينة، وكذلك في تقدير المتوسط والتباين للمجتمع، فمثلا لو سحبت عينة من وحدات زراعية وكان رقم البداية هو 1 وفترة العينة هي 20 وكانت الوحدات الزراعية التي ارقامها 1 أو 21، هي وحدات مجاورة لمصدر المياه الرئيسي وغالبا لا تعاني هذه الوحدات من نقص في مياه الري لذا فإن إنتاجية هذه الوحدات الزراعية أعلى من الوحدات الزراعية الأخرى مما يؤدي إلى تحيز في التقدير.

**أسلوب التقدير:**

الحالة السهلة للتقديرات عندما تكون N=nk عندما تسحب العينة بأسلوب المنتظمة الخطية فإن تقدير المتوسط والمجموع يكون غير متحيز

تقدير المتوسط $\hat{\bar{y}}_{S_y}$

$$\hat{\bar{y}}_{S_y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (35)$$

أما التباين المقدر للمتوسط وهو متحيز فهو $v\left(\hat{\bar{y}}_{S_y}\right)$

$$v\left(\hat{\bar{y}}_{S_y}\right) = \frac{1}{k}\sum_{k=1}^{k}(\bar{y}_k - \bar{y})^2 \quad (36)$$

حيث $\bar{y}_k$ هو متوسط العينة المنتظمة.

التباين المقدر وهو متحيز $\hat{v}\left(\hat{\bar{y}}_{S_y}\right)$

$$\hat{v}\left(\hat{\bar{y}}_{S_y}\right) = \frac{N-n}{2Nn(n-1)}\sum_{i=1}^{n-1}(y_{i+1} - y_i)^2 \quad (37)$$

تقدير المجموع وهو تقدير غير متحيز $\hat{y}_{S_y}$

$$\hat{y}_{S_y} = \frac{N}{n}\sum_{i=1}^{n} y_i \quad (38)$$

أما تقدير التباين للمجموع فهو تقدير متحيز $\hat{v}\left(\hat{y}_{S_y}\right)$

$$\hat{v}\left(\hat{y}_{S_y}\right) = \frac{N^2(N-n)}{2Nn(n-1)}\sum_{i=1}^{n-1}(y_{i+1}-y_i)^2 \qquad (39)$$

في حالة العينة المنتظمة الدائرية يكون المتوسط المقدر للمجتمع غير متحيز، ويمكن استخدام تقديرات التباين الواردة في المعادلات السابقة.

**مثال 3:**

يقوم أحد مصانع الأدوية بفحص محتويات عبوة واحدة من كل 50 عبوة مُصنّعة لمعرفة مدى مطابقتها للمواصفات، فإذا كان عدد العبوات المفحوصة في أحد الأيام (100) عبوة ووجد أن (5) منها غير مطابقة للمواصفات.

1. ما هو عدد العبوات المنتجة في ذلك اليوم؟
2. قدر نسبة العبوات غير المطابقة للمواصفات بحدود ثقة 95%.
3. قدر حجم العينة إذا كان الفرق المسموح به هو 0.05.

**الإجابة**

1. عدد العبوات المنتجة **(N)**

$$1\text{-N} = 100 \times 50 = 5000$$

$$2\text{-P}_{S_y} = \frac{\sum a}{n} = \frac{5}{100} = 0.05$$

$$S^2(P_{S_y}) = \frac{pq}{n-1}\frac{(N-n)}{N} = \frac{(0.05)(0.95)}{99}\frac{(5000-100)}{5000} = 0.00047$$

$$S.E = 0.02 \qquad \text{C.I} = 0.05 \mp .0425 \qquad 3\text{-n} \geq \frac{N_P q}{N_D + pq} \geq \frac{5000(0.0475)}{5000(0.000625) + 0.0475}$$

$$3 - D = \frac{(0.05)^2}{4} = 0.000625 \qquad \text{n} \geq 74.86 \Rightarrow n = 75$$

**مثال 4:**

في دراسة لتقدير المساحة المزروعة بمحصول القمح في إحدى القرى والتي يوجد فيها 200 حيازة زراعية سحبت عينة منتظمة منها حجمها 10 حيازات فكانت النتائج كما يلي:

| الرقم المتسلسل لوحدة المعاينة | مساحة القمح | الرقم المتسلسل لوحدة المعاينة | مساحة القمح |
|---|---|---|---|
| 1 | 12 | 6 | 20 |
| 2 | 150 | 7 | 0 |
| 3 | 27 | 8 | 40 |
| 4 | 50 | 9 | 5 |
| 5 | 13 | 10 | 2 |

**الإجابة**

مساحة القمح المقدرة $\left(\hat{y}_{S_y}\right)$

$$\hat{y}_{S_y} = \frac{N}{n}\sum_{i=1}^{n} y_i = \frac{200}{10}(319) = 6380$$

أما التباين المقدر $\hat{v}\left(\hat{y}_{S_y}\right)$ فهو

$$\hat{v}\left(\hat{y}_{S_y}\right) = \frac{N^2(N-n)}{2Nn(n-1)}\sum_{i=1}^{n-1}(y_{i+1}-y_i)^2$$

$$= \frac{190^2(200)}{2\times 200\times 10\times 9}\sum_{i=1}^{n-1}(y_{i+1}-y_i)^2$$

$$= 211.11\times[(150-12)^2+\cdots+(2-5)^2] =$$

$$= 211.11(19044+15129+529+1369+49$$

$$+400+1600+1225+9) = 8308022.94$$

$$S.E = 2882.36$$

$$S.E\% = 45.18$$

### تكرار العينة المنتظمة

للتخلص من التحيز الموجود في تقدير التباين للعينة المنتظمة يمكن أخذ أكثر من عينة منتظمة في المجتمع الواحد كل عينة حجمها $(n')$ وعدد العينات الإجمالي هو **n** وكل عينة لها رقم بداية مستقل ومجموع أرقام البداية هو m وكل عينة من هذه العينات تسمى مكررة، وفي هذه الحالة يمكن الحصول على تقدير غير متحيز للتباين .

$$n' = \frac{n}{m} \quad (40)$$

$$\hat{\bar{y}}_{S_y} = \frac{1}{m}\sum_{i=1}^{n} \bar{y}_i \quad (41)$$

$$\bar{y}_i = \frac{\sum \bar{y}_i}{n'} \quad (42)$$

أما التباين المقدر فهو $\hat{v}\left(\hat{y}_{S_y}\right)$

$$\hat{v}\left(\hat{y}_{S_y}\right) = \frac{1-f}{m(m-1)}\sum\left(\hat{\bar{y}}_i - \hat{\bar{y}}_{S_y}\right)^2 \quad (43)$$

### مثال 5:

في دراسة لاستطلاعات الرأي حول موضوع دعم المواد التموينية سحبت ثلاث عينات منتظمة (مكررات) من إحدى التجمعات السكنية فكانت النتائج كما يلي:

| رقم العينة | عدد وحدات المعاينة المسحوبة | عدد المؤيدين لسياسة الدعم |
|---|---|---|
| 1 | 30 | 20 |
| 2 | 30 | 25 |
| 3 | 30 | 22 |

إذا علمت أن فترة الانتظام للمكررة الواحدة هي (50) أحسب ما يلي:

1. نسبة المؤيدين للدعم بحدود ثقة 95%.
2. إجمالي عدد وحدات المعاينة (الأشخاص) في المجتمع.

**الإجابة:**

$$\hat{p}_{syw} = \frac{\sum_{i=1}^{n} p_i}{n_s} = \frac{0.67 + 0.833 + 0.733}{3} = 0.745 \qquad N = 1500$$

$$S^2(\hat{p}_{syw}) = \frac{(1500 - 90)}{1500} \frac{(0.2211 + 0.13911 + 0.1957)}{3(2)} = 0.087$$

$$S.E = 0.295 \qquad 1\text{-}C.i = 0.745 \mp 0.578 \qquad 2\text{-}N = 30 \times 50 = 1500$$

**مثال 6:**

أحد مصانع الزيوت النباتية يقوم بفحص العبوة الثالثة والعاشرة من العبوات المصنعة لمعرفة مدى مطابقتها للمواصفات فإذا كانت نتائج الفحوصات للشهر الماضي هي كما يلي:

| العينة (المكررة) | عدد العينات المفحوصة | عدد العينات المطابقة للمواصفات |
|---|---|---|
| 1 | 20 | 18 |
| 2 | 20 | 16 |

1. قدر نسبة العبوات المطابقة للمواصفات بحدود ثقة 95%.
2. ما هو أجمالي عدد العبوات المصنعة خلال نفس الشهر.

**الإجابة:**

$$\hat{p}_{syw} = \frac{\sum_{i=1}^{n} p_i}{n_s} = \frac{.9 + .8}{2} = .85 \qquad N = 200$$

$$S^2(\hat{p}_{syw}) = \frac{200 - 40}{200} \frac{(.9 + 1.6)}{2} = 0.1$$

$$S.E = .316 \qquad 1\text{-}C.I = .85 \mp .6198 \qquad 2\text{-}N = 10 \times 20 = 200$$

**مثال 7:**

في دراسة لتقدير المساحة المزروعة بمحصول البطاطا في إحدى المناطق الزراعية سحبت عينة حجمها 35 حقل بأسلوب المعاينة المنتظمة بهدف تقدير المساحة المزروعة بمحصول البطاطا، وقد تكونت العينة من خمس مكررات (كل مكررة 7 وحدات معاينة) وقد سحبت هذه العينات من أصل 210 حقل مزروع بمحصول البطاطا في المنطقة وكانت النتائج كما يلي:

| المكررة (1) | | المكررة (2) | | المكررة (3) | | المكررة (4) | | المكررة (5) | |
|---|---|---|---|---|---|---|---|---|---|
| رقم العينة | المساحة | رقم العينة | المساحة | رقم العينة | المساحة | رقم العينة | المساحة | رقم العينة | المساحة |
| 3 | 3 | 2 | 20 | 11 | 8 | 25 | 3 | 8 | 15 |
| 33 | 10 | 32 | 8 | 41 | 20 | 55 | 25 | 38 | 30 |
| 63 | 30 | 62 | 4 | 71 | 14 | 85 | 20 | 68 | 20 |
| 93 | 8 | 92 | 10 | 101 | 10 | 115 | 20 | 98 | 10 |
| 123 | 4 | 122 | 15 | 131 | 8 | 145 | 40 | 128 | 35 |
| 153 | 4 | 152 | 12 | 161 | 14 | 175 | 1 | 158 | 5 |
| 183 | 3 | 182 | 30 | 191 | 20 | 205 | 30 | 188 | 25 |

**الإجابة:**

لتقدير المساحة المزروعة بمحصول البطاطا في إحدى المناطق الزراعية ($\hat{A}$)

$$\hat{A} = \frac{\sum A_i}{m}$$

$$\hat{A}_I = \frac{N}{n'} \sum_{i=1}^{\overline{n'}} A'_i$$

$$\hat{A}_1 = \frac{210}{7}(3+10+30+8+4+4+3)$$

$$= 1860$$

$$\hat{A}_2 = 210/7(20+8+4+10+15+12+30)$$

$$= 2970$$

$$\hat{A}_3 = 210/7(8+20+14+10+8+14+20)$$

$$= 2820$$

$$\hat{A}_4 = 210/7(3+25+20+20+40+1+30)$$

$$= 4170$$

$$\hat{A}_5 = 210/7(15+30+20+10+35+5+25)$$

$$= 4200$$

$$\hat{A} = \frac{\sum \hat{A}_i}{m}$$

$$= (1860+2970+2820+4170+4200)/5$$

$$= 3204$$

أما التباين المقدر $\hat{V}(\hat{A})$ فهو

$$\hat{V}(\hat{A}) = \frac{1-f}{m(m-1)}\sum(A-\bar{A})^2 \quad (44)$$

$$f = \frac{35}{210} = 0.1666$$

$$\hat{V}(\hat{A}) = \frac{1-0.1666}{5\times 4}\sum(A-\bar{A})^2$$

$$\bar{A} = \frac{\bar{A}_1 + \bar{A}_2 + \bar{A}_3 + \bar{\boldsymbol{A}}_{\boldsymbol{4}} + \bar{A}_5}{m}$$

$$\bar{A}_1 = \frac{\sum A'_I}{n'} = 8.857, \qquad \bar{A}_2 = 14.143, \bar{A}_3 = 13.428, \bar{A}_4 = 19.857, \bar{A}_5 = 20.000$$

$$\bar{A} = \frac{\bar{A}_1 + \bar{A}_2 + \bar{A}_3 + \bar{A}_4 + \bar{A}_5}{5} = 15.257$$

$$\hat{V}(\hat{A}) = 3.7168, \quad SE = 1.9279,$$

$$SE\% = \frac{SE}{\bar{A}} = 12.636\%$$

# تمارين

1. ما هو الهدف من أخذ اكثر من مكرره في العينة المنتظمة؟
2. في دراسة لتقدير إنتاجية الدونم من محصول البطاطا سحبت عينة في إحدى القرى بأسلوب العينة المنتظمة تتكون من خمس عينات من أصل 50 حقل كما يلي:

| رقم الحقل | معدل النتاج طن/ دونم |
|---|---|
| 7 | 2 |
| 17 | 1.5 |
| 27 | 3 |
| 37 | 2 |
| 47 | 1.5 |

**احسب ما يلي:**

- أكمل الجدول بكتابة أرقام الحقول التي سحبت بالعينة.
- احسب معدل إنتاج الدونم المقدر للقرية.
- احسب طول فترة الانتظام.
- احسب تباين المجتمع المقدر.
- احسب خطأ المعاينة.

3. في دراسة لتقدير متوسط عدد الساعات الضائعة نتيجة تأخر العاملين عن الدوام في أحد المصانع لشهر شباط سحبت 3 عينات منتظمة كل عينة تتكون من 5 عمال من أصل 100 عامل في المصنع فكانت النتائج كما يلي:

| رقم العينة | $\bar{y}_i$ |
|---|---|
| 1 | 1 |
| 2 | 1.5 |
| 3 | 0.5 |

**احسب ما يلي:**

- متوسط تأخر العامل الواحد عن العمل.
- قدر عدد الساعات الضائعة نتيجة تأخر العاملين.
- فترة الانتظام للعينة الثانية.

4. في دراسة لتقدير المساحة المزروعة بمحصول الكوسا سحبت عينة منتظمة بسبعة أرقام بداية وكان حجم العينة الإجمالي 49 حيث كان حجم العينة لكل رقم بداية هو (7) من اصل 225 حيازة موجودة في القرية فكانت النتائج كما يلي:

| Rep No. | المساحة بالدونم |
|---|---|
| 1 | 60 - 50 -14 -10 - 1 - 0 -0 |
| 2 | 0 - 0- 0- 150- 150 - 100- 20 |
| 3 | 0 -25 - 192 - 25 - 0 -13 - 0 |
| 4 | 0 - 50 - 0 -10 - 0 - 0 - 0 |
| 5 | 85- 30 -30 -70 -30 - 35 - 0 |
| 6 | 30 - 0 - 0 - 0- 10- 0- 20 - 70 |
| 7 | 16- 15 - 35 -0 - 0 -17- 0 |

**احسب ما يلي:**

- أحسب مجموع المساحة المزروعة لمحصول الكوسا في القرية.
- أحسب متوسط المساحة المزروعة بالكوسا للحيازة الواحدة بفترة ثقة 90%.

5. في إحدى القرى والتي تحتوي 200 حيازة زراعية سحب منها 10 حيازات بأسلوب العينة المنتظمة بهدف تقدير المساحة الزراعية للقرية.

| الرقم المتسلسل | المساحة الزراعية (دونم) |
|---|---|
| 1 | 12 |
| 2 | 150 |
| 3 | 27 |
| 4 | 50 |
| 5 | 83 |
| 6 | 82 |
| 7 | 100 |
| 8 | 40 |
| 9 | 12 |
| 10 | 50 |

**احسب ما يلي:**

- قدر المساحة الزراعية للقرية.
- احسب خطأ المعاينة للمساحة المقدرة.

6. في دراسة لتقدير نسبة الأشجار المثمرة فعلا في بستان زيتون سحبت الشجرة الثانية والسابعة من كل 10 شجيرات موجودة في البستان والذي يحتوي على 1200 شجرة في سن الأثمار فكانت النتائج كما يلي:

| الرقم المتسلسل | المساحة الزراعية (دونم) |
|---|---|
| 1 | 12 |
| 2 | 150 |
| 3 | 27 |
| 4 | 50 |
| 5 | 83 |
| 6 | 82 |
| 7 | 100 |
| 8 | 40 |
| 9 | 12 |
| 10 | 50 |

**احسب ما يلي:**

- قدر نسبة الأشجار المثمرة فعلا في البستان بحدود ثقة 95%.

7. يقوم أحد مصانع الشامبو بفحص العبوة الثالثة من كل عشر عبوات مصنعة، وفي أحد الأيام قام المصنع بفحص 200 عبوة ووجد منها 10 عبوات غير مطابقة للمواصفات.

**احسب ما يلي:**

- ما هو عدد العبوات المصنعة في ذلك اليوم.
- قدر عدد العبوات غير المطابقة للمواصفات بحدود ثقة 95%.

## الفصل الخامس

# التقدير النسبي
# Ratio estimation

## مقدمة

في بعض المسوح يوجد معلومات عن متغيرات مساعدة متوفرة من خلال المسوح والتعدادات السابقة، بالإضافة إلى المعلومات عن المتغيرات التي توفرها الدراسة، وهذه المعلومات يمكن استخدامها في تقدير المجموع أو المعدل للصفة المدروسة، حيث تزيد فاعلية العينة ودقة التقديرات، ومن الأمثلة على ذلك لو طلب تقدير عدد أشجار الزيتون في منطقة معينة وكان معلوما لدينا أعداد الأشجار المثمرة للمجتمع ككل من خلال تعداد أجري قي سنة سابقة .

ويهدف هذا الأسلوب إلى تحسين دقة النتائج عن طريق الاستفادة من العلاقة بين متغيرات الدراسة (Y) والمتغيرات المساعدة (X)، ويمكن استخدام المعلومات المساعدة عند اختيار العينة، وعند تقسيم المجتمع إلى طبقات، وعند أجراء التقديرات، أما أهم التقديرات التي تستخدم فهي التقدير النسبي والتقدير عن طريق الانحدار الخطي.

## التقدير النسبي

تزيد فاعلية هذا الأسلوب عند وجود علاقة خطية بين المتغير قيد الدراسة (yi) والمتغير المساعد (xi)، كذلك عندما يكون التباين للمتغير قيد الدراسة (yi) يتناسب مع التباين للمتغير المساعد (xi).

الرموز والمصطلحات المستخدمة:

(R) نسبة المتغير Y إلى المتغير X

$$R = \frac{Y}{X} \quad (45)$$

النسبة المقدرة

$$\hat{R} = \frac{\sum_i y_i}{\sum_i x_i} = \frac{\bar{y}}{\bar{x}} \quad (46)$$

حيث $y_i$ هي قيمة الصفة المدروسة للمفردة i في العينة من خلال الدراسة .
حيث $x_i$ هي قيمة الصفة المدروسة للمفردة i في العينة من خلال المعلومات المساعدة.
أما المتوسط المقدر للمجتمع عن طريق التقدير النسبي $(\hat{\bar{Y}}_R)$

$$\hat{\bar{Y}}_R = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} \bar{X} = \frac{\bar{y}}{\bar{x}} \bar{X} \quad (47)$$

حيث $\bar{X}$ هي متوسط المجتمع للمعلومات المساعدة وهذا التقدير هو تقدير متحيز لذلك فإن

$$E(\hat{\bar{Y}}_R) \neq \bar{Y}$$

أما تقدير المجموع بإستخدام التقدير النسبي فهو $(\hat{Y}_R)$

$$(\hat{Y}_R) = \frac{\sum y_i}{\sum x_i} X = \frac{\overline{y}}{\bar{x}} X \quad (48)$$

حيث **x** هي مجموع قيم الصفة للمتغير المساعد

أما التباين المقدر فهو كما يلي:

**1. التباين المقدر للمجموع** $\hat{v}(\hat{y}_R)$

$$\hat{v}(\hat{y}_R) = N^2 \left(\frac{1}{n} - \frac{1}{N}\right) \left[\frac{\sum_{i=1}^{n} (y_i - \hat{R}_{xi})^2}{n-1}\right] \quad (49)$$

**2. التباين المقدر للمتوسط** $\hat{v}\left(\hat{\bar{Y}}_R\right)$

$$\hat{v}\left(\hat{\bar{Y}}_R\right) = \left(\frac{1}{n} - \frac{1}{N}\right) \left[\frac{\sum_{i=1}^{n} (y_i - \hat{R}_{xi})^2}{n-1}\right] \quad (50)$$

**3. التباين المقدر للنسبة** $\hat{V}(\hat{R})$

$$\hat{V}(\hat{R}) = \frac{1}{\bar{X}^2} \left(\frac{1}{n} - \frac{1}{N}\right) \left[\frac{\sum_{i=1}^{n} (y_i - \hat{R}_{xi})^2}{n-1}\right] \quad (51)$$

ويمكن كتابة المعادلات السابقة بالطريقة التالية:

### 1. التباين المقدر للمجموع

$$\hat{V}(\hat{Y}_R) = N^2\left(\frac{1}{n}-\frac{1}{N}\right)\left[S_y^2 + \hat{R}^2 S_x^2 - 2\hat{R}S_{xy}\right] \quad (52)$$

$$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x}) = \frac{1}{n-1}\sum x_i\, y_i - \frac{1}{n}\sum x_i \sum y_i \quad (53)$$

### 2. التباين المقدر للمتوسط

$$\hat{V}\left(\hat{\bar{Y}}_R\right) = \left(\frac{1}{n}-\frac{1}{N}\right)\left[S_y^2 + \hat{R}^2 S_x^2 - 2\hat{R}S_{xy}\right] \quad (54)$$

### 3. التباين المقدر للنسبة

$$\hat{V}(\hat{R}) = \frac{1}{\bar{X}^2}\left(\frac{1}{n}-\frac{1}{N}\right)\left[S_y^2 + \hat{R}^2 S_x^2 - 2\hat{R}S_{xy}\right] \quad (55)$$

* ولتقدير التحيز يمكن استخدام المعادلة التالية:

$$B\left(\hat{\bar{Y}}_R\right) = \bar{Y}\left(\frac{N-n}{N_n}\right)\left[c_x^2 - p_{cx}c_y\right] \quad (56)$$

حيث $B\left(\hat{\bar{Y}}_R\right)$ التحيز للمعدل المقدر بأسلوب التقدير النسبي $c_x$ و $c_y$ هي معاملات التباين لـ $y_i$ و $x_i$ في المجتمع.

$p$: هي معامل الارتباط بين $y_i$ و $x_i$ في المجتمع.

**مثال 1:**

في دراسة لتقدير أعداد الضأن في أحد الأقاليم سحبت عينه من ذلك الإقليم واجري لها مسح شامل فكانت النتائج كما يلي:

| القرية | عدد الضأن من المسح $y_i$ | عدد الضان من التعداد $x_i$ |
|---|---|---|
| 1 | 136 | 9 |
| 2 | 6 | 1 |
| 3 | 351 | 534 |
| 4 | 226 | 790 |
| 5 | 2772 | 2823 |
| 6 | 715 | 997 |
| 7 | 401 | 67 |
| 8 | 428 | 260 |
| 9 | 3404 | 2607 |
| 10 | 758 | 663 |
| 11 | 846 | 1959 |
| 12 | 3706 | 3575 |
| 13 | 7698 | 4504 |
| 14 | 3062 | 2545 |
| 15 | 836 | 1785 |
| 16 | 791 | 1827 |
| 17 | 2700 | 3502 |
| 18 | 3764 | 3216 |
| **المجموع** | 32600 | 31664 |

مع العلم أن عدد الضأن الإجمالي في الإقليم حسب نتائج التعداد الذي اجري في عام 1991 كان 212,722 رأسا وأن أجمالي عدد قرى الإقليم هو 95 قرية والمطلوب تقدير أعداد الضأن لعام 1993 للإقليم والخطأ المعياري النسبي.

**الإجابة:**

أعداد الضأن المقدر لعام 1993 $(\hat{Y}_R)$

$$\hat{Y}_R = \frac{\sum y_i}{\sum x_i} \times$$
$$= \frac{32600}{31664}(212722)$$
$$= 219010$$

**1. التباين المقدر للمجموع**

$$\hat{V}(\hat{Y}_R)$$
$$\hat{V}(\hat{Y}_R) = N^2(\frac{1}{n} - \frac{1}{N})(S_{yi}^2 + \hat{R}^2 S_{xi} - 2\hat{R}S_{yx})$$
$$S_y^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1}$$
$$= 3983472.58$$
$$S_x^2 = 1999555.87$$
$$S_{yx}^2 = \frac{1}{n-1}(\sum x_i y_i - \frac{\sum x_i \sum y_i}{n})$$
$$= \frac{1}{17}(100289031 - 57347022.2)$$
$$= 2526000.5$$

$$\hat{V}(\hat{Y}_R) = 341690746$$
$$\boldsymbol{S.E} = 18484.88$$
$$\boldsymbol{S.E}\,\% = 8.4402\%$$

الفصل السادس

# العينة المتناسبة مع الحجم
# Sampling with varying probabilities

## العينة المتناسبة مع الحجم

هي العينة التي يكون احتمال ظهور كل وحده معاينة في العينة يتناسب مع حجم تلك الوحدة للصفة المدروسة، فعلى سبيل المثال القرية التي يزرع فيها محصول القمح بمساحة 1000 دونم تُعطى احتمال ظهور 10 أضعاف احتمال القرية التي يزرع فيها 100 دونم قمح عند دراسة المساحة المزروعة بهذا المحصول.

## أسلوب اختيار العينة:

- **العينة المتناسبة مع الحجم مع الإرجاع:**

هناك أسلوبين لاختيار وحدات المعاينة

1. **أسلوب التجميع التراكمي:**

ويتلخص هذا الأسلوب بإعطاء كل وحدة معاينة رقم يوازي الصفة التي تحملها فمثلا القرية التي مساحتها 1000 دونم يعتبر وزنها 1000 أي أنها تحتوي على 1000 وحدة معاينة فرضية، والقرية التي مساحتها 100دونم يكون وزنها 100 وهكذا.

وللتوضيح نأخذ المثال التالي:

| القرية | مساحة القرية (دونم) | المساحة التجميعية | الأرقام التي تناظرها |
|---|---|---|---|
| 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 |

ولاختيار ثلاث قرى عينة يتم سحب ثلاثة أرقام عشوائية تقع بين 1و4500 بإستخدام الجدول العشوائي، فإذا اختيرت الأرقام العشوائية 75،2000،4000 فإن قرى العينة تكون القرية 1،3،6.

### 2. طريقة لاهير (Lahir):

في الطريقة السابقة عندما يكون المجتمع كبير فإن عملية سحب العينة تستغرق وقتا طويلا لذا فقد أُقترح طريقة أخرى تؤدي نفس الغرض حيث أنه في هذا الأسلوب لا توجد ضرورة لكتابة المجموع التجميعي، حيث يجري سحب أزواج من الأرقام العشوائية حيث يمثل الرقم الأول من كل زوج رقم وحده المعاينة بحيث يسحب الرقم العشوائي بين 1 و N حيث N تمثل عدد وحدات المعاينة الإجمالية في المجتمع ويمثل الرقم الثاني من كل زوج حجم وحدة المعاينة بحيث يسحب الرقم العشوائي الثاني بين I و M حيث M تمثل حجم اكبر وحدة معاينة موجودة في المجتمع للصفة التي يتم الوزن على أساسها.

**مثال 2:**

لاختيار ثلاث وحدات معاينة من المثال السابق

| القرية | المساحة |
|---|---|
| 1 | 1000 |
| 2 | 700 |
| 3 | 1200 |
| 4 | 500 |
| 5 | 300 |
| 6 | 800 |

**الأرقام العشوائية المختارة:**

(3،500)،(2،1200)،(6،600) يلاحظ في الزوج الثاني أن الرقم العشوائي الثاني (1200) أكبر من حجم القرية (700) لذا يسحب زوج جديد بدلا منه.

## 3. أسلوب التقدير

Pi = احتمالية ظهور الوحدة i من المجتمع
حيث مجموع احتمالات الظهور لجميع وحدات المعاينة للمجتمع =1

$$\sum_{i=1}^{N} P_i = 1 \quad (57)$$

لتقدير المتوسط تستخدم المعادلة التالية

$$\hat{\bar{Y}}_{PPS} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{NP_i} \quad (58)$$

لتقدير النسبة نستخدم المعادلة التالية

$$P_i = \frac{X_i}{X} \quad (59)$$

لتقديـر المجمـوع

$$\hat{Y}_{PPS} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{P_i} \quad (60)$$

$$V_{PPS}(\hat{Y}_{PPS}) = \frac{1}{n}\left[\sum_{i=1}^{n}\frac{{y_i}^2}{P_i} - {y_n}^2\right] \quad (61)$$

$$V_{PPS}(\hat{Y}_{PPS}) = \frac{1}{n(n-1)N^2}\sum_{i=1}^{n}\left(\frac{y_i}{P_i} - \hat{y}_{PPS}\right)^2 = \frac{1}{n(n-1)}\left[\sum_{i=1}^{n}\left(\frac{y_i}{P_i}\right)^2 - n\hat{y}_{PPS}^2\right] \quad (62)$$

**مثال 3:**

في دراسة لتقدير أعداد العاملين في مصانع المواد الغذائية سحبت عينة متناسبة مع الحجم مع الإرجاع حجمها 10 مصانع من اصل 80 مصنعا هي إجمالي عدد المصانع في المحافظة وكانت نتائج المسح كما يلي:

| المصنع | عدد العاملين عام 90 | عدد العاملين من التعداد عام 89 | $pi = xi/x$ |
|---|---|---|---|
| 1 | 20 | 18 | .006 |
| 2 | 30 | 32 | .010666 |
| 3 | 10 | 10 | .00333 |
| 4 | 6 | 6 | .002 |
| 5 | 12 | 10 | .00333 |
| 6 | 50 | 48 | .016 |
| 7 | 4 | 6 | .002 |
| 8 | 80 | 84 | .028 |
| 9 | 8 | 10 | .00333 |
| 10 | 10 | 12 | .004 |

مع العلم أن إجمالي عدد العاملين لجميع المصانع لتعداد سابق كان 3000 عامل وأن عدد العاملين للمصانع الواردة في العينة حسب نتائج التعداد كانت كما هي مبينة في الجدول أعلاه.

فإن متوسط عدد العاملين للمصنع الواحد يساوي $\hat{\bar{y}}_{pps}$

$$\hat{\bar{y}}_{pps} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{N_{P_i}} = \frac{1}{10}\times\sum_{i=1}^{n}\frac{y_i}{N_{P_i}}$$

| $y_i$ | $P_i$ | $\frac{y_i}{N_{P_i}}$ | $\left(\frac{y_i}{P_i} - \hat{y}_{PPS}\right)^2$ |
|---|---|---|---|
| 20 | .006 | 41.666 | 205511.1 |
| 30 | .01066 | 35.158 | 4532.5 |
| 10 | .0033 | 37.538 | 15129.7 |
| 6 | .002 | 37.5 | 14400.0 |
| 12 | .00333 | 45.045 | 518918.6 |
| 50 | .016 | 39.0625 | 60025.0 |
| 4 | .002 | 25.0 | 774400.0 |
| 80 | .028 | 35.714 | 522.4 |
| 8 | .00333 | 30.03 | 228099.5 |
| 10 | .009 | 31.25 | 144400.0 |
| | | 357.9635 | 1965938.8 |

$$\hat{\bar{y}}_{pps} = \frac{1}{10} \times 357.9635 = 35.79635 \approx 36$$

$$\hat{y}_{PPS} = N\hat{\bar{y}}_{pps} = 2880$$

$$\hat{v}_{pps}(\hat{\bar{y}}_{pps}) = \frac{1}{n(n-1)N^2} \sum_{i=1}^{n} \left(\frac{y_i}{P_i} - \hat{y}_{PPS}\right)^2$$

$$= \frac{1}{10 \times 9 \times 80^2} \times 1965938.8 = 3.413$$

$$S.E = 1.84745$$

$$S.E\% = 5.13\%$$

**مثال 4:**

في دراسة لتقدير معدل إنتاج البستان الواحد من محصول الزيتون في إحدى القرى سحبت عينة متناسبة مع الحجم مع الإرجاع بأسلوب الطريقة التجميعية حجمها 15 بستانا وجمع بيانات عن البساتين الواردة قي العينة كما يلي:

| عدد الأشجار المثمرة في البستان $x_i$ | الإنتاج (طن) $y_i$ | $\frac{y_i}{x_i}$ |
|---|---|---|
| 20 | 0,2 | .010 |
| 100 | 1,5 | .015 |
| 120 | 1,8 | .015 |
| 80 | 0,9 | .01125 |
| 50 | 0,5 | .01 |
| 60 | 0,4 | .006666 |
| 90 | 1,2 | .0137 |
| 40 | 0,6 | .015 |
| 150 | 2,2 | .01466 |
| 170 | 2,0 | .01176 |
| 30 | 0,2 | .00666 |
| 70 | 0,8 | .0114 |
| 130 | 1,6 | .012307 |
| 140 | 1,8 | .012857 |
| 130 | 1,4 | .010769 |
| | | .17664166 |

مع العلم أن عدد البساتين في القرية هو120 بستانا وتحتوي على 12000 شجرة زيتون مثمرة.

معدل إنتاج البستان المقدر من الزيتون= $\hat{\bar{y}}_{pps}$

$$\hat{\bar{y}}_{pps} = \frac{x}{nN} \times \sum_{i=1}^{n} \frac{y_i}{x_i}, x = 12000, \quad n = 15, \quad N = 120$$

$$\hat{\bar{y}}_{pps} = \frac{12000}{15 \times 120} \times 0.17664166 = 1.1776/\text{طن}$$

أما التباين المقدر للمتوسط $\hat{v}_{pps}(\hat{\bar{y}}_{pps})$

$$V_{PPS}(\hat{\bar{Y}}_{PPS}) = \frac{1}{n(n-1)}\left[x^2 \sum_{i=1}^{n} \left(\frac{y_i}{x_i}\right)^2 - n\hat{y}_{PPS}^2\right]$$
$$= \frac{1}{15 \times 14}[12000^2 \times 0.0021834 - 20 \times 1.1776^2]$$

**المزايا التي نحصل عليها من استخدام أسلوب العينة المتناسبة مع الحجم:**

أن عملية مقارنة النتائج التي نحصل عليها بأسلوب العينة المتناسبة مع الحجم ومقارنة ذلك بما يمكن الحصول عليه في حالة استخدام العينة العشوائية البسيطة لنفس العينة يعتبر شيء جيد.

التباين المقدر للعينة العشوائية البسيطة $\hat{v}_{pps}(\hat{\bar{y}}_{srs})$ لعينة مسحوبة بأسلوب العينة المتناسبة مع الحجم =

$$\hat{v}_{pps}(\hat{\bar{y}}_{srs}) = N\frac{N-n}{n}\frac{1}{N-1}\left[\frac{1}{nN^2}\sum_{i=1}^{n}\left(\frac{y_i}{P_i}\right)^2 - \frac{1}{N}\left\{\hat{\bar{y}}_{PPS}{}^2 - \hat{v}_{pps}(\hat{\bar{y}}_{PPS})\right\}\right] \quad (63)$$

نسبة الزيادة في الفاعلية المكتسبة من استخدام العينة المتناسبة مع الحجم مقارنة بالعينة العشوائية البسيطة

$$EF = \frac{\hat{v}_{pps}(\hat{\bar{y}}_{srs}) - \hat{v}_{pps}(\hat{\bar{y}}_{pps})}{\hat{v}_{pps}(\hat{\bar{y}}_{pps})} \times 100 \quad (64)$$

$$\hat{v}_{pps}(\hat{\bar{y}}_{srs}) = 80 \times \frac{80-10}{10} \times \frac{1}{80-1}\left[\frac{1}{10\times 80^2}\sum_{i=1}^{n}\left(\frac{y_i}{P_i}\right)^2 - \frac{1}{80}\{35.79635^2 - 3.413\}\right]$$
$$= 7.088886\left[\frac{1}{10\times 80^2} \times 665792.4664 - 15.975\right] = 73.741$$

$$EF = \frac{\hat{v}_{pps}(\hat{\bar{y}}_{srs}) - \hat{v}_{pps}(\hat{\bar{y}}_{pps})}{\hat{v}_{pps}(\hat{\bar{y}}_{pps})} \times 100 = 2060.59$$

- **العينة المتناسبة مع الحجم المسحوبة بانتظام**
  **PPS Systematic sampling**

تعتمد هذه العينة على اختيار وحدات المعاينة بأسلوب الاختيار العينة المنتظمة والتي تكون غالبا فعّالة أكثر من العينة المتناسبة مع الحجم مع الإرجاع بالإضافة الى سهولة سحب العينات واختصار الوقت مقارنة بالأساليب الأخرى.

ويتمثل هذا الأسلوب بإعداد إطار بوحدات المعاينة وقيمة الصفة المراد دراستها لكل وحدة معاينة، بعد ذلك يجري إختار رقم عشوائي يقع بين 1 و K ويرمز له بالرمز R، حيث K=TN/N، وتكون وحدات المعاينة المسحوبة في العينة هي JK+R؛ حيث J= (0, 1, 2, ... n-1).

**ويمكن توضيح الأسلوب بالمثال التالي:**

لتقدير أعداد الضأن والماعز في احد التجمعات تم تبويب بيانات التعداد الذي أجري في عام 1991 والمطلوب سحب عينة حجمها 20% من قرى التجمع بأسلوب العينة المتناسبة مع الحجم المسحوبة بأسلوب المنتظمة.

تم حساب التجميعي للصفة المدروسة $TI = \sum_J^I XJ$ حيث أن XJ هي حجم الصفة المدروسة لوحدة المعاينة I.

| رقم القرية | اعداد الضأن والماعز | العدد التراكمي للضأن والماعز | رقم القرية | اعداد الضأن والماعز | العدد التراكمي للضأن والماعز |
|---|---|---|---|---|---|
| 1 | 2202 | 2202 | 18 | 55 | 50135 |
| 2 | 63 | 2265 | 19 | 2286 | 52421 |
| 3 | 5550 | 7815 | 20 | 1403 | 53824 |
| 4 | 10443 | 18258 | 21 | 2319 | 56143 |
| 5 | 97 | 18355 | 22 | 4367 | 60510 |
| 6 | 1451 | 19806 | 23 | 5258 | 65768 |
| 7 | 186 | 19992 | 24 | 331 | 66099 |
| 8 | 3498 | 23490 | 25 | 1162 | 67261 |
| 9 | 3510 | 27000 | 26 | 199 | 67460 |
| 10 | 4871 | 31871 | 27 | 225 | 67685 |
| 11 | 4949 | 36820 | 28 | 45 | 67730 |
| 12 | 1582 | 38402 | 29 | 1948 | 69678 |
| 13 | 399 | 38801 | 30 | 1357 | 71035 |
| 14 | 6826 | 45627 | 31 | 2115 | 73150 |
| 15 | 1866 | 47493 | 32 | 3137 | 76287 |
| 16 | 2426 | 49919 | 33 | 1252 | 77539 |
| 17 | 161 | 50080 | | | |

حجم العينة n= 7 قرى

$$k = \frac{TN}{n} = \frac{77539}{7} = 11077$$

الرقم العشوائي الأول R=7320
قرى العينة هي: 3،6،10،14،19،23،32.
وقد أجرى مسح شامل لهذه القرى في عام 1992 فكانت النتائج كما يلي:

| القرية | عدد الضأن والماعز عام 92 | عدد الضأن والماعز من التعداد 91 |
|---|---|---|
| 3 | 7424 | 5550 |
| 6 | 1882 | 1451 |
| 10 | 7360 | 4871 |
| 14 | 5476 | 6826 |
| 19 | 2440 | 2286 |
| 23 | 4926 | 5258 |
| 32 | 1482 | 3137 |

لتقدير عدد الضأن والماعز في التجمع لعام 1992

$$\hat{Y} = \sum \frac{y_i}{n_{p_i}} \quad (65)$$

$$p_i = \frac{x_i}{x} \quad (66)$$

$$x = 77539$$

| القرية | npi | yi/npi | Pi |
|---|---|---|---|
| 1 | .501 | 14818 | .0716 |
| 2 | .131 | 14366 | .0187 |
| 3 | .440 | 16727 | 0.0628 |
| 4 | .616 | 8890 | .0880 |
| 5 | .206 | 11845 | 0.0295 |
| 6 | .475 | 10371 | 0.0678 |
| 7 | .283 | 5237 | 0.0405 |
| المجموع | | 82254 | |

إذا كان عدد الضأن والماعز المقرر لعام 1992 هو 82254 رأس.

أما التباين المقدر فيمكن حسابه بالمعادلة التالية:

$$\hat{v}(y) = \frac{1}{n}\sum\left(\frac{y_i}{p_i} - y\right)^2 p_i\,[i - (n-1)p_i] \qquad (67)$$

$$= \frac{1}{7}[(14818 - 82254)^2 \times 4.0716[1 - 6 \times 0.716] + \cdots]$$
$$= 1163877283$$

$$S.E = 3411564, \qquad S.E\% = 41.47\%$$

الفصل السابع

# العينة العنقودية
# Cluster Sample
# (CL)

# مقدمـــة

إن فكرة العينة العنقودية تتمثل بشكل رئيسي بتقسيم المجتمع إلى مجموعات بشكل مناسب، بحيث تكون متقاربة بالحجم وتكون هذه المجموعات متجانسة بالنسبة للصفة المدروسة، وكل مجوعة من هذه المجموعات تسمى عنقود، وتشكل العناقيد المجتمع كاملا دون حذف أو تكرار.

من ميزات العينة العنقودية أنها فعّالة بالنسبة لوحدة التكاليف حيث تعطي دقة اكثر لوحدة الكلفة، كذلك في كثير من الأحيان يصعب توفير إطار حديث بكل مفردة من مفردات المجتمع، ولكن من الممكن توفير إطار بالعناقيد مما يوفر بالجهد والوقت، وفي حالات يصعب إحصاء كل وحدة معاينة لوحدها، فمثلا لمعرفة متوسط استهلاك الفرد من الخبز يحتاج الأمر إلى سحب عينة أُسر وجمع بيانات عن استهلاك الأسرة كاملا وبقسمة الاستهلاك على عدد أفراد الأسرة يتم الحصول على متوسط استهلاك الفرد، كذلك يوجد محدد أخر وهو توفر وسائل النقل حيث أن من مميزات العينة العنقودية أنها تقلل كلفة التنقل بين وحدات المعاينة ، ومن عيوب العينة العنقودية أنها اقل فاعلية من العينة العشوائية إذا كان حجم العينة ثابت وكذلك فان العينة أقل انتشاراً .

**وعند استخدام العينة العنقودية يجب مراعاة ما يلي:**

1. أن يكون حجم العنقود صغير نسبيا وعدد العناقيد كبير وذلك لضمان انتشار أفضل لوحدات المعاينة.
2. عند تكوين العناقيد تؤخذ مفردات المجتمع المتجاورة أو ضمن منطقة معينة في نفس العنقود، حيث تكون غالبا متشابهة للصفة المدروسة.
3. أن يكون حجم العناقيد متقارب قدر الإمكان.
4. يجب أن يكون كل عنقود موضّح ومعّرف لجامع البيانات، له حدود واضحة ويمكن الإستدلال عليه بسهولة.

يمكن أن تكون العناقيد طبيعية مثل أسر أو قرى وغيرها وقد تكون مقسمه من قبل العاملين في المسح عن طريق تقسيم المجتمع إلى مناطق عد حيث كل منطقة تعتبر عنقود، وقد يقسم عن طريق سحب عينة من المجتمع واعتبارها عينة مفتاح وضم وحدات المعاينة المجاورة لها لتكوين العناقيد وفق أسس علمية.

# أسلوب التقديرات

### الرموز المستخدمة

**N**: عدد العناقيد في المجتمع.
**n**: عدد العناقيد في العينة.
$M_i$: عدد وحدات المعاينة الثانوية في العنقود **(i)**.
**M**: إجمالي عدد وحدات المعاينة الثانوية في المجتمع.
$\frac{M}{N} = \overline{M}$ : متوسط عدد وحدات المعاينة الثانوية للعنقود الواحد.
$Y_{ij}$ : قيمة الصفة المراد دراستها لوحدة المعاينة الثانوية **(j)** في العنقود **(i)** في المجتمع.
$\overline{Y}_i$ : متوسط الصفة المراد دراستها للعنقود **(i)**.
$\overline{Y}_N$: متوسط لمتوسطات العناقيد في المجتمع

$$\overline{Y}_N = \frac{1}{N}\sum_{i=1}^{N}\overline{Y}_i \qquad (68)$$

**1. إذا كان حجم العناقيد متساوي فان:**

$$M_i = \frac{M}{N} = \overline{M} \qquad (69)$$

$$\overline{Y}_i = \frac{1}{M}\sum_{J=1}^{M}\overline{Y}_{iJ} \qquad (70)$$

$$\overline{Y}_N = \frac{1}{N}\sum_{i=1}^{N}\overline{Y}_i \qquad (71)$$

المتوسط المقدر للمجتمع $\hat{\bar{Y}}_{CL}$

$$\hat{\bar{Y}}_{CL} = \frac{1}{n}\sum_{i=1}^{n} \bar{Y}_i \quad (72)$$

وتباينه $v\left(\hat{\bar{Y}}_{CL}\right)$ يساوي

$$v\left(\hat{\bar{Y}}_{CL}\right) = \left(\frac{1}{n} - \frac{1}{N}\right) S_b^2 \quad (73)$$

$$S_b^2 = \frac{1}{N-1}\sum_{i=1}^{N} (\bar{Y}_i - \bar{Y})^2 \quad (74)$$

والتباين المقدر $\hat{v}\left(\hat{\bar{Y}}_{CL}\right)$

$$\hat{v}\left(\hat{\bar{Y}}_{CL}\right) = \left(\frac{1}{n} - \frac{1}{N}\right) s_b^2 \quad (75)$$

$$s_b^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\bar{Y}_i - \bar{Y}_{CL})^2 \quad (76)$$

**مثال 4:**

في دراسة لتقدير إنتاجية الشجرة الواحدة من الليمون في إحدى القرى والتي تحتوي على 800 شجرة قسمت إلى عناقيد كل عنقود يتكون من 4 شجرات، وسحبت عينة تتكون من 4 عناقيد فكانت النتائج كما يلي:

| رقم العنقود | إنتاج الشجرة الأولى | إنتاج الشجرة الثانية | إنتاج الشجرة الثالثة | إنتاج الشجرة الرابعة |
|---|---|---|---|---|
| 1 | 20 | 30 | 25 | 15 |
| 2 | 40 | 25 | 30 | 35 |
| 3 | 30 | 20 | 25 | 35 |
| 4 | 25 | 40 | 35 | 20 |

**أجب على ما يلي:**

- قّدر إنتاجية الشجرة الواحدة بحدود ثقة 95 %

$$\hat{\bar{Y}}_{CI} = \frac{1}{n}\sum_{i=1}^{n} \bar{y}_i$$

$$\bar{y}_i = \frac{1}{M_i}\sum_{i=1}^{M_i} \bar{Y}_{iJ}$$

$$\bar{y}_1 = \frac{20 + 30 + 25 + 15}{4} = 22.5$$

$$\bar{y}_2 = 32.5, \quad \bar{y}_3 = 27.5, \quad \bar{y}_4 = 30$$

$$\hat{\bar{Y}}_{CI} = \frac{22.5 + 32.5 + 27.5 + 30}{4} = 28.125$$

$$\hat{v}\left(\hat{\bar{Y}}_{CL}\right) = \left(\frac{1}{4} - \frac{1}{20}\right) s_b^2$$

$$s_b^2 = \frac{1}{4-1}(31.64 + 19.14 + 0.39 + 3.52) = 18.23$$

$$\hat{v}\left(\hat{\bar{Y}}_{CL}\right) = \left(\frac{1}{4} - \frac{1}{20}\right) 18.23 = 4.46635$$

$$S.E = 2.1133$$

$$C.V\% = 7.51\%$$

$$C.I = \hat{\bar{y}} \pm t_{\frac{\alpha}{2}} SE = 28.125 \pm 1.96 \times 2.1133 = (23.98, 32.267)$$

### 1. إذا كان حجم العناقيد غير متساوٍ:

عندما يكون حجم العناقيد غير متساوٍ فأن تقدير متوسط المجتمع $(\hat{\bar{Y}})$ لصفة معينة يحسب وفق المعادلة التالية:

$$\hat{\bar{Y}}_{CL} = \frac{1}{n}\sum_{i=1}^{n} \bar{Y}_i \quad (77)$$

وتقدير تباين المتوسط $v(\hat{\bar{Y}}_{CL})$ بإستخدام المعادلة التالية:

$$v(\hat{\bar{Y}}_{1CL}) = \left(\frac{1}{n} - \frac{1}{N}\right) S_{1b}^2 \quad (78)$$

$$S_{1b}^2 = \frac{1}{N-1}\sum (\bar{Y}_i - \bar{Y}_N)^2 \quad (79)$$

ويمكن الحصول على تقدير لمتوسط المجتمع بإستخدام المعادلة التالية:

$$\hat{\bar{Y}}_{2CL} = \frac{1}{n}\sum_{i=1}^{n} \frac{M_i \bar{Y}_i}{\bar{M}} \quad (80)$$

وتباينها يساوي

$$V(\hat{\bar{Y}}_{2CL}) = \left(\frac{1}{n} - \frac{1}{N}\right) s_{2b}^2 \quad (81)$$

$$s_{2b}^2 = \frac{1}{N-1}\sum_{i=1}^{N} \left(\frac{M_i}{\bar{M}} - \bar{Y}\right)^2 \quad (82)$$

أما التباين المقدر

$$\hat{V}(\hat{\bar{Y}}_{2C}) = \left(\frac{1}{n} - \frac{1}{N}\right) S_{2b}^2 \quad (83)$$

$$S_{2b}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{M_i}{\bar{M}}\bar{Y}_i - \hat{\bar{Y}}_{2CL}^2\right) \quad (84)$$

وهذه التقديرات تعتبر متحيزة.

**مثال:**

في دراسة لتقدير إنتاجية الرأس الواحد سنويا من الأبقار الحلوب سحبت عينة حجمها أربع مزارع أبقار في إحدى المناطق الزراعية والتي تحتوي على 30 مزرعة تحتوي على 1200 رأس من الأبقار الحلوب فكانت نتائج المسح كما يلي:

| رقم المزرعة في العينة | عدد الأبقار الحلوب | إنتاجية الرأس الواحد في المزرعة |
|---|---|---|
| 1 | 50 | 4 |
| 2 | 25 | 3.7 |
| 3 | 60 | 4.2 |
| 4 | 30 | 3.5 |

**الإجابة:**

$$\hat{\bar{Y}}_{2CI} = \frac{1}{n}\sum_{i=1}^{n}\frac{M_i\bar{Y}_i}{\bar{M}} = \frac{1}{4}((50\times 4) + (25\times 3.7) + (60\times 4.2) + (30\times 3.5)/1200/30)$$

$$= \frac{1}{4}\left(\frac{(50\times 4) + (25\times 3.7) + (60\times 4.2) + (30\times 3.5)}{(1200/30)}\right) = 3.46875$$

$$\hat{V}(\hat{\bar{Y}}_{2CI}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{2b}^2$$

$$S_{2b}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{M_i}{\bar{M}}\bar{Y}_i - \hat{\bar{Y}}_{2C}\right)^2$$

$$= \frac{1}{4-1}\left(\frac{50}{40}4 - 3.46875\right)^2 + \left(\frac{25}{40}3.7 - 3.46875\right)^2 + \left(\frac{30}{40}3.5 - 346875\right)^2$$

$$= \frac{1}{3}(2.34 + 1.34 + 8.02 + 0.71) = 4.137$$

$$\hat{V}(\hat{\bar{Y}}_{2CI}) = \left(\frac{1}{4} - \frac{1}{30}\right)4.137 = 0.89635$$

$$S.E = 0.95 \qquad C.V\% = 27.29\%$$

ويمكن الحصول على تقديرات غير متحيزة بإستخدام المعادلة التالية:

$$\bar{y} = \frac{\sum_i^n M_i \bar{y}_i}{n\bar{M}}$$

$$V(\bar{y}) = \frac{(1-f)}{n} S_b^2$$

$$S_b^2 = \frac{1}{N-1} \sum_i^N \frac{M_I}{\bar{M}} \bar{y}_i - \bar{Y})^2$$

أما التباين المقدر $\hat{V}(\bar{y})$

$$\hat{V}(\bar{y}) = \frac{1-f}{n} S_b^2$$

$$S_b^2 = \sum_i^n \frac{(\frac{M_I \bar{y}_i}{\bar{M}} - \bar{y}_n)^2}{n-1}$$

**مثال:**

في دراسة لتقدير متوسط وزن الطفل في دور الحضانة للأطفال في عمر السنتين سحبت عينة تمثل (3) دور حضانة من أصل (20) دار موجودة في المنطقة فكانت النتائج كما يلي:

| رقم الدار في العينة (العنقود) | عدد الأطفال | متوسط وزن الطفل |
|---|---|---|
| 1 | 20 | 12 |
| 2 | 10 | 10 |
| 3 | 30 | 8 |

إذا علمت أن إجمالي عدد الأطفال في جميع دور الحضانة هو (300) طفل.
احسب متوسط وزن الطفل بحدود ثقة 95% وقدر معامل التباين (C.V%).

**الإجابة:**

$M = 300$
$N = 20$
$n = 3$

$$\bar{y} = \sum_{i=1}^{n} M_i \bar{y}_i / n\bar{M} = \frac{580}{60} = 9.67$$

$$\hat{V}(\bar{y}) = \frac{1-f}{n} \sum (\frac{M_i \bar{y}_i}{\bar{M}} - \bar{y})^2 = (\frac{1-\frac{3}{20}}{3})(\frac{1}{2})(5.4289 + 21.8089 + 5.4289) = 4.63$$

$S.E = 2.15$ C.V% $= 22.24\%$

C.I $= \bar{y} \mp t_2$ S.E $9.67 \mp 1.996\ (2.15)$ $(5.456, 13.884)$

الفصل الثامن

# العينـة متعـددة المراحل
# Multistage Sampling

## مقدمـــــة

في كثير من الأحيان لا يتوفر إطار حديث لوحدات المعاينة مثل الحيازات الزراعية أو المساكن وغيرها، ويكون من الصعب إعداد إطار حديث لها، وفي نفس الوقت يتوفر إطار لوحدات المعاينة الاولية مثل القرى أو العناقيد وغيرها، ويعتبر التغير على وحدات المعاينة الاولية ابطأ نسبيا من التغير على وحدات المعاينة الثانوية مثل الأسر.

## ميزات العينة متعددة المراحل

- توفر في الوقت والمال حيث يكتفى بإعداد الإطار بوحدات المعاينة الثانوية لوحدات المعاينة الرئيسية المسحوبة بالعينة.
- توفر في كُلف التنقلات.
- هذا التصميم مرن حيث أنه من الممكن استخدام أسلوب سحب العينات في كل مرحلة مختلف عن أسلوب سحب العينة في المراحل الأخرى.

## حجم وحدات المعاينة الرئيسية

يفضل تقسيم المجتمع إلى وحدات معاينة رئيسية متساوية ومتوسطة الحجم وذلك لانه:

- 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 بين وحدات المعاينة عند تنفيذ المسح.

## أسلوب اختيار وحدات المعاينة الرئيسية

1. إذا كانت وحدات المعاينة الأولية متجانسة يمكن استخدام أسلوب العينة العشوائية البسيطة.
2. إذا وجد تفاوت يمكن تقسيم وحدات المعاينة الرئيسية إلى طبقات وسحب عينة من كل طبقة.
3. إذا كان التفاوت كبير يمكن استخدام العينة المتناسبة مع الحجم.
4. يمكن استخدام العينة المنتظمة ولكن مشكلتها صعوبة الحصول على تقدير غير متحيزللتباين.

## أسلوب سحب وحدات المعاينة الثانوية

- العشوائية البسيطة
- المنتظمة
- الطبقية
- العنقودية
- المتناسبة مع الحجم

## أسلوب التقدير

عدد وحدات المعاينة الرئيسية $N$

عدد وحدات العينة الرئيسية $n$

عدد وحدات المعاينة الثانوية $M_i$

عدد وحدات العينة الثانوية $m_i$

$$M^{\circ} = \sum M_i \qquad m^{\circ} = \sum M_i \tag{85}$$

$$\bar{M} = \frac{M_{\circ}}{N} \qquad \bar{m} = \frac{M_{\circ}}{n} \tag{86}$$

متوسط العنقود $(i)$

$$\bar{Y}_i = \frac{1}{M_i} \sum_{g=1}^{M_i} Y_{ig} \tag{87}$$

متوسط المجتمع

$$\bar{Y} = \frac{\sum M_i \bar{Y}_i}{\sum M_i} \tag{88}$$

المتوسط المقدر للعنقود (i)

$$\bar{y}_i = \frac{1}{m_i}\sum_{g=1}^{M_i} y_{ig} \quad (89)$$

المتوسط المقدر للمجتمع

$$\hat{\bar{Y}} = \frac{1}{n}\sum \frac{M_i \hat{\bar{Y}}_i}{\bar{M}} \quad (90)$$

التباين المقدر للمتوسط

$$V(\hat{\bar{Y}}) = (\frac{1}{n} - \frac{1}{N})\ S^2{}_{\text{ib}} + \frac{1}{n_N}\sum \frac{M_i^2}{\bar{m}^2}(\frac{1}{m_i} - \frac{1}{M_i})\ S_{\text{li}}{}^2 \quad (91)$$

حيث:

$$S_{\text{ib}}{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\frac{m_i \bar{Y}_i}{\bar{M}} - \bar{Y})^2 \quad (92)$$

$$S_{\text{li}}{}^2 = \frac{1}{m_i - 1}\sum_{g=1}^{M_i}(Y_{ig} - \bar{Y}_I)^2 \quad (93)$$

$$\hat{V}(\hat{\bar{Y}}) = (\frac{1}{n} - \frac{1}{N})s_{ib}{}^2 + + \frac{1}{n_N}\sum \frac{M_i^2}{\bar{m}^2}(\frac{1}{m_i} - \frac{1}{M_i})\ s_{\text{li}}{}^2 \quad (94)$$

$$S_{ib}{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\frac{M_i \hat{\bar{Y}}}{\bar{m}} - \hat{\bar{Y}})^2 \quad (95)$$

$$S_{li}{}^2 = \frac{1}{m_i - 1}\sum_{y=1}^{m_i}(y_{ig} - \hat{\bar{Y}}_i)^2 \quad (96)$$

**مثال 1:**

في دراسة لتقدير نسبة المطاعم التي لا تلتزم بالأسعار المقررة من قبل وزارة التموين في إحدى المناطق والتي تحتوي على (40) تجمع سحبت عينة حجمها (3) تجمعات وسحبت عينة من كل تجمع تتكون من (5) مطاعم فكانت النتائج كما يلي:

| التجمع (العنقود) | عدد المطاعم في التجمع $(M_i)$ | عدد المطاعم في العينة $(m_i)$ | عدد المطاعم التي لا تلزم بالتسعيرة $a_i$ |
|---|---|---|---|
| 1 | 20 | 5 | 3 |
| 2 | 15 | 5 | 2 |
| 3 | 30 | 5 | 2 |

قدر نسبة المطاعم التي لا تلتزم بالأسعار المقررة من قبل وزارة التموين بحدود ثقة 95% علما بأن أجمالي عدد المطاعم في المحافظة هو (1000) مطعم.

**الإجابة:**

$$N = 40$$
$$n = 3$$
$$m = 5$$
$$\hat{P}_{csR} = \frac{\sum_{i=1}^{n} M_i\, p_i}{\sum_{i=1}^{n} M_i} = \frac{12 + 6 + 12}{65} = .4615$$
$$S^2(\hat{P}_{csR}) = (\frac{N-n}{N}(\frac{1}{n\bar{M}^2})St^2 + \frac{1}{nN\bar{M}^2}\sum {M_i}^2 (\frac{M_i - m_i}{M_i})(\frac{p_i q_i}{m_i - 1})$$
$${S_t}^2 = \frac{1}{n-1}\sum {M_i}^2 (\hat{p}_i - \hat{P}_{csR})^2 = \frac{1}{2}(7.67 + .852 + 3.41)$$
$$S^2(\hat{P}_{csR}) = (\frac{40-3}{40})(\frac{5.966}{3 \times 625}) + \frac{1}{3 \times 40 \times 625}(18 + 9 + 45) = 0.00318 + .00096$$
$$= 0.00414$$
$$S.E = 0.064357 \qquad C.1 = .4615 \mp 1.96(.064357)$$

**مثال 2:**

في دراسة لتقدير متوسط دخل الأسرة السنوي بالألف دينار في إحدى المدن والتي تحتوي على (48) قطاع سحب منها في العينة (3) قطاعات وسحب من كل قطاع عينة تتكون من أربع أسر فكانت النتائج كما يلي:

| التجمع (العنقود) | عدد الأسر في القطاع $(M_i)$ | $Y_i$ |
|---|---|---|
| 1 | 60 | 2،3،5،1،5،2 |
| 2 | 40 | 3،4،2،3 |
| 3 | 80 | 2،5،1،5،2،2 |

قدر متوسط دخل الأسرة السنوي في المدينة بحدود ثقة 95% إذا علمت أن إجمالي عدد الأسر في المدينة هو (2400) أسرة.

**الإجابة:**

$$N = 40$$
$$n = 3$$
$$m = 4$$
$$M = 2400$$
$$\bar{Y}_{cs2} = \frac{N}{M_n}\sum_{i=1}^{n} M_i\,\bar{y}_i = \frac{48}{2400(3)}(135 + 120 + 160) = 2.7667$$
$$S^2(\bar{Y}_{cs2}) = \frac{N-n}{N}\frac{1}{n\bar{m}^2}S_b^2 + \frac{1}{n_N\bar{m}^2}\sum M_i^2\,(\frac{M_i - m_i}{M_i})(\frac{S_i^2}{m_i})$$
$$S_b^2 = \frac{1}{n-1}\sum_{i=1}^{n}(M_i\,\bar{y}_i - \bar{M}\bar{Y}_{cs2})^2 \qquad \bar{\mathrm{M}} = \frac{2400}{48} = 50$$
$$S_b^2 = \frac{1}{2}(19.8 + 3085.8 + 2612.23) = 816.66412 = 2858.9$$
$$S_1^2 = 0.4167 \qquad S_2^2 = 0.667 \qquad S_3^2 = 0.167$$
$$S^2(\vec{Y}_{cs2}) = \frac{48-3}{48}(\frac{1}{3(50)^2})2858.9 + (\frac{1}{3(48)(50)^2}) \times (350.028 + 240.12 + 255.84)$$
$$= 0.051 + 0.0023 = 0.0533$$

## تمريـــن

في دراسة لتقدير معدل إنتاج الدونم من محصول القمح تم سحب 8 قرى من اصل 50 قرية موجودة في المحافظة وسحب من كل قرية 3 حقول وأجري لكل قرية في كل حقل تجربة حصاد واحدة فكانت النتائج كما يلي:

| القرية | الحقل في القرية | $\hat{Y}_i$ | معدل إنتاج الحقل كغم/دونم |
|---|---|---|---|
| 1 | 10 | 131.3 | 214 |
| | | | 110 |
| | | | 70 |
| 2 | 12 | 25.5 | 28 |
| | | | 30.8 |
| | | | 17.6 |
| 3 | 8 | 48.4 | 57 |
| | | | 39.8 |
| | | | 48.4 |
| 4 | 10 | 10.9 | 15 |
| | | | 10.6 |
| | | | 7 |
| 5 | 6 | 54.3 | 55.5 |
| | | | 53 |
| | | | 54.3 |
| 6 | 4 | 20.5 | 23.4 |
| | | | 34.0 |
| | | | 4.2 |
| 7 | 6 | 20.0 | 17.9 |
| | | | 14.6 |
| | | | 27.6 |
| 8 | 12 | 29.3 | 28 |
| | | | 52 |
| | | | 34 |

## الفصل التاسع

# تطبيقات أسلوب المعاينة

# أولا: مسوح القوى العاملة

## دراسة حالة

**القوى العاملة في الأردن 2001**

**أولاً: أهداف الدراسة**

تهدف الدراسة إلى اتباع مدخل جديد، يتصف بالمرونة، ويمكن من خلاله احتساب معدل البطالة وفقاً لمنهجيات عديدة تتفق ومعايير منظمة العمل الدولية، كما تهدف أيضاً إلى بلورة منهجية ذات خصوصية وطنية.

وهناك أهداف بحثية أخرى تسعى الدراسة إلى تحقيقها في الحقول التالية:

أ. حقل الأفراد العاملين
ب. حقل الأفراد المتعطلين
ج. حقل الأفراد غير الناشطين اقتصادياً

**ثانياً: مستويات النشر والدقة المطلوبة**

سيتم نشر نتائج هذه الدراسة على مستوى المحافظة، والإقليم، والريف والحضر، وعلى مستوى المملكة،كما يمكن أن توفر هذه الدراسة بيانات موثوق بها على مستويات أقل مثل الحضر على مستوى إقليم.

## تقدير حجم العينة وتوزيعها على الطبقات

نظراً لعدم وجود دراسات تحتوي على بيانات تفصيلية عن تباين العينة للخصائص الرئيسية التي تشتمل عليها الدراسة سوى متغير البطالة؛ فقد تم استخدام البيانات المتوافرة في عينة البطالة التي نفذتها دائرة الإحصاءات العامة عام 1998.

### 1. تقدير الخطأ المعياري من دراسات سابقة

تم الاطلاع على نتائج الدراسات السابقة في هذا المجال، ومنها الدراسة التي نفذتها دائرة الإحصاءات العامة عام 1998، وشملت 744 وحدة معاينة أولية* ( عنقود) في المرحلة الأولى، وقد سحبت عينة حجمها 20 أسرة من كل عنقود في المرحلة الثانية لتلك العينة. ويحتوي الجدول (1) على النتائج الرئيسية للجولة الأولى من تلك الدراسة التي استخدمت لتقدير حجم العينة وتوزيعها في هذه الدراسة.

**جدول (1): نسبة البطالة المقدرة، والخطأ المعياري، ومعامل الاختلاف، وتأثير التصميم للجولة الأولى من مسح العمالة والبطالة الذي نفذته دائرة الإحصاءات العامة لعام 1998**

| المحافظة / الإقليم | نسبة البطالة المقدرة | الخطأ المعياري S.E | معامل الاختلاف % C.V | تأثير التصميم Des.Eff. |
|---|---|---|---|---|
| 11 | 0.130 | 0.006 | 4.615 | 2.46 |
| 12 | 0.140 | 0.011 | 7.857 | 1.53 |
| 13 | 0.157 | 0.009 | 5.732 | 2.20 |
| 14 | 0.164 | 0.014 | 8.537 | 0.77 |
| 21 | 0.144 | 0.009 | 6.250 | 2.78 |
| 22 | 0.135 | 0.013 | 9.629 | 1.13 |
| 23 | 0.129 | 0.015 | 11.627 | 1.05 |
| 24 | 0.152 | 0.013 | 8.552 | 0.59 |
| 31 | 0.150 | 0.013 | 8.666 | 1.28 |
| 32 | 0.134 | 0.021 | 15.672 | 1.13 |
| 33 | 0.135 | 0.019 | 14.074 | 1.14 |
| 34 | 0.094 | 0.010 | 10.638 | 0.45 |
| إقليم الوسط | 0.138 | 0.004 | 2.899 | 2.25 |
| إقليم الشمال | 0.142 | 0.007 | 4.930 | 2.21 |
| إقليم الجنوب | 0.133 | 0.008 | 6.015 | 1.10 |
| الحضر | 0.141 | 0.004 | 2.837 | 2.18 |
| الريف | 0.130 | 0.007 | 5.385 | 1.94 |
| إجمالي الدولة | 0.139 | 0.003 | 2.158 | 2.14 |

ويلاحظ أن معامل الاختلاف مرتفع نسبيا في بعض المحافظات على الرغم من أن حجم العينة كبير نسبياً، إذ بلغ 14880 أسرة. ويعزى السبب في ذلك إلى كون العينة لا تهدف إلى توفير نتائج موثوق بها على مستوى المحافظة، والى تركيز التصميم على مستوى الإقليم، والريف والحضر، ومستوى المملكة بمجمله، وإلى كبر عدد الأسر المسحوبة من كل عنقود نسبياً، إذ بلغ 20 أسرة مما أدى إلى زيادة تأثير التصميم، وتقليل فاعلية العينة.

وتختلف الدراسة عن ذلك في أنها توفر المعلومات الموثوق بها على مستوى المحافظة، ليتسنى للجهات الرسمية والباحثين استخدام نتائجها، نظراً لأهميتها في التنمية، ومساهمتها في إيجاد الحلول الملائمة لمشكلتي الفقر والبطالة كما وردت في كتاب التكليف السامي. وعليه تم الاهتمام بمستوى الثقة بحيث لا يزيد معامل الاختلاف للمتغير الرئيسي على مستوى المحافظة على 10%.

### 2. حساب تأثير التصميم

اجريت دراسة لعدة بدائل تم فيها حساب تأثير التصميم عند مستويات مختلفة من عدد وحدات المعاينة للعينة الثانوية المسحوبة من كل عنقود؛ وذلك بهدف الوصول إلى عدد الأسر الذي يمكن سحبه من كل عنقود لتكون العينة أكثر فاعلية. واستخدمت المعادلة التالية لحساب تأثير التصميم في حالة استخدام عدد مختلف عن 20 أسرة في كل عنقود:

$$DEFFn=1+(n-1)*(DEFF20-1)(20-1)$$

حيث:

**DEFFn:** تأثير التصميم عند **n** من العينات (عدد الأسر) المسحوبة في العنقود (البلوك).
**n:** عدد الأسر المسحوبة في العنقود (البلوك).
**DEFF20:** تأثير التصميم عندما يكون عدد الأسر المسحوبة في العينة مساوياً 20 أسرة من كل عنقود (البلوك).

ويدل الرقم 1 في هذه المعادلة على تأثير الطبقية، وهو عادة أقل من 1، واعتباره 1 يؤدي إلى زيادة في تأثير التصميم مما يوسع من حجم العينة المقدرة ، الأمر الذي من شأنه جعل النتائج أكثر دقة، وتخفيض معامل الاختلاف عن 10% وهو الوضع الأحوط.

**جدول (2): تأثير التصميم في حالة سحب 10،12،15،20 أسر من كل عنقود**

| المحافظة / الإقليم | تأثير التصميم في حالة سحب 20 أسرة | تأثير التصميم في حالة سحب 15 أسرة | تأثير التصميم في حالة سحب 12 أسرة | تأثير التصميم في حالة سحب 10 أسر |
|---|---|---|---|---|
| 11 | 2.46 | 2.23 | 2.00 | 1.84 |
| 12 | 1.53 | 1.45 | 1.36 | 1.31 |
| 13 | 2.20 | 2.01 | 1.82 | 1.69 |
| 14 | 0.77 | 0.80 | 0.84 | 0.87 |
| 21 | 2.78 | 2.50 | 2.22 | 2.03 |
| 22 | 1.13 | 1.11 | 1.09 | 1.08 |
| 23 | 1.05 | 1.04 | 1.03 | 1.03 |
| 24 | 0.59 | 0.65 | 0.72 | 0.76 |
| 31 | 1.28 | 1.24 | 1.19 | 1.16 |
| 32 | 1.13 | 1.11 | 1.09 | 1.08 |
| 33 | 1.14 | 1.12 | 1.10 | 1.08 |
| 34 | 0.45 | 0.54 | 0.62 | 0.68 |
| إقليم الوسط | 2.25 | 2.05 | 1.86 | 1.72 |
| إقليم الشمال | 2.21 | 2.02 | 1.83 | 1.70 |
| إقليم الجنوب | 1.10 | 1.08 | 1.07 | 1.06 |
| الحضر | 2.18 | 1.99 | 1.81 | 1.68 |
| الريف | 1.94 | 1.79 | 1.64 | 1.54 |
| إجمالي الدولة | 2.14 | 1.96 | 1.78 | 1.66 |

يلاحظ أن تأثير التصميم يقل كلما انخفض عدد الأسر المسحوبة من كل عنقود، وهذا يرجع إلى اقتراب العينة من العشوائية البسيطة، ولكن عدد العناقيد التي يمكن سحبها عند مستوى دقة معين يزداد، مما يرفع التكاليف. كما قد لا يزيد سحب عدد كبير من الأسر من كل عنقود من دقة البيانات ضمن حدود معينة.

وعليه لا بد من الموازنة بين الكلفة والدقة لتحديد الحجم الأمثل للعينة، ومراعاة عدد وحدات المعاينة الرئيسية والثانوية، بحيث لا يقل عن حجم معين لكل مستوى نشر. وقد بلغ في هذه الدراسة 24 عنقوداً تشتمل على 360 أسرة، وتحوي كل أسرة على عدد من الأفراد، وبذلك يتم التأكيد بأن أعداد المشاهدات المتوقعة لأصغر مستوى نشر هي أعداد كافية يتسنى من خلالها نشر ودراسة المتغيرات الرئيسية في هذه الدراسة. أما فيما يتعلق بالمتغيرات المتوقع بأن يكون عدد مشاهداتها قليلاً؛ فيمكن أن تدرس على مستوى إقليم. في حين يمكن دراسة المفردات التي لها خصائص الندرة في المجتمع على مستوى المملكة. وسيتم حساب معامل التغير للبيانات ونشرها ضمن نتائج هذه الدراسة لمعرفة مستوى الدقة لكل متغير يستخدمه الباحثون أو يبنى عليه قرار.

### 3. تقدير حجم العينة

استخدم لتقدير حجم العينة نتيجة حاصل ضرب عدد وحدات المعاينة المقدر عن طريق العينة العشوائية البسيطة بتأثير التصميم، بحيث يكون معامل الاختلاف (C.V) المتوقع على مستوى المحافظة 10% وكان تقدير هذا الحجم بإستخدام عدة بدائل لعدد وحدات المعاينة الثانوية (عدد الأسر من كل عنقود) كما هو مبين في جدول 3.

**جدول (3): عدد العناقيد وعدد الأسر التي يمكن سحبها بحيث يكون معامل التغير (C.V) مساوياً 10% وعدد الأسر المسحوب من كل عنقود مساوياً 10،12،15،20 أسر**

| المحافظة | 20 أسرة | | 15 أسرة | | 12 أسرة | | 10 أسر | |
|---|---|---|---|---|---|---|---|---|
| | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد |
| 11 | 1640 | 82 | 1500 | 100 | 1344 | 112 | 1230 | 123 |
| 12 | 940 | 47 | 885 | 59 | 840 | 70 | 800 | 80 |
| 13 | 1180 | 59 | 1080 | 72 | 984 | 82 | 910 | 91 |
| 14 | 400 | 20 | 405 | 27 | 432 | 36 | 440 | 44 |

| المحافظة | 20 أسرة | | 15 أسرة | | 12 أسرة | | 10 أسر | |
|---|---|---|---|---|---|---|---|---|
| | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد | عدد الأسر | عدد العناقيد |
| 21 | 1660 | 83 | 1485 | 99 | 1320 | 110 | 1210 | 121 |
| 22 | 720 | 36 | 705 | 47 | 696 | 58 | 690 | 69 |
| 23 | 700 | 35 | 705 | 47 | 696 | 58 | 700 | 70 |
| 24 | 320 | 16 | 360 | 24 | 408 | 34 | 430 | 43 |
| 31 | 720 | 36 | 705 | 47 | 672 | 56 | 660 | 66 |
| 32 | 720 | 36 | 720 | 48 | 708 | 59 | 700 | 70 |
| 33 | 720 | 36 | 720 | 48 | 708 | 59 | 690 | 69 |
| 34 | 440 | 22 | 525 | 35 | 600 | 50 | 660 | 66 |
| **المجموع** | **10160** | **508** | **9795** | **653** | **9408** | **784** | **9120** | **912** |

ويلاحظ انخفاض حجم العينة الكلي كلما قل عدد الأسر المسحوبة من كل عنقود؛ لأن العينة تقترب من العشوائية، ولكن هذا النقص تقابله زيادة في عدد العناقيد مما يعني ارتفاع تكاليف تحديث الإطار، وتكاليف الأعمال الميدانية، مثل كلفة نقل الباحثين من عنقود إلى آخر. ومن ميزات عينة الأسر قليلة العدد في العنقود، أنها أكثر انتشاراً في المجتمع، وأخطاء غير المعاينة فيها الناتجة عن تركز العينة في عدد قليل من العناقيد، وتزيد من مرونة العينة. أما المحافظات الصغيرة، فإنها بخلاف ذلك، إذ يصعب فيها سحب عدد كبير من العناقيد، نظراً لمحدودية هذه العناقيد في المجتمع. وتجدر الإشارة إلى أن عدد الأسر المناسب سحبه من كل عنقود للدراسات الأسرية التي تشتمل على بعض الجوانب الاقتصادية، في ضوء ما حددته التجارب الدولية، يتراوح بين 15-10 أسرة من كل عنقود، مع وجود بعض المرونة التي قد يتطلبها التصميم حسب خصائص الإطار المتوافر، والصفة المراد دراستها.

ولدراسة الوضع، بشكل أفضل، يمكن النظر إلى كسر المعاينة لمعرفة مدى تمثيل العينة للمجتمع، وفقاً لمختلف البدائل، واختيار البديل الأنسب.

| المحافظة | عدد العناقيد في الإطار | 20 أسرة من العنقود | | 15 أسرة من العنقود | | 12 أسرة من العنقود | | 10 أسر من العنقود | |
|---|---|---|---|---|---|---|---|---|---|
| | | عدد عناقيد العينة | كسر المعاينة | عدد عناقيد العينة | كسر المعاينة | عدد عناقيد العينة | كسر المعاينة | عدد عناقيد العينة | كسر المعاينة |
| 11 | 3307 | 82 | 0.0248 | 100 | 0.030 | 112 | 0.033 | 123 | 0.037 |
| 12 | 632 | 47 | 0.036 | 59 | 0.046 | 70 | 0.054 | 80 | 0.062 |
| 13 | 1295 | 59 | 0.046 | 72 | 0.056 | 82 | 0.063 | 91 | 0.070 |
| 14 | 268 | 20 | 0.074 | 27 | 0.101 | 36 | 0.134 | 44 | 0.164 |
| 21 | 1664 | 83 | 0.050 | 99 | 0.059 | 110 | 0.066 | 121 | 0.072 |
| 22 | 452 | 36 | 0.080 | 47 | 0.104 | 58 | 0.128 | 69 | 0.153 |
| 23 | 272 | 35 | 0.129 | 47 | 0.173 | 58 | 0.213 | 70 | 0.257 |
| 24 | 232 | 16 | 0.069 | 24 | 0.103 | 34 | 0.147 | 43 | 0.185 |
| 31 | 414 | 36 | 0.087 | 47 | 0.114 | 56 | 0.135 | 66 | 0.159 |
| 32 | 166 | 36 | 0.217 | 48 | 0.289 | 59 | 0.355 | 70 | 0.422 |
| 33 | 202 | 36 | 0.178 | 48 | 0.238 | 59 | 0.292 | 69 | 0.342 |
| 34 | 208 | 22 | 0.106 | 35 | 0.168 | 50 | 0.240 | 66 | 0.317 |
| **المجموع** | **9112** | **508** | **0.056** | **653** | **0.072** | **784** | **0.086** | **912** | **0.100** |

يلاحظ ازدياد كسر المعاينة بازدياد عدد العناقيد في العينة؛ أي أنه كلما انخفض عدد الأسر المسحوبة من كل عنقود ازداد عدد العناقيد وكسر المعاينة مما يعني انتشاراً أفضل للعينة، كما يعني ضمان وجود عدد كافٍ من الأسر في كل عنقود لسحب عينة منها. فعلى سبيل المثال، إذا كان عدد الأسر في أحد العناقيد هو 15 أسرة بينما المطلوب سحب 20 أسرة في العينة؛ فإن ذلك يسبب مشاكل قد تؤدي إلى وجود تحيز في العينة؛ نتيجة دمج عنقودين بعد عملية سحب العينة، وإلى تشابه الأسر التي تسكن في المنطقة نفسها من ناحية الخصائص الاقتصادية والاجتماعية مما يعني أن زيادة عدد وحدات المعاينة المسحوبة من العنقود عن حد معين لا تُحسّن من دقة النتائج إلا بشكل بسيط جدا؛ أي عدم فاعلية العينة من ناحية الكلفة، وهذا يزيد كذلك من تأثير التصميم.

وبناء على ما تقدم تقرر أن يكون حجم العينة في المرحلة الثانية للمحافظات عمان، والزرقاء، واربد، والبلقاء، والمفرق، والكرك هو 10 أسر من كل عنقود، إذ تمثل هذه المحافظات ما نسبته 85.2% من إجمالي عدد العناقيد في المملكة. وقد زاد عدد العناقيد في كل محافظة على 400 عنقود بينما كان عدد العناقيد في المحافظات المتبقية أقل من 300 عنقود لكل محافظة. لذلك كان عدد الأسر المعتمد للمحافظات الأخرى هو 15 أسرة من كل عنقود. وعليه يكون المجتمع قد قسم إلى محافظات صغيرة، سحب من كل عنقود من عناقيدها 15 أسرة، ومحافظات كبيرة سحب من كل عنقود من عناقيدها 10 أسر.

**جدول( 5): توزيع العينة للمرحلتين بين المحافظات وكسر المعاينة**

| المحافظة | عدد العناقيد في العينة | كسر المعاينة للمرحلة الأولى | عدد الأسر المسحوبة في العينة |
|---|---|---|---|
| 11 | 123 | 0.037 | 1230 |
| 12 | 80 | 0.062 | 800 |
| 13 | 91 | 0.070 | 910 |
| 14 | 27 | 0.101 | 405 |
| 21 | 121 | 0.072 | 1210 |
| 22 | 69 | 0.153 | 690 |
| 23 | 47 | 0.173 | 705 |
| 24 | 24 | 0.103 | 360 |
| 31 | 66 | 0.159 | 660 |
| 32 | 48 | 0.289 | 720 |
| 33 | 48 | 0.238 | 720 |
| 34 | 35 | 0.168 | 525 |
| **المجموع** | **779** | **0.085** | **8935** |

يتضح من الجدول أن حجم العينة هو 779 عنقوداً، وأنها تحتوي على 8935 أسرة. ونظراً لأن عدد وحدات المعاينة الأولية (العناقيد) ينبغي أن يقبل القسمة على رقم أربعة؛ بهدف تقسيم هذه العينة إلى أربعة مكررات، وذلك للحد من التحيز الناتج عن سحب العينة بانتظام، و لجعل العينة أكثر مرونة؛ بحيث يمكن استخدام عينة جزئية منها في حالة وجود حاجة للقيام بدراسة أكثر تفصيلاً لعينة أصغر، وربط نتائجها بهذه الدراسة. وقد أجري تعديل بسيط على توزيع العينة، إذ أصبحت كما هي في الجدول (6)

**جدول(6): عدد العناقيد والأسر المسحوبة في العينة بعد التعديل**

| رمز المحافظة | عدد العناقيد في العينة | عدد الأسر في العينة |
|---|---|---|
| 11 | 124 | 1240 |
| 12 | 80 | 800 |
| 13 | 92 | 920 |
| 14 | 28 | 420 |
| 21 | 120 | 1200 |
| 22 | 68 | 680 |
| 23 | 48 | 720 |
| 24 | 24 | 360 |
| 31 | 68 | 680 |
| 32 | 48 | 720 |
| 33 | 48 | 720 |
| 34 | 40 | 600 |
| **المجموع** | **788** | **9060** |

يلاحظ من الجدول أن حجم العينة الكلي 788 عنقوداً، وأنها تحتوي على 9060 أسرة. ويعد هذا الحجم ممثلاً للمحافظات، والأقاليم، والريف، والحضر على مستوى المملكة، ويمكن اختبار مدى تمثيل العينة بعد استخراج النتائج على مستويات أقل قبل استخدام تلك النتائج لإجراء التحاليل الإحصائية.

## تصميم العينة المستخدم لهذه الدراسة

استخدم في هذه الدراسة أسلوب المعاينة الطبقية العنقودية من مرحلتين. فقد قسمت المملكة إلى29 طبقة،وذلك اعتماداً على نتائج التعداد العام للمساكن والسكان لعام 1994؛ بهدف ضمان أعلى حد من التجانس داخل مفردات المجتمع في الطبقة الواحدة، ولزيادة فاعلية العينة، وبخاصة فيما يتعلق بالمسوح الأسرية الاقتصادية، إذ قسمت المملكة إلى ريف وحضر، وقسم الحضر إلى مدن كبرى، وعددها خمس مدن وبقية الحضر، أما الريف، فبقي كما هو (12طبقة) .

وضماناً لفاعلية العينة بشكل جيد قسمت كل مدينة من المدن الخمس إلى أربع طبقات جزئية حسب الخصائص الاجتماعية والاقتصادية؛ بناء على نتائج التعداد العام للمساكن والسكان إذ أخذت عدة متغيرات لها علاقة بهذا الموضوع، مثل : نوع المسكن، وقيمة المسكن، والمستوى التعليمي لأفراد الأسرة، وغيرها من الخصائص. ويعد هذا التقسيم من ضمن الطبقية الضمنية، كذلك جرى ترتيب التجمعات ضمن 9 فئات حسب عدد السكان في كل تجمع، وذلك لضمان انتشار العينة على مختلف أحجام التجمعات. وهذه أيضا تقع ضمن الطبقية الضمنية، وهذا يؤدي إلى زيادة دقة العينة، وتخفيض معامل الاختلاف إلى مستوى أقل مما هو متوقع، بالإضافة إلى التخفيض الناتج عن حساب تأثير التصميم، إذ اعتبر تأثير الطبقية مساوياً 1، وهو بالتأكيد أقل من ذلك. وعليه، يتوقع أن يكون معامل الاختلاف الفعلي على مستوى المحافظة أقل من 8% وعلى مستوى المملكة حوالي 3%. ويعتبر هذا الرقم جيداً، فقد تراوح معامل الاختلاف لعينة البطالة التي أجرتها دائرة الإحصاءات العامة عام 1998 بين حوالي 5% في محافظة عمان و 16% لمحافظة الطفيلة، وكان حجم العينة لتلك الدراسة هو 14880 أسرة, أي بزيادة قدرها 39.1% عن عينة هذه الدراسة، الأمر الذي يدل على جدوى هذه الدراسة من ناحية الكلفة.

## ثانياً: أسلوب سحب العناقيد من كل طبقة

بعد أن رتبت عناقيد التجمعات داخل كل طبقة تصاعدياً ورتبت عناقيد كل مدينة حسب الخصائص الاقتصادية والاجتماعية لكل عنقود؛ تم حساب عدد الأسر التجميعي في كل طبقة، وعدد الأسر الكلي لكل طبقة، بعد ذلك جرت عملية سحب عينة العناقيد لكل طبقة بإستخدام أسلوب المعاينة المتناسبة مع الحجم ( إذ أعتبر عدد الأسر في كل عنقود هو حجم العنقود) المسحوبة بانتظام، وذلك لتحقيق **هدفين**، هما:
**الهدف الأول** أن تكون العينة موزعة بشكل جيد على فئات التجمعات المختلفة. و**الهدف الثاني** إعطاء أهمية نسبية لكل عنقود تتناسب وحجمه. واستخدم برنامج أعد خصيصاً لعملية سحب العينة بهذه الطريقة.

## ثالثاً: أسلوب سحب عينة الأسر من كل عنقود

بعد سحب عينة العناقيد جرى تحديث هذه العناقيد ميدانياً، وإعداد إطار بالأسر الموجودة في كل عنقود التي ينطبق عليها الشرط في أن تكون أسراً أردنية كما هو وارد في أهداف الدراسة، وذلك لتحديد المجتمع المستهدف في كل عنقود قبل سحب عينة المرحلة الثانية منه، ومن المعروف أن من استخدامات العينة العنقودية، كما هو في أدبيات هذا النوع من العينات، استخدامها في حالة عدم توافر إطار حديث لجميع مفردات المجتمع، فتسحب عينة عناقيد ويجري تحديثها بدلا من تحديث الإطار بمجمله. وهذا أسلوب فيه توفير للنفقات، واستيفاء لشروط الإطار الواجب توافرها لسحب العينات.

بعد تحديث الإطار يتم سحب عينة من الأسر المستهدفة من كل عنقود حسب العدد الذي حدد كما ورد سابقاً، وتجرى عملية السحب بأسلوب العينة المنتظمة بعد أن تكون قد رتبت الأسر في كل عنقود حسب الموقع الجغرافي لها، وضماناً لانتشار العينة على مختلف أجزاء العنقود الذي يؤدي إلى تقليل الارتباط الداخلي **(Interclass correlation)** مما يزيد من دقة النتائج. وفي حالة وجود عدم استجابة في الميدان لعدد من الأسر؛ فلا يجوز سحب عينة بديلة ميدانياً؛ إنما يجب إجراء عدة محاولات بالعودة إلى الأسرة في أوقات مختلفة، أما في حالة تغير صفة الإشغال للمسكن أو إذا اصبح المسكن خالياً، أو مأهولاً من أسرة غير أردنية، فعندئذ تسحب عينة بديلة مكتبياً، ويزود بها الباحث الميداني. ولا يقوم أي من العاملين في الجهاز الميداني بسحب عينات بديلة، ويجري تعديل الأوزان بناءً على نسبة الاستجابة من العينة المسحوبة ممن ينطبق عليهم الشرط، وسيأتي لاحقاً شرح ذلك عند التحدث عن أسلوب التقدير.

## 5. خطوات التقدير

في هذا الباب سيتم عرض أسلوب تقدير النتائج بالاعتماد على تصميم العينة المشار إليه. ولكي تكون تقديرات العينة من مسح معين ممثلة للمجتمع، فإنه من الضروري ضرب البيانات بوزن المعاينة او معامل التمدد (الرفع)، وبالوزن الأساسي لكل عينة أسرة حيث يساوي معكوس احتمالية السحب او الاختيار ( تحسب بضرب الاحتمالات لكل مرحلة معاينة)، وبما أن جميع بيانات المسح سوف تعالج بإستخدام الحاسب الآلي، لذلك من السهل وضع الوزن لكل سجل عينة أسرة في ملفات الكمبيوتر. ونظراً لقدرة البرامج المجدولة على عمل الأوزان للبيانات تلقائياً ( أوتوماتيكياً )؛ فإنه من الواجب الاحتفاظ بمعلومات كل مرحلة سحب في ملف الحاسب الآلي؛ لأن جميع الاحتمالات والأوزان المقابلة سوف تستخدم لحساب الوزن النهائي لكل عنقود، ويعمم هذا الوزن على جميع أسر العنقود، مع الانتباه إلى أن هذه العينة غير موزونة ذاتياً على المستوى الكلي، بل هي موزونة ذاتياً على مستوى الطبقة قبل عملية التحديث مما يعني أن وزن العينة قد يختلف من عنقود إلى آخر حتى داخل الطبقة الواحدة، ويعزى ذلك إلى التغيرات في عدد الأسر في عينة العناقيد أو وجود عدم استجابة مما يستدعي إعادة تعديل الأوزان بناء على تلك التغيرات:

**أولاً: حساب احتمالية ظهور وحدة المعاينة وتشمل المرحلتين الأولى والثانية ما يلي:**

1. احتمالية ظهور العنقود **i** من الطبقة **h** يرمز له بالرمز $(p_{h_i})$ وتساوي:

$$p_{h_i} = \frac{n_h M_{h_i}}{M_h} \qquad (97)$$

حيث:

$n_h$ = عدد وحدات المعاينة الرئيسية المسحوبة من الطبقة **(h)**.

$M_h$ = عدد الأسر في الإطار للطبقة **(h)** من التعداد العام للسكان لعام 1994.

$M_{hi}$ = مجموع عدد الأسر الموجودة في العنقود **i** من الطبقة **h** كما وردت في إطار التعداد قبل التحديث.

2. احتمالية سحب الأسرة j من العنقود i من الطبقة h يرمز لها بالرمز Phij وتساوي:

$$p_{h_{ij}} = \frac{m_{h_i}}{M'_{h_i}} \quad (98)$$

حيث:

*Phij* = احتمالية ظهور الأسرة j من العنقود i في الطبقة h
*mhi* = عدد الأسر المسحوبة من العنقود i في الطبقة h
*M'hi* = عدد الأسر الموجودة في العنقود i في الطبقة h بعد التحديث

**ثانياً: حساب الأوزان**

نظراً لكون العينة غير موزونة ذاتياً على مستوى المملكة والمحافظات، كما أشير في السابق، بل هي موزونة ذاتياً على مستوى الطبقة قبل تحديث وحدات المعاينة الأولية، ولما كانت هذه الأوزان ستختلف وتصبح العينة غير موزونة نتيجة التغيرات التي قد تحدث، مثل: تغير عدد الأسر في العنقود؛ وعدم استجابة بعض الأسر المسحوبة في العينة؛ واختلاف نسبة الأسر الأردنية إلى إجمالي عدد الأسر من عنقود إلى آخر داخل الطبقة الواحدة، فإنه لا بد من حساب وزن نهائي لكل عنقود، وتوزين النتائج حسب هذا الوزن.

الوزن الأساسي للأسرة j من العنقود i في الطبقة h يساوي مقلوب احتمالية ظهور تلك الأسرة في العينة, ويرمز له بالرمز *Whij*

حيث يساوي :

$$W_{h_{ij}} = \frac{M_h M'_{h_i}}{n_h M_{h_i} m_{h_i}} \quad (99)$$

ويمكن ملاحظة أنه إذا كان $m_{hij}$ ثابتاً لكل طبقة (مثلاً 15 أسرة مسحوبة بالعينة من كل عنقود)، و $M'_{hij} = m_{hij}$ (عدد الأسر في القائمة المحدثة للعنقود المسحوب في العينة مساو لعدد الأسر في الإطار)، فإن العينة تكون موزونة ذاتياً ضمن الطبقة. وهذه الأوزان حقيقية، وسوف تختلف قليلاً بسبب الفرق بين عدد الأسر في القائمة المحدثة للعنقود المسحوب في العينة وعدد الأسر في الإطار.

ومن المهم أيضا تعديل الأوزان للأخذ بالحسبان نسبة عدم الاستجابة لكل عنقود. وبما أن الأوزان ستحسب على مستوى كل عنقود مسحوب في العينة، فإنه من المفيد تعديل هذه الأوزان على هذا المستوى، كذلك يوصى بأن تكون عدم الاستجابة في حدها الأدنى، بهدف تقليل التحيز وأخطاء غير المعاينة، ما أمكن، وذلك عن طريق العودة إلى البيت المغلق إغلاقاً مؤقتاً في أوقات مختلفة، وإقناع الرافضين بإعطاء بيانات، كما يوصى بأن تكون البدائل في حدها الأدنى، وأن تؤخذ للحالات التي تحتاج إلى أخذ بدائل. وفي حالة وجود عدد من الاستمارات المنجزة أقل من عدد الأسر المسحوبة من عنقود معين، يوصى بضرب الوزن الأساسي بمعامل التعديل، ومعادلته هي:

$$AdjWi = \frac{mhi}{m''hi}$$

حيث:

**AdjWi**= معامل التعديل للعنقود **i** في الطبقة **h**
**mhi**= عدد الأسر المسحوبة في العينة من العنقود **i** في الطبقة **h**
**m"hi**= عدد الأسر التي جمعت لها استمارات من العنقود **i** في الطبقة **h**

وبذلك يرمز للوزن النهائي لكل أسرة مسحوبة من العنقود **i** في الطبقة **h** بالرمز $W''_{hij}$ ومعادلته هي:

$$W''_{hij} = W_{hij} \times mhi/m''hi$$

**تعديل الأوزان داخل الأسرة الواحدة:**

نظراً لكون الاستمارة تستوفى من المستجيب مباشرة لكل فرد من أفراد الأسرة ممن هم في سن 15 عاماً فأكثر؛ فإنه من المتوقع أن يحدث عدم استجابة لفرد أو أكثر من أفراد الأسرة الواحدة. وعليه فإنه من الضروري إجراء تعديل الأوزان داخل الأسرة الواحدة في حالة وجود عدم استجابة لفرد أو أكثر من أفراد الأسرة؛ بحيث يؤخذ بالاعتبار الحالة العملية لأفراد الأسرة (العاملين والمتعطلين )، والحالة العملية للأفراد الذين لم يستجيبوا داخل كل أسرة.

## ثالثاً: التقديرات

### 1. تقدير المجموع لبعض المتغيرات في الدراسة

في حالة وجود متغيرات يتطلب الأمر الحصول على تقدير المجموع لها، مثل إجمالي عدد المتعطلين، فالتقدير على مستوى الطبقة مثلاً يساوي وزن كل مشاهدة مضروباً بمتغير تلك المشاهدة للصفة المدروسة. وللحصول على تقدير لمجموعة طبقات، مثل التقدير على مستوى المحافظة أو على مستوى إقليم، فإن حاصل جمع التقدير لطبقات المحافظة أو الإقليم يساوي حاصل جمع القيمة المقدرة للمتغير للطبقات التي تمثل المحافظة أو الإقليم.
وفي حالة التقدير على مستوى المملكة، فإن القيمة المقدرة للمتغير تساوي حاصل جمع القيمة المقدرة للصفة المدروسة لجميع الطبقات.

### 2. تقدير المتوسط لبعض متغيرات الدراسة

هناك بعض المتغيرات، مثل متوسط دخل الأسرة على مستوى الطبقة أو المحافظة أو الإقليم، يتم تقديرها كما يلي:
على مستوى الطبقة حيث تقدر بعد حساب متوسط كل عنقود الذي يساوي حاصل قسمة مجموع قيم الصفة المدروسة على عدد المشاهدات في العنقود. وللحصول، بعد ذلك، على المتوسط على مستوى المحافظة أو الإقليم تستخدم المعادلة التالية:

$$\overline{y}_h = \frac{\sum_{i=1}^{n_h} w_{hi} \times \overline{y}_{hi}}{N_h}$$

$\overline{y}_h$ متوسط الطبقة
$w_{hi}$ وزن العنقود **h** في الطبقة **i**
$\overline{y}_{hi}$ متوسط الصفة المدروسة للعنقود i في الطبقة **h**
$N_h$ عدد العناقيد الموجودة في الطبقة **h**

### 3 - الوزن النسبي

نظراً لكون العينة غير موزونة ذاتياً، وفي ضوء وجود صعوبة في الحصول على عدد السكان المقدر لعام 2001 حسب التوزيعات التي يتطلبها المسح؛ فإنه من الصعب التوصل إلى تقديرات على المستوى الكلي، ويلجأ في مثل هذه الحالة إلى استخدام أحد الأساليب الذي يعامل العينة كأنها موزونة ذاتياً مع الاستفادة من ميزة العينة غير الموزونة. لذلك يجري حساب الوزن النسبي لكل مشاهدة بحيث يكون حاصل جمع الأوزان النسبية كافة للمشاهدات في المسح مساوياً لعدد المشاهدات الفعلي الذي يعطي الباحثين مرونة في التحليل والتعامل مع النتائج، سواء بإجراء حساب متوسطات أو نسب أو مجاميع على مستويات أعلى من مستويات النشر للنتائج المنشورة، فمثلا يمكن استخراج نتائج على مستوى الإقليم من خلال تجميع نتائج المحافظات، ما لم تكن منشورة على مستوى الإقليم، وذلك على اعتبار أن العينة موزونة ذاتيـاً.

ويساوي الوزن النسبي لكل أسرة من أسر عنقود معين حاصل قسمة الوزن المعدل للعنقود مقسوماً على حاصل ضرب متوسط الوزن المعدل مضروباً في عدد الاستمارات المكتملة للعنقود.

حيث يساوي متوسط الوزن المعدل حاصل جمع الأوزان المعدلة لجميع العناقيد المسحوبة في العينة مقسوماً على عدد الاستمارات الكلية المنجزة.

# توثيق عينة مسح دخل وإنفاق الأسرة 2014 - 2015

## 1. أهداف الدراسة

يمكن تلخيص الأهداف العامة كما يلي:

- تحديد سلة السلع والخدمات والحصول على الأوزان الترجيحية وغيرها من البيانات، لجمع مؤشرات اسعار المستهلك ومؤشر تكاليف المعيشة ومؤشرات الغلاء.
- دراسة مستويات الأجور وأثرها على أوجـه الإنفاق الاستهلاكي العائلي (كالسكن، والتعليم، والصحة).
- حساب المتوسطات السنوية لإنفاق ودخل الأسر والأفراد، وربط هذه المتوسطات بمختلف المتغيرات الاجتماعية والديموغرافية والجغرافية، كالحالة العملية والحالة التعليمية والنوع الاجتماعي والإقليم وما إلى ذلك.
- توفير البيانات الخاصة بإعداد الميزانية الغذائية المتعلقة باستهلاك الفرد من السعرات الغذائية.
- التعرف على توزيع السكان حسب مستويات الدخل والإنفاق المختلفة، واستخدام بيانات الإنفاق في تحديد انماط الإنفاق لدى السكان ، وتوجهات الإنفاق الاستهلاكي لديهم بحسب الخصائص المختلفة ، كالجنسية والمستوى التعليمي والعملي وما إلى ذلك.
- حساب مرونة الإنفاق (الطلب) على بنود الإنفاق المختلفة.

## 2. مجتمع الدراسة

يغطي هذا المسح الأسر الخاصة (الإماراتية وغير الاماراتية) في مختلف الإمارات إضافة إلى تجمعات العمال، وتم استثناء فئات (الهيئات الدبلوماسية والمستشفيات والفنادق ومساكن الطلبة والزوار غير المقيمين مع الأسرة إقامة دائمة وأفراد الأسرة المسجونين خلال فترة الإسناد).

## 3. إطار المعاينة المستخدم

تم إعداد إطار لمناطق العد في اربع إمارات من الامارات الشمالية وتم تحديث عينة من مناطق العد في كل امارة من تلك الامارات وهي عجمان، أم القيوين، رأس الخيمة، الفجيرة وذلك بإستخدام العينة الأساسية والتي تم تحديثها في ديسمبر2013- في تلك الإمارات، ويوفر الاطار بيانات تفصيلية عن كل اسرة مقيمة في كل منطقة من مناطق العد التي تم تحديثها اضافة الى البيانات الجغرافية التي

سوف تمكن الباحث من الوصول الى الأسرة بكل سهولة ويسر.
وفيما يتعلق بإمارة الشارقة فقد تم اعتماد اطار المساكن الذي تم اعداده خلال تعداد المنشآت الاقتصادية لإمارة الشارقة الذي نفذه المركز الوطني عام 2013، وقد تم تقسيم الامارة الى مناطق عد وجرى سحب عينة من مناطق العد وقد تم اعداد اطار شامل بالأسر الموجودة في مناطق العد المسحوبة بالعينة والبالغ عددها 152 منطقة عد.

### 4. الطبقية في العينة

بهدف الحصول على نتائج موثوقة لكل من الإماراتيين وغير الإماراتيين على مستوى كل إمارة، فقد تم تقسيم كل إمارة إلى طبقتين الأولى طبقة الإماراتيين والثانية طبقة غير الإماراتيين وقد اعتبرت كل منطقة عد يقطنها 50% فأكثر من الأسر الإماراتية طبقة إماراتيين، وكل منطقة عد يقطنها أقل من 50% من الأسر الإماراتية طبقة غير الإماراتيين.

### 5. مستويات النشر

صممت العينة بحيث تسمح نشر نتائج المسح على مستوى الإماراتيين وغير الإماراتيين لكل إمارة في الدولة بمستوى دقة 90%.

### 6. تصميم العينة

استخدم أسلوب المعاينة الطبقية العنقودية متعددة المراحل، حيث قسمت كل امارة الى طبقتين وتم سحب عينة عناقيد من كل طبقة في المرحلة الاولى، ثم سحبت عينة اسر من كل عنقود في المرحلة الثانية.

وقد جرى تقسيم العينة إلى أربع مكررات متماثلة من حيث الحجم والانتشار على مختلف المناطق الجغرافية بحيث يتم تغطية كل مكررة في ربع من أرباع السنة، لذا فانه يتم تقسيم عينة الأسر الموجودة في كل عنقود إلى ثلاثة أجزاء كل جزء يمثل شهر بواقع أربع أسر تغطى شهريا في كل منطقة عد.

## 7. تقدير حجم العينة

تم حساب معامل الاختلاف لبنود الانفاق الرئيسية من مسح دخل وانفاق الأسرة لعام 2007-2008، وقد تم استخدام التباين لكل طبقة من ذلك المسح وتم حساب العينة لكل طبقة بشكل مستقل بحيث يكون مستوى الدقة على مستوى الطبقة الواحدة 90%.

## 8. تخصيص العينة بين الطبقات

نظرا لكون المسح سوف ينفذ لمدة عام فقد تقرر أن يتم تغطية أسرة واحدة في كل منطقة عد شهريا وذلك لضمان انتشار العينة جغرافياً على مدار العام، وقد تم تخصيص (توزيع) العينة بين مختلف الطبقات لكل نوع من أنواع الأسر بطريقة توزيع نيمان مع الأخذ بالاعتبار وجود حد أدنى لكل طبقة لضمان الحصول على الدقة المطلوبة وبشرط أن يقبل حجم العينة القسمة على رقم 12 لكل طبقة لأن العينة تنفذ على مدى 12 شهراً.

## 9. أسلوب سحب العينة

**أولا: سحب عينة وحدات المعاينة الأولية (مناطق العد)**

تم سحب عينة مناطق العد (العناقيد) في المرحلة الاولى بإستخدام اسلوب العينة المتناسبة مع الحجم والمسحوبة بانتظام حيث اعتبر عدد الأسر الخاصة الموجودة في كل منطقة عد وزن تلك المنطقة، وقد تم سحب العينة بعد ترتيب مناطق العد وفقاً لموقعها الجغرافي وذلك لضمان انتشار العينة على مختلف مناطق العد بطريقة جيدة.

**ثانيا: سحب عينة وحدات المعاينة الثانوية (الأسر)**

جرى سحب عينة أسر من كل منطقة عد بأسلوب المعاينة المنتظمة وذلك لضمان انتشار العينة على مختلف مناطق العد الامر الذي يقلل من الارتباط الداخلي ومن تأثير التصميم ويزيد من مستوى دقة العينة، وقد تم سحب 12 أسرة من كل منطقة عد (عنقود).

**ثالثا : سحب عينة بدائل**

نظرا لوجود عدم استجابة متوقع نتيجة تغير صفة اشغال المسكن أو نتيجة كون المسكن أصبح مغلق وقت الزيارة فقد تم أخذ بدائل للعينات حسب الحالة وقد جرى سحب عينة البدائل بشكل مستقل عن العينة المسحوبة. وينبغي التنويه الى انه تم أخذ بدائل حالات محددة مثل حالة كون المسكن غير مأهول بأسرة أو إذا تغيرت صفة اشغال المسكن وذلك بتبديل العينة بعينة أخرى من نفس منطقة العد ومجموعة الجنسية، وتجدر الاشارة إلى أنه في حالة الرفض لم يتم أخذ بديل عن االأسرة التي رفضت الإدلاء بأية بيانات.

وفي حالة كون الأسرة مسافرة (مغلقة مؤقت) أو كان البيت مغلق في الشهر المخصص لزيارة الأسرة فقد تم تأجيل العينة المطلوبة إلى الشهر التالي واختيار اسرة من عينة الشهر التالي ليتم تغطيتها بحيث تكون ضمن من نفس منطقة العد ومجموعة الجنسية.

## 10. أسلوب التقدير

**حساب الأوزان**

**أولاً حساب احتمالية ظهور وحدة المعاينة وتشمل المرحلة الأولى والثانية كما يلي:**

1 - احتمالية ظهور العنقود **i** من الطبقة **h** يرمز له بالرمز **(Phi)** وتساوي:

$$Phi = (nh \times Mhi) / Mh$$

حيث:

**nh=** عدد وحدات المعاينة الرئيسية المسحوبة من الطبقة **(h)**.
**Mh=** عدد الأسر في الإطار للطبقة **(h)** من التعداد العام للسكان لعام 2004.
**Mhi=** مجموع عدد الأسر الموجودة في العنقود **i** من الطبقة **h** كما وردت في إطار التعداد قبل التحديث.

2. احتمالية سحب الأسرة j من العنقود i من الطبقة h يرمز له بالرمز Phij ويساوي

$$Phi = mhi \ / \ M'hi$$

حيث:

Phij= احتمالية ظهور الأسرة j من العنقود i في الطبقة h
mhi= عدد الأسر المسحوبة من العنقود i في الطبقة h
M'hi= عدد الأسر الموجودة في العنقود i في الطبقة h بعد التحديث

**ثانياً : حساب الأوزان**

نظراً لكون العينة غير موزونة ذاتياً، إلا على مستوى الطبقة حيث انها موزونة ذاتياً قبل تحديث وحدات المعاينة الأولية، وبعد عملية التحديث تصبح العينة غير موزونة نتيجة التغيرات التي قد تحدث، مثل: تغير عدد الأسر في العنقود؛ وعدم استجابة بعض الأسر المسحوبة في العينة، واختلاف نسبة الأسر الخاصة إلى إجمالي عدد الأسر من عنقود إلى آخر داخل الطبقة الواحدة، لذا فمن الواجب حساب وزن نهائي لكل عنقود، وتوزين النتائج حسب هذا الوزن .

الوزن الأساسي للأسرة ز من العنقود i في الطبقة h يساوي مقلوب احتمالية ظهور تلك الأسرة في العينة ويرمز له بالرمز Whij حيث يساوي:

$$Whij = ( Mh \ X \ Mhi ) / ( nh \ x \ Mhi \ x \ mhi )$$

ويمكن ملاحظة أنه إذا كان mhij ثابتاً لكل طبقة (مثلاً 8 أسر مسحوبة بالعينة من كل عنقود)، و M'hij = Mhij (عدد الأسر في القائمة المحدثة للعنقود المسحوب في العينة مساو لعدد الأسر في الإطار)، فإن العينة تكون موزونة ذاتياً ضمن الطبقة الواحدة، وسوف تختلف هذه الاوزان قليلاً نتيجة الفرق بين عدد الأسر في القائمة المحدثة للعنقود المسحوب في العينة وعدد الأسر في الإطار.

ومن المهم أيضا تعديل الأوزان للأخذ بالحسبان نسبة عدم الاستجابة لكل عنقود. وبما أن الأوزان ستحسب على مستوى كل عنقود مسحوب في العينة، فإنه من المفيد تعديل هذه الأوزان على هذا المستوى، كذلك يوصى بأن تكون عدم الاستجابة في حدها الأدنى، بهدف تقليل التحيز وأخطاء غير المعاينة "ما أمكن" وذلك عن طريق العودة إلى البيت المغلق إغلاقاً مؤقتاً في أوقات مختلفة، وإقناع الرافضين بإعطاء بيانات، كما يوصى بأن تكون البدائل في حدها الأدنى وأن تؤخذ للحالات التي تحتاج إلى أخذ بدائل. وفي حالة وجود عدد من الاستمارات المنجزة أقل من عدد الأسر المسحوبة من عنقود معين، يوصى بضرب الوزن الأساسي بمعامل التعديل، ومعادلته هي:

$$AdjWi = \frac{mhi}{m"hi}$$

حيث:

**AdjWi=** معامل التعديل للعنقود **i** في الطبقة **h**
**mhi=** عدد الأسر المسحوبة في العينة من العنقود **i** في الطبقة **h**
**m"hi=** عدد الأسر التي جمعت لها استمارات من العنقود **i** في الطبقة **h**

وبذلك يرمز للوزن النهائي لكل أسرة مسحوبة من العنقود **i** في الطبقة **h** يرمز له بالرمز **W''hij** ومعادلته هي:

**W''hij = Whij X mhi/m"hi**

# أسلوب التعويض لمسح دخل وإنفاق الأسرة

**أولاً: قبل اجراء عملية الوزن والتعويض تم اجراء الخطوات التالية**

- التدقيق على بيانات الاستمارات حيث تم اكتشاف عدة أخطاء مثل خطأ برقم منطقة العد، جنسية رب الأسرة، نتيجة الزيارة، مناطق عمل بها 13 أسرة، قيم شاذة وغير منطقية بالإنفاق على السلع الغذائية، وقد تم ارسال كشف بجميع هذه الأخطاء الى فريق التدقيق الفني لتعديلها على برنامج الادخال.
- تم استثناء استمارات منطقة عد (****) والتي استوفي فيها 4 استمارات فقط واستبعدت من جميع الملفات.
- وجدت حالات كانت حالة الزيارة استوفيت كليا ومع ذلك فيها أسابيع بدون إنفاق وقد استخرج كشف بجميع هذه الحالات وارسلت الى فريق تدقيق وتنظيف البيانات لتعديلها.

1. تم تقسيم الاستمارات الى خمس مجموعات بهدف اجراء عملية التعويض وحساب الاوزان كما يلي:

- الاستمارة الخاصة بالمبنى والمسكن
- الاستمارة الخاصة بالخصائص الديموغرافية للأسرة
- الاستمارة الخاصة بدخل الأسرة
- الانفاق على السلع الغذائية
- النفقات الاخرى

**ثانياً: التعويض عن القيم المفقودة**

2. جرى اعداد جداول مقارنة لعدد الأسر المطلوبة في كل منطقة عد حسب مجموعة الجنسية (-1 مواطن، -2 عربي، -3 جنسيات أخرى) ولك استمارة من استمارات المسح (انظر البند الأول) مقارنة مع عدد الأسر التي استوفيت فعليا، وذلك بهدف تحديد عدم الاستجابة الكلي على مستوى كل استمارة ولكل منطقة عد حسب الجنسية.

3. وحيث انه يتم جمع بيانات الانفاق على السلع الغذائية في أربع زيارات خلال الشهر المخصص لتلك الأسرة، فقد تم تحديد الاستمارات التي لم يحضر لها الأربع زيارات فاذا كان عدد الزيارات 3 زيارات تعتبر عدم استجابة جزئي ويجري التعويض عن الأسبوع الذي لم يتم تغطيته، وإذا كان عدد الزيارات 2 او اقل فانه يتم استبعاد الاستمارة وتعتبر عدم استجابة كلية لهذه الاستمارة ويجري التعويض عنها ضمن التعويض عن الاستجابة الكلي.
4. تم التعويض عن عدم الاستجابة الكلي عن طريق تعديل الوزن للأسر التي استوفيت من نفس مجموعة الجنسية ونفس منطقة العد وفقا للأسلوب التالي:

- تم حساب معامل التعديل للاستمارات التي لم تستوفى بقسمة عدد المطلوبة حسب الربع ومجموعة الجنسية على عدد الاستمارات المستوفاة فعلا لكل منطقة عد ومجموعة الجنسية ولك استمارة من الاستمارات التي سبق ذكرها في البند رقم 1 باستثناء استمارة الانفاق على السلع الغذائية.
- تم ضرب معامل التعديل الذي تم حسابه بالوزن الأساسي المحسوب لك جنسية داخل منطقة العد التي جرى لها عملية تعويض، وبذلك أصبح لدينا ملف يحتوي الاوزان المعدلة لك منطقة عد لجميع الاستمارات التي يوجد فيها عدم استجابة كلية باستثناء استمارة الانفاق على السلع الغذائية.

## أسلوب التعويض عن عدم الاستجابة الكلي للإنفاق على السلع الغذائية:

فيما يتعلق بعملية التعويض عن عدم الاستجابة الكلي للإنفاق على السلع الغذائية فقد جرى اختبار وجود فروق معنوية بين متوسطات الانفاق الشهري بين الأشهر لنفس مجموعة الجنسية ونفس الربع من السنة، وفي حال وجود فروق معنوية بين تلك المتوسطات فقد تمت عملية التعويض بنفس الشهر ولنفس فئة الجنسية ولنفس منطقة العد بإستخدام أسلوب التعويض الذي سبق ذكره عن طريق تعديل الوزن، اما في حالة عدم وجود فروق معنوية فقد تم التعويض من نفس الربع بغض النظر عن الشهر.

تم عمل تحليل التباين (ANOVA) بغرض اختبار الفروق بين متوسطات الانفاق الشهري على السلع الغذائية ولكل مجموعة جنسية على حده، وكانت النتيجة كما يلي:

| الربع الأول | | | الربع الثاني | | | الربع الثالث | | | الربع الرابع | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| شهر 1 | شهر 2 | شهر 3 | شهر 1 | شهر 2 | شهر 3 | شهر 1 | شهر 2 | شهر 3 | شهر 1 | شهر 2 | شهر 3 |
| | | | معنوي مع كلا الشهرين | معنوي مع الشهر الأول | معنوي مع الشهر الأول | | | | | | |
| | | | غير معنوي | غير معنوي | غير معنوي | | | | | | |
| | | | غير معنوي | غير معنوي | غير معنوي | | | | | | |

## عملية التعويض عن عدم الاستجابة الجزئي

فيما يتعلق بعدم الاستجابة الجزئي والخاص بالإنفاق على السلع الغذائية فقد تم عمل الخطوات التالية:

- تم استخراج ملف بجميع الاستمارات المستوفاة مع بيانات الانفاق على السلع الغذائية لكل أسبوع وبالتدقيق على بيانات الانفاق تم استبعاد الحالات التي كانت الاستجابة فيها تقل عن 75% (الأسر التي احضرت لها بيانات المصاريف لأسبوعين او أسبوع واحد)، وجرى عمل تعويض للأسر التي احضر لها ثلاثة أسابيع على اعتبار انها استجابة جزئية.

- استخدم أسلوب **(model – based)** بإستخدام اسلوب النسبة للتقدير وذلك للتعويض عن عدم الاستجابة الجزئي للإنفاق على السلع الغذائية ويتلخص هذا الأسلوب بالخطوات التالية:

أ- جرى حساب نسبة الانفاق في كل أسبوع من اجمالي إنفاق كامل الشهر حسب الشهر ومجموعة الجنسية.

ب- جرى التعويض عن الأسبوع الذي ورد فيه عدم استجابة عن طريق احتساب نسبة ما يمثله ذلك الأسبوع من اجمالي إنفاق الشهر وقد تم التعويض من نفس بياناتا الأسرة من خلال بيانات الأسابيع التي يتوفر عنها بيانات وفقا للمعادة التالية:

معامل تعديل **(adjusted factor)** لكل استمارة يساوي 100% مقسوما على نسبة ما يشكله الإنفاق للأسابيع التي أحضرت لها بيانات.

# المختصرات المستخدمة

| المختصر | المعنى |
|---|---|
| srs | العينة العشوائية البسيطة |
| st | العينة الطبقية |
| sy | العينة المنتظمة |
| pps | العينة المتناسبة مع الحجم |
| cl | العينة العنقودية |
| V | التباين |
| ANOVA | تحليل التباين |
| DEFF | تأثير التصميم |
| S.E | الخطأ المعياري |
| C.L | مستوى الثقة |

# المراجـع

## أ – المراجع العربية


- علوان، حسين. (1993) طرق المعاينة، دار الفرقان، عمان.
- رداد، خميس. (2005) معجم مصطلحات العينات، العهد العربي للتدريب والبحوث الاحصائية، عمان.


## ب – المراجع الانجليزية


- Cochran, W.G (1977), Sampling Technique,3rd Ed. John Wiley& Sons, New York.
- Singh, D. and F.S. Chanudhary(1986), Theory and Analysis of Sampling Survey Design, Wiley Fastern Limited, New Delhi.
- Hansen MH, Hurwitz WN and Madow WG (1953), Sampling Survey Methods and Theory, Vol.1, John Wiley& Sons, New York.

@FCSCUAE @FCSAUAE