الفصل الأول

المقدمة
Introduction

1.1 الفكرة الأساسية للمعاينة والتقدير

عندما نبدأ بتنفيذ المعاينة الأساسية، يكون المجتمع محتوياً على عدد معروف ومحدد N من الوحدات مثل مجموعة من الناس أو قطع من الأرض أو عدد من المنازل وغيرها. مع كل وحدة هنالك قيمة مصاحبة لمتغير نرغب في دراسته. وهذه القيم ينظر إليها على أنها ثابتة، وليست متغيراً عشوائياً. وترقم وحدات المجتمع على النحو التالي1,2,....N.

نسحب عينة من وحدات المجتمع ونقوم بمشاهدة وحدات العينة فقط. تحتوي البيانات على قيمة المتغير، ومعلومات أخرى حول الوحدة، فعلى سبيل المثال كل بئر نفط يجري حفره نقوم بتسجيل مقدار النفط في هذا البئر، بالإضافة إلي رقم البئر وموقعه وأي معلومات أخرى حوله. كذلك نقوم بحفظ معلومات حول متغيرات إضافية مثل عمق البئر، ونوع الخام وغيرها. أما في المسوحات التي تجرى على النباتات، فمن المتغيرات الإضافية وجود نباتات أخرى، وكثافة وجودها في المنطقة. أما في حالة استطلاعات الرأي فمن المتغيرات الإضافية التي غالباً ما نسجلها الجنس والعمر ومقدار الدخل وغيرها.

obeikandl.com

تسمى الطريقة التي نسحب بها وحدات العينة من المجتمع بتصميم المعاينة. مع معظم تصاميم المعاينة المعروفة، يُحدد التصميم بواسطة إعطائه لكل عينة S احتمالاً مقداره P(s) لاختيار تلك العينة. فعلى سبيل المثال في المعاينة العشوائية البسيطة وبحجم n وحدة، فإن الاحتمال P(s) يكون متساوياً لجميع العيناتs .

إن الاستنتاجات المعتادة من المعاينة هي تقدير بعض صفات (معلمات) المجتمع مثل المتوسط، أو المجموع الكلي، أو التباين، أو النسبة بعد مشاهدة العينة فقط. بالإضافة إلى ذلك في معظم حالات المعاينة والتقدير ربما نرغب بتقييم دقة التقديرات التي أوجدنها لبعض معلمات المجتمع باستخدام الثقة المصاحبة للتقدير، وفي معظم الحالات نستخدم فترة الثقة.

من وجهة النظر الأساسية للمعاينة، إذا كبرنا حجم العينة إلى أن تكون مساوية لحجم المجتمع N، أي تم مشاهدة جميع وحدات المجتمع، فإن صفات أو معلمات المجتمع ستكون معلومة تماماً، وإن الغموض في التقديرات التي أوجدناها باستخدام العينات يعود سببه إلى أن جزءاً فقط من المجتمع تم مشاهدته باستخدام العينة، بينما تبقى صفات المجتمع ثابتة، يعتمد التقدير على أي عينة من المجتمع جرى سحبها، إذا كان لكل عينة ممكنة يبقى التقدير قريباً من القيمة الحقيقية لصفة المجتمع، ويوجد غموض صغير في إستراتيجية المعاينة؛ لذا فإن هذا النوع من إستراتيجيات المعاينة مطلوب، ومن الناحية الأخرى إذا كانت قيمة التقدير تختلف بشكل كبير من عينة إلى أخرى، سيكون الغموض المصاحب لهذه الطريقة كبيراً.

يمكننا أن نحصل على تقديرات غير متحيزة إلى صفات المجتمع مثل المتوسط أو المجموع الكلي للمجتمع إذا تم اختيار تصميم المعاينة بعناية، وكذلك إذا تم اختيار طريقة التقدير الملائمة. أن يكون التقدير غير متحيز

يعني أن معدل جميع العينات التي يمكن سحبها يساوي القيمة الحقيقية لصفة المجتمع التي يجري تقديرها؛ لذا من خلال تصميم المعاينة وطريقة التقدير يمكننا أن نحصل على تقدير غير متحيز إلى كثافة الطيور في منطقة دراسية معينة بغض النظر فيما إذا كان توزيع الطيور في المجتمع بشكل متساوٍ أو إذا كانت متجمعة في مجموعات قليلة في منطقة الدراسة.

إن استخدام تصميم المعاينة الاحتمالية مثل المعاينة العشوائية البسيطة؛ من شأنه أن يوفر لنا تقديرات غير متحيزة إلى متوسط وإلى المجموع الكلي للمجتمع، كما يوفر لنا تقديرًا غير متحيز إلى التباين، الذي يستخدم للتأكد من موثوقية نتائج المسح بالعينة، كذلك يمكننا أن نحصل على تقديرات غير متحيزة من التصاميم التي تكون الاحتمالات فيها غير متساوية.

بالإضافة إلى هدف الحصول على تقدير غير متحيز من المسح بالعينة، هنالك هدف الدقة أو ما يسمي بالتباين القليل للتقديرات، كذلك نهدف للحصول على تصميم يكون سهل الاستخدام وتكاليف تنفيذه قليلة. لقد أدت هذه الأهداف إلى تطوير تصاميم معاينة كثيرة ومختلفة منها المعاينة العشوائية، ومعاينة الاحتمالات المختلفة، استخدام المعلومات الإضافية، والطبقية، والمنتظمة، والعنقودية وغيرها من تصاميم المعاينة المختلفة التي سنتناول الكثير منها في هذا الكتاب.

2.1 وحدات المعاينة

قد يسهل تحديد وحدات المعاينة ومن ثم وضع كشف بجميع أسماء الوحدات وهو ما يسمى بإطار المعاينة في كثيرٍ من المجتمعات مثل الإنسانية والمعاهد والمنازل وغيرها. بغض النظر عن الصعوبات العملية في تجهيز إطار المعاينة. تكون الوحدات في هذه المجتمعات إما شخصاً، أو منزلاً، أو مستشفى، أو شركة، أو غيرها. ويكون هناك كشف كامل لجميع

الأشخاص، أو المنازل، أو المستشفيات، أو الشركات في المجتمع الذي نرغب في دراسته ليعطينا إطار المعاينة المثالي، وغالباً ما تعتري عملية بناء إطار المعاينة صعوبات كثيرة عند القيام بتنفيذ ذلك ميدانياً، على أن استخدام دليل الهاتف إطارًا للمعاينة مثلاً في المسوحات التي تستخدم الهاتف، تكتنفه بعض المشكلات، منها أنه ليس جميع أفراد المجتمع لديهم هواتف، كذلك هنالك أشخاص أسماؤهم ليست موجودة في دليل الهاتف بناء على طلبهم؛ لذا فإن عينة عشوائية باستخدام دليل الهاتف إطارًا للمعاينة لا يعطي فرصة للعوائل التي لا يوجد لديها هواتف أو أن هواتفها غير موجودة في الدليل بأن تكون في العينة، أما في حالة المسوحات الخاصة بالشركات فقد تكون الكشوفات ناقصة، أو قديمة ولم يجر تحديثها لتشمل الشركات التي أسست بعد إعداد الكشف.

هنالك مجتمعات كثيرة غير واضح ماذا ستكون وحدة المعاينة فيها؟ في المسوحات الخاصة بالمواد الطبيعية، أو المحاصيل الزراعية في منطقة الدراسة، يمكن أن تكون وحدة المعاينة عبارة عن قطعة أرض بمساحة معينة، يتم الحصول عليها من خلال تقسيم المنطقة إلى مجموعة من القطع، على الخارطة ومن ثم نقوم بسحب عينة من القطع، ولكن إذا كان الشخص حراً في اختيار حجم القطع وأشكالها، فمثل هذا الاختيار ربما يؤدي إلى التأثير على تكاليف المسح بالعينة وعلى دقة المقدرات، أما في حالة طرق المعاينة التي تعتمد على اختيار نقطة بطريقة عشوائية في منطقة الدراسة، وتتمحور وحدة المعاينة حولها، فإنه من المحتمل أن تتقاطع وحدات العينة مع بعضها، وهذا سيؤدي إلى أن تكون عدد وحدات المجتمع غير محدودة.

أما في المجتمعات التي تتصف بعدم الاستقرار أو المراوغة مثل مجتمع الحيوانات البرية والبحرية، فإن مشكلة اكتشاف الوحدة يؤدي دوراً مهماً،

وهنالك دالة لاكتشاف الوحدات، لذا فإن دور الوحدة سيتأثر كثيراً بدالة الاكتشاف التي تكون مصاحبة للطريقة التي سيشاهد بها المجتمع، وتحديد المواضع للقيام بمشاهدة الوحدات.

قد يكون المتغير الذي نرغب في دراسته يختلف من منطقة إلى أخرى بشكل مستمر ضمن منطقة الدراسة، فعلى سبيل المثال في حالة المسوحات الخاصة بالبحث عن النفط في منطقة دراسية ما، قد يكون المتغير هو عمق أو كمية النفط في الموضع. المتغيران هنا ليس بالضرورة أن يكونا مصاحبين لمجموعة من الوحدات المحدودة في المنطقة، ولكن بالأحرى يقاسا أو يقدرا لنقطة معينة، أو المجموع لمنطقة جزئية من المجتمع بأي حجم أو شكل.

على الرغم من أن بعض طرق المعاينة التي سنتناولها في هذا الكتاب تتجاوز مسألة إطار المعاينة الذي يقسم المجتمع إلى مجموعة من الوحدات المحددة، ومن ثم نقوم بسحب العينة من هذه المجموعة من الوحدات، إلا أنَّ بعض صفات طرق المعاينة المعروفة مثل المعاينة العشوائية البسيطة، والمعاينة الطبقية مثل عدم تحيز التقدير وتقديرات النسبة وغيرها ستبقى تستعمل.

3.1 أخطاء المعاينة وأخطاء أخرى

تفترض طرق المعاينة الأساسية أن المتغير الذي نرغب بدراسته يمكن قياسه في جميع وحدات العينة وبدون أخطاء، ومن ثم فإن الخطأ في التقدير يكون سببه أن جزءاً من المجتمع فقط جرى سحبه بالعينة، وهذا الخطأ يُسمى خطأ المعاينة (Sampling errors)، ولكن في الحياة العملية وفي حالات المسوحات الحقيقية هنالك أخطاء أخرى (Nonsampling errors) تحدث ليس سببها المعاينة وإنما أسباب لا علاقة لها بالمعاينة، قد لا يرد أحد على الهاتف في حالة المسوحات التي تجرى بالهاتف وذلك لكونهم خارج المنزل، أو يرفضون الإجابة عن بعض الأسئلة الموجودة ضمن الاستبانة لسبب أو لآخر. ربما يكون

هؤلاء الرافضون غير نموذجيين في المجتمع، وهذا سيؤدي لأن تكون العينة غير ممثلة للمجتمع ومن ثم ستكون التقديرات متحيزة في المسوحات الخاصة بالسمك، ربما بعض المواضع التي جرى سحبها في العينة لم يتم مشاهدتها لكون أحوال الطقس سيئة وقت إجراء المسح، أو تكون بعيدة جداً من الشاطئ بحيث يصعب الوصول إليها باستخدام الوسائل المتاحة للمسح.

إن مشكلة عدم الإجابة من المشكلات التي تعتري كثيراً المسوحات، التي مقياسها يكون قلة عدد المستجيبين لبعض الأسئلة الموجودة في الاستبانة، أو رفض الإجابة عن جميع أسئلة الاستبانة لأسباب كثيرة منها: أن الأشخاص لا يعيرون أهمية لمثل هذه الأمور، أو لكونهم مشغولين، أو لكون بعض الأسئلة تتعلق بأمور شخصية لا يحبون الإفصاح عنها للآخرين. يمكننا أن نقلل تأثير عدم الإجابة وذلك من خلال جهود أكبر من قبل الباحثين أو العادِّين منها: زيارة الوحدة مرة أخرى أو عدة مرات، أو إعادة صياغة بعض الأسئلة المحرجة التي تتعلق بالأمور الشخصية، أو استخدام بعض المعلومات أو المتغيرات الإضافية، أو بناء نماذج رياضية خاصة بعدم الإجابة وغيرها.

هنالك أخطاء أخرى مثل أخطاء القياس أو التسجيل التي يمكن أن تحدث لسبب أو لآخر منها أن أجهزة القياس غير دقيقة، أو قديمة، أو لا يعرف العادون كيفية استخدامها بشكل فاعل، لابد لنا من محاولة السيطرة على هذا النوع من الأخطاء قدر المستطاع، هنالك حالات يمكننا أن نبني فيها نموذجاً لأخطاء القياس بصورة مستقلة عن أخطاء المعاينة، لأخطاء عدم الإجابة التي تحدث في المجتمعات البرية والمجتمعات المراوغة، في مسحات الطيور من المألوف ألاَّ يستطيع العاد أن يكتشف جميع الطيور الموجودة بجواره في إحدى وحدات المعاينة، في مسوحات الأسماك لا يمكن صيد جميع الأسماك التي تقع في طريق الشبكة، وليس جميع المشردين الذين لا يجدون

سكناً خاصاً بهم يمكن الوصول إليهم، سنتناول في هذا الكتاب مجموعة من الطرق التي تتعامل مع هذا النوع من المجتمعات منها: طريقة الخط القاطع، وطريقة خط التقاطع، وطريقة الإمساك وإعادة الإمساك وغيرها.

4.1 لمحات من تاريخ المعاينة

إن أول من استخدم المعاينة هو ملك فارس وذلك في عام 440 قبل الميلاد حسب المعلومات المتوافرة لدينا. ولو أن البابليين قد استخدموا نوعاً من المعاينة لتقدير منتجاتهم من القمح والمحاصيل الأخرى التي تُعَد المادة الغذائية الرئيسة لساكني بلاد ما بين النهرين وذلك منذ 3000 سنة من قبل الميلاد. أما الملك الفارسي فقد طلب أن يحصى جيشه الجرار الذي استخدمه لاحتلال بلاد اليونان، والطريقة التي اعتمدت لتقدير أعداد الجند تتلخص في أنه جرى تحديد منطقة معينة وطلب من مجموعة من الجند الوقوف مصطفّين بعضهم جنب بعض، وتم حصر عدد الجنود الذين يمكن أن تحتويهم هذه القطعة، ومن ثم وضع حولها سياجاً، وطلب من جميع الجند المرور من خلال هذه المنطقة، وتم مشاهدة عدد المرات التي امتلأت بها، ومن ثم تم تقدير عدد الجنود من خلال ضرب سعة المنطقة، في عدد المرات التي امتلأت بها هذه المنطقة، وقد تم تقدير عدد الجند بنحو 1,700,000 جندي، لم تذكر لنا كتب التأريخ كيف تم اختيار الجنود الذين استخدموا في المرة الأولى لمعرفة عدد الجنود الذين يستطيعون الوقوف بهذه المنطقة، وكذلك لم نعرف إذا تمت عملية التأكد من دقة التقدير أم لا.

لقد طُورت معظم طرق المعاينة التي نعرفها اليوم في القرن العشرين. كان هنالك نقاش بين مستخدمي طرق المعاينة مع بدايات القرن العشرين حول جدارة وكفاءة العينة العشوائية بالمقارنة بالعينة التي يجري اختيار وحداتها من قبل المستخدم لتكون ممثلة للمجتمع قدر الإمكان. لقد تم تطوير طريقة

المعاينة العشوائية البسيطة خلال العقدين الأولين من القرن العشرين. لقد قام Neyman (1934) بالمقارنة بين الطريقتين، ووضع المفاهيم الأساسية للمعاينة الاحتمالية التي نستطيع من خلالها سحب عينة عشوائية من توزيع احتمالي محدد، لقد طُورت معظم طرق المعاينة المعروفة قبل نهايات العقد الثالث من القرن العشرين، ومن هذه الطرق: المعاينة الطبقية، والمعاينة المنتظمة، والمعاينة العنقودية، والمعاينة متعددة المراحل، والمعاينة المزدوجة، والمعاينة متعددة الوجوه. لقد استخدمت دائرة الإحصاءات الأمريكية المعاينة الاحتمالية لأول مرة في إحصاءات العاطلين عن العمل مع بداية عام 1940م. أما معاينات الاحتمالات غير المتساوية فقد طُورت في أربعينيات وخمسينيات القرن العشرين.

لقد استمرت نظريات وطرق المعاينة في التطور حتى نهايات القرن العشرين وحتى يومنا هذا، قام بعض الباحثين من بينهم: Godambe في خمسينيات القرن العشرين بدراسة الجوانب النظرية من طرق المعاينة، مما ساعد كثيراً في توضيح قضايا الاستدلال في المعاينة، ومهد الطريق لتقديم طرق جديدة للمعاينة، لقد جرى تطوير مجموعة من التصاميم وطرق الاستدلال استجابة لبعض المشكلات الصعبة التي واجهت الباحثين في مختلف فروع المعرفة، أسهمت بتطوير طرق المعاينة. كذلك الاختلاف بين وجهات النظر حول معاينة تعتمد على التصميم، ومعاينة تعتمد على النموذج، أسهمت في تطوير طرق للجمع بين الطريقتين، التطورات الأخيرة حول تحليل البيانات مع فقدان بعضها أسهم في تطوير طرق تحليلية جديدة أُستفيد منها في أوجه متعددة.

5.1 تصاميم وخلاصة المعاينة

يؤدي تصميم العينة دوراً مهماً في تحديد دقة التقديرات؛ لذا فإن الطريقة التي تسحب بها العينة مهمة كما للشكل الرياضي للمقدر من الأهمية.

obeikandl.com

يتضمن تصميم المعاينة: خطة سحب العينة، وطريقة التقدير. كلٌّ منهما يحتوي على مجموعة من المعالم: أولاً لابد أن نعرف وحدات المعاينة (الأولية والثانوية وهكذا)ثُمّ نختار إحدى طرق المعاينة مثل (المعاينة العشوائية البسيطة، أو المعاينة الطبقية، أو المعاينة المنتظمة، وهكذا)؛ نوزع الوحدات بين الطبقات المختلفة بطريقة تضمن التوزيع الأمثل لهذه الوحدات بين الطبقات، ومعاينة باحتمالات مختلفة.

نحاول تقدير معالم المجتمع باستخدام التقديرات الخطية، يمكننا أن نستخدم معلومات إضافية لتقديرات النسبة والانحدار، نحول المقدرات غير الخطية (على سبيل المثال تقدير النسبة) إلى مقدرات خطية قبل حساب التباين. كما أنه لابد من بذل جهد مهم في مسوحات المعاينة للسيطرة أو تقليل أخطاء عدم الإجابة.

6.1 بدائل المعاينة الاحتمالية

فيما يأتي بعض الطرق المعروفة للمعاينة غير الاحتمالية

1. جرى حصر العينة في جزء من المجتمع الذي يمكن الوصول إليه بسهولة. نقوم بسحب العينة من أعلى السيارة التي تحمل منتجاً زراعياً. أو سيارة مفتوحة محملة بالفحم الحجري.
2. نقوم بسحب العينة كيفما اتفق. عندما نريد أن نسحب 10 أرانب من القفص ونستخدمها في التجارب المختبرية، يقوم الباحث بإدخال يده في القفص والأرنب الذي تقع يده عليها يسحبها.
3. يقوم الباحث في المجتمعات الصغيرة وغير المتجانسة بسحب وحدات نموذجية لوحدات المجتمع وتكون قريبة من متوسط المجتمع.
4. تحتوي العينة على متطوعين فقط وهذا النوع من العينات شائع في الدراسات الطبية.

obeikandi.com

إذا توافرت الشروط الملائمة، فإنَّ أياً من الطرق أعلاه يمكن أن تعطي نتائج مفيدة، ولكنها لا يحكمها معطيات نظرية المعاينة القائمة على أساس لا معلمي، باعتبار أنه لا يوجد أي دور لعملية الاختيار العشوائية، وربما تكون الطريقة الوحيدة للتأكد من جودة أو صحة النتائج التي حصلنا عليها باستخدام إحدى هذه الطرق هو بمقارنتها بنتائج المجتمع إن وجدت، أو بنتائج حصلنا عليها عن طريقة معاينة احتمالية نفذت على نفس المجتمع في دراسات أخرى، وحتى إن كانت النتائج لصالح المعاينة غير الاحتمالية لا يوجد ما يضمن أننا سنحصل على نفس النتائج في المستقبل.

obeikandl.com

الفصل الثاني

مفاهيم أساسية للمسح بالعينات

1.2 مقدمة

لقد أصبحت المجتمعات في أيامنا هذه منشغلة بالأرقام، وهنالك رغبة للتعبير عن أي أمر من أمور حياتنا اليومية من "كتب الطبخ إلى الأفكار السياسية" باستخدام الأرقام، كما أن وسائل الإعلام اليومية من إذاعة وتلفاز وصحف ومجلات أصبحت مليئة بالمعلومات الرقمية، والحقيقة أنه لا توجد قضية يمكن طرحها وإثباتها بدون استخدام المعلومات الرقمية.

ومما لا يقبل الشك أن أفراد المجتمع أصبحوا أفضل معرفة من أي وقت مضى، وذلك بسبب استخدام المعلومات الرقمية أكثر فأكثر عن العالم الذي يحيط بهم، وهذا أمر جيد، ومن هنا ظهرت الحاجة إلى معرفة هذه الأرقام وتفسيرها، وهذا يحتاج منا إلى معلومات إحصائية معينة في الإحصاء، وعلى وجه التحديد معرفة الكيفية التي جمعت بها هذه المعلومات، والتأكد من مصداقيتها وتمثيلها إلى المجتمع الذي سحبت منه ومقدار الخطأ و... إلخ.

إن معرفة هذه الأشياء أمر حسن، ولكنه لا يكفي، إذ لا بد من معرفة كيفية تفسير هذه الأرقام أو تحليلها للوصول إلى النتائج المطلوبة، كما نرى أن العينات تمثل حجر الزاوية في كل هذه الأمور، بل هي الأساس الذي يعتمد عليه أي تحليل إحصائي أو أي استنتاجات يمكن الوصول إليها من البيانات التي وجدت غالبيتها باستخدام العينات، لذلك رأينا أن نتعرف على أساسيات

المسح بالعينات حتى يكون القارئ على دراية بهذه الأمور التي تعدّ الأساس لعملية المسح بالعينات.

2.2 تعريفات

1. **المجتمع** (Population): هو مجال الدراسة الذي يحوي مجموعة من العناصر التي نرغب بدراستها والحصول على بعض النتائج حولها. والمجتمع لا يعني فقط مجموعة من الأفراد ولكن قد يكون مجموعة من الحيوانات أو الحقول الزراعية أو الأشجار أو سلع معينة ينتجها مصنع أو ... إلخ.
2. **العينة** (Sample): عبارة عن مجموعة جزئية من المجتمع أو جزء من المجتمع.
3. **المسح الشامل** (Census survey): هو عملية عد أو حصر جميع عناصر المجتمع التي تخضع لصفات أو متغيرات معينة.
4. **وحدة المعاينة** (Sampling Unit): هي جزء مميز من المجتمع أو العنصر الذي تتم ملاحظته وإجراء القياسات وتسجيلها حوله. ووحدة المعاينة قد تكون طالباً في مدرسة أو جامعة، أو رأس غنم في قطيع أو قطعة أرض في قرية ... إلخ.
5. **تقدير** (Estimate): هو عبارة عن قيمة رقمية يجري حسابها من البيانات التي حصلنا عليها باستخدام العينة، وتستخدم بدلاً من معلمة المجتمع (Population Parameter) المراد تقديرها والتي تكون غير معروفة.
6. **عشوائي** (Random): تعني اختيار وحدات من المجتمع بطريقة موضوعية (تسمى طريقة المعاينة العشوائية) مثل استخدام جدول الأعداد العشوائية أو استخدام الحاسب الآلي أو أي طريقة من الطرق العشوائية الأخرى.

obeikandi.com

7. **العينة العشوائية** أو العينة الاحتمالية (Random Sample): هي عبارة عن عينة تم اختيارها **باستخدام** طريقة المعاينة العشوائية.
8. **عشوائية المعاينة** (Random Sampling): تعني بصورة عامة أي طريقة تستخدم العملية العشوائية في اختيار وحدات العينة من المجتمع.
9. **المعاينة بالإرجاع أو دون إرجاع**

(Sampling with Replacement or without Replacement): تُعنى بعملية معاملة الوحدات خلال مدة الاختيار، فإذا سحبنا الوحدة من المجتمع ولم نعدها إلى المجتمع سميت معاينة دون إرجاع، ولكن إذا تم إعادة الوحدة إلى المجتمع مرة ثانية ومن ثم قمنا بسحب وحدة جديدة تسمى معاينة مع الإرجاع.

3.2 لماذا نستخدم العينات وما مجالات استخدامها؟

تستخدم العينات بصورة واسعة -في مجالات الحياة المختلفة- لجمع معلومات تحتاج إليها جهات مختلفة، كالدوائر الحكومية أو المصانع أو الشركات التجارية أو الأفراد ... إلخ وفي الحقيقة كل واحد منا يستخدم العينات بصورة أو بأخرى، فمثلاً: إذا أردنا شراء سلعة معينة فإننا قبل أن نقوم بالشراء نقوم بجمع معلومات عن سعر هذه السلعة لدى بعض المحلات التي تبيعها، ومن ثم نقوم بشراء هذه السلعة، كذلك رب البيت الذي يقوم بشراء فاكهة معينة فإنه يقوم بعملية انتقاء هذه الفاكهة من بين المعروض من الفاكهة بناء على جودتها وسعرها ... إلخ، والأمثلة من هذا النوع كثيرة. ويتضح لنا من ذلك أننا استخدمنا العينات في شراء السلعة أو اختيار الفاكهة.

تستخدم العينات الآن بصورة واسعة في غالبية حقول المعرفة أو مجالات الحياة اليومية؛ ولذلك فليس من السهل أن نعطي تفصيلاً كاملاً عن جميع مجالات استخدام العينات. ولكن هنالك حالات يكون استخدام العينات فيها ضرورياً وواضح المزايا منها:

obeikandi.com

1- عندما نريد الحصول على نتائج دقيقة وذات ثقة عالية، لا سيما إذا كنا مقيدين بمبلغ محدود من المال.
2- عندما تكون الوحدات التي ندرسها ذات تشتت عالٍ بالنسبة إلى المتغيرات التي نريد دراستها.
3- عندما لا نستطيع القيام بالمسح الشامل وذلك لكونه عالي التكاليف.
4- عندما يكون مجال البحث واسعاً جداً والمجتمع غير معروف بصورة كاملة.
5- عندما نحتاج إلى نتائج سريعة لاتخاذ قرارات معينة لا يمكن الحصول عليها من خلال المسح الشامل، لأنه يحتاج إلى وقت طويل للإعداد والتنفيذ.
6- عندما تكون الموارد المالية والبشرية وكذلك الوقت غير كافية للقيام بالمسح الشامل.
7- عندما تتسبب عملية المسح في إتلاف وحدات المجتمع، ومثال ذلك معرفة نسبة البيض التالف من إنتاج مزرعة للدواجن.
8- قد يكون المجتمع غير قابل للعد مثلاً، كمخزون المملكة العربية السعودية من النفط أو مخزون الأردن من الفوسفات، ولمعرفة هذا المخزون يجب أن ننقب جميع الأراضي التابعة لهذين البلدين وهذا أمر غير ممكن.

على الرغم من المزايا الكثيرة لاستخدام العينات إلا أنها لا تخلو من عيوب أهمها:

1- مهما بلغت نتائج العينة من الدقة تبقى تقديرية وهي ليست النتائج الحقيقية للمجتمع.
2- المسح بالعينات لا يغطي إلا مجموعة جزئية من المجتمع وليس المجتمع بأكمله.

obeikandl.com

3- إن استخدام العينات يحتاج إلى أشخاص مدربين ومؤهلين علمياً للقيام بإعداد وتصميم العينات، وبدونهم لا يمكن الاعتماد على نتائج العينة.
4- تخطيط وتنفيذ المسح بالعينات يجب أن يتم بشكل دقيق جداً وإلا ستكون النتائج التي حصلنا عليها مضللة.

4.2 المسح الشامل والمسح بالعينة

إذا قمنا بدراسة جميع وحدات المجتمع بالنسبة لمتغيرات معينة فإن هذا ما يعرف بالمسح الشامل، ومما لا شك فيه أن المسح الشامل يحتاج إلى وقت طويل، وجهد كبير، وتكاليف عالية لتنفيذه، لذلك في كثير من الأحيان يكاد المسح الشامل يكون غير ممكن، وهنالك حالات (حتى ولو توفرت الجهود والتكاليف والوقت الكافي) لا يمكن معها القيام بالمسح الشامل، لعدم إمكانية تطبيقه، كما هو الحال في مثال معرفة مقدار مخزون النفط في السعودية، أو مخزون الفوسفات في الأردن.

فإذا قمنا باختيار مجموعة جزئية من المجتمع ومن ثم نقوم بجمع البيانات حول هذه المجموعة فإن هذا ما يسمى بطريقة المسح بالعينة، ومن الواضح أن المسح بطريقة العينة يحتاج إلى وقت وجهد ومال أقل بكثير من المسح الشامل. وهذا لا يعني بالضرورة أن السبب الاقتصادي هو الذي يقف خلف عملية اختيار المسح بالعينة، ولكن الأهم من ذلك لا بد من التأكد من دقة المعلومات التي حصلنا عليها سواء كان ذلك عن طريقة المسح الشامل أو المسح بالعينة. في بعض الأحيان تكون طريقة المسح بالعينة أكثر دقة من المسح الشامل. كذلك نستطيع التأكد من دقة المعلومات التي حصلنا عليها باستخدام المسح بالعينة. ولكن هذا غير ممكن بالنسبة إلى المسح الشامل حيث لا توجد طريقة للتأكد من صحة أو دقة المعلومات التي حصلنا عليها عن

طريق المسح الشامل، ومن الأمور الثابتة أن طريقة المسح بالعينة أفضل من المسح الشامل من حيث دقة النتائج التي نحصل عليها شريطة أن تكون العينة نفذت بشكل دقيق ومتقن، ولمزيد من المعلومات حول هذه النقطة يراجع Lahiri (1963) و Yates (1953)، Zarkovich (1961) و Mahalanobis (1950).

لقد لخص Cochran (1977) ميزات المسح بالعينة على المسح الشامل بأنها: تقليل تكاليف المسح، وسرعة الحصول على النتائج، ودقة أكبر في النتائج ومدى أوسع، وكذلك سهولة تعديل وتبديل المسح بالعينة.

وكذلك لخص أبو الإحصاء Fisher (1953) ميزات المسح بالعينة على المسح الشامل قائلاً: كنت أدعي أن هنالك أربعة أسباب لتفضيل المسح بالعينة. أما الثلاثة الأولى فهي سهولة التعديل والتبديل وسرعة التنفيذ وقلة التكاليف، أما الرابع فلا أحتاج لإضافة أي سبب آخر.

ولا بد من التأكيد هنا على أنّ المزايا التي ذكرناها لتفضيل المسح بالعينة على المسح الشامل لا يمكن أن تؤتي ثمارها إلا:

1. أن تكون وحدات المعاينة سحبت بطريقة صحيحة من المجتمع.
2. أن تكون طريقة المعاينة التي تم استخدامها ملائمة للمجتمع.
3. أن يكون حجم العينة الذي تم اختياره من المجتمع كافياً.

إذا كانت المعلومات المطلوب جمعها تخص جميع عناصر المجتمع فلا بد من استخدام طريقة المسح الشامل، كما هو الحال في التعدادات السكانية التي تقوم بها الدول كل عشر سنوات، بناء على توصيات الأمم المتحدة.

obeikandi.com

5.2 الخطوات الأساسية للمسح بالعينة

1- تحديد الأهداف

إن الخطوة الأولى التي يجب اتخاذها قبل القيام بالمسح بالعينة هي تحديد الهدف أو الأهداف من المسح، ولا بد من التأكد من أن هذه الأهداف متلائمة مع ما هو متوافر للمسح من أموال وأفراد ووقت.

2- تعريف المجتمع

لا بد أن يُعرف المجتمع الذي سنقوم بدراسته بطريقة واضحة لا تقبل التأويل أو عدم الوضوح، على سبيل المثال إذا أردنا تقدير مصروفات العائلة الشهرية لمدينة مكة المكرمة لا بد من تحديد وبشكل واضح حدود مدينة مكة المكرمة كأن نأخذ مثلاً الحدود البلدية المعتمدة، أو أي طريقة أخرى يمكن من خلالها تحديد حدود المدينة، كذلك لا بد أن يكون المجتمع الذي تسحب منه العينة وهو الذي يسمى **مجتمع المعاينة** مماثلاً للمجتمع الذي نهدف لدراسته وهو الذي يسمى **مجتمع الهدف**؛ ولذلك لا بد من تحديد الحدود السكانية والجغرافية والإدارية لهذا المجتمع.

3- تحديد إطار المعاينة ووحدات المعاينة

من المتطلبات الأساسية للمسح بالعينة تحديد إطار المعاينة، وهو عبارة عن كشف لجميع وحدات المجتمع، كذلك الإشارة إلى الوقت المناسب لجمع البيانات من هذه الوحدات، وتجدر الإشارة هنا إلى أن الإطار هو الذي يؤدي دوراً بارزاً في تحديد كيفية تخطيط وتصميم وبناء المسح بالعينة. كما يُعد الأساس الذي تدور حوله عملية اختيار الوحدات من المجتمع، وكذلك تقدير معلمات المجتمع. كذلك لا بد أن يكون المجتمع، قابلاً للتقسيم إلى وحدات

openkandil.com

تسمى وحدات المعاينة، وهي وحدات متميزة عن بعضها وواضحة وغير متداخلة، وتغطي جميع المجتمع.

4- اختيار إحدى طرق المعاينة

إن اختيار إحدى طرق المعاينة المناسبة يؤدي إلى نتائج نهائية يمكن الاعتماد عليها، لذلك لا بد من إعطاء عناية خاصة لاختيار حجم العينة، وطريقة اختيار الوحدات من المجتمع وتقدير معلمات المجتمع التي هي من الأمور الإحصائية الأساسية التي يجب أن تعطى عناية كافية، كذلك لا بد من الأخذ بعين الاعتبار دقة النتائج والتكاليف عند اختيار طريقة المعاينة المناسبة.

5- تنظيم العمل الميداني

إن الوصول إلى النتائج التي نرجو تحقيقها عن طريق استخدام المسح بالعينة يعتمد إلى حد كبير على دقة تنفيذ وجمع البيانات في الميدان، وعليه فلا بد من إنجاز العمل في الميدان بصورة دقيقة وصحيحة ومخلصة؛ وبناء على التعليمات التي وضعت من قبل مصممي المسح، لذلك لا بد من تدريب العاملين في ميدان جمع البيانات ومراقبتهم في أثناء قيامهم بالعمل، وكذلك مراقبة وتفتيش البيانات التي جمعوها بدقة وبصورة متواصلة في أثناء عملية جمع البيانات وبعد الحصول عليها.

6- تلخيص وتحليل البيانات

إن الخطوة الأخيرة في عملية المسح بالعينة هي تحليل البيانات والحصول على نتائج من العينة ومن ثم تعميمها على المجتمع، وتعد هذه الخطوة أساسية وحيوية. وبما أن نتائج المسح تعد الأساس في عملية وضع السياسات – وهذا

openkandil.com

يعتمد بصورة مباشرة على النتائج والتحليلات التي يتم التوصل إليها من البيانات – لذلك لا بد من إعطاء هذا الجانب أهمية كبيرة.

ويمكن تلخيص هذه الفقرة فيما يأتي:

أ. المحافظة على سرية البيانات.

ب. مراجعة البيانات وحذف الأخطاء الواضحة فيها وإدخالها إلى الحاسب الآلي.

ت. وضع هذه البيانات في جداول.

ث. القيام بمعالجة هذه البيانات بإجراء التحليلات الإحصائية المناسبة.

ج. كتابة تقرير في النتائج والاستنتاجات النهائية.

6.2 المعاينة الاحتمالية والمعاينة الاعتباطية (غير الاحتمالية)

المعاينة الاحتمالية هي الطريقة التي نقوم بها باختيار العينات وفقاً لبعض قوانين الاحتمالات، بحيث إن كل وحدة من وحدات المجتمع يكون لها احتمال محدد بالظهور في العينة. كذلك كل عينة من العينات الممكنة يكون لها احتمال محدد بالظهور، ومما تجدر الإشارة إليه هنا أن غالبية العينات المستخدمة هنا يكون فيها احتمال ظهور أي عنصر من عناصر المجتمع في العينة متساوياً، وكذلك احتمال اختيار عينة من بين العينات الممكنة الظهور متساوية، كما أننا سوف نقوم فقط بدراسة العينات الاحتمالية.

المعاينة الاعتباطية (غير الاحتمالية) هي طريقة لاختيار عينة أو عينات من المجتمع، بحيث إن عملية اختيار وحدات العينة يعتمد بصورة مباشرة على الشخص الذي يقوم بعملية الاختيار أو السحب، إن هذه الطريقة تُعد غير موضوعية أو شخصية، في هذه الطريقة يقوم الشخص المعاين بفحص جميع

oberkampf.com

عناصر المجتمع ومن ثم يقوم باختيار وحدات من المجتمع تكون قريبة من متوسط أو وسط المجتمع، وتستخدم هذه الطريقة للقيام بمعرفة آراء الآخرين في قضية معينة، ولكن لا يمكن بأي حال من الأحوال أن يوصى بهذه الطريقة؛ لأنها تعتمد بالدرجة الأولى على الشخص الذي يقوم بالمعاينة، وأكثر الناس لا يطمئنون لهذا الأسلوب؛ لأنه لا يخلو من التحيز من قبل المعاين، ومع ذلك إذا كان المعاين ذا خبرة ودراية وموضوعية فإنه يمكن الاستفادة من هذا النوع من المعاينة للحصول على نتائج مفيدة، ومع هذا فإنها تعاني من عيب كبير ألا وهو عدم القدرة على قياس درجة دقة التقديرات التي حصلنا عليها باستخدام هذا الأسلوب من المعاينة.

7.2 وحدة المعاينة

لقد سبق أن عرفنا وحدة المعاينة لكنها تحتاج إلى مناقشة بصورة أعمق؛ لأنها تشكل الأساس في عملية المعاينة، فالوحدة قد تكون وحدة طبيعية للمجتمع كالفرد في المنطقة السكنية، وقد تكون بتجمع طبيعي لهذه الوحدات كالعائلة، وقد تكون وحدة مصطنعة كالحقل أو قطعة الأرض ... إلخ. وقبل أن تبدأ بسحب العينة، يجب أن يكون المجتمع مقسماً إلى أجزاء متميزة وواضحة وغير متداخلة، ويجب أن تحدد وحدات المعاينة بحيث إن كل عنصر من عناصر المجتمع يقع في وحدة واحدة فقط من وحدات المعاينة، وخلاف ذلك تكون بعض العناصر لا تقع في أي وحدة ممكن سحبها من المجتمع، فعلى سبيل المثال وحدة المعاينة العائلة يجب أن تحدد بحيث إن أي شخص لا يمكن أن يكون في أكثر من عائلة، ولا يترك أي شخص دون عائلة.

obeikandi.com

8.2 الإطار أو إطار المعاينة

سبق أن عرفنا الإطار بأنه عبارة عن كشف كامل لجميع وحدات المعاينة التي تمثل المجتمع قيد الدرس، وتعد عملية بناء إطار المعاينة من الأمور الصعبة التنفيذ من الناحية الواقعية، وبصورة عامة يعد الإطار مثالياً إذا كان كاملاً وشاملاً ويحتوي على أحدث التطورات الحاصلة في المجتمع بالنسبة إلى جميع وحدات المعاينة؛ لذلك لا بد أن تجري عملية تحديث الإطار باستمرار، وأن يكون الإطار خالياً من الأخطاء ولا يهمل أي وحدة من وحدات المجتمع، ولا يكرر بعض الوحدات أكثرة من مرة واحدة، وهناك نواحٍ مختلفة بالنسبة إلى إطار المعاينة جرى مناقشتها من قبل Yates (1960) و Seal (1962) و Singh (1978) و Mahalanobis (1944) وغيرهم، على أن هنالك بعض العيوب التي يمكن أن نلاحظها في إطار المعاينة منها:

1. قد يكون الإطار غير كامل وذلك عندما تكون بعض وحدات المعاينة قد أُهملت بصورة كاملة أو كررت أكثر من مرة.
2. قد يكون الإطار غير دقيق، وذلك بأن تكون بعض الوحدات قد سجلت بصورة غير صحيحة أو أن بعض الوحدات غير الموجودة سجلت في الإطار.
3. وقد يكون الإطار غير كافٍ عندما لا يحتوي على جميع طبقات المجتمع التي يجب أن تدخل في عملية المسح.
4. قد يكون الإطار قديماً ولم تجر عملية تحديثه على الرغم من أنه كان كاملاً ودقيقاً وكافياً في الوقت الذي أُعد فيه.

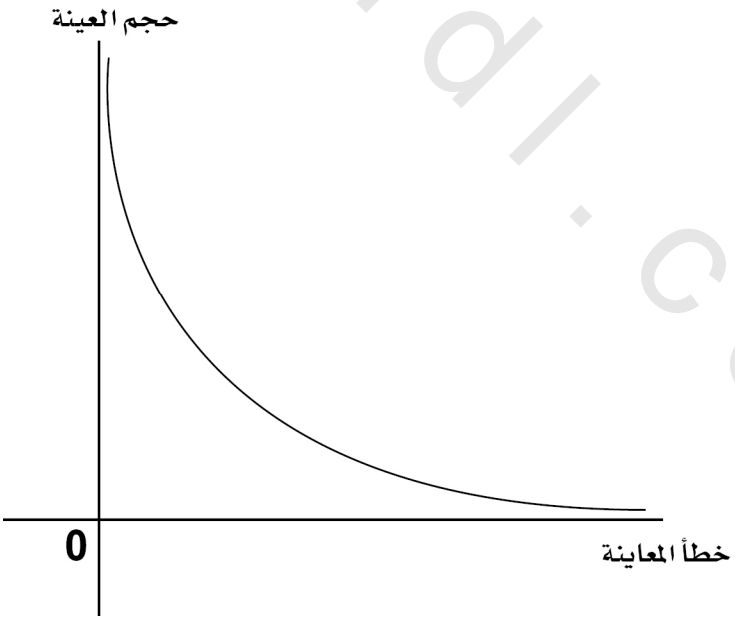
ولا بد من الإشارة هنا إلى إن الإطار صعب التنفيذ بصورة دقيقة وصحيحة، ولكن الخبرة والمثابرة دائماً تؤديان دوراً في عملية بناء الإطار المتكامل.

obeikandi.com

9.2 أخطاء المعاينة والأخطاء الأخرى

1- أخطاء المعاينة

إن الخطأ الذي يحدث بسبب استخدام العينة أو البيانات التي حصلنا عليها من العينة لتقدير معلمات المجتمع يسمى خطأ المعاينة، فمهما كانت الطريقة التي نختار فيها العينة سيبقى هنالك فرق بين معلمات المجتمع (Population Parameters) وما يقابلها من تقدير لهذه المعلمات باستخدام العينة. إن خطأ المعاينة موروثٌ ولا بد منه، كما أن العينة التي يكون مقدار الخطأ فيها قليلاً تعد ممثلة جيدة للمجتمع الذي سحبت منه، ويمكن باستمرار تقليص خطأ المعاينة وذلك بزيادة حجم العينة، في الحقيقة إن حجم العينة يتناسب عكسياً مع الخطأ كما هو موضح في الشكل 1.1 ويلاحظ أنه عندما يكون حجم العينة مساوياً لحجم المجتمع يكون الخطأ صغيراً جداً أو صفراً.



الشكل 1 – 1: العلاقة بين خطأ المعاينة وحجم العينة

obeikandl.com

2- الأخطاء الأخرى

بالإضافة إلى خطأ المعاينة هنالك أخطاء أخرى يمكن إجمالها فيما يأتي:

1. عدم قياس بعض وحدات العينة التي سحبت.
2. أخطاء في الملاحظة قد يكون سببها عدم الدقة في القياس أو عطل بعض أجهزة القياس المستخدمة.
3. أخطاء بشرية مثل خطأ في تدوين بعض الأرقام أو جدولة بعض النتائج بصورة خاطئة.

على الأرجح أن تزداد هذه الأخطاء مع زيادة حجم العينة وهذا على العكس من خطأ المعاينة، بصفة عامة المسح الشامل أكثر عرضة إلى أخطاء من هذا النوع، وفي المقابل لا يوجد خطأ للمعاينة بالنسبة للمسح الشامل.

10.2 جداول الأعداد العشوائية

يُعد مفهوم العشوائية من الأمور المهمة جداً للمعاينة، وخصوصاً إذا أردنا استخدام العينة للوصول إلى نتائج، ومن ثم تعميمها على المجتمع، والسبب في ذلك يعود إلى أننا نريد أن نطبق قوانين الاحتمالات للقيام بالاستنتاج الإحصائي، إن العشوائية لا تعني التخبط في الاختيار، ولكنها تعني طريقة معرفة جيدة وبدقة للقيام بعملية الاختيار لوحدات العينة من المجتمع. والعشوائية نعني بها هنا هي الطريقة المنتظمة التي تعطي كل وحدة في المجتمع احتمالاً معلوماً أو فرصة معلومة في الاختيار من المجتمع، في أكثر طرق المعاينة العشوائية يكون احتمال جميع الوحدات بالاختيار متساوياً، ولكن لاشك أن هناك بعض طرق المعاينة يكون احتمال اختيار الوحدات فيها غير متساو، كما هو الحال بالنسبة إلى طرق المعاينة في الفصل العاشر من هذا الكتاب.

openkandil.com

توجد طرق كثيرة لاختيار الوحدات عشوائياً من المجتمع، ومن الطرق الشائعة الاستخدام هي استخدام جداول الأعداد العشوائية، وهذه الجداول معروفة ومتداولة راجع ملحق رقم (3)، وكذلك الطريقة الأخرى التي أصبحت الآن أكثر استخداماً بسبب انتشار الحاسب الآلي حيث توجد حقائب إحصائية جاهزة يمكن استخدامها في عملية السحب العشوائية مثل MINTAB راجع الملحق رقم (2). على أنه ليس من الصعب كتابة برنامج للحاسب ليقوم بعملية الاختيار العشوائي للوحدات من المجتمع.

إن جداول الأعداد العشوائية المطبوعة عبارة عن مجموعة مكونة من الأرقام 9 ,... ,2 ,1 ,0 اختيرت عشوائياً باستخدام إحدى طرق العشوائية المعروفة، ووضعت على شكل جدول مكونة من صفوف وأعمدة. تحتوي الجداول على نفس العدد من كل رقم من دون أي دورة أو ترتيب معين للأرقام. إن الرقم الموجود في موضع معين حدد بصورة مستقلة عن الأرقام في المواضع الأخرى، وفي العادة ترقم الصفوف والأعمدة في كل صفحة ليسهل استخدامها في جمع الأعمدة المتجاورة مع بعض إذا كانت هناك حاجة لتكون عدداً مؤلفاً في عدة أرقام.

قبل بدء عملية سحب العينة لا بد من تحديد الحد الأعلى لحجم المجتمع؛ وذلك ليكون لدينا عدد كافٍ من الأرقام في كل عدد وذلك لإعطاء جميع الوحدات الفرصة نفسها في الظهور، على سبيل المثال إذا كانت المجموعة التي نسحب منها العينة بين 100 و1000 وحدة. لذلك لا بد من استعمال 3أرقام في العدد الواحد. وإذا كانت المجموعة بين 1000 و10.000 وحدة لا بد من استخدام 4 أرقام في العدد الواحد وإذا كان حجم المجموعة غير معروف نحتاج إلى الحد الأعلى للمجموعة.

لاستخدام جداول الأعداد العشوائية نبدأ في نقطة تسمى نقطة البداية. كثير من الكتاب يقترحون فتح صفحة من جداول الأعداد العشوائية، ومن ثم نقوم برمي القلم داخل الجدول والرقم الذي يقع عليه القلم يكون هو نقطة البداية، يبدو أن هذه الطريقة غير علمية، وقد تؤدي إلى أن جميع المسوحات تبدأ بالمنطقة نفسها في الجدول، كطريقة بديلة يمكن أن تحدد نقطة البداية بسحب أحد الأعمدة وأحد الصفوف بصورة عشوائية، والرقم الذي يلتقي عنده العمود والصف المختاران بصورة عشوائية يمثل نقطة البداية. والخطوة الأخيرة هي تحديد العلاقة بين العدد العشوائي الذي اختير من الجدول وعناصر المجتمع. وهذا يكون إذا سحب الرقم 4256 مثلاً فإنه يعني أن وحدة معينة من وحدات المجتمع اختيرت ونستطيع الآن تحديدها والوصول إليها ومن ثم جمع المعلومات منها.

مثال: لنفترض أننا نريد أن نسحب عينة عشوائية متكونة من 5 وحدات من مجتمع يحتوي على 200 وحدة، ولتكن هذه الوحدات تمثل عائلات واقعة في منطقة سكنية معينة، تبدأ عملية سحب عمود وصف بطريقة عشوائية وحيث إن في الجدول الموجود في الملحق رقم 1 توجد في الصفحة 13 أعمدة كل عمود يحتوي على 4 أرقام أي في الحقيقة لدينا 52 عموداً و51 صفاً في الصفحة، نقوم باختيار عمود من بين الأعمدة وصف من بين الصفوف بطريقة عشوائية. لنفترض أنه ظهر لدينا العمود رقم 27 والصف رقم 34 لذلك يكون الرقم 3 هو نقطة البداية، ولقد رُقّمت وحدات المجتمع 001، 002، 003، ... 200 وعليه سوف تختار كل مرة عدداً مكوناً من 3 أرقام، كما ذكرنا فإن نقطة البداية هي الرقم 3 الواقع عند تلاقي العمود رقم 27 والصف رقم 34. الآن نبدأ التحرك نحو إما يمين أو يسار، كل مرة نأخذ 3 أرقام 360، 260، 406، وهكذا. سوف نهمل الأعداد التي تكون أكبر من 200، ونأخذ الأعداد التي هي أقل أو تساوي 200؛ لأن مجتمعنا يتكون من 200 وحدة. لذلك ستكون

الوحدات رقم 166، 124، 068، 060، 092 في العينة. طبعاً إذا انتهى صف ننتقل إلى الصف الذي يليه وهكذا. إذا انتهت الصفحة نرجع إلى بداية الصفحة وهكذا. من الملاحظ أن كثيراً من الأعداد أُهملت لأنها غير داخلة في المجتمع. يمكن استخدام بعض الطرق الرياضية للتقليل من عدد الأعداد التي تهمل، وبالإمكان سحب هذه العينة بسهولة جداً وذلك باستخدام الحقيبة الإحصائية MINTAB انظر الملحق رقم 2.

obeikandl.com

المراجع العربية

1. وليم كوكوران – ترجمة أنيس كنجو – تقنية المعاينة الإحصائية (1995)، جامعة الملك سعود، الرياض.
2. حسين علوان (1994)، طرق المعاينة، دار الفرقان، عمان.

References

1. Barnett, V. (1991). Sampling Survey Principles and Methods, Edward Arnold, London.
2. Cochran, W. G. (1977). Sampling Technique, 3rd Ed. John Wiley & Sons, New York.
3. Fisher, R. A. (1953). The Design of Experiments, Oliver & Boyd, London.
4. Lahiri, D. B. (1951). A Method of Sample Selection Providing Unbiased Ratio Estimates Bull. Int. Statist., 33, 133, 140.
5. Levy, P. S. and Lemeshow, S (1999). Sampling of Population: Methods and Applications, Wiley, New York.
6. Mahalanobis, P. C. (1944). On the Large-Scale Sample Surveys, Phil. Trans. R. Soc., 231, 329-451.
7. Mahalanobis, P. C. (1950). Cost and Accuracy of Results in Sampling and Complete Enumeration, Bull. Int. Statist. Inst., 32, 210-213.
8. Scheaffer, R. L., Mendenhall, W. and Ott, L. (1996). Elementary Survey Sampling, 5th Ed., Duxbury press, Boston.
9. Seal, K. C. (1962). Use of Outdated Frames in Large Sample Surveys. Bull-Cal. Statist Assoc., 11, 68-84.
10. Singh, D. (1978). Taking Agricuitural Censuses, FAO, Rome.
11. Sturat, A. (1984). the Ideas of Sampling, Revised Edition, Griffin, London.
12. Thompson, S. K (2002). Sampling, 2end Wiley, New York.
13. Yates, F. and Grundy, P. M. (1953). Selection without Replacement from Within Strata Probability Proportional to Size J. R. Statist, Soc. 15 B, 253-261.
14. Yates, F. (1960). Sampling Methods for Censuses and Surveys, Charles Grittin & Co., London.
15. Yates, F. (1981). Sampling Methods for Censuses and Surveys 4th Ed., Grittin, London.
16. Zarkovich, S. S. (1961). Sampling Methods and Censuses, FAO, Rome.

obeikandl.com

obeikandl.com

الفصل الثالث

متطلبات العينة الجيدة

1.3 مقدمة

إن الهدف الرئيس لأي طريقة من طرق المعاينة هو الحصول على عينة يمكن من خلالها الوصول إلى صفات المجتمع تحت الدراسة وبخاصة تلك التي تهمنا بالدرجة الأولى، أو التي يهدف المسح بالعينة الحصول على معلومات عنها. فمثلاً إذا أردنا الحصول على معلومات عن مصروفات العائلة الشهرية على المواد الغذائية في بلد معين أو منطقة معينة، فإن المعلومات التي تهمنا بالدرجة الأولى هي: دخل العائلة الشهري، أو معدل دخل العائلة الشهري، ومعدل مصروف العائلة على المواد الغذائية، ولكن قد تكون هنالك صفات أخرى للمجتمع نريد معلومات عنها مثل معدل عدد أفراد العائلة الواحدة، ومعدل عدد العاملين في العائلة، ومعدل عدد الأطفال دون سن 18 سنة ... إلخ. ولكن ليست هي الهدف الرئيس للمعاينة، ولكن هذا لا يعني أن نصل إلى صفات المجتمع بأي وسيلة وإنما ضمن الشروط والحدود المفروضة أو الموضوعة للبحث، على سبيل المثال قد نحدد وقتاً نهائياً لإنجاز المسح، أو قد تكون ميزانية الدراسة محدودة ... إلخ.

2.3 التحيز

قد يخطر على بال كثير منا أنه إذا أردنا تقدير معدل إنتاج الدونم الواحد من القمح أن نسأل أحد ذوي الخبرة أو أحد الموظفين العاملين في دائرة الزراعة

obeikandl.com

كي يخرج إلى حقول القمح ويختار بعض الحقول ذات الإنتاجية المتوسطة، ويحدد إنتاجية هذه الحقول، ومن ثم نستطيع أن نحدد معدل إنتاج الدونم الواحد من القمح في منطقة معينة بهذه الطريقة. تبدو هذه الطريقة للوهلة الأولى أنها جيدة لأنها تختار حقولاً متوسطة الإنتاجية، ونحن هدفنا تحديد معدل إنتاج الدونم الواحد. إن هذه العينة أو هذه النوعية من المعاينة لا يمكن بأي حال من الأحوال الاعتماد عليها، لأنها لا قيمة لها. ولكن خطؤها الوحيد أنها يمكن أن تكون متحيزة، وهذا يعني أن عملية اختيار الحقول ذات الإنتاجية المتوسطة معرضة للخطأ في الاختيار، وذلك لعدة أسباب منها:

1. قد يكون الخبير أو موظف دائرة الزراعة في المنطقة يميل إلى اختيار الوحدات ذات الإنتاجية العالية، لأنه يرغب أن تظهر منطقته على أنها أفضل المناطق في الإنتاج، أو أنها سياسة عامة من قبل الدولة للميل نحو إظهار أن إنتاجية البلد منخفضة لأسباب سياسية أو أمنية أو ... إلخ. وكذلك فقد يشعر الخبير أو الموظف بأنه ليس من مصلحة الفلاحين إعطاء إنتاجية عالية لحقولهم، لأسباب قد تكون ضريبية أو لمساعدات تقدمها الدولة إلى الفلاحين.
2. حتى ولو سلمنا أن الخبير أو الموظف سوف يقوم بواجبه ويعطي الأرقام الحقيقة مهما كانت الظروف، تبقى هنالك أخطاء غير ممكنة السيطرة عليها، كالأخطاء الشخصية لأنه تبقى العملية اختيارًا شخصيًا. وممكن أن يقع الشخص في الخطأ من دون قصد. وقد يكون خطأ تحديد حقل واحد من الحقول يلغي كل الدقة التي حصل عليها من خلال تحديد الحقول الأخرى.
3. إذا كان الخبير أو الموظف ذا خبرة وموضوعية ويمكن الحصول على نتائج جيدة وخالية من الأخطاء حتى الشخصية منها، فإن هذه الطريقة

obeikandl.com

لا يمكن قياس دقة التقديرات التي حصلنا عليها من خلالها، ومن ثم لا نستطيع الاعتماد على النتائج التي حصلنا عليها.

لا بد أن نفرق بين نوعين من أخطاء المعاينة، الأول سببه التحيز في الاختيار، والثاني سببه الفرق بين وحدات المجتمع الموجودة في العينة والوحدات الأخرى الموجودة في المجتمع ولكن ليست في العينة، ويسمى الأول خطأ التحيز (أو سببه التحيز) ويسمى الثاني خطأ المعاينة العشوائية أو يسمى خطأ المعاينة إذا لا يوجد تحيز في الاختيار، على أن الخطأ بسبب التحيز سوف لا يقل مهما كبر المجتمع أو كبرت العينة، ولكن خطأ المعاينة العشوائية يقل بالمعدل إذا زاد حجم العينة.

3.3 طرق اختيار تسبب التحيز

توجد هنالك طرق عديدة يمكن أن نختار بها وحدات العينة بطريقة غير صحيحة قد تسبب تحيزاً في الاختيار، وأهم هذه الطرق يمكن إجماله بصورة عامة فيما يأتي:

1. الاختيار المتعمد لعينة ممثلة للمجتمع وهذا النوع من التحيز نوقش في الفقرة أعلاه.
2. اختيار يعتمد على بعض الصفات التي تكون لها علاقة أو تكون مرتبطة مع خواص الوحدة ذات الاهتمام، إن كثيراً من الاختيار الاعتباطي يسبب تحيزاً من هذا النوع.
3. اختيار وحدات العينة من غير وجود طريقة واضحة تستعمل في اختيار وحدات العينة من بين وحدات المجتمع عشوائياً. أي عدم وضوح الطريقة العشوائية في الاختيار. لكن الباحث يدعي أنه طبق العشوائية، وربما يسمح الباحث لرغباته بإيجاد نتائج معينة يكون لها تأثير على اختياره

للوحدات، وهذا النوع من التحيز أخطر أنواع التحيز؛ لأن وجوده غير ظاهر للعيان.

4. غالباً ما يعمد الباحثون أو العادون إلى تعويض وحدة ملائمة من المجتمع بدل الوحدة الموجودة في العينة، والتي وجدوا صعوبة ما في الوصول إليها للحصول منها على المعلومات. فيؤخذ المنزل المجاور بديلاً عن المنزل الموجود في العينة عندما لا يكون هنالك جواب من المنزل المعني. هذا يؤدي إلى أن المنازل التي ستكون في العينة هي المأهولة طوال اليوم؛ أي المنازل التي يسكن فيها عائلات مع أطفالهم.
5. الفشل في تغطية كل العينة المختارة. فإذا لم تكن هنالك زيارة ثانية إلى المنازل التي لم نحصل على جواب منها في الزيارة الأولى سيكون هنالك تحيز في الاختيار، حتى ولو لم نحاول التعويض عنها في المرة الأولى، وهذه الظاهرة تظهر جلياً في المسح الذي يُجرى بواسطة البريد، حيث إنه غالباً ما يكون الجواب غير كامل من قبل كل الذين أرسلنا لهم الاستبانات، ويرد غالباً الأشخاص الذين لهم اهتمام بموضوع البحث، وهؤلاء لا يمثلون المجتمع.

4.3 تجنب التحيز في الاختيار

من الواضح أنه إذا كان هنالك تحيز في الاختيار لا يمكن الوصول إلى نتائج موضوعية من العينة؛ إذن المهمة الأولى الأساسية لأي طريقة من طرق المعاينة هي التخلص من مصادر التحيز، إن أبسط طريقة عالمية للتخلص من التحيز أو لتجنب التحيز هي أن نسحب العينة بصورة عشوائية كاملة أو عشوائية مشروطة، تؤدي إلى زيادة دقة العينة ولا تؤدي إلى التحيز في النتائج. وفي بعض الأحيان هنالك ما يسمى بالاختيار المنظم، مثل اختيار أسماء من كشف، على مسافات متساوية من بداية الكشف إلى نهايته، أو اختيار نقاط

على خريطة باستخدام المربعات المتساوية المساحات، هذا النوع من الاختيار ربما يكون مسموحاً به.

إن عشوائية الاختيار لا تعني التخبط في الاختيار، ويمكن الحصول على العينة العشوائية باستخدام بعض الطرق العشوائية الملائمة، كاستخدام جداول الإعداد العشوائية، أو استخدام الحاسب الآلي .... إلخ. ولا بد من التأكيد هنا على أن كلمة عشوائية أو عينة عشوائية يساء استخدامها كثيراً من قبل الناس، ولهذا السبب لا بد أن تشرح بصورة واضحة طريقة سحب العينة العشوائية من المجتمع في جميع الأعمال التي تستخدم العينة العشوائية.

ولأجل منع التحيز الناتج عن الإهمال أو التعمد "إذا كان العمل كبيراً نسبياً" يفضل أن ينفذ بصورة مركزية من قبل دائرة واحدة بطريقة لا تترك أي خيارات أمام العادين أو الباحثين. وكذلك أن ينفذ بطريقة يسهل معها مراقبة العمل الميداني كلما دعت الحاجة إلى ذلك.

5.3 تحيز بسبب سوء تحديد وحدات المعاينة

عندما لا تكون وحدات المعاينة وحدات طبيعية للمجتمع، يجري في الغالب تحديد حدود الوحدات التي اختيرت في الوقت الذي تجري فيه عملية القياسات، إن عملية تقدير إنتاج المحاصيل الحقلية تجري عادة عن طريق اختيار وحدات صغيرة من المساحة للحصول منها على تقدير للمحصول أو أي صفات أخرى للمحصول. وتجري عملية تحديد مواقع هذه المساحات الصغيرة نظرياً من خلال اختيار نقاط عشوائية في نظام الإحداثيات، وهذا يضمن العشوائية في الاختيار، وبما أنه واقعي فمن المستحيل تحديد موضع هذه المساحات في الحقل بصورة دقيقة؛ لذلك لا بد من استخدام طرق تقريبية. ويجب أن تكون المساحات غير صغيرة في هذا النوع من العمل؛ لأن الخطأ في تحديد حدود المساحات يزداد كلما صغرت المساحة، وكذلك فإن أي تغيير

بسيط في موضع الوحدة سوف يؤثر في النتائج إذا كانت المساحات صغيرة. إن الوحدات الصغيرة تعطي نتائج دقيقة جداً إذا كان هناك عاملون ذوي خبرة في العمل الميداني، ولكن ربما تكون النتائج غير دقيقة إذا كان العاملون غير مدربين ولا توجد لديهم خبرة ميدانية.

6.3 تحيز في التقدير

بالإضافة إلى التحيز الذي ينتج عن سوء طريقة الاختيار أو تحيز سببه عملية جمع المعلومات، يوجد هنالك تحيز آخر ناتج من الأسلوب المتبع لتقدير النتائج، ومثال بسيط على هذا النوع من التحيز يحدث عند تقدير النسبة. فعلى سبيل المثال، توجد ثلاثة أنواع من الأراضي المختلفة من حيث الخصوبة، ومن ثم مختلفة من حيث معدل إنتاجية الدونم الواحد من محصول معين، فإذا كان معدل إنتاج هذه الأنواع الثلاثة من الأرض من محصول معين على التوالي 5 طن، 3 طن، 2 طن ومعدل مساحة الحقول الثلاثة على التوالي هي 5 دونم، 10 دونم و15 دونم وعدد الحقول مساوٍ للأنواع الثلاثة. إن معدل إنتاج الدونم الواحد يمكن الحصول عليه باستخدام الوسط الحسابي المرجح $\bar{Y}_w$ طن للدونم الواحد:

$$\bar{Y}_w = \frac{5(5) + 3(10) + 2(15)}{5 + 10 + 15}$$

بينما الوسط الحسابي لإنتاج للدونم الواحد لكل الحقول $\bar{Y}$ طن للدونم الواحد:

$$\bar{Y} = \frac{5 + 3 + 2}{3} = 3.33$$

والنتيجة وجود تحيز في التقدير من قبل الأخير $\overline{Y}$ مقداره حوالي 18%، على أن التحيز بسبب التقدير مقدور عليه من خلال استخدام طرق التقدير الملائمة.

7.3 الحالات التي يسمح بها التحيز

على الرغم من أن تحاشي التحيز يُعَدُّ موضوعاً مهماً وخصوصاً في المسوحات التي يترتب عليها قرارات إدارية معينة، إلا أنه يمكن قبول كمية معينة ثابتة من التحيز في بعض المسوحات حقب. ويعد التحيز في المسوحات التي تجري على حقب منتظمة –والتي يكون الهدف منها تحديد التغير الحاصل من حقبة إلى أخرى وليس القيمة المطلقه– مقبولاً إذا كان صغيراً وثابتاً مع الوقت. كذلك في المسوحات التي يكون هدفها الرئيس المقارنة بين مجموعات مختلفة من المجتمع، يعد التحيز الذي يكون ثابتاً من مجموعة إلى أخرى ذا أهمية قليلة، ويجب أن يتحاشى الباحثون وضع أهمية مبالغ فيها على مصادر التحيز الصغيرة، التي ينتج عنها في الحقيقة خطأ تافه قياساً إلى خطأ المعاينة العشوائية.

8.3 طرق لتقليل أخطاء المعاينة

بعد التخلص من التحيز لا بد من توجيه اهتمامنا إلى أخطاء المعاينة، هذه الأخطاء يجب أن تكون صغيرة حتى نصل إلى الدقة المطلوبة، فإن أبسط الطرق لزيادة الدقة هو زيادة حجم العينة، إذا كانت الأمور الأخرى كما هي عند زيادة حجم العينة، فإن خطأ المعاينة يتناسب عكسياً على الجذر التربيعي لحجم العينة.

إن الحصول على الدقة المطلوبة لا يعتمد فقط على حجم العينة، وإنما أيضاً على التغير لكل وحدة، بصورة أكثر تحديداً على الجزء من التغير

لكل وحدة والذي يدخل ضمن خطأ المعاينة. وباستخدام طرق اختيار ملائمة – والتي تضع قيوداً على العملية العشوائية الكاملة التي لا تسبب تحيزاً في النتائج– يمكن دائماً تقليل خطأ المعاينة بصورة كبيرة رغم بقاء حجم العينة ثابتاً، إن من أبسط أنواع القيود التي تضع على عملية الاختيار العشوائية الكاملة هي (الطبقية) التي يقسم فيها المجتمع إلى طبقات من الوحدات، بحيث تكون الوحدات في كل طبقة متماثلة، وبعد ذلك تسحب عينة عشوائية من كل طبقة. فإذا سحبنا من كل طبقة وحدات متناسبة مع حجم الطبقة سيؤدي هذا إلى إزالة الفروق بين الطبقات المختلفة، ومن ثم التقليل من خطأ المعاينة.

وبالإضافة إلى استخدام الطبقية هناك ثلاثة طرق أخرى يمكن استخدامها لزيادة الدقة، وفي بعض الأحيان تؤدي إلى زيادة كبيرة في الدقة، ومن ثم يمكن تقليل خطأ المعاينة. وهذه الطرق هي:

1. استخدام المعلومات الإضافية المتوافرة. وسوف نتكلم عن هذا الموضوع بالتفصيل في أحد الفصول القادمة.
2. استخدام العينة متعددة المراحل أو العينة العنقودية، وكذلك سوف نتكلم عن هذا النوع من المعاينة في أحد الفصول القادمة.
3. استخدام التوزيع الأمثل. إن هذه الطريقة تضمن لأجزاء مهمة أو أجزاء أكثر تغيراً من المجتمع أن تسحب منها وحدات أكثر. إن التوزيع الأمثل للعينة يعتمد على التغير النسبي للطبقات المختلفة التي تقسم المجتمع، فمثلاً إذا كان مطلوباً أن نحدد عدد العاملين في صناعة معينة فمن الأفضل أن نأخذ عدداً أكبر من المعامل الكبيرة وأقل من المعامل الصغيرة.

obeikandi.com

9.3 اختيار الوحدة

توجد اختيارات كثيرة ومتنوعة في اختيار نوع وحجم وحدات المعاينة في بعض أصناف المواد، وهذا يعطي مجالاً واسعاً لزيادة كفاءة طريقة المعاينة. وبصفة عامة كلما صغر حجم الوحدة من المواد الداخلة في العينة زادت دقة النتائج، فعلى سبيل المثال في المسح الزراعي سيكون أكثر دقة أن نأخذ 10% من جميع الحقول الموجودة في كل وحدة إدارية من أن نأخذ جميع الحقول الموجودة في 10% من الوحدات الإدارية، ويبقى هذا الكلام صحيحاً حتى في حالة استخدام العينة متعددة المراحل.

إن الحاجة إلى وحدات صغيرة موزعة على كل أجزاء المجتمع غالباً ما يتناقض مع المتطلبات الإدارية، واضح لنا أنه من السهل الإعداد لمسح حقول في مناطق صغيرة مثل الوحدات الإدارية أو المقاطعات من الإعداد لمسح بنفس العدد من الحقول موزعة على كل أجزاء البلد. إن الاختبار المتوازن بين هذين الأمرين المتناقضين غالباً ما يكون هو إحدى المشكلات الرئيسة للتخطيط للمسح بالعينة. فإذا قمنا بسحب عدد صغير من الوحدات الكبيرة في العينة فإن ذلك سيؤدي إلى عدم تحديد خطأ المعاينة بصورة جيدة؛ لأنه لا يوجد سوى فرق قليل نسبياً بين الوحدات الكبيرة.

ولا بد من الإشارة هنا إلى أن اختيار طريقة المعاينة لا يعتمد فقط على دقتها بالنسبة إلى الطرق الأخرى، ولكن يعتمد كذلك على مدى ملاءمتها للتطبيق على أرض الواقع، إن من الأهمية بمكان أن تعتمد طريقة المعاينة على ما هو متوافر في المجتمع من المعلومات والوسائل، فمثلاً قد تكون طريقة معينة ممتازة في بلد توجد فيه خرائط، بالمقابل تكون عديمة الفائدة في بلد آخر لا توجد فيه خرائط.

obeikandl.com

المراجع العربية

1. وليم كوكوران – ترجمة أنيس كنجو – تقنية المعاينة الإحصائية (1995)، جامعة الملك سعود، الرياض.
2. حسين علوان (1994)، طرق المعاينة، دار الفرقان، عمان.

References

1. Barnett, V. (1991). Sampling Survey Principles and Methods, Edward Arnold, London.
2. Cochran, W. G. (1977). Sampling Technique, 3rd Ed. John Wiley & Sons, New York.
3. Fisher, R. A. (1953). The Design of Experiments, Oliver & Boyd, London.
4. Lahiri, D. B. (1951). A Method of Sample Selection Providing Unbiased Ratio Estimates Bull. Int. Statist., 33, 133, 140.
5. Levy, P. S. and Lemeshow, S. (1999). Sampling of Population: Methods and Applications, Wiley, NewYork.
6. Mahalanobis, P. C. (1944). On the Large-Scale Sample Surveys, Phil. Trans. R. Soc., 231, 329-451.
7. Mahalanobis, P. C. (1950). Cost and Accuracy of Results in Sampling and Complete Enumeration, Bull. Int. Statist. Inst., 32, 210-213.
8. Scheaffer, R. L., Mendenhall, W.and Ott, L. (1996). Elementary Survey Sampling, 5th Ed., Duxbury press, Boston.
9. Seal, K. C. (1962). Use of Outdated Frames in Large Sample Surveys. Bull-Cal. Statist Assoc., 11, 68-84.
10. Singh, D. (1978). Taking Agricuitural Censuses, FAO, Rome.
11. Sturat, A. (1984) the Ideas of Sampling, Revised Edition, Griffin, London.
12. Thompson, S. K (2002). Sampling, 2end Wiley, New York.
13. Yates, F. and. Grundy, P. M (1953). Selection without Replacement from Within Strata Probability Proportional to Size J. R. Statist, Soc. 15 B, 253-261.
14. Yates, F. (1960). Sampling Methods for Censuses and Surveys, Charles Grittin & Co., London.
15. Yates, F. (1981). Sampling Methods for Censuses and Surveys 4th Ed., Grittin, London.
16. Zarkovich, S. S. (1961). Sampling Methods and Censuses, FAO, Rome.

obeikandl.com

الفصل الرابع

تنفيذ المسح بالعينة

1.4 مقدمة

بعد أن يكون الهدف أو الأهداف الأساسية من المسح قد حددت، ودرجة الدقة المطلوبة من البيانات كذلك حددت، وجميع الخطوات الأساسية مثل تحديد المجتمع ووحدة المعاينة وحجم العينة وطريقة المعاينة التي سوف تستخدم قد حددت أيضاً، تبدأ مرحلة التنفيذ التي قد تبدأ بتساؤلات منها:

1. كيف نطرح الأسئلة بصورة واضحة وغير مضللة؟
2. كيف نشجع الناس للإجابة عن أسئلتنا بدقة؟
3. كيف نتصل بالذين نرغب في مقابلتهم لجمع البيانات منهم؟
4. كيف نضمن أن مجموعة معينة من الناس والذين لهم مصالح معينة من هذا المسح أن تستجيب للإجابة عن الأسئلة، ولا ترفض بحيث تؤثر على النتائج وتجعلها متحيزة؟
5. كيف نقرر ما هو الصحيح والملائم أو المناسب لنسأل عنه؟
6. كيف نستطيع أن نقيم عملنا بحيث نقول: إننا قد حققنا أهدافنا الإحصائية من المسح وخصوصاً عدم التحيز ودقة المعلومات التي حصلنا عليها؟

نلاحظ أن غالبية التساؤلات أعلاه تدور حول المسح بالعينة والتي تخص معرفة آراء الناس حول قضية أو قضايا تهمهم، ولكن في الحقيقة إن عملية تنفيذ أي مسح بالعينة تسبقها تساؤلات عن كيفية التنفيذ، وقبل أن ننتقل في

obeikandi.com

دراستنا إلى التصميم والتحليل الإحصائي، رأينا أن نتفحص مدى المشكلات الواقعية التي سترافق عملية جمع البيانات وكيف نستطيع أن نتعامل معها ونتغلب عليها.

2.4 اختيار وتدريب العاديّن

سوف لا أدخل في تفصيلات كثيرة في هذا الموضوع، ولكن أحب أن أنوه إليه؛ لأن العاديّن لهم دور مهم في عملية المسح بالعينة، لذلك لابد أن يتم اختيارهم بشكل دقيق، وأن يكونوا ذوي خبرة وذكاء ولديهم القابلية للتعامل مع الناس ولديهم بعض المعلومات عن المبادئ الأولية في الحسابات. وبعد اختيارهم لابد من تدريبهم على عملية المسح وإطلاعهم على أهداف المسح، والقيام بإعدادهم بشكل جيد ليقوموا بعملهم على أفضل وجه. جرت العادة على أن تعقد دورات خاصة لتدريبهم وإعدادهم من قبل الدوائر أو الدائرة المشرفة على المسح، ولمزيد من المعلومات يراجع Raj (1972). ولابد من الإشارة في نهاية الحديث عن هذا الموضوع إلى ضرورة الإشراف ومتابعة العاديّن عند القيام بأعمالهم للتأكد من قيامهم بها وفق ما طلب منهم.

3.4 مصادر التشتت والخطأ

يمكن أن تكون نقطة البداية هنا بالسؤال الآتي: ما الخطأ الذي يمكن أن يحدث؟ ولماذا؟ لنفرض أننا نستطيع أن نتخذ قراراً فيما يخص ميزانية المسح. وقبل أن تبدأ عميلة المسح لا بد من التأكد أن البيانات بالشكل التي حددها التصميم يمكن جمعها، وما هي المصاعب التي يمكن أن تواجهنا في عملية تنفيذ المسح؟ يمكن أن نحصل على معلومات من خلال دراسة واختيار الموقع أو مسح أولي قبل بدء المسح الرئيس كي نستطيع أن نعدل على التصميم ليصبح ممكن التنفيذ، وهنالك مصادر متعددة للخطأ في المعاينة

يمكن أن نلخصها فيما يأتي: أخطاء المعاينة، وأخطاء التغطية أو عدم ضم جميع الوحدات، وأخطاء ناتجة عن عدم الإجابة(أخطاء الاستجابة)، وأخطاء المشاهدة، وسوف نتطرق إلى كل مصدر خطأ مما سبق على حده .

1. خطأ المعاينة

هذا النوع من الخطأ لا يمكن التخلص منه، وذلك لأنه نشأ من اختلاف وحدات المعاينة فيما بينها، فمثلاً إذا كنا نجمع بيانات عن دخل الأسرة نلاحظ أن هناك أسرة دخلها مرتفع جداً وأخرى منخفض، وهكذا لا نستطيع التخلص من هذا النوع من الفروقات، لذلك فإن خطأ المعاينة لا بد منه، ويمكن تقليله عن طريق زيادة حجم العينة أو تقسيم المجتمع إلى طبقات أو استخدام بعض المعلومات الإضافية.

2. أخطاء التغطية أو عدم ضم جميع الوحدات

هذا النوع من الأخطاء يحدث عندما تكون لدينا وحدة من وحدات المجتمع لا يمكن أن تظهر في العينة. مثلاً عندما نجري عملية المسح باستخدام الهاتف فقط فإن كل وحدات المجتمع التي ليس لديها تلفون سوف لا تكون لها فرصة في الظهور في العينة، أو أن نقوم بسحب عينة من مدينة ما في وقت تكون هنالك مباراة في كرة القدم في ملعب المدينة، واضح هنا أن الرياضيين والمشاهدين للمباراة سوف لن تكون لهم فرصة في الظهور في العينة، وهذا النوع من الأخطاء يحدث عندما يكون هناك سوء في المطابقة بين إطار المعاينة والمجتمع الذي نهدف إلى دراسته. وكلما زادت الوحدات التي لا يمكن أن يغطيها المسح زاد التحيز وتعرض البحث أو المسح إلى خطر عدم قبول نتائجه أو التشكيك في دقتها وتمثيلها للمجتمع.

obeikandl.com

3. أخطاء بسبب عدم الإجابة (أخطاء الاستجابة)

كما هو واضح من التسمية فإن عناصر من المجتمع التي ظهرت في العينة لم نستطع الحصول منها على المعلومات أو البيانات المطلوبة، وغالباً ما يكون هناك أكثر من متغير أو صفة يتم مشاهدتها في نفس الوحدة، كما هو الحال بالنسبة إلى الاستبانات التي تحتوي على مجموعة من الأسئلة، فعدم الاستجابة قد يتم مع مجموعة من الأسئلة حيث تترك دون جواب، كما هو الحال بالنسبة إلى الاستبانات التي ترسل بواسطة البريد، فمثلاً قد يرفض الشخص الإفصاح عن دخله. إن سبب عدم الاستجابة يحدث لأسباب مختلفة منها: قد يكون لطبيعة المعلومات المطلوبة حقائق أو (رأي شخصي)، كذلك قد يكون السبب هو الطريقة أو حتى الوقت الذي نطلب فيه جمع المعلومات، أو قد تكون هذه العوامل أو الأسباب متداخلة أو مترابطة مع بعضها.

غالباً ما يعد عدم الاستجابة الكلية مقياساً على نجاح أو عدم نجاح المسح، ربما يكون عدم الاستجابة سببها هو القرارات غير الصائبة عن كيف ومتى وأين وماذا نجمع، وما هي المعلومات المطلوبة. لوحظ أن موضوعات الاستبانات المختلفة والطرق المختلفة لجمع البيانات تسبب مستويات مختلفة من عدم الاستجابة، إن فقدان جزء من العينة بسبب عدم الاستجابة سيؤدي إلى تضخم التباين للتقديرات المختلفة، ولكن درجة تأثيرها على التحيز يعتمد على نوع المجموعة التي لم تستجب هل هي تمثل جزءاً معيناً من المجتمع أم لا؟. فإذا كانت تمثل رأي جزء من المجتمع فإن مقدار التحيز سوف يتضخم، إذا كان المتغير أو الصفة التي نجمع عنها المعلومات لها علاقة قوية مع عدم الاستجابة (مثلاً الذين يتقاضون رواتب كبيرة يكونون أكثر تحفظاً في الإجابة عن أسئلة تخص مدخولاتهم) نتوقع أن يكون التحيز كبيراً وخطيراً.

إن من أصعب أنواع عدم الاستجابة التي يصعب التعامل معها هو الشخصي منها والذي سببه رفض التعاون مع العادين. وهذا النوع يتأثر بصورة خاصة بالطريقة التي تستخدم (مثلاً مقابلة شخصية، استعمال البريد ...إلخ) لجمع المعلومات. إن الفشل في إيجاد عنصر من عناصر العينة يعد أحد مصادر عدم الاستجابة، ويمكن علاج هذا النوع من عدم الاستجابة بمحاولة إيجاد هذا العنصر في محاولة لاحقة.

4. أخطاء المشاهدة

يمكن أن نعرف خطأ المشاهدة: بأنه الحصول على معلومات خاطئة من وحدات المعاينة كما يمكن أن تحدث أخطاء المشاهدة من خلال طرق مختلفة:

1. خطأ المقابلة أو خطأ السؤال: عندما يوجه السؤال بطريقة خاطئة أو بطريقة غير واضحة تؤدي إلى أن تكون الإجابة خاطئة.
2. خطأ التسجيل: تكون الإجابة صحيحة ولكنها تسجل بطريقة خاطئة.
3. خطأ النقل أو الإدخال: عندما تدخل المعلومات في الحساب الآلي أو قاعدة البيانات يحدث خطأ ما في أثناء ذلك.

هذه الأخطاء ليس لها علاقة بوحدة المعاينة أو عنصر العينة وقد تكون الإجابة خاطئة عن سؤال معين حتى ولو كان هذا السؤال قد طرح بشكل واضح ودقيق، وأسباب ذلك تعود إلى أنّ السؤال له علاقة بموضوع حساس، أو له تأثير نفسي على الشخص، أو لأنه احتوى على تفاصيل كثيرة، أو لأنه قد يجرم أو يلقي التهمة على الشخص.

في نهاية المطاف لا بد من أن نعرج قليلاً على أخطاء سببها القياسات، وذلك بأن تكون أجهزة القياس غير صالحة أو لا تقرأ القياسات بشكل

obeikandl.com

صحيح، أو لا يحسن الباحث استخدام الأجهزة، هذه أمور مقدور عليها. ولكن تبقى هناك أخطاء عندما توجد قيمة معينة يصعب الحصول عليها من غير أخطاء، فمثلاً عدد نبضات المريض، لنفرض إننا قسنا عدد النبضات فوجدناها 78 فهل يعني هذا بالضرورة أن نبض المريض هو 78؟ فقد يكون الجواب لا لأسباب عديدة منها اختلاف نبض المريض من وقت لآخر تبعاً لحالته الصحية.

4.4 المسح الأولي

يمكن التغلب على كثير من الصعوبات التي تواجهنا في عملية تصميم وتنفيذ المسح بالعينة وحتى المسح الشامل، وذلك من خلال إجراء مسح أولي لبعض عناصر المجتمع قبل أن نقوم بتنفيذ المسح الرئيس للمجتمع. إن المعلومات التي سوف نحصل عليها بهذه الطريقة، يمكن أن تساعدنا لتحديد كثير من العوامل الحرجة التي منها:

1. أخطاء القياسات الممكنة الحدوث.
2. معدل أو نسبة عدم الاستجابة.
3. المواضيع الحساسة أو مصادر عدم الوضوح.
4. مدى الاختلاف بين العادّين.
5. صعوبة الوصول إلى بعض وحدات العينة.
6. مدى التشتت للمتغير الرئيس أو الثانوي.

لا توجد هناك وصفة جاهزة وبسيطة يمكن إعطاؤها لتحديد مقدار المعلومات التي نحتاجها، أو الطريقة التي نستخدمها للحصول على هذه المعلومات الأولية، ويجمع الجميع على ضرورة الحصول على هذه المعلومات والحاجة إليها، ولكن مقدارها يعتمد على مدى وسعة المشكلات الأساسية

التي تحتاج إلى معلومات أولية لها، كذلك تعتمد على مقدار الوقت والمال المتوفرين لها.

نستخدم المعلومات في المسح الأولي في العادة لتعديل وتحسين تصميم المعاينة (على سبيل المثال اختيار حجم العينة)، وتستخدم لتعديل طريقة البحث وطريقة العمل (على سبيل المثال تعديل الأسئلة المطروحة في الاستبانة)، وتعديل التعليمات المعطاة إلى العادّين، والسماح بطرق مختلفة للمتابعة والإشراف، ولا تدخل هذه البيانات التي نحصل عليها هنا ضمن المسح الرئيس فيما عدا بعض طرق المعاينة كالمعاينة المزدوجة أو المعاينة المتكررة.

إن الأداة الرئيسة في عمليات المسح الأولي هي اختبار الموقع Pilot-Survey ويمكن أن يأخذ عدة أشكال، وذلك يعتمد على الغرض من المعاينة. يمكن أن تقتصر الدراسة على أعداد محدودة جداً من عناصر المجتمع الذين سنجرب معهم الطرق المختلفة للحصول على المعلومات المطلوبة، وهذا النوع مهم جداً ويستخدم كثيراً إذا كان موضوع المسح يتعلق بأمور حساسة أو شخصية أو عندما يكون الموضوع معقداً، ومن المحتمل أن يساء فهم الدراسة. إن إجراء مقابلة واسعة وشاملة وعميقة مع مجموعة من عناصر المجتمع يمكن أن يزودنا بمعلومات أولية مفيدة، يمكن الاعتماد عليها لتعديل تصميم المسح أو الأسئلة المطروحة، أو الطريقة التي تتبع للحصول على المعلومات، حيث إن الأسئلة وطريقة المعاينة وكيفية الوصول إلى الوحدات يمكن أن تدرس وتهذب باستخدام ما يسمى بدراسة الموقع، أو اختبار الموقع لجوانب مختلفة من المجتمع.

يمكن أن نقوم بدراسة الموقع أو المسح الأولي من خلال سحب عينة عشوائية مقبولة الحجم قد يكون حجمها بين 50 إلى 500 وحدة، وغالباً ما يكون هذا المسح الأولي مهماً جداً لنجاح المسح بالعينة الرئيس. ويستخدم هذا

المسح الأولي لتقدير نسبة عدم الاستجابة، كذلك للتخلص من أخطاء الاستجابة المحتملة وبطبيعة الحال سوف يساعدنا في تحديد حجم العينة للمسح الرئيس بشكل أكثر دقة، وأيضًا في تحديد عملية المتابعة، ومن الممكن أن نحصل على الخطوط العريضة للطرق التي يمكن لخطأ الاستجابة أن يحصل بها، وعلاقته بأفراد المجتمع الذي سيؤدي إلى تحيز في التقدير بشكل خاص. فإذا كان عدم الاستجابة أو الإجابة الخاطئة لها علاقة بجزء أو أجزاء من أفراد المجتمع فيمكن أن نستخدم في تقدير التكاليف لأنواع مختلفة من طرق جمع البيانات، وكذلك تقدير تكاليف المسح الرئيس، وأخيراً يمكن استخدام المسح الأولي للحصول على تقديرات أولية لبعض صفات المجتمع، أو متغيرات الدراسة.

إن النقطة الأخيرة هنا في جميع المسوحات سواء كانت أولية أو رئيسة لابد أن نقوم بدراسة ومعرفة كيفية الوصول إلى وحدات المجتمع من قبل العادّين، وهل ستكون الطرق ملائمة، وهل توجد وسائط نقل لنقلهم إلى هذه الوحدات؟ وهل يحتاج العادون إلى إشراف مستمر؟ وهل توجد أجزاء من المجتمع لا يمكن الوصول إليها؟

5.4 طرق جمع البيانات

هنالك عنصران رئيسان في عملية التخطيط للمسح بالعينة، الأول هو اختيار طريقة المعاينة، هل نختار العينة العشوائية البسيطة أو إحدى طرق المعاينة الأخرى التي تكون أكثر تعقيداً، أما الثاني فهو كيف سنقوم بتنفيذ المسح بالعينة بمعنى جمع البيانات وفقاً لطريقة المعاينة التي اخترناها؟ هنالك طرق عديدة لجمع البيانات التي يمكن أن نستخدمها بالاعتماد على طبيعة المجتمع وموضوع البحث أو المسح. سوف نتكلم عن أهم هذه الطرق وهي استخدام المعلومات المتوافرة، وطريقة المشاهدة، والمقابلة وجهاً لوجه،

واستخدام البريد، وأخيراً استخدام الهاتف. وسوف نتطرق إلى كل طريقة من الطرق السابقة على حدة.

1. استخدام المعلومات المتوافرة

أصبح من متطلبات الحياة العصرية اليوم أن نقوم بجمع معلومات عن كل شيء بما فيها المعلومات في كل المجالات، الشخصية، والعلمية والاجتماعية، هذه المعلومات مدونة، ويمكن أن تكون متوافرة على مستويات مختلفة من التفصيل وشروط مختلفة للاستعمال. لقد تم جمع هذه البيانات لأغراض إدارية وحكومية، فمثلاً كشف بأسماء وعناوين كل المشتركين مع دوائر الماء، والكهرباء وأسماء المحلات التجارية والشركات المحلية المختلفة والمسوحات الشاملة التي تقوم بها الدولة كل عشر سنين، وتفاصيل عن كل المركبات المرخصة ... إلخ، وكذلك معلومات أكثر دقة وأكثر شخصية يحتفظ فيها في المستشفيات ودوائر الشرطة والضريبة... الخ، وكذلك تحتفظ هيئات مختلفة بمعلومات عن أعضائها ابتداء بالجمعيات الاجتماعية وانتهاء بالمحامين. وأخيراً يحتفظ الأفراد بمعلومات شخصية عنهم وعن أسرهم.

من الناحية المبدئية كل هذه المعلومات يمكن أن تكون مفيدة ولكن استخداماتها محدودة، وذلك لعدة أسباب:

1. قد لا تكون متطابقة مع المجتمع الذي نهدف إلى دراسته.
2. يمكن أن تكون قديمة.
3. يمكن أن تكون مجمعة مع بعضها بشكل يجعلها غير نافعة.
4. قد لا يمكن الحصول عليها لأسباب قانونية.
5. قد لا تتعاون الجمعيات أو الهيئات مع الباحثين لإعطاء معلومات عن أعضائها بالتفصيل الذي يحتاجونه.

opelkanal.com

تعد قيمة المعلومات المتوافرة في السجلات محددة على مستوى وحدة المعاينة ما عدا بعض التفصيلات المحدودة، مثل الجنس والعنوان. بعض الأحيان تعد مفيدة إذا أريد إجراء مسوحات باستخدام العينة داخلياً أي ضمن الدوائر والمؤسسات التي تحتفظ بهذه المعلومات، والتي يمكن أن تزودنا بإطار المعاينة، الذي يعد هو الأساس للمسح بالعينة.

2. طريقة المشاهدة

إنه من الطبيعي أن نحصل على معلومات بالملاحظة أو المشاهدة، والذي يجري من دون الحاجة إلى اتصال مباشر بالوحدات. هذا يعني أننا ننظر ولا نسأل. وهذا الأسلوب شائع بين علماء الطبيعة، ولكن قد تكون هذه الطريقة لا تستخدم بشكل منتظم للحصول على بيانات باستخدام المسح بالعينة، ولكن سوف نتكلم عن هذه الطريقة باختصار آخذين في الحسبان المزايا والعيوب لهذه الطريقة.

أما المزايا فتتضمن الموضوعية والدقة وتجنب خطأ الاستجابة. حتى في الاستبانات الاجتماعية نتجنب ردود فعل الناس حول موضوع أو حالة جديدة وقد تكون الملاحظة المباشرة مفيدة، فعلى سبيل المثال في مسح السوق يمكن أن تقوم شركة ما أنتجت بضاعة جديدة بوضعها في منطقة تجريبية، وملاحظة ردود فعل الزبائن حول البضاعة الجديدة من مناطق معينة بدلاً من مقابلة الزبائن وسؤالهم لمعرفة رأيهم في البضاعة الجديدة، والذين قد يتحرجون من الإجابة. إن البيانات التي تجمع بطريقة الملاحظة المباشرة تتحاشى أو تتخلص من مشكلة عدم المعرفة أو عدم فهم السؤال، وفي بعض الأحيان فإن طريقة الملاحظة أو المشاهدة المباشرة لا يوجد بديل عنها، أو هي الأسلوب الوحيد الذي يمكن أن يستخدم، فمثلاً في حالة حشرات تطفو على

سطح ماء النهر لا بد من الذهاب ومشاهدتها، أو حيوانات برية نريد جمع معلومات عنها، أو أسماك في بحيرة، أو ... إلخ.

أما عيوب هذه الطريقة فأهمها: أنها تحتاج إلى وقت كثير وتكاليف عالية، وقد تؤثر في التصرفات الطبيعية لبعض المخلوقات التي نراقبها لجمع معلومات عنها، وقد تكون غير ممكنة التطبيق، فمثلاً لا يمكن أن نتصور أننا نجلس ونراقب أفراداً معينين لمعرفة آرائهم في قضية معينة.

3. المقابلة وجهاً لوجه

تعد المقابلة وجهاً لوجه من الوسائل الشائعة الاستعمال لجمع البيانات، وبصورة خاصة استطلاعات الرأي، حيث يخرج العادون إلى المجتمع مزودين باستبانة مكونة من مجموعة من الأسئلة، ويقومون بمقابلة الأشخاص الذين ظهرت أسماؤهم بالعينة، ويجمعون المعلومات المطلوبة منهم. ولكي يكون المسح ناجحاً يتطلب إعداداً متقناً وتصميماً جيداً لأسئلة الاستبانة والعمل الميداني، وتقيم المسح من خلال المعاينة الأولية، وكذلك إعداد وتدريب العادين والمحافظة على مستواهم، وكذلك التأكيد على مدى قابليتهم لإجراء المقابلات.

إن من أهم مزايا هذا الأسلوب من جمع المعلومات أنه يقلل نسبة سوء الفهم لأسئلة الاستبانة، وكذلك يشجع على زيادة نسبة الاستجابة، وتقليل خطأ عدم الاستجابة. أما الصعوبات التي تواجهنا باستخدام المقابلة فأهمها عدم استجابة الناس للمقابلة، وكذلك يستخدم بعض أفراد العينة أسلوب الإجابة بنعم، أو الموافقة على كل ما يطرح عليهم من أسئلة، وقد يكون هنالك سؤال أو أسئلة مقبولة في المحادثات الخاصة ومرفوضة عندما تكون بصورة رسمية لجمع البيانات، وكذلك هنالك حالات لإعطاء إجابة خاطئة عن قصد أو إعطاء إجابة يعتقد أنها هي المطلوبة. إن الحالة النفسية للعادّين تُعد

من الأمور المعقدة، حيث إنه قد يكون لها تأثيرات كبيرة ومختلفة على جمع البيانات المطلوبة. إن المقابلة الشخصية أو ما نسميها المقابلة وجهاً لوجه مكلفة جداً بالقياس إلى الطرق الأخرى كاستخدام البريد أو الهاتف أو جمع المعلومات من السجلات، ولكن المميزات الكثيرة لهذا الأسلوب والتي من أهمها تقليل نسبة عدم الاستجابة وسوء الفهم أو سوء تفسير السؤال وربما زيادة سرعة الحصول على المعلومات (لأن السرعة قد تكون مطلوبة في حالات استطلاعات الرأي أو أبحاث السوق) يجعل هذه الطريقة شائعة الاستخدام.

لا بد من التأكيد هنا على أن المقابلة الشخصية لا تعني الوقوف على مفترق الطرق لمقابلة الناس، ولكنها في الغالب تجري من خلال تصميم المسح الذي يتضمن اختيار وحدات المعاينة بطريقة عشوائية من المجتمع باستخدام إحدى طرق العشوائية المعروفة، وتحديد أين نجد هؤلاء الأشخاص لمقابلتهم وكيفية الوصول إليهم، وهذا يتضمن تحديد طريقة للوصول إليهم ووسائل النقل التي تستخدم، وكذلك طريقة للتعامل مع الأشخاص الذين لا نجدهم وقت المقابلة، في حالة عدم الاستجابة فإن عملية إعادة المحاولة لمقابلتهم تكون مكلفة.

والمشكلة الأخيرة التي تواجهنا في هذا الأسلوب هي دور العاد نفسه في التأثير على إجابة الأشخاص، وذلك من خلال توضيح الأسئلة أو التعليق على بعض الأمور المطروحة، ومن الناحية النظرية يجب أن يطرح العاد السؤال بموضوعية وبدون أي تحيز، ويمكن أن يعلق على موضوعات جانبية لشرح السؤال إذا لزم الأمر ذلك، وهذا النوع من الخطأ يسمى خطأ العادين، ويمكن أن يكون هنالك تحيز بسبب الاستجابة من قبل الشخص للعاد، وقد يكون هنالك خطر عدم فهم الجواب من قبل العاد، ويمكن التغلب على هذه المشكلة من خلال تدريب ومتابعة العادين.

4. استخدام البريد

إن البديل الواضح والقريب لأسلوب المقابلة هو استخدام البريد، وذلك عن طريق إرسال الاستبانات إلى الأشخاص بالبريد، والطلب منهم الإجابة عنها، وإعادتها إلى الجهة التي أرسلتها لهم، هنالك طريقة مماثلة، وهي أن توزع الاستبانات من قبل العادين على أعضاء العينة وتركها معهم والطلب منهم إعادتها بواسطة البريد، إن توزيع الاستبانات من قبل العادين له ميزات أهمها أن الاتصال الشخصي قد يحث الأشخاص على ملء هذه الاستبانات، وتعطى الفرصة لتوضيح بعض الصعوبات، ويمكن أن ترفع من نسبة الاستجابة، ولكنها مكلفة كما هو الحال بالنسبة إلى المقابلة الشخصية، ولكنها أحيانا تكون أقل كلفة.

عند استخدام طريقة المسح بالبريد يجب أن نراعي أن هناك رغبة لدى الناس لرمي هذا النوع من الرسائل عند استلامها؛ لذلك لا بد من بذل الجهود لمنع مثل هذه الظاهرة كأن تكتب رسالة معدة بشكل جيد ومصممة بعناية مع الاستبانة، وأن يرسل مغلف بريدي عليه طابع وعنوان لإعادة الاستبانة بعد ملئها إلى الجهة التي أرسلتها؛ لأنه من الصعوبة أن ترمي رسالة معنونة عليها طوابع غير مستخدمة.

هنالك مسوحات تقوم بها دوائر الدولة وبعض الشركات الخاصة، وهذا النوع من المسوحات لا يحتاج إلى وسائل لحث عناصر العينة على الرد وخصوصاً إذا كان يجري على أعضاء هذه المؤسسات أو الدوائر؛ لأنه من السهولة أن تطلب هذه الدوائر من العاملين فيها ضرورة ملء هذه الاستبانات وإعادتها إلى الجهة التي طلبتها. ولكن لا يمكن أن نتصور أن هذه الدوائر سوف تحصل على رد كامل وخصوصاً المسوحات التي تقوم بها دوائر الدولة.

تعاني المسوحات في البريد من انخفاض نسبة الاستجابة أو الرد على الرسائل، وقد تصل إلى أقل من 50% لذلك يحتاج الباحثون إلى إرسال رسائل أخرى للتأكيد على ضرورة ملء الاستبانات التي أرسلت إليهم وإعادتها، وقد يتطلب هذا الأمر أكثر من تأكيد، وفي الغالب لا بد أن ترسل مع هذه التأكيدات استبانات جديدة ومغلفات عليها طوابع وعنوان لإعادتها إلى الجهة التي أرسلتها. فوائد أخرى للمسح بالبريد بالإضافة إلى تقليل التكاليف أيضاً التخلص من خطأ العادّين، وإعطاء فرصة لتغطية موضوعات متفرقة بصورة أعمق وأوسع. ولكنها قد تستغرق عدة شهور.

5. استخدام الهاتف

أصبح الهاتف مع تطور الاتصالات موجوداً بشكل كبير في غالبية المنازل وفي متناول الجميع، لذلك ليست فكرة سيئة أن تحاول تنفيذ المسح باستخدام الهاتف والاتصال بالأشخاص أو المؤسسات لجمع البيانات المطلوبة، وفكرة استخدام الهاتف ليست بالجديدة، فلقد استخدمت في الولايات المتحدة عام 1934 لاستطلاع أراء الناس حول بعض الأمور السياسية، ولكنها فشلت لأن حوالي 60% من المنازل في ذلك الوقت لا يوجد فيها تلفون، وهذا واضح لنا لأن المجتمع لم يغط كاملاً، وهنالك نسبة كبيرة من المجتمع لم تكن لها فرصة في الظهور في العينة، وخصوصاً ذوو الدخول المحدودة والذين لا يمتلكون الهاتف. ولكن الأمر اختلف الآن حيث إن أكثر من 90% من المنازل تحتوي على هواتف في الدول المتقدمة وحتى في بعض الدول النامية.

هناك مزايا واضحة لاستخدام الهاتف منها السرعة، وقلة التكاليف، ولكنها لا تخلو من العيوب، وأهمها فيما يتعلق بمسألة تغطية المجتمع، فحتى مع توافر تلفون في 90% من المنازل فإن العينة لا تكون عشوائية، لأن المنازل التي تحتوي على هواتف لها آراء اجتماعية واقتصادية وسياسة قد تكون

مختلفة. لذلك نواجه هنا مشكلة التحيز بسبب عدم تغطية المجتمع كاملاً، ولا بد من تقييمها ومعرفة تأثرها في كل مسح على انفراد.

جميع الصعوبات التي واجهناها في المقابلة الشخصية سوف نواجهها باستخدام الهاتف ولا توجد حاجة إلى إعادتها بالإضافة إلى أن الناس يبدون رفضاً أكبر لاستخدام الهاتف في جمع المعلومات، كما يعتبر البعض هذا تدخلاً في منازلهم أو حياتهم الخاصة سببه الهاتف، لذلك لا بد من استخدام بعض المهارات من قبل العاد الذي يستخدم الهاتف ليتجاوز هذه الصعوبة. على الرغم من المشكلات التي تعزى لهذه الطريقة في جمع المعلومات لكنها تستخدم بصورة واسعة وذلك لسهولتها وقلة تكاليفها وسرعتها.

6. استخدام البريد الإلكتروني

للقيام بالمسح بالعينة باستخدام البريد الإلكتروني، نقوم بتجهيز كشف بعناوين البريد الإلكتروني للأشخاص الذين نرغب في الحصول على معلومات منهم، وهذا ما يسمى بإطار المعاينة، ومن ثم نقوم بسحب عينة عشوائية من بين العناوين الإلكترونية الموجودة في الكشف. ومن ثم نقوم بإرسال رسالة إلى العناوين الإلكترونية التي سحبت بالعينة تحتوي على الاستبانة. هنالك طريقتان لإرسال الاستبانة إما أن ترسل كملحق، أو نضعها ضمن الرسالة الإلكترونية. ونطلب من الشخص أن يجيب عن الأسئلة ومن ثم يقوم بإعادة إرسال الاستبانة إلى نفس العنوان أو إلى عنوان آخر نحدده في نهاية الرسالة. لابد أن هناك رسالة قصيرة تكون ضمن الرسالة الإلكترونية، نوضح للمستجيب في هذه الرسالة من نحن، وما هي أهداف البحث باختصار شديد، مع شرح مبسط لكيفية تعبئة الاستبانة. ولابد أن نؤكد له أن هذه المعلومات سوف تستخدم لأغراض البحث العلمي، أو لمساعدة الإدارات المعنية في اتخاذ القرارات المناسبة، وأن هذه المعلومات لن تستخدم ضده بأي حال من الأحوال.

obeikandl.com

هناك مجموعة من المصاعب والمشكلات التي تعتري هذا النوع من طرق جمع البيانات نجملها فيما يأتي:

1. هناك صعوبات ومشكلات فنية وتقنية تتعلق بالحاسب الآلي واستخداماته والبرامج المستخدمة مع كل نظام للبريد الإلكتروني وغيرها، فمثلاً في كثير من الأحيان لا نستطيع فتح ملحق معين لسبب أو لآخر، لا أريد الخوض كثيراً في هذا الموضع لأنني لست من المتخصصين في هذا المجال. ولكن لنفترض أن جميع هذه المشكلات تخلصنا منها، وأن المستجيب يستلم الرسالة الإلكترونية ويستطيع الإجابة عنها إلكترونيا.
2. جميع الصعوبات والمشكلات التي واجهتنا في حالة استخدام الهاتف والبريد العادي ستكون منطبقة هنا، فعلى سبيل المثال رفض الشخص للإجابة عن الأسئلة أو تعبئة الاستبانة أو انخفاض عدد المستجيبين (أي نسبة الإجابة قليلة).
3. ينفرد استخدام البريد الإلكتروني ببعض المشكلات الخاصة به منها:
   أ. أن كثيراً من الناس إلى الآن لا يستخدمون الحاسب الآلي ناهيك عن استعمال البريد الإلكتروني.
   ب. غالباً ما يرفض الأشخاص فتح أي رسالة لا يعرفون مرسلها، ويقومون بحذفها، إما لضيق الوقت أو ظناً منهم أنها رسائل تعلن عن أشياء لا رغبة لهم بها.
   ت. من أهم المشكلات التي تواجه استخدام البريد الإلكتروني، هو انتشار الفيروسات عبر البريد الإلكتروني، وما لها من تأثيرات كبيرة على أجهزة الحاسب؛ لذا غالباً ما يقوم الشخص بحذف أي رسالة بريدية لا يعرف مرسلها خصوصاً تلك التي تحمل ملحقاً.

obeikandi.com

ث. قد لا يفتح الشخص بريده الإلكتروني لأيام عديدة، أو لديه كم هائل من الرسائل بحيث لا يستطيع قراءتها، وغالباً ما يقوم بحذف كثيرٍ من الرسائل خصوصاً تلك التي لا يعرف مرسلها.

ج. من الأخطاء الشائعة في استخدام البريد الإلكتروني أن يقوم الباحثون بإرسال الاستبانة إلى عدد كبير من الأشخاص، أو ربما إلى جميع العناوين الموجودة في الكشف، وهذا سيؤدي إلى أن أعداد غير المستجيبين ستكون كبيرة جداً حتى ولو أرسلنا لهم الاستبانة مرات عديدة.

إن من أهم مزايا استخدام البريد الإلكتروني هي: السرعة، وانخفاض التكاليف وسهولة التنفيذ، وإمكانية الرجوع إلى الوحدة مرات عديدة دون أي تكاليف إضافية، وإمكانية إرسال الرسالة مرة أخرى.

7. استخدام الشبكة العنكبوتية (Internet)

كثيراً ما نرى استبانات أو استطلاعات رأي موجودة في كثير من المواقع على الشبكة العنكبوتية (Internet)، هذه الاستطلاعات لا تمثل سوى رأي المستجيبين ولا يمكن أن نعتبرها مسحاً علمياً يمكن تعميم نتائجه على المجتمع الذي نرغب بمعرفة آراء أفراده أو دراسته.

يمكن الجمع بين طريقة البريد العادي أو الهاتف والشبكة العنكبوتية بأن نقوم بإرسال رسائل إلى الأشخاص الذين سحبوا في العينة أو الاتصال بهم هاتفياً، ونطلب منهم الدخول إلى الموقع في الشبكة العنكبوتية لتعبئة الاستبانة. لابد أن يكون الدخول إلى هذا الموقع أو الجزء الذي يحوي الاستبانة محدوداً بحيث يقتصر على الذين خرجت أسماؤهم في العينة. ويمكننا القيام بذلك بأن نعطي لكل شخص كلمة مرور خاصة به.

Obeikandl.com

جميع العيوب التي تواجه طريقة البريد العادي والهاتف تنسحب على استخدام الشبكة العنكبوتية، يضاف إليها عدم معرفة استخدام الحاسب ومن ثم الدخول إلى الموقع على الشبكة على افتراض أن جميع المشكلات الفنية والتقنية المتعلقة بتعبئة الاستبانة وإرسالها جرى حلها بشكل أو بآخر.

لمزيد من المعلومات حول طرق استخدام البريد الإلكتروني والشبكة العنكبوتية يراجع Malhotra (2006) و(2006) Zikmund and Pabin.

6.4 الاستبانة

واضح من خلال مناقشاتنا السابقة أن نجاح أي مسح يتطلب موضوعية في الإجابات من قبل الأشخاص والمؤسسات التي وقع عليهم الاختيار في العينة. وهذا يعتمد بصورة مباشرة على مهارات مصممي الاستبانة وهناك عوامل كثيرة ممكن أن تؤثر في الإجابات عن أسئلة في الاستبانة يتقدمها فهم السؤال، لذلك لا بد أن يكون السؤال واضحاً ودقيقاً، والإجابات المختلفة للسؤال الواحد تلك التي اختيرت أيضا واضحة ومحددة ومختارة بعناية، فمن السهولة قول شيء ولكن من الصعوبة فعله، سوء الفهم قد يأخذ أشكالاً متعددة، فقد يكون السؤال فنياً، أو الموضوع غير مألوف، أو الكلمات غير معروفة. ولنأخذ هذا السؤال:

جهود الحكومة لحفظ التضخم لم تكن ناجحة إلى الآن: يجب أن تعمل شيئاً ما فهل أنت؟

موافق بشدة	موافق	غير موافق	غير موافق بشدة

توجد مشكلات كثيرة في هذا السؤال أولها قد يكون الفرد المسؤول لا يعرف معنى كلمة تضخم، طرح السؤال وجهة نظر سياسة قد لا يريد الفرد أن يقحم نفسه فيها. طرح السؤال بهذه الصيغة قد يفهم منه وقف العمل

بالجهود الحالية أو الاستمرار بالجهود الحالية مع محاولة تعديلها أو تسريعها، وكذلك اختيارات الإجابة عن هذا السؤال لم تعط مجالاً لأن يكون الشخص محايداً أو لا رأي له، وترتيب الإجابات من غير موافق بشدة إلى موافق بشدة قد يكون له تأثير على اختيار الإجابة، وماذا عن مساحة المربعات المخصصة للإجابة حيث إنها غير متساوية المساحة، وقد تؤثر في اختيار الإجابة.

توجد عوامل كثيرة تؤثر في الإجابة منها: أن الأفراد قد يحاولون الإجابة لحماية أنفسهم من أذى قد يصيبهم، أو محاولة الإجابة بشكل يفهم منه أكثر من إجابة. أو أنهم يجيبون لتوقعهم أن السائل يريد منهم إجابة معينة، لذلك لا بد من صياغة الأسئلة والإجابات بشكل يقلل كثيراً من هذه الصعوبات.

جانب آخر هو الأسئلة الشخصية أو الحساسة التي قد لا يرغب الفرد في الإجابة عنها، لأسباب كثيرة فمثلاً إذا سئل عن دخله قد يخشى أن نفرض عليه ضريبة أو قد يخشى الحسد، كثيرٌ من الفلاحين إذا سألناهم عن مقدار الحليب الذي يحصل عليه من بقراته، قد لا يعطينا الإجابة الصحيحة، وللتغلب على هذه المشكلات فيما يخص هذا النوع من الأسئلة يعمد إلى أسلوب الأسئلة غير المباشرة فبدلاً من أن نسأل الشخص عن دخله الشهري فمن الممكن أن نسأله عن نوع المنزل الذي يسكنه، وعن المنطقة التي هو ساكن فيها وعن قراءته وعاداته. إن هذه الأمور في مجملها يمكن أن تعطينا فكرة عن دخله، ويمكن أن نستعمل أسلوب الإجابة العشوائية على السؤال فمثلاً لتقدير معدل عدد السرقات بالنسبة لعدد أفراد المجتمع يمكن أن نسأل: هل ولدت في شهر معين أو سجنت في السنة الماضية بسبب السرقة، بالإضافة إلى هذه الأمور التي تكلمنا عنها، وهناك عوامل أخرى يمكن أن تؤثر في الإجابات ومنها: موضع السؤال في الاستبانة، وعلاقته بالأسئلة

الأخرى، والكلمات التي صيغ بها السؤال، واهتمام الأشخاص بالموضوع المطروح، والحاجة إلى الذاكرة في الإجابة عن السؤال.

إن تصميم الاستبانة موضوع مهم جداً وله تأثير كبير في نتائج المسح، لذا يجب أن يعطى أهمية كبيرة. لمزيد من المعلومات راجع: Grover (1989) و Schuman and Presser (1981) و Homville and Jowell (1978) و Gillham (2000) و Moser and Kalton (1986) و Robson (1993).

7.4 الخطوات الأساسية لتجهيز البيانات

تجميع البيانات وفقاً لطريقة المعاينة المناسبة التي اختيرت لهذا الغرض يُعدُّ أمراً جيداً. ولكن السؤال الآن ماذا نعمل بهذه البيانات؟ الجواب الأولي نستعملها لتقدير معلمات المجتمع مع تقدير دقتها وإجراء بعض الاختبارات الإحصائية، ولكن قبل الانتقال إلى هذه المرحلة هناك بعض القرارات الأساسية عن تنظيم وتسجيل البيانات لا بد من اتخاذها ليتسنى تحليلها، مع التأكيد على الحاجة إلى الحاسب الآلي والحقائب الإحصائية الجاهزة، التي يمكن استخدامها لأغراض تنظيم وتسجيل وتحليل البيانات، فاستخدام الحاسب في أيامنا هذه أصبح ضرورة قصوى في غالبية أعمالنا، ومنها استخدام الحاسب لحفظ البيانات وتحليلها، وفي الحقيقة فإن الكثير من المسوحات لا يمكن إنجازها بدون الحاسب، وبالنسبة إلى تجهيز البيانات فيمكن أن تمر بمراحل هي: تدقيق البيانات، وتجهيزها لكي تدخل في الحاسب، ومن ثم تحليلها وعرضها، ولاستخدام الحاسب أصبح ضرورياً أن ترقم الإجابات، وذلك حتى لا نحتاج إلى طاقة تخزينية كبيرة. فمثلاً السؤال عن الجنس وجوابه ذكر أو أنثى يمكن أن نستخدم 0 و1، هذه الأرقام التي استخدمت لا معنى لها سوى أن الصفر يرمز إلى الذكر والواحد يرمز إلى

الأنثى، ويمكن أن نستخدم أي رقمين آخرين. كذلك السؤال حول المستوى التعليمي الذي وصل إليه الشخص بحيث يمكن أن يكون جوابه.

أمي
يقرأ ويكتب
حاصل على التوجيهي
حاصل على درجة جامعية
دراسات عليا

ويمكن أن ترمز إلى هذه الإجابات المختلفة 4,3,2,1,0 يبدو أن الترقيم هنا له معنى وهو ترتيب طبيعي لمستوى التعليم حيث الصفر يرمز إلى " أمي" و1 يرمز إلى "يقرأ ويكتب" وهكذا.

نقوم بترقيم أو ترميز جميع إجابات الأسئلة الواردة في الاستبانة. وتكون عملية الترقيم هذه هي البداية في تحليل البيانات وهي في الغالب كذلك. لكن الخطوة التي تأتي مباشرة بعد الحصول على البيانات هي عملية تدقيق البيانات، والتأكد من صحة الإجابات، ومحاولة التخلص أو تصحيح بعض الأرقام أو تعديلها، فمثلاً من غير المعقول أن يكون عمر شخص 500 سنة، على الأرجح إن عمره 50 سنة، وبعد التأكد من سلامة البيانات من الأخطاء تبدأ عملية تجهيز البيانات لإدخالها إلى الحاسب الآلي وعملية التجهيز هذه تعد ضرورة لضمان سير عمليات الإدخال بصورة سريعة ودقيقة وبدون أخطاء، ولكن أعمال التدقيق لا تقتصر على عملية التدقيق قبل إدخال البيانات في الحاسب، ولكن تستمر بعد الإدخال للتأكد من سلامة البيانات من أخطاء الإدخال أو حتى بعض الأخطاء التي لم تستطع اكتشافها قبل إدخال البيانات في الحاسب. ويمكن القيام ببعض هذه الأعمال باستخدام الحاسب.

obeikandi.com

لقد وجهت عناية خاصة في السنوات الأخيرة للتعامل مع الإجابات الخاطئة وذلك إحصائياً، ومن خلال محاولة توقُّع الإجابة الممكنة، ومن ثم محاولة مقارنة الأرقام أو المعلومات الموجودة باستخدام الحاسب إحصائياً أو روتينياً للتأكد من صحة الإجابات. وأخيراً لا بد من الإشارة هنا إلى أن أي تحليل تفصيلي للبيانات يحتاج إلى وضع البيانات في جداول بسيطة يسهل فهمها والتعامل معها، وتأخذ أشكالاً مختلفة من أهمها جداول التوزيعات التكرارية للإجابة عن كل سؤال، وكذلك جدول التوزيع الثنائي للإجابة عن سؤالين.

obeikandl.com

تمارين

1. يرغب أحد الباحثين في تقدير معدل المصروفات الشهرية من المياه لإحدى المدن. ناقش فكرة اختيار وحدة المعاينة، وكذلك ناقش كيف يمكن الحصول على إطار للمعاينة؟
2. تنوي دائرة الغابات في إحدى المناطق تقدير نسبة الأشجار التي يزيد قطرها عن 15 سم. إذا كان متوفراً لدى الدائرة خارطة بجميع الغابات الموجودة في المنطقة، ناقش فكرة وحدة المعاينة وكيف يمكن الحصول على إطار للمعاينة؟
3. يحتوى أحد المجتمعات على 8940 وحدة قمنا بترقيمها من 0001 إلى8940. قمنا بسحب عينة عشوائية بحجم 15 وحدة ومن ثم اشرح الطريقة التي استخدمتها لسحب العينة.
4. يرغب مراقب السير في إحدى المدن تقدير نسبة الإطارات التي تُعد غير صالحة للاستعمال والمستخدمة من قبل سائقي السيارات، هل تعتقد أنَّ وحدة المعاينة يجب أن تكون إطارًا أو سيارة أو مجموعة من السيارات واقفة في موقف خاص للسيارات. ما هو إطار المعاينة؟
5. ناقش الطريقة أو الطرق التي يمكن استخدامها لجمع البيانات في الحالات الآتية:

أ. ترغب إحدى مؤسسات الإنتاج التلفزيوني في تقدير نسبة المشاهدين الذين يراقبون البرامج المنتجة
من قبلهم في إحدى البلدان.

ب. ترغب إدارة إحدى الصحف المحلية معرفة رأي قرائها حول تغطيتها الأخبار المحلية والعالمية.

oberikandi.com

ت. تود دائرة الصحة العامة تقدير نسبة الكلاب التي ستقوم بمهاجمة المارة خلال الشهر القادم في أحد الأحياء السكنية.

6. يعتقد كثير من الناس أن السياسة المستخدمة للقبول في كثير من الجامعات العالمية، والتي تعتمد على علامات الطلاب في الدراسة الثانوية غير دقيقة، وقد تكون غير عادلة في حق كثير من الطلبة، صمم استمارة لدراسة سياسة القبول التي يمكن أن تكون بديلة عما هي مستخدمة الآن.
7. سئل خريجو طلبة الثانوية العامة في إحدى الاستبانات: ما هو المساق الذي درست في الثانوية العامة، والذي كان له الدور الكبير في تحديد مستقبلك الدراسي؟ فكانت إجابة 25% من الطلبة رياضيات و25% اللغة الإنجليزية. هل تعتقد أنَّ هذا السؤال جيد، وأنَّ نتائجه يمكن الاعتماد عليها؟
8. ناقش ما هي الصعوبات التي سنواجهها في الحصول على إطار للمعاينة للحالات الآتية:

أ. إذا أردنا دراسة مصروفات طلبة جامعة الملك فهد للبترول والمعادن الشهرية.

ب. إذا أردنا دراسة الحيوانات البرية الموجودة في صحراء الربع الخالي.

ت. إذا أردنا دراسة مستوى معيشة الفلاحين في غور الأردن.

ث. إذا أردنا دراسة نسبة الفواتير المتأخرة الدفع إلى شركة الكهرباء في مدينة الدمام.

9. لمعرفة رد فعل طلبة إحدى الجامعات حول طرح مساق جديد بالحاسب كمتطلب جامعي لجميع طلبة الجامعة، صمم استمارة استبانه لكي نقوم

obeikandl.com

بتوزيعها على طلبة الجامعة لكي نستخدم المعلومات التي حصلنا عليها من هذه الاستبانة لتصويب تدريس هذا المساق في الجامعة.

10. صمم استمارة استبانه لجمع معلومات لدراسة الحالة الاجتماعية لطلبة في إحدى كليات الجامعة، تتضمن محل الميلاد، والمشكلات التي يواجهها الطالب في الحصول على مسكن في المدينة، وكيفية قضاء أوقات فراغه، والهوايات التي يمارسها، ومصروفاته الشهرية (موزعة على السكن والغذاء والكتب ... إلخ).

11. ما هو رأيك في صحة تمثيل العينات الآتية للمجتمعات التي سحبت منها:

أ. أراد أحد الباحثين دراسة الحالة الاجتماعية لطلبة جامعة اليرموك فقام بالوقوف عند البوابة الرئيسة للجامعة الساعة الثامنة صباحاً، وسأل بعض الطلاب الذين دخلوا الجامعة بعض الأسئلة حول أحوالهم الاجتماعية.

ب. أراد أحد الباحثين دراسة الحالة الاقتصادية لعمال أحد المعامل، حصل على كشف بأسماء العمال، وقام باختيار العمال الذين يحملون الأرقام 6, 26, 46.

ت. أراد أحد الأطباء العاملين في مجمع للصناعات الكيماوية دراسة الحالة الصحية لعمال المجمع، فجاء بجهاز للأشعة السينية وفحص من تقدم من العمال للفحص، وأعدَّ تقريراً عن الحالة الصحية لعمال المجمع.

ث. لدراسة أحوال الفلاحين في إحدى القرى، ذهب الباحث إلى مختار القرية وسأله مجموعة أسئلة عن بعض العائلات الساكنة في هذه القرية.

obeikandl.com

12. يطلب من الطلبة اختيار عينة عشوائية واعتباطية من طلبة الفصل لدراسة معدل أطوال الطلبة في الصف، ثم يقارن بين النتائج التي حصل عليها كل طالب، والنتائج الحقيقية لطلبة الفصل.

obeikandl.com

المراجع العربية

1. وليم كوكوران – ترجمة أنيس كنجو – تقنية المعاينة الإحصائية (1995)، جامعة الملك سعود، الرياض.
2. حسين علوان (1994) ، طرق المعاينة، دار الفرقان، عمان.

References

1. Barnett, V. (1991). Sampling Survey Principles and Methods, Edward Arnold, London.
2. Cochran, W. G. (1977). Sampling Technique, 3rd Ed. John Wiley & Sons, New York.
3. Fisher, R. A. (1953). The Design of Experiments, Oliver & Boyd, London.
4. Govindarajulu, Z (1999). Elements of Sampling Survey and Methods, Prentice Hall, New Jersey.
5. Hajek, J (1981). Sampling from Finite Population, Marcel Dekker, New York.
6. Lahiri, D. B. (1951). A Method of Sample Selection Providing Unbiased Ratio Estimates Bull. Int. Statist., 33, 133, 140.
7. Gillham, Bill (2000). Development of Questionnaire, Continuum, London.
8. Groves, R. M (1998). Nonresponse in a Household Interview Survey, Wily, New York.
9. Lavrakas, P.J. (1993). Telephone Survey Methods: Sampling Selection and Supervision, 2nd, Sage, Newbury Park.
10. Lessler, J. T., and Kalsbeek, W. D. (1992). Nonsampling Error in Surveys, Wily, New York.
11. Levy, P. S. and Lemeshow S (1991). Sampling of Population: Methods and Applications, Wiley, New York.
12. Mahalanobis P. C. (1944). On the Large-Scale Sample Surveys, Phil. Trans. R. Soc., 231, 329-451.
13. Mahalanobis, P. C. (1950). Cost and Accuracy of Results in Sampling and Complete Enumeration, Bull. Int. Statist. Inst., 32, 210-213.
14. Malhotra, N. K. (2006). Marketing Research An Applied Orientation, 5 Ed. Prentice Hall, New Jersey.
15. Moser, C. A. and Kalton, G. (1986). Survey Methods in Social Investigation, Aldershot: Gower Publisher, London.
16. Neyman, J. (1959). Bias in Survey due to Nonresponse. In New Developments in Survey Sampling (Johnson, N. L. and Smith H. Eds.) p 712-732, Wiley, New York.
17. Oppenheim, A. N. (1992). Questionnaire Design, Interviewing and Attitude Measurements, New ed, Printer Publisher, London.
18. Robson, C. (1993). Real World Research: A Resource for Social Scientists and Practitioner Researchers, Blackwell, Oxford.
19. Sampath, S. (2001). Sampling Theory and Methods, Alpha Science International Ltd. U. K.
20. Scheaffer, R. L., W. Mendenhall and L. Ott (1996). Elementary Survey Sampling, 5th Ed., Duxbury press, Boston.
21. Schuman, H., and Presser, S. (1981). Questions and Answers in Attitude Surveys, Academic Press New York

obeikandl.com

22. Seal, K. C. (1962). Use of Outdated Frames in Large Sample Surveys, Bull-Cal. Statist Assoc., 11, 68-84.
23. Singh, D. (1978). Taking Agricuitural Censuses, FAO, Rome.
24. Sturat A. (1984). The Ideas of Sampling, Revised Edition, Griffin, London.
25. Thompson, S. K (2002). Sampling, 2end Wiley, New York.
26. Yates, F. and P. M. Grundy, (1953). Selection without Replacement from Within Strata Probability Proportional to Size J. R. Statist, Soc. 15 B, 253-261.
27. Yates, F. (1960). Sampling Methods for Censuses and Surveys, Charles Grittin & Co., London.
28. Yates, F. (1981). Sampling Methods for Censuses and Surveys 4th Ed., Grittin, London.
29. Zarkovich, S. S. (1961). Sampling Methods and Censuses, FAO, Rome.
30. Zikmund, W. G. and Pabin, B. J. (2006). Exploring Marketing Research, 9 Ed., South-Western College Pub.

obeikandl.com

الفصل الخامس

العينة العشوائية البسيطة
Simple Random Sample

1.5 مقدمة

هناك أساليب كثيرة لاختيار عينة عشوائية بسيطة بحجم n من الوحدات من مجتمع يحتوي على n من الوحدات، ومهما كانت التحضيرات الأولية للأعداد والتحضير إلى المسح بطريقة العينة العشوائية البسيطة لابد من سحب هذه العينة من المجتمع بطريقة أو بأخرى.

تعرف العينة العشوائية البسيطة على أنها الطريقة التي نسحب بها n وحدة من وحدات المجتمع التي عددها N وحدة بحيث تكون أي عينة بالحجم n لها الفرصة نفسها بأن تكون هي العينة المسحوبة. إذا كان المجتمع يحتوي على Nمن الوحدات وأردنا أن نسحب عينة حجمها n بدون إرجاع سيكون لدينا

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

عينة ممكنة للسحب، وكل عينة متميزة عن الأخرى، ونعني بالمتميزة أن لا تكون جميع عناصر أي عينة مشابهة لجميع عناصر عينة أخرى، على سبيل المثال، إذا كان لدينا مجتمع يحتوي على 5 عناصر أي N=5 ونريد أن نسحب عينة حجمها 2 أي إن n=2 فيكون لدينا 10 عينات ممكنة السحب أي:

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$$

obeikandi.com

كيف يمكن سحب عينة واحدة بحجم وحدتين من بين العشر عينات الممكنة الظهور؟ إنه من حسن الحظ أن نستطيع سحب هذه العينة، وذلك عن طريق استخدام جداول الأعداد العشوائية أو الحاسب الآلي، وذلك بسحب وحدة واحدة كل مرة من المجتمع وبدون إرجاع إلى أن نسحب n=2 من الوحدات، وهذا يمثل العينة المطلوبة. ولكن السؤال هو هل احتمال سحب هذه العينة مساو لاحتمال سحب جميع العينات الممكنة السحب؟ أثبت Raj (1968) أن العينة التي تسحب وحداتها كل مرة وحدة من المجتمع وبدون إرجاع يكون احتمال سحبها مساوياً لجميع العينات الأخرى الممكنة السحب.

2.5 مثال

لنفترض أن لدينا عائلة مكونة من 4 أفراد، وهذه العائلة تمثل المجتمع الذي ننوي دراسته، وبالتحديد ننوي دراسة أطوال هذه العائلة، فإذا كانت هذه العائلة مؤلفة على وجه التحديد من الأشخاص: أ، ب، ج، د، وقمنا بقياس أطوالهم فوجدناها حسب الترتيب110، 130، 150، 170، تكون البيانات كما يلي:

(أ، 110) (ب، 130) (ج، 150) (د، 170)

إن معدل أطوال المجتمع (العائلة) 140 لنفرض أننا نريد أن نسحب عينة حجمها n=2 سيكون لدينا عدد العينات الممكنة السحب6، الجدول الآتي يمثل كل العينات الممكنة السحب.

جدول1: العينات الممكنة السحب بحجم 2 سحبت من مجتمع عدد وحداته4

رقم العينة	1	2	3	4	5	6
عناصر العينة	(ا،ب)	(ا،ج)	(ا،د)	(ب،ج)	(ب،د)	(ج،د)
قيم العينة	(130، 110)	(150، 110)	(170، 110)	(150، 130)	(170، 130)	(150، 170)

obeikandl.com

لو أننا قمنا بترقيم هذه العينات من 1 إلى 6 كما هو واضح في الجدول، وقمنا بسحب رقم واحد من بين الأرقام الستة لكان احتمال ظهور كل عينة هو $\frac{1}{6}$ أي إن احتمال سحب كل عينة ممكنة مساو إلى احتمال سحب العينات الأخرى، ولكن كما هو واضح لا يمكن تطبيق هذه الطريقة لصعوبتها من الناحية العملية، على سبيل المثال: إذا كان حجم المجتمع N=100 وحجم العينة التي ننوي سحبها n=2 فإن عدد العينات الممكنة الظهور أو السحب هي 4950 عينة.

نلاحظ أنه من غير الممكن أن نكتب جميع العينات الممكنة الظهور. لذلك تم الاستعاضة عن هذه الطريقة بأن نقوم بسحب وحدة واحدة كل مرة من المجتمع وبدون إرجاع. قد يسأل سائل ما هو الغرض إذن من هذا المثال غير الواقعي؟ حيث إنه لا يوجد مجتمع بهذا الصغر ونريد دراسته عن طريق استخدام العينات، للجواب عن هذا التساؤل شقان: الأول هو توضيح الفكرة الأساسية وراء سحب العينة العشوائية البسيطة، والسبب الثاني وهو الأهم أننا سوف نستخدم هذا المثال في هذا الفصل لتوضيح بعض الأفكار الأساسية والمهمة والتي تستخدم كثيراً في طرق المعاينة.

3.5 رموز ومصطلحات

N: عدد وحدات المجتمع الذي ننوي دراسته

n: عدد وحدات العينة التي سحبت من المجتمع

yi: قيمة المتغير y للوحدة i من وحدات المجتمع

$y_1, y_2, \ldots, y_N$: قيم المتغير y لجميع وحدات المجتمع

$y_1, y_2, \ldots, y_n$: قيم المتغير y لوحدات العينة

openkandil.com

$Y = \sum_{i=1}^{N} Yi$: مجموع المتغير y في المجتمع

$\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Yi$: الوسط الحسابي للمجتمع للمتغير y

$S^2 = \frac{1}{N-1} \sum (y_i - \bar{Y})^2$: تباين المجتمع للمتغير y

A: عدد وحدات المجتمع التي تحمل صفة مميزة أو تنتمي إلى فئة معينة

$P = \frac{A}{N}$: نسبة الوحدات في المجتمع التي تحمل صفة مميزة أو تنتمي إلى فئة معينة

4.5 تقدير الوسط الحسابي للمجتمع

كما هو متوقع فإن تقدير الوسط الحسابي للمجتمع $\bar{Y}$ بالاعتماد على البيانات التي سحبت باستخدام العينة العشوائية البسيطة هو الوسط الحسابي للعينة:

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

إن من أهم صفات الوسط الحسابي للعينة $\bar{Y}$ أنه تقدير غير متحيز إلى الوسط الحسابي للمجتمع $\bar{Y}$ أي:

$$E(\bar{y}) = \bar{Y}$$

حيث إن E ترمز إلى التوقع الرياضي.

إن البرهان الرياضي لهذه الميزة ليس بصعب، ولكننا سوف لا نستخدم الأسلوب الرياضي لبرهان هذه الميزة المهمة والميزات الأخرى، يراجع Cochran(1977) أو ترجمة أنيس كنجو(1995) لهذا البرهان الرياضي،

obeikandl.com

وكذلك جميع البراهين الرياضية الأخرى. وبدلاً من الأسلوب الرياضي سوف أستخدم المثال 2.5 لشرح فكرة عدم التحيز للوسط الحسابي للعينة $\bar{y}$. وذلك من خلال دراسة توزيع لكل العينات الممكنة الظهور. وفي الحقيقة عدم التحيز بالنسبة للوسط الحسابي للعينة $\bar{y}$ يعني إذا أخذت جميع العينات الممكنة الظهور أو السحب بحجم n من المجتمع الذي يحتوي على N من الوحدات فإن الوسط الحسابي لأوساط جميع العينات الممكنة الظهور يجب أن يساوي الوسط الحسابي للمجتمع $\bar{Y}$ أي

$$Aver(\bar{y}) = \frac{1}{\binom{N}{n}} = \sum_{i=1}^{\binom{N}{n}} \bar{y}_i = \bar{Y}$$

رقم العينة	1	2	3	4	5	6
$\bar{y}_i$	120	130	140	140	150	160

$$Aver(\bar{y}) = \frac{1}{6} = \sum_{i=1}^{6} \bar{y}_i = \frac{840}{6} = 140$$

أما الوسط الحسابي للمجتمع:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i = \frac{1}{4}\sum_{i=1}^{4} y_i = \frac{840}{6} = 140$$

واضح أن معدل جميع الأوسط الحسابية للعينات الممكنة الظهور مساو للوسط الحسابي للمجتمع $\bar{Y}$. أي أن $\bar{y}$ تقدير غير متحيز إلى $\bar{Y}$. الآن ننتقل إلى تباين $\bar{y}$ حيث إنه يعرف:

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

obeikandl.com

حيث إن f =n/N: يسمى كسر المعاينة، وكذلك (1-f) يسمى معامل التصحيح للمجتمعات المحدودة أو الصغيرة الحجم. بإمكاننا إهمال معامل التصحيح للمجتمعات المحدودة إذا كانت قيمته أقل من 5%.

5.5 تقدير تباين المجتمع

يمكن تلخيص استعمالات تباين الوسط الحسابي للعينة $V(\bar{y})$ بما يأتي:

1- للحصول على دقة تقدير $\bar{y}$ للوسط الحسابي للمجتمع $\bar{Y}$.

2- لمقارنة $\bar{y}$ بتقديرات أخرى للوسط الحسابي للمجتمع $\bar{Y}$.

3- لتحديد حجم العينة التي نحتاجها للوصول إلى الدقة المطلوبة لتقدير الوسط الحسابي للمجتمع $\bar{Y}$ باستخدام الوسط الحسابي للعينة $\bar{y}$.

ليس من السهل معرفة S^2 لذلك لابد من تقديرها من العينة العشوائية البسيطة التي سحبت من المجتمع وذلك باستخدام ما يسمى بتباين العينة s^2 والذي يعرف:

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

إن تباين العينة s^2 تقدير غير متحيز إلى تباين المجتمع S^2 ، والآن نعود إلى مثالنا في 2.5 لإثبات هذه الخاصية. إذا وجدنا تباين كل عينة ممكنة السحب والتي عددها 6 عينات. وأوجدنا الوسط الحسابي لهذه التباينات فإنه يساوي تباين المجتمع أي:

رقم العينة	1	2	3	4	5	6
s_1^2	200	800	1800	200	800	200

obeikandl.com

الآن

$$Aver(s^2) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} s_1^2 = \frac{200 + 800 + ... + 200}{6} = 666.67$$

أما تباين المجتمع

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \overline{Y})^2 = 666.67$$

وهذا يعني أن s^2 تقدير غير متحيز إلى S^2 .

6.5 تقدير تباين الوسط الحسابي للعينة

لقد سبق أن عرفنا تباين $\overline{y}$ بأنه:

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

ولكن تباين المجتمع S^2 غير معروف إذاً لابد من استعمال s^2 تباين العينة الذي هو تقدير غير متحيز إلى S^2 ليصبح تقدير تباين الوسط الحسابي للعينة $\overline{y}$:

$$s^2(\overline{y}) = \frac{1-f}{n} s^2$$

واضح أن $s^2(\overline{y})$ تقدير غير متحيز إلى $V(\overline{y})$ ، لإثبات هذه الخاصية لنعد إلى مثالنا في 2.5 لنجد الجدول الآتي:

رقم العينة	1	2	3	4	5	6
$\overline{y_i}$	120	130	140	140	150	160
$(\overline{y_i} - \overline{Y})^2$	400	100	0	0	100	400

$$V(\overline{y}) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} (\overline{y_i} - \overline{Y})^2 = \frac{1}{6}(1000) = 166.67$$

obeikandl.com

أما تباين $\bar{y}$:

$$V(\bar{y}) = \frac{1-f}{n} S^2 = \frac{1-2/4}{2}(666.67) = 166.67$$

وهذا يثبت أن $s^2(\bar{y})$ تقدير غير متحيز إلى $V(\bar{y})$. ولكن في غالبية كتب الإحصاء يعرف تباين الوسط الحسابي للعينة بما يأتي:

$$V(\bar{y}) = \frac{s^2}{n}$$

حيث إن:

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{Y})^2$$

لو استخدمنا البيانات الموجودة في مثال 2.5 لوجدنا

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{Y})^2 = 500$$

لذلك فإن تباين $\bar{y}$

$$V(\bar{y}) = \frac{s^2}{n} = \frac{500}{2} = 250$$

والسؤال الآن لماذا هذا الاختلاف بين قيمة $V(\bar{y})$ في الحالتين؟ في الحقيقة إن المعادلة $V(\bar{y}) = \frac{s^2}{n}$ وجدت للمجتمعات غير المحدودة أو الكبيرة جداً، ولكن في مثالنا هنا المجتمع محدود وصغير جدً لذلك نجد هذا الاختلاف الواضح. وسوف نستخدم في هذا الكتاب

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

obeikandl.com

حيث إن

$$S^2 = \frac{N}{N-1} s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \overline{Y})^2$$

7.5 فترة ثقة للوسط الحسابي

لإيجاد فترة ثقة للوسط الحسابي $\overline{Y}$ يستخدم الوسط الحسابي للعينة $\overline{y}$ لهذا الغرض، ولإيجاد هذه المدة لابد من معرفة توزيع $\overline{y}$ ، هناك حالتان لمعرفة توزيع $\overline{y}$:

1. إذا كانت $y_1, y_2, ..., y_n$ ، تتبع التوزيع الطبيعي. فإن $\overline{y}$ كذلك يتبع التوزيع الطبيعي أي:

$$\overline{y} \sim N(\overline{Y}, \frac{1-f}{n} S^2)$$

2. إذا كان حجم العينة $n < 30$ فإننا نستطيع أن نستخدم نظرية تقارب التوزيعات الاحتمالية لنجد أن $\overline{y}$ يتبع التوزيع الطبيعي التقريبي أي:

$$\overline{y} \sim N(\overline{Y}, \frac{1-f}{n} S^2)$$

إذا كانت S^2 معلومة فإن فترة $(1-a)100\%$ ثقة الوسط الحسابي $\overline{Y}$ يمكن أن تعرف:

$$\overline{y} - z_{a/2} S \sqrt{\frac{1-f}{n}} \le \overline{Y} \le \overline{y} + z_{a/2} S \sqrt{\frac{1-f}{n}}$$

حيث إن $Z_{a/2}$ يمكن إيجادها باستخدام جدول التوزيع الطبيعي المعياري بحيث

$$P(Z > z_{a/2}) = a/2$$

obeikandl.com

إذا كان $\bar{y}$ تقريباً يتبع التوزيع الطبيعي ففترة الثقة للوسط الحسابي أعلاه تكون تقريبية. وإذا كانت S^2 غير معلومة نستخدم بدلاً منها s^2 ومن ثم نستخدم توزيع t بدلاً من توزيع z لتصبح فترة الثقة للوسط الحسابي كما يأتي:

$$\bar{y} - t_{n-1,a/2}\, s\sqrt{\frac{1-f}{n}} \le \bar{Y} \le +t_{n-1,a/2}\, s\sqrt{\frac{1-f}{n}}$$

حيث إن (n-1) تسمى درجات الحرية، ونستخدم جدول توزيع t لإيجاد قيمة $t_{n-1,a/2}$.

مثال (1): لمعرفة معدل عدد أيام الإجازة الإضافية التي يأخذها عمال مجموعة من المصانع عند نهاية إجازة عيد الفطر وعيد الأضحى، سحبت عينة عشوائية حجمها 100 عامل من 36000 عامل يعملون في هذه المصانع، وسُئلوا عن عدد أيام الإجازة الإضافية التي أخذوها بسبب إجازة عيد الفطر أو الأضحى فكانت إجاباتهم كما يلي:

عدد الأيام y_i	0	1	2	3	4	5	6	7	8
عدد العمال f_i	45	17	16	10	6	2	1	3	0

لتقدير معدل عدد الأيام التي يأخذها العمال إجازة إضافية بسبب إجازة العيدين نستخدم الوسط الحسابي للعينة $\bar{y}$ لنجد أن:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{k} y_i f_i = \frac{1}{100}(1400) = 1.4$$

مع تباين عينة $s^2 = 3.071$. بما أن حجم العينة كبير فإن $\bar{y}$ تقريباً يتبع التوزيع الطبيعي لذلك فإن فترة 95% ثقة للوسط الحسابي $\bar{Y}$ تكون:

$$1.05701 \langle \bar{Y} \langle 1.74299$$

إذا أهملنا معامل التصحيح تصبح الفترة:

$$1.05653 \langle \bar{Y} \langle 1.74347$$

obeikandl.com

8.5 تقدير المجموع الكلي للمجتمع

هنالك حالات كثيرة يكون فيها تقدير مجموع المجتمع الكلي لمتغير معين أكثر أهمية من المعدل، على سبيل المثال مجموع أوزان الحقائب التي سيتم شحنها على متن الطائرة أهم بكثير من معدل وزن الحقيبة الواحدة.

ويمكن تعريف مجموع المجتمع الكلي:

$$Y = \sum_{i=1}^{N} y_i = N\overline{Y}$$

باستخدام نتائج العينة العشوائية البسيطة، وحيث إن N حجم المجتمع معروف فإن:

$$Y = N\bar{y}$$

تقدير غير متحيز إلى المجموع الكلي للمجتمع Y وتباين Y :

$$V(Y) = N^2\, V(\bar{y}) = N^2 \frac{1-f}{n} S^2$$

بما أن S^2 غير معروفة فإننا نستخدم s^2 تباين العينة ليصبح تقدير تباين Y :

$$s^2(Y) = N^2\, s^2(\bar{y}) = N^2 \frac{1-f}{n} s^2$$

كذلك تقدير الخطأ المعياري يصبح:

$$s(Y) = \sqrt{s^2(Y)}$$

إذا كان حجم العينة أكبر من 30 وحدة فإن Y يتبع التوزيع الطبيعي التقريبي أي:

$$\cdot\ N(Y, N^2 \frac{1-f}{n} S^2)\ Y$$

obeikandl.com

وبالطريقة نفسها يمكن إيجاد فترة $(1-\alpha)100\%$ ثقة تقريبية للمجموع الكلي Y:

$$Y \mp z_{\alpha/2} NS\sqrt{(1-f)/n}$$

إذا كانت S^2 غير معروفة نستخدم s^2 بدلاً منها و $t_{n-1,\alpha/2}$ بدلاً من $z_{\alpha/2}$.

9.5 تقدير النسبة

لنفرض أنه يوجد لدينا حالة يهمنا فيها فقط إذا كان المتغير y ينتمي إلى فئة معينة أو يتصف بصفة معينة، فمثلاً في صف يتكون من 40 طالباً يهم المدرس أن يعرف ما نسبة أو عدد الطلاب الراسبين في الامتحان الأول، أو في مصنع يهم قسم مراقبة جودة الإنتاج إذا كانت الوحدة المصنوعة مطابقة للمواصفات أم لا. لذلك نستطيع أن نعيد تعريف المتغير y ليصبح:

إذا كان المتغير y يتصف بصفة مميزة أو يقع ضمن فئة معينة $y_i = 1$

إذا كان خلاف ذلك $y_i = 0$

إذاً

$$Y = \sum_{i=1}^{N} y_i = A; \quad \overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i = \frac{A}{N} = \pi$$

حيث A تمثل عدد الوحدات في المجتمع التي تتصف بصفة مميزة و π تمثل نسبة الوحدات للذين يحملون صفة مميزة أو ينتمون إلى فئة معينة في المجتمع كذلك

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \pi)^2 = \frac{N\pi(1-\pi)}{N-1}$$

الآن نحن بصدد تقدير π واضح أن نسبة الذين يحملون صفة معينة أو ينتمون إلى فئة معينة في العينة تكون تقديراً إلى π أي

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{a}{n} = p$$

obeikandl.com

حيث إن a تمثل عدد الوحدات في العينة للذين يحملون صفة مميزة أو ينتمون إلى فئة معينة. إن برهنة p كتقدير غير متحيز إلى π ، أي $E(p)=\pi$ ليس بالصعب، نستطيع أن نستخدم نفس الطريقة التي استخدمناها لبرهنة أن الوسط الحسابي للعينة $\bar{y}$ تقدير غير متحيز إلى الوسط الحسابي للمجتمع $\bar{Y}$ أو باستخدام الطرق الرياضية راجع Cochran (1977) أو ترجمة أنيس كنجو (1995). أما تباين p فهو

$$V(p)=\frac{1-f}{n}S^2=\frac{N-n}{N-1}\frac{\pi(1-\pi)}{n}$$

ولكن π غير معلومة لذلك فإن S^2 غير معلومة. نستطيع تقدير S^2 باستخدام تباين العينة:

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2=\frac{np(1-p)}{n-1}$$

لذلك فإن تقدير تباين p سيكون:

$$s^2(p)=\frac{1-f}{n-1}p(1-p)$$

إذا أهملنا معامل التصحيح فإن $s^2(p)$ يصبح:

$$s^2(p)=\frac{p(1-p)}{n-1}$$

أما تقدير الخطأ المعياري لتقدير p

$$s(p)=\sqrt{\frac{p(1-p)}{n-1}}$$

بقي أن نجد فترة ثقة للنسبة π إذا كان $n\pi\geq 5$ و $n(1-\pi)\geq 5$ فإن p يتبع التوزيع الطبيعي التقريبي أي أن:

$$p\cdot N(\pi,\frac{1-f}{n}\pi(1-\pi))$$

obeikandl.com

إن فترة ثقة تقريبية للنسبة π فستكون:

$$p \mp z_{\alpha/2} \sqrt{\frac{1-f}{n-1} p(1-p)}$$

10.5 اختيار حجم العينة

يُعد اختيار حجم العينة المناسب مهماً جداً في النتائج النهائية للدراسة التي نقوم بها، حيث إنه كلما كبر حجم العينة كانت التقديرات والنتائج التي نحصل عليها دقيقة وقريبة من معلمات المجتمع. ولكن في المقابل لا نستطيع أن نأخذ حجم العينة كبيراً؛ لأن ذلك يتطلب جهداً ووقتاً وتكاليف مادية كبيرة قد لا يكون بالمستطاع توفيرها لحجم عينة كبير. فإذاً السؤال كيف نختار حجم عينة – وأقصد بالعينة هنا العينة العشوائية البسيطة– من المجتمع لتأمين دقة معينة للنتائج، وفي الوقت نفسه يكون بالإمكان تنفيذها على أرض الواقع وضمن الإمكانيات المتاحة؟ سوف يقتصر كلامنا على اختيار حجم العينة في ثلاث حالات نتناولها بالتفصيل فيما يأتي

1.10.5 اختيار حجم العينة لتقدير الوسط الحسابي للمجتمع

إن زيادة حجم العينة سيؤدي إلى زيادة الدقة للوسط الحسابي للعينة $\bar{y}$ كتقدير للوسط الحسابي للمجتمع $\bar{Y}$ ولكن هذه الزيادة لها حدود في بعض الأحيان لا نستطيع تجاوزها، إن حجم عينة كبيرة يعني إضاعة للموارد المتاحة. كذلك حجم عينة صغيرة على الأرجح سوف يعطينا تقديراً غير دقيق وعلى هذا فمن الناحية المثالية باستطاعتنا أن نحدد الدقة التي نريدها أو الحد الأعلى للتكاليف التي نود صرفها، ومن ثم نحدد حجم العينة المطلوبة لتحقيق هذين الشرطين.

لكن هذا لن يتم بسهولة، حيث إن هناك كثيراً من المشكلات فمثلاً كيف نحدد الدقة التي نحتاجها للتقديرات المختلفة؟ كيف نوازن بين صفات أو متغيرات المجتمع المختلفة التي نريد أن ندرسها؟ كيف نتعامل مع بعض معلمات المجتمع غير المعروفة والتي نحتاجها –مثل تباين المجتمع– والتي ربما تؤثر في الدقة للتقديرات.

سوف تدرس حالة سهلة ولنفرض أننا نهدف لتقدير الوسط الحسابي للمجتمع لمتغير واحد وهو y باستخدام الوسط الحسابي للعينة العشوائية البسيطة. لذلك لا بد من تحديد الخطأ في تقدير الوسط الحسابي للمجتمع $\overline{Y}$ بمقدار معين وليكن d بحيث

$$P\left(\left|\overline{Y}-\overline{y}\right|\le d\right)=1-\alpha$$

حيث إن d رقم محدد سلفاً و(α) صغير و(1-α) تسمى درجة الثقة. وباستخدام التوزيع الطبيعي التقريبي نستطيع أن نحدد الحد الأدنى لحجم العينة لتقدير الوسط الحسابي للمجتمع بما يأتي:

$$n \ge \frac{n_0}{1+n_0/N}$$

حيث إن

$$n_0=\frac{z_{\alpha/2}^2 S^2}{d^2}$$

لكن S^2 تكون في الغالب غير معلومة وهناك أربعة طرق نستطيع أن نقدر بها S^2.

اختبار الموقع (Pilot Studies)

في أغلب الأحيان هذه الدراسة تجري قبل بدء عملية المسح الشامل أو المسح بالعينة، لتخدم أغراضاً متعددة، منها دراسة عدة إطارات للمعاينة إذا

كانت هذه الدراسة تأخذ شكل عينة عشوائية بسيطة يمكن استخدام نتائجها لكي تعطينا فكرة عن قيمة S^2. ولكن إذا كانت هذه الدراسة لا تأخذ بمبدأ العينة الاحتمالية فإذاً يجب التحفظ على استخدام نتائجها لمعرفة أو إعطاء فكرة عن S^2.

أ. استخدام مسوحات سابقة

ليس من غير الشائع أن نجد مسوحات سابقة قامت بها جهات معينة لدراسة المجتمع نفسه لصفات أو متغيرات مماثلة، وهذا يحدث كثيراً في ميدان التربية والطب وعلم النفس. في أغلب الأحيان يمكن استخدام البيانات من هذه الدراسات لتقدير S^2 .

ب. استخدام عينة عشوائية أولية

إن هذه الطريقة تُعَد من أفضل الطرق لتقدير S^2 ، ولكنها قد لا تكون ممكنة من الناحية الإدارية أو من ناحية التكاليف، ويمكن تلخيصها بأن تؤخذ عينة عشوائية بسيطة أولية بحجم n_1 أي بحجم كافٍ للحصول على تقدير مقبول من ناحية الدقة لتباين المجتمع S^2. وبعد تحديد حجم العينة الكلي n بالاعتماد على s_1^2 الذي وجد باستخدام العينة الأولية بحجم n_1. نقوم بسحب عينة عشوائية بسيطة بحجم $(n-n_1)$ ، أي أن الوحدات التي حصلنا عليها باستخدام العينة الأولية سوف تدخل ضمن العينة العشوائية البسيطة النهائية. إن هذه الطريقة من أفضل الطرق، وتعد طريقة موضوعية وممكنة التطبيق، وكذلك يمكن الاعتماد على النتائج التي تم الحصول عليها باستخدام هذه الطريقة.

ت. استخدام بعض المعلومات عن ترتيب المجتمع

في بعض الأحيان نعرف معلومات عن المجتمع تحت الدراسة، يمكن استخدام هذه المعلومات لإعطائنا فكرة عن S^2 فمثلاً:

obeikandl.com

1. عدد الأخطاء المرتكبة في الصفحة الواحدة للكتب ذات الأحجام المتساوية والتي طبعت من قبل دار نشر معينة خلال مدة معينة. أو
2. عدد الأعطال التي تحدث في مسجل صوت من نوع معين خلال السنة الأولى من استعماله.

في كلتا الحالتين نستطيع أن نقول إنهما تتبعان توزيع بواسون، ومن معلوماتنا عن توزيع بواسون فإن $\overline{Y} = S^2$ يمكن استخدام هذه الخاصية، كذلك المعلومات عن متغيرات مشابهة من مسوحات سابقة نستطيع أن نعرف أو نقدر S^2 .

2.10.5 اختيار حجم العينة لتقدير المجموع الكلي للمجتمع

لنفترض أننا نريد أن نقدر حجم العينة n لتقدير مجموع المجتمع الكلي Y باستخدام Y من العينة العشوائية البسيطة بحيث

$$P(\left| Y - Y \right| > d_1) < \alpha$$

حيث إن d_1 هنا تمثل مقدار حد الخطأ لتقدير Y. وباستخدام التوزيع الطبيعي التقريبي نحصل

$$n \geq \frac{n_0}{1 + n_0/N}$$

حيث إن:

$$n_0 = \frac{N^2 z_{\alpha/2}^2 S^2}{d_1^2}$$

بعض الأحيان يعبر عن الدقة لتقدير $\overline{Y}$ أو Y على شكل نسبة. فعلى سبيل المثال يمكن أن نسأل ما هو حجم العينة بحيث إن:

$$P(\left| Y - Y \right| > \pi Y) \leq \alpha$$

obeikandi.com

فمثلاً ممكن أن نريد حجم العينة بحيث نكون متأكدين 95% أن Y سيكون ضمن 2% للمجموع الكلي الحقيقي للمجتمع Y هنا $\pi = 2\%$ و $\alpha = 0.05$ لابد من استبدال d_1 بما يساويها وهو πY إذا كان المتغير y تقريبا يتبع توزيع بواسون فإن:

$$S^2 = \overline{Y} = \frac{Y}{N}$$

وسيكون حجم العينة المطلوب

$$n \geq N\left[1 + (\pi / z_{\alpha/2})^2 Y\right]^{-1}$$

مثال (2): لمعرفة مقدار مبيعات سلسلة من المحلات التجارية مكونة من 243 محلاً. ما هو حجم العينة الذي نحتاجه لتقدير مجموع المبيعات ليكون ضمن 10% من الرقم الحقيقي مع ثقة أو دقة 95٪، من معلوماتنا عن مبيعات السنوات الثلاث الماضية عن مجموع المبيعات والانحراف المعياري للمبيعات بآلاف الريالات هي:

السنة	1	2	3
Y	321.7	366.8	401.0
S	0.826	0.776	0.804

$$n \geq 243\left[1 + \frac{1}{243}\left(\frac{0.10(363.17)}{0.8(1.96)}\right)^2\right]^{-1} = 76$$

3.10.5 اختيار حجم العينة لتقدير النسبة

بإمكاننا استخدام الطريقة التي استخدمناها في حالة تقدير الوسط الحسابي $\overline{Y}$ مع أخذ قيمة $S^2 = \pi(1-\pi)$ لتحديد أو تقدير حجم العينة المطلوب لتقدير النسبة π وإذا أمعنا النظر في تباين تقدير النسبة p

$$V(p) = \frac{1-f}{n} S^2 = \frac{N-n}{N-1}\frac{\pi(1-\pi)}{n}$$

سنجد ان أعلى قيمة لها عندما تكون $\pi = (1-\pi) = 1/2$ ذلك عندما نعطي قيمة محددة لحجم العينة n سوف نقدر π بأقل دقة عندما تكون قيمته قريباً من0.5، إذا أهملنا معامل التصحيح فإن حجم العينة في حده الأدنى يكون:

$$n_0 \geq \frac{z^2_{\alpha/2}\pi(1-\pi)}{d^2}$$

وخلاف ذلك فإن حجم العينة يكون:

$$n \geq \frac{n_0}{1+n_0/N}$$

كذلك سوف نحصل على أكبر قيمة ممكنة لحجم العينة عندما تكون $\pi = (1-\pi) = 1/2$ في حالة تقدير النسبة فإن حجم العينة التي سنختارها عندما تكون قيمة $\pi = 0.5$ سيكون هو الحد الأعلى لحجم العينة المطلوب.

إذا كانت قيمة π غير معروفة في هذه الحالة يمكن وفي أسوأ الحالات أن نأخذ $\pi = 0.5$ أو نحاول معرفة قيمة π بالطرق نفسها التي اتبعناها لمعرفة قيمة S^2 .

مثال (3): أوجد قيمة n بحيث تكون قيمة ضمن $\pm 2\%$ وثقه 95%.

إذا فرضنا $\pi = 0.5$ فإن

$$n_0 \geq \frac{(1.96)^2(0.5)(0.5)}{(0.02)^2} = 2401$$

ولكن إذا كانت $\pi = 0.1$ فإن

$$n_0 \geq \frac{(1.96)^2(0.1)(0.9)}{(0.02)^2} = 865$$

obeikandi.com

تمارين

1. إذا كان لدينا مجتمع يتكون من الوحدات A, B, C, D, E, F وهذه الوحدات تمثل أشخاصاً مصروفاتهم الشهرية بالريال كما يأتي:

الوحدة	A	B	C	D	E	F
المصروف الشهري	1200	1150	1400	1300	2000	1700

أ. أثبت باستخدام جميع العينات العشوائية البسيطة وبحجم 3 وحدات أن الوسط الحسابي للعينة وتباين العينة تقدير غير متحيز للوسط الحسابي ولتباين المجتمع أي أن $E(\bar{y}) = \overline{Y}$ و $E(s^2) = S^2$.

ب. إذا أردنا تقدير نسبة الأشخاص الذين مصروفاتهم الشهرية أكثر من 1400 ريال أثبت أن نسبة العينة التي هي تقدير للنسبة الحقيقية للمجتمع أنها تقدير غير متحيز أي $E(p) = \pi$.

2. سحبت عينة عشوائية بسيطة بحجم وحدتين من مجتمع متكون من $N = 5$، قيم المتغير y_i هي 5 ,1- ,0 ,1. احسب الوسط الحسابي $\bar{y}$ لجميع العينات الممكنة السحب، وأثبت أن تقدير الوسط الحسابي للعينة هو تقدير غير متحيز إلى الوسط الحسابي للمجتمع $\overline{Y}$.

3. للمجتمع في السؤال السابق احسب تباين العينة s^2 لجميع العينات الممكنة السحب وأثبت أنه تقدير غير متحيز لتباين المجتمع S^2.

4. إذا كان لدينا البيانات الآتية التي تمثل علامات N=100 طالب دخلوا الامتحان الأول لمساق مبادئ الإحصاء:

5, 6, 7, 10, 2, 7, 3, 7, 4, 5, 7, 4, 9, 6, 0, 3, 7, 2, 9, 3, 6, 8, 8, 4, 9, 10, 6, 3, 7, 5, 7 6, 8, 7, 5, 2, 4, 8, 4, 7, 4, 4, 7, 7, 4, 5, 8, 5, 4, 6, 7, 10, 4, 5, 7, 4, 1, 7, 9, 8, 6, 5, 8, 2, 2, 4, 7, 5, 10, 4, 6, 8, 4, 7, 8, 9, 3, 10, 7, 4, 7, 5, 3, 6, 6, 4, 9, 7, 7, 10, 5, 0, 1, 4, 5, 7, 8, 3, 2, 6

obeikandi.com

أ. أوجد معدل علامات الطلبة $\bar{Y}$ وتباين S^2 العلامات .

ب. اسحب عينة عشوائية بسيطة حجمها 5 طلاب، ثم أوجد الوسط الحسابي للعينة $\bar{y}$ وتباين. s^2 العينة وتباين $\bar{y}$ وتقدير تباين $\bar{y}$. قارن بين النتائج التي حصلت عليها باستخدام العينة العشوائية البسيطة والنتائج التي حصلت عليها في الفرع أ.

ت. كرر العملية التي قمت بها في الفرع ب وذلك باستخدام أحجام عينات 10, 20 , 40 ثم قارن بين هذه النتائج ونتائج الفرع أ.

ث. اسحب 25 عينة عشوائية بحجم 5 وحدات لكل عينة، ثم أوجد الوسط الحسابي لأوساط العينات والتباين لأوساط العينات. قارن بين النتائج التي سنحصل عليها هنا ونتائجك في الفرع أ.

5. سحبنا عينة عشوائية بسيطة حجمها 12 قطعة أرض زراعية مزروعة بمحصول البطاطا، إذا كانت مساحة كل قطعة 1/4 دونم، قدر إنتاج الأرض الزراعية والتي مساحتها 500 دونم، ثم أوجد فترة 95% ثقة للمجموع الكلي للإنتاج.

4.1, 4.7, 5.2, 3.9, 4.8, 3.8, 4.4, 6.0,5.3, 5.1, 4.9, 3.3

6. سحبنا عينة عشوائية بسيطة حجمها 100 وحدة (مقياس الماء) في منطقة سكنية لتقدير معدل المصروفات الأسبوعية من الماء للعائلة، فحصلنا على $\bar{y} = 5.7$ متر مكعب و $s^2 = 425$. إذا كانت N=10.000 عائلة تسكن في المنطقة. أوجد فترة 90% ثقة للوسط الحسابي للمجتمع $\bar{Y}$. كذلك قدر المجموع الكلي لمصروفات المنطقة الأسبوعية من المياه، وأوجد معامل التغير لتقدير مجموع المياه المصروفة أسبوعياً بالمترات المكعبة.

7. تم سحب عينة عشوائية بحجم 13 عائلة من مدينة فيها 14.000 عائلة، عدد الأشخاص لكل عائلة في العينة هو: 4,6, 5, 2, 7, 6, 3, 8, 10, 9, 3, 6, 4.

obeikandi.com

قدر عدد سكان المدينة، ثم أوجد 95% فترة ثقة لعدد سكان المدينة الحقيقي.

8. لمعرفة آراء طلبة إحدى الجامعات بنظام العلامات المستخدم في الجامعة، سحبت عينة عشوائية بسيطة حجمها 120 طالباً فوجدنا أن 80 طالباً ضد نظام العلامات المستخدم، قدر نسبة الطلبة المعارضين، وأوجد فترة 95% ثقة للنسبة الحقيقية π وكذلك أوجد معامل التغير للنسبة.
9. إذا اعتبرنا البيانات الواردة في السؤال الخامس تمثل عينة أولية من المجتمع، قدر حجم العينة المطلوبة لتقدير معدل إنتاج قطعة الأرض، إذا كان حد الخطأ للتقدير $d = \pm 1.5$ ودرجة الثقة 90%.
10. ما هو حجم العينة الذي يؤمن لنا حد خطأ للتقدير مقداره $d = \pm 0.03$ ودرجة ثقة 99% في السؤال الثامن أعلاه.
11. لتقدير إنتاج أحد بساتين البرتقال الذي يوجد فيه N= 1500 شجرة، قام المزارع بسحب عينة عشوائية بسيطة حجمها n=10 أشجار فوجد البيانات الآتية $\bar{y} = 40.5$ كغم و $s^2 = 40$. قَدِّر الإنتاج الكلي للبستان، ثم أوجد فترة 95% ثقة للمجموع الكلي.
12. إذا اعتبرنا أن البيانات في السؤال الحادي عشر تمثل عينة أولية، أوجِدْ حجم العينة لتقدير مجموع الإنتاج الكلي إذا كان حد الخطأ في تقدير المجموع d= 1000.
13. يبلغ عدد سكان إحدى المدن الصغيرة 5000 شخص. ترغب دائرة الصحة في المدينة أن تقدر نسبة الأشخاص في الفرية الذين تكون فصيلة دمهمAB. ما هو حجم العينة التي يمكن سحبها إذا كنا نرغب بأن يكون تقدير النسبة الحقيقية بخطأ مقداره ±4% وبدرجة ثقة مقدارها 99%. علماً بأننا سحبنا عينة عشوائية أولية فوجدنا نسبة الذين يحملون فصيلة الدم AB حوالي 20%.

obeikandi.com

14. سحبت عينة عشوائية حجمها 5 طلاب من كشف يحتوي على 50 طالباً مسجلين في مادة الإحصاء، الاستبانة التي سلمت لهم تحتوي على سؤالين هما:

السؤال الأول: هل تمتلك سيارة؟ نعم _____ لا _____

السؤال الثاني: ما هو مصروفك بالدولار ألأمريكي؟ _____ دولار.

يعطي الجدول الآتي إجابات الطلبة

الطالب	السؤال الأول	السؤال الثاني
1	لا	300
2	نعم	1500
3	لا	1000
4	لا	500
5	نعم	1200

أ. قدر معدل ومجموع مصروف الطلاب الشهري. ثم أوجد 95% فترة ثقة لمعدل ومجموع مصروفاتهم الحقيقية لكل شهر.

ب. أوجد الخطأ المعياري لتقدير نسبة الطلاب الذين يمتلكون سيارة ثم أوجد 93% فترة ثقة للنسبة الحقيقية للطلاب الذين يمتلكون سيارة.

15. قدر مصروفات الطلبة الشهرية لطلبة إحدى كليات الجامعة.

16. إذا طلبت منك إدارة المطاعم الموجودة في الجامعة القيام باستطلاع آراء الطلبة الذين يتناولون وجبات غذائية في أحد مطاعم الجامعة حول آرائهم في الوجبات الغذائية المقدمة في هذا المطعم. صمم استمارة استبانه ثم نفذها وذلك بجمع البيانات اللازمة لمعرفة آراء الطلبة حول الوجبات الغذائية المقدمة لهم.

obeikandi.com

17. ترغب غرفة تجارة الدمام معرفة آراء أعضائها عن الخدمات التي تقدمها الغرفة إلى أعضائها. إذا طلب منك تنفيذ هذا المشروع، ما الخطوات اللازمة لإنجازه ابتداء من تحديد الأهداف إلى تحليل البيانات؟
18. حدد إحدى المشكلات في مجال عملك أو دراستك، بحيث تستطيع أن تسحب عينة عشوائية بسيطة لتقدير الوسط الحسابي للمجتمع والمجموع الكلي أو النسبة.

obeikandl.com

المراجع العربية

1. سليمان محمد طشطوش (2001)، أساسيات المعاينة الإحصائية، دار الشروق للنشر والتوزيع، عمان.
2. حسين علوان (1994)، طرق المعاينة، دار الفرقان، عمان.
3. هلال عبود البياتي وصبري رديف العاني (1981)، العينات، جامعة بغداد، بغداد.
4. وليد عبدالحميد نوري وعبدالمجيد حمزة الناصر (1981) العينات، وزارة التعليم العالي والبحث العلمي، بغداد.
5. وليم كوكوران –ترجمة أنيس كنجو– تقنية المعاينة الإحصائية (1995)، جامعة الملك سعود، الرياض.

References

1. Barnett, V. (1991). Sampling Survey Principles and Methods, Edward Arnold, London.
2. Benedetto, J. J. and Ferreira, P. J. (2001). Modern Sampling Theory, Birkhauser Boston.
3. Baus, D. (1958). On Sampling with and without Replacement, Sankhya 20, p 287-294.
4. Chaudhuri, A. and Stenger, H. (1992). Survey Sampling: Theory and Methods, Marcel Dekker, New York.
5. Cochran, W. G. (1977). Sampling Technique, 3rd Ed. Wiley, New York.
6. Deming, W. E. (1960). Sampling Design in Business Research. Wiley, New York.
7. Foreman, E. K. (1991). Survey Sampling Principles, Marcel Dekker, New York.
8. Govindarajulu, Z. (1999). Elements of Sampling Theory and Methods, Prentice Hall.
9. Hajek, J (1960). Limiting Distributions in Simple Random Sampling from Finite Population. *Pub. Math. Inst. Hungarian Acad. Sci.* 5, 361-374.
10. Hajek, J (1981). Sampling from Finite Population, Marcel Dekker, New York.
11. Hansen, M.H., and Hurwitz, W. N. (1943). On the theory of sampling from finite popula-ions. Ann. Math. Statist. 14,333-362.
12. Hedayat, A. S. and Sinha, B. K. (1991). Design and Inference in Finite Population Sampling, Wiley, New York.
13. Kish, L (1995). Survey Sampling, Wiley, New York.
14. Lohr, S. L. (1999). Sampling: Design and Analysis, Duxbury, New York.
15. Levy, P. S. and Lemeshow S (1991) Sampling of Population: Methods and Applications, Wiley, New York.
16. Mahalanobis, P. C. (1946). Recent experiments in statistical sampling in the Indian Sta-tistical Institute. 1 Roy. Statist. Soc. A109, 325-378, reprinted in Sankhyii.
17. Murthy, M. N. (1962). Variance and confidence interval estimation Sankhyii 24(B), 1-12.
18. Raj, D (1968). Sampling Theory, McGraw-Hill, New York.
19. Rao, J. N. K. (1985). Conditional inference in survey sampling. Survey Methodology, D,I. Inst. Statist. Math. No.1, 15-31.
20. Sampath, S. (2000). Sampling Theory and Methods, CRC Press.

obeikandl.com

21. Scheaffer, R.L., Mendenhall, W. and Ott L. (1996). Elementary Sampling Survey, 5th ed., Duxbury, New York.
22. Smith, T. M. F. (1976). The Foundation of Survey Sampling: A review, *Journal of Royal Statistics Society* A139, p 183-204
23. Sturat, A. (1984). The Ideas of Sampling, Revised Edition, Griffin, London.
24. Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S., and Ashok, C. (1984). Samplmg Theory of Surveys with Applications, 3d ed., Ames (Iowa): Iowa State University Press.
25. Thompson, S. K. (2002). Sampling, 2end Wiley, New York.
26. Tryfos, P. (1996). Sampling Methods for Applied Research, Wiley, New York.
27. Yates, F. (1981). Sampling Methods for Censuses and Surveys 4th Ed., Grittin, London.

obeikandl.com

الفصل السادس

العينة العشوائية الطبقية

Stratified Random Sample

1.6 مقدمة

إن المسح بطريقة العينة العشوائية الطبقية يُعد من أكثر طرق المعاينة استخداماً. حيث يُقسم المجتمع في المعاينة العشوائية الطبقية الذي يتكون منN من الوحدات إلى K من المجتمعات الجزئية تدعى طبقات. إن المجتمع الجزئي *i* يحتوي على N_i من الوحدات $i=1, 2, \ldots, K$ يجب أن تكون المجتمعات الجزئية غير متداخلة حتى يكونوا المجتمع أي:

$$N= N_1+ N_2+ \ldots+N_k$$

تُعرف العينة العشوائية الطبقية على أنها: العينة التي اختيرت عند تقسيم عناصر المجتمع إلى مجتمعات جزئية وغير متداخلة تدعى طبقات، وبعد ذلك نختار عينة عشوائية بسيطة من كل طبقة. نختار العينة من كل طبقة بصورة مستقلة، حجم العينة الذي يختار من الطبقة i هو n_i و $i=1, 2, \ldots, K$ بحيث $n = n_1+n_2+\ldots+n_k$. إذا العينة اختيرت من كل طبقة بشكل عشوائي عندئذ تسمى العينة العشوائية الطبقية.

إن الهدف الرئيس من الطبقية ليعطينا فهم ودراية أفضل بأجزاء مختلفة من المجتمع ومن ثم نحصل على دقة أعلى. للوصول إلى تحقيق الهدف يجب أن

obeikandl.com

يقسم المجتمع إلى طبقات، ويحدد عدد الطبقات، وكذلك توزع العينة بين الطبقات، وأخيراً تحلل البيانات على أنها من عينة طبقية.

2.6 أسس تقسيم المجتمع إلى طبقات

يمكن تلخيص الأسس التي تتبع لتقسيم المجتمع إلى طبقات بما يأتي:

1. يجب أن تكون الطبقات غير متداخلة، ويجب أن تؤلف مع بعضها المجتمع كاملاً.
2. يجب أن تتم عملية تقسيم المجتمع إلى طبقات بحيث إن كل طبقة تكون متجانسة بالنسبة إلى الصفة التي تجري دراستها.
3. عندما تصعب عملية تقسيم المجتمع إلى طبقات بناء على الصفة التي تجري دراستها يكون تقسيم المجتمع إلى طبقات على أساس ملاءمة التقسيم بالنسبة إلى النواحي الإدارية.
4. إذا أعطيت حدود الدقة إلى بعض المجتمعات الجزئية من الأفضل أن تعامل هذه المجتمعات الجزئية على أنها طبقات.

3.6 مزايا تقسيم المجتمع إلى طبقات

إن تقسيم المجتمع إلى طبقات له مزايا عديدة أهمها:

1. ربما يكون تقسيم المجتمع إلى طبقات مفضلاً ذلك لأنه ملائم من النواحي الإدارية. حيث إن الجهة التي تقوم بالمسح يمكن أن تجد عاملين في مختلف المناطق الإدارية، وبالتالي يؤدي إلى تنظيم ومراقبة ومتابعة أفضل. وبالتالي ربما سيؤدي إلى تقليل تكاليف مشاهدة الوحدة الواحدة.
2. تقسيم المجتمع بناء على الصفات الطبيعية يساعد على تحسين تصميم المعاينة. فمثلاً في حالة مسح الحقول الزراعية ربما تكون هنالك

obeikandl.com

مشكلات معاينة في مناطق السهول تختلف عما هو موجود في الصحراء أو المناطق الجبلية. لذلك يكون من الأفضل أن يقسم المجتمع إلى طبقات بحسب المناطق الجغرافية المختلفة.

3. إذا كان المجتمع مكوناً من عدد كبير من العناصر سوف يؤدي تقسيمه إلى طبقات إلى تقليل التشتت داخل الطبقات، ومن ثم سوف نحصل على تقديرات أكثر دقة باستخدام العينة الطبقية، وذلك من خلال القيام بالتقديرات داخل كل طبقة، ومن ثم جمعها مع بعضها البعض للوصول إلى التقدير النهائي.
4. إن تقسيم المجتمع إلى طبقات يعطي مجموعات مختلفة من المجتمع فرصة للظهور في العينة والتي قد تكون مهمة. ويساعد على اختيار أفضل لأجزاء مختلفة للمجتمع.
5. نحصل على دقة أكبر للتقديرات المختلفة لمعلمات المجتمع، وذلك لأن المجتمع يكون غير متجانس بصورة عامة، ولكن عندما يقسم إلى أجزاء (طبقات) متجانسة مع بعضها، هذا يؤدي إلى الحصول على تقديرات أكثر دقة داخل كل طبقة ومن ثم إلى المجتمع كله.

4.6 رموزومصطلحات

لنفترض أن i يرمز إلى الطبقة وj يرمز إلى وحدات المعاينة في الطبقة i، المصطلحات الآتية تعود إلى الطبقة i

N_i : عدد الوحدات في الطبقة

n_i : عدد الوحدات في العينة المسحوبة من الطبقة

$W_i=N_i/N$: وزن الطبقة

f_i : كسر المعاينة في الطبقة

لنفترض أن y_{ij} يمثل القيمة للوحدة j الموجودة في الطبقة i

obeikandl.com

$Y_i = \sum_{j=1}^{N_i} y_{ij}$: المجموع الكلي للطبقة

$\overline{Y}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}$: الوسط الحسابي للطبقة

$\overline{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$: الوسط الحسابي للعينة المسحوبة من الطبقة

$Y_i = N_i \overline{y}_i$: تقدير المجموع الكلي في الطبقة

$S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (y_{ij} - \overline{Y}_i)^2$: تباين الطبقة

$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_i)^2$: تباين العينة المسحوبة من الطبقة

$Var(\overline{y}_i) = (1 - f_i) \frac{S_i^2}{n_i}$: تباين الوسط الحسابي للعينة المسحوبة من الطبقة

$Var(Y_i) = \frac{N_i^2}{n_i} (1 - f_i) S_i^2$: تباين تقدير المجموع الكلي للطبقة

$Y = \sum_{i=1}^{K} \sum_{j=1}^{N_i} y_{ij}$: المجموع الكلي للمجتمع

$\overline{Y} = \frac{Y}{N}$: الوسط الحسابي للمجتمع

5.6 تقدير الوسط الحسابي والمجموع الكلي للمجتمع

1.5.6 تقدير الوسط الحسابي للمجتمع

إن الوسط الحسابي للعينة الطبقية $\overline{y}_{st}$ هو التقدير البدهي للوسط الحسابي للمجتمع $\overline{Y}$ حيث يعرف

$$\overline{y}_{st} = \sum_{i=1}^{K} \frac{N_i}{N} \overline{y}_i = \sum_{i=1}^{K} W_i \overline{y}_i$$

إنه من السهولة ملاحظة أن $\overline{y}_{st}$ تقدير غير متحيز إلى $\overline{Y}$ يراجع Cochran(1977) أو ترجمة أنيس كنجو (1995) لهذا البرهان الرياضي

obeikandi.com

وكذلك جميع البراهين الرياضية الأخرى في هذا الفصل. أما تباين $\bar{y}_{st}$ فيمكن إيجاده بسهولة

$$Var(\bar{y}_{st}) = \sum_{i=1}^{K} W_i^2 (1 - f_i) S_i^2 / n_i = \sum_{i=1}^{K} \frac{W_i^2 S_i^2}{n_i} - \sum_{i=1}^{K} \frac{W_i S_i^2}{N}$$

إن إيجاد تباين الوسط الحسابي للعينة الطبقية يقتضي معرفة تباين الطبقات S_i^2 وهذا غير ممكن لأنه في الغالب S_i^2 غير معروفة؛ لذلك لا بد من تقدير تباين الوسط الحسابي للعينة الطبقية وذلك باستخدام s_i^2 تباين العينة المسحوبة من الطبقة بدلاً من S_i^2 تباين الطبقة، ليصبح تقدير تباين $\bar{y}_{st}$ كما يأتي:

$$s^2(\bar{y}_{st}) = \sum_{i=1}^{K} W_i^2 (1 - f_i) s_i^2 / n_i = \sum_{i=1}^{K} \frac{W_i^2 s_i^2}{n_i} - \sum_{i=1}^{K} \frac{W_i s_i^2}{N}$$

كذلك فإن تقدير الخطأ المعياري:

$$s(\bar{y}_{st}) = \sqrt{s^2(\bar{y}_{st})}$$

أما معامل التغير إلى $\bar{y}_{st}$ فيمكن إيجاده:

$$CV = s(\bar{y}_{st}) / \bar{y}_{st}$$

وأخيراً يمكن استخدام التوزيع الطبيعي التقريبي إذا كان حجم العينة كبيرًا لإيجاد 100% (1-*a*) فترة ثقة للوسط الحسابي للمجتمع:

$$\bar{y}_{st} \pm z_{\alpha/2}\, s(\bar{y}_{st})$$

2.5.6 تقدير المجموع الكلي للمجتمع

لتقدير المجموع الكلي للمجتمع يمكن إيجاده باستخدام:

$$Y_{st} = N \bar{y}_{st} = \sum_{i=1}^{K} N_i \bar{y}_i$$

obeikandl.com

أما تباين Y_{st} فيمكن إيجاده

$$Var(Y_{st}) = \sum_{i=1}^{K} \frac{N_i^2}{n_i}(1 - f_i)S_i^2$$

كذلك لا بد من تقدير تباين تقدير المجموع الكلي حيث (وكما مر معنا قبل قليل) إن S_i^2 غير معروفة ولا بد من تقديرها باستخدام s_i^2 ليصبح تقدير تباين Y_{st} كما يأتي

$$s^2(Y_{st}) = \sum_{i=1}^{K} \frac{N_i^2}{n_i}(1 - f_i)s_i^2 = \sum_{i=1}^{K} \frac{N_i^2 s_i^2}{n_i} - \sum_{i=1}^{K} N_i s_i^2$$

مثال (1): لتطبيق المعادلات التي وردت في هذا الفصل، لنفترض أن لدينا الجدول الآتي:

الطبقة	n_i	$\bar{y}_i$	s_i^2	N_i	W_i	Y_i
1	40	50	225	700	0.2	35000
2	50	30	64	1750	0.5	52500
3	40	20	36	1050	0.3	21000

الحل:

تقدير المجموع الكلي

$$Y_{st} = \sum_{i=1}^{3} N_i \bar{y}_i = \sum_{i=1}^{3} Y_i = 108500$$

تقدير تباين تقدير المجموع الكلي

$$s^2(Y_{st}) = \sum_{i=1}^{K} \frac{N_i^2 s_i^2}{n_i} - \sum_{i=1}^{K} N_i s_i^2 = 7361200$$

تقدير الخطأ المعياري لتقدير المجموع الكلي

$$s(Y_{st}) = \sqrt{s^2(Y_{st})} = 2173.13$$

openkandil.com

الوسط الحسابي للعينة الطبقية

$$\bar{y}_{st} = \sum_{i=1}^{3} W_i \bar{y}_i = \frac{Y_{st}}{N} = 31$$

تقدير تباين الوسط الحسابي للعينة الطبقية

$$s^2(\bar{y}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 s_i^2}{n_i} - \sum_{i=1}^{K} \frac{W_i^2 s_i^2}{N} = 0.6009$$

تقدير الخطأ المعياري

$$s(\bar{y}_{st}) = \sqrt{s^2(\bar{y}_{st})} = 0.78$$

تقدير معامل التغير إلى $\bar{y}_{st}$

$$CV = \frac{s(\bar{y}_{st})}{\bar{y}_{st}} = \frac{0.78}{31} = 0.025$$

فترة 95% ثقة للوسط الحسابي $\bar{Y}$

$$\bar{y}_{st} \pm z_{\alpha/2} s(\bar{y}_{st}) = 31 \pm 1.96(0.78); (29.47, 32.53)$$

6.6 توزيع حجم العينة بين الطبقات

هنالك عدة طرق لتوزيع حجم العينة n بين الطبقات، ولكن سوف نتناول أهم طريقتين. المعادلات التي أعطيت سابقاً تعتمد على قيمة n_i ، بحيث إن

$$n = \sum_{i=1}^{K} n_i$$

1.6.6 التوزيع المتناسب (Proportional Allocation)

إن فكرة التوزيع المتناسب تقوم على أساس توزيع حجم العينة n بين الطبقات بحيث يكون نصيب كل طبقة من n متناسباً مع حجم الطبقة في

obeikandl.com

المجتمع، أي أن $n_i = nW_i$ وإذا عوضنا قيمة في المعادلة الخاصة بتباين الوسط الحسابي للعينة الطبقية نحصل على

$$V_{prop}(\bar{y}_{st}) = \sum_{i=1}^{K} \frac{W_i S_i^2}{n} - \frac{1}{N} \sum_{i=1}^{K} W_i S_i^2$$

وهذا يسمى تباين الوسط الحسابي للعينة الطبقية باستخدام التوزيع المتناسب.

2.6.6 التوزيع الأمثل (Optimum Allocation)

لإيجاد أفضل توزيع لحجم العينة n في الحقيقة هو عبارة عن تصغير Minimizing دالة التباين $Var(\bar{y}_{st})$ لحجم العينة مع بعض الشروط. ويمكن أن نفترضها دالة لحجم العينة n_i

$$Var(\bar{y}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 S_i^2}{n_i} - \sum_{i=1}^{K} \frac{W_i S_i^2}{N}$$

للحصول على أقل تباين ممكن راجع Cochran (1977) أو ترجمة أنيس كنجو(1995) للبرهان وتفصيلات أخرى. إذا كانت تكاليف مشاهدة الوحدة الواحدة في جميع الطبقات متساوية فإن التوزيع الأمثل لحجم العينة بين الطبقات يكون

$$n_i = n \frac{W_i S_i}{\sum_{i=1}^{K} W_i S_i}$$

إذا عوضنا عن قيمة n_i في تباين الوسط الحسابي للعينة الطبقية نحصل على ما يسمى بالتباين الأمثل

$$V_{opt}(\bar{y}_{st}) = \frac{[\sum_{i=1}^{K} W_i S_i]^2}{n} - \sum_{i=1}^{K} \frac{W_i S_i^2}{N}$$

obeikandi.com

كذلك يمكن الحصول على أفضل (أصغر) n مع كون التباين ثابتاً ومساوياً إلى V

$$n = \frac{[\sum_{i=1}^{K} W_i S_i]^2}{V + \frac{1}{N}\sum_{i=1}^{K} W_i S_i^2}$$

مثال (2): لنعد إلى المثال **(1)** ونقارن بين قيمة التباين للحالات الثلاثة أي توزيع n الحقيقي كما وردت في المثال وباستخدام التوزيع المتناسب والتوزيع الأمثل

الطبقة	التوزيع الحقيقي	التوزيع الأمثل	التوزيع المتناسب
1	40	44	26
2	50	59	65
3	40	27	39
مجموع حجم العينة	130	130	130
التباين	0.6009	0.570	0.6503

أما إذا كانت تكاليف مشاهدة الوحدة في الطبقات المختلفة غير متساوية فلا بد من معرفة دالة التكاليف، وتكاليف مشاهدة الوحدة الواحدة في كل طبقة. إن دالة التكاليف الشائعة الاستعمال في العينة الطبقية هي:

$$C = C_0 + \sum_{i=1}^{K} c_i n_i$$

حيث إن C تمثل إجمالي الميزانية، و C_0 تكاليف ثابتة و c_i تكاليف مشاهدة الوحدة الواحدة في الطبقةi. وباستخدام أسلوب التصغير(Minimizing) لدالة

التباين باستخدام دالة التكاليف نحصل على التوزيع الأمثل لحجم العينة بين الطبقات كما يأتي:

$$n_i = n \frac{W_i S_i / \sqrt{c_i}}{\sum_{i=1}^{K} W_i S_i / \sqrt{c_i}}$$

أما إذا كانت ميزانية المسح للعينة ثابتة أي أن

$$C' = C - C_o = \sum_{i=1}^{K} c_i n_i$$

باستخدام هذه الميزانية لا بد من إيجاد حجم العينة الأمثل وهو

$$n = C' \frac{\sum_{i=1}^{K} W_i S_i / \sqrt{c_i}}{\sum_{i=1}^{K} W_i S_i \sqrt{c_i}}$$

إذا كان التباين ثابتاً ومساوياً إلى V نجد قيمة n التي تعطينا التباين المطلوب

$$n = \frac{[\sum_{i=1}^{K} W_i S_i \sqrt{c_i}][\sum_{i=1}^{K} W_i S_i / \sqrt{c_i}]}{V + \frac{1}{N} \sum_{i=1}^{K} W_i S_i^2}$$

مثال (3) : لنعد إلى المثال (1) ولنفترض أن تكاليف مشاهدة الوحدة الواحدة في الطبقات هي

$c_1 = 1, c_2 = 4, c_3 = 9$ أوجد تقدير التباين المثالي باستخدام هذه العلاقة

$$n_i = n \frac{W_i S_i / \sqrt{c_i}}{\sum_{i=1}^{K} W_i S_i / \sqrt{c_i}}$$

obeikandl.com

وجدنا $n_1 = 70, n_2 = 46, n_3 = 14$. الآن نعوض عن قيمة n_i في المعادلة الآتية لنجد تقديرًا لقيمة التباين الأمثل مع كون تكاليف مشاهدة الوحدة مختلفة من طبقة إلى أخرى.

$$s^2(\bar{y}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 s_i^2}{n_i} - \sum_{i=1}^{K} \frac{W_i s_i^2}{N} = 0.6827$$

لابد من الإشارة في نهاية هذا الجزء إلى أن

$$Var(\bar{y}) \geq V_{prop}(\bar{y}_{st}) \geq V_{opt}(\bar{y}_{st})$$

لمزيد من المعلومات راجع Cochran (1977) أو ترجمة أنيس كنجو(1995).

7.6 تقدير نسبة المجتمع

تقدير النسبة للمجتمع يمكن إيجاده بسهولة باستخدام التعريف الآتي

$y_{ij} = 1$ إذا كان العنصر j الموجود في الطبقة i يحمل صفة مميزة

$y_{ij} = 0$ خلاف ذلك أي إن العنصر j الموجود في الطبقة i لا يحمل الصفة المميزة

سوف نستخدم P_i للتعبير عن النسبة في الطبقة i وP للتعبير عن النسبة في المجتمع وكذلك $\hat{p}_i = \frac{a_i}{n_i}$ للتعبير عن تقدير النسبة في الطبقة i ، حيث إن a_i يمثل مجموع الوحدات من الذين يحملون الصفة المميزة في العينة المسحوبة من الطبقة. نستطيع الآن أن نعرف تقدير النسبة في العينة العشوائية الطبقية بما يأتي

$$\hat{P}_{st} = \sum_{i=1}^{K} W_i \hat{P}_i = \sum_{i=1}^{K} \frac{N_i}{N} \frac{a_i}{n_i}$$

يمكن أن نبرهن بسهولة أن $\hat{p}_{st}$ تقدير غير متحيز إلى نسبة المجتمع P وبتباين

$$Var(\hat{p}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 P_i Q_i}{n_i} - \sum_{i=1}^{K} \frac{W_i P_i Q_i}{N}$$

obeikandl.com

حيث إن $Q_i = 1 - P_i$. ولكن P_i غير معلومة لذلك لابد من تقدير $Var(\hat{p}_{st})$ بما يأتي

$$s^2(\hat{p}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 \hat{p}_i \hat{q}_i}{n_i} - \sum_{i=1}^{K} \frac{W_i \hat{p}_i \hat{q}_i}{N}$$

و معامل التغير لتقدير النسبة في العينة العشوائية الطبقية

$$CV(\hat{p}_{st}) = \frac{s(\hat{p}_{st})}{\hat{p}_{st}}$$

حيث إن

$$s(\hat{p}_{st}) = \sqrt{s^2(\hat{p}_{st})}$$

ويسمى تقدير الخطأ المعياري للعينة الطبقية.

ملاحظة: يمكن بسهولة تقدير مجموع التكرارات للعينة العشوائية الطبقية باستخدام

$$A_{st} = N \hat{p}_{st}$$

تباين مجموع التكرارات

$$Var(A_{st}) = N^2 Var(\hat{p}_{st})$$

تقدير تباين مجموع التكرارات

$$s^2(A_{st}) = N^2 s^2(\hat{p}_{st})$$

معامل التغير لمجموع التكرارات

$$CV(A_{st}) = \frac{s(A_{st})}{A_{st}}$$

مثال (4): استخدام الجدول الآتي لحساب النتائج التي وصلنا إليها فيما يخص تقدير النسبة، ومجموع التكرارات

obeikandl.com

الطبقة	N_i	n_i	W_i	p_i
1	700	40	0.2	0.5
2	1750	50	0.5	0.8
3	1050	40	0.3	0.1
المجموع	3500	130		

الحل:

تقدير النسبة الطبقية

$$\hat{P}_{st} = \sum_{i=1}^{3} W_i \hat{P}_i = 0.53$$

تقدير تباين النسبة الطبقية

$$s^2(\hat{p}_{st}) = \sum_{i=1}^{K} \frac{W_i^2 \hat{p}_i \hat{q}_i}{n_i} - \sum_{i=1}^{K} \frac{W_i \hat{p}_i \hat{q}_i}{N} = 0.0012$$

تقدير الخطأ المعياري للنسبة الطبقية

$$s(\hat{p}_{st}) = \sqrt{s^2(\hat{p}_{st})} = 0.0351$$

معامل التغير للنسبة الطبقية

$$CV(\hat{p}_{st}) = \frac{s(\hat{p}_{st})}{\hat{p}_{st}} = \frac{0.0351}{0.531} = 0.066$$

تقدير مجموع التكرارات

$$A_{st} = N\hat{p}_{st} = 3500(0.53) = 1855$$

تقدير تباين مجموع التكرارات

$$s^2(A_{st}) = N^2 s^2(\hat{p}_{st}) = 15123.13$$

معامل التغير لمجموع التكرارات

$$CV(A_{st}) = \frac{s(A_{st})}{A_{st}} = 0.066$$

obeikandl.com

إذا استخدمنا التوزيع المتناسب لتوزيع n بين الطبقات أي $n_i = nW_i$ نحصل على التباين المتناسب

$$V_{prop}(\hat{p}_{st}) = \sum_{i=1}^{K} \frac{W_i P_i Q_i}{n} - \sum_{i=1}^{K} \frac{W_i P_i Q_i}{N}$$

إذا استخدمنا التوزيع الأمثل وكذلك

$$S_i^2 = \frac{N_i}{N_i - 1} P_i Q_i \approx P_i Q_i$$

وإذا كانت التكاليف متساوية فإن

$$n_i = n \frac{W_i \sqrt{P_i Q_i}}{\sum_{i=}^{K} W_i \sqrt{P_i Q_i}}$$

أما إذا كانت التكاليف غير متساوية فإن

$$n_i = n \frac{W_i \sqrt{P_i Q_i} / \sqrt{c_i}}{\sum_{i=}^{K} W_i \sqrt{P_i Q_i} / \sqrt{c_i}}$$

أخيراً إذا كانت التكاليف متساوية نحصل على التباين الأمثل إلى $\hat{p}_{st}$ كما يأتي

$$V_{opt}(\hat{p}_{st}) = \frac{1}{n} \sum_{i=1}^{K} [W_i \sqrt{P_i Q_i}]^2 - \sum_{i=1}^{K} \frac{W_i P_i Q_i}{N}$$

8.6 اختيار حجم العينة

إن مقدار المعلومات الموجودة في العينة يعتمد على حجم العينة n بما أن $Var(\bar{y}_{st})$ يتناقص عندما يزداد حجم العينة n دعنا نختار الطريقة التي يمكن

obeikandl.com

أن نختار فيها حجم عينة معين لإيجاد كمية ثابتة من المعلومات لتقدير معلمات المجتمع.

1.8.6 اختيار حجم العينة لتقدير الوسطي الحسابي والمجموع الكلي للمجتمع

لنفترض أننا حددنا أن التقدير $\bar{y}_{st}$ يجب أن يكون ضمن d وحدة من الوسط الحسابي للمجتمع $\bar{Y}$ باحتمال تقريبا 0.95، أي إن حد خطأ التقدير للوسط الحسابي الطبقي للعينة هو d ودرجة الثقة 0.95 بعبارة أخرى

$$2\sqrt{\mathrm{Var}(\bar{y}_{st})} = d$$

أو

$$\mathrm{Var}(\bar{y}_{st}) = d^2/4$$

الملاحظ أن المعادلة الأخيرة تعتمد على $S_1^2, S_2^2, \ldots, S_k^2$ وهذه غير معروفة، ولكن يمكن استبدال هذه القيم بـ $s_1^2, s_2^2, \ldots, s_k^2$، إذا كان حجم المجتمع N كبير، كذلك نحن وضعنا: $\mathrm{Var}(\bar{y}_{st}) = d^2/4$ في الحقيقة لا نستطيع حل هذه المعادلة لإيجاد قيمة n ما لم نعلم العلاقة بين n و $n_1, n_2, \ldots, n_k$ وطبعاً هناك عدة طرق لتوضيح العلاقة، منها ما يسمى بالتوزيع المتناسب أي إن $n_i = nW_i$. وباستخدام هذه المعادلة نستطيع حل المعادلة $\mathrm{Var}(\bar{y}_{st}) = d^2/4$. وبالطريقة نفسها نستطيع إيجاد قيمة n لتقدير المجموع الكلي Y وذلك بحل المعادلة $\mathrm{Var}(\bar{y}_{st}) = d^2/4N^2$.

الآن باستطاعتنا إيجاد القيمة التقريبية لحجم العينة n لتقدير $\bar{Y}$ أو Y مع حد لخطأ التقدير مقداره d

$$n \geq \frac{\sum_{i=1}^{K} N_i^2 S_i^2 / W_i}{N^2 D + \sum_{i=1}^{K} N_i S_i^2}$$

obeikandl.com

عندما نقدر $\overline{Y}$ فإن $D = d^2/4$ وعندما نقدر Y فإن $D = d^2/4N^2$

لا بد من معرفة $S_1^2, S_2^2, ..., S_k^2$ وإلا لا بد من تقديرها باستخدام الطرق نفسها التي تكلمنا عنها في الفصل الخامس. إحدى هذه الطرق نقوم فيها باستخدام تباين العينة الأولية المسحوبة من الطبقة i للحصول على s_i^2 و $i = 1, 2, ..., k$.

مثال (5): في المثال (1) لنفترض أنه من التجارب السابقة فإن تباين المجتمع للطبقات الثلاثة $S_1^2 = 300, S_2^2 = 81, S_3^2 = 40$ ، نرغب بتقدير الوسط الحسابي للمجتمع باستخدام الوسط الحسابي للعينة الطبقية، ما هو حجم العينة الذي نحتاجه بحيث إن الحد في مقدار الخطأ للتقدير يساوي 2.

إن حجم العينة المطلوب

$$n \geq \frac{\sum_{i=1}^{K} N_i^2 S_i^2 / W_i}{N^2 D + \sum_{i=1}^{K} N_i S_i^2} \approx 109$$

2.8.6 اختيار حجم العينة لتقدير نسبة المجتمع

لتقدير نسبة المجتمع أولاً نحدد مقدار المعلومات التي نرغب بها، وذلك من خلال تحديد مقدار حد الخطأ في تقدير النسبة، ومن ثم نختار حجم العينة بناء على ذلك إن الطريقة التي تستخدم لإيجاد حجم العينة n عندما يحدد الحد في خطأ التقدير مشابه لتلك التي تكلمنا عنها في الفقرة أعلاه. مع تغير واحد وهو أن $S_i^2 = P_i Q_i$ لذلك تصبح المعادلة كما يأتي

$$n \geq \frac{\sum_{i=1}^{K} N_i^2 P_i Q_i / W_i}{N^2 D + \sum_{i=1}^{K} N_i P_i Q_i}$$

حيث إن $D = d^2/4$.

مثال (6): في المثال (4) إذا علمنا من تجارب سابقة أن $P_1 = P_3 = 0.25$ وكذلك $P_2 = 0.5$ وإذا حددنا مقدار الخطأ في التقدير $d = 0.05$ أوجد حجم العينة

الحل:

$$n \geq \frac{\sum_{i=1}^{K} N_i^2 P_i Q_i / W_i}{N^2 D + \sum_{i=1}^{K} N_i P_i Q_i} = 319$$

9.6 تحديد عدد الطبقات

لتحديد عدد الطبقات هناك نقطتان يجب ملاحظتهما:

1. كيف سينقص $\text{var}(\bar{y}_{st})$ عندما تزداد عدد الطبقات.
2. كيف ستتأثر تكاليف المسح عندما يتغير عدد الطبقات.

النقطة الأولى نوقشت من قبل عدة باحثين منهم Cochran (1961) و Sethi (1963) و Hess et. al. (1966) وتوصلوا إلى عدة نماذج لتحديد عدد الطبقات. أما النقطة الثانية: لقد درس هذا الموضوع من قبل Sethi (1963) واقترح أن زيادة عدد الطبقات أكثر من 6 طبقات قلما يكون مفضلاً.

يمكن إثبات أن تقسيم المجتمع إلى طبقات يؤدي إلى زيادة الكفاءة إذا ارتفع عدد الطبقات. لكن الحد الأعلى لعدد الطبقات يكون بمقدار حجم العينة، بحيث نصل إلى النقطة التي لا نستطيع فيها أن نسحب سوى وحدة واحدة من كل طبقة. قد لا تكون هناك فائدة من زيادة عدد الطبقات في المسوحات الكبيرة إلى الحد الأعلى لأنه سيؤدي إلى زيادة التكاليف.

obeikandl.com

10.6 تقسيم المجتمع إلى طبقات بعد سحب العينة

في بعض الأحيان تظهر مشكلة عند المعاينة ألا وهي رغبتنا في تقسيم المجتمع إلى طبقات حسب متغير مهم، ولكن لا نستطيع وضع وحدات المعاينة في الطبقات المناسبة إلا بعد سحب العينة، على سبيل المثال ربما نرغب في تقسيم المجتمع بناء على الجنس لمعرفة آرائهم في قضية معينة، ونقوم بالمسح باستخدام الهاتف لكن الشخص الذي سيرد على الهاتف لا يمكن معرفة جنسه إلا بعد الاتصال به.

لنفترض أن عينة عشوائية بسيطة سحبت من أفراد المجتمع لغرض معرفة آرائهم في قضية معينة. هنا لا يمكن توزيع العينة بين الطبقتين إلى n_1 رجال و n_2 نساء إلا بعد مقابلة الأفراد الذين سحبوا في العينة. والآن بدلاً من استخدام $\bar{y}$ الوسط الحسابي للعينة العشوائية البسيطة كتقدير إلى الوسط الحسابي للمجتمع $\bar{Y}$ يمكن أن نستخدم $\bar{y}_{st}$ الوسط الحسابي للعينة الطبقية شريطة أن N_i/N معروفة للرجال والنساء. ولكن n_1 , n_2 في هذه الحالة متغيرات لأنهم يتغيرون من مسح إلى آخر حتى ولو كانت n، ثابتة لذلك فإن هذه العينة ليست بالضبط هي عينة عشوائية طبقية حسب تعريف العينة العشوائية الطبقية. ولكن إذا كان N_i/N معروفاً وإذا كانت $n_i > 20$ وحدة من كل طبقة، فإن هذه الطريقة لتقسيم المجتمع إلى طبقات بعد سحب العينة تكون قريبة من دقة المعاينة العشوائية الطبقية مع استخدام التوزيع المتناسب لتوزيع العينة بين الطبقات.

openkanal.com

تمارين

1. ترغب إحدى الشركات في تقدير عدد ساعات العمل التي تضيع بسبب الحوادث التي يتعرض لها العاملون في الشركة، وبما أن العمال والفنيين والإداريين يختلف معدل الحوادث لكل فئة منهم لذلك قرر الباحث سحب عينة عشوائية طبقية بحيث إن كل فئة من العاملين الثلاثة تمثل طبقة. لدينا بيانات من دراسات سابقة تعطينا النتائج الآتية

الطبقة	S_i^2	N_i
العمال	36	132
الفنيون	25	92
الإداريون	9	27

إذا قرر الباحث سحب عينة عشوائية بحجم n=35 حدد توزيع العينة بين الطبقات باستخدام التوزيع الأمثل والتوزيع المتناسب.

2. قام الباحث بسحب عينة عشوائية بسيطة من كل طبقة للسؤال السابق، وحصل على عدد الساعات الضائعة خلال الشهر الماضي كما هي موضحة في الجدول الآتي:

الطبقة	عدد الساعات الضائعة
العمال	8, 0, 24, 20, 0, 16, 12, 0, 5, 6, 7, 8, 2, 5, 4, 3, 2
الفنيون	4, 5, 0, 8, 7, 6, 1, 3, 2, 0, 4, 1, 7, 24
الإداريون	8, 1, 0, 4

أ. قدر معدل عدد الساعات الضائعة خلال الشهر. وأوجد 90% فترة ثقة للوسط الحسابي.

obeikandl.com

ب. استخدم التوزيع الأمثل والمتناسب لتقدير عدد الساعات الضائعة وقارن بين النتائج التي تحصل عليها هنا ونتائج الفرع أعلاه.

ت. استخدم العينة العشوائية البسيطة لتقدير عدد ساعات العمل الضائعة وتقدير التباين، هل تعتقد

أن استخدام العينة الطبقية أفضل من العينة العشوائية البسيطة؟ وضح إجابتك.

3. لنفترض أن البيانات في السؤال الثاني أعلاه تمثل عينة عشوائية أولية، استخدمها لتحديد حجم العينة المطلوب، إذا كان حد الخطأ لتقدير الوسط الحسابي الطبقي d=1.5 ودرجة ثقة 95% .

4. استخدم البيانات المعطاة في السؤال الثاني لتحديد حجم العينة المطلوبة إذا كانت d= 1.5 ودرجة الثقة 90%.

5. إذا كانت نسبة العاملين الذين تتجاوز أعمارهم 50 سنة من السؤال الثاني هي كما يأتي: العمال 20% والفنيون 20% والإداريون 35% قدر نسبة العاملين في الشركة الذين تزيد أعمارهم عن 50 سنة، ثم أوجد فترة 95% ثقة للنسبة الحقيقية.

6. يحتوي مجتمع على 4 طبقات، الجدول الآتي يعطينا معلومات عن الطبقات في المجتمع

الطبقة	n_i	s_i^2	S_i^2	$\bar{y}_i$	N_i
1	25	1.5	1.3	7.4	120
2	20	2.7	2.1	7	100
3	15	0.9	1.1	11	75
4	10	2.1	1.9	9.1	40

أ. أحسب $\bar{y}_{st}$ و Y_{st} و $var(\bar{y}_{st})$ و $var(Y_{st})$ و $s^2(\bar{y}_{st})$ و $s^2(Y_{st})$.

ب. استخدم التوزيع المتناسب والتوزيع الأمثل لتوزيع العينة بين الطبقات، ثم أوجد تباين الوسط الحسابي الطبقي للحالتين، ثم

قارن النتائج التي حصلت عليها هنا مع النتائج التي حصلت عليها في أ.

ت. استخدم البيانات التي لديك لتحديد حجم العينة المطلوب لتقدير المجموع الكلي إذا حدد مقدار حد الخطأ d=400 ودرجة الثقة95%.

7. لنفترض أن تكاليف معاينة الوحدة الواحدة في الطبقات الخمسة الواردة في السؤال السادس هي كما يأتي: $c_1 = 3, c_2 = 2.5, c_3 = 2, c_4 = 1.5$ استخدم التوزيع الأمثل لتوزيع حجم العينة بين الطبقات، وكذلك أوجد تقدير الوسط الحسابي الطبقي وتباينه.

8. إذا كانت تكاليف معاينة الوحدة الواحدة كما وردت في السؤال السابع وتقدير النسبة P_i للطبقة المختلفة هي كما يأتي: $\hat{P}_1 = 0.2, \hat{P}_2 = 0.3, \hat{P}_3 = 0.15, \hat{P}_4 = 0.12$ قدر النسبة الحقيقية باستخدام النسبة الطبقية $\hat{p}_{st}$ ثم أوجد معامل التغير إلى $\hat{p}_{st}$.

9. قام الباحثون الاقتصاديون في أحد البحوث بتصنيف القرى الموجودة في منطقة جغرافية معينة على أربع طبقات، بناءً على ارتفاع القرية عن مستوى سطح البحر. سحبت عينة عشوائية بسيطة حجمها 10 قرى (من كل طبقة) بما فيها القرى غير المأهولة بالسكان، وجرت عملية حصر عدد المساكن الموجودة في كل قرية، وحصلنا على النتائج المبينة في الجدول الآتي:

عدد المساكن في كل قرية ظهرت في العينة

الطبقة	عدد القرى	1	2	3	4	5	6	7	8	9	10
1	1400	43	84	98	0	10	44	0	124	13	0
2	4700	50	147	62	87	84	158	170	104	56	160
3	2500	228	262	110	232	139	178	334	0	63	220
4	14900	17	34	25	34	36	0	25	7	15	31

obeikandl.com

أ. قدر مجموع المساكن الموجودة في المنطقة وكذلك قدر التباين للمجموع الكلي .

ب. قارن التوزيع الحالي لحجم العينة بين الطبقات والتوزيع الأمثل لحجم العينة بين الطبقات.

ت. قدر الزيادة في الدقة الناتجة باستخدام العينة العشوائية الطبقية بالمقارنة بالعينة العشوائية البسيطة .

10. ترغب إحدى المدارس الابتدائية بتقدير علامات الطلبة في القراءة لطلبة الصف السادس الابتدائي، قامت المدرسة بتقسيم الطلاب إلى ثلاث مجموعات، المجموعة الأولى تحتوي على الطلبة سريعي التعلم، والمجموعة الثالثة على الطلبة بطيئي التعلم. تحتوي المجموعة الأولى على 55 طالباً والمجموعة الثانية على80، والمجموعة الثالثة على 65 طالباً. قامت المدرسة بسحب عشوائية طبقية بحجم 30 طالباً موزعة على المجموعات الثلاث، الجدول الآتي يعطينا نتائج الطلاب الذين سُحبوا من المجموعات (الطبقات) الثلاث.

المجموعة الأولى	91 78 85 90 80 77 7559
المجموعة الثانية	81 67 54 56 64 72 59 61 42 49 60 68
المجموعة الثالثة	61 47 45 33 67 56 44 42 50 19

أ. قدر معدل علامات الطلاب للصف السادس ومن ثم أوجد 95% فترة ثقة لمعدل علامات الطلاب الحقيقي.

ب. قدر الفرق بعلامات الطلاب بين المجموعة الأولى والثالثة ومن ثم أوجد 95% فترة ثقة للفرق الحقيقي بين المجموعتين.

11. لنفترض أن المدرسة ترغب في تقدير علامات الطلاب في نهاية العام الدراسي وتريد أن تسحب عينة جديدة وبحجم 30 طالباً، وتوزعها بين

المجموعات الثلاث باستخدام الطريقة المثلى لتوزيع حجم العينة بين الطبقات. لنفترض أن تكلفة جمع البيانات من المجموعات (الطبقات) الثلاث متساوية، استخدم المعلومات في السؤال العاشر لتقدير التباينات.

12. استخدم البيانات في السؤال العاشر لتحديد حجم العينة لتقدير الوسط الحسابي للمجتمع ليكون بين ± 4 ، استخدم التوزيع المتناسب.
13. كرر العملية للسؤال الثاني عشر باستخدام التوزيع الأمثل.
14. ترغب إدارة المدرسة في تقدير نسبة الطلاب الذين حصلوا على علامات أقل من 60 أستخدم البيانات المعطاة في السؤال العاشر لتقدير نسبة الطلاب الحاصلين على علامة أقل من 60، ثم أوجد فترة 95% ثقة للنسبة الحقيقية للطلاب الحاصلين على أقل من 60.
15. يحتوي أحد المجتمعات على 8000 عنصر جرى تقسيمه إلى طبقتين تحتويان على التوالي 5000 و3000 عنصر. جرى سحب عينة عشوائية بسيطة من كلتا الطبقتين بواقع 200 و100 عنصر. الجدول الآتي يلخص النتائج التي حصلنا عليها.

الطبقة	N_i	n_i	$\bar{y}_i$	s_i^2	$\hat{p}_i$
1	5000	200	400	50	0.25
2	3000	100	450	35	0.36

أ. احسب $\bar{y}_{st}$ و $\hat{p}_{st}$ و $s(\bar{y}_{st})$ و $s(\hat{p}_{st})$.

ب. أوجد 95% فترة ثقة للوسط الحسابي $\bar{Y}$ والنسبةP.

16. لنفترض أن البيانات في السؤال الخامس عشر تمثل عينة عشوائية أولية، ونرغب أن نحدد حجم العينة وتوزيعها بين الطبقتين باستخدام التوزيع المتناسب بحيث يكون الوسط الحسابي والنسبة للمجتمع بحدود ∓ 50 و ± 0.7 على التوالي.

obeikandl.com

17. كرر السؤال السادس عشر ولكن باستخدام التوزيع الأمثل لتوزيع لحجم العينة بين الطبقات.
18. لنفترض أن تكاليف مشاهدة العنصر الواحد للطبقتين هي $C_1 = 9$ و $C_2 = 6$ استخدم التوزيع الأمثل والمعلومات المعطاة في السؤال الخامس عشر لتحديد حجم العينة وتوزيعها بين الطبقتين.
19. يبلغ عدد سكان إحدى المدن 500.000 نسمه حوالي 60% من البيض و25% من السود والباقي يمثلون 15% . لنفترض أن التباين في الطبقات الثلاث على النحو الآتي: $S_1^2 = 5000$ و $S_2^2 = 500$ و $S_3^2 = 200$. نرغب في سحب عينة بحجم n=200. استخدم التوزيع الأمثل لتوزيع العينة بين الطبقات الثلاث.
20. قدر نسبة الكراسي المكسورة في قاعات إحدى كليات جامعتك، كذلك قدر معدل سعة القاعة في الكلية.
21. صمم استبانة لاستطلاع رأي طلبة جامعتك في قضية معينة على اعتبار أن الكليات في الجامعة تمثل طبقات مختلفة، ثم نفذ المشروع للحصول على البيانات المطلوبة.
22. اقترح مشروعاً له علاقة بدراستك أو مجال عملك تستخدم فيه العينة العشوائية الطبقية لتقدير الوسط الحسابي للمجتمع والمجموع الكلي وقم بتنفيذه.

obeikandl.com

المراجع العربية

1. سليمان محمد طشطوش. (2001) أساسيات المعاينة الإحصائية، دار الشروق للنشر والتوزيع ، عمان.
2. حسين علوان (1994). طرق المعاينة، دار الفرقان، عمان.
3. هلال عبود البياتي وصبري رديف العاني (1981). العينات، جامعة بغداد، بغداد.
4. وليد عبدالحميد نوري وعبدالمجيد حمزة الناصر (1981). العينات، وزارة التعليم العالي والبحث العلمي، بغداد.
5. وليم كوكوران – ترجمة أنيس كنجو (1995). تقنية المعاينة الإحصائية، جامعة الملك سعود، الرياض.

References

1. Aoyama, H. (1954). A Study of the Stratified Random Sampling, *Ann. Inst. Statist. Math.* 6, 1-36.
2. Armitage, P. (1947). A Comparison of Stratified with Unrestricted Random Sampling from a finite Population. *Biometrika*, 34, 273-280/
3. Barnett, V. (1991). Sampling Survey Principles and Methods, Edward Arnold, London.
4. Bryant, E. C., Hartley, H. O. and Jessen, R. J. (1960). Design and Estimation in Two-way Stratification. *J. Amer. Statist. Assoc.* 55, 175-182.
5. Benedetto, J. J. and Ferreira, P. (2001). Modern Sampling Theory, Birkhauser Boston.
6. Chatterjee, S. (1967). A Note on Optimum Stratification. *Skand. Akt.* 50, 40-44.
7. Chatterjee, S. (1968). Multivariate Stratified Survey. *J. Amer. Stat. Asso.* 63, 530-534.
8. Chaudhuri, A. and Stenger, H. (1992). Survey Sampling: Theory and Methods, Marcel Dekker, New York.
9. Cochran, W. G. (1961). Comparison of Methods for Determined Stratum Boundaries. *Bull. Inst. Stat. Inst.*, 38, 345-358.
10. Cochran, W. G. (1977). Sampling Technique, 3rd Ed. Wiley, New York.
11. Cornell, F. G. (1947). A Stratified Random Sample of a Small Finite Population. *J. Amer. Stat. Asso.* 42, 523-532..
12. Dalenius, T. and Gurney, M. (1951). The Problem of Optimum Stratification II. *Skand. Aktuartidsk* 7: 34, 133-148.
13. Dalenius, T. and Hodges, J. L. (1957). The Choice of Stratification Points. *Skand. Aktuar-tidskr.* 3-4, 198-203.
14. Dalenius, T. and Hodges, J. L. (1959). Minimum Variance Stratification. *J. Amer. Statist. Assoc.* 54, 88-101, correction JASA, 58 (1963), p. 1161.
15. Deming, W. E. (1960). Sampling Design in Business Research. Wiley, New York
16. Ekman, G. (1959). An Approximation useful in Univariate Stratification. *Ann. Math.Statist.* 30,219-229.

obeikandl.com

17. Evans, W. D. (1951). On Stratification and Optimum Allocations. *J. Amer. Statist. Assoc.* 46, 95-104.
18. Foreman, E. K. (1991). Survey Sampling Principles, Marcel Dekker, New York.
19. Fuller, W. A. (1966). Estimation Employing Post Strata. *J. Amer. Statist. Assoc.* 61, 1172-1183.
20. Fuller, W. A. (1970). Sampling with Random Stratum Boundaries. *J. Roy. Stat. Soc.* B32, 209-206.
21. Govindarajulu, Z. (1999). Elements of Sampling Theory and Methods, Prentice Hall.
22. Goodman, R. and Kish, L. (1950). Controlled Selection-a Technique in Probability Sampling. *J. Amer. Statist. Assoc.*, 45, 350-372.
23. Hajek, J (1981). Sampling from Finite Population, Marcel Dekker, New York.
24. Hansen, M. H., and Hurwitz, W. N. (1943). On the Theory of Sampling from Finite Populations. *Ann. Math. Statist.* 14, 333-362.
25. Hedayat, A. S. and Sinha, B. K. (1991). Design and Inference in Finite Population Sampling, Wiley, New York.
26. Hess, I., Sethi, V. K., and Balakrishnan, T. R. (1966). Stratification: A Practical Investigation. *J. Amer. Statist. Assoc.* 61, 74-90.
27. Holt, D. and Smith, T. M. F. (1979). Post-stratification, *J. Roy. Statist. Soc.* A142, 33-46.
28. Kish, L (1995). Survey Sampling, Wiley, New York.
29. Levy, P. S. and Lemeshow S. (1991). Sampling of Population: Methods and Applications, Wiley, New York.
30. Lohr, S. L. (1999). Sampling: Design and Analysis, Duxbury, New York.
31. Mahalanobis, P. C. (1946). Recent Experiments in Statistical Sampling in the Indian Statistical Institute. *J. Roy. Stat. Soc.* A109, 325-378, reprinted in Sankhyii
32. Murthy, M. N. (1962). Variance and Confidence Interval Estimation. Sankhyii 24(B), 1-12.
33. Neyman, J. (1934). On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection, *J. Roy. Stat. Soc.* 97, 558-606.
34. Raj, D. (1968). Sampling Theory, McGraw Hill, New York.
35. Rao, J. N. K. (1985). Conditional Inference in Survey Sampling. Survey Methodology, D. I. *Inst. Statist. Math.* No.1, 15-31.
36. Sampath, S. (2000). Sampling Theory and Methods, CRC Press.
37. Satterthwaite, F. E. (1946). An Approximate Distribution of Estimates of Variance Components, *Biometrics* 2, 110-114.
38. Scheaffer, R.L., Mendenhall, W. and Ott, L. (1996). Elementary Sampling Survey, 5th ed., Duxbury, New York.
39. Serfling, R. J. (1968). Approximately Optimal Stratification, *J. Amer. Statist. Assoc.* 63, 1298-1309.
40. Sethi, V. K. (1963). A Note on Optimum Stratification of Populations for Estimating the Population Means, *Aust. J. Statist.* 5, 20-33.
41. Sirken, M. G. (1972). Stratified Sample Surveys with Multiplicity, *J. Amer. Statist. Assoc.*67, 224-227.
42. Smith, T.M.F. (1976). The Foundation of Survey Sampling: A review, *Journal of Royal Statistics Society* A139, p 183-204.

obeikandl.com

43. Stephan, F. F. (1945). The Expected Value and Variance of the Reciprocal and other Negative Powers of a Positive Bernoulli Variate. *Ann. Math. Statist.* 16, 50-61.
44. Sturat A. (1984). The Ideas of Sampling, Revised Edition, Griffin, London.
45. Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S., and Ashok, C. (1984). Sampling Theory of Surveys with Applications, 3rd ed., Ames (Iowa): Iowa State University Press.
46. Taga, Y. (1967). On Optimum Stratification for the Objective Variable based on Concomitant Variables using Prior Information. *Ann. Inst. Statist. Math.* (Tokyo) 4, 95-102.
47. Thompson, S. K (2002). Sampling, 2end Wiley, New York.
48. Tryfos, P. (1996). Sampling Methods for Applied Research, Wiley, New York.
49. Yates, F. (1981). Sampling Methods for Censuses and Surveys 4th Ed., Grittin, London

obeikandl.com