الجمهورية العربية المتحدة

معهد التخطيط القومي

مذكرة رقم ٧٧٦

دراسات
في
الاحصاءات الزراعية

دكتور
محمد جلال الدين أبو الدهب

يونيو ١٩٦٧

القاهرة
٣ شارع محمد مظهر - بالزمالك

الآراء التي وردت في هـــذه المذكـــــرة
تمثل رأي الكاتب ولا تمثل رأي المعهد ذاته

الاحصـــاءات الزراعيـــــة

تعتبر الاحصاءات الزراعية جزء هام من الاحصاءات الاقتصاديه ، ولذا نجد ان اغلب الـــدول توليها اهتماما كبيرا مادامت الزراعه تمدنا بالغذاء والكساء هذا بجانب انها تمدنا بالمواد الخام التى تقوم عليها كثير من الصناعات . وتعتبر الزراعه حجر الاساس للتقدم الاقتصادى والتطور الصناعى للدول التى تخلفت لسبب او لآخر عن عجلة التقدم . لذا وجب على هذه الدول ان تعطى الاحصـاءات الزراعيه اهتماما كبيرا حتى يمكنها التعرف على مالديها من عناصر الانتاج الزراعى وبذلك يمكنهـــا استغلالها باحسن الطرق لتحقيق اكبر مايمكن من الناتج الزراعى وتكون هذه المعرفه بمثابة مرشـــد لأولى الامر حتى يتمكنوا من وضع السياسات الزراعيه والسياسات الاقتصاديه العامه على اسس سليمـــه حتى يمكن تحقيق الرخاء والرفاهيه لشعوبها .

تعريف الأحصاء الزراعى :

يمكن تعريف الاحصاء الزراعى بصوره عامه على انه جمع للبيانات المختلفه العديده عن مختلـــف النواحى الزراعيه والاقتصاديه . وهذه تشمل بطبيعة الحال بيانات عن عناصر الانتاج الزراعى مثـــل المساحات المتوفره ، والمساحات المزروعه ، والعمال الزراعيون ، الموارد المائيه ، التمويل الزراعى ، للمحاصيل ، الاسعار الزراعيه ، وغير ذلك من البيانات المتعلقه بالقطاع الزراعى . ويعرف الأنتـــاج الزراعى على انه استغلال الارض فى انتاج المحاصيل الزراعيه والفاكهه والحبوب وتربية شتلات الاشجار والخضراوات والزهور ـــ داخل وخارج الصوبات الزجاجيه ـــ كما يتضمن انتاج الماشيه ومستخرجاتهـــا والدواجن ومنتجاتها والعسل والارانب ودواب الحمل وغيرها .

ويمكن تقسيم الأحصاء الزراعى الى قسمين اساسيين :

١ ـــ الاحصاءات الاساسيـــة .

٢ ـــ الاحصاءات الدوريـــه .

فالاولى تشمل الاحصائيات الزراعيه العامه والتى لها صفة الأستقرار والدوام ويدخل ضمنها طرق الاستغلال الزراعى ، الملكيه ، الايجارات ، توزيع الحيازات وما الى ذلك . أما الثانيه فهى التى تتصل بالنشـــاط الزراعى الذى يتناوله التعديل والتغيير كل عام ويدخل تحت هذا القسم الاحصاءات الخاصه بالمساحـــه

المحصوليه ، والمساحات المخصصه للمحاصيل الزراعيه ، والكميات المستخدمه من الاسمده ، الآفات الزراعيه وما الى ذلك .

وتجمع الاحصاءات الاساسيه عن طريق التعداد الزراعى الذى يجمع كل خمس او عشر سنوات اما الاحصاءات الدوريه فتجمع سنويا او على فترات متقاربه . ومن الناحيه العمليه لايمكن وضع حد فاصل بين الاثنين فهناك بعض الاحصاءات والبيانات تجمع بكلتا الطريقتين على حسب الحاجة اليها ومدى اهميتها فمثلا عندما يحدث تغيير جوهرى فى ملكية الاراضى الزراعيه فان هذا يستدعى جمع احصاءات جديده عن الملكيه . وقد يتم هذا عن طريق اخذ عينه للملاك مثلا وعن طريقها يتم تعديل وتغيير بيانات التعداد السابقه ، او عن طريق اجراء تعداد شامل .

الحاجة الى الاحصاءات الزراعيه السليمه :

ترجع الحاجه الى الاحصاءات والبيانات الزراعيه الدقيقه الى انها تستعمل فى كثير من المجالات سواء كان استعمالا حكوميا ، او للقطاع العام ، او للقطاع الخاص من تجار ورجال الصناعه والمزارعين انفسهم ، ولتنظيم عمليات الانتاج والاستهلاك وفى استعمالات اخرى كثيره . ويجب ان تتوفر فى البيانات الزراعيه المعايير التاليه :

١ - الدقه : يجب ان تكون الاحصاءات الزراعيه دقيقه سليمه ويتأتى ذلك بتوضيح ما تهدف اليه هذه الاحصاءات وذلك عن طريق تعريف وتحديد معنى البنود الاحصائيه المطلوب جمعها فمثلا عند جمع بيانات عن المساحه المنزرعه يجب وضع تعريف دقيق محدد لها فقد تكون المساحه المنزرعه بالفعل او المساحه التى حصدت او المساحه القابله للزراعه كما يجب تحديد ما اذا كانت المساحه المطلوبه هى المساحه الكليه بما فى ذلك المصارف والمنافع والطرق او هى المساحه الصافيه بعد استبعاد هذه المنافع .

٢ - الملاءمه : تعتمد بعض الطرق المستخدمه فى جمع البيانات والاحصاءات الزراعية على التخمين او على الخبره الشخصيه لجامع هذه البيانات فى حين انه فى بعض الاحيان تستخدم الطرق الاحصائيه العلميه فى الحصول على بيانات اخرى . لذلك كان من الواجب عند نشر الاحصاءات الزراعيه التنويه عن الطرق والاساليب التى استخدمت فى الحصول على هذه البيانات حتى يمكن اعطاء فكره للافراد والهيئات التى تقوم باستخدامها عن مدى ملاءمتها وصلاحيتها .

٣ - عامل الوقت: يلعب الوقت دورا هاما فى صلاحيه البيانات الاحصائيه الزراعيه للأستخدامات المختلفه. فمثلا تأخير نشر الاحصاءات قد يخرجها من النطاق الزمنى الذى جمعت فيه ويفقدها فائدتها العلميه وتصبح ذات فائده تاريخيه.

٤ - الشمول: يجب مراعاة أن تكون الاحصاءات الزراعيه المنشوره شامله لكل البيانات المتعلقة بالموضوع تحت البحث. فاحيانا قد تخلو بعض الاحصاءات الخاصه بالحيازه من الحيازات التى تقل عن حجم معين. ويتحدد معنى الشمول عن طريق تحديد التعريف الدقيق للبيانات المنشوره.

٥ - المقارنة: من اهم اهداف الاحصاءات بصوره عامه هو امكان اجراء عمليات المقارنة فى الزمان والمكان. وتكون المقارنات مرضيه وصحيحه طالما كانت التعاريف المستخدمه محدده ودقيقة وكذلك طرق جمع هذه البيانات. فمثلا قد تنتج بعض الصعوبات نتيجه استخدام اكثر من طريقه فى تقدير المحاصيل المختلفه. ففى مثل هذه الحاله قد تجرى بعض التعديلات على البيانات حتى يمكن استخدامها فى المقارنات المطلوبه.

التعداد الزراعى

(١) اجرى التعداد الاول الزراعى فى مصر فى سنه ١٩٢٩ وذلك بناء على دعوه الجمعية العموميه للمعهد الدولى للزراعه بروما الذى وضع الاسس والاجراءات التى يمكن اتخاذها فى كل دوله حتى يمكن الاستفاده منها فى المقارنات الدوليه.

(٢) وفى سنة ١٩٣٩ اجرى التعداد الزراعى الثانى وتم ذلك مع الاسترشاد بالاسس والتوصيات التى وضعها المعهد الدولى للزراعه بروما.

(٣) واجرى التعداد الزراعى الثالث فى سنه ١٩٥٠ بناء على توصيات منظمه الاغذيه والزراعه التابعه لهيئه الامم المتحده والتى حلت محل المعهد الدولى للزراعه بروما.

(٤) وفى سنة ١٩٦٠ اجرى التعداد الزراعى الرابع. وقد وضعت منظمه الاغذيه والزراعه برنامج التعداد الزراعى العالمى لسنه ١٩٦٠. وقد اعتمد فى وضع هذا البرنامج على الخبره المكتسبه من السنين الماضيه. ويختلف تعداد سنة ١٩٦٠ عن تعداد سنة ١٩٥٠ فى الثلاث نقط الاساسية الآتية

١ — الزياده العامه فى اغراض البرنامج ٠ وعلى سبيل المثال فالجزء الخاص بملكية الاراضى اعطى اكثر اهميه وخاصه من ناحية ادارة الاراضى ٠

٢ — تعديل البرنامج بحيث يلائم الاوضاع المختلفه والمتباينه لمختلف المناطق والبلاد ٠

٣ — وضع البرنامج امكانيه استخدام طريقه العينات فى جمع البيانات التفصيليه عن بعض اقسام التعداد وحتى يمكن عمل دراسات تحليليه عن مدى دقه البيانات المجموعه بالتعداد نفسه ٠

الخطوط الرئيسيه للتعداد :

يشمل التعداد العشره اقسام التاليه والتى تتعلق بالمواضيع الرئيسيه والتى تسهل المقارنات بالتعدادات العالميه الاخرى :—

١ — الحائز — الحيازه والملكيه
٢ — استخدام الارض
٣ — المحاصيل
٤ — الحيوانات والدواجن
٥ — العمل والزراعة
٦ — المجتمع الزراعى
٧ — القوه الزراعيه
٨ — الرى والصرف
٩ — الاسمده ومخصبات التربه
١٠ — منتجات الغابات والاسماك ٠

مجال التعداد الزراعى :

تشمل بيانات التعداد الزراعى بيانات عن النقط الرئيسيه الآتيه :

١ — عدد الحيازات الزراعيه وصفاتها الرئيسيه كالحجم — نوع الملكية — طرق استخدام المساحات الزراعية — الآلات الزراعية ٠٠٠٠٠٠ وغيرها ٠

٢ — بيانات متعلقه بصفات المجتمع الزراعى ٠

٣ — عدد الحيوانات الزراعيه والدواجن ٠

٤ — عدد العمال الزراعيين وصفات المشتغلين بالزراعه ٠

٥ — حجم الانتاج من المنتجات الزراعيه والحيوانيه الرئيسيه ٠

٦ — منتجات الغابات والاسماك التى يتحصل عليها من الحيازات الزراعية ويجب ان يشمل التعداد جميع الحيازات الزراعيه مهما كان حجمها واينما كانت سواء كانت فى الحضر او الريف ٠

٧ — بيانات عن الرى والصرف واستخدام الاسمده ومخصبات التربه ٠

علاقه التعداد الزراعى بتعداد السكان :

من المرغوب فيه الحصول على بيانات عن الافراد الذين يرتبطون بالزراعه وهؤلاء سواء كانوا يعيشون على الحيازات او الذين لهم علاقه مباشره بالانتاج الزراعى ولكنهم لايعيشون على الحيازات. وعادة يستخدم تعداد السكان والتعداد الزراعى للحصول على بيانات عن هذا الجزء من السكان .

الحيازه : يقصد بالحيازه فى التعداد الزراعى الارض التى تستعمل كليا او جزئيا للانتاج الزراعى وتستغل بشخص واحد "الحائز" وحده او بمساعده آخرين بدون النظر الى الحجم او الموقع وتعتبر الحيوانات المحتفظ بها للاغراض الزراعية — مع عدم وجود ارض زراعيه — من مكونات الحيازة .

هذا وقد تتكون الحيازة من قطعه واحده او اكثر بشرط ان تقع فى قريه واحده واذا كانت القرى صغيرة ومتجاوره ووقعت الحيازه فى اكثر من قريه يمكن اعتبارها حيازه واحده .

ويعرف الحائز بأنه الشخص المدنى او القانونى الذى تقع عليه مسئولية جميع الأعمال فى المزرعة والذى يقوم بالعمل بنفسه او بواسطه اجراء آخرين واذا وجد اكثر من مسئول واحد يعتبروا جميعا حائزا واحدا . والحيازه الجماعيه هى التى تداريه مجموعة من الاشخاص ليس من الضرورى وجود قرابه بينهم ولكنهم يقومون بمقتضى تصريح حكومى بالعمل المشترك ومن امثلة ذلك المزارع الجماعيه فى الاتحاد السوفيتى .

الاحصاءات الزراعيه الدوريه

وتشمل هذه الاحصاءات كل البيانات المتعلقه بالنشاط الزراعى والتى قد يعتريها التغيير على مر الزمان ، اى ان التغيير قد يحدث على فترات اما قصيره او طويله . اى انها قد تجمع كل يوم او اسبوع اوشهر او عام . وتشمل هذه الاحصاءات بيانات عن البنود الاتيه على سبيل المثال لا الحصر .

أولا : المحاصيل الزراعيه والحيوانات الزراعيه والدواجن :

١ — المحاصيل وهذه تشمل : المساحات المزروعه بالمحاصيل المختلفه ، المساحات التى تم حصادها ، متوسط محصول الفدان ، الطرق الزراعيه المستخدمه ... الخ .

٢ — الحيوانات والدواجن وهذه تشمل : الدخل من الانتاج الحيوانى ، الامراض الحيوانيه ومدى نشاط العلاج البيطرى ، انتاج اللبن والبيض السنوى ، عدد الحيوانات .

ثانيا : الاحصاءات الاقتصادية والاجتماعية المرتبطة بالزراعة وهى تشمل بيانات عن السكان الزراعيون والعمال الزراعيون ، الاجور الزراعيه مقسمه الى اجور بالنقد أو اجور عينيه - بيانات عن الالات الزراعيه وقيمتها ، الاسمدة المستخدمه ، مقاومة الافات والامراض النباتيه والحشائش ، كميات مواد الوقود والزيوت وقيمتها ، أثمان المنتجات الزراعيه ، استعمالات المنتجات الزراعيه الصادرات والواردات ممتلكات المزرعة ، وسائل المواصلات للمزارع ، التسليف الزراعى ، ديون المزرعة استصلاح الاراضى ، وما الى ذلك من البيانات.

ويمكن الحصول على البيانات الدوريه السابقة وذلك عن طريق مكاتب الاحصاء الزراعى المنتشرة فى جميع محافظات الجمهورية. وهذه تتم أما عن طريق الحصر الشامل أو عن طريق استخدام طريقة العينات فمثلا فى تقدير المساحات المنزرعه بالمحاصيل المختلفه أو مساحات حدائق الفاكهة تتم عن طريق الحصر الشامل. فتقوم مثلا مصلحة المساحه بقياس المساحه المنزرعه قطنا كل عام وفى السنوات الاخيره ادخلت الطريقه الاحصائيه المعروفه باسم قياس النسبه Ratio Estimate لتقدير المساحة المنزرعه قطنا أو المنزرعه بالمحاصيل الاخرى وتتلخص فى القياس الفعلى لنسبه معينه من المساحه الكليه ويمكن حساب تقدير للمساحة المنزرعه الكليه باستخدام تلك النسبه. وفى حالات اخرى تستخدم طريقه العينات وذلك لتوفير المال والجهد ولتحقيق الدقة المطلوبه فى البيانات المجموعه وتستخدم طريقة العينات فى الجمهوريه العربيه فى مجالات كثيره من القطاع الزراعى. مثل تقدير محصول القطن ، تقدير محصول الارز والبصل والانتاج الحيوانى. وفى السنوات الاخيره تستخدم طريقه العد الكامل (التعداد) مع طريقة العينات للحصول على بيانات أكثر دقه عن القطاع الزراعى. ففى الحصر الشامل للقطاع الزراعى سنة ١٩٦٠ استخدم كلا من طريقتى العد اكامل (التعداد الزراعى) وطريقة العينات وذلك فى الحالات التى وجد أنه من الصعب جمعها عن طريق العد اكامل. والهدف من ذلك تخفيض التكاليف الاجمالية أو عندما يتعذر الحصول على اجابات دقيقه أو لتصحيح أو تعديل البيانات المتحصل عليها من التعداد الكامل. وسواء استخدمت طريقه العد الكامل أو طريقه العينات وجب تجهيز الاستمارات الخاصة بجمع البيانات المطلوبه. وتنقسم هذه الاستمارات الى نوعين :

١- كشف البحث :

وهذه يقوم الباحث بملئها بنفسه أى أنه يقوم بالاتصال المباشر الشخصى بمصدر البيانات سواء كان ذلك المزارع نفسه ، هيئه ، جمعيه ، مؤسسه ، مصلحه حكوميه ، ويتميز هذا النوع بأنه

يمكن للباحث ان يوضح الأسئله الوارده فى الاستماره وازاله اى سوء فهم قد يحدث لدى الشخص الذى لديه البيانات . كذلك يجب على الباحث ان لايؤثر على الاجابه اما بتوجيهها او التعديل فيها .

٢ - صحيفة الاستبيان :

ويختلف هذا النوع عن سابقه فى ان الشخص اوالهيئه التى لديها البيانات تقوم بنفسها بملء الاستماره . وفى العاده تسلم الى مصدر البيانات شخصيا او عن طريق البريد وفى الحالة الاخيره تقل نسبة الاجابات الوارده .

القواعد التى يجب مراعاتها فى تصميم الأستمارات :

١ - سهولة الاسئله مع وضوحها . ويجب ان تستبعد الاسئله الغامضه والتى تتطلب المستويات العاليه من الذكاء للاجابه عليها .

٢ - عدم كثرة الأسئله حتى لايتسرب الملل الى ملء الاستماره سواء كان الباحث نفسه اوالمبحوث . وان تكون الاسئله دقيقه محدده .

٣ - اختيار الاسئله التى تكون الاجابه عليها مختصره ومحدده ويجب الابتعاد عن الاسئلة التى تتطلب الاجابه عليها التفكير العميق او اجراء عمليات حسابيه معقده او التى تعتمد على قوة الذاكره .

٤ - مراعاة عدم تكرار الاسئلة .

٥ - يجب الابتعاد عن الاسئلة المثيرة والتى قد تثير غضب المبحوث والتى قد يكون لها اثر عكسى على نجاح البحث .

٦ - يجب ان توضع الاسئله بالطريقه التى تمكن الباحث من معرفة اذا كان المبحوث يدلى بالفعل بالبيانات الصحيحة .

ويمكن تلخيص اوجه الخلاف بين التعداد الزراعى وطريقه العينه فى الآتى :-

التعداد الزراعي	طريقة العينه
١ — يشمل القطاع الزراعى باجمعه مثل المساحه المنزرعه الكليه ، الحيوانات الزراعية وهكذا .	١ — يقتصر على جزء صغير نسبيا من المجموع الكلى للافراد اوالوحدات المطلوب جمع بيانات عنها .
٢ — يحتاج الى افراد كثيرين من احصائيين وفنيين وعداديين وخلافه .	٢ — يحتاج الى عدد قليل من الافراد ولكن على درجة كبيره من الكفاءة والتدريب .
٣ — يحتاج التعداد الزراعى الى امكانيات كبيره من اموال وفنيين ، وخرائط ووعى احصائى عام وتدريب وما الى ذلك .	٣ — تستخدم طريقه العينات عندما لاتتوفر الاموال اللازمه اوالامكانيات الكبيره وكذلك عند اجراء البحوث الخاصة باجزاء صغيرة من القطاع الزراعى .
٤ — يحتاج لوقت كبير لاجرائه واتمامه كما ان النتائج تتأخر فى العاده .	٤ — لاتحتاج لوقت طويل لاتمام جميع البيانات اللازمه . كما ان النتائج تظهر مباشره .
٥ — احتمال وجود اخطاء كثيره وهذا راجع اما الى حذف بعض الوحدات اوالمزارع عند العد اواخطاء ناتجه عن عدم تفهم الاسئله بالاستمارات اولكبر حجم التبويب .	٥ — يجب اولا توفر الاطار الكامل لكل المفردات اوالوحدات المكونه للمجتمع حتى يمكن تجنب اخطاء التكرار . وتعتبر نتائج طريقه العينات اكثر دقه من نتائج التعداد الكامل .

طريقـــة المعاينـــة
Sampling Method

يلجأ الباحثون الى استخدام طريقه المعاينه عندما تواجههم مشكله الحصول على معلومـــات تتعلق بمجموعات كبيره من الافراد ، المزارع ، الوحدات الانتاجيه ، او أى مجتمع ذو اعداد كبـــيره من الوحدات ، وبذلك يمكنهم الحصول على نتائج تتصف بالدقه العاليه مع خفض تكاليف البحـــث . وقبل البدء فى استخدام طريقه المعاينه يجب ان تتوافر لدى الباحث فكره عن العلاقه بين تكاليـــف البحث والدقه التى يجب ان تكون عليها المعلومات المطلوب الحصول عليها .

وعمليه المعاينه عبارة عن دراسة جزء صغير من مجتمع كبير بحيث يمثل هذا الجزء المجتمـــع كله . وتعتمد دقة النتائج المتحصل عليها من طريقه المعاينه على الطريقه التى اختيرت بها العينـــه وعلى الطريقه التى حسبت بها التقديرات المطلوبه من المعلومات التى جمعت بالعينه . ولأختيار التصميم المناسب للعينه وجب على الباحث استخدام كل المصادر المتاحه له . وتشمل هذه المصادر الفنيـــين والآلات وكل التسهيلات الماديه المتوفره هذا بجانب المعلومات الاحصائية والمعلومات الأخرى المتوفرة عن المجتمع الذى ستأخذ منه العينه بالاضافة الى المعرفه المتوفره عن طرق ونظرية المعاينة .

معيار اختيار التصميم المناسب للعينه :

يوجد فى العاده عدد كبير من الطرق المختلفه لاجراء المعاينه والتى يمكن استخدامها لمختلـــف المشاكل التى تواجه الباحثون . وهناك معايير مختلفة تستخدم للتمييز بين التصميم الجيد والـــردىء للعينه .

حدود بعض المعايير المستخدمه :

اذا اراد باحث الحصول على تقدير متوسط محصول الفدان من القطن مثلا فى محافظة مـــــن المحافظات فهناك عدة طرق للحصول على هذا التقدير . واحدى هذه الطرق عن طريق سؤال احـــد مهندسى الزراعه المتخصصين فى تلك المحافظه والذى له اتصال كبير مع الفلاحين والمزارعين والرجـــال المسئولين فى مختلف المناطق ليعطى لك رأيه وتقديره وحكمه الشخصى عن متوسط محصول الفدان المتوقع . ويعتمد تقديره وحكمه الشخصى على معلوماته وخبرته الشخصيه عما حدث فى الماضى وعن توقعاته لما قـــد يحدث فى المستقبل . وفى كثير من الاحيان تعطى هذه الطريقه نتائج فى غاية من الدقـــــة

ولاتحتاج هذه الطريقه الى تكاليف كبيره . وفى كثير من الاحيان لاتتوافرطرق القياس لذا يكون اعتماد الباحث الكلى على هذه الطريقه . وعند تقييم التقديرات التى تعتمد على الآراء الشخصيه للخبراء فان المعيار المستخدم ينحصر فى مدى الثقة التى تعطى لهذه الآراء . وتنحصر حدود هذه الطريقه فى عدم القدره على تقييم هذه الثقه وفى عدم توفر الاسس الجوهريه للاختبار بين آراء عديد من الخبراء .

قد يلجأ باحث آخر الى ارسال صحائف الاستبيان الى الافراد المختارين فى العينه للحصول على المعلومات التى يرغبها . لنفرض انه يرغب فى تقدير الدخل الزراعى لسكان مدينة اسيوط وأنه استخدم دفتر التليفون كاطار ليختار منه العينه التى يرسل اليها صحائف الاستبيان . وفى العاده فان عدد الاجابات الوارده تكون قليله بالنسبه لحجم العينه المرسله له صحائف الاستبيان . وبالرغم من ذلك فان الباحث يبنى تقديراته لمتوسط الدخل الزراعى على المعلومات التى تحصل عليها . واول قصور لهذه الطريقه هو استعمال دفتر التليفون كاطار لسكان مدينه اسيوط . فهذا اطار غير كامل لكل سكان مدينة اسيوط . لذا تكون العينة المختاره لاتمثل سكان اسيوط تمثيلا صحيحا . اما القصور الثانى فهو قلة عدد الاجابات الوارده ، وذلك لانه لايمكن اعتبار ان صفات الافراد التى لم تصل اجاباتهم هى نفسها صفات الافراد التى وصلت اجاباتهم .

ولتجنب القصور السابق يلجأ بعض الباحثين الى طريقة اخرى تعتمد على الاستجواب الشخصى للباحث او من ينوب عنه للافراد المختارين فى العينه وتعرف بطريقة البحث الميدانى . وبعد اختيار المناطق التى ستجمع منها البيانات يترك للباحث حرية اختيار الافراد الذين سيقوم بزيارتهم بعد تحديد الصفات والمدى الذى يختار فى حدوده (كأن يطلب منه اختيار عشرة افراد اعمارهم مابين ١٨ — ٢٥ سنه) .

قد يكون تصميم العينه فى غاية من الدقه ولكن استخدام الحكم الشخصى فى الاختيار النهائى بالرغم من ان المدى المسموح منه الاختيار قد حدد سابقا سيجعل من الصعب تقييم النتائج المتحصل عليها من العينه بصورة علمية .

وطريقه اخرى — كثيرا ماتستخدم ببعض الباحثين — والتى تتضمن الكثير من التحيز هى اختيار الافراد المثاليين ، المدن المثاليه ، المزارع المثاليه او المتوسطه لتمثل المجتمع . ومصدر التحيز هنا هو انه حتى ولو ان هذه المدن او المزارع او الافراد كانوا مثاليين او يعتبروا متوسطين فى الماضى

فانه لايوجد مايبرر او يؤكد انهم سيبقون على مثاليتهم هذه . وطريقتى البحث السابقتين لهم ايضا نوعين من القصور . الاول هو صعوبة تحديد العناصر المختلفه التى يمكن ان يعزى اليها اخطاء المعاينه المختلفه . اى انه ينقص الباحث فى كلتا الطريقتين المبادىء التى يمكنه استخدامها لتحديد العناصر التى يتسبب عنها الخطأ أو التحيز فى النتائج حتى يمكنه ان يتجنبها عند وضع خطه المعاينه .

والقصور الثانى ينحصر فى عدم امكانية قياس دقه نتائج العينه . ففى الطريقتين السابقتين لايوجد اساس لقياس درجة الثقه التى يمكن اعطائها لتقديرات العينه .

وطرق المعاينه السابقه تشترك معا فى صفة عامه وهى ان احتمال ظهور اى وحده او فرد من المجتمع فى العينه يكون غير معروفا . واحتمال ظهور فرد ما لن يكون معلوما فى كل طرق المعاينة التى لاتستخدم قانون الاحتمالات فى الانتخاب النهائى لمفردات العينة .

المعاينة الاحتمالية
Probability Sampling

مما سبق يتبين أن استخدام الخبره الشخصيه او استخدام طرق المعاينه التى لاتعتمد فى اختيارها على قانون الاحتمالات ستكون نتائجها غير دقيقه وستتضمن كثير من الاخطاء والتحيز . لذا لكى تكون تقديراتنا من العينات ذات دقة معينه وحتى يمكن تحديد الثقة التى تعطى لهذه التقديرات وجب علينا استخدام قانون الاحتمالات فى اختيارنا للوحدات او الافراد الذين يكونون العينه . وتعرف هذه بطريقة المعاينه الاحتماليه .

المعاينه العشوائيه البسيطه Simple Random Sampling :

يقصد بالعينه العشوائيه البسيطه والتى تتكون من عدد (n) من الوحدات بانها عينه مختاره من المجتمع بطريقه تجعل كل مجموعه من الوحدات والتى عددها (n) لها نفس الفرصه اوالاحتمال لتكون هى المختاره كأى مجموعه اخرى . واى طريقه اخرى لاختيار الوحدات (n) من غير ان تعطى نفس الفرصه لكل المجموعات الممكنه لان تختار لايطلق عليها عينه عشوائيه . وطريقه المعاينه العشوائيه البسيطه لاتستعمل فى العاده بكثرة فى الحياه العمليه ولكن لاهميتها الكبرى لنظريه المعاينه ولبساطتها ، ولتشابهها للانواع الاخرى من طرق المعاينه فاننا سنوقم بمناقشتها بالتفصيل .

لنفرض لدينا مجتمع فرضى يتكون من ١٢ مزارعا واننا نرغب فى تقدر متوسط دخلهم وذلك باستخدام عينه . والجدول رقم (١) يبين الافراد ودخلهم السنوى .

(جدول رقم (١) — الدخل السنوى للمجتمع الفرضى)

الافراد	الدخل السنوى بالجنيه
A	٢٠٠
B	٣٤٠
C	٤٦٠
D	٣٢٠
E	٢٨٠
F	٢٤٠
G	٣٨٠
H	٢٢٠
I	١٨٠
J	٣١٢
K	٢٦٠
L	٥٤٠
الدخل الكلى	٣٧٣٢
متوسط الدخل	٣١١

اختيار العينه العشوائيه البسيطه :

هناك عده طرق لاختيار العينه العشوائيه البسيطة فمثلا يمكن كتابة اسماء الافراد على قطع متساويه من الورق بحيث يكون على كل ورقة حرف واحد ، اى ان كل فرد من هذا المجتمع سيكون ممثلا بقطعه واحده من الورق ، ويوضع قطع الورق بعد طيها فى صندوق وخلطها جيدا ثم سحب ورقتين من الاثنى عشر ورقه فتكون بذلك قد اخترنا عينة عشوائيه بسيطه مكونه من فردين ، لنفرض ان الورقتين ظهر عليهما حرفى K & D فيكون مجموع دخلهم مساويا ٥٨٠ جم ومتوسط دخلهم ٢٩٠ جم ويكون عدد السحبات الممكنه وكل منها يتكون من فردين هو ٦٦ سحبه .

عدد السحبات = $ن ق ر = \frac{ن !}{(ن - ر) ! (ر) !}$

فاذا تكررت عملية السحب لعدد كبير من المرات ، واذا حسبت عدد المرات التى تظهر فيها كل من الـ ٦٦ مجموعه ، فانه سيلاحظ ان النسبه التى تظهر فيها المجموعه AB ستكون نفس النسبة التى تظهر فيها المجموعه BC او اى من الـ ٦٦ مجموعه الاخرى . لذا يمكن القول بان كل العينات الممكنه لها نفس الفرصه لان تختار باستعمال طريقه المعاينه العشوائيه البسيطه .

وهناك طريقة ابسط من السابقه لاختيار العينه العشوائيه البسيطه وذلك عن طريق استخدام الجداول العشوائيه .

معرفة احتمال اختيار المفردات:

اذا كانت العينه المختاره تتكون من فرد واحد ، فيكون لكل من الاثنى عشر فردا فى مجتمعنا نفس الفرصه لان يقع عليه الاختيار ولذا يكون احتمال اختيار اى فرد مساويا $\frac{1}{12}$ ، اى ان احتمال ظهور فرد فى عينه عشوائيه بسيطه باى حجم ستكون مساويه ($\frac{n}{N}$) حيث أن (N) هو عدد المفردات فى المجتمع ، (n) عدد المفردات فى العينه .

التقدير من العينة : يعتبر متوسط دخل الفرد فى العينه تقديرا لمتوسط دخل الفرد فى المجتمع فعندما سحبت العينه المكونه من A & K كان متوسط دخل العينه مساويا ٢٩٠ جم وهو تقدير لمتوسط دخل الفرد فى المجتمع . ولتحديد مدى دقه هذا التقدير المبنى على عينه واحده ، فيتعين اختبار كل العينات الممكنه والمكونه من فردين . فاذا حسب متوسط دخل الـ ٦٦ عينه الممكنه ثم اخذ متوسط هذه المتوسطات فاننا نلاحظ انه يساوى متوسط دخل الفرد فى المجتمع اى يساوى ٣١١جم لذا ـ فانه فى المتوسط ـ فان تقدير العينه يساوى بالضبط متوسط المجتمع المقدر . وهذا سيكون صحيحا لمتوسط العينه المقدر من عينه عشوائيه بسيطه بصرف النظر عن حجم العينه المستعمل وعن نوع المجتمع المأخوذه منه العينه . ويطلق على متوسط التقديرات المتحصل عليها من كل العينات الممكنة اصطلاح القيمه المتوقعه لتقدير العينه Expected Value of Sample Estimate

أى ان القيمه المتوقعه لتقدير متوسط الدخل = $\frac{\text{مجموع تقديرات متوسط الدخل من كل العينات الممكنه}}{\text{عدد العينات الممكنه}}$

التحيز الناتج عن المعاينة : يعتبر التحيز مساويا للفرق بين القيمه المتوقعه للتقدير والقيمة الحقيقيه للمجتمع والمقدره بواسطة العينه . فاذا كان التحيز يساوى صفرا ، يقال عن التقدير انه غير متحيز . من هذا يتبين ان متوسط العينه يكون تقدير غير متحيز لمتوسط المجتمع عندما تستخدم طريقة المعاينة العشوائيه البسيطه بصرف النظر عن حجم العينه المستخدمه او نوع المجتمع المأخوذه منه العينه . وليس معنى هذا ان التقدير الغير متحيز يكون خاليا من الخطأ فتقدير المتوسط الذى حصلنا عليه من العينة A & K قد يختلف عن القيمه الحقيقيه لمتوسط دخل المجتمع (تقدير متوسط العينه = ٢٩٠ جم فى حين ان متوسط المجتمع = ٣١١ جم) بالرغم من ذلك فان هذا التقدير يعتبر تقديرا غير متحيز (حيث ان متوسط تقديرات كل العينات الممكنه سيكون مساويا لمتوسط المجتمع).

أى ان القيمه المتوقعه للتقدير = القيمه الحقيقيه للمجتمع .

لنفرض أن $\hat{\theta}$ هى تقدير لـ θ ، فان $\hat{\theta}$ تكون تقديرا غير متحيز لـ θ اذا كان توقع $(\hat{\theta}) = \theta$

$$\text{وتوقع}(\hat{\theta}) = \frac{\text{مجـ}_{1}^{\text{ن}} \hat{\theta}}{\text{ن}}$$

التقديرات المتسقة Consistant Estimates يعتبر تقدير العينه تقديرا متسقا اذا كانت نسبة التقديرات المتحصل عليها من العينه - والتى تختلف عن القيمه المقدره بمقدار اقل من قيمة صغيره محدده - تقترب من ١٠٠ ٪ عندما يزداد حجم العينه .

اى أن $\hat{\theta}$ يكون تقديرا متسقا لـ θ اذا كان احتمال $[\ |\hat{\theta} - \theta| < \text{مقدار صغير}\ ]$ تقترب من ١

عندما يقترب حجم العينه من حجم المجتمع .

وهذا يعنى انه اذا كان حجم العينه كبيرا فان الباحث لن يرتكب خطأ كبيرا اذا استعمل تقديرا مأخوذا من عينة عشوائيه .

الصحه والدقه Accuracy and Precision

يمكن قياس التحيز على انه الفرق المطلق بين صحة التقدير ودقة التقدير .

والمثال الآتى يوضح ما المقصود بصحه التقدير ودقته :

مثال : لنفرض ان طلبه كلية الزراعه عددهم ١٠٠٠ ، منهم ٢٠٠ طالبه ، ٨٠٠ طالب . وانه امكننا الحصول على كشف الحضور ليوم دراسى وكان عدد الحاضرين ٩٠٠ ، منهم ١٥٠ طالبه ، ٧٥٠ طالبا . واننا اعتبرنا ان كشف الحضور هذا هو الاطار الذى يمكن الاختيار منه .

الافراد المسموح لهم بالدخول فى العينه		
طالب		طالبه
٧٥٠	+	١٥٠ = ٩٠٠ الاطار
٩٠٠	+	١٠٠ = ١٠٠٠ = المجتمع

الافراد الغير مسموح لهم بالدخول فى العينة		
طالب		طالبه
٥٠	+	٥٠ = ١٠٠ نقص

لنفرض اننا نريد تقدير نسبة الطالبات الى الطلبه فى المجتمع ولنرمز لها (p) ولـــــذا فاننا سنسحب عينه عشوائيه من الاطار وليكن حجمها ١٠٠ ، وكان عدد الطالبات فى العينـــــه ١٦ طالبه .

فيكون تقدير النسبه من العينه $\hat{P} = \frac{١٦}{١٠٠} = ٠ر١٦$

ولكن قيمه النسبه الحقيقيه من الاطار $p_f = \frac{١٥٠}{٩٠٠} = ٠ر١٦٧$

وقيمة النسبه الحقيقيه للمجتمـــــع $p_u = \frac{٢٠٠}{١٠٠٠} = ٠ر٢٠$

فتكون الدقـــه Precission $= p_f - \hat{p} = ٠ر١٦ - ٠ر١٦٧ = -٠ر٠٠٧$

وتكون الصحـــه Accuracy $= p_u - \hat{p} = ٠ر١٦ - ٠ر٢٠٠ = -٠ر٠٤٠$

ويقاس التحيز Bias = | الدقه - الصحه | = $|p_u - p_f|$ = | -٠ر٠٣٣ | = ٠ر٠٣٣

ويجب التنبيه هنا الى انه توجد مصادر اخرى للتحيز غير تلك الخاصه بالاختلاف بين الاطار المسحوبـــه منه العينه والحجم الحقيقى للمجتمع . وقد يظهر التحيز كنتيجه للتحيز فى الاختيار للعينه او تحـــــيز فى التقدير . فالتحيز فى الاختيار قد ينتج عن اختيار العينه بطريقه متحيزه كما فى حاله استعمال دليـــل تليفونات مدينه اسيوط فى مثالنا السابق . كما ان التحيز فى الاختيار قد يكون ناتجا عن استبـــــدال وحدات غير مختاره بالعينه بدلا من الوحدات المختاره وذلك اما لصعوبة الوصول الى الوحدات المختاره ويمكن التخلص من تحيز الاختيار هذا وذلك عن طريق اختيار العينه بالطريقه التى تضمن عدم ظهور هذا التحيز وتنفيذ خطه البحث بدقة وامانة والنوع الاخر من التحيز والذى يعرف بتحيز التقدير وهـــــذا ناتج كما سبق ان وضحنا نتيجة ان بعض طرق المعاينه تكون تقديراتها متحيزه .

١) المتوسط الحسابى للمجتمع :

وهو عباره عن حاصل قسمة مجموع قسم مفردات المجتمع مقسومه على عدد المفردات فى المجتمـــــع

$$\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{X_1 + X_2 + X_3 \cdots + X_N}{N}$$

بفرض أن $\bar{X}$ = متوسط الحسابى للمجتمع ،

وان N = عدد المفردات فى المجتمع .

والمتوسط الحسابى للعينه عباره عن حاصل قسمة مجموع قيم المفردات الداخله فى العينه مقسومه على عدد المفردات فى العينه .

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

بفرض أن $\bar{x}$ = المتوسط الحسابى للعينه)

وان n = عدد المفردات فى العينه .

٢) وتكون القيمة الكلية للمجتمع عباره عن مجموع قيم المفردات .

$$X = \sum_{i=1}^{N} X_i$$

حيث ان X = القيمة الكليه للمجتمع .

وتكون القيمه الكليه للمجتمع مقدره من العينه عباره عن حاصل ضرب متوسط العينه فى عدد مفردات المجتمع

$$X' = N\bar{x} = \frac{N}{n} \sum_{i=1}^{n} x_i$$

حيث أن X' = القيمه الكلية للمجتمع مقدره من العينه .

٣) وتكون النسبة Proportion للمجتمع وهى حالة خاصه للمتوسط الحسابى للمجتمع أى فى الحاله التى تكون فيها (x_i = o or 1)

$$P = \frac{X}{N}$$

بفرض أن P = النسبة للمجتمع)

X = عدد المفردات ذات الصفه الخاصه فى المجتمع .

وتكون النسبه للعينه (x_i = o or 1)

حيث p = النسبه للعينه $\qquad p = \frac{x}{n}$

x = عدد المفردات ذات الصفة الخاصة فى العينه .

٤) ويكون تباين المجتمع عباره عن متوسط مربعات انحرافات القيم لجميع المفردات عن متوسطه اى أن

$$\text{التباين} = \sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N} = \frac{1}{N}\left(\sum_{i=1}^{N} X_i^2 - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N}\right)$$

ويطلق على الجزر التربيعى لهذه القيمه اصطلاح الانحراف المعيارى (σ) ٠

ويكون تباين العينه عباره عن متوسط مربعات انحرافات القيم للمفردات الداخله فى العينه عن متوسط العينه اى ان

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)$$

ويمكن القول بأن تباين التقدير المحسوب من عينه حجمها (n) هو متوسط مربع انحرافات التقديرات المقدره من كل العينات الممكنه ذات نفس الحجم عن قيمتها المتوقعه ٠ والجذر التربيعى لهذا المقدار هو مايعرف بالخطأ المعيارى وهو يقيس مدى تباين تقديرات العينات عن متوسطها (اى قيمتها المتوقعه) ٠ فاذا كان الخطأ المعيارى لتقديرات العينه صغيرا تكون اغلب التقديرات متركزه حول المتوسط العام لتقديرات كل العينات الممكنه واذا كان كبيرا فان هذا يعنى ان التقديرات المتحصل عليها من العينات الممكنه تكون منتشره حول المتوسط العام لها ٠

فاذا اردنا حساب تباين متوسط العينة ($\sigma^2_{\bar{x}}$) فيتعين علينا ايجاد توزيع متوسط العينه ($\bar{x}$) بمعنى ان نحسب متوسطات كل العينات الممكنة ذات الحجم (n) ثم حساب متوسط هذه المتوسطات الممكنه عن متوسطها العام ٠ ولكن يمكن حساب تباين متوسط العينه بدون ايجاد توزيع متوسط العينه ($\bar{x}$) وذلك باستخدام المعادلات الآتية :

١ — فى حاله المعاينه بدون ارجاع :

$$\sigma^2_{\bar{x}} = \frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}$$

اى ان الوحدات المسحوبه فى اول سحبه لاترجع الى الوحدات الباقية اى انها لن تكون لها فرصة ثانيه لان تختار فى السحبات التاليه ٠

٢ — فى حالة المعاينه مع الارجاع :

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

اى ان الوحدات المسحوبه فى اول سحبه تعود الى الوحدات الباقيه اى انه سيكون امامها فرصه ثانيه وثالثه وهكذا لكى تختار فى السحبات التاليه ٠

فاذا فرض ان لدينا مجتمع يتكون من ثلاث مفردات A & B & C واريد سحب كل العينات الممكنه والتى تتكون من مفردتين ففى الحاله الاولى (بدون ارجاع) سيكون عدد العينات هو ثلاثه (AB & AC & BC) . اما فى الحاله الثانيه (مع الارجاع) فيكون عدد العينات الممكنه هو تسعه (AA, BB, CC, AB & BA, AC & CA, BC, CB)

تمرين : اذا كان لدينا مجتمع مكون من خمس مفردات كما هو مبين بالجدول الاتى اسحب كل العينات الممكنه بدون ارجاع على ان يكون حجم العينه = ٢ ، ٣ ، ٤ ، ٥ المطلوب :

مفرده	X
A	صفر
B	٢
C	٣
D	٥
E	١٥

١) حساب متوسط المتوسطات لكل حاله .

٢) حساب التباين للمتوسط .

٣) حساب النسبه Proportion التى تظهر فيها كل مفرده فى العينات الممكنه ذات الحجم المعين .

هذا مع العلم ان تباين المجتمع σ^2 = ٢٧,٦ .

والجدول رقم (٢) يبين كل العينات الممكن سحبها من المجتمع بالاحجام المختلفه مع حساب القيمه المتوقعه لـ $\bar{x}$ ، p ، تباين المتوسط $\bar{x}$

جدول رقم (٢)

2 = n		3 = n		4 = n		5 = n	
العينه	$\bar{x}$	العينه	$\bar{x}$	العينة	$\bar{x}$	العينة	$\bar{x}$
AB	١ر٠	ABC	١ر٦٧	ABCD	٢ر٥٠	ABCDE	٥ر٠
AC	١ر٥	ABD	٢ر٣٣	ABCE	٥ر٠٠		
AD	٢ر٥	ABE	٥ر٦٧	BCDE	٦ر٢٥		
AE	٧ر٥	ACD	٢ر٦٧	CDEA	٥ر٧٥		
BC	٢ر٥	ACE	٦ر٠٠	DEAB	٥ر٥٠		
BD	٣ر٥	BCD	٣ر٣٣				
BE	٨ر٥	BCE	٦ر٦٦				
CD	٤ر١	BDE	٧ر٣٣				
CE	٩ر٠	CDE	٧ر٦٧				
DE	١٠ر٠	DEA	٦ر٦٧				
المجموع	٥٠ر٠	المجموع	٥٠ر٠	المجموع	٢٥ر٠	المجموع	٥ر٠
$E(\bar{x})$	٥ر٠	$E(\bar{x})$	٥ر٠	$E(\bar{x})$	٥ر٠	$E(\bar{x})$	٥ر٠
$p = \frac{n}{N}$	$\frac{٤}{١٠} = \frac{٢}{٥}$	$p = \frac{n}{N}$	$\frac{٦}{١٠} = \frac{٣}{٥}$	$P = \frac{n}{N}$	$\frac{٤}{٥}$	$p = \frac{n}{N}$	١
$\sigma^2_{\bar{x}}$	١٠ر٣٦	$\sigma^2_{\bar{x}}$	٤ر٦	$\sigma^2_{\bar{x}}$	١ر٧٢٥	$\sigma^2_{\bar{x}}$	٢٧ر٦

طرق تقليل خطأ المعاينة العشوائيه :

تميز طريقة المعاينه العشوائيه بانه يمكن تصغير خطأ العينه (اى زياده الدقه) وذلك بطريقة سهله بسيطه وذلك بزياده حجم العينه . وذلك لان خطأ المعاينه العشوائية يتناسب عكسيا مع الجزر التربيعى لعدد مفردات العينه . ويجب التنبيه هنا الى ان الدقه لاتعتمد فقط على حجم العينة ولكنها تتوقف بجانب ذلك على التباين لكل وحده . لذا فاننا نجد ان نظريه المعاينه قد امدتنا بعده طرق تمكننا من اختيار العينه عشوائيا مع وضع بعض القيود التى تؤدى الى التقليل من تباين الوحدات وبالتالى تؤدى الى تصغير حجم العينه . وابسط هذه القيود مايسمى بالتقسيم الى طبقات Stratification اذا ماقسم مجتمع ما الى اجزاء بحيث تكون وحدات كل جزء متجانسه بقدر الامكان . ومن ثم تجرى المعاينه لكل طبقه بطريقه عشوائيه ، فيقال فى هذه الحاله ان المجتمع مقسم الى طبقات Strata .

ومن الطرق المستخدمه ايضا لتقليل خطأ المعاينه طريقه المعاينه المتعدده المراحل Cluster Sampling وفيها يقسم المجتمع الى عدد من وحدات المعاينة للمرحله الاولى والتى يختار منها عددا من الوحدات والتى تحتوى على الوحدات المراد دراسه خواصها وتسمى بوحدات المرحله الثانيه ثم تجرى على هذه الوحدات عمليه المعاينه وقد تتعدد المراحل اذا احتاج الامر الى ذلك .

اخطاء غير المعاينه Non -Sampling Errors

يطلق على الفرق بين التقدير المحسوب من عينه والقيمه التى يحصل عليها عند اجراء تعداد شامل اصطلاح اخطاء المعاينه . وهناك اخطاء كثيره مختلفه تظهر فى كلا من النتائج المتحصل عليها سواء من التعداد الشامل او من المعاينه . ويمكن التحكم فى هذه الاخطاء عند التحقق من طبيعتها ومصدرها . ومن هذه الاخطاء مانتج عن خطأ فى التعاريف المستخدمه او خطأ فى تصميم صحائف الاستبيان او كشوف البحث وما الى ذلك من اخطاء التصميم . ونوع آخر من الخطأ قد يكون نتيجه سوء فهم المبحوث للاسئله او عدم وضوح الاسئله نفسها وقد ينتج عن عدم الادلاء بالاجابات الصحيحه لعدم توفر الوعى الاحصائى بين المبحوثين كما ان هناك اخطاء تظهر نتيجه عمليات التبويب والعمليات المكتبيه وعمليات تصنيف البيانات واخطاء الجمع والتحليل وما الى ذلك .

الطرق المستخدمه للتحكم فى خطأ المعاينه :

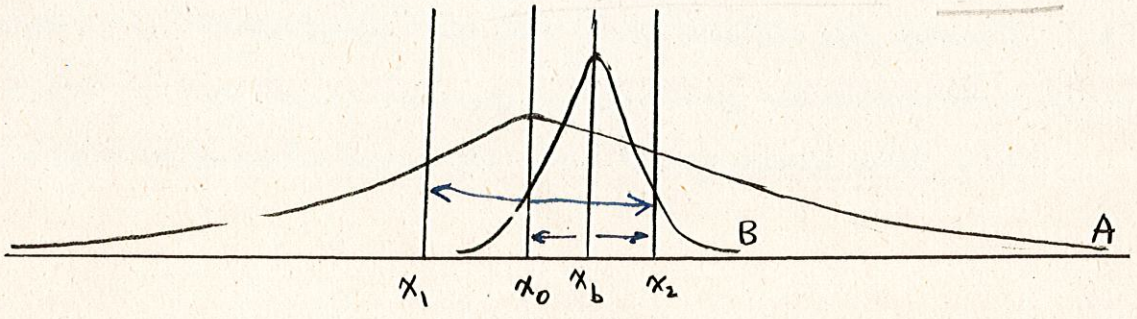
مما سبق يتبين انه يمكن التحكم فى حجم خطأ المعاينه عند استعمال الطريقه البسيطه للمعاينه العشوائيه وذلك بزياده حجم العينه . وكذلك تبين لنا ان خطأ المعاينه بطبيعه الحال يعتمد على تباين المجتمع والى درجه ماعلى حجم المجتمع الكلى ، ولكن الباحث لايمكنه التحكم فيهم عند استعماله طريقه المعاينه البسيطه . وهناك طرق اخرى يمكن بها التحكم فى خطأ المعاينة ، ويتوقف اختيارنا لاى من هذه الطرق على مدى الدقه المرغوبه فى نتائج العينه او على التكاليف التى يمكن انفاقها للحصول على هذه النتائج او على كليهما .

ومن المعلوم ان الباحث يمكنه من زياده كفاءة العينه وذلك عن طريق استخدامه كل المعلومات التى يمكن توفرها فى صوره خرائط ، قوائم ، معلومات التعداد او عن طريق استخدام الطرق التى تجمع وحدات المعاينه فى مجاميع تكون الوحدات بها اكثر تجانسا او باستخدام طرق المعاينه المرحليه ، او طرق اخرى للمعاينه . ويقال ان تصميم العينه ذو كفاءة اكثر من غيره اذا كانت نتائجه اكثر كفاءه لوحده التكاليف .

طرق التقدير المختلفة :

ان مناقشتنا السابقه كانت تدور حول التقديرات الغير متحيزه بمعنى ان القيمه المتوقعه للتقدير تكون مساويه للقيمه المقدره . وفى الواقع فاننا نهتم اكثر بالخطأ الذى قد يظهر عند حساب التقديرات من العينه عن الخطأ فى متوسط التقديرات المتحصل عليها من كل العينات الممكنه . فاذا كان الخطأ المعيارى كبيرا فكون التقدير غير متحيز يكون غير ذات موضوع .

والشكل رقم (١) يبين تلك الحاله ، وفيه المنحنى (A) يمثل توزيع تقديرات العينه . وفيه القيمه المتوقعه لتقديرات العينه تساوى القيمه الحقيقيه (x_0) ويلاحظ ان عددا قليلا من تقديرات العينة يكون قريبا من (x_0) اى فى المدى مابين (x_1) ، (x_2)



شكل رقم (١)

واذا نظرنا الى المنحنى (B) ويمثل توزيع تقديرات العينه وفيه متوسط تقديرات العينات ($\bar{x}_b$) لايساوى القيمه الحقيقيه للمجتمع ($\bar{x}_o$) فنلاحظ ان عدد كبيرا من تقديرات العينات يكون قريبا من ($\bar{x}_o$) بالرغم من ان طريقه التقدير متحيزه . وفى هذه الحاله يكون من الافضل الحصول على تقدير جيد من مجموعه من التقديرات المتحيزه والممثله بـ (B) عن تقدير من مجموعة التقديرات الممثله بـ (A) والتى تعتبر غير متحيزه .

مثال : لنفرض انه يمكننا الاختيار بين طريقتين لتقدير متوسط دخل المجتمع المكون من الاثنى عشره فردا فى جدول رقم (٣) وفيه يظهر مستوى الدخل الحالى ودخل تعداد ١٩٥٠ ، وان حجم العينه = ٣ افراد . وان متوسط دخلهم لسنه ١٩٥٠ = ٣٤٤ جنيها ومتوسط دخلهم الحالى = ٣١١ جنيها .

جدول رقم (٣)

الافراد	الدخل من تعداد ١٩٥٠	الدخل الحالى
A	٢٣٦	٢٠٠
B	٣٧٠	٣٤٠
C	٤٩٠	٤٦٠
D	٣٥٦	٣٦٠
E	٣١١	٢٨٠
F	٢٧٥	٢٤٠
G	٤١٢	٣٨٠
H	٢٥٤	٢٢٠
I	٢١٣	١٨٠
J	٣٤٥	٣١٢
K	٢٩٣	٢٦٠
L	٥٧٣	٥٤٠
الدخل الكلى	٤١٢٨	٣٧٣٢
متوسط الدخل	٣٤٤	٣١١

لنفرض اننا سنقوم بتقدير متوسط دخل الفرد وذلك بحساب متوسط دخل العينه فى الوقت الحاضر مقسومه بمتوسط دخل نفس الافراد لسنة ١٩٥٠ ويضرب هذه النسبه فى متوسط دخل المجتمع لسنة ١٩٥٠ والمعروف من التعداد . لنفرض ان العينه تتكون من الثلاث افراد (B) ودخله فى جدول (٢) = ٣٤٠ جنيه ، (C) ودخله فى جدول (٢) = ٤٦٠ جنيه ، (J) ودخله فى جدول (٢) = ٣١٢ جنيه ومتوسط دخلهم = ٣٧٠٫٦٧ جنيها ومتوسط دخل الثلاثه افراد تبعا لتعداد ١٩٥٠ هو ٤٠١٫٦٧ جنيها .

فيكون تقدير النسبه للدخل الحالى

$$= \frac{٣٧٠٫٦٧}{٤٠١٫٦٧} \times ٣٤٤ = ٣١٧٫٤٥ \text{ جنيها}$$

ويلاحظ ان هذا التقدير افضل من التقدير المتحصل عليه من العينه • ومن هذا يمكن القول بأن استعمالنا للعينه هنا كان لتقدير التغير النسبى فى متوسط الدخل وليس لتقدير مستوى الدخل نفسه • اى اننا قدرنا من العينه النسبه بين متوسط الدخل الحالى الى متوسط دخل ١٩٥٠ •

واذا حسبنا هذا التقدير لكل العينات الممكنه والتى حجمها يساوى ٣ افراد • فاننا سنجد أن متوسط هذه التقديرات لن يساوى ٣١١ جنيها اى القيمه الحقيقيه لمتوسط الدخل الحالى ولذا يكون هذا التقدير متحيزا •

ويمكن قياس تباين تقدير النسبة من القيمه الحقيقيه للمجتمع وذلك يأخذ متوسط مجموع مربع الفروق بين تقدير النسبه والقيمه الحقيقيه للمجتمع (بدلاً من متوسط المتوسطات) لكل من العينات الممكنه • ويعرف هذا المقياس بمتوسط مربع الخطأ Mean Square Error • والفرق بينه وبين التباين ينحصر فى ان التباين يقاس بأخذ متوسط مجموع مربعات الفروق من القيمة المتوقعه • فالتقديرات الغير متحيزه يكون تباينها مساويا لمتوسط مربع الخطأ ولكن للتقديرات المتحيزه فانهما يختلفان • وبزياده حجم العينه فان خطأ التحيز لتقدير النسبه سيكون صغيرا لدرجه انه يمكن اهماله من هذا يتضح • من الناحيه العمليه — ان تقدير النسبه كثيرا مايفضل على التقدير الغير متحيز البسيط • ولكن هذا يتوقف على ما اذا كان هناك ارتباط بين دخول سنه ١٩٥٠ والدخول الحاليه •

متوسط مربع الخطأ Mean Square Error :

يعتبر الخطأ المعيارى مقياسا غير كافيا لقياس صحه التقدير فى الحالات التى يكون التحيز فيها كبيرا بالنسبه للخطأ المعيارى • والسبب فى هذا راجع الى ان الخطأ المعيارى يقيس الفروق بين تقديرات العينه وقيمتهم المتوقعه لكل العينات الممكنه • فاذا كان التقدير متحيزا فتكون القيمه المتوقعه لاتساوى القيمه الحقيقيه • ومتوسط مربع الخطأ لتقدير ما يساوى تباين التقدير زائد مربع التحيز •

فيكون متوسط مربع الخطأ لمتوسط العينه هو

$$MSE_{\bar{x}} = \sigma^2_{\bar{x}} + (\bar{x} - \hat{\bar{x}})^2$$

حيث ان $MSE_{\bar{x}}$ = متوسط مربع الخطأ لمتوسط العينه ،

$(\bar{x} - \hat{\bar{x}})$ = التحيز الناتج عن استخدام $\bar{x}$ لتقدير $\hat{\bar{x}}$ •

في الصوره العامه : اذا كانت $\hat{\theta}$ تقدير لـ θ فان التباين لـ $\hat{\theta}$ يساوى

$$\sigma^2_{\hat{\theta}} = Var(\hat{\theta}) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_j - E(\hat{\theta}))^2$$

حيث ان N = عدد العينات الممكنه من المجتمع بالحجم المعين .

$E(\hat{\theta})$ = القيمه المتوقعه لـ $(\hat{\theta})$.

ويكون متوسط مربع الخطأ Mean Square Error يساوى

$$MSE_{\hat{\theta}} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_j - \theta)^2$$

حيث ان $MSE_{\hat{\theta}}$ = متوسط مربع الخطأ للتقدير $(\hat{\theta})$.

فاذا كانت القيمه المتوقعه لـ $(\hat{\theta})$ تساوى القيمه الحقيقيه للمجتمع (θ) فان $(\hat{\theta})$ تكون تقديرا غير متحيزا لـ (θ) . اى ان

$$\sigma^2_{\hat{\theta}} = MSE_{\hat{\theta}}$$

ولاثبات ذلك فاننا نتبع الخطوات الاتيه

$$\sum_{j=1}^{N}(\hat{\theta}_j - \theta)^2 = \sum_{j=1}^{N}\left\{[\hat{\theta}_j - E(\hat{\theta})] + [E(\hat{\theta}) - \theta]\right\}^2$$

$$= \sum_{j=1}^{N}[\hat{\theta}_j - E(\hat{\theta})]^2 + 2\sum_{j=1}^{N}[\hat{\theta}_j - E(\hat{\theta})][E(\hat{\theta}) - \theta] + \sum_{j=1}^{N}(E(\hat{\theta}) - \theta)^2$$

$$= \sum_{j=1}^{N}[\hat{\theta}_j - E(\hat{\theta})]^2 + 2[E(\hat{\theta}) - \theta]\sum_{j=1}^{N}[\hat{\theta}_j - E(\hat{\theta})] + N[E(\hat{\theta}) - \theta]^2$$

والمقدار $\sum_{j=1}^{N}(\hat{\theta}_j - E(\hat{\theta}))$ = صفرا

$$\therefore \sum_{j=1}^{N}(\hat{\theta}_j - \theta)^2 = \sum_{j=1}^{N}(\hat{\theta}_j - E(\hat{\theta}))^2 + N(E(\hat{\theta}) - \theta)^2$$

ولكى نستكمل مربع الخطأ يأتى

$$MSE_{\hat{\theta}} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_j - \theta)^2 = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_j - E(\hat{\theta}))^2 + \frac{N(E(\hat{\theta}) - \theta)^2}{N}$$

$$= \sigma^2_{\hat{\theta}} + [E(\hat{\theta}) - \theta]^2$$

والمقدار $(E(\hat{\theta}) - \theta)^2$ = صفرا اذا كانت $\hat{\theta}$ تقديرا غير متحيزا لـ θ .

من هذا يتبين ان متوسط مربع الخطأ MSE لتقدير مايساوى تباين هذا التقدير زائد مربع التحيز.

مثال : لنفرض لدينا مجتمع فرضى يتكون من خمسه مزارعين ، والجدول رقم (٤) يبين المساحه بالفدان المملوكه لكل منهم فى سنة ١٩٥٢ ، سنة ١٩٥٥ . والمطلوب تقدير متوسط المساحه المملوكه لهذا المجتمع وذلك باستخدام تقدير النسبه . وذلك باستخدام عينه حجمها = ٢ .

جدول رقم (٤)

الفرد	المساحه سنة ١٩٥٢ x_0	المساحه سنة ١٩٥٥ x
A	١	صفر
B	١	٢
C	٤	٣
D	٦	٥
E	١٢	١٥
المجموع	٢٤	٢٥
المتوسط البسيط	٤٨ر٤ فدان	٥ر٠

فيكون تقدير النسبه =

$$\hat{\theta}_2 = \frac{\bar{x}}{\bar{x}_0} \bar{X}_0$$

حيث ان $\bar{x}$ متوسط العينه لسنة ١٩٥٥ ،

$\bar{x}_0$ متوسط العينه لسنة ١٩٥٢ ،

$\bar{X}_0$ متوسط المجتمع الحقيقى لسنة ١٩٥٢

ان عدد العينات (حجمها = ٢) الممكن الحصول عليها من هذا المجتمع المكون من خمسة مزارعين هو عشرة عينات . والجدول رقم (٥) يبين المتوسط لسنتى ١٩٥٢ ، ١٩٥٥ ، تقدير النسبه للعشره عينات .

جـدول رقـــم (٥)

الرقم مسلسل	العينـــه	متوسـط ١٩٥٢ $\bar{x}_0$	متوسـط ١٩٥٥ $\bar{x}$	$\bar{x}^2$	$\hat{\theta}_2 = \frac{\bar{x}}{\bar{x}_0}\chi_0$	$(\hat{\theta}-\theta)^2$
١	AB	١ر٠	١ر٠	١ر٠	٤ر٨٠	٠ر٤
٢	AC	٢ر٥	١ر٥	٢ر٢٥	٢ر٨٨	٤ر٤٩
٣	AD	٣ر٥	٢ر٥	٦ر٢٥	٣ر٤٣	٢ر٤٦
٤	AE	٦ر٥	٧ر٥	٥٦ر٢٥	٥ر٥٤	٠ر٢٩
٥	BC	٢ر٥	٦ر٥	٦ر٢٥	٤ر٨٠	٠ر٤
٦	BD	٣ر٥	٣ر٥	١٢ر٢٥	٤ر٨٠	٠ر٤
٧	BE	٦ر٥	٨ر٥	٧٢ر٢٥	٦ر٢٨	١ر٦٤
٨	CD	٥ر٠	٤ر٠	١٦ر٠٠	٣ر٨٤	١ر٣٦
٩	CE	٨ر٠	٩ر٠	٨١ر٠٠	٥ر٤٠	٠ر١٦
١٠	DE	٩ر٠	١٠	١٠٠ر٠٠	٥ر٢٣	٠ر١١
	المجمـوع	٤٨ر٠	٥٠ر٠	٣٥٣ر٥٠	٤٧ر١٠	١٠ر٦٤
	متوسط المتوسطات $E(\hat{\theta})$	٤ر٨	٥ر٠	———	٤ر٧١	———

من هذا يتبين ان $E(\hat{\theta}_2)$ لتقدير النسبه = ٤ر٧١ وهذه لاتساوى (θ) القيمة الحقيقية والتى = ٥ ، ويكون التحيز مساويا ٥ — ٤ر٧١ = ٠ر٢٩

ويمكن حساب تباين ومتوسط مربع الخطأ لكل من $\bar{x}$ ، $\hat{\theta}_2$ من البيانات السابقه مع العـلم بـأن

$$\sum_{j=1}^{10}(\hat{\theta}_j)^2 = ٢٣١ر٦٢٣٨$$

وذلك كما هو موضح بالجدول رقم (٦) .

جـــدول رقـــم (٦)

	التبايـــــن = σ^2	متوسط مربع الخطأ = MSE	التحـيز
المتوسط البسيط $\bar{x}$	$\frac{1}{10}(٥٣٥,٥ - \frac{(٥٠)^2}{10}) = ١٠,٣٥$	$\frac{1}{10}(٥٣٥,٥ - \frac{(٥٠)^2}{10}) = ١٠,٣٥$	صفـــر
تقدير النسبـــه $\hat{\theta}_2$	$\frac{1}{10}(٢٣١,٦٢٣٨ - \frac{(٤٧,١)^2}{10}) = ٠,٩٧٨٢$	$\frac{1}{10}(١٠,٦٤) = ١,٠٦٤٠$	٠,٠٨٥٨

من هذا يتبين ان مربع التحيز $(٠,٢٩)^2$ تساوى تقريبا ٠,٠٨٥٨ وهى الفرق بين التباين ومتوسـط مربع الخـطأ .

مما سبق يتبين أن تقدير النسبة يمثل احدى طرق المعاينة التى يمكن بها استخدام المعلومات المتوفرة من المجتمع حتى يمكن تقليل خطأ المعاينة .

والطريقة الثانية من طرق المعاينة والتى يمكن بها تقليل خطأ المعاينة تعرف بطريقة المعاينة الطبقية Stratified Sampling وتستعمل أساسا لتقليل تباين التقدير (σ^2) وذلك عن طريق استخدام المعلومات المتوفرة عن المجتمع الاصلى فمثلا قد تستعمل المعلومات السابقة عن المجتمع لتقسيمه الى مجموعات بحيث تكون المفردات داخل كل مجموعه أكثر تجانسا من المفردات فى المجتمع الكلى . وبالتالى اذا سحبت العينة من كل مجموعة على حدة وذلك باستعمال المعاينة العشوائية البسيطة ، فنكون بذلك قد حققنا تمثيل كل مجموعة وفى نفس الوقت نكون قد حققنا أيضا شرط المعاينة الاحتمالية اذا كان اختيارنا للمفردات داخل كل مجموعة عشوائيا . ويطلق على المجموعات التى تسحب منها العينة اصطلاح طبقات .

ويجب التنبيه هنا الى أن تقسيم المجتمع الى طبقات يجب أن يتم على أساس الصفة أو الخاصية المراد قياسها . أى بمعنى أنه اذا لم يتم تقسيم المجتمع الى الطبقات بحيث أن تكون المفردات داخل كل طبقة متجانسة بالنسبة للخاصية أو الصفة المراد دراستها وقياسها ، فانه لن تتحقق الفائدة المرجوة من استخدام طريقة المعاينة الطبقية .

المعاينة الطبقية المتناسقة Proportionate Stratified Sampling

ويقصد بالمعاينة الطبقية المتناسقة تلك الطريقة التى يقسم فيها المجتمع الى طبقات وان تمثل كل طبقة فى العينة بنفس النسبة التى تمثل بها كل الطبقات . ولتوضيح ذلك نفرض مجتمع يتكون من ١٢ مفرده (A, B, C, D, E, F, G, H, I, J, K, L) واننا قمنا بتقسيمهم الى اربعة طبقات كالاتى : طبقة (١) وبها ثلاث مفردات (A , J , D) طبقة (٢) بها (L , G , F) طبقة (٣) (I , H , C) طبقة (٤) بها (E , K , B) فاذا قمنا بسحب فرد من كل طبقة حتى تتكون لدينا عينة مكونة من أربعة افراد . فبذلك نكون

قد اتبعنا طريقة المعاينة الطبقية المتناسقة . وبالاضافة الى أن طريقة المعاينة الطبقية سوف تعمل على تقليل خطأ المعاينة فان تقديراتنا بطريقة المعاينة الطبقية المتناسقة ستكون أيضا غير متحيزة وذلك لان احتمال ظهور اى مفرده من أى من الطبقات الاربع يكون متساويا .

المعاينة الطبقية الغير متناسقة Disproportionate Stratified Sampling

وفى هذه الحالة فان الطريقة التى يقسم بها المجتمع الى طبقات لن تسمح بان تمثل كل طبقة فى العينة بنفس النسبة التى تمثل بها كل الطبقات . ولتوضيح ذلك لنفرض اننا قمنا بتقسيم المجتمع السابق المكون من اثنى عشره مفرده الى اربعة طبقات بحيث ان تتكون الطبقات الاربع كالاتى :

الطبقة	المفردات
١	A, J, D, L
٢	G, F, I, H
٣	C, E, K
٤	B

فاذا سحبنا كل العينات الممكنة والتى حجمها أربعة مفردات (عدد العينات الممكنة فى هذه الحالة =٤٨ عينة) فاننا سنلاحظ ان المفرده (B) ستظهر فى كل العينات ، اى بمعنى ان فرصة ظهورها فى العينة ستكون اكيدة ، كذلك فان احتمال ظهور مفرده من المفردات (A, J, D, L, G, F,I,H) ستكون $\frac{1}{4}$ فى حين انها كانت $\frac{1}{3}$ فى حالة المعاينة الطبقية المتناسقة .

فاذا فرض اننا اردنا أن نقدر متوسط العينة فيمكن تقدير ذلك بطريقتين :

١ـ تقدير المتوسط الغير مرجح : وفى هذه الحالة يكون

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4}$$

حيث أن $\bar{x}$ المتوسط الغير مرجح .

X_1 المفردة المسحوبة من الطبقة رقم (١) وبالمثل X_2 ، X_3 ، X_4 .

ويعتبر هذا التقدير متحيزا وذلك لعدم تساوى احتمال ظهور كل مفردة فى العينة المسحوبة .

٢ـ تقدير المتوسط المرجح : وفى هذه الحالة ترجح كل مفردة مسحوبة من طبقة ما بعدد المفردات الموجودة فى الطبقة . ويكون

$$\bar{X} = \frac{4X_1 + 4X_2 + 3X_3 + X_4}{12}$$

ويلاحظ اننا قسمنا على عدد المفردات الكلى والذى يساوى حاصل جمع الاوزان التى استعملت فى توجيهنا ويعتبر هذا التقدير غير متحيزا .

مثال : لنفرض انه من الممكن تقسيم المجتمع الفرضى السابق والمكون من خمسة مزارعين الى ثلاثة مجموعات اى طبقات بحيث ان الطبقة الاولى تحتوى على (A , B) والطبقة الثانية تحتوى على (C , D) والطبقة الثالثة على (E) . والجدول الاتى يبين المساحة المملوكة بالفدان لكل فرد .

الفرد	المساحة بالفدان
A	صفر
B	٢
C	٣
D	٥
E	١٥
المجموع	٢٥
المتوسط	٥

الطبقة	الافراد
١	A, B
٢	C, D
٣	E

والمطلوب سحب كل العينات التى تتكون من ثلاثة افراد على أن تحتوى العينة فردا من كل طبقة ٠
ثم احسن التقديرات الاتية :

١ـ احسب التقدير المرجح

$$\hat{\theta}_3 = \frac{2X_1 + 2X_2 + X_3}{5}$$

والاوزان هنا عبارة عن عدد الافراد فى كل طبقة ٠

٢ـ احسب القيمة المتوقعة للتقدير $\hat{\theta}_3$

٣ـ احسب تباين التقدير $\sigma^2_{\hat{\theta}_3}$ ، ومتوسط مربع الخطأ $MSE_{\hat{\theta}_3}$

٤ـ احسب المتوسط الغير مرجح

$$\bar{X} = \frac{X_1 + X_2 + X_3}{3}$$

٥ـ احسب القيمة المتوقعة للمتوسط $\bar{X}$

٦ـ احسب تباين التقدير $\sigma^2_{\bar{X}}$ ومتوسط مربع الخطأ $MSE_{\bar{X}}$

رقم العينة	العينة	$\hat{\theta}_3$	$\bar{X}$	$\hat{\theta}_3^2$	$[\bar{X} - E(\bar{X})]^2$	$[\hat{\theta}_3 - E(\hat{\theta})]^2$
١	ACE	4.2	6.00	17.64	0.45	0.64
٢	ADE	5.0	6.67	25.00	0.00	0.00
٣	BCE	5.0	6.67	25.00	0.00	0.00
٤	BDE	5.8	7.33	33.64	0.44	0.64
المجموع		20.0	26.67	101.28	0.89	1.28
القيمة المتوقعة		5.0	6.67			

من الجدول السابق يتبين ان القيمة المتوقعة للتقدير $\hat{\theta}_3 = 5$ وهى تساوى متوسط المجتمع الاصلى ، اى ان المتوسط المرجح كان غير متحيزا ، أما بالنسبة للمتوسط $\bar{X}$ فنجد أن قيمته المتوقعة = ٦٫٦٧ وهذه لا تساوى متوسط المجتمع الحقيقى ، اى ان المتوسط $\bar{X}$ الغير مرجح يكون تقديرا متحيزا ٠

ويكون التباين للتقدير $\hat{\theta}_3$ المرجح مساويا

$$\sigma^2_{\hat{\theta}_3} = \frac{1}{4}\left[101.28 - \frac{400}{4}\right] = \frac{1}{4}\left[1.28\right] = 0.32$$

ويكون متوسط مربع الخطأ $MSE_{\hat{\theta}_3}$ للتقدير المرجح $\hat{\theta}_3$ مساويا

$$MSE_{\hat{\theta}_3} = \frac{\sum_{j=1}^{4} (\hat{\theta}_{j3} - \theta)^2}{4} = \frac{(0.8)^2+(0)^2+(0)^2+(0.8)^2}{4} = \frac{1.28}{4} = 0.32$$

من هذا يتبين أن $\sigma^2_{\hat{\theta}_3} = MSE_{\hat{\theta}_3}$ أى ان التقدير $\hat{\theta}_3$ تقديرا غير متحيزا

ويكون تباين المتوسط $\bar{x}$ الغير مرجح مساويا

$$\sigma^2_{\bar{x}} = \frac{1}{4}\sum_{j=1}^{4}\left(\bar{X}_j - E(\bar{X})\right)^2 = \frac{(0.67)^2+(0)^2+(0)^2+(0.66)^2}{4} = \frac{0.89}{4} = 0.22$$

ويكون متوسط مربع الخطأ $MSE_{\bar{x}}$ للتقدير $\bar{x}$ الغير مرجح مساويا .

$$MSE_{\bar{X}} = \frac{1}{4}\sum_{j=1}^{4}\left(\bar{X}_j - \bar{\bar{X}}\right)^2 = \frac{(1)^2+(1.67)^2+(1.67)^2+(2.33)^2}{4} = \frac{12.0067}{4} = 3.00$$

من هذا يتبين أن $\sigma^2_{\bar{x}}$ لا تساوى $MSE_{\bar{x}}$ للمتوسط $\bar{x}$ الغير مرجح لذا فانه يعتبر تقديرا متحيزا .

مما سبق يمكن أن نستنتج الآتى :

١ - فى بعض الحالات يكون من الأصلح استعمال طريقة المعاينة البسيطة بالاضافة الى أنها قد تكون هى الطريقة المثلى بمعنى انه اذا استعملت طريقة المعاينة الطبقية فان ذلك سيؤدى الى كبر تباين العينة .

٢ - من الممكن تقليل حجم تباين التقدير المتحصل عليه من العينة اذا أمكن تقسيم مفردات المجتمع فى مجموعات بحيث يكون الاختلاف بين المفردات داخل كل مجموعة أصغر ما يمكن ، وأن يكون الاختلاف بين المجاميع اكبر ما يمكن .

٣ - تستعمل طريقة المعاينة الطبقية - بفائدة كبرى - فى الحالات التى يتحيز فيها المجتمع بوجود مفردات ذات قيم شاذة والتى يمكن عن طريق المعاينة الطبقية عزلها فى طبقات منفصلة .

نقط التشابه بين المعاينة الطبقية والمعاينة البسيطة :

١- كل مفردة فى المجتمع يكون لها نفس الاحتمال لان تظهر فى العينة وذلك فى حالة المعاينة البسيطة والمعاينة الطبقية المتناسقة .

٢- تعتبر التقديرات المقدرة من المعاينة الطبقية تقديرات متسقة اذا ما قدرت هذه التقديرات بعناية . ويجب أن نتذكر هنا الى انه عند استعمال طريقة المعاينة الغير متناسقة فانه يكون من الضرورى الترجيح باوزان الطبقات والا ستكون التقديرات المتحصل عليها متحيزة جدا .

ومن بين المعلومات التى يمكن الباحث ان يستخدمها فى تقسيمه للمجتمع الى طبقات الخبرة المكتسبة لبعض الافراد اللذين لهم صلة بموضوع البحث . فاذا استعملت الخبرة الشخصية فى تقسيم المجتمع الى طبقات ، فان النتائج المتحصل عليها من العينة لن تكون متحيزة نتيجة لاستخدام الخبرة الشخصيه ، بل أن تباين التقديرات قد يقل اذا كانت الخبرة المستخدمة جيدة وكذلك فان تصحيح العينة سيتبع المعاينة الاحتماليه بشرط ان العينة ستحسب من كل طبقة بالطريقة العشوائية .

المعاينة المرحلية Cluster Sampling

لقد بينا فى الأجزاء السابقة انه يمكن زيادة دقة النتائج المتحصل عليها باستعمال طرق المعاينة وذلك بتغيير تصميم العينة وكذلك بتغيير طريقة التقدير دون الحاجة الى تغيير حجم العينة ولاحظنا أيضا انه لو كان لدينا طريقتين للمعاينة وكل منهم تستخدم عينة بنفس الحجم فاننا سنفضل بطبيعة الحال الطريقة التى تعطى لنا التقديرات والنتائج الاكثر دقة . ولكن هذا الاستنتاج سيكون صحيحا فقط اذا كانت تكاليف الوصول الى الوحدات الاولية فى كلتا الطريقتين واحدا . وذلك لاننا نسعى دائما الى الحصول على أقصى درجة من الدقة لكل وحدة تكاليف فاذا كانت تكاليف الوحدة الاولية فى طريقة ما أقل بكثير من تكاليفها فى طريقة اخرى ، فانه يمكننا فى هذه الحالة تكبير حجم العينة للطريقة الاولى عنه فى الطريقة الثانية . وعندما يأخذ الباحث فى اعتباره تكاليف المعاينة فان ذلك يؤدى الى التفكير فى طريقة المعاينة المرحلية . كذلك قد لا تتوفر فى كثير من الاحيان المعلومات الكافية واللازمة لتقسيم المجتمع الى طبقات تكون فيها المفردات اكثر تجانسا وهذا يدعو أيضا الى التفكير فى استعمال طريقة المعاينة المرحلية .

ولتفهم ما هو المقصود من طريقة المعاينة المرحلية دعنا نفرض اننا نرغب فى تقدير الدخل الفردى الزراعى فى ج . ع . م . وأنه يمكننا اتباع احدى الطريقتين الاتيتين :

١ — لنفرض انه لدينا قائمة باسماء جميع المزارعين فى ج . ع . م واننا سحبنا عينة مكونة من ١٥٠٠ مزارع من القائمة وذلك باستعمال طريقة المعاينة العشوائية . وباتباع هذه الطريقة سنجد ان المزارعين المختارين فى العينة سيكونون منتشرون فى أغلب مراكز وقرى الجمهورية .

٢ — لنفرض انه لدينا قائمة باسماء جميع المراكز فى ج . ع . م واننا سنقوم بسحب عينة من هذه المراكز وليكن حجمها ٢٠ مركزا ، واننا سنقوم بسحب عينة المزارعين المكونة من ١٥٠٠ مزارع من عينة المراكز المختارة . بهذه الطريقة سنتمكن من حصر عملنا فقط فى المراكز المختارة .

وكما سنرى فى هذا الجزء فان العينة المختارة بالطريقة الاولى اى بالطريقة العشوائية البسيطة سيكون خطأ المعاينة فيها أصغر من تلك المختارة بالطريقة الثانية ولكننا فى نفس الوقت نجد أن المزارعين المختارين سيكونون منتشرين على طول الجمهورية وعرضها وبالتالى فتكاليف

اجراء البحث ستكون مرتفعة للغاية ما دامت مصاريف الانتقال والاشراف ستكون مرتفعة . اما الطريقة الثانية فهى مثال توضيحى للمعاينة المرحلية وواضح ان تكاليف اجراء البحث ستكون أقل بكثير مما هو عليه فى الطريقة الاولى . فاذا امكن تقسيم المجتمع الكلى الى مجموعات ، وأن عينة من هذه المجموعات قد سحبت لتمثل المجتمع ، وأن هذه المجموعات ستعامل كوحدات معاينة Sampling Units ثم اختيرت الوحدات الاولية Elementary Units من المجموعات المختارة ، فاننا نكون قد اتبعنا طريقة المعاينة المرحلية .

ولتوضيح طريقة المعاينة المرحلية فاننا سنفرض انه يمكننا تقسيم المجتمع الفردى المكون من الاثنى عشرة فردا والمبين فى جدول (١) فى صفحة (١٤) الى ثلاثة مجموعات بحيث أن تحتوى كل مجموعة على أربعة افراد . (متوسط الاثنى عشرة فردا = ٣١١ جنيها)

مجموعة	الافراد	الدخل السنوى بالجنيه
I	A	٢٠٠
	F	٢٤٠
	H	٢٢٠
	I	٢٨٠
	المجموع	٨٤٠
	المتوسط	٢١٠
II	D	٣٢٠
	E	٢٨٠
	J	٣١٢
	K	٢٦٠
	المجموع	١١٧٢
	المتوسط	٢٩٣

مجموعه	الافـراد	الدخل السنوى بالجنيه
III	B	٣٤٠
	C	٤٦٠
	G	٣٨٠
	L	٥٤٠
	المجموع	١٧٢٠ جنيها
	المتوسط	٤٣٠ "

لنفرض اننا سنقوم باختيار عينة عشوائية مكونة من وحدة معاينة واحدة (مجموعة واحدة) من هذا المجتمع ، ونستعمل متوسط العينة المكونة من الاربعة أفراد كتقدير لمتوسط الدخل للاثنى عشرة فردا فى المجتمع . بطبيعة الحال سيكون لدينا فقط ثلاثة عينات ممكنة وتكـــون تقديراتهم كالاتـــى :

مجموعـه	متوسط الدخل بالجنيه
I	٢١٠
II	٢٩٣
III	٤٣٠
المجمـوع	٩٣٣
المتوسطات	٣١١

ويلاحظ ان متوسط المتوسطات لكل العينات الممكنة (أى القيمة المتوقعة للمتوسط) مساوية لمتوسط المجتمع الحقيقى والذى يساوى ٣١١ جنيها ، أى ان هذا التقدير غير متحيزا)

وبحساب التباين لهذا المتوسط فنجد أنه يساوى ٨٢٢٨ر٦٧٠ ويلاحظ كبر حجم هـــذا التباين ويرجع السبب الى ذلك ان تقسيم افراد المجتمع الى المجموعات الثلاثة كان بطريقة جعلت دخل الافراد متجانسا داخل كل مجموعة ٠ أى ان وحدات المعاينة (أى المجموعات) كانت أقرب فى تكوينها كالطبقات فى حالة المعاينة الطبقيـة ٠

يمكن تقسيم المجتمع الفرضى هذا الى مجموعات بطريقة أخرى حتى يمكننا من الحصول على تباين للتقدير (المتوسط فى هذه الحالة) أصغر من التباين المتحصل عليه فى التقسيـم السابق والتقسيم الجديد سيكون كالاتى ؛

	الافـــراد	الدخل السنوى بالجنيـه
I	A	٢٠٠٠
	D	٣٢٠٠
	J	٣١١٢
	L	٥٤٠٠
	المجموع	١٣٧١٢
	المتوسط	٣٤٣٠
II	F	٢٤٠٠
	G	٣٨٠٠
	H	٢٢٠٠
	I	١٨٠٠
	المجموع	١٠٢٠٠
	المتوسط	٢٥٥٠

III	B	٣٤٠
	C	٤٦٠
	E	٢٨٠
	K	٢٦٠
	المجموع	١٣٤٠ جنيها
	المتوسط	٣٥٥ "

وتكون تقديرات الثلاثة عينات الممكنة هى كالاتـــى :

مجموعـــه	متوسط الدخل بالجنيـه
I	٣٤٣
II	٢٥٥
III	٣٣٥
المجموع	٩٣٣
متوسط المتوسطات	٣١١

وهنا أيضا يكون متوسط المتوسطات للثلاثة عينات الممكنة (أى القيمة المتوقعة للمتوسط) مساويه لمتوسط المجتمع الحقيقى الذى يساوى ٣١١ جنيها • وبحساب تباين المتوسط فنلاحظ انه يساوى ١٥٧٨٫٦٧ أى أنه أقل بكثير من تباين المتوسط السابق • ويرجع ذلك أن تقسيـــم المجتمع فى هذه الحالة الى مجموعات كان بطريقة جعلت دخل الفرد داخل كل مجموعـــــه يخالف دخل الفرد الاخر بدرجة كبيرة •

أى انه فى التقسيم الاول فان المجموعات كانت تحتوى على أفراد متجانسة الدخـل ، ولذا كان التباين بين متوسطات كل العينات الممكنة كبيرا • اما فى حالة التقسيم الثانى فـان المجموعات كانت تحتوى على افراد غير متجانسة الدخل ولذا أمكن تقليل التباين بين متوسطـات

كل العينات الممكنة . أى ان أساس التقسيم الذى يجعل المعاينة الطبقية ذات كفاءة عالية فانه نفسه يقلل من كفاءة المعاينة المرحلية .

ومن سوء الحظ — فانه من الناحية العملية — قد يستحيل جعل المفردات داخل المجموعة الواحدة غير متجانسة بالدرجة التى تتواجد بها فى المجتمع الاصلى أو قد يتكلف ذلك نفقات طائلة . فمثلا نجد أن الافراد ذوى الدخول العالية يسكنون فى العادة فى مناطق معينة من المدن فى حين ان الافراد ذوى الدخول المنخفضة يسكنون فى مناطق اخرى . ففى هذه الحالة اذا امكن سحب عينة من كل منطقة فى المدينة فتكون بذلك قد مثلنا المجتمع الكلى للمدينة فى العينة . ولكن هذا يستلزم قوائم باسماء السكان فى كل منطقة حتى يمكن وضع الافراد ذوى الدخول المتقاربة فى قائمة واحدة معا ، وقد تكون نفقات هذه التجهيزات كبيرة لدرجة أن يصبح من المحال اجراء البحث .

التحيز واخطاء غير المعاينة
Bias and Nonsampling Enors

سنناقش فى هذا القسم اسباب تحيز نتائج المعاينة ، والطرق التى يمكن بها التحكم فى أو التخلص من هذا التحيز . وستشمل هذه المناقشة طرق المعاينة الغير عشوائية وتحيزها كذلك التحيز الذى قد يظهر فى المعاينة العشوائية ، كذلك مشاكل التى تنتج عن عدم استجابة المبحوثين للباحثين ، واخطاء غير المعاينة الاخرى . وفيما يلى شرح مبسط لاهم انواع التحيز ومصادرها وكيفية التخلص منها :

١ـ التحيز الناتج عند اختيار وحدات المعاينة باحتمالات معروفة ولكنها مختلفة : ان احد مصادر التحيز فى تصميم العينة يظهر نتيجة استخدام الاحتمالات المختلفة عند اختيار وحدات المعاينة . وفى كثير من الاحيان تكون هذه الاحتمالات المختلفة معروفة ، وفى هذه الحالة يمكن التخلص من التحيز وذلك عن طريق تعديل طريقة الاختيار او عن طريق تعديل طريقة التقدير من العينة .

مثال : تقدير صفات وحدات عائلية من عينة من الافراد : لنفرض اننا نرغب فى تقدير نسبة عائلات مدرسة فى مدينة ما واللذين يملكون مزارعهم . قد يتبادر الينا من اول وهلة انه يمكننا سحب عينة من تلاميذ المدرسة ومعرفة ما اذا كانت عائلة أى من التلاميذ تملك عائلاتهم مزارعهم الى عدد التلاميذ فى العينة فاننا نحصل بذلك على تقدير لنسبة العائلات التى تملك مزارعها . وفى كثير من الاحيان فان النتيجة المتحصل عليها قد تكون متحيزة بالرغم من استخدام طريقة المعاينة المناسبة .

لنفرض أن الجدول رقم (٧) يمثل عدد التلاميذ فى المدارس فى مدينة ما وتوزيعهم على العائلات ، كذلك يبين عدد العائلات التى تملك مزارعها . ولنفرض ايضا اننا سنتبع طريقة المعاينة البسيطة اى ان كل تلميذ سيعطى نفس الفرصة للظهور فى العينة .

جدول رقم (٧)

عدد التلاميذ في العائلة	العائلات				التلاميذ	
	عدد	النسبة المئوية لمجموع العائلات	يملكون مزارعهم		عدد	النسبة المئوية لمجموع التلاميذ
			عدد	النسبة المئوية لمجموع العائلات		
٤	٥٠٠٠	١٠	٥٠٠	١٠	٢٠٠٠٠	٢٠
٣	١٠٠٠٠	٢٠	٣٠٠٠	٣٠	٣٠٠٠٠	٣٠
٢	١٥٠٠٠	٣٠	٦٠٠٠	٤٠	٣٠٠٠٠	٣٠
١	٢٠٠٠٠	٤٠	١٢٠٠٠	٦٠	٢٠٠٠٠	٢٠
المجموع	٥٠٠٠٠	١٠٠	٢١٥٠٠	١٠٠	١٠٠٠٠٠	١٠٠

من الجدول السابق نلاحظ أن ٢٠٪ من تلاميذ المدارس قد اتوا من عائلات ذات اربعة تلاميذ في المدرسة . وبالتالي فاذا كانت العينة المسحوبة جيدة ، فاننا نتوقع ان ٢٠٪ من التلاميذ في العينة سيأتون من هذه العائلات ، لذا فان ٢٠٪ من العائلات في العينة سيكون لديها اربعة تلاميذ في حين ان ١٠٪ فقط من عائلات المدرسة في المدينة لها اربعة تلاميذ . ومن الجهة الاخرى فان ٢٠ ٪ فقط من تلاميذ العينة سيأتون من عائلات ذات تلميذ واحد في حين أن ٤٠ ٪ من العائلات في المدينة تتبع هذا النوع . لذا يكون تقديرنا لنسبة العائلات التي تملك مزارعها الى المجموع الكلي للعائلات متحيزا . وذلك لان العائلات الكبيرة الحجم مثلت اكثر في العينة في حين ان تمثيل العائلات الصغيرة كان بدرجة أقل . اي بمعنى ان المجموعات ذات الملكية المنخفضة كانت ممثلة في العينة بدرجة أكبر في حين ان المجموعات ذات الملكية العالية مثلت في العينة بدرجة اقل . ويمكن في مثالنا هذا ان نتجنب هذا التحيز او التخلص منه كلية وذلك باستخدام احدى الثلاث طرق الاتية :

أ ـ المعاينة من قائمة بعائلات المدارس : يمكن تجنب هذا التحيز وذلك عن طريق تجهيز قائمة باسماء العائلات في المدارس وبذلك نكون قد تمكننا من عدم ظهور اسم العائلة الواحدة اكثر من

مرة فى قائمة التلاميذ . وبذلك تكون لكل عائلة نفس الفرصة لان تسحب فى العينة وبالتالى ستكون النتائج المتحصل عليها غير متحيزة .

ب — يمكن التخلص من التحيز باستخدام الاوزان المناسبة : قد يكون من الصعب تجهيز قائمة باسماء العائلات ، وفى هذه الحالة يمكن سحب العينة من قائمة تلاميذ المدارس ويتحصل منها على تقديرات متسقة . وذلك بالاخذ فى الاعتبار الاحتمالات التى تظهر بها العائلات المختلفة فى العينة .

لنفرض ان (x_i) تمثل الصفة المراد دراستها فى العائلة والفرد (i) المسحوب من العينة ، ان (a_i) هو عدد الافراد فى العائلة اللذين لهم الفرصة لان يظهروا فى العينة ، فيحسب التقدير الغير متحيز لاجمالى الصفة المدروسة لكل العائلات فى المجتمع كالاتى :

$$\bar{x} = \frac{N}{n} \sum_{i=1}^{n} \frac{x_i}{a_i}$$

حيث أن $\bar{x}$ تقدير لاجمالى الصفة المدروسة لكل العائلات فى المجتمع ،

N عدد التلاميذ فى المجتمع ،

n عدد التلاميذ فى العينة .

ويلاحظ أننا قسمنا على (a_i) فى تقديرنا السابق — اى عدد الافراد فى العائلة اللذين لهم الفرصة لان يظهروا فى العينة — حتى تحدث عملية موازنة للاحتمالات المختلفة لظهور الافراد فى العينة . فمثلا العائلات المكونة من ثلاثة افراد وكل منهم له الفرصة فى الظهور فى العينة ، وبذلك يكون لمثل هذه العائلات فرصة للظهور فى العينة بثلاثة افراد أى يكون احتمال تمثيلها فى العينة ثلاثة امثال احتمال تمثيل العائلات ذات الفرد الواحد .

مثال : لنفرض أن عينة مكونة من خمسة تلاميذ (A, B, C, D, E) قد سحبت من المجتمع الفرض الذى يتكون من ٢٠ تلميذا . والجدول رقم (٨) يبين العائلات التى تملك مزارعها (حيث أن $x_i = 1$) ، العائلات التى لا تملك مزارعها (حيث أن $x_i = 0$) وكذلك الدخل السنوى للعائلة بالجنيه ، وعدد التلاميذ فى كل عائلة .

جدول رقم (٨)

التلميذ	صفات العائلة: تملك المزرعة x_i	دخل العائلة السنوى بالجنيه	عدد التلاميذ فى العائلة (a_i)
A	0	600	4
B	0	150	1
C	1	300	2
D	1	400	1
E	1	200	4

ويكون تقدير عدد العائلات التى تملك مزارعها مساويا

$$x' = \frac{N}{n} \sum_{i=1}^{n} \frac{x_i}{a_i} = N \frac{\frac{x_1}{a1} + \frac{x_2}{a2} + \frac{x_3}{a3} + \frac{x_4}{a4} + \frac{x_5}{a5}}{n}$$

$$= 20 \frac{\frac{0}{4} + \frac{0}{1} + \frac{1}{2} + \frac{1}{1} + \frac{1}{4}}{5} = 7$$

أى أن تقدير عدد العائلات فى المجتمع المأخوذة منه العينة (المجتمع = ٢٠) واللذين يملكون مزارعهم هو ٧ عائلات ، وبالمثل يكون تقدير اجمالى الدخل لهذه العائلات مساويا

$$x' = 20 \frac{\frac{600}{4} + \frac{150}{1} + \frac{300}{2} + \frac{400}{1} + \frac{200}{4}}{5} = 3600$$

حـ - يمكن تجنب التحيز وذلك باقران بيانات العائلة الى فرد فى العائلة : فمثلا فى مثالنا السابق يمكن أن نقرن بيانات العائلة بالتلميذ الاكبر فى العائلة ، وحيث ان كل عائلة لديها تلميذ واحد كبير لذا فان كل عائلة سيكون لها نفس الفرصة لان تسحب فى العينة فعند اختيار

العينة فاننا سنهمل كل التلاميذ المختارين اللذين لا يعتبرون اكبر الابناء فى عائلاتهم وسنستعمل فى تقديرنا فقط البيانات الخاصة بالتلاميذ الاكبر سنا فى العائلة .

فمثلا فى مثالنا السابق البسيط لنفرض أن المعلومات الاضافية الاتية كانت معلومة لنا عن العينة المكونة من الخمسة افراد (A, B, C, D, E) .

التلميذ	هل اكبر ابناء العائلة
A	نعم
B	نعم
C	لا
D	نعم
E	لا

وهذا فاننا فى هذه الحالة سنستعمل فى تقديرنا بيانات الدخل للعائلات الممثلة باكبر التلاميذ (A, B, D) اى ان حجم العينة سيكون فى هذه الحالة ثلاثة افراد بدلا من خمسة افراد ، وتكون كل عائلة فى هذه الحالة لها نفس الفرصة لان تظهر فى العينة ولذا لا تكون هناك حاجة الى استعمال اوزان خاصة فى تقديرنا وتستعمل طرق التقدير بالطرق العادية تبعا لطريقة المعاينة المستخدمة .

٢ - الاستجابة الاختيارية لصحائف الاستبيان المرسلة بالبريد : يعتقد بعض الباحثين انه لا يوجد اى سبب يدعو الى الاعتقاد الى ان هناك اختلاف بين الافراد اللذين يستجيبون للبحث بطريق البريد عن الافراد اللذين لا يستجيبون ، ويعتبرون أن الافراد اللذين يستجيبون يمثلون عينة عشوائية من المجتمع تحت الدراسة . ولكن فى الواقع فان الخبرات السابقة تثبت عدم صحة هذا الاعتقاد ، أى ان صفات الافراد او المؤسسات أو وحدات الانتاج التى تهتم بالاستجابة تختلف عن اللذين لا يستجيبوا . ويرجع ذلك لعدم التجانس بين افراد المجتمع الواحد .

ويمكن تقليل التحيز - ولكن لا يمكن التخلص منه - وذلك بارسال صحائف الاستبيان الى مجموعات مختلفة من المجتمع ، مثل مجموعات منفصلة حسب السن والجنس ، مجموعات منفصلة حسب الوظائف ، ومجموعات اخرى من المجتمع تحت الدراسة والتى تكون اكثر تجانسا للاستجابة لصحائف الاستبيان المرسلة اليهم بالبريد .

٣- التحيز الناتج عن اختيار طريقة المعاينة بعد الحصول على نتائج العينة : فى بعض الاحيان يلجأ بعض الباحثون الى اختبار جزء من العينة المسحوبة للحصول على معلومات تساعدهم فى وضع قراراتهم الخاصة بطرق المعاينة التى سيتبعونها . وهذه الطريقة قد يتسبب عنها ظهور التحيز فى نتائج العينة . وقد تكون هذه الطريقة مفيدة ولها تأثير حسن على النتائج اذا ما استعملت بطريقة مناسبة ولكنها فى احيان اخرى تعطى نتائج متحيزة . وبالمثل قد يقرر الباحث اى طرق التقدير سيقوم باستخدامها بعد دراسة جزء من العينة ، ففى بعض الاحيان تساعد هذه الطريقة على تحسين النتائج) ومدى اعتماد الباحث عليها وفى احيان اخرى ينتج عنها نتائج متحيزة .

وأنه من الصعب وضع المبادئ العامة لتحديد الحالات التى يمكن فيها استخدام المعلومات المتحصل عليها من العينة لوضع القرارات الخاصة بطرق المعاينة والتقدير ، ولكن فيما يلى سنستعرض بعض الامثلة . والقاعدة الاساسية التى يجب اتباعها هى انه يجب على الباحث أن يتجنب وضع قراراته لطرق المعاينة أو طرق التقدير على اساس نتائج جزء من العينة ما لم يكن متأكدا من ان هذا لن يتسبب فى تحيز النتائج .

أ - اختيار عينة اخرى بعد النظر الى نتائج العينة : من الطرق التى يتسبب عنها تحيز النتائج ان يقوم الباحث باختبار نتائج عينة مبدئية ليقرر ما اذا كانت هذه النتائج مقبولة ام لا . فاذا كانت مقبولة فانه يستخدم نتائج هذه العينة ، اما اذا كانت غير مقبولة فانه يهمل هذه النتائج ويسحب عينة اخرى وقد تكون بطريقة اخرى على أمل انه قد يحصل على نتائج تشابه ما يتوقع . وبهذه الطريقة فيمكن للباحث أن يحصل على اى من النتائج المرغوبة أو انه يستطيع اثبات اى نقطة أو مسألة .

ب - أختيار حجم العينة بعد دراسة نتائج عينة اولية : قد يلجأ الباحث - بعد دراسة نتائج عينة أولية - الى زيادة حجم العينة وخاصة اذا كان تباين العينة كبيرا ، على امل أن يستطيع ان يقلل من حجم التباين . قد ينتج من زيادة حجم العينة - تحت هذه

الظروف أن تصبح النتائج متحيزة وخاصة اذا كان حجم العينة الاولية صغير لدرجة ان تقدير تباينها يكون غير موثوق به ، ولكن اذا كان حجم العينة الاولية كبيرا لدرجة أن تقدير معامل الاختلاف ($\frac{ك}{\bar{س}}$) لها لا يزيد عن ١٥ ٪ ، فانه يمكن فى هذه الحالة استخدام تلك الطريقة ويكون الباحث فى هذه الحالة متاكدا انه حتى ولو ظهر بعض التحيز فانه سيكون صغيرا بالنسبة للخطأ المعيارى لتقدير العينة .

جـ - تحديد طريقة التقدير بعد دراسة نتائج العينة : من الطرق المقبولة والتى يمكن فيها للباحث ان يستعمل نتائج العينة ينحصر فى اختيار طريقة التقدير . ويمكن اتباع هذه الطريقة فقط اذا حدد الباحث مقدما انه سيقوم باستخدام عدة طرق للتقدير ليختار من بينها . لذا فانه يحسب من العينة تقديرات التباين لكل طرق التقدير المختارة ثم يختار طريقة التقدير التى تعطى له اصغر تباين .

المعاينة الطبقية العشوائية البسيطة
Stratified Simple Random Sampling

يقصد بطريقة المعاينة الطبقية العشوائية البسيطة تلك الطريقة التى يقسم فيها المجتمع الكلى الى طبقات تكون المفردات داخل كل طبقة أكثر تجانسا ويكون اختيار العينة من كل طبقة بطريقة المعاينة العشوائية البسيطة .

لنفرض أن عدد الطبقات التى يقسم المجتمع الكلى اليها هو (L) ، وان عدد المفردات الكلى فى المجتمع هو (N) ، وان (N_h) تمثل عدد المفردات فى الطبقة (h) . فيكون

$$N = \sum_{h=1}^{L} N_h = N_1 + N_2 + \dots + N_L$$

ويرمز لحجم العينة المسحوبة من الطبقة (h) بالرمز (n_h) ، ويكون حجم العينة الكلى (n) والمسحوب من كل الطبقات مساويا .

$$n = \sum_{h=1}^{L} n_h$$

ويرمز لقيمة الصفة (X) للمفردة (i) فى الطبقة (h) بالرمز (X_{hi}) ، فمثلا لو كانت عائلات المزارعين تمثل المجتمع المراد سحب منه العينة وان الصفة المراد دراستها هى دخلهم الزراعى ، واذا كان المجتمع يتكون من عشرة عائلات وانه أمكن تقسيمه الى ثلاثة طبقات ، بحيث ان عائلتين كانت فى الطبقة الاولى ، أربعة عائلات فى الطبقة الثانية ، أربعة عائلات فى الطبقة الثالثة فتكون

$$N = 10 ,$$
$$N_1 = 2 ,$$
$$N_2 = 4 ,$$
$$N_3 = 4 ,$$

وتمثل X_{11} ، X_{12} دخل العائلتين فى الطبقة الاولى ، وتمثل X_{21} ، X_{22} ، X_{23} ، X_{24} دخل الاربعة عائلات فى الطبقة الثانية ، وتمثل X_{31} ، X_{32} ، X_{33} ، X_{34} دخل الاربعة عائلات فى الطبقة الثالثة . ويكون الدخل الكلى للعائلات فى الطبقة الاولى مساويا

$$X_h = \sum_{i=1}^{N_h} X_{hi}$$

$$X_1 = X_{11} + X_{12}$$

ويمكن وضع ذلك فى الصورة الاتية

$$X_1 = \sum_{i=1}^{N_1} X_{1i}$$

ويكون مجموع دخل العائلات فى الطبقة (h) فى الصورة العامة مساويا

$$X_h = \sum_{i=1}^{N_h} X_{hi}$$

ويكون مجموع دخل العائلات فى كل الطبقات أى فى المجتمع مساويا

$$X = \sum_{h=1}^{L} X_h = \sum_{h=1}^{L} \sum_{i=1}^{N_h} X_{hi}$$

ويكون مجموع دخل العائلات فى الطبقة (h) من العينة مساويا

$$x_h = \sum_{i=1}^{n_h} x_{hi}$$

ويكون مجموع دخل العائلات فى العينة من كل الطبقات مساويا

$$x = \sum_{h=1}^{L} x_h = \sum_{h=1}^{L} \sum_{i=1}^{n} x_{hi}$$

ويكون متوسط المجتمع الكلى مساويا

$$\bar{X} = \frac{X}{N} = \frac{\sum_{h=1}^{L} X_h}{N}$$

ومتوسط الطبقة (h) يرمز له

$$\bar{X}_h = \frac{X_h}{N_h} = \frac{\sum_{i=1}^{N_h} X_{hi}}{N_h}$$

ومتوسط العينة من الطبقة (h) يرمز له

$$\bar{x}_h = \frac{x_h}{n_h} = \frac{\sum_{i=1}^{n_h} x_{hi}}{n_h}$$

ويكون تباين المفردات فى الطبقة (h) مساويا

$$\sigma_h^2 = \frac{\sum_{i=1}^{N_h} (X_{hi} - \bar{X}_h)^2}{N_h}$$

ويقاس التباين بين مفردات الطبقة (h) بمقياس آخر (S_h^2) ويكون مساويا

$$S_h^2 = \frac{\sum_{i=1}^{N_h} (X_{hi} - \bar{X}_h)^2}{N_h - 1} = \frac{N_h}{N_h - 1} \sigma_h^2$$

ويقاس التباين النسبى Rel-Variance كالاتـــى

$$V_h^2 = \frac{S_h^2}{\bar{X}_h^2}$$

ويكون تباين المجتمع الكلى مساويـــا

$$\sigma^2 = \frac{\sum_{h=1}^{L} \sum_{i=1}^{N_h} (X_{hi} - \bar{X})^2}{N}$$

وكذلـــك

$$S^2 = \frac{\sum_{h=1}^{L} \sum_{i=1}^{N_h} (X_{hi} - \bar{X})^2}{N - 1} = \frac{N}{N - 1} \sigma^2$$

ويكون التباين النسبى للمجتمع مساويا

$$V^2 = \frac{S^2}{\bar{X}^2}$$

تمرين : لنفرض لدينا مجتمع مكون من أثنى عشرة مزارعا ، وانه قسم الى ثلاثة طبقات بحيث تشمل الطبقة الأولى خمسة مزارعين والثانية ثلاثة مزارعين ، والثالثة أربعة مزارعين ، وأمام كل مزارع مبين المساحة التى يملكها بالفدان كما هو مبين فيما يلى :

Stratum 1

المزارع	المساحة بالفدان
X_{11}	6
X_{12}	10
X_{13}	2
X_{14}	4
X_{15}	8

Stratum 2

المزارع	المساحة بالفدان
X_{21}	9
X_{22}	18
X_{23}	12

Stratum 3

المزارع	المساحة بالفدان
X_{31}	20
X_{32}	26
X_{33}	16
X_{34}	26

والمطلوب حساب كل من :

١ — متوسط المساحة $(\bar{X}_i)$ لكل طبقة ،

٢ — متوسط المساحة للمجتمع $\bar{X}$ ،

٣ — تباين كل من الطبقات الثلاثة σ_1^2 ، σ_2^2 ، σ_3^2 على الترتيب

٤ — تباين كل من الطبقات الثلاثة S_1^2 ، S_2^2 ، S_3^2 على الترتيب

٥ — تباين المجتمع σ^2 ، تباين المجتمع S^2 .

الحل : الجدول رقم (٩) يبين طريقة حساب كل من القيم المطلوبة .

جـــــدول رقم (٩)

Stratum 3			Stratum 2			Stratum 1		
X^2	المساحة (X)	رقم الـزارع	X^2	المساحة (X)	رقم الـزارع	X^2	المساحة (X)	رقم الـزارع
400	20	1	81	9	1	36	6	1
676	26	2	324	18	2	100	10	2
256	16	3	144	12	3	4	2	3
676	26	4				16	4	4
						64	8	5
2008	88	المجموع	549	39	المجموع	220	30	المجموع
	22	$(\bar{X}_3)$		13	$(\bar{X}_2)$		6	$(\bar{X}_1)$
$\frac{1}{4}(2008-\frac{(88)^2}{4})=18$		σ_3^2	$\frac{1}{3}(549-\frac{(39)^2}{3})=14$		σ_2^2	$\frac{1}{5}(220-\frac{(30)^2}{5})=8$		σ_1^2
$\frac{1}{3}(2008-\frac{(88)^2}{4}=24$		S_3^2	$\frac{1}{2}(549-\frac{(39)^2}{3}=21$		S_2^2	$\frac{1}{4}(220-\frac{(30)^2}{5})=10$		S_1^2

متوسط المجتمع $(\bar{X}) = \frac{88+39+30}{12} = \frac{157}{12} = 13.08$

تباين المجتمع $\sigma^2 = \frac{1}{12}(2777 - \frac{(157)^2}{12}) = 60.24$

تباين المجتمع $S^2 = \frac{1}{11}(2777 - \frac{(157)^2}{12}) = 65.72$

أسس التقسيم الى طبقات :

لقد رأينا سابقا أن المعاينة الطبقية تساعد على زيادة ثقة الباحث فى النتائج المتحصل عليها من العينة . ويتوقف مقدار الزيادة فى دقة النتائج المتحصل عليها من المعاينة الطبقية على درجة التجانس بين مفردات الطبقة الواحدة ، أى على مقدار التباين فى الصفة المقدرة والذى يمكن ارجاعه الى الاختلاف بين الطبقات . وهذا بالتالى يتوقف على الطريقة التى قسم بها المجتمع الى طبقات .

وعند تحديد الحدود الفاصلة بين الطبقات المختلفة يجب استخدام كل المعلومات المتوفرة والتى يمكن بها وضع كل مفردة من مفردات المجتمع فى طبقات تختلف عن بعضها بالنسبة للصفة المراد قياسها أو بالنسبة لتكاليف جمع البيانات . ويجب أن يتبع اختيار المفردات من الطبقات طريقة المعاينة الاحتمالية أى أنه لا يسمح بأى حال من الأحوال أن يعتمد على الخبرة الشخصية فى اختيار وحدات المعاينة .

اختيار العينة الطبقية :

بعد تقسيم المجتمع الى طبقات وتحديد حجم العينة المطلوب سحبها من كل طبقة ، فان العينة تختار بنفس الطريقة التى تتبع فى حالة استخدام طريقة المعاينة البسيطة فقط مع معاملة كل طبقة على أنها مجتمع قائم بذاته . أى أنه اذا سحبنا العينات العشوائية البسيطة ($n_1, n_2, n_3, \ldots\ldots, n_L$) من الطبقات (L) على الترتيب ، فانه يكون لدينا فى النهاية عينة طبقية عشوائية بسيطة . وفى بعض الأحيان فان العينات تسحب من الطبقات بحيث يكون كسر المعاينة Sampling Fraction $\left(\frac{n}{N}\right)$ متساوى فى كل الطبقات وبذلك نحصل على ما يطلق عليه بالعينة الطبقية المتناسبة . وليس من الضرورى أن يكون كسر المعاينة متساويا لكل الطبقات . ويكون كسر المعاينة للطبقة (h) مساويا

$$f_h = \frac{n_h}{N_h}$$

جدول رقم (١٠) - انتاج اللبن اليومى

الطبقه	الانتاج اليومى بالرطل	عدد المنتجين فى كل طبقه	اجمالى الانتاج بالرطل	متوسط الانتاج بالرطل	الانحراف المعيارى بالرطل
١	٥٠٠٠ فاكثر	٥٣٨	٥٩٣٤٠٠٠	١١٠٢٩٫٧	٩٠٠٠
٢	١٠٠٠ - ٤٩٩٩	٤٧٥٦	٨٤٦٤٠٠٠	١٧٧٩٫٦	١٢٠٠
٣	اقل من ١٠٠٠	٣٠٩٦٤	٦٣١١٠٠٠	٢٠٣٫٨	٣٠٠
الاجمالى لكل الطبقات		٣٦٢٥٨	٢٠٧٠٩٠٠٠	٥٧١٫٢	١٦٨٤

فاذا فرض اننا اردنا سحب عينه من هذا المجتمع وليكن حجمها الكلى (n) مساويا ٢٥٠٠ منتجا ، فالمطلوب توزيع حجم العينه (n) حسب قاعدة الحجم الامثل على الثلاث طبقات - اى المطلوب حساب (n_1) ، (n_2) ، (n_3) . كذلك احسب التباين الامثل (s^2_{opt}) لهذا المجتمع .

الحل : لايجاد العينه (n_1) من الطبقه الاولى باستخدام قاعدة الحجم الامثل فيكون

$$n_1 = \frac{N_1 S_1}{\sum N_h S_h} n ,$$

$$n_2 = \frac{N_2 S_2}{\sum N_h S_h} n ,$$

$$n_3 = \frac{N_3 S_3}{\sum N_h S_h} n .$$

الطبقه	N_h	S_h	$N_h S_h$	$n_h = \frac{N_h S_h}{\sum N_h S_h} n$
١	٥٣٨	٩٠٠٠	٤٨٤٢٠٠٠	٦١٠
٢	٤٧٥٦	١٢٠٠	٥٧٠٧٢٠٠	٧١٩
٣	٣٠٩٦٤	٣٠٠	٩٢٨٩٢٠٠	١١٧١
المجموع	٣٦٢٥٨	—	١٩٨٣٨٤٠٠	٢٥٠٠

يلاحظ ان (n_1) حجم العينه من الطبقه الاولى يساوى ٦١٠ وحده ، وهذا اكبر من حجم الطبقه الاولى (N_1) والذى يساوى ٥٣٨ وحده . لذا فاننا نعيد توزيع حجم العينه الكلى (n) والمساوى ٢٥٠٠ وحده وذلك باخذ (n_1) للطبقه الاولى مساويا ٥٣٨ وحده وهو حجم الطبقه كلها ، وبعد ذلك يوزع عدد الوحدات الباقيه والمساويه ١٩٦٢ وحده على الطبقتين الثانيه والثالثه حسب قاعدة الحجم الامثل كالآتى :

الطبقه	N_h	S_h	$N_h S_h$	$n_h = \frac{N_h S_h}{\sum N_h S_h} n$	$N_h S_h^2$
١	٥٣٨	٩٠٠٠	٤٨٤٢٠٠	٥٣٨	—
٢	٤٧٥٦	١٢٠٠	٥٧٠٧٢٠٠	٧٤٧	٦٨٤٨٦٤٠٠٠٠
٣	٣٠٩٦٤	٣٠٠	٩٢٨٩٢٠٠	١٢١٥	٢٧٨٦٧٦٠٠٠٠
المجموع	٣٦٢٥٨	—	—	—	٩٦٣٥٤٠٠٠٠٠

نلاحظ من الجدول السابق :

١ - ان حجم العينه من الطبقه الاولى (n_1) مساويا لحم الطبقه ٠٥٣٨

٢ - ان حجم العينه من الطبقة الثانية (n_2) مساويا

$$n_2 = \frac{N_2 S_2}{(N_2 S_2 + N_3 S_3)} (n - n_1)$$

$$n_2 = \frac{5707200}{14996400} \times 1962 = \underline{747}$$

وبالمثل تكون (n_3) مساويـــه

$$n_3 = \frac{N_3 S_3}{(N_2 S_2 + N_3 S_3)}$$

$$= \frac{9289200}{14996400} \quad 1962 = \underline{1215}$$

٣ - يمكن حساب التباين الامثل ($\sigma^2(opt)$) وذلك بتطبيق المعادلة الآتيـه :

$$\sigma^2(opt) = \frac{1}{N^2} \left[\frac{(\Sigma N_h S_h)^2}{n} \Sigma N_h S_h^2 \right]$$

وبعد اسقاط القيم الخاصة بالطبقة الاولى ، وتكون القيم الداخله فى حسابنا هى :

$$(\Sigma N_h S_h)^2 = (5707200 + 9289200)^2 = 211549012960000$$

$$\Sigma N_h S_h^2 = 9635400000$$

$$n = 1962$$

$$N^2 = 1314642564$$

فيكون بذلـك

$$\sigma^2(opt) = \underline{74.6}$$

$$\left[\frac{\sum N_{45} N_{45}^2}{n} - \frac{\left(\sum N_{45} N_{45}\right)^2}{n} \right] \frac{1}{N_{45}^2}$$

ثانيا : مع اخذ التكاليف فى الاعتبار :

يمكن تعديل الطريقه السابقة لتوزيع الحجم الامثل للعينه على الطبقات المختلفة حتى يمكننا الاخذ فى الاعتبار التكاليف المختلفة للمعاينه من الطبقات المختلفة فاذا كانت تكاليف المعاينه تختلف فيما بينها اختلافا كبيرا فان حجم العينه لن يكون مقياسا جيدا للتكاليف الكليه للمعاينه . فمثلا لو فرض اننا قسمنا منتجى اللبن فى مثالنا السابق داخل كل طبقة الى مجموعتين بحيث تشمل المجموعة الاولى المنتجين اللذين تقع مزارعهم بالقرب من المدن والمجموعة الثانية تشمل اولئك اللذين تقع مزارعهم بعيدا عن المدن . وان تكاليف الحصول على المعلومات اللازمة لاجراء البحث من منتجى المجموعة الثانيه - بفرض الاستعانه بباحثين لعمل المقابلة الشخصية - كان اربعه اضعاف التكاليف اللازمة للحصول على المعلومات من المجموعة الاولى ، لذا فانه يترتب على ذلك انه يجب الاخذ فى الاعتبار كلا من تكاليف المعاينه والانحراف المعيارى فى مختلف الطبقات . وفى بعض الاحيان قد نلاحظ انه للحصول على المعلومات اللازمة من صغار المنتجين قد يتكلف الباحث اكثر من الحصول على المعلومات من كبار المنتجين او بالعكس . فاذا كانت الاختلافات فى التكاليف كبيرة (اى تساوى ١٠٪ او اكثر من المارت) فانه يتأتى ان يأخذ فى الاعتبار هذه الاختلافات بالاضافة الى العوامل الاخرى عند تقسيم المجتمع الى طبقات مختلفه .

سنفرض هنا ان تكاليف الحصول على المعلومات اللازمة من وحدة واحده من العينة ستكون ثابتة لطبقه معينة بصرف النظر عن عدد الوحدات المسحوبة من هذه الطبقة . ولنفرض ايضا ان التكاليف الكلية للحصول على المعلومات - دون الاخذ فى الاعتبار التكاليف الثابتة للبحث - سيرمز لها بالرمز (c) وان (c_1) ترمز لتكاليف الوحده من الطبقة الاولى وان (c_2) ترمز لتكاليف الوحده من الطبقة الثانية وهكذا اى ان التكاليف الكلية تساوى .

$$C = C_1 n_1 + C_2 n_2 + C_3 n_3 + \cdots\cdots + C_L n_L$$

$$= \sum_{h=1}^{L} C_h n_h$$

ونحصل على التوزيع الأمثل للعينه على الطبقات المختلفة والذى يعطى الحد الاقصى للمعلومات اللازمة لكل وحده من النقود بتطبيق القاعدة الآتية :

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

وفى حالة ما اذا كانت التكاليف الكلية للبحث محددة فيكون الحجم الكلى للعينه (n) مساويا

$$n = \frac{C}{\sum N_h S_h \sqrt{C_h}} \sum \frac{N_h S_h}{\sqrt{C_h}}$$

ويكون التباين الأمثل $\sigma^2(opt)$ فى حالة استخدام الحجم الامثل للعينه مأخوذا فى الاعتبار تكاليف البحث ،

$$\sigma^2(opt) = \frac{1}{nN^2} \sum N_h S_h \sqrt{C_h} \sum \left(\frac{N_h S_h}{\sqrt{C_h}}\right) - \frac{\sum N_h S_h^2}{N^2}$$

ويجب ادخال التعديلات اللازمة - والسابق الاشارة اليها - فى حالة ما اذا كان الحجم الامثل للعينه (n_h) من الطبقة (h) اكبر من حجم الطبقة نفسها (N_h) .

مثال لنفرض ان تكاليف الحصول على المعلومات من كل من منتجى اللبن ـــــــى الجدول رقم (١٠) السابق هى كالآتى :

قرش واحد لكل استمارة بحث من منتجى اللبن فى الطبقة الاولى ،

ربعة قروش " " " " " " " " الثانيه ،

تسعة قروش " " " " " " " " الثالثه .

احسب :

١) تكاليف البحث الكليه اذا كان حجم العينه المسحوبه من كل من الطبقات الثلاث هو: $n_1 = 244$, $n_2 = 288$, $n_3 = 468$ اى ان حجم العينه الكلى $n = 1000$.

٢) باستخدام التكاليف الكلية المتحصل عليها من الخطوه الاولى احسب الحجم الامثـــل للعينه الكليه ، مع توزيعها على الطبقات الثلاثه مع الاخذ فى الاعتبار تكاليـــف الحصول على المعلومات من منتجى اللبن فى كل طبقه .

٣) قارن بين التباين الامثل(σ^2_{opt}) لطريقتى توزيع الحجم الامثل للعينه (اى بـــدون اخذ التكاليف فى الاعتبار ، وبأخذها فى الاعتبار) .

الحـل :

١ ـ $c_1 = 1$, $c_2 = 4$, $c_3 = 9$ وحيث ان $n_1 = 244$, $n_2 = 288$, $n_3 = 468$

فتكون التكاليف الكلية للبحث مساويه

$$C = \sum_{h=1}^{3} c_h n_h = (244 \times 1) + (288 \times 4) + (468 \times 9) = 5608$$

٢ ـ حيث ان

$$n = \frac{C}{N_h S_h \sqrt{c_h}} \quad \frac{N_h S_h}{\sqrt{c_h}}$$

$$n = \frac{C}{\sum N_h S_h \sqrt{c_h}} \sum \frac{N_h S_h}{\sqrt{c_h}}$$

Strata	N_h	S_h	$N_h S_h$	$\sqrt{c_h}$	$N_h S_h \sqrt{c_h}$	$\frac{N_h S_h}{\sqrt{c_h}}$
1	538	9000	4842000	1	4842000	4842000
2	4756	1200	5707200	2	11414400	2853600
3	30964	300	9289200	3	27867600	3096400
Total	-	-	19836400	-	44124000	10792000

$$\sum_{h=2}^{3} \frac{N_h S_h}{\sqrt{c_h}} = \frac{N_2 S_2}{\sqrt{c_2}} + \frac{N_3 S_3}{\sqrt{c_3}} = 2853600 + 3096400 = 5950000.$$

ويكون الحجم الامثل للعينه (n) الكليه من كل الطبقات مساويا

$$n = \frac{5608}{44124000} \quad (10792000) = 1372$$

ويكون حجم العينه من الطبقه الاولى(n_1) مساويا

$$n_1 = \frac{\frac{N_1 S_1}{\sqrt{c_1}}}{\sum \frac{N_h S_h}{\sqrt{c_h}}} \; n$$

$$= \frac{4842000}{10792000} \times 1372 = 616$$

ولكن نلاحظ ان $n_1 = 616$ وهذه اكبر من $N_h = 538$ حجم الطبقة الاولى الكلى . لذا فاننا - حسب القاعدة - نأخذ (n_1) حجم العينه من الطبقه الاولى مساويا لحجم الطبقة الاولى اى ان $n_1 = 538$. ويقسم حجم العينه المتبقى والمساوى 834 منتجا على الطبقتين الثانية والثالثة بنفس القاعدة فقط بعد استبعاد القيم الخاصة بالطبقة الاولى . وبذا تكون

$$n_2 = \frac{2853600}{5950000} \; 834 = 400 \; ,$$

$$n_3 = \frac{3096400}{5950000} \; 834 = 434 \; .$$

فيكون التوزيع الامثل للعينه الكليه على الطبقات الثلاث كالآتى :

$n_1 = 538$ منتجا

$n_2 = 400$ منتجا

$n_3 = 434$ منتجا

$n = 1372$ (العينه الكليه) منتجا

٣ - أ - ويحسب التباين الامثل لتوزيع حجم العينه ($n = 1000$) اى بدون اخـــذ التكاليف فى اعتبارنا بالطريقة الآتيه :

$$\sigma^{2(opt)} = \frac{1}{N^2}\left[\frac{\left(\sum N_h S_h\right)^2}{n} - \sum N_h S_h^2\right]$$

$$= \frac{1}{(36258)^2}\left[\frac{(19838408)^2}{1000} - 532131 \times 10^9\right]$$

$$= 258.89$$

ب - ويحسب التباين الامثل لتوزيع حجم العينه ($n = 1372$) مع اخذ تكاليـــف البحث فى اعتبارنا بالطريقة الآتيـــة :

$$\sigma^{2(opt)} = \frac{1}{n N^2}\sum N_h S_h \sqrt{c_h}\sum\left(\frac{N_h S_h}{\sqrt{c_h}}\right) - \frac{\sum N_h S_h^2}{N^2}$$

ويجب فى هذه الحاله اسقاط كل القيم الخاصة بالطبقة الاولى مع التنبيه بان:

$N = 36258$ وهذا يساوى الحجم الكلى للطبقات الثلاث ،

$n = 834$ وهذا يساوى $n_2 + n_3$ فقط حيث اننا اسقطنا حجم (n_1) من حسابنا .

Strata	N_h	$N_h S_h \sqrt{c_h}$	$\frac{N_h S_h}{\sqrt{c_h}}$	$N_h S_h^2$
2	4746	11414400	2853600	6848640000
3	30964	27867600	3096400	278676000000
Total	-	39282000	5950000	285524640000

$$6^2(\text{opt}) = \frac{1}{834(36258)^2}(39282000)(5950000) - \frac{285524640000}{(36258)^2}$$

$$= 205.85$$

References

1- Ezekiel, Mordecai, Methods of Correlation Analysis. John Wiley and Sons, Inc., New York, 1930.

2- Li, Jerome C.R., Introduction to Statistical Inference. Edwards Brothers, Inc., Ann Arbor , Michigan, 1961.

3- Madow , W.G. , M.H.Hansen , and W.N. Hurwity , Sample Survey Methods and Theory. Volume I, John Wiley and Sons, Inc., New York, Second Printing 1956.

4- Snedecor, G.W., Statistical Methods. The Iowa State College Press , Ames , Iowa , 1959.

5- Zarkovich, S. S., Sampling Methods and Censuses. Volume I, F.A.O. of the U.N. , 1961.