وزارة التعليم العالي والبحث العلمي

الجامعة المستنصرية

كلية الإدارة والاقتصاد

مناهج واخلاقيات البحث العلمي

السنة الدراسية 2023-2022

القسم الإحصاء

المرحلة الرابعة

الفصل الدراسي الأول

الدراسة الصباحية والمسائية

استاذ المادة ليث فاضل سيد حسين

الفصل الثاني: جمع وتصنيف البيانات والمعلومات

تعد البيانات والمعلومات الجزء الأساسي لبناء البحث العلمي ، وعلى الباحث تحليل تلك البيانات لانجاز بحثه.أن عملية جمع البيانات تتم بأكثر من طريقة منها الاستبيان ، المراسلة فضلاً عن دراسة المجتمع الاحصائي وهو المجتمع الذي يؤطر مشكلة البحث وتتم من خلاله سحب عينة تتصف بأنها تمثل ذلك المجتمع خير تمثيل الأمر الذي يستوجب دراسة العينات وطرائق المعاينة الاحتمالية وغير الاحتمالية.

البيانات والمعلومات النظرية:

يتم الحصول على البيانات والمعلومات النظرية من خلال الاطلاع على الكتب والرسائل والاطاريح والدوريات والمجلات العلمية المتعلقة بالظاهرة التي يروم الباحث دراستها مستفيداً من المكاتب او مواقع الانترنت .

البيانات والمعلومات العملية والميدانية:

يتم الحصول على البيانات والمعلومات العملية والميدانية من خلال مصادر متعددة سبق وان تناولت مجتمع الدراسة لأي سبب من الاسباب ، وكذلك زيارة مجتمع الدراسة ميدانياً للحصول على المعلومات مباشرةً.

تحديد مصادر البيانات والمعلومات:

بعد تحديد عنوان البحث ومنهجيته هي العودة الى الكتب والمصادر النظرية لجمع المواد اللازمة لانجاز البحث ، إذ يجب تحديد المصادر المهمة من خلال مراجعة البحوث والدراسات والكتب التي تقترب بشكل او باخر من موضوع البحث ، وملاحظة نوع المصادر المستعملة في تلك الدراسات والكتب ، واستشارة أهل الخبرة والاختصاص والمشرف في تحديد المصادر المفيدة للبحث فضلاً عن ذلك مسؤول المكتبة.

في البداية وبعد الحصول على المصادر على الباحث قراءة هذه المصادر قراءة اولية عادة ما تكون سريعة يطلع من خلالها على العناوين للكتب والرسائل والاطاريح والبحوث ، والمحتوى والمقدمة ومنهج البحث والهدف من الدراسة ، وبعد المطالعة من قبل الباحث يحدد اهم المصادر ليقوم بقراءتها بشكل تفصيلي.

تدوين المعلومات:

يقصد بالتدوين نقل النصوص من مصادرها المختلفة الى اوراق خاصة بالبحث كخطوة اولى ، ليقوم بعد ذلك بكتابة بحثه ، ولغرض تدوين المعلومات نستعمل النظامين التاليين:

1- نظام الملف او الاضبارة او الدوسيه.
2- نظام البطاقات.

النموذج في ادناه نموذج جمع نص بطريقة الدوسيه:

الباب: الفصل: المبحث: الموضوع:		
النص المقتبس (المنقول)		
المؤلف: عنوان المصدر: الجزء: الطبعة: الناشر مكان وزمان الطبع: الصفحات:		

نظام الاضبارة للاقتباس

وهي اضبارة خاصة مقسمة من الداخل باحجام تتناسب مع الحجم المتوقع جمعه من كل فصل ومبحث وموضوع ، مما يتيح لهُ التطوير المستمر والاضافة ، ويبدأ بجمع النصوص داخل هذه الصفحات مشيراً الى مصادرها ، وهذا النظام غير معمول بهِ كثيراً.

اما نظام البطاقات فهي عبارة عن مجموعة من الأوراق السميكة المتساوية الحجم ، مساحتها 12سم × 7.5سم قد تزيد او تنقص احياناً حسب رأي الباحث ، وترقم هذه البطاقات حفاظاً على تسلسلها ، وينبغي ان تضم البطاقة الواحدة سوى اقتباس واحد فقط ، وينبغي ان ترتب البطاقات حسب عناوين او فصول او مباحث الدراسة ، لا حسب المصادر المستعملة ، ولأجل عدم إضاعة الوقت والجهد والفكرة ينبغي ان يخصص ظرف او درج لكل فصل او مبحث توضع بداخلهِ البطاقات ذات العلاقة ، والاتي نموذج لبطاقة اقتباس:

الموضوع: عنوان المصدر: الجزء: الطبعة: المؤلف: مكان وزمان الطبع: الصفحات: الناشر:		
النص المقتبس (المنقول)		

نظام البطاقة للاقتباس

ان المدرسة الحديثة هي التي تستخدم نظام البطاقات ، ولا ينصح الباحثون الا بها ، فهي اكثر سهولة وضبطاً ودقة ، وايسر في التصنيف وأنفع لدى استخدام المعلومات ويسهل معها التقديم والتأخير والتعديل.

ملاحظة: مهما كان النظام المستخدم في الاقتباس ، فينبغي وضع النص المنقول بين علامتي اقتباس " " ، في حالة نقل نص دون تغيير ، اما اذا كانت الفكرة قد تم استيعابها وتفريغها بأسلوب الباحث فلا نضع علامتي اقتباس ، وفي الحالتين يشار الى المصدر المقتبس منه.

كتابة المسودة (الكتابة الاولية):

بعد الفراغ من تجميع البطاقات حسب فصولها ومباحثها ، يتم نقل ما جمع في البطاقات الى أوراق البحث الأولية (المسودة) مرتباً حسب ما يقتضيه السياق تاركاً مجالاً بين الاسطر لاجل إضافة معلومات قد تستجد مستقبلاً ، وليس المقصود بالكتابة هنا النقل المباشر لما دون في البطاقات ، اذ يقتضي البحث التعديل والحذف والشرح والتعليق والمناقشة وحسب الحاجة لكل حالة ، ويجب ان تجمع البطاقات المترابطة عند نقلها في فقرات قائمة بذاتها تظهر مستقلة عند الكتابة ، والمقصود بالفقرة هنا مجموعة من الجمل بينها اتصال وثيق تهدف الى ابراز فكرة معينة.

أساليب جمع البيانات (Data Collection Methods):

أن أي بحث علمي يستند في تحليلهِ الى الطريقة الاحصائية يحتاج الى بيانات ومعلومات حول موضوع البحث قيد الدراسة ، ويمكن الحصول على هذهِ البيانات والمعلومات من مصدرين هما:

1) المصادر التأريخية (Historical Sources):

وهي البيانات المحفوظة والمتجمعة لدى أجهزة ومؤسسات و دوائر الدولة المختلفة نتيجة لمسوحات قامت بها هذهِ الجهات أو هيئات معينة لاغراض خاصة بها أو تجمعت لديها بحكم وظائفها الادارية والفنية. مثال على ذلك، البيانات المتجمعة عن تعدادات السكان في العراق ، ، احصاءات الطلبة المتخرجين من الجامعات العراقية وغيرها من الأمثلة الاخرى.

وتقسم المصادر التاريخية الى نوعين:

النوع الأول/ المصادر الأولية : وهي البيانات التي يقوم بتفريغها وتبويبها نفس الجهة التي قامت بجمعها ، كالنشرات التابعة لوزارة التخطيط وما ينشره الباحثون.

النوع الثاني/ المصادر الثانوية : وهي البيانات التي يقوم بتفريغها وتبويبها جهة أخرى غير التي قامت بجمع البيانات الأولية ، كالنشرات التابعة للأمم المتحدة والجداول الإحصائية المتوفرة في كتب ودراسات الباحثين.

ومهما كانت نوع المصادر يجب التأكد من مقدار الثقة بالجهات التي قامت بجمع البيانات والتأكد من دقتها ومدى صلاحيتها حالياً كونها جمعت في فترات سابقة.

2) المصادر الميدانية (Field Sources):

في حالة تعذر الحصول على البيانات من المصادر التأريخية نلجأ الى الميدان للحصول على تلك البيانات اي الحصول على البيانات من مصادرها الاصلية ، عن طريق الاتصال بالافراد الموجودة لديهم البيانات ، وان تعذر ذلك يقوم الباحث بجمعها بنفسه عن طريق الملاحظة والمشاهجة وهذا يتطلب مستوى عالي من الخبرة والتدريب ، ولاجل الحصول على البيانات من مصادرها الميدانية يتسخدم الباحث احدى الطرائق الاتية:

المقابلة الشخصية: وتمتاز هذه الطريقة بدقة البيانات نتيجة الاتصال المباشر بين الباحث وافراد مجتمع البحث ، وتجاوز عيب الغموض في بعض الأسئلة ، ولكن يعاب على هذه الطريقة **ان الباحث قد يؤثر على الافراد باتجاه إجابات معينة** كما ان هذه الطريقة **تتطلب الكثير من الجهد والوقت**.

الاتصال الهاتفي: تمتاز هذه الطريقة بسرعة جمع المعلومات وقلة التكاليف ، ولكن يعاب عليها **اقتصارها على الافراد الذين يملكون هاتف مما قد يضيع فرصة الحصول على اجابات من معظم افراد مجتمع البحث** ، اضافة الى **تخوف بعض الافراد من اعطاء الاجابات لشخص مجهول**.

المراسلة: وفقاً لهذه الطريقة يتم ارسال استمارة الاسئلة الى الافراد بالبريد ، وميزات هذه الطريقة هو تلافي التأثير الشخصي للباحث وكذلك توفر الوقت الكافي امام الافراد للاجابة عن الاسئلة بهدوء ، ولكن يعاب عليها عدم حصول الافراد على توضيح للاسئلة الغامضة وكذلك يعاب عليها انهُ قد يتولى الاجابة عن استمارة الاسئلة شخص من خارج المجتمع قيد البحث.

وهنالك طرائق اخرى كالتسجيل او طريقة المشاهدة والتجربة والاتي توضيح لكل منهما:

طريقة التسجيل: وهي الحصول على البيانات من الافراد بموجب تشريع قانوني تصدره الدولة ، مثل تسجيل الولادات والوفيات والزواج ... الخ.

طريقة المشاهدة والتجربة: وفيها يتم الحصول على البيانات التي تنتج عن التجارب التي تجري في الكيمياء والفيزياء والطب وحركة المرور ... الخ ، حيث يتم الحصول على هذه البيانات بمساعدة الوسائل الاحصائية.

ان اختيار الباحث لطريقة واحدة او اكثر من الطرائق السابقة الذكر يعتمد على الاسلوب المتبع في جمع البيانات وكذلك قد يستخدم اكثر من طريقة لاجل ضمان الحصول على بيانات اكثر دقة وتمثيلاً لمجتمع البحث.

بغية الحصول على البيانات عن ظاهرة أو مجموعة ظواهر التي يتطلبها البحث او الدراسة هنالك اسلوبان يمكن من خلالهما جمع البيانات وهي **أسلوب التسجيل الشامل** الذي يوفر المعلومات عن كافة مفردات المجتمع الاحصائي ، و **أسلوب التسجيل عن طريق العينات**.

اسلوب التسجيل الشامل (Census):

بموجب هذا الأسلوب يتم جمع البيانات عن كافة مفردات المجتمع الإحصائي هذا يعني بأن أسلوب التسجيل الشامل يمتاز بما يلي:

1) **يجب أن يكون المجتمع الاحصائي محدداً** ، أي يمكن ملاحظة كل مفردة فيهِ. مثال على ذلك، التعداد العام للسكان.
2) يعد أفضل أسلوب في جمع البيانات كونهُ **يجهز الباحث ببيانات كاملة عن كافة مفردات المجتمع الإحصائي (مجتمع الدراسة)**.

الا أن اسلوب التسجيل الشامل يحتاج الى وقت وجهد موارد مادية ، وبشرية كبيرة في انجاز مهمة جمع البيانات ، فضلاً عن احتمال الوقوع في اخطاء نتيجة التعامل مع مفردات المجتمع الاحصائي بشكل كامل.

اسلوب العينات (Samples):

طالما ان العينة هي مجموعة جزئية من المجتمع الاحصائي هذا يعني بأن أسلوب التسجيل عن طريق العينات يمتاز بما يلي:

1) لا نحتاج الى وقت و جهد موارد مادية ، وبشرية كبيرة.
2) يمكن استخدامهُ في حالة المجتمعات الاحصائية الغير محدودة.
3) إمكانية الحصول على بيانات تتضمن صفات أكثر عن المفردة الاحصائية (معلومات أكثر عن الظاهرة قيد الدراسة).
4) إمكانية اختبار دقة النتائج المستحصل عليها.

وسيتم تعريف كل من المجتمع والعينة بالتفصيل لاحقاً.

1) الاستبيان (Questionnaire):

تعريف: استمارة الاستبيان: عبارة عن استمارة يتم من خلالها جمع البيانات من مفردات مجتمع الدراسة او من عينة البحث من خلال مواجهة الباحث الشخصية للمفردة الاحصائية او عن طريق المراسلة.

انواع استمارة الاستبيان:

النوع الاول/ استمارة الكشف: وهنا يقوم الباحث بطرح الاسئلة الموجودة في الاستمارة على افراد المجتمع ويثبت الاجابة بنفسه ، ومن سلبيات هذه الاستمارة التحيز والتلاعب بالاجابات باتجاه مصلحة الباحث الخاصة.

النوع الثاني/ الاستمارة الاعتيادية: وهي تلك الاستمارة التي يقوم الباحث بتوزيعها على افراد المجتمع ليقوموا بتثبيت الاجابات بانفسهم بصورة مباشرة ، وهذه الاستمارة تتجاوز عيوب الاستمارة السابقة.

شروط تصميم استمارة الاستبيان : ويجب مراعاة النقاط التالية عند تصميم الاستمارة:

1) اعداد مقدمة ايضاحية تكتب في بداية استمارة الاستبيان.
2) ان تكون فقرات استمارة الاستبيان متسلسلة وغير مبعثرة بحيث ان كل صنف منها يحقق غرض معين.
3) ان تكون الاسئلة متوسطة العدد.
4) ان تكون الاسئلة واضحة المعنى.
5) ان تهيء الاسئلة بالشكل الذي يجب ان تكون اجابة الفرد عليها محددة وقصيرة.
6) ان يؤخذ بنظر الاعتبار ظروف تفريغ وتصنيف وتبويب وترميز الاجابات.
7) عدم طرح الاسئلة باسلوب متحيز باتجاه اجابات معينة.

طرائق اختبار صحة استمارة الاستبيان:

بعد الانتهاء من وضع الاسئلة وفق الشروط التي تطرقنا لها ، وقبل توزيع الاستمارة على افراد المجتمع ، ينبغي على الباحث قراءة هذه الاستمارة لمرات عديدة لاجل الحذف والاضافة والتعديل عليها ، ومن ثم اختبار مدى صحة هذه الاستمارة من خلال:

1) عرض الاستمارة على اخصائيين في اللغة العربية لاجل تصحيحها لغوياً.
2) عرض الاستمارة على اخصائيين بمشكلة البحث بما فيهم المشرف على البحث.
3) عرض الاستمارة على اخصائيين في الاحصاء والرياضيات لاجل اعادة ترتيب الاسئلة من اجل عملية التصنيف والترتيب.
4) توزيع الاستمارة على عينة صغيرة جداً من افراد المجتمع الذين تعنيهم المشكلة لاخذ ارائهم ومقترحاتهم.

تصنيف وتبويب البيانات

ان البيانات المستحصل عليها بخصوص الظاهرة المعنية تسمى بالبيانات الخام او الاولية وتلك البيانات تكون غير منظمة ويصعب على الباحث تكوين فكرة عن الظاهرة المدروسة كذلك يصعب الأعتماد عليها لاغراض التحليل الاحصائي ، لذا فإن الخطوة الهامة التي تلي عملية جمع البيانات هي **عملية مراجعة وتصنيف البيانات وتبويبها.**

1) **مراجعة البيانات:** بعد اتمام عملية جمع البيانات يجب مراجعة تلك البيانات وتدقيقها لغرض التأكد من مطابقتها وتكاملها لمتطلبات الدراسة.
2) **تصنيف البيانات:** بعد التأكد من تكامل و وضوح ودقة البيانات التي تم الحصول عليها ، نبدأ بعملية تصنيف على أساس الظواهر التي جمعت عنها البيانات حيث يتم فرز البيانات على هيئة مجموعة فقد يكون التصنيف على أساس ظاهرة العمر او الجنس او المهنة او الوزن او الطول او محل الاقامة وغيرها.
3) **تبويب البيانات:** بعد اتمام عملية تصنيف البيانات نبدأ بعملية تبويب البيانات ويقصد بتبويب البيانات تفريغ البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعنية يعود الى مستوى معين لتلك الظاهرة.

والهدف من عملية التبويب هو ابراز البيانات وتوضيحها في اضيق حيز ممكن كي نتمكن من تكوين فكرة عنها.

وهناك اربعة اشكال للتبويب هي:

1) **التبويب الزمني:** عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها يعود لوحدة زمنية معينة ، كاليوم او الاسبوع او الشهر او السنة.
2) **التبويب الجغرافي:** عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة جغرافية معينة ، او تقسيم اداري معين.
3) **التبويب الكمي:** عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة كمية معينة ، كوحدات الوزن ، الطول ، العمر ، المسافة ، ... الخ.
4) **التبويب على اساس صفة معينة:** عبارة عن عملية تجميع للبيانات المصنفة وترتيبها في جداول خاصة على اساس ان كل جمع منها يشترك بصفة معينة ، كالجنس، الحالة الاجتماعية، عنوان الوظيفة، القومية.

مجتمع البحث (Research Population)

بعد الانتهاء من تحديد مشكلة البحث يجب معرفة مجتمع الدراسة الذي يمثل مصدر البيانات ، فيمكن تعريف مجتمع البحث ويسمى ايضاً بالمجتمع الاحصائي كما يلي:

تعريف: المجتمع الإحصائي (**Statistical Population**): هو جميع مفردات الظاهرة موضوع البحث او الدراسة وتلك المفردات يفترض ان تشترك بصفة او صفات معينة وقد تكون تلك المفردات كائن حي او أي شيء آخر ، والمجتمع الاحصائي قد يكون محدد او غير محدد ، على سبيل الأمثلة (الافراد ، بحيرة اسماك ، ... الخ) وهناك ارتباط وثيق ومباشر بين مشكلة البحث ومجتمع البحث.

مثال(1): العلاقة بين معدلات التخرج لطلبة كلية الإدارة والاقتصاد / الجامعة المستنصرية والعبئ الدراسي.

المجتمع الاحصائي : طلبة كلية الإدارة والاقتصاد / الجامعة المستنصرية.

مثال(2): القصص الخاصة بالأطفال وماتقدمه من معلومات وتسلية.

المجتمع الاحصائي: القصص الخاصة بالأطفال.

هناك حالات معينة يستوجب فيها دراسة المجتمع الاحصائي بأكمله ، ولا يمكن الاكتفاء باختيار عينة منه.

1- عندما يتطلب الامر جمع البيانات من كل فرد من افراد المجتمع ، على سبيل المثال تعداد السكان.
2- عندما يكون المجتمع محدد وصغير أي تتراوح المفردات من 15 الى 25 مفردة مثلاً.

ان أنواع المجتمع الاحصائي تنقسم الى قسمين وكما يلي:

المجتمع المتجانس (Homogeneous Population**)**

هو المجتمع الذي يتميز بتماثل الخصائص لدى جميع مفرداته ، على سبيل المثال ، ان تكون الدراسة عن صعوبة مادة اللغة الإنكليزية للصف الثالث المتوسط من وجهة نظر الطلبة ، فنلاحظ ان المجتمع الاحصائي هو طلبة الصف الثالث المتوسط في عدد من المدارس الحكومية للبنين متجانس ، وهذا يعني ان جميع مفردات المجتمع قيد الدراسة تنطبق عليهم نفس الخصائص.

فالخصائص هنا :

1- جميع مفردات المجتمع الاحصائي ذكور.
2- جميع الطلبة يدرسون نفس المادة.
3- جميع الطلبة يدرسون في نفس المرحلة.
4- جميع المدارس التي اختيرت هي حكومية.

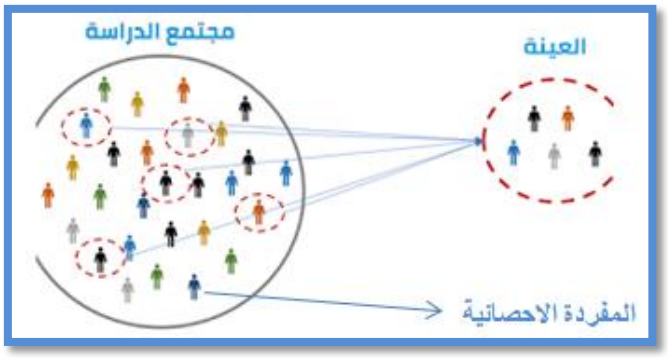
المجتمع غير المتجانس (Heterogeneous Population)

هو المجتمع الذي تتفاوت فيه الخصائص لدى جميع مفرداته ، على سبيل المثال ، ان تكون الدراسة سلوك الأشخاص الذين يرتادون مول النخيل من وجهة نظر المتسوقين ، فنلاحظ ان المجتمع الاحصائي هو في حالة متباينة (غير متجانس) ، وهذا يعني ان جميع مفردات المجتمع قيد الدراسة لا تنطبق عليهم نفس الخصائص.

فالخصائص هنا :

1- مفردات المجتمع الاحصائي ذكور واناث.
2- اختلاف في اعمار الاشخاص.
3- اختلاف الثقافة بين الاشخاص.
4- اختلاف في كمية الاستهلاك.

تعريف: العينة (**Sample**): عبارة عن مجموعة جزئية من المجتمع الاحصائي ، بحيث يتم اختيارها بطريقة تضمن تمثيلها لهُ بشكل دقيق، ان عملية الاختيار تتم وفقاً لأساليب المعاينة الاحصائية ، لكي نضمن تمثيلها للمجتمع الاحصائي لجميع مفرداتهِ أفضل تمثيل ، والشكل رقم (1) يبين أهم المفاهيم التي تم تعريفها.



الشكل رقم (1): يوضح مفاهيم المجتمع الاحصائي ، المفردة الإحصائية ، العينة

تعريف: المفردة الاحصائية (**Statistical Unit**): اصغر وحدة في المجتمع الاحصائي.

مثال(1): عند اجراء شخص ما لتحليل الدم لمعرفة صنف دمه ، فان المجتمع الاحصائي هو كامل الدم ، الا ان ما يحدث هو اخذ جزء من هذا الدم وتسمى عينة ليجري عليها عملية التحليل ، ثم يتم تمثيل المجتمع من خلال هذه العينة أي ان نتيجة التحليل لهذا الجزء اذا كانت (A+) مثلاً ، فيقال ان هذا الشخص صنف دمه بشكل كامل هو (A+).

مثال(2): عند اجراء شخص ما لفحص ماء بئر معين ، فان المجتمع الاحصائي هو كامل البئر ، الا ان ما يحدث هو اخذ جزء من هذا الماء وتسمى عينة ليجري عليها عملية الفحص ، ثم يتم تمثيل المجتمع من خلال هذه العينة أي ان نتيجة الفحص لهذا الجزء اذا كان (القدح المملوء بالماء عذب) مثلاً ، فيقال ان هذا البئر ذو ماء عذب.

أي ان العينة في المثالين أعلاه مثلت المجتمع ، وهكذا فان العينة يجب ان تكون ممثلة للمجتمع من حيث توفر الخصائص ، حيث يمكن في هذه الحالة الاكتفاء فقط باختبار العينة وتعميم النتائج على كامل المجتمع.

هناك شروط يجب ان تتوفر في العينة وهي:

1- ان يكون حجم العينة ملائماً بمعنى انهُ يسمح باحتواء كافة الخصائص المتوفرة في المجتمع الأصلي ، فعلى سبيل المثال اذا كان المجتمع يتكون من (10000) مفردة ، فان اختيار عينة بحجم (10) مفردات غير مقبول.
2- ان يتم اختيار المفردات التي تتكون منها العينة وفقاً لنظام معين بما يضمن ان تحتوي على الخصائص الموجودة في المجتمع.

<u>تحديد حجم عينة البحث:</u> يتوقف حجم العينة على مجموعة من العوامل أهمها ما يلي:

1- حجم المجتمع الأصلي.
2- المجتمع الأصلي ومدى تجانسه (متجانس ام غير متجانس).
3- مستوى او درجة الدقة المطلوبة.
4- الإمكانات المتاحة امام الباحث.

ان عملية تحديد واختيار العينة تتمثل في المعادلات الرياضية المعتمدة على نظرية الاحتمالات ، فيجب تحديد المجتمع بدقة من خلال الحصول على البيانات الدقيقة الصادرة من جهات موثوقة سواء كان ذلك في القطاع الحكومي او القطاع الخاص وحسب مشكلة البحث ، ثم نحدد الخصائص المتعلقة بمشكلة البحث والتي تتوفر في المجتمع.

على سبيل المثال ، لو أراد باحث تحديد أي الجنسين لديه رغبة في الانضمام الى دورات تدريبية قبل الالتحاق بالجامعة.

يجب معرفة توزيع المجتمع وفقاً لخاصية الجنس (ذكور، واناث) ، ولنفترض ان النسبة للذكور في هذا المجتمع كانت 60% ، هذا يعني ان نسبة الاناث ما تبقى من ذلك المجتمع وهي (40%).

عند اختيار عينة من هذا المجتمع التي تتكون من الذكور والاناث وفقاً للنسب المحددة بمعنى ان لكل 6 مفردات من الرجال يقابله 4 مفردات من الاناث لهذا المجتمع.

ان احتمال اختيار أي مفردة هو نسبة حجم العينة (n) على حجم المجتمع (N) هذا يعني:

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

حيث ان: P : احتمال اختيار أي مفردة ، n : حجم العينة ، N : حجم المجتمع.

مثال: ماهو احتمال اختيار مفردة إحصائية من مجتمع حجمة (5000) ، اذا أراد باحث اختيار عينة بحجم (500).

الحل:

$$P = \frac{n}{N} = \frac{500}{5000} = 0.1$$

- ⇐ ان تحديد حجم العينة المطلوب والمناسب للدراسة مهم جدا للحصول على نتائج يمكن الاعتماد عليها ، على هذا الأساس يتضح ان العلاقة عكسية بين حجم العينة والخطأ ، بمعنى ان ازدياد حجم العينة يؤدي الى تقليل مقدار الخطأ والعكس صحيح ، ان ازدياد حجم العينة غير المدروس وفق أسس صحيحة لا يعني بالضرورة تقليل الخطأ ، فعلى الباحث الموازنة ما بين ازدياد حجم العينة وما بين تقليل الخطأ بالشكل الذي ينتج عنه تحديد حجم العينة المناسب.
- ⇐ ان النقطة الأساسية التي أكد عليها الباحثين في مختلف مجالات البحث العلمي هو اختيار حجم عينة كبير ، فيعد حجم العينة (30) مفردة او اكثر كبير ، وعدا ذلك فان حجم العينة يكون صغير.
- ⇐ ان حجم العينة وبصورة عامة يعتمد على نوعية الدراسة والتحليل والغرض منها ، فمثلاً في البحوث التي يكون فيها السيطرة النوعية عالياً فان حجوم عينات صغيرة يكون مناسباً.
- ⇐ حجم العينة الذي يتراوح ما بين (30) مفردة الى (500) مفردة يعتبر مناسباً ولجميع أنواع الدراسات والبحوث.
- ⇐ ان حجم العينة في موضوع تحليل الانحدار يفضل ان يكون (10) اضعاف عدد متغيرات الدراسة، على سبيل المثال اذا كانت الدراسة تحتوي على 5 متغيرات فيفضل ان لا يقل حجم العينة عن (50) مفردة.

يعتبر تحديد حجم المجتمع قيد الدراسة الذي يتضمن المفردات الإحصائية والتي ينبغي ان تكون (تشكل) العينة من المسائل الأساسية في عملية تصميم العينة ، وذلك لتجنب اخذ عينة بحجم صغير يكون تقديرها للمجتمع غير دقيق وبالتالي غير مفيد.

ان تحديد حجم العينة يعتمد على الدقة المستهدفة ، فإذا كان لدينا مستوى معنوية $\boxed{\alpha = 0.05}$ (وتمثل مقدار الخطأ)، والتي يعبر عنها بحجم الخطأ المسموح بهِ في إيجاد التقديرات فهذا يعني ان درجة الثقة ستكون $\boxed{1 - \alpha = 0.95}$ (وتمثل مقدار الثقة)، ان درجة الثقة هي عبارة عن احتمال يقع ضمنه تقدير معلومية المجتمع.

⇐ يمكن إيجاد حدود الثقة لمتوسط المجتمع الاحصائي الذي يتبع التوزيع الطبيعي القياسي بالصيغة التالية:

$$\mathbf{L.L < \mu < U.L}$$

$$\mathbf{L.L = \bar{x} - Z_{\alpha} * S.E}$$

$$\mathbf{U.L = \bar{x} + Z_{\alpha} * S.E}$$

$$\mathbf{S.E = \frac{s}{\sqrt{n}}}$$

$$\mathbf{S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}}$$

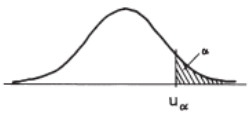
$$\mathbf{S_{xx} = \sum(X_i - \bar{X})^2 = \sum X_i^2 - n\bar{X}^2}$$

اما طريقة إيجاد القيمة الجدولية $\mathbf{Z_{\alpha}}$ ، تعتمد على مستوى المعنوية α ، والجدول الآتي:

Table 4

PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

The table gives the 100α percentage points, u_α, of a standardised Normal distribution where $\alpha = \frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-u^2/2}\, du$. Thus u_α is the value of a standardised Normal variate which has probability α of being exceeded.



α	u_α	α	u_α	α	u_α	α	u_α	α	u_α	α	u_α
.50	0.0000	.050	1.6449	.030	1.8808	.020	2.0537	.010	2.3263	.050	1.6449
.45	0.1257	.048	1.6646	.029	1.8957	.019	2.0749	.009	2.3656	.010	2.3263
.40	0.2533	.046	1.6849	.028	1.9110	.018	2.0969	.008	2.4089	.001	3.0902
.35	0.3853	.044	1.7060	.027	1.9268	.017	2.1201	.007	2.4573	.0001	3.7190
.30	0.5244	.042	1.7279	.026	1.9431	.016	2.1444	.006	2.5121	.00001	4.2649
.25	0.6745	.040	1.7507	.025	1.9600	.015	2.1701	.005	2.5758	.025	1.9600
.20	0.8416	.038	1.7744	.024	1.9774	.014	2.1973	.004	2.6521	.005	2.5758
.15	1.0364	.036	1.7991	.023	1.9954	.013	2.2262	.003	2.7478	.0005	3.2905
.10	1.2816	.034	1.8250	.022	2.0141	.012	2.2571	.002	2.8782	.00005	3.8906
.05	1.6449	.032	1.8522	.021	2.0335	.011	2.2904	.001	3.0902	.000005	4.4172

مثال: جد حدود الثقة لمتوسط المجتمع الاحصائي μ ، إذا علمت ان المجتمع يتبع التوزيع الطبيعي القياسي N(0,1) ، وان حجم العينة n=100 ، والوسط = 24.1 ، والتباين = 2.56 ، وبمستوى معنوية وفق الحالات التالية:

$\alpha = 0.05 \Leftarrow$

$\alpha = 0.01 \Leftarrow$

$\alpha = 0.1 \Leftarrow$

مستعيناً بالجداول الإحصائية الخاصة بالتوزيع الطبيعي القياسي.

الحل (1):

$$\bar{x} - Z_{\alpha} * \frac{s}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha} * \frac{s}{\sqrt{n}}$$

$$\alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025 \Longrightarrow \boxed{Z_{\alpha} = 1.96} \qquad (\text{from table}(4))$$

$$24.1 - (1.96) * \frac{\sqrt{2.56}}{\sqrt{100}} < \mu < 24.1 + (1.96) * \frac{\sqrt{2.56}}{\sqrt{100}}$$

$$24.1 - (1.96) * \frac{1.6}{10} < \mu < 24.1 + (1.96) * \frac{1.6}{10}$$

$$24.1 - 0.3136 < \mu < 24.1 + 0.3136$$

$$\mathbf{23.7864 < \mu < 24.4136}$$

الحل (2):

$$\alpha = 0.01 \Longrightarrow \frac{\alpha}{2} = 0.005 \Longrightarrow \boxed{Z_{\alpha} = 2.5758} \qquad (\text{from table}(4))$$

$$24.1 - (2.5758) * \frac{\sqrt{2.56}}{\sqrt{100}} < \mu < 24.1 + (2.5758) * \frac{\sqrt{2.56}}{\sqrt{100}}$$

$$24.1 - (2.5758) * \frac{1.6}{10} < \mu < 24.1 + (2.5758) * \frac{1.6}{10}$$

$$24.1 - 0.4121 < \mu < 24.1 + 0.4121$$

$$\mathbf{23.6879 < \mu < 24.5121}$$

واجب: المطلب الثالث واجب على الطالب.

كلما كانت الفترة لحدود الثقة متقاربة كلما كان التقدير للمعلمة أفضل ، فعلى سبيل المثال لتقدير $\mu = 25$ (بشكل افتراضي وضعنا قيمة للمعلمة والتي عادة ما تكون مجهولة يتم تقديرها من بيانات العينة) فان التقدير بفترة من 20 الى 30 يكون بالطبع افضل من التقدير بفترة من 10 الى 40.

كلما كان الخطأ المعياري (S.E) أقل كلما زادت دقة التقدير.

كلما كان حجم العينة كبير كلما قل مقدار الخطأ.

ان متوسط الوسط الحسابي $\mu_{\bar{x}}$ يساوي متوسط المجتمع μ ، وللبرهنة على ذلك لنفترض انا لدينا مجتمع احصائي مكون من $(N = 3)$ مفردات هي $(x = 8,3,6)$ ، يراد سحب عينة مكونة من مفردتين (عدم تكرار المفردات في العينات) ، فيكون لدينا المعلومات عن المجتمع والعينات في الجدول الآتي:

المجتمع	**8, 3, 6**		
العينات	8,3	8,6	3,6

$$\mu = \frac{8+3+6}{3} = \frac{17}{3} = 5.667$$

$$\bar{x}_1 = \frac{8+3}{2} = \frac{11}{2} = 5.5$$

$$\bar{x}_2 = \frac{8+6}{2} = \frac{14}{2} = 7$$

$$\bar{x}_3 = \frac{3+6}{2} = \frac{9}{2} = 4.5$$

$$P(1) = P(2) = P(3) = \frac{1}{3}$$

$$\mu_{\bar{x}} = E(\bar{x}_i) = \sum_{i=1}^{k} \bar{x}_i * P(i)$$

$$\mu_{\bar{x}} = \frac{1}{3} * 5.5 + \frac{1}{3} * 7 + \frac{1}{3} * 4.5 = \frac{1}{3} * (5.5 + 7 + 4.5) = \frac{17}{3} = 5.667$$

$$\mu = \mu_{\bar{x}} = 5.667$$

مثال: اثبت ان متوسط الوسط الحسابي $\mu_{\bar{x}}$ يساوي متوسط المجتمع μ ، اذا علمت ان (التكرار في المفردات مسموح به) :

المجتمع	0, 3			
العينات	0,0	0,3	3,0	3,3
متوسط العينات	0	1.5		3
الاحتمال	0.25	0.5		0.25

$$\mu = \frac{0+3}{2} = 1.5$$

$$\mu_{\bar{x}} = 0 * 0.25 + 0.5 * 1.5 + 0.25 * 3 = 0 + 0.75 + 0.75 = 1.5$$

$$\mu = \mu_{\bar{x}} = 1.5$$

طريقة إيجاد حجم العينة n في حالة التوزيع الطبيعي بوسط مجهول μ ، وحجم المجتمع غير معلوم ، وتباين معلوم σ^2.

ان القيمة d عبارة عن القيمة المطلقة لانحراف الوسط الحسابي للعينة عن وسطها الحسابي للمجتمع ، أي ان:

$$d = |\bar{x} - \mu|$$

ومن صيغة التوزيع الطبيعي نجد ان :

$$P(|\bar{x} - \mu| \leq Z\sigma) = 1 - \alpha$$

نحصل على:

$$d = Z\sigma$$

حيث ان:

$$\sigma = \frac{S}{\sqrt{n}}$$

$$d = Z\frac{S}{\sqrt{n}}$$

بتربيع الطرفين نحصل على:

$$d^2 = \frac{Z^2 * S^2}{n}$$

$$\boxed{n = \frac{Z^2 * S^2}{d^2}}$$

حيث ان:

S^2: يمثل التباين للمجتمع يكون معلوم من خلال نتائج مسوحات سابقة ، او باجراء مسح تجريبي بتعبئة عدد من الاستمارات من المجتمع قيد الدراسة.

Z^2: مربع القيمة الجدولية للتوزيع الطبيعي القياسي.

d^2: مربع القيمة المطلقة لانحراف الوسط الحسابي للعينة عن وسطها الحسابي للمجتمع ، وتعطى بالسؤال.

مثال: جد حجم العينة ، اذا علمت ان :

$$\alpha = 0.05 \Leftarrow$$

$$S^2 = 90.3 \Leftarrow$$

$$d = 1.5 \Leftarrow$$

الحل:

$$n = \frac{Z^2 * S^2}{d^2}$$

$$\alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025 \Longrightarrow \boxed{Z_\alpha = 1.96} \qquad (\text{from table}(4))$$

$$n = \frac{(1.96)^2 * (90.3)}{(1.5)^2} = \frac{346.8965}{2.25} = 154.1762 \cong \boxed{154}$$

طريقة إيجاد حجم العينة n في حالة التوزيع الطبيعي بوسط مجهول μ وتباين معلوم σ^2 ، وحجم المجتمع معلوم.

$$n = \frac{N * Z^2 * S^2}{Nd^2 + Z^2 * S^2}$$

مثال: جد حجم العينة ، اذا علمت ان :

$$N = 6420 \Leftarrow$$

$$\alpha = 0.05 \Leftarrow$$

$$S^2 = 81 \Leftarrow$$

$$d = 2 \Leftarrow$$

الحل:

$$n = \frac{N * Z^2 * S^2}{Nd^2 + Z^2 * S^2}$$

$$\alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025 \Longrightarrow \boxed{Z_\alpha = 1.96} \qquad (\text{from table}(4))$$

$$n = \frac{(6420) * (1.96)^2 * (81)}{(6420)(2)^2 + (1.96)^2 * (81)} = \frac{1997708.832}{25680 + 311.1696}$$

$$n = \frac{1997708.832}{25991.1696} = 76.8611 \cong \boxed{77}$$

طريقة إيجاد حجم العينة n في حالة التوزيع ثنائي الحدين بوسط مجهول μ وتباين معلوم σ^2 ، وان وسط وتباين المجتمع عبارة عن نسبة (خاصية في المجتمع)

من الأمثلة على المجتمعات الإحصائية التي يعبر عنها بنسب ، نسبة المتزوجين ، نسبة وحدات الإنتاج الصالحة ، نسبة النجاح ... الخ ، ويتم الافتراض من ان توزيع هذه المجتمعات هو مقارب للتوزيع الطبيعي.

$$\boxed{n = \frac{Z^2 * p * q}{d^2}}$$

p : نسبة النجاح

$q = 1 - p$: نسبة الفشل48

مثال: جد حجم العينة ، اذا علمت ان :

$$p = 15\% \Leftarrow$$

$$\alpha = 0.05 \Leftarrow$$

$$d = 0.05 \Leftarrow$$

الحل:

$$\boxed{n = \frac{Z^2 * p * q}{d^2}}$$

$$p = 15\% = \frac{15}{100} = 0.15$$

$$q = 1 - p = 85\% = \frac{85}{100} = 0.85$$

$$\alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025 \Longrightarrow \boxed{Z_\alpha = 1.96} \qquad (\text{from table}(4))$$

$$n = \frac{(1.96)^2 * 0.15 * 0.85}{(0.05)^2} = \frac{0.4898}{0.0025} = 195.92 \cong 196$$

طريقة إيجاد حجم العينة n في حالة التوزيع ثنائي الحدين بوسط مجهول μ وتباين معلوم σ^2 ، وان حجم المجتمع معلوم، وان وسط وتباين المجتمع عبارة عن نسبة (خاصية في المجتمع)

على افتراض ان توزيع المجتمعات هو مقارب للتوزيع الطبيعي.

$$\boxed{n = \frac{N * Z^2 * p * q}{Nd^2 + Z^2 * p * q}}$$

مثال: جد حجم العينة ، اذا علمت ان :

$N = 10000 \Leftarrow$

$p = 10\% \Leftarrow$

$\alpha = 0.1 \Leftarrow$

$d = 0.02 \Leftarrow$

الحل:

$$n = \frac{N * Z^2 * p * q}{Nd^2 + Z^2 * p * q}$$

$$p = 10\% = \frac{10}{100} = 0.1$$

$$q = 1 - p = 90\% = \frac{90}{100} = 0.9$$

$$\alpha = 0.1 \Longrightarrow \frac{\alpha}{2} = 0.05 \Longrightarrow \boxed{Z_\alpha = 1.6449} \qquad (\text{from table}(4))$$

$$n = \frac{10000 * (1.6449)^2 * 0.1 * 0.9}{10000 * (0.02)^2 + (1.6449)^2 * 0.1 * 0.9} = \frac{2435.1264}{4 + 0.2435}$$

$$n = \frac{2435.1264}{4.2435} = 573.8486 \cong \boxed{574}$$

تحديد حجم العينة وفقاً للامكانيات المالية المتاحة

في حالة جمع البيانات باستعمال استمارة الاستبيان ، فالامر هنا يتطلب إمكانيات مالية متاحة ، ونفقات عامة كالتحليل والطباعة للاستمارة ، ونفقات ملئ الاستمارة ،إذ يتطلب الامر مراعاة هذه الإمكانيات المتوفرة لدى الباحث فيمكن تحديد حجم العينة المتمثل بعدد الاستمارات التي تخص البحث قيد الدراسة وفق الصيغة التالية:

$$n = \frac{c - c_0}{c_1}$$

حيث ان:

c : الإمكانيات المالية المتاحة لدى الباحث.

c_0 : كلفة التحليل والطباعة لاستمارة الاستبيان.

c_1 : كلفة ملئ الاستمارة.

مثال: أراد باحث جمع بيانات باستعمال استمارة الاستبيان وكان لديه مبلغ متاح قدره = 25000 دينار ، وان كلفة ملئ الاستمارة الواحدة هو 250 دينار ، وان كلفة التحليل والطباعة 6000 دينار ، المطلوب: جد عدد الاستبانات؟

الحل:

$$\mathrm{n} = \frac{\mathrm{c} - \mathrm{c_0}}{\mathrm{c_1}}$$

$$\mathrm{n} = \frac{25000 - 6000}{250} = \frac{19000}{250} = \boxed{76}$$

العينة العشوائية البسيطة (Simple Random Sampling)

وهي تلك العينة المختارة بشكل عشوائي يضمن لان تمتلك اية مفردة من مفردات المجتمع الاحصائي نفس الفرصة في الظهور ضمن مفردات العينة ويراعى عند استخدام هذا النوع من العينات ان يكون المجتمع الاحصائي متجانس من حيث الصفة او الصفات ذات العلاقة بالبحث او الدراسة.

وفيما يلي وصف لأسلوب المعاينة العشوائية البسيطة:

بافتراض أن المجتمع الاحصائي متجانس ومحدود وأن عدد مفرداتهِ يساوي $(\mathbf{N})$.

نستنتج الاتي:

احتمال اختيار أي مفردة من المجتمع $= \frac{\mathbf{1}}{\mathbf{N}}$

عدد العينات العشوائية البسيطة $= \mathbf{r} = \mathbf{C_n^N} = \frac{\mathbf{N!}}{\mathbf{n!(N-n)!}}$

احتمال سحب أي عينة $= \frac{\mathbf{1}}{\mathbf{r}}$

عدد مرات ظهور مفردة معينة في جميع العينات $= \mathbf{r^*} = \mathbf{C_{n-1}^{N-1}}$

احتمال ظهور مفردة معينة في جميع العينات $= \frac{\mathbf{r^*}}{\mathbf{r}}$

حيث ان:

$$\mathbf{N! = N(N-1)(N-2)\dots 3.2.1}$$

على سبيل الأمثلة:

$$\mathbf{6! = 6.5.4.3.2.1 = 720}$$

مع ملاحظة ان:

$$\mathbf{1! = 1}$$

$$\mathbf{0! = 1}$$

<u>مثال (1):</u> مجتمع احصائي متجانس عدد مفرداتهِ يساوي $(\mathbf{4})$ ، لتكن (احمد ، طارق ، خالد ، حسن) ، يراد اختيار عينة عشوائية بسيطة قوامها $(\mathbf{3})$ مفردات.

المطلوب:

1) ما هو احتمال اختيار المفردة (خالد) من هذا المجتمع.
2) ما هو عدد العينات العشوائية البسيطة التي يمكن اختيارها من هذا المجتمع.
3) ما هو احتمال سحب اي عينة.
4) ما هو عدد مرات ظهور المفردة (خالد) في جميع العينات.
5) ما هو احتمال ظهور المفردة (خالد) في جميع العينات.

الاجابة: $(\mathbf{n = 3})$ ، $(\mathbf{N = 4})$

احتمال اختيار أي مفردة من المجتمع = $\frac{1}{N}$

احتمال اختيار المفردة (خالد) = $\frac{1}{4}$

عدد العينات العشوائية البسيطة $= r = C_n^N = \frac{N!}{n!(N-n)!}$

عدد العينات العشوائية البسيطة $= r = C_3^4 = \frac{4!}{3!(4-3)!} = \frac{24}{6(1)} = 4$

احتمال سحب أي عينة = $\frac{1}{r} = \frac{1}{4}$

عدد مرات ظهور مفردة معينة في جميع العينات $= r^* = C_{n-1}^{N-1}$

عدد مرات ظهور المفردة (خالد) في جميع العينات $= r^* = C_2^3 = \frac{3!}{2!(3-2)!} = \frac{6}{2(1)} = 3$

احتمال ظهور المفردة (خالد) في جميع العينات = $\frac{r^*}{r} = \frac{3}{4}$

والجدول الآتي للتوضيح:

المفردات الاحصائية				تسلسل العينة	العينات العشوائية البسيطة		
احمد	طارق	خالد	حسن	الأولى	احمد	طارق	خالد
احمد	طارق	خالد	حسن	الثانية	احمد	طارق	حسن
احمد	طارق	خالد	حسن	الثالثة	احمد	خالد	حسن
احمد	طارق	خالد	حسن	الرابعة	طارق	خالد	حسن

العينة العشوائية الطبقية (Stratified Random Sampling)

تعد العينة العشوائية الطبقية أفضل أنواع العينات وأكثرها دقة في تمثيل المجتمع الاحصائي الغير متجانس ، حيث أنهُ في كثير من الاحوال تكون مفردات المجتمع الاحصائي غير متجانسة من حيث الصفة أو الصفات المدروسة. ففي حالة كون المجتمع الاحصائي ذو مفردات غير متجانسة لا يجوز سحب عينة عشوائية بسيطة تمثل هذا المجتمع.

فعلى سبيل المثال اذا اريد دراسة دخل الاسرة فأننا نجد ان هناك اسر ذات دخول عالية واخرى ذات دخول متوسطة واخرى ذات دخول منخفضة ، إذاً المجتمع الاحصائي هنا غير متجانس من حيث الصفة المدروسة ولا يجوز سحب عينة عشوائية بسيطة لأننا سنحصل على تقدير لمتوسط الدخل يكون منحازاً لاحدى الفئات الثلاثة ، وعليهِ يجب تقسيم المجتمع الاحصائي الى ثلاث طبقات (فئات ، مجموعات).

الطبقة الأولى تضم الاسر ذات الدخول المرتفعة

الطبقة الثانية تضم الاسر ذات الدخول المتوسطة

الطبقة الثالثة تضم الاسر ذات الدخول المنخفضة

وبعد ذلك يتم سحب عينة عشوائية بسيطة من كل طبقة يتناسب حجمها وحجم الطبقة في المجتمع ، وان مجموع حجوم العينات العشوائية الثلاثة تؤلف (تكون) العينة العشوائية الطبقية.

وفيما يلي وصف لأسلوب المعاينة العشوائية الطبقية ، بافتراض أن المجتمع الاحصائي مؤلف من $(\mathbf{N})$ من المفردات والذي يمكن تجزئتهِ الى $[\mathbf{L}]$ من الطبقات حجمها $[\mathbf{N_1, N_2, \dots, N_L}]$ ، حيث ان:

$$\sum_{\mathbf{i=1}}^{\mathbf{L}} \mathbf{N_i} = \mathbf{N_1} + \mathbf{N_2} + \dots + \mathbf{N_L} = \mathbf{N}$$

وان تقسيم المجتمع الاحصائي الى $[\mathbf{L}]$ من الطبقات يتم على اساس اشتراك مفردات كل طبقة بصفة او عدة صفات مشتركة.

فإذا أُريد اختيار عينة طبقية عشوائية ذات حجم $(\mathbf{n})$ فان هناك عدة طرائق لاختيار هذهِ العينة منها ما تسمى بطريقة التوزيع المتناسب **(Proportional allocation)** وبموجب هذهِ الطريقة يتم سحب عينة عشوائية بسيطة من كل طبقة يتناسب حجمها وحجم الطبقة في المجتمع ، فإذا كانت نسبة كل طبقة في المجتمع (وزن الطبقة) يساوي:

$$\mathbf{W_h} = \frac{\mathbf{N_h}}{\mathbf{N}} \quad , \mathbf{h = 1, 2, \dots, L}$$

حيث أن:

$\mathbf{W_h}$: وزن الطبقة.

$\mathbf{h}$: يمثل دليل (مؤشر) لتسلسل الطبقة .

$$\sum_{\mathbf{h=1}}^{\mathbf{L}} \mathbf{W_h} = \mathbf{W_1} + \mathbf{W_2} + \cdots + \mathbf{W_L} = \mathbf{1}$$

وأن حجم العينة العشوائية البسيطة المسحوبة من الطبقة الاولى هو

$$\mathbf{n_1 = W_1 * n = \frac{N_1}{N} * n}$$

وأن حجم العينة العشوائية البسيطة المسحوبة من الطبقة الثانية هو

$$\mathbf{n_2 = W_2 * n = \frac{N_2}{N} * n}$$

⋮

وأن حجم العينة العشوائية البسيطة المسحوبة من الطبقة الاخيرة هو

$$\mathbf{n_L = W_L * n = \frac{N_L}{N} * n}$$

وبذلك فأن حجم العينة العشوائية الطبقية يساوي:

$$\mathbf{\sum_{i=1}^{L} n_i = n_1 + n_2 + \ ... + \ n_L = n}$$

مما سبق نلاحظ ان:

$$\mathbf{\frac{n_h}{n} = \frac{N_h}{N}}$$

اي ان نسبة مساهمة الطبقة **h** في العينة تساوي نسبة مساهمة الطبقة **h** في المجتمع.

والجدول الآتي يوضح اهم خطوات سحب عينة عشوائية طبقية:

تسلسل الطبقة	N_h	$W_h = \frac{N_h}{N}$	$n_h = n * W_h$	$\frac{N_h}{N} = \frac{n_h}{n}$
1	N_1	$W_1 = \frac{N_1}{N}$	$n_1 = n * W_1$	$\frac{N_1}{N} = \frac{n_1}{n}$
2	N_2	$W_2 = \frac{N_2}{N}$	$n_2 = n * W_2$	$\frac{N_2}{N} = \frac{n_2}{n}$
3	N_3	$W_3 = \frac{N_3}{N}$	$n_3 = n * W_3$	$\frac{N_3}{N} = \frac{n_3}{n}$
⋮	⋮	⋮	⋮	⋮
L	N_L	$W_L = \frac{N_L}{N}$	$n_L = n * W_L$	$\frac{N_L}{N} = \frac{n_L}{n}$
المجموع	N	1	n	

مثال (2): مجتمع احصائي مؤلف من (2200) اسرة ، يراد دراسة متوسط الدخل للأسرة، علماً ان هذا المجتمع الاحصائي يضم اسر ذات دخول مرتفعة وذات دخول متوسطة وذات دخول منخفضة ، فإذا علمت ان:

عدد الأسر ذات الدخول المرتفعة يساوي (700)

عدد الاسر ذات الدخول المتوسطة يساوي (900)

عدد الاسر ذات الدخول المنخفضة يساوي (600)

يراد سحب عينة عشوائية طبقية ذات حجم (110) ، المطلوب: ما هو حجوم عينات الطبقات باعتماد طريقة التوزيع المتناسب.

الحل: من معطيات السؤال نستخرج الآتي:

$$N_3 = 600 \text{ ، } N_2 = 900 \text{ ، } N_1 = 700 \text{ ، } n=110 \text{ ، } N=2200$$

$$n_1 = W_1 * n = \frac{N_1}{N} * n = \frac{700}{2200} * 110 = \frac{770}{22} = \boxed{35}$$

$$n_2 = W_2 * n = \frac{N_2}{N} * n = \frac{900}{2200} * 110 = \frac{990}{22} = \boxed{45}$$

$$n_3 = W_3 * n = \frac{N_3}{N} * n = \frac{600}{2200} * 110 = \frac{660}{22} = \boxed{30}$$

نلاحظ ان:

$$\sum_{i=1}^{L} N_i = N_1 + N_2 + \ldots + N_L = N \quad \Rightarrow 700 + 900 + 600 = 2200$$

$$\sum_{i=1}^{L} n_i = n_1 + n_2 + \ldots + n_L = n \quad \Rightarrow 35 + 45 + 30 = 110$$

$$\sum_{h=1}^{L} W_h = W_1 + W_2 + \cdots + W_L = 1 \Rightarrow \frac{7}{22} + \frac{9}{22} + \frac{6}{22} = \frac{22}{22} = 1$$

$$\left[\frac{n_h}{n} = \frac{N_h}{N}\right] \Rightarrow \frac{35}{110} = \frac{700}{2200} = \boxed{0.318}$$

$$\frac{30}{110} = \frac{600}{2200} = \boxed{0.273} \quad , \quad \frac{45}{110} = \frac{900}{2200} = \boxed{0.409}$$

العينة العشوائية المنتظمة (Systematic Random Sampling)

بموجب هذا النوع من المعاينة يتم تقسيم مفردات المجتمع الاحصائي البالغ عددها (N) الى عدد من المجاميع كل مجموعة تحتوي على (K) من المفردات حيث ان $\mathbf{k = \frac{N}{n}}$ ، علماً بأن مفردات المجتمع **مرتبة وفق نظام معين** كأن يكون ترتيباً تصاعدياً او تنازلياً أو مثلاً ترتيب الدور السكنية حسب تسلسلها في شارع معين. وعند اختيار مفردات العينة المنتظمة يتم اولاً اختيار مفردة واحدة بصورة عشوائية من المجموعة الاولى وعلى ضوء تسلسل المفردة المختارة من المجموعة الاولى يتم اختيار مفردة من المجموعة الثانية بعد اضافة العدد (K) على تسلسل المفردة الاولى وهكذا يضاف العدد (K) على تسلسل المفردة الثانية لنحصل على المفردة الثالثة وهكذا وصولاً لأخر مفردة. بمعنى اخر يتم اختيار مفردة من المجموعة الاولى بطريقة عشوائية اما بقية المفردات يتم اختيارها على ابعاد متساوية من المفردة الاولى.

مثال (3): تم ترتيب **(24)** طالب حسب تسلسل درجاتهم تنازلياً ، يراد اختيار عينة عشوائية منتظمة بحجم **(6)** طلاب للتعرف على اسباب انخفاض مستواهم في الامتحان. المطلوب تحديد تسلسل هؤلاء الطلبة.

الحل: من معطيات السؤال نلاحظ بأن:

n=6 ، N=24

يتم تقسيم الطلبة الى (6) مجاميع كل منها تحتوي على اربعة طلاب وكمايلي:

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
13	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**

ملاحظة: 1. الرقم يشير الى تسلسل الطالب وليس درجتهِ حيث ان الطالب رقم **(1)** لهُ اعلى درجة في الامتحان.

2. تم تحديد عدد الطلبة في كل مجموعة كالآتي: $\mathbf{k = \frac{N}{n} = \frac{24}{6} = 4}$.

ب- يتم اختيار طالب بصورة عشوائية من المجموعة الاولى وليكن مثلاً الطالب ذو تسلسل رقم (3).

ج- اما بقية الطلبة فنحصل عليهم بإضافة **(k=4)** الى التسلسل (3) لنحصل على الطالب الثاني في العينة وهو التسلسل (7) وهكذا ... وصولاً لأخر طالب في العينة وهو التسلسل (23).

إذاً العينة العشوائية المنتظمة تكون على اساس ان الطالب الذي تم اختياره عشوائياً من المجموعة الاولى والذي يحمل التسلسل (3):

[3 7 11 15 19 23]

ملاحظة: عدد العينات العشوائية المنتظمة التي يمكن اختيارها من المجتمع الاحصائي يساوي عدد مفردات كل مجموعة **(K)**.

العينة متعددة المراحل (Cluster Random Sampling)

بموجب اسلوب المعاينة المتعددة المراحل يتم تقسيم المجتمع الاحصائي الى وحدات اولية ثم يتم اختيار عينة عشوائية بسيطة من هذهِ الوحدات كمرحلة اولى ، ثم يتم تقسيم الوحدات الاولية المختارة الى وحدات اصغر تدعى بالوحدات الثانوية ويتم اختيار عينة عشوائية بسيطة من الوحدات الثانوية لكل وحدة اولية كمرحلة ثانية ، ثم تقسم الوحدات الثانوية المختارة الى وحدات اصغر يتم اختيار عينة عشوائية بسيطة منها كمرحلة ثالثة ، وتستمر عملية التقسيم والاختيار لحين الوصول الى المفردات التي يتم جمع البيانات منها.

فمثلاً عند إجراء دراسة لتقدير متوسط استهلاك العائلة العراقية لمادة السكر فإن الوحدة الاحصائية التي يمكن الحصول على البيانات منها هي العائلة العراقية.

فعند اختيار عينة عنقودية يتم تقسيم العراق الى **محافظات** كوحدات اولية نختار منها عشوائياً عينة من المحافظات كمرحلة اولى ثم نقسم المحافظات المختارة في المرحلة الاولى الى **اقضية** كوحدات ثانوية يتم اختيار عينة عشوائية بسيطة منها كمرحلة ثانية ثم تقسم الوحدات الثانوية المختارة في المرحلة الثانية الى **نواحي** يتم اختيار عينة عشوائية بسيطة منها كمرحلة ثالثة ثم تقسم النواحي المختارة في المرحلة الثالثة الى **محلات سكنية** يتم اختيار عينة عشوائية بسيطة منها كمرحلة رابعة ثم تقسم المحلات السكنية المختارة الى **ازقة** ومنها يتم اختيار عينة عشوائية بسيطة والتي تقسم بدورها الى **الدور السكنية** التي يتم اختيار عينة عشوائية بسيطة منها وبهذا يتم الوصول الى العائلة التي منها يتم جمع البيانات.

العينة العمدية: **(Purposive Sample)**

وهي العينة المختارة بشكل متعمد نعتقد مسبقاً ان مفردات هذهِ العينة هي خير من يمثل مجتمع الدراسة.

فمثلاً عند دراسة السبل الكفيلة للارتقاء برياضة كرة القدم فمن الافضل اختيار عينة من المختصين برياضة كرة القدم وبشكل عمدي (مقصود) كون أن هذهِ العينة هي ذات خبرة بشؤون هذهِ الرياضة.

العينة الحصصية: **(Quota Sample)**

بموجب هذا النوع من المعاينة يتم تقسيم المجتمع الاحصائي الى عدة طبقات بالاستناد الى معايير تتعلق بطبيعة الدراسة ثم يتم اختيار عينة عمدية من كل طبقة (بشكل غير عشوائي) يتناسب حجمها وحجم الطبقة في المجتمع ومجموع حجوم هذهِ العينات العمدية يمثل حجم العينة الحصصية.