المحـــور الأول: نظريـــة المعاينـــة

تهتم نظرية العينات بدراسـة العلاقة بين المجتمع والعينات المسـحوبة منه فيما يسـمى بالاستدلال الإحصائي، statistical inference. هناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي ومن أشهر هذه الطرق هي العينة العشوائية

مجتمع الدراسة:

أي مجموعات من المفردات تشـترك في صـفه أو صـفات وتكون موضـوع دراسـة أو بحث فإن هذه المجموعة يطلق عليها إحصـائيا مجتمع الدراسـة (أو اختصـاراً المجتمع Population). والمجتمع قد يكون مجموعة ما من البشـر أو أشـجار أنواع معينه من الفاكهة أو الحيوانات الزراعية أو إنتاج دولة ما لسلع معينه خلال فترة زمنية محدده...الخ.

والمجتمع قد يكون محدوداً إذا كان يمكن حصـر عدد أفراده مثل سـكان مدينة ما أو طلاب مرحلة تعليمية معينة، وقد يكون المجتمع غير محدود (لانهائي) إذا كان لا يمكن حصـر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار وعند دراسـة صـفة ما أو صـفات معينه لمجتمع ما، فإن البيانات الإحصـائية عن تلك الصـفة أو الصـفات تجمع بأحد أسلوبين

أساليب جمع البيانات الإحصائية

أولاً: أسـلوب الحصـر الشـامل (census) وفيه تجمع البيانات عن كل مفرده من مفردات المجتمع، وهذا الأسـلوب يتطلب وفرة في الوقت والمال والمجهود الفني وتزداد هذه المتطلبات وتتضـاعف كلما ازداد حجم المجتمع (عدد أفراد المجتمع). وهذا الأسـلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضـخمه مثل تعدادات السـكان والتعدادات الصناعية والتعدادات الزراعية.

الثاني: أسـلوب المعاينة (Sampling method) وفيه يتم جمع البيانات عن جزء من مفردات المجتمع يختـار بطريقـة أوبأخرى ويطلق عليـه عينـه **(Sample)** ثم بعـد ذلـك يتم

تعميم نتائج الدراسة على المجتمع بأكمله. أي أن أسلوب المعينة يقصد به دراسة خصائص المجتمع من خلال دراسة عينه مسحوبة منه، ونجاح هذا الأسلوب يعتمد على أن تحمل العينة أقصى درجة من دقة التمثيل للمجتمع المسحوبة منه.

من الأفضل في بعض الحالات الحصول على معلومات دقيقة عن طريق التعداد التام أو الحصر الشامل لجميع عناصر المجتمع، لكن لاستخدام أسلوب المعاينة فوائد جمة مقارنة بالتعداد الشامل يرد بيانها في الفقرة التالية.

مزايا أسلوب المعاينة

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عدديه منها:

* يؤدي استخدام العينات العشوائية إلى خفض تكاليف الدراسات الميدانية بسبب صغر حجم العينة بالنسبة إلى حجم المجتمع وهوما يؤدي إلى تخفيض الأعباء الإدارية والفنية التي تتطلبها أي دراسة ميدانية.

* يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتتضح أهمية الوقت عندما نقوم بدراسة ظاهرة تتغير بمرور الوقت، وعينة قد يترتب على دراسة تلك الظاهرة في المجتمع كله بجمع البيانات من جميع مفردات المجتمع أن يمر وقت بديل فتكون البيانات والنتائج وقت ظهورها غير مطابقة لواقع المجتمع وتصبح النتائج ذات قيمه محدودة بعد أن فقدت عنصر المطابقة مع واقع الظاهرة وتوزيعها الحالي لمجتمع، والتعدادات الدورية للسكان وبسبب ضخامة حجم العمل بها تستغرق وقتاً طويلاً حتى تصبح نتائجها جاهزة ومنشورة وقد يطول هذا الوقت إلى أكثر من ثلاث أو أربع سنوات حتى مع استخدام أحدث أجهزة الحاسبات الأليه الضخمة، ويكون على الباحثين مستخدمي هذه النتائج مراعاة الوقت الذي ينقض بين تاريخ إجراء التعداد وتاريخ نشر نتائجه وتعديل هذه النتائج في حدود ذلك.. وهذا دفع الكثير من الدول إلى تعزيز نتائج

التعدادات الدورية للسكان بنتائج تعدادات تجري بين كل تعدادين متتاليين على أساس العينة.

* في المجتمعات غير المحدودة (اللانهائية) مثل مجتمع الكائنات الحية في البحار والمحيطات لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة.

* أيضاً هناك بعض الاختبارات لابد وأن تتم بأسلوب المعاينة لأن إجراء مثل هذه الاختبارات على أساس الحصر الشامل يؤدي إلى تلف المادة المختبرة أو هلاكها. فاختبار صلاحية شحنه من المفرقعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينه.

أقسام العينات

تنقسم العينات عادة إلى قسمين رئيسين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لكل قسم منها.

العينات العشوائية

هي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة.. من أهم أنواع العينات العشوائية مايلي.

العينة العشوائية البسيطة

العينة العشوائية البسيطة هي طريقة المعاينة التي يكون فيها احتمال اختيار أي مفردة متساوأي أن المجتمع ككل يعامل بنفس الطريقة ولا يجرى عليه أي تقسيمات مختلفة حيث أن الوحدات المكونة لهذا المجتمع تعامل كلها باحتمالات متساوية مما يجعل المعادلات الرياضية والإحصائية المستخدمة لتقدير معالم المجتمع أبسط ما يمكن وتعرف هذه المعاينة بأسماء عديدة من أهم هذه الأسماء انتشارا العينة غير المقيدة وعينة تكافؤ الفرص.

شــروط اختيــار العينـة:

وجود إطار للمجتمع يكون حديثا وشاملا لكل مفردات المجتمع يرقم فيها وحدات المجتمع من 1 إلى N.

تحديد حجم العينة.

يتم اختيار كل مفردة من مفردات العينة مستقلة عن اختيار المفردات الأخرى أي يكون لكل مفردة من مفردات المجتمع الأصلي فرصة متساوية مع غيرها من المفردات في أن تكون ضمن مفردات العينة.

طــــرق سحــــب العينــــــة:

1- القرعـــة:

الطريقـــة الأولـــى:يقوم الباحث بإعداد قائمة بها جميع العينات المحتمل تكوينها من مجتمع البحث فمثلا لوكان لدينا مجتمع مكون من 6 مفردات (A,B,C,D,E,F) وأردنا معرفة العينات الممكن تكوينها من هذا المجتمع بحيث يكون حجم كل منها مفردتين فقط، إن عدد العينات الممكن سحبها يتم حسابه كالتالي:

أولا : في حالـة عـدم إعـادة المفـردة قبـل سحـب الـتي تليـها(بدون إرجــاع):

في هذه الحالة يتم استبعاد المفردة أوالعينة في كل مرة قبل سحب الثانية وهنا يمكن استخدام فكرة التوافيق حيث يتم توفيق عدد 2 مفردة وهم حجم المجموعة الواحدة من بين 6 مفردات وهم حجم المجتمع كله في الصورة التالية:

$$C_N^n = \frac{N!}{n!(N-n)!} = \frac{6!}{2!(6-2)!} = \frac{6*5*4!}{2*4!} = \frac{30}{2} = 15$$

حيث N تمثل حجم المجتمع وعددها 6 كما أن n تمثل حجم العينة أوالمجموعة الواحدة وعددها 2 في مثالنا

رقم العينة	مفردات العينة	رقم العينة	مفردات العينة	رقم العينة	مفردات العينة
1	AB	6	BC	11	CE
2	AC	7	BD	12	CF
3	AD	8	BE	13	DE
4	AE	9	BF	14	DF
5	AF	10	CD	15	EF

ثانيـــا: في حالـة إعـادة المفـردة قبـل سحب الـتي تليـها(مع الإرجـاع)

في هذه الحالة يتم إعادة المفردة أوالعينة أوالمجموعة المسحوبة في كل مرة قبل سحب الثانية وبالتالي يظل حجم المجتمع ثابت في كل مرة ولا ينقص وفي هذه الحالة تستخدم فكرة الأسس في الصورة التالية:

$$N^n=6^2=6*6=36$$

رقم العينة	مفردات العينة	رقم العينة	مفردات العينة	رقم العينة	مفردات العينة	رقم العينة	مفردات العينة
1	AA	10	BD	19	DA	28	ED
2	AB	11	BE	20	DB	29	EE
3	AC	12	BF	21	DC	30	EF
4	AD	13	CA	22	DD	31	FA
5	AE	14	CB	23	DE	32	FB
6	AF	15	CC	24	DF	33	FC
7	BA	16	CD	25	EA	34	FD
8	BB	17	CF	26	EB	35	FE

FF	36	EC	27	CE	18	BC	9

بعد ذلك يقوم الباحث بتسجيل رقم كل عينة محتملة في قصاصة من الورق أوكرة من الكرات أوبطاقة من البطاقات ثم تخلط حتى يكون السحب عشوائيا تماما وتعطى فرصا متساوية لكل مجموعة في الظهور في الدارسة ثم يتم السحب ويقرأ الرقم فيقع الاختيار على العينة التي تحمل هذا الرقم المختار، فمثلا لوقام الباحث بسحب قصاصة تحمل الرقم 5 لكانت المجموعة المؤلفة من المفردات *(A,E)*هي العينة التي تمثل المجتمع.

كثيرا ما يتعذر على الباحث إتباع الطريقة السابقة في اختيار العينة العشوائية البسيطة خصوصا في حالة كثرة عدد مفردات مجتمع البحث فمثلا لوكان حجم المجتمع 100 مفردة وكان حجم العينة المطلوبة 3 مفردات فإن عدد العينات التي يمكن سحبها تكون:

في حالة سحب العينات مع عدم الإرجاع هو:

$$C_N^n = \frac{N!}{n!(N-n)!} = \frac{100!}{3!(100-3)!} = \frac{100*99*98*97!}{3*2*97!} = \frac{970200}{6} = 161700$$

عينة

في حالة سحب العينات مع الإرجاع هو:

$$N^n = 100^3 = 100*100*100 = 1000000$$

فهل يعقل أن يقوم الباحث بكتابة مليون قصاصة ورق لكي يسحب منهم عينة تحتوي على 3 مفردات فقط ؟ الإجابة بالطبع تكون بالنفي

الطريقـــة الثانيـــة: هي أن يقوم الباحث بترقيم كل مفردة من مفردات المجتمع وتسجيل هذه الأرقام في قصاصات أوبطاقات أوكرات وخلطها خلطا جيدا ثم يسحب منها العدد المطلوب الذي يمثل حجم العينة وفي هذه الحالة يجب أن يفرق الباحث بين سحب المفردات مع الإرجاع

أومع عدم الإرجاع بعدها يتم سحب إحدى القصاصات ويسجل رقمها ثم يقوم بالسحب مرة أخرى ويتم تسجيل رقمها وهكذا إلى أن يتم اختيار العدد المطلوب.

فمثلا في حالة المثال السابق حيث يتكون المجتمع من 6 مفردات، يتم إعطاء كل مفردة رقم مسلسل:A=1، B=2، C=3، D=4، E=5، F=6 ، ثم يكتب كل رقم في قصاصة من الورق وتخلط القصاصات جيدا ثم يتم سحب قصاصة واحدة ويقرأ رقمها وليكن الرقم 4 وبالرجوع إلى قائمة المفردات نجد أن الرقم 4 يمثلD ومن ثم يكون أول مفردة في العينة المطلوبة هي المفردة D فإذا كان السحب مع عدم الإرجاع يتم استبعاد هذه المفردة فيكون مجتمع الدراسة الجديد هو*(A,B,C,E,F)* أما إذا كان السحب مع الإرجاع فيكون مجتمع الدراسة الجديد هو*(A,B,C,D,E,F)* ويتم تكرار السحب مرة أخرى وليكن الرقم الذي تحمله القصاصة الجديدة هوالرقم 2 حيث تحمله المفردة B فتكون المفردة B هى المفردة الثانية في العينة فإذا كان حجم العينة المطلوبة اثنان فنكون بذلك قد سحبنا عدد المفردات المطلوب إدخالها والعينة المختارة في الدراسة هي DB .

يصعب إتباع الطريقين السابقين عمليا خاصة إذا كان عدد مفردات مجتمع البحث كبيرا جدا نظرا لصعوبة عملية إعداد القصاصات أوالبطاقات أوالكرات وتسجيل أرقام مفردات المجتمع عليها ثم خلطها وسحب العدد المطلوب منها.

2- جـدول الأرقـام العشوائيـــة:

هي أعداد صـحيحة مكونة من الأرقام 0,1,2,3, ... 8,9 تشـكل أرقاما من رقم 1 حتى رقم 100 ألف تقريبا موضـوعة في جداول بغير ترتيب وإنما بصـورة عشـوائية مكونة من صفوف وأعمدة وهذه الجداول موجودة على شكل كتيب.

يتم التعامل مع الجدول من خلال اختيار نقطة بداية بشـكل عشـوائي ومن ثم يتم قراءة الأرقام التي يكون عدد منازلها مســـاويا لعدد منازل الأرقام المعطاة لوحدات المجتمع وفي

أي من الاتجاهات الأربعة، وهذه الطريقة تشترط وجود أسماء أفراد المجتمع مرقمين في تسلسل، فيقوم الباحث بفتح أي صفحة من كتيب الأرقام العشوائية ويضع رأس القلم على أي رقم فإن كان الرقم موافقا لرقم أحد أفراد المجتمع يختار ذلك الفرد ضمن العينة ثم يذهب للرقم الذي يليه ثم الذي يليه وهكذا.

أمثلة:

مثال 1: تحتوي كلية العلوم الاقتصادية بجامعة خميس مليانة 3500 طالب، ونريد أن نختار عشرة طلبة لإجراء دراسة معينة، إذا أردنا أن نسحب عينة عشوائية بسيطة عدد أفراها 10 أفراد من مجتمع عدد أفراده 3500 فرد فإننا هنا نستخدم جدول الأعداد العشوائية حسب الخطوات التالية:

1- إعطاء أرقام متسلسلة لجميع أفراد المجتمع بحيث يتكون كل رقم من أربعة أرقام، لماذا لأن مجتمع الدراسة مكون من 3500 فرد والرقم 3500 هذا مكون من أربعة أرقام، وعليه عند إعطاء أرقام متسلسلة لجميع أفراد المجتمع يكون كالتالي، الفرد الأول في العينة يعطى الرقم 0001 والثاني 0002 والثالث 0003 وهكذا حتى تصل إلى الرقم 3500.

2- نستخدم جدول الأعداد العشوائية، ثم نختار صفّا وعمودا بطريقة عشوائية، وليكن العمود الثالث والسطر الثاني من الصفحة الأولى من جدول الأعداد العشوائية، ولنفرض أن القراءة تكون من اليمين إلى اليسار ومن الأعلى إلى الأسفل ونقرأ الأرقام الأربعة الأولى من اليسار حسب عدد منازل حجم مجتمع الدراسة كما ذكرنا آنفاً.

3- إذا كان الرقم من ضمن الأعداد المتسلسلة السابقة نأخذه، أما إذا لم يكن نذهب إلى العدد الذي يليه وهكذا حتى نحصل على الحجم المطلوب حسب حجم العينة وهو 10 أفراد.

ومن الخطوات السابقة نحصل على الأرقام التالية:

-3961-0087-3897-8501-3416-9715-1616-9960-3185-5296-5586-9814
-9775-0216 -6665-1925-1545-1032-6473-2290-0061-9399-5694-9129
.1859-9161-5159-6714-5414-8789-3128-6499-1555-6552-6334

أي عدد يزيد عن 3500 أوإذا تكرر نهمله وهكذا حتى نحصل على العدد المطلوب للعينة، بالتالي الأشخاص الذين يتم اختيارهم هم أصحاب الأرقام:

.0216 -1925 -1545 -1032 -2290 -0061 -0087 -3416 -1616 -3185

مثـــــال 2: لدينا قائمة مكونة من 100 شخص ونرغب في اختيار عينة من 10 أشخاص، نحتاج هنا إلى رقم مكون من 3 خانات فقط نستخدم جدول الأعداد العشوائية، نختار العمود الرابع والسطر الثاني، ولنفرض أن القراءة تكون من اليمين إلى اليسار ومن الأعلى إلى الأسفل ونقرأ الأرقام الثلاثة الأولى من اليسار.

-021 -666 -192 -154 -103 -647 -229 -006- 939 -318 -529-558-981-982
.633-977

نلاحظ أن معظم الأرقام تهمل وقد ينتهي الجدول ولا نحصل على الأرقام المطلوبة، بالتالي نستعمل الطريقة التالية: كل رقم يزيد عن أكبر رقم متسلسل لأفراد المجتمع أي 100 نقوم بطرحه من الرقم العشوائي أوطرح مضاعفاته 100-200-300...فتصبح الأرقام السابقة على النحوالتالي:

-021 -066 -092 -054 -003 -047 -029-006- 039 -018 -029-058-081-082
.033-077

فيكون أفراد العينة المطلوبة هم: -047 -029-006- 039 -018 -029-058-081-082
.003

العينـــة العشوائيـــة المنتظمـة (Systematic Sample)

اختيار هذه العينة يتطلب وجود إطار للمجتمع كما في حالة العينة العشوائية البسيطة بحيث يعطى لكل مفرده من مفردات المجتمع رقماً متسلسلاً داخل الإطار، ثم نختار مفردات العينة من الإطار بحيث يكون الرقم المتسلسل لكل مفردة يبعد بعداً ثابتاً منتظماً عن رقم المفردة السابقة لها وكذلك رقم المفردة اللاحقة لها. فمثلاً إذا كان لدينا مجتمعاً حجمه 2000 مفرده

ونريد اختيار عينه منتظمة حجمها 100 مفرده فإننا نقسم الإطار إلى فترات منتظمة طول كل فترة $\frac{N}{n} = \frac{2000}{100} = 20$ مفردة ومن داخل مفردات الفترة الأولى (1 – 20) يختار مفرده واحدة عشوائياً ولتكن رقم 14 مثلاً وبناء على رقم تلك المفردة يتحدد باقي أرقام المفردات الأخرى بشكل منتظم وفقا لمتوالية عددية حدها الأول هو14، وأساسها يساوي عدد مفردات كل فترة " 20 "، أي أن مفردات العينة هي: 34، 54، ....، 1974، 1994.

والعينة المنتظمة كثيرة الاستعمال في التطبيقات العملية لقلة تكاليفها وقلة الأخطأء التي ترتكب في اختيار مفردات العينة فضلاً عن سهولة إجرائها. ولكن أهم عيوب المعاينة المنتظمة هوعدم صلاحيتها إذا ما وجدت علاقة دورية مع ترتيب العناصر في القائمة وكان طول الفترة بين عناصر العينة مساوياً لطول الدورة أوإحدى مضاعفاتها.

العينـــة الطبـقيــة العشوائيــــة (Stratified random sampling)

تعتبر العينة العشـــوائية الطبقية أفضـــل أنواع العينات وأكثرها دقة في تمثيل المجتمع الاحصـــائي غير المتجانس حيث أنه في كثير من الأحوال تكون مفردات المجتمع الاحصـــائي غير متجانسة من حيث الصفة أوالصفات المدروسة

يلجأ إليها الباحث في حالة ما إذا كان مجتمع الدراسـة منقسـما إلى طبقات طبيعية وتكون لدينا الرغبة في تمثيل جميع هذه الطبقات في العينة بحيث أن التجانس أوالتقارب داخل كل طبقة من طبقات مجتمع الدراسـة أكبر من التجانس داخل المجتمع ككل (أي أن التشـتت داخل المجتمع ككل أكبر من التشـتت داخل كل فئة من فئاته على حده) في هذه الحالة يجب على الباحث مراعاة أن الطبقة داخل العينة بنفس نســـبة وجودها داخل المجتمع، وتتلخص الطريقة بتحديد حجم العينات الجزئية المتناسبة من كل طبقة على أساس المعادلة

حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

بعد أن يتم تحديد عدد المفردات التي يجب سـحبها من كل طبقة للدخول في العينة فإن هذه المفردات يتم سـحبها عشـوائياً من داخل الطبقة (عينة عشـوائية بسـيطة أومنتظمة) ومجموع هذه المفردات تكوّن العينة الطبقية العشوائية.

فعلى سبيل المثال إذا أريد دراسة دخل الأسرة فإننا نجد أن هناك أسر ذات دخول عالية وأخرى ذات دخول متوسطة وأخرى ذات دخول منخفضة إذن المجتمع الاحصائي هنا غير متجانس من حيث الصفة المدروسة ولا يجوز سحب عينة عشوائية بسيطة لأننا سنحصل على تقدير متوسط الدخل يكون منحازا لإحدى الفئات الثلاث، وعليه يجب تقسيم المجتمع الاحصائي إلى ثلاث فئات الأولى تضم الأسر ذات الدخول المرتفعة، والثانية تضم الأسر ذات الدخول المتوسطة والثالثة تضم الأسر ذات الدخول المنخفضة، وبعد ذلك يتم سحب عينة عشوائية بسيطة أومنتظمة من كل مجموعة يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع حجوم العينات العشوائية الثلاث تؤلف حجم العينة العشوائية الطبقية.

مثـال 1: إذا كانت طبقات أحد المجتمعات تحتوي العناصر كما في الجدول التالي:

الطبقة الأولى	الطبقة الثانية	الطبقة الثالثة	الطبقة الرابعة	الطبقة الخامسة
500	400	280	200	220

وأراد باحث اختيار عينة حجمها 160 من هذا المجتمع، فما حجم العينة في كل طبقة.

الحـل:

المجتمع الكلي = 500 + 400+280+200+220=1600   العينة من الطبقة الأولى =

$\frac{160}{1600}\times 500 = 50$ العينة من الطبقة الثانية = $\frac{160}{1600}\times 400 = 40$ العينة

من الطبقة الثالثة = $\frac{160}{1600}\times 280 = 28$ العينة من الطبقة الرابعة =

$\frac{160}{1600}\times 200 = 20$ العينة من الطبقة الخامسة = $\frac{160}{1600}\times 220 = 22$

مثـال 2: بفرض أن مجتمع الدراسة مكون من 5000 مزرعة تمور، وهذه المزارع مصنفة

حسب حجم الإنتاج كما يلي:

حجم الإنتــاج	صغير	متوسط	كبير
عدد المــزارع	1500	2500	1000

فإذا كان حجم العينة المطلوب ســحبها هو1000 مزرعة، وفقا لطريقة التقســيم المتناســب تسحب عينة تتناسب مع الوزن النسبي لكل طبقة في المجتمع كما هومبين بالجدول التالي:

حجم عينة كل طبقة	نسب الطبقات	عدد المزارع	نوع المزرعة
$n_1 = n \times 0.3$ $= 1000 \times 0.3 = 300$	1500/5000 = 0.3	1500	حجم صغير
$n_2 = n \times 0.5$ $= 1000 \times 0.5 = 500$	2500/5000 = 0.5	2500	حجم متوسط
$n_3 = n \times 0.2$ $= 1000 \times 0.2 = 200$	1000/5000 = 0.2	1000	حجم كبير
$n = 1000$	1.0	N=5000	عدد المزارع

إذا تسحب عينة عشوائية من المزارع ذات الحجم الصغير حجمها 300 مزرعة، وعينة عشوائية من المزارع ذات الحجم المتوسط حجمها 500 مزرعة، وعينة عشوائية من المزارع ذات الحجم الكبير حجمها 200 مزرعة.

العينـــة العنقوديـــة (Cluster Sample)

يلجأ إليها الباحث عندما يكون مجتمع الدراســة كبير جداً ومتناثراً على مســاحات شــاسـعة تكلف الكثير من الوقت والجهد في التنقل بينها عند جمع البيانات، أيضــاً في حالة عدم وجود إطار يضم جميع مفردات المجتمع فيستحيل الاختيار العشوائي مباشر من المجتمع، لهذا يلجأ الباحث إلى أخذ العينة على مراحل متعددة متتالية.

يتم تقسيم المجتمع الاحصائي إلى وحدات أولية ثم يتم سحب عينة عشوائية بسيطة من هذه الوحدات الأولية كمرحلة أولى ثم يتم تقسيم الوحدات الأولية المختارة إلى وحدات أصغر تدعى

بالوحدات الثانوية ويتم اختيار عينة عشوائية بسيطة من كل وحدة من الوحدات الثانوية، ثم تقسيم الوحدات الثانوية المختارة إلى وحدات أصغر ويتم اختيار عينة عشوائية منها كمرحلة ثالثة وتستمر عملية التقسيم والاختيار لحين الوصول إلى المفردات التي يتم جمع البيانات منها وكأننا نتحدث عن عنقود.

وعلى سبيل المثال فعند إجراء دراسة لتقدير متوسط استهلاك العائلة في الجزائر لمادة السكر، إذن الوحدة الاحصائية التي يمكن الحصول على بيانات منها هي العائلة الجزائرية، فعند اختيار عينة عنقودية يتم تقسيم الجزائر إلى ولايات كوحدات أولية لاختيار عينة عشوائية من الولايات كمرحلة أولى ثم تقسيم الولايات المختارة في المرحلة الأولى إلى دوائر كمرحلة ثانية يتم اختيار عينة عشوائية منها كمرحلة ثانية ثم تقسيم الوحدات الثانوية المختارة في المرحلة الثانية إلى بلديات ويتم اختيار عينة عشوائية منها كمرحلة ثالثة ثم تقسيم البلديات المختارة في المرحلة الثالثة إلى نواحي يتم اختيار عينة عشوائية منها كمرحلة رابعة ثم تقسيم النواحي المختارة في المرحلة الرابعة إلى محلات سكنية التي يتم اختيار عينة عشوائية منها وبهذا يتم الحصول على العائلات التي منها يتم عملية جمع البيانات.

العينات غير العشوائية:

هي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالباً يتدخل الباحث في عملية الاختيار بصورة أوبأخرى... ومن أهم أنواع العينات غير العشوائية:

(أ) **العينة العمدية أوالمقصودة**: **Purposive sample**

يلجأ الباحث إلى هذه الطريقة فيما إذا كان مجتمع الدراسة كبير جداً وكانت إمكانياته لا تسمح له إلا بدراسة عينة حجمها صغير جداً بالنسبة لمجتمع الدراسة، في هذه

الحالة يتعمد الباحث اختيار مفردات معينة كعينة لمجتمع الدراسة يرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلاً مقبولاً لمجتمع الدراسة.

مثلاً إذا أراد باحث دراسة خصائص اقتصادية أواجتماعية معينه عن ريف دولة ما، وكانت إمكانياته المالية والإدارية لا تسمح له بعينة سوى سكان قرية واحدة، فإنه في هذه الحالة إذا ما تم اختيار القرية عشوائياً من بين آلاف القرى بتلك الدولة فإن الصدفة قد تأتي بقرية بعيدة في خصائصها (من حيث الظاهرة موضوع الدراسة) عن خصائص معظم قرى تلك الدولة... كأن تأتي بالصدفة قرية ساحلية معظم سكانها من الصيادين أوقرية قريبة من مشروع صناعي ضخم يستوعب في قواه العاملة معظم سكانها.. هذه القرية أوتلك قد يأخذ النمط المعيشي لسكناها طابعاً خاصاً – نابعاً عن ظروفها الخاصة – بعيداً عن النمط المعيشي المعتاد لبقية القرى، لذلك فأي منها لا يمكن أن يعطي تمثيلاً مقبولا لريف تلك الدولة. لهذا فإن الباحث وعلى ضوء خبراته السابقة يتعمد اختيار قرية معينة يرى أنها – من وجهة نظره الشخصية- يمكن أن تمثل الريف. وهذه الطريقة غير علميه وغالباً يتم اللجوء إليها في حالة البحوث التمهيدية.

(ب) العينة الحصصيه: Quota sample

وهي نوع خاص من العينات غير العشوائية وتستخدم كثيراً في معاينة الرأي العام (على سبيل المثال عمليات استطلاعية الرأي العام التي يقوم بها معهد جالوب قبل إجراء انتخابات الرئاسة في الولايات المتحدة الأمريكية)... في هذه الطريقة يقسم المجتمع موضوع الدراسة إلى طبقات بالنسبة إلى صفات أوخصائص معينة ويتم العمل على تمثيل كل طبقة منها في العينة بنسبة وجودها في المجتمع الأصلي (وعلى سبيل المثال في حالة دراسة الدخل لمنطقة ما ورؤى أن يكون حجم العينة المطلوبة 100 فرد مثلاً عندما يريد الباحث أن يقوم جامعوا البيانات بالحصول على البيانات من 20 موظفاً، 45 من العمال الحرفيين، 35 من

ذوي الأعمال الحرة.. وتترك الحرية لجامعي البيانات في اختيار الأفراد المطلوبة فيها حدود المواصفات الموضوعة لكل طبقة من الطبقات المذكورة.

واضح أنه رغماً من أن هذه الطريقة في ظاهرها مماثلة للعينة الطبقية العشوائية.. إلا أنه في الحالة الأخيرة (العينة الطبقية العشوائية) يكون اختيار المفردات عشوائياً من داخل كل طبقة ولا يترك لجامع البيانات حرية اختيار المفردات من كل طبقة والذي قد يترتب عليه تميزاً كبيراً.

عموماً.. يلجأ الباحث إلى العينة الحصصيه إذا كان من المرغوب فيها اظهار النتائج في وقت قصير مع التغاضي عن توافر درجة دقة عالية بتلك النتائج.

توزيعات المعاينة: Sampling Distributions

نفرض أننا أخذنا عينه حجمها n من مجتمع ما، ثم سحبنا منها بعض المقاييس الإحصائية مثل المتوسط الحسابي، التباين،... فإن كل مقياس من هذه المقاييس يعتبر متغير عشوائي في ذاته يختلف من عينه إلى أخرى – هذا المتغير العشوائي يخضع لتوزيع معين – هذا التوزيع يسمى بتوزيع العينة. فمثلاً نقول أن توزيع المعاينة للمتوسط الحسابي وهوعبارة عن توزيع جميع المتوسطات الحسابية للعينات المأخوذة من نفس هذا المجتمع ذات الحجم n، وكذلك فإن توزيع المعاينة للتباين هوتوزيع جميع التباينات المحسوبة من عينات لها نفس الحجم n ومأخوذة من نفس المجتمع، وهكذا....

توزيعات المعاينة للأوساط: Sampling Distributions of Means

نفرض أننا سحبنا عينه حجمها n من مجتمع لانهائي، القيمة المتوقعة له تساوي μ والانحراف المعياري هو σ فإن المتوسط الحسابي $\overline{X}$ يخضع لتوزيع ما، متوسط هذا التوزيع وانحرافه المعياري هما

$$\mu_{\overline{X}} = \mu \quad ، \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

وفي الحالة التي يكون فيها المجتمع الأصلي المسحوبة منه العينة مجتمع طبيعي (ويرمز له بالرمز $N(\mu, \sigma^2)$ فإن توزيع المعاينة للمتوسط الحسابي $\bar{X}$ يكون في هذه الحالة توزيع طبيعي أيضاً له نفس المتوسط الأصلي μ ولكن انحرافه المعياري يساوي $\sigma/\sqrt{n}$، أي بمعنى أن

$$X \sim N(\mu,\sigma^2) \quad \Rightarrow \quad \bar{X} \sim N(\mu,\frac{\sigma^2}{n})$$

ومن ثم يكون

$$z = \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \sim N(0,1)$$

أما إذا كان المجتمع غير طبيعي فإن $\bar{X}$ لا تخضع للتوزيع الطبيعي ولكنها تتوزع توزيع يكون قريباً من التوزيع الطبيعي لقيم n الكبيرة $(n \geq 30)$ حيث أن

$$z = \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \xrightarrow[n \to \infty]{as} N(0,1)$$

وتعتبر النتيجة السابقة الهامة جداً في الإحصاء وخاصة في التطبيقات العلمية وتسمى نظرية النهاية المركزية Central Limit Theorem والتي تنص على أنه في حالة العينات الكبيرة الحجم فإن المتوسط الحسابي $\bar{X}$ يخضع للتوزيع الطبيعي بالمعاملات μ و $\frac{\sigma^2}{n}$، حيث أن μ, σ^2 هما متوسط وتباين المجتمع الأصلي بغض النظر عن شكل توزيع المجتمع الأصلي.
ومن ثم فإنه لقيم n الكبيرة تتحقق العلاقة بصرف النظر عن توزيع المجتمع الأصلي.

كذلك فإنه إذا كان $\bar{X}_1$ هوالمتوسط الحسابي لعينه عشوائية مسحوبة من مجتمع لانهائي متوسطه هو μ_1 وانحرافه المعياري هو σ_1، وكان $\bar{X}_2$ هوالمتوسط الحسابي لعينة عشوائية مسحوبة من مجتمع لانهائي آخر متوسط μ_2 وانحرافه المعياري σ_2 وكانت العينتين مستقلتين فإن المجموع الجبري لمتوسط العينتين يخضع لتوزيع المعاينة بالمعاملات

$$\mu_{(\bar{x}\pm\bar{x}_2)} = \mu_1 \pm \mu_2 \quad \text{and} \quad \sigma^2_{(\bar{x}_1\pm\bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

حيث n_1, n_2 هما حجم العينة الأولى والثانية.

وإذا كان المجتمعين الأصـــليين طبيعيين فإن $(\bar{x}_1 \pm \bar{x}_2)$ يخضـــع لتوزيع طبيعي أيضـــاً بالبارامترات المعطاة وعليه فإنه في هذه الحالة

$$z = \frac{(\bar{x}_1 \pm \bar{x}_2) - (\mu_1 \pm \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

أما إذا كـان أحـد المجتمعين أوكليهمـا لا يتوزع توزيعاً طبيعياً فإن $(\bar{x}_1 \pm \bar{x}_2)$ لا يتوزع توزيعاً طبيعياً كذلك، ولكن لقيم n_1, n_2 الكبيرة فإنه طبقاً لنظرية النهاية المركزية الســـابقة فإن $(\bar{x}_1 \pm \bar{x}_2)$ يتوزع توزيعاً قريباً من التوزيع الطبيعي وبذلك يمكننا اســـتخدام نفس العلاقة في حالة العينات الكبيرة.

توزيع المعاينة للتباين: Sampling Distribution of The Variance

إذا كان $s^2 = \dfrac{\sum (x_i - \bar{x})^2}{n-1}$ هوتباين عينه عشـــوائية حجمها n مأخوذة من مجتمع متوسطه μ وتباينه σ^2 وعزمه الرابع حول المتوسط هو μ_4 فإن

$$\mu_{S^2} = \sigma^2 \qquad \text{and} \qquad \sigma^2{}_{S^2} = \frac{\mu_4 - \sigma^2}{n-1}$$

وإذا كان المجتمع طبيعي فإن $\mu_4 = 3\sigma^2$ وبالتالي فإن

$$\sigma_{s^2}^2 = \left(\frac{2}{n-1}\right)\sigma^2$$

نلاحظ هنا أن s^2 لا تتوزع طبيعي حتى ولوكان المجتمع طبيعي، ولكنه يتوزع توزيع قريب من التوزيع الطبيعي وذلك لقيم n الكبيرة $(n \geq 100)$. أما إن كان المجتمع الأصـــلي يخضـــع للتوزيع الطبيعي فإن المتغير $(n-1)s^2/\sigma^2$ يخضـــع لتوزيع يســـمى توزيع مربع كاي χ^2 بعدد درجات حرية يساوي $n-1$. أي أن

$$\frac{(n-1)s^2}{\sigma^2} \quad \sim \quad \chi^2(n-1)$$

ويعتبر توزيع مربع كاي من التوزيعات الهامة في الإحصاء التطبيقي ودالة كثافته هي

$$f(y) \;=\; y^{\frac{v-1}{2}} \; e^{-y/2} \quad , \quad y>0$$

حيث v هي عدد درجات الحريه للتوزيع وتعتبر هي المعامل الوحيد له ويتضـــح من شـــكل الدالة أنها داله متصلة وتقع بأكملها فوق النصف الموجب لمحور السينات، منحنى هذه الداله غير متماثل ويعتبر من المنحنيات موجبة الالتواء ويقل التواؤه (وبالتالي يقترب من التماثل) كما زادت درجات الحريه v. وتكون القيمة المتوقعة لهذا التوزيع هي v وتباينه هو $2v$ أي بمعنى أن

$$E(y)=\mu_v=v$$

$$V(y)=\sigma^2=2v$$

فإذا كان s_1^2 هوتباين عينه عشـــوائية حجمها n_1 مســـحوبة من مجتمع طبيعي $N(\mu_1,\sigma_1^2)$، وكان s_2^2 هوتباين عينه عشوائية أخرى حجمها n_2 ومسحوبة من مجتمع طبيعي آخر $N(\mu_2,\sigma_2^2)$ وكانت العينتان مستقلتان فإن المتغير:

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \quad \sim \quad F(n_1-1,n_2-1)$$

حيث أن $F(n_1-1,n_2-1)$ تســـمى بتوزيع F بدرجتي الحريه n_1-1 و n_2-1 ودالة الكثافة الإحتماليه للمتغير y الذي يخضع لتوزيع F بدرجتي الحريه v_1,v_2 تعطى بالصورة:

$$f(y) \;=\; \frac{y^{\frac{v_1}{2}-1}}{(v_1y+v_2)^{\frac{v_1+v_2}{2}}} \quad , \quad y>0$$

وكما يتضح من الداله في أن المنحنى يقع بالكامل في النصف الموجب لمحور السينات كما في حالة توزيع χ^2، وهوأيضــاً غير متماثل وموجب الالتواء ولكن يقترب من التماثل كلما زادت درجات الحريه v_1,v_2.

ذكرنا سـابقاً أنه إذا كان $\bar{X}$ هوالمتوسـط الحسـابي لعينه حجمها n مأخوذة من مجتمع طبيعي بالمعاملات μ,σ^2 فإن

$$z=\frac{\sqrt{n}\,(\bar{x}-\mu)}{} \quad \sim \quad N(0,1)$$

هذا إذا كانت σ معلومة، ولكن في حالة ما إذا كانت قيمة σ غير معلومة فإننا نستخدم بدلا منها الانحراف المعياري للعينة s ولكن في هذه الحالة يصبح المتغير $\frac{\sqrt{n}\,(\bar{x}-\mu)}{s}$ يخضع لتوزيع يعرف بتوزيع t ستيودنت $t-$student بدرجات حريه $n-1$، أي أن

$$t=\frac{\sqrt{n}\,(\bar{x}-\mu)}{s} \quad \sim \quad t(n-1)$$

دالة الكثافة لتوزيع t بدرجات حريه v تعطي بالصورة:

$$f(t)= \quad (1+\frac{t^2}{v})^{-\frac{v+1}{2}}$$

وهوتوزيع متماثل حول محور y وهويشـــبه في ذلك المنحنى الطبيعي القياســـي $N(0,1)$ ولكنه أقل تحدباً من التوزيع الطبيعي القياسـي ولكنه يقترب من التوزيع الطبيعي كلما زادت درجات الحريه.

وإذا كان $\bar{X}_2$ و S_1^2 هما المتوســـط الحســـابي والتباين لعينه حجمها n_1 مأخوذة من مجتمع طبيعي متوسط هو μ_1 وكان $\bar{X}_2$ و S_2^2 هما المتوسط الحسابي والتباين لعينه أخرى حجمها n_2 ومأخوذة من مجتمع طبيعي آخر له المتوسط μ_2 وكانت العينتان مستقلتان فإن المتغير

$$t=\frac{(\bar{X}_2 \pm \bar{X}_2)-(\mu_1 \pm \mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

حيث أن $S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ يسمى بالتباين المشترك للعينتين The Pooled Variance.

سلسلة تمارين حول المعاينة

تمريــن 1:

إذا كان لدينا مجتمع يتكون من المفردات التالية: 2, 1, 3

المطلــوب:

1. إيجاد متوسط وتباين المجتمع.
2. تحديد عدد العينات العشــوائية البســيطة من الحجم 2 التي يمكن ســحبها من هذا المجتمع في كل من الحالات التالية:

- إذا كان السحب بإرجاع. - إذا كان السحب بدون إرجاع.

3. أوجد توزيع المعاينة للوسط الحسابي في الحالات التالية:

- إذا كان السحب بإرجاع. - إذا كان السحب بدون إرجاع.

4. أوجد متوســط وتباين الوســط الحســابي للعينة باســتخدام توزيع المعاينة في الحالات السابق تحديدها في السؤال الثالث مع التحقق من الإجابة المتحصل عليها.

تمريــن 2:

إذا كان العدد الكلي للمصانع في مدينة ما 5 مصانع والبيانات التالية تمثل عدد العاملين في هذه المصانع.

$X_1 = 40$ ، $X_2 = 42$ ، $X_3 = 48$ ، $X_4 = 56$ ، $X_5 = 64$

المطلـوب:

- أوجد توزيع المجتمع، ثم احسب الوسط الحسابي للمجتمع μ وتباين المجتمع δ^2
- أوجد كل العينات $n = 2$ الممكن سحبها في حالة السحب مع الارجاع.
- أوجد توزيع المعاينة للوسط الحسابي للعينة، ثم احسب وسطه الحسابي، وتباينه.

تمريــن 3:

نفس التمرين السابق مع طريقة السحب مع عدم الإرجاع

تمريــن 4:

في دراسة لأرصدة عملاء بنك ما تبين أنها تتبع التوزيع الطبيعي بـ $\mu = 13600DA$ و $\sigma = 600DA$ إذا قمنا بسحب 60 عينة حجم كل منها 9 حسابات من مجموع الحسابات المفتوحة وعددها. 6000

- أحسب $\mu_{\bar{X}}$ و $\sigma_{\bar{X}}$ في الحالتين : إذا كان السحب بإرجاع وإذا كان السحب بدون إرجاع ؟
- ما هي نسبة وعدد العينات التي يكون فيها $\bar{X}$ محصورا بين 13600 و 13800؟ أقل من 13800 ؟

تمريــن 5:

إذا كان لدينا متغير عشوائي X يتبع التوزيع الطبيعي بمتوسط $\mu_X = 80$ وتباين $\sigma_X^2 = 49$، أوجد توزيع المعاينة لمتوسط العينة من الحجم 25 مسحوبة من المجتمع من الحجم 1000، ثم أوجد $P(\bar{X} \geq 78)$.

تمريــن 6:

إذا كان عدد عاملي مصنع ما 1500 وعلمت أن أعمارهم تتوزع توزيعا طبيعيا بمتوسط قدره 45 سنة، وانحراف معياري قدره 7 سنوات، فإذا سحبنا مع عدم الإرجاع من هذا المجتمع عينة بها 16 عاملا فما احتمال أن يكون متوسط العمر لهذه العينة أكبر من 48 سنة ؟

تمريــن 7:

إذا علمت أن درجات 420 طالبا في امتحان في مادة الإحصاء تتوزع توزيعا طبيعيا بمتوسط قدره 68 وتباين قدره 25، فإذا سحبنا من هؤلاء الطلبة مع عدم الإرجاع عينة عشوائية تشمل 100 طالب، فما احتمال أن يتراوح الوسط الحسابي لهذه العينة بين 67 و69 درجة ؟

تمريــن 8:

أنتج مصنع للتونة 5000 علبة في الشهر وكان متوسط وزن العلبة 223 غراما بانحراف معياري 2.25 غرام، فإذا سحبنا من هذا الانتاج عينة عشوائية تحتوي عل 100 علبة مع عدم الارجاع، فما احتمال أن يكون الوسط الحسابي لأوزان العينة أقل من 222.5 غرام؟

تمريـــن 9:

إذا كان درجات طلبة الجامعة لمقياس الذكاء تتبع توزيع طبيعي بمتوســـط $\mu = 100$ وانحراف معياري $\sigma = 75$، تم اختيار25 طالباً عشـــوائيا وبدون إرجاع من بين طلبة الجامعة، أوجد التالي:

1. إذا كان عدد الطلبة المســـجلين في الجامعة 60000 طالب، ما هوعدد الطلبة الذين تزيد درجة مقياس ذكائهم عن 130.
2. عدد الطلبة في العينة المتوقع أن تزيد درجة مقياس ذكائهم عن 130.
3. احتمال أن يكون متوسط الذكاء المحسوب من العينة أكبر من.125
4. احتمال أن يكون متوسط الذكاء المحسوب من العينة أصغر من 80
5. احتمال أن يكون متوسط الذكاء المحسوب من العينة محصورا بين 70 و130
6. سوف يتم إعداد برنامج خاص للطلبة الذين يشكلون نسبة الخمسة في المائة الأولى لمقياس الذكاء، ما هي درجة مقياس الذكاء المقابلة لهذه النسبة؟

الحـــــــــــــــل

تمريـــن 1:

1. لإيجاد متوسط وتباين المجتمع:

نوجد أولاً مجموع قيم المفردات ومجموع مربعات المفردات (البيانات في صـــورة جدول كما يلي):

i	X_i	X_i^2
1	2	4

2	1	1
3	3	9
المجموع	$\sum X_i = 6$	$\sum X_i^2 = 14$

من هذه البيانات نحسب الوسط الحسابي للمجتمع كما يلي: $\mu = \frac{\sum X_i}{N} = \frac{6}{3} = 2$

بينما تباين المجتمع يحسب كما يلي: $\sigma^2 = \frac{1}{N}\left(\sum X_i^2 - N\bar{X}^2\right) = \frac{1}{3}\left(14 - \frac{6^2}{3}\right) = \frac{2}{3}$

2. عدد العينات الممكنة: سنرمز لعدد العينات الممكنة بالرمز M:

- إذا كان السحب بإرجاع: $M = N^n = 3^2 = 9$
- إذا كان السحب بدون إرجاع: $M = C_N^n = \frac{N!}{n!(N-n)!} = \frac{3!}{2!(3-2)!} = \frac{6}{2} = 3$

3. توزيع المعاينة للوسط الحسابي:

- في حالة السحب بدون إرجاع: نوجد أولاً جميع العينات الممكنة في هذه الحالة، ثم نحسب الوسط الحسابي لكل عينة كما يلي: حيث أن عدد العينات الممكن سحبها هي:

$\bar{x}$	1.5	2	2.5
العينة	(1.2)	(1.3)	(2.3)

توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ هو:

$\bar{x}$	1.5	2	2.5
$f(\bar{x})$	1/3	1/3	1/3

- في حالة السحب بإرجاع: نوجد أولاً جميع العينات الممكنة، ثم نحسب الوسط الحسابي لكل عينة:

العينة	$\bar{X}$	العينة	$\bar{X}$
(2.3)	2.5	(1.1)	1
(3.1)	2	(1.2)	1.5
(3.2)	2.5	(1.3)	2
(3.3)	3	(2.1)	1.5
		(2.2)	2

توزيع المعاينة في هذه الحالة هو:

$\bar{x}$	1	1.5	2	2.5	3
$f(\bar{x})$	1/9	2/9	3/9	2/9	1/9

4. متوسط وتباين الوسط الحسابي للعينة:

- في حالة السحب بدون إرجاع:

$\bar{x}$	$f(\bar{x})$	$\bar{x}.f(\bar{x})$	$(\bar{x}-\mu_{\bar{x}})$	$(\bar{x}-\mu_{\bar{x}})^2$	$(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$
1.5	1/3	0.5	0.5-	0.25	0.0833
2	1/3	0.67	0	0	0
2.5	1/3	0.83	0.5	0.25	0.0833
$\sum$	*1*	2		0.5	0.1666

*متوسط التوزيع العيني لـ $\bar{X}$:

$$\mu_{\bar{X}} = \sum \bar{x} f(\bar{x}) = (1.5*\frac{1}{3}) + (2*\frac{1}{3}) + (2.5*\frac{1}{3}) = 2 = \mu$$

وهويساوي نفس قيمة μ كما يجب أن يكون.

*تباين التوزيع العيني لـ $\bar{X}$:

$$\delta^2{}_{\bar{X}} = \sum (\bar{x}-\mu_{\bar{X}})^2 f(\bar{x}) = (1.5-2)^2 * (\frac{1}{3})$$
$$+ (2-2)^2 * (\frac{1}{3}) + (2.5-2)^2 * (\frac{1}{3}) = \frac{0.5}{3} = 0.1666$$

وهي تساوي $\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$. **للتحـــــقق**: $\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{2/3}{2}\left(\frac{3-2}{3-1}\right) = 0.1666$

- في حالة السحب مع الإرجاع:

$\bar{x}$	$f(\bar{x})$	$\bar{x}.f(\bar{x})$	$(\bar{x}-\mu_{\bar{x}})$	$(\bar{x}-\mu_{\bar{x}})^2$	$(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$
1	1/9	1/9	1-	*1*	1/9
1.5	2/9	1/3	0.5-	*0.25*	0.5/9
2	3/9	2/3	0	0	0
2.5	2/9	5/9	0.5	0.25	0.5/9
3	1/9	3/9	1	1	1/9
$\sum$	*1*	2	0	2.5	1/3

*متوسط التوزيع العيني لـ $\bar{X}$:

$$\mu_{\bar{X}} = \sum \bar{x} f(\bar{x}) = (1*\frac{1}{9}) + (1.5*\frac{2}{9}) + (2*\frac{3}{9}) + (2.5*\frac{2}{9}) + (3*\frac{1}{9}) = 2 = \mu$$

وهويساوي نفس قيمة μ كما يجب أن يكون.

*تباين التوزيع العيني لـ $\bar{X}$:

$$\delta^2{}_{\bar{X}} = \sum(\bar{x}-\mu_{\bar{X}})^2 f(\bar{x}) = (1-2)^2*(\frac{1}{9}) + (1.5-2)^2*(\frac{2}{9})$$

$$+(2-2)^2*(\frac{3}{9}) + (2.5-2)^2*(\frac{2}{9}) + (3-2)^2*(\frac{1}{9}) = \frac{1}{3} = 0.333$$

وهي تساوي $\frac{\sigma^2}{n}$. **للتحـــــقق**: $\frac{\sigma^2}{n} = \frac{2}{3}\frac{1}{2} = \frac{2}{6} = \frac{1}{3} = 0.333 = \sigma^2_{\bar{X}}$

تمريــن 2:

*التوزيع الاحتمالي للمجتمع

X	40	42	48	56	64
$f(X)$	1/5	1/5	1/5	1/5	1/5

* الوسط الحسابي للمجتمع هو:

$$\mu = \sum x f(x) = (40*\frac{1}{5}) + (42*\frac{1}{5}) + (48*\frac{1}{5}) + (56*\frac{1}{5}) + (64*\frac{1}{5}) = 50$$

* تباين المجتمع هو:

$$\delta^2 = \sum x^2 f(x) - \mu^2 = \{(40)^2*(\frac{1}{5}) + (42)^2*(\frac{1}{5}) + (48)^2*(\frac{1}{5}) + (56)^2*(\frac{1}{5}) + (64)^2*(\frac{1}{5})\} - 50^2 = 80$$

أو

$$\delta^2 = \sum(x-\mu)^2 f(x) = (40-50)^2*(\frac{1}{5}) + (42-50)^2*(\frac{1}{5}) + (48-50)^2*(\frac{1}{5}) + (56-50)^2*(\frac{1}{5}) + (64-50)^2*(\frac{1}{5}) = 80$$

*عدد العينات الممكن سحبها في حالة السحب مع الإرجاع مع العلم أن $n = 2$ هو

$N^n = 5^2 = 25$ والجدول التالي يوضح العينات

العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$
(40.40)	40	(42.40)	41	(48.40)	44	(56.40)	48	(64.40)	52
(40.42)	41	(42.42)	42	(48.42)	45	(56.42)	49	(64.42)	53
(40.48)	44	(42.48)	45	(48.48)	48	(56.48)	52	(64.48)	56
(40.56)	48	(42.56)	49	(48.56)	52	(56.56)	56	(64.56)	60
(40.64)	52	(42.64)	53	(48.64)	56	(56.64)	60	(64.64)	64

* توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ في حالة السحب مع الإرجاع هو:

$\bar{x}$	40	41	42	44	45	48	49	52	53	56	60	64
$f(\bar{x})$	1/25	2/25	1/25	2/25	2/25	3/25	2/25	4/25	2/25	3/25	2/25	1/25

*حساب الوسط الحسابي $\mu_{\bar{x}}$ والتباين $\delta^2{}_{\bar{x}}$ لتوزيع المعاينة

$\bar{x}$	$f(\bar{x})$	$\bar{x}.f(\bar{x})$	$(\bar{x}-\mu_{\bar{x}})^2$	$(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$
40	1/25	40/25	100	100/25
41	2/25	82/25	81	162/25
42	1/25	42/25	64	64/25
44	2/25	88/25	36	72/25
45	2/25	90/25	25	50/25
49	3/25	144/25	4	12/25
48	2/25	98/25	1	2/25
52	4/25	208/25	4	16/25
53	2/25	106/25	9	18/25
56	3/25	168/25	36	108/25
60	2/25	120/25	100	200/25
64	1/25	64/25	196	196/25
$\sum$	1	1250/25	1000	1000/25

$$\mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = 50 = \mu$$

$$\delta^2{}_{\bar{x}} = \sum (\bar{x}-\mu_{\bar{x}})^2 f(\bar{x}) = \frac{1000}{25} = 40 = \frac{80}{2} = \frac{\delta^2}{n}$$

تمريــن 3:

*عدد العينات الممكن سحبها في حالة السحب مع عدم الإرجاع مع العلم أن $n=2$ هو

$C_N{}^n = C_5{}^2 = \frac{5!}{2!*(5-2)!} = \frac{5*4}{2} = 10$ والجدول التالي يوضح العينات:

العينة	$\bar{X}$	العينة	$\bar{X}$
(40.42)	41	(42.56)	49
(40.48)	44	(42.64)	53
(40.56)	48	(48.56)	52
(40.64)	52	(48.64)	56

60	(56.64)	45	(42.48)

* توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ في حالة السحب مع الإرجاع هو:

$\bar{x}$	41	44	45	48	49	52	53	56	60
$f(\bar{x})$	*1/10*	*1/10*	*1/10*	*1/10*	*1/10*	*/102*	*1/10*	*1/10*	*1/10*

*حساب الوسط الحسابي $\mu_{\bar{x}}$ والتباين لتوزيع المعاينة $\delta^2{}_{\bar{x}}$

$\bar{x}$	$f(\bar{x})$	$\bar{x}.f(\bar{x})$	$(\bar{x}-\mu_{\bar{x}})^2$	$(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$
41	0.1	*4.1*	81	8.1
44	0.1	4.4	36	3.6
45	0.1	4.5	25	2.5
48	0.1	4.8	4	0.4
49	0.1	4.9	1	0.1
52	0.2	10.4	4	0.8
53	0.1	5.3	9	0.9
56	0.1	5.6	36	3.6
60	0.1	6.0	100	10
$\sum$	1	50	296	30

$$\mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = 50 = \mu$$

$$\delta^2{}_{\bar{x}} = \sum (\bar{x}-\mu_{\bar{x}})^2 f(\bar{x}) = 30 = \frac{80}{2} * \frac{5-2}{5-1} = \frac{\delta^2}{n} * \frac{N-n}{N-1} = 30$$

تمريــن 4:

لدينا $n = 9$ و $N = 6000$ إذن $\frac{n}{N} = \frac{9}{6000} = 0.0015 \prec 0.05$

إذن إذا كان السحب مع الإرجاع أوبدون إرجاع فإن:

$$\mu_{\bar{x}} = \mu = 13600 \text{ و } \sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \frac{600}{3} = 200$$

$$P(13600 \le \overline{X} \le 13800) = P(\frac{13600-13600}{200} \le Z \le \frac{13800-13600}{200})$$

$$= P(0 \le Z \le 1) = 0.84134 - 0.5000 = 0.34134$$

عدد العينات هو: $0.34134*60 \approx 20$

$$P(\bar{X} \le 13800) = P(Z \le \frac{13800-13600}{200}) = P(Z \le 1) = 0.84134 = 0.84134$$

عدد العينات هو: $0.84134*60 \approx 50$

تمريـــن 5:

لدينا $n = 25$ و $N = 1000$ إذن $\frac{n}{N} = \frac{25}{1000} = 0.025 \prec 0.05$

إذن إذا كان السحب مع الإرجاع أوبدون إرجاع فإن:

$$\mu_{\bar{x}} = \mu_x = 80 \text{ و } \sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{49}{25}} = \frac{7}{5} = 1.4$$

بما أن المتغيرة X تتبع التوزيع الطبيعي فإن متوسط العينات $\bar{X}$ يتبع هوأيضا التوزيع الطبيعي حتما بغض النظر عن حجم العينة سواء كان صغيرا أم كبيرا.

إذن: $X \mapsto N(80,49) \Rightarrow \bar{X} \mapsto N(80, \frac{49}{25})$

- إيجاد $P(\bar{X} \ge 78) = ?$

$$P(\bar{X} \ge 78) = 1 - P(Z \le \frac{78-80}{7/5}) = 1 - P(Z \le 1.4285) = 1 - 0.92364 = 0.07636$$

تمريـــن 6:

المتغير العشوائي محل الدراسة X هوعمر العامل حيث $X \to N(45.49)$

والاحتمال المطلوب $P(\bar{X} \succ 48) = ?$

بما أن مجتمع أعمار العمال تتوزع توزيعا طبيعيا إذن توزيع المعاينة للوسط الحسابي للعينة سيكون توزيعا طبيعيا بغض النظر عن حجم العينة، وسيكون الوسط الحسابي والتباين لهذا التوزيع كما يلي:

$$\delta^2{}_{\bar{x}} = \frac{\delta^2}{n} = \frac{49}{16} \ \mu_{\bar{x}} = \mu = 45$$

لم نستخدم معامل التصحيح بالرغم من أن السحب كان مع عدم الإرجاع لأن:

$$\frac{n}{N} = \frac{16}{1500} = 0.01 \prec 0.05$$

نحسب القيمة المعيارية كما يلي: $Z = \frac{\overline{X} - \mu_{\overline{X}}}{\delta_{\overline{X}}} = \frac{48-45}{\sqrt{49/16}} = \frac{12}{7} = 1.71$

بالتالي فإن:

$$P(\overline{X} \succ 48) = P(Z \succ 1.71) = 0.5 - 0.4564 = 0.0436$$

أي أنه إذا سحبنا من هذا المجتمع مع عدم الارجاع كل العينات العشوائية الممكنة ذات

الحجم $(n = 16)$ ، فستكون نسبة العينات التي وسطها الحسابي أكبر من 48 هي

%4.36

تمريـــن 7:

المتغير العشوائي محل الدراسة X هودرجة الطالب حيث $X \to N(68.25)$

والاحتمال المطلوب هو ? = $P(67 \le \overline{X} \le 69)$

بما أن مجتمع درجة الطالب تتوزع توزيعا طبيعيا إذن توزيع المعاينة للوسط الحسابي للعينة

سيكون توزيعا طبيعيا، وسيكون الوسط الحسابي والتباين لهذا التوزيع كما يلي:

$$\mu_{\overline{x}} = \mu = 68$$

$$\delta^2{}_{\bar{x}} = \frac{\delta^2}{n}\frac{N-n}{N-1} = \frac{25}{100}\frac{420-100}{420-1} = 0.1909$$

اسخدمنا معامل التصحيح لأن السحب كان مع عدم الإرجاع لأن:

$$\frac{n}{N} = \frac{100}{420} = 0.24 \succ 0.05$$

نحسب القيمتين المعياريتين كما يلي:

$$Z_1 = \frac{\overline{X} - \mu_{\overline{X}}}{\delta_{\overline{X}}} = \frac{67-68}{\sqrt{0.1909}} = \frac{12}{7} = -2.29$$

$$Z_2 = \frac{\overline{X} - \mu_{\overline{X}}}{\delta_{\overline{X}}} = \frac{69-68}{\sqrt{0.1909}} = 2.29$$

بالتالي فإن:

$$P(67 \le \overline{X} \le 69) = P(-2.29 \le Z \le 2.29) = 0.4890 + 0.4890 = 0.9780$$

تمريـــن 8:

المتغير العشوائي محل الدراسة X هووزن علبة التونة والمجتمع هومجموعة أوزان كل العلب وتوزيعه الاحتمالي غير معروف وحجم العينة $(n=100)$ أي $(n \succ 30)$، إذن بالرغم من أن توزيع المجتمع مجهول فإن توزيع المعاينة للوسط الحسابي للعينة سيكون توزيعا قريبا من التوزيع الطبيعي (نظرية النهاية المركزية)، وسيكون الوسط الحسابي والتباين لهذا التوزيع كما يلي:

$$\delta^2_{\overline{x}} = \frac{\delta^2}{n} = \frac{(2.25)^2}{100} = 0.0506 \; \mu_{\overline{x}} = \mu = 223$$

لم نستخدم معامل التصحيح لأن السحب كان مع عدم الإرجاع لأن:

$$\frac{n}{N} = \frac{100}{50000} = 0.002 \prec 0.05$$

نحسب القيمة المعيارية كما يلي:

$$Z_1 = \frac{\overline{X} - \mu_{\overline{X}}}{\delta_{\overline{X}}} = \frac{222.5 - 223}{\sqrt{0.0506}} = \frac{-0.5}{0.225} = -2.22$$

بالتالي فإن:

$$P(\overline{X} \prec 222.5) = P(Z \prec -2.22) = 0.50.4868 = 0.0132$$

تمريـــن 9:

ســـنرمز لمقياس الذكاء بالرمز X، وبالتالي المتغير العشـــوائي $\overline{X}$ يتبع توزيع طبيعي بمتوســـط $\mu_{\overline{X}} = 100$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{75}{\sqrt{25}} = 15$ أي أن:

$$\overline{X} \sim N(100, 225)$$

والآن نوجد المطلوب:

1. عدد الطلبة في الجامعة الذين تزيد درجة مقياس ذكائهم عن 130.

نوجد أولاً نسبة أفراد المجتمع الذين تزيد درجة مقياس ذكائهم عن 130 والتي سنرمز لها بالرمز P_A كما يلي:

$$P_A = P(X > 130) = P\left(Z > \frac{130-100}{75}\right) = 1 - \Phi(0.4) = 1 - 0.6554 = 0.3446$$

إذا، عدد الطلبة في المجتمع الذين تزيد درجة مقياس ذكائهم عن 130 والتي سنرمز لها بالرمز N_A هي:

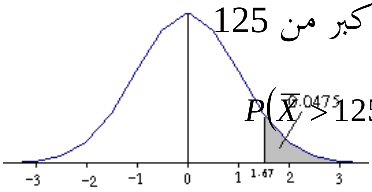
$$N_A = N\,P(X > 130) = (60000)(0.3446) = 20676$$

حيث N هي حجم المجتمع.

2. عدد الطلبة في العينة المتوقع أن يزيد مقياس ذكائهم عن 130: حيث n هي حجم العينة

سنرمز للعدد المتوقع في العينة بالرمز M: $M = n\,P(X > 130) = (25)(0.3446) = 8.615 \cong 9$

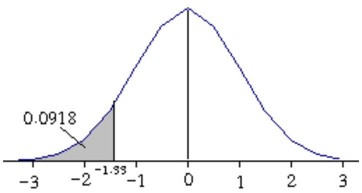
3. احتمال أن يكون متوسط الذكاء المحسوب من العينة أكبر من 125



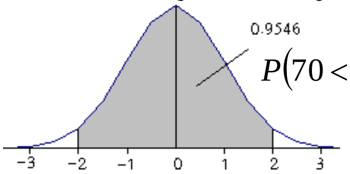
$$P(\bar{X} > 125) = P\left(Z > \frac{125-100}{15}\right) = 1 - \Phi(1.67) = 1 - .9525 = 0.0475$$

4. احتمال أن يكون متوسط الذكاء المحسوب من العينة أصغر من 80

$$P(\bar{X} < 80) = P\left(Z < \frac{80-100}{15}\right) = \Phi(-1.33) = 0.0918$$



5. احتمال أن يكون متوسط الذكاء المحسوب من العينة محصورا بين 70 و130



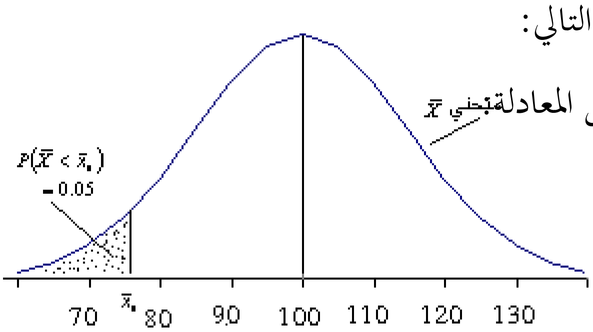
$$P(70 < \bar{X} < 130) = P\left(\frac{70-100}{15} < Z < \frac{130-100}{15}\right) = P(-2 < Z < 2)$$

$$= \Phi(2) - \Phi(-2) = 0.97725 - 0.02275 = 0.9545$$

6. درجة مقياس الذكاء المقابلة لنسبة الخمسة في المائة الأولي؟

أي نريد إيجاد القيمة $\bar{x}_\circ$ للمتغير $\bar{X}$ التي تحقق المعادلة $P(\bar{X} < \bar{x}_\circ) = 0.05$ كما هوموضح في الشكل التالي: $P(\bar{X} < \bar{x}_\circ) = 0.05$

أي $P\left(Z < \frac{\bar{x}_\circ - 100}{15}\right) = 0.05$ نريد حل المعادلة



من جدول التوزيع الطبيعي نجد أن:

$$P(Z < -1.645) = 0.05$$

أي أن: $\frac{\bar{x}_{\circ} - 100}{15} = -1.645$

ومنها نجد أن: $\bar{x}_{\circ} = 100 - (15)(1.645) = 75.325$

وهذا يعني أن نســـبة الطلبة في الجامعة الذين تقل درجة مقياس الذكاء لهم عن 75.325 تساوي 5% من أجمالي الطلبة.

تمريــن 1:

لدينا مجتمعين A و B حيث $A = \{1.3.4\}$ و $B = \{2.5\}$

المطلــوب:

5. كوّن توزيع المعاينة D للفرق $D = B - A$

6. أحسب المتوسط لـ D و A و B ، ماذا تستنتج؟

7. أحسب التباين لـ D و A و B ، ماذا تستنتج؟

8. كوّن توزيع المعاينة S للمجموع $S = B + A$

9. أحسب المتوسط لـ S ، ماذا تستنتج؟

10. أحسب التباين لـ S ، ماذا تستنتج؟

تمريــن 2:

لدينا مجتمع يتكون من المفردات التالية $U_1 = \{3.7.8\}$ والمجتمع $U_2 = \{2.4\}$

تحقق أن :

$$\mu_{U_1 - U_2} = \mu_{U_1} - \mu_{U_2}$$

$$\sigma^2_{U_1 - U_2} = \sigma^2_{U_1} + \sigma^2_{U_2}$$

تمريــن 3:

متوسط مدة حياة مصباح من إنتاج مصنع A 1400 ساعة بانحراف معياري 200 ساعة، ومن إنتاج مصنع B 1200 ساعة بانحراف معياري 100 ساعة، نكوّن عينتين من إنتاج المصنعين حجم كل منهما 125

- أحسب الفرق المتوقع بين متوسط عمر المصباح في العينتين.
- أحسب الانحراف المعياري لهذا الفرق.

تمريــن 4:

في سباق العدوبالتناوب تقدم فريق من ثلاث لاعبين متوسط سرعة كل منهم دقيقتين و30 ثانية بانحراف معياري 15 ثانية.

- أحسب المدة المتوقعة للفريق في السباق.
- أحسب الإنحراف المعياري.

تمريــن 5:

ليكن المجتمع 1.3.5.6.10 نسحب منه عينة بالإرجاع وبدون إرجاع مكونة من مفردتين.

- أحسب تباين المجتمع.
- أحسب القيمة المتوقعة لتباين العينة المسحوبة بالإرجاع وبدون إرجاع من خلال متوسط تباينات العينات الممكنة.
- قارن بين تباين المجتمع والقيمة المتوقعة لتباين العينة.

تمريــن 6:

حصل مرشح معين على 46 بالمائة من الأصوات في انتخاب معين. ما هواحتمال أن يحصل المرشح

على الأغلبية في مكتب ما محدد عشوائيا؟

- إذا كان عدد الناخبين في المكتب 200.
- إذا كان عدد الناخبين في المكتب 1000.

تمريــن 7:

أوجد احتمال بأنه ضمن 200 إصابة في حوادث المرور على الطريق يوجد:

- أقل من 30 بالمئة حادث للذكور.
- أكثر من 80 بالمئة حادث للذكور.
- بين 40 و60 بالمئة حادث للذكور.

وذلك علما أنه في كل خمسة أشخاص مصابين توجد امرأة.

تمريــن 8:

من بين وحدات القطع المنتجة في مصنع توجد 3 بالمئة فاسدة، يقوم مشتري باستلام علبة تحتوي 500 قطعة مباشرة من المصنع

- ما هواحتمال أن يجد على الأقل 1 بالمئة قطعة فاسدة في العلبة.
- ما هواحتمال أن يجد على الأكثر 5 بالمئة قطعة فاسدة في العلبة.
- ما هواحتمال أن يجد على الأكثر قطعة فاسدة في العلبة.