



وزارة التعليم العالي والبحث العلمي

جامعة البغليالي بونعامه خميس مليانة

كلية العلوم الاقتصادية، التجارية وعلوم التسيير



مطبوعة

في مقياس

إحصاء 3

موجهة لطلبة السنة الثانية ليسانس

مدعمة بأمثلة تطبيقية

من إعداد الدكتور:

بوكريطة عبد القادر

السنة الجامعية

2021-2020

البسمة

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

قائمة المحتويات

قائمة المحتويات

1 العينات

2 1- مصطلحات ومفاهيم

6 2- أساليب جمع البيانات الاحصائية

6 1-2 الحصر الشامل:

7 2-2 المعاينة:

8 3- أقسام المعاينة

8 1-3 المعاينة العشوائية

8 1-1-3 العينة العشوائية البسيطة

13 2-1-3 العينة العشوائية المنتظمة

13 3-1-3 العينة الطبقية العشوائية

15 4-1-3 العينة العنقودية

16 2-3 العينات غير العشوائية:

16 1-2-3 العينة الميسرة:

16 2-2-3 العينة العمدية أو القصدية:

17 3-2-3 العينة الحصصية:

18 نظرية المعاينة

19 1- توزيعات المعاينة للأوساط:

19 1-1- متوسط الوسط الحسابي للعينة $\mu_{\bar{X}}$:

21 2-1- تباين الوسط الحسابي للعينة $\sigma_{\bar{X}}^2$

23 3-1- طبيعة توزيع الوسط الحسابي للعينة $\bar{X}$

24 2- توزيعات المعاينة لتباين العينات:

24 1-2- متوسط تباينات العينات $E(S_{\bar{X}}^2)$:

قائمة المحتويات

26	2-2- تباين تباينات العينات $V(S_X^2)$
26	2-3- طبيعة توزيع المعاينة للتباين S_X^2 :
27	2-4- توزيع النسبة بين تبايني عينتين :
28	3- توزيعات المعاينة للفروق والمجاميع
28	3-1- توزيعات المعاينة للفروق والمجاميع للمتوسط :
29	3-2- توزيعات المعاينة للفروق والمجاميع للتباين :
30	3-3- توزيع الفرق بين وسطين
31	4- توزيعات المعاينة للنسبة وللفرق بين نسبتيين
31	4-1- توزيع المعاينة للنسبة
33	4-2- توزيع المعاينة للفرق بين نسبتيين
34	نظرية التقدير
35	1- التقدير بنقطة
36	1-1- طريقة الإمكان الأكبر
42	1-2- طريقة العزوم
45	1-3- طريقة المربعات الصغرى
49	2- التقدير بمجال
50	2-1- مجال الثقة للمتوسط الحسابي للمجتمع μ_X
50	2-1-1- مجال الثقة للمتوسط الحسابي في حالة توزيع المجتمع طبيعي
50	2-1-1-1- حالة تباين المجتمع معلوماً (σ_X^2 معلومة)
54	2-1-1-2- حالة تباين المجتمع غير معلوم (σ_X^2 مجهولة)
54	الحالة الأولى: العينات الصغيرة
58	الحالة الثانية: العينات الكبيرة
60	2-1-2- مجال الثقة للمتوسط الحسابي في حالة المجتمع غير طبيعي أو مجهول التوزيع
60	2-1-2-1- حالة تباين المجتمع معلوماً (σ_X^2 معلومة)

قائمة المحتويات

60	الحالة الأولى: العينات الصغيرة
63	الحالة الثانية: العينات الكبيرة
64	2-2-1-2 حالة تباين المجتمع غير معلوم (σ_X^2 مجهولة)
66	2-2 مجال الثقة لتباين المجتمع σ_X^2 :
66	1-2-2 متوسط المجتمع معلوما (μ_X معلوم)
67	2-2-2 إذا كان متوسط المجتمع مجهولا (μ_X مجهول)
69	3-2 مجال الثقة للنسبة
72	نظرية إختبار الفرضيات
73	1- مفاهيم أساسية في نظرية الإختبار
75	2- إختبارات الفرضيات حول معالم المجتمع الإحصائي
76	1-2-1 إختبارات الفرضيات حول المتوسط الحسابي لمجتمع طبيعي μ_X
76	1-1-2 تباين المجتمع معلوم σ_X^2
77	2-1-2 تباين المجتمع غير معلوم أي مجهول σ_X^2
77	1-2-1-2 العينات الصغيرة
79	2-2-1-2 العينات الكبيرة
80	2-2 إختبارات الفرضيات حول مقارنة وسطين حسابيين μ_X
81	1-2-2 العينتان كبيرتان: ($n_1 \geq 30 : n_2 \geq 30$)
81	1-1-2-2 الانحرافان المعياريان للمجتمعان معلومان
82	2-1-2-2 الانحرافان المعياريان للمجتمعان متساويان و مجهولان $\sigma_1 = \sigma_2 = \sigma$
84	3-1-2-2 الانحرافان المعياريان للمجتمعان غير متساويان و مجهولان
85	2-2-2 العينتان صغيرتان ($n_1 < 30 : n_2 < 30$)
86	3- إختبار الفرضيات المتعلقة بالنسبة والفرق بين نسبتيين
86	1-3-1 إختبار الفرضيات يتعلق بالنسبة P
88	2-3-2 إختبار الفرضيات يتعلق بالفرق بين نسبتيين

قائمة المحتويات

4- إختبارات تتعلق بالتباين والمقارنة بين تباينين _____ 91

1-4 إختبارات تتعلق بالتباين _____ 91

2-4 إختبارات تتعلق بمقارنة تباينين. _____ 92

قائمة المراجع. _____ 94

المحور الأول

العينات

S a m p l e s

العينات

Samples

من أهم أهداف البحث بطريقة العينات هو الحصول على أكبر قسط من الدقة بأقل قدر من التكاليف ولذلك فإن الباحث يجب أن يوازن بين مستوى الدقة الذي يرغب في الحصول عليه والتكاليف اللازمة للحصول على هذه الدرجة من الدقة.

فالمعينة هي أن نأخذ مجموعة صغيرة من المفردات تسمى "العينة" من المجموعة الكبيرة التي نود بحثها والتي تسمى "المجتمع الأصلي"، وخواص العينة هذه هي تقريبا نفس خواص المجتمع الأصلي (الأطرقجي، 1980، صفحة 211).

1- مصطلحات ومفاهيم

مفهوم علم الإحصاء

يشار لعلم الإحصاء بأنه مجموعة الطرق العلمية القياسية التي يمكن توظيفها لجمع البيانات والمعلومات والمعطيات الإحصائية عن الظواهر، وتبويبها وتلخيصها وتقييمها والخروج من خلالها باستنتاجات حول مجموع وحدات المجتمع اعتمادا على جزءا صغير من هذا المجتمع، وهذا الجزء يدعى بالعينة.

مجتمع الدراسة: Population

هو مجموعة من المفردات تشترك في صفه أو صفات وتكون موضوع دراسة أو بحث فإن هذه المجموعة يطلق عليها إحصائيا مجتمع الدراسة أو اختصاراً المجتمع ويرمز لحجمه عادة بالرمز N ، والمجتمع قد يكون مجموعة ما من البشر أو أشجار أو أنواع معينة من الفاكهة أو الحيوانات أو إنتاج دولة ما لسلع معينة خلال فترة زمنية محددة... الخ. ويمكن تصنيف المجتمعات إلى نوعين:

المجتمع المحدود:

يكون المجتمع محدوداً إذا كان يمكن حصر عدد أفراده مثل سكان مدينة ما أو طلاب مرحلة تعليمية معينة.

المجتمع غير المحدود:

قد يكون المجتمع غير محدود (لانهائي) إذا كان لا يمكن حصر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار، كما في حالة عدد الملاحظات أو التجارب العلمية أو عدد المحاضرات التي تلقى في المدارس في كافة أنحاء العالم.

الفرد:

تطلق كلمة فرد في علم الإحصاء للدلالة على الكائن الواحد سواء أكان هذا الكائن إنسانا أو حيوانا أو شيئا، وهو الذي يرجع في أصله إلى مجموعة من الأفراد المشابهة له في المظهر،

العيينة:

العيينة و نقصد بها مجموعة المفردات التي تشكل جزء من المجتمع الإحصائي يتم اختيارها وفق قواعد وأصول معينة تسمى أساليب المعاينة (هرمز، 1990، صفحة 471)، ثم بعد ذلك يتم تعميم نتائج الدراسة على المجتمع بأكمله

والعيينة تطلق على ذلك الجمع الذي يضم عددا كبيرا أو قليلا من الأفراد المتغيرة في الشكل أو اللون القياس لكنها تعود إلى أصل واحد وهي متشابهة في إحدى الصفات على الأقل مثل مجموعة من التلاميذ، مجموعة من الأفلام... إلخ، (حليمي، 1985، صفحة 19)

المتغير Variable

هو مقدار كمي أو وصفي ويستخدم لقياس خاصية أو مميزة معينة لعناصر المجتمع أو العينة، من الأمثلة عليه الطول، عدد الأطفال، فصيلة دم الشخص، المستوى التعليمي... إلخ.

المعلمة parameter

هي شيء يميز المجتمع الإحصائي كله مثل متوسط الدخل الشهري للأسر في دولة معينة... إلخ، المعلمة هي قيم محددة وثابتة

الإحصاء statistics

هي سمة أو خاصية رقمية تصف عينة من المجتمع مثل متوسط الدخل الشهري لعينة مكونة من 100 أسرة في دولة ما... إلخ، الإحصاء هي قيم غير ثابتة لأن عينات مختلفة تنتج إحصائيات مختلفة أصناف المعطيات الإحصائية:

تصنف المعطيات الإحصائية إلى صنفين رئيسين هما:

أ- المعطيات الكمية Quantitative data

وهي التي تعبر بشكل رقمي عن ظاهرة معينة، ويطلق عليها أحيانا بالمعطيات المقاسة *measured data* ، وتمثل أية نشاط أو فعالية على وفق المقدار المنجز ، فنقيس الانتاج بالطن أو الكيلو وأجزائه وما شابه ، والتعبير عن السعر بالدينار أو الدولار وأجزائها وعن الزمن بالساعة والدقيقة... إلخ .

وعندما تشتمل قيم هذه المعطيات على كسور يطلق عليها بالمتغيرات المستمرة أو المتصلة *Continuous* أما عندما تكون قيم المعطيات عبارة عن أعداد صحيحة من دون كسور فتسمى بالمتغيرات المتقطعة *discrete* *variables* حيث يكون تمثيلها بيانياً عبارة عن نقاط منفصلة.

ب- المعطيات النوعية *Qualitative data*

وهي المعطيات التي تصف ظاهرة معينة بشكل غير رقمي كالجنس (ذكور - إناث)، التحصيل الدراسي (دكتوراه - ماجستير - بكالوريا... الخ)، كما ويمكن تنظيم وحدات الظاهرة حسب اشتراكها في الصفة مثل ممتاز، جيد جداً، جيد... الخ.

أقسام المعطيات الإحصائية:

علم الاحصاء ينقسم إلى نوعين هما:

أ- الاحصاء الوصفي *Descriptive Statistics*

وهو ما يتعلق بطرق جمع وتحليل المعطيات ووصفها لتكون بصيغة ذات مدلول من دون التعامل مع تعميم النتائج.

ب- الاحصاء الاستدلالي *Inferential Statistics*

ويختص بطرق تحليل وتفسير واستخلاص الاستنتاجات بالاعتماد على جزء (عينة) من المجتمع للتوصل إلى قرارات تخص المجتمع الإحصائي، وعليه فإن الاحصاء الاستدلالي يتعامل مع التعميم والتنبؤ والتقدير، وتتسم الاستنتاجات في بعض الحالات بعدم التأكد عندها يتم قياسها باستخدام الاحتمالات (البلداوي، 2009، صفحة

(17)

أساليب اختيار العينة العشوائية *Random Selection Method*

أ- الاختيار بالارجاع: (*Selection With Replacement*)

وهو يعني أننا حين نختار مفردة من المجتمع فإننا نعيدها ثانية إلى المجتمع ليتم اختيار المفردة الثانية، وقد تظهر المفردة نفسها أو غيرها .

ب- الاختيار بدون إرجاع (*Selection Without Replacement*)

وهو يعني أنه عند اختيارنا للمفردة الأولى فإننا لا نلجأ إلى إعادة ثانيها إلى المجتمع وإنما نختار مفردة مما تبقى من المجتمع وهكذا، (البلداوي، 2009، صفحة 24) ومن الناحية العملية فإن جميع مسوحات العينة تعتمد على أسلوب الاختيار بدون إرجاع ، لذا سيكون التركيز على هذا الأسلوب في دراستنا للعينات .

وعند دراسة صفة ما أو صفات معينه لمجتمع ما، فإن البيانات الإحصائية عن تلك الصفة أو الصفات تجمع بأحد أسلوبين

وسائل جمع البيانات

- أ- الاستبيان: أداة لجمع البيانات عن الظاهرة أو المشكلة المراد بحثها (أسئلة تقدم الى المختبر)
- ب- المقابلة: من أفضل الوسائل (كون طبيعة الافراد الرغبة في التحدث أكثر من الكتابة)
- ج- الملاحظة: من الوسائل المهمة التي يتم بموجبها جمع البيانات يحصل عليها الباحث بالفحص المباشر
- د- الاختبارات والمقاييس: تعد من الوسائل الاساسية والمهمة وهي عبارة عن ادوات صممت لوصف وقياس عينة من افراد المجتمع (مقاييس اللياقة البدنية الحركية والاتجاهات النفسية)

مصادر جمع البيانات الاحصائية

يمكن الحصول على المعلومات من مصدرين هما:

أولاً: المصدر المباشر (الميداني)

وهو عبارة عن البيانات المجموعة من أفراد المجتمع الاحصائي كله أو جزء منه (عينة إحصائية) بالاتصال المباشر (المقابلة) أو غير المباشر مثل البريد أو الهاتف أو استخدام كلا الطريقتين حسب طبيعة المشكلة محل الدراسة (عبد الرحمن بري و إبراهيم هندي، الصفحات 12-13)

أ - المقابلة الشخصية

وتتم هذه المقابلة بواسطة أشخاص مدربين على القيام بهذه الأعمال ويقوم الباحث المدرب بطرح أسئلة محددة ومعدة مسبقاً على الشخص المقصود ويسجل الإجابة عن هذه الأسئلة .
ومن مميزات المقابلة الشخصية الحصول على معلومات دقيقة ويستطيع الباحث الذي يقوم بطرح الأسئلة توضيح أي غموض أو التباس قد تكون موجودة في الأسئلة .

وأما عيوبها فهي التكلفة العالية والتحيز الناتج عن تأثير جامع البيانات على الشخص المبحوث سواء كان بقصد

أم بغير قصد

ب - الهاتف

ويستخدم كوسيلة أيضاً مباشرة وهو غير مكلف لكنه غير متوفر لدى الجميع مما يجعل عملية جمع البيانات مقتصرة على من يملكونه وهي أهم عيوب هذه الطريقة

ج- المراسلة

ويتم جمع المعلومات عن طريق إرسال استمارة إحصائية إلى الشخص المبحوث عبر البريد ، ومن مميزاتها التكلفة القليلة ولكن يعاب عليها احتمال عدم رد الإستمارة إلى الجهة المصدرة لها .

ثانيا: المصدر غير المباشر (التاريخي):

هي بيانات معدة مسبقا عن ظاهرة ما وباستطاعة الباحث الرجوع إليها وأخذ المعلومات المطلوبة مثل دائرة الإحصاءات العامة ودائرة الأحوال المدنية والوزارات والمؤسسات الخاصة والمؤسسات العامة والمصادر غير المباشرة تشمل الوثائق والمطبوعات والنشرات الإحصائية التي تصدرها الهيئات في البلاد المختلفة وكذلك الهيئات الدولية مثل هيئة الأمم المتحدة.

وكمثال على المصادر التاريخية يمكن أخذ المعلومات عن حالات الوفيات والولادة والزواج والطلاق من سجلات دائرة الأحوال المدنية دون الرجوع إلى الوحدات الأصلية . أما مميزات هذا المصدر للمعلومات أنه يوفر الوقت والجهد والمال أما عيوبه فمن المحتمل أن تكون البيانات غير دقيقة (عوض و عزام، 2000، الصفحات 13-14).

2- أساليب جمع البيانات الإحصائية

وعند دراسة صفة ما أو صفات معينه لمجتمع ما، فإن البيانات الإحصائية عن تلك الصفة أو الصفات تجمع بأحد الأسلوبين:

1-1 الحصر الشامل:

يعتبر أسلوب الحصر الشامل أفضل أسلوب في جمع البيانات كونه يوفر للباحث بيانات كاملة عن كافة مفردات مجتمع الدراسة (المشهداني و هرمز، 1989، صفحة 30)، تجمع فيه البيانات عن كل مفردة من مفردات المجتمع الإحصائي التي تشترك بخاصية أو مجموعة خصائص معينة كمثال مجتمع طلبة وطالبات جامعة الجليلي بونعامة خميس مليانة، مجتمع الأسر الساكنة في ولاية عين الدفلى وغيرها.

من مميزات هذا الأسلوب هو دقة النتائج المتحصل عليها والوثوق في كفاءتها نظرا لجمع البيانات من كل فرد شمله البحث من دون ترك مفردة أو حالة، كذلك تجنب أخطاء التعميم التي تنتج من استخدام بيانات مأخوذة من عينة محددة من المجتمع وتطبيق نتائجها على المجتمع كله، كما تتفادى هذه الطريقة الأخطاء الشائعة والناجمة في غيرها من الطرائق (طريقة العينة) خاصة خطأ التحيز وخطأ الصدفة.

أما من عيوبه فهو باهظ التكاليف ويحتاج إلى إمكانيات طائلة، كما يستغرق وقتا طويلا وتبذل فيه جهود كبيرة في جمع البيانات وتصنيفها، كما يتطلب وفرة في الوقت والمال والمجهود الفني وتزداد هذه المتطلبات وتتضاعف كلما ازداد حجم المجتمع (عدد أفراد المجتمع)، بالإضافة إلى أنه يحتاج إلى جهاز إداري وفني ضخم ومدرّب للقيام به.

وهذا الأسلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضخمة مثل تعدادات السكان والتعدادات الصناعية والتعدادات الزراعية.

من الأمثلة عليه في الجزائر الإحصاء العام للسكن والسكان، الإحصاء العام الفلاحي، الإحصاء الاقتصادي

2-2 المعاينة:

أسلوب المعاينة 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 مثلاً البيانات الخاصة بالضرائب فمن المستحسن أن يتم حصر جميع الأفراد (المصري، 1970، صفحة 19)، 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 عديدة منها:

* يؤدي استخدام أسلوب الحصر الشامل إلى ارتفاع التكاليف فهي تفوق بكثير تكاليف إجراء البحث على أساس العينة، وإذا أدخلنا عامل دقة النتائج في الاعتبار فإننا نجد أن مستوى الدقة التي نحصل عليها من العينة بالنسبة لوحدة من التكاليف يكون أكبر في حالة العينة منه في حالة الحصر الشامل فمثلاً إذا كانت دقة الحصر الشامل هو 100% وتكلفته مثلاً مليون دينار فعند أخذ عينة من ربع مفردات المجتمع فقد ينخفض مستوى الدقة إلى 90% مثلاً لكن تنخفض التكاليف إلى الربع (المصري، 1970، الصفحات 17-18)، فأسلوب العينات يؤدي إلى خفض تكاليف الدراسات الميدانية بسبب صغر حجم العينة بالنسبة إلى حجم المجتمع وهو ما يؤدي إلى تخفيض الأعباء الإدارية والفنية التي تتطلبها أي دراسة ميدانية.

* يتعذر في بعض الأحيان تحديد جميع مفردات المجتمع لإجراء حصر شامل مثل دراسة أذواق المستهلكين لسلعة معينة ففي هذه الحالة يصعب علينا تحديد كل المستهلكين لها لذا نلجأ على أسلوب المعاينة (الصياد و محمد ربيع، 1983، صفحة 106).

* يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتوضح 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 الآلية الضخمة، ويكون على الباحثين مستخدمي هذه النتائج مراعاة الوقت الذي ينقضي 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 (اللانهاية) مثل مجتمع الكائنات الحية في البحار والمحيطات والطيور لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة (البلداوي، 2009، صفحة 22).

* 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 المفرعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينة وليس سحب الدم كله. (عوض و عزام، 2000، صفحة 15).

3- أقسام المعاينة

تنقسم العينات عادة إلى قسمين رئيسيين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لها

1-3 المعاينة العشوائية

هي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفردة 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 درجة عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة.. من أهم أنواع العينات العشوائية ما يلي.

1-1-3 العينة العشوائية البسيطة

العينة العشوائية البسيطة تعد الأساس لباقي أنواع العينات العشوائية وتستخدم عندما يكون المجتمع متجانساً من حيث الغرض أو الصفة التي تتعلق بها الدراسة (البلداوي، 2009، صفحة 23)

العينة العشوائية البسيطة هي طريقة المعاينة التي يكون فيها احتمال اختيار أي مفردة متساو $(1/N)$ أي أن المجتمع ككل يعامل بنفس الطريقة ولا يجري 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.

من شروط اختيار العينة البسيطة وجود إطار للمجتمع يكون حديثا وشاملا لكل مفردات المجتمع يرقم فيها وحدات المجتمع من 1 إلى N ، وتحديد حجم العينة، كذلك يتم اختيار كل مفردة من مفردات العينة مستقلة عن اختيار المفردات الأخرى أي يكون لكل مفردة من مفردات لمجتمع الأصلي فرصة متساوية مع غيرها من المفردات في أن تكون ضمن مفردات العينة.

وبذلك فلكل عينة حجمها n احتمال الاختيار نفسه من بين العينات الممكنة أي $1/C_N^n$ (البلداوي، 2009،

صفحة 23)

طرق سحب العينة العشوائية البسيطة:

هناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي ومن أشهر هذه الطرق هي:

القرعة:

الطريقة الأولى: يقوم الباحث بإعداد قائمة بها جميع العينات المحتمل تكوينها من مجتمع البحث فمثلا لو كان 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:

الحالة الأولى: السحب بدون إرجاع

في هذه الحالة يتم استبعاد المفردة أو العينة في كل مرة قبل سحب الثانية وهنا يمكن استخدام فكرة التوافق حيث يتم توفير 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 أو المجموعة 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

الحالة الثانية: السحب مع إرجاع

في هذه الحالة يتم إعادة المفردة أو العينة أو المجموعة 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
9	BC	18	CE	27	EC	36	FF

بعد ذلك يقوم الباحث بتسجيل رقم كل عينة محتملة في قصاصة من الورق أو كرة من الكرات أو بطاقة 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 لو قام 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 لو كان 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:

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

في حالة سحب العينات مع الإرجاع هو:

$$N^n = 100^3 = 100*100*100 = 1000000 \text{ عينة}$$

فهل يعقل أن يقوم الباحث بكتابة مليون قصاصة ورق لكي يسحب منهم عينة تحتوي على 3 مفردات

فقط؟ الإجابة بالطبع تكون بالنفي

الطريقة الثانية:

هي أن يقوم الباحث بترقيم كل مفردة من مفردات المجتمع وتسجيل هذه الأرقام في قصاصات أو بطاقات أو كرات وخلطها خلطا جيدا ثم يسحب منها العدد المطلوب الذي يمثل حجم العينة وفي هذه الحالة يجب أن يفرق الباحث بين سحب المفردات مع الإرجاع أو مع 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 يمثل D 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 القصاصات الجديدة هو الرقم 2 حيث تحمله المفردة B فتكون المفردة B هي 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 الطريقتين السابقتين عمليا خاصة إذا كان عدد مفردات مجتمع البحث كبيرا جدا نظرا لصعوبة عملية إعداد القصاصات أو البطاقات أو الكرات وتسجيل أرقام مفردات المجتمع عليها ثم خلطها وسحب العدد المطلوب منها.

جدول الأرقام العشوائية:

هي أعداد صحيحة مكونة من الأرقام 0,1,2,3, ... 8,9 تشكل أرقاما من رقم 1 حتى رقم 100 ألف تقريبا موضوعة في جداول بغير ترتيب وإنما بصورة عشوائية مكونة من صفوف وأعمدة وهذه الجداول موجودة على شكل كتيب.

يتم التعامل مع الجدول من خلال اختيار نقطة بداية بشكل عشوائي ومن ثم يتم قراءة الأرقام التي يكون عدد منازلها مساويا 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 أفرادها 10 أفراد من مجتمع عدد أفراد 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 صفًا وعمودًا 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 آنفًا.

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:

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

أي عدد يزيد عن 3500 أو إذا 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

مثال 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.

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

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 أو طرح مضاعفاته 100-200-300... فتصبح الأرقام السابقة على النحو التالي:

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

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

2-1-3 العينة العشوائية المنتظمة (Systematic Sample)

اختيار هذه العينة يتطلب وجود إطار للمجتمع كما في حالة العينة العشوائية البسيطة بحيث يعطى لكل مفردة 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 مفردة ونريد اختيار عينة منتظمة حجمها 100 مفردة فإننا نقسم الإطار إلى فترات منتظمة طول كل فترة $20 = \frac{2000}{100} = \frac{N}{n}$ مفردة ومن داخل مفردات الفترة الأولى (1 - 20) يختار مفردة واحدة عشوائياً ولتكن رقم 14 مثلاً وبناء على رقم تلك المفردة يتحدد باقي أرقام المفردات الأخرى بشكل منتظم وفقاً لمتوالية عددية حدها الأول هو 14، 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 الأخطاء التي ترتكب في اختيار مفردات العينة فضلاً عن سهولة إجراءاتها. ولكن أهم عيوب المعاينة المنتظمة هو عدم صلاحيتها إذا ما وجدت علاقة دورية مع ترتيب العناصر في القائمة وكان طول الفترة بين عناصر العينة مساوياً لطول الدورة أو إحدى مضاعفاتها.

3-1-3 العينة الطبقية العشوائية (Stratified random sampling)

تعتبر العينة العشوائية الطبقية أفضل أنواع العينات وأكثرها دقة في تمثيل المجتمع الإحصائي غير المتجانس حيث أنه في كثير من الأحوال تكون مفردات المجتمع الإحصائي غير متجانسة من حيث الصفة أو الصفات المدروسة. وتناسب هذه العينة المجتمعات التي يمكن تقسيمها إلى مجتمعات تتصف بدرجة كبيرة من التجانس داخل كل مجموعة ودرجة كبيرة من التباين بين المجموعات بعضها البعض (المصري، 1970، صفحة 27) يلجأ إليها الباحث في حالة ما إذا كان مجتمع الدراسة منقسماً إلى طبقات طبيعية وتكون لدينا الرغبة في تمثيل جميع هذه الطبقات في العينة بحيث أن التجانس أو التقارب داخل كل طبقة من طبقات مجتمع الدراسة أكبر من التجانس داخل المجتمع ككل (أي أن التشتت داخل المجتمع ككل أكبر من التشتت داخل كل فئة من فئاته على حده) في هذه الحالة يجب على الباحث مراعاة أن الطبقة داخل العينة بنفس نسبة وجودها داخل المجتمع.

وتتلخص الطريقة بتحديد حجم العينات الجزئية المتناسبة من كل طبقة على أساس المعادلة:

$$\text{حجم العينة الطبقية} = (\text{حجم الطبقة} \div \text{حجم المجتمع}) \times \text{حجم العينة}$$

بعد أن يتم تحديد عدد المفردات التي يجب سحبها من كل طبقة للدخول في العينة فإن هذه المفردات يتم سحبها عشوائياً من داخل الطبقة (عينة عشوائية بسيطة أو منتظمة) 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 منحازاً لإحدى الفئات الثلاث. وعليه يجب تقسيم المجتمع الاحصائي إلى ثلاث فئات الأولى تضم الأسر ذات الدخل المرتفعة، والثانية تضم الأسر ذات الدخل المتوسطة والثالثة تضم الأسر ذات الدخل المنخفضة، وبعد ذلك يتم سحب عينة عشوائية بسيطة أو منتظمة من كل مجموعة يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع أحجام العينات العشوائية الثلاث تؤلف حجم العينة العشوائية الطبقية.

مثال 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.

الحل:

$$\text{المجتمع الكلي} = 500 + 400 + 280 + 200 + 220 = 1600 \quad \text{العينة من الطبقة الأولى} = 500 \times \frac{160}{1600} = 50$$

$$\text{العينة من الطبقة الثانية} = 400 \times \frac{160}{1600} = 40 \quad \text{العينة من الطبقة الثالثة} = 280 \times \frac{160}{1600} = 28$$

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

مثال 2:

بفرض أن مجتمع الدراسة مكون من 5000 مزرعة تمر، وهذه المزارع مصنفة حسب حجم الإنتاج كما يلي:

حجم الإنتاج	صغير	متوسط	كبير
عدد المزارع	1500	2500	1000

فإذا كان حجم العينة المطلوب سحبها هو 1000 مزرعة، وفقاً لطريقة التقسيم المتناسب تسحب عينة تتناسب مع الوزن النسبي لكل طبقة في المجتمع كما هو مبين بالجدول التالي:

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

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.

4-1-3 العينة العنقودية (Cluster Sample)

يلجأ إليها الباحث عندما يكون مجتمع الدراسة كبير جداً ومتناثراً على مساحات شاسعة تكلف الكثير من الوقت والجهد في التنقل بينها عند جمع البيانات، أيضاً 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 الإحصائي 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 وكأنا نتحدث عن عنقود.

وعلى سبيل المثال فعند إجراء دراسة لتقدير متوسط استهلاك العائلة في الجزائر لمادة السكر، إذن الوحدة الإحصائية 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.

2-3 العينات غير العشوائية:

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

1-2-3 العينة الميسرة: Consenience sample

يتطلب اختيار العينة الميسرة كما يفهم من الاسم ضمناً جمع معلومات من أعضاء المجتمع الموجودين في ظروف مريحة لجمع تلك المعلومات، من الأمثلة على ذلك:

أن يسأل الباحث مئة شخص الذين يقابلهم قبل غيرهم في الطريق، أو حين يسأل الإعلامي أول من يصادفه في الشارع، أو كما في الاستطلاعات الفورية للرأي العام حيث يتم اختيار المفردات من المراكز التجارية والشوارع... إلخ، دون التقيد بمحددات علمية لتوصيف العينة.

مما تقدم فإنه نجد أن وحدة العينة هنا قد اختارت نفسها أو اختيرت بواسطة الباحث لأنها متاحة فقط وليس لأي سبب آخر، وهنا لا يمكن أن يقول الباحث أن عينته ممثلة للمجتمع لأن معظم وحدات مجتمع البحث لم تتح لها الفرصة لاختيارها في العينة (فهومي، 2005، صفحة 114)

2-2-3 العينة العمدية أو القصدية: Purposive sample

في مثل هذه الحالة يختار الباحث مفردات العينة بطريقة تجعلها تمثل في نظره المجتمع الاحصائي، بالتالي فإن هذه العينة لا يمكن أن تكون غير متحيزة لأنها تتحكم فيها التأثيرات الشخصية (الأطرقجي، 1980، صفحة 286) 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 يعتمد الباحث اختيار مفردات معينة معينة لمجتمع الدراسة يرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلاً مقبولاً لمجتمع الدراسة.

مثلاً إذا أراد باحث دراسة خصائص اقتصادية أو اجتماعية معينة 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 أو قرية قريبة من مشروع صناعي ضخم يستوعب في قواه العاملة معظم سكانها.. هذه القرية أو تلك قد يأخذ النمط المعيشي لسكانها طابعاً خاصاً بعيداً عن النمط المعيشي المعتاد لبقية القرى، لذلك فأي منها لا يمكن أن يعطي تمثيلاً مقبولاً لريف تلك الدولة، لهذا

فإن الباحث وعلى ضوء خبراته السابقة يعتمد اختيار قرية معينة يرى أنها — من وجهة نظره الشخصية — يمكن أن تمثل الريف، وهذه الطريقة غير علمية وغالباً يتم اللجوء إليها في حالة البحوث التمهيدية.

على الرغم من اعتقاد الباحث أن اختياره لوحدة العينة بدقة على أساس خبرته ومعرفته بخصائص مجتمع البحث إلا أنه ليس هناك ما يؤكد أن هذه العينة ممثلة لمجتمع البحث وهذا عيب جوهري خاصة إذا ما أراد الباحث تعميم نتائج بحثه على مجتمع البحث، وهذا عيب ينطبق على جميع العينات غير الاحتمالية. (فهيم، 2005، صفحة 115)

3-2-3 العينة الحصصية: Quota sample

هي عبارة عن تقسيم المجتمع الاحصائي إلى طبقات بالنسبة لصفات أو أغراض معينة تتحدد بغرض البحث الذي نقوم به والنتائج المطلوبة ثم يسحب بطريقة شخصية وحدات من جميع هذه الطبقات بحيث تتناسب أحجامها مع أحجام هذه الطبقات في المجتمع الأصلي (الأطرقجي، 1980، صفحة 288)

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 أو خصائص 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 جامعو 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 الموضوعية لكل طبقة من الطبقات المذكورة.

واضح أنه رغماً من أن هذه الطريقة في ظاهرها مماثلة للعينة الطباقية العشوائية، إلا أنه في الحالة الأخيرة (العينة الطباقية 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 الحصصية إذا كان من المرغوب فيها إظهار النتائج في وقت قصير مع التغاضي عن توافر درجة دقة عالية بتلك النتائج حيث تكون ذات فائدة عملية في المراحل الأولى للبحث وهي مناسبة بشكل كبير الدراسات الاستطلاعية (فهيم، 2005، صفحة 116).

المحور الثاني

نظرية المعاينة

Sampling Theory

نظرية توزيعات المعاينة

Sampling Distributions Theory

تهتم نظرية المعاينة بدراسة العلاقة بين المجتمع والعينات المسحوبة منه فيما يسمى بالاستدلال الإحصائي، *statistical inference*، وهي علاقات تعتبر ذات أهمية أساسية في تقويم القيم التي تميز المجتمع أي مؤشرات المجتمع مثل المتوسط الحسابي والوسيط والتباين إلخ...، انطلاقاً من مؤشرات العينة (حليمي، 1985، صفحة 186)

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

1- توزيعات المعاينة للأوساط: *Sampling Distributions of Means*1-1- متوسط الوسط الحسابي للعينة $\mu_{\bar{X}}$:

نظرية 1:

إذا كانت X متغيرة عشوائية تمثل مجتمع ما و $\bar{X}$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من هذا المجتمع بالإرجاع أو بدون إرجاع فإن: $E(\bar{X}) = \mu_{\bar{X}} = \mu_X$

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n (X_i)\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu_x = \frac{1}{n} n \mu_x = \mu_x$$

ملاحظة:

- إذا سحبنا جميع العينات من الحجم n من مجتمع ما حجمه N فإنه وبالضرورة سنجد أن متوسط المجتمع يكون مساوياً لمتوسط متوسطات العينات المسحوبة منه بإرجاع أو بدون إرجاع.

- إذا سحبنا عينة من مجتمع ما فإننا نتوقع أن يكون متوسط العينة مساوي لمتوسط المجتمع، لذلك يستخدم متوسط العينة لتقدير متوسط المجتمع إذا كان هذا الأخير مجهولاً، ونكتب: $\hat{\mu}_x = \bar{X}$ ونقول إن الإحصائية $\bar{X}$ هي مقدرة لمعلمة متوسط المجتمع μ_x .

مثال:

إذا كان لدينا مجتمع يتكون من المفردات التالية: 1.2.3.4.5 وسحبنا جميع العينات العشوائية البسيطة من الحجم 2 التي يمكن سحبها من هذا المجتمع في الحالتين مع أو بدون إرجاع.

- أوجد توزيع المعاينة للوسط الحسابي وأحسبه في الحالتين مع أو بدون إرجاع.

الحل:

$$\mu_x = \frac{\sum X_i}{N} = \frac{1+2+3+4+5}{5} = 3 \quad \text{الوسط الحسابي للمجتمع هو:}$$

الحالة الأولى: توزيع المعاينة للوسط الحسابي وحسابه في حالة السحب مع الإرجاع

$$N^n = 5^2 = 25 \quad \text{عدد العينات الممكنة سحبها إذا كان السحب بإرجاع هي:}$$

نوجد أولاً جميع العينات الممكنة في هذه الحالة، ثم نحسب الوسط الحسابي لكل عينة كما يلي:

العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$
(1.1)	1	(2.1)	1.5	(3.1)	2	(4.1)	2.5	(5.1)	3
(1.2)	1.5	(2.2)	2	(3.2)	2.5	(4.2)	3	(5.2)	3.5
(1.3)	2	(2.3)	2.5	(3.3)	3	(4.3)	3.5	(5.3)	4
(1.4)	2.5	(2.4)	3	(3.4)	3.5	(4.4)	4	(5.4)	4.5
(1.5)	3	(2.5)	3.5	(3.5)	4	(4.5)	4.5	(5.5)	5

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

$\bar{X}$	1	1.5	2	2.5	3	3.5	4	4.5	5
$f(\bar{X})$	1/25	2/25	3/25	4/25	5/25	4/25	3/25	2/25	1/25

متوسط التوزيع العيني لـ $\bar{X}$:

$$\mu_{\bar{X}} = \sum \bar{X} f(\bar{X}) = (1 * \frac{1}{25}) + (1.5 * \frac{2}{25}) + + (4.5 * \frac{2}{25}) + (5 * \frac{1}{25}) = 3 = \mu_x$$

وهو يساوي نفس قيمة μ_x كما يجب أن يكون.

الحالة الثانية: توزيع المعاينة للوسط الحسابي وحسابه في حالة السحب بدون الإرجاع

عدد العينات الممكنة سحبها إذا كان السحب بإرجاع هي:

$$C_N^n = \frac{N!}{n!(N-n)!} = \frac{5!}{2!(5-2)!} = \frac{20}{2} = 10$$

نوجد أولاً جميع العينات الممكنة في هذه الحالة، ثم نحسب الوسط الحسابي لكل عينة كما يلي:

العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$	العينة	$\bar{X}$
(1.2)	1.5	(1.4)	2.5	(2.3)	2.5	(2.5)	3.5	(3.5)	4
(1.3)	2	(1.5)	3	(2.4)	3	(3.4)	3.5	(4.5)	4.5

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

$\bar{X}$	1.5	2	2.5	3	3.5	4	4.5
$f(\bar{x})$	1/10	1/10	2/10	2/10	2/10	1/10	1/10

متوسط التوزيع العيني لـ $\bar{X}$:

$$\begin{aligned} \mu_{\bar{X}} = \sum \bar{x} f(\bar{x}) &= (1.5 * \frac{1}{10}) + (2 * \frac{1}{10}) + (2.5 * \frac{2}{10}) + (3 * \frac{2}{10}) + (3.5 * \frac{2}{10}) + (4 * \frac{1}{10}) \\ &+ (4.5 * \frac{1}{10}) = \frac{30}{10} = 3 = \mu_x \end{aligned}$$

2-1- تباين الوسط الحسابي للعينات $\sigma_{\bar{X}}^2$

نظرية 2:

إذا كانت X متغيرة عشوائية تمثل مجتمع ما حجمه N و $\bar{X}$ متغيرة عشوائية تمثل متوسط عينة حجمها n

مسحوبة من هذا المجتمع فإن تباين $\bar{X}$ أي $\sigma_{\bar{X}}^2$ يكتب كما يلي:

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \quad - \text{ في حالة السحب مع الإرجاع فإن:}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \left(\frac{N-n}{N-1} \right) \quad - \text{ في حالة السحب بدون إرجاع فإن:}$$

ملاحظة:

- في حالة السحب مع الإرجاع فإن $\sigma_{\bar{X}}^2$ يتأثر طردياً بتباين المجتمع وعكسياً بحجم العينة أي كلما كان حجم العينة كبيراً والمجتمع أكثر تجانساً كان التقدير أدق.

- في حالة السحب بدون إرجاع فإن المقدار $\left(\frac{N-n}{N-1} \right)$ والذي يسمى بمعامل الإرجاع يصبح مهماً إذا

كان يقترب من الواحد أي إذا كان حجم العينة صغير جداً بالمقارنة مع حجم المجتمع

$$\left(\frac{n}{N} < 0.05 \right) \Rightarrow \left(\frac{N-n}{N-1} \right) \approx 1$$

- تباين توزيع المعاينة للمتوسطات $\sigma_{\bar{X}}^2$ في حالة المعاينة بدون إرجاع أقل منه في حالة السحب مع الإرجاع

أي أنه المعاينة بدون إرجاع تعطي تقديراً أكثر دقة لمعلمة المجتمع μ_X أي:

$$n > 1 \Rightarrow \left(\frac{N-n}{N-1} \right) < 1 \Rightarrow \left(\left(\frac{\sigma_X^2}{n} \left(\frac{N-n}{N-1} \right) \right) \leq \left(\frac{\sigma_X^2}{n} \right) \right)$$

مثال:

(نفس المثال السابق) إذا كان لدينا مجتمع يتكون من المفردات التالية: 1.2.3.4.5 وسحبنا جميع

العينات العشوائية البسيطة من الحجم 2 التي يمكن سحبها من هذا المجتمع في الحالتين مع أو بدون إرجاع.

- أحسب تباين توزيع المعاينة للوسط الحسابي في الحالتين مع أو بدون إرجاع مع التحقق.

الحل:

تباين المجتمع يحسب كما يلي:

$$\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2 = \frac{1}{5} ((1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2) = \frac{10}{5} = 2$$

الحالة الأولى: تباين الوسط الحسابي للعينات في حالة السحب مع الإرجاع

من خلال جدول العينات الممكنة نحسب تباين الوسط الحسابي لكل عينة كما يلي:

$\bar{X}$	$(\bar{X} - \mu_{\bar{X}})^2$	$\bar{X}$	$(\bar{X} - \mu_{\bar{X}})^2$	$\bar{X}$	$(\bar{X} - \mu_{\bar{X}})^2$	$\bar{X}$	$(\bar{X} - \mu_{\bar{X}})^2$	$\bar{X}$	$(\bar{X} - \mu_{\bar{X}})^2$
1	$(1-3)^2 = 4$	1.5	2.25	2	1	2.5	0.25	3	0
1.5	2.25	2	1	2.5	0.25	3	0	3.5	0.25
2	1	2.5	0.25	3	0	3.5	0.25	4	1
2.5	0.25	3	0	3.5	0.25	4	1	4.5	2.25
3	0	3.5	0.25	4	1	4.5	2.25	5	4

جدول توزيع المعاينة لتباين الوسط الحسابي للعينات:

$\bar{X}$	1	1.5	2	2.5	3	3.5	4	4.5	5
$f(\bar{X})$	1/25	2/25	3/25	4/25	5/25	4/25	3/25	2/25	1/25
$(\bar{X} - \mu_{\bar{X}})^2$	$(1-3)^2 = 4$	2.25	1	0.25	0	0.25	1	2.25	4
$f(\bar{X})(\bar{X} - \mu_{\bar{X}})^2$	4/25	4.5/25	3/25	1/25	0	1/25	3/25	4.5/25	4/25

تباين التوزيع العيني لـ $\bar{X}$:

$$\sigma_{\bar{X}}^2 = \sum (\bar{X} - \mu_{\bar{X}})^2 f(\bar{X}) = (1-3)^2 * \left(\frac{1}{25}\right) + (1.5-3)^2 * \left(\frac{2}{25}\right) + \dots + (5-3)^2 * \left(\frac{1}{25}\right) = \frac{25}{25} = 1$$

وهي تساوي $\frac{\sigma_X^2}{n}$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} = \frac{2}{2} = 1 \quad \text{للتحقق: من خلال النظرية}$$

الحالة الثانية: تباين الوسط الحسابي للعينات في حالة السحب بدون إرجاع

جدول توزيع المعاينة لتباين الوسط الحسابي للعينات:

$\bar{X}$	1.5	2	2.5	3	3.5	4	4.5
$(\bar{X} - \mu_{\bar{X}})^2$	$(1.5 - 3)^2 = 2.25$	1	0.25	0	0.25	1	2.25
$f(\bar{X})$	1/10	1/10	2/10	2/10	2/10	1/10	1/10
$f(\bar{X})(\bar{X} - \mu_{\bar{X}})^2$	2.25/10	1/10	0.25/10	0	0.25/10	1/10	2.25/10

تباين التوزيع العيني لـ $\bar{X}$:

$$\sigma_{\bar{X}}^2 = \sum (\bar{X} - \mu_{\bar{X}})^2 f(\bar{X}) = (1.5 - 3)^2 * \left(\frac{1}{10}\right) + (2 - 3)^2 * \left(\frac{1}{10}\right) + \dots + (4.5 - 3)^2 * \left(\frac{1}{10}\right) = \frac{7.5}{10} = 0.75$$

$$\frac{\sigma_X^2}{n} \left(\frac{N - n}{N - 1} \right) \quad \text{وهي تساوي}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \left(\frac{N - n}{N - 1} \right) = \frac{2}{2} \left(\frac{5 - 2}{5 - 1} \right) = \frac{3}{4} = 0.75 \quad \text{للتحقق: من خلال النظرية}$$

3-1- طبيعة توزيع الوسط الحسابي للعينة $\bar{X}$

نظرية 3:

- إذا كان المجتمع موزع طبيعياً بمتوسط μ_X وتباين σ_X^2 فإن متوسط العينة المسحوبة من هذا المجتمع عشوائياً

يتبع أيضاً التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu_X$ وتباين $\sigma_{\bar{X}}^2$

$$X \sim N(\mu_X, \sigma_X^2) \Rightarrow \bar{X} \sim N(\mu_{\bar{X}}, \sigma_{\bar{X}}^2) \Rightarrow Z = (\bar{X} - \mu_{\bar{X}}) / \sigma_{\bar{X}} \sim N(0, 1)$$

- هذا إذا كانت σ_X^2 معلومة، ولكن في حالة ما إذا كانت قيمة σ_X^2 غير معلومة فإننا نستخدم بدلاً منها

الانحراف المعياري للعينة S_x^2 ولكن في هذه الحالة يصبح المتغير T يخضع لتوزيع يعرف بتوزيع t ستودنت

بدرجات حرة $(n - 1)$ ، حيث:

$$T = (\bar{X} - \mu_{\bar{X}}) / S_x / \sqrt{n - 1} \sim t_{(1 - (\alpha/2), n - 1)}$$

$$X \sim N(\mu_X, \sigma_X^2) \Rightarrow \bar{X} \sim N(\mu_{\bar{X}}, \sigma_{\bar{X}}^2) \quad \text{and} \quad \sigma_X^2 = ?$$

$$T = (\bar{X} - \mu_{\bar{X}}) / S_x / \sqrt{n - 1} \sim t_{(1 - (\alpha/2), n - 1)}$$

- حسب نظرية النهاية المركزية فإنه إذا كان المجتمع غير طبيعي أو مجهول التوزيع بمتوسط μ_X وتباين σ_X^2 فإن

متوسط العينة المسحوبة من هذا المجتمع عشوائياً $\bar{X}$ لا تخضع للتوزيع الطبيعي ولكنها تتوزع توزيع يكون قريباً من

التوزيع الطبيعي أي يؤول إلى التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu_X$ وتباين $\sigma_{\bar{X}}^2$ إذا كان حجم العينة كبيراً $(30 \leq n)$

$$X \sim ?(\mu_X, \sigma_X^2) \text{ and } (n \geq 30) \Rightarrow \bar{X} \sim N(\mu_{\bar{X}}, \sigma_{\bar{X}}^2) \\ \Rightarrow Z = (\bar{X} - \mu_{\bar{X}} / \sigma_{\bar{X}}) \sim N(0,1)$$

ملاحظة:

- وتعتبر النتيجة السابقة هامة جداً في الإحصاء وخاصة في التطبيقات العلمية وتسمى نظرية النهاية المركزية

Central Limit Theorem

- في حالة السحب مع الإرجاع فإن: $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$
- في حالة السحب بدون إرجاع فإن: $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \left(\frac{N-n}{N-1} \right)$
- في حالة السحب بدون إرجاع و $\left(\frac{n}{N} < 0.05 \right)$ فإن: $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$

2- توزيعات المعاينة لتباين العينات: **Sampling Distributions of Variance**

1-2- متوسط تباينات العينات $E(S_{\bar{X}}^2)$:

نظرية 4:

إذا كانت X متغيرة عشوائية تمثل مجتمع ما حجمه N و S^2 متغيرة عشوائية تمثل تباين عينة حجمها n مسحوبة من هذا المجتمع فإن القيمة المتوقعة لتباين العينة $E(S^2)$ تكتب كما يلي:

- في حالة السحب مع الإرجاع فإن: $E(S^2) = \sigma_X^2 \left(\frac{n-1}{n} \right)$
- في حالة السحب بدون إرجاع فإن: $E(S^2) = \sigma_X^2 \left(\frac{n-1}{n} \right) \left(\frac{N}{N-1} \right)$

مثال:

(نفس المثال السابق) إذا كان لدينا مجتمع يتكون من المفردات التالية: **1.2.3.4.5** وسحبنا جميع

العينات العشوائية البسيطة من الحجم 2 التي يمكن سحبها من هذا المجتمع في الحالتين مع أو بدون إرجاع.

- أحسب القيمة المتوقعة لتباين العينة من خلال متوسط تباينات العينات في الحالتين مع أو بدون إرجاع، ثم قارن بينه وبين تباين المجتمع.

الحل:

الحالة الأولى: حساب متوسط تباينات العينات $E(S^2)$ في حالة السحب مع الإرجاع

نوجد أولاً جميع العينات الممكنة في هذه الحالة، ثم نحسب التباين لكل عينة كما يلي:

العينة	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(2.1)	(2.2)	(2.3)	(2.4)
$\bar{X}$	1	1.5	2	2.5	3	1.5	2	2.5	3
S^2	0	0.25	1	2.25	4	0.25	0	0.25	1
العينة	(2.5)	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)	(4.1)	(4.2)	(4.3)
$\bar{X}$	3.5	2	2.5	3	3.5	4	2.5	3	3.5
S^2	2.25	1	0.25	0	0.25	1	2.25	1	0.25
العينة	(4.4)	(4.5)	(5.1)	(5.2)	(5.3)	(5.4)	(5.5)	Σ	المتوسط
$\bar{X}$	4	4.5	3	3.5	4	4.5	5	75	75/25=3
S^2	0	0.25	4	2.25	1	0.25	0	25	25/25=1

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{2} \left((1.5-1)^2 + (1.5-2)^2 \right) = \frac{0.25+0.25}{2} = 0.25$$

توزيع المعاينة لتباين العينات S^2 هو:

S^2	0	0.25	1	2.25	4
$f(s^2)$	5/25	8/25	6/25	4/25	2/25

متوسط التوزيع العيني لـ S^2 :

$$E(S^2) = \sum S^2 f(s^2) = (0 * \frac{5}{25}) + (0.25 * \frac{8}{25}) + (1 * \frac{6}{25}) + (2.25 * \frac{4}{25}) + (4 * \frac{2}{25}) = \frac{25}{25} = 1$$

$$E(S^2) = \sigma_X^2 \left(\frac{n-1}{n} \right) = 2 \left(\frac{1}{2} \right) = 1 \quad \text{للتحقق: من خلال النظرية}$$

الحالة الثانية: حساب متوسط تباينات العينات $E(S^2)$ في حالة السحب بدون إرجاع

نوجد أولاً جميع العينات الممكنة في هذه الحالة، ثم نحسب التباين لكل عينة كما يلي:

العينة	(1.2)	(1.3)	(1.4)	(1.5)	(2.3)	(2.4)	(2.5)	(3.4)	(3.5)	(4.5)
$\bar{X}$	1.5	2	2.5	3	2.5	3	3.5	3.5	4	4.5
S^2	0.25	1	2.25	4	0.25	1	2.25	0.25	1	0.25

توزيع المعاينة لتباين العينات S^2 هو:

S^2	0.25	1	2.25	4	Σ
$f(S^2)$	4/10	3/10	2/10	1/10	1

متوسط التوزيع العيني لـ S^2 :

$$E(S^2) = \sum S^2 f(s^2) = (0.25 * \frac{4}{10}) + (1 * \frac{3}{10}) + (2.25 * \frac{2}{10}) + (4 * \frac{1}{10}) = \frac{12.5}{10} = 1.25$$

$$E(S^2) = \sigma_x^2 \left(\frac{n-1}{n} \right) \left(\frac{N}{N-1} \right) = 2 \left(\frac{1}{2} \right) \left(\frac{5}{4} \right) = \frac{10}{8} = 1.25 \quad \text{للتحقق: من خلال النظرية}$$

2-2- تباين تباينات العينات $(S_{\bar{X}}^2)$

نظرية 5:

إذا كانت X متغيرة عشوائية تمثل مجتمع ما حجمه N و $\bar{X}$ متغيرة عشوائية تمثل متوسط عينة حجمها n

مسحوبة من هذا المجتمع فإن تباين $S_{\bar{X}}^2$ أي $V(S_{\bar{X}}^2)$ يكتب كما يلي:

$$V(\hat{S}^2) = \frac{\mu_4}{n} + \frac{3-n}{n(n-1)} \sigma^4 \quad \text{تباين } \hat{S}^2 \text{ هو كالتالي:}$$

$$\mu_4 = E\left(\left[(X - \mu)^4\right]\right) \quad \text{حيث:}$$

3-2- طبيعة توزيع المعاينة للتباين $S_{\bar{X}}^2$: Sampling Distribution of The Variance

نظرية 6:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu, \sigma^2)$ ،

وكان S_x^2 تباين هذه العينة وكانت σ_x^2 معلومة فإن: $\frac{n \cdot S_x^2}{\sigma_x^2}$ يتبع توزيع كاي تربيع درجة حريته $(n-1)$.

مثال:

إذا كانت $X_1, X_2, \dots, X_{25}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_1, 25)$ ، و

$$P(S_x^2 \leq c) = 0.90 \quad \text{حيث } c \text{ فأوجد}$$

الحل:

$$P(S_x^2 \leq c) = P(nS_x^2 / \sigma_x^2 \leq nc / \sigma_x^2) = P(\chi^2_{(24)} \leq c) = 0.90$$

و باستعمال الجدول الاحصائي لتوزيع مربع كاي نجد $c = 33.196$

4-2- توزيع النسبة بين تبايني عينتين: Distribution of $\frac{S_1^2}{S_2^2}$

نظرية 7:

إذا كانت S_x^2 تباين عينة عشوائية $X_1, X_2, \dots, X_{n_1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت S_y^2 تباين عينة عشوائية $Y_1, Y_2, \dots, Y_{n_2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض، فإن:

$$F = \frac{\left[S_x^2 \frac{n_1}{n_1 - 1} \right] \frac{1}{\sigma_X^2}}{\left[S_y^2 \frac{n_2}{n_2 - 1} \right] \frac{1}{\sigma_Y^2}} = \frac{\frac{\hat{S}_x^2}{\sigma_X^2}}{\frac{\hat{S}_y^2}{\sigma_Y^2}} \sim F_{(n_1 - 1, n_2 - 1)}$$

يتبع توزيع فيشر درجة حريته $[(n_1 - 1), (n_2 - 1)]$.

مثال:

إذا كانت $X_1, X_2, \dots, X_{16}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_X, \sigma_X^2)$ ، وتباينها S_x^2 ، و $Y_1, Y_2, \dots, Y_{13}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_Y, \sigma_Y^2)$ ، وتباينها S_y^2 .

$$- \text{أوجد } c \text{ حيث : } P\left(\frac{S_x^2}{S_y^2} \geq c\right) = 0.01$$

الحل:

باستعمال الجدول الاحصائي لتوزيع فيشر الذي درجة حريته

$$F[(n_1 - 1), (n_2 - 1)] = [(16 - 1), (13 - 1)] = [(15), (12)] = 4.01$$

نجد $c = 4.01$

3- توزيعات المعاينة للفروق والمجاميع

3-1- توزيعات المعاينة للفروق والمجاميع للمتوسط:

نظرية 8:

- ليكن لدينا مجتمعان X و Y ، متوسطهما الحسابي μ_x و μ_y ، تباينهما σ_x^2 و σ_y^2 ، نسحب من كل مجتمع عينة عشوائية بحيث نسحب عينة عشوائية x حجمها n_1 من المجتمع X و عينة ثانية y حجمها n_2 من المجتمع Y والعينتان مستقلتان فإن:

$$\mu_{x+y} = \mu_x + \mu_y \quad \text{و} \quad \mu_{x-y} = \mu_x - \mu_y$$

مثال:

ليكن المجتمعان X و Y سحبنا منهما العينتين $x: 1.3.5.7$ و $y: 2.4.6$

$$\mu_{x+y} = \mu_x + \mu_y \quad \text{and} \quad \mu_{x-y} = \mu_x - \mu_y \quad \text{تحقق أن}$$

حالة المجموع

المجموع		العينة x			
μ_{x+y}		1	3	5	7
العينة	2	3	5	7	9
	4	5	7	9	11
y	6	7	9	11	13

$$\mu_y = \bar{y} = \frac{2+4+6}{3} = 4, \quad \mu_x = \bar{x} = \frac{1+3+5+7}{4} = 4$$

$$\mu_{x+y} = \frac{3+5+7+.....+9+11+13}{12} = \frac{96}{12} = 8$$

حالة الفرق

الفرق		العينة x			
μ_{x-y}		1	3	5	7
العينة	2	1-	1	3	5
	4	3-	1-	1	3
y	6	5-	3-	1-	1

$$\mu_{x-y} = \frac{(-1)+1+3+.....+(-3)+(-1)+1}{12} = \frac{0}{12} = 0$$

2-3- توزيعات المعاينة للفروق والجاميع للتباين:

نظرية 9:

- ليكن لدينا مجتمعان X و Y ، متوسطهما الحسابي μ_X و μ_Y ، تباينهما σ_X^2 و σ_Y^2 ، نسحب من كل مجتمع عينة عشوائية بحيث نسحب عينة عشوائية x حجمها n_1 من المجتمع X وعينة ثانية y حجمها n_2 من المجتمع Y والعينتان مستقلتان فإن:

$$\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 \quad \text{و} \quad \sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$$

مثال:

نفس المثال السابق ليكن المجتمعان X و Y سحبنا منهما العينتين $x: 1.3.5.7$ و $y: 2.4.6$

$$\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 \quad \text{and} \quad \sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2 \quad \text{تحقق أن}$$

حالة المجموع

المجموع		العينة x			
μ_{x+y}		1	3	5	7
العينة	2	3	5	7	9
	4	5	7	9	11
	6	7	9	11	13

$$\sigma_x^2 = \frac{(1-4)^2 + (3-4)^2 + (5-4)^2 + (7-4)^2}{4} = \frac{20}{4} = 5$$

$$\sigma_y^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = \frac{8}{3}$$

$$\sigma_{x+y}^2 = \frac{(3-8)^2 + (5-8)^2 + (7-8)^2 + \dots + (11-8)^2 + (13-8)^2}{8} = \frac{92}{12} = \frac{23}{3}$$

حالة الفرق

الفرق		العينة x			
μ_{x-y}		1	3	5	7
العينة	2	1-	1	3	5
	4	3-	1-	1	3
	6	5-	3-	1-	1

$$\sigma_{x-y}^2 = \frac{((-1)-0)^2 + (1-0)^2 + (3-0)^2 + \dots + ((-1)-0)^2 + (1-0)^2}{8} = \frac{92}{12} = \frac{23}{3}$$

3-3- توزيع الفرق بين وسطين

أ- توزيع المجتمعين طبيعي

نظرية 10:

إذا كانت $X_1, X_2, \dots, X_{n1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت $Y_1, Y_2, \dots, Y_{n2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض، نرسم لوسط العينة الأولى ب $\bar{X}$ و لوسط العينة الثانية ب $\bar{Y}$ فإن المتغير $\bar{X} - \bar{Y}$ يخضع لتوزيع طبيعي وسطه $\mu_X - \mu_Y$ و تباينه $(\sigma_1^2/n_1) + (\sigma_2^2/n_2)$ أي أن:

$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \sim N(0;1)$$

مثال:

إذا كانت أجور العمال اليومية في القطاع الفلاحي في بعض المستثمرات في ولاية عين الدفلى تتبع التوزيع الطبيعي $N(1350;900)$ وبالمقارنة فإن أجور العمال اليومية في القطاع الفلاحي في بعض المستثمرات في ولاية الوادي تتبع التوزيع الطبيعي $N(1190;700)$ وأخذت عينة من عمال ولاية عين الدفلى حجمها 25 عامل متوسط أجورهم اليومية $\bar{X}$ ، وعينة من عمال ولاية الوادي حجمها 20 عامل متوسط أجورهم اليومية $\bar{Y}$.

المطلوب: أوجد $P(\bar{X} - \bar{Y} \leq 150)$

الحل:

$$\begin{aligned} P(\bar{X} - \bar{Y} \leq 150) &= P\left(\frac{(\bar{X} - \bar{Y}) - (1350 - 1190)}{\sqrt{(900/25) + (700/20)}} \leq \frac{150 - (1350 - 1190)}{\sqrt{(36) + (35)}}\right) \\ &= P\left(Z \leq \frac{-10}{\sqrt{71}}\right) = P(Z \leq -1.186) = 0.1190 \end{aligned}$$

ب- توزيع المجتمعين غير طبيعي

نظرية 11:

إذا كانت $X_1, X_2, \dots, X_{n1}$ تمثل قياسات عينة عشوائية وسطها هو $\bar{X}$ مسحوبة من مجتمع وسطه μ_1 وتباينه σ_1^2 ، وكانت $Y_1, Y_2, \dots, Y_{n2}$ تمثل قياسات عينة عشوائية حجمها n_2 وسطها هو $\bar{Y}$ مسحوبة

من مجتمع وسطه μ_2 وتباينه σ_2^2 ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض وكان حجم العينتين كبير ($30 \leq n_1, n_2$) فإن: (أبو صالح و عوض، 1983، الصفحات 144-145)

- σ_1^2, σ_2^2 معلومتين

$$\frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \sim N(0;1)$$

- $\sigma_1^2 = \sigma_2^2$ مجهولتين

$$\frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{S \sqrt{(1/n_1) + (1/n_2)}} \sim t_{(1-(\alpha/2), n_1+n_2-2)}$$

حيث التباين المجمع للعينتين معا S^2 يسمى بالتباين المشترك للعينتين **The Pooled Variance** معرف بالعلاقة:

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

ملاحظة

إذا كان أحد المجتمعين أو كليهما لا يتوزع توزيعاً طبيعياً فإن $(\bar{x} - \bar{y})$ لا يتوزع توزيعاً طبيعياً كذلك، ولكن لقيم $(30 \leq n_1, n_2)$ الكبيرة فإنه طبقاً لنظرية النهاية المركزية السابقة فإن $(\bar{x} - \bar{y})$ يتوزع توزيعاً قريباً من التوزيع الطبيعي وبذلك يمكننا استخدام نفس العلاقة في حالة العينات الكبيرة.

4- توزيعات المعاينة للنسبة والفرق بين نسبتي

1-4- توزيع المعاينة للنسبة

عند دراسة مجتمع إحصائي يتكون من وحدات إحصائية تحمل صفة معينة ما a ووحدات إحصائية لا

تحمل هذه الصفة α ومجموعهما يمثل المجتمع أي $N = N_a + N_\alpha$

نرمز ب P للنسبة في المجتمع التي تحمل صفة معينة ما a حيث: $P = N_a/N$

نرمز ب q للنسبة في المجتمع التي لا تحمل هذه الصفة حيث: $q = N_\alpha/N = (N - N_a)/N$

$$\frac{N_a}{N} + \frac{N_\alpha}{N} = \frac{N_a + N_\alpha}{N} = 1 \quad \text{أي} \quad p + q = 1$$

عند سحب عينة عشوائية مؤلفة من n وحدة من هذا المجتمع فنلاحظ أن عدد الوحدات الإحصائية التي

تحمل الصفة a هي n_a

نرمز بـ P' للنسبة في العينة التي تحمل صفة معينة ما a حيث: $P' = n_a/n$

نرمز بـ q' للنسبة في العينة التي لا تحمل هذه الصفة حيث: $q' = n_{\alpha}/n = (n - n_a)/n$

$$\frac{n_a}{n} + \frac{n_{\alpha}}{n} = \frac{n_a + n_{\alpha}}{n} = 1 \quad \text{أي} \quad p' + q' = 1$$

إن النسبة P' في العينة قد تختلف عن النسبة P في المجتمع

نعلم أنه وفي حالة العينات الكبيرة يمكن تقريب التوزيع الثنائي بالتوزيع الطبيعي، فإذا كان المتغير العشوائي

هو P' و كانت العينة كبيرة و حجمها n فإن:

$$E\left(\frac{n_a}{n}\right) = E(P') = \frac{1}{n} E(n_a) = \frac{1}{n} \cdot n \cdot p = p \quad \text{ومنه} \quad E(X) = \mu_X = n \cdot p$$

$$V\left(\frac{n_a}{n}\right) = V(P') = \frac{1}{n^2} V(n_a) = \frac{1}{n^2} \cdot n \cdot p \cdot q = \frac{p \cdot q}{n} \quad \text{ومنه} \quad V(X) = n \cdot p \cdot q \quad \text{كذلك}$$

و إذا كانت n كبيرة ووفقا لنظرية النهاية المركزية فإن التوزيع التقريبي لـ P' هو:

$$n_a \sim N(nP, nPq) \Rightarrow Z = \frac{(n_a - nP)}{\sqrt{nPq}} \sim N(0,1)$$

$$\frac{n_a}{n} \sim N\left(P, \frac{Pq}{n}\right) \Rightarrow Z = \frac{(n_a/n - P)}{\sqrt{Pq/n}} \sim N(0,1)$$

$$P' \sim N\left(P, \frac{Pq}{n}\right) \Rightarrow Z = \frac{(P' - P)}{\sqrt{Pq/n}} \sim N(0,1)$$

نظرية 12:

لتكن $X_1, X_2, \dots, X_N$ تمثل قياسات مجتمع ما موزع طبيعيا حيث P نسبة المفردات في المجتمع

ذات صفة معينة، ولتكن P' متغيرة عشوائية تمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة من ذات

المجتمع و كانت n كبيرة ، فإننا نحصل على توزيع للإحصائية P'

حيث:

$$E(P') = \mu_{P'} = P$$

$$V(P') = \sigma_{P'}^2 = \frac{P(1-P)}{n} = \frac{Pq}{n}$$

فإن:

$$P' \sim N\left(P, \frac{Pq}{n}\right)$$

$$Z = \frac{(P' - E(P'))}{\sqrt{V(P')}} \sim N(0,1) \Rightarrow Z = \frac{(P' - P)}{\sqrt{Pq/n}} \sim N(0,1)$$

مثال:

في مستشفى الأمراض القلبية المستعصية فإن احتمال نجاح العملية هو 0.6 وأخذت عينة من 25 مريض من المرضى، فأوجد احتمال أن يكون نسبة الشفاء أكبر من 0.65.

الحل:

نعلم أن:

$$E(P') = \mu_{P'} = P = 0.60$$

$$V(P') = \sigma_{P'}^2 = \frac{P(1-P)}{n} = \frac{0.6 \times 0.6}{100} = \frac{0.36}{100} = 0.0036$$

فإن:

$$P' \sim N(0.6, 0.0036)$$

$$Z = \frac{(P' - E(P'))}{\sqrt{V(P')}} = \frac{(P' - 0.6)}{\sqrt{0.0036}} \sim N(0,1)$$

$$Z = P(P' \geq 0.65) = P\left(\frac{(P' - 0.60)}{\sqrt{0.0036}} \geq \frac{(0.65 - 0.60)}{\sqrt{0.0036}}\right)$$

$$P\left(Z \geq \frac{0.05}{0.06}\right) = 1 - P(Z < 0.833) = 1 - 0.7967 = 0.2033$$

4-2- توزيع المعاينة للفرق بين نسبتي

نظرية 13:

إذا أخذنا عينتان مستقلتان حجمهما n_1, n_2 من مجتمعين الأول يتبع توزيع ذي الحدين $b(1, P_1)$ ، والثاني يتبع توزيع ذات الحدين $b(1, P_2)$ ، وكانت n_1, n_2 كبيرتين، فإن الفرق بين النسبتين من العينتين يخضع تقريبا لتوزيع طبيعي

$$E (P_1' - P_2') = \mu_{P_1' - P_2'} = P_1 - P_2 \quad \text{وسطه :}$$

$$V (P_1' - P_2') = \sigma_{(P_1' - P_2')}^2 = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2} = \frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2} \quad \text{وتباينه :}$$

فإن:

$$P_1' - P_2' \sim N \left(P_1 - P_2, \frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2} \right)$$

$$Z = \frac{(P_1' - P_2' - E (P_1' - P_2'))}{\sqrt{V (P_1' - P_2')}} \sim N (0,1)$$

$$Z = \frac{((P_1' - P_2') - (P_1 - P_2))}{\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}} = \frac{((P_1' - P_2') - (P_1 - P_2))}{\sqrt{(P_1(1-P_1)/n_1) + (P_2(1-P_2)/n_2)}} \sim N (0,1)$$

المحور الثالث

نظرية التقدير

Estimation Theory

نظرية التقدير

Estimation Theory

ذكرنا في المحور السابق بأن العينة **sample** هي جزء من المجتمع **Population** وطريقة اختيار هذا الجزء يسمى طريقة المعاينة **sampling method** والغاية الرئيسية من دراسة العينات هوللاستدلال منها على خواص المجتمع الذي تعود اليه هذه العينات.

فإذا كان لدينا المتغير العشوائي X فان توزيعه الاحتمالي (أو كثافة احتماله) تعتمد على ثابت θ واحد أو أكثر لا تعرف قيمتها، وهذه الثوابت تسمى معالم **Parameters** وفي هذا المحور سندرس طرق تقدير معالم المجتمع من مقاييس الاحصائيات التي تحسب من العينة حيث اننا نحتاج لحساب قيمة أو إحصائية **Statistic** من العينة لكل معلمة من معالم المجتمع (أو دالة كثافة الاحتمال).

هذا وكل قيمة تحسب من العينة تسمى تقديرا **Estimate** أما الطريقة التي استخدمت في التقدير فتسمى مقدرة **Estimator** فالتقدير غير ثابت من عينة إلى أخرى عند استخدام نفس الطريقة بينما المقدر يكون ثابتا إلا إذا تغيرت طريقته، هذا والتقدير إما أن يكون : تقدير المعلمة بنقطة

Point Estimation أو تقدير المعلمة بفترة Interval Estimation

1- التقدير بنقطة Point Estimation

إذا حسبت قيمة مفردة من العينة كتقدير لمعلمة من المجتمع فالطريقة تسمى تقدير النقطة لأن نقطة واحدة فقط من فضاء العينة قد استخدم تقدير للمعلمة.

مثال: إن قيمة الوسط الحسابي للعينة $\bar{X}$ هو تقدير نقطة للمعلمة μ_X الوسط الحسابي للمجتمع الذي تعود إليه هذه العينة ، كما أن التباين S_X^2 للعينة هو تقدير نقطة لتباين المجتمع σ_X^2 ، كذلك النسبة P' هو تقدير لنسبة المجتمع P .

خصائص المقدر الجيد هي :

عدم التحيز Unbiasedness : التحيز هو ذلك الفرق بين مقدرة ما ووسط توزيعها، فإذا كان هذا الفرق يختلف عن الصفر نقول عن ذلك المقدر بأنه متحيز.

إن المقدر $\hat{\theta}$ يعتبر مقدرا غير متحيز اذا كان توقعه يساوي قيمة المعلمة θ أي $E(\hat{\theta}) = \theta$

مثال

الاتساق **Consistency** : يكون المقدّر متسقاً إذا كانت قيمته لا تختلف اختلافاً جوهرياً عن

قيمة المعلمة الحقيقية للمجتمع بزيادة حجم العينة أي أن $\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \varepsilon) = 1$ ، حيث أن ε هو الفرق بين المقدّر والمعلمة .

الكفاءة **Efficiency** : إن كفاءة المقدّر غير المتحيزة $\hat{\theta}_1$ إلى المقدّر غير المتحيز $\hat{\theta}_2$ هونسبة

تباين المقدّر $\hat{\theta}_2$ إلى تباين المقدرة $\hat{\theta}_1$ أي $e(\hat{\theta}_1) = V(\hat{\theta}_2) / V(\hat{\theta}_1)$ ومن التعريف أعلاه يتضح بأن المقدّر الأقل تبايناً هو الأعلى كفاءة.

مثال:

الوسط الحسابي للعينة $\bar{X}$ هو مقدّر غير متحيز لـ μ_X وكذلك الوسيط للعينة هو مقدّر غير متحيز لـ μ_X

ولكن تباين $\bar{X}$ هو $\frac{\sigma_X^2}{n}$ بينما تباين الوسيط هو $\frac{\pi}{2} \frac{\sigma_X^2}{n}$ ، لذا فإن كفاءة الوسط الحسابي هو

$$eff(\bar{X}) = \frac{V(M_e)}{V(\bar{X})} = \frac{(\pi/2) \times (\sigma_X^2 / n)}{(\sigma_X^2 / n)} = \frac{\pi}{2}$$

أي أن الوسط الحسابي $\bar{X}$ هو أكفأ تقديراً من الوسيط.

الكفاية **Sufficiency** :

يكون المقدّر $\hat{\theta}$ مقدراً كافياً للمعلمة θ إذا كان قد شمل كل المعلومات ذات العلاقة به θ المتوافرة في العينة. فعند سحب عينة عشوائية في مجتمع يتوزع توزيعاً طبيعياً فإن $\bar{X}$ هو مقدّر كافٍ لـ μ_X لأنه لا يمكن إضافة أي شيء على $\bar{X}$ لجعله مقدراً أحسن لـ μ_X لأن $\bar{X}$ يحوي على جميع المعلومات المتعلقة بـ μ_X من العينة.

هذا وهناك عدة طرق مهمة لايجاد تقدير المعلمة بنقطة وهي :

1-1- طريقة الإمكان الأكبر **Maximum Likelihood Method**

تعتبر هذه الطريقة من أهم طرق التقدير النقطي وخاصة في حالة العينات الكبيرة وتعزى هذه الطريقة لرونالد فيشر، وهي تعطي مقدرات كافية إن وجدت، إلا أنها قد تكون متحيزة أحياناً، ولكن في الحالة عندما تقول n إلى ما لا نهاية (حيث n حجم العينة) فإنها تعطي مقدرات غير متحيزة ولها أقل تباين ومتسقة دوماً، كما يؤول توزيعها إلى التوزيع الطبيعي مع زيادة حجم العينة.

نفرض أن المجتمع الإحصائي $\{A, X, f(x, \theta)\}$ حيث θ الوسيط المراد تقديره، ولنفرض أن متغيرات القيم الملاحظة في العينة من الشكل: $X_1, X_1, \dots, X_n \sim I.I.D$ ، نقصد بـ $I.I.D$ أي مستقلة ومتماثلة التوزيع، عندئذ تابع الكثافة المشترك للعينة ولنرمز له بالرمز $L(\theta)$ هو:

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

وفي طريقة الإمكانية العظمى يتم اختيار قيمة θ التي تجعل تابع الكثافة المشترك والذي يسمى بتابع الإمكانية العظمى أو تابع الإمكان أكبر ما يمكن، ويسمى المقدّر في هذه الحالة بالمقدّر المبني على التوقع الأعظمي أو مقدّر الإمكان الأكبر.

فإذا كانت $T = T(x_1, x_2, \dots, x_n)$ هي قيمة θ التي تعظم الدالة $L(\theta) = \prod_{i=1}^n f(x_i, \theta)$ ، أي

إذا كان $L(\theta) = \max_{\theta} L(\theta)$ عندئذ نقول عن T إنه مقدّر مبني على التوقع الأعظمي للوسيط θ وبما أن T يجب أن يحقق النهاية العظمى لتابع الإمكانية العظمى فإنه يمكن الحصول على هذه النهاية عادة عن طريق الاشتقاق بالنسبة للوسيط θ ثم مساواة الناتج بالصفر.

بمعنى آخر: إذا كان $L(\theta)$ تابعا قابلا للمفاضلة بالنسبة θ فإن الشرط اللازم لكي يبلغ $L(\theta)$ نهايته

$$\frac{\partial L}{\partial \theta} = 0 \text{ العظمى هو:}$$

ولما كان $L(\theta)$ هو إما جداء توابع احتمال أو جداء توابع كثافة، فإنه موجب دوما وبالتالي يكون

$$\log L(\theta) \text{ معرفا دوما، وهو تابع متزايد باضطراد.}$$

وبالحالة العامة يكون الوصول إلى أعظمية $\log L(\theta)$ أسهل من الوصول إلى أعظمية $L(\theta)$ لأنه

بأخذ اللوغاريتم تحول الدالة إلى حاصل جمع وهذا يسهل العمليات الرياضية في حين أن $L(\theta)$ عبارة عن جداء دوال.

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta) \text{ وبشكل مكافئ إن تغيرات التابع}$$

$$\log L(\theta) = \log \prod_{i=1}^n f(x_i, \theta) \text{ تماثل تغيرات التابع:}$$

وقيمة θ التي تجعل هذا التابع في نهايته العظمى تعطى من حل المعادلة

$$\frac{\partial}{\partial \theta} \left[\sum_{i=1}^n \log f(x_i, \theta) \right] = 0$$

والمقدر T المحسوب بهذه الطريقة يسمى: المقدر المبني على التوقع الأعظمي. (اسبر ديب، 2009،

الصفحات 297-299)

مثال 1:

في المجتمع الطبيعي $X \sim N(\mu, \sigma^2)$ وبفرض μ هو الوسيط المجهول، أوجد المقدر الأعظمي المبني على التوقع الأعظمي للوسيط μ

الحل:

طريقة 1: الخطوات كما يلي:

$$f(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad * \text{ دالة الكثافة من الشكل}$$

$$f(x_i, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \quad \text{وبالتالي:}$$

$$\log f(x, \mu) = \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2}(x_i - \mu)^2 \quad *$$

$$\sum_{i=1}^n \log f(x, \mu) = \sum_{i=1}^n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \quad * \text{ نأخذ المجموع}$$

$$\sum_{i=1}^n \log f(x, \mu) = n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial}{\partial \mu} \sum_{i=1}^n \log f(x, \mu) = 0 + \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \quad * \text{ نشتق بالنسبة للوسيط } \mu:$$

$$\frac{\partial}{\partial \mu} \sum_{i=1}^n \log f(x, \mu) = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \quad * \text{ نساوي المشتق بالصفر ونحل المعادلة الناتجة:}$$

$$\Rightarrow \sum_{i=1}^n x_i - n\mu = 0 \quad \Rightarrow \mu = \frac{1}{n} \sum_{i=1}^n x_i$$

بالتالي المقدر المبني على التوقع الأعظمي للوسيط μ في المجتمع الطبيعي المذكور هو $T = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$

طريقة 2: الخطوات كما يلي:

$$L(\mu) = \prod_{i=1}^n f(x, \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

$$\log L(\mu) = n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial}{\partial \mu} \log L(\mu) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)$$

$$\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \Rightarrow \sum_{i=1}^n x_i - n\mu = 0 \Rightarrow \mu = \frac{1}{n} \sum_{i=1}^n x_i$$

و بمساواتها بالصفر يكون:

$$T = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

هو المقدّر المبني على التوقع الأعظمي للوسيط μ في المجتمع الطبيعي المذكور .

مثال 2:

في المجتمع الطبيعي $X \sim N(\mu, \sigma^2)$ وبفرض μ و σ^2 مجهولين.

أوجد المقدرين المبنيين على التوقع الأعظمي للوسيطين μ و σ^2

الحل:

تابع الامكانية العظمى:

$$L(\mu, \sigma^2) = \prod_{i=1}^n f(x, \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}} \right)^n \left(\frac{1}{\sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}} \right)^n (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

$$\log L(\mu, \sigma^2) = \log \left(\frac{1}{\sqrt{2\pi}} \right)^n - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

وبالاشتقاق الجزئي مرة بالنسبة لـ μ وأخرى بالنسبة لـ σ^2 يكون:

$$\frac{\partial}{\partial \mu} \log L(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)$$

$$\frac{\partial}{\partial \sigma^2} \log L(\mu, \sigma^2) = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2$$

و بمساواة المعادلتين الأخيرتين بالصفر نحصل على:

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$

هو المقدّر المبني على التوقع الأعظمي للوسيط μ في المجتمع الطبيعي المذكور .

$$T_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

هو المقدّر المبني على التوقع الأعظمي للوسيط σ^2 في المجتمع الطبيعي وهو مقدّر متحيز .

مثال 3:

لتكن لدينا $X_1, X_2, \dots, X_n$ عينة من متغيرات عشوائية تتبع التوزيع البواسوني $X \sim P(m)$

(El-saadi, 1997, p. 37) حيث:

$$f(x, m) = P_m [X = x] = e^{-m} \frac{m^x}{x!}$$

وبفرض m هو الوسيط المجهول، أوجد المقدّر الأعظمي المبني على التوقع الأعظمي للوسيط m

الحل: الخطوات كما يلي:

$$L(m) = \prod_{i=1}^n f(x_i, m) = \prod_{i=1}^n e^{-m} \frac{m^{x_i}}{x_i!} = e^{-nm} \frac{m^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

$$\log L(m) = \left(\sum_{i=1}^n x_i \right) \log m - nm - \sum_{i=1}^n \log(x_i!)$$

$$\begin{cases} \frac{\partial}{\partial m} \log L(m) = 0 \\ \frac{\partial^2}{\partial m^2} \log L(m) < 0 \end{cases} \Rightarrow \begin{cases} \frac{\sum_{i=1}^n x_i}{m} - n = 0 \Rightarrow \hat{m} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x} \\ -\frac{\sum_{i=1}^n x_i}{m^2} < 0 \end{cases}$$

$T = \hat{m} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$ هو المقدّر المبني على التوقع الأعظمي للوسيط m في المجتمع البواسوني.

مثال 4:

لتكن لدينا $x_1, x_2, \dots, x_n$ عينة من متغيرات عشوائية تتبع التوزيع الأسّي حيث:

$$f(x, \lambda) = \lambda e^{-\lambda x}, \quad 0 \leq x \leq \infty, \lambda > 0$$

وبفرض λ هو الوسيط المجهول، أوجد المقدّر الأعظمي المبني على التوقع الأعظمي للوسيط λ

الحل: الخطوات باختصار كما يلي:

$$L(\lambda) = \prod_{i=1}^n f(x_i, \lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n * e^{-\lambda \sum_{i=1}^n x_i}$$

$$\log L(\lambda) = n \log \lambda - \lambda \cdot \sum_{i=1}^n x_i$$

$$\begin{cases} \frac{\partial}{\partial \lambda} \log L(\lambda) = 0 \\ \frac{\partial^2}{\partial \lambda^2} \log L(\lambda) < 0 \end{cases} \Rightarrow \begin{cases} n \left(\frac{1}{\lambda} \right) - \sum_{i=1}^n x_i = \frac{n}{\lambda} - n\bar{x} = 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{x}} \\ -\frac{n}{\lambda^2} < 0 \end{cases}$$

$T = \hat{\lambda} = \frac{1}{\bar{X}}$ هو المقدّر المبني على التوقع الأعظمي للوسيط λ في المجتمع الأصلي.

مثال 5:

نعلم من الفرضيات أن الأخطاء تتوزع توزيعاً طبيعياً: $\varepsilon_i \sim N(0, \sigma_\varepsilon^2), \forall i=1, \dots, n$

و نموذج الانحدار الخطي البسيط هو: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, \dots, n$

و نموذج الانحدار الخطي البسيط المقدّر هو: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \quad i = 1, \dots, n$

إذن دالة كثافة الأخطاء التي تتبع التوزيع الطبيعي تكتب على الشكل التالي:

$$f(\varepsilon_i) = \left(2\pi\sigma_\varepsilon^2\right)^{-\frac{1}{2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma_\varepsilon^2}\right) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \varepsilon_i^2}$$

المطلوب: أوجد المقدرات بطريقة المعقولة العظمى المبنية على التوقع الأعظمي لكل من: $\beta_1, \beta_0, \sigma_\varepsilon^2$

الحل:

لدينا تابع الامكانية العظمى $L(\varepsilon_1, \dots, \varepsilon_n)$ حيث:

$$L(\varepsilon_1, \dots, \varepsilon_n) = \prod_{i=1}^n f(\varepsilon_i) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma_\varepsilon}\right)^n \exp\left(-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \varepsilon_i^2\right)$$

نعلم أن: $\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i, \quad i = 1, \dots, n$ وعليه:

$$L(\beta_0, \beta_1, \sigma_\varepsilon^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma_\varepsilon}\right)^n \exp\left(-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n (Y_i - \beta_0 + \beta_1 X_i)^2\right)$$

$$\log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = \log\left(\frac{1}{\sqrt{2\pi}}\right)^n - \frac{n}{2} \log \sigma_\varepsilon - \frac{1}{2} \sum_{i=1}^n \left(\frac{Y_i - \beta_0 + \beta_1 X_i}{\sigma_\varepsilon}\right)^2$$

$$\log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = \log\left(\frac{1}{\sqrt{2\pi}}\right)^n - \frac{n}{2} \log \sigma_\varepsilon - \frac{1}{2} \sum_{i=1}^n \left(\frac{Y_i - \beta_0 + \beta_1 X_i}{\sigma_\varepsilon}\right)^2$$

نقوم بتقدير معالم النموذج وذلك بتعظيم لوغاريتم دالة المعقولة العظمى و هذا بالاشتقاق الجزئي مرة بالنسبة

لـ β_0 وأخرى بالنسبة لـ β_1 و أخرى بالنسبة σ_ε^2 يكون:

$$\frac{\partial}{\partial \hat{\beta}_0} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = -\frac{2}{2\hat{\sigma}_\varepsilon^2} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-1) \quad (1)$$

$$\frac{\partial}{\partial \hat{\beta}_1} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = -\frac{2}{2\hat{\sigma}_\varepsilon^2} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-X_i) \quad (2)$$

$$\frac{\partial}{\partial \hat{\sigma}_\varepsilon^2} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = \frac{-n}{2\hat{\sigma}_\varepsilon^2} + \frac{1}{2\hat{\sigma}_\varepsilon^4} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = \left(-\frac{n}{2\hat{\sigma}_\varepsilon^2} \right) + \left(\sum_{i=1}^n \hat{\varepsilon}_i^2 / 2\hat{\sigma}_\varepsilon^4 \right) \quad (3)$$

و بمساواة المعادلات الثلاث بالصفر نحصل على:

$$T_1 = \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad \text{من المعادلة الأولى:}$$

$\hat{\beta}_0$ هو المقدّر المبني على التوقع الأعظمي للوسيط β_0 في النموذج وهو مقدّر غير متحيز.

$$T_2 = \hat{\beta}_1 = \left(\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) / \sum_{i=1}^n (X_i - \bar{X})^2 \right) \quad \text{من المعادلة الثانية:}$$

$\hat{\beta}_1$ هو المقدّر المبني على التوقع الأعظمي للوسيط β_1 في النموذج وهو مقدّر غير متحيز.

$$T_3 = \hat{\sigma}_\varepsilon^2 \Leftrightarrow -n\hat{\sigma}_\varepsilon^2 = -\sum_{i=1}^n \hat{\varepsilon}_i^2 \Rightarrow \hat{\sigma}_\varepsilon^2 = \left(\sum_{i=1}^n \hat{\varepsilon}_i^2 / n \right) \quad \text{من المعادلة الثالثة:}$$

$\hat{\sigma}_\varepsilon^2$ هو المقدّر المبني على التوقع الأعظمي للوسيط σ_ε^2 في النموذج.

$$\frac{\partial^2}{\partial \hat{\beta}_0^2} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = -\frac{n}{\hat{\sigma}_\varepsilon^2} < 0$$

$$\frac{\partial^2}{\partial \hat{\beta}_1^2} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = -\frac{\sum_{i=1}^n X_i^2}{\hat{\sigma}_\varepsilon^2} < 0$$

$$\frac{\partial^2}{\partial \hat{\sigma}_\varepsilon^2} \log L(\beta_0, \beta_1, \sigma_\varepsilon^2) = \frac{n}{2\hat{\sigma}_\varepsilon^4} - \frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{\hat{\sigma}_\varepsilon^6}$$

2-1- طريقة العزوم Method of Moments

تعتبر طريقة العزوم أقدم طرق التقدير وكان أول من أشار إليها كارل بيرسون، تقوم فكرة هذه الطريقة في:

* تقدير معالم المجتمع على حساب العزوم من دالة كثافة الاحتمال للمجتمع (قد تكون العزوم حول الصفر أو

التوقع الرياضي أو قد نختار العزوم k الأولى).

* حساب العزوم المقابلة لها من بيانات العينة العشوائية التي يتم سحبها من المجتمع محل الدراسة.

* مساواة العزوم النظرية بما يقابلها من العزوم الفعلية (عاروري، 2015، الصفحات 83-84):

في المجتمع الاحصائي $\{A, X, f(x, \theta)\}$ نعلم أن العزوم من الرتبة i حول الصفر تعطى بالشكل:

$$\alpha_i = E(X^i) = \sum_i x_i p_i \quad \text{or} \quad \int_{-\infty}^{+\infty} x^i f(x) dx$$

والعزوم حول التوقع الرياضي ومن ذات الرتبة هي:

$$\mu_i = E[(X - m)^i] = \sum_i (x_i - m)^i p_i \quad ; \quad m = E(X) \quad \text{or} \quad = \int_{-\infty}^{+\infty} (x - m)^i f(x) dx$$

تعريف:

بفرض أن $X_i, i = \overline{1, n}$ متغيرات القيم الملاحظة لعينة عشوائية حجمها n ، عندئذ نعرف العزم الابتدائي حول

$$\alpha_i^s = \frac{1}{n} \sum_{i=1}^n X_i^s \quad ; s = \overline{1, k} \quad \text{لهذه العينة بالشكل:}$$

$$\text{ومنه ينتج: } \hat{\alpha}_1 = \frac{1}{n} \sum_{i=1}^n X_i^1, \quad \hat{\alpha}_2 = \frac{1}{n} \sum_{i=1}^n X_i^2, \quad \hat{\alpha}_3 = \frac{1}{n} \sum_{i=1}^n X_i^3, \dots$$

وكل منها يعتبر متغيراً إحصائياً وتعتبر مقدرًا غير متحيز ومعتقولة للوسيط θ بالاستناد على النظرية التالية

نظرية:

عزوم العينة $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_s$ هي مقدرات غير متحيزة ومعتقولة لعزوم المجتمع $\alpha_1, \alpha_2, \dots, \alpha_s$

على الترتيب، وذلك من أجل قيمة محددة لـ s وبشرط وجود $\hat{\alpha}_s$ حيث لدينا (اسبر ديب، 2009، الصفحات

:302-303):

$$E(\hat{\alpha}_s) = E\left(\frac{1}{n} \sum_i X_i^s\right) = \frac{1}{n} \sum_i E(X_i^s) = \frac{1}{n} \cdot n E(X^s) = E(X^s) = \alpha_s$$

وهذا يؤكد صفة عدم التحيز.

$$V(\hat{\alpha}_s) = V\left(\frac{1}{n} \sum_i X_i^s\right) = \frac{1}{n^2} \sum_i V(X_i^s) = \frac{1}{n^2} \cdot n V(X^s) = \frac{V(X^s)}{n} \xrightarrow{n \rightarrow \infty} 0$$

وهذا يؤكد صفة الإتساق.

مثال 1:

لتكن لدينا $X_1, X_2, \dots, X_n$ عينة من متغيرات عشوائية تتبع التوزيع الطبيعي $X \sim N(\mu, \sigma^2)$

المطلوب:

1- إيجاد مقدر الوسيط المجهول μ اعتماداً على طريقة العزوم.

2- إيجاد مقدر الوسيط المجهول σ^2 اعتماداً على طريقة العزوم.

الحل:

1- بما أن μ هو التوقع الرياضي في التوزيع الطبيعي فإنه من المناسب اختيار العزم الأول حول الصفر، وكنا قد وجدنا سابقاً أن العزم الأول حول الصفر هو:

$$\alpha_1 = E(X_i) = \mu \quad \text{- في المجتمع الطبيعي هو:}$$

$$\hat{\alpha}_1 = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{- وفي العينة هو:}$$

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} \quad \text{ينتج لدينا:} \quad \hat{\alpha}_1 = \alpha_1 \quad \text{وبوضع}$$

$$T = \bar{X} \quad \text{أي المقدّر المطلوب هو:}$$

وهو ذات المقدّر الذي حصلنا عليه بطريقة الامكانية العظمى وهو مقدّر متنسق وغير متحيز.

2- بما أن σ^2 يمثل التباين في التوزيع الطبيعي فإنه من المناسب اختيار العزم الثاني حول التوقع الرياضي أي:

$$\mu_2 = E(X - \mu)^2 = \sigma^2 \quad \text{- في المجتمع هو:}$$

$$\hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{- وفي العينة هو:}$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{ينتج لدينا:} \quad \hat{\mu}_2 = \mu_2 \quad \text{وبوضع}$$

$$T = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{أي أن المقدّر المطلوب هو:}$$

وهو مقدّر التباين في التوزيع الطبيعي بطريقة العزوم وهو نفسه الذي حصلنا عليه بطريقة الامكانية العظمى

وهو مقدّر متنسق، لكنه متحيز.

مثال 2:

استخدم طريقة العزوم لتقدير الوسيط θ في المجتمع الاحصائي الذي تابع كثافته من الشكل:

$$f(x, \theta) = \theta x^{\theta-1}; (0 < x < 1, \theta > 0)$$

الحل:

$$\alpha_1 = E(X) = \int_0^1 x \cdot f(x, \theta) dx \quad \text{- العزم الأول حول الصفر للمجتمع هو:}$$

$$= \int_0^1 x \cdot \theta x^{\theta-1} dx = \theta \int_0^1 x^{\theta} dx = \frac{\theta}{\theta+1} [x^{\theta+1}]_0^1 = \frac{\theta}{\theta+1}$$

$$\hat{\alpha}_1 = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} \quad - \text{العزم الأول حول الصفر للعينة هو:}$$

$$\frac{\theta}{\theta+1} = \bar{X} \Rightarrow \theta = \frac{\bar{X}}{1-\bar{X}} \quad \text{وبوضع: } \hat{\alpha}_1 = \alpha_1 \quad \text{ينتج لدينا:}$$

$$T = \frac{\bar{X}}{1-\bar{X}} \quad \text{أي أن المقدّر المطلوب هو:}$$

3-1- طريقة المربعات الصغرى Method of Least Square

يمكن نمذجة العلاقة بين المتغيرين Y_i و X_i على الشكل : $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, i=1, \dots, n$

حيث : Y_i يسمى بالمتغير المُفسَّر أو التابع و X_i بالمتغير المُفسِّر أو المستقل، β_0 و β_1 هما معلمتا النموذج.

أما ε_i فيمثل الخطأ في تفسير Y_i ، ومنه يمكن كتابته انطلاقاً من العلاقة: $\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i$

ويرجع وجود حد الخطأ إلى إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج

وحدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الاقتصادية ويرجع ذلك أيضاً إلى الصياغة الرياضية غير السليمة للنموذج.

فرضيات النموذج:

$$E(\varepsilon_i) = 0, \forall i = 1, \dots, n \quad *$$

$$Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_\varepsilon^2, \forall i = 1, \dots, n \quad *$$

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2), \forall i = 1, \dots, n \quad *$$

$$Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0, \forall i \neq j, i, j = 1, \dots, n \quad *$$

$$Cov(X_i, \varepsilon_i) = 0, \forall i = 1, \dots, n \quad *$$

إن طريقة المربعات الصغرى تحاول إيجاد أحسن تصحيح خطي بتدنته مربعات الانحراف (بين المشاهدات

الفعلية والمقدرة) $\sum_{i=1}^n \hat{\varepsilon}_i^2$ ، حيث: $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$ ، وهذا ما يمكن كتابته رياضياً بـ :

$$\text{Min} \sum_{i=1}^n \hat{\varepsilon}_i^2 = \text{Min}_{\hat{\beta}_0, \hat{\beta}_1} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

والشرط اللازم لتدنته هذه العلاقة هو أن تكون المشتقات الجزئية بالنسبة $\hat{\beta}_1, \hat{\beta}_0$ معدومة أي :

$$\begin{cases} \frac{\partial}{\partial \hat{\beta}_0} \sum_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0 \\ \frac{\partial}{\partial \hat{\beta}_1} \sum_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0 \end{cases}$$

بعد حل جملة المعادلتين السابقة نتحصل على تقدير معلمتي النموذج:

$$\begin{cases} \hat{\beta}_1 = \frac{n \sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n \sum_i X_i^2 - \left(\sum_i X_i \right)^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \end{cases}$$

ويكون النموذج المقدر بطريقة المربعات الصغرى (OLS) كما يلي: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

خصائص مقدرات المربعات الصغرى

إن المقدرات المتحصل عليها لكل من β_0, β_1 بطريقة المربعات الصغرى هي تقديرات نقطية $\hat{\beta}_0, \hat{\beta}_1$ حيث

توزيع المعاينة لهذه المقدرات هو:

- متوسطات هذه المقدرات هي:

$$* E(\hat{\beta}_0) = \beta_0 \text{ ومنه نقول أن } \hat{\beta}_0 \text{ مقدرة لـ } \beta_0.$$

$$* E(\hat{\beta}_1) = \beta_1 \text{ ومنه نقول أن } \hat{\beta}_1 \text{ مقدرة لـ } \beta_1.$$

- وتباينات هذه المقدرات هي:

$$* \text{تباين } \hat{\beta}_0: \text{var}(\hat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \sigma_\varepsilon^2 = \left(\frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \bar{X})^2} \right) \sigma_\varepsilon^2$$

$$* \text{تباين } \hat{\beta}_1: \text{var}(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

نلاحظ أن تباين كل مقدر غير معروف لأنه يرتبط بتباين الأخطاء النظري σ_ε^2 ، فينبغي في هذه الحالة تقدير

تباين الأخطاء للحصول على تباين البواقي:

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n-2} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2}$$

من النتائج التي توصلنا إليها عن مقدرات المربعات الصغرى يمكن أن نقول أن شكل التوزيع الاحتمالي

الخاص بالمقدرات $\hat{\beta}_0, \hat{\beta}_1$ هو:

$$\hat{\beta}_0 \sim N \left[\beta_0, \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \sigma_\varepsilon^2 \right], \quad \hat{\beta}_1 \sim N \left[\beta_1, \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right]$$

من المعادلتين يتبين أنه كلما زاد التباين σ_e^2 كلما زاد تباين المقدرات $\hat{\beta}_0, \hat{\beta}_1$ ، وكلما كان انتشار قيم X

أكبر كلما قل تباين $\hat{\beta}_0, \hat{\beta}_1$

ومن خواص مقدرات المربعات الصغرى $\hat{\beta}_0, \hat{\beta}_1$ ، لدينا:

أ- خاصية عدم التحيز:

$$E(\hat{\beta}_0) = \beta_0 \text{ ومنه نقول أن } \hat{\beta}_0 \text{ مقدرة غير متحيزة لـ } \beta_0.$$

$$E(\hat{\beta}_1) = \beta_1 \text{ ومنه نقول أن } \hat{\beta}_1 \text{ مقدرة غير متحيزة لـ } \beta_1.$$

ب- أفضل مقدر خطي غير متحيز BLUE: تنطلق هذه الفكرة من نظرية *Gauss-Markov* والتي تقول "من بين المقدرات الخطية وغير المتحيزة، تكون مقدرتا المربعات الصغرى العادية $\hat{\beta}_0$ و $\hat{\beta}_1$ أفضل مقدرتين خطيتين وغير متحيزتين، حيث أن لها أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية وغير المتحيزة الأخرى".

ج- خاصية الاتساق: نقول عن $\hat{\beta}_1$ بأنه مقدر متسق إذا كان:

- توزيع المعاينة لـ $\hat{\beta}_1$ يقترب من القيمة الحقيقية β_1 كلما كانت $n \rightarrow \infty$ ، ونقول أن النهاية الاحتمالية

$$\text{للمقدر } \hat{\beta}_1 \text{ هي } \beta_1 \text{ ونكتب: } p \lim_{n \rightarrow \infty}(\hat{\beta}_1) = \beta_1$$

- يجب أن تكون قيمتا التحيز والتباين تقتربان أو تساويان الصفر كلما اقترب n مما لا نهاية أي:

$$1. \lim_{n \rightarrow \infty} E(\hat{\beta}_1) = p \lim_{n \rightarrow \infty}(\hat{\beta}_1) = \beta_1$$

$$2. \lim_{n \rightarrow \infty} \text{var}(\hat{\beta}_1) = p \lim_{n \rightarrow \infty} \text{var}(\hat{\beta}_1) = 0$$

وبتحقق هذين الشرطين، نقول عن المقدر $\hat{\beta}_1$ بأنه مقدر متسق للمعلمة الحقيقية β_1 (شيخي، 2011،

الصفحات 19-29).

مثال 1:

X_i	3	4	1	5	7	20
Y_i	2	9	5	8	16	40
X_i^2	9	16	1	25	49	100
$(X_i - \bar{X})$	1-	0	3-	1	3	0
$(Y_i - \bar{Y})$	6-	1	3-	0	8	0
$X_i Y_i$	6	36	5	40	112	199
$(X_i - \bar{X})(Y_i - \bar{Y})$	6	0	9	0	24	39
$(X_i - \bar{X})^2$	1	0	9	1	9	20

$\hat{Y}_i$	6.05	8	2.15	9.95	13.85	40
$\hat{\varepsilon}_i = (Y_i - \hat{Y}_i)$	-4.05	1	2.85	1.95-	2.15	0
$\hat{\varepsilon}_i^2 = (Y_i - \hat{Y}_i)^2$	16.4025	1	8.1225	3.8025	4.6225	33.95

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{39}{20} = 1.95$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 8 - (1.95)4 = 0.2$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i = 0.2 + 1.95X_i$$

$$\sum_{i=1}^5 \hat{\varepsilon}_i^2 = \sum_{i=1}^5 (Y_i - \hat{Y}_i)^2 = 33.95$$

لدينا:

$$\hat{\sigma}_{\varepsilon}^2 = \frac{\sum_{i=1}^{12} \hat{\varepsilon}_i^2}{n-2} = \frac{\sum_{i=1}^{12} (Y_i - \hat{Y}_i)^2}{n-2} = \frac{33.95}{5-2} = 11.31666$$

إذا مقدرة التباين هي:

$$\text{var}(\hat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \sigma_{\varepsilon}^2 = \left(\frac{1}{5} + \frac{(4)^2}{20} \right) 11.316 = 11.316$$

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_{\varepsilon}^2}{\sum_{i=1}^{12} (X_i - \bar{X})^2} = \frac{11.316}{20} = 0.5658$$

مثال 2

البيانات التالية تعبر عن السعر وكمية الفراولة الذي تم بيعها في أحد أسواق الفواكه والخضار لمدينة عين الدفلى في

مدة 12 يوم إذا رمزنا للسعر بـ X وللكمية بـ Y

باستخدام المعادلة التالية:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, 2, 3, \dots, 12$$

$$\bar{X} = 55, \quad \sum_{i=1}^{12} (X_i - \bar{X})(Y_i - \bar{Y}) = -3150$$

$$\bar{Y} = 100, \quad \sum_{i=1}^{12} (X_i - \bar{X})^2 = 2625, \quad \sum_{i=1}^{12} (Y_i - \hat{Y}_i)^2 = 538$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{-3150}{2650} = -1.2$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 100 - (-1.2)55 = 166$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i = 166 - 1.2X_i$$

$$\text{var}(\hat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right) \sigma_\varepsilon^2 = \left(\frac{1}{12} + \frac{(55)^2}{2650} \right) \sigma_\varepsilon^2 = 1.2248 \sigma_\varepsilon^2$$

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{12} (X_i - \bar{X})^2} = \frac{\sigma_\varepsilon^2}{2625} = 0.00038 \sigma_\varepsilon^2$$

$$\sum_{i=1}^{12} \hat{\varepsilon}_i^2 = \sum_{i=1}^{12} (Y_i - \hat{Y}_i)^2 = 538 \quad \text{لدينا:}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{12} \hat{\varepsilon}_i^2}{n-2} = \frac{\sum_{i=1}^{12} (Y_i - \hat{Y}_i)^2}{n-2} = \frac{538}{12-2} = 53.8 \quad \text{إذا مقدرة التباين هي:}$$

ومنها يمكن الحصول على مقدرات تباين $\hat{\beta}_1, \hat{\beta}_0$

$$\text{var}(\hat{\beta}_0) = 1.2248 * 53.8 = 65.89$$

$$\text{var}(\hat{\beta}_1) = 0.00038 * 53.8 = 0.0204$$

2- التقدير بمجال Interval Estimation

يفضل أحيانا إعطاء تقدير للمعلمة المجهولة θ بحيث لا تكون بعيدة عن الحقيقة و الواقع و باحتمال معين و الوسيلة المستخدمة في البحث عن تقديرات موثوقة باحتمال معين ضمن مجال محدد ، و يطلق على هذا المجال الذي يحتوي فيه التقدير إسم مجال التقدير

فمن أجل الحصول على مجال التقدير للمعلمة المجهولة θ علينا أن نشكل مجالا من الشكل $\hat{\theta} \pm k$ حيث

تتعلق $\hat{\theta}$ بالعينة الخاصة المختارة و يتحدد العدد k بواسطة توزيع المعاينة بالإحصاء $\hat{\theta}$

إن قولنا أن التقدير $\hat{\theta}$ مساو تماما للمعلمة θ يعني أن:

$$\hat{\theta} - k < \theta < \hat{\theta} + k$$

و يمكن و بواسطة توزيع المعاينة للإحصاء $\hat{\theta}$ أن نحدد قيمة k بحيث يكون الاحتمال :

$$P(\hat{\theta} - k < \theta < \hat{\theta} + k)$$

مساويا لقيمة معينة تم الباحث و لتكن $1 - \alpha$ و عليه يمكن أن نعبر عن هذا المجال بما يلي :

$$\text{prob} [\theta \in (T_1, T_2)] = 1 - \alpha \quad \text{أي} \quad \text{prob} (T_1 < \theta < T_2) = 1 - \alpha$$

حيث أن: θ المعلمة المراد تقديرها، (T_1, T_2) يسمى مجال ثقة، T_1 حد الثقة الأدنى، T_2 حد الثقة الأعلى، $1-\alpha$ معامل الثقة حيث $0 < \alpha < 1$ ويحدد مسبقا، $(1-\alpha) \times 100\%$ درجة الثقة للمعلمة θ .

1-2 مجال الثقة للمتوسط الحسابي للمجتمع μ_X

1-1-2 مجال الثقة للمتوسط الحسابي في حالة توزيع المجتمع طبيعي

نفرض أن $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و أن

$\bar{X}$ يمثل الوسط الحسابي لقياسات هذه العينة لايجاد فترة الثقة للمعلمة μ

1-1-1-2 حالة تباين المجتمع معلوما (σ_X^2 معلومة)

لنفرض أن المطلوب هو إنشاء مجال ثقة للمتوسط الحسابي μ_X للتوزيع الطبيعي إذا كانت σ_X^2 معلومة

بما أن المتوسط الحسابي للعينة $\bar{X}$ كمقدر للمعلمة μ_X للتوزيع الطبيعي (التباين معلوم) له خصائص جيدة مثل الكفاءة و الكفاية وعدم التحيز فإننا سوف نستخدمه لإنشاء مجال ثقة للمتوسط الحسابي للمجتمع μ_X ضمن مجال محدد بمستوى ثقة $1-\alpha$ من خلال العلاقة :

$$\text{prob} [\mu_X \in (T_1, T_2)] = P(T_1 \leq \mu_X \leq T_2) = 1-\alpha$$

نسمي:

T_1 : حد الثقة الأدنى ، T_2 : حد الثقة الأعلى ، $1-\alpha$: مستوى الثقة

$[T_1, T_2]$: مجال الثقة لمعلمة متوسط المجتمع μ_X

رأينا سابقا ، أن المتوسط الحسابي لعينة عشوائية حجمها n من توزيع طبيعي أو أن حجم العينة كبير و

ذلك بغض النظر عن نوع المجتمع المدروس وسطه الحسابي μ_X و تباينه σ_X^2 يتبع أيضا التوزيع الطبيعي بمتوسط

حسابي $\mu_{\bar{X}}$ ولكن بتباين $\sigma_{\bar{X}}^2 = \sigma_X^2 / n$

و بالتالي فإن الكمية المحورية المناسبة هي: $Z = (\bar{X} - \mu_{\bar{X}}) / (\sigma_{\bar{X}})$

$\mu_X = \mu_{\bar{X}}$ avec et sans remise

$$\sigma_{\bar{X}}^2 = \begin{cases} \frac{\sigma_X^2}{n} & \text{avec remise} \\ \frac{\sigma_X^2}{n} \times \frac{N-n}{N-1} & \text{sans remise} \end{cases}$$

$$\frac{n}{N} \leq 0.05 \Rightarrow \frac{N-n}{N-1} \simeq 1 \Rightarrow \frac{\sigma_X^2}{n} \times \frac{N-n}{N-1} \simeq \frac{\sigma_X^2}{n}$$

$$Z = (\bar{X} - \mu_{\bar{X}}) / (\sigma_{\bar{X}}) \Rightarrow Z = (\bar{X} - \mu_X) / (\sigma_X / \sqrt{n})$$

وهي دالة في المقدّر $\bar{X}$ و المعلمة μ_X و توزيعهما هو التوزيع الطبيعي المعياري الذي لا يعتمد على المعلمة μ_X

و من ثم فإن احتمال أن تقع هذه القيمة بين القيمتين: $Z_{(1-\frac{\alpha}{2})}$ و $-Z_{(1-\frac{\alpha}{2})}$ يساوي $1-\alpha$

$$p\left(-Z_{(1-(\alpha/2))} \prec Z \prec Z_{(1-(\alpha/2))}\right) = 1 - \alpha$$

وبالتعويض عن Z نحصل على :

$$p\left(-Z_{(1-(\alpha/2))} \prec (\bar{X} - \mu_{\bar{X}})/(\sigma_{\bar{X}}) \prec Z_{(1-(\alpha/2))}\right) = 1 - \alpha$$

وبعد ضرب كل طرف من أطراف هذه المتباينة بالانحراف المعياري $\sigma_{\bar{X}}$ نجد:

$$p\left(-Z_{(1-(\alpha/2))}\sigma_{\bar{X}} \prec \bar{X} - \mu_X \prec Z_{(1-(\alpha/2))}\sigma_{\bar{X}}\right) = 1 - \alpha$$

ثم نطرح من جميع الأطراف المقدار $\bar{X}$

$$p\left(-\bar{X} - Z_{(1-(\alpha/2))}\sigma_{\bar{X}} \prec -\mu_X \prec -\bar{X} + Z_{(1-(\alpha/2))}\sigma_{\bar{X}}\right) = 1 - \alpha$$

وبعد ضرب كل أطراف هذه المتباينة بإشارة ناقص وترتيبها نجد:

$$p\left(\bar{X} - Z_{(1-(\alpha/2))}\sigma_{\bar{X}} \prec \mu_X \prec \bar{X} + Z_{(1-(\alpha/2))}\sigma_{\bar{X}}\right) = 1 - \alpha$$

وهذا ما أشرنا إليه سابقا في الصورة العامة كما يلي:

$$p(T_1 \prec \mu_X \prec T_2) = 1 - \alpha$$

حيث أن:

$$T_1 = \bar{X} - Z_{(1-(\alpha/2))}\sigma_{\bar{X}} \quad : T_1 \text{ حد الثقة الأدنى هو:}$$

$$T_2 = \bar{X} + Z_{(1-(\alpha/2))}\sigma_{\bar{X}} \quad : T_2 \text{ حد الثقة الأعلى هو:}$$

$$\mu_X \in [\bar{X} \pm Z_{(1-(\alpha/2))}\sigma_{\bar{X}}]_{1-\alpha} \quad : [T_1, T_2] \text{ مجال الثقة لمعلمة متوسط المجتمع } \mu_X \text{ هو:}$$

$$\mu_X \in [\bar{X} - Z_{(1-(\alpha/2))}\sigma_{\bar{X}} , \bar{X} + Z_{(1-(\alpha/2))}\sigma_{\bar{X}}]_{1-\alpha}$$

والمعلمة θ محل التقدير هنا تمثل المتوسط الحسابي للمجتمع μ_X ، ومن مجال الثقة السابق لـ μ_X نقول

أننا واثقون بنسبة $(1-\alpha)\%$ بأن متوسط المجتمع μ_X ينتمي إلى المجال

$$[\bar{X} - Z_{(1-(\alpha/2))}\sigma_{\bar{X}} , \bar{X} + Z_{(1-(\alpha/2))}\sigma_{\bar{X}}]$$

مثال 1:

أخذت العينة 3.7 ، 4.1 ، 3.2 ، 4 ، 5 من مجتمع طبيعي غير محدود

$$X \sim N(\mu_X, \sigma_X^2) \text{ حيث } \sigma_X^2 = 4.$$

من خلال هذه العينة أوجد مجال الثقة لمتوسط المجتمع μ_X عند مستوى ثقة 99%، 95%، 90%؟

الحل:

$$\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_{\bar{X}}, \bar{X} + Z_{(1-(\alpha/2))} \sigma_{\bar{X}} \right]_{1-\alpha} \text{ نعلم أن مجال الثقة لـ } \mu_X \text{ هو كالتالي:}$$

لدينا المجتمع غير محدود معناه حجم العينة مهممل أمام حجم المجتمع أي $0 \leq 0.05 \approx \frac{5}{\infty} = \frac{n}{N}$ و هذا معناه أن

$$\frac{N-n}{N-1} \approx \frac{N}{N} \approx 1 \text{ معامل الإرجاع يؤول إلى الواحد حيث:}$$

$$\sigma_{\bar{X}} = (\sigma_X / \sqrt{n}) = (2 / \sqrt{5}) = 0.8944 \text{ بالتالي ففي الحالتين إذا كان السحب بإرجاع أو بدون إرجاع فإن:}$$

$$\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n}, \bar{X} + Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} \right]_{1-\alpha} \text{ يصبح مجال الثقة لـ } \mu_X \text{ كالتالي:}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{3.7 + 4.1 + 3.2 + 4 + 5}{5} = \frac{20}{5} = 4$$

* تحديد قيم معاملات الثقة $Z_{(1-(\alpha/2))}$ حسب قيم مستوى المعنوية $\alpha = (1, 5, 10)\%$

$$Z_{(1-(\alpha/2))} = Z_{(1-(0.01/2))} = Z_{(1-0.005)} = Z_{(0.995)} = 2.58 \quad \alpha = 0.01 \quad *$$

$$Z_{(1-(\alpha/2))} = Z_{(1-(0.05/2))} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96 \quad \alpha = 0.05 \quad *$$

$$Z_{(1-(\alpha/2))} = Z_{(1-(0.1/2))} = Z_{(1-0.05)} = Z_{(0.95)} = 1.645 \quad \alpha = 0.10 \quad *$$

* تحديد مجال الثقة لـ μ_X حسب قيم مستوى الثقة $(0.99, 0.95, 0.90) = (1-\alpha)$

$$\mu_X \in \begin{cases} [4 - (2.58 * 0.8944), 4 + (2.58 * 0.8944)]_{0.99} \\ [4 - (1.96 * 0.8944), 4 + (1.96 * 0.8944)]_{0.95} \\ [4 - (1.645 * 0.8944), 4 + (1.645 * 0.8944)]_{0.90} \end{cases}$$

$$\mu_X \in \begin{cases} [1.692, 6.307]_{0.99} \\ [2.246, 5.753]_{0.95} \\ [2.528, 5.471]_{0.90} \end{cases}$$

نلاحظ أن مجال الثقة يتسع بزيادة قيم معامل الثقة $Z_{(1-(\alpha/2))}$ أي عند تصغير مستوى المعنوية $\alpha\%$ $\Leftrightarrow$

زيادة قيم معامل الثقة $\Leftrightarrow$ توسيع مجال الثقة $\Leftrightarrow$ زيادة الدقة في التقدير.

من مجالات الثقة السابقة لـ μ_X نقول أننا:

* واثقون بنسبة 99% بأن متوسط المجتمع μ_X ينتمي إلى المجال [1.692 , 6.307]

* واثقون بنسبة 95% بأن متوسط المجتمع μ_X ينتمي إلى المجال [2.246 , 5.753]

* واثقون بنسبة 90% بأن متوسط المجتمع μ_X ينتمي إلى المجال [2.528 , 5.471]

مثال 2:

في دراسة لإنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا، سحبنا من مجتمع المنتجين وعددهم 4519 منتج عينة عشوائية حجمها 100 منتج يشتغلون بمناطق مختلفة في الولاية، متوسط الإنتاجية المحصل عليها هي 385 قنطار في الهكتار.

- أوجد مجال الثقة لمتوسط إنتاجية فلاحي الولاية لمحصول البطاطا μ_X بمستوى الثقة 0.95 إذا علمت أن الانحراف المعياري لإنتاجية فلاحي الولاية ككل هو 100 قنطار/هكتار لمحصول البطاطا ؟

الحل:

نعلم أن مجال الثقة لـ μ_X هو كالتالي: $\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_{\bar{X}} , \bar{X} + Z_{(1-(\alpha/2))} \sigma_{\bar{X}} \right]_{1-\alpha}$

لدينا مجتمع محدود وحجمه 4519 و العينة حجمها 50 بالتالي كسر المعاينة يساوي إلى

معناه حجم العينة مهملة أمام حجم المجتمع و هذا معناه أن معامل الإرجاع $\frac{n}{N} = \frac{100}{4519} = 0.022 \leq 0.05$

يؤول إلى الواحد حيث: $\frac{N-n}{N-1} = \frac{4519-100}{4519-1} = \frac{4419}{4518} = 0.978 \approx 1$ ، بالتالي ففي الحالتين إذا كان السحب

بإرجاع أو بدون إرجاع فإن: $\sigma_{\bar{X}} = \left(\sqrt{\sigma_X^2 / n} \right) = \left(\sqrt{10000/100} \right) = \left(\sqrt{100} \right) = 10$

يصبح مجال الثقة لـ μ_X كالتالي: $\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} , \bar{X} + Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} \right]_{1-\alpha}$

* تحديد قيم معامل الثقة علما أن مستوى الثقة: $1-\alpha = 0.95$ هو:

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = 1.96$$

* تحديد مجال الثقة لـ μ_X

$$\mu_X \in [385 - (1.96 * 10) , 385 + (1.96 * 10)]_{0.95}$$

$$\mu_X \in [365.4 , 404.6]_{0.95}$$

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا μ_X ينتمي إلى

المجال [365.4 , 404.6]

2-1-1-2 حالة تباين المجتمع غير معلوم (σ_X^2 مجهولة)

في هذه الحالة فإننا نواجه معلمتين إثنين مجهولتين μ_X و σ_X^2 ، إذ غالبا ما نفكر في إيجاد تقدير للمتوسط في مجتمع إحصائي تباينه σ_X^2 مجهول، و يختلف الآن الموضوع عن الحالة السابقة إذ يجب تقدير تباين هذا المجتمع σ_X^2 من خلال تباين العينة المسحوبة S_X^2 لكي نتمكن من تقدير المتوسط الحسابي للمجتمع μ_X

الحالة الأولى: العينات الصغيرة

إذا كان تباين المجتمع مجهول والمجتمع طبيعي وحجم العينة صغير $30 < n$ ففي هذه الحالة نعتمد في تحديد مجال الثقة لمتوسط المجتمع μ_X على توزيع ستيودنت وليس على التوزيع الطبيعي بالرغم من أن توزيع المجتمع الأصلي هو التوزيع الطبيعي وهذا نظرا لكون تباين المجتمع مجهولا σ_X^2 وسوف نقوم بتقديره من خلال تباين

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{حيث: العينة المسحوبة } S_X^2$$

$$\hat{S}_x^2 = \frac{n}{n-1} S_X^2 \quad \text{حيث: } \hat{S}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{و} \quad S_x^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

تقدير σ_X^2 باستعمال تباين العينة S_X^2 :

من النظريات السابقة التي تطرقنا لها في محور نظرية المعاينة نعلم أن:

- إذا كان السحب بإرجاع:

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \quad \text{و} \quad E(S_X^2) = \sigma_X^2 \frac{n-1}{n} \quad \text{، ومنه فإن تقدير تباين المجتمع } \sigma_X^2 \text{ من خلال تباين العينة}$$

$$\hat{\sigma}_X^2 = \hat{S}_X^2 = S_X^2 \frac{n}{n-1} \quad \text{المسحوبة } S_X^2 \text{ يكون كالتالي:}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} = \frac{\hat{\sigma}_X^2}{n} = \frac{S_X^2 \frac{n}{n-1}}{n} = \frac{S_X^2}{n-1} \Rightarrow \sigma_{\bar{X}} = \frac{S_X}{\sqrt{n-1}} \quad \text{أي أن:}$$

- إذا كان السحب بدون إرجاع:

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \frac{N-n}{N-1} \quad \text{و} \quad E(S_X^2) = \sigma_X^2 \cdot \frac{n-1}{n} \cdot \frac{N}{N-1} \quad \text{، ومنه فإن تقدير تباين المجتمع } \sigma_X^2 \text{ من}$$

$$\text{خلال تباين العينة المسحوبة } S_X^2 \text{ يكون كالتالي: } \hat{\sigma}_X^2 = S_X^2 \frac{n}{n-1} \frac{N-1}{N} \quad \text{أي أن:}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \cdot \frac{N-n}{N-1} = \frac{\hat{\sigma}_X^2}{n} \cdot \frac{N-n}{N-1} = \frac{S_X^2 (n/n-1)(N-1/N)}{n} \cdot \frac{N-n}{N-1}$$

$$= S_X^2 \cdot \frac{n}{n-1} \cdot \frac{N-1}{N} \cdot \frac{N-n}{n \cdot (N-1)} = \frac{S_X^2}{n-1} \cdot \frac{N-n}{N} \Rightarrow \sigma_{\bar{X}} = \frac{S_X}{\sqrt{n-1}} \cdot \frac{\sqrt{N-n}}{\sqrt{N}}$$

$$\sigma_{\bar{X}}^2 = \begin{cases} \frac{\sigma_X^2}{n} = \frac{\hat{\sigma}_X^2}{n} = \frac{s_x^2}{n-1} & \text{avec remise} \\ \frac{\sigma_X^2}{n} \cdot \frac{N-n}{N-1} = \frac{\hat{\sigma}_X^2}{n} \cdot \frac{N-n}{N-1} = \frac{s_x^2}{n-1} \cdot \frac{N-n}{N} = & \text{sans remise} \\ \frac{n}{N} \leq 0.05 \Rightarrow \frac{N-n}{N} \simeq 1 \Rightarrow \frac{s_x^2}{n-1} \cdot \frac{N-n}{N} \simeq \frac{s_x^2}{n-1} & \end{cases}$$

تحديد الكمية المحورية المناسبة T :

توزيع t هو توزيع لمتغير يمثل حاصل قسمة متغير يتبع التوزيع الطبيعي المعياري على الجذر التربيعي لمتغير

مستقل منه يتبع توزيع χ^2 مقسوما على درجات حريته فإن: $T = \frac{Z}{\sqrt{\chi^2/n-1}}$ يتبع توزيع t بدرجات

حرية $n-1$ حيث: $Z = \left(\frac{\bar{X} - \mu_X}{\sigma_{\bar{X}}} \right) \sim N(0,1)$ و $\frac{nS_X^2}{\sigma_X^2} \sim \chi_{(n-1)}^2$ ، إذن:

$$T = \frac{Z}{\sqrt{\chi^2/n-1}} = \left(\left(\frac{\bar{X} - \mu_X}{\sigma_{\bar{X}}} \right) / \left(\left(\frac{\sqrt{n}S_X}{\sigma_X} \right) / \sqrt{n-1} \right) \right)$$

$$T = \begin{cases} \left[\left(\bar{X} - \mu_X / (\sigma_X / \sqrt{n}) \right) / \left(\left(\frac{\sqrt{n}S_X}{\sigma_X} \right) / \sqrt{n-1} \right) \right] & \text{avec remise} \\ \left[\left(\bar{X} - \mu_X / \left((\sigma_X / \sqrt{n}) \cdot \left(\sqrt{\frac{N-n}{N-1}} \right) \right) \right) / \left(\left(\frac{\sqrt{n}S_X}{\sigma_X} \right) / \sqrt{n-1} \right) \right] & \text{sans remise} \end{cases}$$

$$T = \begin{cases} \bar{X} - \mu_X / (S_X / \sqrt{n-1}) & \sim t_{(n-1)} \quad \text{avec remise} \\ \bar{X} - \mu_X / \left((S_X / \sqrt{n-1}) \cdot \left(\sqrt{\frac{N-n}{N-1}} \right) \right) & \sim t_{(n-1)} \quad \text{sans remise} \end{cases}$$

تحديد مجال الثقة لمتوسط المجتمع باستعمال توزيع ستودنت:

الكمية المحورية المناسبة T لها توزيع t بدرجات حرية $n-1$ و هي المناسبة في حالة تباين المجتمع مجهول

والتوزيع طبيعي وحجم العينة صغير، وعليه فإن:

$$P\left(-t_{(1-(\alpha/2), n-1)} \prec T \prec t_{(1-(\alpha/2), n-1)}\right) = 1 - \alpha$$

حيث أن $t_{(1-(\alpha/2), n-1)}$ هي قيمة t الجدولة و بالتعويض عن T نحصل على ما يلي:

-إذا كان السحب بإرجاع:

$$P\left(-t_{(1-(\alpha/2), n-1)} \prec \bar{X} - \mu_X / (S_x / \sqrt{n-1}) \prec t_{(1-(\alpha/2), n-1)}\right) = 1 - \alpha$$

وبعد ضرب كل طرف من أطراف هذه المتباينة $(S_x / \sqrt{n-1})$ نجد:

$$P\left(-t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} \prec \bar{X} - \mu_X \prec t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1}\right) = 1 - \alpha$$

ثم نطرح من جميع الأطراف المقدار $\bar{X}$ و بعد ضربها (-1) وترتيب المتباينة نجد :

$$P\left(\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} \prec \mu_X \prec \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1}\right) = 1 - \alpha$$

حيث أن:

$$\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} \quad \text{حد الثقة الأدنى هو:}$$

$$\bar{X} + t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} \quad \text{حد الثقة الأعلى هو:}$$

$$\mu_X \in \left[\bar{X} \pm t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1}\right]_{1-\alpha} \quad \text{مجال الثقة لمعلمة متوسط المجتمع } \mu_X \text{ هو:}$$

$$\mu_X \in \left[\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} , \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1}\right]_{1-\alpha}$$

من مجال الثقة لـ μ_X نقول أننا واثقون بنسبة $(1-\alpha)\%$ بأن متوسط المجتمع μ_X ينتمي إلى المجال

$$\left[\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1} , \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot S_x / \sqrt{n-1}\right]$$

-إذا كان السحب بدون إرجاع:

$$P\left(-t_{(1-(\alpha/2), n-1)} \prec \bar{X} - \mu_X / \left(S_x / \sqrt{n-1}\right) \left(\sqrt{N-n/N-1}\right) \prec t_{(1-(\alpha/2), n-1)}\right) = 1 - \alpha$$

وبعد ضرب كل طرف من أطراف هذه المتباينة $(S_x / \sqrt{n-1}) \left(\sqrt{N-n/N-1}\right)$ نجد:

$$P\left(-t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \prec \bar{X} - \mu_X \prec t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}}\right) = 1 - \alpha$$

ثم نطرح من جميع الأطراف المقدار $\bar{X}$ و بعد ضربها (-1) وترتيب المتباينة نجد :

$$P\left(\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \prec \mu_X \prec \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}}\right) = 1 - \alpha$$

حيث أن:

$$\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \quad \text{حد الثقة الأدنى هو:}$$

حد الثقة الأعلى هو: $\bar{X} + t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}}$

مجال الثقة لمعلمة متوسط المجتمع μ_x هو: $\mu_x \in \left[\bar{X} \pm t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha}$

$\mu_x \in \left[\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} , \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha}$

من مجال الثقة لـ μ_x نقول أننا واثقون بنسبة $(1-\alpha)\%$ بأن متوسط المجتمع μ_x ينتمي إلى المجال

$\left[\bar{X} - t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} , \bar{X} + t_{(1-(\alpha/2), n-1)} \cdot \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]$

مثال 1: نفس المثال السابق مع σ_x^2 مجهولة (المجتمع طبيعي وتباين المجتمع مجهول وحجم العينة صغير)

أخذت العينة 3.7 ، 4.1 ، 3.2 ، 4 ، 5 من مجتمع طبيعي $X \sim N(\mu_x, \sigma_x^2)$ حيث

σ_x^2 مجهولة.

من خلال هذه العينة أوجد مجال الثقة لمتوسط المجتمع μ_x عند مستوى ثقة 90% ، 95% ، 99% ؟

الحل:

نعلم أن مجال الثقة لـ μ_x في حالة المجتمع طبيعي وتباين المجتمع مجهول وحجم العينة صغير هو كالتالي:

$\mu_x \in \left[\bar{X} - t_{(1-(\alpha/2), (n-1))} \sigma_{\bar{X}} , \bar{X} + t_{(1-(\alpha/2), (n-1))} \sigma_{\bar{X}} \right]_{1-\alpha}$

$$\sigma_{\bar{X}}^2 = \begin{cases} \frac{\sigma_x^2}{n} = \frac{s_x^2}{n-1} & \text{avec remise} \\ \frac{\sigma_x^2}{n} \cdot \frac{N-n}{N-1} = \frac{s_x^2}{n-1} \cdot \frac{N-n}{N-1} & \text{sans remise} \\ \frac{n}{N} \leq 0.05 \Rightarrow \frac{N-n}{N-1} \approx 1 \Rightarrow \frac{s_x^2}{n-1} \cdot \frac{N-n}{N-1} \approx \frac{s_x^2}{n-1} \end{cases}$$

لدينا المجتمع غير محدود معناه حجم العينة مهملة أمام حجم المجتمع أي $0 \leq \frac{n}{N} \approx 0.05$ و هذا معناه

أن معامل الإرجاع يؤول إلى الواحد حيث: $\frac{N-n}{N-1} \approx \frac{N}{N} \approx 1$ ، بالتالي ففي الحالتين إذا كان السحب بإرجاع أو

بدون إرجاع فإن مجال الثقة لـ μ_x هو كالتالي:

$\mu_x \in \left[\bar{X} - t_{(1-(\alpha/2), (n-1))} S_x / \sqrt{n-1} , \bar{X} + t_{(1-(\alpha/2), (n-1))} S_x / \sqrt{n-1} \right]_{1-\alpha}$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{3.7 + 4.1 + 3.2 + 4 + 5}{5} = \frac{20}{5} = 4$$

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^5 (x_i - \bar{x})^2 = \frac{(3.7-4)^2 + (4.1-4)^2 + (3.2-4)^2 + (4-4)^2 + (5-4)^2}{4}$$

$$= \frac{0.09 + 0.01 + 0.64 + 0 + 1}{4} = \frac{1.74}{4} = 0.435$$

$$\sigma_{\bar{x}} = \left(\sigma_X / \sqrt{n} \right) = \left(s_X^2 / \sqrt{n-1} \right) = \left(0.435 / \sqrt{5-1} \right) = 0.2175$$

* تحديد قيم معاملات الثقة $t_{(1-(\alpha/2), n-1)}$ حسب قيم مستوى المعنوية α (1, 5, 10)%

$$t_{(1-(\alpha/2), n-1)} = t_{(1-(0.01/2), 5-1)} = t_{((1-0.005), 5-1)} = t_{(0.995, 4)} = 4.596 \quad ddl (n-1=4) \quad \alpha = 0.01 \quad *$$

$$t_{(1-(\alpha/2), n-1)} = t_{(1-(0.05/2), 5-1)} = t_{((1-0.025), 5-1)} = t_{(0.975, 4)} = 2.776 \quad ddl (n-1=4) \quad \alpha = 0.05 \quad *$$

$$t_{(1-(\alpha/2), n-1)} = t_{(1-(0.1/2), 5-1)} = t_{((1-0.05), 5-1)} = t_{(0.950, 4)} = 2.132 \quad ddl (n-1=4) \quad \alpha = 0.10 \quad *$$

* تحديد مجال الثقة لـ μ_X حسب قيم مستوى الثقة $(1-\alpha)$ (0.99, 0.95, 0.90)

$$\mu_X \in \begin{cases} [4 - (4.596 * 0.2175) , 4 + (4.596 * 0.2175)]_{0.99} \\ [4 - (2.776 * 0.2175) , 4 + (2.776 * 0.2175)]_{0.95} \\ [4 - (2.132 * 0.2175) , 4 + (2.132 * 0.2175)]_{0.90} \end{cases}$$

$$\mu_X \in \begin{cases} [3.00037 , 4.99963]_{0.99} \\ [3.39622 , 4.60378]_{0.95} \\ [3.53629 , 4.46371]_{0.90} \end{cases}$$

نلاحظ أن مجال الثقة يتسع بزيادة قيم معامل الثقة $t_{(1-(\alpha/2), n-1)}$ أي عند تصغير مستوى المعنوية $\alpha\%$

$\Leftrightarrow$ زيادة قيم معامل الثقة $\Leftrightarrow$ توسيع مجال الثقة $\Leftrightarrow$ زيادة الدقة في التقدير.

من مجالات الثقة السابقة لـ μ_X نقول أننا:

* واثقون بنسبة 99% بأن متوسط المجتمع μ_X ينتمي إلى المجال [3.00037 , 4.99963]

* واثقون بنسبة 95% بأن متوسط المجتمع μ_X ينتمي إلى المجال [3.39622 , 4.60378]

* واثقون بنسبة 90% بأن متوسط المجتمع μ_X ينتمي إلى المجال [3.53629 , 4.46371]

الحالة الثانية: العينات الكبيرة

إذا كان تباين المجتمع مجهول والمجتمع طبيعي وحجم العينة كبير $n \geq 30$ ففي هذه الحالة من المفروض نعتمد في تحديد مجال الثقة لمتوسط المجتمع μ_X على توزيع ستيودنت وليس على التوزيع الطبيعي بالرغم من أن توزيع المجتمع الأصلي هو التوزيع الطبيعي وهذا نظرا لكون تباين المجتمع مجهولا σ_X^2 وسوف نقوم بتقديره من خلال تباين العينة المسحوبة S_X^2 ولكن وطبقا لنظرية النهاية المركزية فإن توزيع ستيودنت يؤول إلى التوزيع

الطبيعي بشرط أن يكون حجم العينة n كبيراً، فعندما تزداد n فإن S_X^2 تقترب من σ_X^2 و يقترب التوزيع الجديد من التوزيع الطبيعي المعياري.

مجال الثقة لمعلمة متوسط المجتمع μ_X هو:

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{S_X}{\sqrt{n-1}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بإرجاع:}$$

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{S_X}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

مثال 1:

في دراسة لإنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا، سحبنا من مجتمع المنتجين وعددهم 4519 منتج عينة عشوائية حجمها 100 منتج يشتغلون بمناطق مختلفة في الولاية، متوسط الإنتاجية المحصل عليها هي 385 قنطار في الهكتار.

- أوجد مجال الثقة لمتوسط إنتاجية فلاحي الولاية لمحصول البطاطا μ_X بمستوى الثقة 0.95 إذا علمت أن توزيع المجتمع طبيعي والانحراف المعياري لإنتاجية فلاحي الولاية ككل هو 100 قنطار/هكتار لمحصول البطاطا ؟

الحل:

بما أن توزيع المجتمع طبيعي فإن مجال الثقة لـ μ_X هو كالتالي:

$$\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_{\bar{X}} , \bar{X} + Z_{(1-(\alpha/2))} \sigma_{\bar{X}} \right]_{1-\alpha}$$

لدينا مجتمع محدود وحجمه 4519 و العينة حجمها 50 بالتالي كسر المعاينة يساوي إلى

$$\frac{n}{N} = \frac{100}{4519} = 0.022 \leq 0.05 , \text{ معناه حجم العينة مهمل أمام حجم المجتمع و هذا معناه أن معامل الإرجاع}$$

$$\text{يؤول إلى الواحد حيث: } \frac{N-n}{N-1} = \frac{4519-100}{4519-1} = \frac{4419}{4518} = 0.978 \approx 1 , \text{ بالتالي ففي الحالتين إذا كان السحب}$$

$$\text{إرجاع أو بدون إرجاع فإن: } \sigma_{\bar{X}} = \left(\sqrt{\sigma_X^2 / n} \right) = \left(\sqrt{10000/100} \right) = \left(\sqrt{100} \right) = 10$$

$$\text{يصبح مجال الثقة لـ } \mu_X \text{ كالتالي: } \mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} , \bar{X} + Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} \right]_{1-\alpha}$$

* تحديد قيم معامل الثقة علماً أن مستوى الثقة : $1-\alpha = 0.95$ هو:

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = 1.96$$

* تحديد مجال الثقة لـ μ_X

$$\mu_X \in [385 - (1.96 \times 10), 385 + (1.96 \times 10)]_{0.95}$$

$$\mu_X \in [365.4, 404.6]_{0.95}$$

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحي ولاية عين الدفلى لمحصل البطاطا μ_X ينتمي إلى

$$[365.4, 404.6]$$

2-1-2 مجال الثقة للمتوسط الحسابي في حالة توزيع المجتمع غير طبيعي أو مجهول التوزيع

نفرض أن $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع غير طبيعي أو مجهول التوزيع و

أن $\bar{X}$ يمثل الوسط الحسابي لقياسات هذه العينة ليجاد فترة الثقة لمعلمة المجتمع μ_X :

1-2-1-2 حالة تباين المجتمع معلوما (σ_X^2 معلومة)

الحالة الأولى: العينات الصغيرة

إذا كان تباين المجتمع معلوم والمجتمع غير طبيعي وحجم العينة صغير $n < 30$ ففي هذه الحالة لا نستطيع الاعتماد في تحديد مجال الثقة لمتوسط المجتمع μ_X على توزيع ستودنت لأن المجتمع ليس طبيعي ولا يمكن استخدام نظرية النهاية المركزية لأن حجم العينة صغير و لكن و نظرا لكون تباين المجتمع معلوما σ_X^2 سوف نقوم بتقدير مجال الثقة لمتوسط المجتمع μ_X باستخدام نظرية تشيبيشيف كما يلي:

نظرية تشيبيشيف

في كثير من الظروف العملية عندما يكون التوزيع الاحتمالي مجهولا يكون الحد الاعلى المحسوب من متباينة تشيبيشيف مفيداً جداً بالرغم من أنه ربما يكون غير دقيق.

نفرض أن X متغير عشوائي بدالة احتمالية $f(x)$ بوسط وتباين محددين هما على التوالي μ و σ_X^2

ولنفرض أن k عدد موجب، عندئذ (هرمز، 1990، الصفحات 477-479):

$$P[(X - \mu) < k\sigma] \geq 1 - \frac{1}{k^2} \quad \text{or} \quad P[(X - \mu) \geq k\sigma] \leq \frac{1}{k^2}$$

$$\sigma^2 = E(X - \mu)^2 = \int_{-\infty}^{+\infty} (X - \mu)^2 f(x) dx$$

$$= \int_{-\infty}^{\mu - k\sigma} (X - \mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (X - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (X - \mu)^2 f(x) dx$$

$$\geq \int_{-\infty}^{\mu - k\sigma} (X - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (X - \mu)^2 f(x) dx \quad \dots (*)$$

- واضح في التكامل الأول من (*) أن $x \leq \mu - k\sigma$ وذلك يعني $\mu - x \geq k\sigma$

- وأنه في التكامل الثاني من (*) أن $x \geq \mu + k\sigma$ وذلك يعني $x - \mu \geq k\sigma$

$$\begin{aligned}
 \sigma^2 &\geq \int_{-\infty}^{\mu-k\sigma} (k\sigma)^2 f(x) dx + \int_{\mu+k\sigma}^{+\infty} (k\sigma)^2 f(x) dx \\
 &= k^2 \sigma^2 \left[\int_{-\infty}^{\mu-k\sigma} f(x) dx + \int_{\mu+k\sigma}^{+\infty} f(x) dx \right] \\
 &= k^2 \sigma^2 [P(X \leq \mu - k\sigma) + P(X - \mu \geq k\sigma)] \\
 &= k^2 \sigma^2 [P(X - \mu \leq -k\sigma) + P(X - \mu \geq k\sigma)] \\
 &= k^2 \sigma^2 P[|X - \mu| \geq k\sigma] \\
 \sigma^2 &\geq k^2 \sigma^2 P[|X - \mu| \geq k\sigma] \quad \dots(**)
 \end{aligned}$$

وبقسمة طرفي (**) على $k^2 \sigma^2$ نحصل على:

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2} \Rightarrow 1 - P[|X - \mu| \geq k\sigma] \geq 1 - \frac{1}{k^2}$$

أي أن: $P[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$ وبفرض أن أي أن $c = k\sigma > 0$ عندئذ فإن هذه المتباينة

$$P[|X - \mu| < c] \geq 1 - \frac{\sigma^2}{c^2} \Rightarrow P[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2} \quad \text{ستكون كالتالي:}$$

من المتباينة: $P[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$ نستطيع إيجاد مجال الثقة لمتوسط المجتمع كما يلي:

$$\begin{aligned}
 \mu_X &\in [\bar{X} - k \cdot \sigma_{\bar{X}}, \bar{X} + k \cdot \sigma_{\bar{X}}]_{1-\alpha} \\
 \mu_X &\in [\bar{X} - k \cdot \sigma_{\bar{X}}, \bar{X} + k \cdot \sigma_{\bar{X}}]_{1-\frac{1}{k^2}}
 \end{aligned}$$

إذن، مجال الثقة لمعلمة متوسط المجتمع μ_X هو:

$$\mu_X \in \left[\bar{X} \pm k \cdot \frac{\sigma_X}{\sqrt{n}} \right]_{1-(1/k^2)} \quad * \text{ إذا كان السحب بإرجاع:}$$

$$\mu_X \in \left[\bar{X} \pm k \cdot \frac{\sigma_X}{\sqrt{n}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-(1/k^2)} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

مثال 1:

ليكن لدينا معطيات حول إنتاجية 1500 أرض زراعية مختلفة لمحصول الحبوب بمتوسط حسابي μ وانحراف معياري قدره 5 قنطار/هكتار، أخذنا عينة عشوائية بسيطة حجمها 20 أرض زراعية مزروعة حبوبا فكان متوسط مردوديتها هو 35 قنطار/هكتار.

المطلوب: إيجاد فترة الثقة لمتوسط مردودية الحبوب عند مستوى ثقة 95%.

الحل:

بما أن المجتمع لا يخضع للتوزيع الطبيعي وحجم العينة صغير ($n = 20 < 30$) إذا لا يمكننا استخدام توزيع ستودنت وإنما نستخدم متراجحة تشيبيشيف كما يلي:

لدينا من المعطيات: $\sigma_X = 5$, $\bar{X} = 35$, $n = 20$, $1 - \alpha = 0.95$

$$P[|X - \mu| < c] \geq 1 - \frac{\sigma^2}{c^2} \Rightarrow P[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2}$$

نعلم أن مجال الثقة لـ μ_X هو كالتالي: $\mu_X \in [\bar{X} - k \cdot \sigma_{\bar{X}}, \bar{X} + k \cdot \sigma_{\bar{X}}]_{1-\alpha=1-(1/k^2)}$

لدينا مجتمع حجمه 1500 و العينة حجمها 20 بالتالي كسر المعاينة يساوي إلى

معناه حجم العينة مهمل أمام حجم المجتمع و هذا معناه أن معامل الإرجاع $\frac{n}{N} = \frac{20}{1500} = 0.013 \leq 0.05$

يؤول إلى الواحد حيث: $\frac{N-n}{N-1} = \frac{1500-20}{1500-1} = \frac{1480}{1499} = 0.987 \approx 1$ ، بالتالي ففي الحالتين إذا كان السحب

بإرجاع أو بدون إرجاع فإن: $\sigma_{\bar{X}} = (\sqrt{\sigma_X^2/n}) = (\sqrt{25/20}) = (\sqrt{1.25}) = 1.118$

يصبح مجال الثقة لـ μ_X كالتالي: $\mu_X \in [\bar{X} \pm k \cdot \frac{\sigma_X}{\sqrt{n}}]_{1-(1/k^2)}$

* تحديد قيم معامل الثقة علما أن مستوى الثقة: $1 - \alpha = 0.95$ هو:

$$1 - \alpha = 1 - \frac{1}{k^2} \Rightarrow 0.95 = 1 - \frac{1}{k^2} \Rightarrow k = \sqrt{20} = 4.472$$

* تحديد مجال الثقة لـ μ_X

$$P\left(\bar{X} - k \frac{\sigma_X}{\sqrt{n}} \leq \mu \leq \bar{X} + k \frac{\sigma_X}{\sqrt{n}}\right) = 1 - \alpha = 1 - \frac{1}{k^2}$$

$$P\left(35 - \sqrt{20} \frac{5}{\sqrt{20}} \leq \mu_X \leq 35 + \sqrt{20} \frac{5}{\sqrt{20}}\right) = 0.95$$

$$\mu_X \in [30, 40]_{0.95}$$

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحى ولاية عين الدفلى لمحصول البطاطا μ_X ينتمي إلى

المجال [30 , 40]

الحالة الثانية: العينات الكبيرة

إذا كان تباين المجتمع معلوم والمجتمع غير طبيعي وحجم العينة كبير $n \geq 30$ ففي هذه الحالة نستخدم في تحديد مجال الثقة لمتوسط المجتمع μ_X على نظرية النهاية المركزية فإن توزيع المجتمع المجهول يؤول إلى التوزيع الطبيعي المعياري.

مجال الثقة لمعلمة متوسط المجتمع μ_X هو:

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{\sigma_X}{\sqrt{n}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بإرجاع:}$$

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{\sigma_X}{\sqrt{n}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

مثال 1:

في دراسة لأوزان خرفان التسمين الموجه للاستهلاك في عيد الأضحى المبارك، سحبنا عينة عشوائية حجمها 250 خروف متوسط الوزن المحصل عليه هو 34.2 كغ.

- أوجد مجال الثقة لمتوسط وزن الخرفان μ_X بمستوى الثقة 0.88 إذا علمت أن الانحراف المعياري لوزن الخرفان للمجتمع ككل هو 6.25 كغ ؟

الحل:

لدينا مجتمع حجمه غير معروف بالتالي ففي الحالتين إذا كان السحب بإرجاع أو بدون إرجاع فإن:

$$\sigma_{\bar{X}} = \left(\sqrt{\sigma_X^2 / n} \right) = \left(\sqrt{39.0625 / 250} \right) = \left(\sqrt{0.15625} \right) = 0.39528$$

$$\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} , \bar{X} + Z_{(1-(\alpha/2))} \sigma_X / \sqrt{n} \right]_{1-\alpha} \quad \text{نعلم أن مجال الثقة لـ } \mu_X \text{ هو كالتالي:}$$

$$* \text{ تحديد قيم معامل الثقة علما أن مستوى الثقة : } 1-\alpha = 0.88 \text{ هو:}$$

$$Z_{(1-(\alpha/2))} = Z_{(0.94)} = 1.555$$

$$* \text{ تحديد مجال الثقة لـ } \mu_X$$

$$\mu_X \in [34.2 - (1.555 * 0.395) , 34.2 + (1.555 * 0.395)]_{0.88}$$

$$\mu_X \in [33.585 , 34.814]_{0.88}$$

نقول أننا واثقون بنسبة 88% بأن متوسط وزن خرفان التسمين الموجه للاستهلاك في عيد الأضحى μ_X ينتمي

إلى المجال [33.585 , 34.814]

2-2-1-2 حالة تباين المجتمع غير معلوم (σ_X^2 مجهولة)

إذا كان تباين المجتمع غير معلوم والمجتمع غير طبيعي وحجم العينة كبير $n \geq 30$ ففي هذه الحالة يجب إيجاد تقدير للمتوسط في مجتمع إحصائي تباينه مجهول، بالتالي يجب تقدير تباين هذا المجتمع من خلال تباين العينة المسحوبة لكي يتمكن من تقدير المتوسط الحسابي للمجتمع
نعتمد في تحديد مجال الثقة لمتوسط المجتمع μ_X على نظرية النهاية المركزية فإن توزيع المجتمع المجهول يؤول إلى التوزيع الطبيعي المعياري و مجال الثقة لمعلمة متوسط المجتمع μ_X هو:

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{S_x}{\sqrt{n-1}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بإرجاع:}$$

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

مثال 1:

أوجد حدي الثقة الأدنى و الأعلى بمستوى ثقة 95% لمجتمع حجمه 1500 مجهول التوزيع والتباين، إذا علمت أنه تم سحب عينة من هذا المجتمع بدون إرجاع حجمها $n=150$ فكان متوسطها الحسابي هو $\bar{x}=12$ و انحرافها المعياري $S_x=2$

الحل:

لدينا مجتمع حجمه 1500 وهو مجهول التباين والتوزيع والسحب بدون إرجاع بالتالي كسر المعاينة يساوي إلى $\frac{n}{N} = \frac{150}{1500} = 0.1 > 0.05$ ، و هذا معناه أن معامل الإرجاع لا يؤول إلى الواحد فإن مجال الثقة لـ μ_X هو

$$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \frac{S_x}{\sqrt{n-1}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\alpha} \quad \text{كالتالي:}$$

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = 1.96 \quad * \text{ تحديد قيم معامل الثقة:}$$

$$\mu_X \in \left[12 \pm \left(1.96 * \frac{2}{\sqrt{150-1}} \sqrt{\frac{1500-150}{1500-1}} \right) \right]_{0.95} \quad * \text{ تحديد مجال الثقة لـ } \mu_X :$$

نقول أننا واثقون بنسبة 95% بأن μ_X ينتمي إلى هذا المجال.

جدول يلخص مجالات الثقة لمتوسط المجتمع μ_X

توزيع المجتمع	تباين المجتمع	حجم العينة	نوع السحب	مجال الثقة
	معلوم	لا يهم	إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))} \cdot \frac{\sigma_X}{\sqrt{n}} \right]_{1-\alpha}$
			بدون إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma_X}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$
المجتمع طبيعي	غير معلوم	$n < 30$	إرجاع	$\mu_X \in \left[\bar{X} \pm t_{\left(1-\frac{\alpha}{2}, n-1\right)} \cdot \frac{S}{\sqrt{n-1}} \right]_{1-\alpha}$
			بدون إرجاع	$\mu_X \in \left[\bar{X} \pm t_{\left(1-\frac{\alpha}{2}, n-1\right)} \cdot \frac{S}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$
	معلوم	$n \geq 30$	إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \right]_{1-\alpha}$
			بدون إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$
المجتمع غير طبيعي	معلوم	$n < 30$	إرجاع	$\mu_X \in \left[\bar{X} \pm k \cdot \frac{\sigma_X}{\sqrt{n}} \right]_{1-(1/k^2)}$
			بدون إرجاع	$\mu_X \in \left[\bar{X} \pm k \cdot \frac{\sigma_X}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \right]_{1-(1/k^2)}$
		$n \geq 30$	إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma_X}{\sqrt{n}} \right]_{1-\alpha}$
	غير معلوم	$n \geq 30$	بدون إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma_X}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$
			إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \right]_{1-\alpha}$
		$n \geq 30$	بدون إرجاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$

2-2 مجال الثقة لتباين المجتمع σ_X^2 :

نفرض أن $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و نرغب

في إيجاد مجال الثقة لتباين المجتمع σ_X^2 وهنا لدينا حالتين (هرمز، 1990، الصفحات 631-633):

1-2-2 متوسط المجتمع معلوما (μ_X معلوم)

إذا كان متوسط المجتمع معلوما فإن: $S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_X)^2$ يمثل تباين العينة على أساس

متوسط المجتمع μ_X .

كما هو معلوم فإن المقدار: $\chi^2 = \frac{n \cdot S_x^2}{\sigma_X^2}$ سوف يتوزع توزيع مربع كاي بدرجة حرية n

$$\chi^2 = \frac{n \cdot S_x^2}{\sigma_X^2} = \sum_{i=1}^n \left(\frac{x_i - \mu_X}{\sigma_X} \right)^2 \sim \chi^2_{(n)}$$

لذلك يمكننا أن نكتب:

$$P \left(\chi^2_{\left(\frac{\alpha}{2}, n\right)} < \frac{n \cdot S_x^2}{\sigma_X^2} < \chi^2_{\left(1-\frac{\alpha}{2}, n\right)} \right) = 1 - \alpha$$

و بتقسيم كافة أطراف المتباينة السابقة على $n \cdot S_x^2$ ثم نأخذ مقلوب هذه المتباينة المزدوجة ، نصل إلى مجال

$$P \left(\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n\right)}} < \sigma_X^2 < \frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}} \right) = 1 - \alpha$$

الثقة لمعلمة تباين المجتمع σ_X^2 هو:

وهذا يعني أنه بالنسبة إلى مجال الثقة لتباين المجتمع المجهول σ_X^2 في مجتمع إحصائي فإن مستوى الثقة هو: $1 - \alpha$

$$\cdot \frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}} \text{ ، و حد الثقة الأعلى هو: } \frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n\right)}} \text{ ، حد الثقة الأدنى هو:}$$

$$\left[\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n\right)}} , \frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}} \right]_{1-\alpha}$$

مجال الثقة لمعلمة تباين المجتمع σ_X^2 هو:

ومنه فإن مجال الثقة للانحراف المعياري هو :

$$P \left(\sqrt{\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n\right)}}} < \sigma_X < \sqrt{\frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}}} \right) = 1 - \alpha$$

مثال:

إذا علمت أن تباين عينة عشوائية ذات حجم $n = 20$ مسحوبة من مجتمع طبيعي $N(15, \sigma_X^2)$

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2 = \frac{1}{20} \sum_{i=1}^{20} (x_i - 15)^2 = 16 \text{ هو}$$

المطلوب: أوجد مجال الثقة لتباين المجتمع بمستوى ثقة 0.95؟

الحل:

$$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \text{ and } 1 - \frac{\alpha}{2} = 0.975$$

لدينا: $n = 20$ من جدول توزيع مربع كاي بدرجة حرية

$$\chi^2_{\left(1 - \frac{\alpha}{2}, n\right)} = \chi^2_{\left(1 - \frac{0.05}{2}, 20\right)} = \chi^2_{(0.975, 20)} = 34.1696$$

$$\chi^2_{\left(\frac{\alpha}{2}, n\right)} = \chi^2_{\left(\frac{0.05}{2}, 20\right)} = \chi^2_{(0.025, 20)} = 9.59078$$

وعليه فإن:

$$\frac{n \cdot S_x^2}{\chi^2_{\left(1 - \frac{\alpha}{2}, n\right)}} = \frac{20 \times 16}{34.1696} = 9.365 \text{ حد الثقة الأدنى هو:}$$

$$\frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}} = \frac{20 \times 16}{9.59078} = 33.365 \text{ و حد الثقة الأعلى هو:}$$

$$P\left(\frac{n \cdot S_x^2}{\chi^2_{\left(1 - \frac{\alpha}{2}, n\right)}} < \sigma_x^2 < \frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n\right)}}\right) = 1 - \alpha$$

$$\sigma_x^2 \in [9.365, 33.365]_{0.95} \text{ معناه: } P(9.365 < \sigma_x^2 < 33.365) = 1 - \alpha$$

وهذا يعني أن: أي أننا واثقون بنسبة 95 % من أن معلمة تباين المجتمع σ_x^2 تنتمي إلى هذا المجال.

2-2-2 إذا كان متوسط المجتمع مجهولا (μ_x مجهول)

إذا كان متوسط المجتمع μ_x مجهولا فإن: $S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ يمثل تباين العينة على

$$\hat{S}_x^2 = \frac{n}{n-1} S_x^2 \text{ و } \bar{x} \text{ أساس متوسط العينة، حيث:}$$

$$\hat{S}_x^2 = \frac{n}{n-1} S_x^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

كما هو معلوم فإن المقدار: $\chi^2 = \frac{n \cdot S_x^2}{\sigma_x^2}$ سوف يتوزع توزيع مربع كاي بدرجة حرية $n-1$

$$\chi^2 = \frac{n \cdot S_x^2}{\sigma_x^2} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x} \right)^2 \sim \chi^2_{(n-1)}$$

لذلك يمكننا أن نكتب:

$$P \left(\chi^2_{\left(\frac{\alpha}{2}, n-1\right)} < \frac{n \cdot S_x^2}{\sigma_x^2} < \chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)} \right) = 1 - \alpha$$

حد الثقة الأدنى هو: $\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}}$ ، و حد الثقة الأعلى هو: $\frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}}$ ،

$$\left[\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}}, \frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}} \right]_{1-\alpha}$$

مجال الثقة لمعلمة تباين المجتمع σ_x^2 هو:

ومنه فإن مجال الثقة للانحراف المعياري هو :

$$P \left(\sqrt{\frac{n \cdot S_x^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}}} < \sigma_x < \sqrt{\frac{n \cdot S_x^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}}} \right) = 1 - \alpha$$

مثال:

إذا علمت أن تباين عينة عشوائية ذات حجم $n=17$ مسحوبة من مجتمع طبيعي متوسطه مجهول

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{17} \sum_{i=1}^{17} (x_i - \bar{x})^2 = 12 \text{ هو } X \sim N(\mu_x, \sigma_x^2)$$

المطلوب: أوجد مجال الثقة لتباين المجتمع بمستوى ثقة 0.90؟

الحل:

$$1 - \alpha = 0.90 \Rightarrow \alpha = 0.10 \Rightarrow \frac{\alpha}{2} = 0.05 \text{ and } 1 - \frac{\alpha}{2} = 0.95$$

لدينا:

من جدول توزيع مربع كاي بدرجة حرية $n-1=17-1=16$ لدينا:

$$\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)} = \chi^2_{\left(1-\frac{0.10}{2}, 16\right)} = \chi^2_{(0.95, 16)} = 26.2962$$

$$\chi^2_{\left(\frac{\alpha}{2}, n-1\right)} = \chi^2_{\left(\frac{0.10}{2}, 17-1\right)} = \chi^2_{(0.05, 16)} = 7.96165$$

وعليه فإن:

$$\frac{n \cdot S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}} = \frac{17 \times 12}{26.2962} = 7.757 \quad \text{حد الثقة الأدنى هو:}$$

$$\frac{n \cdot S_X^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}} = \frac{17 \times 12}{7.96165} = 25.6228 \quad \text{و حد الثقة الأعلى هو:}$$

$$P\left(\frac{n \cdot S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}} < \sigma_X^2 < \frac{n \cdot S_X^2}{\chi^2_{\left(\frac{\alpha}{2}, n-1\right)}}\right) = 1-\alpha \quad \text{مجال الثقة لمعلمة تباين المجتمع } \sigma_X^2 \text{ هو:}$$

$$\sigma_X^2 \in [7.757, 25.6228]_{0.90} \quad \text{معناه: } P(7.75 < \sigma_X^2 < 25.62) = 1-\alpha$$

أي أننا واثقون بنسبة 90 % من أن معلمة تباين المجتمع σ_X^2 تنتمي إلى هذا المجال.

3-2 مجال الثقة للنسبة

إذا كان المطلوب هو تقدير النسبة p فإننا نفترض أن المجتمع له التوزيع الثنائي و المشكلة هي تقدير معلمة هذا التوزيع p باعتماد p' و الذي يمثل النسبة في العينة المسحوبة من المجتمع محل الدراسة و الخاضع للتوزيع الثنائي و ذلك ضمن مجال محدد بمعامل الثقة $1-\alpha$ أي احتمال وقوع تلك القيمة p ضمن المجال

$$P[p \in (T_1, T_2)] = 1-\alpha \quad \text{أي: } (T_1, T_2)$$

نعلم أنه و في حالة العينات الكبيرة يمكن تقريب التوزيع الثنائي بالتوزيع الطبيعي، فإذا كان المتغير العشوائي (عدد مرات النجاح في العينة) هو X و كانت العينة كبيرة و حجمها n فإن:

$$E\left(\frac{X}{n}\right) = E(\hat{p}) = \frac{1}{n} E(X) = \frac{1}{n} \cdot n \cdot p = p \quad \text{ومنه } E(X) = \mu_X = n \cdot p$$

$$V\left(\frac{X}{n}\right) = V(\hat{p}) = \frac{1}{n^2} V(X) = \frac{1}{n^2} \cdot n \cdot p \cdot q = \frac{p \cdot q}{n} \quad \text{ومنه } V(X) = n \cdot p \cdot q \quad \text{كذلك}$$

$$Z = \frac{(X - n \cdot p)}{\sqrt{n \cdot p \cdot q}} \quad \text{و إذا كانت } n \text{ كبيرة و وفقا لنظرية النهاية المركزية فإن التوزيع التقريبي لـ } X \text{ هو:}$$

$$P(-Z_{(1-(\alpha/2))} < Z < Z_{(1-(\alpha/2))}) = 1-\alpha \quad \text{له تقريبا توزيع طبيعي معياري و بالتالي فإن:}$$

و بالتعويض بقيمة Z نحصل على $1 - \alpha = P \left(-Z_{(1-(\alpha/2))} \leq \frac{(X - n.p)}{\sqrt{n.p.q}} \leq Z_{(1-(\alpha/2))} \right)$

و بالتالي فإن: $1 - \alpha = P \left(p' - Z_{(1-(\alpha/2))} \sqrt{p.q/n} \leq p \leq p' + Z_{(1-(\alpha/2))} \sqrt{p.q/n} \right)$

رغم توفر المعلومات عن p' و n فإنه لا يمكن إيجاد قيمتي المجال و ذلك لعدم معرفتنا القيمة p و q (وهما

مجهولان) الموجودة في الانحراف المعياري $\sigma_p^2 = \frac{p.q}{n}$ و لهذا يجب البحث عن قيمة p هذه .

في حقيقة الأمر هناك عدة طرق تستخدم من أجل إزالة هذه الصعوبة نذكر أبسط طريقة و هي طريقة تقريب

p بواسطة p'

نحن نعلم بأن p' يمكن أن تستعمل كمقدر فعال و بدون تحيز ل p و هذه الخواص هي التي تقودنا إلى تقدير

p (نسبة المجتمع) ضمن المجال المذكور أعلاه و بناء على ذلك فإنه من المنطقي أن نستعمل p' أيضا من

أجل تقدير p و من ثم q الموجودة تحت الجذر أي في عبارة

الانحراف المعياري $\sqrt{\frac{p.q}{n}}$ فإن الخطأ المرتكب سيكون صغيرا جدا من أجل قيمة كبيرة ل n

و يجدر الإنتباه إلى أن تغير $\sqrt{p.q}$ طفيف جدا من أجل قيم p قريبة من 0.5 لذلك فإنه وبإحلال التقدير

النقطي $p' = \frac{X}{n}$ مكان p و إدخال p' داخل الجذر التربيعي يمكننا من الوصول إلى مجال الثقة للنسبة p

كما يلي:

$$P \left(p' - Z_{(1-(\alpha/2))} \sqrt{p'(1-p')/n} \leq p \leq p' + Z_{(1-(\alpha/2))} \sqrt{p'(1-p')/n} \right) = 1 - \alpha$$

و بالتالي فإن :

حد الثقة الأدنى هو: $\left(p' - Z_{(1-(\alpha/2))} \sqrt{p'(1-p')/n} \right)$ ،

و حد الثقة الأعلى هو: $\left(p' + Z_{(1-(\alpha/2))} \sqrt{p'(1-p')/n} \right)$ ،

مجال الثقة لمعلمة النسبة في المجتمع p هو:

$$\left[p' - Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{p'(1-p')}{n}} , p' + Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{p'(1-p')}{n}} \right]_{1-\alpha}$$

إن المجال السابق و هو $(1-\alpha) \times 100\%$ مجال الثقة ل p محسوبا من عينة عشوائية حجمها n حيث $(n \geq 30)$ تم سحبها من مجتمع ثنائي .

ملاحظة :

بفرض أن عملية السحب كانت تتم بدون إرجاع و من مجتمع صغير (منته) ففي هذه الحالة يجب أن نأخذ

مؤشر عدم الإعادة بعين الاعتبار $\left(\frac{N-n}{N-1}\right)$ و عندها فإن العلاقة السابقة تصبح كما يلي

$$\text{prob} \left(P' - Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{P'(1-P')}{n}} \sqrt{\left(\frac{N-n}{N-1}\right)} < p < P' + Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{P'(1-P')}{n}} \sqrt{\left(\frac{N-n}{N-1}\right)} \right) = 1-\alpha$$

المحور الرابع

نظرية إختبار الفرضيات

Hypothesis Testing Theory

نظرية إختبار الفرضيات

Hypothesis Testing Theory

إن إختبار الفروض طريقة عامة وشاملة لاتخاذ قرار بقبول أو رفض فرضية ما حول توزيع المجتمع، ويتم ذلك عن طريق سحب عينة من المجتمع والحصول منها على تقدير الثابت الإحصائي المفترض، وفي غالب الأحوال سيكون هناك فرق بين القيمة المفروضة للثابت الإحصائي وبين قيمته المقدرة من العينة ويشكل هذا الفرق محور الدراسة عند نظرية الإختبار

1- مفاهيم أساسية في نظرية الإختبار

الفرضية الإحصائية statistical hypothesis

الفرضية الإحصائية بمفهومها العام تعني إدعاء أو إفتراض حول توزيع متغير عشوائي أو متغيرات عشوائية، وفي الغالب فإن الفرضية الإحصائية تكون حول معلمة من معالم المجتمع ويكون المطلوب اختبار صحة هذا الادعاء أو التخمين وتساعدنا هذه الفرضية في اتخاذ القرارات الرشيدة.

أنواع الفروض الإحصائية

يمكن التمييز بين أنواع مختلفة من الفروض من زوايا مختلفة فالفرض قد يكون بسيط أو مركب.

أ- الفرض البسيط

الفرض البسيط هو الذي يحدد قيمة واحدة فقط لمعلمة المجتمع المجهولة، مثلا الفرضيتين

$$H_0 : \mu_X = 100, H_1 : \mu_X = 95$$

ب- الفرض المركب

الفرض المركب هو الذي يحدد أية قيمة لمعلمة المجتمع في مجموعة من القيم المعلمة المجتمع المجهولة، مثلا

$$H_0 : \mu_X \leq 100, H_1 : \mu_X > 65$$

أصناف الفروض الإحصائية

أما إذا نظرنا إلى الفرض الذي يجري إختباره فإننا نقسم الفروض إلى صنفين **فرضية العدم** و**الفرضية البديلة**.

$$H_0 : \theta = \theta_0 \quad \text{null hypothesis} \quad \text{أ- فرضية العدم}$$

تسمى الفرضية التي يجري إختبارها بفرضية العدم وتصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو

عدم وجود تغير، أي تعكس عدم الاختلاف وبالتالي تظهر فيها عبارة التساوي، ويرمز لها بـ H_0 .

ب- الفرضية البديلة H_1 alternative hypothesis

إن الفرضية التي تمثل البديل لفرضية العدم في حالة رفضها تسمى الفرضية البديلة، وهو الفرض الذي يجب أن يكون صحيحا اذا كان فرض العدم غير صحيح، ويرمز لها بـ H_1

حالات الفرضية البديلة:

إن الفرضية البديلة H_1 يمكن أن تأخذ أحد الأشكال الثلاثة التالية :

أ - الشكل الأول: $H_1: \theta > \theta_0$

ومن هنا نجد أن منطقة الرفض تقع بكاملها على يمين منطقة القبول .

ب - الشكل الثاني: $H_1: \theta < \theta_0$

هنا نجد أن منطقة الرفض تقع بكاملها على يسار منطقة القبول .

يعرف هذين النوعين من الفرضية البديلة بالفرض أحادي الطرف

ج - الشكل الثالث: $H_1: \theta \neq \theta_0$

هنا نجد أن منطقة الرفض تقع على يمين ويسار منطقة القبول بشكل متناظر ويعرف هذا النوع من الفرضية البديلة بالفرض ثنائي الطرف.

إختبار الفرضية

إختبار الفرضية هو قاعدة أو إجراء يؤدي إلى رفض أوعدم رفض فرضية العدم ويتم إختبار الفرضية بتقسيم فضاء العينة لكل النتائج الممكنة للتجربة العشوائية لجزئين غير متداخلين أحدهما للنتائج التي إذا حدثت لا نرفض فرضية العدم والأخرى للنتائج إذا حدثت نرفضه . ويسمى هذان القسمان منطقة القبول ومنطقة الرفض بالترتيب . يطلق على القيمة التي تفصل بين منطقة الرفض ومنطقة القبول القيمة الحرجة قد تكون متواجدة على جانب واحد ويسمى إختبار أحادي الجانب وقد تكون متواجدة على الجانبين ويسمى إختبار ثنائي الجانب، وفقا لارتفاع وانخفاض مستوى المعنوية تتغير كل من منطقة الرفض والقبول .

أنواع الأخطاء

هناك نوعان من الأخطاء : خطأ من النوع الأول وخطأ من النوع الثاني (العلي و صبح، 2020، صفحة

(69) .

أ- خطأ من النوع الأول

إذا رفضنا فرضية العدم بينما هي فرضية صحيحة فإننا هنا نكون قد إرتكبنا خطأ نسميه خطأ من النوع الأول ونرمز إلى إحتمال الخطأ من النوع الأول بـ α وهو مستوى الدلالة α : هو إحتمال رفض فرض العدم بالرغم من أنه في الواقع صحيح .

ب- خطأ من النوع الثاني

قد يحدث ألا نرفض فرضية العدم بينما هي في الحقيقة خطأ هنا فإننا نكون قد إرتكبنا خطأ نسميه خطأ من النوع الثاني. ونرمز إلى إحتمال الخطأ من النوع الثاني بـ β حيث β : هو إحتمال قبول فرض العدم بالرغم من أنه في الواقع غير صحيح .
ونلخص ما سبق إلى إبراز علاقة القرار الإحصائي بالفرضية بالجدول الموالي :

حقيقة الفرضية	القرار	قبول الفرضية	رفض الفرضية
صحيحة	قرار صحيح واحتماله $1-\alpha$	خطأ من النوع الأول واحتماله α	
غير صحيحة	خطأ من النوع الثاني واحتماله β	قرار صحيح واحتماله $1-\beta$	

وهذا ما يمكن أن نعبر عنه بما يلي : نتيجة القرارات الممكنة

النتيجة	القرارات الممكنة	نوع الفرضية	
		الفرضية البديلة	فرضية العدم
القرار صائب	قبول فرضية العدم	غير صحيح	صحيح
القرار خاطئ	رفض فرضية العدم	غير صحيح	صحيح
القرار خاطئ	قبول فرضية العدم	صحيح	غير صحيح
القرار صائب	رفض فرضية العدم	صحيح	غير صحيح

2- إختبارات الفرضيات حول معالم المجتمع الإحصائي

إن عملية المقارنة لكمية ما تم تقديرها باعتماد المعلومات التي تقدمها العينة بقيمة ثابتة محددة مسبقا وتم إستيقائها من المجتمع ومن الأمور الشائعة جدا في الحياة العملية وخاصة في المصانع التي تعمل على مراقبة الإنتاج من وقت لآخر من خلال إعتداد العينات .

فالمسألة تتمثل إذا في مقارنة قيمة صفة ما من الصفات محل الدراسة ولتكن θ والتي يمكن إستخلاص قيمتها من عينة يتم سحبها من المجتمع المعني بالدراسة مع قيمة ثابتة تعتبر معيارية ولتكن θ_0 وذلك من خلال إفتراض فرضية العدم H_0 وفريضة بديلة لها H_1 ضمن مستوى دلالة محدد سلفا

1-2- إختبارات الفرضيات حول المتوسط الحسابي لمجتمع طبيعي μ_X

يتم إجراء إختبارات الفرضيات حول المتوسط الحسابي للمجتمع في حالتين مختلفتين، أننا نعرف قيمة الانحراف المعياري، أننا لا نعرف الانحراف المعياري للمجتمع وعندها نعمل على تقديره باستخدام الانحراف المعياري للعينة

1-1-2 تباين المجتمع معلوم σ_X^2

نظرية 1:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و كان تباين المجتمع σ_X^2 معلوما فإن دالة الاختبار هي (أبو صالح و عوض، 1983، الصفحات 167-168):

$$Z = \frac{\bar{X} - \mu_0}{\sigma_x / \sqrt{n}} \quad \text{و إذا كانت } H_0: \mu = \mu_0$$

$$* \text{ وكانت } H_1: \mu \neq \mu_0$$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $|Z| \geq Z_{(1-(\alpha/2))}$

$$\text{أي: } Z \leq -Z_{(1-(\alpha/2))} \quad \text{or} \quad Z \geq Z_{(1-(\alpha/2))}$$

$$* \text{ وكانت } H_1: \mu > \mu_0$$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z > Z_{(1-(\alpha/2))}$

$$* \text{ وكانت } H_1: \mu < \mu_0$$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z < -Z_{(1-(\alpha/2))}$

مثال:

إذا كانت $X_1, X_2, \dots, X_{49}$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N(\mu, 16)$

وكان $\sum_{i=1}^{49} X_i = 441$ و أردنا اختبار $H_0: \mu = \mu_0 = 7$ مقابل $H_0: \mu \neq \mu_0 \neq 7$ على مستوى الدلالة $\alpha = 0.05$ فإننا نجد :

$$Z = \frac{\bar{x} - \mu_0}{\sigma_x / \sqrt{n}} = \frac{(441/49) - 7}{4/\sqrt{49}} = 3.5$$

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = +1.96$$

ولما كانت: $(Z = 3.5) > (Z_{(0.975)} = 1.96)$ ، فإننا نرفض $H_0: \mu = \mu_0 = 7$ على مستوى الدلالة $\alpha = 0.05$

2-1-2 تباين المجتمع غير معلوم أي مجهول σ_x^2

عندما يكون تباين المجتمع مجهولا فإننا نستخدم القيمة التقديرية له s المحسوبة من المشاهدات في العينة

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2$$

والمعرف كما يلي:

فإذا كان عدد الوحدات الإحصائية للعينة أكبر من 30 وحدة فإن المؤشر t يخضع لقانون التوزيع الطبيعي
أما إذا كان عدد الوحدات الإحصائية للعينة أصغر من 30 وحدة فإن المؤشر t يخضع لقانون توزيع ستودنت ذو درجات حرية مساويا إلى $n - 1$

1-2-1-2 العينات الصغيرة

إذا كان تباين المجتمع مجهول والمجتمع طبيعي وحجم العينة صغير $30 < n$ ففي هذه الحالة نعتد في تحديد قيمة دالة الاختبار T على توزيع ستودنت وليس على التوزيع الطبيعي بالرغم من أن توزيع المجتمع الأصلي هو التوزيع الطبيعي وهذا نظرا لكون تباين المجتمع مجهولا σ_x^2 وسوف نقوم بتقديره من خلال تباين العينة

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ و نعلم أن: } \hat{S}_x^2 = \frac{n}{n-1} S_x^2$$

قيمة دالة الاختبار T هو:

$$T = \frac{\bar{x} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{x} - \mu_0}{\hat{S}_x / \sqrt{n}}$$

* إذا كان السحب بإرجاع:

$$T = \frac{\bar{X} - \mu_0}{(S_x / \sqrt{n-1}) \cdot (N - n / N - 1)} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

نظرية 2:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و كان تباين المجتمع σ_X^2 مجهولاً فإن دالة الإختبار هي:

$$T = \frac{\bar{X} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{X} - \mu_0}{\hat{S}_x / \sqrt{n}} : t_{n-1}$$

و إذا كانت $H_0 : \mu = \mu_0$

* وكانت $H_1 : \mu \neq \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $|T| \geq t_{(1-(\alpha/2), n-1)}$

أي: $T \leq -t_{(1-(\alpha/2), n-1)} \text{ or } T \geq t_{(1-(\alpha/2), n-1)}$

* وكانت $H_1 : \mu > \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $T > t_{(1-(\alpha/2), n-1)}$

* وكانت $H_1 : \mu < \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $T < -t_{(1-(\alpha/2), n-1)}$

مثال:

إذا كانت $X_1, X_2, \dots, X_{26}$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و كان

$$\bar{X} = 65 \text{ و } S_x^2 = 9$$

وأردنا إختبار $H_0 : \mu = \mu_0 = 60$ مقابل $H_1 : \mu \neq \mu_0 \neq 60$ على مستوى الدلالة $\alpha = 0.05$ فإننا نجد

$$T = \frac{\bar{X} - \mu_0}{s_x / \sqrt{n-1}} = \frac{65 - 60}{3 / \sqrt{26-1}} = \frac{5}{0.6} = 8.33$$

$$t_{(1-(\alpha/2), n-1)} = t_{(0.975, 25)} = +2.06$$

ولما كانت: $(T = 8.33) > (t_{(0.975, 25)} = 2.06)$ ، فإننا نرفض $H_0 : \mu = \mu_0 = 60$ على مستوى الدلالة

$$\alpha = 0.05$$

2-2-1-2 العينات الكبيرة

إذا كان تباين المجتمع مجهول والمجتمع طبيعي وحجم العينة كبير $n \geq 30$ ففي هذه الحالة من المفروض نعتمد في تحديد قيمة دالة الاختبار T على توزيع ستيودنت وليس على التوزيع الطبيعي بالرغم من أن توزيع المجتمع الأصلي هو التوزيع الطبيعي وهذا نظرا لكون تباين المجتمع مجهولا σ_X^2 وسوف نقوم بتقديره من خلال تباين العينة المسحوبة S_X^2 ولكن وطبقا لنظرية النهاية المركزية فإن توزيع ستيودنت يؤول إلى التوزيع الطبيعي بشرط أن يكون حجم العينة n كبيرا، فعندما تزداد n فإن S_X^2 تقترب من σ_X^2 و يقترب التوزيع الجديد من التوزيع الطبيعي المعياري.

قيمة دالة الاختبار Z هو:

$$Z = \frac{\bar{X} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{X} - \mu_0}{\hat{S}_x / \sqrt{n}} \quad * \text{ إذا كان السحب بإرجاع:}$$

$$Z = \frac{\bar{X} - \mu_0}{(S_x / \sqrt{n-1}) \cdot (N - n / N - 1)} \quad * \text{ إذا كان السحب بدون إرجاع:}$$

نظرية 3:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$ و كان تباين المجتمع σ_X^2 مجهولا و وحجم العينة كبيرا $n \geq 30$ فإن دالة الاختبار هي:

$$Z = \frac{\bar{X} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{X} - \mu_0}{\hat{S}_x / \sqrt{n}}$$

و إذا كانت $H_0: \mu = \mu_0$

* وكانت $H_1: \mu \neq \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $|Z| \geq Z_{(1-(\alpha/2))}$

أي: $Z \leq -Z_{(1-(\alpha/2))} \text{ or } Z \geq Z_{(1-(\alpha/2))}$

* وكانت $H_1: \mu > \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z > Z_{(1-(\alpha/2))}$

* وكانت $H_1: \mu < \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z < -Z_{(1-(\alpha/2))}$

مثال:

إذا كانت $X_1, X_2, \dots, X_{82}$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N(\mu, \sigma^2)$

وكان $\bar{x} = 115$ و $S_x^2 = 25$ وحجم العينة $n = 82$

وأردنا اختبار $H_0: \mu = \mu_0 = 100$ مقابل $H_0: \mu \neq \mu_0 \neq 100$ على مستوى الدلالة $\alpha = 0.05$ فإننا نجد :

$$Z = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n-1}} = \frac{115 - 100}{5 / \sqrt{82-1}} = \frac{15}{5/9} = \frac{15.9}{5} = 27$$

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = +1.96$$

ولما كانت: $(Z = 27) > (Z_{(0.975)} = 1.96)$ ، فإننا نرفض $H_0: \mu = \mu_0 = 100$ على مستوى الدلالة

$$\alpha = 0.05$$

2-2 إختبارات الفرضيات حول مقارنة وسطين حسابيين μ_X

في كثير من الأحيان يطلب منا دراسة متوسطي عینتين لمعرفة ما إذا كان الفرق بينهما كبيراً بدرجة يمكننا من القول بأنهما مسحوبتين من مجتمعين مختلفين ، أم الفرق صغير بدرجة يمكن إرجاعها للصدفة .

ليكن لدينا مجتمعين X و Y كبيرين كبراً كافياً لهما متوسطين حسابيين μ_X و μ_Y و تباينين σ_1^2 و σ_2^2 نسحب عینتين عشوائيتين x و y من المجتمعين X و Y ، العينة الأولى بحجم n_1 لها متوسط حسابي قدره $\bar{X}$ و العينة الثانية بحجم n_2 لها متوسط حسابي قدره $\bar{Y}$

المطلوب هو معرفة ما إذا كان هناك فرق جوهري بين الوسطين المجهولين μ_X و μ_Y اعتماداً على الوسطين $\bar{X}$ و $\bar{Y}$ المستخلصين من العینتين العشوائيتين x و y المسحوبتين من المجتمعين X و Y إن الحالات التي يمكن أن تعترضنا عند إجراء الإختبار تتمثل بما يلي

- الإنحرافين المعياريين للمجتمعين مقدرين .
- الإنحرافين المعياريين للمجتمعين متساويين .
- حجم العينات المسحوبة كبير .
- حجم العينات المسحوبة صغير .

و هذه الحالات تتشابه في خطوات الإختبار و تختلف عن بعضها في بعض النقاط و لذا

في البداية سنكتفي كخطوة أولى بالتطرق إلى الحالة العامة لهذا النوع من الإختبارات .

1-2-2 العينتان كبيرتان: ($n_1 \geq 30 : n_2 \geq 30$)

1-1-2-2 الانحرافان المعياريان للمجتمعان معلومان

إذا كان حجم العينتين كبيراً فإننا نعلم حسب نظرية توزيع المعاينات أن قوانين الاحتمال لكل من $\bar{X}$ و $\bar{Y}$ هي:

$$\bar{Y} \sim N \left(\mu_2, \sigma_2^2 / \sqrt{n_2} \right) , \quad \bar{X} \sim N \left(\mu_1, \sigma_1^2 / \sqrt{n_1} \right)$$

ومنه فإن قانون الاحتمال للمتغير العشوائي الناتج عن الفرق $(\bar{X} - \bar{Y})$ هو:

$$(\bar{X} - \bar{Y}) \sim N \left(\mu_1 - \mu_2 ; \sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)} \right)$$

ويتضمن إختبار مقارنة الوسطين الإختبار بين الفرضيتين التاليتين :

$$H_0 : \mu_1 - \mu_2 = 0 \Rightarrow \mu_1 = \mu_2 \quad \text{فرضية العدم} :$$

$$H_1 : \mu_1 - \mu_2 \neq 0 \Rightarrow \mu_1 \neq \mu_2 \quad \text{الفرضية البديلة} :$$

فإذا كانت الفرضية الإبتدائية صحيحة أي إذا كان $\mu_1 = \mu_2$ فيكون قانون الاحتمال للفرق هو

$$(\bar{X} - \bar{Y}) \sim N \left(0 ; \sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)} \right)$$

و مؤشر الإختبار Z يخضع لقانون التوزيع الطبيعي المعياري $N(0,1)$ و يعرف بالعلاقة التالية :

$$Z = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}}$$

نظرية 4:

إذا كانت $X_1, X_2, \dots, X_{n1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت

$Y_1, Y_2, \dots, Y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين

مستقلتين عن بعضهما البعض ، وكانت σ_1^2 و σ_2^2 معلومتين

أردنا إختبار الفرضية $H_0 : \mu_1 = \mu_2$ أي أن $H_0 : \mu_1 - \mu_2 = 0$ فإن دالة الإختبار هي:

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}}$$

* فإذا كانت $H_1 : \mu_1 \neq \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$Z \leq -Z_{(1-(\alpha/2))} \quad \text{or} \quad Z \geq Z_{(1-(\alpha/2))} \quad \text{أي} \quad |Z| = \left| \frac{(\bar{X} - \bar{Y})}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}} \right| \geq Z_{(1-(\alpha/2))}$$

* وإذا كانت $H_1: \mu_1 > \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(\bar{x} - \bar{y})}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} > Z_{(1-(\alpha/2))}$$

* وكانت $H_1: \mu_1 < \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(\bar{x} - \bar{y})}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} < -Z_{(1-(\alpha/2))}$$

مثال:

إذا كانت $x_1, x_2, \dots, x_{25}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, 25)$ ، وكانت $y_1, y_2, \dots, y_{16}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, 36)$ ، وكانت هاتين العينتين مستقلتين، وكانت $\bar{x} = 8.9$ و $\bar{y} = 11.2$ وأردنا اختبار الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية

$H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$

الحل:

$$Z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} = \frac{8.9 - 11.2}{\sqrt{(25/20) + (36/16)}} = \frac{-2.3}{\sqrt{3.5}} = -1.229$$
 دالة الاختبار هي:

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = +1.96$$

ولما كانت: $(|Z| = |-1.223|) < (Z_{(0.975)} = 1.96)$ ، فإننا نقبل $H_0: \mu_1 = \mu_2$ على مستوى الدلالة

$\alpha = 0.05$

2-1-2-2 الانحرافان المعياريان للمجتمعان متساويان و مجهولان $\sigma_1 = \sigma_2 = \sigma$

عندما يكون الانحرافان المعياري للمجتمعين متساويين و مجهولين أي $\sigma_1 = \sigma_2 = \sigma$ فإن :

$$\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)} = \sigma \sqrt{(1/n_1) + (1/n_2)}$$

$$S^2 = \frac{(S_1^2 n_1) + (S_2^2 n_2)}{n_1 + n_2 - 2}$$
 نعلم أن التقدير النقطة ل σ^2 هو:

$$s_2^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \quad \text{و} \quad s_1^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 \quad \text{حيث}$$

وتكون الكمية $\sigma \sqrt{(1/n_1) + (1/n_2)}$ مقدرة بواسطة $S \sqrt{(1/n_1) + (1/n_2)}$

نظرية 5:

إذا كانت $x_1, x_2, \dots, x_{n_1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت $y_1, y_2, \dots, y_{n_2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض ، وكانت $\sigma_1^2 = \sigma_2^2$ ولكنهما مجهولتين وأردنا إختبار الفرضية $H_0 : \mu_1 = \mu_2$ أي أن $H_0 : \mu_1 - \mu_2 = 0$ فإن دالة الإختبار هي:

$$T = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{S \sqrt{(1/n_1) + (1/n_2)}} ; t_{(1-(\alpha/2), n_1+n_2-2)}$$

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

* فإذا كانت $H_1 : \mu_1 \neq \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|T| = \left| \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} \right| \geq t_{(1-(\alpha/2), n_1+n_2-2)}$$

أي: $T \leq -t_{(1-(\alpha/2), n_1+n_2-2)}$ or $T \geq t_{(1-(\alpha/2), n_1+n_2-2)}$

* وإذا كانت $H_1 : \mu_1 > \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} > t_{(1-(\alpha/2), n_1+n_2-2)}$$

* وكانت $H_1 : \mu_1 < \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} < -t_{(1-(\alpha/2), n_1+n_2-2)}$$

مثال:

أخذنا عينتين من توزيعين مستقلين $x_1, x_2, \dots, x_{11}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت $y_1, y_2, \dots, y_9$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت $\bar{x} = 11.3$ و $\bar{y} = 10.7$ مع $\sigma_1^2 = \sigma_2^2$ و $S_1^2 = 9$ و $S_2^2 = 12$ ، وأردنا إختبار الفرضية $H_0 : \mu_1 = \mu_2$ مقابل الفرضية $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$

الحل:

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} = \frac{(11 \times 9) + (9 \times 12)}{11 + 9 - 2} = \frac{207}{18} = 11.5 \text{ إذن } \sigma_1^2 = \sigma_2^2$$

دالة الاختبار هي:

$$T = \frac{(\bar{x}_1 - \bar{x}_2)}{S \sqrt{(1/n_1) + (1/n_2)}} = \frac{11.3 - 10.7}{\sqrt{11.5 \times [(1/11) + (1/9)]}} = \frac{0.6}{1.52} = 0.3936$$

$$t_{(1-(\alpha/2), n_1+n_2-2)} = t_{(0.975, 11+9-2)} = t_{(0.975, 18)} = 2.101$$

ولما كانت: $(T = 0.39) < (t_{(0.975, 18)} = 2.101)$ ، فإننا نقبل $H_0: \mu_1 = \mu_2$ على مستوى الدلالة

$$\alpha = 0.05$$

3-1-2-2 الانحرافان المعياريان للمجتمعان غير متساويان و مجهولان

عندما يكون الانحرافان المعياري للمجتمعين غير متساويين و مجهولين أي $\sigma_1 \neq \sigma_2$ فإن :

$$\sigma_{(\bar{x}-\bar{y})} = \sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}$$

فإذا كان حجم العينتين كبيراً فإن المقدار: $S_{(x-y)} = \sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}$ يمثل

تقدير نقطي للانحراف المعياري $\sigma_{(\bar{x}-\bar{y})}$

$$s_y^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \quad \text{و} \quad s_x^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 \quad \text{حيث}$$

نظرية 6:

إذا كانت $x_1, x_2, \dots, x_{n_1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت

$y_1, y_2, \dots, y_{n_2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين

مستقلتين عن بعضهما البعض، وكانت $\sigma_1^2 \neq \sigma_2^2$ ولكنهما مجهولتين وأردنا اختبار الفرضية $H_0: \mu_1 = \mu_2$

أي أن $H_0: \mu_1 - \mu_2 = 0$ فإن دالة الاختبار هي:

$$T = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}}; t_{(1-(\alpha/2), n_1+n_2-2)}$$

* فإذا كانت $H_1: \mu_1 \neq \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|T| = \left| \frac{(\bar{x} - \bar{y})}{\sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}} \right| \geq t_{(1-(\alpha/2), n_1+n_2-2)}$$

أي: $T \leq -t_{(1-(\alpha/2), n_1+n_2-2)} \quad \text{or} \quad T \geq t_{(1-(\alpha/2), n_1+n_2-2)}$

* وإذا كانت $H_1: \mu_1 > \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{\sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}} \succ t_{(1-(\alpha/2), n_1+n_2-2)}$$

* وكانت $H_1: \mu_1 < \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{\sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}} \prec -t_{(1-(\alpha/2), n_1+n_2-2)}$$

مثال:

(نفس المثال السابق) أخذنا عينتين من توزيعين مستقلين $X_1, X_2, \dots, X_{11}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت $y_1, y_2, \dots, y_9$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت $\bar{x} = 11.3$ و $\bar{y} = 10.7$ مع $\sigma_1^2 \neq \sigma_2^2$ و $S_1^2 = 9$ و $S_2^2 = 12$ ، وأردنا اختبار الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$

الحل:

نعلم أن $\sigma_1^2 \neq \sigma_2^2$ إذن دالة الاختبار هي:

$$T = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{(S_x^2/n_1 - 1) + (S_y^2/n_2 - 1)}} = \frac{11.3 - 10.7}{\sqrt{(9/11 - 1) + (12/9 - 1)}} = \frac{0.6}{\sqrt{0.9 + 1.5}} = \frac{0.6}{1.5491} = 0.387$$

$$t_{(1-(\alpha/2), n_1+n_2-2)} = t_{(0.975, 11+9-2)} = t_{(0.975, 18)} = 2.101$$

ولما كانت: $(T = 0.387) < (t_{(0.975, 18)} = 2.101)$ ، فإننا نقبل $H_0: \mu_1 = \mu_2$ على مستوى الدلالة

$$\alpha = 0.05$$

2-2-2 العينتان صغيرتان ($n_1 < 30 : n_2 < 30$)

إذا كان حجم العينتين صغيرا فإن تقدير σ_1 و σ_2 بواسطة S_x و S_y يصبح غير دقيق و يختلف كثيرا عن القيم الحقيقية ل σ_1 و σ_2 .

ضمن هذه الشروط فإن الإختبار السابق لم يعد قابلا للتطبيق، إذ لم يعد يسمح بالتمييز فيما إذا كان الفرق المشاهد بين الوسطين الحسابين $\bar{X}$ و $\bar{Y}$ للعينتين يمكن أن يعزى إلى الفرق الحقيقي بين μ_X و μ_Y أو أن يعزى إلى الفرق بين الانحرافين σ_1 و σ_2 ، وإن لم يجد الباحثين في مجال الإحصاء طريقة أو أداة تمكن من التمييز بين السببين المذكورين ، لكن يمكن اعتماد ما يلي :

-إذا كان الانحرافان المعياريان للمجتمعين متساويين و مجهولين $\sigma_1 = \sigma_2 = \sigma$ فإن التباين المشترك S تعتبر تقديرا نقطيا ل σ والكمية:

$$T = \frac{|\bar{x} - \bar{y}|}{S \sqrt{(1/n_1) + (1/n_2)}}; t_{(1-(\alpha/2), n_1+n_2-2)}$$

$$S^2 = \frac{(S_1^2 n_1) + (S_2^2 n_2)}{n_1 + n_2 - 2}$$

* فإذا كانت $H_1: \mu_1 \neq \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|T| = \left| \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} \right| \geq t_{(1-(\alpha/2), n_1+n_2-2)}$$

أي: $T \leq -t_{(1-(\alpha/2), n_1+n_2-2)}$ or $T \geq t_{(1-(\alpha/2), n_1+n_2-2)}$

* وإذا كانت $H_1: \mu_1 > \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} > t_{(1-(\alpha/2), n_1+n_2-2)}$$

* وكانت $H_1: \mu_1 < \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S \sqrt{(1/n_1) + (1/n_2)}} < -t_{(1-(\alpha/2), n_1+n_2-2)}$$

3- إختبار الفرضيات المتعلقة بالنسبة والفرق بين نسبتي

1-3 إختبار الفرضيات يتعلق بالنسبة P

لنفرض أننا بصدد دراسة مجتمع إحصائي يتكون من وحدات إحصائية تحمل الصفة a بنسبة P ووحدات إحصائية لا تحمل هذه الصفة بنسبة q ، وأن النسبة $P = N_a / N$ مجهولة.

لتقدير P تسحب عينة عشوائية مؤلفة من n وحدة من هذا المجتمع فنلاحظ أن نسبة الوحدات الإحصائية التي تحمل الصفة A تساوي p' وبفرض أن n_a يمثل عدد الوحدات الإحصائية الذين يحملون الصفة a في العينة فإن: $P' = n_a / n$

إن النسبة P' في العينة قد تختلف عن النسبة P في المجتمع واعتمادا على القيمة P' فإننا نستطيع إختبار فيما إذا كانت النسبة P يمكن إعتبارها مساوية للنسبة P_0 المحددة مسبقا أم لا .

إن n_a يخضع لقانون التوزيع الثنائي فإذا كان حجم العينة كبيرا وكانت P_0 ليست قريبة من الصفر أو الواحد فيمكننا تقريب قانون التوزيع الثنائي بالتوزيع الطبيعي، أما إذا كان حجم العينة كبيرا وكانت P قريبة من الصفر أو الواحد فيمكننا تقريب قانون التوزيع الثنائي بالتوزيع البواسوني

نظرية 7:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b(1, P)$ و كان حجم العينة كبيرا $n \geq 30$ و أردنا اختبار $H_0: P = P_0$ فإن دالة الاختبار هي:

$$\frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}}$$

* فإذا كانت $H_1: P \neq P_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z_{(1-(\alpha/2))}$ أي: $\left| \frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}} \right| \geq Z_{(1-(\alpha/2))}$

$$\frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}} \leq -Z_{(1-(\alpha/2))} \quad \text{or} \quad \frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}} \geq Z_{(1-(\alpha/2))}$$

* وإذا كانت $H_1: P > P_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z_{(1-(\alpha/2))}$ أي: $\frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}} > Z_{(1-(\alpha/2))}$

* وكانت $H_1: P < P_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z_{(1-(\alpha/2))}$ أي: $\frac{P' - P_0}{\sqrt{P_0(1 - P_0)/n}} < -Z_{(1-(\alpha/2))}$

مثال:

نعلم أن حوالي ثلاثة من عشرة أشخاص يفضلون نوعا معينا من الشاي، قام مكتب الدعاية لشركة الشاي هذه بدعاية لمنتجها وللحكم على مدى نجاح هذه الدعاية أخذت عينة حجمها 250 شخصا فوجدو من بينهم 35 شخص يفضل هذا النوع من الشاي
فهل هذه البيانات تعطي دليلا واضحا ازدياد نسبة الأشخاص الذين يفضلون هذا النوع من الشاي.

الحل:

نلاحظ أن هذا مثال على توزيع الثنائي (ذو الحدين) ولك لأن كل شخص إما أن يستهلك هذا النوع من الشاي أو لا يستهلكه وجميع عناصر العينة مستقلة عن بعضها البعض والفرضية المراد اختبارها هي:

$$H_0: P = P_0 = 0.3 \text{ مقابل } H_1: P > 0.3 \text{ على مستوى الدلالة } \alpha = 0.05$$

$$\frac{P' - P_0}{\sqrt{P_0(1-P_0)/n}} = \frac{(35/250) - 0.3}{\sqrt{0.3(1-0.3)/250}} = \frac{0.16}{0.0289} = 5.52$$

$$Z_{(1-(\alpha/2))} = Z_{(0.975)} = +1.96$$

ولما كانت: $(Z = 5.52) > (Z_{(0.975)} = 1.96)$ ، فإننا نرفض $H_0: P = P_0 = 0.3$ على مستوى الدلالة

$\alpha = 0.05$ أي أن الحملة الدعائية زادت من نسبة الأشخاص المستهلكين لهذه النوعية من الشاي

2-3- اختبار الفرضيات يتعلق بالفرق بين نسبتي

لنفرض أننا بصدد دراسة مجتمعان إحصائيان X و Y يتكون كل منهما من وحدات إحصائية بعضها منهم يحملون الصفة a و لتعتبر P_1 و P_2 نسبة الوحدات الإحصائية الذين يحملون الصفة a في المجتمعين و بالتالي فإن الوحدات الإحصائية التي لا تحمل هذه الصفة تقدر نسبتها ب q_1 و q_2 على التوالي

$$q_1 = 1 - P_1 \text{ و } q_2 = 1 - P_2 \text{ و لنفرض أن النسبتين } P_1 \text{ و } P_2 \text{ مجهولتين}$$

لتقدير هاتين النسبتين P_1 و P_2 نسحب عينتين عشوائيتين X و Y من المجتمعين X و Y ، العينة

الأولى بحجم n_1 لها نسبة قدرها P_1' و العينة الثانية بحجم n_2 لها نسبة قدرها P_2'

و لنعمل على معرفة ما إذا كان هناك فرق جوهري بين النسبتين المجهولتين P_1 و P_2 اعتماداً على النسبتين P_1'

و P_2' المستخلصين من العينتين X و Y المسحوبتين من المجتمعين X و Y

-إذا كان حجم العينين كبيراً فإن قانون التوزيع الإحتمالي لكل من P_1' و P_2' هو :

$$P_2' \sim N\left(p_2, \sqrt{(P_2(1-P_2)/n_2)}\right) \text{ and } P_1' \sim N\left(p_1, \sqrt{(P_1(1-P_1)/n_1)}\right)$$

وهذا يعني أن التوقع الرياضي للنسبة P_1' يساوي إلى P_1 عندما يكون حجم العينة كبيرا و كما أن الانحراف المعياري للنسبة P_1' هو $\sqrt{(P_1(1-P_1)/n_1)}$ و نفس المبدأ ينطبق على النسبة P_2' ، وعندها يكون قانون التوزيع الإحتمالي للمتغير العشوائي $(P_1' - P_2')$

$$(P_1' - P_2') \sim N\left(P_1 - P_2; \sqrt{(P_1(1-P_1)/n_1 + P_2(1-P_2)/n_2)}\right)$$

إن الانحراف المعياري يتعلق بالمؤشرين المجهولين P_1 و P_2 و يتم تقييمهما بواسطة التقدير النقطي التالي : لتكن

$$P' = (n_1 P_1' + n_2 P_2') / (n_1 + n_2)$$

و الانحراف المعياري للفرق $(P_1' - P_2')$ يساوي

$$\sigma_{(P_1' - P_2')} = \sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)} = \sqrt{P'(1-P')\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

و يصبح قانون توزيع $(P_1' - P_2')$ كما يلي $(P_1' - P_2') \sim N\left(P_1 - P_2; \sqrt{P'(1-P')(1/n_1 + 1/n_2)}\right)$

ويتضمن إختبار مقارنة النسبتين الإختبار بين الفرضيتين التاليتين :

$$H_0 : P_1 - P_2 = 0 \Rightarrow P_1 = P_2 \quad \text{— فرضية العدم}$$

$$H_1 : P_1 - P_2 \neq 0 \Rightarrow P_1 \neq P_2 \quad \text{— الفرضية البديلة}$$

فإذا كانت الفرضية الإبتدائية صحيحة أي إذا كان $P_1 = P_2$ فيكون قانون الإحتمال للفرق بين النسبتين يتبع

$$(P_1' - P_2') \sim N\left(0; \sqrt{P'(1-P')(1/n_1 + 1/n_2)}\right)$$

و مؤشر الإختبار Z يعرف بالعلاقة التالية : $Z = \frac{(P_1' - P_2') - 0}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}}$

و يخضع لقانون التوزيع الطبيعي المعياري $N(0;1)$

نظرية 8:

إذا كانت $x_1, x_2, \dots, x_{n1}$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b(1, P_1)$ ، وكانت

$y_1, y_2, \dots, y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b(1, P_2)$ ، وكانت هاتين العينتين

مستقلتين عن بعضهما البعض ، وكان حجم العينتين n_1, n_2 كبيرا

و أردنا إختبار الفرضية $H_0 : P_1 = P_2$ أي أن $H_0 : P_1 - P_2 = 0$ فإن دالة الإختبار هي:

$$\frac{(P_1' - P_2') - (P_1 - P_2)}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}}$$

$$P' = \frac{n_1 P'_1 + n_2 P'_2}{n_1 + n_2}$$

* فإذا كانت $H_1: P_1 \neq P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|Z| = \left| \frac{(P'_1 - P'_2)}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}} \right| \geq Z_{(1-(\alpha/2))}$$

$$Z \leq -Z_{(1-(\alpha/2))} \quad \text{or} \quad Z \geq Z_{(1-(\alpha/2))} \quad \text{أي:}$$

* وإذا كانت $H_1: P_1 > P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(P'_1 - P'_2)}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}} > Z_{(1-(\alpha/2))}$$

* وكانت $H_1: P_1 < P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(P'_1 - P'_2)}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}} < -Z_{(1-(\alpha/2))}$$

مثال:

في دراسة أجريت على الشباب المتقدمين للخدمة الوطنية فإنهم يتقدمون لفحص طبي لبيان مدى لياقتهم البدنية، ومن خلال الخبرات السابقة فنسبة شباب المدينة وشباب الريف الغير لائقين صحيا يكون متساوي، أمهلت لجنة القبول هؤلاء الشباب فترة من الزمن للمعالجة وبفرض أن شباب المدينة المرضى الذين تم استدعاؤهم هم 3516 وشباب الريف هم 1517 وبعد فترة المعالجة تم رفض 1023 من شباب المدن و 413 من شباب الريف، وبفرض أن P_1 هي نسبة شباب المدن المرفوضين لأسباب صحية وأن P_2 هي نسبة شباب الريف المرفوضين لأسباب صحية.

اختبر الفرضية $H_0: P_1 = P_2$ مقابل الفرضية $H_1: P_1 > P_2$

الحل:

$$P'_1 = \frac{n_a}{n_1} = \frac{1023}{3516} = 0.2909 \quad n_1 = 3516$$

$$P'_2 = \frac{n_b}{n_2} = \frac{413}{1517} = 0.2722 \quad n_2 = 1517$$

$$P' = \frac{n_1 P'_1 + n_2 P'_2}{n_1 + n_2} = \frac{1023 + 413}{3516 + 1517} = \frac{1436}{5033} = 0.2853$$

$$\begin{aligned}
 Z &= \frac{(P'_1 - P'_2) - (P_1 - P_2)}{\sqrt{(P'(1-P')/n_1) + (P'(1-P')/n_2)}} \\
 &= \frac{0.2909 - 0.2722}{\sqrt{\frac{(0.2853 * 0.7147)}{3516} + \frac{(0.2853 * 0.7147)}{1517}}} \\
 &= \frac{0.0187}{\sqrt{0.00005799 + 0.0001344}} = 1.348 \\
 Z_{(1-(\alpha/2))} &= Z_{(0.975)} = +1.96
 \end{aligned}$$

ولما كانت: $(Z = 1.348) < (Z_{(0.975)} = 1.96)$ ، فإننا نقبل $H_0: P_1 = P_2$ على مستوى الدلالة $\alpha = 0.05$ أي أن فترة المعالجة تؤدي إلى أن تكون نسبة المرفوضين لأسباب صحية من شباب المدن متساوية لنسبة المرفوضين لأسباب صحية من شباب الريف.

4- إختبارات تتعلق بالتباين والمقارنة بين تباينين

1-4 إختبارات تتعلق بالتباين

نظرية 9:

إذا كانت $X_1, X_2, \dots, X_n$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu, \sigma^2)$ ، وأردنا إختبار الفرضية $H_0: \sigma^2 = \sigma_0^2$

* فإذا كانت $H_1: \sigma^2 \neq \sigma_0^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا لم تتحقق المتباينة:

$$\frac{n.S^2}{\chi^2 [1 - (\alpha/2); n - 1]} \leq \sigma_0^2 \leq \frac{n.S^2}{\chi^2 [(\alpha/2); n - 1]}$$

* وإذا كانت $H_1: \sigma^2 > \sigma_0^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{n.S^2}{\chi^2 [1 - \alpha, n - 1]} > \sigma_0^2$$

* وكانت $H_1: \sigma^2 < \sigma_0^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{n.S^2}{\chi^2 [\alpha, n - 1]} < \sigma_0^2$$

مثال:

ينتج مصنع صايدال على سبيل المثال فقط دواء يحتوي مادة فعالة وهي محددة بشكل دقيق، ولدراسة مدى دقة المصنع في إضافة هذه الكمية إلى كل حبة من حبوب العلاج قام المسؤولون بتصنيع 25 حبة كعينة ووجدوا أن الانحراف المعياري لهذه العينة يساوي 1.45 مغ

أراد مسؤولو مصلحة المراقبة والجودة اختبار الفرضية $H_0: \sigma^2 = 1.65$ مقابل الفرضية $H_0: \sigma^2 > 1.65$ عند مستوى معنوية $\alpha = 0.05$

الحل:

$$\frac{nS^2}{\chi^2 [1-\alpha, n-1]} = \frac{25 \times 1.45^2}{\chi^2 [0.95, 25-1]} = \frac{25 \times 1.45}{36.415} = 1.4434$$

ولما كانت: $(\sigma_0^2 = 1.65) < 1.4434$ ، فإننا نقبل $H_0: \sigma^2 = 1.65$ على مستوى الدلالة $\alpha = 0.05$ ، وهذا معناه أن مقدار كمية التغير لم يزد عن الحد المقرر مبدئياً وهو $\sigma^2 = 1.65$.

2-4 إختبارات تتعلق بمقارنة تباينين .

لنفرض أننا بصدد دراسة مجتمعان إحصائيان X و Y تباينهما σ_1^2 و σ_2^2 مجهولان و نرغب باختبار فيما إذا كان هناك فرق جوهري بين التباينين المجهولين σ_1^2 و σ_2^2 اعتماداً على تبايني S_1^2 لعينة عشوائية X حجمها n_1 تم سحبها من المجتمع X و S_2^2 لعينة عشوائية Y حجمها n_2 تم سحبها من المجتمع Y هناك فرضية إضافية تتعلق بهذا الإختبار هي أن المتغير العشوائي يخضع في كلا المجتمعين للتوزيع الطبيعي، كذلك فعوضاً عن تشكيل الفرق $d = S_1^2 - S_2^2$ و اختبار فيما إذا كان d يختلف جوهرياً عن الصفر و فإن من المفضل أن نأخذ العلاقة التالية: $F = S_1^2 / S_2^2$ و بطبيعة الحال فمن الملائم أن يكون التباين الأكبر في البسط.

نظرية 10:

إذا كانت $X_1, X_2, \dots, X_{n1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكانت $Y_1, Y_2, \dots, Y_{n2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N(\mu_2, \sigma_2^2)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض.

وأردنا اختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ أي أن: $H_0: \sigma_1^2 - \sigma_2^2 = 0$

* فإذا كانت $H_1: \sigma_1^2 \neq \sigma_2^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا لم تتحقق المتباينة:

$$F \left[\frac{\alpha}{2}; n_2 - 1, n_1 - 1 \right] \leq \frac{S_2^2}{S_1^2} \leq F \left[1 - \frac{\alpha}{2}; n_2 - 1, n_1 - 1 \right]$$

* وإذا كانت $H_1: \sigma_2^2 > \sigma_1^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{S_2^2}{S_1^2} > F [1 - \alpha; n_2 - 1, n_1 - 1]$$

* وكانت $H_1: \sigma_2^2 < \sigma_1^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{S_2^2}{S_1^2} < F [\alpha; n_2 - 1, n_1 - 1]$$

و قيمة F النظرية و التي نحصل عليها بالرجوع إلى جدول فيشر سنيديكور تبعا لمستوى الدلالة α و لعدد

درجات الحرية: $\nu_1 = n_1 - 1$ و $\nu_2 = n_2 - 1$

مثال:

أراد أحد الباحثين دراسة فيما إذا كانت الدروس الخصوصية تؤدي إلى رفع التحصيل العلمي لطلاب

البكالوريا، فقام باختيار عينة حجمها 22 طالب وقسمهم عشوائيا إلى مجموعتين، الأولى مكونة من $n_1 = 9$

طلاب تم تقديم دروس خصوصية لهم ووجد أن التباين في النقاط هو $S_1^2 = 36$ ، أما المجموعة الثانية والمكونة

من $n_2 = 13$ طالب فلم يتم تقديم دروس خصوصية لهم ووجد أن التباين في النقاط هو $S_1^2 = 68$

وأردنا اختبار الفرضية أن تقديم الدروس الخصوصية يؤدي إلى تقليل التباين في النتائج فإننا نختبر الفرضية

$H_0: \sigma_1^2 = \sigma_2^2$ مقابل الفرضية $H_1: \sigma_1^2 > \sigma_2^2$ عند مستوى معنوية $\alpha = 0.05$

الحل:

$$\frac{S_2^2}{S_1^2} = \frac{68}{36} = 1.888 \text{ دالة الاختبار هي:}$$

$$F [1 - \alpha; n_2 - 1, n_1 - 1] = F [1 - 0.05; 13 - 1, 9 - 1] = F [0.95; 12, 8] = 3.28$$

ولما كانت: $1.888 < (F [0.95; 12, 8] = 3.28)$ ، فإننا نقبل $H_0: \sigma_1^2 = \sigma_2^2$ على مستوى الدلالة

$\alpha = 0.05$ ، وهذا معناه أن الطلاب الذين تلقوا الدروس الخصوصية المقدمة لم تحسن كما ينبغي النتائج و المستوى

بالمقارنة مع الطلاب الآخرين الذين لم يتلقوا دروسا خصوصية.

قائمة المراجع .

- 1- ابراهيم محمد العلي، و فتاة صبح. (2020). محاضرات في الاحصاء الرياضي. سوريا: كلية الاقتصاد جامعة تشرين.
- 2- أمير حنا هرمز. (1990). الإحصاء الرياضي. جامعة الموصل: كلية الادارة والاقتصاد.
- 3- بدر الدين المصري. (1970). مذكرات في الاحصاء. الاسكندرية، مصر: دار الجامعات المصرية.
- 4- جلال الصياد، و عبد الحميد محمد ربيع. (1983). مبادئ الطرق الاحصائية. المملكة العربية السعودية: تهامة.
- 5- عبد الحميد عبد المجيد البلداوي. (2009). أساليب الاحصاء للعلوم الاقتصادية وإدارة الأعمال مع استخدام برنامج spss. عمان الاردن: دار وائل للنشر.
- 6- عبدالقادر حلبي. (1985). مدخل إلى الإحصاء. الجزائر: ديوان المطبوعات الجامعية.
- 7- عدنان ماجد عبد الرحمن بري، و محمود محمد إبراهيم هندي. (بلا تاريخ). مبادئ الإحصاء والإحتمالات مع حل الأمثلة باستخدام ميكروسوفت إكسل.
- 8- فتحي أحمد عاروري. (2015). المعاينة الإحصائية طرقها واستخداماتها. عمان، الأردن: شركة دار الأكاديميون للنشر والتوزيع.
- 9- مبارك اسبر ديب. (2009). مبادئ في الاحتمالات والاحصاء. اللاذقية، سوريا: جامعة تشرين.
- 10- محمد شامل بهاء الدين فهمي. (2005). الاحصاء بلا معاناة المفاهيم مع التطبيقات باستخدام برنامج spss. المملكة العربية السعودية: الادارة العامة للطباعة والنشر.
- 11- محمد شبيخي. (2011). طرق الاقتصاد القياسي محاضرات وتطبيقات. دار الحامد.
- 12- محمد صبحي أبو صالح، و عدنان محمد عوض. (1983). مقدمة في الاحصاء. إربد، الأردن: دار جون وايلي و أبنائه.

13- محمد علي الأطرقجي . (1980). الوسائل التطبيقية في الطرق الاحصائية. بيروت، لبنان: دار الطليعة للطباعة والنشر.

14-محمود حسن المشهداني، و أمير حنا هرمز. (1989). الإحصاء. بغداد: بيت الحكمة.

15-منصور عوض، و صبري عزام. (2000). مبادئ الإحصاء. عمان، الأردن: دار صفاء للنشر والتوزيع.

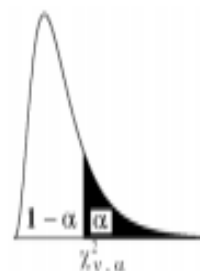
16- Nadjia El-saadi .(1997) .Travaux diriges de stat-maths .ecole superieure de banque -16

A graph of a normal distribution curve. The horizontal axis is labeled with 0 at the center and z to the right. The area under the curve to the right of z is shaded in gray.

[illegible]

Percentage Points of the χ^2 Distribution; $\chi^2_{v,\alpha}$

$$P(\chi^2 > \chi^2_{v,\alpha}) = \alpha$$



	α														
v	0.001	0.005	0.010	0.025	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.975	0.990	0.995	0.999
1	10.83	7.88	6.63	5.02	3.84	2.71	1.32	0.45	0.10	0.02					
2	13.82	10.60	9.21	7.38	5.99	4.61	2.77	1.39	0.58	0.21	0.10	0.05	0.02	0.01	
3	16.27	12.84	11.34	9.35	7.81	6.25	4.11	2.37	1.21	0.58	0.35	0.22	0.11	0.07	0.02
4	18.47	14.86	13.28	11.14	9.49	7.78	5.39	3.36	1.92	1.06	0.71	0.48	0.30	0.21	0.09
5	20.52	16.75	15.09	12.83	11.07	9.24	6.63	4.35	2.67	1.61	1.15	0.83	0.55	0.41	0.21
6	22.46	18.55	16.81	14.45	12.59	10.64	7.84	5.35	3.45	2.20	1.64	1.24	0.87	0.68	0.38
7	24.32	20.28	18.48	16.01	14.07	12.02	9.04	6.35	4.25	2.83	2.17	1.69	1.24	0.99	0.60
8	26.12	21.95	20.09	17.53	15.51	13.36	10.22	7.34	5.07	3.49	2.73	2.18	1.65	1.34	0.86
9	27.88	23.59	21.67	19.02	16.92	14.68	11.39	8.34	5.90	4.17	3.33	2.70	2.09	1.73	1.15
10	29.59	25.19	23.21	20.48	18.31	15.99	12.55	9.34	6.74	4.87	3.94	3.25	2.56	2.16	1.48
11	31.26	26.76	24.72	21.92	19.68	17.28	13.70	10.34	7.58	5.58	4.57	3.82	3.05	2.60	1.83
12	32.91	28.30	26.22	23.34	21.03	18.55	14.85	11.34	8.44	6.30	5.23	4.40	3.57	3.07	2.21
13	34.53	29.82	27.69	24.74	22.36	19.81	15.98	12.34	9.30	7.04	5.89	5.01	4.11	3.57	2.62
14	36.12	31.32	29.14	26.12	23.68	21.06	17.12	13.34	10.17	7.79	6.57	5.63	4.66	4.07	3.04
15	37.70	32.80	30.58	27.49	25.00	22.31	18.25	14.34	11.04	8.55	7.26	6.26	5.23	4.60	3.48
16	39.25	34.27	32.00	28.85	26.30	23.54	19.37	15.34	11.91	9.31	7.96	6.91	5.81	5.14	3.94
17	40.79	35.72	33.41	30.19	27.59	24.77	20.49	16.34	12.79	10.09	8.67	7.56	6.41	5.70	4.42
18	42.31	37.16	34.81	31.53	28.87	25.99	21.60	17.34	13.68	10.86	9.39	8.23	7.01	6.26	4.90
19	43.82	38.58	36.19	32.85	30.14	27.20	22.72	18.34	14.56	11.65	10.12	8.91	7.63	6.84	5.41
20	45.31	40.00	37.57	34.17	31.41	28.41	23.83	19.34	15.45	12.44	10.85	9.59	8.26	7.43	5.92
21	46.80	41.40	38.93	35.48	32.67	29.62	24.93	20.34	16.34	13.24	11.59	10.28	8.90	8.03	6.45
22	48.27	42.80	40.29	36.78	33.92	30.81	26.04	21.34	17.24	14.04	12.34	10.98	9.54	8.64	6.98
23	49.73	44.18	41.64	38.08	35.17	32.01	27.14	22.34	18.14	14.85	13.09	11.69	10.20	9.26	7.53
24	51.18	45.56	42.98	39.36	36.42	33.20	28.24	23.34	19.04	15.66	13.85	12.40	10.86	9.89	8.08
25	52.62	46.93	44.31	40.65	37.65	34.38	29.34	24.34	19.94	16.47	14.61	13.12	11.52	10.52	8.65
30	59.70	53.67	50.89	46.98	43.77	40.26	34.80	29.34	24.48	20.60	18.49	16.79	14.95	13.79	11.59
40	73.40	66.77	63.69	59.34	55.76	51.81	45.62	39.34	33.66	29.05	26.51	24.43	22.16	20.71	17.92
50	86.66	79.49	76.15	71.42	67.50	63.17	56.33	49.33	42.94	37.69	34.76	32.36	29.71	27.99	24.67
60	99.61	91.95	88.38	83.30	79.08	74.40	66.98	59.33	52.29	46.46	43.19	40.48	37.48	35.53	31.74
70	112.32	104.21	100.43	95.02	90.53	85.53	77.58	69.33	61.70	55.33	51.74	48.76	45.44	43.28	39.04
80	124.84	116.32	112.33	106.63	101.88	96.58	88.13	79.33	71.14	64.28	60.39	57.15	53.54	51.17	46.52
90	137.21	128.30	124.12	118.14	113.15	107.57	98.65	89.33	80.62	73.29	69.13	65.65	61.75	59.20	54.16
100	149.45	140.17	135.81	129.56	124.34	118.50	109.14	99.33	90.13	82.36	77.93	74.22	70.06	67.33	61.92

t-distribution										
	Confidence Level									
	60%	70%	80%	85%	90%	95%	98%	99%	99.8%	99.9%
	Level of Significance									
	2 Tailed	0.40	0.30	0.20	0.15	0.10	0.05	0.02	0.01	0.001
	1 Tailed	0.20	0.15	0.10	0.075	0.05	0.025	0.01	0.005	0.0005
df										
1		1.376	1.963	3.133	4.195	6.320	12.69	31.81	63.67	—
2		1.060	1.385	1.883	2.278	2.912	4.271	6.816	9.520	19.65
3		0.978	1.250	1.637	1.924	2.352	3.179	4.525	5.797	9.937
4		0.941	1.190	1.533	1.778	2.132	2.776	3.744	4.596	7.115
5		0.919	1.156	1.476	1.699	2.015	2.570	3.365	4.030	5.876
6		0.906	1.134	1.440	1.650	1.943	2.447	3.143	3.707	5.201
7		0.896	1.119	1.415	1.617	1.895	2.365	2.999	3.500	4.783
8		0.889	1.108	1.397	1.592	1.860	2.306	2.897	3.356	4.500
9		0.883	1.100	1.383	1.574	1.833	2.262	2.822	3.250	4.297
10		0.879	1.093	1.372	1.559	1.813	2.228	2.764	3.170	4.144
11		0.875	1.088	1.363	1.548	1.796	2.201	2.719	3.106	4.025
12		0.873	1.083	1.356	1.538	1.782	2.179	2.682	3.055	3.930
13		0.870	1.079	1.350	1.530	1.771	2.160	2.651	3.013	3.852
14		0.868	1.076	1.345	1.523	1.761	2.145	2.625	2.977	3.788
15		0.866	1.074	1.341	1.517	1.753	2.131	2.603	2.947	3.733
16		0.865	1.071	1.337	1.512	1.746	2.120	2.584	2.921	3.687
17		0.863	1.069	1.333	1.508	1.740	2.110	2.567	2.899	3.646
18		0.862	1.067	1.330	1.504	1.734	2.101	2.553	2.879	3.611
19		0.861	1.066	1.328	1.500	1.729	2.093	2.540	2.861	3.580
20		0.860	1.064	1.325	1.497	1.725	2.086	2.529	2.846	3.552
21		0.859	1.063	1.323	1.494	1.721	2.080	2.518	2.832	3.528
22		0.858	1.061	1.321	1.492	1.717	2.074	2.509	2.819	3.505
23		0.857	1.060	1.319	1.489	1.714	2.069	2.500	2.808	3.485
24		0.857	1.059	1.318	1.487	1.711	2.064	2.493	2.797	3.467
25		0.856	1.058	1.316	1.485	1.708	2.060	2.486	2.788	3.451
26		0.856	1.058	1.315	1.483	1.706	2.056	2.479	2.779	3.435
27		0.855	1.057	1.314	1.482	1.703	2.052	2.473	2.771	3.421
28		0.855	1.056	1.313	1.480	1.701	2.048	2.468	2.764	3.409
29		0.854	1.055	1.311	1.479	1.699	2.045	2.463	2.757	3.397
30		0.854	1.055	1.310	1.477	1.697	2.042	2.458	2.750	3.386
40		0.851	1.050	1.303	1.468	1.684	2.021	2.424	2.705	3.307
50		0.849	1.047	1.299	1.462	1.676	2.009	2.404	2.678	3.262
60		0.848	1.045	1.296	1.458	1.671	2.000	2.391	2.661	3.232
70		0.847	1.044	1.294	1.456	1.667	1.994	2.381	2.648	3.211
80		0.846	1.043	1.292	1.453	1.664	1.990	2.374	2.639	3.196
90		0.846	1.042	1.291	1.452	1.662	1.987	2.369	2.632	3.184
100		0.845	1.042	1.290	1.451	1.660	1.984	2.365	2.626	3.174
∞		0.842	1.036	1.282	1.440	1.645	1.960	2.327	2.576	3.091

