وزارة التعليـــم العـــالي والبحـــث العلمـــي

جامعـــة الجيلالي بونعامة خميـــس مليانـــة

كليـــة العلوم الاقتصادية، التجارية وعلـوم التسيير

قسـم العلــوم الاقتصادية

مطبوعـــة

فـــي مقيـــاس

إحصــــــــــــاء 3

موجهـة لطلبـة السنـة الثانيـة ليســانس شعبـة العلـوم الاقتصاديـة

مدعمة بأمثلة تطبيقية

من إعـــداد الدكتـور:

بوكـريطـة عبد القــادر

أستـــاذ محاضـــر صنف "ب"

السنـــة الجامعيـــة

2021-2020

البسملة

بسم الله الرحمن الرحيم

قائمـــة المحتويـــات

العيـنـــات 1

1- مصطلحـات ومفاهيـم 2

2- أساليـب جمـع البيانـات الاحصائيـة 6

2-1 الحصـر الشامـل: 6

2-2 المعاينـة: 7

3- أقسـام المعاينـة 8

3-1 المعاينـة العشوائيـة 8

3-1-1 العينـة العشوائيـة البسيطـة 8

3-1-2 العينـــة العشوائيـــة المنتظمـة 13

3-1-3 العينـــة الطبـقيـة العشوائيـــة 13

3-1-4 العينـــة العنقوديـــة 15

3-2 العينـات غيـر العشوائيـة: 16

3-2-1 العينـة الميسـرة: 16

3-2-2 العينة العمديـة أو القصديـة: 16

3-2-3 العينـة الحصصيـة: 17

نظريـــة المعاينـــة 18

1- توزيعـات المعاينـة للأوسـاط: 19

1-1- متوسـط الوسـط الحسـابي للعينـة $\mu_{\bar{X}}$: 19

1-2- تبايــن الوسـط الحسـابي للعينـة $\sigma_{\bar{X}}^2$ 21

1-3- طبيعـة توزيـع الوسـط الحسـابي للعينـة $\bar{X}$ 23

2- توزيعـات المعاينـة لتبايــن العينـات: 24

2-1- متوسـط تباينـــات العينـات $E\left(S_{\bar{X}}^2\right)$: 24

قائمة المحتويات

2-2- تبايــن تبايـنــات العينـات $V\left(S_{\bar{X}}^{2}\right)$ 26
2-3- طبيعـة توزيـع المعاينــة للتبايـن $S_{\bar{X}}^{2}$: 26
2-4- توزيــع النسبـة بيـن تبايـنـي عينتيـن: 27
3- توزيعــات المعاينــة للفــروق والمجاميـع 28
3-1- توزيعــات المعاينــة للفــروق والمجاميـع للمتوسـط: 28
3-2- توزيعــات المعاينــة للفــروق والمجاميـع للتبايــن: 29
3-3- توزيــع الفــرق بين وسطيـن 30
4- توزيعــات المعاينــة للنسبــة وللفــرق بيـن نسبتيـن 31
4-1- توزيع المعاينــة للنسبـة 31
4-2- توزيع المعاينــة للفرق بيـن نسبتيـن 33
نظريــــة التقـديــــر 34
1- التقديــر بنقطــة 35
1-1- طريقـة الإمكـان الأكبـر 36
1-2- طريقـة العـزوم 42
1-3- طريقـة المربعـات الصغـرى 45
2- التقديــر بمجــال 49
2-1 مجـال الثقـة للمتوسـط الحسابي للمجتمـع μ_X 50
2-1-1 مجـال الثقـة للمتوسـط الحسابي في حالـة توزيع المجتمـع طبيعــي 50
2-1-1-1 حالـة تبايــن المجتمع معلومـا (σ_X^2 معلومة) 50
2-1-1-2 حالـة تبايــن المجتمع غيـر معلــوم (σ_X^2 مجهولــة) 54
الحالـة الأولى: العينــات الصغيـرة 54
الحالـة الثانيـة: العينــات الكبيــرة 58
2-1-2 مجـال الثقـة للمتوسـط الحسابي في حالـة المجتمـع غير طبيعــي أو مجهــول التوزيـع 60
2-1-2-1 حالـة تبايــن المجتمع معلومـا (σ_X^2 معلومة) 60

الحالة الأولى: العينــات الصغيـرة ______ 60

الحالة الثانيـة: العينــات الكبيـرة ______ 63

2-2-1-2 حالـة تبايـن المجتمع غير معلـوم (σ_X^2 مجهولـة) ______ 64

2-2 مجـال الثقـة لتبايـن المجتمـع σ_X^2 : ______ 66

2-2-1 متوسـط المجتمـع معلومـا (μ_X معلـوم) ______ 66

2-2-2 إذا كان متوسـط المجتمـع مجهـولا (μ_X مجهـول) ______ 68

2-3 مجـال الثقـة للنسبـة ______ 70

نظريــة إختبــار الفرضيــات ______ 72

1- مفاهيــم أساسيـة في نظريـة الإختبــار ______ 73

2- إختبــارات الفرضيــات حــول معالــم المجتمــع الإحصــائي ______ 75

2-1- إختبـارات الفرضيـات حول المتوسـط الحسـابي لمجتمع طبيعـي μ_X ______ 76

2-1-1 تبايـن المجتمـع معلـوم σ_X^2 ______ 76

2-1-2 تبايــن المجتمــع غيــر معلـوم أي مجهـول σ_X^2 ______ 77

2-1-2-1 العينــات الصغيـرة ______ 77

2-1-2-2 العينــات الكبيـرة ______ 79

2-2 إختبـارات الفرضيـات حــول مقارنـة وسطيـن حسابييـن μ_X ______ 80

2-2-1 العينتـان كبيرتـان: $(n_1 \geq 30 : n_2 \geq 30)$ ______ 81

2-2-1-1 الإنحرافـان المعياريـان للمجتمعـان معلومـان ______ 81

2-2-1-2 الإنحرافـان المعياريـان للمجتمعـان متساويـان و مجهـولان $\sigma_1 = \sigma_2 = \sigma$ 82

2-2-1-3 الإنحرافـان المعياريـان للمجتمعـان غيـر متساويـان و مجهـولان ______ 84

2-2-2 العينتـان صغيرتـان $(n_1 \prec 30 : n_2 \prec 30)$ ______ 85

3- إختبــار الفرضيــات المتعلقـة بالنسبـة والفـرق بيـن نسبتيــن ______ 86

3-1- إختبـار الفرضيـات يتعلـق بالنسبـة P ______ 86

3-2- إختبـار الفرضيـات يتعلـق بالفـرق بيـن نسبتيـن ______ 88

4- إختبــارات تتعلــق بالتبايــن والمقارنــة بين تباينيــن ______ 91

4-1 إختبــارات تتعلــق بالتبايــن______ 91

4-2 إختبــارات تتعلــق بمقارنــة تباينيــن.______ 92

قائمة المراجع. ______ **94**

المحـــور الأول

العيـنـــات

S a m p l e s

العينـــات

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.

فالمعاينة 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 (**الأطرقجـي، 1980، صفحة 211**).

1- مصطلحـات ومفاهيـم

مفهوم علم الإحصاء

يشار لعلم الإحصاء بأنه مجموعة الطرق العلمية القياسية التي يمكن توظيفها لجمع البيانات والمعلومات والمعطيات الاحصائية 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 معينه من الفاكهة أو الحيوانات أو إنتاج دولة ما لسلع معينه خلال فترة زمنية محددة...الخ.

ويمكن تصنيف المجتمعات إلى نوعين:

المجتمـع المحـدود:

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

المجتمـع غير المحـدود:

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،

العينــة:

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

والعينة 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 الأقلام ...إلخ، (**حليمـي، 1985، صفحة 19**)

المتغيـر *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.

المعلمـة *parameter*

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

الإحصـاءة *statistics*

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

أصنـاف المعطيـات الإحصائيـة:

تصنف المعطيات الإحصائية إلى صنفين رئيسين هما:

أ- المعطيـات الكميـة *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 وأجزاءه 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 بيانيا عبارة عن نقاط منفصلة.

ب- المعطيـات النوعيـة *Qualitative data*

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

أقســام المعطيــات الإحصائيــة:

علم الاحصاء ينقسم إلى نوعين هما:

أ- الاحصــاء الوصفــي *Descriptive Statistics*

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

ب- الاحصـــاء الاستدلالي *Inferential Statistics*

ويختص بطرق تحليل وتفسير واستخلاص الاستنتاجات بالاعتماد على جزء (عينة) من المجتمع للتوصل الى قرارات تخص المجتمع الاحصائي، وعليه فان الاحصاء الاستدلالي يتعامل مع التعميم والتنبؤ والتقدير، وتتسم الاستنتاجات في بعض الحالات بعدم التأكد عندها يتم قياسها باستخدام الاحتمالات (**البلــداوي، 2009، صفحة 17**)

أساليــب اختيــار العينــة العشوائيــة *Random Selection Method*

أ- الاختيــار بالارجــاع: **(** *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-الاختيــار بــدون إرجــاع *(Selection Without Replacement)*

وهو يعني أنه عند اختيارنا للمفرده الأولى فإننا لا نلجأ إلى إعادتها ثانيه إلى المجتمع وأنما نختار مفرده مما تبقى من المجتمع وهكذا، (**البلــداوي، 2009، صفحة 24**) ومن الناحية العملية فان جميع مسوحات العينة تعتمد على أسلوب الاختيار بدون إرجاع ، لذا سيكون التركيز على هذا الاسلوب في دراستنا للعينات .

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

وسائـل جمـع البيانـات

أ- الاستبيان: اداة لجمع البيانات عن الظاهرة او المشكلة المراد بحثها (اسئلة 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)

مصـادر جمـع البيانـات الاحصائيـة

يمكن الحصول على المعلومات من مصدرين **هما**:

أولا: المصـدر المباشـر (الميدانــي)

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

أ – المقابلـة الشخصيـة

وتتم هذه المقابلة بواسطة أشخاص مدربين على القيام بهذه الأعمال ويقوم الباحث المدرب بطرح أسئلة محددة ومعدة مسبقا 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 – الهاتــف

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

ج- المراسلـة

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: المصـدر غيـر المباشـر (التاريخـي):

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 (**عـوض و عـزام، 2000، الصفحات 13-14**).

2- أساليـب جمـع البيانـات الاحصائيـة

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

2-1 الحصـر الشامـل:

يعتبر أسلوب الحصر الشامل أفضل أسلوب في جمع البيانات كونه يوفر للباحث بيانات كاملة عن كافة مفردات مجتمع الدراسة (**المشهدانـي و هرمـز، 1989، صفحة 30**)، تجمع فيه البيانات عن كل مفرده من مفردات المجتمع الاحصائي التي تشترك بخاصية أو مجموعة خصائص معينة كمثال مجتمع طلبة وطالبات جامعة الجيلالي 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 الى جهاز إداري وفني ضخم ومدرب للقيام به.

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

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

2-2 المعاينــة:

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 مثلا 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.

مزايـا أسلـوب المعاينـة

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، و إذا 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 فمثلا إذا كانت دقة الحصر الشامل هو 100% وتكلفته مثلا مليون دينار فعند أخذ عينة من ربع مفردات المجتمع فقد ينخفض مستوى الدقة إلى 90% مثلا 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 وهوما 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 (**الصيـاد و محمـد ربيـع، 1983، صفحة 106**).

* يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتتضح أهمية الوقت عندما نقوم بدراسة ظاهرة تتغير بمرور الوقت، وعينة قد يترتب على دراسة تلك الظاهرة في المجتمع كله بجمع البيانات من جميع مفردات المجتمع أن يمر وقت بديل فتكون البيانات والنتائج وقت ظهورها غير مطابقة لواقع المجتمع وتصبح النتائج ذات قيمه محدودة بعد أن فقدت عنصر المطابقة مع واقع الظاهرة وتوزيعها

الحالي لمجتمع، والتعدادات الدورية للسكان وبسبب ضخامة حجم العمل بها تستغرق وقتاً طويلاً حتى تصبح نتائجها جاهزة ومنشورة وقد يطول هذا الوقت إلى أكثر من ثلاث أو أربع سنوات حتى مع استخدام أحدث أجهزة الحاسبات 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 (اللانهائية) مثل مجتمع الكائنات الحية في البحار والمحيطات والطيور لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة (**البلـداوي، 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 المفرقعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينة وليس سحب الدم كله. (**عـوض و عـزام، 2000، صفحة 15**).

3- أقســام المعاينــة

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

3-1 المعاينــة العشوائيــة

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

3-1-1 العينـة العشوائيـة البسيطـة

العينة العشوائية البسيطة تعد الأساس لباقي أنواع العينات العشوائية وتستخدم عندما يكون المجتمع متجانسا من حيث الغرض أو الصفة التي تتعلق بها الدراسة (**البلـداوي، 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 يجرى عليه أي تقسيمات مختلفة حيث أن الوحدات المكونة لهذا المجتمع تعامل كلها

باحتمالات متساوية مما يجعل المعادلات الرياضية والإحصائية المستخدمة لتقدير معالم المجتمع أبسط ما يمكن وتعرف هذه المعاينة بأسماء عديدة من أهم هذه الأسماء انتشارا العينة غير المقيدة وعينة تكافؤ الفرص.

من شــروط اختيــار العينــة 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 $1/C_N^n$ (**البلـداوي، 2009، صفحة 23**)

طـــرق سحـــب العينـــة العشوائية البسيطة:

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

القرعـــة:

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

الحالة الأولى: السحـب بدون إرجــاع

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

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

حيث N تمثل حجم المجتمع وعددها 6 كما أن n تمثل حجم العينة 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 الثانيــة: السحـب مـع إرجــاع

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 الدراسة 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 لوكان حجم المجتمع 100 مفردة وكان حجم العينة المطلوبة 3 مفردات فإن عدد العينات التي يمكن سحبها تكون:

في حالة سحب العينات مع عدم الإرجاع هو:

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

في حالة سحب العينات مع الإرجاع هو:

$N^n=100^3=100*100*100=1000000$ عينة

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

الطريقـــة الثانيـــة:

هي أن يقوم الباحث بترقيم كل مفردة من مفردات المجتمع وتسجيل هذه الأرقام في قصاصات 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 يمثلD ومن ثم يكون أول مفردة في العينة المطلوبة هي المفردة D فإذا كان السحب مع عدم الإرجاع يتم استبعاد هذه المفردة فيكون مجتمع الدارسة الجديد هو*(A,B,C,E,F)* أما إذا كان السحب مع الإرجاع فيكون مجتمع الدارسة الجديد هو*(A,B,C,D,E,F)* ويتم تكرار السحب مرة أخرى وليكن الرقم الذي تحمله القصاصة 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 الطريقين السابقين 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.

جـــدول الأرقـام العشوائيـــة:

هي أعداد صحيحة مكونة من الأرقام 8,9 ... ,0,1,2,3 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.

أمثلــة:

مثــال 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 أفراده 3500 فرد فإننا هنا نستخدم جدول الأعداد العشوائية حسب الخطوات التالية:

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

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

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

ومن الخطوات السابقة نحصل على الأرقام التالية:

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

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

<u>3185</u>-<u>1616</u>-<u>3416</u>- <u>0087</u>-<u>0061</u>-<u>2290</u>-<u>1032</u>-<u>1545</u>-<u>1925</u>-<u>0216</u>.

مثــال 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 -<u>021</u> -666 -192 -154 -103 -647 -229 -<u>006</u>- 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 العينـــة العشوائيـــة المنتظمة (*Systematic Sample*)

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

فمثلاً إذا كان لدينا مجتمعاً حجمه 2000 مفرده ونريد اختيار عينه منتظمة حجمها 100 مفرده فإننا نقسم الإطار إلى فترات منتظمة طول كل فترة $\frac{N}{n}=\frac{2000}{100}=20$ مفردة ومن داخل مفردات الفترة الأولى $(1-20)$ يختار مفرده واحدة عشوائياً ولتكن رقم 14 مثلاً وبناء على رقم تلك المفردة يتحدد باقي أرقام المفردات الأخرى بشكل منتظم وفقا لمتوالية عددية حدها الأول 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 إجرائها.

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 العينـــة الطبـقيــة العشوائيـــة (*Stratified random sampling*)

تعتبر العينة العشوائية الطبقية أفضل أنواع العينات وأكثرها دقة في تمثيل المجتمع الاحصائي غير المتجانس حيث أنه في كثير من الأحوال تكون مفردات المجتمع الاحصائي 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 (**المصـري، 1970، صفحة 27**)

يلجأ إليها الباحث في حالة ما إذا كان مجتمع الدراسة منقسما 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:

حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × 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 منحازا لإحدى الفئات الثلاث.

وعليه يجب تقسيم المجتمع الاحصائي إلى ثلاث فئات الأولى تضم الأسر ذات الدخول المرتفعة، والثانية تضم الأسر ذات الدخول المتوسطة والثالثة تضم الأسر ذات الدخول المنخفضة، وبعد ذلك يتم سحب عينة عشوائية بسيطة أو منتظمة من كل مجموعة يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع حجوم العينات العشوائية الثلاث تؤلف حجم العينة العشوائية الطبقية.

مثـــال 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.

الحـــل:

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

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

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

مثـــال 2:

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

حجم الإنتـــاج	صغير	متوسط	كبير
عدد المـــزارع	1500	2500	1000

فإذا كان حجم العينة المطلوب سحبها هو 1000 مزرعة، وفقا 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 العينـــة العنقوديـــة (*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 الاحصائي إلى وحدات أولية ثم يتم سحب عينة عشوائية بسيطة من هذه الوحدات الأولية كمرحلة أولى ثم يتم تقسيم الوحدات الأولية المختارة إلى وحدات أصغر تدعى بالوحدات الثانوية ويتم اختيار عينة عشوائية بسيطة من كل وحدة من الوحدات الثانوية، ثم تقسيم الوحدات الثانوية المختارة إلى وحدات أصغر ويتم اختيار عينة عشوائية منها كمرحلة ثالثة وتستمر عملية التقسيم والاختيار لحين الوصول إلى المفردات التي يتم جمع البيانات منها وكأننا نتحدث عن عنقود.

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

3-2 العينـات غيـر العشوائيـة:

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:

3-2-1 العينـة الميسـرة: Consenience sample

يتطلب اختيار العينة الميسرة كما يفهم من الاسم ضمنا 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 (**فهمـي، 2005، صفحة 114**)

3-2-2 العينة العمديـة أو القصديـة: *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 (**الأطرقجـي، 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 يتعمد الباحث اختيار مفردات معينة كعينة لمجتمع الدراسة يرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلاً مقبولاً لمجتمع الدراسة.

مثلاً إذا أراد باحث دراسة خصائص اقتصادية أو اجتماعية معينه عن ريف دولة ما، وكانت إمكانياته المالية والإدارية لا تسمح له بعينة سوى سكان قرية واحدة، فإنه في هذه الحالة إذا ما تم اختيار القرية عشوائياً من بين آلاف القرى بتلك الدولة فإن الصدفة قد تأتي بقرية بعيدة في خصائصها (من حيث الظاهرة موضوع الدراسة) عن خصائص معظم قرى تلك الدولة... كأن تأتي بالصدفة قرية ساحلية معظم سكانها من الصيادين أو قرية قريبة من مشروع صناعي ضخم يستوعب في قواه العاملة معظم سكانها.. هذه القرية أو تلك قد يأخذ النمط المعيشي لسكناها طابعاً خاصاً بعيداً عن النمط المعيشي المعتاد لبقية القرى، لذلك فأي منها لا يمكن أن يعطي تمثيلاً مقبولا لريف تلك الدولة، لهذا

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

على الرغم من اعتقاد الباحث أن اختياره لوحدات 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. (**فهمي، 2005، صفحة 115**)

3-2-3 العينـة الحصصيـــة: *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 (**الأطرقجـي، 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 جامعوا البيانات بالحصول على البيانات من 20 موظفاً، 45 من العمال الحرفيين، 35 من ذوي الأعمال الحرة.. وتترك الحرية لجامعي البيانات في اختيار الأفراد المطلوبة فيها حدود المواصفات الموضوعة لكل طبقة من الطبقات المذكورة.

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

عموماً.. يلجأ الباحث إلى العينة 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 (**فهمـي، 2005، صفحة 116**).

المحـــور الثانـي

نظريـــــة المعاينـــــة

Sampling Theory

نظريـــة توزيعـــات المعاينـــــــة

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...، انطلاقا من مؤشرات العينة (**حليمـي، 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- توزيعـــات المعاينــة للأوســاط: *Sampling Distributions of Means*

1-1- متوســط الوســط الحســابي للعينــة $\mu_{\bar{X}}$:

نظريــة 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: $E\left(\bar{X}\right)=\mu_{\bar{X}}=\mu_X$

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

ملاحظــة:

- إذا سحبنا جميع العينات من الحجم n من مجتمع ما حجمه N فإنه وبالضرورة سنجد أن متوسط المجتمع يكون مساويا 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 مجهولا، ونكتب: $\hat{\mu}_X = \bar{X}$ ونقول إن الإحصائية $\bar{X}$ هي مقدرة لمعلمة متوسط المجتمع μ_x.

مثـال:

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.

الحـل:

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

الحالـة الأولى: توزيع المعاينة للوسط الحسابي وحسابه في حالة السحب مع الارجاع

عدد العينات الممكنة سحبها إذا كان السحب بإرجاع هي : $N^n = 5^2 = 25$

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 $\bar{X}$:

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

وهو يساوي نفس قيمة μ_x كما يجب أن يكون.

الحالـة الثانيـة: توزيع المعاينة للوسط الحسابي وحسابه في حالة السحب بدون الارجاع

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 $\bar{X}$:

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

2-1- تبايـــن الوســط الحســابي للعينـــة $\sigma_{\bar{X}}^2$

نظريـــة 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:

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

ملاحظـــة:

- في حالة السحب مع الارجاع فإن $\sigma_{\bar{X}}^2$ يتأثر طرديا بتباين المجتمع وعكسيا بحجم العينة أي كلما كان حجم العينة كبيرا والمجتمع أكثر تجانسا كان التقدير أدق.
- في حالة السحب بدون إرجاع فإن المقدار $\left(\frac{N-n}{N-1}\right)$ والذي يسمى بمعامل الإرجاع يصبح مهملا إذا كان يقترب من الواحد أي إذا كان حجم العينة صغير جدا بالمقارنة مع حجم المجتمع

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

- 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 تقديرا أكثر دقة لمعلمة المجتمع μ_X أي:

$$n \succ 1 \Rightarrow \left(\frac{N-n}{N-1}\right) \prec 1 \Rightarrow \left(\left(\frac{\sigma_X^2}{n}\left(\frac{N-n}{N-1}\right)\right) \le \left(\frac{\sigma_X^2}{n}\right)\right)$$

مثـال:

(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.

الحـل:

تباين المجتمع يحسب كما يلي:

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

الحالـة الأولى: تباين الوسط الحسابي للعينات في حالة السحب مع الارجاع

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 $\bar{X}$:

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

وهي تساوي $\frac{\sigma_X^2}{n}$

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

الحالــة الثانيــة: 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 $\bar{X}$:

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

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

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

1-3- طبيعـة توزيـع الوسـط الحسـابي للعينـة $\bar{X}$

نظريــة 3:

- إذا كان المجتمع موزع طبيعيا بمتوسط μ_x وتباين σ_X^2 فإن متوسط العينة المسحوبة من هذا المجتمع عشوائيا يتبع أيضا التوزيع الطبيعي بمتوسط $\mu_{\bar{x}}=\mu_x$ وتباين $\sigma_{\bar{X}}^2$

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

- هذا إذا كانت σ_X^2 معلومة، ولكن في حالة ما إذا كانت قيمة σ_X^2 غير معلومة فإننا نستخدم بدلا منها الانحراف المعياري للعينة S_x^2 ولكن في هذه الحالة يصبح المتغير T يخضع لتوزيع يعرف بتوزيع t ستيودنت بدرجات حريه $(n-1)$ ، حيث:

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

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

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

- 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 عشوائيا $\bar{X}$ 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 كبيرا $(30 \le n)$

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

ملاحظـــة:

– وتعتبر النتيجة السابقة هامة جداً في الإحصاء وخاصة في التطبيقات العلمية وتسمى نظرية النهاية المركزية *Central Limit Theorem*

– في حالة السحب مع الإرجاع فإن: $\sigma_{\bar{X}}^2 = \dfrac{\sigma_X^2}{n}$

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

– في حالة السحب بدون إرجاع و $\left(\dfrac{n}{N} \prec 0.05\right)$ فإن: $\sigma_{\bar{X}}^2 = \dfrac{\sigma_X^2}{n}$

2- توزيعـــات المعاينـــة لتبايـــن العينـــات: *Sampling Distributions of Variance*

2-1- متوســط تباينـــات العينـــات $E\left(S_{\bar{X}}^2\right)$:

نظريـــة 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 $E\left(S^2\right)$ تكتب كما يلي:

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

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

مثـــال:

(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.

الحـل:

الحالــة الأولى: حساب متوسط تباينات العينات $E\left(S^2\right)$ في حالة السحب مع الارجاع

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)	$\sum$	**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 :

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

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

الحالــة الثانيـــة: حساب متوسط تباينات العينات $E\left(S^2\right)$ 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**	$\sum$
$f(S^2)$	4/10	3/10	2/10	1/10	**1**

متوسط التوزيع العيني لـ S^2 :

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

للتحـــــقق: من خلال النظرية

$$E\left(S^2\right)=\sigma_X^2\left(\frac{n-1}{n}\right)\left(\frac{N}{N-1}\right)=2\left(\frac{1}{2}\right)\left(\frac{5}{4}\right)=\frac{10}{8}=1.25$$

2-2- تبايـــن تباينـــات العينـــات $V\left(S_X^2\right)$

نظريـــة 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_X^2 أي $V\left(S_X^2\right)$ يكتب كما يلي:

تباين $\hat{S}^2$ هو كالتالي:

$$V\left(\hat{S}^2\right)=\frac{\mu_4}{n}+\frac{3-n}{n(n-1)}\sigma^4$$

حيث:

$$\mu_4=E\left(\left[(X-\mu)^4\right]\right)$$

3-2- طبيعـــة توزيـــع المعاينـــة للتبايـــن S_X^2 : ***Sampling Distribution of The Variance***

نظريـــة 6:

إذا كانت $x_1, x_2 \ldots\ldots \quad x_n$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu, \sigma^2\right)$، وكان S_x^2 تباين هذه العينة وكانت σ_X^2 معلومة فإن: $\frac{n.S_X^2}{\sigma_X^2}$ يتبع توزيع كاي تربيع درجة حريته $(n-1)$.

مثـال:

إذا كانت $x_1, x_2 \ldots\ldots \quad x_{25}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_1, 25\right)$، و تباينها S_x^2، فأوجد c حيث : $P\left(S_x^2 \leq c\right)=0.90$

الحـل:

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

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

4-2- توزيــع النسبــة بيــن تباينــي عينتيــن: Distribution of $\frac{S_1^2}{S_2^2}$

نظريـــة 7:

إذا كانت S_x^2 تباين عينة عشوائية $x_1, x_2 \ldots\ldots \quad x_{n1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_1, \sigma^2{}_1\right)$، وكانت S_y^2 تباين عينة عشوائية $y_1, y_2 \ldots\ldots \quad y_{n2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_2, \sigma^2{}_2\right)$، وكانت هاتين العينتين مستقلتين عن بعضهما البعض، فإن:

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

يتبع توزيع فيشر درجة حريته $\left[(n_1-1).(n_2-1)\right]$.

مثـــال:

إذا كانت $x_1, x_2 \ldots\ldots \quad x_{16}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_X, \sigma_X^2\right)$ تباينها S_x^2، و $y_1, y_2 \ldots\ldots \quad y_{13}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_Y, \sigma_Y^2\right)$ تباينها S_y^2.

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

الحـــل:

باستعمال الجدول الاحصائي لتوزيع فيشر الذي درجة حريته

$$F\left[(n_1-1)\text{،}(n_2-1)\right] = \left[(16-1)\ (13-1)\right] = \left[(15)\ (12)\right] = 4.01$$

نجد $c = 4.01$

3- توزيعـــات المعاينــة للفــروق والمجاميع

3-1- توزيعـــات المعاينــة للفــروق والمجاميع للمتوسـط:

نظريــة 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 Yسحبنا منهما العينتين $x: 1.3.5.7$ و $y: 2.4.6$

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

حالــة المجمـــوع

المجمـــوع		العينــــة x			
μ_{x+y}		1	3	5	7
العينـة	2	3	5	7	9
	4	5	7	9	11
y	6	7	9	11	13

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

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

حالــة الفــرق

الفـــرق		العينــــة x			
μ_{x-y}		1	3	5	7
العينـة	2	1-	1	3	5
	4	3-	1-	1	3
y	6	5-	3-	1-	1

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

3-2- توزيعــات المعاينــة للفــروق والمجاميـع للتبايــن:

نظريــة 9:

– ليكن لدينا مجتمعان X و Y، متوسطهما الحسابي μ_X و μ_Y، تباينهما σ_X^2 وσ_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:

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

مثال:

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

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

حالـة المجمــوع

المجمــوع		العينـــة x			
μ_{x+y}		1	3	5	7
العينـة	2	3	5	7	9
	4	5	7	9	11
y	6	7	9	11	13

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

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

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

حالـة الفــرق

الفــرق		العينـــة x			
μ_{x-y}		1	3	5	7
العينـة	2	1-	1	3	5
	4	3-	1-	1	3
y	6	5-	3-	1-	1

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

3-3- توزيــع الفـرق بين وسطيـن

أ– توزيــع المجتمـعيـن طبيعـي

نظريــة 10:

إذا كانت $x_1, x_2 ...... x_{n1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_1,\sigma^2{}_1\right)$، وكانت $y_1, y_2 ...... y_{n2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_2,\sigma^2{}_2\right)$، وكانت هاتين العينتين مستقلتين عن بعضهما البعض، نرمز 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 $\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)$ أي أن:

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

مثــال:

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 هو $\bar{y}$.

المطلوب: أوجد $P\left(\bar{x}-\bar{y}\le 150\right)$

الحــل:

$$P\left(\bar{x}-\bar{y}\le 150\right)=P\left(\frac{(\bar{x}-\bar{y})-(1350-1190)}{\sqrt{(900/25)+(700/20)}}\le\frac{150-(1350-1190)}{\sqrt{(36)+(35)}}\right)$$

$$=P\left(Z\le\frac{-10}{\sqrt{71}}\right)=P\left(Z\le -1.186\right)=0.1190$$

ب– توزيــع المجتمـعيـن غير طبيعـي

نظريــة 11:

– إذا كانت $x_1, x_2 ...... x_{n1}$ تمثل قياسات عينة عشوائية وسطها هو $\bar{x}$ مسحوبة من مجتمع وسطه μ_1 وتباينه $\sigma^2{}_1$، وكانت $y_1, y_2 ...... y_{n2}$ 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 $(30 \le n_1, n_2)$ فإن: (**أبو صالـح و عـوض، 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:

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

ملاحظـــة

إذا كان أحد المجتمعين أوكليهما لا يتوزع توزيعاً طبيعياً فإن $(\bar{x}-\bar{y})$ لا يتوزع توزيعاً طبيعياً كذلك، ولكن لقيم $(30 \le n_1, n_2)$ 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- توزيعـــات المعاينـــة للنسبـــة وللفـــرق بيـــن نسبتيـــن

4-1- توزيع المعاينـــة للنسبـــة

عند دراسة مجتمع إحصائي يتكون من وحدات إحصائية تحمل صفة معينة ما a ووحدات إحصائية لا تحمل هذه الصفة $\bar{a}$ ومجموعهما يمثل المجتمع أي $N = N_a + N_{\bar{a}}$

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: $q = N_{\bar{a}}/N = (N - N_a)/N$

$$\frac{N_a}{N} + \frac{N_{\bar{a}}}{N} = \frac{N_a + N_{\bar{a}}}{N} = 1 \qquad \text{أي} \qquad p + q = 1$$

عند سحب عينة عشوائية مؤلفة من n وحدة من هذا المجتمع فنلاحظ أن عدد الوحدات الإحصائية التي تحمل الصفة a هي n_a

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

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

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

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:

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

كذلك $V(X)=n.p.q$ ومنه

$$V\left(\frac{n_a}{n}\right)=V(P')=\frac{1}{n^2}V(n_a)=\frac{1}{n^2}.n.p.q=\frac{p.q}{n}$$

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

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

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

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

نظريــة 12:

لتكن $X_N \quad X_1,X_2......$ 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:

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

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

فإن:

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

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

مثـــال:

في مستشفى الامراض 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 .

الحـــل:

نعلم أن:

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

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

فإن:

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

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

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

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

2-4- توزيع المعاينـــة للفـرق بيـن نسبتيـن

نظريـــة 13:

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

وسطه : $E\left(P_1'-P_2'\right)=\mu_{P_1'-P_2'}=P_1-P_2$

وتباينه : $V\left(P_1'-P_2'\right)=\sigma^2_{\left(P_1'-P_2'\right)}=\dfrac{P_1\left(1-P_1\right)}{n_1}+\dfrac{P_2\left(1-P_2\right)}{n_2}=\dfrac{P_1q_1}{n_1}+\dfrac{P_2q_2}{n_2}$

فإن:

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

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

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

المحـــور الثالـــث

نظريـــــة التقـديـــــر

Estimation Theory

نظريـــــة التقـديـــر

Estimation Theory

ذكرنا في المحور السابق بأن العينة ***sample*** هي جزء من المجتمع ***Population*** وطريقة اختيار هذا الجزء يسمي 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 فتسمي 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- التقديـــر بنقطـــة *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.

مثـــال: إن قيمة الوسط الحسابي للعينة $\overline{X}$ هوتقدير 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' هوتقدير لنسبة المجتمع P .

خصائص المقدر الجيد هي :

عدم التحيز *Unbiasedness* : التحيز هو ذلك الفرق بين مقدرة ما ووسط توزيعها، فإذاكان 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 $E\left(\hat{\theta}\right)=\theta$

مثال

الاتســاق Consistency : يكون المقدر متسقا اذا كانت قيمته لا تختلف اختلاف جوهرية عن قيمة المعلمة الحقيقية للمجتمع بزيادة حجم العينة أي أن $\lim_{n\to\infty} P\left(\left|\hat{\theta}-\theta\right| \prec \varepsilon\right)=1$ ، حيث أن ε هوالفرق بين المقدر والمعلمة .

الكفــاءة Efficiency : إن كفاءة المقدر غير المتحيزة $\hat{\theta}_1$ إلى المقدر غير المتحيز $\hat{\theta}_2$ هونسبة تباين المقدر $\hat{\theta}_2$ إلى تباين المقدرة $\hat{\theta}_1$ أي $e\left(\hat{\theta}_1\right)=V\left(\hat{\theta}_2\right)/V\left(\hat{\theta}_1\right)$

ومن التعريف أعلاه يتضح بأن المقدر الأقل تباينا هوالأعلى كفاءة.

مثــال:

الوسط الحسابي للعينة $\bar{X}$ هومقدر غير متحيز ل μ_X وكذلك الوسيط للعينة هومقدر غير متحيز لـ μ_X ولكن تباين $\bar{X}$ هو $\frac{\delta_X^2}{n}$ بينما تباين الوسيط هو $\frac{\pi}{2}\frac{\delta_x^2}{n}$ ، لذا فان كفاءة الوسط الحسابي هو

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

أي أن الوسط الحسابي $\bar{X}$ هوأكفأ تقديرا من الوسيط.

الكفايــة Sufficiency :

يكون المقدر $\hat{\theta}$ مقدرا كافيا للمعلمة θ اذا كان قد شمل كل المعلومات ذات العلاقة به θ المتوافرة في العينة. فعند سحب عينة عشوائية في مجتمع يتوزع توزعا طبيعيا فأن $\bar{X}$ هو مقدر كاف لـ μ_X لأنه لا يمكن إضافة أي شيء على $\bar{X}$ لجعله مقدرا أحسن لـ μ_X لأن $\bar{X}$ يحوي على جميع المعلومات المتعلقة ب μ_X من العينة.

هذا وهناك عدة طرق مهمة لايجاد تقدير المعلمة بنقطة وهي :

1-1- طريقة الإمكـان الأكبـر 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 أحيانا، ولكن في الحالة عندما تؤول n إلى ما لا نهاية (حيث n حجم العينة) فإنها تعطي مقدرات غير متحيزة ولها أقل تباين ومتسقة دوما، كما يؤول توزيعها إلى التوزيع الطبيعي مع زيادة حجم العينة.

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

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

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

فإذا كانت $T=T(x_1,x_2,\ldots\ldots,x_n)$ 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 مقدر مبني على التوقع الأعظمي للوسيط θ وبما أن T يجب أن يحقق النهاية العظمى لتابع الامكانية 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 العظمى هو: $\frac{\partial L}{\partial\theta}=0$.

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

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

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

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**)

مثـال1:

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

الحـل:

طريقـة 1: الخطوات كما يلي:

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

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

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

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

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

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

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

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

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

طريقـة 2: الخطوات كما يلي:

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

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

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

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

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

مثـال 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

الحــل:

تابع الامكانية العظمى:

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

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

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

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 المقدر 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 المقدر المبني على التوقع الأعظمي للوسيط σ^2 في المجتمع الطبيعي وهو مقدر متحيز.

مثـال3:

لتكن لدينا $x_1,x_2,.......,x_n$ عينة من متغيرات عشوائية تتبع التوزيع البواسوني $x \sim P(m)$

حيث: **(El-saadi, 1997, p. 37)**

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

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

الحـــل: الخطوات كما يلي:

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

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

$$\begin{cases}\dfrac{\partial}{\partial m}\log L\,(m)=0\\[2mm] \dfrac{\partial^2}{\partial m^2}\log L\,(m)\prec 0\end{cases}\Rightarrow\begin{cases}\dfrac{\sum_{i=1}^{n} x_i}{m}-n=0\Rightarrow \hat{m}=\dfrac{\sum_{i=1}^{n} x_i}{n}=\bar{x}\\[2mm] -\dfrac{\sum_{i=1}^{n} x_i}{m^2}\prec 0\end{cases}$$

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

مثـال 4:

لتكن لدينا $x_1,x_2,........,x_n$ عينة من متغيرات عشوائية تتبع التوزيع الأسي حيث:

$$f\ (x,\lambda)=\lambda e^{-\lambda x},\qquad 0\le x\le\infty\quad ,\lambda\succ 0$$

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

الحـــل: الخطوات باختصار كما يلي:

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

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

$$\begin{cases}\dfrac{\partial}{\partial\lambda}\log L\,(\lambda)=0\\[2mm] \dfrac{\partial^2}{\partial\lambda^2}\log L\,(\lambda)\prec 0\end{cases}\Rightarrow\begin{cases}n\left(\dfrac{1}{\lambda}\right)-\sum_{i=1}^{n} x_i=\dfrac{n}{\lambda}-n\bar{x}=0\Rightarrow\hat{\lambda}=\dfrac{1}{\bar{x}}\\[2mm] -\dfrac{n}{\lambda^2}\prec 0\end{cases}$$

$T=\hat{\lambda}=\dfrac{1}{\bar{x}}$ هو المقدر المبني على التوقع الأعظمي للوسيط λ في المجتمع الأسي.

مثـال 5:

نعلم من الفرضيات أن الأخطاء تتوزع توزيعا طبيعيا: $\varepsilon_i \sim N\left(0,\sigma_\varepsilon^2\right) \quad ,\forall i=1....n$

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

و نموذج الانحدار الخطي البسيط المقدر هو: $\hat{Y_i}=\hat{\beta}_0+\hat{\beta}_1 X_i, \quad i=1,\ ...,\ .n$

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

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

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

الحـل:

لدينا تابع الامكانية العظمى $L\left(\varepsilon_1,...,\varepsilon_n\right)$ حيث:

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

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

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

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

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

نقوم بتقدير معالم النموذج وذلك بتعظيم لوغاريتم دالة المعقولية العظمى و هذا بالاشتقاق الجزئي مرة بالنسبة لـ β_0 وأخرى بالنسبة لـβ_1 و أخرى بالنسبة σ_ε^2 يكون:

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

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

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

و بمساواة المعادلات الثلاث بالصفر نحصل على:

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

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

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

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

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

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

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

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

$$\frac{\partial^2}{\partial \hat{\sigma}_1^{\,2}} \log L\left(\beta_0, \beta_1, \sigma_\varepsilon^{\,2}\right) = \frac{n}{2\hat{\sigma}_\varepsilon^{\,4}} - \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^{\,2}}{\hat{\sigma}_\varepsilon^{\,6}}$$

1-2- طريقــة العــزوم *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:

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

والعزوم حول التوقع الرياضي ومن ذات الرتبة هي:

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

تعريــف:

بفرض أن $i=\overline{1.n}$ ، X_i متغيرات القيم الملاحظة لعينة عشوائية حجمها n ، عندئذ نعرف العزم الابتدائي حول الصفر ومن الرتبة s ، لهذه العينة بالشكل: $;s=\overline{1,k}$ $\quad \alpha_i^s=\frac{1}{n}\sum_{i=1}^{n}X_i^s$

ومنه ينتج: $\hat{\alpha}_1=\frac{1}{n}\sum_{i=1}^{n}X_i^1$ ، $\hat{\alpha}_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ ، $\hat{\alpha}_3=\frac{1}{n}\sum_{i=1}^{n}X_i^3$ ،...،إلخ.

وكل منها يعتبر متغيرا إحصائيا وتعتبر مقدرا غير متحيز ومعقول للوسيط θ بالاستناد على النظرية التالية

نظريــة:

عزوم العينة $\hat{\alpha}_1,\hat{\alpha}_2,......,\hat{\alpha}_s$ هي مقدرات غير متحيزة ومعقولة لعزوم المجتمع $\alpha_1,\alpha_2,......,\alpha_s$ على الترتيب، وذلك من أجل قيمة محددة لـ s وبشرط وجود $\hat{\alpha}_s$ حيث لدينا (**اسبر ديب، 2009، الصفحات 302-303**):

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

وهذا يؤكد صفة عدم التحيز.

$$V(\hat{\alpha}_S)=V\left(\frac{1}{n}\sum_i X_i^s\right)=\frac{1}{n^2}\sum_i V\left(X_i^s\right)=\frac{1}{n^2}.nV\left(X^s\right)=\frac{V\left(X^s\right)}{n}\xrightarrow[n\to\infty]{}0$$

وهذا يؤكد صفة الإتساق.

مثــال1:

لتكن لدينا $x_1,x_2,........,x_n$ عينة من متغيرات عشوائية تتبع التوزيع الطبيعي $X\sim N\left(\mu,\sigma^2\right)$

المطلوب:

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

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

الحــل:

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

- في المجتمع الطبيعي هو: $\alpha_1=E\left(X_i\right)=\mu$

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

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

أي المقدر المطلوب هو: $T = \bar{X}$

وهو ذات المقدر الذي حصلنا عليه بطريقة الامكانية العظمى وهو مقدر متسق وغير متحيز.

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

- في المجتمع هو: $\mu_2 = E\left(X - \mu\right)^2 = \sigma^2$

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

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

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

وهو مقدر التباين في التوزيع الطبيعي بطريقة العزوم وهو نفسه الذي حصلنا عليه بطريقة الامكانية العظمى وهو مقدر متسق، لكنه متحيز.

مثـــال 2:

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

$$f\left(x,\theta\right) = \theta . x^{\theta - 1} \quad ;\left(0 \prec x \prec 1 \quad , \quad \theta \succ 0\right)$$

الحـــل:

- العزم الأول حول الصفر للمجتمع هو: $\alpha_1 = E\left(X\right) = \int_0^1 x . f\left(x,\theta\right) dx$

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

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

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

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

3-1- طريقـة المربعـات الصغـرى Method of Least Square

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

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

أما ε_i فيمثل الخطأ في تفسير Y_i، ومنه يمكن كتابته انطلاقا 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 أيضا إلى الصياغة الرياضية غير السليمة للنموذج.

فرضيات النموذج:

* $E(\varepsilon_i) = 0 \qquad , \forall i = 1.....n$
* $Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_\varepsilon^{\,2} \qquad , \forall i = 1.....n$
* $\varepsilon_i \sim N\left(0, \sigma_\varepsilon^{\,2}\right) \qquad , \forall i = 1.....n$
* $Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0 \qquad , \forall i \neq j \qquad i, j = 1.....n$
* $Cov(X_i, \varepsilon_i) = 0 \qquad , \forall i = 1......n$

إن طريقة المربعات الصغرى تحاول إيجاد أحسن تصحيح خطي بتدنئة 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 رياضيا بـ :

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

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

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

بعد حل جملة المعادلتين السابقة نتحصل على تقدير معلمتي النموذج:

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

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

خصائـص مقـدرات المربعات الصغـرى

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

- متوسطات هذه المقدرات هي:

* $E\left(\hat{\beta}_0\right) = \beta_0$ ومنه نقول أن $\hat{\beta}_0$ مقدرة لـ β_0.

* $E\left(\hat{\beta}_1\right) = \beta_1$ ومنه نقول أن $\hat{\beta}_1$ مقدرة لـ β_1.

- وتباينات هذه المقدرات هي:

* تبايـن $\hat{\beta}_0$:

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

* تبايـن $\hat{\beta}_1$:

$$\operatorname{var}(\hat{\beta}_0) = \frac{\sigma_\varepsilon^2}{n} + \bar{X}^2 \operatorname{var}(\hat{\beta}_1)$$

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

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

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

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

أ- خاصية عدم التحيز:

$E\left(\hat{\beta}_0\right) = \beta_0$ ومنه نقول أن $\hat{\beta}_0$ مقدرة غير متحيزة لـ β_0.

$E\left(\hat{\beta}_1\right)=\beta_1$ ومنه نقول أن $\hat{\beta}_1$ مقدرة غير متحيزة لـ β_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 $n\to\infty$، ونقول أن النهاية الاحتمالية للمقدر $\hat{\beta}_1$ هي β_1 ونكتب : $\underset{n\to\infty}{p\lim}(\hat{\beta}_1)=\beta_1$
- يجب أن تكون قيمتا التحيز والتباين تقتربان أو تساويان الصفر كلما اقترب n مما لا نهاية أي:

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

$$2.\lim_{n\to\infty}\mathrm{var}(\hat{\beta}_1)=\underset{n\to\infty}{p\lim}\,\mathrm{var}(\hat{\beta}_1)=0$$

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**
$\left(X_i-\bar{X}\right)$	1-	0	3-	1	3	**0**
$\left(Y_i-\bar{Y}\right)$	6-	1	3-	0	8	**0**
X_iY_i	6	36	5	40	112	**199**
$\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)$	6	0	9	0	24	**39**
$\left(X_i-\bar{X}\right)^2$	1	0	9	1	9	**20**
$\hat{Y}_i$	6.05	8	2.15	9.95	13.85	**40**
$\hat{\varepsilon}_i=\left(Y_i-\hat{Y}_i\right)$	-4.05	1	2.85	1.95-	2.15	**0**
$\hat{\varepsilon}_i^2=\left(Y_i-\hat{Y}_i\right)^2$	16.4025	1	8.1225	3.8025	4.6225	**33.95**

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

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

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

لدينا:

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

إذا مقدرة التباين هي:

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

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

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

مثـال 2

البيانات التالية تعبر عن السعر وكمية الفراولة الذي تم بيعها في أحد أسواق الفواكه والخضار لمدينة عين الدفلى في مدة 12يوم إذا رمزنا للسعر بـ X وللكمية ب Y

باستخدام المعادلة التالية:

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

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

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

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

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

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

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

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

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

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

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

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

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

2- التقديــــر بمجــــال *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: $\hat{\theta}-k\prec\theta\prec\hat{\theta}+k$

و يمكن و بواسطة توزيع المعاينة للإحصاء θ أن نحدد قيمة k بحيث يكون الإحتمال :

$$P\left(\hat{\theta}-k\prec\theta\prec\hat{\theta}+k\right)$$

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

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

حيث أن: θ المعلمة المراد تقديرها، $\left(T_1,T_2\right)$ يسمى مجال ثقة ، T_1 حد الثقة الأدنى، T_2 حد الثقة الأعلى، $1-\alpha$ معامل الثقة حيث $0\prec\alpha\prec1$ ويحدد مسبقا، $(1-\alpha)\times100\%$ درجة الثقة للمعلمة θ .

1-2 مجــال الثقــة للمتوســط الحسابي للمجتمـــع μ_X

1-1-2 مجــال الثقــة للمتوســط الحسابي في حالـة توزيع المجتمــع طبيعـــي

نفرض أن $x_n \quad ......x_1, x_2$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N\left(\mu,\sigma^2\right)$ و أن $\bar{X}$ يمثل الوسط الحسابي لقياسات هذه العينة لايجاد فترة الثقة للمعلمة μ

1-1-1-2 حالـة تبايــن المجتمع معلومـا (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 :

$$prob\left[\mu_X \in \left(T_1,T_2\right)\right]=P\left(T_1\leq\mu_X\leq T_2\right)=1-\alpha$$

نسمي:

T_1: حـد الثقـة الأدنى ، T_2: حـد الثقـة الأعلى ، $1-\alpha$: مستوى الثقـة

$\left[T_1,T_2\right]$: مجــال الثقــة 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: $Z=\left(\bar{X}-\mu_{\bar{X}}\right)/\left(\sigma_{\bar{X}}\right)$

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

$$\sigma_{\bar{X}}^2=\begin{cases}\dfrac{\sigma_X^2}{n} & avec\ \ remise\\[2mm] \dfrac{\sigma_X^2}{n}\times\dfrac{N-n}{N-1} & sans\ \ remise\\[2mm] \dfrac{n}{N}\leq 0.05\Rightarrow\dfrac{N-n}{N-1}\simeq 1 \quad\Rightarrow\quad \dfrac{\sigma_X^2}{n}\times\dfrac{N-n}{N-1}\cong\dfrac{\sigma_X^2}{n}\end{cases}$$

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

وهي دالة في المقدر $\bar{X}$ 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: $Z_{\left(1-\frac{\alpha}{2}\right)}$ و $-Z_{\left(1-\frac{\alpha}{2}\right)}$ يساوي $1-\alpha$

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

وبالتعويض عن Z نحصل على :

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

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:

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

حيث أن:

T_1: حـد الثقـة الأدنى هو: $T_1=\bar{X}-Z_{(1-(\alpha/2))}\sigma_{\bar{X}}$

T_2: حـد الثقـة الأعلى هو: $T_2=\bar{X}+Z_{(1-(\alpha/2))}\sigma_{\bar{X}}$

$[T_1,T_2]$: مجـال الثقــة لمعلمة متوسط المجتمع μ_X هو: $\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))}\sigma_{\bar{X}}\right]_{1-\alpha}$

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

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

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

مثـال1:

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

حيث $X \sim N\left(\mu_X, \sigma_X^2\right)$ حيث $\sigma_X^2 = 4$.

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

الحـــل:

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

لدينا المجتمع غير محدود معناه حجم العينة مهمل أمام حجم المجتمع أي $\frac{n}{N} = \frac{5}{\infty} \approx 0 \leq 0.05$ و هذا معناه أن معامل الإرجاع يؤول إلى الواحد حيث: $\frac{N-n}{N-1} \simeq \frac{N}{N} \simeq 1$

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

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

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

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

$$Z_{(1-(\alpha/2))} = Z_{(1-(0.01/2))} = Z_{(1-0.005)} = Z_{(0.995)} = 2.58 \qquad \alpha = 0.01 \quad *$$
$$Z_{(1-(\alpha/2))} = Z_{(1-(0.05/2))} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96 \qquad \alpha = 0.05 \quad *$$
$$Z_{(1-(\alpha/2))} = Z_{(1-(0.1/2))} = Z_{(1-0.05)} = Z_{(0.95)} = 1.645 \qquad \alpha = 0.10 \quad *$$

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

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

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

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 لــ μ_X نقول أننا:

* واثقون بنسبة 99% بأن متوسط المجتمع μ_X ينتمي إلى المجال $[1.692 \quad , \quad 6.307]$
* واثقون بنسبة 95% بأن متوسط المجتمع μ_X ينتمي إلى المجال $[2.246 \quad , \quad 5.753]$

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

مثـــال 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 ؟

الحـــل:

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

لدينا مجتمع محدود وحجمه 4519 و العينة حجمها 50 بالتالي كسر المعاينة يساوي إلى $\frac{n}{N} = \frac{100}{4519} = 0.022 \leq 0.05$، معناه حجم العينة مهمل أمام حجم المجتمع و هذا معناه أن معامل الإرجاع 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: $\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))}\,\sigma_X/\sqrt{n} \quad , \quad \bar{X} + Z_{(1-(\alpha/2))}\,\sigma_X/\sqrt{n}\right]_{1-\alpha}$

* 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 لــ μ_X

$$\mu_X \in \left[385 - (1.96 * 10) \quad , \quad 385 + (1.96 \times 10)\right]_{0.95}$$
$$\mu_X \in [365.4 \quad , \quad 404.6]_{0.95}$$

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا μ_X ينتمي إلى المجال $[365.4 \quad , \quad 404.6]$

2-1-1-2 حالـة تبايــن المجتمع غيـر معلــوم (σ_X^2 مجهولـــة)

في هذه الحالة فإننا نواجه معلمتين إثنتين 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: العينــات الصغيــرة

إذا كان تبايـن المجتمـع مجهــول والمجتمـع طبيعـي وحجـم العينـة صغيـر $n \prec 30$ 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 العينة المسحوبة S_X^2 حيث: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

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

تقديــر σ_X^2 باستعمــال تبايــن العينـــة S_X^2:

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

- **إذا كــان السحــب بإرجــاع:**

$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$ و $E\left(S_X^2\right) = \sigma_X^2 \frac{n-1}{n}$ ، ومنه فإن تقدير تباين المجتمع σ_X^2 من خلال تباين العينة المسحوبة S_X^2 يكون كالتالي: $\hat{\sigma}_X^2 = \hat{S}_X^2 = S_X^2 \frac{n}{n-1}$

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

- **إذا كــان السحــب بدون إرجــاع:**

$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}\frac{N-n}{N-1}$ و $E\left(S_X^2\right) = \sigma_X^2 . \frac{n-1}{n} . \frac{N}{N-1}$ ، ومنه فإن تقدير تباين المجتمع σ_X^2 من خلال تباين العينة المسحوبة S_X^2 يكون كالتالي: $\hat{\sigma}_X^2 = S_X^2 \frac{n}{n-1}\frac{N-1}{N}$ أي أن:

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

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

$$\sigma_{\bar{X}}^2=\begin{cases}\dfrac{\sigma_X^2}{n}=\dfrac{\hat{\sigma}_X^2}{n}=\dfrac{s_x{}^2}{n-1} & avec\ remise\\ \dfrac{\sigma_X^2}{n}.\dfrac{N-n}{N-1}=\dfrac{\hat{\sigma}_X^2}{n}.\dfrac{N-n}{N-1}=\dfrac{s_x{}^2}{n-1}.\dfrac{N-n}{N}= & sans\ remise\\ \dfrac{n}{N}\le 0.05\Rightarrow\dfrac{N-n}{N}\simeq 1\ \Rightarrow\ \dfrac{s_x^2}{n-1}.\dfrac{N-n}{N}\cong\dfrac{s_x^2}{n-1} & \end{cases}$$

تحديـد الكميـة المحوريـة المناسبـة 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: $T=\dfrac{Z}{\sqrt{\chi^2/n-1}}$ يتبع توزيع t بدرجات حرية $n-1$ حيث: $Z=\left(\dfrac{\bar{X}-\mu_X}{\sigma_{\bar{X}}}\right)\sim N\,(0,1)$ و $\dfrac{nS_X^2}{\sigma_X^2}\sim\chi^2_{(n-1)}$ ، إذن:

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

$$T=\begin{cases}\left[\left(\bar{X}-\mu_X/(\sigma_X/\sqrt{n})\right)\middle/\left(\left(\dfrac{\sqrt{n}\,S_x}{\sigma_X}\right)\middle/\sqrt{n-1}\right)\right] & avec\ remise\\ \left[\left(\bar{X}-\mu_X\middle/\left((\sigma_X/\sqrt{n}).\left(\sqrt{\dfrac{N-n}{N-1}}\right)\right)\right)\middle/\left(\left(\dfrac{\sqrt{n}\,S_x}{\sigma_X}\right)\middle/\sqrt{n-1}\right)\right] & sans\ remise\end{cases}$$

$$T=\begin{cases}\bar{X}-\mu_X/\left(S_x/\sqrt{n-1}\right) & \sim t_{(n-1)}\quad avec\ remise\\ \bar{X}-\mu_X\middle/\left(\left(S_x/\sqrt{n-1}\right).\left(\sqrt{\dfrac{N-n}{N-1}}\right)\right) & \sim t_{(n-1)}\quad sans\ remise\end{cases}$$

تحديد مجال الثقة لمتوسط المجتمع باستعمال توزيع ستيودنت:

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 المجدولة و بالتعويض عن T نحصل على ما يلي:

-إذا كــان السحــب بإرجــاع:

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

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

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

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

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

حيث أن:

حـد الثقـة الأدنى هو: $\bar{X}-t_{(1-(\alpha/2),n-1)}.S_x/\sqrt{n-1}$

حـد الثقـة الأعلى هو: $\bar{X}+t_{(1-(\alpha/2),n-1)}.S_x/\sqrt{n-1}$

مجـال الثقـــة لمعلمة متوسط المجتمع μ_X هو: $\mu_X \in \left[\bar{X} \pm t_{(1-(\alpha/2),n-1)}.S_x/\sqrt{n-1}\right]_{1-\alpha}$

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

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

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

-إذا كــان السحــب بدون إرجــاع:

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

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

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

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

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

حيث أن:

حـد الثقـة الأدنى هو: $\bar{X}-t_{(1-(\alpha/2),n-1)}.\frac{S_x}{\sqrt{n-1}}\frac{\sqrt{N-n}}{\sqrt{N-1}}$

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

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

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

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

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

مثـال1: 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 $X\sim N\left(\mu_X,\sigma_X^2\right)$ حيث σ_X^2 مجهولة.

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

الحـــل:

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

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

$$\sigma_{\bar{X}}^2=\begin{cases}\frac{\sigma_X^2}{n}=\frac{s_x^{\ 2}}{n-1} & avec\ \ remise\\ \frac{\sigma_X^2}{n}\cdot\frac{N-n}{N-1}=\frac{s_x^{\ 2}}{n-1}\cdot\frac{N-n}{N-1}= & sans\ \ remise\\ \frac{n}{N}\le 0.05\Rightarrow\frac{N-n}{N-1}\simeq 1\quad\Rightarrow & \frac{s_x^2}{n-1}\cdot\frac{N-n}{N-1}\cong\frac{s_x^2}{n-1}\end{cases}$$

لدينا المجتمع غير محدود معناه حجم العينة مهمل أمام حجم المجتمع أي $\frac{n}{N}=\frac{5}{\infty}\approx 0\le 0.05$ 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:

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

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

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

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

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

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

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

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

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

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

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

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

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 لــ 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 $[3.53629\ ,\ 4.46371]$

الحالـة الثانيـة: العينـــات الكبيـــرة

إذا كان تبايـن المجتمـع مجهـول والمجتمـع طبيعـي وحجـم العينـة 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 كبيرا، فعندما تزداد n فإن S_X^2 تقترب من $\sigma_X{}^2$ و يقترب التوزيع الجديد من التوزيع الطبيعي المعياري.

مجـال الثقــة لمعلمة متوسط المجتمع μ_X هو:

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

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

مثـال1:

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 ؟

الحـــل:

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

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

لدينا مجتمع محدود وحجمه 4519 و العينة حجمها 50 بالتالي كسر المعاينة يساوي إلى $\frac{n}{N} = \frac{100}{4519} = 0.022 \leq 0.05$، معناه حجم العينة مهمل أمام حجم المجتمع و هذا معناه أن معامل الإرجاع 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: $\mu_X \in \left[\bar{X} - Z_{(1-(\alpha/2))}\sigma_X/\sqrt{n} \quad , \quad \bar{X} + Z_{(1-(\alpha/2))}\sigma_X/\sqrt{n} \right]_{1-\alpha}$

* 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 لــ μ_X

$$\mu_X \in [385-(1.96*10) \quad , \quad 385+(1.96\times10)]_{0.95}$$

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

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا μ_X ينتمي إلى المجال $[365.4 \quad , \quad 404.6]$

2-1-2 مجـال الثقـة للمتوسـط الحسابي في حالـة توزيع المجتمـع غير طبيعـــي أو مجهــول التوزيـع

نفرض أن $x_1, x_2 ...... \quad x_n$ تمثل قياسات عينة مسحوبة من مجتمع غير طبيعي أو مجهول التوزيع و أن $\bar{x}$ يمثل الوسط الحسابي لقياسات هذه العينة لايجاد فترة الثقة لمعلمة المجتمع μ_X :

1-2-1-2 حالـة تبايـــن المجتمع معلومـا (σ_X^2 معلومة)

الحالـة الأولى: العينــــات الصغيــرة

إذا كان تبايـن المجتمـع معلوم والمجتمـع غير طبيعـي وحجـم العينـة صغيـر $n \prec 30$ ففي هذه الحالة لا نستطيع الاعتماد في تحديد مجال الثقة لمتوسط المجتمع μ_X على توزيع ستيودنت 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:

نظريــة تشيبيشيـف

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

نفرض أن X متغير عشوائي بدالة احتمالية $f(x)$ بوسط وتباين محدودين هما على التوالي μ و σ_X^2 ولنفرض أن k عدد موجب، عندئذ (**هرمـز، 1990، الصفحات 477–479**):

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

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

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

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

- 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 $x-\mu \geq k\sigma$

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

$$= k^2\sigma^2 \left[\int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu+k\sigma}^{+\infty} f(x)dx \right]$$

$$= k^2\sigma^2 \left[P(X \leq \mu - k\sigma) + P(X - \mu \geq k\sigma) \right]$$

$$= k^2\sigma^2 \left[P(X - \mu \leq -k\sigma) + P(X - \mu \geq k\sigma) \right]$$

$$= k^2\sigma^2 P\left[|X - \mu| \geq k\sigma \right]$$

$$\sigma^2 \geq k^2\sigma^2 P\left[|X - \mu| \geq k\sigma \right] \quad \cdots(**)$$

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

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

أي أن: $P\left[|X - \mu| \prec k\sigma \right] \geq 1 - \frac{1}{k^2}$ وبفرض أن أي أن $c = k\sigma \succ 0$ عندئذ فإن هذه المتباينة ستكون كالتالي:

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

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

$$\mu_X \in \left[\bar{X} - k.\sigma_{\bar{X}} \quad , \quad \bar{X} + k.\sigma_{\bar{X}} \right]_{1-\alpha}$$

$$\mu_X \in \left[\bar{X} - k.\sigma_{\bar{X}} \quad , \quad \bar{X} + k.\sigma_{\bar{X}} \right]_{1-\frac{1}{k^2}}$$

إذن، مجـال الثقـة لمعلمة متوسط المجتمع μ_X هو:

* إذا كان السحب بإرجاع:

$$\mu_X \in \left[\bar{X} \pm k . \frac{\sigma_X}{\sqrt{n}} \right]_{1-\left(1/k^2\right)}$$

* إذا كان السحب بدون إرجاع:

$$\mu_X \in \left[\bar{X} \pm k . \frac{\sigma_X}{\sqrt{n}} \frac{\sqrt{N-n}}{\sqrt{N-1}} \right]_{1-\left(1/k^2\right)}$$

مثـال 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.

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

الحــل:

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

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

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

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

لدينا مجتمع حجمه 1500 و العينة حجمها 20 بالتالي كسر المعاينة يساوي إلى $\frac{n}{N} = \frac{20}{1500} = 0.013 \leq 0.05$، معناه حجم العينة مهمل أمام حجم المجتمع و هذا معناه أن معامل الإرجاع 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: $\sigma_{\bar{X}} = \left(\sqrt{\sigma_X^2/n}\right) = \left(\sqrt{25/20}\right) = \left(\sqrt{1.25}\right) = 1.118$

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

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

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

* تحديد مجال الثقة لــ μ_X

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

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

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

نقول أننا واثقون بنسبة 95% بأن متوسط إنتاجية فلاحي ولاية عين الدفلى لمحصول البطاطا μ_X ينتمي إلى المجال $[30 \quad , \quad 40]$

الحالـة الثانيـة: العينـــات الكبيـــرة

إذا كان تبايـن المجتمـع معلوم والمجتمـع غير طبيعـي وحجـم العينـة كبير $n \geq 30$ ففي هذه الحالة نعتمد في تحديد مجال الثقة لمتوسط المجتمع μ_X على نظرية النهاية المركزية فإن توزيع المجتمع المجهول يؤول إلى التوزيع الطبيعي المعياري.

مجـال الثقــة لمعلمة متوسط المجتمع μ_X هو:

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

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

مثـال 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 ؟

الحـــل:

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

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

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

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

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

* تحديد مجال الثقة لـــ μ_X

$$\mu_X \in \left[34.2-(1.555*0.395) \quad , \quad 34.2+(1.555\times 0.395)\right]_{0.88}$$
$$\mu_X \in [33.585 \quad , \quad 34.814]_{0.88}$$

نقول أننا واثقون بنسبة 88% بأن متوسط وزن خرفان التسمين الموجه للاستهلاك في عيد الأضحى μ_X ينتمي إلى المجال $[33.585 \quad , \quad 34.814]$

2-2-1-2 حالــة تبايــن المجتمع غير معلـــوم (σ_X^2 مجهولـة)

إذا كان تبايـن المجتمـع غير معلوم والمجتمـع غير طبيعـي وحجـم العينـة 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 نتمكن 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 مجــال الثقــة لمعلمة متوسط المجتمع μ_X هو:

* إذا كان السحب بإرجاع:

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

* إذا كان السحب بدون إرجاع:

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

مثــال1:

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 $S_x = 2$

الحــل:

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

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

* تحديد قيم معامل الثقة:

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

* تحديد مجال الثقة لـ μ_X:

$$\mu_X \in \left[12 \pm \left(1.96 * \frac{2}{\sqrt{150-1}} \sqrt{\frac{1500-150}{1500-1}} \right) \right]_{0.95}$$

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

جدول يلخص مجالات الثقة لمتوسط المجتمع μ_X

توزيـع المجتمـع	تبايـن المجتمـع	حجـم العينــة	نـوع السحـب	مجـال الثقــة
المجتمـع طبيعــي	معلـوم	لا يهـم	بإرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{(1-(\alpha/2))}\cdot\frac{\sigma_X}{\sqrt{n}}\right]_{1-\alpha}$
			بدون إرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right]_{1-\alpha}$
	غيـر معلـوم	$n<30$	بإرجـاع	$\mu_X \in \left[\bar{X} \pm t_{\left(1-\frac{\alpha}{2},n-1\right)}\cdot\frac{S}{\sqrt{n-1}}\right]_{1-\alpha}$
			بدون إرجـاع	$\mu_X \in \left[\bar{X} \pm t_{\left(1-\frac{\alpha}{2},n-1\right)}\cdot\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\right]_{1-\alpha}$
		$n \geq 30$	بإرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n-1}}\right]_{1-\alpha}$
			بدون إرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\right]_{1-\alpha}$
المجتمـع غيـر طبيعــي	معلـوم	$n<30$	بإرجـاع	$\mu_X \in \left[\bar{X} \pm k\cdot\frac{\sigma_X}{\sqrt{n}}\right]_{1-\left(1/k^2\right)}$
			بدون إرجـاع	$\mu_X \in \left[\bar{X} \pm k\cdot\frac{\sigma_X}{\sqrt{n}}\frac{\sqrt{N-n}}{\sqrt{N-1}}\right]_{1-\left(1/k^2\right)}$
		$n \geq 30$	بإرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma_X}{\sqrt{n}}\right]_{1-\alpha}$
			بدون إرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right]_{1-\alpha}$
	غير	$n \geq 30$	بإرجـاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n-1}}\right]_{1-\alpha}$

	معلــوم			بدون بإرجــاع	$\mu_X \in \left[\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]_{1-\alpha}$

2-2 مجــال الثقــة لتبايــن المجتمــع σ_X^2 :

نفرض أن $x_1, x_2 ...... x_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N\left(\mu, \sigma^2\right)$ و نرغب في إيجاد مجال الثقة لتباين المجتمع σ_X^2 وهنا لدينا حالتين (**هرمـز، 1990، الصفحات 631–633**):

1-2-2 متوســط المجتمــع معلومــا (μ_X معلــوم)

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

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

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

لذلك يمكننا أن نكتب:

$$P\left(\chi^2_{\left(\frac{\alpha}{2},n\right)} \prec \frac{n . S_X^2}{\sigma_X^2} \prec \chi^2_{\left(1-\frac{\alpha}{2},n\right)}\right) = 1-\alpha$$

و بتقسيم كافة أطراف المتباينة السابقة عل $n\,.S_X^2$ ثم نأخذ مقلوب هذه المتباينة المزدوجة ، نصل إلى مجـال الثقـــة لمعلمة تباين المجتمع σ_X^2 هو:

$$P\left(\frac{n\,.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n\right)}} \prec \sigma_X^2 \prec \frac{n\,.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n\right)}}\right)=1-\alpha$$

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

، حـد الثقـة الأدنى هو: $\dfrac{n\,.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n\right)}}$ ، و حـد الثقـة الأعلى هو: $\dfrac{n\,.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n\right)}}$.

مجـال الثقـــة لمعلمة تباين المجتمع σ_X^2 هو:

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

ومنه فإن مجال الثقة للإنحراف المعياري هو :

$$P\left(\sqrt{n\,.S_X^2 \big/ \chi^2_{\left(1-\frac{\alpha}{2},n\right)}} \prec \sigma_X \prec \sqrt{n\,.S_X^2 \big/ \chi^2_{\left(\frac{\alpha}{2},n\right)}}\right)=1-\alpha$$

مثـــال:

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

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

المطلـوب: أوجد مجال الثقة لتباين المجتمع بمستوى ثقة 0.95؟

الحـل:

لدينا: $1-\alpha=0.95 \Rightarrow \alpha=0.05 \Rightarrow \dfrac{\alpha}{2}=0.025 \quad and \quad 1-\dfrac{\alpha}{2}=0.975$

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

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

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

وعليه فإن:

حـد الثقـة الأدنى هو: $\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n\right)}}=\frac{20\times16}{34.1696}=9.365$

و حـد الثقـة الأعلى هو: $\frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n\right)}}=\frac{20\times16}{9.59078}=33.365$

مجـال الثقــة لمعلمة تباين المجتمع σ_X^2 هو: $P\left(\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n\right)}}\prec\sigma_X^2\prec\frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n\right)}}\right)=1-\alpha$

وهذا يعني أن: $P\left(9.365\prec\sigma_X^2\prec33.365\right)=1-\alpha$ معناه: $\sigma_X^2\in[9.365\ ,\ 33.365]_{0.95}$

أي أننا واثقون بنسبة 95 % من أن معلمة تباين المجتمع σ_X^2 تنتمي إلى هذا المجال.

2-2-2 إذا كان متوسـط المجتمـع مجهـولا (μ_X مجهـول)

إذا كان متوسط المجتمـع μ_X مجهولا فإن: $S_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ يمثل تباين العينة على أساس متوسط العينة $\bar{x}$، و $\hat{S}_x^2=\frac{n}{n-1}S_X^2$ حيث:

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

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

$$\chi^2=\frac{n.S_X^2}{\sigma_X^2}=\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_X}\right)^2\sim\chi^2_{(n-1)}$$

لذلك يمكننا أن نكتب:

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

حـد الثقـة الأدنى هو: $\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}$ ، و حـد الثقـة الأعلى هو: $\frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}$ ،

مجــال الثقــة لمعلمة تباين المجتمع σ_X^2 هو:

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

ومنه فإن مجال الثقة للإنحراف المعياري هو :

$$P\left(\sqrt{\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}} \prec \sigma_X \prec \sqrt{\frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}}\right)=1-\alpha$$

مثـال:

إذا علمت أن تباين عينة عشوائية ذات حجم $n=17$ مسحوبة من مجتمع طبيعي متوسطه مجهول $X \sim N\left(\mu_X ,\sigma_X^2\right)$ هو $S_x{}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{17}\sum_{i=1}^{17}(x_i-\bar{x})^2=12$

المطلـوب: أوجد مجال الثقة لتباين المجتمع بمستوى ثقة 0.90؟

الحـل:

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

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:

حـد الثقـة الأدنى هو:

$$\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}=\frac{17\times 12}{26.2962}=7.757$$

و حـد الثقـة الأعلى هو:

$$\frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}=\frac{17\times 12}{7.96165}=25.6228$$

مجــال الثقــة لمعلمة تباين المجتمع σ_X^2 هو:

$$P\left(\frac{n.S_X^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}} \prec \sigma_X^2 \prec \frac{n.S_X^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}\right)=1-\alpha$$

وهذا يعني أن: $P\left(7.75 \prec \sigma_X^2 \prec 25.62\right)=1-\alpha$ معناه: $\sigma_X^2 \in [7.757 \quad , \quad 25.6228]_{0.90}$

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

3-2 مجــال الثقــة للنسبــة

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 (T_1,T_2) أي: $P\left[p \in (T_1,T_2)\right]=1-\alpha$

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:

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

كذلك $V(X)=n.p.q$ ومنه

$$V\left(\frac{X}{n}\right)=V(\hat{p})=\frac{1}{n^2}V(X)=\frac{1}{n^2}.n.p.q=\frac{p.q}{n}$$

و إذا كانت n كبيرة و و فقا لنظرية النهاية المركزية فإن التوزيع التقريبي ل X هو: $Z=\dfrac{(X-n.p)}{\sqrt{n.p.q}}$

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

و بالتعويض بقيمة Z نحصل على

$$P\left(-Z_{(1-(\alpha/2))} \prec \frac{(X-n.p)}{\sqrt{n.p.q}} \prec Z_{(1-(\alpha/2))}\right)=1-\alpha$$

و بالتالي فإن:

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

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 الإنحراف المعياري $\sigma_p^2=\dfrac{p.q}{n}$ 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 بناءا على ذلك فإنه من المنطقي أن نستعمل p' أيضا من أجل تقدير p و من ثم q الموجودة تحت الجذر أي في عبارة

الإنحراف المعياري $\sqrt{\dfrac{p.q}{n}}$ 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 $p'=\dfrac{X}{n}$ مكان p و إدخال p' داخل الجذر التربيعي يمكننا من الوصول إلى مجال الثقة للنسبة p كما يلي:

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

و بالتالي فإن :

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

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

مجـال الثقــة لمعلمة النسبة في المجتمع p هو:

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

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 .

ملاحظـــة :

بفرض أن عملية السحب كانت تتم بدون إرجاع و من مجتمع صغير (منته) ففي هذه الحالة يجب أن نأخذ مؤشر عدم الإعادة بعين الإعتبار $\left(\dfrac{N-n}{N-1}\right)$ و عندها فإن العلاقة السابقة تصبح كما يلي

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

المحــــور الرابــع

نظريـــة إختبــار الفرضيــات

Hypothesis Testing Theory

نظريــة إختبــار الفرضيــات

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- مفاهيــم أساسيــة في نظريــة الإختبــار

الفرضيــة الإحصائيــة ***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.

أنــواع الفــروض الإحصائيــة

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

أ- الفـرض البسيـط

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

$$H_0 : \mu_X = 100 \quad ، \quad H_1 : \mu_X = 95$$

ب- الفـرض المركـب

الفرض المركب هو الذي يحدد أية قيمة لمعلمة المجتمع في مجموعة من القيم المعلمة المجتمع المجهولة، مثلا الفرضيتين $H_0 : \mu_X \leq 100$ ، $H_1 : \mu_X \succ 65$

أصنــاف الفــروض الإحصائيــة

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

أ- فرضيــة العــدم ***null hypothesis*** $H_0 : \theta = \theta_0$

تسمى الفرضية التي يجري إختبارها بفرضية العدم وتصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو عدم وجود تغير، أي تعكس عدم الإختلاف وبالتالي تظهر فيها عبارة التساوي، ويرمز لها بـ H_0.

ب- الفرضيـة البديلـة ***alternative hypothesis*** H_1

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

حـالات الفرضيـة البديلـة:

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

أ -الشكـل الأول: $H_1 : \theta \succ \theta_0$

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

ب – الشكـل الثانـي: $H_1 : \theta \prec \theta_0$

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

يعرف هذين النوعين من الفرضية البديلة بالفرض أحادي الطرف

ج – الشكـل الثالـث: $H_1 : \theta \neq \theta_0$

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

إختبـار الفرضيـة

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 .

أنـواع الأخطـاء

هناك نوعان من الأخطاء : خطأ من النوع الأول وخطأ من النوع الثاني (**العلي و صبوح، 2020، صفحة 69**) .

أ- خطـأ مـن النـوع 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 بــ α وهو مستوى الدلالة

α: هو إحتمال رفض فرض العدم بالرغم من أنه في الواقع صحيح .

ب- خطـأ مـن النـوع الثانـــي

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 ب 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- إختبـارات الفرضيـات حـول معالـم المجتمـع الإحصـائي

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 لأخر من خلال إعتماد 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

2-1- إختبـارات الفرضيـات حول المتوســط الحســابي لمجتمـع طبيعـي μ_X

يتم إجراء إختبارات الفرضيات حول المتوسط الحسابي للمجتمع في حالتين مختلفتين، أننا نعرف قيمة الإنحراف المعياري، أننا لا نعرف الإنحراف المعياري للمجتمع وعندها نعمل على تقديره باستخدام الإنحراف المعياري للعينة

2-1-1 تبايـن المجتمـــع معلـوم σ_X^2

نظريـــة 1:

إذا كانت $X_1, X_2 ...... \quad X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N\left(\mu,\sigma^2\right)$ و كان تباين المجتمع σ_X^2 معلوما فإن دالة الاختبار هي (**أبو صالـح و عـوض، 1983، الصفحات 167−168**):

$$Z = \frac{\overline{x} - \mu_0}{\sigma_x / \sqrt{n}}$$ 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: $Z \leq -Z_{(1-(\alpha/2))} \quad or \quad Z \geq Z_{(1-(\alpha/2))}$

* وكانت $H_1 : \mu \succ \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z \succ Z_{(1-(\alpha/2))}$

* وكانت $H_1 : \mu \prec \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z \prec -Z_{(1-(\alpha/2))}$

مثـــال:

إذا كانت X_{49} ......X_1, X_2 تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N(\mu,16)$

و كان $\sum_{i=1}^{49} X_i = 441$ 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: $(Z = 3.5) \succ (Z_{(0.975)} = 1.96)$، فإننا نرفض $H_0 : \mu = \mu_0 = 7$ على مستوى الدلالة $\alpha = 0.05$

2-1-2 تبايــن المجتمــع غيــر معلــوم أي مجهــول σ_X^2

عندما يكون تباين المجتمع مجهولا فإننا نستخدم القيمة التقديرية له s المحسوبة من المشاهدات في العينة والمعرف كما يلي:

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

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 العينـــات الصغيـرة

إذا كان تبايـن المجتمـع مجهـول والمجتمـع طبيعـي وحجـم العينـة صغيـر $n \prec 30$ ففي هذه الحالة 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 حيث: $S_x{}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ و نعلم أن: $\hat{S}_x{}^2 = \frac{n}{n-1} S_X^2$

قيمة دالة الاختبار T هو:

* إذا كان السحب بإرجاع:

$$T = \frac{\bar{x} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{x} - \mu_0}{\hat{S}_x / \sqrt{n}}$$

* إذا كان السحب بدون إرجاع: $T = \dfrac{\bar{x} - \mu_0}{\left(S_x / \sqrt{n-1}\right).\left(N - n / N - 1\right)}$

نظرية 2:

إذا كانت $X_1, X_2 ...... \quad X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N\left(\mu, \sigma^2\right)$ و كان تباين المجتمع σ_X^2 مجهولا فإن دالة الاختبار 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: $T \leq -t_{(1-(\alpha/2), n-1)} \quad or \quad T \geq t_{(1-(\alpha/2), n-1)}$

* وكانت $H_1 : \mu \succ \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $T \succ t_{(1-(\alpha/2), n-1)}$

* وكانت $H_1 : \mu \prec \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $T \prec -t_{(1-(\alpha/2), n-1)}$

مثـــال:

إذا كانت $X_1, X_2 ...... \quad X_{26}$ تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N\left(\mu, \sigma^2\right)$ و كان $\bar{x} = 65$ و $S_X^2 = 9$

وأردنا اختبار $H_0 : \mu = \mu_0 = 60$ مقابل $H_0 : \mu \neq \mu_0 \neq 60$ 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: $\left(t_{(0.975,25)} = 2.06\right) \prec \left(T = 8.33\right)$، فإننا نرفض $H_0 : \mu = \mu_0 = 60$ على مستوى الدلالة $\alpha = 0.05$

2-2-1-2 العينـــات الكبيـــرة

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 $\sigma_X{}^2$ و يقترب التوزيع الجديد من التوزيع الطبيعي المعياري.

قيمة دالة الاختبار Z هو:

* إذا كان السحب بإرجاع:
$$Z = \frac{\bar{x} - \mu_0}{S_x / \sqrt{n-1}} = \frac{\bar{x} - \mu_0}{\hat{S}_x / \sqrt{n}}$$

* إذا كان السحب بدون إرجاع:
$$Z = \frac{\bar{x} - \mu_0}{\left(S_x / \sqrt{n-1}\right).\left(N - n / N - 1\right)}$$

نظرية 3:

إذا كانت $X_1, X_2 ...... \quad X_n$ تمثل قياسات عينة مسحوبة من مجتمع طبيعي $N\left(\mu, \sigma^2\right)$ 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: $Z \leq -Z_{(1-(\alpha/2))} \quad or \quad Z \geq Z_{(1-(\alpha/2))}$

* وكانت $H_1 : \mu \succ \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z \succ Z_{(1-(\alpha/2))}$

* وكانت $H_1 : \mu \prec \mu_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $Z \prec -Z_{(1-(\alpha/2))}$

مثــال:

إذا كانت X_{82} ...... X_1, X_2 تمثل قياسات عينة عشوائية مسحوبة من مجتمع طبيعي $N\left(\mu,\sigma^2\right)$

و كان $\bar{x}=115$ و $S_X^2=25$ وحجـم العينـة $n=82$

وأردنا اختبار $H_0:\mu=\mu_0=100$ مقابل $H_0:\mu\neq\mu_0\neq100$ 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: $\left(Z_{(0.975)}=1.96\right)\succ\left(Z=27\right)$، فإننا نرفض $H_0:\mu=\mu_0=100$ على مستوى الدلالة $\alpha=0.05$

2-2 إختبــارات الفرضيــات حــول مقارنــة وسطيــن حسابييــن μ_X

في كثير من الأحيان يطلب منا دراسة متوسطي عينتين لمعرفة ما إذا كان الفرق بينهما كبيرا بدرجة تمكننا من القول بأنهما مسحوبتين من مجتمعين مختلفين ، أم الفرق صغير بدرجة يمكن إرجاعها للصدفة .

ليكن لدينا مجتمعين X و Y كبيرين كبرا كافيا لهما متوسطين حسابيين μ_X و μ_Y و تباينين σ_1^2 و σ_2^2 نسحب عينتين عشوائيتين x و y من المجتمعين X و Y ، العينة الأولى بحجم n_1 لها متوسط حسابي قدره $\overline{X}$ و العينة الثانية بحجم n_2 لها متوسط حسابي قدره $\overline{y}$

المطلوب هو معرفة ما إذا كان هناك فرق جوهري بين الوسطين المجهولين μ_X و μ_Y إعتمادا على الوسطين $\overline{X}$ و $\overline{y}$ المستخلصين من العينتين 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 العينتــان كبيرتــان: $(n_1 \geq 30 : n_2 \geq 30)$

1-1-2-2 الإنحرافـــان المعياريـــان للمجتمعـــان معلومـــان

إذا كان حجم العينتين كبيرا فإننا نعلم حسب نظرية توزيع المعاينات أن قوانين الإحتمال لكل من $\bar{x}$ و $\bar{y}$ هي:

$$\bar{y} \sim N\left(\mu_2 ، \sigma_2 / \sqrt{\ \ _2}\right) \quad ، \quad \bar{x} \sim N\left(\mu_1 ، \sigma_1 / \sqrt{\ \ _1}\right)$$

ومنه فإن قانون الإحتمال للمتغير العشوائي الناتج عن الفرق $(\bar{x} - \bar{y})$ هو:

$$(\bar{x} - \bar{y}) \sim N\left(\mu_1 - \mu_2 ; \sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}\right)$$

ويتضمن إختبار مقارنة الوسطين الإختبار بين الفرضيتين التاليتين :

فرضية العدم : $H_0 : \mu_1 - \mu_2 = 0 \Rightarrow \mu_1 = \mu_2$

الفرضية البديلة : $H_1 : \mu_1 - \mu_2 \neq 0 \Rightarrow \mu_1 \neq \mu_2$

فإذا كانت الفرضية الإبتدائية صحيحة أي إذا كان $\mu_1 = \mu_2$ فيكون قانون الإحتمال للفرق هو

$$(\bar{x} - \bar{y}) \sim N\left(0 ; \sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}\right)$$

و مؤشر الإختبار Z يخضع لقانون التوزيع الطبيعي المعياري $N\ (0,1)$ و يعرف بالعلاقة التالية :

$$Z = \frac{(\bar{x} - \bar{y}) - 0}{\sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}}$$

نظرية 4:

إذا كانت $x_1, x_2 ...... \quad x_{n1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_1, \sigma^2{}_1\right)$ ، وكانت $y_1, y_2 ...... \quad y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_2, \sigma^2{}_2\right)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض ، وكانت σ_1^2 و σ_2^2 معلومتين

أردنا اختبار الفرضية $H_0 : \mu_1 = \mu_2$ أي أن $H_0 : \mu_1 - \mu_2 = 0$ فإن دالة الاختبار هي:

$$\frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}}$$

* فإذا كانت $H_1 : \mu_1 \neq \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$Z \leq -Z_{(1-(\alpha/2))} \quad or \quad Z \geq Z_{(1-(\alpha/2))} \qquad \text{أي:} \quad |Z| = \left| \frac{(\bar{x} - \bar{y})}{\sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}} \right| \geq Z_{(1-(\alpha/2))}$$

* وإذا كانت $H_1 : \mu_1 \succ \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(\bar{x}-\bar{y})}{\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}} \succ Z_{(1-(\alpha/2))}$$

* وكانت $H_1 : \mu_1 \prec \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(\bar{x}-\bar{y})}{\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}} \prec -Z_{(1-(\alpha/2))}$$

مثــال:

إذا كانت $x_1, x_2 ...... \quad x_{25}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N(\mu_1, 25)$ ، وكانت $y_1, y_2 ...... \quad y_{16}$ 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 $\alpha = 0.05$

الحــل:

دالة الاختبار هي: $Z = \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}} = \frac{8.9-11.2}{\sqrt{(25/20)+(36/16)}} = \frac{-2.3}{\sqrt{3.5}} = -1.229$

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

ولما كانت: $\left(Z_{(0.975)}=1.96\right) \prec \left(|Z| = |-1.223|\right)$، فإننا نقبل $H_0 : \mu_1 = \mu_2$ على مستوى الدلالة $\alpha = 0.05$

2-1-2-2 الإنحرافــان المعياريــان للمجتمعــان متساويــان و مجهــولان $\sigma_1 = \sigma_2 = \sigma$

عندما يكون الإنحرافان المعياري للمجتمعين متساويين و مجهولين أي $\sigma_1 = \sigma_2 = \sigma$ فإن :

$$\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)} = \sigma\sqrt{(1/n_1)+(1/n_2)}$$

نعلم أن التقدير النقطي ل σ^2 هو: $S^2 = \frac{\left(S_1^2 n_1\right)+\left(S_2^2 n_2\right)}{n_1+n_2-2}$

حيث $s_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_i-\bar{x})^2$ و $s_2^2 = \frac{1}{n_2}\sum_{i=1}^{n_2}(y_i-\bar{y})^2$

وتكون الكمية $\sigma\sqrt{(1/n_1)+(1/n_2)}$ مقدرة بواسطة $S\sqrt{(1/n_1)+(1/n_2)}$

نظريــة 5:

إذا كانت $x_1, x_2 ...... \quad x_{n1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_1, \sigma^2{}_1\right)$ ، وكانت $y_1, y_2 ...... \quad y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_2, \sigma^2{}_2\right)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض ، وكانت $\sigma_1^2 = \sigma_2^2$ ولكنهما مجهولتين وأردنا اختبار الفرضية $H_0 : \mu_1 = \mu_2$ أي أن $H_0 : \mu_1 - \mu_2 = 0$ فإن دالة الاختبار هي:

$$T = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{S\sqrt{(1/n_1) + (1/n_2)}} ; t_{(1-(\alpha/2), n_1+n_2-2)}$$

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

* 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: $T \leq -t_{(1-(\alpha/2), n_1+n_2-2)} \quad or \quad T \geq t_{(1-(\alpha/2), n_1+n_2-2)}$

* وإذا كانت $H_1 : \mu_1 \succ \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S\sqrt{(1/n_1) + (1/n_2)}} \succ t_{(1-(\alpha/2), n_1+n_2-2)}$$

* وكانت $H_1 : \mu_1 \prec \mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T = \frac{(\bar{x} - \bar{y})}{S\sqrt{(1/n_1) + (1/n_2)}} \prec -t_{(1-(\alpha/2), n_1+n_2-2)}$$

مثــال:

أخذنا عينتين من توزيعين مستقلين $x_1, x_2 ...... \quad x_{11}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_1, \sigma_1^2\right)$ ، وكانت $y_1, y_2 ...... \quad y_9$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_2, \sigma_2^2\right)$ ، 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 اختبار الفرضية $H_0 : \mu_1 = \mu_2$ مقابل الفرضية $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$

الحــل:

نعلم أن $\sigma_1^2 = \sigma_2^2$ إذن $S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} = \frac{(11 \times 9) + (9 \times 12)}{11 + 9 - 2} = \frac{207}{18} = 11.5$

دالة الاختبار هي:

$$T=\frac{(\overline{x}_1-\overline{x}_2)}{S\sqrt{(1/n_1)+(1/n_2)}}=\frac{11.3-10.7}{\sqrt{11.5\times[(1/11)+(1/9)]}}=\frac{0.6}{1.52}=0.3936$$

$$t_{(1-(\alpha/2),n_1+n_2-2)}=t_{(0.975,11+9-2)}=t_{(0.975,18)}=2.101$$

ولما كانت: $\left(t_{(0.975.18)}=2.101\right)\succ(T=0.39)$، فإننا نقبل $H_0:\mu_1=\mu_2$ على مستوى الدلالة

$\alpha=0.05$

3-1-2-2 الإنحرافــان المعياريــان للمجتمعــان غيــر متساويــان و مجهــولان

عندما يكون الإنحرافان المعياري للمجتمعين غير متساويين و مجهولين أي $\sigma_1\neq\sigma_2$ فإن :

$$\sigma_{(\overline{x}-\overline{y})}=\sqrt{\left(\sigma_1^2/n_1\right)+\left(\sigma_2^2/n_2\right)}$$

فإذا كان حجم العينتين كبيرا فإن المقدار: $S_{(x-y)}=\sqrt{\left(S_x^{\,2}/n_1-1\right)+\left(S_y^{\,2}/n_2-1\right)}$ يمثل

تقدير نقطي للإنحراف المعياري $\sigma_{(\overline{x}-\overline{y})}$

حيث $s_x^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(x_i-\overline{x})^2$ و $s_y^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(y_i-\overline{y})^2$

نظريــة 6:

إذا كانت $x_1,x_2......\quad x_{n1}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_1,\sigma^2{}_1\right)$، وكانت $y_1,y_2......\quad y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_2,\sigma^2{}_2\right)$، 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:

$$T=\frac{(\overline{x}-\overline{y})-(\mu_1-\mu_2)}{\sqrt{(S_x/n_1-1)+(S_y/n_2-1)}};t_{(1-(\alpha/2),n_1+n_2-2)}$$

* فإذا كانت $H_1:\mu_1\neq\mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|T|=\left|\frac{(\overline{x}-\overline{y})}{\sqrt{(S_x/n_1-1)+(S_y/n_2-1)}}\right|\geq t_{(1-(\alpha/2),n_1+n_2-2)}$$

أي: $$T\leq -t_{(1-(\alpha/2),n_1+n_2-2)}\quad or\quad T\geq t_{(1-(\alpha/2),n_1+n_2-2)}$$

* وإذا كانت $H_1:\mu_1\succ\mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T=\frac{(\bar{x}-\bar{y})}{\sqrt{(S_x/n_1-1)+(S_y/n_2-1)}}\succ t_{(1-(\alpha/2),n_1+n_2-2)}$$

* وكانت $H_1:\mu_1\prec\mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T=\frac{(\bar{x}-\bar{y})}{\sqrt{(S_x/n_1-1)+(S_y/n_2-1)}}\prec -t_{(1-(\alpha/2),n_1+n_2-2)}$$

مثــال:

(نفس المثال السابق) أخذنا عينتين من توزيعين مستقلين $x_1,x_2......\quad x_{11}$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_1,\sigma_1^2\right)$، وكانت $y_1,y_2......\quad y_9$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu_2,\sigma_2^2\right)$ ، 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 $\alpha=0.05$

الحــــل:

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: $\left(t_{(0.975.18)}=2.101\right)\prec\left(T=0.387\right)$، فإننا نقبل $H_0:\mu_1=\mu_2$ على مستوى الدلالة $\alpha=0.05$

2-2-2 العينتــان صغيرتــان $(n_1\prec 30:n_2\prec 30)$

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 الإنحرافين σ_1 و σ_2 ، وإن لم يجد الباحثين في مجال الإحصاء طريقة أو أداة تمكن من التمييز بين السببين المذكورين ، لكن يمكن إعتماد ما يلي :

-إذا كان الإنحرافان المعياريان للمجتمعين متساوين و مجهولين $\sigma_1=\sigma_2=\sigma$ فإن التباين المشترك S تعتبر تقديرا نقطيا ل σ والكمية:

$$T=\frac{|\bar{x}-\bar{y}|}{S\sqrt{(1/n_1)+(1/n_2)}};t_{(1-(\alpha/2),n_1+n_2-2)}$$

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

* 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: $$T\leq -t_{(1-(\alpha/2),n_1+n_2-2)}\quad or\quad T\geq t_{(1-(\alpha/2),n_1+n_2-2)}$$

* وإذا كانت $H_1:\mu_1\succ\mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T=\frac{(\bar{x}-\bar{y})}{S\sqrt{(1/n_1)+(1/n_2)}}\succ t_{(1-(\alpha/2),n_1+n_2-2)}$$

* وكانت $H_1:\mu_1\prec\mu_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$T=\frac{(\bar{x}-\bar{y})}{S\sqrt{(1/n_1)+(1/n_2)}}\prec -t_{(1-(\alpha/2),n_1+n_2-2)}$$

3- إختبــار الفرضيــات المتعلقــة بالنسبــة والفــرق بيــن نسبتيــن

3-1- إختبــار الفرضيــات يتعلــق بالنسبــة 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 فبإمكاننا تقريب قانون التوزيع الثنائي بالتوزيع الطبيعي، أما إذا كان حجم العينة كبيرا وكانت P قريبة من الصفر أو الواحد فبإمكاننا تقريب قانون التوزيع الثنائي بالتوزيع البواسوني

نظريــة 7:

إذا كانت $X_1, X_2 ...... \quad X_n$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b(1, P)$ و وكان 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: $\left|\frac{P' - P_0}{\sqrt{P_0 (1 - P_0)/n}}\right| \geq Z_{(1-(\alpha/2))}$ أي:

$$\frac{P' - P_0}{\sqrt{P_0 (1 - P_0)/n}} \leq -Z_{(1-(\alpha/2))} \quad or \quad \frac{P' - P_0}{\sqrt{P_0 (1 - P_0)/n}} \geq Z_{(1-(\alpha/2))}$$

* وإذا كانت $H_1 : P \succ P_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $\frac{P' - P_0}{\sqrt{P_0 (1 - P_0)/n}} \succ Z_{(1-(\alpha/2))}$

* وكانت $H_1 : P \prec P_0$

فإننا نرفض H_0 على مستوى الدلالة α إذا كانت: $\frac{P' - P_0}{\sqrt{P_0 (1 - P_0)/n}} \prec -Z_{(1-(\alpha/2))}$

مثــال:

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.

الحــل:

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:

$H_0 : P = P_0 = 0.3$ مقابل $H_1 : P \succ 0.3$ على مستوى الدلالة $\alpha = 0.05$

$$\frac{P' - P_0}{\sqrt{P_0\left(1 - P_0\right)/n}} = \frac{\left(35/250\right) - 0.3}{\sqrt{0.3\left(1 - 0.3\right)/250}} = \frac{0.16}{0.0289} = 5.52$$

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

ولما كانت: $\left(Z_{(0.975)} = 1.96\right) \succ \left(Z = 5.52\right)$، 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

3-2- إختبـار الفرضيـات يتعلـق بالفـرق بيـن نسبتيــن

لنفرض أننا بصدد دراسة مجتمعان إحصائيان X و Y يتكون كل منهما من وحدات إحصائية بعضا 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 $q_1 = 1 - P_1$ و $q_2 = 1 - P_2$ و لنفرض أن النسبتين P_1 و P_2 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 إعتمادا على النسبتين P_1' و P_2' المستخلصين من العينيتن x و y المسحوبتين من المجتمعين X و Y

-إذا كان حجم العينينن كبيرا فإن قانون التوزيع الإحتمالي لكل من P_1' و P_2' هو :

$$P_1' \sim N\left(p_1.\sqrt{\left(P_1\left(1 - P_1\right)/n_1\right)}\right) \quad and \quad P_2' \sim N\left(p_2.\sqrt{\left(P_2\left(1 - P_2\right)/n_2\right)}\right)$$

وهذا يعني أن التوقع الرياضي للنسبة P_1' يساوي إلى P_1 عندما يكون حجم العينة كبيرا و كما أن الإنحراف المعياري للنسبة P_1' هو $\sqrt{\left(P_1\left(1 - P_1\right)/n_1\right)}$ و نفس المبدأ ينطبق على النسبة P_2' ، وعندها يكون قانون التوزيع الإحتمالي للمتغير العشوائي $\left(P_1' - P_2'\right)$

$$\left(P_1' - P_2'\right) \sim N\left(P_1 - P_2 ; \sqrt{\left(P_1\left(1-P_1\right)/n_1 + P_2\left(1-P_2\right)/n_2\right)}\right)$$

إن الإنحراف المعياري يتعلق بالمؤشرين المجهولين P_1 و P_2 و يتم تقييمهما بواسطة التقدير النقطي التالي : لتكن P' تمثل تقديرا نقطيا لـ P فإن أفضل تقدير هو في هذه الحالة : $P' = \left(n_1 P_1' + n_2 P'\right)/\left(n_1 + n_2\right)$

و الإنحراف المعياري للفرق $\left(P_1' - P_2'\right)$ يساوي

$$\sigma_{(P_1'-P_2')} = \sqrt{\left(P'(1-P')/n_1\right) + \left(P'\left(1-P'\right)/n_2\right)} = \sqrt{P'(1-P')\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

و يصبح قانون توزيع $\left(P_1' - P_2'\right)$ كما يلي $\left(P_1' - P_2'\right) \sim N\left(P_1 - P_2 ; \sqrt{P'(1-P')(1/n_1 + 1/n_2)}\right)$

ويتضمن إختبار مقارنة النسبتين الإختبار بين الفرضيتين التاليتين :

- فرضية العدم $H_0 : P_1 - P_2 = 0 \Rightarrow P_1 = P_2$
- الفرضية البديلة $H_1 : P_1 - P_2 \neq 0 \Rightarrow P_1 \neq P_2$

فإذا كانت الفرضية الإبتدائية صحيحة أي إذا كان $P_1 = P_2$ فيكون قانون الإحتمال للفرق بين النسبتين يتبع التوزيع الطبيعي المعياري أي: $\left(P_1' - P_2'\right) \sim N\left(0 \quad ; \sqrt{P'(1-P')(1/n_1 + 1/n_2)}\right)$

و مؤشر الإختبار Z يعرف بالعلاقة التالية : $Z = \dfrac{\left(P_1' - P_2'\right) - 0}{\sqrt{\left(P'(1-P')/n_1\right) + \left(P'(1-P')/n_2\right)}}$

و يخضع لقانون التوزيع الطبيعي المعياري $N(0;1)$

نظرية 8:

إذا كانت $x_1, x_2 \ldots\ldots \quad x_{n1}$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b\left(1, P_1\right)$، وكانت $y_1, y_2 \ldots\ldots \quad y_{n2}$ تمثل قياسات عينة مسحوبة من توزيع ذات الحدين $b\left(1, P_2\right)$ ، وكانت هاتين العينتين مستقلتين عن بعضهما البعض ، وكان حجم العينـتين n_1, n_2 كبيرا

و أردنا اختبار الفرضية $H_0 : P_1 = P_2$ أي أن $H_0 : P_1 - P_2 = 0$ فإن دالة الاختبار هي:

$$\frac{\left(P_1' - P_2'\right) - \left(P_1 - P_2\right)}{\sqrt{\left(P'(1-P')/n_1\right) + \left(P'(1-P')/n_2\right)}}$$

$$P' = \frac{n_1 P_1' + n_2 P_2'}{n_1 + n_2}$$

* فإذا كانت $H_1 : P_1 \neq P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$|Z|=\left|\frac{(P_1'-P_2')}{\sqrt{\left(P'(1-P')/n_1\right)+\left(P'(1-P')/n_2\right)}}\right|\geq Z_{(1-(\alpha/2))}$$

أي:

$$Z\leq -Z_{(1-(\alpha/2))} \quad or \quad Z\geq Z_{(1-(\alpha/2))}$$

* وإذا كانت $H_1: P_1 \succ P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(P_1'-P_2')}{\sqrt{\left(P'(1-P')/n_1\right)+\left(P'(1-P')/n_2\right)}}\succ Z_{(1-(\alpha/2))}$$

* وكانت $H_1: P_1 \prec P_2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{(P_1'-P_2')}{\sqrt{\left(P'(1-P')/n_1\right)+\left(P'(1-P')/n_2\right)}}\prec -Z_{(1-(\alpha/2))}$$

مثـــال:

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 $H_0: P_1 \succ P_2$

الحــــل:

$$P_1'=\frac{n_a}{n_1}=\frac{1023}{3516}=0.2909 \qquad n_1=3516$$

$$P_2'=\frac{n_b}{n_2}=\frac{413}{1517}=0.2722 \qquad n_2=1517$$

$$P'=\frac{n_1P_1'+n_2P_2'}{n_1+n_2}=\frac{1023+413}{3516+1517}=\frac{1436}{5033}=0.2853$$

$$Z=\frac{\left(P_1'-P_2'\right)-\left(P_1-P_2\right)}{\sqrt{\left(P'\left(1-P'\right)/n_1\right)+\left(P'\left(1-P'\right)/n_2\right)}}$$

$$=\frac{0.2909-0.2722}{\sqrt{\frac{(0.2853*0.7147)}{3516}+\frac{(0.2853*0.7147)}{1517}}}$$

$$=\frac{0.0187}{\sqrt{0.00005799+0.0001344}}=1.348$$

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

ولما كانت: $\left(Z_{(0.975)}=1.96\right)\succ\left(Z=1.348\right)$، 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- إختبــارات تتعلـق بالتبايــن والمقارنــة بين تباينيـن

1-4 إختبــارات تتعلـق بالتبايـن

نظريـــة 9:

إذا كانت $x_1,x_2......\ x_n$ تمثل قياسات عينة مسحوبة من توزيع طبيعي $N\left(\mu,\sigma^2\right)$، وأردنا اختبار 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:

$$\frac{nS^2}{\chi^2\left[1-(\alpha/2);n-1\right]}\leq\sigma_0^2\leq\frac{nS^2}{\chi^2\left[(\alpha/2);n-1\right]}$$

* وإذا كانت $H_1:\sigma^2\succ\sigma_0^2$، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{nS^2}{\chi^2\left[1-\alpha\text{،}n-1\right]}\succ\sigma_0^2$$

* وكانت $H_1:\sigma^2\prec\sigma_0^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{nS^2}{\chi^2\left[\alpha\text{،}n-1\right]}\prec\sigma_0^2$$

مثــال:

ينتج مصنع صايدال على سبيل المثال فقط دواءا 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 ووجدو أن الانحراف المعياري لهذه العينة يساوي 1.45 مغ

أراد مسؤولو مصلحة المراقبة والجودة اختبار الفرضية $H_0 : \sigma^2 = 1.65$ مقابل الفرضية $H_0 : \sigma^2 \succ 1.65$ عند مستوى معنوية $\alpha = 0.05$

الحـل:

دالة الاختبار هي: $\dfrac{n.S^2}{\chi^2[1-\alpha ، n ، -1]} = \dfrac{25\times1.45^2}{\chi^2[0.95 \ \ 25-1]} = \dfrac{25\times1.45}{36.415} = 1.4434$

ولما كانت: $\left(\sigma_0^2 = 1.65\right) \prec 1.4434$، 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 مبدئيا وهو $\sigma^2 = 1.65$.

2-4 إختبـارات تتعلـق بمقارنـة تباينيـن .

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 إعتمادا 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

هناك فرضية إضافية تتعلق بهذا الإختبار هي أن المتغير العشوائي يخضع في كلا المجتمعين للتوزيع الطبيعي، كذلك فعوضا عن تشكيل الفرق $d = S_1^2 - S_2^2$ و اختبار فيما إذا كان d يختلف جوهريا 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.

نظريـة 10:

إذا كانت $x_1, x_2 ...... \quad x_{n1}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_1, \sigma^2{}_1\right)$، وكانت $y_1, y_2 ...... \quad y_{n2}$ تمثل قياسات عينة عشوائية مسحوبة من توزيع طبيعي $N\left(\mu_2, \sigma^2{}_2\right)$، 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:

$$F\left[\frac{\alpha}{2};n_2-1,n_1-1\right]\le\frac{S_2^2}{S_1^2}\le F\left[1-\frac{\alpha}{2};n_2-1,n_1-1\right]$$

* وإذا كانت $H_1:\sigma_2^2\succ\sigma_1^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{S_2^2}{S_1^2}\succ F\left[1-\alpha;n_2-1,n_1-1\right]$$

* وكانت $H_1:\sigma_2^2\prec\sigma_1^2$ ، فإننا نرفض H_0 على مستوى الدلالة α إذا كانت:

$$\frac{S_2^2}{S_1^2}\prec F\left[\alpha;n_2-1,n_1-1\right]$$

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 $\nu_2=n_2-1$

مثـــال:

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 $H_1:\sigma_1^2\succ\sigma_2^2$ عند مستوى معنوية $\alpha=0.05$

الحـــل:

دالة الاختبار هي: $\frac{S_2^2}{S_1^2}=\frac{68}{36}=1.888$

$$F\left[1-\alpha;n_2-1,n_1-1\right]=F\left[1-0.05;13-1,9-1\right]=F\left[0.95;12,8\right]=3.28$$

ولما كانت: $\left(F\left[0.95;12,8\right]=3.28\right)\prec 1.888$ ، فإننا نقبل $H_0:\sigma_1^2=\sigma_2^2$ على مستوى الدلالة $\alpha=0.05$ ، وهذا معناه أن الطلاب الذين تلقو الدروس الخصوصية المقدمة لم تحسن كما ينبغي النتائج و المستوى بالمقارنة مع الطلاب الآخرين الذين لم يتلقو دروسا خصوصية.

قائمة المراجع .:

1- ابراهيم محمد العلي، و فتاة صبوح. (**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 عبدالحميد 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 حليمي. (**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 شيخي. (**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.

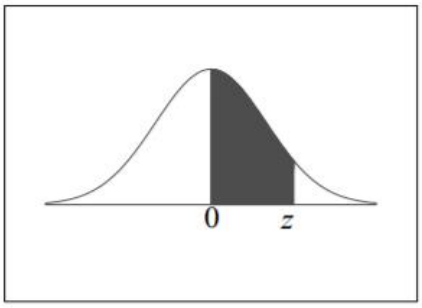
13- محمد علي الأطرقجـي. (**1980**). *الوسائل التطبيقية في الطرق الاحصائية*. بيروت، لبنان: دار الطليعة للطباعة والنشر.

14-محمود حسن المشهدانـي، و أمير حنا هرمـز. (**1989**). *الإحصاء*. بغداد: بيت الحكمة.

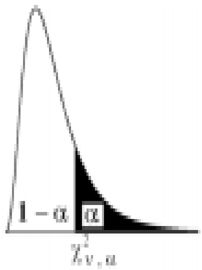
15-منصور عـوض، و صبري عـزام. (**2000**). *مبادئ الاحصاء*. عمان، الأردن: دار صفاء للنشر والتوزيع.

16- Nadjia El-saadi .(1997) .*Travaux diriges de stat-maths* .ecole superieue de banque -16

Standard Normal Distribution Table



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998



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	0.002	0.001
1 Tailed	0.20	0.15	0.10	0.075	0.05	0.025	0.01	0.005	0.001	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	26.30
3	0.978	1.250	1.637	1.924	2.352	3.179	4.525	5.797	9.937	12.39
4	0.941	1.190	1.533	1.778	2.132	2.776	3.744	4.596	7.115	8.499
5	0.919	1.156	1.476	1.699	2.015	2.570	3.365	4.030	5.876	6.835
6	0.906	1.134	1.440	1.650	1.943	2.447	3.143	3.707	5.201	5.946
7	0.896	1.119	1.415	1.617	1.895	2.365	2.999	3.500	4.783	5.403
8	0.889	1.108	1.397	1.592	1.860	2.306	2.897	3.356	4.500	5.039
9	0.883	1.100	1.383	1.574	1.833	2.262	2.822	3.250	4.297	4.780
10	0.879	1.093	1.372	1.559	1.813	2.228	2.764	3.170	4.144	4.586
11	0.875	1.088	1.363	1.548	1.796	2.201	2.719	3.106	4.025	4.437
12	0.873	1.083	1.356	1.538	1.782	2.179	2.682	3.055	3.930	4.318
13	0.870	1.079	1.350	1.530	1.771	2.160	2.651	3.013	3.852	4.221
14	0.868	1.076	1.345	1.523	1.761	2.145	2.625	2.977	3.788	4.141
15	0.866	1.074	1.341	1.517	1.753	2.131	2.603	2.947	3.733	4.073
16	0.865	1.071	1.337	1.512	1.746	2.120	2.584	2.921	3.687	4.015
17	0.863	1.069	1.333	1.508	1.740	2.110	2.567	2.899	3.646	3.965
18	0.862	1.067	1.330	1.504	1.734	2.101	2.553	2.879	3.611	3.922
19	0.861	1.066	1.328	1.500	1.729	2.093	2.540	2.861	3.580	3.884
20	0.860	1.064	1.325	1.497	1.725	2.086	2.529	2.846	3.552	3.850
21	0.859	1.063	1.323	1.494	1.721	2.080	2.518	2.832	3.528	3.820
22	0.858	1.061	1.321	1.492	1.717	2.074	2.509	2.819	3.505	3.792
23	0.857	1.060	1.319	1.489	1.714	2.069	2.500	2.808	3.485	3.768
24	0.857	1.059	1.318	1.487	1.711	2.064	2.493	2.797	3.467	3.746
25	0.856	1.058	1.316	1.485	1.708	2.060	2.486	2.788	3.451	3.725
26	0.856	1.058	1.315	1.483	1.706	2.056	2.479	2.779	3.435	3.707
27	0.855	1.057	1.314	1.482	1.703	2.052	2.473	2.771	3.421	3.690
28	0.855	1.056	1.313	1.480	1.701	2.048	2.468	2.764	3.409	3.674
29	0.854	1.055	1.311	1.479	1.699	2.045	2.463	2.757	3.397	3.660
30	0.854	1.055	1.310	1.477	1.697	2.042	2.458	2.750	3.386	3.646
40	0.851	1.050	1.303	1.468	1.684	2.021	2.424	2.705	3.307	3.551
50	0.849	1.047	1.299	1.462	1.676	2.009	2.404	2.678	3.262	3.496
60	0.848	1.045	1.296	1.458	1.671	2.000	2.391	2.661	3.232	3.460
70	0.847	1.044	1.294	1.456	1.667	1.994	2.381	2.648	3.211	3.435
80	0.846	1.043	1.292	1.453	1.664	1.990	2.374	2.639	3.196	3.417
90	0.846	1.042	1.291	1.452	1.662	1.987	2.369	2.632	3.184	3.402
100	0.845	1.042	1.290	1.451	1.660	1.984	2.365	2.626	3.174	3.391
∞	0.842	1.036	1.282	1.440	1.645	1.960	2.327	2.576	3.091	3.291