

جامعة الزائر 3

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

مطبوعة مقدمة تحت عنوان:

محاضرات في الإحصاء 3
مدعمة بأمثلة و تمارين محلولة
موجهة لطلبة السنة الثانية جذع مشترك ليسانس

إعداد :

الدكتورة : سموك نوال

السنة الجامعية 2019 - 2020

فهرس المحتويات

الصفحة	العنوان
III	فهرس المحتويات
VIII	فهرس الجداول
IX	فهرس الأشكال
X	مقدمة
الفصل الأول: اختيار العينات	
12	تمهيد
13	1- مصطلحات ومفاهيم
14	2- أساليب المعاينة
14	1-2. أساليب المعاينة العشوائية
15	1-1-2. العينة العشوائية البسيطة
16	2-1-2. العينة العشوائية الطبقية
17	3-1-2. العينة العشوائية المنتظمة
18	4-1-2. العينة العشوائية العنقودية
19	2-2. أساليب المعاينة غير العشوائية
19	1-2-2. أسلوب الاختيار عن طريق الصدفة
20	2-2-2. أسلوب المعاينة القصدية
20	3-2-2. أسلوب المعاينة الحصصية
20	3- تحديد حجم العينة وأساليب توزيعه على الطبقات

20	3-1. اتجاهات تحديد حجم العينة
22	3-2. أساليب توزيع حجم العينة على الطبقات
24	4- سلسلة تمارين غير محلولة
	الفصل الثاني: توزيع المعاينة
27	تمهيد
28	1- مصطلحات ومفاهيم
28	2- توزيع المعاينة للمتوسط الحسابي
34	3- نوع توزيع المعاينة للمتوسط
34	3-1. نوع توزيع المعاينة للمتوسط عندما يكون للمجتمع توزيع طبيعي
34	3-2. نوع توزيع المعاينة لمتوسط العينة عندما يكون المجتمع له توزيع غير طبيعي
36	4- توزيع المعاينة للنسبة
36	4-1. المتوسط والخطأ المعياري للنسبة
37	4-2. نوع توزيع المعاينة للنسبة P في العينة
38	5- توزيع المعاينة للتباين
40	6- توزيع المعاينة للفروق والمجاميع
40	6-1. الفروق والمجاميع للمتوسط الحسابي والتباين والنسبة لتوزيع المعاينة
41	6-2. طبيعة توزيع المعاينة للمجموع أو الفرق بين متوسطين
42	7- توزيع المعاينة للنسبة بين تباينين
42	7-1. توزيع فيشر
43	7-2. خواص توزيع فيشر

45	8- سلسلة تمارين محلولة
49	9- سلسلة تمارين غير محلولة
	الفصل الثالث: نظرية التقدير
53	تمهيد
54	1- مصطلحات ومفاهيم
54	2- شروط التقدير الجيد
54	1-2. عدم التحيز
54	2-2. الاتساق
55	2-3. الكفاءة
55	2-4. الكفاية
55	3- طرق التقدير
55	1-3. التقدير النقطي
56	2-3. التقدير بفترة
57	1-2-3. تفسير فترة الثقة
59	2-2-3. فترة الثقة لمتوسط مجتمع طبيعي تباينه معلوم
60	3-2-3. فترة الثقة للمتوسط في حالة المجتمعات الكبيرة
61	4-2-3. فترة الثقة للمتوسط في العينات الكبيرة من مجتمعات مجهولة التباين
62	5-2-3. فترة الثقة للمتوسط في العينات الصغيرة من مجتمعات مجهولة التباين
62	1-5-2-3. توزيع ستيودنت
63	2-5-2-3. خواص توزيع ستيودنت

66	3-3. تقدير النسبة
66	1-3-3. تقدير النسبة بنقطة
67	2-3-3. تقدير النسبة بفترة
68	3-3-3. تحديد حجم العينة اللازم لتقدير فترة النسبة P
70	4-3. فترات الثقة حول تباين المجتمع
70	1-4-3. توزيع كاي تربيع
71	2-4-3. خواص توزيع كاي تربيع
73	5-3. فترات الثقة للنسبة بين تباينين
75	4- سلسلة تمارين محلولة
78	5- سلسلة تمارين غير محلولة
الفصل الرابع: اختبار الفرضيات	
82	تمهيد
83	1- مصطلحات ومفاهيم
83	2- عناصر الاختبار الإحصائي
83	1-2. الفرضية الصفرية (فرضية العدم)
83	2-2. الفرضية البديلة
84	3-2. إحصاء الاختبار
84	4-2. منطقة اتخاذ القرار
85	5-2. الخطأ من النوع الأول والخطأ من النوع الثاني
86	1-5-2. الخطأ من النوع الأول
86	2-5-2. الخطأ من النوع الثاني

86	2-6. قوة إحصاء الاختبار
87	3- خطوات اختبار الفرضيات
88	3-1. صياغة الفرضية الصفرية والفرضية البديلة
88	3-2. تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)
88	3-3. تحديد قاعدة القرار
88	3-4. حساب القيمة الجدولية والفعلية لإحصاء الاختبار
88	3-5. المقارنة واتخاذ القرار
88	4- اختبار الفرضيات المتعلقة بالمتوسط
89	4-1. الاختبار ثنائي الاتجاه للمتوسط
90	4-2. الاختبار أحادي الاتجاه من اليمين للمتوسط
92	4-3. الاختبار أحادي الاتجاه من اليسار للمتوسط
94	5- اختبار الفرضيات المتعلقة بالنسبة
99	6- اختبار الفرضيات المتعلقة بتباين المجتمع
102	7- سلسلة تمارين محلولة
105	8- سلسلة تمارين غير محلولة
109	قائمة المراجع

فهرس الجداول

الصفحة	عنوان الجدول	رقم الجدول
40	توزيع الفروق والمجاميع للأوساط الحسابية والتباينات ونسب المعاينة	1-2
59	أهم مستويات الثقة ومعاملاتها في حالة استخدام التوزيع الطبيعي	1-3
85	منطقتا الرفض والقبول للفرضية الصفرية في الحالات الثلاث	1-4
86	الخطأ من النوع الأول والخطأ من النوع الثاني	2-4
94	قاعدة القرار اختبار النسبة في الحالات الثلاث	3-4
100	قاعدة القرار لاختبار النسبة في الحالات الثلاث	4-4

فهرس الأشكال

الصفحة	عنوان الشكل	رقم الشكل
43	منحنى توزيع فيشر	1-2
58	مساحة فترة الثقة وحديها عند مستوى معنوية $1 - \alpha$	1-3
63	منحنى تقريبي توزيع r إلى التوزيع الطبيعي	2-3
70	منحنى كاي تربيع	3-3
89	منطقة الرفض في حالة الاختبار ثنائي الاتجاه عند مستوى معنوية 0.05	1-4
91	منطقة الرفض في حالة الاختبار أحادي الاتجاه من اليمين عند مستوى معنوية 0.05	2-4
93	منطقة الرفض في حالة الاختبار أحادي الاتجاه من اليسار عند مستوى معنوية 0.01	3-4
95	منطقة الرفض في حالة الاختبار أحادي الاتجاه من اليمين عند مستوى معنوية 0.05	4-4
97	منطقة الرفض في حالة اختبار أحادي الاتجاه من اليسار عند مستوى معنوية 0.1	5-4
98	منطقة الرفض في حالة اختبار النسبة ثنائي الاتجاه عند مستوى معنوية 0.05	6-4
101	اختبار التباين أحادي الاتجاه من اليمين عند مستوى معنوية 0.05	7-4

مقدمة:

عزيزي الطالب :

بعد دراستك لأهم مقاييس النزعة المركزية ومقاييس التشتت في مقياس الإحصاء 1 أو ما يعبر عنه بالإحصاء الوصفي خلال السداسي الأول أين كنت طالبا بالسنة الأولى لتتناول بعدها المبادئ الأساسية المتعلقة بالاحتمالات وبعض توزيعاتها في مقياس الإحصاء 2 أو ما يعبر عنه بالإحصاء الرياضي في السداسي الثاني من نفس السنة، أضع بين يديك هذه المطبوعة التي هي عبارة عن مجموعة من الدروس والتمارين في الإحصاء 3 أو ما يسمى بالإحصاء الاستدلالي والموجهة إلى طلبة السنة الثانية جذع مشترك بكلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير بالدرجة الأولى، وإلى كل التخصصات التي تتناول هذا المقياس في مقرراتها.

وقد حرصت في تقديمها على بساطة العرض، بأسلوب سهل ومختصر يعتمد على أمثلة وتمارين محلولة، بعيدا عن البراهين والعمليات الرياضية المعقدة، وذلك بوضع العلاقة ذات الصلة وشروط استخدامها وكيفية تطبيقها.

ولقد تضمنت هذه المطبوعة أربعة فصول، خصص الفصل الأول للعينات واختيارها حيث تم تناول مفهومها وأنواعها وطرق اختيارها، أما الفصل الثاني فقد شمل موضوع توزيع المعاينة لكل من المتوسط والنسبة والتباين والفروق والمجاميع، أما الفصل الثالث فقد تضمن موضوع نظرية التقدير، ليتم تخصيص الفصل الرابع لاختبار الفرضيات.

وإذ أقوم بتقديم هذا العمل لطلبتنا في الجذع المشترك من السنة الثانية ولبعض التخصصات المتواجدة ضمن مقرراتها، فإنني أهيب بزملائي الأساتذة وطلبتي الأعزاء أن لا يبخلوا علي بنصائحهم من خلال تنبيهنا لأي خطأ تخلل هذا العمل، أو ملاحظات أو إضافات من شأنها المساهمة في الارتقاء به نحو الأحسن. وفي الأخير آمل أن أكون قد وفقت في تقديم هذا العمل، وأن يساهم في إثراء مكتباتنا الجامعية، وأن يكون مفيدا لطلبتنا الأعزاء، والله من وراء القصد وهو وليّ كل توفيق.

د/ سموك نوال

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

جامعة الجزائر 3

الفصل الأول

اختيار العينات

- ✓ مصطلحات ومفاهيم
- ✓ أساليب المعاينة العشوائية
- ✓ أساليب المعاينة غير العشوائية
- ✓ تحديد حجم العينة وأساليب توزيعه على الطبقات
- ✓ سلسلة تمارين غير محلولة

تمهيد:

يتم جمع البيانات الإحصائية المتعلقة بمشكلة أو ظاهرة ما من طرف الباحث وفق أسلوبين، يتمثل الأسلوب الأول في الحصر الشامل وذلك بالتوجه لجميع مفردات المجتمع الإحصائي دون استثناء بغية الحصول على بيانات شاملة، حيث يطلق على هذا النوع بأسلوب المسح الشامل والذي يتطلب تكاليف مرتفعة ويستلزم وقتاً وجهداً فائقاً، كما أنه يحتاج إلى كوادرن فنية متخصصة وذات الخبرة والجهد والتكلفة التي تصاحب عملية تحليل البيانات بعد جمعها، ولهذا تتجه معظم البحوث الإحصائية إلى أسلوب آخر يدعى أسلوب العينة لعدة اعتبارات أهمها⁽¹⁾:

- ✓ يوفر أسلوب العينة للباحثين الجهد والتكلفة.
- ✓ توفير الوقت، فغالبا ما يكون الباحث مجبرا على جمع البيانات خلال فترة محددة.
- ✓ إمكانية الحصول على بيانات ومعلومات وفيرة لاقتصار الباحث على مجموعة جزئية من المجتمع.
- ✓ سهولة المتابعة في الحصول على ردود وافية ومتكاملة ودقيقة من خلال متابعة العينة بحجمها الصغير الممثل لأفراد المجتمع ومتابعة أجوبتها وردودها.
- ✓ سيطرة الباحث على حجم محدد من أفراد المجتمع المتمثل بالعينة مما يؤدي إلى سيطرته على البيانات والمعلومات المجمعة وبالتالي الدقة في التعامل مع البيانات وتجميعها.
- ✓ صعوبة واستحالة الوصول إلى كل وحدات المجتمع في بعض الأحيان.

(1) عامر إبراهيم قنديلجي، منهجية البحث العلمي، دار اليازوري، عمان، الأردن، 2012، ص ص 188 - 190.

1. مصطلحات ومفاهيم:

الإحصاء علم يهتم بالمعلومات والبيانات ويهدف إلى تجميعها وتبويبها وتنظيمها وتحليلها واستخلاص النتائج منها بل وتعميم نتائجها واستخدامها في اتخاذ القرارات.

لقد أدى التقدم المذهل في تكنولوجيا المعلومات واستخدام الحاسبات الآلية إلى مساعدة الدارسين والباحثين ومتخذي القرارات في الوصول إلى درجات عالية ومستويات متقدمة من التحليل ووصف الواقع ومتابعته ثم إلى التنبؤ بالمستقبل.

ولم تعد بالتالي البحوث الاقتصادية والاجتماعية والإدارية وغيرها في وقتنا المعاصر، وفي ظل التقدم التكنولوجي الهائل في كافة ميادين حياتنا اليومية تكفي بمجرد عرض المشاكل ودراسة الظواهر وتحديد الأسباب واستخلاص النتائج واتخاذ القرارات بطريقة سطحية مجردة عن أسلوب الإقناع والتقدير والقياس.

لقد أصبح الاتجاه العام في مثل هذه البحوث والدراسات هو استخدام طرق القياس الكمية ووسائل الإقناع الإحصائية وذلك لتحديد الخصائص وإبراز الاتجاهات العامة في الظواهر الاجتماعية والإدارية، وتحليل العلاقات المتشابكة والمتبادلة بين الظواهر على أساس موضوع غير متميز.

وعلم الإحصاء يعطي للباحثين في مجال العلوم الاقتصادية والاجتماعية والإدارية، العديد من الطرق والأساليب اللازمة لضرورة القيام بالدراسات والبحوث الاقتصادية والاجتماعية والإدارية والجغرافية على أساس من القياس لحركة العديد من المتغيرات المحددة للظواهر موضوع الدراسة.

وتستخدم كلمة الإحصاء لتشير إلى عملية جمع البيانات الكمية والأساليب المستعملة في معالجة تلك البيانات، وقد نعني بهذه الكلمة أيضا عملية استخلاص بعض الاستنتاجات من دراسة عينة صغيرة لصياغة تعميمات يمكن تطبيقها على مجتمعات أكبر حجما.

فبحوث الرأي العام على سبيل المثال تقوم على مقابلة ودراسة عينة صغيرة من أفراد المجتمع ولكن نتائجها تستخدم في الاستدلال على اتجاهات الرأي العام في المجتمع ككل، وبذلك يمكن القول بان الإحصاء يشير إلي طرق تنظيم وتلخيص البيانات وإلى الأساليب التي تستخدم في تحليل وتفسير النتائج واستخلاصها والتي يمكن تعميمها على مجتمع الدراسة.

فالإحصاء هو علم يبحث في طريق جمع الحقائق الخاصة بالظواهر العلمية الاجتماعية التي تتمثل في حالات أو مشاهدات متعددة، وفي كيفية تسجيل هذه الحقائق في صورة قياسية رقمية، وتلخيصها بطريقة يسهل بها معرفة اتجاهات الظواهر وعلاقات بعضها ببعض، ويبحث أيضاً في دراسة هذه العلاقات والاتجاهات واستخدامها في فهم حقيقة الظواهر ومعرفة القوانين التي تسير تبعاً لها.

ومن هنا يتضح أن الإحصاء لا غنى عنه لأي باحث في شتى المجالات المختلفة إذ اعتمد في بحثه على الأسلوب العلمي، أي أن الإحصاء هو عصا الباحث التي تقوده إلى الطريق الصحيح، وهي الأداة التي

تساعده على تفسير الظواهر التي يدرسها وتوضيح النتائج التي يحصل عليها ودلالات البيانات والأرقام التي يحصل عليها.

- ◀ **المجتمع:** هو مجموعة عناصر لها خاصية أو عدة خصائص مشتركة تميزها عن غيرها من العناصر الأخرى، وقد تكون منتهية أو غير منتهية.
- ◀ **العينة:** هي مجموعة جزئية مأخوذة من المجتمع تشترك في خاصية أو خصائص معينة ويشترط أن تكون ممثلة له.
- ◀ **المعاينة:** هي الكيفية أو العمليات التي تسمح بانتقاء مجموعة فرعية من المجتمع.
- ◀ **العشوائية:** أن تكون لكل مفردات العينة نفس فرصة الظهور أو الاختيار، كما لا تكون على علم مسبق بالنتائج المتوصل إليها.

2. أساليب المعاينة:

توجد عدة طرق وأساليب يتم من خلالها اختيار العينة المراد دراستها والتي يمكن أن تنتظم في مجموعتين رئيسيتين تتمثلان في:

- المعاينة العشوائية (الاحتمالية).
- المعاينة غير العشوائية (غير الاحتمالية).

2-1. أساليب المعاينة العشوائية:

يتوفر هذا الأسلوب على العديد من المزايا تلخص أهمها فيما يلي⁽¹⁾:

- حذف التحيز:

هذا الأسلوب يؤكد أن مفردات العينة تم اختيارها بدون تحيز، ولكن إذا تم اختيار مفردات العينة تحكيمياً، فهناك دائماً إمكانية لوجود تحيز، وعلى الرغم من أن الاختيار العشوائي لا يضمن بأن تكون العينة ممثلة للمجتمع، إلا أنه يحذف مخاطر الاختيار المتحيز.

- تحديد الثقة:

أسلوب العينات العشوائية يضع الأساس الإحصائي لتحديد الثقة المقترنة بالاستنتاج الإحصائي، والاستنتاج الإحصائي لا يمكن تنفيذه إذا كان اختيار مفردات العينة بطريقة أخرى خلاف الطريقة العشوائية.

- التحكم في خطأ المعاينة:

أسلوب العينات العشوائية يسمح بالتحكم في خطأ المعاينة من خلال اختيار حجم العينة، ولكن مع الطرق الغير عشوائية في اختيار العينة فإنه لا يمكن أن يتحقق مستوى مقبول من خطأ المعاينة.

(1) جورج كانافوس، دون ميلر، ترجمة سلطان محمد عبد المجيد، محمد توفيق البلقيني، الإحصاء للتجاربيين، مدخل حديث، دار المريخ، الرياض، المملكة العربية السعودية، 2004، ص ص 33 - 34.

وتتمثل أساليب المعاينة العشوائية فيما يلي:

2-1-1. العينة العشوائية البسيطة:

حيث تؤخذ العينة بشكل يعطي لأي عنصر من عناصر المجتمع نفس الفرصة لأن يكون ضمن العينة، وتستخدم في حالة المجتمعات المتجانسة والمحدودة، ويتم اختيار أفراد العينة بكتابة أرقام أفراد المجتمع على بطاقات متشابهة وخطها بشكل جيد ثم تتم عملية السحب عن طريق القرعة، أو باستخدام جداول الأرقام العشوائية أو باستخدام الحاسب الآلي...

ويتم استخدام جداول الأرقام العشوائية كما يلي:

نعطي كل عناصر المجتمع الإحصائي أرقاماً متسلسلة من 1 إلى N، حيث N هو عدد وحدات المجتمع، ويكون لكل عنصر من عناصر المجتمع نفس منزلة حجم المجتمع أي يكون له نفس عدد الأرقام، فإذا كان حجم المجتمع 56 مثلاً، فهذا العدد يتكون من رقمين، فإن كل وحدة من وحدات المجتمع الإحصائي تعطي عدداً يتكون من رقمين بدءاً من 01 إلى غاية 56.

وإذا كان حجم المجتمع هو 325 مثلاً فإن كل وحدة من وحدات المجتمع تأخذ عدداً مكوناً من ثلاثة

أرقام كما يلي: 001، 002، 003، ...، 010، 011، ...، 099، 100، 101، ...، 323، 324، 325.

نأخذ صفحة من صفحات الجداول العشوائية ونختار عموداً أو سطراً بحيث يكون عدد أرقامه يساوي عدد أرقام حجم المجتمع، ونأخذ كل الأرقام المحصورة ضمن المجال ونلغي البقية، وفي حالة نفاذ أعداد العمود أو السطر دون الانتهاء من حجم العينة المفترض فإننا ننتقل إلى العمود أو السطر الموالي، وهكذا إلى غاية حصر العدد المطلوب من حجم العينة.

مثال 1:

نريد اختيار عينة عشوائية تتكون من 10 طالبة من الجذع المشترك من قسم علوم التسيير المتكون من 300 طالبا بطريقة العينة العشوائية البسيطة، وباستخدام جداول الأرقام العشوائية. بين كيف يتم ذلك؟

الحل :

1- نحدد قائمة أسماء هؤلاء الطلبة.

2- نعطي لكل طالب عدداً من الأعداد التالية: 001، 002، 003، ...، 010، 011، ...، 099، 100، 101، ...، 298، 299، 300.

3- من الجدول الأول نختار السطر الخامس، فنجد أن العدد الأول هو 739، وهو مرفوض لأنه يقع ضمن المجال المختار (أكبر من 299).

العدد الذي يليه: 440 مرفوض أيضا، وهكذا إلى غاية العدد الخامس 032 نقبله، ونختار الطالب ذا العدد 32، نستمر في العملية لتكون مفردات العينة هم الطلبة ذوا الأعداد التالية: 32، 148، 238، 24، 56، 128، 104، 271، 180، 151.

2-1-2. العينة العشوائية الطبقية:

وتستخدم في حالة المجتمعات المحدودة وغير المتجانسة، حيث يقسم المجتمع إلى مجموعات متجانسة تسمى طبقات ويتم اختيار عينة عشوائية بسيطة من كل طبقة (جامعة ← كليات)، (مجتمع ← طبقات (موظفين، عمال، فلاحين)).

فهذه الطريقة تستخدم في الحالات التي تكون فيها نتائج البحث تعتمد على المتغيرات المفسرة مثلا، العمر، الجنس، الخبرة، الدخل، طبيعة النشاط، حجم المؤسسة. ففي هذه الحالة يتم تقسيم المجتمع الإحصائي إلى مجموعات جزئية تعتمد على هذه الصفات وتسمى بالطبقات، ثم باستخدام طريقة العينة العشوائية يتم اختيار عينة جزئية يتناسب حجمها مع حجم الطبقة، وتشكل مجموعات العينات الجزئية المختارة ما يسمى بالعينة الطبقية. ويتم اختيار عدد مفردات كل طبقة حسب العلاقة:

$$\text{عدد مفردات كل طبقة} = \frac{\text{حجم الطبقة}}{\text{حجم المجتمع}} \times \text{حجم العينة}^{(1)}$$

مثال 2:

بالعودة إلى معطيات المثال السابق نريد اختيار 30 طالبا على أن تكون من الطلبة والطبات، علما أن عدد الطلبة الذكور هو 200، وعدد الطالبات الإناث هو 100 طالبة. ما هو عدد الطلبة والطبات التي تتشكل منها العينة؟

الحل:

$$\begin{aligned} \text{عدد الطلبة الذكور} &= 30 \times \frac{200}{300} = 20 \text{ طالبا} \\ \text{عدد الطالبات الإناث} &= 30 \times \frac{100}{300} = 10 \text{ طالبات} \end{aligned}$$

وبالتالي تتكون العينة المطلوبة من 20 طالبا و 10 طالبات، وحجمها هو 30.

ملاحظة:

قد تكون الطبقات المكونة معبرا عنها بالنسب المئوية لكل منها، في هذه الحالة يتم حساب عدد مفردات كل طبقة بالاعتماد على تلك النسب.

(1) فايز جمعة صالح النجار وآخرون، أساليب البحث العلمي، منظور تطبيقي، دار الحامد، عمان، الأردن، 2009، ص 96 بتصريف.

مثال 3:

في مثالنا السابق نفترض أن نسبة الطلبة الذكور هي 60% ونسبة الإناث هي 40%.
- فما هو عدد الطلبة الذكور والطالبات في عينة حجمها 30؟

الحل:

$$\begin{aligned} \text{عدد الطلبة الذكور} &= 30 \times \frac{60}{100} = 18 \text{ طالبا} \\ \text{عدد الطالبات الإناث} &= 30 \times \frac{40}{100} = 12 \text{ طالبة} \end{aligned}$$

نلاحظ أن: $30 = 12 + 18$.

مثال 4:

إذا أردنا اختيار عينة من 20 طالبا من قسمين من أقسام كلية العلوم الاقتصادية والتجارية وعلوم التسيير التي تتألف من أربعة أقسام، وكان 30% من الطلبة من قسم علوم التسيير، 25% من قسم العلوم المالية والمحاسبية، 25% من قسم العلوم الاقتصادية، 20% من قسم العلوم التجارية.

الحل:

$$\begin{aligned} - \text{ عدد الطلبة من قسم علوم التسيير} &= 20 \times \frac{30}{100} = 6 \text{ طالبة} \\ - \text{ عدد الطلبة من قسم العلوم المالية والمحاسبية} &= 20 \times \frac{25}{100} = 5 \text{ طالبة} \\ - \text{ عدد الطلبة من قسم العلوم الاقتصادية} &= 20 \times \frac{25}{100} = 5 \text{ طالبة} \\ - \text{ عدد الطلبة من قسم العلوم التجارية} &= 20 \times \frac{20}{100} = 4 \text{ طالبة} \end{aligned}$$

نلاحظ أن حجم العينة المطلوب هو: $20 = 4 + 5 + 5 + 6$ طالبا.

3-1-2. العينة العشوائية المنتظمة:

يتم اختيارها من خلال تحديد مجتمع الدراسة ووضع أفرادها في قائمة بشكل عشوائي وإعطاء كل منهم رقما، ثم يتم تحديد قاعدة الاختيار وفق قسمة حجم المجتمع على حجم العينة من أجل الحصول على طول الفترة، وبعد ذلك يتم انتقاء أحد الأرقام عشوائيا من بين الأرقام التي تساوي أو أقل من طول الفترة، ليتم اعتباره كعنصر أول من مفردات العينة ويشعر في إضافة طول الفترة له للحصول على المفردة الثانية، وهكذا نستمر في إضافة العدد الثابت إلى غاية الوصول إلى العدد الممثل لحجم العينة المطلوب. وبهذه الطريقة نكون قد حصلنا على متتالية حسابية حدها الأول هو الرقم المختار عشوائيا في البداية، وأساسها هو طول الفترة، وعدد حدودها هو حجم العينة⁽¹⁾.

(1) فايز جمعة صالح النجار وآخرون، مرجع سابق، ص 95.

مثال 5:

نريد اختيار عينة عشوائية حجمها 10 وفق طريقة العينة العشوائية المنتظمة من المجتمع الإحصائي السابق ذي الحجم 300. كيف يتم ذلك؟

الحل:

1 - باعتبار أن القائمة مرتبة عشوائيا، فإن طول الفترة يتم حسابه كما يلي:

$$\text{طول الفترة} = \frac{300}{100} = 30$$

2 - يتم اختيار الرقم الأول عشوائيا على أن يكون أقل من أو يساوي 30، وليكن العدد الأول من الصف الأول من جدول الأرقام العشوائية وهو 11، أي الطالب ذو الرقم 11.

3 - يشرع في إضافة 30 في كل مرة ليتم الحصول على الطلبة الذين تتألف منهم هذه العينة كما يلي: 11، 41، 71، 101، 131، 161، 191، 221، 251، 281.

2-1-4. العينة العشوائية العنقودية (المتعددة المراحل):

تستخدم في حالة المجتمعات الكبيرة أو لما تكون مفرداتها متباعدة جغرافيا، حيث يقسم المجتمع إلى مجموعات وتختار من هذه المجموعات عينة عشوائية بسيطة ثم نأخذ جميع الأفراد في المجموعات المختارة فتسمى عينة عشوائية عنقودية ذات مرحلة واحدة، أما إذا اخترنا عينة عشوائية بسيطة من الأفراد من كل مجموعة مختارة فتسمى عينة عشوائية عنقودية ذات مرحلتين وهكذا.

أما اختيار مفردات العينة المراد دراستها فإنه يتم بنفس الطريقة المنتهجة في أسلوب العينة العشوائية البسيطة في حالة المجموعة الواحدة، أو العينة العشوائية الطباقية في حالة مجموعتين فأكثر⁽¹⁾.

مثال 6:

في دراسة بعنوان تطبيقات تكنولوجيا المعلومات في إدارة الجماعات المحلية الجزائرية، نريد اختيار عينة عشوائية عنقودية من بلدية ما من بلديات الوطن الجزائري تتكوّن من 100 موظف إداري.
- بين كيف يتم ذلك؟

الحل:

أولا: يتم تقسيم المجتمع الإحصائي الممثل في بلديات الوطن الجزائري إلى مجموعات حسب الجهات، أي، وسط، غرب، شرق، جنوب.
في المرحلة الأولى يتم اختيار جهة من بين الجهات الأربع بشكل عشوائي، ولنفرض أننا حصلنا عن طريق القرعة على منطقة الشرق الجزائري.

(1) محمد عبد الفتاح الصرفي، الدليل التطبيقي للباحثين، دار وائل، عمان، الأردن، 2002، ص 201.

في المرحلة الثانية وبعد حصر ولايات الشرق الجزائري، يتم اختيار ولاية من بين ولايات الشرق الجزائري وبشكل عشوائي، ولنفرض أننا حصلنا على ولاية المسيلة.

في المرحلة الثالثة وبعد تحديد قائمة بلديات ولاية المسيلة، يتم اختيار بلدية من بينها بشكل عشوائي، ليتم تعيين البلدية ميدان الدراسة، وفي الأخير يتم اختيار 100 موظف من موظفيها على أن يكون الاختيار أيضا بشكل عشوائي.

والمخطط الموالي يلخص ذلك.

(الجزائر ← (وسط، غرب، شرق، جنوب) ← مر 1 (شرق) ← مر 2 (ولايات) ← مر 3 (بلديات)

2-2. أساليب المعاينة غير العشوائية (غير الاحتمالية):

رغم أهمية الأساليب العشوائية للمعاينة وتبني العديد من الباحثين لاستخدامها، فإنه توجد بعض البحوث التي تتجه إلى اختيار أساليب غير عشوائية خاصة في البحوث النوعية، حيث يسعون إلى التركيز على جوانب أخرى مهمة في عينات البحث النوعي مثل⁽¹⁾:

- غزارة البيانات والمعلومات المتوفرة عند أفراد العينة.
- قربهم من الأحداث والموضوعات المعنية بالبحث.
- استعدادهم للتعاون وإعطاء المعلومات الوافية.

فأساليب المعاينة غير العشوائية إذن تقوم على مبدأ عدم تحكم الباحث في اختيار مفردات العينة، وعدم معرفة مفردات مجتمع البحث، وهذا ما يؤدي إلى عدم تساوي الفرصة لمفردات مجتمع البحث لتكون ضمن مفردات العينة، وبالتالي عدم تمثيل المجتمع، وتعميم النتائج المتوصل إليها من خلال تلك العينة على المجتمع.

والمعاينة غير العشوائية هي الأخرى تتضمن مجموعة من الأساليب سنأتي على ذكر أهمها وأكثرها استعمالا بإيجاز فيما يلي⁽²⁾:

2-2-1. أسلوب الاختيار عن طريق الصدفة:

حيث لا يخضع اختيار الباحث مفردات العينة لأي اعتبار وإنما يكون على سبيل المصادفة، كأن يشرع الباحث في توزيع الاستبيان على مجموعة من العاملين بمؤسسة ما وذلك بتسليم الاستمارة على من يلتقيه صدفة أثناء زيارته الميدانية، أو كأن يقف أمام مدخل المؤسسة محل الدراسة ويختار الثلاثين الأوائل الذين يلتقي بهم عند دخولهم.

(1) عامر إبراهيم قنديلجي، مرجع سابق، ص 205.

(2) محمد عبد الفتاح الصيرفي، مرجع سابق، ص 211 بتصرف.

2-2-2. أسلوب المعاينة القصدية:

هي عينة يتم اختيار أفرادها بشكل مقصود ومستهدف لتوفر بعض الخصائص فيهم بما يخدم أهداف الدراسة، ويلجأ إلى هذا النوع من العينات عند توفر البيانات اللازمة للدراسة لدى فئة محددة من المجتمع الأصلي للدراسة.

فعلى سبيل المثال: عندما يريد الباحث أن يدرس سلوك المستهلكين لمنتجات مؤسسة ما، فإنه يحدد المستهلكين الحقيقيين لمنتجات تلك المؤسسة ويتجنب جمع البيانات من أشخاص لا يقبلون على انتقاء منتجاتها.

2-2-3. أسلوب المعاينة الحصصية:

سميت بالحصصية أو الحصصية لأن مجتمع البحث يقسم إلى فئات طبقاً لصفاته الرئيسية، وتمثل كل فئة بنسبة وجودها في المجتمع، ويتجلى الفرق بين هذا النوع من المعاينة والمعاينة العشوائية الطبقية في كون أن اختيار مفردات هذه الأخيرة يتم بشكل عشوائي فضلاً على عدم محدودية مفردات المجتمع الأصلي في حالة المعاينة الحصصية.

وعلى العموم فعلى الباحث أن يتوخى الحذر في تفسيره لنتائج العينة عند اللجوء إلى أساليب المعاينة غير العشوائية، كما ينبغي عليه تبرير أسباب اختياره لأي منها. ولا شك في أن نتائج الدراسة المتوصل إليها وفق طرق العينات العشوائية تكون هي الأكثر مصداقية في مختلف الدراسات الإحصائية لخلوها من عنصر التحيز⁽¹⁾.

3. تحديد حجم العينة وأساليب توزيعه على الطبقات:**3-1. اتجاهات تحديد حجم العينة:**

قبل الشروع في عملية اختيار العينة، يحتاج الباحث إلى تحديد حجم العينة المناسب حتى تزوده بالبيانات والمعلومات التي يعتمد عليها في تعميم النتائج على المجتمع كله، وهناك اتجاهان يمكن السير فيهما لتحديد حجم العينة⁽²⁾.

(1) موريس انجرس، ترجمة بوزيد صحراوي وآخرون، منهجية البحث في العلوم الإنسانية، دار القصة، الجزائر، 2004، ص ص 311 - 313 بتصرف.

(2) سعيد التل وآخرون، مناهج البحث العلمي، تصميم البحث والتحليل الإحصائي، دار المسيرة، عمان، الأردن، 2007، ص ص 105 - 106، بتصرف.

أ. الاتجاه الأول:

حيث يعتمد الباحث عند تحديد حجم العينة المطلوب على خبراته السابقة في هذا المجال، أو قد يسترشد الباحث برأي وخبرة الآخرين وهذا الأسلوب في اختيار العينة يفيد الباحثين الذين لا يميلون إلى استخدام الأسلوب الرياضي في اختيار العينة.

ففي الدراسات المسحية تكون من النسبة 20% من أفراد المجتمع الكلي إذا كان عدد أفراد هذا المجتمع معتدلاً (ما بين 500 إلى 1000)، وتقل هذه النسبة كلما كبر حجم المجتمع الأصلي لتصل إلى حوالي 5%. و في الدراسات التجريبية ذات المعالجة الواحدة (متغير مستقل واحد) يكون حجم العينة الواحدة مناسباً إذا زاد عدد أفرادها عن 30 فرداً (لكل مستوى من مستويات هذه المعالجة). أما في الدراسات التجريبية ذات المعالجتين اثنان أو أكثر، فإنه من المستحسن أن لا يقل عدد أفراد الخلية الواحدة في التصميم الإحصائي عن خمسة أفراد.

ب. الاتجاه الثاني:

و يأخذ بعين الاعتبار بعض القواعد الاحتمالية لتحديد حجم العينة، كاحتمال الوقوع في الخطأ من النوع الأول واحتمال حدوث الخطأ من النوع الثاني، بالإضافة إلى بعض الأمور الأخرى المرتبطة بالتكلفة وبعض المقاييس الإحصائية.

وهناك فكرة خاطئة عند البعض بأنه كلما كبر حجم المجتمع يجب أن يزيد حجم العينة المسحوبة منه، وهذا خطأ شائع لأنه في مجتمع متجانس الصفات والخصائص تكفي عينة صغيرة لدراسته.

إن التباين بين أفراد المجتمع، وليس حجم المجتمع هو العامل الحاسم في تقرير حجم العينة، فكلما كبر التباين بين مفردات المجتمع كلما استوجب ذلك أن يكون حجم العينة كبيراً بغض النظر عن حجم ذلك المجتمع.

كلما زاد حجم العينة قل الخطأ المعياري للمعينة، وكلما نقص حجمها زاد الخطأ المعياري للمعينة.

ومن أهم المبادئ التي يجب مراعاتها عند تحديد حجم العينة ما يلي⁽¹⁾:

1. تحديد الهدف من اختيار العينة، وبناء على هذا الهدف تعطى حدود الخطأ المسموح به، ونوع القرار المتوقع اتخاذه بناء على البيانات التي ستوفرها العينة.
2. إعطاء علاقة تربط بين حجم العينة المطلوب ومدى الدقة المطلوبة من العينة (حدود الخطأ المسموح به) وتختلف هذه العلاقة باختلاف طبيعة الإحصائي (Statistic) الذي سيستعمل لتقدير معلمة التوزيع (Parameter). ومثل هذه العلاقة قد تحتوي على بعض معالم التوزيع المجهولة، ومن هنا لا بد أولاً من تقديرها بناء على خبرة سابقة.

(1) سعيد التل وآخرون، مرجع سبق ذكره، ص 106 - 107.

3. في بعض الحالات، مثل العينة الطبقية، تحتاج إلى تقدير حجم العينة في كل طبقة، وبالتالي يكون حجم العينة الكلي هو تلك الحجوم الجزئية.
4. في بعض الحالات قد نقيس أكثر من متغير واحد على نفس الفرد من أفراد العينة، ويجب أن يتم حساب العينة المطلوبة بناء على التقدير المطلوب لكل متغير، مع ملاحظة أن الحجوم الناتجة قد تختلف عن الحسابات المبنية على متغير عن تلك المبنية على متغير آخر.
5. يجب ربط حجم العينة مع التكلفة المتاحة للدراسة والزمن، وما شابه ذلك من عوامل قد تؤثر في حجم المجتمع.
6. معرفة حجم المجتمع الأصلي مسبقاً، وإن لم يكن هذا الحجم معروفاً، فيمكن تقديره بإجراءات إحصائية معينة.

3-2. أساليب توزيع العينة على الطبقات:

يتم توزيع العينة على الطبقات وفقاً⁽¹⁾:

- أسلوب التخصيص المتساوي: ويقوم على اختيار أعداد متساوية من المفردات من كل طبقة، بغض النظر عن حجمها وتجانس مفرداتها أمام الظاهرة موضوع الدراسة، ورغم بساطة هذا الأسلوب، إلا أنه غير دقيق لاختلاف فرص الاختيار من طبقة أو فئة لأخرى.
- أسلوب التخصيص النسبي: ويقوم على اختيار عدد من المفردات من كل طبقة أو فئة تبعاً لوزنها النسبي في المجتمع الأصلي.
- أسلوب نيومان: يستخدم هذا الأسلوب لتخفيض حجم التباين، حيث يؤخذ في الاعتبار تباين كل طبقة فيكون حجم العينة في الطبقة يتناسب طردياً مع الانحراف المعياري لتلك الفئة أو الطبقة، وهذا قصد جعل تصميم العينة أكثر فعالية من أسلوب التخصيص النسبي.
- يستخدم أسلوب نيومان في العادة، عندما يكون الانحراف المعياري مختلفاً من طبقة أو فئة لأخرى، وعندما يكون حجم العينة محددًا سلفاً وتكلفتها ثابتة لمختلف الطبقات أو الفئات. وبناءً على ذلك يقدر حجم العينة للطبقة، حسب المعادلة التالية:

$$n_h = \frac{N_h S_h}{\sum_{h=1}^I N_h S_h}$$

حيث I: عدد الطبقات.

N_h : عدد وحدات المعاينة في الطبقة n.

h: عدد العينات لجميع الطبقات.

S_h : الانحراف المعياري على مستوى كل طبقة.

(1) عبد المجيد قدي، أسس البحث العلمي في العلوم الاقتصادية، دار الأبحاث، الجزائر، 2009، ص ص 88 - 89.

يمكن الحصول على الانحراف المعياري على مستوى كل طبقة من خلال تعداد سابق أو يتم تقديره من عينات سابقة.

- أسلوب التخصيص الأمثل: يقوم هذا الأسلوب على الأخذ في الحسبان درجة تجانس الظاهرة داخل كل طبقة أو فئة، وبالتالي يتم زيادة حجم وعدد المفردات الطبقة أو الفئة التي تقل فيها درجة التجانس ضمن العينة ، ويهدف هذا الأسلوب إلى تخفيض التباين لأقل قدر ممكن بكلفة محددة أو تقليل الكلفة أقل ما يمكن بمستوى دقة معين حيث يدخل عامل الكلفة في توزيع العينات على الطبقات.

يستخدم هذا الأسلوب عادة، عند وجود تفاوت في كلفة جمع المعلومات بين الطبقات أو الفئات (ككلفة جمع المعلومات من المناطق البعيدة وكلفة جمعها من المناطق القريبة).
من بين المعادلات للتخصيص الأمثل:

$$n_h = \frac{N_h S_g / C_h}{\sum_{h=1}^l N_h S_h / C_h}$$

علما بأن n_h هي تكلفة إحصاء وحدة المعاينة الواحدة للطبقة h .

4 - سلسلة تمارين غير محلولة:

التمرين الأول:

أجب عن الأسئلة التالية:

1. حدد الفرق بين الإحصاء الوصفي والإحصاء الرياضي ؟
2. ما المقصود بالمعاينة العشوائية ؟ وما أهميتها ؟
3. لماذا يلجأ إلى أسلوب العينة في معظم الدراسات الإحصائية ؟

التمرين الثاني:

ما هو أسلوب المعاينة المناسب في كل حالة مما يلي مع التبرير:

1. دراسة برامج المؤسسات الصناعية الجزائرية اتجاه المحافظة على البيئة المحيطة.
2. الوقوف بمدخل الجامعة واستجواب الطلبة الخمس الذين يتم لقاءهم حول موضوع معين.
3. استجواب الرجال والنساء المتسوقين بمحل سوبرات وسط المدينة حول اتجاهاتهم نحو أثر ارتفاع الأسعار على القدرة الشرائية بالوطن.
4. دراسة تتعلق بتجار التجزئة لمنتجات مؤسسة ما في ضوء إستراتيجية تسويقية ما.
5. دراسة حول مستوى رضا العاملين بقسم الأطفال بأحد مستشفيات ولاية المسيلة الموجودة في الشرق الجزائري.

التمرين الثالث:

نرغب في سحب عينة عشوائية بسيطة حجمها 20 من مجتمع إحصائي عدد مفرداته 205. بالاستعانة بجدول الأرقام العشوائية، حدد أرقام مفردات العينة التي يمكن سحبها من هذا المجتمع الإحصائي.

التمرين الرابع:

نريد اختيار عينة عشوائية منتظمة من مجتمع إحصائي حجمه 600 وبافتراض أن مفرداته متوفرة ضمن قائمة مرتبة، حدد أرقام مفردات العينة المسحوبة منه ذات الحجم 30.

التمرين الخامس:

عند دراسة اتجاهات طلبة ليسانس بكلية العلوم الاقتصادية بجامعة المسيلة نحو تخصصات الماستر المتاحة بها بالأقسام الأربعة، وبعد الرجوع إلى السجلات الرسمية وبافتراض أنه تم الحصول على البيانات المدونة في الجدول أدناه:

التخصص	علوم التسيير	مالية ومحاسبة	علوم تجارية	علوم اقتصادية	المجموع
عدد الطلبة	280	260	220	240	1000

وأردنا سحب عينة عشوائية منه بحيث يكون عدد مفرداتها منه هو 10% من مجموع مفردات المجتمع، بحيث تكون كل التخصصات ممثلة في العينة.

- ما هو تبريرك لإمكانية اختيار عينة عشوائية من المجتمع الإحصائي؟
- ما هو نوع هذه العينة العشوائية؟
- ما هو حجم العينة الكلي المطلوب؟
- ما هو حجم العينة الممثل للمجتمع الإحصائي في كل قسم من أقسام الكلية؟

الفصل الثاني

توزيع المعاينة

- ✓ مصطلحات ومفاهيم
- ✓ توزيع المعاينة للمتوسط
- ✓ توزيع المعاينة للنسبة
- ✓ توزيع المعاينة للتباين
- ✓ توزيع المعاينة للفروق والمجاميع
- ✓ توزيع المعاينة للنسبة بين تباينين
- ✓ سلسلة تمارين محلولة
- ✓ سلسلة تمارين غير محلولة

تمهيد:

يقوم الباحث باختيار الأسلوب المناسب للعينة العشوائية وتحديد المقياس أو المقاييس التي ستم دراستها في العينة، وذلك بدراسته للخاصية أو الصفة التي تتعلق بالعينة من أجل الاستدلال من خلالها على خاصية أو صفة معينة تقابلها في المجتمع التي أخذت منه تلك العينة.

هذا الاتجاه في الدراسة هو الفكرة الأساسية التي يقوم عليها الاستدلال الإحصائي، وفي هذا الفصل سيتم تناول بعض المقاييس التي تتصف بها العينة كالمتوسط الحسابي والانحراف المعياري والنسبة والاسترشاد من خلالها على المقاييس المقابلة لها في المجتمع بالإضافة إلى نوع التوزيع لكل منها.

1. مصطلحات ومفاهيم⁽¹⁾:

- ✓ معالم المجتمع: تتمثل في الخصائص المتعلقة بالمجتمع مثل المتوسط، التباين...
- ✓ إحصائية العينة: تتمثل في الخصائص المتعلقة بالعينة مثل: المتوسط، التباين...
- ✓ المعاينة: الكيفية أو العمليات التي تسمح بانتقاء مجموعة فرعية من المجتمع.
- ✓ المعاينة النفاذية: تكون المعاينة نفاذية عندما يكون السحب بدون إرجاع لأن المجتمع يتناقص مع تكرار ومواصلة السحب، إذ يستحيل أن تظهر مفردة في العينة أكثر من مرة، وفي هذه الحالة لا تكون نتائج السحب مستقلة.
- ✓ المعاينة غير النفاذية: هي تلك المعاينة التي يكون فيها السحب مع الإرجاع، وتسمى غير نفاذية لأنها لا تؤدي إلى نفاذ وزوال مفردات المجتمع، كما أن المفردة يمكن أن تظهر أكثر من مرة في العينة، وهنا تكون متغيرات العينة مستقلة ولها نفس التوزيع.
- ✓ توزيع المعاينة: يعبر توزيع المعاينة عن قيم المقياس المحسوبة (متوسط حسابي أو نسبة أو تباين...) لكل عينة من العينات العشوائية التي لها نفس الحجم (n)، والتي يمكن سحبها من المجتمع قيد الدراسة، فهو إذن توزيع احتمالي لجميع القيم الممكنة لإحصاءة ما في كل العينات العشوائية الممكنة والتي لها نفس الحجم.
- ✓ خطأ المعاينة: هو الفرق بين القيمة المحسوبة لإحصائية العينة وقيمة معلمة المجتمع المناظرة.
- ✓ الخطأ المعياري: هو الانحراف المعياري لقيم الإحصاءة الممكنة في كل العينات العشوائية الممكنة التي من نفس الحجم.

2. توزيع المعاينة للمتوسط الحسابي:

من أجل تحديد العلاقة بين المتوسط الحسابي للمجتمع والمتوسط الحسابي لتوزيع المعاينة في حالتي السحب بدون إرجاع والسحب بالإرجاع نعرض المثال التالي⁽²⁾:

مثال 1:

مجتمع يتكون من 5 مفردات هي (2، 4، 6، 8، 10)، قمنا بسحب جميع العينات المكونة من مفردتين والمطلوب هو:

(1) جورج كانافوش، دون ميلر، ترجمة سلطان محمد عبد المجيد، محمد توفيق البلقيني، مرجع سابق، ص 271 بتصريف.
- د. ليونارد، ج. كازمير، ترجمة مصطفى جلال مصطفى، الإحصاء التجاري، الدار الدولية للاستثمارات الثقافية، مصر، 2004، ص 80 بتصريف.

(2) المثال مقتبس من:

محمد حسين محمد رشيد، منى عطا الله الشويلات، مبادئ الإحصاء والاحتمالات ومعالجتها باستخدام برنامج SPSS، دار صفاء، عمان، الأردن، 2012، ص ص 378 - 383 بتصريف.

1. تحديد العينات الممكن سحبها بدون إرجاع.
2. تحديد العينات الممكن سحبها إذا كان السحب بإرجاع.
3. حساب المتوسط الحسابي والتباين للمجتمع.
4. تكوين توزيع المعاينة لمتوسطات العينات الممكنة ثم حساب الوسط والتباين لتوزيع المعاينة في حالة السحب بدون إرجاع.
5. الإجابة على الفرع السابق في حالة السحب بالإرجاع.
6. مقارنة الإجابات في الأسئلة رقم 4 و رقم 5 مع النتائج المتحصل عليها في السؤال رقم 3.
7. اشتقاق العلاقة بين كل من توزيع المعاينة لكل من المتوسط الحسابي والتباين في حالتي السحب بدون إرجاع والسحب بالإرجاع.

الحل:

1- لإيجاد عدد العينات التي يمكن سحبها بدون إرجاع نستخدم التوفيقية:

$$C_2^5 = 5.4.3.2.1 / 2!3! = 10$$

العينات الممكن سحبها هي:

$$(4, 2), (6, 2), (8, 2), (10, 2), (4, 4), (6, 4), (8, 4), (10, 4), (6, 6), (8, 6), (10, 6), (8, 8), (10, 8), (4, 10), (6, 10), (8, 10), (10, 10)$$

2- في هذه الحالة تم السحب مع الإرجاع وبالتالي نستخدم القائمة لتحديد عدد العينات الممكن سحبها:

$$5^2 = 25$$

العينات الممكن سحبها هي:

$$(2, 2), (2, 4), (2, 6), (2, 8), (2, 10), (4, 2), (4, 4), (4, 6), (4, 8), (4, 10), (6, 2), (6, 4), (6, 6), (6, 8), (6, 10), (8, 2), (8, 4), (8, 6), (8, 8), (8, 10), (10, 2), (10, 4), (10, 6), (10, 8), (10, 10)$$

3- حساب المتوسط الحسابي والتباين للمجتمع:

$$\mu = \frac{2+4+6+8+10}{5} = 6 \text{ المتوسط الحسابي}$$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \text{ التباين للمجتمع}$$

$$= \frac{(2-6)^2 + (4-6)^2 + \dots + (10-6)^2}{5} = 8$$

4- لتكوين توزيع المعاينة لمتوسطات العينات الممكنة في حالة السحب بدون إرجاع نقوم أولاً

بحساب المتوسط الحسابي لكل من العينات العشر المختلفة:

$$\bar{X} = \frac{\sum X}{n}$$

يتم حساب المتوسط الحسابي لكل عينة وفق العلاقة:

فمثلاً: تم حساب المتوسط الحسابي للعينة الأولى (4، 2) في الجدول.

$$\frac{2+4}{2} = \frac{6}{2} = 3$$

العينة	(2,4)	(2,6)	(2,8)	(2,10)	(4,6)	(4,8)	(4,10)	(6,8)	(6,10)	(8,10)
المتوسط الحسابي	3	4	5	6	5	6	7	7	8	9

وانطلاقاً من قيم المتوسط الحسابي في كل عينة نقوم بتكوين توزيع المعاينة له وحساب متوسطه

وتباينه كما يلي:

قيمة المتوسط \bar{X}_i	التكرارات f_i	$\bar{X}_i f_i$	\bar{X}_i^2	$\bar{X}_i^2 f_i$
3	1	3	9	9
4	1	4	16	16
5	2	10	25	50
6	2	12	36	72
7	2	14	49	98
8	1	8	64	64
9	1	9	81	81
المجموع	10	60	/	390

حساب متوسط المتوسطات:

$$\mu_{\bar{x}} = \frac{60}{10} = 6$$

أما قيمة التباين فهي:

$$\sigma^2 = \frac{\sum \bar{X}_i^2 f_i}{\sum f_i} - \mu_{\bar{x}}^2 = \frac{390}{10} - 6^2 = 3$$

5- تكوين توزيع المعاينة لمتوسطات العينات في حالة السحب بالإرجاع:

المتوسط الحسابي	العينة	المتوسط الحسابي	العينة
7	(8 ، 6)	2	(2 ، 2)
8	(10 ، 6)	3	(2 ، 4)
5	(8 ، 2)	4	(6 ، 2)
6	(4 ، 8)	5	(8 ، 2)
7	(6 ، 8)	6	(10 ، 2)
8	(8 ، 8)	3	(2 ، 4)
9	(10 ، 8)	4	(4 ، 4)
6	(10 ، 2)	5	(6 ، 4)
7	(4 ، 10)	6	(8 ، 4)
8	(6 ، 10)	7	(10 ، 4)
9	(8 ، 10)	4	(2 ، 6)
10	(10 ، 10)	5	(4 ، 6)
		6	(6 ، 6)

بعد الحصول على المتوسطات الحسابية نستخدم الجدول الموالي من أجل تحديد قيمتي المتوسط الحسابي والتباين في حالة السحب بالإرجاع:

$\bar{X}_i^2 f_i$	\bar{X}_i^2	$\bar{X}_i f_i$	التكرارات f_i	قيمة المتوسط \bar{X}_i
4	4	2	1	2
18	9	6	2	3
48	16	12	3	4
100	25	20	4	5
180	36	30	5	6
196	49	28	4	7
192	64	24	3	8
162	81	18	2	9
100	100	10	1	10
1000	/	150	25	المجموع

$$\mu_{\bar{X}} = \frac{150}{25} = 6$$

حساب متوسط المتوسطات:

$$\sigma^2 = \frac{\sum \bar{X}_i^2 f_i}{\sum f_i} - \mu_{\bar{X}}^2 = \frac{1000}{25} - 6^2 = 4$$

أما قيمة التباين فهي:

6- مقارنة نتائج حالتي السحب:

العينة من الإرجاع	العينة دون إرجاع	المجتمع	
6	6	6	المتوسط الحسابي
4	3	8	التباين

من خلال الجدول السابق يتضح بأن قيمة المتوسط الحسابي في المجتمع هي نفسها عند توزيع المعاينة في كل من الحالتين سواء السحب بالإرجاع أو بدون إرجاع.

أما بالنسبة للتباين فإن قيمه جاءت مختلفة في الحالات الثلاث، أي في المجتمع وحالتي السحب أيضا.

7- اشتقاق العلاقة لكل من المتوسط والتباين في حالتي السحب

أولاً: بالنسبة للمتوسط الحسابي

المتوسط الحسابي لتوزيع المعاينة في حالة السحب بالإرجاع أو بدون إرجاع:

$$\bar{X} = u_{\bar{X}} = \mu$$

ثانياً: بالنسبة للتباين

أ- في حالة السحب مع الإرجاع:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

ب- في حالة السحب دون الإرجاع:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \times \left(\frac{N-n}{N-1} \right)$$

التحقق:

أولاً: بالنسبة لمتوسط توزيع المعاينة للمتوسط في حالتي السحب بإرجاع أو بدون إرجاع فإن القيمة ثابتة

ومساوية لمتوسط المجتمع، وبالتالي فإن: $\mu = \mu_{\bar{X}} = \bar{X}$

ثانياً: بالنسبة لتباين توزيع المعاينة للمتوسط:

أ- في حالة السحب بالإرجاع: $\sigma_{\frac{2}{X}} = 4$

وتباين المجتمع $\sigma^2 = 8$

وباستخدام العلاقة: $\sigma_{\frac{2}{X}} = \frac{\sigma^2}{n}$ والتعويض نجد:

$$\sigma_{\frac{2}{X}} = \frac{\sigma^2}{n} = \frac{8}{2} = 4$$

ب- في حالة السحب بدون إرجاع: $\sigma^2 = 8$ و $\sigma_{\frac{2}{X}} = 3$

وباستخدام العلاقة: $\sigma_{\frac{2}{X}} = \frac{\sigma^2}{n} \times \left(\frac{N-n}{N-1}\right)$

والتعويض نجد: $\sigma_{\frac{2}{X}} = \frac{\sigma^2}{n} \times \left(\frac{N-n}{N-1}\right) = \frac{8}{2} \times \left(\frac{5-2}{5-1}\right) = 3$

ومما سبق نستخلص ما يلي:

– \bar{X} قيمة ثابتة في العينة الواحدة، ولكن إذا كان لدينا عددا كبيرا من العينات فإن \bar{X} يعتبر متغيرا

عشوائيا والتوزيع الاحتمالي للمتغير العشوائي \bar{X} يطلق عليه توزيع المعاينة للمتوسط، وبالتالي

فإن إحصاءات توزيع المعاينة للمتوسط هي:

$\mu_{\bar{X}}$: المتوسط الحسابي لتوزيع المعاينة.

$\sigma_{\bar{X}}$: الانحراف المعياري لتوزيع المعاينة (نسميه الخطأ المعياري للمتوسط).

– متوسط توزيع المعاينة للمتوسط مساو لمتوسط المجتمع سواء كانت المعاينة غير نفاذية أو نفاذية

(السحب بالإرجاع أو بدون إرجاع) ومهما كان حجم العينة المسحوبة.

ملاحظات:

● في حالة المجتمعات الكبيرة فإن صيغة الانحراف المعياري سواء كانت المعاينة غير نفاذية

أو نفاذية تكون كما يلي: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

أي هي نفسها صيغة الانحراف المعياري لما تكون المعاينة غير نفاذية، حيث أنها تعتمد فقط على

حجم العينة، وأن قيمة $\sigma_{\bar{X}}$ تقل مع زيادة حجم العينة.

● يطلق على النسبة $\frac{n}{N}$ معدل الاستقصاء.

عندما تكون قيمة معدل الاستقصاء أقل من 0.05 أي $\left(\frac{n}{N} < 0.05\right)$ فإنه يتم إهمال قيمة معامل

الإرجاع $\left(\frac{N-n}{N-1}\right)$ في علاقة التباين⁽¹⁾.

(1) د. ليونارد، ج. كازمير، ترجمة مصطفى جلال مصطفى، الإحصاء التجاري، الدار الدولية للاستثمارات الثقافية، مصر، 2004،

3. نوع توزيع المعاينة للمتوسط:

سيتم هنا تناول طبيعة توزيع المعاينة للمتوسط لما يكون المجتمع طبيعياً أو غير طبيعي.

3-1. نوع توزيع المعاينة للمتوسط عندما يكون للمجتمع توزيع طبيعي:

إذا كان المجتمع الذي تتم منه المعاينة له توزيع طبيعي بمتوسط μ وتباين σ^2 فإن متوسط العينة المسحوبة أيضاً يتبع التوزيع الطبيعي بمتوسط μ وانحرافاً معيارياً $\frac{\sigma}{\sqrt{n}}$ ، ونكتب:

$$X \sim N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{ومنه: } z = \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}} \sim N(0,1)$$

مثال 2:

يختار مراجع عينة عشوائية حجمها 16 من مجتمع يتكون من 1000 حساب مدين يفترض أنها تتوزع طبيعياً حيث متوسط قيمة الحساب المدين للمجتمع هو 26000 دج بانحراف معياري 4500 دج. - ما هو احتمال أن يكون متوسط العينة أقل من 28250 دج؟

الحل:

$$\mu = 26000, \sigma = 4500, N = 1000, n = 16$$

$$P(\bar{X} < 28250) = ?$$

$$n = 16 : n/N = 16/1000 = 0.016 < 0.05 \quad (\text{نهمل معامل الإرجاع})$$

$$\Rightarrow \sigma_m = \sigma/\sqrt{n} = 4500/\sqrt{16} = 1125$$

$$X \sim N(26000, 4500) \Rightarrow \bar{X} \sim N(26000, 1125)$$

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{28250 - 26000}{1125} = 2$$

$$P(\bar{x} < 28250) = P(Z < 2) = 0.5 + P(0 < Z < 2)$$

$$= 0.5 + 0.4772 = 0.9772$$

أي أن احتمال أن يكون متوسط العينة أقل من 28250 هو: 0.9772

3-2. نوع توزيع المعاينة لمتوسط العينة عندما يكون المجتمع له توزيع غير طبيعي:

تصادفنا العديد من الحالات التي لا يمكننا فيها تحديد طبيعة توزيع المجتمع، الأمر الذي يحولنا عن تحديد توزيع المعاينة لمتوسط العينة، ومع ذلك فقد تمكن علماء الإحصاء من إثبات أن توزيع المعاينة لـ \bar{X}

هو التوزيع الطبيعي في حالة العينات ذات الأحجام الكبيرة أيا كان توزيع المجتمع، وهذه النتيجة الحاسمة تعرف باسم نظرية النهاية المركزية⁽¹⁾.

نظرية النهاية المركزية

إذا كان حجم العينة كبيرا ($n \geq 30$)، فإن توزيع المعاينة للمتوسط الحسابي \bar{X} يقترب من التوزيع الطبيعي، بغض النظر عن شكل المجتمع الأصلي ذي المتوسط μ والانحراف المعياري σ ، ولذلك يمكن حساب احتمال أن يكون \bar{X} لعينة عشوائية من خلال القيمة الإحصائية لـ Z وفق العلاقة الشهيرة:

$$\bar{X} \approx N(\mu, \sigma_{\bar{X}}) \text{ ، ويكون: } Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

مثال 3:

شركة مختصة في صناعة الحافلات وتنتج 70 كغ لوزن الراكب بانحراف معياري 10 كغ، إذا علمت أنها قامت بتصميم حافلة لنقل العمال حمولتها القصوى هي 2880 كغ وتتسع إلى 36 عاملا، فما هو احتمال أن تحمل هذه الحافلة أكبر من حمولتها؟

الحل:

نلاحظ أن :

$$\bar{X} = \frac{2880}{36} = 80$$

المطلوب حساب: $P(\bar{x} > 80) = ?$

$$\mu = 70, \sigma = 10, X \sim ?, n = 36.$$

توزيع المجتمع مجهول لكن حجم العينة 36 أكبر من 30. وبالاستناد على نظرية النهاية المركزية نستنتج بأن متوسطات أوزان العمال تقترب من التوزيع الطبيعي.

$$Z = \frac{\bar{x} - \mu}{\sigma/n} = \frac{80 - 70}{24/\sqrt{36}} = 2.5$$

$$P(\bar{X} > 80) = P(Z > 2.5) = 0.5 - 0.4938 = 0.0062$$

(1) للاستزادة أكثر حول الفكرة الأساسية لنظرية النهاية المركزية يمكن الرجوع إلى: أنيس إسماعيل كنجو، الإحصاء والاحتمال، مكتبة العبيكان، الرياض، المملكة العربية السعودية، 2000.

4. توزيع المعاينة للنسبة:

في العديد من الحالات تكون المعلمة الأساسية هي النسبة p ، ومن أمثلة ذلك نسبة الفواتير التي بها أخطاء، نسبة المكالمات التليفونية التي تتجاوز حداً قياسيًّا، نسبة شيكات العملاء التي بدون رصيد، نسبة الوحدات المرتجعة من العملاء، والنسبة p' في العينة هي أفضل إحصاءة يمكن استخدامها للاستدلال عن النسبة P في المجتمع.

4-1. المتوسط والخطأ المعياري للنسبة:

إذا كان لدينا مجتمع ما نسبة صفة معينة فيه هي p ، وتم سحب عينة منه ذات حجم n نسبة نفس تلك الصفة فيها هي p' ، فإن قيمة كل من المتوسط الحسابي والانحراف المعياري لتوزيع المعاينة يكونان كما يلي:

$$\mu_{p'} = p, \sigma_{p'} = \sqrt{\frac{pq}{n}}$$

حيث:

$\mu_{p'}$ = متوسط توزيع المعاينة للنسبة.

p : نسبة النجاح في المجتمع.

q : نسبة الفشل في المجتمع. $q = 1 - p$

$\sigma_{p'}$: الانحراف المعياري للإحصائية p'

مثال 4:

حدد كلا من المتوسط والخطأ المعياري لنسبة العينة في كل حالة مما يلي:

أ. عينة حجمها 100، سحبت من مجتمع نسبته 0.5

ب. عينة حجمها 20 سحبت من مجتمع نسبته 0.5

ج. ما هو تعليقك على النتيجتين؟

الحل:

أ- متوسط العينة للنسبة في الحالة أ :

لدينا: $\mu_{p'} = p = 0.5$ (متوسط العينة للنسبة)

- الانحراف المعياري لنسبة العينة

$$\sigma_{p'} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.5(1-0.5)}{100}} = 0.05$$

ب- متوسط العينة في الحالة ب:

لدينا: $\mu_{p'} = p = 0.5$ (متوسط العينة للنسبة)

- الانحراف المعياري لنسبة العينة:

$$\sigma_{p'} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.5(1-0.5)}{20}} = 0.11$$

ج- نلاحظ بأن قيمة متوسط النسبة لم تتغير بتغير حجم العينة في نفس المجتمع، بينما ارتفعت قيمه الانحراف المعياري بانخفاض حجم العينة.

4-2. نوع توزيع المعاينة للنسبة p في العينة:

النظرية الموالية توضح طبيعة توزيع p'.

إذا تم أخذ عينات عشوائية من الحجم n ($n > 30$) من مجتمع ذي الحدين بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ ، فإن توزيع المعاينة لـ p' (والتي تمثل نسبة النجاح)، يتوزع طبيعياً بمتوسط حسابي $\mu_{p'}$ وانحراف معياري $\sigma_{p'}$ ، حيث أن:

$$\mu_{p'} = p, \sigma_{p'} = \sqrt{\frac{pq}{n}}$$

وتكون قيمة Z تساوي: $Z = \frac{p' - p}{\sqrt{pq/n}}$

فإن: $p' \approx N(P, \sigma_{p'})$

و $p' = \frac{x}{n}$ هي قيمة المتغير العشوائي القياسي.

مثال 5:

إذا كان احتمال نجاح الطالب بدون ديون في كلية العلوم الاقتصادية هو 0.9، أخذت عينة عشوائية حجمها 49 طالبا من أولئك الذين يدرسون في الكلية، ما هو احتمال أن تزيد نسبة هؤلاء الطلبة بدون ديون على 80%؟

الحل:

لدينا: $\mu_{p'} = p = 0.9$ (متوسط العينة للنسبة).

أما الانحراف المعياري فيحسب كما يلي:

$$\sigma_{p'} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.9(1-0.9)}{49}} = 0.042$$

وبما أن حجم العينة $n = 49 > 30$ وبالاعتماد على النظرية السابقة فإنه يمكن التقريب للتوزيع الطبيعي

ويكون لدينا:

$$P(p' > 0.8) = p\left(Z > \frac{p' - p}{\sigma_{p'}}\right)$$

$$= p\left(Z > \frac{0.8 - 0.9}{0.042}\right) = p(Z > -2.38)$$

$$P(p' > 0.8) = P(Z > -2.38) = 0.5 + P(0 < Z < 2.38)$$

$$= 0.5 + 0.4913 = 0.9913$$

5. توزيع المعاينة للتباين:

بالعودة للمثال⁽¹⁾ رقم (1) من هذا الفصل سيتم حساب متوسط تباينات جميع العينات المسحوبة في حالتها المعاينة النفاذية وغير النفاذية ومقارنة متوسط التباينات في كل حالة منها مع تباين المجتمع: أولاً: في حالة المعاينة النفاذية (السحب دون إرجاع)

$$S^2 = \sum \frac{(x_i - \bar{x})^2}{2} \text{ : العلاقة التالية: } n \text{ الحجم وفق العلاقة التالية:}$$

العينه	المتوسط الحسابي	
(2 ، 4)	3	1
(2 ، 6)	4	4
(2 ، 8)	5	9
(2 ، 10)	6	16
(4 ، 6)	5	1
(4 ، 8)	6	4
(4 ، 10)	7	9
(6 ، 8)	7	1
(6 ، 10)	8	4
(8 ، 10)	9	1
المجموع	/	50

$$E(S^2) = \frac{\sum S^2}{25} = \frac{50}{10} = 5$$

(1) سامح جزماتي، الاحتمالات والإحصاء، مديرية الكتب والمطبوعات الجامعية، جامعة حلب، دمشق، 1989، ص ص 27 - 28.

ثانيا: في حالة المعاينة غير النفاذية (السحب مع الإرجاع)

	المتوسط الحسابي	العينة
0	2	(2، 2)
1	3	(2، 4)
4	4	(6، 2)
9	5	(8، 2)
16	6	(10، 2)
1	3	(2، 4)
0	4	(4، 4)
1	5	(6، 4)
4	6	(8، 4)
9	7	(10، 4)
4	4	(2، 6)
1	5	(4، 6)
0	6	(6، 6)
1	7	(8، 6)
4	8	(10، 6)
9	5	(8، 2)
4	6	(4، 8)
1	7	(6، 8)
0	8	(8، 8)
1	9	(10، 8)
16	6	(10، 2)
9	7	(4، 10)

4	8	(6 ، 10)
1	9	(8 ، 10)
0	10	(10 ، 10)
100	/	المجموع

$$E(S^2) = \frac{\sum S^2}{25} = \frac{100}{25} = 4$$

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

6. توزيع المعاينة للفروق والمجاميع:

في الحالة التي تكون عندنا عينتين مأخوذتين من مجتمعين مختلفين فإنه يمكن استنتاج كل من الفروق والمجاميع وتوزيعها لعدة خصائص تتعلق بهما كما يلي:

6-1. الفروق والمجاميع للمتوسط الحسابي والتباين والنسبة لتوزيع المعاينة:

إذا كانت x_1, x_2, \dots, x_n عينة مأخوذة من توزيع متوسطه الحسابي μ_1 وانحرافه المعياري σ_1 أو له صفة معينة نسبتها هي p_1 كانت y_1, y_2, \dots, y_n عينة أخرى مأخوذة من توزيع آخر متوسطه الحسابي μ_2 وانحرافه المعياري σ_2 ، وله صفة معينة نسبتها p_2 .

فإنه يمكن الحصول على توزيع الفروق والمجاميع للأوساط الحسابية والتباينات ونسب المعاينة وفق

العلاقات التالية⁽¹⁾:

الجدول 2-1: توزيع الفروق والمجاميع للأوساط الحسابية والتباينات ونسب المعاينة

النسبة	التباين	المتوسط الحسابي	
$\mu_{p_1 - p_2} = P_1 - P_2$	$\mu_{\bar{x} - \bar{y}}^2 = \sigma_x^2 + \sigma_y^2$	$\mu_{\bar{x} - \bar{y}} = \mu_1 - \mu_2$	الفروق
$\mu_{p_1 + p_2} = P_1 + P_2$	$\sigma_{\bar{x} + \bar{y}}^2 = \sigma_x^2 + \sigma_y^2$	$\mu_{\bar{x} + \bar{y}} = \mu_1 + \mu_2$	المجاميع

المصدر: من تلخيص الباحثة اعتمادا على عدة مراجع

(1) دلال القاضي وآخرون، الإحصاء للإداريين والاقتصاديين، دار الحامد، عمان، الأردن، 2005، ص ص 2015 - 2019 بتصرف.

مثال 6:

أخذت عينة حجمها 30 وحدة من توزيع متوسط الحسابي 75 وتباينه 25، وأخذت عينة ثانية حجمها 60 من توزيع آخر مستقل متوسطه الحسابي 60 وتباينه 15.

- أوجد الفروق والمجاميع للمتوسط الحسابي لتوزيع المعاينة للتوزيعين.

الحل:

أولاً: بالنسبة للفروق

$$\mu_{\bar{x}-\bar{y}} = 75 - 60 = 15 \quad \text{المتوسط الحسابي:}$$

أما الفروق بين الانحرافين:

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{25}{30} + \frac{15}{60}} = 1.04$$

ثانياً: بالنسبة للمجاميع

$$\mu_{\bar{x}+\bar{y}} = 75 + 60 = 135 \quad \text{المتوسط الحسابي:}$$

أما الفروق بين الانحرافين:

$$\sigma_{\bar{x}+\bar{y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{25}{30} + \frac{15}{60}} = 1.04$$

6-2. طبيعة توزيع المعاينة لمجموع أو الفرق بين متوسطين:

إذا كانت العينتان المسحوبتان من مجتمعين طبيعيين أو حجم كل منهما يساوي أو يفوق 30 مفردة فإن الفرق بين متوسطي العينتين أو مجموعهما يتبع أو يقترب من التوزيع الطبيعي⁽¹⁾، حيث:

$$(\bar{x} - \bar{y}) \rightarrow N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \Rightarrow \frac{(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \rightarrow N(0,1)$$

مثال 7:

تم سحب عینتين عشوائيتين من شركتين مختلفتين لإنتاج الأدوية، وكانت الأجور المدفوعة من قبل الشركتين تتبع التوزيع الطبيعي، وأن معدل الأجور المدفوعة من الشركة الأولى إلى 25 عاملاً تساوي 35000 دج بانحراف معياري 4000 دج، أما معدل الأجور لـ 40 عاملاً فهو 28000 دج بانحراف معياري 6000 دج.

(1) محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل باستخدام SPSS، دار المسيرة، عمان، الأردن، 2004، ص 192، بتصرف.

- ما هو احتمال أن يزيد معدل الأجور في الشركة الأولى بـ 8000 دج على الأقل عن الشركة الثانية؟

الحل:

نلخص معطيات المثال في الجدول الموالي:

الشركة الثانية	الشركة الأولى	
$\mu_2 = 28000$	$\mu_1 = 35000$	المتوسط الحسابي
$\sigma_2 = 6000$	$\sigma_1 = 4000$	الانحراف المعياري
40	25	حجم العينة

$$\mu_{\bar{x}-\bar{y}} = \mu_1 - \mu_2 = 35000 - 28000 = 7000$$

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4000^2}{25} + \frac{6000^2}{40}} = 981.83$$

$$p(\bar{x} - \bar{y} > 8000) =$$

$$= p[Z > (8000) - (7000)/981.83] = p[Z > 1.02]$$

$$= 0.5 - 0.3461 = 0.1539$$

7. توزيع المعاينة للنسبة بين تباينين:

عند المقارنة بين تباينين مجتمعين فإننا نحتاج إلى النسبة بين تباينين عينتين مأخوذتين منهما، وهذا يمهّد لدراسة أحد التوزيعات الاحتمالية المتصلة المستخدمة كثيرا وهو توزيع فيشر.

7-1. توزيع فيشر⁽¹⁾:

يعتبر توزيع فيشر أحد التوزيعات الاحتمالية المتصلة المهمة والمستخدم في اختبار الفرضيات وفي تحليل التباين، حيث يعرف متغيره العشوائي بالدالة الاحتمالية التالية:

$$f(F) = \frac{CF^{\left(\frac{v_1-2}{2}\right)}}{(v_2+v_1F)^{\left(\frac{v_1+v_2}{2}\right)}}; F > 0$$

(1) معتوق أحمد، الإحصاء الرياضي والنماذج الإحصائية، ديوان المطبوعات الجامعية، الجزائر، 2007، ص ص 44 - 45.

حيث ν_1, ν_2 هما درجات الحرية و c ثابت يعتمد عليهما ويتحدد ليجعل المساحة تحت منحني التوزيع تساوي 1.

فهذا التوزيع يتضمن عددين لدرجات الحرية، وحيث أن ν_2 لا تظهر إلا في المقام فإنه يعتبر درجات حرية المقام، ويعتبر درجات حرية البسط ونكتب: $X \sim F_{\nu_1, \nu_2}$.

7-2. خواص توزيع فيشر:

- أحادي المنوال، وملتوي إلى اليمين قليلاً.
- موجب لجميع قيم F .
- كلما زادت درجات الحرية ν_1, ν_2 يقترب توزيع فيشر من التوزيع الطبيعي.

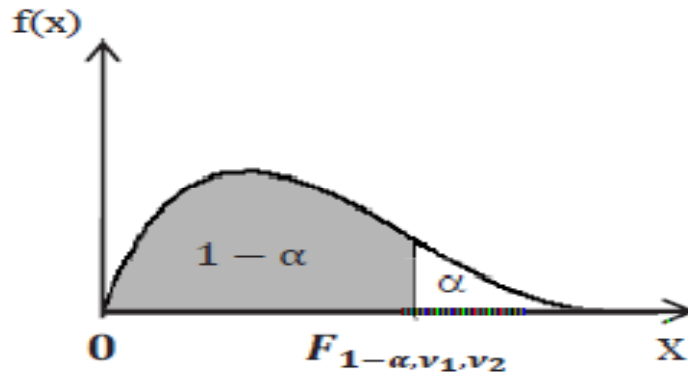
توجد جداول خاصة بتوزيع فيشر لإيجاد قيمة F التي على يسارها مساحة ما بدرجات حرية ν_1, ν_2 . حيث قيم ν_1 توجد في السطر الأفقي أعلى الجدول أما قيم ν_2 فتوجد على العمود الأيسر من الجدول ونجد المساحات أو الاحتمالات إما بالعمود المحاذي لعمود قيم ν_2 أو كمساحة مستقلة بجدول خاص قيمتها أعلى ذلك الجدول.

كما يمكن حساب قيمة F بالاعتماد على العلاقة التالية:

$$F = \frac{x_1^2/\nu_1}{x_2^2/\nu_2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$$

حيث أن x_1^2 هي قيمة توزيع كاي تربيع بدرجة حرية ν_1 ، x_2^2 هي قيمة توزيع كاي تربيع بدرجة حرية ν_2 ، أما S_1^2 و S_2^2 فهما تباين العينتين ذات الحجمين n_1 و n_2 على الترتيب.

الشكل 2-1: منحني توزيع فيشر



والنظرية الموالية تتناول المعاينة للنسبة بين تباينين:

ليكن S_1^2 تباين عينة عشوائية حجمها n_1 من مجتمع $N(\mu_1, \sigma_1^2)$ وليكن S_2^2 تباين عينة عشوائية حجمها n_2 من مجتمع $N(\mu_2, \sigma_2^2)$ مستقل عن الأول، فإن:

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{s_1^2\sigma_2^2}{s_2^2\sigma_1^2} \sim F_{1-\alpha, v_1, v_2}$$

ملاحظة: بما أن جداول توزيع فيشر لا تتضمن بعض المساحات الصغيرة فإنه يمكن الاستعانة بالقواعد التالية:

$$-F_{1-\alpha, v_1, v_2} = \frac{1}{F_{\alpha, v_1, v_2}}$$

$$-F_{p, 1, v} = t_{1-(p/2), v}^2$$

$$-F_{p, v, \infty} = \frac{\chi_{1-\alpha, v}^2}{v}$$

مثال:

أوجد: $F_{0.95, 9, 7}$ ؛ $F_{0.025, 11, 10}$ ؛ $F_{0.01, 11, 15}$ ؛ $F_{0.05, 7, 10}$

الحل:

بالنظر إلى جدول فيشر نحصل على ما يلي:

$$F_{0.95, 9, 7} = 3.68$$

$$F_{2.025, 11, 10} = 0.283$$

$$F_{0.01, 11, 15} = \frac{1}{F_{0.99, 15, 11}} = \frac{1}{4.25} = 0.235$$

$$F_{0.05, 7, 10} = \frac{1}{F_{0.05, 10, 7}} = \frac{1}{3.64} = 0.275$$

8 - سلسلة تمارين محلولة:

التمرين 1:

إذا كان المتغير العشوائي X له التوزيع الطبيعي $N(3;4)$..

1- ما هو احتمال أن تقع X بين القيمتين 3 و 5؟.

2- ما هو احتمال أن يكون X أكبر من 1؟

حل التمرين 1:

لدينا: $Z = \frac{X - \mu}{\delta}$ والدالة Φ هي دالة التوزيع الطبيعي المعياري (قيمها تستخرج من الجدول مباشرة).

1- احتمال أن تقع X بين القيمتين 3 و 5:

$$\begin{aligned} p(3 < X < 5) &= p\left(\frac{3-3}{2} < Z < \frac{5-3}{2}\right) \\ &= p(0 < Z < 1) = \Phi(1) - \Phi(0) \\ &= 0.8413 - 0.5 = 0.3413 \end{aligned}$$

2- احتمال أن يكون X أكبر من 1:

$$\begin{aligned} p(X > 1) &= p\left(Z > \frac{1-3}{2}\right) \\ &= p(Z > -1) = 1 - p(Z \leq -1) \\ &= 1 - \Phi(-1) = 1 - 1 + \Phi(1) \\ &= 0.8413 \end{aligned}$$

التمرين 2:

مجتمع إحصائي يخضع للتوزيع الطبيعي، يتكون من المفردات التالية: $(2; 3; 4)$ ، نسحب بإرجاع عينة من

هذا المجتمع ذات حجم $n = 2$.

1- أحسب الوسط والتباين للمجتمع؟

2- أحسب عدد العينات الممكنة؟

3- أحسب معالم العينة $(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$ بطريقة الجداول؟

حل التمرين 2:

1- حساب الوسط والتباين للمجتمع:

$$\mu = \frac{\sum X_i}{N} = \frac{2 + 3 + 4}{3} = 3$$

$$\delta^2 = \frac{\sum (X_i - \mu)^2}{N} \approx 0.67$$

2- حساب عدد العينات الممكنة:

العينات الممكنة ذات الحجم $n=2$ من مجتمع حجمه 3 عددها: $N^n = 3^2 = 9$ 3- أحسب معالم العينة $(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$ بطريقة الجداول:

العينات الممكنة			المتوسطات الممكنة للعينة (معاينة بالإرجاع) \bar{x}_i		
(2, 2)	(2, 3)	(2, 4)	2	2,5	3
(3, 2)	(3, 3)	(3, 4)	2,5	3	3,5
(4, 2)	(4, 3)	(4, 4)	3	3,5	4

- القيمة المتوقعة لمتوسط العينة $E(\bar{X})$ هي متوسط قيمها وهي: $E(\bar{X}) = \mu_x = \frac{\sum \bar{x}_i}{9} = 3$ - تباين المتوسطات الممكنة للعينة δ_x^2 :

$$\delta_x^2 = \frac{\sum (\bar{x}_i - \mu_x)^2}{9} = 0.33$$

التمرين 3:

1- إذا كان ضغط الدم لمجموعة من الأفراد يتبع التوزيع الطبيعي بمتوسط يساوي 90 وانحراف معياري يساوي 9، سحبنا عينة حجمها 36 فردا من هذه المجموعة.

ما هو احتمال أن يكون متوسط ضغط الدم في العينة أكبر من 86؟

2- إذا علم أن نسبة البيض التالف الذي ينتجه أحد مراكز إنتاج الدواجن هي 0.03 ولها توزيع طبيعي، اشترى شخص 400 بيضة من إنتاج هذا المركز.

- ما هو احتمال أن يجد من بينها 20 بيضة على الأقل تالفة؟

حل التمرين 3:

1- إيجاد احتمال أن يكون متوسط ضغط الدم في العينة أكبر من 86:

لدينا: $\mu = 90; n = 36; \delta = 9$

$$\mu_{\bar{x}} = \mu = 90$$

$$\delta_{\bar{x}} = \frac{\delta}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5$$

$$\begin{aligned} p(\bar{x} > 86) &= p\left(\frac{\bar{x} - \mu_{\bar{x}}}{\delta_{\bar{x}}} > \frac{86 - 90}{1.5}\right) \\ &= p(z > -2.67) = 1 - p(z \leq -2.67) \\ &= 1 - \Phi(-2.67) = \Phi(2.67) = 0.9962 \end{aligned}$$

2- إيجاد احتمال أن يجد 20 بيضة على الأقل تالفة:

لدينا: $p = 0.03; n = 400$

$$\mu_{p'} = p = 0.03$$

$$\delta_{p'}^2 = \frac{pq}{n} = \frac{(0.03)(0.97)}{400}$$

$$\delta_{p'} = \sqrt{\frac{(0.03)(0.97)}{400}} = 0.0085$$

$$\begin{aligned} p(x \geq 20) &= p\left(p' \geq \frac{20}{400}\right) = p(p' \geq 0.05) \\ &= p\left(z \geq \frac{0.05 - 0.03}{0.0085}\right) = p(z \geq 2.35) \\ &= 1 - p(z < 2.35) = 1 - 0.9906 = 0.0094 \end{aligned}$$

التمرين 4:

إذا كان لدينا البيانات التالية:

المجتمع	متوسط المجتمع	تباين المجتمع
1	40	18
2	35	25

- تم سحب عينتين مستقلتين من المجتمعين حيث $n_1 = 12$ و $n_2 = 10$. مع العلم أن المجتمعين طبيعيين.

- أحسب التوزيع العيني للفرق بين متوسطي العينتين $(\mu_{\bar{x}_1 - \bar{x}_2}; \delta_{\bar{x}_1 - \bar{x}_2}^2)$

حل التمرين 4:

- حساب التوزيع العيني للفرق بين متوسطي العينتين $(\mu_{\bar{x}_1 - \bar{x}_2}; \sigma_{\bar{x}_1 - \bar{x}_2}^2)$:

$$\mu_{\bar{x}_1 - \bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2 = 40 - 35 = 5$$

$$\sigma_{\bar{x}_1 - \bar{x}_2}^2 = \sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{18}{12} + \frac{25}{10} = 4$$

$$\bar{x}_1 - \bar{x}_2 \rightarrow N(5; 4)$$

9 - سلسلة تمارين غير محلولة:

التمرين 1:

إذا كان لدينا مجتمع يتكون من المفردات التالية (0، 2، 4، 6) نسحب عينة من هذا المجتمع ذات الحجم $n = 2$.

- أوجد متوسط وتباين المجتمع؟
- أحسب عدد العينات الممكنة ومعالم العينة $(\mu_{\bar{x}}; \delta_{\bar{x}}^2)$ (باستعمال طريقة الجداول)، وقارن بين معالم المجتمع ومعالم العينة في الحالات التالية:
- في حالة السحب بالإرجاع؟
- في حالة السحب بدون إرجاع؟

التمرين 2:

تدرس شركة طيران إمكانية السماح بحمولة يدوية للزبون مجانية، وقد وجد أن الوزن المتوسط للحمولة بالكيلوغرام هو $\mu = 5$ وانحراف معياري $\delta = 0.5$ ، إذا أخذت عينة من 100 راكب.

- ما هو المتوسط المتوقع للحمولة اليدوية $\mu_{\bar{x}}$ في العينة والانحراف المعياري $\delta_{\bar{x}}$ ؟
- أحسب احتمال أن يكون الوزن الإجمالي للأمتعة:
- أ- محصور بين 500 و 515 كلغ؟
- ب- أقل من 515 كلغ؟

التمرين 3:

1- أوجد توزيع المعاينة للمتوسط الحسابي لعينة من الحجم 3 مسحوبة من المجتمع التالي: 4، 6، 8، 5، 7، و ذلك في الحالات التالية:

- السحب بالإرجاع؟
- السحب بدون إرجاع؟

2- إذا كان لدينا متغير عشوائي x يتبع التوزيع الطبيعي بمتوسط $\mu = 80$ وتباين $\delta^2 = 49$.

- أوجد توزيع المعاينة لمتوسط العينة من الحجم 25 مسحوبة من المجتمع؟
- أحسب $P(\bar{x} > 78)$ ؟

3- إذا كان لدينا متغيران عشوائيان مستقلان X_1 و X_2 حيث X_1 يتبع التوزيع الطبيعي بمتوسط يساوي 12 و تباين يساوي 40، و X_2 يتبع التوزيع الطبيعي بمتوسط يساوي 8 و تباين يساوي 30.

- ما هو التوزيع العيني للفرق بين متوسطي عينتين الأولى من المجتمع (1) حجمها 20 والثانية من المجتمع (2) حجمها 15؟

- ما هو احتمال أن لا يزيد الفرق بين المتوسطين عن 4 ؟

التمرين 4:

- ما هو الفرق بين المعلمة والإحصاءة ؟
- ما هو الفرق بين الخطأ المعياري والانحراف المعياري ؟
- هل يمكن أن يكون الانحراف المعياري لمتوسط أو نسبة العينة في حالة المعاينة النفاذية هو نفسه في حالة المعاينة غير النفاذية ؟ علل.

التمرين 5:

مصنع ينتج كراسي تركز على قاعدة دائرية، اعتماداً على التجارب السابقة فإن مفتش الرقابة على العملية الإنتاجية مقتنع بما يلي:

- متوسط قطر القاعدة الدائرية هو 5 سم.
- الانحراف المعياري لها 0.005 سم.
- توزيع العملية الإنتاجية هو التوزيع الطبيعي.

يهتم الفاحص بالمحافظة على متوسط قطر العملية الإنتاجية عند 5 سم، ولتحقيق ذلك تسحب عينات عشوائية بصفة دورية، حجم كل منها 9 كراسي وذلك في محاولة لاكتشاف الانحرافات عن الطبيعية المشار إليها.

أ. حدد توزيع المعاينة لـ \bar{X} .

ب. بفرض أن الفاحص سحب عينة عشوائية من 9 كراسي وقيست أقطارها ووجد أن: $\bar{X} = 5.004$ سم.

- ما هي إمكانية (احتمال) أن متوسط القطر في تلك العينة العشوائية سيكون على الأقل 5.004 سم ؟

ج. ما هو حجم العينة التي يجب سحبها لتحقيق خطأ معياري لمتوسط العينة يساوي 0.001 ؟

د. في الجزء (ج)، لماذا يتفضل الفاحص أن يكون الخطأ المعياري لمتوسط العينة يساوي 0.001 على أن يكون الخطأ المعياري كما حصلت عليه في الجزء (أ).

التمرين 6:

سجلت إدارة كلية العلوم الاقتصادية بجامعة المسيلة 1000 طالبا في السنة الثانية، وبناء على نتائج السنوات السابقة تبين أن معدلات الطلبة في مادة الإحصاء 3 تتبع التوزيع الطبيعي واقتُرحت الإدارة أن تكون نسبة الطلبة المتحصلين على المعدل في مادة الإحصاء 60%.

أولاً: إذا تم اختيار فوج من قسم علوم التسيير مكون من 30 طالبا.

- فما هو توزيع المعاينة للنسبة في هذا الفوج ؟

- ما هو احتمال أن لا تقل نسبة طلبة هذا الفوج المتحصلين على المعدل في هذه المادة على 70%؟

ثانيا: إذا علمت أن إدارة قسم علوم التسيير والتي يبلغ عدد الطلبة فيها من نفس السنة 300 طالبا تقترح نفس النسبة (60%)، وباعتبار هؤلاء الطلبة يمثلون مجتمعا جديدا.

- أجب على نفس الأسئلة السابقة.

- ما هو تعليقك على النتائج؟

التمرين 7:

إذا كانت نسبة النجاح لطلبة السنة أولى في كلية العلوم الاقتصادية هي 85%، وكانت نسبة النجاح في كلية العلوم الإنسانية هي 80%، وتم سحب عينة عشوائية حجمها 100 طالب من كلية العلوم الاقتصادية وعينة أخرى حجمها 90 طالبا من كلية العلوم الإنسانية.

- أوجد احتمال أن تزيد نسبة النجاح في كلية العلوم الاقتصادية عن نسبة النجاح في كلية العلوم الإنسانية بمقدار 10% على الأكثر.

الفصل الثالث نظرية التقدير

- ✓ مصطلحات ومفاهيم
- ✓ شروط التقدير الجيد
- ✓ التقدير النقطي
- ✓ التقدير بفترة
- ✓ سلسلة تمارين محلولة
- ✓ سلسلة تمارين غير محلولة

تمهيد:

تهدف دراسة العينات وتوزيعاتها عن طريق إحصاءاتها المحددة إلى الاستدلال على المعالم المناظرة لها في المجتمع، إلا أن تلك القيم التي يتم الحصول عليها عن طريق المعاينة لا تعكس بالضرورة القيم الحقيقية المناظرة لها فيه، مما يدفع الباحث إلى تقدير تلك المعالم أو المجالات التي تقع ضمنها بدرجة ثقة معينة، لهذا سيتناول هذا الفصل شروط التقدير الجيد و أنواعه وبعض نظرياته.

1. مصطلحات ومفاهيم:

- **التقدير:** هو عملية استنتاج أو تقدير المعلمة من الإحصاءة المناظرة لها.
- **التقدير الجيد:** هو ذلك التقدير الأقرب من غيره إلى معلمة المجتمع.
- **التقدير المنحاز:** إذا ابتعدت قيمة إحصاءة العينة المقدررة على معلمة المجتمع نقول بأن التقدير متحيز.

2. شروط التقدير الجيد:

بصورة عامة فإن التقدير يعتبر جيدا إذا توفر على الخواص التالية:

1. عدم التحيز.
2. الاتساق.
3. الكفاءة.
4. الكفاية.

وفيما يلي شرح لهذه الخواص⁽¹⁾.

1-2. عدم التحيز:

يقال للتقدير $\hat{\theta}$ أنه تقدير غير متحيز للمعلمة θ إذا كانت القيمة المتوقعة للتقدير $\hat{\theta}$ تساوي قيمة

$$\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta$$

فمن خلال نظرية توزيع المعاينة وجدنا أن:

$$\bar{X} = \mu_{\bar{X}} = \mu$$

أي أن متوسط توزيع المعاينة للأوساط يساوي متوسط المجتمع فهو إذن تقدير غير متحيز، بينما

الانحراف المعياري للعينة هو تقدير متحيز للانحراف المعياري للمجتمع، حيث أن: $\sigma_{\bar{X}} \neq \sigma$.

2-2. الاتساق:

يقال بأن $\hat{\theta}$ مقدر متسق لمعلمة المجتمع إذا كانت $\hat{\theta}$ تؤول إلى θ ، (أي تقترب منها) كلما زاد حجم

العينة.

(1) محمد حسين محمد رشيد، منى عطا الله الشويلات، مرجع سابق، ص ص 410 - 415 يتصرف.

سليمان محمد طشوش، أساسيات الإحصاء الرياضي، دار اليازوري، دروب للنشر، حمادة للدراسات الجامعية، عمان، الأردن، 2012،

فمن علاقة التباين: $\sigma_x^2 = \sigma^2/n$ نجد بأن تباين العينة يقترب من الصفر عندما يقترب حجم العينة من ما لا نهاية، ولهذا نقول بأن σ_x^2 تقدير متنسق لتباين المجتمع σ^2 .

2-3. الكفاءة:

إذا كان لدينا مقدرين $\hat{\theta}_1$ ، $\hat{\theta}_2$ لمعلمة المجتمع θ وكان المقدران غير متحيزين ومتسقين، فإن الاختيار بينهما يكون على أساس التباين الأقل، حيث يتم حساب الكفاءة النسبية عن طريق العلاقة التالية:

$$\frac{\text{var}(\hat{\theta}_1)}{\text{var}(\hat{\theta}_2)} = \text{الكفاءة النسبية}$$

فإذا كانت قيمة الكفاءة النسبية أقل من الواحد الصحيح فإن $\hat{\theta}_1$ يكون مقدرًا ذي تباين أقل من $\hat{\theta}_2$ ويكون هو الأكفأ، وإذا كانت النسبة أكبر من الواحد فإن $\hat{\theta}_2$ يكون هو الأكفأ. فالتباين الأقل هو الأكثر تركيز وأقل تشتتًا وبالتالي هو الأكثر جودة.

2-4. الكفاية:

يقال بأن مقدر $\hat{\theta}$ هو مقدر كاف إذا استخدم في حساب قيمته كل معلومات المعلمة θ والتي تتوفر في العينة. فبالنسبة لمقاييس النزعة المركزية الثلاث: الوسيط والمنوال والوسط الحسابي، نجد أن الوسيط لا يستخدم في حسابه سوى قيم المفردات التي تقع في وسط العينة، كذلك الأمر بالنسبة للمنوال حيث يتم الاعتماد على القيم الأكثر شيوعًا عند حسابه، أما الوسط الحسابي فإنه يأخذ بعين الاعتبار كل القيم عند حسابه، ولهذا يعتبر المتوسط الحسابي مقدرًا كافيًا لمعلمة المجتمع μ .

3. طرق التقدير:

يمكن أن نميز بين طريقتين كلاسيكيتين للتقدير وهما التقدير النقطي والتقدير بفترة أو مجال.

3-1. التقدير النقطي:

وهي أبسط طرق التقدير، وتتمثل في تقدير المعلمة بقيمة واحدة، كأن نقول أن متوسط دخل أفراد أسرة ما هو 30000 دج، فيكون $\bar{X} = 30000$ ، هو تقدير نقطي لمتوسط المجتمع⁽¹⁾.

مثال 1: (2)

لتقدير عدد الكلمات في أحد الكتب أخذت عينة حجمها 10 صفحات عشوائيًا من ذلك الكتاب، وتم عد الكلمات فيها، فكان كما يلي: 283، 317، 303، 291، 297، 300، 305، 295، 310، 309

(1) معتوق أحمد، مرجع سابق، ص 105 بتصرف.

(2) عدنان عوض، الإحصاء التطبيقي، الشركة العربية المتحدة، القاهرة، مصر، 2009، ص ص 7 - 8.

1. كم تقدر وسط عدد الكلمات في الصفحة الواحدة ؟
2. إذا كان الكتاب يحتوي على 400 صفحة فكم تقدر عدد الكلمات فيه ؟

الحل:

$$\bar{X} = \frac{\sum X_i}{n} \text{ لدينا حجم العينة } n = 10, \text{ أما وسطها فيحسب من العلاقة:}$$

$$\bar{X} = (283 + 317 + 303 + 291 + 297 + 300 + 305 + 295 + 310 + 309)/10 = 301$$

ومنه يتم تقدير متوسط عدد الكلمات في الصفحة الواحدة بـ 301 كلمة، أي أن $\mu = 301$.
أما عدد كلمات الكتاب فهو:

$$.400 \times 301 = 120400$$

مثال 2:

من معطيات المثال السابق، قدر نقطياً قيمة الانحراف المعياري.
لقد تم الاعتماد على العلاقة $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ في حساب الانحراف المعياري لتوزيع المعاينة للأوساط الحسابية، وعند حساب الانحراف المعياري للمعاينة فيتم الاعتماد على العلاقة:

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

ولكن عند تقدير الانحراف المعياري للمجتمع فإنه يتم استخدام العلاقة:

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

فلقد تمت القسمة على $n - 1$ لأنه تم حساب أحد المقاييس من العينة اللازمة لحساب الانحراف المعياري المقدر، وهذا المقياس هو \bar{X} ، وبصفة عامة أنه عند تقدير أي مقياس يجب ملاحظة عدد المقاييس التي حسبت من العينة والتي تلزم لحساب ذلك المقياس لأخذه في الاعتبار لتحديد درجات الحرية.

$$\begin{aligned} \hat{S} &= \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} \\ &= \sqrt{\frac{(283-301)^2 + (317-301)^2 + (303-301)^2 + \dots + (309-301)^2}{n-1}} \\ &= \sqrt{\frac{989}{9}} = 9.98 \end{aligned}$$

3-2. التقدير بفترة:

عموما لا يمكن توقع الحصول على تقدير لمعلمة المجتمع بدون خطأ مهما كان التقدير جيدا، ومع أن دقة التقدير تزداد مع زيادة حجم العينة فإنه لا يوجد ما يبرر إمكانية الحصول على تقدير لمعلمة المجتمع

بدون خطأ، ولهذا يتم إعطاء مجال أو فترة يتوقع وجود معلمة المجتمع داخلها، تلك الفترة يطلق عليها فترة التقدير أو فترة الثقة، فتقدير الفترة إذن يحدد مدى معين من القيم نتوقع أن تقع ضمنه معلمة المجتمع.

3-2-1. تفسير فترة الثقة:

تعتبر فترة الثقة من الأدوات القوية التي تعطي معلومات عن المعلمة المجهولة (مثل μ) باستعمال العينة⁽¹⁾، ولذلك نعطي بعض الملاحظات حول طبيعة فترة الثقة، ونكتفي بالشرح عن فترة ثقة 95%.

مثال 3:

نفرض أننا أخذنا عينة عشوائية حجمها n من مجتمع طبيعي تباينه معلوم أي $\bar{X} \sim N(\mu, \sigma)$. من جدول التوزيع الطبيعي نجد أن:

$$P(-1.96 \leq Z \leq +1.96) = 0.95$$

إذن:

$$P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq +1.96\right) = 0.95$$

وبعد إجراء العمليات الحسابية وتبسيط هذه العلاقة نجد أن:

$$P(\bar{X} + 1.96 \sigma/\sqrt{n} \leq \pi \leq \bar{X} - 1.96 \sigma/\sqrt{n}) = 0.95$$

وهكذا نكون قد حددنا طرفي أو حدي المجال الذي يقع ضمنه المتوسط الحسابي للمجتمع، أي أن الحد

الأدنى للفترة هو $\bar{X} - 1.96 \sigma/\sqrt{n}$ بينما الحد الأعلى لها هو $\bar{X} + 1.96 \sigma/\sqrt{n}$.

1- قبل دراسة العينة وتسجيل المشاهدات وإيجاد قيمة الوسط الحسابي فإن:

$$P(\bar{X} - 1.96 \sigma/\sqrt{n}, \bar{X} + 1.96 \sigma/\sqrt{n})$$

هي فترة نهايتها متغيرتان عشوائيتان، تمثلان فترة عشوائية تحاول تقدير المعلمة المجهولة μ .

2- إن تفسير الاحتمال:

$$P(\bar{X} - 1.96 \sigma/\sqrt{n} \leq \mu \leq \bar{X} + 1.96 \sigma/\sqrt{n}) = 0.95$$

على أنه التكرار النسبي لمحاولات المعاينة الكثيرة المتكررة يحدد أن 95% من فترات الثقة ستحتوي

μ ، وأن 5% منها لا تحويها.

(1) محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء، دار المسيرة، الأردن، ط2، 2005، ص ص 212 - 214

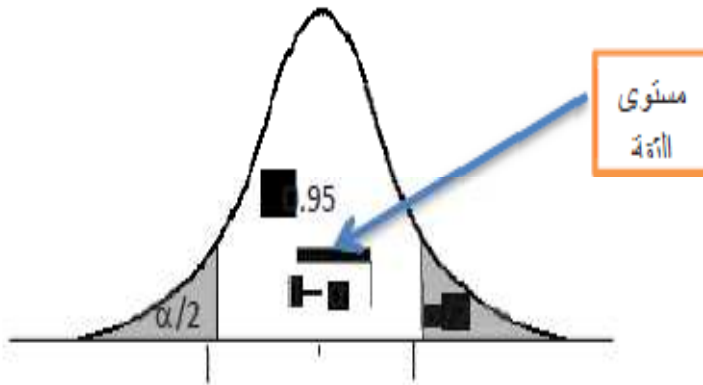
3- حالما نحسب المتوسط الحسابي من العينة ونجد قيمته \bar{X} فإن فترة الثقة:
 $(\bar{X} - 1.96 \sigma/\sqrt{n}, \bar{X} + 1.96 \sigma/\sqrt{n})$ تسمى فترة ثقة 95% للمعلمة μ .

4- وعند التطبيق فنحن لا نعلم فيما إذا كانت فترة الثقة 95% التي حصلنا عليها من عينة معينة، تحوي القيمة المجهولة μ أو لا تحويها، ولكن اعتمادا على التكرار النسبي فقد أمكننا استعمال كلمة ثقة بعد حساب الفترة.

وبأخذ هذه الملاحظات بعين الاعتبار فإن تفسير فترة الثقة 95% للمعلمة μ يكون كما يلي : إذا أخذنا 100 عينة عشوائية ذات الحجم n وفي كل مرة نحسب \bar{X} وتحسب فترة الثقة فإننا نتوقع أن 95 فترة تحوي μ الحقيقية، و5 فترات فقط تخطئ μ الحقيقية فلا تحتويها.

ويمكن توضيح مساحة فترة الثقة وحديها في مثالنا هذا بيانيا في الشكل الموالي:

الشكل 3-1: مساحة فترة الثقة وحديها عند مستوى معنوية $1 - \alpha$



ونكتب: $1 - \alpha = p = 0.95 \Rightarrow \alpha = 0.05$

أي أن: مستوى المعنوية + مستوى الثقة = 1

ويرمز عادة بالرمز α لمستوى المعنوية وبـ $(1 - \alpha)$ لمستوى الثقة.

ويمكن كتابة العلاقة التي تحدد فترة الثقة في العبارة التالية:

فترة الثقة = التقدير بنقطة \pm معامل الثقة \times الخطأ المعياري للتقدير

$$\theta = T \mp C_C \sigma^T$$

حيث C_C يدعى معامل الثقة، وفي مثالنا هذا قيمته هي: $Z_{1-\alpha/2} = 1.96$

والجدول الموالي يوضح أهم مستويات الثقة ومعاملاتها في حالة استخدام التوزيع الطبيعي في التقدير:

الجدول 3-1: أهم مستويات الثقة ومعاملاتها في حالة استخدام التوزيع الطبيعي

مستويات الثقة				الرمز	
0.90	0.95	0.98	0.99	$1 - \alpha$	مستوى الثقة
0.10	0.05	0.02	0.01	α	مستوى المعنوية
0.45	0.475	0.49	0.495	$0.5 - \alpha / 2$	الاحتمال
1.645	1.96	2.33	2.58	$Z_C = Z_{0.5-\alpha/2}$	معامل الثقة

وتلخص خطوات إيجاد حدود فترة الثقة في ما يلي:

1. تعيين مستوى الثقة.
2. استنتاج معاملات الثقة بالنسبة للمستويات الشهيرة مثل: (0.99، 0.95، 0.90)، أما في الحالات الأخرى فيتم حساب المساحة (الاحتمال) بطرح نصف مستوى المعنوية من 0.5 أو 1 وذلك حسب جدول التوزيع الطبيعي المستعمل وإيجاد قيمة معامل الثقة المقابلة للاحتمال.
3. حساب حدود مجال الثقة.

وسيتم تنفيذ هذه الخطوات عند حساب فترة الثقة في الحالات الموالية:

3-2-2. فترة الثقة لمتوسط مجتمع طبيعي تباينه معلوم:

المثال السابق المتعلق بشرح فترة الثقة يندرج ضمن هذه الحالة التي تشمل على النظرية الموالية: إذا أخذت عينة عشوائية حجمها n من مجتمع طبيعي $N(\mu, \sigma^2)$ ، وكانت σ^2 معلومة. فإن فترة ثقة $(1 - \alpha)$ للمعلمة μ هي:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} \right]$$

حيث:

\bar{X} : متوسط العينة.

$(1 - \alpha)$: مستوى الثقة.

$Z_{0.5-\frac{\alpha}{2}}$: معامل الثقة.

مثال 4:

شركة مختصة في صناعة العجلات المطاطية تخضع أقطارها للتوزيع الطبيعي بانحراف معياري 1 سم، أخذت عينة عشوائية حجمها 16 عجلة فكان الوسط الحسابي لأقطار هذه العجلات هو 40 سم، أوجد فترة ثقة 99% لمعدل أقطار العجلات التي تنتجها الشركة؟

الحل:

متوسط العينة هو $\bar{X} = 40$ ، ومنه تقدير μ هو 40.

1. مستوى الثقة هو 99%.

$$2. \quad 1 - \alpha = 0.99 \Rightarrow \frac{\alpha}{2} = 0.005 \Rightarrow 0.5 - 0.005 = 0.4950$$

ومنه معامل الثقة هو: $Z_{0.5-\frac{\alpha}{2}} = 2.58$

3. فترة الثقة هي:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} \right]$$

بعد التعويض نجد:

$$\left[40 - 2.58 \times \frac{1}{\sqrt{16}} ; 40 + 2.58 \times \frac{1}{\sqrt{16}} \right]$$

أي أن فترة الثقة هي: [39.355 ; 40.25]

3-2-3. فترة الثقة للمتوسط في حالة المجتمعات الكبيرة:

نستطيع من خلال النظرية السابقة تقدير فترة الثقة للمتوسط إذا كانت المعاينة من مجتمع طبيعي

تباينه معلوم، لكن إذا لم يكن المجتمع خاضعا للتوزيع الطبيعي، فكيف يتم تقدير تلك الفترة؟

إن نظرية النهاية المركزية تعطينا أيضا الحل عندما يتوفر شرط كبر حجم العينة⁽¹⁾، ولذلك نستطيع

استعمال النظرية الموالية:

نظرية: إذا أخذت عينة عشوائية حجمها n من مجتمع ليس بالضرورة طبيعيا $N(\mu, \sigma^2)$ وكانت σ^2 معلومة،

وحجم العينة كبير ($n \geq 30$) فإن فترة ثقة $(1 - \alpha)$ للمعلمة μ هي تقريبا:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \sigma / \sqrt{n} \right], n \geq 30$$

(1) د. ليونارد، ج. كازمير، ترجمة مصطفى جلال مصطفى، الإحصاء التجاري، الدار الدولية للاستثمارات الثقافية، مصر، 2004، ص 84 - 85 بتصرف.

مثال 5:

أخذت عينة عشوائية حجمها 100، متوسطها الحسابي 52 من مجتمع تباينه 25. أوجد فترة ثقة 98% لوسط المجتمع μ ؟

الحل:

متوسط العينة هو: $\bar{X} = 52$, $n = 100$, $1 - \alpha = 0.98$, $\sigma^2 = 25$
المجتمع مجهول التوزيع أي أن: $X \sim ?$

وما أن حجم العينة أكبر من 30 فسيتم الاعتماد على نظرية النهاية المركزية، ونحصل على تقريب لفترة الثقة كما يلي:

1. مستوى الثقة هو 98%.

$$1 - \alpha = 0.98 \Rightarrow \frac{\alpha}{2} = 0.01 \Rightarrow 0.5 - 0.01 = 0.4900 \quad 2.$$

ومنه معامل الثقة هو: $Z_{0.5-\frac{\alpha}{2}} = 2.33$

3. فترة الثقة هي:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \sigma/\sqrt{n} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \sigma/\sqrt{n} \right]$$

وبعد التعويض نجد:

$$\left[52 - 2.33 \times \frac{5}{\sqrt{100}} ; 52 + 2.33 \times \frac{5}{\sqrt{100}} \right]$$

أي أن فترة الثقة هي: [50.835 ; 53.165]

3-2-4. فترة الثقة للمتوسط في العينات الكبيرة من مجتمعات مجهولة التباين:

غالبًا ما يكون تباين المجتمع σ^2 مجهولًا وحجم العينة كبيرًا، لذلك نستعمل الانحراف المعياري للعينة S بدل الانحراف المعياري للمجتمع⁽¹⁾، وفي هذه الحالة تكون فترة الثقة:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right]$$

مثال 6:

أخذت عينة من مصابيح كهربائية حجمها 64، ووجد بأن متوسط أعمارها الاقتصادي 800 ساعة بانحراف معياري 40 ساعة. أوجد فترة ثقة 95% لمعدل أعمار المصابيح ؟.

(1) سالم عيسى بدر، عماد غصاب عبابنة، مبادئ الإحصاء الوصفي والاستدلالي، دار المسيرة، عمان، الأردن، 2007، ص 284

الحل:

لدينا:

$$n = 64 ; \bar{X} = 800 ; S = 40 ; 1 - \alpha = 0.95. X \sim ?$$

بما أن حجم العينة كبير فإنه يمكن الاستناد إلى نظرية النهاية المركزية والاعتماد على الانحراف المعياري للعينة في تحديد مجال الثقة لمتوسط أعمار المصابيح عند مستوى 95% كما يلي:

$$1 - \alpha = 0.95 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 0.5 - 0.025 = 0.4750 \quad .2$$

$$Z_{0.5-\frac{\alpha}{2}} = 1.96 \quad \text{ومنه معامل الثقة هو:}$$

3. فترة الثقة هي:

$$\left[\bar{X} - Z_{0.5-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} ; \bar{X} + Z_{0.5-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right]$$

وبعد التعويض نجد:

$$\left[800 - 1.96 \frac{40}{\sqrt{64}} ; 800 + 1.96 \frac{40}{\sqrt{64}} \right]$$

أي أن فترة الثقة هي: [790.2 ; 809.8]

3-2-5. فترة الثقة للمتوسط في العينات الصغيرة من مجتمعات مجهولة التباين:

إن تحديد فترة الثقة للمتوسط في العينات الصغيرة والمأخوذة من مجتمعات طبيعية مجهولة التباين يقودنا إلى الحديث على توزيع احتمالي يدعى توزيع ستيودنت.

3-2-5-1. توزيع ستيودنت:

عندما يكون حجم العينة صغيرا وتباين المجتمع σ^2 غير معلوم، فإنه لا يمكن تقريب $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ بالتوزيع الطبيعي المعياري، ولكن عندما تكون المعاينة من مجتمع طبيعي فإن توزيع $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ هو توزيع ستيودنت (1).

تعريف:

توزيع ستيودنت هو أحد التوزيعات الاحتمالية المتصلة تكمن تطبيقاته في نظرية العينات واختبار الفرضيات كبديل للتوزيع الطبيعي في العينات الصغيرة المسحوبة من مجتمعات مجهولة التباين، ودالة كثافته الاحتمالية تعطى بالعلاقة التالية:

(1) محمد حسين محمد رشيد، منى عطا الله الشويلات، مرجع سابق، ص ص 368 - 369 بتصرف.

معنوق أحمد، مرجع سابق، ص ص 40 - 41 بتصرف.

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

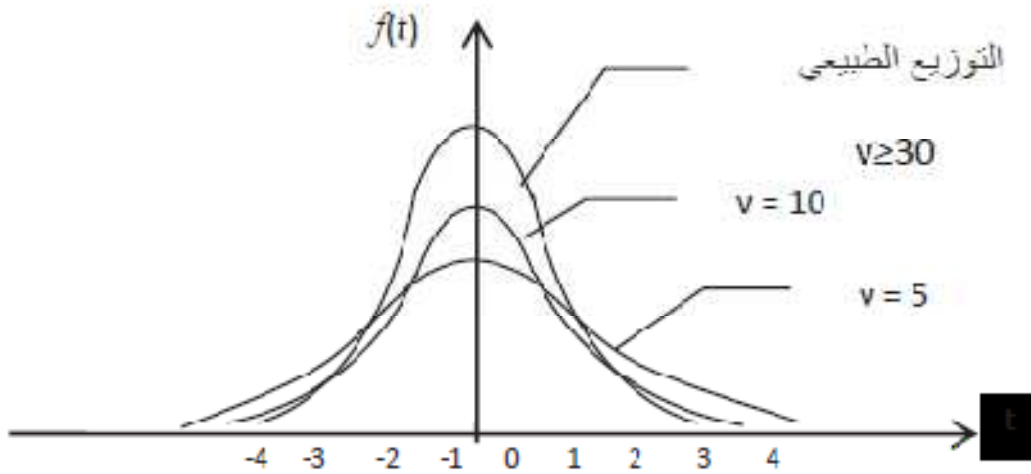
حيث تسمى v درجات الحرية، و c ثابت يعتمد على v ليجعل المساحة تحت المنحني تساوي 1.

$$T \sim t_{1-\frac{\alpha}{2}, v}$$

3-2-5-2. خواص توزيع t

1. شكله يشبه الجرس.
2. أحادي المنوال، حيث له قيمة واحدة تقابل ($t = 0$)
3. شكله يشبه التوزيع الطبيعي إلا أنه منخفض عنه.
4. يتقارب طرفيه من الصفر عندما $t \rightarrow \mp \infty$ ، إلا أنه أبداً من تقارب طرفي التوزيع الطبيعي.
5. يعتمد توزيع t على حجم العينة، فعندما يقل حجم العينة يزداد ارتفاع منحنى t عن المحور الأفقي، وكلما زاد حجم العينة اقترب طرفي المنحنى من المحور الأفقي، وإذا كان حجم العينة أكبر من أو يساوي 30 فيؤول توزيع t إلى التوزيع الطبيعي والشكل الموالي يوضح ذلك.

الشكل 3-2: تقريب توزيع t إلى التوزيع الطبيعي



أما كيفية إيجاد قيمة t باستخدام توزيع ستودنت فيتم كما يلي:

1. تحديد درجة الحرية v انطلاقاً من حجم العينة عن طريق العلاقة: $v = n - 1$.
2. تحديد مستوى الثقة المراد ($1 - \alpha$) واستنتاج المساحة $\left(1 - \frac{\alpha}{2}\right)$ والبحث عليها في السطر الأفقي أعلى الجدول.
3. التقاطع بين قيمة درجات الحرية v في العمود الأيسر والمساحة $\left(1 - \frac{\alpha}{2}\right)$ في السطر العلوي يحددان قيمة t .

وبعد إيجاد القيمة الجدولية لمعامل الثقة $t_{1-\frac{\alpha}{2};v}$ يتم تحديد فترة الثقة بالاعتماد على النظرية التالية:

نظرية:

إذا أخذت عينة عشوائية صغيرة حجمها n من مجتمع طبيعي فإن فترة ثقة $(1 - \alpha)$ للمعلمة μ هي:

$$\left[\bar{X} - t_{1-\frac{\alpha}{2};v} \frac{S}{\sqrt{n}} ; \bar{X} + t_{1-\frac{\alpha}{2};v} \frac{S}{\sqrt{n}} \right]$$

\bar{X} : المتوسط الحسابي المحسوب من العينة.

n : حجم العينة.

v : درجة الحرية = $n - 1$.

α : مستوى المعنوية.

S : الانحراف المعياري المحسوب من العينة.

μ : متوسط المجتمع المجهول.

مثال 7:

أخذت عينة عشوائية حجمها 16 من مجتمع طبيعي فوجد أن متوسطها الحسابي 15.5 وانحرافها المعياري 2.4 ، أوجد فترة ثقة 95% لمعدل المجتمع ؟.

الحل:

لدينا: $n = 16, \bar{X} = 15.5 ; S = 2.4 ; 1 - \alpha = 0.95. X \sim N (\mu, ?)$

مجتمع طبيعي
حجم العينة صغير
تباين المجتمع مجهول
1. مستوى الثقة هو 95%.

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{1-\frac{0.05}{2};16-1}$$

$$1 - \alpha = 0.95 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - 0.025 = 0.975 \quad 2.$$

$$t_{1-\frac{0.05}{2};16-1} = 2.131 \quad \text{ومنه معامل الثقة هو:}$$

3. فترة الثقة هي:

$$\left[\bar{X} - t_{1-\frac{\alpha}{2};v} \frac{S}{\sqrt{n}} ; \bar{X} + t_{1-\frac{\alpha}{2};v} \frac{S}{\sqrt{n}} \right]$$

وبعد التعويض نجد:

$$\left[15.5 - 2.131 \times \frac{2.4}{\sqrt{16}} ; 15.5 + 2.131 \times \frac{2.4}{\sqrt{16}} \right]$$

أي أن فترة الثقة هي: [14.221 ; 16.778]

مثال 8:

بلغ عدد أفواج طلبة السنة الأولى جذع مشترك في كلية العلوم الاقتصادية 40 فوجا، وكانت معدلات الطلبة لـ 5 أفواج من السنة الأولى جذع مشترك في مادة الإحصاء كالاتي:

$$10.2 ; 9.8 ; 9.9 ; 10.3 ; 10.1$$

1. قدر نقطيا متوسط معدلات الطلبة للأفواج الخمسة.
2. أوجد فترة ثقة 90% لمعدلات أفواج الطلبة من تلك السنة بافتراض أنها تتبع التوزيع الطبيعي.

الحل:

$$\text{لدينا: } n = 5, \bar{X} = ? ; S = ? ; 1 - \alpha = 0.95. X \sim N(\mu, ?)$$

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{1-\frac{\alpha}{2}; n-1} \quad \left\{ \begin{array}{l} \text{مجتمع طبيعي} \\ \text{حجم العينة صغير} \\ \text{تباين المجتمع مجهول} \end{array} \right.$$

1. التقدير النقطي لمتوسط معدلات الأفواج الخمسة في مادة الإحصاء:

$$\bar{X} = \frac{\sum X_i}{n} \quad \text{لدينا حجم العينة } n = 5, \text{ أما وسطها فيحسب من العلاقة:}$$

$$\bar{X} = (10.1 + 10.3 + 9.9 + 9.8 + 10.2)/5 = 10.06$$

ومنه يتم تقدير متوسط معدلات أفواج السنة الأولى في مادة الإحصاء بـ 10.06، أي أن $\mu =$

10.06.

2- تحديد فترة الثقة:

عند تقدير الانحراف المعياري للمجتمع فإنه يتم استخدام العلاقة:

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

$$= \sqrt{\frac{(10.1-10.06)^2 + (10.3-10.06)^2 + (9.9-10.06)^2 + (9.8-10.06)^2 + (10.2-10.06)^2}{5-1}}$$

$$= \sqrt{\frac{0.172}{4}} = 0.207$$

وحيث أن قيمة معدل الاستقصاء هي: $\frac{n}{N} = \frac{5}{40} = 0.125 > 0.05$

فإنه يتم اعتماد معامل الإرجاع في علاقة الانحراف المعياري كما هو موضح في فترة الثقة في

الخطوة الثالثة من خطوات تحديد فترة الثقة.

1. مستوى الثقة هو 90%.

$$1 - \alpha = 0.90 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1 - 0.05 = 0.95 \quad .2$$

ومنه معامل الثقة هو: $t_{1-\frac{0.10}{2}; 5-1} = 2.132$

3. فترة الثقة هي:

$$\left[\bar{X} - t_{1-\frac{\alpha}{2}; v} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}; \bar{X} + t_{1-\frac{\alpha}{2}; v} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \right]$$

وبعد التعويض نجد:

$$\left[10.06 - 2.132 \frac{0.207}{\sqrt{5}} \sqrt{\frac{40-5}{40-1}}; 10.06 + 2.132 \frac{0.207}{\sqrt{5}} \sqrt{\frac{40-5}{40-1}} \right]$$

$$[10.06 - 2.132 (0.09)(0.947); 10.06 + 2.132 (0.092)(0.947)]$$

أي أن فترة الثقة هي: [9.874 ; 10.245]

3-3. تقدير النسبة:

إن عملية تقدير نسبة وجود خاصية أو صفة في مجتمع ما عادة ما تتم من خلال تقدير نسبة تلك الخاصية أو الصفة في عينة عشوائية مأخوذة من هذا المجتمع، فتقدير النسبة أيضا إما أن يكون نقطيا أو بفترة.

3-3-1. تقدير النسبة بنقطة:

رأينا في الفصل السابق من توزيع المعاينة للنسبة أن نسبة النجاح في العينة تقابل نفس النسبة من النجاح في المجتمع، ولهذا صار تقدير نسبة وجود ظاهرة في مجتمع ما يتم على أساس نسبة وجود تلك الظاهرة في عينة عشوائية مأخوذة من ذلك المجتمع.

فإذا أردنا تقدير نسبة الطلبة الذين يمتلكون الحاسوب النقال في الجامعة الجزائرية فإنه يمكن أن نختار كلية من جامعة ما عشوائيا ثم نقوم بحساب نسبة الطلبة الذين يمتلكون الحاسوب النقال بالكلية، وتستعمل تلك النسبة في العينة كتقدير نقطي للنسبة في الجامعة الجزائرية.

فإذا كانت لدينا عينة عشوائية مأخوذة من مجتمع برنولي $B(1, P)$ ، وكانت قيم مفرداتها هي: $X_1, X_2,$

$$X_3, \dots, X_n \text{ فإن نسبة النجاح في العينة تكون كما يلي: } p' = \frac{\sum X_i}{n}$$

وتمثل تلك النسبة في العينة التقدير النقطي للنسبة في المجتمع⁽¹⁾.

(1) محمد صبحي أبو صالح، عدنان محمد عوض، مرجع سابق، ص ص 222 - 223 بتصرف.

مثال 9:

لتقدير نسبة المرضى الراضين عن الخدمات المقدمة لهم من طرف الطاقم الطبي في إحدى المستشفيات تمت مقابلة 80 مريضاً فوجد من بينهم 50 مريضاً عبروا عن رضاهم عن نوع تلك الخدمات.
- ما هي نسبة المرضى الراضين في ذلك المستشفى؟

الحل:

نسبة المرضى الراضين عن الخدمات الصحية في العينة هي:

$$p' = \frac{50}{80} = 0.625 = 62.5\%$$

3-3-2. تقدير النسبة بفترة:

رغم أن التوزيع الاحتمالي المستعمل في النسب هو التوزيع الثنائي إلا أن كثرة الحسابات وتعقدها عند تحديد فترة ثقة لنسبة المجتمع غير المعروفة جعل العديد من المراجع تلجأ إلى استخدام التوزيع الطبيعي كتقريب للتوزيع الثنائي، وهذا عندما تكون العينات كبيرة الحجم وعندما لا تكون نسبة النجاح قريبة جداً من الصفر أو الواحد والنظرية الموالية توضح ذلك⁽¹⁾:

نظرية:

إذا كانت $p' = \frac{x}{n}$ نسبة النجاح في عينة عشوائية حجمها n وكانت n كبيرة فإن فترة الثقة $(1 - \alpha)$ التقريبية لنسبة النجاح هي كما يلي:

ملاحظات:

1 - إذا كان المجتمع محدوداً و N حجم N وقيمة معدل الاستقصاء أقل من 0.05 فإنه يتم ضرب قيمة الانحراف المعياري في جذر معامل الإرجاع وتكون عبارة فترة الثقة للقسم كما يلي:

$$\left[p' - Z_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} \sqrt{\frac{N-n}{N-1}} ; p' + Z_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} \sqrt{\frac{N-n}{N-1}} \right]$$

2 - إن تقريب التوزيع من الثنائي إلى الطبيعي يؤدي إلى وجود خطأ، ولهذا يقترح استخدام معامل التصحيح $\frac{1}{2n}$ عند حساب قيمة الاحتمال، بإضافته إلى يمين مترابحة الاحتمال في حالة حساب الاحتمال لما يكون أكبر، وطرحه لمل يكون اتجاه مترابحة الاحتمال أقل.

(1) شيزل، شيلر، سرينيفيسان، ترجمة محمد علي عبد الناصر، مصطفى جلال مصطفى، الاحتمالات والإحصاء، الدار الدولية للاستشارات الثقافية، مصر، 2004، ص 91 بتصرف.

مثال 10:

أخذت عينة عشوائية حجمها 200 طالبا من كلية العلوم الاقتصادية من قسم الجذع المشترك فوجد أن 40 منهم متحصلون على شهادة البكالوريا تخصص علوم الطبيعة والحياة.

1. قدر نسبة طلبة الجذع المشترك الحاصلين على شهادة البكالوريا تخصص علوم الطبيعة والحياة .
2. أوجد فترة ثقة 99% للنسبة الحقيقية للطلبة الحاصلين على شهادة البكالوريا شعبة علوم الطبيعة والحياة.

الحل:

1 - تقدير نسبة طلبة الجذع المشترك الحاصلين على شهادة البكالوريا تخصص علوم الطبيعة والحياة:

$$p' = \frac{40}{200} = 0.2 = 20\%$$

2 - إيجاد فترة الثقة:

بما أن حجم العينة أكبر من 30 فإنه يمكن التقريب للتوزيع الطبيعي وسنتبع نفس الخطوات السابقة في تحديد فترة الثقة كما يلي:

1. مستوى الثقة هو 99%.

$$2. \quad 1 - \alpha = 0.90 \Rightarrow \frac{\alpha}{2} = 0.005 \Rightarrow 0.5 - 0.005 = 0.4950$$

ومنه معامل الثقة هو: $Z_{0.5-\frac{\alpha}{2}} = 2.58$

3. فترة الثقة هي:

$$\left[p' - Z_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} ; p' + Z_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} \right]$$

وبعد التعويض نجد:

$$\left[0.2 - 2.58 \sqrt{\frac{0.2(1-0.2)}{200}} ; 0.2 + 2.58 \sqrt{\frac{0.2(1-0.2)}{200}} \right]$$

أي أن فترة الثقة هي: [0.127 ; 0.273]

3-3-3. تحديد حجم العينة اللازم لتقدير فترة النسبة P:

قبل الشروع في اختيار العينة المناسبة فإنه يجب تحديد الحد الأدنى المطلوب من حجم العينة للحصول على فترة ثقة بطول معلوم أو بخطأ محدد لا نريد تجاوزه.

فعند توفر شروط التقريب للتوزيع الطبيعي فإن طول فترة الثقة هو: $Z^2_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}}$

ويمكن أن نميز بين حالتين⁽¹⁾:

الحالة الأولى: إذا كانت قيمة النسبة P معلومة

نرمز لقيمة الخطأ المسموح به عند مستوى ثقة $(1 - \alpha)$ بالرمز d ويكون لدينا:

$$Z_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} \leq d \Rightarrow Z^2_{0.5-\frac{\alpha}{2}} \sqrt{\frac{p'(1-p')}{n}} \leq d^2$$

$$\Rightarrow n \geq \frac{Z^2_{0.5-\frac{\alpha}{2}} p'(1-p')}{d^2}$$

الحالة الثانية: إذا كانت قيمة النسبة P غير معلومة

إذا لم تكن لدينا أي فكرة عن قيمة النسبة وأردنا تحديد الحد الأدنى لحجم العينة عند مستوى ثقة $(1 - \alpha)$ فإننا نعطي للقيمة $p'(1-p')$ أعظم قيمة لها من خلال أكبر قيمة يمكن أن تأخذها النسبة وهي $p' = \frac{1}{2}$ ،

فتكون أعظم قيمة لـ $p'(1-p')$ هي $\frac{1}{4}$ وبالتالي يتحدد حجم العينة كما يلي:

$$n \geq \frac{Z^2_{0.5-\frac{\alpha}{2}}}{4d^2}$$

مثال 11:

نريد تقدير نسبة الطلبة الذين يتم توجيههم إلى السنة الثانية في كلية العلوم الاقتصادية وفقا لرغباتهم.

- ما هو عدد الطلبة الذين يجب اختيارهم كي نكون واثقين بنسبة 98% أن الخطأ في تقدير هذه

النسبة لا يزيد عن 0.04 في كل حالة مما يلي:

1. إذا لم نكن على علم سبق عن هذه النسبة.

2. إذا كانت نسبة أولئك الطلبة هي 80%.

الحل:

الحالة الأولى:

$$1 - \alpha = 0.98 \Rightarrow \frac{\alpha}{2} = 0.01 \Rightarrow 0.5 - 0.01 = 0.4900 \quad \text{لدينا:}$$

$$Z_{0.5-\frac{\alpha}{2}} = 1.645 \quad \text{ومنه معامل الثقة هو:}$$

كما لدينا أيضا: $d = 0.04$ (نصف طول الفترة)

- بما أن النسبة غير معلومة فإننا نأخذ أسوأ حالة وهي لما تكون قيمتها هي: 0.5 ويكون عدد

الطلبة هو:

$$\begin{aligned} n &\geq \frac{Z^2_{0.5-\frac{\alpha}{2}}}{4d^2} \Rightarrow n \geq \frac{1.645^2}{4(0.04)^2} \\ &\Rightarrow n \geq 164.5 \\ &\Rightarrow n = 165 \end{aligned}$$

(1) عدنان عوض، مرجع سابق، ص ص 23 - 25 بتصرف.

الحالة الثانية:

بما أن النسبة P معلومة فإن حجم العينة يحسب كما يلي:

$$n \geq \frac{z^2_{0.5-\frac{\alpha}{2}}}{d^2} p'(1-p') \Rightarrow n \geq \frac{1.645^2}{0.04^2} 0.8(1-0.8)$$

$$\Rightarrow n \geq 270.6$$

$$\Rightarrow n = 271$$

3-4. فترات الثقة حول تباين المجتمع:

رأينا فيما سبق عند تقدير المتوسط الحسابي في حالة العينات الصغيرة المأخوذة من مجتمعات مجهولة التباين أن التباين المقدر يكون وفق العلاقة التالية:

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

وهذا ما يمهد لدراسة أحد التوزيعات الاحتمالية المتصلة والمتمثل في توزيع كاي تربيع.

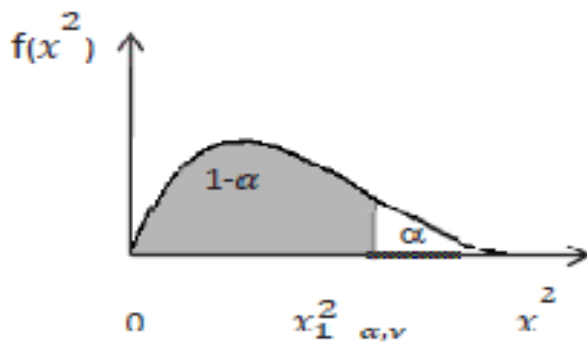
3-4-1. توزيع كاي تربيع:

إذا تم سحب عينات عشوائية حجمها n من مجتمع طبيعي تباينه σ^2 وتم حساب تباين توزيع المعاينة S^2 لكل عينة ⁽¹⁾، فإنه يتم الحصول على قيم للإحصاءة S^2 ، ولندرة استخدام توزيع المعاينة لـ S^2 في الواقع وبدلاً من ذلك سوف نستخدم توزيع المتغير العشوائي x^2 ، ويطلق عليه مربع كاي، ويمكن حساب قيمته من كل عينة بالمعادلة:

$$x^2 = \frac{(n-1)S^2}{\sigma^2}$$

ويطلق على توزيع هذا المتغير توزيع مربع كاي بدرجة حرية $(n-1)$ ، حيث يرمز لها بالرمز v والشكل الموالي يوضح منحنى كاي تربيع:

الشكل 3-3: منحنى كاي تربيع



(1) جورج كانافوس، دون ميلر، ترجمة سلطان محمد عبد المجيد، محمد توفيق البلقيني، مرجع سابق، ص ص 302 - 305 بتصرف

ولإيجاد المساحات تحت المنحنى أو إيجاد القيم التي يقع إلى يسارها أو يمينها مساحة معينة نستعمل جدول كاي تربيع حيث يمثل العمود الأيسر درجات الحرية والخط الأفقي المساحات إلى يسار قيمة كاي تربيع، أما قيم كاي تربيع فهي في داخل الجدول، ونستخدم التعبير $x_{1-\alpha, v}^2$ ليعبر عن قيمة كاي تربيع عند مستوى ثقة $(1 - \alpha)$ بدرجة حرية v .

3-4-2. خواص توزيع كاي تربيعي:

يلاحظ بأن المتغير X^2 يمكن أن يكون سالبا، وبالتالي لا يمتد إلى يسار الصفر وإذا كانت درجات الحرية أكبر من 2 فإن لتوزيع x^2 منوال واحد ويكون ملتويا نحو اليمين وكلما زادت درجات الحرية قل الالتواء.

مثال 12:

إذا كان x^2 يخضع لتوزيع كاي تربيع في عينة حجمها 16، أوجد:

1. قيمة X^2 التي يكون إلى يسارها 0.95 من المساحة، أي: $p(X^2 \leq x^2) = 0.95$

2. قيمة X^2 التي يكون إلى يمينها 0.05 من المساحة، أي: $p(X^2 > x^2) = 0.05$

3. قيمة X^2 التي يكون إلى يسارها 0.99 من المساحة، أي: $p(X^2 \leq x^2) = 0.99$

4. قيمة X^2 التي يكون إلى يمينها 0.01 من المساحة، أي: $p(X^2 \leq x^2) = 0.01$

الحل:

$$N = 20 \Rightarrow v = 20 - 1 = 19$$

فدرجة الحرية إذن هي 19 أما قيم كاي تربيع في كل حالة هي كما يلي:

1. قيمة X^2 التي يكون إلى يسارها 0.95 من المساحة هي: $x_{0.95, 19}^2 = 30.144$. 2. قيمة X^2 التي

يكون إلى يمينها 0.05 من المساحة هي نفسها قيمة كاي تربيع إلى يسار 0.95 أي أن $(p(X^2 > x^2) = 0.05) = (p(X^2 \leq x^2) = 0.95)$

ونكتب أيضا: $x_{0.95, 19}^2 = 30.144$

3. قيمة X^2 التي يكون إلى يسارها 0.99 من المساحة هي: $x_{0.99, 19}^2 = 36.191$

4. قيمة X^2 التي يكون إلى يمينها 0.01 من المساحة هي نفسها قيمة كاي تربيع إلى يسار 0.99 أي

أن $(p(X^2 > x^2) = 0.01) = (p(X^2 \leq x^2) = 0.99)$

ونكتب أيضا: $x_{0.99, 19}^2 = 36.191$

نظرية:

إذا كانت $X_1, X_2, X_3 \dots X_n$ عينة عشوائية من توزيع طبيعي متوسطه μ وتباينه σ^2 فإن فترة ثقة $(1 - \alpha)$

α هي:

$$\left[\frac{(n-1)S^2}{x^2_{1-\frac{\alpha}{2}; n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{x^2_{\frac{\alpha}{2}; n-1}} \right]$$

حيث: S^2 هو تباين العينة أي: $S^2 = \sum(X_i - \bar{X})^2 / n - 1$

ملاحظة:

لإيجاد فترة ثقة $(1 - \alpha)$ للتباين S^2 في النظرية السابقة، نجد أنه من خلال جدول توزيع x^2 فإن

النقطتين $x^2_{1-\frac{\alpha}{2}; n-1}, x^2_{\frac{\alpha}{2}; n-1}$ تحصران بينهما مساحة $(1 - \alpha)$ تحت توزيع x^2 بدرجة حرية $(n - 1)$ ، أي أن:

$$P\left(x^2_{\frac{\alpha}{2}; n-1} \leq \frac{(n-1)S^2}{\sigma^2} \leq x^2_{1-\frac{\alpha}{2}; n-1}\right) = 1 - \alpha$$

وبعد القيام بالعمليات الرياضية والتبسيط الجبري على المقدار نجد:

$$x^2_{\frac{\alpha}{2}; n-1} \leq \frac{(n-1)S^2}{\sigma^2} \leq x^2_{1-\frac{\alpha}{2}; n-1}$$

نأخذ مقلوب المتباينة فيكون عندنا:

$$\frac{1}{x^2_{1-\frac{\alpha}{2}; n-1}} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{x^2_{\frac{\alpha}{2}; n-1}}$$

وبعد الضرب في $(n - 1)S^2$ نحصل على:

$$\frac{(n-1)S^2}{x^2_{1-\frac{\alpha}{2}; n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{x^2_{\frac{\alpha}{2}; n-1}}$$

مثال 13:

أخذت عينة عشوائية حجمها 10 فوجد أن متوسطها الحسابي 0.7 وانحرافها المعياري 0.4، أوجد

فترة ثقة 90% للانحراف المعياري.

الحل:

$$n = 10 ; S = 0.4 ; 1 - \alpha = 0.90$$

1. مستوى الثقة هو 90%.

$$1 - \alpha = 0.90 \Rightarrow \begin{cases} \frac{\alpha}{2} = 0.005 \\ 1 - \frac{\alpha}{2} = 0.95 \end{cases} \quad .2$$

ومنه قيمتي كاي تربيع عند درجة حرية (10 - 1) هما:

$$\begin{cases} x_{1-\frac{\alpha}{2};n-1}^2 = x_{0.95;9}^2 = 16.91 \\ x_{\frac{\alpha}{2};n-1}^2 = x_{0.05;9}^2 = 3.325 \end{cases}$$

3. فترة الثقة هي:

$$\left[\frac{(n-1)S^2}{x_{1-\frac{\alpha}{2};n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{x_{\frac{\alpha}{2};n-1}^2} \right]$$

وبعد التعويض نجد:

$$\frac{(10-1)0.4^2}{16.91} \leq \sigma^2 \leq \frac{(10-1)0.4^2}{3.325}$$

أي أن: تباين المجتمع محصور بين: [0.085 ; 0.433]

ومنه فترة ثقة 90% للانحراف المعياري للمجتمع هي: [0.29 ; 0.66].

3-5. فترات الثقة للنسبة بين تباينين:

إذا كانت $X_1, X_2, X_3 \dots X_n$ عينة عشوائية من مجتمع $N(\mu_1, \sigma_1^2)$ وكانت $Y_1, Y_2, Y_3 \dots Y_n$ عينة عشوائية من مجتمع $N(\mu_2, \sigma_2^2)$ مستقل عن الأول فإن فترة الثقة $(1 - \alpha)$ للنسبة σ_1^2/σ_2^2 هي ⁽¹⁾:

$$\frac{S_1^2}{S_2^2} \frac{1}{F_{1-\frac{\alpha}{2}, v_1, v_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} F_{1-\frac{\alpha}{2}, v_2, v_1}$$

مثال 14:

أخذت عينة عشوائية حجمها 9 من مجتمع $N(\mu_1, \sigma_1^2)$ فأعطت التباين $S_1^2 = 18$ وأخذت عينة عشوائية حجمها 11 من مجتمع $N(\mu_2, \sigma_2^2)$ مستقل عن الأول فأعطى التباين $S_2^2 = 12$. أوجد فترة ثقة 90% للنسبة σ_1^2/σ_2^2 .

الحل:

1. مستوى الثقة هو 90%.

$$1 - \alpha = 0.90 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1 - 0.05 = 0.95 \quad .2$$

(1) سليمان محمد طشطوش، أساسيات الإحصاء الرياضي، دار البيازوري، دروب للنشر، حمادة للدراسات الجامعية، عمان الأردن، 2012،

وبالنظر إلى جدول فيشر نستخرج القيم التالية ;

$$F_{0.95, 8, 10} = 3.07 \quad ; \quad F_{0.95, 10, 8} = 3.35$$

3. فترة الثقة هي:

$$\frac{S_1^2}{S_2^2} \frac{1}{F_{1-\frac{\alpha}{2}, v_1, v_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} F_{1-\frac{\alpha}{2}, v_2, v_1}$$

بعد التعويض نجد:

$$\frac{18}{12} \cdot \frac{1}{3.07} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{18}{12} \cdot 3.35$$

أي أن فترة الثقة للنسبة $\frac{\sigma_1^2}{\sigma_2^2}$ هي: [0.488 ; 5.02]

4 - سلسلة تمارين محلولة:

التمرين 1:

ضمن خطتها لإصلاح حركة المرور في المدينة قامت ولاية جيجل بإجراء مسح ميداني لتحديد حجم الحركة عبر تقاطع رئيسي في المدينة خلال فترات الصباح، باختيار يوم السبت لثمانية أسابيع متوالية تم عد المركبات التي تمر عبر التقاطع بين الساعة السابعة والساعة التاسعة صباحا ووجد أن متوسط عدد المركبات للعينة يساوي 1500 والانحراف المعياري للعينة يساوي 300.

المطلوب: أوجد مجال الثقة لمتوسط عدد المركبات في مجتمع الحركة عبر التقاطع باعتبار أن توزيع المجتمع طبيعي؟ (مع العلم أن مستوى الثقة يساوي 99%)

حل التمرين 1:

- إيجاد مجال الثقة لمتوسط عدد المركبات في مجتمع الحركة عبر التقاطع باعتبار أن توزيع المجتمع طبيعي:
لدينا عموما:

$$IC_{\mu} = \left[\bar{x} \pm z_c \delta_x \right]$$

$$IC_{\mu} = \left[\bar{x} \pm z_c \frac{\delta}{\sqrt{n}} \right]$$

- بما أن حجم العينة صغيرة $n < 30$ وعدم معرفة الانحراف المعياري للمجتمع نستخدم التوزيع ستودنت t لإيجاد مجال الثقة، ومنه يصبح مجال الثقة كما يلي:

$$IC_{\mu} = \left[\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n-1}} \right]$$

- ولدينا: $\bar{x} = 1500; t_{0.005, 7} = 3.499; S = 300; n = 8$ بالتعويض نجد:

$$IC_{\mu} = \left[\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n-1}} \right] = \left[1500 - (3.499) \left(\frac{300}{\sqrt{7}} \right); 1500 + (3.499) \left(\frac{300}{\sqrt{7}} \right) \right]$$

$$IC_{\mu} = [(1103 .25); (1896 .75)]$$

التمرين 2:

إذا كان لدينا متغيران عشوائيان x_1 و x_2 يتبعان التوزيع الطبيعي بتباين متساوي ومتوسطات مختلفة، تم سحب عينتان عشوائيتان مستقلتان من المجتمعين محل الدراسة وكانت لدينا البيانات التالية:

المجتمع 02	المجتمع 01	
$n_2 = 10$	$n_1 = 15$	حجم العينة
$\bar{x}_2 = 30$	$\bar{x}_1 = 40$	متوسط العينة
$s_2^2 = 16$	$s_1^2 = 18$	تباين العينة

- أوجد مجال الثقة للفرق بين متوسطي المجتمعين $\mu_1 - \mu_2$ مع العلم أن مستوى الثقة هو 95%.

حل التمرين 2:

- لدينا مجال الثقة للفرق بين متوسطي المجتمعين $\mu_1 - \mu_2$ ، وفي حالة تباين المجتمعين متساويين ومجهولين ($\sigma_1^2 = \sigma_2^2$)، يكتب كما يلي:

$$IC_{\mu_1 - \mu_2} = \left[\bar{x}_1 - \bar{x}_2 \pm t_{n_1 + n_2 - 2; \frac{\alpha}{2}} (S_p) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right]$$

- إذن سوف نحسب التباين المشترك S_p^2 أولاً حيث يساوي:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(14)(18) + (9)(16)}{23} \approx 17.22$$

$$t_{23; 0.025} = 2.069$$

- مجال الثقة هو:

$$\begin{aligned} IC_{\mu_1 - \mu_2} &= \left[\bar{x}_1 - \bar{x}_2 \pm t_{n_1 + n_2 - 2; \frac{\alpha}{2}} (S_p) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] \\ &= \left[(40 - 30) \pm (2.069)(4.15) \sqrt{\frac{1}{15} + \frac{1}{10}} \right] \\ &= [(10 - 3.5); (10 + 3.5)] = [(6.5); (13.5)] \end{aligned}$$

التمرين 3:

1- لتقدير نسبة عملاء بنك معين في منطقة ما، أخذت عينة عشوائية مكونة من 200 شخص فكان عدد عملاء البنك 120 شخصا.

- المطلوب: أوجد مجال الثقة لنسبة عملاء البنك في المنطقة؟ (مع العلم أن $1-\alpha = 95\%$)
($z = 1.96$) .

2- إذا علمنا أن عدد سكان تلك المنطقة يقدر ب 20000 نسمة.

- المطلوب: أوجد مجال تقدير لعدد عملاء البنك في تلك المنطقة؟ (يرمز لعدد عملاء البنك بالرمز A).

حل التمرين 3:

1- إيجاد مجال الثقة لنسبة عملاء البنك:

$$p' = \frac{120}{200} = 0.6$$

لدينا:

ولدينا مجال الثقة للنسبة في المجتمع يكتب كما يلي:

$$IC_p = [p' \pm z_c \delta_{p'}]$$

$$IC_p = \left[p' \pm z_c \sqrt{\frac{pq}{n}} \right] \quad \delta_{p'} = \sqrt{\frac{pq}{n}}$$

ولدينا: ، ومنه يصبح مجال الثقة كما يلي:

وبما أن p مجهولة ونريد إيجاد مجال الثقة لها ومنه لحساب $\delta_{p'}$ نستبدل p بدلالة p' النسبة في العينة وبذلك يصبح مجال الثقة يكتب كما يلي:

$$IC_p = \left[p' \pm z_c \sqrt{\frac{p'q'}{n}} \right] = \left[(0.6) \pm (1.96) \sqrt{\frac{(0.6)(0.4)}{200}} \right]$$

$$= [(0.532); (0.668)]$$

2- إيجاد مجال تقدير لعدد عملاء البنك يرمز للمجال بـ IC_A :

$$IC_A = [(0.532)(20000); (0.668)(20000)]$$

$$IC_A = [(10640); (13360)]$$

5 - سلسلة تمارين غير محلولة:

التمرين 1:

أجب عن الأسئلة التالية:

- أ. فيما يتجلى الفرق بين التقدير النقطي والتقدير بفترة؟
- ب. ما هي شروط التقدير الجيد؟
- ج. متى يتم استخدام توزيع ستودنت لإيجاد فترة الثقة لوسط المجتمع غير المعلوم؟ وما هي الحالة التي يمكن فيها استخدام التوزيع الطبيعي رغم مجهولية تباين المجتمع؟
- د. ماذا يقصد بدرجات الحرية؟ وضح بمثال.

التمرين 2:

- مجتمع طبيعي حجمه 800 وتباينه 625، تم سحب عينة عشوائية منه حجمها 25 فكان وسطها هو 100.
- أ. أوجد فترات الثقة لوسط المجتمع غير المعلوم عند المستويات التالية: 90%، 95%، 99%.
 - ب. ما هو طول الفترة في كل حالة مما سبق؟
 - ج. ما هو تعليقك على النتائج الثلاث؟

التمرين 3:

- ترغب إدارة جامعة المسيلة في تقييم جودة التعليم من مجموع الأساتذة البالغ عددهم 1389 أستاذًا، حيث قامت خلية جودة التعليم بالجامعة بتوزيع 100 استمارة ذات سلم ليكارت الخماسي على عينة عشوائية من الأساتذة فوجد أن متوسط العينة هو 2.85 بانحراف معياري 1.05.
- أوجد فترة ثقة 95% لوسط أساتذة الجامعة غير المعلوم.

التمرين 4:

- شركة مختصة في نقل المسافرين تريد تقدير المدة المستغرقة من طرف حافلاتها في خط محدد فقامت بمعاينة عشوائية لإحدى الحافلات وتم تسجيل الفترات التي تقضيها خلال كل رحلة بالدقائق، فكانت كما يلي:
- 92, 97, 104, 103, 96, 98, 102, 100
- أ. قدر نقطيا معدل الفترة التي تستغرقها حافلاتها خلال الرحلة.
 - ب. قدر نقطيا الانحراف المعياري للفترات المستغرقة خلال الرحلة الواحدة.
 - ج. بافتراض أن فترات الرحلات تتبع التوزيع الطبيعي، أوجد فترة ثقة 90% لمعدل فترة جميع الرحلات التي تقضيها حافلاتها من نفس الخط.

التمرين 5:

في عينة عشوائية حجمها 144 عاملا من مصنع به 3600 عاملا، وجد أن 94 عاملا يفضلون التقاعد المسبق.

1. أوجد فترة ثقة 95% لنسبة العاملين الذين يفضلون التقاعد المسبق.
2. إذا كنا نرغب في الحصول على نصف طول فترة الثقة السابقة فما هو حجم العينة المناسب لنفس المعطيات.

التمرين 6:

تريد إدارة مصنع لإفارج التحكم في جودة الإنتاج من خلال ضبط أوزان أكياس الإسمنت ودراسة التغيرات التي قد تطرأ على أوزان الأكياس المختلفة، فقامت بأخذ عينة من الأكياس حجمها 100 كيسا، حيث وجدت أن الانحراف المعياري لها هو 0.7 كغ.

- المطلوب: إيجاد فترة ثقة 90% للانحراف المعياري لأوزان أكياس الإسمنت بالمصنع.

التمرين 7:

من بين المؤسسات العاملة في قطاع معين وعددها 194 مؤسسة، تم سحب عينة (عينة عشوائية بسيطة بدون إرجاع) من 40 مؤسسة فوجد أن عدد العمال في المؤسسات كما يلي:

91	168	171	53	114	37	126	12	71	95
33	43	158	137	2	115	99	190	32	140
81	147	68	78	11	86	127	64	57	194
131	141	93	25	105	26	79	23	69	101

1. قدر متوسط عدد العمال للمؤسسة العاملة في القطاع؟
2. تقرر منح إعفاء ضريبي لمؤسسات القطاع التي تشغل 135 عاملا أو أكثر. قدر نسبة المؤسسات المستفيدة من الإجراء؟.
3. أحسب $(\delta_{p'})$ خطأ المعاينة للمقدر المستخدم في السؤال 2 ؟
4. أعط مجال ثقة لنسبة مؤسسات القطاع المستفيدة من الإعفاء؟ (مستوى معنوية = 5%).
5. كم يجب أن يكون حجم العينة إذا أردنا أن يكون تباين المقدر لنسبة المؤسسات المستفيدة $(\delta_{p'}^2) = 0.01$ ؟.

التمرين 8:

نسحب من مجتمع الأعمار في بلد معين عينة عشوائية بالإرجاع حجمها 410، متوسط القيم المحصلة 38.1 سنة.

1. أوجد مجال الثقة لمتوسط الأعمار في البلد μ بمستوى ثقة 0.95، إذا علمت أن تباين المجتمع هو $\sigma^2 = 36$ ؟

2. كيف يكون المجال إذا كان الانحراف المعياري للمجتمع مجهولاً، علماً أن الانحراف المعياري للعينة $S = 6.2$ ؟

التمرين 9:

1- لتقدير التفرقة بين الأجرور بين الإناث والذكور وجد لعينتين من 80 ذكر و60 أنثى أن متوسطي الدخل يساويان 20600 دج بانحراف معياري 3000 دج للذكور و19700 دج بانحراف معياري 2500 دج للإناث.
- أحسب مجال الثقة للفرق بين الدخل بافتراض التوزيع طبيعي ومستوى الثقة هو 95% ؟

2- إذا استبدلنا في السؤال الأول الأوساط الحسابية بالنسب في العينتين مثلاً أن نسبة الذكور ذوي الدخل أكثر من 10000 دج تساوي 40% وللإناث تساوي 30%، بحيث لم يتم تغيير حجم العينتين.
- أحسب مجال الثقة للفرق في النسب، التوزيع طبيعي ومستوى الثقة هو 95% ؟

التمرين 10:

لاحظ أستاذ بخبرته أن وسط درجات الطلاب في مادة الإحصاء 75 علامة وبانحراف معياري 9 علامات، إذا رغب الأستاذ في تطوير أسلوب تدريس مادة الإحصاء ومن تم تقدير الوسط الحسابي للعلامات وفق الأسلوب الجديد بحيث يكون متأكد بنسبة 95% أن الخطأ في التقدير الناتج يساوي 3 علامات.
- كم طالباً يحتاج الأستاذ لإخضاعهم للتجربة ؟

الفصل الرابع اختبار الفرضيات

- ✓ مصطلحات ومفاهيم
- ✓ عناصر الاختبار الإحصائي
- ✓ خطوات اختبار الفرضيات
- ✓ اختبار الفرضيات المتعلقة بالمتوسط
- ✓ اختبار الفرضيات المتعلقة بالنسبة
- ✓ اختبار الفرضيات المتعلقة بالتباين
- ✓ سلسلة تمارين محلولة
- ✓ سلسلة تمارين غير محلولة

تمهيد:

رأينا في الفصل السابق المتعلق بموضوع التقدير أن الباحث لا تكون لديه معلومات عن معلمة المجتمع التي يريد تقديرها نقطياً أو ضمن مجال بمستوى ثقة محدد فيلجأ إلى شواهد ومعلومات من العينة لتحديد تلك التقديرات الخاصة بالمعالم.

وفي هذا الفصل سيتم تناول موضوع اختبار الفرضيات الذي يعتبر أحد فروع الإحصاء الذي يهدف إلى اتخاذ القرار بشأن القيمة المعلنة أو المختبرة لمعلمة المجتمع من خلال فحص فرضيات حولها، انطلاقاً من الشواهد والأدلة التي تقدم من العينة.

1. مصطلحات ومفاهيم:

- **الفرضية الإحصائية:** الفرضية الإحصائية⁽¹⁾ هي عبارة عن معلمة أو أكثر من معالم المجتمع الإحصائي تكون قابلة للاختبار وبالتالي تكون صحتها أو عدم صحتها بحاجة إلى قرار.
- **اختبار الفروض:** يشير اختبار الفروض إلى قبول أو رفض ما لخاصية غير معلومة للمجتمع مثل أحد المعالم أو شكل توزيع المجتمع.
- **المنطقة الحرجة:** هي مجموعة قيم إحصاء الاختبار التي تؤدي إلى رفض الفرضية الصفرية، حيث أن كل حد من حدود المنطقة الحرجة يسمى قيمة حرجة لإحصاء الاختبار.
- **مستوى الدلالة (مستوى المعنوية):** هو احتمال رفض الفرضية الصفرية، وهو صحيح ويرمز له بالرمز α فعند استخدامنا لمستوى دلالة 0.05 (5%) فهذا يعني أننا سوف نرفض الفروض عندما يجب أن نقبله وأنها سوف نكون واثقين بنسبة 95% في أننا سنأخذ القرار السليم.

2. عناصر الاختبار الإحصائي:

يشتمل اختبار أي فرضية إحصائية على مجموعة من العناصر تتمثل فيما يلي:

1-2. الفرضية الصفرية (فرضية العدم):

وهي التي تصاغ بشكل نفي وجود علاقة بين متغيرين أو فروق بين مجموعتين أو أثر لمتغير مستقل على متغير آخر تابع، وهي الفرضية التي يجري عليها الاختبار الإحصائي ويبنى القرار على عدم صحتها ويرمز لها بالرمز: H_0 ⁽²⁾.

مثال 1:

- ✓ لا توجد فروق ذات دلالة إحصائية في استخدام تكنولوجيا المعلومات تعزى لحجم المؤسسة.
- ✓ لا توجد علاقة بين ترويج مبيعات المؤسسة إلكترونياً وارتفاع المبيعات.
- ✓ لا يوجد أثر ذو دلالة إحصائية لاستخدام تكنولوجيا المعلومات على تطوير إستراتيجية الترويج بالمؤسسة محل الدراسة.

2-2. الفرضية البديلة:

وهي الفرضية المعاكسة للفرضية الصفرية والتي يدعمها الباحث حيث تصاغ بشكل وجود علاقة بين متغيرين أو فروق بين مجموعتين أو أثر لمتغير مستقل على متغير آخر تابع ويرمز لها بالرمز: H_1 .

(1) محمد صبحي أبو صالح، عدنان محمد عوض، مرجع سابق، ص 233.

د. ليونارد، ج. كازمير، ترجمة مصطفى جلال مصطفى، مرجع سابق، ص 103 - 104 بتصرف.

(2) محمد خير سليم أبو زيد، التحليل الإحصائي لبيانات باستخدام برمجية SPSS، دار صفاء، عمان، الأردن، 2010، ص 241 بتصرف.

فعدم تحقق الفرضية الصفرية H_0 يؤدي إلى رفضها وبالتالي قبول الفرضية البديلة H_1 .

مثال 2:

- ✓ توجد فروق ذات دلالة إحصائية في استخدام تكنولوجيا المعلومات تعزى لحجم المؤسسة.
- ✓ توجد علاقة بين ترويج مبيعات المؤسسة إلكترونيا وارتفاع المبيعات.
- ✓ يوجد أثر ذو دلالة إحصائية لاستخدام تكنولوجيا المعلومات على تطوير إستراتيجية الترويج.

مثال حول الفرضية الصفرية والبديلة:

إذا كان معدل الطلبة في مادة الإحصاء هو 10 وأردنا اختبار معدل عينة من أولئك الطلبة من أجل اتخاذ قرار يتعلق بمعدلاتهم في مادة الإحصاء، فإن الفرضيات تصاغ كما يلي:

$$H_0 : \mu = 10$$

أما الفرضية البديلة فإنها يمكن أن تأخذ ثلاثة أشكال، وذلك حسب ما يقترحه الباحث:

$$H_1 : \begin{cases} \mu \neq 10 \\ \mu > 10 \\ \mu < 10 \end{cases}$$

ففي الفرضية الصفرية يقترح الباحث عدم وجود فروق بين معدل العينة في مادة الإحصاء والمعدل العام للطلبة في هذه المادة، بينما الفرضية البديلة تتضمن وجود فروق (اختلاف) بين معدل الطلبة في العينة والمعدل العام، أو أن المعدل أكبر من 10، أو أقل من 10.

2-3. إحصاء الاختبار:

هو إحصاء (اقتران تعين قيمته من العينة)، أي يتم حساب قيمته من خلال بيانات العينة مثل: (الإحصائي Z ، والإحصائي t ، والإحصائي F)، وتستخدم قيمته للوصول إلى قرار رفض أو قبول الفرضية الصفرية H_0 ، واختيار إحصائي الاختبار يعتمد على شكل التوزيع والفرضيات المفحوصة⁽¹⁾.

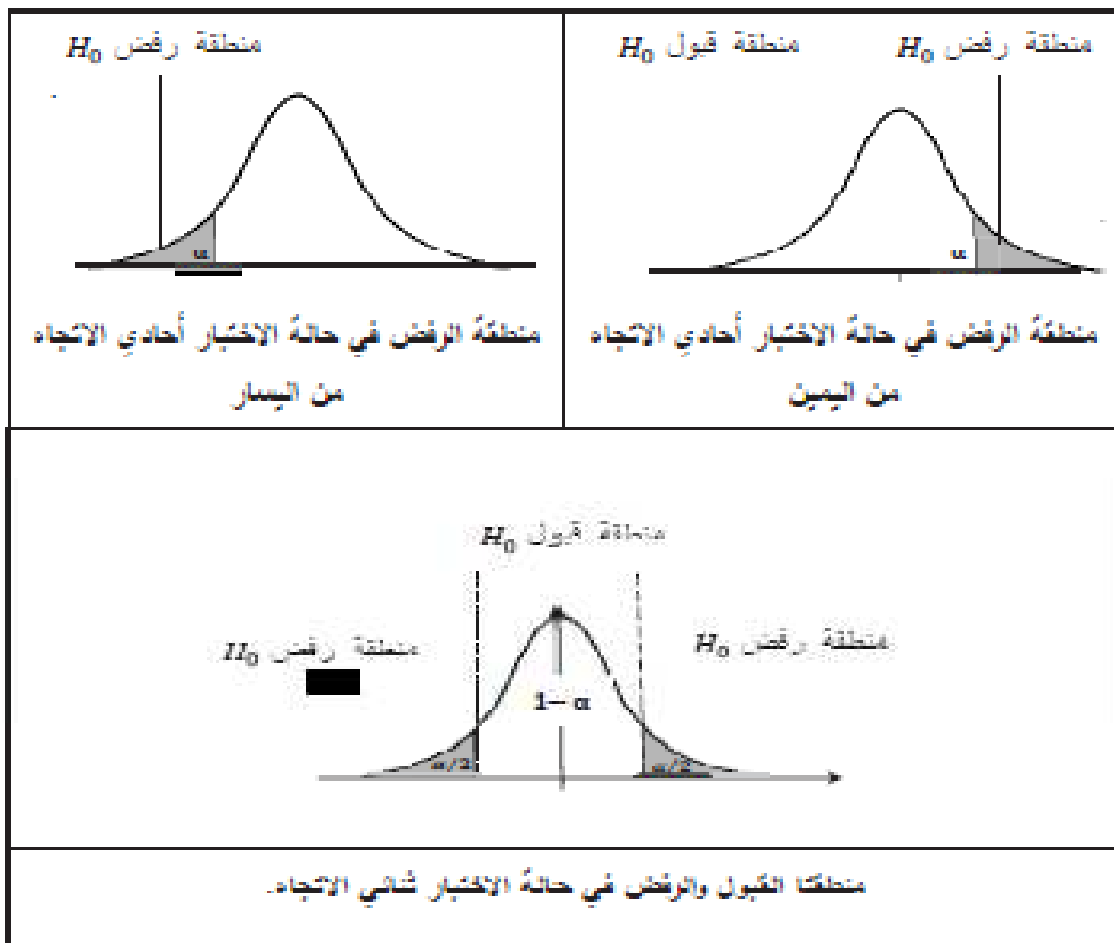
2-4. منطقة اتخاذ القرار:

تقسم منطقة اتخاذ القرار إلى منطقتين، إحداهما تحوي القيم التي تدعم الفرضية البديلة وتسمى منطقة الرفض (المناطق الحرجة)، وهي مجموعة من القيم التي تستخدمها والتي تمكننا من رفض الفرضية الصفرية، وحدودها يطلق عليها القيم الحرجة، وهي قيم تناظر مستوى الدلالة، والتي تمثل أعلى احتمالية لرفض الفرضية الصفرية وهي صحيحة صدفة.

(1) سالم عيسى بدر، عماد غصاب عيابة، مرجع سابق، ص 300 بتصرف.

والمنطقة الأخرى تحتوي على القيم التي تدعم الفرضية الصفرية وتسمى منطقة القبول، ففوق قيمة إحصاء الاختبار في منطقة الرفض يؤدي إلى رفض فرضية العدم وبالتالي قبول الفرضية البديلة⁽¹⁾.
فمنطقة الرفض إذن تتضمن قبول الفرضية البديلة والتي يعبر عنها إما باتجاه واحد على اليمين أو على اليسار، أو باتجاهين على جانبي القيمتين الحرجتين.
والشكل الموالي يوضح منطقتي الرفض والقبول للفرضية الصفرية H_0 في الحالات الثلاث:

الجدول 1-4: منطقتا الرفض والقبول للفرضية الصفرية H_0 في الحالات الثلاثة



2-5. الخطأ من النوع الأول والخطأ من النوع الثاني:

كل قرار يبنى على نتائج عينة يكون معرضاً للخطأ، وفي اختبار الفرضيات هناك حالتان بالنسبة للفرضية الصفرية هما: إما أن تكون صحيحة، وإما تكون غير صحيحة، كما أنه يوجد نوعان من القرارات المتخذة من طرف الباحث بشأن الفرضية الصفرية، هما: رفض الفرضية الصفرية أو عدم رفضها، وهذا ما يجعل اتخاذ القرار ينطوي على نوعين من الأخطاء في اختبارات الفروض الإحصائية يتمثلان فيما يلي⁽²⁾:

(1) سالم عيسى بدر، عماد غصاب عباينة، مرجع سابق، ص 298 بتصرف

(2) عبد الحميد عبد المجيد البلداوي، الأساليب الإحصائية التطبيقية، دار الشروق للنشر والتوزيع، عمان، الأردن، 2004، ص 282

2-5-1. الخطأ من النوع الأول:

هو الخطأ الذي يحدث عند رفض فرضية العدم عندما تكون صحيحة، والاحتمال الذي يحدث نتيجة الخطأ الأول يرمز له بالرمز α ، فالاسم البديل للخطأ من النوع الأول هو مستوى المعنوية.

2-5-2. الخطأ من النوع الثاني:

هو الخطأ الذي يحدث عند قبول (عدم رفض) فرضية العدم عندما تكون خاطئة، والاحتمال الذي يحدث نتيجة الخطأ الثاني يرمز له بالرمز β .
والجدول الموالي يوضح هذين النوعين من الأخطاء.

الجدول 4-2: الخطأ من النوع الأول والخطأ من النوع الثاني

خاطئ	صحيح	واقع الفرضية القرار
قرار صحيح $1 - \beta$	خطأ من النوع الأول α	رفض H_0
خطأ من النوع الأول β	قرار صحيح $1 - \alpha$	قبول H_0

ويشار إلى أن تصغير الخطأ من النوع الأول يترتب عليه زيادة في الخطأ من النوع الثاني، ولتخفيض كل من α و β يتم زيادة حجم العينة.

2-6. قوة إحصاء الاختبار:

ذكرنا بأنه في حالة قبولنا لفرضية العدم H_0 وهي غير صحيحة سيؤدي إلى وقوعنا في الخطأ من النوع الثاني، ويعتبر هذا القرار غير سليم، أما القرار السليم فهو عند رفض فرضية العدم H_0 لما تكون خاطئة وهنا تتجلى قوة الاختبار.

فقوة الاختبار إذن تقاس بمدى قدرته على رفض الفرضية الصفرية لما تكون هذه الأخيرة خاطئة. وعلى هذا الأساس فإن احتمال الوقوع في الخطأ الثاني يعتمد على الابتعاد من H_0 ، وأيضا على حجم العينة n والانحراف المعياري σ وعلى مستوى المعنوية α ونوع الاختبار إن كان من جانب واحد أو من جانبيين⁽¹⁾.

(1) عبد الحميد عبد المجيد البلداوي، مرجع سابق، ص 285 - 286

مثال 3:

نريد اختبار الفرضية:

$$H_0 : \mu = 25 \quad \longleftrightarrow \quad H_1 : \mu \neq 25$$

عند مستوى معنوية 0.05 وباستعمال الوسط الحسابي لعينة حجمها 100 أخذت من مجتمع انحرافه المعياري 12.

المطلوب:

- إيجاد احتمال الوقوع في الخطأ من النوع الثاني β عندما يكون $\mu = 25.6$

الحل:

المطلوب حساب:

$$\beta = P(\text{صحيحة } H_1 / \text{عدم رفض } H_0)$$

$$S_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{100}} = 1.2 \quad \bar{X} : \text{الخطأ المعياري لمتوسط العينة}$$

بما أن حجم العينة أكبر من 30 فإنه يمكن الاعتماد على نظرية النهاية المركزية وبالتالي استعمال Z . فعند مستوى دلالة $\alpha = 0.05$ نرفض H_0 عندما يقع \bar{X} خارج المجال: $25 \pm (1.96) \cdot (1.2)$

$$\begin{cases} \bar{X} > 22,648 \\ \bar{X} < 27,352 \end{cases} \text{ أي أن قرار عدم رفض } H_0 \text{ هو:}$$

إذن:

$$\beta = P(\text{صحيحة } H_1 : \text{عدم رفض } H_0) \Rightarrow$$

$$\beta = P(22.648 < X < 27.352) \Rightarrow$$

$$\beta = P\left(\frac{22.648-25.6}{1.2} < Z < \frac{27.352-25.9}{1.2}\right) \Rightarrow$$

$$\beta = P(-2.46 < Z < 1.46) = 0.4931 + 0.4279 = 0.921$$

نلاحظ أن قيمة احتمال الوقوع في الخطأ الثاني β جاءت مرتفعة جدا لأن القيمة 25.6 في الفرضية البديلة قريبة جدا من القيمة 25 في الفرضية الصفرية.

3. خطوات اختبار الفرضيات:

يتم اختبار الفرضيات كأسلوب إحصائي من خلال الخطوات التالية⁽¹⁾:

3-1. صياغة الفرضية الصفرية والفرضية البديلة:

في هذه الخطوة يركز الباحث على اتجاه الاختبار ليتمكن من صياغة الفرضية البديلة، والفرضية الصفرية عادة ما تأخذ شكل المساواة في حين أن الفرضية البديلة تأخذ أشكالاً مختلفة ويتحدد اتجاهها من خلال القراءة الجيدة للمشكلة وتحديد الاتجاه عن طريق العبارات التي تتضمنها مثل:
تختلف، لا تساوي، ...

تزيد، أكبر من، أعلى من، تفوق، تتجاوز...

تنقص، أقل من، أدنى، لا تفوق، لا تتجاوز،...

3-2. تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم):

يساعدنا مستوى الدلالة α (وهي قيمة الخطأ من النوع الأول) في تحديد النقاط الحرجة وبالتالي تحديد مناطق الرفض للفرضية الصفرية من خلال رسم المنحنى، حيث يعبر عن الاحتمال α بينياً بمنطقة الرفض لفرضية العدم H_0 في حالة الاختبار أحادي الاتجاه، فتضلل مساحة واحدة على اليمين أو على اليسار حسب اتجاه الاختبار، أو تضلل المساحتين $\frac{\alpha}{2}$ على جانبي منطقة قبول فرضية العدم H_0 .

3-3. تحديد قاعدة القرار:

تعتمد قاعدة القرار على تحديد طبيعة الإحصاءة بناء على طبيعة التوزيع، فبعد معرفة طبيعة التوزيع تتحدد طبيعة الإحصاءة مثل (الإحصائي Z والإحصائي t والإحصائي F) ، ثم يتم تحديد قاعدة القرار من خلال المجال الذي يتحدد بناء على نوع الاختبار الإحصائي.

3-4. حساب القيمة الجدولية والفعلية لإحصاءة الاختبار:

تستنتج القيمة الجدولية للإحصاءة بالاعتماد على الجداول الإحصائية المتوفرة (جدول التوزيع الطبيعي، جدول توزيع ستيودنت، جدول فيشر...) ، أما القيمة الفعلية فتحسب انطلاقاً من المعطيات المتوفرة في المسألة.

3-5. المقارنة واتخاذ القرار:

مقارنة القيمتين الجدولية والمحسوبة واتخاذ القرار إما برفض الفرضية الصفرية في حالة عدم انتماء القيمة المحسوبة لمجال قبول فرضية العدم وبالتالي قبول الفرضية البديلة، أو بقبولها في حالة العكس.

(1) سليمان محمد طشوش، مرجع سابق، ص ص 163 - 164 بتصريف

سالم عيسى بدر، عماد غصاب عبابنة، مرجع سابق، ص ص 298 - 301 بتصريف

4. اختبار الفرضيات المتعلقة بالمتوسط:

رأينا بأن اختبار الفرضيات يتوقف على اتجاه الفرضية البديلة، إما أن تكون متجهة لليمين فيكون اختبار أحادي الاتجاه لليمين أو كما يطلق عليها الفرضية البديلة ذات الذيل الأعلى، وإما أن تكون متجهة نحو اليسار فنكون بصدد اختبار أحادي الاتجاه من اليسار، أو كما يطلق عليها الفرضية البديلة ذات الذيل الأدنى، وإما أن تكون الفرضية البديلة غير موجهة فيكون الاختبار ثنائي الاتجاه .

4-1. الاختبار ثنائي الاتجاه للمتوسط:

لتوضيح هذا النوع من الاختبار والكيفية التي يتم بها تنفيذ خطواته نضع المثال الموالي:

مثال 4:

وجد في دراسة سابقة أن معدل حضور الطلبة للمحاضرات بإحدى كليات جامعة المسيلة هو 70 طالبا، وأن توزيعها يقترب من التوزيع الطبيعي.
نريد اختبار فرضية أن معدل الحضور قد تغير خلال هذا الموسم، فتم اختيار 81 طالبا عشوائيا فوجد أن معدل الحضور هو 73 طالبا بانحراف معياري قدره 10 طلاب.
- باستعمال مستوى دلالة 0.05 هل تؤيد تلك الفرضية ؟

الحل:**الخطوة الأولى: صياغة الفرضيات**

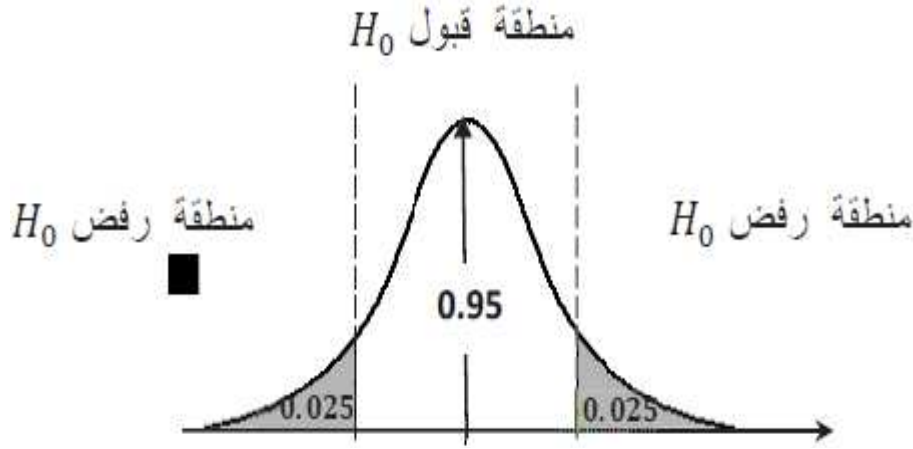
بما أن المسألة لم تتضمن اتجاه محدد للتغير، فإنه يتم استعمال فرضية بديلة ذات اتجاهين، وتكون صياغة الفرضيتين كما يلي:

$$H_0 : \mu = 70 \quad \longleftrightarrow \quad H_1 : \mu \neq 70$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)

مستوى الدلالة هو 0.05، وبما أن الاختبار ثنائي الاتجاه فإن منطقة الرفض ستكون على جانبيين كما هو موضح في الشكل الموالي:

الشكل 4-1: منطقة الرفض في حالة الاختبار ثنائي الاتجاه عند مستوى معنوية 0.05



الخطوة الثالثة: تحديد قاعدة القرار

بما توزيع المجتمع طبيعي فإننا سنحدد قاعدة القرار بناء على الإحصاءة Z ، وكون الاختبار ثنائي الاتجاه فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} -Z_{0.5-\alpha/2} \leq Z_C \leq Z_{0.5-\alpha/2} & \Rightarrow \text{عدم رفض } H_0 \\ Z_C < -Z_{0.5-\alpha/2} \vee Z_C > +Z_{0.5-\alpha/2} & \Rightarrow \text{رفض } H_0 \end{cases}$$

- الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاءة الاختبار

أولاً: القيمة الجدولية

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 0.5 - 0.025 = 0.4750$$

$$\Rightarrow Z_{0.5-\alpha/2} = 1.96$$

ثانياً: القيمة الفعلية

$$Z_C = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{73 - 70}{10/\sqrt{81}} = 2.7$$

الخطوة الخامسة: المقارنة واتخاذ القرار

$$\text{نلاحظ أن: } 2.7 > 1.96 \text{ أي أن: } Z_C > +Z_{0.5-\alpha/2}$$

وهذا يعني رفض الفرضية الصفرية وقبول الفرضية البديلة على أساس أن معدل الحضور قد تغير

هذا الموسم.

4-2. الاختبار أحادي الاتجاه من اليمين للمتوسط:

لتوضيح هذا النوع من الاختبار والكيفية التي يتم بها تنفيذ خطواته نضع المثال الموالي:

مثال 5:

بلغ معدل تسجيل مترشحي البكالوريا في كلية العلوم الاقتصادية بجامعة المسيلة 11.5 وفي السنوات الأخيرة زاد الطلب على التسجيل في هذه الكلية فرفعت الإدارة معدل القبول للدراسة بها. تم اختيار عينة عشوائية من الطلبة المسجلين بالكلية هذا العام حجمها 49 طالبا فأعطى معدل التسجيل 11.80 بانحراف معياري 1.5.

- قم باختبار الفرضية التي تقول بأن معدل القبول قد ارتفع عند مستوى معنوية 5%.

الحل:

الخطوة الأولى: صياغة الفرضيات

بما أن المسألة تضمنت اتجاها محددًا وهو ارتفاع معدل قبول التسجيل بالكلية، فإنه يتم استعمال فرضية بديلة ذات اتجاه واحد ومن اليمين (أكبر من)، وتكون صياغة الفرضيتين كما يلي:

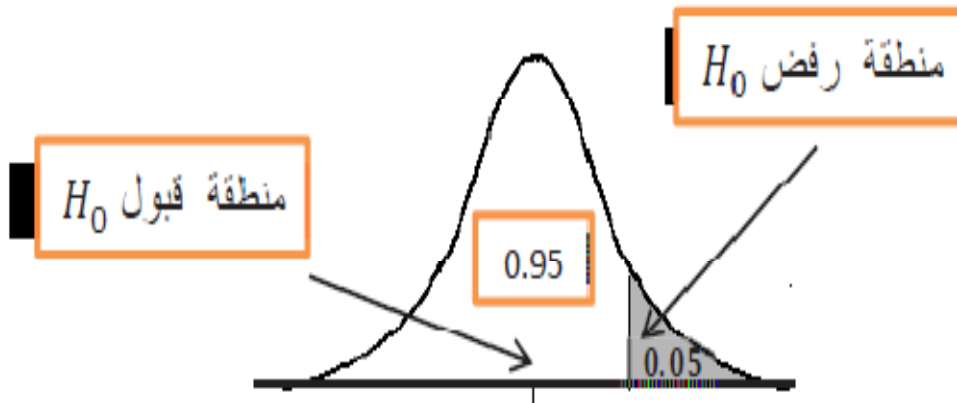
$$H_0: \mu = 11.5 \longleftrightarrow H_1: \mu > 11.5$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)

مستوى الدلالة هو 0.05، و بما أن الاختبار أحادي الاتجاه من اليمين فإن منطقة الرفض ستكون على جانب واحد ومن اليمين كما هو موضح في الشكل الموالي:

الشكل 4-2: منطقة الرفض في حالة الاختبار أحادي الاتجاه من اليمين

عند مستوى معنوية 0.05



الخطوة الثالثة: تحديد قاعدة القرار

بما أن حجم العينة أكبر من 30 فإنه يتم التقريب للتوزيع الطبيعي وسنحدد قاعدة القرار بناء على الإحصاء Z ، وكون الاختبار أحادي الاتجاه من اليمين فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} Z_C \leq Z_{0.5-\alpha} & \Rightarrow H_0 \text{ رفض} \\ Z_C > +Z_{0.5-\alpha} & \Rightarrow H_0 \text{ رفض} \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاء الاختبار

أولاً: القيمة الجدولية

$$\alpha = 0.05 \Rightarrow 0.5 - 0.05 = 0.4500$$

$$\Rightarrow Z_{0.5-\alpha} = 1.645$$

ثانياً: القيمة الفعلية

$$Z_C = \frac{\bar{X} - \mu}{S_{\bar{X}}} = \frac{11.80 - 11.50}{1.5/\sqrt{49}} = 1.4$$

الخطوة الخامسة: المقارنة واتخاذ القرار

$$\text{نلاحظ أن: } 1.4 < 1.645 \text{ أي أن: } Z_C < +Z_{0.5-\alpha}$$

وهذا يعني عدم رفض الفرضية الصفرية H_0 ورفض الفرضية البديلة على أساس أن معدل التسجيل بالكلية لم يرتفع.

3-4. الاختبار أحادي الاتجاه من اليسار للمتوسط:

لتوضيح هذا النوع من الاختبار والكيفية التي يتم بها تنفيذ خطواته نضع المثال الموالي:

مثال 6:

لاحظت إدارة محطة بنزين أن معدل فترة انتظار السيارة للتزود بالوقود هو 4 دقائق، وأرادت أن تقلص من هذه الفترة فقامت بإعادة تنظيم صفوف الانتظار، فأخذت عينها حجمها 9 سيارات وبعد أن دونت فترات الانتظار وجدت أن معدل فترات الانتظار هو 3 دقائق بانحراف معياري 30 ثانية.

المطلوب:

اختبر قرار إدارة المحطة لتخفيض فترات الانتظار عند مستوى معنوية 0.01 بافتراض أنها تتبع

التوزيع الطبيعي.

الحل:

الخطوة الأولى: صياغة الفرضيات

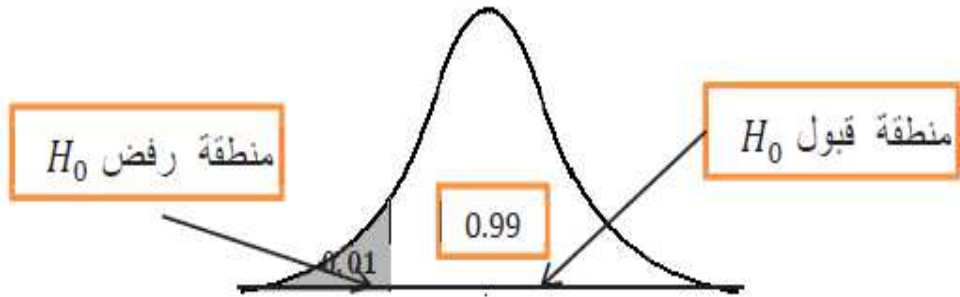
بما أن الإدارة تهدف إلى تخفيض فترة انتظار السيارة للتزود بالوقود، فإنه يتم استعمال فرضية بديلة ذات اتجاه واحد ومن اليسار (أصغر من)، وتكون صياغة الفرضيتين كما يلي:

$$H_0 : \mu = 4 \longleftrightarrow H_1 : \mu < 4$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم H_0)

مستوى الدلالة هو 0.01، وبما أن الاختبار أحادي الاتجاه من اليسار فإن منطقة الرفض ستكون على جانب واحد ومن اليسار كما هو موضح في الشكل الموالي:

الشكل 3-4: منطقة الرفض في حالة الاختبار أحادي الاتجاه من اليسار عند مستوى معنوية 0.01



الخطوة الثالثة: تحديد قاعدة القرار

المجتمع طبيعي التوزيع وحجم العينة صغير (أقل من 30) وتباين المجتمع مجهول وبالتالي فإن المتغيرة تتبع توزيع ستودنت وستحدد قاعدة القرار بناء على الإحصاء t ، وكون الاختبار أحادي الاتجاه من اليسار فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} t_C \geq -t_{0.99, 10-1} & \Rightarrow \text{عدم رفض } H_0 \\ t_C < -t_{0.99, 10-1} & \Rightarrow \text{رفض } H_0 \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاء الاختبار t

أولاً: القيمة الجدولية

$$\alpha = 0.01 \Rightarrow 1 - 0.01 = 0.99$$

$$\Rightarrow -t_{0.99, 10-1} = -2.821$$

ثانياً: القيمة الفعلية

$$t_c = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{3-4}{0.5/\sqrt{9}} = -6.02$$

الخطوة الخامسة: المقارنة واتخاذ القرار

$$t_c < -t_{0.99,10-1} \text{ أي أن: } -6.02 < -2.821$$

وهذا يعني عدم رفض الفرضية الصفرية H_0 ورفض الفرضية البديلة على أساس أن القرار بإعادة تنظيم صفوف الانتظار أدى إلى تخفيض فترات انتظار السيارات للتزود بالبنزين.

5. اختبار الفرضيات المتعلقة بالنسبة:

إن الخطوات المتبعة في اختبار الفرضيات المتعلقة بالوسط الحسابي هي نفسها التي سيتم استخدامها في اختبار الفرضيات المتعلقة بالنسبة، حيث نهتم بدراسة النسبة p ، وهي النسبة في العينة التي تملك الخاصية المطلوبة.

وكما ذكرنا آنفاً أن توزيع المعاينة للنسبة يتبع إما التوزيع الثنائي أو التوزيع الهندسي الزائد، لكن عندما يكون حجم العينة كبيراً ومع توفر الشروط ($n.p > 5$, $n.q > 5$)، وبالاعتماد على نظرية النهاية المركزية فإنه يمكن تقريب توزيع المعاينة للنسبة من التوزيع الطبيعي. لذلك سوف نستخدم التوزيع الطبيعي دائماً في اختبار الفرضيات حول النسبة (P). وتكون قيمة الإحصاء Z كما أشرنا لها في فصل توزيع المعاينة كما يلي:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

أما الفرضية الصفرية فتكون كما يلي:

$$H_0 : P = p_0$$

في حين أن الفرضية البديلة تأخذ الحالات الثلاث كما رأينا في اختبار الفرضيات المتعلقة بالوسط، وذلك حسب نوع الاختبار، تبعاً لذلك يتم استخراج قاعدة القرار كما يلي:

الجدول 4-3: قاعدة القرار لاختبار النسبة في الحالات الثلاث

نوع الاختبار	الفرضية البديلة	قاعدة القرار
أحادي الاتجاه من اليمين	$H_1: P > p_0$	$\begin{cases} Z_C \leq +Z_{0.5-\alpha} \Rightarrow H_0 \text{ عدم رفض} \\ Z_C > +Z_{0.5-\alpha} \Rightarrow H_0 \text{ رفض} \end{cases}$
أحادي الاتجاه من اليسار	$H_1: P < p_0$	$\begin{cases} Z_C > -Z_{0.5-\alpha} \Rightarrow H_0 \text{ عدم رفض} \\ Z_C \leq -Z_{0.5-\alpha} \Rightarrow H_0 \text{ رفض} \end{cases}$
ثنائي الاتجاه	$H_1: P \neq p_0$	$\begin{cases} -Z_{0.5-\alpha/2} \leq Z_C \leq +Z_{0.5-\alpha/2} \Rightarrow H_0 \text{ عدم رفض} \\ Z_C < -Z_{0.5-\alpha/2} \vee Z_C > +Z_{0.5-\alpha/2} \Rightarrow H_0 \text{ رفض} \end{cases}$

مثال 7:

أبدت إدارة التسويق في مؤسسة ما فكرا جديدا يسمى "البيع بروح الفريق" وفيه تستخدم الهاتف في عرض السلع على الزبائن قبل البيع، يدعي فريق البيع أنه يمكنه زيادة نسبة مكالمات البيع الناجحة إلى أكثر من 18%، وهي نسبة البيع الحالية والمحققة، ارتأت الإدارة أن تستخدم فريق البيع ولكن في البداية اشترطت أن تكون نسبة مكالمات البيع الناجحة يجب أن تتعدى النسبة الحالية للمبيعات.

- في اختبار ما، جرب فريق البيع مع عينة من 100 مكالمات بيع نجح في إتمام 22 عملية بيع.
- هل هذه العينة تظهر أن هناك تحسنا قد حدث مع فريق البيع عند مستوى معنوية 0.05؟

الحل:

الخطوة الأولى: صياغة الفرضيات

بما أن المسألة تضمنت اتجاها محددًا وهو زيادة نسبة مكالمات البيع الناجحة، فإنه يتم استعمال فرضية بديلة ذات اتجاه واحد ومن اليمين (أكبر من)، وتكون صياغة الفرضيتين كما يلي:

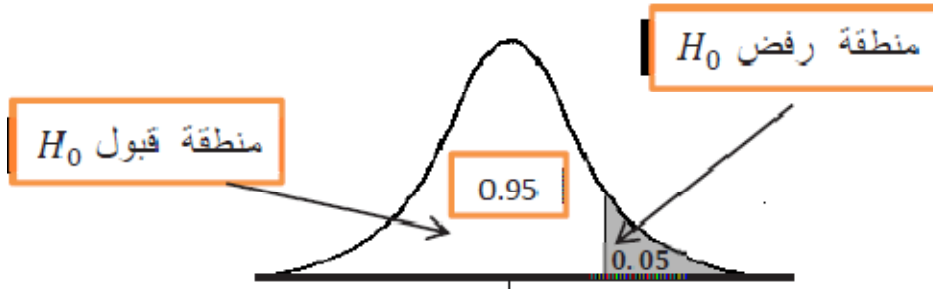
$$H_0 : P = 0.18 \longrightarrow H_1 : P > 0.18$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)

مستوى الدلالة هو 0.05، وبما أن الاختبار أحادي الاتجاه من اليمين فإن منطقة الرفض ستكون على

جانب واحد ومن اليمين كما هو موضح في الشكل الموالي:

الشكل 4-4: منطقة الرفض في حالة اختبار النسبة أحادي الاتجاه من اليمين عند مستوى معنوية 0.05



الخطوة الثالثة: تحديد قاعدة القرار

بما أن حجم العينة أكبر من 30 فإنه يتم التقريب للتوزيع الطبيعي وسنحدد قاعدة القرار بناء على الإحصاء Z ، وكون الاختبار أحادي الاتجاه من فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} Z_c \leq +Z_{0.5-\alpha} & \Rightarrow \text{عدم رفض } H_0 \\ Z_c > +Z_{0.5-\alpha} & \Rightarrow \text{رفض } H_0 \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاء الاختبار

أولاً: القيمة الجدولية

$$\alpha = 0.05 \Rightarrow 0.5 - 0.05 = 0.4500$$

$$\Rightarrow Z_{0.5-\alpha} = 1.645$$

ثانياً: القيمة الفعلية

$$\hat{p} = \frac{22}{100} = 0.22$$

$$Z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.22 - 0.18}{\sqrt{\frac{0.18(1-0.18)}{100}}} = 1.05$$

الخطوة الخامسة: المقارنة واتخاذ القرار

$$\text{نلاحظ أن: } 1.05 < 1.645 \text{ أي أن: } Z_c < +Z_{0.5-\alpha}$$

وهذا يعني عدم رفض الفرضية الصفرية H_0 ورفض الفرضية البديلة أي أن الفكر الجديد في البيع

المنتج من قبل إدارة التسويق لم يؤد إلى تحسن حقيقي.

مثال 8:

توصلت دراسة بإحدى الجامعات أن 40% من الطلبة يغيرون تخصصاتهم بعد السنة الأولى من

تسجيلهم ولقد تم أخذ عينة عشوائية من 100 طالب وتبين أن 38 طالباً غيروا تخصصهم بعد تلك السنة.

فهل يمكن الادعاء بأن هناك انخفاض معنوي في نسبة الطلاب الذين يغيرون تخصصهم عند مستوى معنوية 10%.

الحل:

الخطوة الأولى: صياغة الفرضيات

بما أن المسألة تضمنت اتجاهها محددًا وهو الانخفاض في نسبة الطلاب الذين يغيرون تخصصهم، فإنه يتم استعمال فرضية بديلة ذات اتجاه واحد ومن اليسار (أصغر من)، وتكون صياغة الفرضيتين كما يلي:

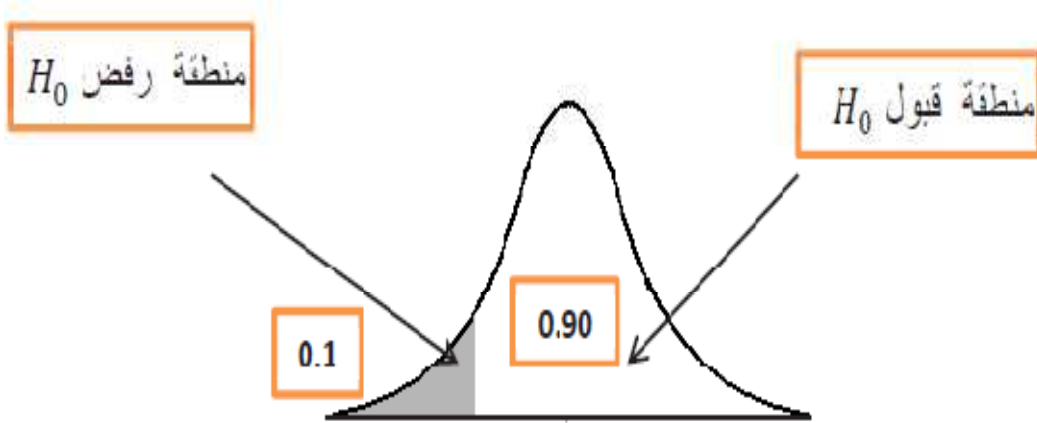
$$H_0 : P = 0.40 \longleftrightarrow H_1 : P < 0.40$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)

مستوى الدلالة هو 0.10، وبما أن الاختبار أحادي الاتجاه من اليسار فإن منطقة الرفض ستكون على جانب واحد ومن اليسار كما هو موضح في الشكل الموالي:

الشكل 4-5: منطقة الرفض في حالة اختبار النسبة أحادي الاتجاه من اليسار

عند مستوى معنوية 0.1



الخطوة الثالثة: تحديد قاعدة القرار

بما أن حجم العينة أكبر من 30 فيتم التقريب للتوزيع الطبيعي وسنحدد قاعدة القرار بناء على الإحصاءة Z ، وكون الاختبار أحادي الاتجاه من اليسار فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} Z_c \geq -Z_{0.5-\alpha} & \Rightarrow H_0 \text{ رفض} \\ Z_c < -Z_{0.5-\alpha} & \Rightarrow H_0 \text{ رفض} \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاء الاختبار

أولاً: القيمة الجدولية

$$\alpha = 0.10 \Rightarrow 0.5 - 0.10 = 0.4000$$

$$\Rightarrow -Z_{0.5-\alpha} = 1.28$$

ثانياً: القيمة الفعلية

$$\hat{p} = \frac{38}{100} = 0.38$$

$$Z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.38 - 0.40}{\sqrt{\frac{0.40(1-0.40)}{100}}} = -0.408$$

الخطوة الخامسة: المقارنة واتخاذ القرار

$$\text{نلاحظ أن: } -0.408 > -1.28 \text{ أي أن: } Z_c > -Z_{0.5-\alpha}$$

وهذا يعني قبول الفرضية الصفرية ورفض الفرضية البديلة أي أنه لا يمكن الادعاء بأن هناك انخفاض معنوي في نسبة الطلاب الذين يغيرون تخصصهم عند مستوى معنوية 10%.

مثال 9:

تبين شركة الحضنة للحليب من دراسات سابقة أن 80% من المستهلكين يفضلون أكياس الحليب العادي على أكياس حليب البقر، وفي استقصاء جديد تم على 500 مستهلك وجد أن 420 يفضلون أكياس الحليب العادي.

- عند مستوى ثقة 95%، هل بيانات العينة تؤيد ادعاء الشركة؟

الحل:

الخطوة الأولى: صياغة الفرضيات

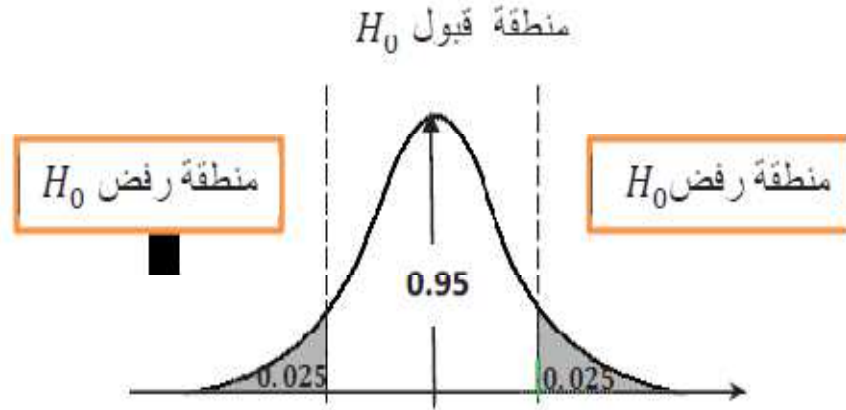
بما أن المسألة لم تتضمن اتجاهها محددًا حيث أشارت معطيات المثال إلى تحقق النسبة على مستوى العينة أم لا، فإنه يتم استعمال فرضية بديلة ذات اتجاهين، وتكون صياغة الفرضيتين كما يلي:

$$H_0 : P = 0.80 \longleftrightarrow H_1 : P \neq 0.80$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم) مستوى الدلالة هو 0.05، وبما أن الاختبار ثنائي الاتجاه فإن منطقة الرفض ستكون على جانبيين كما هو موضح في الشكل الموالي:

الشكل 4-6: منطقة الرفض في حالة اختبار النسبة ثنائي الاتجاه

عند مستوى معنوية 0.05



الخطوة الثالثة: تحديد قاعدة القرار

بما أن حجم العينة أكبر من 30 فإنه يتم التقريب للتوزيع الطبيعي وسنحدد قاعدة القرار بناء على الإحصاء Z ، وكون الاختبار ثنائي الاتجاه فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} -Z_{0.5-\alpha/2} \leq Z_C \leq +Z_{0.5-\alpha/2} & \Rightarrow \text{عدم رفض } H_0 \\ Z_C < -Z_{0.5-\alpha/2} \vee Z_C > +Z_{0.5-\alpha/2} & \Rightarrow \text{رفض } H_0 \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولة والفعلية لإحصاء الاختبار

أولاً: القيمة الجدولية

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 0.5 - 0.025 = 0.4750$$

$$\Rightarrow Z_{0.5-\alpha/2} = 1.96$$

ثانياً: القيمة الفعلية

$$\hat{p} = \frac{420}{500} = 0.84$$

$$Z_C = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.84 - 0.80}{\sqrt{\frac{0.80(1-0.80)}{100}}} = 1$$

الخطوة الخامسة: المقارنة واتخاذ القرار

نلاحظ أن: $-1.96 > 1 > +1.96$ أي أن: $-Z_{0.5-\alpha/2} \leq Z_C \leq +Z_{0.5-\alpha/2}$

وهذا يعني قبول الفرضية الصفرية H_0 ورفض الفرضية البديلة أي أن ادعاء الشركة بأن 80% من المستهلكين يفضلون أكياس الحليب العادي على أكياس حليب البقر هو ادعاء صحيح.

6. اختبار الفرضيات المتعلقة بتباين المجتمع:

لاختبار الفرضيات حول التباين نسلك نفس الخطوات المتبعة في اختبار الفرضيات السابقة، غير أن مجال قبول أو رفض الفرضية الصفرية يتحدد بالاعتماد على توزيع كاي تربيع بدرجات حرية $(n - 1)$ ،

$$\text{وتكون دالة الاختبار هي: } x^2 = (n - 1) \frac{S^2}{\sigma_0^2}$$

σ_0^2 : هي قيمة σ^2 المحددة بالفرضية الصفرية (المبدئية)

n : حجم العينة.

S^2 : تباين العينة

أما الفرضية الصفرية H_0 فتكون كما يلي: $H_0: \sigma^2 = \sigma_0^2$

في حين أن الفرضية البديلة تأخذ الحالات الثلاث كما رأينا في اختبار الفرضيات المتعلقة بالوسط،

أو بالنسبة، وذلك حسب نوع الاختبار، تبعا لذلك يتم استخراج قاعدة القرار كما يلي:

الجدول 4-7: قاعدة القرار لاختبار التباين في الحالات الثلاث

قاعدة القرار	الفرضية البديلة	نوع الاختبار
$\begin{cases} X^2 \leq x_{1-\alpha, v}^2 \Rightarrow H_0 \text{ رفض} \\ X^2 > x_{1-\alpha, v}^2 \Rightarrow H_0 \text{ رفض} \end{cases}$	$H_1: \sigma^2 > \sigma_0^2$	أحادي الاتجاه من اليمين
$\begin{cases} X^2 > x_{1-\alpha, v}^2 \Rightarrow H_0 \text{ رفض} \\ X^2 \leq x_{1-\alpha, v}^2 \Rightarrow H_0 \text{ رفض} \end{cases}$	$H_1: \sigma^2 < \sigma_0^2$	أحادي الاتجاه من اليسار
$\begin{cases} x_{\frac{\alpha}{2}; n-1}^2 \leq X^2 \leq x_{1-\frac{\alpha}{2}; n-1}^2 \Rightarrow H_0 \text{ رفض} \\ X^2 < x_{\frac{\alpha}{2}; n-1}^2 \vee X^2 > x_{1-\frac{\alpha}{2}; n-1}^2 \Rightarrow H_0 \text{ رفض} \end{cases}$	$H_1: \sigma^2 \neq \sigma_0^2$	ثنائي الاتجاه

مثال 10:

تدعي لجنة مختصة أن منتجات شركة صناعة العطور لها مدة استعمال بعد إنتاجها يزيد انحرافها المعياري عن شهرين، ولاختبار هذا الادعاء أخذت عينة حجمها 20 فكان متوسط مدة استعمالها هو 3 سنوات وانحرافها المعياري 3 أشهر.

- هل النتائج المتوصل إليها من خلال العينة تفقد هذا الادعاء أم تؤيده عند مستوى معنوية 5% ؟

الحل:

الخطوة الأولى: صياغة الفرضيات

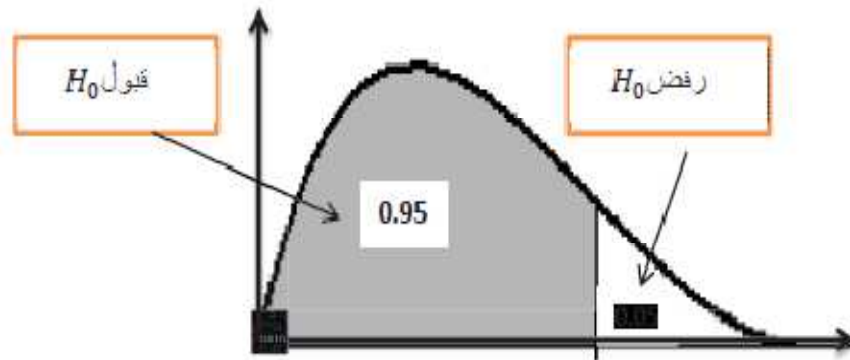
بما أن المسألة تضمنت اتجاهها محددًا وهو زيادة الانحراف المعياري لمدة استعمالها عن شهرين فإن الفرضية البديلة ذات اتجاه واحد ومن اليمين (أكبر من)، وتكون صياغة الفرضيتين كما يلي:

$$H_0: \sigma^2 = \sigma_0^2 \longleftrightarrow H_1: \sigma^2 > \sigma_0^2$$

الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم)

مستوى الدلالة هو 0.05، وبما أن الاختبار أحادي الاتجاه من اليمين فإن منطقة الرفض ستكون على جانب واحد ومن اليمين كما هو موضح في الشكل الموالي:

الشكل 4-7: اختبار التباين أحادي الاتجاه من اليمين عند مستوى معنوية 0.05



الخطوة الثالثة: تحديد قاعدة القرار

سيتم بناء قاعدة القرار على الإحصاءة χ^2 ، وكون الاختبار أحادي الاتجاه من اليمين فإن قاعدة القرار

تكون كما يلي:

$$\begin{cases} X^2 \leq x_{0.95, 19} & \Rightarrow \text{عدم رفض } H_0 \\ X^2 > x_{0.95, 19} & \Rightarrow \text{رفض } H_0 \end{cases}$$

الخطوة الرابعة: حساب القيمة الجدولية والفعلية لإحصاء الاختبار

أولاً: القيمة الجدولية

$$n = 20 \Rightarrow v = 20 - 1 = 19$$

فدرجة الحرية إذن هي 19، أما قيمة كاي تربيع المناظرة لها عند مستوى ثقة 0.95 من الجدول هي:

$$x_{0.95,19}^2 = 30.14$$

ثانياً: القيمة الفعلية

$$x^2 = (n - 1) \frac{s^2}{\sigma_0^2} = (20 - 1) \frac{3^2}{2^2} = 42.75$$

الخطوة الخامسة: المقارنة واتخاذ القرار

نلاحظ أن: $42.75 + 30.144$ أي أن: $X^2 > x_{1-\alpha, v}^2$

وهذا يعني رفض الفرضية الصفرية H_0 وقبول الفرضية البديلة أي أن النتائج المتوصل إليها من

خلال العينة تؤيد أدعاء اللجنة عند مستوى معنوية 5%.

7 - سلسلة تمارين محلولة:

التمرين 1:

اشترت شركة خمسة منتجات من مورد يقول أن متوسط عمر المنتج هو 1050 ساعة، ولما قامت الشركة باستعمال عينة المنتجات عاشت عدد 964، 1082، 1136، 825، 863 ساعة، ولم تكن هذه النتائج مرضية بالنسبة للشركة، المطلوب إرسال شكوى للمورد تحلل النتائج وتقدم التوصيات.

- هل يمكن القول بصحة زعم المورد بمستوى معنوية 0.05؟

حل التمرين 1:

- نلاحظ بأن الاختبار هنا هو اختبار المتوسط، وبما أن النتائج لم تكن مرضية بالنسبة للشركة فهذا يعني أن الاختبار هو اختبار أحادي الاتجاه من اليسار ومنه الفرضيات تكتب من الشكل التالي:

$$\begin{cases} H_0 : \mu = 1050 \\ H_1 : \mu < 1050 \end{cases}$$

ومنه قاعدة القرار تكتب من الشكل التالي:

$$\begin{cases} RH_0 \text{ si } \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} < -z_{1-\alpha} \\ \bar{RH}_0 \text{ sinon.} \end{cases}$$

بما أن عدد المشاهدات للعينة هو 5، ولدينا δ الانحراف المعياري للمجتمع مجهول، إذن نحتاج إلى الانحراف المعياري للعينة S عند حساب δ_x . ونستخدم كذلك توزيع ستودنت t ومنه تصبح قاعدة القرار تكتب كما يلي:

$$\begin{cases} RH_0 \text{ si } \frac{\bar{X} - \mu_0}{S / \sqrt{n-1}} < -t_{n-1; 1-\alpha} \\ \bar{RH}_0 \text{ sinon.} \end{cases}$$

- لدينا: $n = 5; \bar{X} = \frac{\sum x_i}{n} = 974; S \approx 134.66$

ومنه بالتعويض في متغيرة القرار t_c نجد:

$$t_c = \frac{974 - 1050}{134.66 / \sqrt{4}} \approx -1.13$$

$$-t_t = -t_{4, 1-\alpha} = -2.132$$

- القرار: بما أن $t_c = -1.13 > -t_t = -2.132$ فإننا نقبل الفرضية الصفرية H_0 ، بمعنى أننا يمكن أن نجزم بصحة ادعاء المورد، وبالتالي لا يمكن رفع شكوى حول عدم صحة ادعائه.

التمرين 2:

لدينا اختبار لقياس قدرة الطفل على تعلم كلمات جديدة، والمطلوب معرفة ما إذا كان متوسط درجات هذا الاختبار لأطفال إحدى المجتمعات يختلف عنه لأطفال مجتمع آخر، افترض أن درجات أطفال هاتين المجتمعين تتبع التوزيع الطبيعي بحيث لهما نفس التباين، أعطى الاختبار لمجموعتين عشوائيتين مستقلتين من الأطفال بعمر 4 سنوات البيانات التالية:

المجموعة الثانية	المجموعة الأولى
$n_2 = 8$	$n_1 = 10$
$\bar{X}_2 = 97$	$\bar{X}_1 = 95$
$S_2^2 = 36$	$S_1^2 = 40$

- اختبر الفرض القائل بتساوي قدرات أطفال المجموعتين على تعلم كلمات جديدة، بمستوى معنوية $\alpha = 0.05$.؟

حل التمرين 2:

لدينا الاختبار هو اختبار الفرق بين متوسطين ومنه فإن الفرضيتين تكتب من الشكل التالي:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

- وبما أن العينات صغيرة فإن الإحصائية في هذه الحالات بافتراض أن المجتمعين طبيعيين، وأن تباين المجتمع الأول يساوي تباين المجتمع الثاني ولكنه مجهول (بمعنى أن $\sigma_1^2 = \sigma_2^2$ ولكن قيمة هذا التباين غير معروفة) وأن العينتين مستقلتان فإن إحصائية الاختبار تأخذ الشكل التالي والتي لها توزيع t بدرجات حرية تساوي $n_1 + n_2 - 2$:

$$T' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \rightarrow t_{n_1 + n_2 - 2}$$

والقرار يكتب من الشكل التالي:

$$\begin{cases} RH_0 \dots Si \dots T' > t_{n_1 + n_2 - 2} \\ \bar{RH}_0 \dots Sinon \end{cases}$$

- بالتعويض بالقيم في متغيرة القرار نجد:

$$T' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$= \frac{95 - 97}{\sqrt{\frac{(9)(40) + (7)(36)}{16} \left(\frac{1}{10} + \frac{1}{8} \right)}} \approx -0.68$$

$$t_{n_1 + n_2 - 2} = t_{16; 0.025} = 2.12$$

- القرار: بما أن $T' = -0.68 < t_t = 2.12$ فإننا نقبل فرضية العدم ونستنتج أن الفرض القائل بتساوي قدرات أطفال المجموعتين على تعلم كلمات جديدة صحيح.

التمرين 3:

سحبنا من مجتمعين طبيعيين عينتين حجم الأولى $n_1 = 51$ وحجم الثانية $n_2 = 67$ ومستقلتين، وجدنا النتائج التالية:

$$. S_1^2 = 13 ; ; S_2^2 = 11 ; ; \bar{X}_1 = 101 ; ; \bar{X}_2 = 96$$

- كيف يمكن إجراء اختبار تساوي تبايني المجتمعين بمستوى معنوية $\alpha = 0.05$.

حل التمرين 3:

الغرض من الاختبار هو تأكيد أو نفي تساوي تباين مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين. تكتب الفرضيات في حالة الاختبار الثنائي كما يلي:

$$H_0 : \sigma_1^2 = \sigma_2^2 \leftrightarrow H_1 : \sigma_1^2 \neq \sigma_2^2$$

بما أن المجتمعين طبيعيين وحجم العينتين كبير يعني أكبر من 30، فإننا نعلم في الاختبار على متغيرة القرار T والتي تساوي:

$$T = \frac{1}{2} \ln \left(\frac{\hat{S}_1^2}{\hat{S}_2^2} \right) / \sqrt{\frac{1}{2} \left(\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1} \right)} \approx N(0; 1)$$

$$\hat{S}_1^2 = S_1^2 \frac{n_1}{n_1 - 1} = 13 \frac{51}{50} = 13.26$$

$$\hat{S}_2^2 = S_2^2 \frac{n_2}{n_2 - 1} = 11 \frac{67}{66} \approx 11.17$$

$$T = \frac{1}{2} \ln \left(\frac{13.26}{11.17} \right) / \sqrt{\frac{1}{2} \left[\left(\frac{1}{50} \right) + \left(\frac{1}{66} \right) \right]} \approx 0.65$$

$$\alpha = 0.05 \Rightarrow z_t = 1.96$$

- القرار: بما أن $T = 0.65 < Z_t = 1.96$ فإننا نقبل الفرضية الصفرية، أي نقبل بتساوي التباينين.

8 - سلسلة تمارين غير محلولة :

التمرين 1:

1. ما هو المبدأ العام لاختبار الفروض ؟
2. ماذا يقصد بالخطأ من النوع الأول والخطأ من النوع الثاني ؟
3. ماذا يقصد بقوة الاختبار ؟

التمرين 2:

شركة مصغرة مختصة في صناعة الأثاث المنزلي، يريد صاحبها التزود بأعمدة خشبية سمكها 4 سم، حيث تبين له من خلال خبرته السابقة أن الأقل سمكا غير ملائمة والأكثر سمكا أكثر تكلفة، فقام بأخذ عينة عشوائية من 36 عمودا من مورد فوجد أن متوسط سمكها هو 3.8 سم بانحراف معياري 0.2.

- بم تتصح صاحب الشركة إذا كان يرغب في اتخاذ القرار بشأن شراء الأعمدة من هذا المورد عند مستوى معنوية 5% ؟

التمرين 3:

إذا كان أحد مصانع الأغذية ينتج نوعا معينا من الألبان حيث متوسط وزن العبوة هو 0.5 كغ وذلك بانحراف معياري 36 غ، حيث أن أوزان العبوات تتبع التوزيع الطبيعي، ولمراقبة التزام المصنع بمعايير الجودة تم اختيار عينة من 25 عبوة فوجد أن متوسط الأوزان يساوي 0.48 كغ.

- هل ترى أن هناك عيبا في الإنتاج أدى إلى انخفاض متوسط الأوزان للعبوات عند مستوى معنوية 1% ؟

التمرين 4:

بلغت نسبة الطلبة المسجلين بالمركز المكثف للغات بإحدى جامعات الجزائر 20%، وفي محاولة من إدارة المعهد لرفع هذه النسبة شرعت في توعية الطلاب بمختلف الوسائل الترويجية للمعهد ومهامه، ثم قامت بأخذ عينة عشوائية من 1000 طالب فوجدت بأن عدد المسجلين بالمعهد هو 220 طالبا.

- ما هو تقييمك لمحاولة إدارة المعهد عند مستوى معنوية 0.05.

التمرين 5:

تم أخذ أربع قراءات بجهاز معين فكانت قيمها كما يلي: 51، 51، 55، 59.
 اختبر الفرضية الصفرية: $H_0: \sigma = 0.7$ مقابل الفرضية البديلة: $H_1: \sigma \neq 0.7$ عند مستوى دلالة 0.10.

التمرين 6:

إذا كان من المعروف أن جسم الإنسان البالغ يحتاج يومياً في المتوسط 800 مليغرام من الكالسيوم لكي يقوم بوظائفه خير قام. ويعتقد احد علماء التغذية أن الأفراد ذوي الدخل المنخفض لا يستطيعون تحقيق هذا المتوسط، ولاختبار ذلك تم اختيار عينة من 50 شخصاً بالغاً من بين ذوي الدخل المنخفض فكان متوسط ما يتناوله من كالسيوم يومياً هو 755,3 مليغرام والانحراف المعياري هو 239,3 مليغرام. فهل تدل هذه النتائج على أن متوسط ما يتناوله الأشخاص البالغون من ذوي الدخل المنخفض من كالسيوم يقل عن 800 مليغرام؟ استخدم مستوى معنوية 0,05.

التمرين 7:

البيانات التالية تمثل نتائج تجربة أجريت على عشرين شخصاً لاختبار مدى فعالية نظام خاص من الغذاء لتخفيف الوزن، حيث تم قياس أوزانهم قبل البدء في تطبيق هذا النظام، وبعد إتباع هذا النظام الخاص لمدة ثلاثة شهور.

92	103	120	89	93	107	94	90	110	96	قبل
84	95	103	76	85	104	87	85	96	90	بعد
123	111	90	95	123	105	110	86	94	86	قبل
107	102	83	89	109	95	102	80	84	78	بعد

المطلوب:

هل تستطيع أن تستنتج أن نظام الغذاء كان فعالاً في تخفيف الوزن مستخدماً مستوى دلالة $\alpha = 0.05$ ؟

التمرين 8:

أجريت التجربة التالية لمقارنة نوعين من الأعلاف، خصص لكل علف 10 أبقار وكان معدل الزيادة في أوزان الأبقار كما يلي:

32	30	34	38	35	32	26	29	34	31	علف (1)
29	31	26	29	30	29	28	24	26	28	علف (2)

- إذا كان المجتمعين يتبعان التوزيع الطبيعي، ولهما نفس التباين. هل هناك فرق معنوي بين نوعي العلف، وذلك بمستوى معنوية $(\alpha = 0,05)$ ؟

قائمة المراجع

أولاً: المراجع باللغة العربية

1. أنيس إسماعيل كنجو، الإحصاء والاحتمال، مكتبة العبيكان، الرياض، المملكة العربية السعودية، 2000.
2. جلال الصياد، عبد الحميد محمد ربيع، مبادئ الطرق الإحصائية، الطبعة الأولى، دار الناشر تهامة، جدة- المملكة العربية السعودية 1983.
3. جورج كانافوس، دون ميلر، ترجمة سلطان محمد عبد المجيد، محمد توفيق البلقيني، الإحصاء للتجاربيين، مدخل حديث، دار المريخ، الرياض، المملكة العربية السعودية، 2004.
4. دليونارد، ج كازمير، ترجمة مصطفى جلال مصطفى، الإحصاء التجاري، الدار الدولية للاستشارات الثقافية، مصر، 2004.
5. دلال القاضي وآخرون، الإحصاء للإداريين والاقتصاديين، دار الحامد، عمان، الأردن، 2005.
6. كمال سلطان محمد سالم، الإحصاء الاحتمالي، الدار الجامعية، الإسكندرية، مصر، 2004.
7. محمد أمانى موسى، التحليل الإحصائي للبيانات، مركز تطوير الدراسات العليا القاهرة، مصر 2007.
8. محمد حسين محمد رشيد، منى عطا الله الشويلات، مبادئ الإحصاء والاحتمالات ومعالجتها باستخدام برنامج SPSS، دار صفاء، عمان، الأردن، 2012.
9. محمد خير سليم أبو زيد، التحليل الإحصائي للبيانات باستخدام برمجية SPSS، دار جرير، دار صفاء، عمان، الأردن، 2010.
10. محمد راتول، الإحصاء الوصفي، ديوان المطبوعات الجامعية، ط2، الجزائر، 2006.
11. محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل باستخدام SPSS، دار المسيرة، عمان، الأردن، 2004.
12. محمد عبد الفتاح الصيرفي، الدليل التطبيقي للباحثين، دار وائل، عمان، الأردن، 2002.
13. معتوق أمحمد، الإحصاء الرياضي والنماذج الإحصائية، ديوان المطبوعات الجامعية، الجزائر، 2007.
14. موراي شبيجل، ترجمة مصطفى جلال مصطفى، الإحصاء، الدار الدولية للاستشارات الثقافية، مصر، 2004.
15. موريس أنجرس، ترجمة بوزيد صحراوي وآخرون، منهجية البحث في العلوم الإنسانية، دار القصة، الجزائر، 2004.
16. سالم عيسى بدر، عماد غصاب عابنة، مبادئ الإحصاء الوصفي والاستدلالي، دار المسيرة، عمان، الأردن، 2007.

17. سامح جزماتي، الاحتمالات والإحصاء، مديرية الكتب والمطبوعات الجامعية، جامعة حلب، دمشق، 1989.
18. سعيد التل وآخرون، مناهج البحث العلمي، تصميم البحث والتحليل الإحصائي، دار المسيرة، عمان، الأردن، 2007.
19. سليمان محمد طشطوش، أساسيات الإحصاء الرياضي، دار اليازوري، دروب للنشر، حمادة للدراسات الجامعية، عمان الأردن، 2012.
20. شبيزل، شيلر، سرينيفيسان، ترجمة محمد علي عبد الناصر، مصطفى جلال مصطفى، الاحتمالات والإحصاء، الدار الدولية للاستثمارات الثقافية، مصر، 2004.
21. صالح بوعيدك، مطبوعة محاضرات الإحصاء الرياضي، كلية العلوم الاقتصادية، جامعة المسيلة، 2006/2005.
22. عامر إبراهيم قندليجي، منهجية البحث العلمي، دار اليازوري، عمان، الأردن، 2012.
23. عبد الحفيظ محمد فوزي مصطفى، الاستدلال الإحصائي (1) نظرية التقدير، مجموعة النيل العربية، مصر، 1999.
24. عبد الحميد عبد المجيد البلداوي، الأساليب الإحصائية التطبيقية، دار الشروق للنشر والتوزيع، عمان، الأردن، 2004.
25. عبد المجيد قدي، أسس البحث العلمي في العلوم الاقتصادية، دار الأبحاث، الجزائر، 2009.
26. عدنان عباس حميدان، مطانيوس مخول، فريد جاعوني، عمار ناصر آغا، الإحصاء التطبيقي، منشورات جامعة، دمشق 2006.
27. عدنان عوض، الإحصاء التطبيقي، الشركة العربية المتحدة بالتعاون مع جامعة القدس المفتوحة، مصر، 2009.
28. عزام صبوي، الإحصاء الرياضي، دار صفاء، عمان، الأردن، 2010.
29. فايز جمعة صالح النجار وآخرون، أساليب البحث العلمي، منظور تطبيقي، دار الحامد، عمان، الأردن، 2009.

ثانيا: المراجع باللغة الأجنبية

1. Amany Mousa: Cairo (2005), Statistical Data Analysis, Center for Advancement of Postgraduate Studies and Research, Faculty of Engineering, Cairo University
2. C. Reder, Probabilités et Statistiques, Cours et Exercices, IUP2- MIAGE, Bordeaux I, 2002-2003.
3. F. Balabdaoui-Mohr et O. Wintenberger, Statistique Mathématique, <https://ceremade.deuphine.fr/~fadoua/statmathpoly2009.pdf>
4. Freund, J (2001) Modern Elementary Statistics 10th Edition, Printice Hall.
5. Hamburg, Morris (1989) Statistical Analysis for Decision Making, Harcoort Brace Jovanwiche, USA. GLEN COWN, Statistical Data Analysis, Clarendon, OXFORD, 1998.
6. Keller, G and Waracck, B (2001): Statistics for Management and Economics 6th Edition Duxbury.
7. Mills, Frederick (1955), Statistical Methods, Holt, Rinehart and Winston, New York, USA.
8. Neeter, J, Wasserman, Whitmare, (1993): Applied Statistics. 4th Edition, Louise Richardson.
9. N V Nagendram, Probability and Statistical Applications - Distributions, https://www.researchgate.net/publication/268870344_Probabili_and_Statistical_Applications_-_Distributions?enrichId=rgreq-36c95422c5fb776a6216f014d071617b-XXX&enrichSource=Y292ZXJQYWdlOzI2ODg3MDMD0NDtBUzoxNjg4MTg?jg4ODM3MTJAMTQxNzI2MDkzOTcwMQ%3D&el=1_x_3.
10. Wonnacott, Thomas and Wonnacott, Ronald (1990) Introductory Statistics for Business and Economics, John Wiley & soon, New York, USA.