الفصل السابع

الاستقراء الإحصائي

٧ـ١ مقدمة : يواجه كل منا في حياته اليومية قرارات شخصية وحالات تتطلب القيام بتنبؤ ات حول المستقبل . ونجد الحكومة مثلاً تهتم بالتنبؤ بعدد الطلاب في مختلف المراحل خلال السنوات القادمة ، وتهتم مؤسسات التجارة الخارجية بمعرفة تقلبات الأسعار العالمية خلال فترة زمنية قادمة ، وتريد شركة صناعية لتشكيل المعادن استخدام نتائج تجربة معينة لاستقراء ما إذا كان نوع جديد من الفولاذ أكثر مقاومة لتغيرات درجة الحرارة من نوع آخر . وترغب ربة البيت بمعرفة أي من المنظفين A أو B أكثر فعالية في غسالتها . ومن حسن الطالع أنه يمكن أن نبني مثل هذه الاستقراءات على نتف من المعلومات الواقعية المتوفرة لنا والتي ندعوها بالملاحظات أو المعلومات الإحصائية أو البيان الإحصائي .

وفي الكثير من الحالات الخاصة تكون المعلومات المتوفرة غزيرة . منسجمة في ظاهرها . ومن نواح عديدة على درجة كبيرة من الوضوح بحيث لا يكون القرار أو التنبؤ المتخذ بعناية أفضل بكثير من مجرد التخمين الشخصي . وعلى الوجه الآخر نجد أن التحليلات الشخصية للمعلومات من قبل علماء ومهندسين تؤدي غالباً إلى آراء متعارضة حول النتائج المستخلصة لتجربة . وبينما يميل العديد من الناس للشعور بقدرتهم الذاتية على القيام باستقراءات جيدة إلا أن التجربة تدل على أنه ليس باستطاعة الإنسان إجهاد ذهنه بكمية كبيرة من الأرقام وتحليل وموازنة النتف المتوفرة من المعلومات للوصول إلى إستقراء جيد . وهكذا يصبح وضع أدوات ونظم للاستقراء أمراً مرغوباً . وهذا هو

هدف الإحصاء الرياضي .

وقد رأينا أن هدف الإحصاء هو القيام باستقراءات حول المجتمع بدءاً من معلومات توفرها لنا العينة . وإلى المدى الذي تتميز فيه المجتمعات بمقاييس وصفية رقمية تدعى الوسطاء ، فإن الاستقراء الإحصائي يهتم بتوفير استقراءات جيدة حول هذه الوسطاء . ونذكر من هذه الوسطاء ، على سبيل المثال ، المتوسط ، الانحراف المعياري ، المساحة تحت منحني كثافة احتمالية فوق أو تحت قيمة معينة للمتحول العشوائي ، أو المساحة بين قيمتين لهذا المتحول . ويمكن ترجمة المسائل التي نواجهها في التطبيق العملي إلى مسائل إستقراء حول وسيط أو أكثر .

وتقع طرق القيام باستقراء حول وسيط في صنفين . إذ يمكن إتخاذ قرارات تتعلق بقيمة الوسيط . كما وجدنا مثلاً في مسألة الكشف بطريقة العينة لقبول أو رفض وسقة من البضاعة ، واختبار الفرضية الذي وصفناه في الفصل الخامس . كما يمكن تقدير الوسيط أي التنبؤ بقيمته .

ولا يكون عرضنا لهدف الاستقراء الإحصائي وأنواعه كاملاً دون الإشارة لمقياس جودة الطرق الاستقرائية . إذ يمكن تعريف الكثير من الطرق الموضوعية للقيام باستقراء بالإضافة إلى الطرق الشخصية القائمة على البداهة . ولا بد من مقياس لجودة كل طريقة بحيث يمكن مقارنة هذه الطرق ببعضها والقيام بمفاضلة فيما بينها . هذا بالإضافة إلى أننا نريد التعبير عن جودة استقراء معين في حالة فيزيائية معينة . فالتنبؤ بأن السعر العالمي لمادة نستوردها سيكون 80 دولاراً في الشهر القادم ليس كافياً ويجب ألا تشكل حافزاً للشراء أو عدم الشراء . فالسؤال الأساسي هو ما إذا كان هذا التقدير صحيحاً في حدود دولار أو دولارين أو عشرة دولارات زيادة أو نقصاناً . ويحوي الاستقراء الاحصائي لحالة معينة عنصرين أساسيين لا غنى عنهما معاً وهما (أ) الاستقراء و(ب) مقياس لجودة هذا الاستقراء .

والسؤال الذي يطرح نفسه هو أية طريقة نفضل في الاستقراء ، التقدير أم

اختبار الفرضيات ؟ والجواب هو أن الأمر يعود إلى طبيعة المسألة المطروحة وإلى التفضيل الشخصي . ولدينا بالنسبة للتقدير نوعان من التقدير هما التقدير النقطي والتقدير المجالي وسنتعرض لهما بالإضافة إلى اختبار الفرضيات في الفقرات التالية :

٧ـ٢ أنواع التقدير : يمكن تصنيف أساليب التقدير إلى نوعين : التقدير النقطي والتقدير المجالي . فلنفرض أننا نرغب في تقدير معدل طالب معين في جامعة دمشق فيمكن إعطاء التقدير على شكل رقم بمفرده ، مثلاً ، 72 . أو يمكن تقدير المعدل بأنه يقع ضمن مجال معين ، بين 70 و 75، على سبيل المثال . فالنوع الأول يسمى بالتقدير النقطي لأنه يمكن تمثيل العدد الوحيد الذي يمثل التقدير بنقطة على محور موجه . ويحوي النوع الثاني نقطتين وبالتالي يعرف مجالاً فوق المحور الموجه ولهذا يسمى بالتقدير المجالي .

في حالة التقدير النقطي نستخدم المعلومات المتوفرة في عينة للوصول إلى عدد واحد أو نقطة تكون تقديراً للوسيط المراد تقديره . وتسمى القاعدة التي تخبرنا عن كيفية حساب التقدير من المعلومات المتوفرة في العينة ، تسمى بالمقدِّر ونعبر عن المقدِّر ، بصورة عامة ، على شكل علاقة . فمتوسط العينة :

$$\bar{x} = \sum_{i}^{n} x_i / n$$

هو تقدير لمتوسط المجتمع μ . ويوضح تماماً طريقة حساب التقدير بدءاً من مقادير العينة $x_1, x_2, \dots, x_n$. وعلى الوجه الآخر يستخدم المقدِّر المجالي مقادير العينة لحساب نقطتين أو عددين يُفترض أن القيمة الحقيقية للوسيط تقع بينهما .

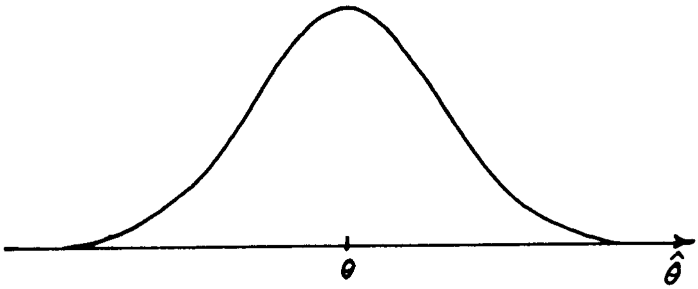
ونوضح المنطق المستخدم في حساب جودة تقدير نقطي بالتشبيه . فالتقدير النقطي يشبه من نواح عديدة الإطلاق على هدف من مسدس . فالمقدِّر في توليده للتقديرات يشبه المسدس . والطلقة هي تقدير معين ، والوسيط المراد تقديره هو الهدف . وسحب عينة من المجتمع وتقدير الوسيط على أساسها يكافيء إطلاق طلقة واحدة على الهدف .

ولنفرض أن رجلاً أطلق طلقة واحدة على الهدف وأصابه تماماً فهل نستنتج أنه رام ماهر ؟ ومن الواضح أن الجواب لا ، فلا يقبل أي منا ، لو طُلب منه ذلك . أن يخاطر ويمسك الهدف بيده عندما يطلق الرجل طلقته الثانية . وعلى الوجه الآخر لو شاهدنا الرجل يطلق عشرة الآف طلقة على التتالي فتصيب الهدف جميعها . فقد نجد . ولقاء تعويض عادل ، من تتبلور في نفسه الثقة بمهارة الرامي بحيث يتبرع بأن يثبت الهدف بيده عند الطلقة 10001 . والنقطة التي نريد إثارتها أصبحت واضحة . إذ لا نستطيع تقييم جودة طريقة في التقدير على أساس تقدير واحد . ولا بد من مراقبة النتائج عند تطبيق الطريقة بصورة متكررة عدداً كبيراً من المرات . وعندئذ نلاحظ مدى قرب الطلقات وهي تتوزع حول مركز الهدف . وبما أن التقديرات هي أعداد فيمكن تقييم جودة المقدِّر بإقامة توزيع تكراري للتقديرات التي نحصل عليها من عينات متكررة ونلاحظ مدى قرب مركز هذا التوزيع من قيمة الوسيط أو مدى تمركز هذا التوزيع حول قيمة الوسيط .

وتوضح نتائج تجربة قذف قطعة زهر ، في الفقرة (٦-٢) الفكرة المعروضة هنا ، فلدينا هنـاك 100 عينة حجم كل منها n = 5 مسحوبة من مجتمع قذف قطعة زهر متوازنة ونعلم أن 3.5 = مـ و 1.71 = سـ في هذا المجتمع . ومتوسط كل عينة يمثل تقديراً لِـ مـ. ونلاحظ من الشكل (٦-٢) أن التوزيع التكراري لمتوسطات هذه العينات يتجمع حول المتوسط 3.5 = مـ. وبينما يقدم الشكل (٦-٢) معلومات أولية حول جودة التقدير . نريد في الحقيقة الشكل النظري لتوزيع التقديرات ، أي الشكل الموافق لما لا نهاية له من العينات ، أو بعبارة أخرى التوزيع الاحتمالي للتقدير . ومن حسن الحظ أن مثل هذا العمل ليس صعباً جداً . وتتوفر طرق رياضية لاشتقاق التوزيع الاحتمالي لتقدير ولكن مثل هذه الطرق فوق مستوى هذا الكتاب . والطريقة الأخرى المفيدة هي إستخدام العقول الالكترونية لسحب عدد كبير جداً من العينات ثم حساب التقدير

المطلوب من كل عينة وتسجيل النتائج على شكل توزيع تكراري .

لنفرض إذن أننا نرغب في تقدير وسيط لمجتمع سنرمز له بـِ θ وسنرمز لتقدير θ بـِ $\hat{\theta}$. وبمثال المسدس قائم في الذهن تتوضح أمامنا الخاصة المرغوبة للتقدير . إذ نريد أن يتمركز توزيع التقدير حول الوسيط المراد تقديره كما يبين الشكل (٧ ـ ١) . وبالإضافة إلى ذلك نرغب أن يكون انتشار (أو تشتت) التوزيع صغيراً قدر الإمكان . وبعبارة أخرى نريد أن تكون قيمة توقع التقدير

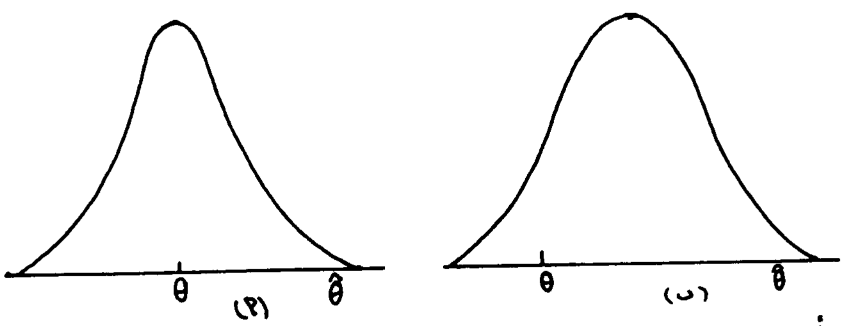


شكل ٧ـ١ توزيع تقدير

مساوية لقيمة الوسيط المراد تقديره أي :

$$E(\hat{\theta}) = \theta \qquad (١)$$

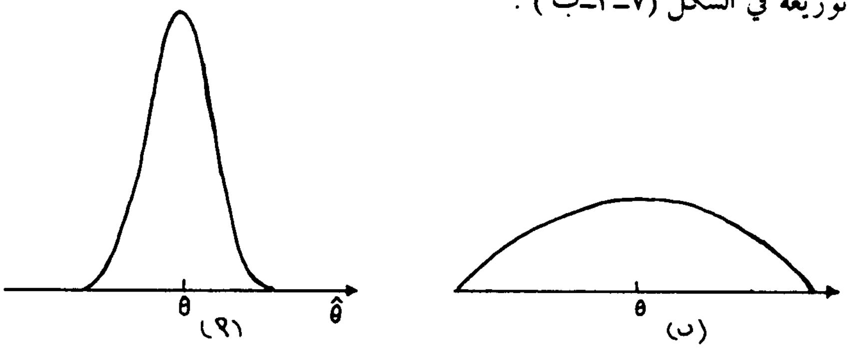
وتدعى التقديرات التي تحقق هذه الخاصة بالتقديرات المنصفة . وإذا لم يحقق التقدير هذه الخاصة قلنا أنه تقدير منحاز . ويبين الشكلان (٧ـ٢ـا) و(٧ـ٢ـب) توزيعي تقديرين أحدهما منحاز والآخر منصف .



(ب) (أ)

شكل ٧ ـ ٢ توزيعي تقديرين أحدهما منصف والآخر منحاز

كما نرغب بأن يكون تشتت تقدير أو انحرافه المعياري أصغر ما يمكن . وهكذا يكون التقدير المبين توزيعه في الشكل (٧ـ٣ـأ) أفضل من ذلك المبين توزيعه في الشكل (٧ـ٣ـب) .



(ب) (أ)

شكل ٧ـ٣ مقارنة تشتت تقديرين

ونحلل بنفس الطريقة جودة تقدير مجالي . إذ نسحب عينات لها نفس الحجم من المجتمع بصورة متكررة ونحسب في كل مرة التقدير المجالي . وستولّد هذه الطريقة عدداً كبيراً من المجالات (بدلاً من النقاط في حالة التقدير النقطي) والطريقة الجيدة للتقدير المجالي أو المقدِّر المجالي الجيد هو المقدِّر الذي يحتوي القيمة الحقيقية للوسيط بنسبة عالية أي أن نسبة عالية من بين المجالات التي يولّدها هذا المقدِّر عند سحب العينات المتكررة تحوي القيمة الحقيقية للوسيط . وتسمى هذه النسبة بأمثال الثقة في التقدير بينما يسمى التقدير نفسه بمجال الثقة .

واختبار « أفضل » تقدير يشمل المقارنة بين طرق مختلفة للتقدير . وهذه المهمة هي من مهام الإحصاء النظري وهي فوق مستوى هذا الكتاب .

٧ـ٣ **التقدير النقطي لمتوسط مجتمع** : تقود الكثير من المسائل التطبيقية إلى تقدير متوسط مجتمع μ . فنحن نهتم مثلاً بالمعدل العام لطلبه كلية ، بالقوة الوسطية لنوع جديد من الفولاذ ، بالنسبه الوسطية للوفيات في طبقة إجتماعية معينة ، وفي معدل الطلب على بضاعة جديدة مطروحة للاستهلاك . ويخدم

تقدير μ كتطبيق عملي جداً لمسألة الاستقراء الإحصائي ، وكتوضيح ممتاز لمبادىء التقدير التي ناقشناها في الفقرة السابقة . ويتوفر العديد من المقدِّرات لمتوسط مجتمع μ ، ومنها وسط العينة والمعدل الحسابي لأكبر وأصغر قيمة في العينة ، ومتوسط العينة $\bar{x}$. وكل من هذه المقدِّرات ستولّد عند تكرار سحب العينة توزيعاً ، يعتمد على المجتمع الذي نسحب منه العينة وطبيعة المسألة التطبيقية المطروحة ، وسيكون لكل من هذه التوزيعات محاسنه ومساوئه . ومع أن حساب الوسط ومعدل القيمتين الكبرى والصغرى في العينة أسهل ، إلا أن متوسط العينة $\bar{x}$ أفضل منهما لأن تشتته من أجل العديد من المجتمعات يكون أصغر . بالإضافة إلى أنه ، وبصرف النظر عن تكوين المجتمع ، منصف دائماً .

وتبرز ثلاث حقائق هامة من دراسة التوزيع الاحتمالي لِ $\bar{x}$ عند تكرار سحب عينة عشوائية حجمها n من مجتمع متوسطه μ وتشتته σ^2 وذلك بصرف النظر عن طبيعة هذا المجتمع :

(١) قيمة توقع $\bar{x}$ أو متوسط $\bar{x}$ يساوي متوسط المجتمع μ أي $E(\bar{x}) = \mu$

(٢) الإنحراف المعياري لِ $\bar{x}$ هو :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \qquad (2)$$

حيث N عدد قياسات المجتمع و n حجم العينة المسحوبة من هذا المجتمع .

وإذا كانت N كبيرة جداً فإن $\sqrt{\frac{N-n}{N-1}}$ تكون مساوية تقريباً للواحد و $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ وفي حالة مجتمع لا نهائي يكون :

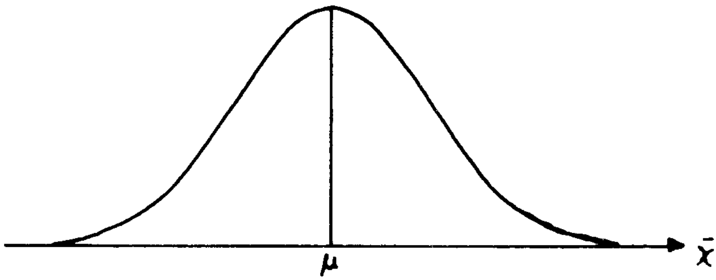
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad (3)$$

وذلك كما رأينا في نهاية الفصل الرابع .

(٣) عندما تكون n كبيرة بكفاية فإن التوزيع التقريبي لِ $\bar{x}$ هو التوزيع الطبيعي وذلك وفقاً لنظرية النهاية المركزية . (بفرض μ و σ كميات منتهية) .

وهكذا فإن $\bar{x}$ تقدير منصف لِ μ بإنحراف معياري متناسب طرداً مع الانحراف المعياري للمجتمع σ ومتناسب عكساً مع الجذر التربيعي لحجم

العينة n. وبالاضافة إلى معرفتنا بمتوسط التوزيع الاحتمالي لِ $\bar{x}$ وبانحرافه المعياري ، فإنّ نظرية النهاية المركزية تقدم معلومات حول شكل هذا التوزيع . فعندما يكون حجم العينة n كبيراً بكفاية سيكون شكل التوزيع الاحتمالي لِ $\bar{x}$ هو تقريباً شكل التوزيع الطبيعي المبين في الشكل (٧ـ٤) .



شكل ٧ـ٤ توزيع $\bar{x}$ من أجل n كبيرة .

ولنفرض ، مع حفظ هذه النتائج في الذهن ، أننا سحبنا عينة واحدة حجمها n = 5 من مجتمع قذف قطعة زهر الذي وصفناه في الفقرة (٦ـ٢) ، ولنحسب متوسط هذه العينة $\bar{x}$. فما هو مدى جودة مثل هذا التقدير لِ μ. أي كم سيكون مدى إنحرافها أو حيدانها عن القيمة الحقيقية لِ μ وهي 3.5 ؟ وبينما لا نستطيع الجزم بأن $\bar{x}$ ستقع ضمن مسافة معينة من μ، إلا أن نظرية تشيبيشيف تقول بأنه إذا سحبنا العديد من العينات من المجتمع فإن ثلاثة أرباعها على الأقل سيكون لها متوسط واقع ضمن مجال يمتد بمقدار $2\sigma_{\bar{x}}$ إلى يمين ويسار متوسط توزيع $\bar{x}$ أي μ. ونعلم أن $\sigma = 1.71$ وبالتالي :

$$2\sigma_{\bar{x}} = 2\frac{\sigma}{\sqrt{n}} = \frac{2(1.71)}{\sqrt{5}} = \frac{3.42}{2.24} = 1.53 .$$

وأكثر من ذلك فإننا نتوقع أن يكون توزيع متوسط العينة هو تقريباً التوزيع الطبيعي ، وفي هذه الحالة فإن 95% من التقديرات الناتجة عن عينات نسحبها على التوالي ستقع ضمن مجال يمتد بمقدار $2\sigma_{\bar{x}}$ أو 1.53 إلى يمين ويسار μ. ونظرة فاحصة على الشكل (٦ـ٢) ستؤكد هذه النتيجة .

والكمية $2\sigma_{\bar{x}}$ هي حد تقريبي لخطأ التقدير . ونقصد بذلك أن ثلاثة أرباع التقديرات على الأقل ، وفي الغالب ، 95% منها سيحيد عن المتوسط بأقل من $2\sigma_{\bar{x}}$. ويبدو اختيار الكمية $2\sigma_{\bar{x}}$ كحد للخطأ بدلاً من $3\sigma_{\bar{x}}$ أمراً معقولاً من أجل معظم المسائل التطبيقية .

مثال ٧ـ١ لنفرض أننا نرغب في تقدير متوسط الانتاج اليومي في منشأة للصناعات الكيميائية . وقد سجلنا الانتاج اليومي لفترة $n = 50$ يوماً فكان متوسط هذه العينة وانحرافها المعياري بالأطنان :

$$\bar{x} = 871$$

$$S = 21$$

والمطلوب تقدير متوسط الانتاج μ.

التقدير الأفضل هو $\bar{x} = 871$ طن يومياً . وحدود الخطأ على هذا التقدير هي :

$$\pm 2\sigma_{\bar{x}} = \pm 2\frac{\sigma}{\sqrt{n}} = \frac{2\sigma}{\sqrt{50}}$$

ومع أن σ مجهول إلا أنه يمكن اعتبار s كقيمة تقريبية له ، أي اعتبار s تقديراً لـ σ. وهكذا تكون حدود خطأ التقدير بصورة تقريبية مساوية لِ :

$$2s/\sqrt{50} = \frac{2(21)}{\sqrt{50}} = 5.94 .$$

وسنشعر إلى حد ما بالثقة بأن التقدير 871 هو في حدود 5.94 طن من القيمة الحقيقية لمتوسط الإنتاج .

ويجدر التنويه بنقطتين أولهما هي أن استخدام 2σ كحدّ للخطأ بدلاً من $2\sigma_{\bar{x}}$ هو خطأ شائع بالنسبة للمبتدئين يجب الانتباه إليه . والنقطة الثانية المهمة هي إستخدام s كتقريب لِ σ. فسيكون مثل هذا التقريب جيداً عندما يكون n كبيراً ، مثلاً 30 أو أكثر . أما إذا كان حجم العينة صغيراً فهناك طريقتان متوفرتان ، فأحياناً ستقدم لنا خبرتنا المستمدة من البيانات الإحصائية لتجارب سابقة تقديراً جيداً لِ σ وعندما لا تتوفر مثل هذه الخبرة يمكن اللجوء إلى

الطرق الإحصائية الخاصة بالعينات الصغيرة مما سنناقشه في الفصل القادم .

٧ـ٤ التقدير المجالي لمتوسط مجتمع : يمكن إستخدام نتائج الفقرة السابقة للحصول بسهولة على المقدّر المجالي أو « مجال الثقة » لمتوسط مجتمع . ومن الممكن أن يقع $\bar{x}$ إما فوق أو تحت متوسط المجتمع ، ولكننا لا نتوقع أن يحيد $\bar{x}$ عن μ بأكثر من $2\sigma_{\bar{x}}$ تقريباً ، وهكذا إذا اخترنا ($\bar{x}-2\sigma_{\bar{x}}$) لتكون الطرف الأيسر من مجال الثقة ويدعى الحد الأدنى للثقة ، و ($\bar{x}+2\sigma_{\bar{x}}$) لتكون الطرف الأيمن للمجال أو الحد الأعلى للثقة . فإن هذا المجال يصبح $(\bar{x}-2\sigma_{\bar{x}}, \bar{x}+2\sigma_{\bar{x}})$ وسيكون احتمال أن يتضمن هذا المجال القيمة الحقيقية لِ μ كبيراً .

وفي الحقيقة إذا كان n كبيراً وتوزيع $\bar{x}$ هو تقريباً التوزيع الطبيعي فيمكننا أن نتوقع ، بصورة تقريبية ، أن 95% من المجالات التي نحصل عليها عند سحب عينات متكررة ، ستتضمن القيمة الحقيقية لِ μ

وتدعى مجالات الثقة هذه بمجالات الثقة من أجل العينات الكبيرة (أو حدود الثقة) ذلك لأنه يجب أن تكون العينة كبيرة بكفاية حتى نتمكن من تطبيق نظرية النهاية المركزية واعتبار أن التوزيع التقريبي لِ $\bar{x}$ هو التوزيع الطبيعي وهذا ما يمكننا أصلاً من استنتاج مجال الثقة الموافق لأمثال ثقة يساوي 0.95 . إذ نتمكن عندئذ من كتابة :

$$P(-c\sigma_{\bar{x}} \leqslant \bar{x} \leqslant c\sigma_{\bar{x}}) = P\left(-c \leqslant \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \leq c\right)$$

$$= P(-c \leqslant z \leqslant c) = .95 \qquad (4)$$

وقد وجدنا في التمرين (٢) من الفصل السابق وباستخدام جدول التوزيع الطبيعي أن $c = 1.96$ واعتبرنا في المناقشة السابقة أن $c = 2$. وبدقة أكبر نقول إذن أن 95% مجال ثقة لِ $\bar{x}$ هو المجال $(\bar{x}-1.96\sigma_{\bar{x}}, \bar{x}+1.96\sigma_{\bar{x}})$.

وبنفس الطريقة إذا أردنا استنتاج مجال الثقة الموافق لأمثال ثقة 90 . نكتب :

$$P(-c\sigma_{\bar{x}} \leq \bar{x} \leq c\sigma_{\bar{x}}) = P(-c \leqslant z \leqslant c) = .90$$

وقد وجدنا في التمرين (٢) من الفصل السابق أن $c = 1.645$ أي أن 90% مجال

ثقة لِ $\bar{x}$ هو : $(\bar{x} - 1.645\,\sigma_{\bar{x}} \,,\, \bar{x} + 1.645\,\sigma_{\bar{x}})$

وبصورة عامة لحساب $(1-\alpha)$ مجال ثقة لِ $\bar{x}$ نكتب :

$$P(-c\,\sigma_{\bar{x}} \leqslant \bar{x} \leqslant +c\,\sigma_{\bar{x}}) = P(-c \leqslant Z \leqslant c) = 1-\alpha \quad (5)$$

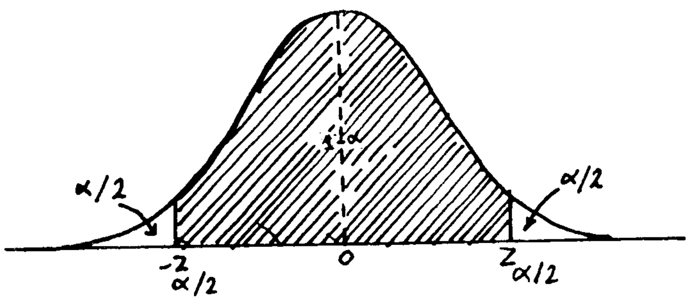
أو :

$$P(0 \leqslant z \leqslant c) = \tfrac{1}{2}(1-\alpha) = .5 - \frac{\alpha}{2}$$

وهذا يعني أن :

$$P(z > c) = \frac{\alpha}{2}$$

أي أن c هي قيمة z في جدول التوزيع الطبيعي التي يقع $\frac{\alpha}{2}$ من المساحة الكلية تحت منحني الكثافة الطبيعي إلى يمينها ونرمز لمثل هذه النقطة عادة بِ $Z_{\alpha/2}$. وهي موضحة في الشكل (٧ـ٥) . وهكذا يكون الـ $(1-\alpha)$ مجال ثقة لـ $\bar{x}$ هو



الشكل ٧ـ٥ موقع $Z_{\frac{\alpha}{2}}$

$(\bar{x} - Z_{\alpha/2}\,\sigma_{\bar{x}} \,,\, \bar{x} + Z_{\alpha/2}\,\sigma_{\bar{x}})$ ، وبما أن $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ فنكتب هذا المجال عادة على الشكل :

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

مثال ٧ـ٢ أحسب 90% مجال ثقة لمتوسط المجتمع المذكور في المثال (٧ـ١) .

$Z_{\alpha/2}$ الموافقة لأمثال ثقة 90% هي 1.645 ومنه يكون المجال المطلوب :

$$\bar{x} \pm 1.645 \frac{\sigma}{\sqrt{n}}$$

وإذا وضعنا s بدلاً من σ نحصل بصورة تقريبية على المجال :

$$871 \pm (1.645)\,\frac{21}{\sqrt{50}}$$

أو :

$$871 \pm 4.89.$$

وهكذا نقدر أن متوسط الإنتاج اليومي μ يقع ضمن المجال من 866.11 إلى 875.89 طناً . وأمثال الثقة 90 . تعني أن 90% من مجالات الثقة الناتجة عن سحب عينات بصورة متكررة سيتضمن μ.

ونلاحظ أن عرض مجال الثقة يزداد مع ازدياد أمثال الثقة مما يتفق مع ما نتوقعه بالبداهة . فمن المؤكد أنه إذا رغبنا بثقة أكبر في مسألة احتواء المجال لِ μ ، فسينبغي زيادة عرض المجال . ونقدم في الجدول (٧ـ١) حدود الثقة الموافقة لبعض من أمثال الثقة الأكثر استخداماً في التطبيق العملي .

واختيار أمثال الثقة الذي سنستخدمه في حالة معينة متروك للمجرب ويعتمد على درجة الثقة التي يريد أن يضعها في تقديره . وقد أصبح بحكم المعتاد في الكثير من التجارب اختيار 0.95 . كأمثال ثقة مع أنه لا توجد أسس منطقية معينة لاستخدامه الواسع النطاق في التطبيقات العملية .

ونلاحظ أنه عندما نضع حدوداً لخطأ تقدير نقطي فإننا نقيم عمنياً مجال ثقة أو نصل إلى تقدير مجالي . وسيلاحظ القارىء الفرق الدقيق بين التقديرين المجالي والنقطي . فعندما يكون الوسيط المراد تقديره هو متوسط مجتمع يقع التقدير النقطي في منتصف مجال الثقة . وعلى أي حال فليس من الضروري أن يكون منتصف التقدير المجالي ، على الدوام ، أفضل تقدير نقطي . وفي العديد من الحالات ليس الأمر كذلك . ومن وجهة النظر العملية ، نقول أنه توجد صلة وشيجة بين التقديرين النقطي والمجالي ، واختيار أيهما في مسألة معينة يعتمد على تفضيل المجرب .

٧ـ٥ التقدير من عينات كبيرة : يمهد تقدير متوسط مجتمع ، الذي ناقشناه

جدول ٧ـ١ حدود الثقة لِـ μ

الحد الأعلى للثقة	الحد الأدنى للثقة	$z_{\alpha/2}$	أمثال الثقة
$\bar{x} + 1.645\ \sigma/\sqrt{n}$	$\bar{x} - 1.645\ \sigma/\sqrt{n}$	1.645	.90
$\bar{x} + 1.96\ \sigma/\sqrt{n}$	$\bar{x} - 1.96\ \sigma/\sqrt{n}$	1.96	.95
$\bar{x} + 2.58\ \sigma/\sqrt{n}$	$\bar{x} - 2.58\ \sigma/\sqrt{n}$	2.58	.99

في الفقرتين السابقتين ، الطريق من أجل مسائل تقدير أخرى سنناقشها في هذا الفصل . وهناك خيط يوحد فيما بينها ماراً عبرها جميعاً . وادراك مثل هذا الخيط سيسهل على المبتديء فهم الطرق المتبعة من أجل كل منها . إذ نفترض تحقق الشروط التالية من أجل كل مسائل التقدير التي نناقشها في هذا الفصل وهي أن كل تقدير نقطي $\hat{\theta}$ لوسيط θ يجب أن يكون منصفاً أي أن $E(\hat{\theta}) = \theta$

وأن الإنحراف المعياري للتقدير $\hat{\theta}$ معروف بحيث نستطيع وضع حدود لخطأ التقدير تساوي $2\sigma_{\hat{\theta}}$ وأخيراً نفترض في كل حالة ندرسها هنا أن التقدير النقطي يتبع بصورة تقريبية التوزيع الطبيعي ، وذلك عندما يكون حجم العينة n كبيراً بكفاية ، وذلك وفقاً لنظرية النهاية المركزية . وبالتالي فإن إحتمال أن يكون خطأ التقدير أقل من $1.96\sigma_{\hat{\theta}}$ سيكون على وجه التقريب 0.95

وسنفترض من أجل التقدير المجالي أن حجم العينة كبير بحيث يكفي من جهة لتطبيق نظرية النهاية المركزية والقول بأن التوزيع التقريبي للتقدير النقطي هو التوزيع الطبيعي ، كما يكفي من جهة أخرى لتقديم تقدير جيد لأي وسيط آخر غير معروف نحتاجه (مثل الوسيط σ) . وعندئذ تكون مجالات الثقة الموافقة لأية أمثال ثقة $(1-\alpha)$ هي :

$$\hat{\theta} \pm z_{\alpha/2}\, \sigma_{\hat{\theta}} \qquad (6)$$

٧ـ٦ تقدير الفرق بين متوسطين : توازي مسألة تقدير الفرق بين متوسطين في أهميتها مسألة تقدير متوسط مجتمع . فقد نرغب في مقارنة طريقتين في التعليم فنقسم الطلبة عشوائياً إلى زمرتين ونخضع كل منهما لطريقة في التعليم ثم نقوم

بالاستقراء حول الفرق بين تحصيلي الطلبة تحت كل من الطريقتين مقاساً بواسطة اختبار معين .

أو قد نرغب في مقارنة معدلي الانتاج في شركة للصناعة الكيميائية عند استخدام المواد الأولية الواردة من كل من اثنين من الممولين A و B . فنأخذ عينة من الانتاج اليومي في كل حالة ونستخدم المعلومات التي تقدمها هاتان العينتان للقيام باستقراء يتعلق بالفرق بين معدلي الانتاج .

ونفترض في كل من هذين المثالين وجود مجتمعين الأول متوسطه μ_1 وتشتته σ_1^2 والثاني متوسطه μ_2 وتشتته σ_2^2. سحبنا عينة عشوائية من n_1 قياساً من المجتمع I وعينة من n_2 قياساً من المجتمع II ونفرض أن العينتين مسحوبتان بحيث تكون كل منهما مستقلة عن الأخرى . ثم نحسب تقديرات وسطاء المجتمعين وهي $\bar{x}_1$ ، S_1^2 من العينة الأولى و $\bar{x}_2$ ، S_2^2 من العينة الثانية . والتقدير النقطي للفرق بين متوسطي المجتمعين $\mu_1 - \mu_2$ هو $\bar{x}_1 - \bar{x}_2$.

وإذا سحبنا زوجاً من العينات بصورة متكررة من المجتمعين ، بحيث يبقى حجم العينة الأولى n_1 وحجم العينة الثانية n_2 ، ثم حسبنا في كل مرة التقدير $\bar{x}_1 - \bar{x}_2$ ، فسنحصل على توزيع لهذا التقدير . وكما نعلم من خواص التوقع والتشتت في الفقرة (٤ـ٥) فإن :

$$E(x_1 - x_2) = E(x_1) + E(-x_2) = E(x_1) - E(x_2) = \mu_1 - \mu_2 \quad (7)$$

وبما أن العينتين مستقلتان فإن $\bar{x}_1$ و $\bar{x}_2$ مستقلان ويمكننا كتابة :

$$V(\bar{x}_1 - \bar{x}_2) = V(\bar{x}_1) + V(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad (8)$$

ومنه :

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (9)$$

وبالإضافة إلى ذلك فإنه إذا كانت n_1 و n_2 كبيرتين ، مثلاً 30 أو أكثر ، فإن نظرية النهاية المركزية تقول بأن التوزيع التقريبي لكل من $\bar{x}_1$ و $\bar{x}_2$

هو التوزيع الطبيعي . وإحدى الخواص الهامة للتوزيع الطبيعي ، والتي تُبرهن في الإحصاء النظري ، هي أنه إذا كان كل من المتحولين المستقلين y_1 و y_2 يتبع التوزيع الطبيعي فإن التركيب $c_1 y_1 + c_2 y_2$ حيث c_1 و c_2 أعداد ثابتة يتبع أيضاً التوزيع الطبيعي . وإذا اعتبرنا $c_1 = 1$ و $c_2 = -1$ فإن الفضل $y_1 - y_2$ يتبع التوزيع الطبيعي . وهكذا نجد أنه يمكن اعتبار التوزيع التقريبي لِـ $\bar{x}_1 - \bar{x}_2$ هو التوزيع الطبيعي .

وحدود خطأ التقدير $\bar{x}_1 - \bar{x}_2$ هي إذن $2\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. ولإقامة $(1-\alpha)$ مجال ثقة من أجل $(\mu_1 - \mu_2)$ نكتب :

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10)$$

ومن أجل n_1 و n_2 أكبر من 30 يمكن استبدال σ_1^2 و σ_2^2 بتقديريهما S_1^2 و S_2^2 كما نجدهما من العينة . .

مثال ٧ـ٣ قمنا بمقارنة نوعين من إطارات السيارات باختبار عملي يتضمن عينة حجمها $n = 100$ إطار من كل نوع . وسجلنا عدد الأميال التي يخدمها الإطار حتى إهترائه وفقاً لمقاييس محددة سلفاً . وكانت نتائج الاختبار بالأميال كما يلي :

$$\bar{x}_1 = 26400 \qquad \bar{x}_2 = 25100.$$

$$S_1^2 = 1440000 \qquad S_2^2 = 1960000$$

والمطلوب تقدير الفرق بين متوسطي العمر في النوعين ووضع حدود لخطأ التقدير .

التقدير النقطي لِـ $(\mu_1 - \mu_2)$ ، حيث μ_1 هو العمر الوسطي لإطار من النوع الأول و μ_2 العمر الوسطي للإطار من النوع الثاني ، هو $(\bar{x}_1 - \bar{x}_2)$ أي :

$$x_1 - x_2 = 26400 - 25100 = 1300$$

ومنه :

$$\sigma_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \cdot \sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} = \sqrt{\frac{1440000}{100}+\frac{1960000}{100}}$$

$$= \sqrt{34000} = 184$$

ونتوقع أن تكون حدود الخطأ حوالي $2\sigma_{(\bar{x}_1-\bar{x}_2)}$ أو 368 ميلاً . وهكذا يبدو النوع الأول متفوقاً على النوع الثاني .

مثال ٧ـ٤ : ضع مجال ثقة من أجل الفرق $\mu_1-\mu_2$ المذكور في المثال السابق . مستخدماً أمثال ثقة تساوي 0.99 .

نعلم أن $Z_{\alpha/2} = .99$ ومنه يكون مجال الثقة المطلوب :

$$(\bar{x}_1 - x_2) \pm 2.58 \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

ومستخدمين نتائج المثال (٧ـ٣) نجد أن مجال الثقة هو :

$$1300 \pm 2.58\,(184)$$

أي أن حد الثقة الأدنى هو 825 وحد الثقة الأعلى 1775 . أي أنه يمكن تقدير الفرق بين متوسطي العمرين بأنه يقع بين هذين الحدين .

٧ـ٧ **تقدير متوسط مجتمع ثنائي** : إن أفضل تقدير نقطي لوسيط المجتمع الثنائي p هو أيضاً ما تمليه البداهة وهو متوسط عدد النجاحات خلال n تكراراً . أي أن التقدير $\hat{p}$ هو :

$$\hat{p} = \frac{x}{n},$$

حيث x عدد النجاحات و n عدد التكرارات . وبعبارة « أفضل تقدير » نقصد أن التقدير $\hat{p}$ منصف وتشتته بالمقارنة مع تقديرات منصفة أخرى هو أصغر ما يمكن .

ووفقاً لنظرية النهاية المركزية فإن التوزيع التقريبي لِـ x هو التوزيع الطبيعي بمتوسط يساوي np وتشتت يساوي np q . وذلك من أجل n كبيرة بكفاية . وبما أن $\frac{1}{n}$ عدد ثابت فوفقاً لخاصة التوزيع الطبيعي المذكورة

في الفقرة السابقة يكون التوزيع التقريبي لِ $\hat{p} = \frac{1}{n}(x)$ هو أيضاً التوزيع الطبيعي بمتوسط يساوي :

$$E(\hat{p}) = E\left(\frac{x}{n}\right) = \frac{1}{n} E(x) = \frac{1}{n} \cdot np = p . \qquad (11)$$

وتشتت يساوي :

$$\sigma_{\hat{p}}^2 = \sigma_{\frac{x}{n}}^2 = \frac{1}{n^2}\sigma_x^2 = \frac{npq}{n^2} = \frac{pq}{n} . \qquad (12)$$

ومنه :

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} \qquad (13)$$

وحدود خطأ التقدير النقطي هي :

$$2\sqrt{\frac{pq}{n}} \qquad (14)$$

ولحساب $(1-\alpha)$ مجال ثقة لِ p نكتب من أجل n كبيرة بكفاية :

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{pq}{n}} \qquad (15)$$

والصعوبة التي نواجهها هنا هي حساب $\sigma_{\hat{p}}$ الذي يعتمد على p وعلى $q = 1 - p$ وهما مجهولان . وإذا وضعنا $\hat{p}$ و $\hat{q}$ بدلاً من p و q في عبارة الانحراف المعياري $\sqrt{\frac{pq}{n}}$ فإن الخطأ المرتكب سيكون صغيراً جداً من أجل n كبيرة . وفي الحقيقة يتغير الانحراف المعياري ببطء شديد عندما تتحول p . ويمكن ملاحظة ذلك بوضوح من الجدول (٧-٢) حيث سجلنا $\sqrt{pq}$ من أجل قيم متعددة لِ p . ويجدر الانتباه إلى أن تغير $\sqrt{pq}$ طفيف جداً من أجل قيم لِ p قريبة من 0.5 .

جدول ٧-٢ بعض القيم لِ

p	$\sqrt{pq}$
.5	.5
.4	.49
.3	.46
.2	.40
.1	.30

مثال ٧ـ٥ انتجت عينة عشوائية من 100 ناخب في منطقة معينة $x = 59$ ناخباً للمرشح A . والمطلوب تقدير نسبة الناخبين في المنطقة الذين يفضلون A ووضع حدود لخطأ التقدير .

التقدير النقطي هو :

$$\hat{p} = \frac{x}{n} = \frac{59}{100} = .59$$

وحدود خطأ التقدير هي حوالي :

$$2\sigma_{\hat{p}} = 2\sqrt{\frac{pq}{n}} \doteq 2\sqrt{\frac{(.59)(.4)}{100}} = .096$$

و 95% مجال ثقة من أجل p هو :

$$\hat{p} \pm 1.96\sqrt{\frac{pq}{n}}$$

أو على وجه التقريب :

$$.59 \pm 1.96\ (.048)$$

وهكذا نقدّر أن تقع p ضمن المجال 496. إلى 684. بأمثال ثقة هي 0.95 .

٧ـ٨ تقدير الفرق بين وسيطي مجتمعين ثنائيين: لنفرض المجتمعين الثنائيين I وسيطه p_1 و II ووسيطه p_2 على التوالي . ولنسحب عينتين عشوائيتين من المجتمعين ونحسب منهما التقديرين $\hat{p}_1$ و $\hat{p}_2$ وليكن حجم الأولى n_1 وحجم الثانية n_2 . فالتقدير النقطي $(\hat{p}_1 - \hat{p}_2)$ للفرق $(p_1 - p_2)$ هو تقدير منصف أي أن :

$$E(\hat{p}_1 - \hat{p}_2) = (p_1 - p_2) \qquad (16)$$

وانحرافه المعياري هو :

$$\sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \qquad (17)$$

وحدود خطأ هذا التقدير هي ، من أجل n_1 و n_2 كبيران بكفاية ، حوالي :

$$2\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \qquad (18)$$

ويمكن تعويض p_1 ، p_2 بـ $\hat{p}_1$ و $\hat{p}_2$ على الترتيب .

و مجال ثقة من أجل الفرق $(p_1 - p_2)$ هو :

(19)

وذلك بفرض أن n_1 و n_2 كبيران بكفاية . ويمكننا هنا أيضاً تعويض p_1 و p_2 بـ $\hat{p}_1$ و $\hat{p}_2$ على الترتيب .

مثال ٧ـ٦ تصنع شركة محلولين لمكافحة الذباب I و II وترغب في مقارنة فاعليتهما . ولهذا الغرض استخدمت غرفتين من نفس الحجم تحوي كل منهما 1000 ذبابة وإحداهما عولجت بالمحلول I والثانية بنفس الكمية من المحلول II . وقد وُجد أن المحلول أهلك 825 ذبابة والمحلول الثاني أهلك 760 ذبابة . والمطلوب تقدير الفرق بين قدرتي المحلولين على إبادة الذباب عند استخدامهما في نفس الشروط المحيطية .

إن تعرّض كل ذبابة لهذا المحلول هي تكرار مستقل من مجتمع ثنائي وإذا فرضنا أن احتمال النجاح (أي هلاك الذبابة) هو p_1 بالنسبة للمحلول I و p_2 بالنسبة للمحلول II فإن المطلوب هو تقدير الفرق $(p_1 - p_2)$ وبالاستناد إلى العلاقة (16) نجد :

$$\hat{p}_1 - \hat{p}_2 = .825 - .760 = .065.$$

وحدود خطأ هذا التقدير هي حوالي :

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

ومن أجل أمثال للثقة تساوي .95 نجد :

$$\mp 2 \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{2}} = 2 \sqrt{\frac{(.825)(.175)}{1000} + \frac{(.76)(.24)}{1000}} = .036$$

أي .

$$.065 \pm .036.$$

وهكذا نقدر أن الفرق في القدرة على إبادة الذباب $p_1 - p_2$ يقع ضمن المجال

.029 إلى 0.101 . وثقتنا بمثل هذا التقدير ناتجة عن معرفتنا بأنه إذا أعدنا نفس التجربة مراراً وتكراراً وفي كل مرة حسبنا تقديراً مجالياً فإن 95% تقريباً من هذه المجالات ستحوي القيمة الحقيقية للفرق $p_1 - p_2$.

٧ـ٩ اختيار حجم العينة : تصميم تجربة هو في الأساس خطة لشراء كمية من المعلومات يمكن أن تكون ، مثلها مثل أي سلعة أخرى ، بأسعار مختلفة وفقاً للطريقة التي اعتمدنا هافي الحصول على المعلومات . فبعض القياسات يمكن أن تحتوي كمية كبيرة من المعلومات حول الوسيط المدروس بينما تحوي قياسات أخرى قدراً أقل ، أو أنها لا تحوي أيه معلومات بالمرَّة . وبما أن الانتاج الوحيد للبحث العلمي هو المعلومات الجديدة فإن ذلك يحدونا إلى جعل عملية شرائها بأقل كلفة ممكنة .

وتؤثر طريقة أو أسلوب أخذ العينة أو نوع التصميم التجريبي في كمية المعلومات التي نحصل عليها لقاء ملاحظة واحدة . وهي تتحكم مع حجم العينة n في الكمية الإجمالية للمعلومات المتوفرة في العينة . وبصورة عامة ، سنهتم هنا بأبسط حالات سحب العينة وهي العينة العشوائية البسيطة مأخوذة من مجتمع كبير نسبياً ، وسنركز انتباهنا على اختيار حجم العينة n .

وفي طليعة ما يواجه الباحث العلمي عند تخطيط تجربته هو اختيار حجم العينة . وربما كان أكثر التساؤلات تواتراً في المسائل الاحصائية هو : « كم قياساً يجب أن تتضمن العينة ؟ » ومن سوء الحظ لا يستطيع الاحصائي أن يجيب على هذا السؤال دون أن يعرف كمية المعلومات التي يرغب المجرب في شرائها . ومن المؤكد أن الكمية الإجمالية التي تحتويها العينة من المعلومات ستؤثر في مقياس جودة الاستقراء ولا بد من أن يحددها المجرب . وبالاشارة إلى مسألة التقدير ، بصورة خاصة ، فإننا نريد معرفة مدى الدقة التي يرغبها المجرب للتقدير المطلوب . ويمكن التعبير عن ذلك بوضع حد لخطأ التقدير .

فمثلاً لنفرض أننا نرغب في تقدير المعدل اليومي لإنتاج كيميائي

(مثال ٧ـ١) ، ونرغب في أن يكون خطأ التقدير 10 طن باحتمال 0.95 . وبما أن 95% تقريباً من متوسطات العينات سيقع ضمن مجال $2\sigma_{\bar{x}}$ على يمين ويسار μ . فالمطلوب هو أن يكون $2\sigma_{\bar{x}}$ مساوياً 10 طن . (أنظر الشكل ٧ـ٦) وهكذا نكتب :

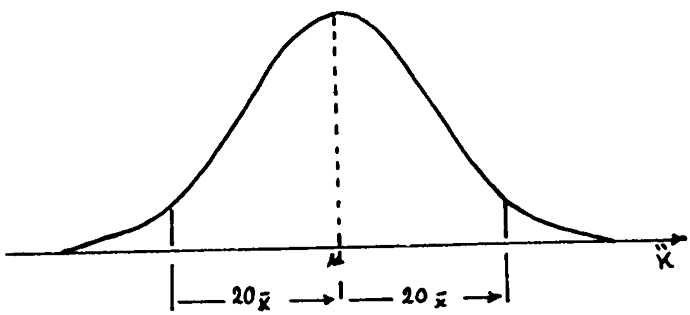
$$2\sigma_{\bar{x}} = 10$$

أو :

$$2\sigma/\sqrt{n} = 10$$

وبحل هذه المعادلة من أجل n نجد :

$$n = \sigma^2/25$$



شكل ٧ـ٦ التوزيع التقريبي لـ $\bar{x}$ من أجل عينات كبيرة .

وسيلاحظ القارىء أنه ما لم نكن نعرف الانحراف المعياري σ لا يمكننا الحصول على قيمة عددية لـ n . وإذا لم نكن نعرف من خبرة سابقة قيمة σ فلا بد من استخدام قيمة تقديرية له هي s الانحراف المعياري للعينة . وفي مثالنا هنا يمكن استخدام قيمة s كما وجدناها في المثال ٧ـ١ وهي s = 21 بدلاً من σ. وعندئذ نجد :

$$n = \frac{\sigma^2}{25} = \frac{(21)^2}{25} = 176.4$$

أو :

$$n = 177$$

وباستخدام عينة حجمها n = 177 سنكون على ثقة كبيرة (باحتمال يساوي

تقريبا 95.) بأن تقديرنا سيقع ضمن $2\sigma_{\bar{x}}=10$ طن من المعدل الحقيقي للانتاج اليومي .

وفي الواقع يمكن أن نتوقع كون خطأ التقدير أقل بكثير من 10 طن . ووفقاً للقاعدة التجريبية فإن احتمال أن يكون خطأ التقدير أقل من $\sigma_{\bar{x}}=5$ طناً يساوي تقريباً 0.68 . ومع أن هذه الطريقة في اختيار حجم العينة تقريبية إلا أنها أفضل ما يتوفر لنا وهي على وجه التأكيد أفضل من اختيار حجم العينة معتمدين على البداهة .

ونلخص الآن الطريقة العامة لاختيار حجم العينة في مسائل التقدير التي تستخدم عينات كبيرة الحجم . فلا بد أن يحدد المجرب حدّ الخطأ الذي يرغبه التقدير أو أمثال ثقة $(1-\alpha)$ للتقدير المجالي . فمثلاً إذا كان الوسيط θ والحد المرغوب B ، فنكتب :

$$Z_{\alpha/2}\,\sigma_{\hat{\theta}} = B$$

وهذه المعادلة تحوي n وبحلّها نحصل على القيمة المطلوبة لِ n . وسنوضح ببعض الأمثلة :

مثال ٧ـ٧ يمكن أن يأخذ رد فعل شخص لمنشط في تجربة نفسية أحد شكلين A أو B . فإذا رغب المجرب في تقدير الاحتمال p بأن رد الفعل سيكون من النوع A ، فكم شخصاً يجب أن تشمل التجربة ؟ ولنفرض أننا سنرضى بخطأ تقدير لا يتجاوز 04. وذلك باحتمال يبلغ 90. كما أننا نتوقع أن تكون النسبة p في مكان ما إلى جوار 0.6 .

بما أن أمثال الثقة $1-\alpha$ تساوي 0.9 فإن $\alpha=.10$ ، $\frac{\alpha}{2}=.05$.

و $Z_{\alpha/2}=1.645$ حيث $Z_{\alpha/2}$ هي قيمة المتحول الطبيعي المعياري التي يقع على يمينها $\alpha/2$ من المساحة الكلية تحت منحني الكثافة الطبيعي . والمطلوب هو أن يكون :

$$1.645\,\sigma_{\hat{p}} = .04$$

$$1.645\sqrt{\frac{pq}{n}} = .04$$

وبما أن تشتت $\hat{p}$ يعتمد على p المجهولة ، فسنضع القيمة التي ضمّنها المجرب بصورة تقريبية لِ p وهي 0.6 . وعندئذ نجد :

أو :

$$1.645\sqrt{\frac{(.6)(.4)}{n}} = .04$$

$$n = 406$$

مثال ٧ــ٨ يرغب مجرب في مقارنة فعالية طريقتين لتدريب مهني لعمال صناعيين على إنجاز عملية تجميع جهاز معين . وهكذا يقسم عدداً من العمال إلى زمرتين متساويتين تتلقى الأولى الطريقة 1 في التدريب وتتلقى الثانية الطريقة 2 . وستنجز كل زمرة عملية التجميع ونسجل الزمن اللازم من أجل كل فرد فيها .

ونتوقع أن تكون القياسات للفترة الزمنية اللازمة في كل من الزمرتين في حدود ثماني دقائق . فإذا أردنا أن يكون تقديرنا للفرق بين متوسطي الفترة الزمنية للتجميع صحيحاً في حدود دقيقة واحدة وذلك باحتمال 0.95 ، فكم من العمال يجب أن تحتوي كل زمرة ؟

بوضع حد الخطأ مساوياً 1 نجد :

$$2\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} = 1$$

وكما فرضنا فإن التشتت ضمن كل من الزمرتين هو نفسه أي $\sigma_1^2 = \sigma_2^2 = \sigma^2$

وبما أن المدى الذي نتوقعه وهو 8 دقائق يساوي تقريباً أربعة أمثال σ (أنظر القاعدة التجريبية في الفصل الثاني) فيمكن بصورة تقريبية تخمين قيمة σ على أنها تساوي 2 . وبالتبديل نجد :

ومنه :

$$2\sqrt{\frac{4}{n} + \frac{4}{n}} = 1$$

$$n = 32$$

أي أن كل زمرة يجب أن تحوي 32 عاملاً .

٧ـ١٠ الاختبار الاحصائي لفرضية : هدف الاختبار الاحصائي هو اختبار فرضية تتعلق بقيم وسيط أو أكثر . ويحتوي الاختبار الاحصائي على أربعة عناصر :

(١) الفرضية الابتدائية .

(٢) إحصاء الاختبار .

(٣) منطقة الرفض .

(٤) الفرضية البديلة .

وتحديد هذه العناصر الأربعة يعرف اختباراً معيناً وتغيير عنصر أو أكثر يؤدي إلى اختبار جديد .

الفرضية الابتدائية ، ونرمز لها بـ H_0 ، تعرض الفرضية التي سيجري اختبارها . أي أنها تحدد قيماً افتراضيةلوسيط أو أكثر من وسطاء المجتمع . فقد نرغب مثلاً في اختيار الفرضية بأن متوسط المجتمع يساوي 50 ، أو أن متوسطي مجتمعين μ_1 و μ_2 متساويان .

وقرارنا برفض أو قبول فرضية إبتدائية مبني على المعلومات التي تحويها عينة سحبناها من المجتمع المدروس . وتُستخدم قيم العينة لحساب عدد واحد يأخذ دور صانع القرار ويدعى بإحصاء الاختبار . ونقسم مجموعة كل القيم التي يمكن أن يأخذها إحصاء الاختبار إلى مجموعتين أو منطقتين إحداهما تدعى منطقة الرفض وتدعى الأخرى منطقة القبول . ونرفض الفرضية الابتدائية إذا وقعت القيمة التي يأخذها إحصاء الاختبار ، محسوباً من العينة التي بين أيدينا ، في منطقة الرفض . ونقبل الفرضية إذا وقعت هذه القيمة في منطقة القبول .

وتخضع طريقة اتخاذ القرار هذه لنوعين من الخطأ ، إذ يمكن أن نرفض الفرضية الابتدائية بينما هي في الواقع صحيحة ، أو يمكن أن نقبل H_0 وهي في الحقيقة غير صحيحة وإنما الصحيح هو فرضية أخرى بديلة . ويدعى هذان

الخطآن بالخطأ من النوع الأول I والخطأ من النوع الثاني II على الترتيب . أي أن رفض فرضية صحيحة (أو الرفض الخاطيء) هو الخطأ من النوع الأول وقبول فرضية غير صحيحة (أي القبول الخاطيء) هو الخطأ من النوع الثاني . ويوضح الجدول (٧ـ٣) الحالتين الممكنتين للفرضية الابتدائية (صحيحة أو خاطئة) ونوعي القرار الممكنين (رفض أو قبول) بالإضافة إلى نوعي الخطأ الذي يمكن إرتكابه .

جدول ٧ـ٣ جدول القرارات

الفرضية الابتدائية		القرار
مخطئة	صحيحة	
قرار صحيح	الخطأ من النوع I	الرفض
الخطأ من النوع II	قرار صحيح	القبول

ونقيس جودة اختبار احصائي لفرضية باحتمالي الخطأ من النوع الأول I والخطأ من النوع الثاني II ، ونرمز لهما بـ α و β على التوالي . ويمكن التعبير عن الاحتمال α بأنه احتمال أن يقع إحصاء الاختبار في منطقة الرفض علماً أن H_0 صحيحة . وتدعى α أيضاً بمستوى أهمية الاختبار . ومن الواضح إذن أن زيادة حجم منطقة الرفض سيزيد من قيمة α وفي نفس الوقت يؤدي إلى تناقص β وذلك من أجل حجم ثابت للعينة n . أما تخفيض حجم منطقة الرفض فيؤدي إلى تناقص α وزيادة β. وعند زيادة حجم العينة n فمن الطبيعي أن توفر لنا العينة الأكبر قدراً أكبر من المعلومات نتخذ على ضوئها قرارنا ، وهذا يؤدي إلى تناقص كل من α و β .

وتختلف قيمة β وفقاً لما تكون عليه القيمة الصحيحة لوسيط المجتمع . فلنفرض أننا نرغب في اختبار أن القيمة الصحيحة لوسيط مجتمع ثنائي هي $p_0 = .4$ (واستخدمنا الرمز p_0 للدلالة على قيمة الوسيط p تحت الفرضية الابتدائية H_0) . ولنفرض أن الفرضية H_0 غير صحيحة وأن p تساوي حقاً لقيمة

بديله p_1 مثلاً . فما هي القيمة التي سيكون من الأسهل علينا تحرّيها : القيمة البديلة $p_1 = .4001$ أو القيمة البديلة $p_1 = 1.0$ ؟ ومن المؤكد أنه إذا كانت القيمة الصحيحة لِـ p مساوية للواحد فإن نتيجة كل تكرار ستكون نجاحاً وستقدم نتائج أية عينة نسحبها دلالة قوية تدعم رفضنا للفرضية $H_0: p_0 = .4$.
وعلى الوجه الآخر تقع $p_1 = .4001$ قريبة جداً من $p_0 = 4$ بحيث أنه سيكون من الصعب جداً تحرّي مثل هذا الفرق البسيط ما لم تكن العينة كبيرة جداً . وبعبارة أخرى فإن احتمال قبول H_0 وهو β سيعتمد على الفرق بين القيمة الصحيحة لِـ p والقيمة المفروضة p_0 . ويدعى المنحني البياني لِـ β كتابع في الكمية $(p - p_0)$ بالمنحني المميز العملياتي من أجل الاختبار الاحصائي . ويلاحظ القارىء أن المنحني المميز العملياتي الخاص بخطة الكشف على وسقة من البضاعة بطريقة العينة والذي رأيناه في الفصل الخامس هو في الحقيقة منحن يعبر عن العلاقة التابعية بين β و p .

وبما أن منطقة الرفض محددة وتبقى ثابتة من أجل اختبار معين ، فإن α ستبقى ثابتة . وعند زيادة حجم العينة n تتناقص β وذلك مهما كانت القيمة البديلة للوسيط موضع الاختبار . وهذا يعني أنه سيوافق كل حجم للعينة أي كل قيمة لِـ n منحنياً مميزاً عملياتياً . ويوضح عدد من تمارين الفصل الخامس مثل هذه الخاصة .

وبالرغم من أن α تحدد حجم منطقة الرفض إلا أن شكل هذه المنطقة بحيث نحصل على أفضل اختبار ممكن لا يتحدد بقيمة α ، ويقدم الاحصاء الرياضي الشكل الأفضل الذي يمكن أن تتخذه منطقة الرفض . وما سنذكره عبر هذا الكتاب التطبيقي من اختبارات هي الاختبارات التي تبرهن البحوث النظرية في الإحصاء أنها أفضل الاختبارات التي يمكن أن تتوفر لنا . ومع معرفة شكل منطقة الرفض يصبح السير النموذجي لعملية الاختبار هي أن يختار المجرب أولاً قيماً لِـ α و β تقيس مدى المخاطرة التي يمكنه التساهل بها وذلك بالنسبة

لكل من نوعي الخطأ . كما يحدد مدى إنحراف القيمة الحقيقية للوسيط عن القيمة المفترضة ، هذا الانحراف الذي يمكن اعتباره مهماً من وجهة النظر العملية ، وبالتالي يرغب المجرب في أن يكون الاختبار قادراً على كشفه . وبعدها يختار شكل منطقة الرفض وفقاً للقواعد التي يقدمها الاحصاء النظري والتي تتوقف بصورة خاصة على الفرضية البديلة H_1 . ثم يحدد حجم منطقة الرفض أو حدودها وفقاً للقيمة α التي وقع اختياره عليها من قبل. وأخيراً يختار حجم العينة n الضروري (في ضوء القيمة التي اختارها للانحراف بين القيمتين الحقيقية والمفترضة) للوصول إلى القيمة التي اختارها منذ البداية لـ β. ويمكن القيام بهذه الخطوة الأخيرة بالعودة إلى المنحنيات المميزة العملياتية الموافقة لاختباره والمرسومة من أجل قيم مختلفة لـِ n .

وبما أن القرار المتخذ هو إما برفض الفرضية أو قبولها فإن الكمية $1 - \beta$ تمثّل احتمال رفض الفرضية H_0 علماً أن الفرضية H_1 البديلة هي الفرضية الصحيحة . وتدعى الكمية $1 - \beta$ قوة الاختبار الاحصائي . وهو مقياس المفاضلة بين اختبارات من نفس مستوى الأهمية α . أي أن الاختبار الأفضل من بين اختبارات لها نفس مستوى الأهمية α هو الاختبار الأكثر قوة .

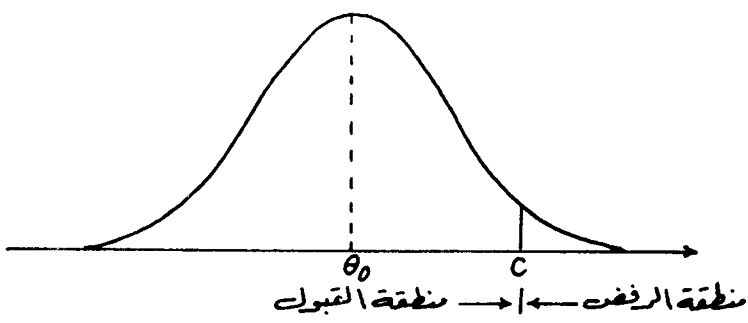
٧ـ١١ الاختبار الاحصائي من أجل عينات كبيرة الحجم : في حالة عينات كبيرة الحجم ، يقوم اختبار الفرضيات المتعلقة بوسطاء المجتمع التي نوقشت في الفقرات (٧ـ٣) إلى (٧ـ٩) ، على إحصاء اختبار يتبع التوزيع الطبيعي . ولهذا السبب يمكن أن نأتي عليها جميعاً من خلال مناقشة عامة واحدة . وسنرمز بصورة عامة للوسيط المدروس بِـ θ حيث يمثل θ : μ ، $(\mu_1 - \mu_2)$ ، p أو $(p_1 - p_2)$. ونوضح كلاً من هذه الاختبارات بمثال .

لنفرض الآن أننا نرغب في اختبار فرضية تتعلق بوسيط θ وأن التقدير النقطي $\hat{\theta}$ لهذا الوسيط يتبع التوزيع الطبيعي بمتوسط يساوي θ وإنحراف معياري يساوي $\sigma_{\hat{\theta}}$. وإذا كانت الفرضية الابتدائية $H_0 : \theta = \theta_0$

فسيتوزع التقدير $\hat{\theta}$ وفق التوزيع الطبيعي بمتوسط يساوي θ_0 وذلك كما يبين الشكل (٧ـ٧) . ولنفرض أننا نهتم من وجهة النظر العملية برفض H_0 عندما يكون θ أكبر من θ_0 ، فعندئذ تكون الفرضية البديلة $H_1 : \theta > \theta_0$ ويمكن البرهان على أن أفضل اختبار هو ذلك الذي يرفض H_0 عندما يكون $\hat{\theta}$ كبيراً ، ونقصد بكبر $\hat{\theta}$ هو أن تكون المسافة بينه وبين θ_0 مساوية لعدد من الانحرافات المعيارية $\sigma_{\hat{\theta}}$. ويبين الشكل (٧ـ٧) موقع منطقة الرفض. وتسمى القيمة c التي تفصل منطقة الرفض عن منطقة القبول بالقيمة الحرجة لإحصاء الاختبار .

ويحدد مستوى أهمية الاختبار α موقع النقطة الحرجة c لأن احتمال أن يكون $\hat{\theta} \geqslant c$ علماً أن H_0 صحيحة أي احتمال رفض الفرضية H_0 مع أنها صحيحة يجب أن يساوي α . أي أن المساحة تحت منحني الكثافة الطبيعي الذي متوسطه θ_0 وانحرافه المعياري $\sigma_{\hat{\theta}}$، (أي المنحني الطبيعي كما تحدده الفرضية H_0 التي نفترض أنها صحيحة) والواقعة إلى يمين النقطة c يجب أن يساوي α . ونعبر عن ذلك عادة بكتابة : $P(\hat{\theta} \geqslant c \mid H_0 \text{ صحيحة}) = \alpha$. وعلى سبيل المثال ، إذا كان $\alpha = .05$ فإن $c = \theta_0 + 1.645\,\sigma_{\hat{\theta}}$. وعندما تقع منطقة الرفض في أحد ذيلي المنحني دون الآخر فإننا نقول أن الاختبار وحيد الذيل .

وإذا كنا من وجهة النظر العملية ، أيضاً نرفض H_0 لمجرد أن لا يكون θ مساوياً لِ θ_0 أي سواء أكان θ أكبر أم أصغر من θ_0 فعندها نعبر عن



شكل ٧ـ٧ توزيع $\hat{\theta}$ عندما تكون H_0 صحيحة

الفرضيتين الابتدائية والبديلة على الشكل :

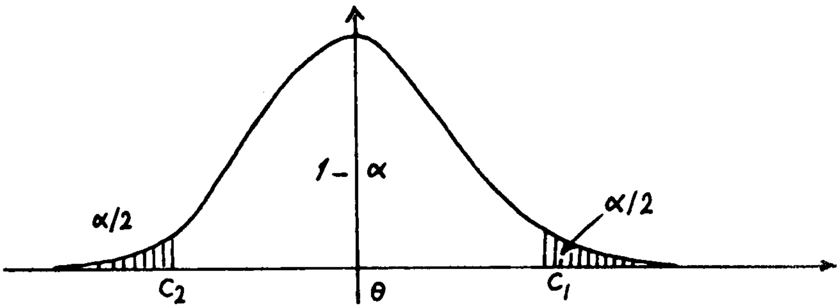
$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

ويمكن القول بأن أفضل اختبار يتوفر لنا عملياً هو الاختبار الذي تتوضع فيه منطقة الرفض ، على التساوي ، في ذيلي المنحني . ويسمى مثل هذا الاختبار عندئذ بالاختبار الثنائي الذيل . وعندها نقتطع مساحة تساوي $\alpha/2$ من كل ذيل أي نحدد نقطتين c_1 و c_2 بحيث يكون :

$P(\hat{\theta} \geqslant c_1 \mid H_0$ صحيحة$) = \alpha/2$ و $P(\hat{\theta} \leqslant c_2 \mid H_0$ صحيحة$) = \alpha/2$.

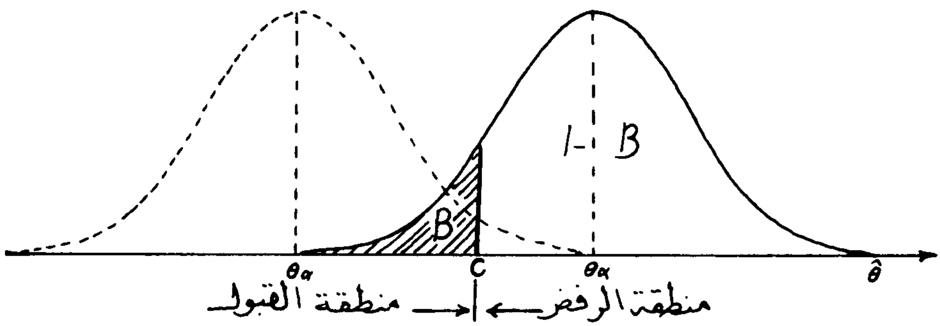
ونرفض الفرضية H_0 إذا كانت قيمة $\hat{\theta}$ إلى يمين c_1 أو إلى يسار c_2 ونقبلها فيما عدا ذلك . ويوضح الشكل (٧ــ٨) منطقة الرفض وتوزيع $\hat{\theta}$ عندما تكون H_0 صحيحة أي توزيع $\hat{\theta}$ تحت الفرضية H_0 .



شكل ٧ــ٨ توزيع $\hat{\theta}$ عندما تكون H_0 صحيحة في حالة اختبار ثنائي الذيل

ويوضح الشكل (٧ــ٩) كيفية حساب β أو القوة $(1 - \beta)$ للاختبار الوحيد الذيل المذكور أعلاه . فعندما تكون H_0 غير صحيحة و $\theta = \theta_1$ وليس θ_0 فإن توزيع اختبار الإحصاء $\hat{\theta}$ سيكون الآن التوزيع الطبيعي ولكن بمتوسط يساوي θ_1 وبنفس التشتت $\sigma^2_{\hat{\theta}}$. أي أن التوزيع الجديد تحت الفرضية H_1 سيكون نفس المنحني بعد سحبه إلى اليمين بحيث يتمركز حول النقطة θ_1 بدلاً من θ_0 .

ويبين الشكل (٧ــ٩) التوزيع تحت H_0 بخط متقطع والتوزيع تحت H_1 بخط متصل



شكل ٧ـ٩ توزيع $\hat{\theta}$ عندما تكون H_0 غير صحيحة و $\theta = \theta_1$.

ولحساب β وهو احتمال قبول H_0 علماً أن H_1 هي الصحيحة نحسب احتمال أن يكون $\hat{\theta} < c$ ولكن على أساس أن توزيع $\hat{\theta}$ هو التوزيع الطبيعي بمتوسط يساوي θ_1 وتشتت $\sigma_{\hat{\theta}}^2$. وهذا الاحتمال هو المساحة المخططة بخطوط مائلة تحت المنحني المتصل في الشكل (٧ـ٩) . وتكون المساحة فوق منطقة الرفض وتحت هذا المنحني أي المساحة الواقعة إلى يمين c هي $(1-\beta)$ أو قوة الاختبار ، وبعبارة أخرى تمثل هذه المساحة احتمال رفض H_0 عندما تكون غير صحيحة . ويتضح من هذا الشكل ما ذكرناه سابقاً من أن تصغير منطقة الرفض أي تناقص α يؤدي بنفس الوقت إلى زيادة β وتخفيض $(1-\beta)$. وأن β ، وبطبيعة الحال $(1-\beta)$ ، يعتمد على المسافة بين النقطتين θ_0 ، و θ_1 على محور الفواصل أي على الفرق بين القيمة المفترضة لِ θ والقيمة البديلة . وبصورة مشابهة نحسب β من أجل الاختبار الثنائي الذيل .

ويلاحظ القاريء أن جميع التقديرات النقطية التي نوقشت في الفقرات السابقة تحقق متطلبات الاختبارين اللذين وصفناهما أعلاه شريطة أن يكون حجم العينة n كبيراً بالقدر الكافي لتطبيق نظرية النهاية المركزية واعتبار التوزيع التقريبي للتقدير النقطي θ هو التوزيع الطبيعي ، والذي يمكننا بالإضافة إلى ذلك ، وعندما لا يكون التشتت σ^2 معروفاً ، من استخدام تشتت العينة s^2 بدلاً من σ^2 . وهكذا يمكننا اختبار فرضيات حول μ ، p ، $(\mu_1-\mu_2)$ ،

و $(p_1 - p_2)$.

ولتبسيط الاجراءات التطبيقية الخاصة بالاختبار نستخدم المتحول المعياري :

$$Z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}$$

كإحصاء الاختبار . ومن أجل الاختبار الثنائي الذيل ، مثلاً ، بمستوى من الأهمية $\alpha = .05$ نرفض الفرضية الابتدائية إذا كان $z < -1.96$ أو $z > 1.96$.

أما من أجل الاختبار الوحيد الذيل المذكور أعلاه فإننا نرفض H_0 إذا كان $Z > 1.645$.

مثال ٧ـ٩ بالإشارة إلى المثال (٧ـ١) في الفقرة (٧ـ٣) ، اختبر الفرضية بأن معدل الانتاج اليومي هو $\mu = 880$ طناً ضد الفرضية البديلة بأن μ هي أقل أو أكثر من 880 طناً في اليوم . أي :

$$H_0 : \mu = 880$$
$$H : \mu \neq 880$$

وتحوي العينة المستخدمة $n = 50$ قياساً وفيها $\bar{x} = 871$ و $s = 21$.

التقدير النقطي لِـ μ هو $\bar{x}$ .. أي أي إحصاء الاختبار هو :

$$Z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

وباستخدام s كتقريب لِـ σ نجد :

$$Z = \frac{871 - 880}{21/\sqrt{50}} = -3.03$$

ومن أجل $\alpha = .05$ تكون منطقة الرفض هي $Z > 1.96$ أو $Z < -1.96$:

وبما أن القيمة المحسوبة لِـ z تقع ضمن منطقة الوفض فإننا نرفض الفرضية بأن $\mu = 800$ ونستنتج أنه أقل من 880 طناً .

وتتضح العلاقة بين الاختبار الإحصائي المبني على إحصاء اختبار يتبع التوزيع الطبيعي بمستوي معين من الأهمية α وبين الـ $(1-\alpha)$ مجال ثقة الذي رأيناه في الفقرة (٧ـ٥) . فالمجال $\bar{x} \pm 1.96\sigma/\sqrt{n}$ أو تقريباً 871 ± 5.82 يعني أنه عند تكرار العينات فإن $(1-\alpha)$ من المجالات الناتجة (مجال واحد من كل

عينة) سوف يحوي القيمة الحقيقية μ . وبملاحظة أن $\mu = 880$ لا تقع ضمن المجال سنميل إلى رفض القيمة $\mu = 880$ كقيمة ممكنة ونستنتج أن متوسط الانتاج اليومي هو ، في الحقيقة ، أقل من 880 طناً .

ويوضح المثال التالي حساب β وبالتالي قوة الاختبار $(1-\beta)$ من أجل الاختبار الاحصائي الوارد في المثال السابق (٧ـ٩) .

مثال ٧ـ١٠ بالإشارة إلى المثال السابق أحسب الاحتمال β وقوة الاختبار $1-\beta$ علماً أن القيمة الحقيقية لِ μ هي 870 طناً في اليوم .

تقع منطقة القبول في الاختبار الذي وصفناه في المثال السابق ضمن المجال $\mu_0 \pm 1.96\sigma_{\bar{x}}$ أي المجال :

$$880 \pm 1.96\ (21/\sqrt{50})$$

أو :

874.18 إلى 885.82

واحتمال قبول H_0 علماً أن $\mu = 870$ يساوي إلى المساحة تحت منحني الكثافة الطبيعي الخاص بإحصاء الاختبار $\bar{x}$ فوق المجال (874.18, 885.82) ، ولكن كما تحدده الفرضية البديلة H_1 أي بمتوسط يساوي 870 . وهذا يعني أن :

$$\beta = P(874.18 < \bar{x} < 885.82)$$

وإذا اعتبرنا $\sigma_{\bar{x}} = 21/\sqrt{n} = 2.97$ نكتب :

$$\beta = P\left(\frac{874.18-870}{2.97} < \frac{\bar{x}-870}{2.97} < \frac{885.82-870}{2.97}\right)$$

$$= P\left(\frac{4.18}{2.97} < Z < \frac{15.82}{2.97}\right) = P(1.41 < Z < 5.32)$$

$$= .5 - .4207 = .0793 = .08$$

أي أن احتمال قبول H_0 علماً أن القيمة الحقيقية لِ μ تساوي 870 هو تقريباً .08 وقوة الاختبار أي احتمال رفض الفرضية عندما تكون غير صحيحة هو

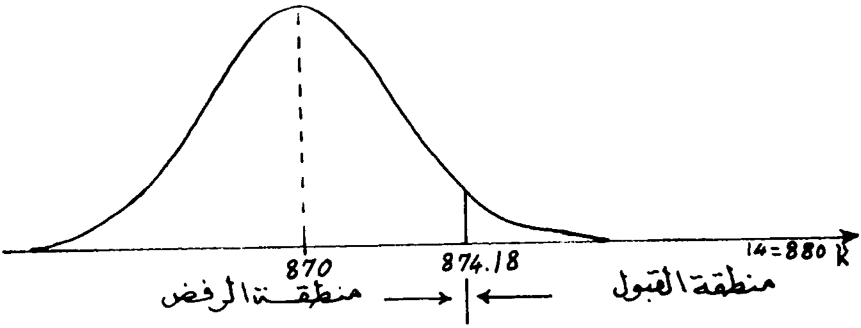
$$1 - \beta = 1 - .08 = .92$$

وللسهولة يمكن حساب β كما يلي : نعلم أن إحصاء الاختبار هو $\frac{\bar{x}-880}{2.97}$. وأن منطقة القبول هي بين 1.96ـ إلى 1.96 ومنه :

$$\beta = P\left(-1.96 < \frac{\bar{x}-880}{2.97} < 1.96\right) = P\left(-1.96 < \frac{\bar{x}-870-10}{2.97} < 1.96\right)$$

$$= P\left(-1.96 + \frac{10}{2.97} < \frac{\bar{x}-870}{2.97} < 1.96 + \frac{10}{2.97}\right) = P(1.41 < Z < 5.32) = .0793$$

ذلك لأنه وفقاً للفرضية H_1 يكون $\frac{\bar{x}-870}{2.97}$ متحولاً طبيعياً معيارياً (انظر الشكل ٧ـ١٠) .



الشكل ٧ـ١٠ حساب β في المثال (٧ـ١٠) .

مثال ٧ـ١١ من المعروف أن واحداً من عشرة على وجه التقريب من المدخنين يفضلون النوع A من السجاير . وبعد حملة دعائية لهذا النوع في منطقة معينة أخذنا عينة من 200 مدخن بقصد اختبار فعالية الحملة الدعائية . وتبين لنا أن 26 مدخناً من بينهم يفضلون A فهل تقدم هذه النتيجة دلالة كافية على أن الحملة الدعائية كانت ذات فائدة ؟

نفرض أن التجربة تحقق شروط التجربة الثنائية . ويمكن وضع الفرضية على الشكل التالي حيث p هو احتمال تفضيل مدخن من المنطقة المدروسة للسيجارة A :

$$H_0 : p = .10$$

وذلك ضد الفرضية البديلة :

$$H_1 : P > .10$$

ومن الطبيعي أن تكون الفرضية البديلة كذلك لأننا نهتم فقط بالكشف عن أية زيادة في p منطلقين في ذلك من الاعتقاد بأن حملة الدعاية إذا لم تكن مفيدة فهي على أي حال لا يمكن أن تكون ضارة فتخفض من قيمة p . وأفضل اختبار لمثل هذه الفرضية هو الاختبار الناتج عن وضع كامل منطقة الرفض في الذيل الأيمن لتوزيع إحصاء الاختبار .

وكما نعلم فإن التقدير النقطي لِ p هو $\hat{P} = \frac{x}{n}$ وهكذا يكون إحصاء الاختبار هو :

$$Z = \frac{\hat{P} - P_0}{\sigma_{\hat{P}}} = \frac{\hat{P} . P_0}{\sqrt{P_0 q_0 / n}}$$

وبضرب الصورة والمخرج بِـ n نجد :

$$Z = \frac{x - n P_0}{\sqrt{n P_0 q_0}}$$

وهو نفس إحصاء الاختبار المستخدم في المثال (٦ـ٨) . ومن الطبيعي أن نستخدم القيمة p_0 من أجل p في عبارة الانحراف المعياري لـ $\hat{p}$ وهي $\sqrt{npq}$ طالما أننا فرضنا أن $p = p_0$. والتقريب المرتكب في مثل هذا التبديل سيكون طفيفاً بحيث لا يؤثر بنتيجة الاستقراء من أجل n كبيرة .

وإذا اخترنا مستوى أهمية الاختبار مساوياً لِ $\alpha = .05$ فسنرفض H_0 عندما يكون $z > 1.645$ وبحساب z نجد :

$$Z = \frac{26 - 20}{\sqrt{200(.10)(.90)}} = 1.41$$

وهذا القيمة لا تقع ضمن منطقة الرفض وبالتالي فإننا لا نرفض الفرضية أي نقبلها . ونلاحظ أن قبول الفرضية يعني فقط أننا لم نجد الدلالة الكافية لرفضها .

مثال ٧ـ١٢ يريد مكتب الدراسات في جامعة دراسة ما إذا كانت الحالة الاقتصادية للطالب ذات أثر في قدراته الدراسية . ولهذه الغاية تمَّ تقسيم الطلاب إلى فئتين وأخذنا عينة عشوائية من مائة طالب من كل فئة . وكان متوسط أحد

العينتين مساوياً $\bar{x}_1 = 67.5$ وانحرافها المعياري $S_1^2 = 225$ ومتوسط العينة الأخرى وانحرافها المعياريهما $\bar{x}_2 = 63.5$ و $S_2^2 = 250$ على الترتيب . فهل تقدم هذه المعلومات الإحصائية دلالة كافية على وجود فرق بين الفئتين فيما يتعلق بالقدرات الدراسية ؟

إذا رمزنا للمعدل العام ضمن مجتمع الفئة الأولى بِـ μ_1 وللمعدل العام ضمن مجتمع الفئة الثانية بِـ μ_2 فيمكن وضع الفرضية على الشكل :

$$H_0 : \mu_1 - \mu_0 = D_0$$

وذلك ضد الفرضية البديلة :

$$H_1 : \mu_1 - \mu_0 \neq D_0$$

حيث D_0 في مثالنا هنا هو الصفر . ونعلم أن $\bar{x}_1 - \bar{x}_2$ هو تقدير نقطي لِـ $(\mu_2 - \mu_1)$. وتوزيعه التقريبي ، من أجل n كبير ، هو التوزيع الطبيعي بمتوسط يساوي $\mu_2 - \mu_1$ وتشتت يساوي $\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}$ وهكذا نأخذ إحصاء الاختبار :

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

حيث σ_1^2، σ_2^2 معروفان أو أن حجم العينة كبير بحيث يكون S_1^2 و S_2^2 تقريبين جيدين لِـ σ_1^2 و σ_2^2 على الترتيب .

وفي مثالنا هنا نجد :

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}} = \frac{67.5 - 63.5}{\sqrt{\frac{225}{100} + \frac{250}{100}}} = 1.83$$

ومن أجل مستوى من الأهمية $\alpha = .05$ تكون منطقة الرفض هي $Z > 1.96$ أو $Z < -1.96$. وبما أن قيمة z لا تقع ضمن منطقة الرفض فإننا لا نرفض الفرضية الابتدائية أي نقبلها .

مثال ٧ــ١٣ تبين سجلات مستشفى أن 52 رجلاً من عينة من 1000 رجل يقابلها 23 امرأة من أصل 1000 إمرأة ممن كانوا يعانون من مرض في القلب .

هل تقدم هذه المعلومات الاحصائية دلالة كافية على أن نسبة الإصابة بمرض القلب أكبر بين الرجال ؟

سنفترض أن عدد الداخلين إلى المستشفى من مرضى القلب يتبع التوزيع الثنائي سواء من الرجال أو من النساء بوسيطين هما على الترتيب p_1 و p_2 .

ونريد اختبار الفرضية :

$$H_0: (p_1 - p_2) = D_0$$

(وفي مثالنا هنا $D_0 = 0$) . وكما نعلم فإن التوزيع التقريبي للتقدير النقطي $\hat{p}_1 - \hat{p}_2$ هو التوزيع الطبيعي بمتوسط يساوي $(p_1 - p_2)$ وانحراف معياري .

$$\sigma_{(\hat{p}_1-\hat{p}_2)} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

ويمكن عندئذ إستخدام :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{(\hat{p}_1-\hat{p}_2)}}$$

كإحصاء اختبار من أجل الفرضية H_0 . ومع استخدام قيم تقريبية مناسبة لكل من p_1 و p_2 اللذين يظهران في عبارة $\sigma_{(\hat{p}_1-\hat{p}_2)}$ نجد الحالتين التاليتين :

الحالة I : إذا فرضنا أن p_1 يساوي p_2 أي أن :

$$H_0 : p_1 = p_2 = p$$

ويكون أفضل تقدير للقيمة المشتركة p مستخدمين قياسات العينتين هو :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

وإحصاء الاختبار هو :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

أو :

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

الحالة II : وعلى الوجه الآخر إذا فرضنا أن D_0 لا يساوي الصفر ، أي أن :

$$H_0 : (p_1 - p_2) = D_0$$

حيث $D_0 \neq 0$ ، فعندئذ يكون أفضل تقديريــن لِـ p_1 و p_2 هما $\hat{p}_1$ و $\hat{p}_2$ على الترتيب . واختبار الإحصاء هو :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$$

ومعظم المسائل التي نواجهها في التطبيق تكون من الحالة I . وفي مثالنا هنا نختبر الفرضية :

$$H_0 : p_1 - p_2 = 0$$

ضد البديل :

$$H_1 : p_1 - p_2 \neq 0$$

وباستخدام الاختبار الثنائي الذيل بمستوى من الأهمية $\alpha = .05$ تكون منطقة الرفض هي $Z > 1.96$ أو $Z < -1.96$.

والتقدير المشترك من العينتين لِـ p والذي نحتاجه للتبديل في عبارة $\sigma_{(\hat{p}_1 - \hat{p}_2)}$ هو :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{52 + 23}{1000 + 1000} = .0375$$

وإحصاء الاختبار هو :

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.52 - .023}{\sqrt{(.0375)(.9625)\left(\frac{1}{1000} + \frac{1}{1000}\right)}}$$

وبما أن هذه القيمة تقع ضمن منطقة الرفض فإننا نرفض الفرضية $p_1 = p_2$ ونستنتج أن البيان الإحصائي يقدم دلالة كافية للقول بأن النسبة المئوية لمرضى القلب ممن يدخلون المستشفى من الرجال أكبر من النسبة المئوية لمرضى القلب ممن يدخلون المستشفى من النساء .

تمارين

١ ـ أعرض نظرية النهاية المركزية . ما هو دور هذه النظرية في الاستقراء الإحصائي ؟

٢ ـ متوسط العمر في عينة عشوائية من مائة مصباح كهربائي هو 1280 ساعة والانحراف المعياري 142 ساعة . والمطلوب تقدير متوسط عمر مجتمع المصابيح الكهربائية الذي سحبنا منه العينة ووضع حدود لخطأ التقدير .

٣ ـ لنفرض أن متوسط المجتمع في التمرين السابق هو في الحقيقة 1285 ساعة بانحراف معياري $\sigma = 150$ ساعة . فما هو احتمال أن يتجاوز متوسط عينة عشوائية حجمها $n = 100$ ال 1300 ساعة ؟

٤ ـ اختبرنا نوعاً جديداً من المصابيح الكهربائية الخاصة بآلات التصوير لتقدير p احتمال أن ينتج مصباحاً جديداً الإضاءة اللازمة وفي الوقت المناسب . وقد وجدنا من بين 1000 من هذه المصابيح أن 920 قد عملت وفقاً للمواصفات المطلوبة . والمطلوب تقدير p ووضع حدود لخطأ التقدير .

٥ ـ باستخدام أمثال ثقة 90. ضع مجال ثقة حول متوسط عمر المصابيح الكهربائية في التمرين الثاني .

٦ ـ ضع 99. مجال ثقة حول الاحتمال p في التمرين الرابع .

٧ ـ اختبرنا عينة عشوائية من 400 أنبوب الكتروني خاص بأجهزة الراديو فوجدنا من بينها 40 أنبوباً عاطلاً عن العمل . ضع 90. مجال ثقة حول النسبة الحقيقية للأنابيب العاطلة؟

٨ ـ يرغب مستشفى بتقدير عدد الأيام التي يحتاجها وعلاج مرضى سنهم بين 25 و 34 . ووجدت إدارة المستشفى أن متوسط عدد الأيام في عينة عشوائية من 500 مريض من هذه الفترة من العمر هو 5.4 يوماً . وبانحراف معياري 1.3 يوماً . والمطلوب تقدير متوسط الإقامة في المستشفى لمجتمع المرضى الذي جاءت منه العينة ، ووضع حدود خطأ لهذا التقدير .

٩ ـ قمنا بتجربة لمقارنة عمق اختراق حفارتي مناجم هيدروليكيتين . وكانت بنية الصخور ومدة الحفر واحدة للحفارتين . وقد وجدنا أن متوسط اختراق الحفارة A هو 10.8 بوصة بانحراف معياري يساوي 1.2 بوصة وذلك في عينة من 50 حفرة . وكان المتوسط والانحراف المعياري الخاص باختراق الحفارة B هما على الترتيب 9.1 و 1.6 بوصة ، وذلك في عينة من 80 حفرة . والمطلوب تقدير الفرق في معدل الاختراق الوسطي ووضع حدود لخطأ التقدير .

١٠ ـ ضع 90 مجال ثقة من أجل الفرق المذكور في التمرين السابق .

١١ ـ متوسط عينة من 43 قياساً هو 26.3 وانحرافها المعياري 1.9 . ضع 98% حدود ثقة من أجل متوسط المجتمع .

١٢ ـ حضر كيميائي دواء يهدف ، عند تطبيقه على نوع معين من الحشرات ، إلى قتل 60% منها . فكم يجب أن يكون حجم العينة المستخدمة إذا رغب في أن يكون على 95% من الثقة بأنه في حدود 02. من النسبة الحقيقية للحشرات الهالكة .

١٣ ـ تبين الخبرة السابقة أن الانحراف المعياري للدخل السنوي لعمال النسيج هو 400 ل.س. فكم يجب أن يكون حجم العينة التي نأخذها من عمال النسيج إذا رغبنا بتقدير متوسط المجتمع في حدود 50 ل.س ، وباحتمال 95% في أن يكون تقديرنا هذا صحيحاً ؟ وإذا علمنا أن متوسط العينة في هذه المسألة هو 4800 ل.س فالمطلوب وضع 95% حدود ثقة من أجل متوسط المجتمع .

١٤ ـ كم ناخباً يجب أن تضم عينة جمعناها لتقدير نسبة الناخبين الذين يفضلون مرشحاً معيناً للرئاسة إذا رغبنا في أن يكون التقدير صحيحاً في حدود .005 ؟ ولنفرض أن النسبة الحقيقية يجب أن تقع في جوار 0.5 .

١٥ ـ في اقتراع أجريناه بين طلبة الجامعات وافق 300 من بين 500 طالب منظم في صفوف الاتحاد الوطني للطلبة على المسألة المطروحة بينما وافق عليها 64 من بين 100 من الطلاب غير الاتحاديين. قـدّر الفرق في نسبة قبول المسألة بين الطلبة الاتحاديين وغير الاتحاديين وضع حدوداً على خطأ التقدير .

١٦ ـ بالإشارة إلى التمرين السابق كم يجب أن يكون عدد الطلاب الذين نقابلهم من كل من الفئتين إذا رغبنا أن يكون تقدير الفرق صحيحاً في حدود .05 ؟ ونفرض أن للعينتين نفس الحجم وأنه يمكن اعتبار 0.6 كتقريب لكلي النسبتين .

١٧ ـ من كل من مجتمعين طبيعيين بنفس المتوسط وبانحراف معياري 6.4 و 7.20 على الترتيب ، سحبنا عينتين مستقلتين حجم كل منهما $n = 64$. ما هو احتمال أن يتجاوز الفرق بين متوسطي العينتين بالقيمة المطلقة 0.60 .

١٨ ـ عرف α و β من أجل اختبار إحصائي لفرضية ..

١٩ ـ تتوزع الأجور اليومية في صناعة طبيعية توزيعاً طبيعياً بمتوسط يساوي 13.2 ل.س وانحراف معياري 2.50 ل.س. إذا كانت إحدى شركات هذه الصناعة تستخدم 40 عاملاً ومتوسط الأجور اليومية التي تدفعها لهم هو 12.20 ل.س. فهل يمكن وبمستوى من الأهمية $\alpha = .01$ اتهام الشركة بأنها تدفع أجوراً متدنية ؟

٢٠ ـ طبقنا طريقتين في تعليم القراءة على مجموعتين من طلبة المدارس الابتدائية تحوي كل منهما 50 طالباً . وكانت نتيجة اختبار القراءة في نهاية الفترة الدرسية في المجموعتين على الترتيب : $\bar{x}_1 = 74$ ، $\bar{x}_2 = 71$ ، $S_1 = 9$ ، $S_2 = 10$. فهل تقدم هذه النتائج دلالة على وجود فرق يُذكر بين الطريقتين ؟

(استخدم $\alpha = .10$).

٢١ ـ يقدم مصنع للغسالات الكهربائية نموذجاً معيناً من الغسالات في ثلاثة ألوان A ، B ، أو C . ولاحظت إدارة المصنع أن 400 من بين الألف غسالة التي بيعت أولاً من اللون A . فهل تستنتج أن الزبائن يفضلون اللون A ؟ علل الإجابة .

٢٢ ـ بالإشارة إلى التمرين الثاني اختبر الفرضية بأن متوسط عمر المصابيح هو 1300 ساعة ضد الفرضية البديلة بأنه أقل من ذلك . استخدم $\alpha = .05$.

٢٣ ـ بالإشارة إلى التمرين التاسع . هل تقدم المعلومات الإحصائية دلالة كافية على وجود فرق في متوسط عمق الاختراق الذي تنتجه كل من الحفارتين ؟

٢٤ ـ بالإشارة إلى التمرين (٢٢) ما هو احتمال عدم رفض الفرضية الابتدائية $\mu_0 = 1300$ ساعة إذا كان متوسط العمر الحقيقي مساوياً 1290 ساعة ؟

٢٥ ـ بالعودة إلى التمرين السابع . لنفرض أننا اختبرنا عشر عينات في كل منها 400 أنبوب وحسبنا من كل عينة مجال ثقة من أجل p . فما هو احتمال ألا يحوي القيمة الحقيقية لِـ p مجال واحد من هذه المجالات ؟ ما هو احتمال أن يكون واحد على الأقل من هذه المجالات غير متضمن القيمة الحقيقية لِـ p ؟

٢٦ ـ بالعودة إلى المثال (٧ـ١٠) أحسب β من أجل قيم بديلة متعددة لِـ μ. (مثلاً $\mu = 873$ ، $\mu = 875$ ، $\mu = 877$).واستخدم القيم الثلاث الناتجة لِـ β مع تلك المحسوبة في المثال (٧ـ١٠) لرسم المنحني المميز العمليـاتي للاختبار الإحصائي .

٢٧ ـ أعد الخطوات المطلوبة في التمرين السابق من أجل عينة حجمها $n = 20$ وقارن المنحنيين الناتجين .