الفصل الثاني: العيّنة وتقديراتها الإحصائية

تمهيد:

التقدير هو تقدير معالم المجتمع الإحصائي (أو التوزيع الاحتمالي) والتي غالباً ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها -عادة- من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة، أو تقدير متوسط عمر الناخب... وغيرها.

وهناك نوعان (أو أسلوبان) للتقدير يسمى الأول تقدير النقطة (أو القيمة الواحدة)، ويسمى الثاني تقدير الفترة (أو فترة التقدير أو الثقة).

ففي حالة تقدير النقطة نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. وكمثال آخر لو أخذنا نسبة الناخبين في العينة الذين يؤيدون مرشحا معيناً كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين.

أما في تقدير الفترة أو فترة التقدير فنحصل على مدى أو فترة تتحدد بحدين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات. فمثلاً : إذا قدرنا أن الوسط الحسابي لأعمار الناخبين يتراوح بين: 40 + 6,40 - 6 سنة أي يتراوح بين 34 سنة كحد أدنى و 46 سنة كحد أعلى نكون قد حصلنا على تقدير فترة للوسط الحسابي لأعمار الناخبين في المجتمع - ونلاحظ أن هذه الفترة (34, 46) تحتوي على عدد لا نهائي من الأعمار، بمعنى أن العدد لا يقتصر فقط على الأعداد الصحيحة والتي تشمل السنوات، ولكنها

تشمل أيضا كسور السنوات، والأيام والشهور، والساعات... الخ وسوف نرى كيف نحدد فترة التقدير هذه في بعض الحالات.

وتتميز تقديرات الفترة بالإضافة إلى أنها تحتوي على عدد كبير جداً من القيم، بأنه يمكن حساب احتمال أن يكون التقدير صحيحاً، وبالتالي فإنه يمكن معرفة مدى دقة التقديرات. لذا فإن فترات التقدير تسمى أيضاً " فترات الثقة " لأن هذه الفترات تعتمد في تكوينها الإحصائي على درجات أو مستويات ثقة معينة مثل 95% أو 99 % وغيرها، بمعنى أن احتمال أن تكون فترة التقدير صحيحة هو 0.95 أو 0.99 وهكذا...

فإذا كان متوسط أعمار الناخبين يتراوح ما بين 46 و 34 سنة، ودرجة الثقة هي 95 % فإن هذا معناه أنه لو تكررت التجربة مائة مرة، فإن التقدير سيكون محصوراً بين هذين الرقمين في95 من الحالات (أي احتمال أن يكون صحيحاً هو 95%).

1-أنواع التقدير:

يوجد لدينا نوعان من التقدير الإحصائي:-التقدير بنقطة و- التقدير بفترة

-التقدير بنقطة: المقصود بهذا النوع هو تقدير معلمة المجتمع المجهولة بإحصائية نحسب قيمتها من بيانات عينة عشوائية مسحوبة من هذا المجتمع.

نقوم بتقدير المعلمة بقيمة واحدة فقط مثلا عند سحب عينة من مجتمع ما و قمنا بحساب:

الوسط الحسابي للعينة فإنّه يعتبر تقدير لنقطة معلمة المجتمع μ. وتباين العينة S^2 فإنه يعتبر تقدير لنقطة لمعلمة المجتمع σ^2. نسبة العينة فإنها تعتبر تقديرا لنقطة لمعلمة المجتمع P.

التقدير لنقطة يعطي قيمة واحدة لمعلمة المجتمع الذي نريد تقديرها، لا يمكننا الحكم على مدى دقة هذه القيمة في تمثيل المعلمة الحقيقية في المجتمع. يفضل الاعتماد على النوع الثاني من التقدير الإحصائي ألا و هو التقدير لفترة.

-التقدير لفترة :عبارة عن تقدير معلمة المجتمع بمدى (فترة) معين من القيم باحتمال معين و هذا الاحتمال يسمى مستوى الثقة و يرمز له بالرمز $(1-\alpha)\%$. لذلك سميت الفترة بفترة الثقة. حيث مستوى الثقة يعطينا مدى ثقتنا أن هذه المعلمة ستقع بين الحد الأدنى و الحد الأعلى للفترة.

2-فترة تقدير الوسط الحسابي للمجتمع في حالة العينات الكبيرة (فترة الثقة للوسط):

إذا كان المجتمع المسحوب منه العينة ذا توزيعاً طبيعياً وتباينه معروفاً أو كانت العينة كبيرة (أي حجمها ثلاثون مفردة أو أكثر) فإن توزيع المعاينة للوسط الحسابي يكون توزيعاً طبيعياً، وطالما نحن نتكلم عن تقدير متوسط المجتمع فإن أول ما نفكر فيه هو الوسط الحسابي للعينة. وفترة التقدير (أو الثقة) للوسط الحسابي للمجتمع تأخذ الشكل التالي :

تقدير متوسط المجتمع = الوسط الحسابي للعينة الخطأ المعياري للوسط

حيث: تقدير الوسط الحسابي للمجتمع، هو الوسط الحسابي للعينة، هو الخطأ المعياري للوسط، + تشير للجمع فنحصل على الحد الأعلى لفترة التقدير، - تشير للطرح فنحصل على الحد الأدنى، ولكن احتمال أن يكون هذا الكلام صحيحا هو 68.26 % فقط، أي أن درجة الثقة هنا لا تتعدى 68.26 % فإذا أضفنا وطرحنا ضعف الخطأ المعياري يرتفع الاحتمال إلى 95.44% أي ترتفع درجة الثقة إلى 95.44 % وفي هذه الحالة تأخذ فترة الثقة الشكل التالي : وإذا أضفنا وطرحنا ثلاثة أمثال الخطأ المعياري يصبح الاحتمال 99.72 % أي ترتفع درجة الثقة إلى 99.72% وتأخذ فترة الثقة الشكل التالي :

أي أنه بزيادة درجة الثقة يزيد طول الفترة. ومما سبق نلاحظ ما يلي:

1- أن هناك علاقة وثيقة بين درجة الثقة والرقم أو " المعامل المضروب في الخطأ المعياري فهو إما 1 أو 2 أو 3 على حسب درجة الثقة 68.26% أو 95.44% أو 99.72 % ولذلك فإن هذا المعامل هو الذي يسمى " معامل الثقة ". فبناء على درجة الثقة المطلوبة يتحدد معامل الثقة.

2- أن درجات ومعاملات الثقة التي ذكرناها تخص التوزيع الطبيعي، وأن المعاملات 1 أو 2 أو 3 ما هي إلا الدرجة المعيارية (Z) والتي نحصل عليها من جدول التوزيع الطبيعي المعياري وذلك بقسمة درجة الثقة (أو الاحتمال) على 2 (حيث أن المساحة موزعة بالتساوي على يمين ويسار الوسط) ثم بالكشف في المساحات تحت المنحنى الطبيعي المعياري عن خارج القسمة (أو أقرب رقم له) فنحصل على Z المقابلة وهذا يرجع إلى أن توزيع المعاينة للوسط هو التوزيع الطبيعي.

3- يمكن الحصول على فترات تقدير بأي درجة ثقة أخرى (غير الثلاث التي ذكرناها) وذلك بقسمة درجة الثقة المطلوبة - كما ذكرنا - على 2 ثم الكشف في المساحات حتى نحصل على Z المناسبة.

ويمكن تلخيص خطوات تقدير الوسط الحسابي للمجتمع فيما يلي :

أ – احسب الوسط الحسابي للعينة .

ب- احسب الخطأ المعياري للوسط

جـ - أضرب الخطأ المعياري للوسط في معامل الثقة المناسب (أو الدرجة المعيارية) حسب درجة الثقة المطلوبة

د- اطرح حاصل الضرب السابق من الوسط الحسابي للعينة فنحصل على الحد الأدنى لفترة التقدير، واجمع حاصل الضرب مرة أخرى على الوسط الحسابي للعينة فتحصل على الحد الأعلى لفترة التقدير.

4- يمكن كتابة أشهر وأهم درجات ومعاملات الثقة (للتوزيع الطبيعي) في الجدول التالي (مع ملاحظة أن 95%، 99% هي أشهرها على الإطلاق)

مثال تطبيقي:

لوأردنا معرفة متوسط الدخل اليومي لمجموعة من الناخبين في دولة ما، فإن ذلك يبدو أمرا صعباً من الناحية العملية نظراً لكبر حجم مجتمع الناخبين، إضافة إلى طول الوقت والتكاليف. لذا فإن الأسلوب العلمي المتبع في حالة كهذه هو اختيار عينة عشوائية نستطيع من خلال معرفة نتائجها تقدير متوسط دخول الناخبين في هذه الدولة.

فلو سحبت عينة عشوائية من مجموع مجتمع الناخبين في دولة ما حجمها 100 ناخب فإذا كان الوسط الحسابي والانحراف المعياري للدخل اليومي للناخبين بالعينة هما على الترتيب 90 دولار و 25 دولار، أوجد فترة تقدير للوسط الحسابي للدخل اليومي لمجموع الناخبين في هذه الدولة بدرجة ثقة 95% ؟

الحل:

بما أن فترة تقدير الوسط الحسابي للمجتمع هي:

والمعلومات المعطاة هي:

حجم العينة n = 100

الوسط الحسابي للعينة

والانحراف المعياري للعينة S = 25

وحيث أن درجة الثقة هي 95 % فإن : Z = 1.96 حسب ما هو موضح في الجدول السابق. وبالتالي فإن فترة تقدير الوسط الحسابي للدخل اليومي لمجتمع الناخبين بدرجة ثقة 95 % هي: أي أن الوسط الحسابي للدخل اليومي لمجتمع الناخبين يتراوح بين 85.1 دولاراً كحد أدنى، 94.9 كحد أعلى، وذلك بدرجة ثقة 95 %.

2- فترة تقدير النسبة للمجتمع (أو فترة الثقة للنسبة):

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر السياسية، وبالذات الوصفية منها كقياس اتجاهات الرأي العام، وقياس نسبة قتلى الحروب، ونسبة الدول التي أوفت بالتزاماتها في المنظمات الدولية أو الإقليمية... وغيرها ونظراً لأنه من الصعوبة بمكان في كثير من الأحيان حساب هذه النسبة مباشرة من المجتمع، فإننا غالبا ما نلجأ لتقدير هذه النسبة من عينة عشوائية مسحوبة من هذا المجتمع.

فلو افترضنا أن نسبة المؤيدين للسياسة الاقتصادية التي تنتهجها دولة ما هي P وأن العينة العشوائية كبيرة بدرجة كافية وأن نسبة مؤيدي هذه السياسة في العينة هي فإن خطوات تقدير النسبة في المجتمع تكون كما يلي :

أ- احسب النسبة في العينة .

ب- احسب الخطأ المعياري للنسبة

جـ- اضرب الخطأ المعياري للنسبة في معامل الثقة المناسب Z (حسب درجة الثقة المطلوبة) والتي نحصل عليها من جدول التوزيع الطبيعي المعياري (أو من الجدول الذي يحوي أهم درجات ومعاملات الثقة والذي ذكرناه آنفاً). أي نحسب:

د – للحصول على الحد الأدنى لتقدير النسبة نطرح حاصل الضرب (السابق) من نسبة العينة و للحصول على الحد الأعلى نجمع حاصل الضرب مع النسبة في العينة فنحصل على فترة تقدير النسبة. وبالتالي فإن فترة تقدير النسبة تكون في شكلها النهائي.

ولتوضيح هذه الخطوات بشيء من التفصيل نورد المثال التالي :

مثال تطبيقي: عينة عشوائية حجمها 144 ناخباً سحبت من إحدى المدن فوجد أن عدد المؤيدين في العينة لمرشح معين هو 60 ناخباً، أنشئ فترة تقدير لنسبة المؤيدين لهذا المرشح في المدينة كلها بدرجة ثقة 95 %.

الحل:

1- نحسب أولاً نسبة المؤيدين للمرشح في العينة التي نحصل عليها بقسمة عدد المؤيدين له على العدد الكلي للعينة (حجم العينة)

وحيث أن درجة الثقة المطلوبة هي 95 % فإن معامل الثقة المناسب هو : Z = 196 وفترة تقدير نسبة المؤيدين لهذا المرشح في المدينة تأخذ الشكل التالي :

وبالتعويض عن حجم العينة n = 144

والنسبة في العينة

ومعامل الثقة $Z = 1.96$

أي أن نسبة المؤيدين للمرشح في المدينة تتراوح بين 0.50 , 0.34 وذلك بدرجة ثقة 95 % بمعنى آخر أن نسبة مؤيدي هذا المرشح في هذه المدينة لا تتجاوز 50 % كحد أعلى، وبالتالي ففرصته في الفوز كمرشح قد لا تكون كبيرة وذلك بدرجة ثقة 95 % بمعنى أن هذا الحكم لا تتجاوز نسبة الخطأ فيه 5 %.

- تحديد حجم العينة لتقدير الوسط الحسابي للمجتمع :

يعتبر تحديد حجم العينة المناسب من المشاكل المهمة والشائعة التي تواجه الباحثين في مختلف المجالات، وبالذات عند دراسة الظواهر السياسية، ويختلف تحديد حجم العينة باختلاف الهدف من التقدير.

فإذا كان المطلوب هو تقدير الوسط الحسابي للمجتمع، فإن فترة تقدير الوسط هي كما سبق وأن أوضحنا:

ومنها نجد أن حجم العينة يأخذ الشكل التالي :

حيث:

Z هو معامل الثقة (أو الدرجة المعيارية) المقابل لدرجة الثقة المطلوبة، ونحصل عليا من جدول التوزيع الطبيعي المعياري.

هو تباين المجتمع (أو هو مربع الانحراف المعياري).

e هو أقصى خطأ مسموح به في تقدير الوسط، وهو عادة ما يحدده الباحث، وتتوقف على أهمية الموضوع أو الظاهرة السياسية المراد دراستها، ومدى الدقة المطلوبة في التقدير، ويسمى اختصاراً "الخطأ في تقدير الوسط".

ولتوضيح كيفية تحديد حجم العينة المناسب عند تقدير الوسط الحسابي للمجتمع، نأخذ المثال التالي :

مثال تطبيقي:

إذا كانت دخول الأفراد اليومية في إحدى دول العالم النامية تتبع التوزيع الطبيعي بانحراف معياري دولاراً، فما هو حجم العينة المناسب لتقدير متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقدير متوسط الدخل اليومي 5 دولارات، وذلك بدرجة ثقة 99 % ؟

الحل:

درجة الثقة 99 % أي أن : Z = 2.58

أقصى خطأ مسموح به هو 5 دولارات، أي أن : e = 5

والانحراف المعياري للمجتمع :

فإن حجم العينة مقرباً لأقرب عدد صحيح هو : 60 فرداً

أي أنه يجب على الباحث أن يأخذ عينة لا يقل حجمها عن 60 فرداً حتى يكون لديه تقديراً دقيقاً عن متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقديره لمتوسط الدخل عن خمس دولارات، وذلك بدرجة ثقة 99 %.

-تحديد حجم العينة لتقدير النسبة في المجتمع :

وبالطريقة نفسها يمكن تحديد حجم العينة اللازمة للحصول على درجة ثقة معينة عند تقدير النسبة في المجتمع بافتراض أن اقصى خطأ في التقدير مسموح به هو e تبعاً للمعادلة التالية:

حيث :

Z هو معامل الثقة (أو الدرجة المعيارية) المقابل لدرجة الثقة ونحصل عليه من جدول التوزيع الطبيعي المعياري.

P هي النسبة في المجتمع (أو تقدير لها).

1 - P هي النسبة المكملة. بمعنى إذا كانت نسبة المؤيدين 60 % فإن نسبة غير المؤيدين 40 %

e أقصى خطأ في التقدير مسموح به. "أو الخطأ في تقدير النسبة".

أي أن حجم العينة المناسب في هذه الحالة يساوي حاصل ضرب مربع z في النسبة، ثم في النسبة المكملة مقسوماً على مربع الخطأ المسموح به كما في المثال التالي:

مثال تطبيقي: يدعي أحد مراكز استطلاعات الرأي العام أنه عند دراسته لاتجاهات آراء الناخبين لاثنين من المتنافسين على أحد مقاعد السلطة التشريعية بأن نتائج دراسته هي من الدقة بحيث لا يتعدى نسبة الخطأ في التقدير 2%، فما هو حجم العينة المناسب التي نستطيع من خلالها الحكم على مدى صحة إدعاء هذا المركز بافتراض أن نسبة المؤيدين للمرشح هي 50 % وذلك بدرجة ثقة 95 %.

الحل:

بما أن درجة الثقة 95 % فإن: Z = 1.96 بافتراض أن نسبة المؤيدين للمرشح هي

$$P= 0.5$$

وبالتالي فإن النسبة المكملة 1 - P هي:

$$P = 1 - 0.5 = 0.5 -1$$

وحيث أن أقصى خطأ مسموح به هو:

$$e = 0.02$$

أي أن حجم العينة المناسب الذي يعطي درجة الدقة المطلوبة هو 2401 ناخب. بمعنى آخر فإن على هذا المركز أن يستطلع حجم عينة لا يقل عددها عن هذا العدد.