المعهد العربي للتدريب والبحوث الإحصائية

دورة متوسطة حول تصميم عينات المسوح الإحصائية

العينة العشوائية البسيطةSimple Random Sample SRS

إعداد:

نايف عابد

nayif@pcbs.gov.ps nayifabed@yaoo.com

أيار – 2022

المحتويات

- خصائص العينة العشوائية البسيطة بإرجاع او بدون إرجاع
- تقديرات المتوسط والمجموع والنسبة المئوية عند استخدام العينة العشوائية البسيطة
- تقدير التباين في حالة العينة العشوائية البسيطة بإرجاع أو بدون إرجاع
- حجم العينة في حالة العينة العشوائية البسيطة
- تطبيقات عملية لسحب عينة عشوائية بسيطة وتقدير التباين باستخدام برنامج SPSS

خصائص تصميم العينة العشوائية البسيطة

- ❖يجب أن يتوفر إطار معاينة محدد عدد وحداته N وهي مرقمة من 1 إلى N
- ❖تعتمد في الأساس على استخدام جدول الأرقام العشوائية
- ❖هناك نوعان من العينة العشوائية البسيطة:
  - ❑مع الإرجاع (SRSWR (with replacement
  - ❑بدون الإرجاع (SRSWOR(without replacement
- ❖احتمال سحب أي عنصر في العينة n من المجتمع N هو احتمال متساوي n/N

❖تعتبر العينة العشوائية البسيطة إحدى طرق سحب العينات ذات الاحتمال المتساوي

Equal Probability Selection Method "EPSEM"

❖عدد العينات التي حجمها **n** والتي يمكن سحبها بدون إرجاع من مجتمع محدد عدد

وحداته الكلية = **N** هو توافيق عدد **n** من **N** أو تكتب بالشكل ${}_N C_n$ أو $\binom{N}{n}$

$$ {}_N C_n = \frac{N!}{n!(N-n)!} $$

❖لنفرض أن **K** عدد العينات التي حجمها **n** وهو ${}_N C_n$ ، وكل عينة سنحسب منها

المتوسط $\bar{x}_{ni}$

❖سيكون لدينا عدد **K** من المتوسطات $\bar{x}_{n1}, \bar{x}_{n2}, \ldots\ldots\ldots \bar{x}_{nK}$

❖حسب نظرية النزعة المركزية Central Limit Theorem فإن المتوسطات

$\bar{x}_{n1}, \bar{x}_{n2}, ...\bar{x}_{nk}$ تتوزع توزيعاً طبيعياً ، المتوسط هو $\bar{X}$ والتباين هو $V(\bar{X})$

❖متوسط جميع المتوسطات من كل العينات التي حجمها n يساوي متوسط المجتمع،

بمعنى:

$$\bar{X} = \frac{\bar{x}_{n1} + \bar{x}_{n2} + ... + \bar{x}_{nK}}{K}$$

❖في الواقع العملي لا يتم سحب سوى عينة واحدة حجمها n ويتم إجراء التقديرات

بناء عليها.

تقديرات المتوسط والمجموع والنسبة المئوية الناتجة عن عينة عشوائية بسيطة

تقدير المتوسط

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{n}$$

تقدير المجموع

$$x = \frac{N}{n}\sum_{1}^{n} x_i = N\bar{x}$$

تقدير النسبة المئوية

$$p = \frac{\sum_{1}^{n} x_i}{n}$$

- N عدد الوحدات في المجتمع
- n عدد الوحدات في العينة
- x_i قيمة المتغير للعنصر i^{th} في العينة من (i=1,2,..n)

تقدير التباين في العينة العشوائية البسيطة

اذا كانت عناصر العينة: $i=1, 2, \ldots, n$

المتغير x الذي يراد دراسته $x_1, x_2, \ldots, x_n$

وكان متوسط العينة $\bar{x}$

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{n}$$

وكان تباين المتغير *x* في العينة

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

فان تباين متوسط العينة هو

$$V(\bar{x}) = (1-\frac{n}{N})\frac{s^2}{n} = (1-f)\frac{s^2}{n}$$

حيث *f* يسمى كسر المعاينة sample fraction، و $(1-f)$ هو تصحيح المجتمع

المحدود Finite Population Correction ***fpc***

تباين المتوسط في العينة العشوائية البسيطة بدون ارجاع

$$V(\bar{x}) = (1 - f) \times \frac{s^2}{n}$$

الخطأ المعياري في العينة العشوائية البسيطة بدون إرجاع

$$SE(\bar{x}) = \sqrt{V(\bar{x})} = \sqrt{1 - f} \times \frac{s}{\sqrt{n}}$$

معامل التباين (الخطأ النسبي) في العينة العشوائية البسيطة بدون إرجاع

$$CV(\bar{x}) = \frac{SE(\bar{x})}{\bar{x}} * 100\% = \sqrt{\frac{(1-f)}{n}} * \left(\frac{s}{\bar{x}}\right) * 100\%$$

تباين المجموع في العينـة العشوائيـة البسيطة بدون إرجاع

$$V(x) = N^2(1 - f)\frac{s^2}{n}$$

الخطأ المعياري في العينـة العشوائيـة البسيطة بدون إرجاع

$$SE(x) = \sqrt{V(x)} = \sqrt{N^2(1 - f)\frac{s^2}{n}}$$

معامل التباين (الخطأ النسبي) في العينـة العشوائيـة البسيطة بدون إرجاع

$$CV(x) = \frac{SE(x)}{x} * 100$$

تباين النسبة المئوية في العينة العشوائية البسيطة بدون إرجاع

$$V(p) = (1 - f)\frac{pq}{n-1} \qquad\qquad s^2 = p(1-p) = pq$$

الخطأ المعياري في العينة العشوائية البسيطة بدون إرجاع

$$SE(p) = \sqrt{V(p)} = \sqrt{1 - f\,\frac{pq}{n-1}}$$

معامل التباين (الخطأ النسبي) في العينة العشوائية البسيطة بدون إرجاع

$$CV(p) = \frac{SE(p)}{p} \times 100\%$$

فترة الثقة:

100(1−α) % فترة الثقة لتقدير المتوسط

الحد الادنى لفترة الثقة

القيمة الجدولية لفترة ثقة
100 (1−α) %

$$C_L = \bar{x} - Z_{\frac{\alpha}{2}} * SE(\bar{x})$$

$$C_U = \bar{x} + Z_{\frac{\alpha}{2}} * SE(\bar{x})$$

الحد الاعلى لفترة الثقة

طول فترة الثقة = $2 * Z_{\frac{\alpha}{2}} * SE(\bar{x})$

بنفس الطريقة يتم حساب فترة الثقة للمجموع والنسبة المئوية

حجم العينة:

1- بالاعتماد على النسبة المئوية:

من اجل تقدير حجم العينة بالاعتماد على النسبة P ينبغي معرفة ما يلي:

- نسبة انتشار الظاهرة الرئيسية التي سيتم دراستها في المسح P لمعرفة تباين التقدير S^2
- معامل الثقة الذي سيعبر عن حدود فترة الثقة t
- الخطأ الهامشي المطلق في تقدير المؤشر E

وهذه المعلومات السابقة تلزم لاستخدام معادلة حجم العينة المطلق وهي:

$$n = \frac{t^2 * S^2}{E^2}$$

ويتم احتساب قيمة S^2 من بيانات سابقة باستخدام المعادلة:

$$s^2 = p(1-p) = pq$$

مثال:

يراد دراسة ظاهرة معينة في منطقة محددة من خلال دراسة مسحية ، إذا علمت من بيانات سابقة أن نسبة انتشار هذه الظاهرة 30% بين الأسر، قدر حجم عينة مناسب من الأسر على افتراض أن العينة عشوائية بسيطة واخترنا فترة ثقة بحدود 95% مع خطأ هامشي مطلق لتقدير المؤشر قيمته 4%.

الحل:

$$n = \frac{t^2 * P(1-P)}{E^2}$$

بما أن فترة الثقة بحدود 95% فان قيمة t تساوي 1.96 وبالتعويض بالمعادلة السابقة:

$$n = \frac{(1.96)^2 * 0.3(0.7)}{(0.04)^2}$$

$n = 504.21$

$= 504$

يتم تقريب حجم العينة لعدد صحيح.

إذا حجم العينة المطلوب لهذه الدراسة هو 504 أسرة على اعتبار أنها عينة عشوائية بسيطة أو منتظمة.

حجم العينة:

1- بالاعتماد على تقدير المتوسط:

من اجل تقدير حجم العينة بالاعتماد على المتوسط $\overline{x}$ينبغي معرفة ما يلي:

- متوسط القيمة الرئيسية التي سيتم دراستها في المسح $\overline{x}$ومعرفة تباين التقدير S^2
- معامل الثقة الذي سيعبر عن حدود فترة الثقة t
- الخطأ الهامشي المطلق في تقدير المؤشر E

وهذه المعلومات السابقة تلزم لاستخدام معادلة حجم العينة المطلق وهي:

$$n = \frac{t^2 * S^2}{E^2}$$

ويتم احتساب قيمة S^2 باستخدام المعادلة:

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

مثال:

يراد دراسة متوسط انفاق الاسرة في منطقة محددة من خلال دراسة مسحية ، إذا علمت من بيانات سابقة أن متوسط الانفاق للاسرة كان 1000 دولار من بيانات سابقة.

قدر حجم عينة مناسب من الأسر على افتراض أن العينة عشوائية بسيطة واخترنا فترة ثقة بحدود 95% مع خطأ هامشي مطلق لتقدير المؤشر قيمته 4%.

وكان الانحراف المعياري لقيم الدخل من بيانات سابقة =500

الحل:

$$n = \frac{t^2 * S^2}{E^2}$$

بما أن فترة الثقة بحدود 95% فان قيمة t تساوي 1.96 وبالتعويض بالمعادلة السابقة:

$$n = \frac{1.96^2 * 500^2}{(0.04 * 1000)^2}$$

n = 600.25

= 600

يتم تقريب حجم العينة لعدد صحيح.

إذا حجم العينة المطلوب لهذه الدراسة هو 600 أسرة على اعتبار أنها عينة عشوائية بسيطة أو منتظمة.

تطبيقات عملية على برنامج SPSS