كيفية حساب حجم العينة n من مجتمع طبيعي حجمه N عنصراً

د. إبراهيم محمد العلي
أستاذ في كلية الاقتصاد بجامعة تشرين

I- مقدمة :

إن حجم العينة العشوائية n التي يجب سحبها من مجتمع الدراسة يرتبط بعدة أمور هي:

1- نوع المعاينة: المعاينة البسيطة- المعاينة الطبقية- المعاينة العنقودية ...الخ . ونظراً لتعقيدات الموضوع وتشعباته فإننا هنا سنقتصر على حساب حجم العينة n في المعاينة العشوائية البسيطة فقط .

2- طريقة السحب: وهناك طريقتان للسحب هما : السحب مع الإعادة- السحب بدون إعادة .

3- نوع المؤشر المدروس: ونقصد به المؤشر الذي نريد تقديره من معلومات العينة وهو غالباً ما يكون أحد المؤشرين التاليين:

- متوسط المؤشر في المجتمع μ والذي سنقدره بمتوسط العينة $\bar{x}$ المعرف بالعلاقة: $\bar{x} = \frac{\sum x_i}{n}$.
- نسبة خاصة معينة في المجتمع R والتي سنقدرها بالنسبة في العينة r المعرف بالعلاقة: $r = \frac{m}{n}$ ،حيث m عدد المتصفين بالخاصة المدروسة في العينة .

4- درجة تجانس المجتمع: ويعبر عنه من خلال تباين المجتمع σ^2 ، فإذا كان التجانس ضعيفاً فإن التباين σ^2 يكون كبيراً ، وعندها فإن حجم العينة n يجب أن يكون كبيراً ، أي أن حجم العينة n يتناسب طرداً مع التباين σ^2 . أما إذا كان تباين المجتمع σ^2 مجهولاً فيتم تقديره من خلال تباين العينة s^2 المعرف بالعلاقة: $s^2 = \frac{\sum(\mathrm{x_i}-\bar{\mathrm{x}})^2}{\mathrm{n}-1}$

5- مقدار الدقة المطلوبة في التقدير: ويرمز له بـ d ، وهو الحد الأعلى للخطأ المسموح به عند تقدير المؤشرات: ويتم تحديده مسبقاً من قبل المسؤولين عن البحث حسب نوع المؤشر المدروس .

فإذا كان المؤشر المطلوب تقديره هو متوسط المجتمع μ فإننا نضع مقدار الدقة d مساوياً لأي عدد مطلق، (فمثلاً إذا كنا نريد تقدير متوسط الدخل فيمكن أن نضع ل.س d=200 أو أي عدد آخر) . وهذا يعني أنه عند تقدير متوسط الدخل يجب أن يكون الخطأ المرتكب فيه (الخطأ المعياري) أقل أو يساوي المقدار المحدد للدقة d .

أما إذا كان المطلوب تقدير النسبة R في المجتمع ، فإن الدقة تحدد على شكل نسبة مئوية مثل d=6% أو d=3% أو d=2% . وهذا يعني أنه عند تقدير النسبة R يجب أن يكون الخطأ المرتكب فيه (الخطأ المعياري) أقل أو يساوي النسبة المحددة للدقة d .

6- الخطأ المعياري للمؤشر المدروس: وهو يحسب لكل مؤشر على حدة ويبرهن في نظرية العينات والاحتمالات على ما يلي:

- إن الخطأ المعياري المرتكب في تقدير متوسط المجتمع μ من خلال متوسط العينة $\bar{x}$ والذي سنرمز له بـ $\sigma_{\bar{x}}$ يتأثر بطريقة السحب، وهو يساوي حسب حالتي السحب ما يلي:

(1) (في حالة السحب مع الإعادة)
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{S}{\sqrt{n}}$$

(2) (في حالة السحب بدون إعادة)
$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N}} * \frac{\sigma}{\sqrt{n}} \approx \sqrt{\frac{N-n}{N}} * \frac{s}{\sqrt{n}}$$

حيث:s هو الانحراف المعياري للعينة ، ويعتبر تقديراً غير متحيز للانحراف المعياري في المجتمع.

- إن الخطأ المعياري المرتكب في تقدير النسبة R في المجتمع من خلال النسبة في العينة r والذي سنرمز له بـ σ_r يساوي حسب حالتي السحب ما يلي:

(3) (في حالة السحب مع الإعادة)
$$\sigma_r = \sqrt{\frac{R*(1-R)}{n}} \approx \sqrt{\frac{r*q}{n}}$$

(4) (في حالة السحب بدون إعادة)
$$\sigma_r = \sqrt{\frac{N-n}{N}} * \sqrt{\frac{R*(1-R)}{n}} \approx \sqrt{\frac{N-n}{N} * \sqrt{\frac{r.q}{n}}}$$

حيث r هي النسبة في العينة أو في أية عينة تجريبية و q=1-r .

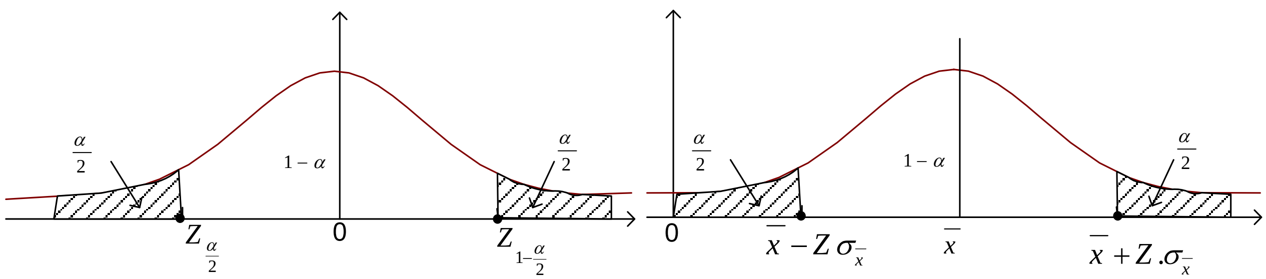
7- مستوى الدلالة المطلوب $\propto$: ويتم تحديده مسبقاً من قبل المسؤولين عن البحث وهو يعبر عن احتمال أن يكون التقدير المأخوذ من العينة مرفوضاً، أو أن يكون مقبولاً باحتمال $(1-\propto)$، ويسمى الاحتمال $(1-\propto)$ باحتمال الثقة في التقدير، لأنه يضمن لنا أن يكون المؤشر المطلوب تقديره واقعاً ضمن مجال محدد يسمى مجال الثقة المقابل للاحتمال $(1-\propto)$. وعندما نحدد مستوى الدلالة $\propto$

نوزعه على طرفي المجال ونضع لكل طرف $\frac{\propto}{2}$. وعندها تتحدد لدينا من الجداول القيمتان العدديتان لمتحول التوزيع الطبيعي المعياري $Z_{\frac{\propto}{2}}$ و $Z_{1-\frac{\propto}{2}}$ الفاصلتان بين منطقتي القبول والرفض , وبسبب التناظر يكون لدينا $Z_{1-\frac{\propto}{2}} = -Z_{\frac{\propto}{2}}$ ، لذلك نأخذ القيمة $Z_{1-\frac{\propto}{2}}$ ونقوم بإنشاء مجال الثقة المقابل لذلك . وتكون منطقة الرفض واقعة خارج مجال الثقة .

ولإنشاء مجال الثقة لمتوسط المجتمع μ نجعل مركزه متوسط العينة $\bar{x}$ ونصف طوله يساوي $(Z * \sigma_{\bar{x}})$ فنحصل على مجال الثقة الذي يحقق احتمال الثقة المطلوب وهو المجال التالي:

$$P\left[\bar{x} - Z_{1-\frac{\propto}{2}} * \sigma_{\bar{x}} \leq \mu \leq \bar{x} + Z_{1-\frac{\propto}{2}} * \sigma_{\bar{x}}\right] = 1-\propto \qquad (5)$$

وهو يأخذ الشكل البياني التالي :



الشكل (2) الشكل المعياري لمجال الثقة   الشكل (1) الشكل العام لمجال الثقة للمتوسط

ولإنشاء مجال الثقة للنسبة R في المجتمع نجعل مركزه النسبة في العينة r ونصف طوله $(Z * \sigma_r)$ فنحصل على مجال الثقة الذي يحقق احتمال الثقة $(1-\propto)$ التالي:

$$P = \left[r - Z_{1-\frac{\propto}{2}} * \sigma_r \leq R \leq r + Z_{1-\frac{\propto}{2}} * \sigma_r\right] = 1-\propto \qquad (6)$$

II- كيفية حساب حجم العينة (في حالة السحب مع الإعادة): لمتوسط المجتمع μ وللنسبة فيه R :

لحساب حجم العينة n في المعاينة العشوائية البسيطة علينا أن نراعي جميع الشروط المذكورة أعلاه ، وخاصة شرط الدقة d (الحد الأعلى للخطأ المسموح به عند التقدير) ، وهنا علينا أن نحسب حجم العينة n لكل مؤشر على حدة (لمتوسط المجتمع μ وللنسبة فيه R).

1- إذا كان المطلوب تقدير متوسط المجتمع الطبيعي μ ، وكان السحب مع الإعادة، وكان مقدار الدقة المطلوبة d , فإننا نجعل نصف طول مجال الثقة للمتوسط μ أصغر أو يساوي مقدار الدقة المحدد d فنحصل على ما يلي:

$$Z_{1-\frac{\propto}{2}} * \sigma_{\bar{x}} \leq d \qquad (7)$$

وبما أن $\left(\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right)$ نجد أن :

$$Z^2 * \sigma_{\bar{x}}^2 \leq d^2$$

$$Z^2 * \frac{\sigma^2}{n} \leq d^2$$

ومنها نجد أن :

$$n \geq \frac{Z^2 * \sigma^2}{d^2} \approx \frac{Z^2 S^2}{d^2}$$ (8) (وهو الحد الأدنى لحجم العينة مع الإعادة)

حيث أن: s^2 هو تباين العينة أو أي عينة تجريبية عشوائية والذي يؤخذ كتقدير لتباين المجتمع σ^2 .

مثال (1): لتقدير متوسط الوزن لطلاب الكلية نريد سحب عينة عشوائية منهم مع الإعادة. والمطلوب حساب الحجم اللازم لهذه العينة لتحقيق دقة لاتتجاوز d=3 kg وباحتمال 95% ، إذا علمت أنه تم تقدير تباين المجتمع σ^2 من خلال تباين عينة تجريبية فكان $S^2 = 560$.

ولحساب حجم هذه العينة نطبق العلاقة (8) ونلاحظ أن مستوى الدلالة $\propto = 0.05$ وأن Z= 1.96

ومنها نجد أن :

$$n \geq \frac{(1.96)^2(560)}{(3)^2} = 239.03 \approx 240$$

وهو حجم العينة الكافي نظرياً لتحقيق الشروط المفروضة على المسألة من حيث الدقة واحتمال الثقة .

علماً بأنه يجب تقريب حجم العينة إلى الأعلى دائماً .

ملاحظة : بعد سحب هذه العينة وإجراء الحسابات المطلوبة نقوم بحساب تباينها الحقيقي S^2 ، ثم نعود ونحسب حجم العينة من جديد للتأكد من أن حجم العينة المسحوبة يحقق الشروط المفروضة على المسألة المدروسة . فإذا كان حجم العينة المسحوبة أقل من حجمها الجديد ، فهذا يعني أنها لا تحقق الشروط المفروضة على المسألة لأسباب تتعلق بعشوائية السحب أو بصحة البيانات أو بغير ذلك ، ولتصحيح ذلك نقوم بزيادة حجم العينة إلى الحجم المطلوب ونعيد الحسابات من جديد .

2- إذا كان المطلوب تقدير النسبة R في المجتمع، وكان السحب مع الإعادة ، فإننا نجعل نصف طول مجال الثقة أصغر أو يساوي نسبة الدقة d والمحددة كنسبة مئوية صغيرة مثل (2% أو 3% ....) .

$$Z_{1-\frac{\propto}{2}} * \sigma_r \leq d \qquad (9)$$

وبما أن $\left(\sigma_r^2 = \frac{R(1-R)}{n}\right)$ نجد أن:

$$Z^2 * \sigma_r^2 \leq d^2$$

$$d^2 \geq Z^2 \frac{R(1-R)}{n} \approx Z^2 * \frac{r*q}{n}$$

$$n \geq \frac{Z^2 * r * q}{d^2}$$ (10) (وهذا الحد الأدنى لحجم العينة مع الإعادة)

حيث أن: r هي النسبة في العينة أو في أي عينة تجريبية أو سابقة و q=1 - r .

مثـال (2) : لنفرض إننـا نريـد حسـاب حجـم العينـة n المطلـوب سـحبها مـع الإعـادة لتقـدير نسـبة المدخنين في مجتمع ما بدقة قدرها d= 0.04 وباحتمال ثقة 95% ، أي أن $\left(Z_{1-\frac{\alpha}{2}} = 1.96\right)$ ، علماً بأن المعلومات السابقة تشير إلى أن نسبتهم r = 0.25 ، للجواب نطبق العلاقة (10) فنجد أن:

$$n \geq \frac{(1.96)^2(0.25)(0.75)}{(0.04)^2} = \frac{0.7203}{0.0016} = 450.19 \approx 451$$

وبعد سحب هذه العينة وحساب النسبة r فيها نعود ونحسب حجم العينة من جديد للتأكد من أن حجم العينة المسحوبة يحقق الشروط المفروضة على المسألة المدروسة .

III– كيفية حساب حجم العينة (في حالة السحب بدون إعادة): لمتوسط المجتمع μ ثم للنسبة فيه R :

1– إذا كان المطلوب تقدير متوسط مجتمع طبيعي μ وكان السحب بدون إعادة، وكان مقدار الدقة المطلوبة d . فإننا نجعل نصف طول مجال الثقة أصغر أو يساوي مقدار الدقة d فحصل على ما يلي:

$$\mathrm{Z}_{1-\frac{\alpha}{2}} * \sigma_{\bar{x}} \leq \mathrm{d} \qquad (11)$$

$$Z^2 * \sigma_{\bar{x}}^2 \leq d^2$$

وبما أن: $\sigma_{\bar{x}}^2 = \frac{N-n}{N} * \frac{\sigma^2}{n}$ نجد أن

$$Z^2 * \frac{N-n}{N} * \frac{\sigma^2}{n} \leq d^2$$

ومنها نجد أن :

$$n * N * d^2 \geq Z^2 * N\sigma^2 - Z^2 * n\sigma^2$$

$$nNd^2 + Z^2 * n\sigma^2 \geq Z^2 * N\sigma^2$$

$$n(Nd^2 + Z^2\sigma^2) \geq Z^2 * N * \sigma^2$$

$$n \geq \frac{N*Z^2*\sigma^2}{Nd^2+Z^2*\sigma^2} \qquad (12)$$

وباستبدال σ^2 بتباين العينة s^2 أو أي عينة تجريبية نحصل على ما يلي:

$$n \geq \frac{NZ^2*S^2}{Nd^2+Z^2S^2} \qquad (13)$$

وهو الحد الأدنى لحجم العينة (في حالة السحب بدون إعادة)

مثـال (3): نريد تقدير متوسط دخل الأسرة في الشهر في مجتمع مؤلف من 1000 =N أسرة ، وبدقة لاتتجاوز ل.س500 = d وباحتمال ثقة 95% . والمطلوب حساب حجم العينة n اللازم لذلك، إذا علمت أن تباين الدخل في ذلك المجتمع $(\sigma^2 = 7562000)$ وإن السحب بدون إعادة . لحساب حجم العينة n نلاحظ أن مستوى الدلالة $\alpha = 0.05$ وأن $Z_{1-\frac{\alpha}{2}} = 1.96$ وبتطبيق العلاقة (12) نحصل على أن:

$$n \geq \frac{1000(1.96)^2(7562000)}{1000(500)\text{^}2 + (1.96)^2(7562000)} = \frac{29050179200}{279050179.2} = 104.10 \approx 105$$

وهو حجم العينة n الكافي نظرياً لتحقيق شروط المسألة في الدقة واحتمال الثقة، مع ملاحظة أنه يجب تقريب حجم العينة إلى الأعلى دائماً .

ملاحظة: إذا كان تباين المجتمع σ^2 مجهولاً ، فإننا نعمل على إيجاد تقدير له من خلال سحب عينة تجريبية بحجم معقول ثم حساب تباينها S^2 وتعويضه في العلاقة (13) فنحصل على الحجم المطلوب. وأخيراً نذكر إلى أنه بعد سحب العينة وإجراء الحسابات اللازمة من بياناتها، يجب أن نتأكد من إنها تحقق الشروط المفروضة على المسألة .

ملاحظة: من مقارنة العلاقتين (8) و (13) نجد أن حجم العينة n المحسوبة بنفس الشروط في حالة السحب بدون إعادة أصغر من حجمها n_0 في حالة السحب مع الإعادة، وهما يرتبطان بالعلاقة :

$n = \frac{N*n_0}{N+n_0}$.

تعليق: إن العلاقة (13) هي العلاقة الأساسية لحساب حجم العينة في حالة السحب بدون إعادة ، ولكن بعض المؤلفين والباحثين قام بتحويرها إلى أشكال أخرى . فقام أحدهم بتقسيم البسط والمقام على s^2 ، فحصل على أن:

$$n \geq \frac{NZ^2}{N\left(\frac{d^2}{S^2}\right)+Z^2} \approx \frac{NZ^2}{Ne^2+Z^2}$$

(14) حيث أن $\left(e^2 = \frac{d^2}{S^2}\right)$

وقام بعضهم بتقسيم البسط والمقام على (Z^2s^2) فحصل على أن :

$$n \geq \frac{N}{N\left(\frac{d^2}{Z^2S^2}\right)+1} \approx \frac{N}{NE^2+1}$$

(15) حيث أن $\left(E^2 = \frac{d^2}{Z^2*S^2}\right)$

وقام بعضهم بتقسيم البسط والمقام على $\bar{x}^2$ وادخل معامل الاختلاف CV في المعادلة (13) فحصل على أن:

$$n \geq \frac{NZ^2\frac{S^2}{\bar{x}^2}}{N\left(\frac{d^2}{\bar{x}^2}\right)+Z^2*\frac{S^2}{\bar{x}^2}} = \frac{NZ^2*CV^2}{N\delta^2+Z^2*CV^2}$$

(16) حيث أن ($CV = \frac{S}{\bar{x}}$ وأن $\delta = \frac{d}{\bar{x}}$ الدقة النسبية)

وقام بعضهم بتقسيم البسط والمقام على حجم المجتمع N (أو على Nd^2) فحصل على أن:

$$n \geq \frac{Z^2\sigma^2}{d^2+\frac{Z^2\sigma^2}{N}} \approx \frac{Z^2s^2}{d^2+\frac{Z^2S^2}{N}} \tag{17}$$

ومنها استنتج أنه عندما يكون حجم المجتمع N كبيراً فإن المقدار ($\frac{Z^2S^2}{N}$)يصبح مهملاً وإن العلاقة السابقة(16) تأخذ الشكل التالي :

$$n \geq \frac{Z^2S^2}{d^2} \tag{18}$$

وهي نفس العلاقة (8) لحالة السحب مع الإعادة . ومنها يمكننا أن نستنتج مايلي:

نتيجة هامة. عندما يكون حجم المجتمع N كبيراً فإن حجم العينة n في حالة السحب بدون إعادة يصبح متقارباً مع حجمها في حالة السحب مع الإعادة، وعندها يمكننا حساب حجم العينة في كلتا حالتي السحب من العلاقة :

$$n \geq \frac{Z^2 S^2}{d^2} \qquad : \qquad (n << N) \qquad (19)$$

حيث s^2 هو تباين العينة أو أي عينة تجريبية .

علماً بأنه كان يمكننا الحصول على هذه النتيجة من قيمة الخطأ المعياري للمتوسط المعرف في العلاقة (2) . حيث نجد أنـه عندما يكـون N كبيـراً فيكـون $(n << N)$ وعنـدها فـإن كسـر معامـل التصـحيح يصبح قريباً من الواحد أي أن :

$$\frac{N-n}{N} \approx 1 \qquad : \qquad (n << N)$$

وعندها نجد أن الخطأ المعياري لحالة السحب بدون إعادة يصبح متقارباً من الخطأ مع الإعادة أي أن :

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N}} * \frac{\sigma}{\sqrt{n}} \approx \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}} \qquad : \qquad n << N \qquad (20)$$

2- إذا كان المطلوب تقدير النسبة R في المجتمع وكان السحب بدون إعادة ، فإننا نجعل نصف طول مجال الثقة أصغر أو يساوي نسبة الدقة المحددة d فنجد أن :

$$Z_{1-\frac{\alpha}{2}} * \sigma_r \leq d \qquad \sigma_r = \sqrt{\frac{N-n}{N}} * \sqrt{\frac{R(1-R)}{n}} \quad \text{حيث أن}$$

$$Z^2 * \frac{N-n}{N} * \frac{R(1-R)}{n} \leq d^2$$

$$n * Nd^2 \geq N * Z^2 R(1-R) - nZ^2 * R(1-R)$$

$$n\left(Nd^2 + Z^2 R(1-R)\right) \geq N * z^2 * R(1-R)$$

ومنها نجد أن:

$$n \geq \frac{N * Z^2 * R(1-R)}{Nd^2 + Z^2 * R(1-R)} \qquad (21)$$

وباستبدال النسبة R بالنسبة r في العينة أو في أي عينة تجريبية نحصل على أن :

$$n \geq \frac{N * Z^2 * r * q}{Nd^2 + Z^2 * r * q} \qquad : \quad q=1-r \text{ :حيث أن} \qquad (22)$$

وهي العلاقة الأساسية لحساب حجم العينـة n عند تقدير النسبة R ، ويمكن تحويرها إلى أشكال أخرى كما رأينا في حالـة المتوسط من خـلال علاقات مشابهة للعلاقات (14)، (15)، (16)، (17) .فمثلاً لو قسمنا البسط والمقام على (Nd^2) , **نحصل على العلاقة الآتية:** (23)

$$n \geq \frac{(Z/d)^2 * r * q}{1 + (1/N)(Z/d)^2 * r * q}$$

ملاحظـة: عند حساب حجم العينـة ، غالبـاً ما يتم حساب النسبة r من أي عينـة تجريبيـة وقد تكون قيمتها الحقيقية 20% أو 30% أو 40% أو غير ذلك ، وبعد تعويضـها في العلاقـة (22) نحصـل علـى حجم العينـة المطلوب . ولكن عندما تكون النسبة المدروسة متوازنـة في المجتمع المدروس (كنسبة الإناث في المجتمع

العام التي تأخذ حوالي (R=0.50) فإن حجم العينة n المحسوبة من تلك العلاقة يبلغ أكبر قيمة له . وذلك لأن الجداء ((R*(1-R)) يأخذ أكبر قيمة له عندما تكون (R=0.50) , ويساوي; R*(1-R)(0.50)(0.50)=(0.50)(0.25) .

وبناءً على هذه الخاصة يقوم بعض الباحثين بحساب حجم n من العلاقة (22) بافتراض أن النسبة متوازنة في المجتمع ويضعون فيها (r = 0.50) و (q= 0.50) ويعوضون قيم باقي المؤشرات d, N, Z, بقيمها المحددة ، فيحصلون على أكبر حجوم للعينات المقابلة لإحجام مختلفة لمجتمعات متوازنة (وذلك في حالة r= 0.50) .وبناءً على ذلك قام بعضهم بتطبيق العلاقة (22) وحسب أحجام العينات المقابلة لمستويات معينة من الدقة d وعند مستوى دلالة محدد $\propto = 0.05$ والمقابل لـ Z= 1.96 فحصل على الجدول التالي :

الجدول (1) أحجام العينات(بدون إعادة)المقابلة لأحجام مختلفة من مجتمعات متوازنة وحسب قيم مختلفة للدقة d

حجم المجتمع N	نسبة الدقة =d			
	%5	%3	%2	%1
50	44	48	49	50
100	79	91	96	99
150	108	132	141	148
200	132	168	185	196
250	151	203	226	244
300	168	234	267	291
400	196	291	343	384
500	217	340	414	475
750	254	440	571	696
1000	278	516	706	906
2000	322	696	1091	1655
5000	357	879	1622	3288
10000	370	964	1936	4899
100000	383	1056	2345	8762
1000000	384	1066	2395	9513
10000000	384	1067	2400	9595

المصدر : **sounders lewis & thornhl, 2009**

مثال (4): لتوضيح كيفية حساب أحجام هذه العينات نفترض أننا نريد أن نسحب بدون إعادة عينة بحجم n من مجتمع طبيعي (كلية الاقتصاد) حجمه معلوم N = 1000 لتقدير نسبة الإناث فيه وبدقة محددة d=0.03 وباحتمال ثقة قدره 95% (أي بمستوى دلالة $\propto = 0.05$ وهو يقابل قيمة حرجة تساوي $Z_{1-\frac{\propto}{2}} = 1.96$).. وللحصول على أكبر حجم للعينة المطلوبة نفترض أن هذا المجتمع متوازن جنسياً ونضع نسبة الإناث R=0.50 ونعوض في العلاقة (21) فنحصل على أن:

(انظر الجدول): $n \geq \frac{1000(1.96)^2(0.50)(0.50)}{1000(0.03)^2+(1.96)^2(0.50)(0.50)} = \frac{960.4}{1.8604} = 516.23 \approx 517$

ولكن إذا علمنا من شؤون الطلاب أن نسبة الإناث قي ذلك المجتمع R = 0.40 ، فإننا نكون في مجتمع غير متوازن من حيث الجنس . وعندها نقوم بتعويض هذه القيمة في العلاقة (21) فنحصل على أن حجم العينة المطلوب يساوي :

$$n \geq \frac{1000(1.96)^2(0.40)(0.60)}{1000(0.03)^2 + (1.96)^2(0.40)(0.60)} = \frac{921.984}{1.821984} = 506$$

وهو حجم العينة الكافي لتحقيق الشروط المفروضة على هذه المسألة .

ولكن إذا كنا نريد تقدير نسبة المدخنين في المجتمع المذكور وبنفس الشروط السابقة وكانت هذه النسبة مجهولة فإننا نكون في مجتمع غير متوازن من حيث خاصة التدخين، لذلك يجب أن نبحث عن تقدير أولي لنسبة المدخنين فيه، ولهذا نسحب عينة تجريبية بحجم معقول (وليكن n=100) ونستخلص منها تقدير نسبة المدخنين الأولية في المجتمع ، ولنفترض إنها كانت تساوي r= 0.35 . وعندها نقوم بحساب حجم العينة من العلاقة (22) فنجد أن:

$$n \geq \frac{1000(1.96)^2(0.35)(0.65)}{1000(0.03)^2 + (1.96)^2(0.35)(0.65)} = \frac{873.964}{1.77396} = 492.48 \approx 493$$

وهو حجم العينة الكافي لتحقيق شروط المسألة وهي أن يكون احتمال الثقة أكبر من 95% وأن تكون نسبة الخطأ المعياري (المرتكب) في تقدير النسبة لا تزيد عن مقدار الدقة المحددة بـ 3%. وكل ذلك بفرض أن جميع عناصر هذه العينة ستستجيب للمعاينة وتعطينا البيانات الصحيحة عن أحوالها .

أما إذا كانت الاستجابة غير كاملة فيجب أن نقسم حجم العينة n على نسبة الاستجابة p ، فنحصل على حجم العينة $\grave{n}$ اللازم للمعاينة من العلاقة : $\left(\grave{n} = \frac{n}{p}\right)$.

وفي جميع الأحوال يجب على الباحث أن يقوم بإعادة حساب حجم العينة $\grave{n}$ اللازم للمعاينة اعتماداً على بيانات العينة النهائية المستخدمة في البحث نفسه. وذلك للتأكد من حجمها ومن صلاحيتها ومن أنها تحقق الشروط المفروضة على البحث.

ملاحظة: لحساب حجم العينة الكلية في المعاينة الطبقية يمكننا استخدام نفس العلاقات وتطبيقها على المجتمع الطبقي ، فنحصل على حجم العينة الكلية n ، ثم نقوم بتوزيعها على الطبقات حسب أحجامها النسبية من المجتمع .كما يمكن اعتبار كل طبقة مجتمعاً مستقلاً وحساب أحجام العينات الطبقية وذلك بتطبيق نفس العلاقات السابقة على كل طبقة ، ثم دمج هذه العينات الطبقية فنحصل على العينة الكلية .

كما يمكن أخذ التكلفة بعين الاعتبار ...الخ . وهناك أمور كثيرة أخرى لا مجال للبحث فيها الآن .