المعاينة العشوائية البسيطة

1. التعريف

المعاينة العشوائية البسيطة هي طريقة سحب عينة مكونة من n وحدة من بين N وحدة من وحدات المجتمع محل الدراسة بحيث يكون لكل عينة من العينات الممكن اختيارها فرصة متساوية (احتمال متساو) في الظهور. ويلجأ إليها في حالة ما إذا كان المجتمع المراد دراسته ليس كبيراً ويحمل قدراً من التجانس بين وحداته بالنسبة للصفة أو الظاهرة موضوع الدراسة. يتم اختيار العينة العشوائية البسيطة بطريقة تعطي لكل مفردة من مفردات المجتمع نفس الفرصة في الظهور. بعبارة أخرى، يكون لكل وحدة من وحدات المجتمع نفس الاحتمال للانتماء للعينة.

2. الشروط

يشترط لتطبيق هذه الطريقة في اختيار العينة توفر الاعتبارات التالية:

- التوفر على قاعدة معاينة (إطار للمعاينة) حديثة وشاملة. ويقصد هنا بإطار المعاينة، اللائحة أو القائمة الشمولية لجميع وحدات المجتمع المدروس (بدون تكرار).
- أن يكون المجتمع متجانسا من حيث الظاهرة أو الخصائص التي نهدف الى دراستها من خلال البحث الاحصائي عن طريق العينة.

3. المزايا والعيوب

3.1. المزايا

- لا تتطلب معرفة معلومات أو خصائص خاصة لأفراد المجتمع (المعلومات المساعدة).
- سهولة الصيغ والمعادلات الرياضية المعتمدة لحساب احتمالات الانتماء للعينة والمقدرات المؤسسة للاستدلال الإحصائي وتحديد حجم العينة.

3.2 العيوب

- التكلفة العالية بالنظر للشرط الأساسي المتمثل في ضرورة توفر قاعدة المعاينة الشاملة والمحينة.
- تعطي دائما عينات يغلب عليها طابع التشتت الجغرافي.
- تطبيقها صعب في المجتمعات الكبيرة.
- دقة التقدير ضعيفة إذا كان المجتمع مقسما إلى فئات وطبقات بداخلها قدر كبير من التجانس.

4. تقدير المعلمات

تقتضي هذه العملية استعمال المشاهدات المحصل عليها بواسطة العينة لتقدير، حسب تصميم معين للمعاينة، المعلمات الإحصائية الأساسية على مستوى المجتمع المدروس. ويتعلق الأمر بصفة عامة بالمعلمات المرتبطة بالنزعة المركزية كالمتوسط الحسابي والوسيط والنسبة وبمقاييس التشتت الإحصائي كالتباين والانحراف المعياري ومعامل التغيير.

يطرح تقدير المعلمات الإحصائية عن طريق العينة عادة مسالتين هامتين:

- اختيار المقدر، الذي يعتمد غالبا على صيغة المعلمة المراد تقديرها.
- جودة المقدر، ويعبر عنها بواسطة التحيز أو التباين أو الفرق التربيعي المتوسط أو معامل التغيير.

المتوسط الحسابي

لدينا : $\bar{Y} = \frac{1}{N}\sum_{1}^{N} y_i$

المتوسط الحسابي للمجتمع المدروس. تجدر الإشارة إلى أن هذه المعلمة حقيقية إلا أنها غير معروفة القيمة.

مقدر المتوسط الحسابي للمجتمع المدروس هو المتوسط الحسابي للعينة أي $\bar{\ }$.

نكتب:

$$\hat{\bar{Y}} = \bar{y}$$

$$\bar{y} = \frac{1}{n}\sum_{1}^{n} y_i$$

$$\hat{\bar{Y}} = \frac{1}{n}\sum_{1}^{n} y_i$$

♣ **<u>التحيز</u>**

$$E\,(\widehat{\bar{Y}}) = E\left(\frac{1}{n}\sum_{1}^{n} y_i\right)$$

$$E\,(\widehat{\bar{Y}}) = \frac{1}{n}E\left(\sum_{1}^{n} y_i\right)$$

$$E\,(\widehat{\bar{Y}}) = \frac{1}{n}\sum_{1}^{n} E(y_i)$$

$$E\,(\widehat{\bar{Y}}) = \frac{n}{n}E(y_i)$$

$$E\,(\widehat{\bar{Y}}) = E(y_i) = \sum_{1}^{N} y_i\, p_i$$

حيث $_i$ هو احتمال سحب الوحدة i.

$$E\,(\widehat{\bar{Y}}) = \sum_{1}^{N} y_i \frac{1}{N}$$

$$E\,(\widehat{\bar{Y}}) = \bar{Y}$$

♣ **التباين**

انعدام التحيز يطمئننا بأن متوسط المتوسطات الحسابية لجميع العينات الممكنة يساوي المتوسط الحسابي للمجتمع $(\bar{Y})$ والذي نتوخى تقديره انطلاقا من المشاهدات المحصل عليها من خلال العينة. إلا أن هذا الأمر غير كاف للتأكد من جودة المقدر. إذ ينبغي التعرف على كيفية توزيع وتشتت المتوسطات الحسابية لكافة العينات الممكنة حول المتوسط الحسابي للمجتمع. وتجدر الإشارة في هذا الصدد إلى أن أحسن وسيلة للتأكد من ذلك هو قياس مؤشر التشتت أي تباين مقدر المتوسط الحسابي للمجتمع $V(\hat{\bar{Y}})$.

<u>الحالة الأولى: السحب مع الاحلال (مع الإعادة)</u>

$$V\left(\hat{\bar{Y}}\right) = V(\bar{y}) = V\left(\frac{1}{n}\sum_{1}^{n} y_i\right)$$

$$V\left(\hat{\bar{Y}}\right) = \frac{1}{n^2} V\left(\sum_{1}^{n} y_i\right)$$

$$V\left(\hat{\bar{Y}}\right) = \frac{1}{n^2}\left(\sum_{1}^{n} V(y_i)\right)$$

$$V\left(\hat{\bar{Y}}\right) = \frac{n}{n^2} V(y_i)$$

$$V\left(\hat{\bar{Y}}\right) = \frac{n}{n^2} V(y_i)$$

$$V\left(\hat{\bar{Y}}\right) = \frac{\sigma^2}{n}$$

علما أن:

$\sigma^2 = \frac{1}{n}\sum_1^N (y_i - \bar{Y})^2$ تمثل تباين الظاهرة المدروسة في المجتمع.

<u>الحالة الثانية: السحب بدون احلال (بدون إعادة)</u>

$$V\left(\hat{\bar{Y}}\right) = V(\bar{y}) = V\left(\frac{1}{n}\sum_{1}^{n} y_i\right)$$

$$V\left(\hat{\hat{Y}}\right) = (1 - f)\,\frac{\sigma^2}{n}$$

علما أن:

$f = \frac{n}{N}$ تمثل كسر المعاينة

ملاحظة أساسية: إذا كان حجم المجتمع كبيرا بالنظر إلى حجم العينة، تتساوى صيغة تباين المتوسط الحسابي في المجتمع في حالة السحب مع الإعادة مع السحب بدون إعادة ، أي:

$$V\left(\hat{\hat{Y}}\right) = \frac{\sigma^2}{n}$$

وعموما، سواء في حالة السحب مع الاحلال أو في حالة السحب بدون إحلال، هناك حالتان:

الحالة التي يكون فيها تباين المجتمع σ^2 معروفا. وهي حالة نادرة الحصول في الواقع. في هذه الحالة نعوض تباين المجتمع المعروف بقيمته ونقوم بحساب تباين المتوسط الحسابي للمجتمع.

الحالة التي يكون فيها تباين المجتمع غير معروف. وهي الحالة الغالبة، وفيها نقوم بتقدير تباين المجتمع إما من خلال عينة تجريبية تسبق البحث أو من خلال معطيات البحث نفسها. إلا أن تقدير التباين من خلال العينة $\left(\widehat{\boldsymbol{\sigma^2}}\right)$ يعتبر متحيزا قليلا للتباين في المجتمع σ^2.

حيث:

$$E\left(\widehat{\sigma^2}\right) \neq \sigma^2$$

$$E\left(\widehat{\sigma^2}\right) = S^2 = \frac{1}{N-1}\sum_{1}^{N}(y_i - \bar{Y})^2$$

$$S^2 = \frac{N}{N-1}\left[\!\left[\frac{1}{N}\sum_{1}^{N}(y_i - \bar{Y})^2\right]\!\right]$$

$$S^2 = \frac{N}{N-1}\sigma^2$$

$$\sigma^2 = \frac{N-1}{N}S^2$$

اذا قمنا بتعويض σ^2 ب $\frac{N-1}{N} S^2$ في صيغة تباين المتوسط الحسابي، نحصل على:

$$V\left(\hat{\bar{Y}}\right) = \frac{N-n}{N-1}\frac{\sigma^2}{n}$$

$$V\left(\hat{\bar{Y}}\right) = \frac{N-n}{N-1}\frac{N-1}{N}\frac{S^2}{n}$$

$$V\left(\hat{\bar{Y}}\right) = \frac{N-n}{N}\frac{S^2}{n}$$

وعليه، فإن المقدر غير المتحيز لتباين المتوسط الحسابي للمجتمع هو:

$$V\left(\hat{\bar{Y}}\right) = (1-f)\frac{S^2}{n}$$

إذا كان حجم المجتمع كبيرا بالنظر إلى حجم العينة، فان $f \rightarrow 0$ وبالتالي:

$$V\left(\hat{\bar{Y}}\right) = \frac{S^2}{n}$$

♣ **تقدير التباين**

للحصول على تقدير لتباين المتوسط الحسابي للمجتمع، نعوض المعلمة s^2 بتقديرها المحصل عليه من العينة S^2 وبالتالي، نحصل على:

$$\widehat{V\left(\hat{\bar{Y}}\right)} = \frac{s^2}{n}$$

بحيث:

$$s^2 = \frac{1}{n-1}\sum_{1}^{n}(y_i - \bar{y})^2$$

♣ التقدير بمجال الثقة

بالاضافة الى التقدير النقطي للمعلمات الاحصائية والتي بواسطته يتم تحديد مستويات نقطية (تقديرات) لتقدير المعلمة المعنية، يتوفر مصمم العينة على أدوات منهجية تمكنه من تحديد مجال ثقة للمعلمة.

بالنسبة للمتوسط الحسابي للمجتمع، فان مجال الثقة بمستوى (1-α) هو:

$$\bar{y} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} < \bar{Y} < \bar{y} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$$

المجموع

نعتبر المعلمة Y مجموع المتغير y المحسوب على مستوى كافة وحدات المجتمع.

$$Y = \sum_{1}^{N} y_i$$

$$Y = N \left[\!\left[\frac{1}{N} \sum_{1}^{N} y_i \right]\!\right]$$

$$Y = N\bar{Y}$$

يتم تقدير المعلمة Y التي تمثل مجموع الظاهرة في المجتمع بالمعلمة:

$$\hat{Y} = N\hat{\bar{Y}}$$

$$\hat{Y} = N\bar{y}$$

$$\hat{Y} = \sum_{1}^{n} \frac{N}{n} y_i$$

نلاحظ في هذه الصيغة الوزن $\frac{\ }{n}$ والذي يطبق على جميع وحدات العينة (معاملات الترجيح).

باستعمال الصيغ والمعادلات المعتمدة في الفقرة السابقة والخاصة بالمتوسط الحسابي للمجتمع، نصل بسهولة إلى الاستنتاجات التالية:

- $N_{\bar{y}}$ تقدر بدون تحييز y مجموع الظاهرة في المجتمع.

- في حالة السحب مع الإعادة: $V(\hat{Y}) = N^2 \frac{S^2}{n}$ و $\widehat{V(Y)} = N^2 \frac{s^2}{n}$

- في حالة السحب بدون إعادة: $V(\hat{Y}) = N^2(1-f)\frac{S^2}{n}$ و $\widehat{V(Y)} = N^2(1-f)\frac{s^2}{n}$

النسبة

يمكن اعتبار المعلمة الاحصائية المرتبطة بالنسبة المئوية كحالة خاصة للمتوسط الحسابي. فهي تقيس، بالنسبة لمجتمع معين، النسبة المئوية للأفراد الذين تتوفر فيهم خاصية معينة. (مثال نسبة الأشخاص الذين يتوفرون على سيارة).

نلاحظ بكل سهولة أن النسبة هي المتوسط الحسابي لمتغير مؤشر y يأخذ القيم:

$y_i = 1$ إذا توفرت الخاصية في الفرد i

$y_i = 0$ في حالة العكس، أي اذا لم تتوفر الخاصية المدروسة في الفرد i

تجدر الاشارة الى أن المتغير y يخضع لتوزيع برنولي بمعلمة p.

نضع:

$A = \sum_1^N y_i$ هو مجموع وحدات المجتمع التي تتوفر فيها الخاصية المدروسة.

$a = \sum_1^n y_i$ هو مجموع وحدات العينة التي تتوفر فيها الخاصية المدروسة.

نحصل على النسبة وهي المعلمة المراد تقديرها على مستوى المجتمع المدروس، وذلك بحساب المتوسط الحسابي للمتغير y على مستوى جميع وحدات المجتمع.

$$P = \frac{1}{N}\sum_1^N y_i = \frac{A}{N} = \bar{Y}$$

نحصل على النسبة p وهي المعلمة المراد تقديرها على مستوى الحسابي للمتغير y على مستوى جميع وحدات العينة.

$$p = \frac{1}{n}\sum_1^n y_i = \frac{a}{n} = \bar{y}$$

مع هذه الصيغ، نجد المعادلات المفصلة في الفقرة السابقة والمتعلقة بالمتوسط الحسابي للمجتمع $\bar{Y}$.

♣ التحيز

تعتبر p النسبة المحسوبة في العينة مقدرا غير متحيز للنسبة التي نريد معرفتها (تقديرها) على مستوى المجتمع المدروس.

$$E(\widehat{P}) = E(p) = E(\bar{y}) = \bar{Y} = P$$

♣ التباين

السحب بإحلال:

$$V(\hat{P}) = V(p) = V(\bar{y}) = \frac{S^2}{n}$$

السحب بدون إحلال

$$V(\hat{P}) = V(p) = V(\bar{y}) = (1 - f)\ \frac{S^2}{n}$$

بحيث: $\sigma^2 = P(1 - P)$

♣ تقدير التباين

للحصول على تقدير لتباين النسبة P ، نعوض المعلمة S^2 بتقديرها s^2 المحسوبة في العينة. ونحصل إذا على:

$$\widehat{V(\hat{P})} = \frac{s^2}{n}$$

بحيث: $s^2 = p(1-p)$

مع افتراض أن حجم المجتمع كبيرا بالنظر إلى حجم العينة وبالتالي ($f \rightarrow 0$).

المعدل

نريد تقدير معدل R (مثال إنتاجية زراعة الحبوب في أرض فلاحية).

المعلمة R هي مقسوم متغيرين كميين. في حالة المثال أعلاه، R هو مقسوم إنتاج الحبوب (متغير Y) على المساحة المزروعة بالحبوب (متغير X).

$$Y = \sum_1^N y_i \quad et \quad X = \sum_1^N x_i$$

نضع :

$$R = \frac{Y}{X}$$

$$R = \frac{\sum_1^N y_i}{\sum_1^N x_i}$$

$$R = \frac{\bar{Y}}{\bar{X}}$$

التحيز

نضع المقدر $\hat{R}$:

$$\hat{R} = \frac{\bar{y}}{\bar{x}}$$

يعتبر هذا المقدر متحيزا، ذلك أن:

$$E(\hat{R}) \neq R$$

B هو تحيز المقدر.

$$B = E(\hat{R}) - R$$

الخطأ التربيعي المتوسط هو:

$$EQM = V(\hat{R}) + B^2$$

$\hat{R}$ هو مقدر تقريبا غير متحير إذا كان

$$\bar{X} = \bar{x}$$

♣ **التباين**

$$V\left(\hat{R}\right) = (1 - f)\frac{1}{n}\frac{1}{\bar{X}^2}\left[\!\left[S_y{}^2 + R^2 S_x{}^2 - 2RS_{xy}\right]\!\right]$$

♣ **تقدير التباين**

$$\widehat{V\left(\hat{R}\right)} = (1 - f)\frac{1}{n}\frac{1}{\bar{x}^2}\left[\!\left[s_y{}^2 + \hat{R}^2 s_x{}^2 - 2\hat{R}s_{xy}\right]\!\right]$$

5. حجم العينة

العوامل المؤثرة في تحديد حجم العينة هي:

- مستوى درجة الدقة والثقة بالنتائج. حيث، كلما زاد حجم العينة، كلما زادت دقة نتائج البحث.
- درجة التعميم التي ينشدها الباحث من نتائج البحث.
- مدى التجانس أو التباين في خصائص مجتمع الدراسة الأصلي. فكلما كانت خصائص المجتمع الأصلي متجانسة كلما كان حجم العينة المطلوبة صغيرا. وعلى العكس، كلما كان تشتت الظواهر المدروسة كبيرا وبالتالي كان تجانس المجتمع ضعيفا، كلما تطلب الأمر عينات من حجم أكبر
- تصميم المعاينة المعتمد.

♣ المتغير الكمي

نتوخى معرفة مستوى ظاهرة معينة على مستوى مجتمع معين وذلك عن طريق عينة عشوائية. يطرح هنا السؤال المتعلق بما هو حجم العينة الذي يمكن من تقدير المعلمة الاحصائية المعنية مع مراعاة هامش الخطأ الذي يجب أن لا يتعدى مستوى محدد.

ستكون نقطة الانطلاقة في تحديد حجم العينة اذا من دراسة مكونات مجال الثقة للمقدر.

لدينا:

$$\bar{y} - z_{(1-\frac{\alpha}{2})}\sqrt{\frac{N-n}{N}\frac{S^2}{n}} < \bar{Y} < \bar{y} + z_{(1-\frac{\alpha}{2})}\sqrt{\frac{N-n}{N}\frac{S^2}{n}}$$

نعتبر r هو هامش الخطأ.

$$r = z_{(1-\frac{\alpha}{2})}\widehat{\sigma}_{\hat{\bar{Y}}}$$

$$r = z_{(1-\frac{\alpha}{2})}\sqrt{\frac{N-n}{N}\frac{S^2}{n}}$$

$$r^2 = z^2\frac{N-n}{N}\frac{S^2}{n}$$

$$r^2 = z^2(1 - \frac{n}{N})\frac{S^2}{n}$$

$$r^2 = z^2 \frac{S^2}{n} - z^2 \frac{S^2}{N}$$

$$n = \frac{z^2 S^2}{r^2} - \frac{z^2 S^2 n}{r^2 N}$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

علما أن:

$$n_0 = \frac{z^2 S^2}{r^2}$$

اذا كان حجم المجتمع كبيرا بالنظر إلى حجم العينة وبالتالي $f \rightarrow 0$، فان:

$$n = n_0 = \frac{z^2 S^2}{r^2}$$

- ♣ **المتغير الكيفي**

$$p - z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p(1-p)}{n}} < P < p + z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p(1-p)}{n}}$$

$$n = \frac{z^2 p(1-p)}{r^2}$$