الفصل الثاني عشر

العيـنـــات

"Samples"

الفصل الثاني عشر

العينـــات

"Samples"

يعد إستخدام العينات من الأمور العادية في مجال البحوث والدراسات العلمية سواء الاجتماعية أو الطبيعية. والعينة هي عبارة عن مجموعة جزئية من الأفراد أو المشاهدات أو الظواهر التي تشكل مجتمع الدراسة الأصلي. فبدلا من إجراء البحث أو الدراسة على كامل مفردات المجتمع يتم إختيار جزء من تلك المفردات بطريقة معينة، سنأتي على ذكرها لاحقا، وعن طريقة دراسة ذلك الجزء يمكن تعميم النتائج التي تم الحصول عليها من مجتمع الدراسة الأصلي.

وكما سنرى لاحقا فأن اختيار العينة بشكل دقيق ومضبوط سوف يعطي نتائج مشابهة إلى حد كبير لعملية دراسة كامل المجتمع. وفي هذه الحالة فأن اختيار عينة لإجراء الدراسة عليها قد يكون مفضلا على دراسة كامل مجتمع الدراسة الأصلي نظرا لما في ذلك من توفير للوقت والمال والجهد المبذول.

أسباب اللجوء إلى إستخدام العينات:

إن إجراء البحث على كامل مجتمع الدراسة الأصلي يكون مفضلا في معظم الحالات على إختيار عينة وإجراء الدراسة عليها نظرا لما يعطيه دراسة كامل المجتمع من نتائج أقرب للواقع وأكثر قابلية للتعميم. إلا أن هناك أسبابا عدة قد تدفع الباحث إلى الاعتماد على العينة بدلا من إجراء دراسته على كامل مجتمع الدراسة الأصلي، ومن ضمن تلك الأسباب ما يلي:

أ- **كلاسيكياً:** توفيرا للجهد والوقت والمال، ففي حالة كون مجتمع الدراسة الأصلي كبيرا ومتباعدا جغرافيا فإن ذلك يتطلب لكلفة عالية وجهدا كبيرا ووقتا طويلا من الباحث.

ب- **حديثاً:** بالإضافة إلى ما تقدم فأن الباحث يلجأ إلى إستخدام العينة لأسباب أخرى منها:

1- توسيع مجال البحث لأن الباحث عندما يتعامل مع عدد محدود، فبالتأكيد سيتناول أمور أخرى في مجال البحث العلمي يرتأي أهميتها.

2- قياس دقة النتائج، حيث للعينات هناك معايير إحصائية على ضوءها يمكن إعطاء الثقة بالنتائج التي توصل إليها الباحث، ولقياس دقة النتائج هناك مقياس معروف يدعى بالخطأ القياسي (Standard Error) وكلما كان هذا الخطأ صغيرا، كلما اعتدت نتائج تلك العينة والعكس بالعكس.

3- هناك بحوث تشترط أن يلجأ الباحث لإستخدام العينة، فهناك بعض أنواع الأبحاث التي تكون فيها عناصر مجتمع الدراسة الأصلي متجانسة بشكل كبير وبالتالي فأن النتائج نفسها يتم الحصول عليها سواء أجريت الدراسة على كامل المجتمع أو على أجزاء منه، ومن هذه الأمثلة الواضحة في هذا المجال فحص الدم. كما وأن هناك عدم الإمكانية لإجراء الدراسة على كامل عناصر المجتمع، مثل بعض أنواع الأطعمة المنتجة كالألبان والمشروبات كالعصير وبعض السلع الكهربائية كالتلفاز تقوم معظم المصانع بإختيار عينات من

الإنتاج بشكل دوري ويتم فحص تلك العينات للتأكد من سلامتها ومطابقتها للمواصفات المحددة. وفي مثل هذه الحالات قد يكون من غير المجدي أن يتم إجراء الدراسة أو الفحص على كامل المنتجات نظرا لأن الوحدات التي تخضع للفحص غير صالحة وبالتالي لا يمكن بيعها لاحقا.

ومن جانب آخر فهناك العديد من الدراسات التي لا يمكن فيها حصر كامل عناصر مجتمع الدراسة الأصلي، ومن الأمثلة على ذلك دراسة المدمنين على المخدرات فقد لا تتوفر معلومات عن كامل المدمنين في الدولة أو قد تكون المعلومات سرية ولا يمكن الإباحة عن هذه الفئة.

أساليب إختيار العينات:

وهناك أساليب عديدة لإختيار العينات أكثرها إنتشارا في العلوم الإدارية، ومن هذه الأساليب:

أ – العينة العشوائية البسيطة:Simple Random Sample

وهي عينة تختار من المجتمع الإحصائي بحيث إن كل مفردة من مفردات المجتمع لها نفس الفرصة في الظهور في العينة. ولتحديد حجم العينة المختارة يواجه المحددات التالية:

1- مقدار الخطأ المسموح به ويطلق عليه (d) ويعني مقدار الدقة المطلوبة.

2- إحتمال الوقوع في هكذا خطأ، بمعنى ما هو الاحتمال أن أقع به، ويطلق عليه هذا الاحتمال (w) وكذلك المساحة تحت التوزيع

الطبيعي والتي تكافئ هذا الاحتمال ولها قيمة تقرأ من الجداول الإحصائية وقيمتها (z).

3- مقدار تجانس المجتمع، حيث يمكن قياس هذا المقدار من التجانس عادة بمقياس أو معيار الانحراف المعياري والذي يرمز له (□).

4- وأخيرا هي عدد مفردات المجتمع الإحصائي والذي يرمز له بالرمز (N) وعليه فأن أحدث قانون يمكن أن يمكن أن يحسب مقدار العينة البدائية هو:

$$N0 = (Zd/d)^2$$

حيث إن (n0) تمثل عدد مفردات العينة بادئ ذي بدأ و (Z) تقرأ من الجداول الإحصائية، فعندها يتم مقابلة إحتمال الوقوع في الخطأ.

والرمز (□) هو الانحراف المعياري والذي يقيس مقدار التشتت أما الرمز (d) فهو الخطأ المسموح به.

إما عندما يريد الباحث أن يحصل على حجم العينة النهائية، فأنه يحتاج إلى استخدام الصيغة الإحصائية التالية:

$$n = \frac{n0}{1 + \frac{n□}{N}}$$

حيث تشير الرموز التالية:

n0 = حجم العينة البدائي

n = حجم العينة المطلوب

N = حجم المجتمع

مثال / المطلوب إيجاد حجم عينة لتقدير متوسط مجتمع عدد مفرداته (5000) بحيث أن الخطأ المسموح به هو 5%، فإذا علمت أن هناك تقديرا مسبقا لتباين المجتمع (□) هو 0.2.

الحل:

$n0 = (Zd/d)^2$ $\quad Z = 1.96 \rightarrow$ من الجدول الإحصائي تحت 0.05

القيمة الاستطلا

$= (2.0.2/0.05)^2 = (8)^2 = 6$

إما حجم العينة اللازم فهو:

$$N = \frac{n0}{1 + \frac{n□}{N}}$$

$$= \frac{64}{1 + \frac{64}{5000}} = \frac{64}{1 + 0.0128} = \frac{64}{1.0128} = 63$$

*ملاحظة مهمة: إن تقدير أو تحديد حجم العينة يتوقف أيضا على الهدف المراد من الدراسة، فإذا كان المطلوب تقدير متوسط المجتمع، والذي يسمى (□ أو ($\bar{Y}$) بواسطة مفردات العينة، أي أن:

$$\bar{y} = □\,^\wedge = \bar{Y}^\wedge = \frac{y1 + y2 + \cdots + □n}{n}$$ (متوسط العينة)

أما الصيغة الإحصائية لتقدير العدد الكلي (المجموع الكلي) فهي كالآتي:

$Y = y1+y2+\ldots+yn$

$Y = N\bar{Y}$

وعليه فأن تقدير المجموع الكلي هو: $Y^\wedge = N\bar{y}$

وبناءا على ذلك فأن الصيغة الإحصائية لحجم العينة البدائي المطلوب هي:

$$n0 = (Zn\sigma/d)^2$$

أما الصيغة الإحصائية للحجم النهائي فهي:

$$n = \frac{n0}{1 + \frac{n□}{N}}$$

كما أن بالإمكان أن يكون الهدف (هدف الباحث) لتقدير النسبة "Proportion" أي نسبة عدد الذين يحملون صفة معينة، فأن الصيغة الإحصائية لتقدير تلك النسبة سوف يكون على النحو التالي:

عدد الأشخاص ممن يحملون صفة معينة=A $\quad P = \frac{A}{□}$ نسبة المجتمع

N=تمثل مفردات المجتمع الكلي

وعليه فأن تقدير نسبة المجتمع والتي تحسب عادة من العينة هي:

$$Y = P\^{}=a/n$$

حيث أن (a) تمثل أو تشير إلى عدد الأشخاص ممن يحملون تلك الصفة من الذين وقعوا ضمن العينة، ويشير الرمز (n) إلى حجم العينة التي أختيرت من المجتمع.

وعليه فأن الصيغة الإحصائية لتقدير نسبة حجم العينة كالآتي:

$$n_0 = \frac{Z^2 \psi \square}{\sigma 2}$$

حيث أن: $\psi = \frac{a}{n}$ ، ع ψ -1 =،

أما حجم العينة النهائي فهو:

$$n = \frac{n0}{1 + \frac{n\square}{N}}$$

ب – أسلوب العينة الطبقية العشوائية البسيطة:

هذا النوع من العينات أو هذه الطريقة في إختيار العينة منتشرة بكثرة في العلوم الإدارية حيث أن المجتمعات المطلوب دراستها قد تكون غير متجانسة ولذلك يقسم المجتمع الإحصائي إلى عدة طبقات "Strata" وفقا لخصائص معينة يحددها الباحث مقدما، وقد تكون كل طبقة كمجتمع قائم بذاته، وتكون المفردات داخل الطبقة "Stratum" متجانسة ويقوم الباحث عادة بإختيار عينة عشوائية بسيطة من كل طبقة وبالتالي فأن العينة الكلية والتي هي مجموع العينات الفرعية تدعى بالعينة الطبقية.

ولتوضيح ذلك، فلو أمكن تقسيم المجتمع إلى "L" من الطبقات بحيث أن كل طبقة تحتوي على عدد من المفردات N1،N2،N3،...... ،Nh،......، NL

بحيث أن عدد مفردات المجتمع الكلي هي:

N = N1 + N2 +……..+Nh+…….NL

ومن هذه الطبقات نختار عينات عشوائية بحيث أن العينة النهائية:

n = n1 + n2+…..nL

n= العينة الطبقية العشوائية البسيطة

n1، n2، …. nL = عينات فرعية

إما كيفية إختيار العينات العشوائية البسيطة ومن كل طبقة على إنفراد فيتم على النحو التالي:

أ – **التوزيع التناسبي:** إن الفكرة الأساسية هي أن نختار عينة عشوائية بسيطة من أية طبقة بنسبة وجود تلك الطبقة في المجتمع الإحصائي (موضوع البحث)

وهذا يعتمد على وزن الطبقة "Stratum Weight"

$$\frac{nh}{\square h} = \frac{n}{\square}$$

$$nh = n.\ \frac{Nh}{\square}$$

nh = عدد مفردات العينة الفرعية لأية طبقة

n = عدد مفردات العينة الطبقية النهائية

Nh = عدد مفردات الطبقة

N = عدد مفردات المجتمع

ب – التوزيع الأمثل "Optimum Allocation"

في هذا النوع من طرق اختيار العينات العشوائية الطبقية يؤخذ بنظر الاعتبار مدى التجانس لكل طبقة على حده، وتؤخذ أيضا الكلفة "Cost" بنظر الإعتبار، وعليه فأن حجم العينة المختارة بهذه الطريقة دقيقة جدا بحيث يجعل تباين متوسط العينة أصغر ما يمكن وتصبح الصيغة الإحصائية لحجم العينة الفرعي على النحو التالي:

$$nh = n.\ \frac{\frac{Nh\sigma h}{\sqrt{Ch}}}{\sum_{h=1}^{L} \frac{\square h\sigma h}{\sqrt{Ch}}}$$

حيث أن "Nh" تمثل عدد مفردات الطبقة، " $\sigma\square$" تمثل الانحراف المعياري للطبقة h ، "Ch" تمثل كلفة المعاينة للطبقة h.

حيث أن:

$$nh = n.\frac{\frac{Nh\sigma h}{\sqrt{Ch}}}{\frac{N1\sigma 1}{\sqrt{C1}}+\frac{N2\sigma 2}{\sqrt{C_2}}+\cdots+\frac{NL\sigma L}{\sqrt{CL}}}$$

كما يمكن كتابة الصيغة نفسها بالشكل التالي:

$$nh = n.\frac{\frac{wh\sigma h}{\sqrt{Ch}}}{\frac{w1\sigma 1}{\sqrt{C1}}+\frac{w2\sigma 2}{\sqrt{C_2}}+\cdots+\frac{wL\sigma L}{\sqrt{CL}}}$$

حيث أن: $wh = Nh/N$

$Wh =$ تمثل وزن الطبقة

إما الكلفة الكلية فتحسب كالآتي:

$$C = C0+C1n1+C2n2+\ldots+CLnL$$

$$C = C0 + \sum_{h=1}^{L} C_h nh$$

حيث تمثل الرموز التالية:

$C0 =$ الكلفة البدائية الأولية

$C =$ الكلفة الكلية ، $Ch =$ كلفة المعاينة

مثـال: درس باحـث فـي الإدارة العامـة مجتمعـا وحصـل علـى الخلاصـة الإحصائية التالية:

Strata	Nh	σh	Ch
I	500	20	4
II	800	25	4
III	1200	30	9
Total	2500		

حيث تمثل " σ" التجانس في كل طبقـة (الانحـراف المعيـاري) و "Ch" تمثل كلفة المعاينة.

فما هي أجزاء المعاينة الفرعية من كل طبقـة، إذا علمـت أن الباحـث كـان محددا بعينة حجمها "100" ؟

الحل:

بالطريقـة التقليديـة والتـي لا تأخذ بنظـر الإعتبـار التجـانس داخـل الطبقـة الواحدة، ولا الكلفة، أي تستخدم التوزيع التناسبي

الطبقة الأولى $n1 = 100 \times \frac{500}{2500} = 20$

$n2 = 100 \times \frac{800}{2500} = 32$

$n3 = 100 \times \frac{1200}{2500} = 48$

في حين أن الطريقة المعاصرة تأخذ بنظر الاعتبار التجانس والكلفة.

$$n1 = \frac{500 \times \frac{200}{\sqrt{4}}}{\frac{500 \times 20}{2} + \frac{800 \times 25}{2} + \frac{1200 \times 30}{3}} \times 100$$

$$= \frac{5000}{5000 + 10000 + 12000} \times 100 = \frac{5000}{27000} \times 100 = 19$$

$$n2 = \frac{800 \times \frac{25}{2}}{27000} \times 100 = \frac{10000}{27000} \times 100 = 38$$

$$n3 = \frac{1200 \times \frac{30}{3}}{27000} \times 100 = \frac{12000}{27000} \times 100 = 43$$

توزيع نيمان "Neyman Allocation"

ومن الجدير بالذكر أن هناك نوع آخر من الأخذ في الاختبار للعينة يدعى توزيع نيمان وفي هذه الحالة يثبت نيمان الكلفة وتصبح حجوم العينات الفرعية وفق الصيغة التالية:

$$Nh = n.\frac{Nh\sigma h}{\sum_{h=1}^{L} Nh\sigma h}$$

$$Nh = n.\frac{wh\sigma h}{\sum_{h=1} wh\sigma h}$$

ومن المؤشرات الإحصائية التي تحسب بعد اختيار العينة العشوائية الطبقية هي:

أ – الوسط الحسابي للعينة الطبقية ويرمز له بالرمز ($\bar{y}$ statified) ($\bar{y}$st)

حيث يمكن استخراج ذلك وفق الصيغة الإحصائية التالية:

$$\bar{y}st = \frac{\sum_{h=1} nh\bar{y}}{n}$$

وبعبارة أخرى فأن الوسط الحسابي للعينة الطبقية هو:

$$\bar{y}st = \frac{n1\bar{y}1 + n2y2 + \cdots + nLyL}{n}$$

ولمعرفة دقة هذا المقياس نحسب التباين له ومن ثم الخطأ المعياري له حيث:

$$Var(\bar{y}t) = \sum_{h=1}^{L} \frac{wh^2 \sigma h^2}{h(1 - fh)}$$

حيث يشير الرمز (h^2) إلى مربع تباين العينة الطبقية و(nh) يشير إلى حجم العينة المختارة من تلك الطبقة. و(fh) كسر المعاينة ويمكن استخراجه كالآتي:

$$fh = nh/Nh$$

والرمز (wh^2) يشير إلى مربع وزن الطبقة.

كما ويجدر بالذكر بأنه كسر المعاينة قد يهمل أحيانا وخصوصا إذا كانت قيمته الحسابية صغيرة.

ب – كما يمكن تقدير المجموع الكلي وفق الصيغة الإحصائية الآتية:

$$S.e(\bar{y}st.) = \sqrt{(Var(\bar{y}st.))}$$

حيث يشير الرمز "S.e" إلى الخطأ المعياري للتقدير (Standard Error)

إختبار الفرضيات في المجال الكمي الإداري"Testing Hypothesis"

على الباحث بعد أن يتم جمع البيانات وتوزيعها في جداول واحتسابها من تلك الجداول نسبة مئوية أو متوسطات أو أية علاقات وصفية أخرى تنتهي عند هذه المرحلة مهمة الإحصاء الوصفي "Descriptive Statistics" وتبدأ مرحلة الإحصاء الاستدلالي أو الاستدلال الإحصائي "Statistics Inference" وتتلخص مهمة الاستدلال الإحصائي في موضوعين أساسيين هما:

أ – إختبار الفرضيات Testing of Hypothesis

ب – التقدير Estimations، والتقدير نوعان:

الأول: التقدير النقطي "point Estimate"

الثاني: التقدير الفئوي "Interval Estimate"، وهذا يشير إلى الحد الأدنى للفئة والحد الأعلى لها.

لذلك فأن المهمة الأولى ينبغي على الباحث إجراءها هي إختبار الفرضيات.

إذن ما المقصود بالفرضية الإحصائية (Statistical Hypothesis) ؟

يقصد بالفرضية الإحصائية، هي إدعاء أو تصريح حول معلمة أو معالم المجتمع وتوضع عادة الفرضية الأساس أو ما تسمى بفرضية العدم "Null Hypothesis" ويرمز لها بالرمز "H0". وتوضع فرضية أخرى بديلة تدعى "Alternative Hypothesis" ويرمز لها "H1" وعلى ضوء معطيات العينة يمكن حساب ما يسمى بالمختبر الإحصائي "Test Statistic" وبواسطة هذه الغحصاءة يتم رفض H0 أو قبولها.

ومما يجدر ذكره بأن فرضية الأساس "H0"بشكل تكون:

أ - حيادية.

ب – حسن النية.

جـ - موضوعية.

د – يأمل الباحث رفضها.