تقنيات الاستقصاء

المعاينة – تحديد حجم العينة

تقدير حجم العينة

عينة صغيرة جدًا

لن تؤدي
الغرض

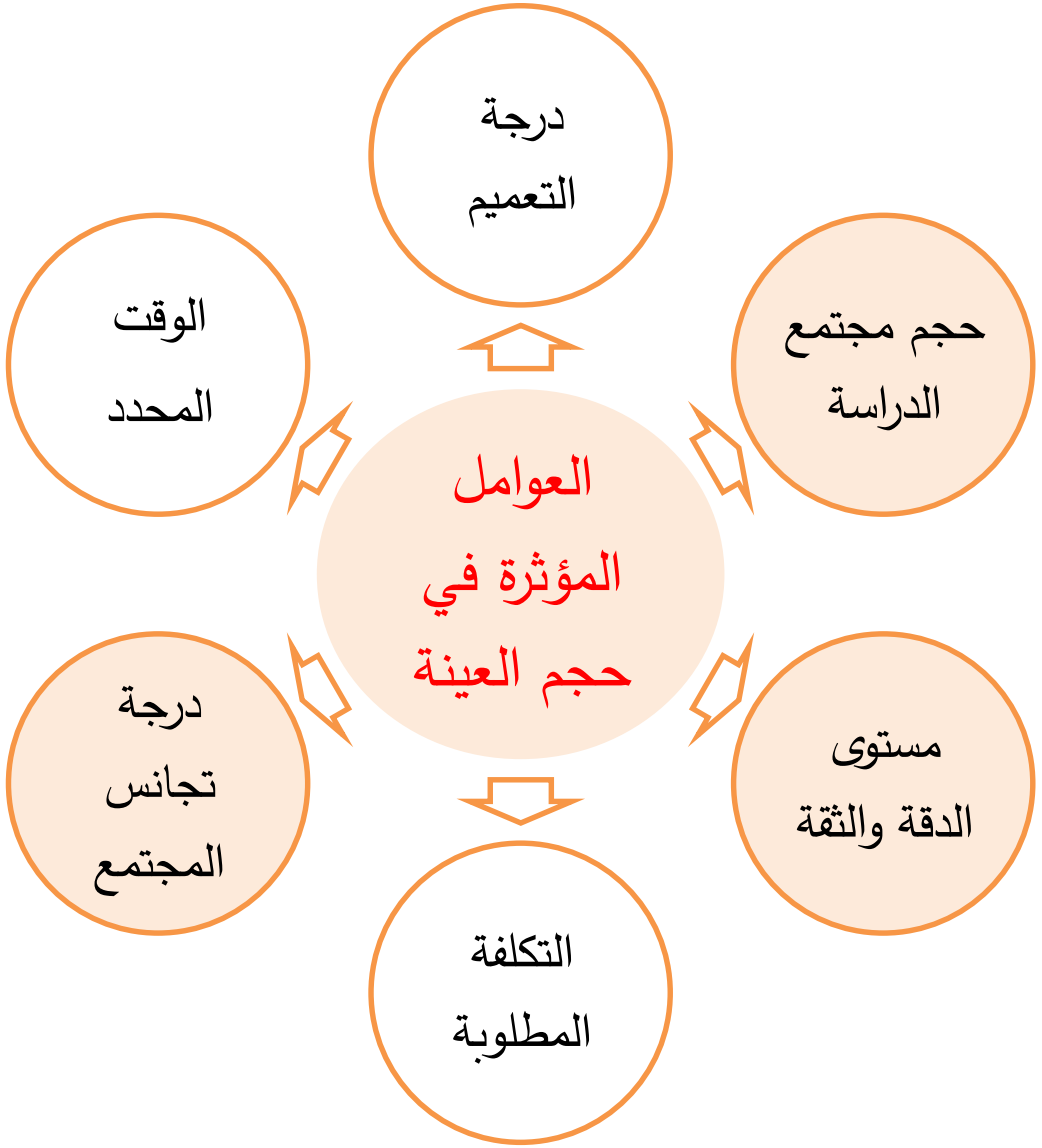
عينة كبيرة جدًا

تتطلب الكثير
من الموارد

لا العينة الصغيرة و لا العينة الكبيرة هي الجيدة والمطلوبة

هناك عوامل مؤثرة على حجم العينة

حجم عينة الدراسة يتأثر بعدة عوامل أهمها:



حجم المجتمع

كلما كان المجتمع كبيرا، احتاج الباحث إلى عينة أكبر لتمثيله

هناك علاقة طرديه بين حجم العينة وحجم مجتمع الدراسة.

مستوى الدقة المطلوب Accuracy

يعرف بأنه درجة قرب القيم المقاسة أو المحسوبة من القيمة الحقيقية أو الفعلية للمتغير الذي نقيسه

يرتبط بمفهوم آخر هو الثبات (Reliability) الذي يقصد به مدى اتساق النتائج عند إعادة القياس بنفس الأداة.

يقدر بنسبة مئوية أو بقيمة حقيقية مرتبطة بوحدة القياس (على سبيل المثال، ± 5 أو 3 ±)،

إذا كان الباحث يرى أن 70% من الطلبة في العينة غير راضين عن النقل الجامعي مع معدل دقة ± 5

هذا يعني أن ما بين 65 % و 75 % من الطلبة في المجتمع غير راضين عن النقل الجامعي.

التفريق بين المفهومين

المفهوم	التعريف	الهدف	المثال
الدقة	مدى صحة النتائج وقربها من القيمة الحقيقية	قياس الصواب	إذا كان متوسط رضا الطلبة في العينة قريب من المتوسط الحقيقي لجميع الطلبة
الثبات	مدى اتساق النتائج عند إعادة القياس بنفس الأداة	قياس الاستقرار	إذا كررنا الاستبيان مرتين وظهرت نتائج متقاربة

في التحليل الإحصائي، مستوى الدقة يقاس من خلال: قرب المتوسط الحسابي للعينة ($\bar{X}$) من المتوسط الحقيقي للمجتمع (μ)

كلما كان الفرق بينهما صغيرا، كانت الدقة أعلى.

مثال

اخذنا عينة عشوائية بسيطة من 50 عاملا من مجتمع قدره 300 عامل في ورشة صناعية. أظهرت أن متوسط الإنتاج اليومي للعامل هو 50 وحدة ($\bar{X}$=50)، فاذا كان لدينا إحصائية تقول أن المتوسط الحقيقي للإنتاج (μ) في هذا المصنع يقع بين 40 و60 وحدة.

هذا يعني أن تقديرنا للمدى الذي نتوقع أن يوجد فيه المتوسط الحقيقي لإنتاج المجتمع هو

$$\mu = 50 \pm 10$$

كلما ضاق التقدير اقتربنا من الدقة والعكس صحيح

في المثال السابق المتعلق بمعدل الإنتاج في المصنع:

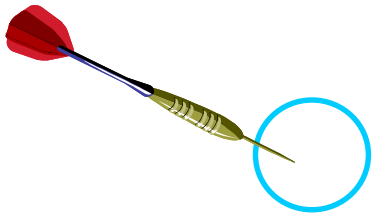
تقديرنا للمتوسط الحقيقي للإنتاج سيكون أكثر دقة حينما نتوقع أنه يقع بين 45 و 55 وحدة

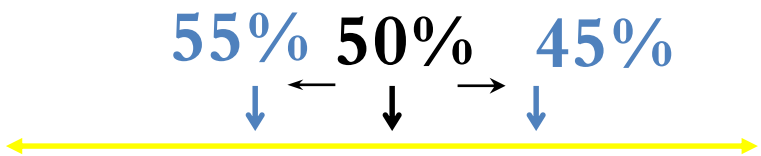
$\mu = 50 \pm 5$

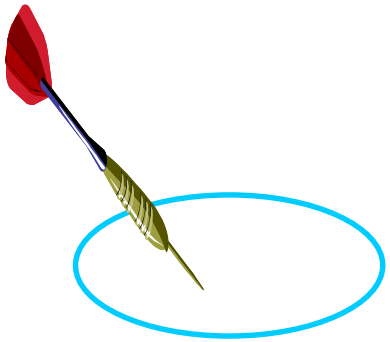
بالمقارنة بتوقعنا الأساسي وهو بين 40 و60 وحدة

$\mu = 50 \pm 10$

ولكننا سنكون أكثر ثقة في تقديرنا الأساسي أكثر من تقديرنا الثاني







تدرجة تجانس المجتمع

القاعدة الذهبية في البحث العلمي أنه كلما كبر حجم العينة نحصل على مستوى أفضل من الدقة والثقة.

ولكن هذا يرتبط بتجانس توزيع هذه الصفات في المجتمع

نحتاج إلى عينات كبيرة في حالة عدم التجانس

نحتاج الى عينات أصغر في حالة التجانس

الانحراف المعياري	حجم العينة
10	96
12	138
14	188
16	246
18	311
20	384

تأثير الانحراف المعياري على حجم العينة في حال ثبات العوامل الأخرى

يقدر حجم العينة بثلاث طرق:

(1) المعادلات الرياضية (الحسابات اليدوية)

(2) الجداول التي تحدد حجم العينة أهمها جدول Uma Sekaran

(3) البرامج الاحصائية

تقدير حجم العينة باستخدام المعادلات الرياضية

هناك عدد من المعادلات الرياضية لتحديد حجم العينة المناسب وهى :

- $n=PQ(Z)^2/d^2$

حيث: n حجم العينة

P نسبة توافر الخاصية في المجتمع وقيمتها تتراوح عادة بين (30 إلى 60) وفى حالة عدم معرفة تلك النسبة يستخدم نسبة (50%)

Q النسبة المكملة $Q = 100 - P$

$Z\alpha = 1.96$ لمستوى ثقة95%

$Z\alpha = 2.58$ لمستوى ثقة 99%

d خطأ المعاينة سواء عند (0.05 أو 0.01)

- عند افتراض نسبة توافر الخاصية في المجتمع (50%)، والدرجة المعيارية (1.96) ، وخطأ المعاينة 0.05 فأن حجم العينة يكون (384) فرد

- $n=4pQ/25$

حيث: n حجم العينة

- P نسبة توفر الخاصية في المجتمع

وفى حالة عدم معرفة تلك النسبة يستخدم نسبة (50%)

Q النسبة المكملة $Q = 100-P$

- عندما تكون $P=30$ فإن $Q=70$ وبذلك يصبح حجم العينة (336) فرد

- وفى حالة $P=50$ فإن $Q=50$ ويصبح حجم العينة (400) فرد

- $n = (Z_{a/2})^2 \, s^2 / d^2$

حيث: s الانحراف المعياري

d خطأ المعاينة

$Z_{a/2}$ الدرجة المعيارية

لنفترض أن:

- مستوى الثقة = 95% $Z\alpha/2=1.96$
- الانحراف المعياري $s = 10$
- الحد الأقصى للخطأ $d=2$

بإدخال القيم في المعادلة سنجد العينة تساوي 97

تقدير حجم العينة باستخدام الجداول:

هناك جداول يستخدمها الباحثين لتحديد حجم العينة أهمها جدول Uma Sekaran الذي يسمح للباحث بمعرفة حجم العينة عند احجام معينة للمجتمع.

N	S	N	S	N	S
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	175	2000	322
55	48	320	181	2200	327
60	52	340	191	2400	331
65	56	360	196	2600	335
70	59	380	205	2800	338
75	63	400	210	3000	341
80	66	420	217	3500	346
85	70	440	226	4000	351
90	73	460	242	4500	354
95	76	480	248	5000	357
100	80	500	260	6000	361
110	86	550	265	7000	364
120	92	600	274	8000	367
130	97	650	278	9000	368
140	103	700	169	10000	370
150	108	750	186	15000	375
160	113	800	201	20000	377
170	118	850	214	30000	379
180	123	900	234	40000	380
190	127	950	254	50000	381
200	132	1000	269	75000	382
210	136	1100	285	1000000	384

الجدول التالي يبين حجم العينة المناسب عند مستويات مختلفة من مجتمع الدراسة الأصلي:

Uma Sekaran, Research Methods for Business ,2012 John Wiley & Sons Ltd

تقدير حجم العينة باستخدام البرامج الاحصائية

- قبل إدخال البيانات في أي برنامج، يحتاج الباحث إلى تحديد بعض العناصر:
  - مستوى الدلالة (α): غالبا 0.05 (أي 5%)
  - نوع الاختبار الإحصائي T-test، ANOVA، Chi-square، الارتباط... إلخ.
  - حجم المجتمع في بعض الحالات.
- البرنامج الأكثر شيوعا في تحديد حجم العينة في البحوث الاكاديمية هو G*Power
- برنامج SPSS لا يحتوي على أداة مدمجة لحساب حجم العينة مثل G*Power

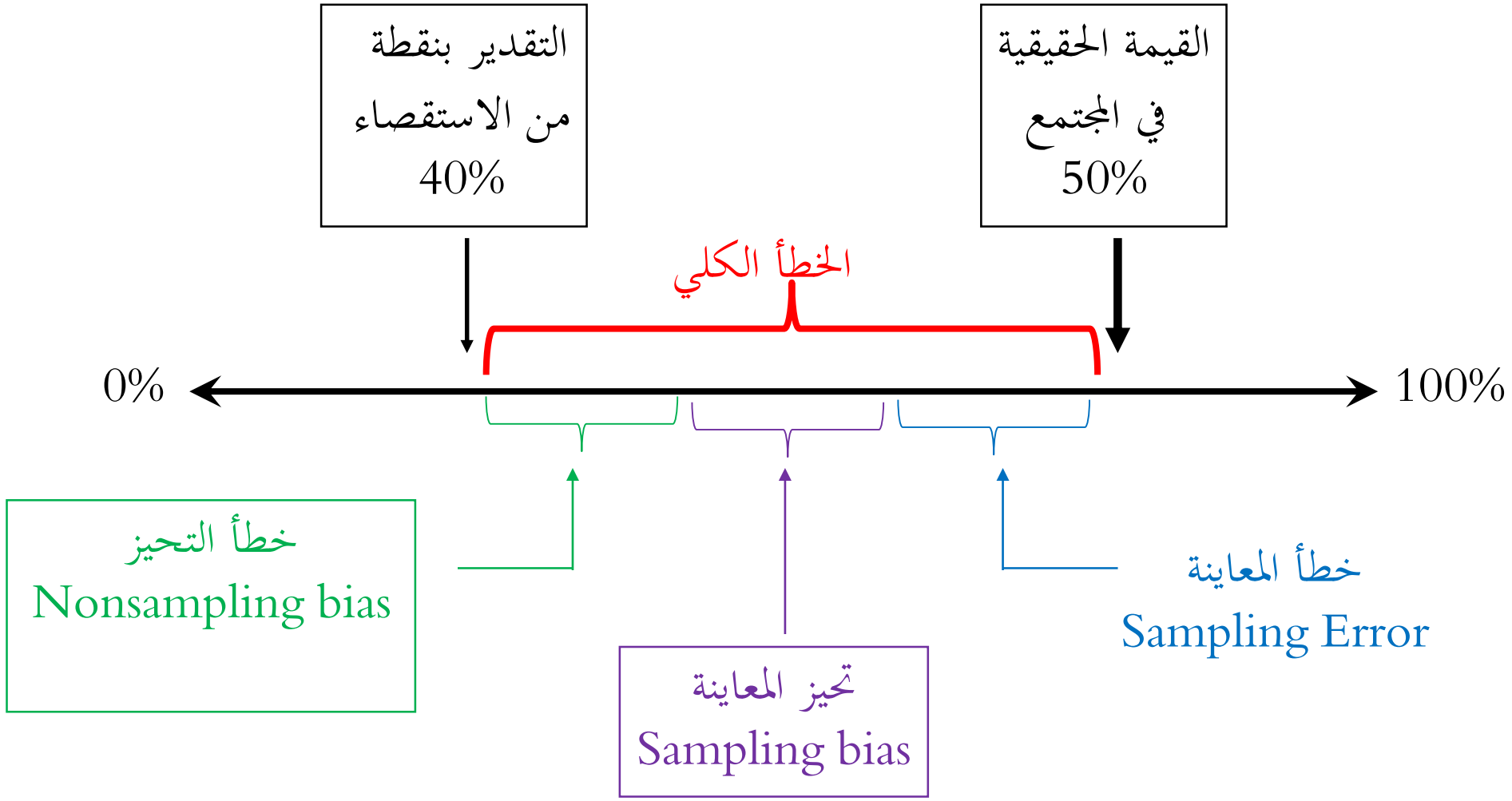
المعاينة: الخطأ والتحيز

القيم التي نحصل عليها من الدراسة الاستقصائية لا تكون مطابقة تماما للقيم الفعلية في المجتمع **(فجوة)**. لماذا؟

هذه الفجوة يمكن أن تؤثر على دقة وموثوقية النتائج المستخلصة

هناك عوامل تساهم في الفجوة بين القيم المستخلصة من الدراسة الاستقصائية والقيم الفعلية في المجتمع.

مكونات الخطأ الكلي



خطأ المعاينة Sampling Error

يسمى كذلك بـ (خ صدفة أو الخطأ العشوائي)

يشير إلى الفرق بين القيمة المقدرة من الع قيمة في المجتمع.

هذا النوع من الخطأ يمكن أن يحدث حتى في حالة اختيار ا، لأنه يعود إلى التباين الطبيعي في البيانات.

اختيار العينات بأفضل أساليب المعاينة لا يضمن أن تكون العينة المختارة ممثلة للمجتمع

مثال :

إذا كان متوسط نسبة الذكاء في مجتمع (100) وأخذنا عينتين حجمهما 100 وطبقنا الاختبار على كل منهما

فإذا حصلنا على متوسط لنسبة ذكاء في العينة الأولى (105)

وحصلنا على متوسط نسبة ذكاء في العينة الثانية (95)

العينة الأولى يزيد متوسطها بمقدار (5) والثانية يقل بمقدار (5) عن متوسط المجتمع

هذا الاختلاف في متوسطات العينات يسمى خطأ المعاينة ،فهو لايرجع للباحث ، ولا يرجع إلى عيب في أسلوب اختيار العينة، ولكن يرجع للصدفة التي تحدث كلما احترنا عينة عشوائية من مجتمع

خطأ التحيز (غير المرتبط بالمعاينة) Nonsampling Bias

يشير إلى الأخطاء التي تحدث نتيجة عوامل غير مرتبطة باختيار العينة:

- ✓تكون الأسئلة غامضة (مضللة) ومن ثم تكون إجابات المبحوثين غير صحيحة
- ✓عند عدم جمع البيانات من بعض الأفراد أو جمع بيانات أكثر من مرة من نفس الأفراد
- ✓عند عدم وجود إطار معاينة سليم عند اختيار العينة

موجود حتى لو تم اختيار العينة بشكل صحيح، أو حتى ولو تم اجراء الدراسة على المجتمع

مثال

نفترض أنك أجريت استطلاعا لقياس رضا العملاء عن منتج معين، باستخدام استبيان يتضمن أسئلة متعددة. إذا كانت هناك مشكلة في صياغة الأسئلة (مثلاً، إذا كانت الأسئلة مضللة أو معقدة)، قد يؤدي ذلك إلى إجابات غير دقيقة.

نفترض أن نتائج الاستقصاء تشير إلى أن 80% من العملاء راضون، لكن بعد إجراء مراجعة دقيقة، تكتشف أن 20% من المشاركين فسروا الأسئلة بشكل خاطئ، مما أدى إلى تقديرات غير صحيحة.

إذا كان الرقم الحقيقي للراضين هو 70%، فإن خطأ التحيز هنا هو 10.%

تحيز المعاينة Sampling Bias

✓ اختيار عينة غير تمثيلية، أي احتمالية عدم تساوي الاختيار لكل وحدة عينة

✓ فشل تحليل الترجيح لعينة احتمالية غير متساوية

مثال1

➢ **عينة غير تمثيلية**

- اختيار الأسر القريبة من الطريق
- استخدام نسبة عينة مختلفة في مناطق مختلفة

➢ الفشل في إجراء الترجيح الإحصائي

مثال2

تخيل أنك ترغب في دراسة متوسط دخل الأسر في مدينة ما. إذا قمت بجمع بياناتك فقط من منطقة غنية، فإن العينة المختارة لن تمثل المدينة بالكامل.

اذا افترضنا أن متوسط دخل الأسر في المنطقة الغنية هو 300000 دينار، بينما متوسط دخل الأسر في المدينة بالكامل هو 60,000 دينار.

هنا، التحيز في اختيار العينة قد يعكس زيادة مقدارها 240000
دينار عن المتوسط الحقيقي.

باختصار...

خطأ المعاينة

- لا يمكن تجنبه إذا كانت العينة أقل من 100% من المجتمع
- يمكن التحكم فيه عن طريق اختيار حجم العينة وطريقة اختيار العينة المناسبين
- يمكن حسابه بدقة بعد وقوع الحدث

التحيز

- يتضمن التحيز المرتبط والغير مرتبط بالمعاينة
- يرجع إلى أخطاء يمكن تجنبها
- لا يمكن قياسها بدقة
- يتطلب التحكم والوقاية ويتطلب اهتماما دقيقًا