



# المعهد العربي للتدريب والبحوث الإحصائية

دورة متوسطة حول تصميم عينات المسوح الإحصائية

10

العينات المعقدة/المركبة complex samples

إعداد:

نايف عابد

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

# مفهوم العينة المعقدة : Complex Sample

يمكن تصنيف العينات من ناحية التصميم إلى نوعين:

❖ العينة العشوائية البسيطة Simple Random Sample

❖ العينة المركبة/المعقدة Complex Sample

بشكل عام فإن العينة التي تصميمها مختلف عن العينة العشوائية البسيطة Simple

Random Sample مع الإرجاع (WR) أو بدون الإرجاع (WOR) يمكن اعتبارها

عينة معقدة/مركبة Complex

# العينات المعقدة او المركبة complex samples قد تحتوي على تصميم عينة كالتالي :

❖ **التقسيم الطبقي Stratification**: عند تقسيم المجتمع المستهدف إلى أجزاء

متجانسة وغير متقاطعة (طبقات) وعند سحب عينة من كل طبقة.

❖ **العقدة (التقسيم إلى عناقيد) Clustering**: حيث تكون وحدات المعاينة في

المجتمع ضمن مجموعات تحتويها، مثل المدارس تحتوي على الطلاب، المستشفيات

تحتوي على الأطباء أو المرضى، المناطق الجغرافية تحتوي على الأسر او

الأفراد....الخ

❖ **تعدد المراحل Multiple Stage:** في العينة متعددة المراحل تكون المرحلة الأولى هي اختيار العناقيد، ويكون هناك مرحلة ثانية يتم اختيار عينة من كل عنقود في المرحلة الأولى، ويمكن إضافة مرحلة ثالثة وهكذا. مثل اختيار عينة مناطق عد ، وفي المرحلة الثانية يتم اختيار عينة من الأسر، وفي المرحلة الثالثة يتم اختيار عينة من الأفراد

❖ **عدم تساوي الاحتمالات Unequal selection probabilities:** وهذا يحصل عندما يتم سحب العينة بطريقة الاحتمال المتناسب مع الحجم -probability-proportional-to-size (PPS)، ويكون هناك مقياس لحجم وحدات المعاينة measure of size

❖ **العينة المنتظمة:** عند سحب عينة منتظمة بفترة ثابتة، وهناك من يصنفها أنها عشوائية إذا كانت نقطة البداية عشوائية

**مثال 1: عينة منشآت تم تصميمها كالتالي:**

عينة طبقية عشوائية بسيطة ، حيث تم تقسيم الإطار إلى طبقات حسب :

• المنطقة الجغرافية : هناك منطقتان رئيسيتان

• النشاط الاقتصادي على الحد الثاني ISIC2

حجم العينة كان 10000 منشأة تم توزيعها بطريقة متناسبة مع حجم الطبقة

**مثال 2: عينة مسح إنفاق واستهلاك الأسرة.**

هي عينة طبقية عنقودية ذات مرحلتين:

• المرحلة الأولى : اختيار مناطق عد جغرافية من إطار مناطق العد

المرحلة الثالثة: اختيار عينة من الاسر من كل منطقة عد تم اختيارها في المرحلة الأولى

الطبقات للمرحلة الأولى وهي:

المحافظة: حيث يوجد 16 محافظة

نوع التجمع السكاني: حضر، ريف

حجم عينة العناقيد = 500

حجم العينة من كل عنقود 12 اسرة

حجم العينة الكلي = 6000 اسرة

### مثال 3: استطلاع رأي حول موضوع العنف الأسري

عينة طبقية عنقودية ذات ثلاثة مراحل :

•المرحلة الأولى : اختيار مناطق عد جغرافية

•المرحلة الثانية: عينة من الاسر من كل منطقة عد تم اختيارها في المرحلة الاولى

المرحلة الثالثة : اختيار عينة من الافراد ذوي الاعمار 18سنة واكثر بطريقة جداول كيش العشوائية

الطبقات للمرحلة الاولى وهي:

المحافظة: حيث يوجد 16محافظة

نوع التجمع السكاني: حضر، ريف

حجم عينة العناقيد = 500

حجم العينة من كل عنقود 12اسرة

حجم العينة الكلي = 6000 اسرة

# جداول كيش Kish grid

هي جداول تم تصميمها لضمان اختيار عشوائي لفرد من الأسرة من ضمن مجموعة من الأفراد.

ترتيب الأفراد يمكن ان يعتمد اما على العلاقة برب الأسرة او على العمر او على التعليم او على أي متغير يعطي توزيع وترقيم للأفراد.

رقم الفرد في الاسرة	ترتيب الاسرة في العينة											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	2	1	2	1	2	1	2	1	2	1	2
3	1	2	3	1	2	3	1	2	3	1	2	3
4	1	2	3	4	1	2	3	4	1	2	3	4
5	1	2	3	4	5	1	2	3	4	5	3	5
6	1	2	3	4	5	6	1	2	3	4	5	6
7	1	2	3	4	5	6	7	1	3	5	6	7
8	1	2	3	4	5	6	7	8	2	4	6	8
9	1	2	3	4	5	6	7	8	9	1	4	9
10	1	2	3	4	5	6	7	8	9	10	1	10

مثال : اذا كان رقم الاسرة في منطقة العد = 7 ونريد اختيار لافرد من الفئة العمرية 18 سنة فأكثر وكان

عدد افراد الاسرة المؤهلين 5 وتم ترتيب الافراد حسب العمر من الاكبر الى الاصغر.

حسب الجدول فالفرد الذي تم اختياره في الاسرة هو الفرد رقم 2

رقم الفرد في الاسرة	ترتيب الاسرة في العينة											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	2	1	2	1	2	1	2	1	2	1	2
3	1	2	3	1	2	3	1	2	3	1	2	3
4	1	2	3	4	1	2	3	4	1	2	3	4
5	1	2	3	4	5	1	2	3	4	5	3	5
6	1	2	3	4	5	6	1	2	3	4	5	6
7	1	2	3	4	5	6	7	1	3	5	6	7
8	1	2	3	4	5	6	7	8	2	4	6	8
9	1	2	3	4	5	6	7	8	9	1	4	9
10	1	2	3	4	5	6	7	8	9	10	1	10



## اثر تصميم العينة $Deff$ : Design Factor

من اجل قياس مدى كفاءة تصميم العينة ، يكون بمقارنة التباين في حال العينة المركبة/المعقدة مع العينة

العشوائية البسيطة ولنفس حجم العينة

$$Deff = \frac{V(\theta)_{Complex}}{V(\theta)_{SRS}}$$

يمكن حساب قيمة معامل التصميم Design Factor ويسمى  $Deft$  وهو عبارة عن الجذر التربيعي لأثر

تصميم العينة

$$Deft = \sqrt{Deff} = \sqrt{\frac{V(\theta)_{Complex}}{V(\theta)_{SRS}}}$$

# التقدير في حالة العينة المركبة/المعقدة :

تقدير هارفتس تومسون The Horvitz–Thompson Estimator يمكن تعميمه ليشمل الكثير من انواع تصميم العينة المركبة/معقدة

$$\hat{y}_{HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} = \sum_{i=1}^n y_i w_i$$

تقدير المجموع من العينة

يمكن الاعتماد على تقدير المجموع ( هارفتس تومسون The Horvitz–Thompson Estimator ) لتقدير

المتوسط بقسمة تقدير المجموع على الأوزان .

$$\sum_{i=1}^n w_i$$

# تقدير التباين في حالة العينة المركبة/المعقدة :

يعتمد على تصميم العينة المركبة/المعقدة، حيث يمكن حساب التباين بشكل مباشر للتقديرات.  
وهناك أسلوبان شائعان لتقدير التباين هما:

1- Taylor series linearization methods

2- replicate methods او re-sampling methods

مثل طرق Jack-knife او bootstrap

هناك طريقة ثالثة أيضا من خلال ما يسمى generalized variance functions.

تمارين تطبيقية على برنامج SPSS