



المعهد العربي للتدريب والبحوث الإحصائية

دورة حول تصميم عينات المسوح الإحصائية

6

العينة الطبقية Stratified Sample

مفهوم العينة الطبقية:

❖ يتم تقسيم المجتمع إلى مجتمعات جزئية (طبقات) بحيث تكون مفردات كل جزء أو طبقة متجانسة فيما بينها ومختلفة عن الطبقات الأخرى، ويتم التعامل مع كل طبقة وكأنها مجتمع مستقل نسحب منه عينة عشوائية ذات حجم معين.

❖ لذلك يسبق هذا النوع من العينات تقسيم المجتمع (الإطار) إلى طبقات أو مجتمعات جزئية.

❖ الهدف من التقسيم الطبقي هو تقليل التباين وتحسين جودة التقديرات الناتجة (تخفيض اخطاء المعاينة)

مبادئ تقسيم المجتمع إلى طبقات:

عند تقسيم المجتمع إلى طبقات يجب مراعاة النقاط التالية:

- أن يكون مجموع وحدات المعاينة لجميع الطبقات مساوي لمجموع وحدات المعاينة للمجتمع ويجب أن لا يكون تداخل بين الطبقات.
- أن تكون وحدات المعاينة في داخل كل طبقة متجانسة بالنسبة للصفة المدروسة.
- هناك حالات عملية تؤدي إلى ضرورة استخدام أسلوب العينة الطبقية، فمثلا إذا كانت نتائج المسح الاحصائي مطلوبة حسب مستويات معينة للنشر.
- استخدام الطرق العلمية عند تقسيم المجتمع إلى فئات إذا كان المتغير الأصلي الذي يعتمد عليه في التقسيم متغير متصل.

مميزات العينة الطبقية:

1. في العينة الطبقية يكون المجتمع متجانسا في كل طبقة ويكون المجتمع ممثل بشكل جيد حيث تؤخذ عينات من مختلف الطبقات، والتي قد يكون لبعضها أهمية خاصة.
2. استخدام العينة الطبقية فعال اكثر من العينات الأخرى خاصة في حالة وجود مجتمع غير متجانس وفي حالة وجود قيم متطرفة لبعض وحدات المعاينة.
3. يمكن استخدام العينة الطبقية للحصول على نتائج على مستويات نشر معينة.

يمكن استخدام العينة الطبقيّة في الحالات التالية:

- وجود توزيع طبقي في المجتمع مثل التقسيم الجغرافي كالأقاليم والمحافظات أو طبقات المدن والقرى والمخيمات وكذلك طبقات الأغنياء والفقراء...حجم المنشآت صغيرة وكبيرة وغير ذلك، أو أي تقسيم للمجتمع يكون فيه أثر اجتماعي اقتصادي.

- وجود مستويات تفصيلية تطلب تقديرات منفصلة.

- توفر المعلومات اللازمة لإنشاء الطبقات لكل وحدات المجتمع الحاجة لاستخدام طرق

معينة مختلفة ضمن أجزاء مختلفة من المجتمع

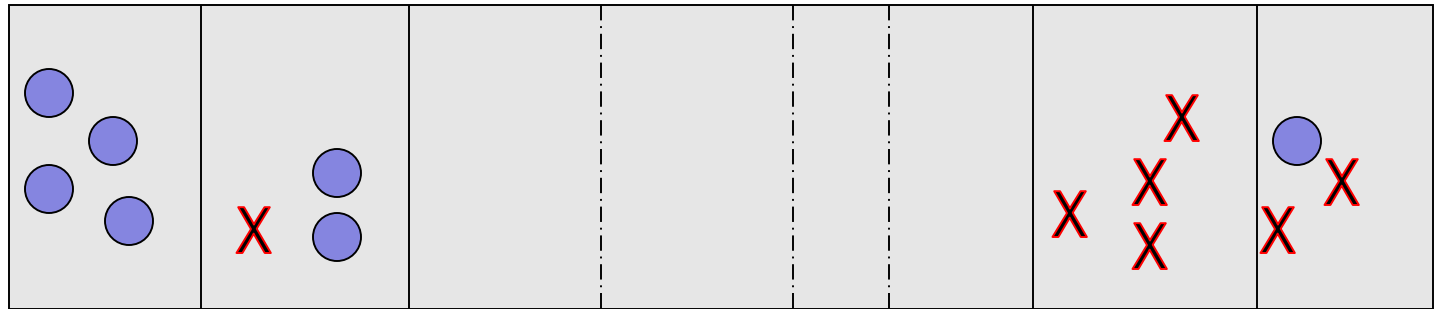
يكون تباين الطبقات أقل عندما:

- يكون متغير الدراسة مرتبط بمتغير إنشاء الطبقة

- تكون الطبقة متجانسة داخليا فيما يتعلق بمتغيرات الدراسة

الإطار

$$N = N_1 + N_2 + \dots + N_H$$



1

2

h

H

حجم الطبقة

N_1

N_2

N_h

N_H

تصميم العينة الطبقية

- تقسيم الإطار حسب متغير لتشكيل الطبقات (أو عدة متغيرات)
- بعد تقدير حجم عينة على المستوى الإجمالي يتم توزيع العينة على كافة الطبقات الفعلية بطريقة معينة
- على الأقل يتم سحب وحدتان معاينة من كل طبقة
- يمكن استخدام أي عينة احتمالية لسحب العينة من كل طبقة (عشوائية بسيطة، منتظمة، عنقودية، PPS... الخ)
- لتقدير قيمة لكل المجتمع يتم تجميع التقديرات من كل الطبقات

أمثلة على التقسيم الطبقي

• في المسوح الأسرية

- حضر/ريف

- المحافظة/المنطقة

• مسوح المنشآت

- النشاط الاقتصادي

- عدد العاملين

- حجم الإنتاج

• المسوح الزراعية

- نوع الحيازة الزراعية (نباتية /حيوانية)

- حجم الحيازة

الرموز المستخدمة

H	عدد الطبقات الكلي
N_h	حجم المجتمع في الطبقة h
n_h	حجم العينة في الطبقة h
f_h	كسر المعاينة (نسبة المعاينة) في الطبقة h $\frac{n_h}{N_h}$
X_h	مجموع المجتمع في الطبقة h
\bar{X}_h	متوسط المجتمع في الطبقة h
\bar{x}_h	متوسط العينة في الطبقة h
$X_{h,i}$	قيمة المتغير X لوحدة المعاينة i^{th} في الطبقة h ($i=1, 2, \dots, n_h$)
$X_{h,i}$	قيمة المتغير X لوحدة المعاينة i^{th} في الاطار في الطبقة h ($i=1, 2, \dots, N_h$)

• عدد وحدات الإطار الكلي $N = N_1 + N_2 + \dots + N_H$

تقديرات في حالة الطبقية (العشوائية البسيطة في كل طبقة) Stratified SRS

$$n = \sum_{h=1}^H n_h$$

حجم العينة الكلي = مجموع العينات في كل الطبقات

$$\bar{x}_h = \frac{\sum_{i=1}^{n_h} x_{hi}}{n_h}$$

تقدير المتوسط في العينة في الطبقة h

تقدير المجموع

للطبقة h
• للمجتمع

$$X_h = \sum_{i=1}^{N_h} X_{hi}$$

$$x_h = n_h \bar{x}_h$$

• للعينة

لكل الطبقات:
• للمجتمع

$$X = \sum_{h=1}^H X_h$$

$$x = \sum_{h=1}^H x_h$$

• للعينة

تقدير المتوسط

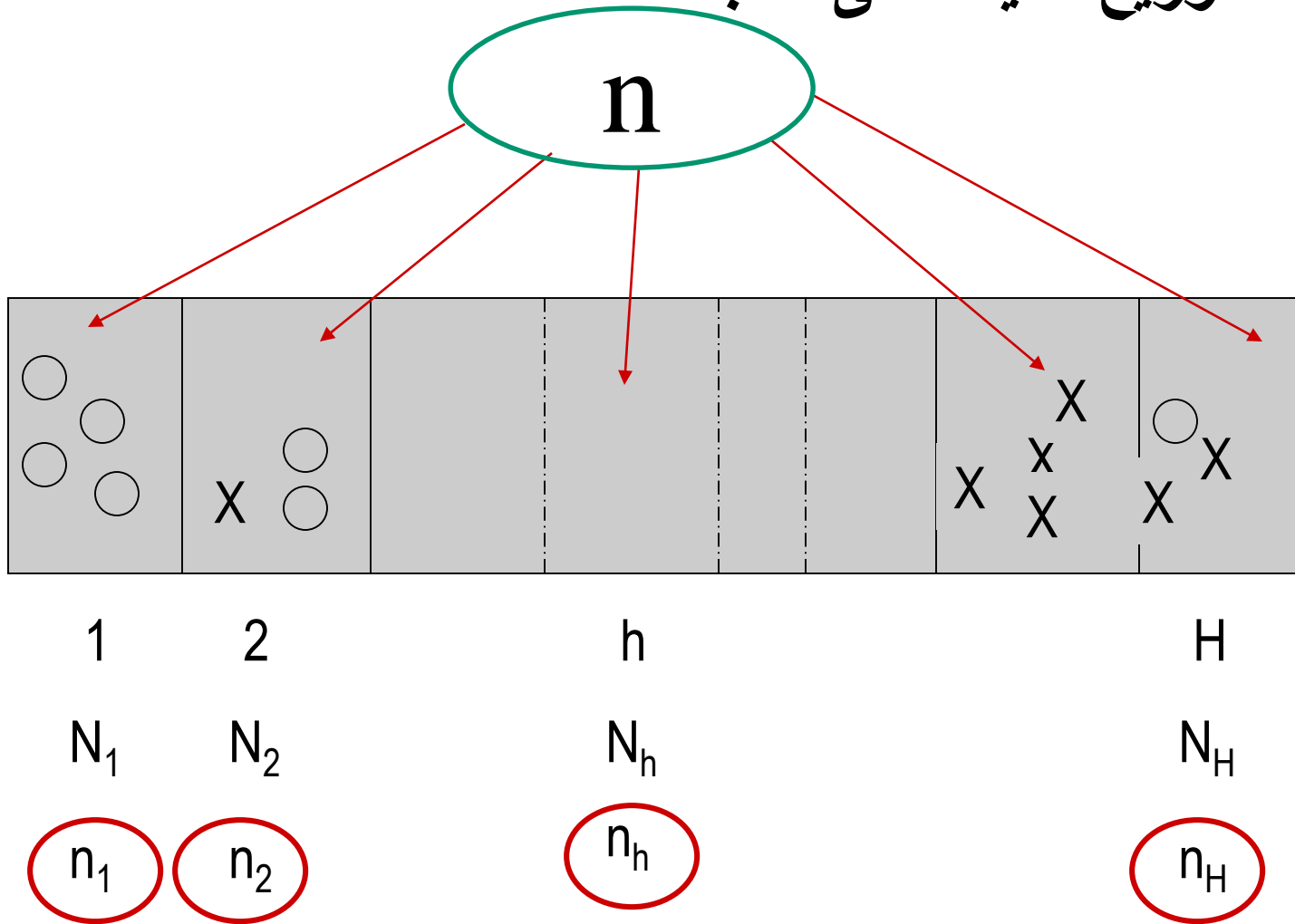
$$\bar{X}_h = \frac{X_h}{N_h}$$

$$\bar{x}_h = \frac{x_h}{n_h}$$

$$\bar{X} = \frac{X}{N}$$

$$\bar{x} = \frac{x}{n}$$

توزيع العينة على الطبقات



كم حجم العينة في كل طبقة???

توزيع العينة على الطبقات

غير متناسب مع الحجم

متناسب مع الحجم

توزيع متساوي

توزيع نيمان

التوزيع الامثل

توزيعات اخرى

1. التوزيع المتناسب مع الحجم: ويستخدم هذا التوزيع في حالة كون الطبقات مختلفة من حيث الحجم الطبقة.

هذا أسلوب من أكثر الأساليب شيوعا نظرا لسهولة استخدامه، عندما لا توجد معلومات عدا عدد وحدات المعاينة في كل طبقة فيمكن استخدام المعادلة التالية لتقدير حجم العينة في

$$n_h = n \left(\frac{N_h}{N} \right) \quad \text{الطبقة (h)}$$

مثال: استخدم التوزيع المتناسب في توزيع عينة مقدارها 35 على الطبقات

حجم الطبقة	الطبقة
200	1
150	2

$$n \left(\frac{N_h}{N} \right) = 35 \left(\frac{200}{350} \right) = 20 \quad \bullet \text{حجم عينة الطبقة الأولى:}$$

$$n \left(\frac{N_h}{N} \right) = 35 \left(\frac{150}{350} \right) = 15 \quad \bullet \text{حجم عينة الطبقة الثانية:}$$

التوزيعات غير المتناسبة مع الحجم

2. التوزيع المتساوي Uniform or Equal Allocation: وتستخدم هذه الطريقة في حالة كون الطبقات متقاربة في الحجم ويتم هنا إعطاء كل طبقة عدد متساوي من العينة.

$$n_h = \left(\frac{n}{H}\right)$$

H عدد الطبقات

n حجم العينة الكلي

n_h حجم العينة في الطبقة h

مثال :

إذا كان حجم العينة 2000 وحدة معاينة وعدد الطبقات يساوي (8) فإن حجم

العينة في كل طبقة يساوي $250 = 8 \div 2000$

3. توزيع نيمان: ويستخدم هذا التوزيع لتقليل التباين داخل الطبقة بزيادة حجم العينة للطبقة الأكثر تبايناً لتقديرات المؤشرات الرئيسية.

ويقدر حجم العينة للطبقة (h) حسب المعادلة التالية:

$$n_h = n \frac{N_h S_h}{\sum_{h=1}^H N_h S_h}$$

S_h الانحراف المعياري على مستوى كل طبقة فيمكن الحصول عليه من خلال تعداد سابق أو يمكن تقديره من عينات سابقة.

4. التوزيع الأمثل :Optimum Allocation

- على نفس المعادلة لتوزيع نيمان اذا تم اعتبار التكلفة في المعادلة ، بحيث تتوزع العينة حسب التباين وحسب التكلفة C_h في الطبقة h

$$n_h = n \times \frac{N_h \times (S_h / \sqrt{c_h})}{\sum N_h \times (S_h / \sqrt{c_h})}$$

4. توزيعات اخرى:

(a) توزيع الجذر التربيعي Square-root allocation

$$n_h = n \frac{\sqrt{N_h}}{\sum \sqrt{N_h}}$$

(b) توزيع متوسط التوزيعين المتساوي والمتناسب مع الحجم

$$n_h = \left(\frac{n_{h1} + n_{h2}}{2} \right)$$

حيث:

$$n_{h1} = n \left(\frac{N_h}{N} \right) \quad n_{h2} = \left(\frac{n}{H} \right)$$

تمرين: استخدم توزيع الجذر التربيعي وتوزيع متوسط التوزيعين المتساوي والمتناسب مع الحجم في توزيع عينة مقدارها 35 على الطبقات

حجم الطبقة	الطبقة
200	1
150	2

حساب التباين في حالة العينة الطبقية العشوائية البسيطة

تباين المتوسط للعينة الطبقية :

$$V(\bar{x}_{st}) = \sum (1 - f_h) [(W_h^2 S_h^2) / n_h]$$

تباين المتوسط للعينة الطبقية ذات التوزيع المتناسب مع الحجم

$$V(\bar{x}_{st}) = (1 - f) \sum (W_h^2 S_h^2) / n_h$$

$$W_h = \frac{N}{N_h}$$

اثر تصميم العينة The design effect

Deff = التباين للتقدير حسب تصميم العينة / التباين للتقدير حسب العينة العشوائية

البسيطة بدون ارجاع ولنفس حجم العينة

$$v(\bar{x}_h) / v(\bar{x}_{sr}) = \text{Deff}$$

في العينة الطبقية ذات التوزيع المتناسب مع الحجم تكون قيمة Deff اقل من 1

تمارين تطبيقية على برنامج SPSS