المعهد العربي للتدريب والبحوث الإحصائية

Arab Institute for Training and Research in Statistics

دورة تدريبية متوسطة المدى حول

تصميم العينات ومنهجيات المسوح في الاحصاءات الرسمية

10 مايو / ايار – 5 تموز / يوليو 2022

توزيعات المعاينة ونظرية النهاية المركزية

محاضرة رقم 2

د. حسان أبو حسان

www.aitrs.org
info@aitrs.org

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- توزيع المعاينة **Sampling distribution**

ما هو إلا جدول توزيع احتمالي للإحصاء المحسوب من العينة.
وتجدر الإشارة إلى أنه:

- ✓ المقاييس أو المؤشرات المحسوبة من بيانات العينة كالوسط الحسابي $\bar{X}$ و التباين S^2 و النسبة $\hat{P}$ تسمى إحصاء العينة (Sample Statistic).

- ✓ أما المقاييس أو المؤشرات المحسوبة من بيانات المجتمع كالوسط الحسابي μ و التباين σ^2 و النسبة P تسمى معالم المجتمع (Population parameters).

- ○ العينة = الجزء الذي يتم اختياره من المجتمع بطريقة عشوائية و هذه العينة تستخدم لتوفير الوقت و الجهد.

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- توزيع معاينة الوسط الحسابي للعينة $\bar{X}$

❖ هو عبارة عن توزيع احتمالي للوسط الحسابي للعينة و لتوضيح مفهوم توزيع المعاينة للوسط الحسابي نفترض أن لدينا مجتمع حجمه N مفردة هي:

$$X_1, X_2, X_3, X_4, \dots, X_N$$

❖ نريد سحب عينة من هذا المجتمع حجمها n و مفرداتها :

$$x_1, x_2, x_3, \dots, x_n$$

▸ عدد العينات ذات الحجم n و التي يمكن اختيارها من المجتمع ذات الحجم N و خصوصا إذا كان السحب بدون إرجاع:

$$NCn = \text{عدد العينات}$$

د. حسان أبو حسان

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- توزيع معاينة الوسط الحسابي للعينة $\bar{X}$

❖ تذكير:

$$N_{C_n} = \frac{N!}{n!(N-n)!}$$

مثال:

إذا كان لدينا مجتمع حجمه 5 مفردات و نريد أخذ عينة حجمها 3 مفردات فإن عدد العينات الممكنة:

$$5_{C_3} = \frac{5!}{3!(5-3)!} = \frac{5!}{3!*2!} = \frac{5*4*3!}{3!*2*1} = 10$$ عينات

○ باستخدام الآلة الحاسبة:

◄ نكتب 5 ثم نضغط على n_{C_r} ثم 3 ثم = النتيجة =10

عدد العينات في حال السحب مع الارجاع:

$$N^n$$

مثال: أفترض ان لدينا مجتمع حجمة 3 سحبت كل العينات الممكنة والتي حجمها 2 فما هو عدد العينات الممكنة اذا كان السحب مع الارجاع:

عدد العينات الممكنة $3^2 = 9$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **توزيع معاينة الوسط الحسابي للعينة** $\overline{X}$

✓ **قبل الخوض في شرح توزيعات المعاينة نتذكر أن:**

- **أهم مقاييس المجتمع و التي تسمى معلمات:**
- **1) الوسط الحسابي μ:**

$$\mu = \frac{\sum X}{N}$$

- **2) التباين σ^2:**

$$\sigma^2 = \frac{1}{N}\left[\sum X^2 - \frac{\left(\sum X\right)^2}{N}\right]$$

- **3) الانحراف المعياري σ = الجذر التربيعي الموجب للتباين:**

$$\sigma = \sqrt{\sigma^2}$$

- **4) النسبة P:**

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

(X: عدد المفردات التي تتوافر فيهم الخاصية محل الدراسة

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- توزيع معاينة الوسط الحسابي للعينة $\bar{X}$

✓ قبل الخوض في شرح توزيعات المعاينة نتذكر أن:

- أهم مقاييس العينة و التي تسمى إحصاءات:
- 1) الوسط الحسابي للعينة $\bar{X}$:

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

- 2) التباين S^2:

$$S^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]$$

- 3) الانحراف المعياري S = الجذر التربيعي الموجب للتباين:

$$S = \sqrt{S^2}$$

- 4) النسبة P:

$$\hat{P} = \frac{x}{n}$$

(x: عدد المفردات التي تتوافر فيهم الخاصية في العينة)

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- توزيع معاينة الوسط الحسابي للعينة $\bar{X}$
  - لتوضيح مفهوم توزيع المعاينة للوسط الحسابي $\bar{X}$ سوف نتناول المثال التالي:

➢ **مثال 1:**

بافتراض أن لدينا مجتمع حجمه 4 مفردات هي 6، 3، 5، 2 و أردنا سحب عينة من هذا المجتمع حجمها مفردتين أوجد:

1. الوسط الحسابي و التباين للمجتمع.
2. إيجاد جميع العينات الممكن سحبها من المجتمع.
3. تكوين توزيع المعاينة $\bar{X}$
4. حساب كل من الوسط الحسابي ل $\bar{X}$ و مقارنته بالوسط الحسابي للمجتمع.
5. أحسب التباين للوسط الحسابي و كذلك الانحراف المعياري للوسط (الخطأ المعياري ل $\bar{X}$)

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

1. الوسط الحسابي للمجتمع

$$\mu = \frac{\sum x}{N} = \frac{2+5+3+6}{4} = 4$$

2. التباين:

X	X²
2	4
5	25
3	9
6	36
المجموع = 16	المجموع = 74

$$\sigma^2 = \frac{1}{N}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{N}\right] = \frac{1}{4}\left[74 - \frac{(16)^2}{4}\right] = \frac{10}{4} = 2.5$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

2- عدد العينات الممكن سحبها (السحب بدون إرجاع):

$$N_{C_n} = 4_{C_2} = \frac{4!}{2!(4-2)!} = \frac{4!}{2!*2!} = \frac{4*3*2!}{2!*2*1} = 6 \text{ عينات}$$

3- لتكوين توزيع معاينة $\bar{X}$ نحسب أولا الأوساطة الحسابية لكل العينات الممكن سحبها من هذا المجتمع و عددها 6 كالتالي كما في جدول رقم 1 ثم نكون جدول توزيع احتمالي ل $\bar{X}$ (جدول2):

جدول 1

العينة	$\bar{X}$
(2،5)	3.5
(2،3)	2.5
(2،6)	4
(5،3)	4
(5،6)	5.5
(3،6)	4.5
$\sum$	$\sum \bar{X}$ = 24

جدول 2

$\bar{X}$	$P(\bar{X})$
3.5	1/6
2.5	1/6
4	2/6
5.5	1/6
4.5	1/6

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

4- الوسط الحسابي ل $\overline{X}$ و يرمز له بالرمز $\mu_{\overline{X}}$ حيث:

$$\mu_{\overline{X}} = \frac{\sum \overline{X}}{\text{عدد العينات}} = \frac{24}{6} = 4$$

- **ملاحظة: الوسط الحسابي ل $\overline{X}$ دائما يساوي الوسط الحسابي للمجتمع أي أن:**

$$\mu_{\overline{X}} = \mu$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

5- تباين الوسط الحسابي ل$\overline{X}$ و يرمز له بالرمز $\sigma^2_{\overline{X}}$ حيث:

- إذا كان لدينا مجتمع غير محدود $\left(\frac{n}{N} \leq 0.05\right)$ ◄ $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$

- إذا كان لدينا مجتمع محدود $\left(\frac{n}{N} > 0.05\right)$ ◄ $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} * \left(\frac{N-n}{N-1}\right)$

- حيث أن $\frac{N-n}{N-1}$ يسمى **بمعامل التصحيح** و يستخدم إذا كان حجم المجتمع محدود.

لدينا في التمرين : $\frac{n}{N} = \frac{2}{4} = 0.50 > 0.05$ ◄ لدينا مجتمع محدود ◄ $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} * \left(\frac{N-n}{N-1}\right)$

إذن: $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} * \left(\frac{N-n}{N-1}\right) = \frac{2.5}{2} * \left(\frac{4-2}{4-1}\right) = \frac{2.5}{2} * \frac{2}{3} = 0.833$

و بالتالي فإن الانحراف المعياري ل $\overline{X}$ أو ما يسمى بالخطأ المعياري ل $\overline{X}$

$$\sigma_{\overline{X}} = \sqrt{\sigma^2_{\overline{X}}} = \sqrt{0.833} = 0.913$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

من المثال السابق نستنتج ما يلي:

1- $\mu_{\overline{X}} = \mu$

2- التباين ل $\overline{X}$

- إذا كان $\frac{n}{N} \leq 0.05$ **فإن:** $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$

- إذا كان $\frac{n}{N} > 0.05$ **فإن:** $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} * \left(\frac{N-n}{N-1}\right)$

3- الخطأ المعياري ل $\overline{X}$

- إذا كان $\frac{n}{N} \leq 0.05$ **فإن:** $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

- إذا كان $\frac{n}{N} > 0.05$ **فإن:** $\sigma_{\overline{X}} = \sqrt{\sigma^2_{\overline{X}}} = \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **ملحوظة:**
- **في حالة عدم معلومية σ للمجتمع فإننا نستخدم بدلا الانحراف المعياري للعينة S و بالتالي فإن الخطأ المعياري $\sigma_{\overline{X}}$ يصبح $S_{\overline{X}}$**

◄ يصبح الخطأ المعياري كالآتي:

- ✓ **إذا كان** $\frac{n}{N} \leq 0.05$ **فإن:** $S_{\overline{X}} = \frac{S}{\sqrt{n}}$
- ✓
- ✓ **إذا كان** $\frac{n}{N} > 0.05$ **فإن:** $S_{\overline{X}} = \frac{S}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$

- **إذا لم يحدد حجم المجتمع N فإننا نعتبر حجم المجتمع غير محدود و بالتالي فإن** $\frac{n}{N} \leq 0.05$ **◄ الخطأ المعياري سيكون:**

- ✓ **إذا كانت σ معلومة فإن:** $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

أو

- ✓ **إذا كانت σ غير معلومة فإن:** $S_{\overline{X}} = \frac{S}{\sqrt{n}}$

- **تمرين 2**

إذا كان لدينا مجتمع يتكون من 5 مفردات و أردنا سحب عينة عشوائية من هذا المجتمع حجمها 3 مفردات مع ملاحظة أن مفردات المجتمع هي 1، 2، 3، 4، 5.

1) أحسب الوسط الحسابي و التباين لمفردات المجتمع.
2) كون توزيع معاينة $\bar{X}$ إذا كان السحب بدون إرجاع.
3) أحسب الوسط الحسابي للعينات التي تم سحبها و قارن بينه و بين الوسط الحسابي للمجتمع.
4) أحسب الخطأ المعياري ل $\bar{X}$.

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

1) الوسط الحسابي للمجتمع

$$\mu = \frac{\sum x}{N} = \frac{1+2+3+4+5}{5} = 3$$

X	X²
1	1
2	4
3	9
4	16
5	25
$\sum \bar{X} = 15$	55

2) تباين المجتمع σ^2:

$$\sigma^2 = \frac{1}{N}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{N}\right]$$

$\rightarrow$

$$= \frac{1}{5}\left[55 - \frac{(15)^2}{5}\right] = 2$$

$$\Rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{2} = 1.414$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- **الحل:**

2) توزيع المعاينة ل $\bar{X}$

نحسب أولا عدد العينات:

$$N_{C_n} = 5_{C_3} = \frac{5!}{3!*2!} = 10$$ **عينات**

ثم نحسب الأوساط الحسابية للعينات كالتالي:

العينة	$\bar{X}$
(1،2،3) →	2
(1،2،4) →	2.33
(1،2،5) →	2.67
(1،3،4) →	2.67
(1،3،5) →	3
(1،4،5) →	3.33
(2،3،4) →	3
(2،3،5) →	3.33
(2،4،5) →	3.67
(3،4،5) →	4
∑	30

- **الحل:**

3) الوسط الحسابي ل $\mu_{\bar{X}} = \bar{X}$

$$\mu_{\bar{X}} = \frac{\sum \bar{X}}{\text{عدد العينات}} = \frac{30}{10} = 3$$

4) الخطأ المعياري ل $\sigma_{\bar{X}} = \bar{X}$

بما أن:

$$\frac{n}{N} = \frac{3}{5} = 0.60 > 0.05$$

يمكن حساب الخطأ المعياري على الشكل التالي:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{1.414}{\sqrt{3}} \sqrt{\frac{5-3}{5-1}} = 0.577$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)
  - توزيع معاينة الوسط الحسابي للعينة $\overline{X}$ يتبع تقريبا توزيع طبيعي بمتوسط يساوي μ و تباين يساوي $\sigma_{\overline{X}}^{2}$ عندما يكون حجم العينة كبيرا (يعتبر حجم العينة كبير إذا كان (n≥30)).
  - و بالتالي إذا كان: $n \geq 30$

فإن:

$$\overline{X} \sim N(\mu, \sigma_{\overline{X}}^{2})$$

- يمكننا تحويل التوزيع الطبيعي ل $\overline{X}$ إلى توزيع طبيعي معياري Z باستخدام التحويلة:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)

- تفيد نظرية النهاية المركزية في أنه يمكن استخدام جدول التوزيع الطبيعي المعياري حتى إذا كان المجتمع المسحوب منه العينة غير طبيعي بشرط ان يكون حجم العينة المسحوبة أكبر أو = 30.

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)
- توزيع المعاينة للنسبة في العينة ($\hat{P}$)
- نسبة الصفة في المجتمع (P) نحصل عليها بقسمة عدد المفردات التي تتوافر فيهم هذه الصفة في المجتمع (X) على حجم المجتمع الكلي (N) أي:

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

- أما نسبة الصفة في العينة $\hat{P}$ نحصل عليها بقسمة عدد المفردات التي تتوافر فيهم هذه الصفة في العينة (x) على حجم العينة الكلي (n) أي أن:

$$\hat{P} = \frac{x}{n}$$

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)
  - توزيع معاينة $\hat{p}$ تقريبا هو التوزيع الطبيعي بمتوسط = P و تباين $\sigma^2_{\hat{P}}$ وذلك إذا كان حجم العينة المسحوب كبيرا أي n*P ≥ 5 و n*(1- P) ≥5 و بالتالي فإن:

$$\hat{P} \sim N(P, \sigma^2_{\hat{P}}), \qquad P = \frac{X}{N}$$

$$\sigma^2_{\hat{P}} = \frac{P*(1-P)}{n} \Rightarrow \frac{n}{N} \leq 0.05$$ مجتمع غير محدود

$$\sigma^2_{\hat{P}} = \frac{P*(1-P)}{n}\left(\frac{N-n}{N-1}\right) \Rightarrow \frac{n}{N} > 0.05$$ مجتمع محدود

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)
- و بالتالي فإن الانحراف المعياري ل $\hat{P}$ أو ما يسمى بالخطأ المعياري ل $\hat{P}$:
  - الحالة 1:

إذا كانت P للمجتمع معلومة:

$$\sigma_{\hat{P}} = \sqrt{\frac{P(1-P)}{n}} \Rightarrow \frac{n}{N} \leq 0.05$$ ← مجتمع غير محدود

$$\sigma_{\hat{P}} = \sqrt{\frac{P*(1-P)}{n}} * \sqrt{\frac{N-n}{N-1}} \Rightarrow \frac{n}{N} > 0.05$$ ← مجتمع محدود

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)

- الحالة 2:

<u>إذا كانت P للمجتمع غير معلومة:</u>

$$\sigma_{\hat{P}} = \sqrt{\frac{\hat{P}*(1-\hat{P})}{n}} \Rightarrow \frac{n}{N} \leq 0.05$$ ⟹ مجتمع غير محدود

$$\sigma_{\hat{P}} = \sqrt{\frac{\hat{P}*(1-\hat{P})}{n}} * \sqrt{\frac{N-n}{N-1}} \Rightarrow \frac{n}{N} > 0.05$$ ⟹ مجتمع محدود

الباب الثاني: توزيعات المعاينة و نظرية النهاية المركزية

- نظرية النهاية المركزية Central Limit Theorem (CLT)

- إذا كان حجم المجتمع N غير معلوم فإننا نعتبر المجتمع غير محدود أي أن $\frac{n}{N} \leq 0.05$ و بالتالي فإن:

إذا كانت P معلومة ⟹

$$\sigma_{\hat{P}} = \sqrt{\frac{P(1-P)}{n}}$$

إذا كانت P غير معلومة ⟹

$$\sigma_{\hat{P}} = \sqrt{\frac{\hat{P}*(1-\hat{P})}{n}}$$

مثال:

إذا كانت نسبة المعيب في إنتاج إحدى الماكينات هو 10% سحبت عينة عشوائية مكونة من 50 وحدة. أحسب احتمال أن يكون بها نسبة معيب قدرها 5% وأقل.

الحل

عندما يكون حجم العينة كبير فان التوزيع الاحتمالي للنسبة في العينة $\hat{p}$ سوف يكون تقريبا التوزيع الطبيعي بمتوسط

$$\mu_{\hat{p}} = P$$

$$\mu_{\hat{p}} = 0.10$$

و انحراف معياري

$$\sigma_{\hat{p}} = \sqrt{\frac{P(1-P)}{n}}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{0.1 \times 0.9}{50}}$$

$$\sigma_{\hat{p}} = 0.042$$

و احتمال أن يكون نسبة المعيب 5%

$$P(\hat{p} \leq 0.05) = P(Z \leq \frac{\hat{p} - P}{\sqrt{\frac{P(1-P)}{n}}})$$

$$= P(Z \leq \frac{0.05 - 0.1}{\sqrt{\frac{0.1 \times 0.9}{50}}})$$

$$= P(Z \leq \frac{-0.05}{0.042})$$

$$= P(Z \leq -1.19)$$

$$= 0.117$$

تمرين: مجتمع مكون من 2، 4، 6، سحبت كل العينات الممكنة والتي حجمها 2 ، إذا كان السحب بدون إرجاع أوجد ما يلي:

1- الوسط الحسابي لتوزيع المعاينة للوسط واثبت أنه يساوي متوسط المجتمع
2- الانحراف المعياري (الخطأ المعياري) لتوزيع المعاينة للوسط

تمرين: إذا كان نسبة نجاح الطالب في اختبار معين يساوي 0.7 سحب عينة حجمها 5 طلاب أوجد:

1- الوسط الحسابي لتوزيع المعاينة لنسبة النجاح في الاختبار.
2- تباين توزيع المعاينة لنسبة النجاح.