المحاضرة الثانية:

المحور الثاني: توزيع المعاينة:

نفترض أنه لدينا مجتمعا من المفردات N يتبع توزيعا احتماليا معينا و أننا بصد سحب عينة حجمها n من هذا المجتمع، بالطبع ليس معنى هذا أن هناك عينة واحد يمكن سحبها ولكن يكون أمامنا عدد كبير من العينات التي يمكن سحبها من هذا المجتمع والتي حجم كل منها هو n من المفردات.

لإيجاد عدد العينات المكنة والمسحوبة من المجتمع N هناك حالتين في حالة السحب بالإرجاع وفي حالة السحب بدون إرجاع:

- في حالة السحب بالإرجاع فإن عدد العينات الممكنة يساوي N^n.
- في حالة السحب بدون إرجاع $C_N^n = \frac{N!}{(N-n)!n!}$

1– توزيع المعاينة لمتوسط حسابي من مجتمع طبيعي التوزيع:

عند أخذ جميع العينات الممكنة ذات الحجم n المجتمع حجمه N ثم حساب الوسط الحسابي لكل منها نجد أن كل قيمة من قيم الاوساط الحسابية للعينات تختلف عن الأخرى لذلك فإن سلوكها يأخذ سلوك متغير عشوائي يرمز له بالرمز نسمي توزيع المتغير العشوائي بتوزيع المتوسط الحسابي للعينة .

نظرية:

إذا كان متغير عشوائي يمثل مجتمع ما و $\bar{X}$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من ذات المجتمع، فإن القيمة المتوقعة لمتوسط العينة $E(\bar{X})$ يكتب كما يلي: $E(\bar{X}) = \mu_{\bar{x}} = \mu$

أما تباين الاوساط الحسابية للعينات العشوائية والمتساوية الحجم، فنرمز له بالرمز σ^2 وهو يحسب بعدة قوانين حسب الحالة الموافقة له، والجدول الموالي يلخص أهم الحالات كالآتي:

<table>
<tr><td></td><td>توزيع الوسط الحسابي</td><td colspan="2">قيمة تباين الوسط الحسابي</td><td colspan="2">قيمة تباين الوسط الحسابي في حالة السحب بدون ارجاع</td></tr>
<tr><td>σ^2 التباين معلوم</td><td>$\bar{X} \rightarrow N(\mu, \sigma_{\bar{x}}^2)$</td><td colspan="2">$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$</td><td colspan="2">$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$</td></tr>
<tr><td>σ^2 التباين</td><td>$\bar{X} \rightarrow N(\mu, \sigma_{\bar{x}}^2)$</td><td>n≥30</td><td>$\sigma_{\bar{x}}^2 = \frac{S^2}{n}$</td><td>n≥30</td><td>$\sigma_{\bar{x}}^2 = \frac{S^2}{n}\left(\frac{N-n}{N-1}\right)$</td></tr>
</table>

مجهول	$\overline{X} \rightarrow t(n-1)$	n<30	$\sigma_{\bar{x}}^2 = \frac{S^2}{n-1}$	n<30	$\sigma_{\bar{x}}^2 = \frac{S^2}{n-1}\left(\frac{N-n}{N-1}\right)$

حيث:

n: حجم العينة

μ: تباين المجتمع وهو نفسه تباين الوسط الحسابي $\overline{X}$

$\sigma_{\bar{x}}^2$: تباين القسط الحسابي $\overline{X}$

σ^2: تباين المجتمع.

S^2: تباين العينة، حيث $S^2 = \frac{(X-\bar{X})^2}{n-1}$

$\left(\frac{N-n}{N-1}\right)$: معامل الارجاع

لايجاد توزيع المتغير العشوائي $\overline{X}$ نستعمل التوزيع الطبيعي المعياري z حيث: $Z = \frac{\bar{X}-\mu}{\sigma_{\bar{x}}}$ في جميع الحالات ما عدا حالة واحدة وهي إذا كان σ^2 (تبيان المجتمع) مجهول وكان n<30، حيث في هذه الحالة $\overline{X}$ يتبع توزيع ستيودنت t حيث: $T = \frac{\bar{X}-\mu}{\sigma_{\bar{x}}} \rightarrow t_{(n-1)}$

مثال: إذا كان عدد سائقي سيارات الأجرة في مدينة ما هو 1500 سائق، وعلمت أن أعمارهم تتوزع توزيعا طبيعيا بمتوسط حسابي قدره 45 سنة، وانحراف معياري قدره 7 سنوات، فإذا سحبنا مع عدم الإرجاع من هذا المجتمع عينة عشوائية بها 16 سائقا. أوجد:

- أوجد توزيع المعاينة لمتوسط أعمار سائقي سيارات الأجرة؟
- أحسب احتمال أن يكون متوسط العمر لهذه العينة أكبر من 48 سنة؟

توزيع المعاينة لمتوسط حسابي من مجتمع غير طبيعي التوزيع:

إذا سحبت عينات عشوائية بسيطة ذات الحجم n، حيث $n \geq 30$ من مجتمع غير طبيعي توقعه μ وتباينه σ^2، فإن المقدار $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ يؤول إلى التوزيع الطبيعي المعياري وذلك حسب نظرية النهايات المركزية.

نظرية النهايات المركزية:

إذا سحبت عينة عشوائية حجمها n من مجتمع غير طبيعي وسطه μ وتباينه σ^2 فإن توزيع المعاينة لـ $\overline{X}$ يقترب من التوزيع الطبيعي بتوقع μ وتباين $\frac{\sigma}{\sqrt{n}}$ وذلك كلما كانت $n \geq 30$.

مثال: علمت أن متوسط الإنتاج اليومي لمؤسسة ما هو 30 وحدة بانحراف معياري 12 وحدة، سحبت منه عينة عشوائية حجمها 49 وحدة، فأوجد:

-توزيع المعاينة للوسط الحسابي للعينة.

- احتمالي أن يكون متوسط إنتاج العينة أقل من 35 وحدة؟

الحل:

لدينا: $\mu = 30$، $\sigma = 12$، $n=49$، كما أن توزيع المجتمع مجهول وبالتالي نعتمد على حجم العينة في تحديد توزيع المعاينة لوسط العينة $\overline{X}$ و $n=49 \geq 30$ ومنه:

$\overline{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$ ومنه $Z \rightarrow N(0,1)$ ، حيث: $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$

$$\overline{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right) \rightarrow \bar{X} \rightarrow N\left(30, \frac{(12)^2}{49}\right) \rightarrow \bar{X} \rightarrow N\left(30, \frac{(12)^2}{7}\right)$$

حساب $P(\bar{X} < 35)$

$$P(\bar{X} < 35) = P\left(\frac{\bar{X}-\mu}{\sigma_{\bar{X}}} < \frac{35-30}{12/_7}\right) = P(Z < 2.92) = Q(2.92)$$
$$= 0.9982$$

توزيع المعاينة لنسبة العينة $\widehat{p}$:

يحتاج الباحث في اغلب الدراسات لمعرفة نسبة ظاهرة معينة في المجتمع محل الدراسة، كنسبة المدخنين في مدينة ما، نسبة الذكور في جامعة ما، نسبة الوحدات التالفة في إنتاج مصنع معين، ... الخ، ففي كل هذه الحالات نجد أن المجتمع محل الدراسة منقسم إلى قسمين، قسم تتوافر فيه الظاهرة محل الدراسة (الخاصية المدروسة) والقسم الثاني لا تتوافر فيه الظاهرة المدروسة، ومجتمعات من هذا النوع يكون فيها المتغير وصفيا أي نوعيا لا نستطيع قياسه كميا، ونتعامل في هذا النوع من المجتمعات مع نسبة الظاهرة محل الدراسة في المجتمع، ويرمز لها بالرمز P ويطلق عليها نسبة المجتمع وهي تحسب كتالي:

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

X: عدد مفردات المجتمع التي تتحقق فيها الظاهرة المدروسة

N: العدد الكلي لمفردات المجتمع.

وبالتالي فإن P تمثل إحتمال ظهور هذه الظاهرة في المجتمع، ويرمز لاحتمال عدم ظهور هذه الظاهرة في المجتمع بالرمز q، حيث أن حدث ظهور الظاهرة وحدث عدم ظهورها هما حدثان لبعضهما البعض أي p+q=01.

وتعتبر النسبة P من أهم معالم المجتمع التي يرغب الباحث في معرفتها ليستطيع وصف المجتمع محل الدراسة وصفا جيدا، ولكن في الكثير من الأحيان لا نستطيع تحديد نسبة المجتمع لعدم توافر بيانات لدينا عن كل مفردات المجتمع، ولذلك نقوم بالاستدلال عليها، أي استنتاجها باستخدام نسبة الظاهرة محل الدراسة في العينة العشوائية المسحوبة من هذا المجتمع، ويرمز لنسبة العينة بالرمز $\hat{p}$ وتحسب كما يلي:

$\hat{\mathrm{p}} = \frac{x}{\mathrm{n}}$ / حيث: x تمثل عدد مفردات العينة التي تتحقق فيها الظاهرة المدروسة و n تمثل العدد الكلي لمفردات العينة.

نسبة العينة $\hat{p}$ ، كأي إحصائية تتغير قيمتها من عينة إلى أخرى وبالتالي فهي متغير عشوائي له توزيع احتمالي يطلق توزيع المعاينة لنسبة العينة، مما يعني أن سلوك هذه النسب هو سلوك متغير عشوائي، وبالتالي فإن لهذا المتغير توزيع احتمالي متوسطه:

$$\mu_{\hat{\mathrm{p}}} = E(\hat{p}) = p$$

أما تباينه فيكون حسب الحالات التالية:

	التوزيع الاحتمالي	تباين النسبة في العينة

P معلومة	$\hat{p} \to N\left(p, \sigma_{\hat{p}}^2\right)$	$\sigma_{\hat{p}}^2 = \frac{pq}{n}$
P مجهولة		$\sigma_{\hat{p}}^2 = \frac{\hat{p}\hat{q}}{n}$

حيث:

P: نسبة تواجد الصفة في المجتمع الإحصائي.

q: عدم توافر الصفة في المجتمع الإحصائي مع p=1-q.

$\hat{p}$: نسبة تواجد الصفة في العينة العشوائية.

$\hat{q}$: عدم توافر الصفة في العينة العشوائية مع $\hat{p} = 1 - \hat{q}$

وسوف نتطرق إلى مختلف حالات توزيع المعاينة للنسبة كما هو موضح في الجدول التالي:

توزيع النسبة للعينة ($\hat{p}$)		
	التوزيع الاحتمالي	حساب احتمال المتغير العشوائي
السحب بالارجاع و n≥30	$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$	$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$
السحب بدون ارجاع و n≥30	$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)\right)$	$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)}} \sim N(0,1)$

توزيع التباين للعينة (S^2)	
توزيع المعاينة (التوزيع الاحتمالي)	حساب احتمال المتغير العشوائي .
$c = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$	$c = \frac{(n-1)S^2}{\sigma^2}$

توزيع النسبة بين تباين عينتين $\left({S_1^2}/{S_2^2}\right)$	
توزيع المعاينة (التوزيع الاحتمالي)	حساب احتمال المتغير العشوائي .
$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - n_2 - 1)$	$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$

مثال1: مصنع ينتج عادة 25% عبوات كبيرة الحجم سحبت من إنتاجه عينة حجمها 2200 عبوة.

- اوجد توزيع النسبة للعبوات.

- احسب احتمال أن المصنع ينتج أقل من 22% من العبوات الكبيرة في فترة إجراء البحث في حالة السحب وبدون إرجاع علما أن حجم المجتمع N=8000.

مثال2: سحبت عينة عشوائية بحجم 20 مشاهدة من مجتمع طبيعي التوزيع تباينه يساوي 9

احسب احتمال أن يزيد تباين العينة عن 15.

مثال3: سحبت عينة حجمها 13 من مجتمع طبيعي تباينه 9 وسحبت عينة أخرى حجمها 21 من مجتمع طبيعي تباينه 25 مستقل عن المجتمع الأول.

- أوجد احتمال النسبة بين تبايني العينتين أقل من 0.8.

توزيع المعاينة للفرق بين متوسطي عينتين مستقلتين:

إذا كان لدينا المجتمع الأول $X_1 \to N(\mu_1, \sigma_1^2)$ وسحبت منه عينة عشوائية حجمها n_1، ومجتمع ثاني $X_2 \to$ $N(\mu_2, \sigma_2^2)$، سحبت منه عينة حجمها n_2، وكان المجتمعين مستقلين، وعرفنا المتغيرين $\bar{X}_1$ و $\bar{X}_2$فإن $\bar{X}_1 - \bar{X}_2$ يمثل متغيرات هعشوائيا للفرق بين متوسطي العينتين، وسطه الحسابي يكون كالآتي:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\overline{X_1}} - \mu_{\overline{X_2}} = \mu_1 - \mu_2$$

إذا تمت المعاينة من مجتمعين مستقلين، ولمعرفة توزيع المعاينة للفرق، فإن ذلك يتوقف على طبيعة توزيع المجتمعين (توزيع طبيعي أو غير طبيعي) وكذلك معلومية أو مجهولية تبايني المجتمعين وحجم العينتين المسحوبتين أكبر أو يساوي 30 أو أصغر تماما من 30 وسوف نتطرق إلى مختلف هذه الحالات كما هو موضح في الجدول التالي:

الحالة	توزيع المعاينة (التوزيع الاحتمالي)	حساب احتمال المتغير العشوائي
تبايني المجتمعين معلومين σ_1^2 و σ_2^2	$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ **التباين:** $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
تبايني المجتمعين مجهولين $\sigma_1^2 = ?$ و $\sigma_2^2 = ?$ و $n_1 \geq 30$ و $n_2 \geq 30$	$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})$ **التباين:** $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$
تبايني المجتمعين مجهولين $\sigma_1^2 = ?$ و $\sigma_2^2 = ?$ و $n_1 < 30$ و $n_2 < 30$	$\bar{X}_1 - \bar{X}_2 \sim t(n_1 + n_2 - 2)$ **التباين:** $\sigma^2_{\bar{X}_1 - \bar{X}_2} = S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$	$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim t(n_1 + n_2 - 2)$ **بحيث:** $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$

حيث:

n_1 و n_2: حجم العينيتن الأولى والثانية على التوالي

σ_1^2 و σ_2^2: تبايني المجتمعين على التوالي.

S_p^2 : الوسط المرجح لتباين العينة الأولى وتباين العينة الثانية، ويتم الترجيح بدرجة الحرية الخاصة بكل عينة، ويطلق على هذا المقدار التباين المشترك.

S_1^2 و S_2^2: تبايني العينتين على التوالي.

مثال:

إذا كان $X_1 \rightarrow N(9300, 250000)$، سحبت من مجتمعه عينة حجمها 25 مشاهدة.

و $X_2 \rightarrow N(9140, 160000)$ ، سحبت من مجتمعه عينة حجمها 20 مشاهدة.

المطلوب:

- أحسب احتمال أن يكون متوسط العينة الأولى أكبر من متوسط العينة الثانية بـ 395؟
- أحسب احتمال أن يكون متوسط العينة الأولى أقل من متوسط العينة الثانية؟

لدينا: $\sigma_1^2 = 250000$ ، $\mu_1 = 9300$ ، $n_1 = 25$

$\sigma_2^2 = 160000$ ، $\mu_2 = 9140$ ، $n_2 = 20$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(\overline{X}_1 - \overline{X}_2) - (9300 - 9140)}{\sqrt{\frac{250000}{25} + \frac{160000}{20}}} = \frac{(\overline{X}_1 - \overline{X}_2) - 160}{\sqrt{18000}}$$

أ إيجاد إحتمال أن يكون متوسط العينة الأولى أكبر من متوسط العينة الثانية بـ **395** أي إيجاد $p(\overline{X}_1 - \overline{X}_2 \succ 395)$

$$p(\overline{X}_1 - \overline{X}_2 \succ 395) = p\left(Z \succ \frac{395 - 160}{\sqrt{18000}}\right)$$

$$= p(Z \succ 1.75)$$

$$= 1 - \phi(1.75)$$

$$= 1 - 0.9599$$

$$= 0.0401$$

ب –إيجاد إحتمال أن يكون متوسط العينة الأولى أقل من متوسط العينة الثانية، أي إيجاد $p(\overline{X}_1 \prec \overline{X}_2)$

$$p(\overline{X}_1 \prec \overline{X}_2) = p(\overline{X}_1 - \overline{X}_2 \lhd 0)$$
$$= p\left(Z \prec \frac{0-160}{\sqrt{18000}} \right)$$
$$= p(Z \prec -1.19)$$
$$= 1-\phi(1.19)$$
$$= 1-0.8830$$
$$= 0.117$$

مثال 2: سحبت عينة عشوائية حجمها 17من مجتمع إحصائي وسطه الحسابي 32 وعينة عشوائية أخرى مستقلة عن الأولى حجمها 27 من مجتمع إحصائي وسطه الحسابي 30 وكان تباين العينة الأولى هو 6 وتباين العينة الثانية هو 9.

المطلوب: إذا كان تباين المجتمعين متساويين، فأوجد ما يلي:

- التوزيع الاحتمالي للفرق بين متوسطي العينتين؟
- احتمال أن يكون الفرق بين متوسطي العينتين أقل من 4؟

الحل:

لدينا: $n_1 = 17$ $\mu_1 = 32$ $S_1^2 = 6$

$n_2 = 27$ $\mu_2 = 30$ $S_2^2 = 9$

أ – إيجاد التوزيع الإحتمالي للفرق بين متوسطي العينتين:

بما أن تبايني المجتمعين مجهولين ومتساويين وحجم العينتين صغير فإننا نستخدم توزيع ستيودنت بدرجة حرية $v = n_1 + n_2 - 2$ ، أي:

$$\overline{X}_1 - \overline{X}_2 \rightarrow t_{(n_1+n_2-2)} \Rightarrow \overline{X}_1 - \overline{X}_2 \rightarrow t_{(17+27-2)}$$
$$\Rightarrow \overline{X}_1 - \overline{X}_2 \rightarrow t_{(40)}$$

ب –إيجاد إحتمال أن يكون الفرق بين متوسطي العينتين أقل من **4** أي إيجاد $p(\overline{X}_1 - \overline{X}_2 \prec 4)$

$$T = \frac{(\overline{X}_1 - \overline{X}_1) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث:

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-1}=\frac{(17-1)6+(27-1)9}{(17+27-2)}$$

$$\Rightarrow S_p^2=8.25$$

وعليه:

$$T=\frac{(\overline{X}_1-\overline{X}_1)-(32-30)}{\sqrt{(8.25)\left(\frac{1}{17}+\frac{1}{27}\right)}}=\frac{(\overline{X}_1-\overline{X}_1)-2}{\sqrt{(8.25)\left(\frac{1}{17}+\frac{1}{27}\right)}}$$

$$p\left(\overline{X}_1-\overline{X}_2\prec 4\right)=p\left(T\prec\frac{4-2}{\sqrt{0.791}}\right)$$

$$=p(T\prec 2.249)$$

$$=1-p(T\succ 2.249)$$

$$=1-0.01$$

$$p\left(\overline{X}_1-\overline{X}_2\prec 4\right)=0.99$$