ثالثا: توزيعات المعاينة

لنفرض أننا أخذنا من مجتمع ما عينة حجمها n، ثم حسبنا من هذه العينة بعض المقاييس الإحصائية مثل المتوسط الحسابي، التباين، النسبة... فإن كل مقياس إحصائي من هذه المقاييس يعتبر متغير عشوائي في حد ذاته يختلف من عينة لأخرى ويخضع لتوزيع معين، هذا التوزيع يسمى بتوزيع العينة، فمثلا نقول أن توزيع المعاينة للمتوسط الحسابي وهو عبارة عن توزيع جميع المتوسطات الحسابية للعينات المأخوذة من نفس هذا المجتمع ذات الحجم n، وكذلك فإن توزيع المعاينة للتباين هو توزيع جميع التباينات المحسوبة من عينات لها نفس الحجم n ومأخوذة من نفس المجتمع، وهكذا...

3-1- توزيع المعاينة للمتوسطات

3-1-1- توزيع المعاينة للمتوسط الحسابي $\bar{X}$ عند المعاينة من مجتمع طبيعي

إذا سحبنا عينة عشوائية وحسبنا متوسطها الحسابي سوف نجده قيمة ثابتة، أما إذا سحبنا عدة عينات (بعض العينات) بنفس الحجم من نفس المجتمع فمن المتوقع أن يأخذ المتوسط الحسابي قيما مختلفة في هذه العينات، أما إذا قمنا بسحب كل العينات الممكنة والتي لها نفس الحجم n وحسبنا المتوسطات الحسابية لهذه العينات، فإن المتوسط الحسابي العام $\bar{X}$ لمتوسطات هذه العينات يجب أن يساوي تماما الوسط الحسابي للمجتمع μ، ويسمى توزيع المتوسطات الحسابية بتوزيع المعاينة للمتوسط الحسابي.

فيما يلي سوف نذكّر بنظرية جدّ مهمة بدون برهان والمتعلقة بهذا التوزيع مع استعراض الحالات المختلفة وكيفية حساب الاحتمالات لهذا المتغير العشوائي.

نظرية (1- 9): توزيع المعاينة للمتوسط الحسابي $\bar{X}$ بالنسبة لعينات عشوائية ذات الحجم n مأخوذة من مجتمع يتوزع توزيعا طبيعيا متوسطه μ وتباينه σ_X^2 يكون مساويا لمتوسط المجتمع μ (سواء كان السحب بالإرجاع أو بدون إرجاع)، أما تباينه $\sigma_{\bar{X}}^2$ فيكون:

في حالة السحب بالإرجاع (بالإعادة)

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$$

في حالة السحب بدون إرجاع (بدون إعادة)

$$\sigma^2 = \frac{\sigma_X^2}{n} . \left(\frac{N-n}{N-1}\right)$$

ملاحظات:

- N يمثل حجم المجتمع وقد يكون محدود الحجم (منته) أو غير محدود (غير منته)، فإذا كان حجم المجتمع محدود وكان السحب بالإرجاع، فيعتبر المجتمع هنا غير منته أو غير محدود الحجم.
- القيمة $\left(\frac{N-n}{N-1}\right)$ تسمى بمعامل التصحيح أو الانتهاء.
- لمعرفة نوعية السحب نستعين بالقاعدة التالية:

سحب بدون إرجاع $\frac{\mathbf{n}}{\mathbf{N}} \geq \mathbf{0.05}$

سحب بالارجاع $\frac{\mathbf{n}}{\mathbf{N}} < \mathbf{0.05}$

سنتكلم الآن عن الحالات المختلفة لإيجاد توزيع المعاينة للمتوسط الحسابي $\bar{}$.

❖الحالة الأولى: تباين المجتمع الطبيعي معلوم (σ^2 معلوم)

عندما يكون مجتمع موزع توزيعا طبيعيا بمتوسط μ وتباين σ^2 معلوم أي: $X \sim N(\mu, \sigma^2)$ فإنه وبغض النظر عن حجم العينة n التي تسحب من هذا المجتمع، سوف يكون توزيع المتوسط الحسابي للعينة هو: $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

ولحساب الاحتمالات للمتغير العشوائي $\bar{}$ يتم إيجاد القيم المعيارية له (أي تحويل قيمته إلى Z المعيارية) حيث:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2 / n}} \sim N(0, 1)$$

مثال (1- 10):

إذا كان $X \sim N(12, 9)$ فأوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها **40** سحبت من المجتمع X، ثم أوجد قيمة الاحتمال $P(\bar{x} \leq 13)$.

الحل:

بما أنّ $X \sim N(12, 9)$ فإن $\bar{X} \sim N(12, \frac{9}{40})$

ولحساب قيمة الاحتمال $P(\bar{x} \leq 13)$ نحول $\bar{}$ إلى Z المعيارية: $Z = \frac{\bar{X} - 12}{\sqrt{9/40}}$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{x} \leq 13) = P\left(Z \leq \frac{13 - 12}{\sqrt{9/40}}\right) = P(Z \leq 2.1) = \emptyset(2.1) = 0.982136$$

❖ **الحالة الثانية: تباين المجتمع الطبيعي مجهول (σ^2 مجهول)**

في هذه الحالة نستبدل تباين المجتمع ^{2} بتباين العينة S^2 المسحوبة منه. ويجب أن نميز بين حالتين:

- **في حالة العينات الكبيرة** ($n \geq$):

يكون توزيع المتوسط الحسابي للعينة هو:

$$\bar{X} \sim N(\mu, \frac{S^2}{n})$$

ويكون التوزيع الطبيعي المعياري هو:

$$Z = \frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim N(0,1)$$

علما أنّ تباين العينة ^{2} يحسب بالعلاقة التالية: $S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$

مثال (1- 11):

إذا كان $X \sim N(12, \sigma^2)$ فأوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها **40** مشاهدة من المجتمع X وتباينها يساوي **16**، ثم أو جد قيمة الاحتمال $P(\bar{x} \leq 13)$.

الحل:

بما أن تباين المجتمع الطبيعي مجهول وحجم العينة كبير، فإن توزيع المتوسط الحسابي للعينة يتبع التوزيع الطبيعي ويكون:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\right) = N\left(12, \frac{16}{40}\right)$$

وبالتالي يكون التوزيع الطبيعي المعياري هو:

$$Z = \frac{\bar{X} - 12}{\sqrt{16/40}} \sim N(0,1)$$

أما الاحتمال المطلوب فيكون:

$$P(\bar{x} \leq 13) = P\left(Z \leq \frac{13-12}{\sqrt{16/40}}\right) = P(Z \leq 1.58) = \emptyset(1.58) = 0.942947$$

- **في حالة العينات الصغيرة** ($n < 30$):

في هذه الحالة توزيع المتوسط الحسابي للعينة سوف يتبع توزيع ستودنت بدرجة حرية $n-1$ (راجع توزيع ستودنت النظرية 1- 7):

$$T=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}\sim t(n-1)$$

مثال (1- 12):

إذا كان $X\sim N(10,\sigma^2)$ فأوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها 20 مشاهدة من المجتمع X وتباينها يساوي 9، ثم أو جد قيمة الاحتمال $P(\bar{x}\leq 11)$.

الحل:

بما أن تباين المجتمع الطبيعي مجهول وحجم العينة صغير، فإن توزيع المتوسط الحسابي للعينة يتبع توزيع ستودنت بدرجة حرية $(n-1=19)$:

$$T=\frac{\bar{X}-10}{\sqrt{9/20}}\sim t(19)$$

أما الاحتمال المطلوب فيكون:

$$P(\bar{x}\leq 11)=P\left(T\leq\frac{11-10}{\sqrt{9/20}}\right)=P(T\leq 1.5)=1-P(T>1.5)=0.95$$

❖ **الحالة الثالثة: حجم العينة كبير $(n\geq 30)$ والسحب بدون إرجاع من مجتمع محدود حجمه N.**

في هذه الحالة ندخل معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ على تباين العينة. ويجب أيضا أن نميز بين حالتين:

- **تباين المجتمع الطبيعي معلوم (σ^2 معلوم)**

في هذه الحالة يكون توزيع متوسط العينة هو:

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\frac{N-n}{N-1}\right)$$

أما التوزيع الطبيعي المعياري فيكون:

$$Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}\frac{N-n}{N-1}}}\sim N(0,1)$$

مثال (1- 13):

تم سحب عينة عشوائية بدون إرجاع حجمها **50** من مجتمع يتوزع توزيعا طبيعيا حجمه **500** بمتوسط $\mu = 20$ وتباين ${}^{2} = 25$. والمطلوب أوجد توزيع المتوسط الحسابي للعينة ثم أحسب $P(\bar{x} > 18)$.

الحل:

تباين المجتمع الطبيعي معلوم والسحب بدون إرجاع، وبذلك يكون توزيع المتوسط الحسابي للعينة هو:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\frac{N-n}{N-1}\right) = N\left(20, \frac{25}{50}\frac{500-50}{500-1}\right) = N(20, 0.45)$$

وبالتالي التوزيع المعياري يكون:

$$Z = \frac{\bar{X} - 20}{\sqrt{0.45}} \sim N(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x} > 18) = P\left(Z > \frac{18-20}{\sqrt{0.45}}\right) = P(Z > -2.98) = 1 - P(Z \leq -2.98) = 1 - \emptyset(-2.98)$$
$$= 1 - 0.001441 = 0.998559$$

- **تباين المجتمع الطبيعي مجهول (σ^2 مجهول)**

نستبدل تباين المجتمع بتباين العينة ^{2} ، وبالتالي يصبح توزيع متوسط العينة هو:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\frac{N-n}{N-1}\right)$$

أما التوزيع الطبيعي المعياري فيكون:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}\frac{N-n}{N-1}}} \sim N(0,1)$$

مثال (1- 14):

تم سحب عينة عشوائية بدون إرجاع حجمها **50** من مجتمع يتوزع توزيعا طبيعيا حجمه **500** بمتوسط $\mu = 20$. والمطلوب، إذا كان تباين العينة ${}^{2} = 36$ أوجد توزيع المتوسط الحسابي للعينة ثم أحسب $P(\bar{x} > 19.5)$.

الحل:

تباين المجتمع الطبيعي مجهول والسحب بدون إرجاع، نستبدل في هذه الحالة تباين المجتمع بتباين العينة، وبذلك يكون توزيع المتوسط الحسابي للعينة هو:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\frac{N-n}{N-1}\right) = N\left(20, \frac{36}{50}\frac{500-50}{500-1}\right) = N(20, 0.65)$$

وبالتالي التوزيع المعياري يكون:

$$Z = \frac{\bar{X}-20}{\sqrt{0.65}} \sim N(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x} > 19.5) = P\left(Z > \frac{19.5-20}{\sqrt{0.65}}\right) = P(Z > -1.94) = 1 - P(Z \leq -1.94) = 1 - \emptyset(-1.94)$$
$$= 1 - 0.026190 = 0.973810$$

3-1-2- توزيع المعاينة للمتوسط $\bar{}$ عند المعاينة من مجتمع غير طبيعي (نظرية النهاية المركزية)

ليس بالضرورة أن يكون المجتمع دائما موزعا توزيعا طبيعيا، فهناك مجتمعات مجهولة التوزيع (توزيعها غير معروف) وفي هذه الحالة نكون أمام التساؤل التالي:

إذا كان لدينا مجتمع متوسطه μ وتباينه σ^2، وأخذت منه جميع العينات العشوائية ذات الحجم n فما هو توزيع المتوسط الحسابي $\bar{X}$ حتى لو لم يكن المجتمع موزعا توزيعا طبيعيا؟

للإجابة على هذا السؤال، نستعين بنظرية النهاية المركزية (نظرية التقارب) بدون برهان.

نظرية (1- 10): إذا سحبت عينة حجمها n من مجتمع إحصائي متوسطه μ وتباينه σ^2، فإنه بغض النظر عن توزيع المجتمع يكون توزيع المتوسط الحسابي للعينة $\bar{X}$ هو بالتقريب يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu$ وتباين $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$ وبالتالي فإن المتغير Z له التوزيع الطبيعي المعياري:

$$Z = \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

ملا

إن هذه النظرية مهمة جدا لأنها تبين توزيع المعاينة للمتوسط الحسابي يتبع التوزيع الطبيعي إذا كان حجم العينة كبيرا ($n \geq 30$) بغض النظر عن شكل المجتمع. أما إذا كان حجم العينة صغيرا ($n < 30$) فالتقريب من التوزيع الطبيعي يكون مقبولا فقط عندما يكون توزيع المجتمع المسحوبة منه تلك العينة طبيعيا (تباين المجتمع معلوم).

مثال (1- 15):

إذا اخترنا $n = 100$ من مجتمع مجهول التوزيع بمتوسط $\mu = 40$ وتباين $^2 = 4$. فأوجد توزيع المتوسط

الحل:

بما أن $30 < \mathbf{n = 100}$ فإن توزيع المتوسط الحسابي للعينة $\bar{}$ يكون طبيعيا بمتوسط $\mathbf{40} = {}_{\bar{X}}$ وتباين $\boldsymbol{\sigma}^2_{\bar{X}} = \frac{\boldsymbol{\sigma}^2}{\mathbf{n}} = \frac{\mathbf{4}}{\mathbf{100}}$ وبالتالي:

$$\bar{\mathbf{X}} \sim \boldsymbol{N}\left(\boldsymbol{\mu}, \frac{\boldsymbol{\sigma}^2}{\boldsymbol{n}}\right) = \boldsymbol{N}(\mathbf{40}, \frac{\mathbf{4}}{\mathbf{100}})$$

أما التوزيع الطبيعي المعياري فيكون:

$$\boldsymbol{Z} = \frac{\bar{\boldsymbol{X}} - \boldsymbol{\mu}}{\sqrt{\boldsymbol{\sigma}^2/\boldsymbol{n}}} = \frac{\bar{\boldsymbol{X}} - \mathbf{40}}{\sqrt{\mathbf{4}/\mathbf{100}}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$$