

الباب 1

مراجعة الاحتمالات والمتغيرات العشوائية

1.1 مقدمة

كلمة (احتمال) هي كلمة دارجة ينطق بها الكثير من الناس، فمثلاً إذا سألت شخص ما عن شيء يريد ان ان يفعله وغير متأكد من ذلك فانه يقول احتمال افعل كذا. كذلك بعض خبراء الأرصاد الجوية حينما يتكلمون عن الطقس يقولون من المحتمل سقوط أمطار اليوم، احتمال ارتفاع في درجات الحرارة، وبعض خبراء البورصة يقولون احتمال ارتفاع قيمة الأسهم المتداولة في سوق المال لشركة معينة، خلال هذا اليوم، واحتمال نجاح طالب، واحتمال إصابة نوع معين من الفاكهة بنوع من البكتريا، وهكذا، يكثر نطق الأفراد بها وربما يجهلون معناها. فماذا تعني كلمة احتمال؟ يقصد بهذه الكلمة فرصة حدوث أو وقوع حادث معين، وتستخدم الاحتمالات في كثير من النواحي التطبيقية، مثل المجالات الاقتصادية، والتجارية، والزراعية، والطبية، والسلوكية، وغيرها، خاصة عند اتخاذ القرار في دراسات الجدوى، والتنبؤ بسلوك الظواهر المختلفة، ولكي يمكن فهم موضوع الاحتمال، وأهميته في النواحي التطبيقية، نقوم بعرض بعض المفاهيم الخاصة بالاحتمالات.

2.1 المتغير العشوائي: Random Variable

المتغير العشوائي عبارة عن دالة معرفة على فضاء العينة بحيث تنقل فضاء العينة الى مجموعة الأعداد الحقيقية أو جزء منها. أي ان المتغير العشوائي دالة مجال تعريفها فضاء العينة S ومداها مجموعة الأعداد الحقيقية أو جزء منها. ويرمز للمتغير العشوائي بشكل عام بحرف من الحروف الأبجدية الكبيرة $X, Y, Z, \dots$ ويرمز للقيم التي يأخذها المتغير بالحروف الأبجدية الصغيرة، $x, y, z, \dots$ ويمكن التعبير عن ذلك رياضياً كما يلي

$$X: S \rightarrow R$$

وينقسم المتغير العشوائي إلى قسمين هما:

1- المتغيرات العشوائية المنفصلة Discrete Random Variables

2- المتغيرات العشوائية المتصلة Continuous Random Variables

3.1 المتغيرات العشوائية المنفصلة

المتغيرات العشوائية المنفصلة أو المتقطعة هي المتغيرات القابلة للعد أو هي المتغيرات التي تكوين قيمها محدودة وتكون قابلة للعد أو هو الذي يأخذ قيم بينية، ومتباعدة. ومن أمثلة هذه المتغيرات:

1- عدد الأولاد الذكور في الأسرة المكونة من أربع أولاد

$$X = \{x = 0, 1, 2, 3, 4\}.$$

2- عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق

$$Y = \{y = 0, 1, 2, 3, \dots\}.$$

3- عدد مرات استخدام نوع معين من الأسمدة خلال الدورة الزراعية.

4- عدد الوحدات التالفة من إنتاج مزرعة معينة تنتج 200 وحدة كل موسم.

5- عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر.

1.3.1 التوزيع الاحتمالي للمتغير العشوائي المنفصل

المقصود بالتوزيع الاحتمالي للمتغير العشوائي المنفصل، هو التوزيع الذي يبين احتمالات حدوث القيم التي يمكن أخذها للمتغير، والتي ترتبط باحتمالات النتائج الممكنة في فراغ العينة، وبمعنى آخر هو التكراري النسبي للقيم التي يمكن أن يأخذها المتغير. فإذا كان المتغير العشوائي المنفصل X يأخذ القيم،

$$X = \{X_1, X_2, \dots, X_n\}$$

، وكان $P(X = X_i) = p_X(X_i)$ هو احتمال أن المتغير العشوائي يأخذ القيمة X_i ، فإنه، يمكن تكوين جدول التوزيع الاحتمالي للمتغير العشوائي X ، وهو جدول مكون من عمودين، الأول به القيم الممكنة للمتغير العشوائي $X = \{X_1, X_2, \dots, X_n\}$ ، والثاني به القيم الاحتمالية لهذا المتغير $P(X = X_i) = p(X_i)$ ، أي أن

الإحصاء الإداري ٢

جدول التوزيع الاحتمالي للمتغير العشوائي المنفصل

القيمة الاحتمالية للمتغير العشوائي	قيمة المتغير العشوائي
$p(X_1)$	X_1
$p(X_2)$	X_2
$\vdots$	$\vdots$
$p(X_n)$	X_n
1	Σ

وتسمى الدالة $p(x)$ دالة الكتلة الاحتمالية للمتغير العشوائي المنفصل وتعرف كما يلي:

$$p(x) = \begin{cases} Pr[X=x_i], & i = 1, 2, \dots \\ 0, & o.w. \end{cases}$$

ولابد لهذه الدالة ان تحقق الشروط الاتية:

$$0 \leq p(x) \leq 1, \forall x \quad (1)$$

$$\sum_x p(x) = 1 \quad (2)$$

حيث ان المجموع مأخوذ على جميع قيم x فى النطاق ويكون نطاق تعريف دالة الكتلة الاحتمالية $p(x)$ إما محدوداً أو يمكن عده.

2.3.1 بعض التوزيعات الاحتمالية المنفصلة الهامة

في كثير من النواحي التطبيقية، تتبع بعض الظواهر توزيعات احتمالية خاصة، وهي التوزيعات التي يمكن حساب احتمالات قيم المتغير عن طريق معادلة رياضية، تسمى بدالة الاحتمال $f(x)$ ، وهذه المعادلة لها معالم معينة، تسمى بمعالم المجتمع الذي ينسب له هذا التوزيع، وهذه المعالم ما هي إلا حقائق ثابتة مجهولة، وهي الأساس في حساب القيم الاحتمالية للتوزيع الاحتمالي للمجتمع محل الدراسة. ومن أهم التوزيعات التي سيتم دراستها في هذا المقرر، توزيع ثنائي الحدين، والتوزيع البواسون.

3.3.1 توزيع ذات الحدين

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح، والأخرى تسمى بحالة الفشل، ومن أمثلة ذلك:

- عند إعطاء مريض نوع معين من الأدوية، لها نتيجتان: (استجابة للدواء، أو عدم استجابة).
- عند فحص عبوة بداخلها نوع معين من الفاكهة، لها نتيجتان (الوحدة إما أن تكون سليمة، أو تكون معيبة).

- عند إلقاء قطعة عملة، لها نتيجتان (ظهور الوجه الذي يحمل الصورة، أو الوجه الذي يحمل الكتابة).

- نتيجة الطالب في الاختبار (نجاح، رسوب).

شكل التوزيع الاحتمالي ثنائي الحدين إذا كررت محاولة n من المرات، بحيث أن كل محاولة لها نتيجتان فقط متنافيتان هما: النتيجة محل الاهتمام " حالة نجاح " وتتم باحتمال ثابت في كل محاولة هو p النتيجة الأخرى "حالة فشل " وتتم باحتمال ثابت أيضا هو $q = 1 - p$ وبافتراض أن هذه المحاولات مستقلة، بمعنى أن نتيجة كل محاولة ليس لها علاقة بنتيجة المحاولة الأخرى، وإذا كان المتغير العشوائي X يعبر عن عدد حالات النجاح "عدد النتائج محل الاهتمام" في n محاولة، فإن مدي المتغير العشوائي X والذي يعبر عن عدد حالات النجاح هو:

$$X = \{0, 1, 2, \dots, n\}$$

، ومن ثم يحسب الاحتمال

$$P(X = x) = f(x)$$

بتطبيق المعادلة التالية

$$f(x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \dots, n \quad (1-1)$$

حيث أن $\binom{n}{x}$ هي عدد طرق اختيار x من n مع إهمال الترتيب، وتحسب كما يلي:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n(n-1)(n-2)\dots(n-x+1)}{x(x-1)(x-2)\dots(3)(2)(1)}.$$

خصائص التوزيع:

يمكن إيجاد خصائص التوزيع مباشرة من المسألة حتى قبل حساب التوزيع الاحتمالي، وهذه الخصائص وهذه الخصائص التالي:

$$(1) \text{ متوسط التوزيع هو } \mu = np.$$

$$(2) \text{ تباين التوزيع هو } \sigma^2 = npq.$$

$$(3) \text{ تباين التوزيع هو } \sigma = \sqrt{npq}.$$

“““

مثال (1-1)

إذا كان من المعلوم أن نسبة الشفاء من مرض معين باستخدام نوع معين من العقاقير الطبية هو

الإحصاء الإداري ٢

0.60 ، فإذا تناول هذا العقار 5 مصابين بهذا المرض. إذا عرف المتغير العشوائي X بأنه عدد الذين استجابوا (حالات الشفاء) لهذا العقار. المطلوب:

(1) ما هو نوع المتغير؟

(2) احسب الوسط الحسابي، والانحراف المعياري لعدد حالات الاستجابة.

(3) حدد شكل التوزيع.

(4) اكتب شكل دالة الاحتمال لهذا المتغير.

احسب الاحتمالات التالية:

(1) ما احتمال استجابة 3 مرضى لهذا العقار؟

(2) ما هو احتمال استجابة مريض واحد على الأقل؟

(3) ما هو احتمال استجابة 2 مرضى على الأكثر؟

الحل

عدد حالات الاستجابة X متغير كمي منفصل ، ومدى هذا المتغير في هذه الحالة هو:

$$X = \{0, 1, 2, 3, 4, 5\}$$

$$n = 5, p = 0.6, q = 1 - p = 1 - 0.4$$

شكل دالة الاحتمال:

$$\begin{aligned} f(x) &= \binom{n}{x} p^x q^{n-x}, \\ &= \binom{5}{x} (0.6)^x (0.4)^{5-x}, x = 0, 1, 2, 3, 4, 5 \end{aligned}$$

حساب الاحتمالات:

حساب احتمال عدم استجابة اي مريض مرضى لهذا الدواء:

$$\begin{aligned} P(X = 0) &= f(0) = \binom{5}{0} 0.6^0 0.4^{5-0}, \\ &= . \end{aligned}$$

حساب احتمال استجابة مريض واحد لهذا الدواء:

$$P(X = 1) = f(1) = \binom{5}{1} 0.6^1 0.4^{5-1},$$

$$=$$

حساب احتمال استجابة 2 مرضى لهذا الدواء:

$$P(X = 2) = f(2) = \binom{5}{2} 0.6^2 0.4^{5-2},$$

$$=$$

حساب احتمال استجابة 3 مرضى لهذا الدواء:

$$P(X = 3) = f(3) = \binom{5}{3} 0.6^3 0.4^{5-3},$$

$$= \frac{5(4)(3)}{3(2)(1)} (0.216)(0.16) = 0.3456.$$

حساب احتمال استجابة 4 مرضى لهذا الدواء:

$$P(X = 4) = f(4) = \binom{5}{4} 0.6^4 0.4^{5-4},$$

$$=$$

تحديد شكل التوزيع: يتحدد شكل التوزيع ثنائي الحدين وفقا لقيمة احتمال النجاح p كما يلي:
إذا كان $P = 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون متماثل. إذا كان $P > 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون موجب الالتواء. إذا كان $P < 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون سالب الالتواء. وحيث أن $P = 0.6 > 0.5$ فإن توزيع عدد حالات الاستجابة سالب الالتواء.
حساب احتمال استجابة 5 مرضى لهذا الدواء:

$$P(X = 5) = f(5) = \binom{5}{5} 0.6^5 0.4^{5-5},$$

$$=$$

حساب احتمال استجابة مريض واحد على الأقل:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - f(0) = 1 - \binom{5}{0} 0.6^0 0.4^{5-0} = 0.95976.$$

حساب احتمال استجابة 2 مريض على الأكثر:

$$P(X \leq 2) = f(0) + f(1) + f(2)$$

حساب الوسط الحسابي هو:

$$\mu = np = 5(0.6) = 3.0$$

تباين عدد حالات الاستجابة هو:

$$\sigma^2 = npq = 5(0.6)(0.4) = 1.2$$

ومن ثم يأخذ الانحراف المعياري الصورة التالية:

$$\sigma = \sqrt{npq} = \sqrt{5(0.6)(0.4)} = \sqrt{1.2}.$$

1.3.3.1 توزيع بواسون Poisson Distribution

هو توزيع منفصل يكثر استخدامه في الحالات التي تقع فيها الأحداث وفقا لمعدلات زمنية، وكذلك في حالة الأحداث نادرة الوقوع، ومن أمثلة ذلك:

(1) عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر.

(2) عدد مرات ري نوع معين من المحاصيل الزراعية خلال الموسم.

(3) عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق.

(4) عدد مرات زيارة المريض للطبيب كل سنة.

(5) عدد مرات تناول الأسرة للحوم الحمراء خلال الأسبوع.

(6) عدد أخطاء الطباعة لكل صفحة من صفحات الكتاب.

وهكذا الأمثلة كثيرة

شكل التوزيع الاحتمالي البواسوني إذا كان متوسط عدد مرات وقوع حادث وفقا لمعدل زمني معين هو λ ، وكان المتغير العشوائي X يعبر عن عدد مرات وقوع الحادث وفقا لهذا المعدل، فإن الدالة الاحتمالية للتوزيع تعطى في الصورة

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots$$

الإحصاء الإداري ٢

خصائص التوزيع:

يمكن إيجاد خصائص التوزيع مباشرة من المسألة حتى قبل حساب التوزيع الاحتمالي، وهذه الخصائص وهذه الخصائص التالي:

$$(1) \text{متوسط التوزيع هو } \mu = \lambda$$

$$(2) \text{تباين التوزيع هو } \sigma^2 = \lambda$$

$$(3) \text{تباين التوزيع هو } \sigma = \sqrt{\lambda}$$

مثال (1-2)

إذا كان من المعلوم أن عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر تتبع توزيع بواسون بمتوسط 3 وحدات شهريا، إذا عرف المتغير العشوائي X بأنه عدد الوحدات التي تستهلكها الأسرة خلال الشهر من هذه السلعة.

المطلوب:

- ما هو نوع المتغير العشوائي؟

- اكتب شكل دالة الاحتمال $f(x)$ لهذا المتغير. - احسب الاحتمالات التالية:

(1) احتمال أن الأسرة تستهلك وحدتين خلال الشهر؟

(2) احتمال أن أسرة ما تستهلك وحدة واحدة على الأقل خلال الشهر؟

(3) احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر؟

(4) احسب الوسط الحسابي، والانحراف المعياري لعدد الوحدات المستهلكة

(5) حدد شكل التوزيع

الحل

عدد الوحدات التي تستهلكها الأسرة X متغير كمي منفصل ، ومجال هذا المتغير في هذه الحالة هو $x = \{0, 1, 2, \dots\}$

بما أن متوسط عدد الوحدات التي تستهلكها الأسرة خلال الشهر هو $\lambda = 3$ ، إذا دالة الاحتمال هي

$$f(x) = \frac{e^{-3} 3^x}{x!}, x = 0, 1, 2, \dots$$

حساب الاحتمالات حساب احتمال أن أسرة ما تستهلك وحدتين خلال الشهر، $f(2)$

$$f(2) = \frac{e^{-3} 3^2}{2!} = 0.22404 >$$

الإحصاء الإداري ٢

احتمال أن أسرة ما تستهلك وحدة واحدة على الأقل خلال الشهر هو:

$$P(X \geq 1) = 1 - f(0) = 1 - \frac{e^{-3}3^0}{0!} = 1 - 0.5 = 0.5$$

احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر هو:

$$\begin{aligned} P(X \leq 3) &= f(0) + f(1) + f(2) + f(3) \\ &= \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} \\ &= 0.6474. \end{aligned}$$

4.3.1 المتغيرات العشوائية المتصلة

المتغير العشوائي المتصل، هو الذي يأخذ قيما متصلة، ويأخذ عدد لانهائي من القيم الممكنة له داخل مجاله، او هو المتغير الذي تأخذ قيمه مدى معين تتأرجح بداخله او هو المتغير الغير قابل للعد فإذا كان X متغير عشوائي مستمر، ويقع في المدى $(-\infty, \infty)$ ، أي أن: $X = \{x: -\infty \leq x \leq \infty\}$ ، فإن للمتغير X عدد لانهائي من القيم تقع بين الحدين الأدنى والأعلى $(-\infty, \infty)$ ، ومن الأمثلة على المتغيرات الكمية المتصلة ما يلي:

1- كمية الألبان التي تنتجها البقرة في اليوم باللتر:

$$X = \{x: 10 \leq x \leq 40\}$$

2- المساحة المنزرعة بالأعلاف في المملكة بالألف هكتار

$$X = \{x: 1000 \leq x \leq 1500\}$$

3- فترة صلاحية حفظ الاغذية المحفوظة بالأيام،

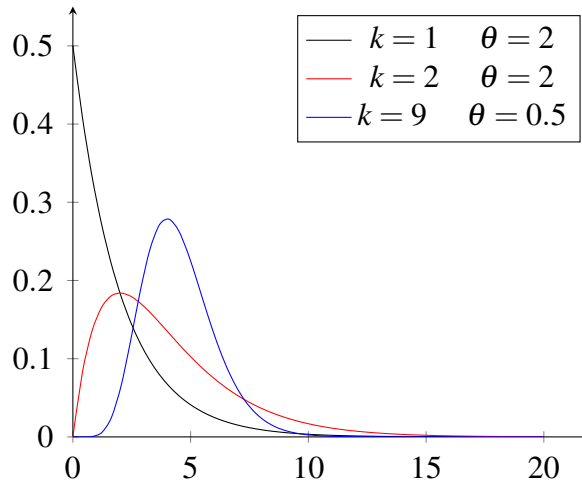
$$X = \{x: 1 \leq x \leq 4\}$$

4- وزن الجسم بالكيلوجرام للأعمار من (30 - 40)،

$$X = \{x: 55 \leq x \leq 80\}$$

5.3.1 التوزيع الاحتمالي للمتغير العشوائي المتصل

المقصود بالتوزيع الاحتمالي للمتغير العشوائي المتصل هو التوزيع الذي يبين احتمالات حدوث القيم التي يمكن أن يأخذها المتغير العشوائي داخل فترة معينة و عند تمثيل بيانات المتغير الكمي المستمر في شكل مدرج تكراري النسبي، نجد أن شكل هذا المدرج هو أقرب وصف لمنحنى التوزيع الاحتمالي للمتغير المستمر، وكلما ضاقت الفترات بين مراكز الفئات، يمكن الحصول على رسم دقيق للمنحنى الخاص بدالة احتمال المتغير المستمر، كما هو مبين بالشكل التالي:



والمساحة أسفل المنحنى تعبر عن مجموع الاحتمالات الكلية، ولذا تساوي هذه المساحة الواحد الصحيح، وتسمى الدالة $f(x)$ بدالة كثافة الاحتمال Probability Density Function(pdf)، وبفرض المتغير العشوائي المتصل يقع في المدى: $X = \{x : -\infty < x < \infty\}$ وتعرف كما يلي:

$$f(x) = \begin{cases} P[X=x], & -\infty \leq x \leq \infty \\ 0, & o.w. \end{cases}$$

ولابد لهذه الدالة ان تحقق الشروط الاتية:

$$1 - 0 \leq f(x) \leq 1$$

$$2 - \int_{-\infty}^{\infty} f(x) dx = 1 \quad (1-2)$$

ويكون نطاق تعريف دالة الكثافة الاحتمالية متصلا، ونقول إن المتغير العشوائي X في هذه الحالة متغير متصل.

مثال (3-1)

أوجد قيمة الثابت C لكي تكون هذه الدالة دالة كثافة احتمالية:

$$f(x) = \begin{cases} Cx, & 0 \leq x \leq 1 \\ 0, & o.w. \end{cases}$$

الحل

لكي تكون هذه الدالة دالة كثافة احتمالية يجب أن يكون

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

وبالتعويض عن قيمة الدالة $f(x)$ بالقيمة المعطاه في المثال نجد أن

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= \int_{-\infty}^{\infty} Cx dx \\ &= C \int_0^1 x dx \\ &= C \left. \frac{x^2}{2} \right|_0^1 = \frac{1}{2} \end{aligned}$$

ومن هنا نجد أن قيمة $C = 2$ وبالتعويض في الدالة المعطاه نجد أن دالة الكثافة الاحتمالية هي

$$f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & o.w. \end{cases}$$

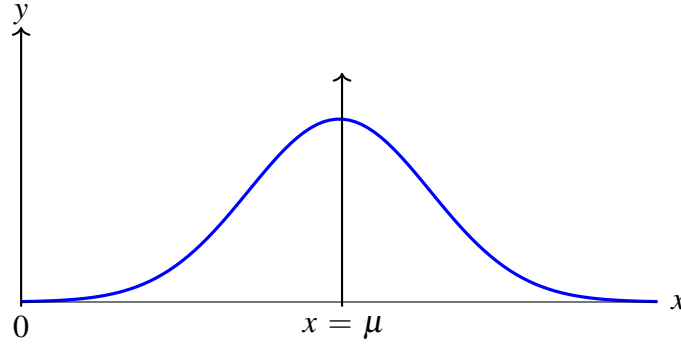
6.3.1 بعض التوزيعات الاحتمالية المتصلة الهامة

هناك بعض التوزيعات الاحتمالية المتصلة الخاصة، ولها دوال كثافة احتمال محددة، وفيما يلي بعض هذه التوزيعات:

1.6.3.1 التوزيع الطبيعي

يعتبر التوزيع الطبيعي من التوزيعات المتصلة المشهورة بل يعتبر من أهم التوزيعات الاحتمالية في علم الإحصاء لأنه يمثل كثيراً من الظواهر الطبيعية وغير الطبيعية في حياتنا اليومية مثل درجة الحرارة ورواتب الموظفين وضغط الدم وغيرها من الظواهر الطبيعية المتصلة (الغير قابلة

للعـد). منحنى التوزيع الطبيعي يشبه المنحني الجرسى او الناقوسى كما فى الشكل:



خصائص التوزيع الطبيعي سنقول إن المتغير العشوائى X يخضع للقانون المعتدل (أو قانون جاوس) بالبارامترين (μ, σ^2) ، إذا كانت دالة الكثافة الاحتمالية للمتغير العشوائى X

$$f(x) = \frac{1}{\sqrt{\pi} \sigma} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \quad -\infty < x < \infty, (-\infty < \mu < \infty, \sigma > 0) \quad (1-3)$$

يعتمد التوزيع الطبيعي على معلمتين أساسيتين هما المتوسط μ والتباين σ^2 ويرمز له بالرمز $N(\mu, \sigma^2)$ ومن أهم خصائص التوزيع الطبيعي ما يأتى

خصائص التوزيع الطبيعي

(1) متوسط التوزيع هو μ وتباين التوزيع هو σ^2 .

(2) التوزيع الطبيعي متماثل حول الخط المار بالمنتصف أى حول المتوسط الذى يقسم المنحنى الى قسمين متساويين أى ان الوسط يساوى الوسيط يساوى المنوال.

(3) المساحة الكلية تحت المنحنى فى التوزيع الطبيعي تساوى واحد صحيح.

(4) منحنى التوزيع الطبيعي يقترب طرفاه من محور x لكن لا يمسّه.

(5) 68% من المساحة تحت التوزيع الطبيعي تنحصر بين $\mu - \sigma, \mu + \sigma$.

(6) 95% من المساحة تحت التوزيع الطبيعي تنحصر بين $\mu - 2\sigma, \mu + 2\sigma$.

(7) 95% من المساحة تحت التوزيع الطبيعي تنحصر بين $\mu - 3\sigma, \mu + 3\sigma$.

2.6.3.1 التوزيع الطبيعي القياسى (المعيارى)

إذا كان متوسط التوزيع الطبيعي العادى $\mu = 0$ وتباينه $\sigma^2 = 1$ فإن التوزيع يسمى بالتوزيع الطبيعي القياسى (المعيارى) ويكتب فى الصورة $N(0, 1)$ وخواص التوزيع الطبيعي القياسى هي نفسها خواص التوزيع الطبيعي العادى الا ان التوزيع الطبيعي متماثل حول المتوسط و التوزيع الطبيعي

الإحصاء الإداري ٢

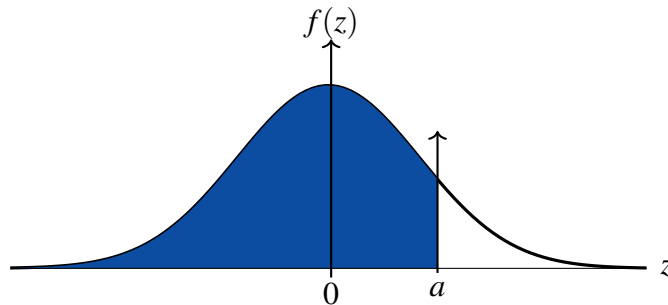
القياسي متمائل حول محور الصادات وبذلك يمكن القول بان التوزيع الطبيعي القياسي حالة خاصة من التوزيع الطبيعي. حساب المساحة تحت المنحنى في التوزيع الطبيعي

(1) منحنى التوزيع الطبيعي القياسي متمائل حول صفر وبالتالي ما يقال على النصف الموجب ينطبق تماماً على النصف السالب بسبب التماثل.

(2) يتم حساب الاحتمال (النسبة) في التوزيع الطبيعي بحساب المساحة وذلك من جداول التوزيع الطبيعي.

(3) جميع قيم الاحتمالات في الجدول كلها محسوبة على الصور

- (1) $P(-\infty \leq Z \leq a) = P(Z < a) = \phi(a)$
- (2) $P(a \leq Z \leq b) = P(a < Z < b) = \phi(b) - \phi(a)$
- (3) $P(Z \geq a) = P(Z > a) = 1 - \phi(a)$
- (4) $P(-\infty \leq Z \leq -a) = P(Z < -a) = 1 - \phi(a)$
- (5) $P(-a \leq Z \leq b) = P(-a < Z < b) = \phi(b) - 1 + \phi(a)$
- (6) $P(-a \leq Z \leq -b) = P(-a < Z < -b) = \phi(b) - \phi(a)$
- (7) $P(Z \geq -a) = P(Z > -a) = \phi(a)$



- (1) $P(Z < 0) = \phi(0) = 0.5$
- (2) $P(Z < 1.5) = \phi(1.5) = 0.9332$
- (3) $P(Z > 0.8) = 1 - \phi(0.8) = 1 - 0.7881$
- (4) $P(1 < Z < 0.8) = \phi(2.4) - \phi(1) = 0.9918 - 0.8413$
- (5) $P(-0.31 < Z < -2.85) = \phi(2.85) - \phi(0.31) =$
- (6) $P(-1.34 < Z < 3.4) = \phi(3.4) - 1 + \phi(1.34) =$
- (7) $P(Z > -1.85) = \phi(1.85) =$

3.6.3.1 الدرجة المعيارية

العلاقة بين التوزيع الطبيعي والتوزيع الطبيعي القياسي تعطى في الصورة

$$Z = \frac{X - \mu}{\sigma}$$

وتسمى هذه العلاقة بالدرجة المعيارية فعند استخدام هذه العلاقة فاننا نحول من التوزيع الطبيعي الى التوزيع الطبيعي القياسي.

مثال (1-5)

إذا كان الوسط الحسابي لمجموعة من المفردات $\mu = 50$ والانحراف المعياري $\sigma = 10$ فأوجد:

(1) الدرجة المعيارية المناظرة للدرجة الخام 60.

(2) الدرجة المعيارية المناظرة للدرجة الخام 45.

(3) الدرجة الخام المناظرة للدرجة المعيارية 1.5.

الحل

حسب المعطيات فإن $\mu = 50$ و $\sigma = 10$ والتعويض الذي نستخدم هنا هي

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 50}{10}$$

(1) الدرجة المعيارية المناظرة للدرجة الخام $X = 50$

$$Z = \frac{X - \mu}{\sigma} = \frac{60 - 50}{10} = 1$$

(2) الدرجة المعيارية المناظرة للدرجة الخام $X = 45$

$$Z = \frac{X - \mu}{\sigma} = \frac{45 - 50}{10} = -0.5$$

(3) الدرجة الخام المناظرة للدرجة المعيارية $Z = 1.5$

$$1.5 = \frac{X - \mu}{\sigma} = \frac{X - 50}{10}$$

اي ان

$$X = 15 + 50 = 65$$

مثال (1-6)

اذا كان دخل الف اسرة في مدينة ما، يتبع توزيعاً طبيعياً متوسطه $\mu = 1800$ ريال وانحرافه المعياري $\sigma = 300$ أوجد

(1) احتمال الحصول على دخل اسرة اكبر من 2400 ريال.

(2) احتمال الحصول على دخل اسرة اقل من 2550 ريال.

(3) احتمال الحصول على دخل اسرة ينحصر بين 1650 ريال وبين 2250 ريال.

(4) اوجد عدد الاسر التي يزيد دخلها عن 1500 ريال.

الحل

لايجاد هذه الاحتمالات لابد ان نحول من القيمة الخام الى القيمة المعيارية (القياسية) للتحويل نستخدم العلاقة الاتية:

$$Z = \frac{X - \mu}{\sigma}$$

حيث ان $\mu = 1800$ ، $\sigma = 300$

(1) احتمال الحصول على دخل اسرة اكبر من 2400 ريال.

$$Z = \frac{2400 - 1800}{300} = 2$$

$$P(X > 2400) = P(Z > 2) = 1 - P(Z < 2) = 1 - 0.9772 = 0.0228.$$

(2) احتمال الحصول على دخل اسرة اقل من 2550 ريال.

$$Z = \frac{2550 - 1800}{300} = 1.5$$

$$P(X < 2550) = P(Z < 1.5) = 0.9332.$$

(3) احتمال الحصول على دخل اسرة ينحصر بين 1650 ريال وبين 2250 ريال.

$$Z_1 = \frac{1650 - 1800}{300} = -0.5, \quad Z_2 = \frac{2250 - 1800}{300} = 1.5$$

$$P(1650 < X < 2250) = P(-0.5 < Z < 1.5)$$

$$= P(Z < 1.5) - P(Z < -0.5) = 0.9332 - 0.3085 = 0.6247.$$

(4) اوجد عدد الاسر التي يزيد دخلها عن 1500 ريال. اولا نوجد احتمال دخل اسرة يزيد على 1500

$$Z = \frac{1500 - 1800}{300} = -1$$

$$P(X > 1500) = P(Z > -1) = 0.8413.$$

وحيث ان العدد الكلي للاسر في العينة هو $n = 1000$ لذلك فان عدد الاسر التي يزيد دخلها عن 1500 هو $1000(0.8413) = 841$ اسرة.

4.6.3.1 توزيع t

وهو معروف بتوزيع t للطالب (*Student's distribution*) حيث نشره تحت اسم طالب البريطاني وليم جوست *William Gosset* في عام 1908. وأصبحت له تطبيقات عديدة في الإحصاء، ولعل شهرته تحت اسم توزيع t للطالب، أكثر من شهرة صاحبه، وربما يجهل الكثيرون اسم صاحب هذا التوزيع. وتكتب دالة الكثافة الاحتمالية لتوزيع t في الصورة

$$f(t) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi v}} \left(1 + \frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, \quad v > 2 \quad (1-4)$$

7.3.1 خواص توزيع t

(1) توزيع t متماثل حول الخط المار بالمنتصف اي حول المتوسط والذي يساوي صفر ويقسم المنحنى الى قسمين متساويين.

(2) منحنى توزيع t يشبه الى حد ما منحنى التوزيع الطبيعي الا انه اكثر تفرطحاً.

(3) المساحة الكلية تحت المنحنى في توزيع t تساوي واحد صحيح.

الإحصاء الإداري ٢

(4) منحنى توزيع t يقترب طرفاه من محور x لكن لا يمسه.

(5) يتم حساب الاحتمال (النسبة) في توزيع t بحساب المساحة وذلك من الجداول .

(6) يرمز أحيانا لتوزيع t بالرمز $t(1 - \alpha, v)$ حيث $1 - \alpha$ هي المساحة المطلوبة و $v = n - 1$ تسمى درجة الحرية و أحيانا يرمز لتوزيع t بالرمز $t_\alpha(v)$.

مثال (7-1)

اوجد كلا مما يأتي:

$$t_{0.01}(7), \quad t_{0.05}(7), \quad t_{0.01}(8), \quad t_{0.1}(8)$$

الحل

$$t_{0.01}(7) = 2.998, \quad t_{0.05}(7) = 1.895$$

$$t_{0.01}(8) = 2.896, \quad t_{0.1}(8) = 1.397$$

مثال (8-1)

اوجد قيمة α في كلا مما يأتي:

$$t_\alpha(10) = 1.372, \quad t_\alpha(5) = 4.032, \quad t_\alpha(7) = 1.415, \quad t_\alpha(8) = 2.896.$$

الحل

$$t_\alpha(10) = 1.372, \quad \alpha = 0.10$$

$$t_\alpha(5) = 4.032, \quad \alpha = 0.005$$

$$t_\alpha(7) = 1.415, \quad \alpha = 0.1$$

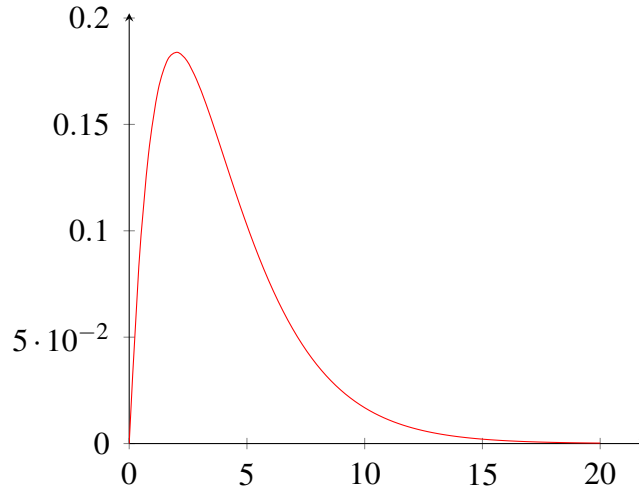
$$t_\alpha(8) = 2.896, \quad \alpha = 0.01.$$

1.7.3.1 توزيع مربع كاي χ^2 distribution

يعتبر من التوزيعات الاحتمالية المتصلة التي لها اهمية كبيرة في التطبيقات الاحصائية ومنحنى هذا التوزيع غير متماثل بل هو منحنى ملتو التواء موجب أي نحو جهة اليمين . الدالة الاحتمالية التي تصف هذا التوزيع هي

$$f(\chi^2) = \frac{1}{(\frac{v}{2} - 1)! 2^{\frac{v}{2}}} e^{-\frac{1}{2}\chi^2} (\chi^2)^{\frac{v}{2}-1}; \quad 0 < \chi^2 < \infty \quad (1-5)$$

حيث ان ν (تمثل معلمة التوزيع ويطلق عليها) درجات الحرية للتوزيع) ويشار للتوزيع بشكل مختصر كما يلي $\chi^2 \sim \chi^2_\nu$.



خواص توزيع χ^2

- (1) قيمة المتغير العشوائي هنا هي كمية موجبة دائما .
- (2) الوسط الحسابي لقيم المتغير العشوائي في هذا التوزيع هي (ν) .
- (3) ان تباين قيم المتغير العشوائي في هذا التوزيع هو (2ν) .
- (4) يمكن الاعتماد على جداول خاصة بهذا التوزيع في ايجاد قيم مربع كاي النظرية وعند مستوى معنوية محدد ودرجة حرية معينة.
- (5) لإيجاد قيمة الاحتمال المقابل لقيمة معينة وحسب جدول مربع كاي يجب ان تكون الصيغة كالآتي

$$P(\chi^2 > \chi^2_{(\nu, \alpha)}) = \alpha$$

حيث α تمثل قيمة الاحتمال.

مثال (9-1)

إذا كان X متغير عشوائي يتبع توزيع مربع كاي بدرجة حرية $\nu = 10$ اوجد الوسط الحسابي والتباين ثم اوجد

$$P(\chi^2 > 12.55), \quad P(\chi^2 < 20.48), \quad P(15.99 < \chi^2 < 29.59).$$

الحل

حيث ان الوسط الحسابي لتوزيع مربع كاي يمثل درجة الحرية فإن $\mu = \nu = 10$ والتباين

$$\sigma^2 = 2v = 2(10) = 20$$

$$P(\chi^2 > 12.55) = 0.25$$

$$P(\chi^2 < 20.48) = 1 - 0.025 = 0.975$$

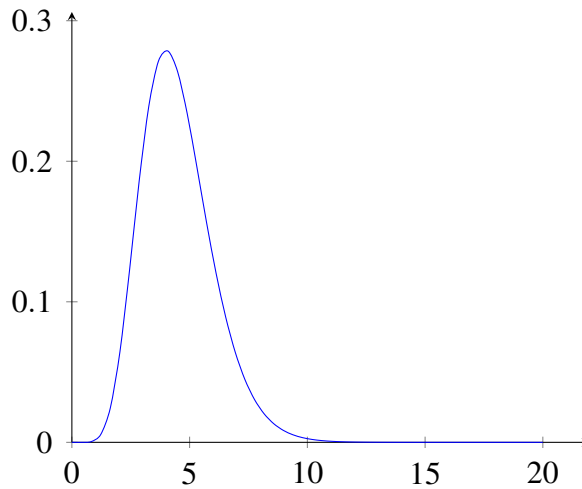
$$P(15.99 < \chi^2 < 29.59) = 0.100 - 0.001 = 0.099.$$

2.7.3.1 توزيع F

ان توزيع F مشتق من توزيعين مستقلين كل منهما يمثل توزيع مربع كاي وعليه يعتبر احد توزيعات المعاينة ويستند بالأساس على التوزيع الطبيعي وله استخدامات كثيرة في المجالات الاحصائية التطبيقية وبالذات في مجال تصميم وتحليل التجارب . فعلى افتراض ان هناك متغيرين الاول هو $\chi_1^2 \sim \chi_{v_1}^2$ والثاني $\chi_2^2 \sim \chi_{v_2}^2$ فإن F يعبر عنها بالشكل التالي

$$F = \frac{\chi_1^2/v_1}{\chi_2^2/v_2} \sim F(v_1, v_2)$$

اي ان المتغير العشوائي (F) يسلك وفق توزيع F بدرجة حرية هي v_1 و v_2 للبسط والمقام على التوالي.



خواص توزيع F

(1) قيمة المتغير العشوائي هنا هي كمية موجبة دائما .

(2) الوسط الحسابي لقيم المتغير العشوائي في هذا التوزيع هي $(\frac{v_2}{v_2 - 2})$ بشرط ان $v_2 > 2$.

(3) ان تباين قيم المتغير العشوائي في هذا التوزيع هو

$$\frac{2 v_2^2 (v_1 + v_2 - 2)}{v_1 (v_2 - 2)^2 (v_2 - 4)}$$

بشرط ان $v_2 > 4$.

(4) يمكن الاعتماد على جداول خاصة بهذا التوزيع للحصول على قيم نظرية وذلك عند مستوى معنوية محدد ودرجة حرية معينة للبسط ودرجة حرية معينة للمقام .

$$F(\alpha, v_1, v_2) = \frac{1}{F(1-\alpha, v_2, v_1)} \quad (5)$$

مثال (10-1)

اوجد قيمة كلاً مما يأتي

$$F(0.1, 7, 9), \quad F(0.05, 10, 15), \quad F(0.01, 8, 12).$$

الحل

$$F(0.1, 7, 9) = 2.88, \quad F(0.01, 8, 12) = 4.5, \quad F(0.05, 10, 15) = 2.54.$$

مثال (11-1)

اوجد قيمة كلاً مما يأتي

$$F(0.99, 5, 7) \quad F(0.99, 8, 12), \quad F(0.95, 10, 15).$$

الحل

$$F(0.99, 5, 7) = 7.46, \quad F(0.99, 8, 12) = 4.5, \quad F(0.95, 10, 15) = 2.554.$$

الباب 2

توزيعات المعاينة الاحصائية

1.2 مقدمة

الاستدلال الإحصائي هو احد وظائف علم الإحصاء، وتظهر اهميته في كثير من النواحي التطبيقية التي يكون فيها معالم توزيع الظاهرة في المجتمع محل الدراسة غير معلومة، او النواحي التي يكون فيها التوزيع ذا معالم معروفة ويحتاج الباحث إلى اتخاذ قرارات حول صحة قيم هذه المعالم وذلك في الوقت الذي يصعب فيه إجراء حصر شامل لجميع مفردات المجتمع من اجل إمكانية حساب القيم الحقيقية لهذه المعالم. لذا يقوم الباحث بسحب عينة عشوائية من المجتمع محل الدراسة من اجل دراسة خصائص هذا المجتمع.

1.1.2 مفهوم المجتمع الاحصائي

المجتمع الإحصائي هو مجموعة من القراءات أو المفردات أو المشاهدات تتصف بخاصية معينة او بمجموعة من الخواص تصف هذا المجتمع. ونظراً لأن أية خاصية تختلف قيمتها من مشاهدة لآخرى لذا تسمى هذه القيمة متغيراً إحصائياً. فمثلاً دخل مجموعة من الأسر أو درجات الطلاب في مقرر دراسي او عينة من سكان مدينة ما اعمار ... الخ.

2.1.2 مفهوم العينة الإحصائية

هي جزء الذي يتم اختياره من المجتمع بحيث يمثل المجتمع الأصلي ويحقق أغراض البحث ويغني الباحث عن مشقات دراسة المجتمع الأصلي.

3.1.2 أنواع العينات الإحصائية

تتمثل فيما يلي:

4.1.2 العينات (العشوائية) الاحتمالية:

تقسم الى عدة أنواع تتمثل فيما يلي:

(1) العينة العشوائية البسيطة:

هي مجموعة جزئية من المجتمع الاحصائي لها نفس الفرصة لتُختار كعينة من ذلك المجتمع، أي بمعنى أن جميع أفراد المجتمع لهم فرصة في أن يُختاروا، ويرجع ذلك إلى أن المجتمع متجانس إذا اختيرت منه عينة وبأي طريقة تستطيع تمثيله وتظهر فيها جميع خصائصه وسماته.

(2) العينة العشوائية الطبقية:

المعينة الطبقية يُلجأ إليها في حالة معرفة التركيب النسبي للمجتمع الأصلي عندما يكون هذا المجتمع مكونا من عدة طبقات بينها اختلاف واضح من حيث أحد أو مجموعة من الخصائص فتختار طريقة العينة الطبقية حرصا على أن تمثل جميع تلك الطبقات في العينة المختارة وعادة تكون العينة الطبقية متباينة فيما بينها ومتجانسة في داخلها، مثال ذلك، سوق ملابس به عدة أقسام: قسم الأطفال، قسم الصبيان، قسم الرجال، قسم النساء وغيرها، فهذه الأقسام هي عبارة عن طبقات.

(3) العينة العشوائية العنقودية:

وهي تختلف عن المعينة الطبقية في مبدأ العناقيد الذي يحدد أن تكون العناقيد متباينة في داخلها متجانسة فيما بينها أي عكس العينة الطبقية. لنأخذ نفس المثال في العينة الطبقية لكن هنا يكون شكل السوق بدون أقسام أي جميع الملابس توجد في محل واحد به الأطفال، الصبيان، الرجال النساء، وهذا ما نعني به متباينة في داخلها. أما متجانسة فيما بينها كأن تكون هنالك عدة أسواق بهذا الشكل. وبالتالي يمكنك أن تأخذ جميع أغراضك من محل واحد. وهذا ما يحدث في حالة العينة العنقودية عنقود واحد تجد فيه جميع أفراد المجتمع ولا تحتاج أن تذهب لكل العناقيد أي يمكنك الاستغناء عن البقية لأنها تحمل نفس الخصائص وهذا لا يحدث في العينة الطبقية حيث تقسم الطبقات على أساس خاصية واحدة محددة لا تتوفر في الطبقات الأخرى لذا لا بد عليك المرور على كل الطبقات (الأقسام) لتجد كل ما تحتاج إليه ولا تستطيع أن تستغني عن أي طبقة أو قسم، حسب المثال.

(4) العينة العشوائية المنتظمة:

العينة المنتظمة يكون اختيار الوحدات منها على أساس تقسيم العدد الكلي للمجتمع على حجم العينة المطلوبة، ومن ثم توزيع وحدات المجتمع الأصلي وبشكل متساوٍ ومنتظم على الرقم الناتج من ذلك التقسيم مثلا: إذا كان العدد الكلي للمجتمع هو (3000) طالب وطالبة مثلا وهو رقم يمثل عدد الطلبة في كلية ما، وكانت العينة المطلوبة هي (150) طالب وطالبة فقط فيكون

الإحصاء الإداري ٢

توزيع الوحدات. الكلية الأصلية للمجتمع على الشكل الآتي $20 = 3000 \div 150$ وعلى هذا الأساس يتحدد اسم الطالب الأول بحيث يكون أقل من الرقم (20) وليكن (3) مثلاً ويختار عشوائياً، ثم يبدأ الباحث بتوزيع العينة على بقية الأسماء بالشكل الآتي: أول رقم هو (3)، أما الرقم الثاني فهو $(20 + 3 = 23)$. والثالث هو 43 ثم 63 ثم 83 وهكذا حتى نصل إلى 2883 وبهذا المنطق أعطينا فرصة لكل فرد من أفراد المجتمع المتمثل بما مجموعه (3000) طالب وطالبة أن يكونوا ضمن أفراد العينة وبشكل منتظم.

5.1.2 العينات غير الاحتمالية:

تقسم إلى عدة أنواع تتمثل فيما يلي

(1) العينة العارضة:

يتجه الباحث في هذا النوع من العينات إلى اختيار الحالات التي تصادف الباحث مثل الأشخاص الذين يقابلهم الباحث قبل غيرهم في الطريق.

(2) العينة العمدية (القصدية):

يختار الباحث في هذا النوع من العينات، حالات يعتقد أنها تمثل المجتمع في الجانب الذي يتناوله البحث، كأن يختار الباحث منطقة حسب اعتقاده هي الأكثر ملائمة للقيام بالبحث فيها. وتوفر هذه الطريقة على الباحث الكثير من الوقت والجهد الذي يبذله في اختيار العينة إلا أنها تستلزم معرفة المعالم الاحصائية بالنسبة للمجتمع الأصلي خاصة بالنسبة للوحدات التي يرغب الباحث في اختيارها وهو أمر لا ييسر في كل الأحوال.

(3) العينة الحصصية:

يكثر استخدام هذا النوع خاصة في استطلاعات الرأي العام لما يتميز به من سرعة حيث يقسم الباحث المجتمع إلى طبقات أو فئات حسب خصائص معينة ويعمل على تمثيل كل فئة من فئات العينة بالنسبة لوجودها في المجتمع، ويترك للباحث حرية اختيار مفردات الحصة بشرط أن يلتزم بالحدود العددية أو النوعية للعينة، إنما يخشى منه عدم تمثيل المجتمع الأصلي تمثيلاً صحيحاً.

6.1.2 مفهوم المعاينة

هو عمل احصائي منظم مبني على اسس علمية ويقوم على مبدأ شمول جزء من المجتمع الاحصائي وتختار المفردات في الغالب باعتماد أحد اساليب المعاينة الاحتمالية، أو شمول جميع وحدات المجتمع وإخضاعها للملاحظة من خلال المسح الشامل، والمعاينة أيضاً هو اختيار جزء من وحدات المجتمع بطرق معينة تضمن تمثيل المجتمع الأصلي بجميع وحداته خير تمثيل ويسمى الجزء المأخوذ بالعينة.

7.1.2 توزيعات المعاينة

توزيع المعاينة للإحصائية هو التوزيع الاحتمالي لهذه الإحصائية المحسوب لكل العينات الممكنة والمأخوذة من المجتمع الإحصائي المدروس أيا كان حجمه وأيا كانت طريقة السحب. إذا أخذنا عينات متكررة من مجتمع ما وقمنا بقياس متوسط لكل عينة، فإننا نجد أن معظم هذه المتوسطات تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات بتوزيع المعاينة للمتوسط الحسابي، ولكن توزيع المعاينة للمتوسط الحسابي له أيضا متوسط حسابي يعبر عنه بالرمز $\mu_{\bar{X}}$ وانحراف معياري (خطأ معياري) يعبر عنه بالرمز $\sigma_{\bar{X}}$. فتوزيع المعاينة للمتوسطات الحسابية هو عبارة عن التوزيع الاحتمالي للمتوسطات الحسابية لعدد كبير من العينات العشوائية المتساوية الحجم ومن مجتمع إحصائي واحد.

8.1.2 معالم المجتمع

معلمة المجتمع هي عبارة عن خاصية أو مقياس يتم حسابها من المجتمع محل الدراسة. أي أن المعالم هي مقاييس تحدد خصائص المجتمع (التوزيع). أهم معالم المجتمع هي: المتوسط الحسابي للمجتمع يساوي

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

التباين يساوي

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

الانحراف المعياري (الجذر التربيعي للتباين) يساوي

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

النسبة تساوي: $P = \frac{X}{N}$ حيث X عدد المفردات التي تتوافر فيهم الخاصية محل الدراسة.

9.1.2 إحصاءات العينة

إحصاءة العينة هي عبارة عن خاصية أو مقياس يتم حسابها من العينة المسحوبة من المجتمع محل الدراسة. أي أن الإحصاءة هي دالة في بيانات العينة. أهم إحصاءات العينة هي:

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

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

الانحراف المعياري (الجذر التربيعي للتباين) يساوي

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

النسبة تساوي: $\bar{P} = \frac{X}{n}$ حيث X عدد المفردات التي تتوافر فيهم الخاصية محل الدراسة.

10.1.2 أنواع طرق المعاينة

تنقسم المعاينة حسب طرق الحصول عليها الى قسمين هما:

1.10.1.2 المعاينة مع الارجاع

هي المعاينة التي يتم فيها اختيار كل عنصر من المجتمع أكثر من مرة، ويمكن اعتبار هذا المجتمع الذي تتم فيه المعاينة بالإرجاع مجتمع غير منته أو مجتمع غير محدود طالما انه يمكن أن نسحب منه عينة أيا كان حجمها دون استنفاد المجتمع. وهناك قاعدة لمعرفة هل المعاينة تمت بالإرجاع أو بدون الارجاع وهي ان المجتمع غير محدود وتمن المعاينة بالارجاع اذا كان $n < 0.05N$ ، حيث N حجم المجتمع و n حجم العينة. إذا كان السحب بالإرجاع عدد العينات يساوي N^n .

مثال (1-2)

لنفترض ان لدينا مجتمع حجمه 3 سحبت كل العينات الممكنة والتي حجمها 2 فما هو عدد العينات الممكنة اذا كان السحب مع الارجاع. عدد العينات يساوي: $N^n = 3^2 = 9$.

2.10.1.2 المعاينة بدون إرجاع:

هي المعاينة التي يتم فيها اختيار كل عنصر من المجتمع مرة واحدة فقط، ويمكن اعتبار هذا المجتمع الذي تمت فيه المعاينة بدون إرجاع بالمجتمع المنته أو المجتمع المحدود. : المجتمع محدود أي أن المعاينة بدون إرجاع اذا تحقق ما يل $n \geq 0.05N$.
نفترض أن لدينا مجتمع حجمه مفردة هي:

$$X_1, X_2, \dots, X_N$$

نريد سحب عينة من هذا المجتمع حجمها n ومفرداتها:

$$X_1, X_2, \dots, X_n$$

عدد العينات ذات الحجم n والتي يمكن اختيارها من المجتمع ذات الحجم N هو

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

مثال (2-2)

إذا كان لدينا مجتمع حجمه 5 مفردات ونريد أخذ عينات بدون إرجاع حجمها 3 مفردات فإن عدد العينات الممكنة:

$$\binom{N}{n} = \frac{N!}{n(N-n)!} = \binom{5}{3} = \frac{5!}{3!(2)!} = 10.$$

مثال (3-2)

مجتمع مكون من عشرة مفردات ، اختيرت منه عينات حجم كل منها 3 مفردات ، وحسب من كل عينة مقياس إحصائي معين (وليكن الوسط الحسابي) . فكم مشاهدة يتكون مجتمع هذا المقياس ؟
الحل

عدد مفردات المجتمع الأصلي $N = 10$

عدد مفردات العينة $n = 3$

وبالتالي فإن : عدد العينات التي يمكن اختيارها مع الإرجاع هو :

$$N^n = 10^3 = 1000$$

وعدد العينات التي يمكن اختيارها بدون إرجاع هو :

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = 120$$

وحيث إننا نحسب المقياس لكل عينة فإن مجتمع هذا المقياس يتكون من 1000 مشاهدة في حالة الاختيار مع الإعادة أو يتكون من 120 مشاهدة في حالة الاختيار بدون إعادة وبالمقارنة نجد أن عدد مفردات مجتمع المقياس أكبر بكثير من عدد مفردات المجتمع الأصلي.

2.2 توزيعات المعاينة للمتوسطات

إذا ما أخذنا عدداً من العينات المختارة عشوائياً من مجتمع ما وحسبنا المتوسط الحسابي لكل عينة فإننا نحصل على قيمة واحدة من كل عينة وبذلك يكون لدينا مجموعة من القيم للمتوسطات الحسابية. فإذا كان عدد العينات كبيراً فإنه يمكننا أن نضع هذه القيم في جدول تكراري يمكن تمثيله بمنحنى مستمر ويسمى هذا التوزيع بالتوزيع العيني. وإذا ما اختيرت العينات بطريقة عشوائية وبعبدة عن التحيز فإننا نجد أن القيم المستخرجة من العينات المختلفة تتجمع حول القيمة الحقيقية لمتوسط المجتمع. ومن ثم فإنه يمكن أن نأخذ قيمة محسوبة من عينة كتقدير للقيمة الحقيقية للمجتمع. وعلى ذلك فإنه يمكن اعتبار المتوسط الحسابي للعينة كتقدير للمتوسط الحسابي للمجتمع وكذلك يمكن اعتبار تباين العينة وانحرافها المعياري كتقديرين لتباين المجتمع وانحرافه المعياري.

3.2 نظرية النهاية المركزية

إذا كان لدينا مجتمع غير محدود مفرداته X يتبع توزيعاً طبيعياً متوسطه μ وانحرافه المعياري σ . وسحبنا من هذا المجتمع عينات عشوائية حجم كل منها $n \geq 30$. فإن الوسط الحسابي لهذه العينات $\bar{X}$ يتبع توزيعاً طبيعياً له الخصائص التالية:

$$\begin{aligned} E(\bar{X}) &= \mu_{\bar{X}} = \mu \\ V(\bar{X}) &= \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \end{aligned} \quad (2-1)$$

إذا كان X_i متغير عشوائي يخضع للتوزيع الطبيعي متوسطه μ وتباينه σ^2 وكان $\bar{X}$ يمثل المتوسط الحسابي للعينات ذات الحجم n والمسحوبة من مجتمع محدود او منته (المعاينة مع عدم الارجاع) فإن توزيع المعاينة للمتوسطات الحسابية له الخصائص التالية:

$$\begin{aligned} \mu_{\bar{X}} &= \mu \\ \sigma_{\bar{X}}^2 &= \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \\ \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}, \end{aligned}$$

حيث $\frac{N-n}{N-1}$ تمثل معامل التصحيح.

مثال (4-2)

بافتراض أن لدينا مجتمع حجمه 4 مفردات هي 2, 5, 3, 6 وأردنا سحب عينة من هذا المجتمع حجمها

مفردتين أو وجد:

(1) المتوسط الحسابي والتباين للمجتمع.

(2) المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية.

(3) التباين لتوزيع المعاينة للمتوسطات الحسابية.

الحل

(1) حساب المتوسط الحسابي للمجتمع

$$\mu = \frac{\sum_{i=1}^N X_i}{N} = \frac{16}{4} = 4.$$

(2) حساب تباين المجتمع

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} = 2.5.$$

(3) حساب المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية

$$\mu_{\bar{X}} = \mu = 4.$$

حساب التباين لتوزيع المعاينة للمتوسطات الحسابية لدينا: $\frac{n}{N} = \frac{2}{4} = 0.5 > 0.05$ اذن المجتمع محدود وعليه:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) = \frac{2.5}{2} \left(\frac{4-2}{4-1} \right) = 0.833.$$

وبالتالي فان الانحراف المعياري يساوي:

$$\sigma_{\bar{X}} = \sqrt{\sigma_{\bar{X}}^2} = \sqrt{0.833}.$$

مثال (2-5)

سحبت عينة عشوائية حجمها $n = 10$ من مجتمع طبيعي معدله $\mu = 70$ وتباينه $\sigma^2 = 40$ فأوجد القيمة المتوقعة للمتوسط (معدل المتوسط) الحسابي وتباينه وانحرافه المعياري.

الحل

الإحصاء الإداري ٢

القيمة المتوقعة للمتوسط (معدل المتوسط) الحسابي وتباينه وانحرافه المعياري

$$\begin{aligned}\mu_{\bar{X}} &= \mu = 70 \\ \sigma_{\bar{X}}^2 &= \frac{\sigma^2}{n} = \frac{40}{10} = 4 \\ \sigma_{\bar{X}} &= \sqrt{4} = 2.\end{aligned}$$

1.3.2 توزيع المعاينة للوسط الحسابي اذا كان تباين المجتمع معلوم

نفرض ان لدينا مجتمع متوسطه μ وتباينه σ^2 ونفرض ان $\bar{X}$ تمثل متوسط عينة حجمها n مسحوبة من ذلك المجتمع فإن $\bar{X}$ يخضع تقريباً لتوزيع طبيعي وسطه وتباينه $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ وذلك عندما تكون n كبيرة نسبياً أي أن

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

مثال (2-6)

إذا كان متوسط الدخل الشهري للأسر في إحدى المدن هو 6000 دولار، والانحراف المعياري هو 1000 دولار، واختيرت عينة عشوائية حجمها 100 من هذه المدينة فأوجد :

(1) القيمة المتوقعة (الوسط الحسابي) والتباين والانحراف المعياري لمتوسط دخل هذه الأسر.

(2) احتمال أن يتراوح متوسط دخل الأسرة في العينة بين 5950, 6150.

الحل

(1) القيمة المتوقعة (الوسط الحسابي) والتباين والانحراف المعياري لمتوسط دخل هذه الأسر.

$$\begin{aligned}E(\bar{X}) &= \mu_{\bar{X}} = \mu = 6000 \\ V(\bar{X}) &= \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} = \frac{1000^2}{100} \\ \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{n}} = \frac{1000}{\sqrt{100}} = 100.\end{aligned}$$

(2) احتمال أن يتراوح متوسط دخل الأسرة في العينة بين 5950, 6150. أي إيجاد احتمال المقدار:

$$P(5950 \leq \bar{X} \leq 6150)$$

حجم العينة $n = 100$ وباستخدام نظرية النهاية المركزية، فإن $\bar{X}$ يتوزع تقريباً توزيعاً طبيعياً بمعالم μ , $\frac{\sigma}{\sqrt{n}}$ ويكتب عادة

$$\bar{X} \sim N(6000, \frac{1000^2}{100})$$

وعليه فإن الاحتمال المطلوب

$$P(5950 \leq \bar{X} \leq 6150)$$

$$Z_{1\bar{X}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{5950 - 6000}{100} = -0.5$$

$$Z_{2\bar{X}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{6000 - 6150}{100} = 1.5.$$

ويكون الاحتمال المطلوب هو

$$\begin{aligned} P(5950 \leq \bar{X} \leq 6150) &= P(-0.5 \leq Z \leq 1.5) \\ &= P(Z \leq 1.5) - 1 + P(Z \leq 0.5) \\ &= 0.9332 - 1 + 0.6915 = 0.6247. \end{aligned}$$

مثال (7-2)

إذا كان معدل الذكاء في مجتمع ما، له توزيع طبيعي بمتوسط $\mu = 100$ وانحراف معياري $\sigma = 10$ ، فإذا إختارنا عينة حجمها $n = 16$ فرداً من هذا المجتمع، فما هو احتمال أن يقل متوسط الذكاء لهذه العينة عن 105 درجة.

الحل

$$\mu = 100, \sigma = 10, n = 16$$

أولاً: نوجد معدل الوسط الحسابي وتباينه وانحرافه (الخطأ) المعياري

$$\begin{aligned} \mu_{\bar{X}} &= \mu = 100, \\ \sigma_{\bar{X}}^2 &= \frac{\sigma^2}{n} = \frac{100}{16} = 6.25, \\ \sigma_{\bar{X}} &= 2.5. \end{aligned}$$

ثانياً: المطلوب هو إيجاد الاحتمال التالي:

$$P(\bar{X} \leq 105)$$

نستخدم التعويض

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{105 - 100}{2.5} = 2$$

ويكون الاحتمال المطلوب هو

$$P(\bar{X} \leq 105) = P(Z \leq 2) = 0.9772.$$

ملاحظة

المقدار $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ يحتوى على متغير واحد فقط هو $\bar{X}$ ، والذي يختلف من عينة لأخرى حيث حجم العينة n ثابت، وكل من μ, σ معلومتين. ولكن إذا كانت σ مجهولة وهذا يحدث فى معظم التطبيقات العملية فإن σ^2 تقدر من مشاهدات العينة بالمقدار S^2 الذى يعطى بالصيغة التالية:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (2-2)$$

وإذا كانت العينة حجمها n كبيراً $n \geq 30$ فإن S^2 تقترب من σ^2 كذلك تعامل S^2 كما لو كانت σ^2 ويصبح المقدار السابق كالتالى:

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

مثال (2-8)

إذا كان معدل الذكاء فى مجتمع ما، له توزيع طبيعى بمتوسط $\mu = 100$ اختيرت عينة عشوائية حجمها $n = 36$ فرداً من هذا المجتمع، فوجد ان الانحراف المعياري لها $S = 10$.

(1) أوجد المعدل والتباين والخطأ المعياري لمتوسط هذه العينة.

(2) ما هو احتمال أن يقع متوسط الذكاء لهذه بين 95 و 105 درجة.

الحل

$$\mu = 100, S = 10, n = 36$$

اولاً: نوجد معدل الوسط الحسابي وتباينه وانحرافه (الخطأ) المعياري

$$\mu_{\bar{X}} = \mu = 100, \quad \sigma_{\bar{X}}^2 = \frac{S^2}{n} = \frac{100}{36} = 2.8, \quad \sigma_{\bar{X}} = 1.67$$

ثانياً: المطلوب هو إيجاد الاحتمال التالي

$$P(95 \leq \bar{X} \leq 105)$$

نستخدم التعويض

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{95 - 100}{1.67} = -2.99$$

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{105 - 100}{1.67} = 2.99.$$

ويكون الاحتمال المطلوب هو

$$\begin{aligned} P(95 \leq \bar{X} \leq 105) &= P(-2.99 \leq Z \leq 2.99) \\ &= P(Z \leq 2.99) - 1 + P(Z \leq 2.99) \\ &= 0.9986 - 1 + 0.9986 = 0.9972. \end{aligned}$$

2.3.2 توزيع المعاينة للوسط الحسابي اذا كان تباين المجتمع مجهول و كانت

n صغيرة

إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية صغيرة الحجم ($n < 30$) مأخوذة من مجتمع طبيعي متوسطة μ وتباينة σ^2 مجهول فان المتغير العشوائي $\bar{X}$ له توزيع t بمتوسط μ وتباين s^2 أي ان المتغير

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}.$$

يخضع لتوزيع t بدرجة حرية $v = n - 1$.

مثال (2-9)

إذا كان X متغيراً عشوائياً يمثل أوزان أكياس الدقيق التي تنتجها إحدى المؤسسات، وكان X خاضعاً لتوزيع طبيعي متوسطه 50 جراماً أخذت عينة حجمها 9 أكياس من انتاج هذه المؤسسة ، ووجد ان الانحراف المعياري لأوزان هذه الاكياس في العينة يساوي 1 جرام، أوجد

(1) المتوسط والتباين والخطأ المعياري لمتوسط أوزان اكياس الدقيق في العينة.

(2) احتمال أن يزيد الوسط الحسابي لهذه العينة عن 51 جرام.

(1) حساب الوسط الحسابي لتوزيع المعاينة للمتوسطات من العلاقة

$$\mu_{\bar{X}} = \mu = 50,$$

وتباين المتوسط يعطى من العلاقة

$$\sigma_{\bar{X}}^2 = \frac{S^2}{n} = \frac{1}{9},$$

الخطأ المعياري لمتوسط العينة هو

$$\sigma_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{1}{\sqrt{9}} = \frac{1}{3}.$$

(2) وحيث أن حجم العينة صغير، $n = 9 < 30$ وتباين المجتمع σ^2 مجهول. فإن المقدار

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{51 - 50}{\frac{1}{3}} = 3$$

يخضع لتوزيع t بدرجات حرية

$$v = n - 1 = 9 - 1 = 8$$

لذلك نحسب الاحتمال كالتالى:

باستخدام جدول t نجد ان

$$P(\bar{X} \geq 51) = P(T \geq 3) = 0.01.$$

مثال (10-2)

إذا كان X متغيراً عشوائياً يمثل وزن معلبات غذائية تنتجها إحدى المؤسسات وكان X خاضعاً لتوزيع طبيعى متوسطه 150 جراماً، أخذت عينة حجمها 9 معلبات من انتاج هذه المؤسسة ، ووجد ان الانحراف المعياري لأوزان هذه المعلبات في العينة يساوى 5 جرام، احسب احتمال أن يزيد الوسط الحسابي لهذه العينة عن 155 جرام.

الحل

بما أن حجم العينة صغير، $n = 9 < 30$ وتباين المجتمع σ^2 مجهول. فإن المقدار $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ يخضع

لتوزيع t بدرجات حرية $v = n - 1 = 9 - 1 = 8$ لذلك نحسب الاحتمال كالتالى:

$$\begin{aligned} P(\bar{X} \geq 155) &= P\left(\frac{\bar{X} - \mu}{S/\sqrt{n}} \geq \frac{155 - 150}{5/\sqrt{9}}\right) \\ &= P(T \geq 3) \end{aligned}$$

وباستخدام جدول t نجد ان

$$P(\bar{X} \geq 155) = P(T \geq 3) = 0.01$$

4.2 توزيع المعاينة للفروق بين المتوسطات الحسابية

فيما سبق تناولنا توزيعات المعاينة لبعض المقاييس المحسوبة من عينات مختارة من مجتمع واحد مثل توزيع المعاينة للمتوسط الحسابى $\bar{X}$ فى حالة العينات الكبيرة والصغيرة، والآن نرض أنه لدينا مجتمعين الأول بمتغير عشوائى X_1 وتوقع μ_1 وتباين σ_1^2 وأختيرت منه عينة عشوائية حجمها n_1 بمتوسط $\bar{X}_1$ وتباين S_1^2 والمجتمع الثانى بمتغير عشوائى X_2 وتوقع μ_2 وتباين σ_2^2 وأختيرت منه عينة عشوائية حجمها n_2 بمتوسط $\bar{X}_2$ وتباين S_2^2 فإنه لدينا الحالات الآتية:

1.4.2 توزيع المعاينة للفرق بين متوسطين حسابيين لمجتمعين طبيعيين بتباينين

معلومات

إذا كان المجتمعان طبيعيان وكان تباين كلاً منهما هو σ_1^2, σ_2^2 معلومتين فإن الفرق بين متوسطى العينتين $\bar{X}_1 - \bar{X}_2$ له توزيع طبيعى بمتوسط $\mu_1 - \mu_2$ وتباين $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ أى أن:

$$\bar{X}_1 - \bar{X}_2 \approx N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}) \quad (2-3)$$

أي ان متوسط الفرق بين متوسطي المجتمعين يعطى من العلاقة

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2.$$

وتباين الفرق بين المتوسطين يعطى من العلاقة:

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

كما يعطى الخطأ المعياري للفرق بين المتوسطين من العلاقة

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ويمكن استخدام التعويض

$$Z_{\bar{X}_1 - \bar{X}_2} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

للتحويل من التوزيع الطبيعي الى التوزيع الطبيعي القياسي اي ان

$$Z_{\bar{X}_1 - \bar{X}_2} \sim N(0, 1).$$

ونوضح طريقة استخدام العلاقة السابقة فى حساب الاحتمالات من خلال المثال التالى:

مثال (11-2)

إذا كان متوسط أوزان العبوة الواحدة لحدى السلع من مؤسسة (أ) هو $\mu_1 = 5$ كجم، بتباين $\sigma_1^2 = 0.5$ وكان متوسط أوزان العبوة الواحدة لنفس السلعة من المؤسسة (ب) هو $\mu_2 = 4$ كجم، بتباين $\sigma_2^2 = 0.25$ أخذت عينة $n_1 = 20$ وحدة من المؤسسة (أ) وعينة أخرى $n_2 = 10$ من وحدات المؤسسة (ب)،

والمطلوب حساب

(1) المتوسط والتباين والانحراف المعياري للفرق بين متوسطي المجتمعين.

(2) احتمال أن يكون بين متوسطى العينتين $P(\bar{X}_1 - \bar{X}_2 < 0.75)$.

الحل

$$n_1 = 20, n_2 = 10, \mu_1 = 5, \mu_2 = 4, \sigma_1^2 = 0.5, \sigma_2^2 = 0.25.$$

(1) - متوسط الفرق بين متوسطي المجتمعين

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 = 5 - 4 = 1$$

تباين الفرق بين متوسطي المجتمعين هو

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{0.5}{20} + \frac{0.25}{10} = 0.05$$