الفصل الخامس

التوزيع الطبيعي

(٥ ـ ١) مقدمة

رأينـــا في الفقـــرة (٣ ـ ٥) أن المتغيرات العشـــوائية المستمـرة تولـد فضاء عينـة مستمــرا، بمعنى أن نقاطــه تكـون متراصـة بعضهـا إلى بعض كنقـاط محور موجـه. وبالتـــالي فإنها، بالإضـافة إلى كـونها لا نهائية في عـددهـا، غير قابلـة للعد. وكأمثلـة تقليدية على متغيرات عشـوائية مستمرة، نذكر أطـوال البشر وأوزانهم، وأخطاء القياس في تجربة مخبريـة، وعمـر مصباح كهربائي، إلخ. كما رأينا في تلك الفقـرة أنه للحصول على نمـوذج احتمالي لمتغير عشوائي مستمـر، X، نبدأ بـاختيار منحن مستمـر يمثل ما سميناه بـدالة الكثـافـة الاحتمـالية، وأن مثل هـذه الدالة، ولنرمز لها بـ $f(x)$، يجب أن تحقق شرطين:

١ ـ $f(x) \geq 0$، مهما يكن x،

٢ ـ المساحة تحت $f(x)$ تساوي الواحد تماما.

وعنـدئـذ يكـون احتمال أي حـادثة عددية مثل $a < X < b$، حيث a، b عـددان محددان، هـو المسـاحة تحت منحنى الكثـافـة فـوق الفترة (a, b) مـن محور السينات. وكنتيجة لذلك نجد أن احتمال أن يفترض المتغير X قيمة معينة، a مثلا، أي $P(X = a)$، هو المسـاحـة تحت المنحنى فوق النقطة a، وهي صفر. وهكـذا فإن مثل

هذا الحل لمشكلة إيجاد نموذج احتمالي لفضاء عينة مستمر يحتم علينا القول إن احتمال أن يكون لمتغير عشوائي مستمر قيمة معينة هو احتمال يساوي الصفر. وهذا تعبير واقعي عن استحالة توصل الإنسان إلى أجهزة قياس دقيقة بصورة مطلقة.

وبينما تتخذ منحنيات الكثافة أشكالا مختلفة نلاحظ أن عددا كبيرا من المتغيرات العشوائية التي نواجهها في حياتنا العامة لها منحني كثافة، أو منحني تكرار، له تقريبا شكل الجرس، أو، كما نعبر عن ذلك إحصائيا، له بصورة تقريبية شكل منحني التكرار الطبيعي، أو شكل التوزيع الطبيعي.

وبصورة عامة لنفرض أننا لاحظنا، في مجتمع القياسات لظاهرة معينة، ميلا واضحا إلى التناظر والاعتدال، بمعنى أن القياسات المتطرفة التي تمثل فرط زيادة أو فرط نقصان، هي قياسات نادرة. ويزداد تكرار ظهور القياس في ذلك المجتمع كلما اقتربت قيمة القياس من المتوسط. فالقيمة المتوسطة في المجتمع والقيم المجاورة لها هي القياسات الأكثر تواترا، بينما تكون القياسات البعيدة عن المتوسط زيادة أو نقصانا نادرة الظهور. وبعبارة أخرى، لنفرض أن الوسطية والاعتدال هي السائدة في مجتمع القياسات لظاهرة معينة، فعندئذ نقول إن النموذج الاحتمالي المناسب لهذه الظاهرة هو نموذج «التوزيع الطبيعي». وقد برزت تسمية «الطبيعي» في القرن الثامن عشر في سياق نظرية «أخطاء القياسات» عندما وجد أنه في تجربة يسير كل شيء فيها سيرا طبيعيا (normally)، ستكون أخطاء القياسات خاضعة للتوزيع الاحتمالي الذي يتخذ منحنى الكثافة فيه شكل الجرس (أو شكل منحنى جاوس). وتجدر هنا ملاحظة أنه عندما تتوافر كفاءة المجرب ومقدرته على إجراء القياسات بصورة سليمة، وتتوافر إلى جانب ذلك سلامة الأجهزة المستخدمة، وسلامة الظروف التي تتم تحتها التجربة، فإن الأخطاء ستتذبذب بصورة قريبة من التناظر بين أخطاء بالزيادة وأخطاء بالنقصان، وستكون الأخطاء الفاحشة بالزيادة أو بالنقصان نادرة، بينما تتمركز معظم نتائج القياسات حول القيمة الحقيقية، التي تشكل المتوسط، وقريبا منها. وينبغي ألا تملي التسمية أي شكل من أشكال خصوصية هذا التوزيع لعلوم الطبيعة، فهو يلعب، في

الـواقع، دورا أعم من ذلك بكثير وأوسع، وهو بين التـوزيعات الاحتمالـيـة، بمختلف أنـواعهـا وسمياتها، علم بـارز، إليـه تستند، بصـورة رئيسـة، العـديـد من الطـرق الإحصـائية، وبـدونه تضيق الحلبـة الواسعـة لتطبيقات الإحصـاء في الحياة المعـاصرة. وسنجـد فيما يسمى «نظـريـة النهـايـة المركـزيـة» أن مجمـوع عـدد كبير من المركَبـات العشـوائية، هـو دائما متغير عشـوائي ينزع، تحت شـروط عـامة جدا، إلى الخضـوع للتوزيـع الطبيعي، وذلـك بصرف النظر عن طبائع تلـك المركبات العشوائيـة التي تمثل كل منها متغيرا عشـوائيا لـه توزيعه الاحتمالي الخاص. وقـد رأينا في الفصـل الثـاني أن المعـايير الإحصائية المهمة يعبر عنهـا بدلالـة مجمـوع متغيرات، فمثلا، $n\bar{x} = \sum_{1}^{n} x_i$ و $(n-1)\,S^2 = \sum_{1}^{n}(x_i - \bar{x})^2$، كـما رأينا في الفصل الرابـع أن عدد النجاحات، X، في تجربة ثنائية ما هـو إلا مجموع عينة حجمها n مأخوذة من مجتمع بيرنوللي. وهذا يشير بوضوح إلى الأهمية الخاصة لهذا التوزيع في مباحث الإحصاء.

(٥ ـ ٢) دالة الكثافة الاحتمالية للتوزيع الطبيعي

تعرف دالة الكثافة الإحتمالية للتوزيع الطبيعي كما يلي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad \begin{array}{l} -\infty < x + \infty \\ -\infty < \mu < +\infty \\ 0 < \sigma < +\infty \end{array}$$

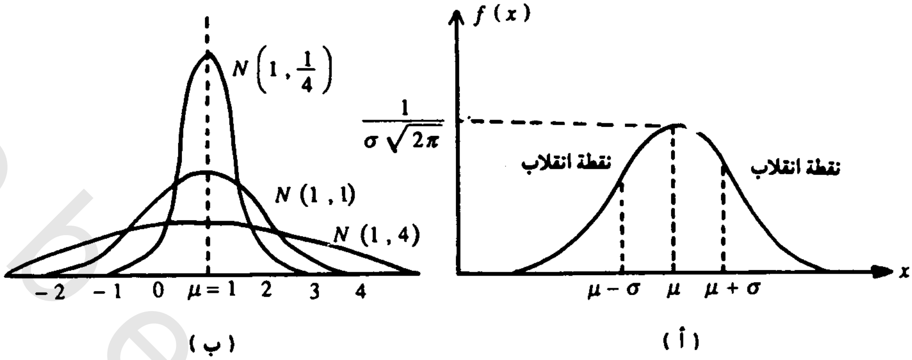
وهي دالة منحن له شكل الجرس (انظر الشكل ٥ ـ ١ (أ)) حيث:

π عدد ثابت يساوي تقريبا 3.1416،

e عدد ثابت يساوي تقريبا 2.7183،

μ عدد ثابت يمكن أن يكون أي عدد حقيقي،

σ عدد ثابت يمكن أن يكون أي عدد حقيقي موجب.



شكل (٥ ـ ١)

والدالة أعلاه لا تحدد منحنيا واحدا بعينه و إنما تحدد الشكل العام لعائلة من المنحنيات. إذ كلما حددنا لـ μ قيمة ولـ σ قيمة نحصل على منحن محدد تماما. ولذلك يسمى كل من الثابتين μ ، σ مَعلمة.

ويمكن البرهان على أن المعلمة μ تمثل متوسط التوزيع الاحتمالي، أي $E(X) = \mu$، وأن المعلمة σ تمثل الانحراف المعياري للتوزيع الاحتمالي، أي $V(X) = \sigma^2$. وللمنحنيات الطبيعية المختلفة متوسطات مختلفة، وانحرافات معيارية مختلفة، إلا أن المتوسط μ والانحراف المعياري σ لمنحن طبيعي معين محددان تمام التحديد وثابتان. وهكذا نجد أن تحديد قيمة لـ μ وقيمة لـ σ يحدد تماما منحنيا، وعلى العكس كل منحن من عائلة المنحنيات الطبيعية (منحنيات جاوس أو المنحنيات على شكل جرس) تحدد تماما قيمة لـ μ وقيمة لـ σ. وهذا يلقي بعض الضوء على سبب تسمية μ و σ مَعلمات. ويُبرهَن في الحساب التكاملي أن المساحة تحت المنحنى

$$e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

يساوي $\sigma\sqrt{2\pi}$ تماما. وبالتالي فإن المساحة تحت المنحنى الطبيعي $f(x)$، كما عرفناه أعلاه تساوي الواحد تماما.

ونلاحظ أن المنحني متناظر حول المستقيم $X = \mu$ الموازي للمحور الرأسي. لأن الدالة f تأخذ القيمة نفسها في نقطتين متناظرتين بالنسبة إلى النقطة $x = \mu$، فلو حسبنا، مثلا، $f(\mu + a)$ و $f(\mu - a)$ لوجدنا:

$$f(\mu + a) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left[\frac{\mu + a - \mu}{\sigma}\right]^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\frac{a^2}{\sigma^2}}$$

$$f(\mu - a) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left[\frac{\mu - a - \mu}{\sigma}\right]^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\frac{a^2}{\sigma^2}}$$

فالنقطة $x = \mu$ على المحور الأفقي هي النقطة التي يتمركز عندها التوزيع $(E(X) = \mu)$ ، وينتشر على جانبيها بصورة متناظرة .

ومن دراستك السابقة للدالة الرأسية e^{-t} ، مثلا، تذكر أنه إذا كان الأس t موجبا دوما، كما هو الحال في الدالة $f(x)$ هنا حيث $t = \frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2$ ، فإن أكبر قيمة لـ e^{-t} تساوي الواحد، وهي القيمة الموافقة لـ $t = 0$ ، (في مثالنا $x = \mu$) . وتتناقص قيمة e^{-t} مع تزايد t وتنتهي إلى الصفر (أي تتقارب إلى المحور الأفقي) عندما تزداد t إلى اللانهاية . وهكذا فإن دالة الكثافة $f(x)$ تبلغ نهايتها العظمى عند $x = \mu$ وتكون قيمتها عندئذ :

$$f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{\mu - \mu}{\sigma}\right)^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^0 = \frac{1}{\sigma\sqrt{2\pi}}$$

وللمنحنى $f(x)$ نقطتا إنقلاب عند $x = \mu + \sigma$ و $x = \mu - \sigma$ (أنظر الشكل ٥ ـ ١ (أ)) . لنتصور في الشكل ٥ ـ ١ (أ) أن المنحنى عبارة عن سلك رفيع وشديد المرونة . فإذا ضغطنا على القمة سينتشر السلك انتشارا أوسع على جانبي μ ، أي يأخذ شكلا أكثر انبساطا باعتبار أن المساحة تحت السلك يجب أن تبقى دائما ثابتة ومساوية للواحد . وإذا رفعنا القمة إلى أعلى فسيقل انبساط المنحنى ويتضاءل انتشاره على جانبي μ . ونرى في الشكل ٥ ـ ١ (ب) تمثيلا يوضح الفكرة . وقد استخدم الرمز $N(\mu, \sigma^2)$ للدلالة على توزيع طبيعي بمتوسط μ وتباين σ^2، وهكذا يعني $N\left(1, \frac{1}{4}\right)$ توزيعا طبيعيا متوسطه $\mu = 1$ وتباينه $\sigma^2 = \frac{1}{4}$ ، وللمنحنيات الثلاثة في الشكل (٥ ـ ١) ب المتوسط نفسه وهو 1 . وعندما ارتفعت قمة المنحنى $N(1, 1)$ تضاءل انتشاره على جانبي المتوسط $\mu = 1$ وبالتالي قل σ^2 من 1 إلى $\frac{1}{4}$ ، وعلى العكس عندما انخفضت قمة المنحنى، اتسع

انتشـاره على جـانبي μ وازداد تباينه من 1 إلى 4 . ولو عــدنا إلى قيمـة $f(x)$ العظمـى وهي $f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$ لتبين لنـا أن القيم الصغــيرة لـ σ تعني قمـة مرتفعة، أي توزيعـا أقل انتشارا حـول متوسطـه، وأن القيم الكبيرة لـ σ تعني قمـة منخفضة، أي توزيعا أكثر انتشارا على جـانبي المتوسط . ولما كان التباين، كما نعلم من الفصلين الثاني والرابع، مقياسا لمدى انتشار التوزيع على جانبي المتوسط، فإن هذه الملاحظة توضح أن σ^2 يمثل تبـاين التـوزيع الأمر الـذي ذكـرناه منـذ قليل كنتيجـة يمكن إثبـاتها رياضيـا باستخدام الحساب التكاملي وبطرق تعتبر فوق مستوى هذا الكتاب .

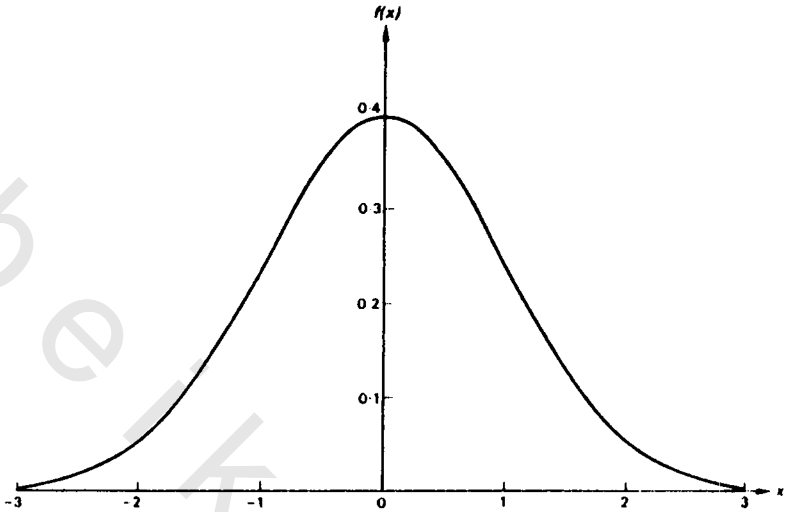
ومن بين أسرة المنحنيات الطبيعية :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad \begin{array}{l} -\infty < x + \infty \\ -\infty < \mu < +\infty \\ 0 < \sigma < +\infty \end{array}$$

سنختار منحنيـا خاصا هـو ذلك المنحنى الذي يكون متـوسطه $\mu = 0$ وانحرافه المعياري $\sigma = 1$. وتمييـزا لهذا المنحنى، الذي سيلعب دورا هـامـا في تطبيقات التـوزيع الطبيعي سنطلق عليــه اسم المنحنى الطبيعي المعيـاري. وإذا استخــدمنـا الحرف Z للمتغير الطبيعي المعياري فستصبح معادلة المنحنى أعلاه بعد وضع $\mu = 0$ ، $\sigma = 1$ على الشكل

$$f(Z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}Z^2} \quad ; \quad -\infty < Z < +\infty \, .$$

ونجد في الشكل (٥ ـ ٢) الرسم البياني لهذا المنحنى. وتجدر ملاحظة أنه متناظر بالنسبـة إلى المحـور الرأسـي. ومـا دامت المساحة تحت المنحنى بكامله من $Z = -\infty$ إلى $Z = +\infty$ هـي الواحد تمامـا فالمساحة على اليمين من $Z = 0$ تساوي المساحة على اليسار من $Z = 0$ وكل منهما تساوي النصف .



شكل (٥ ـ ٢) المنحنى الطبيعي المعياري

تمارين (٥ ـ ١)

١) اكتب دالة الكثافة للتوزيع الطبيعي مفترضا القيم التالية للمتوسط والتباين:

أ ـ المتوسط يساوي 3 ، والتباين يساوي 4 .

ب ـ المتوسط يساوي 0 ، والتباين يساوي 5 .

جـ ـ المتوسط يساوي 2–، والتباين يساوي 1 .

د ـ المتوسط يساوي 6– ، والتباين يساوي 10 .

حدد في كل حالة أين تقع قمة المنحنى وحاول أن تخطط رسما تقريبيا له .

٢) متغير عشوائي X دالة كثافته الاحتمالية :

$$f(x) = \frac{1}{3\sqrt{2\pi}} \; e^{-\frac{1}{18}x^2} \quad ; \quad -\infty < x < +\infty .$$

ما متوسطه و انحرافه المعياري؟

٣) متغير عشوائي X دالة كثافته الاحتمالية :

$$f(x) = c\, e^{-\frac{(x-4)^2}{6}} \quad ; \quad -\infty < x < +\infty \, .$$

ما قيمة c ؟

٠ (٥ ـ ٣) المساحات تحت منحنى الكثافة الطبيعي

ذكرنا أن معادلة منحنى الكثافة الطبيعي ، كما وردت في مستهل الفقرة السابقة ، لا تمثل منحنيا واحدا ، بل عائلة من المنحنيات لا حصر ولا عد لأعضائها . ووضع جدول للمساحات خاص بكل منها أمر غير ممكن . وسنجد الآن أنه يمكن وضع جدول واحد كاف لحساب المساحات تحت أي منحنى كثافة طبيعي . وأسهل طريقة لتحقيق ذلك هي أن نحسب المساحات الواقعة ضمن عدد محدد من الانحرافات المعيارية على جانبي المتوسط . وبما أن المنحنى متناظر يمكن التبسيط بإقامة جدول للمساحات تحت المنحنى بين μ والنقاط x الواقعة على اليمين من μ . وإذا فرضنا نقطة x أكبر من μ فإن المسافة بين x و μ هي $x - \mu$ ، وإذاعبرنا عنها بدلالة الانحراف المعياري σ ، ولنفرض أنها تساوي Z مرة الانحراف المعياري σ ، فيمكننا أن نكتب $x - \mu = Z\sigma$ ، وإذا قسنا المسافات على محور الفواصل بوحدة قياس تساوي σ (وعندها يكون $\sigma = 1$ حكما) فإن قيمة المسافة $x - \mu$ مقيسة بالوحدة الجديدة تصبح Z أي تساوي $\frac{x-\mu}{\sigma}$. وهكذا نكتب المتغير الجديد Z بدلالة المتغير X على الشكل :

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

ونلاحظ أنه يوافق كل قيمة لـ X قيمة واحدة لـ Z والعكس بالعكس . وأن Z ليس إلا القيمة المعيارية لـ X . وفي الواقع ، لو حسبنا $E(Z)$ و $V(Z)$ لوجدنا :

$$E(Z) = E\left[\frac{1}{\sigma}(X-\mu)\right] = \frac{1}{\sigma}\, E(X-\mu) = \frac{1}{\sigma}\left[E(X) - \mu\right] = 0$$

$$V(Z) = V\left(\frac{1}{\sigma} \times - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}\; V(X) = \frac{\sigma^2}{\sigma^2} = 1 \, .$$

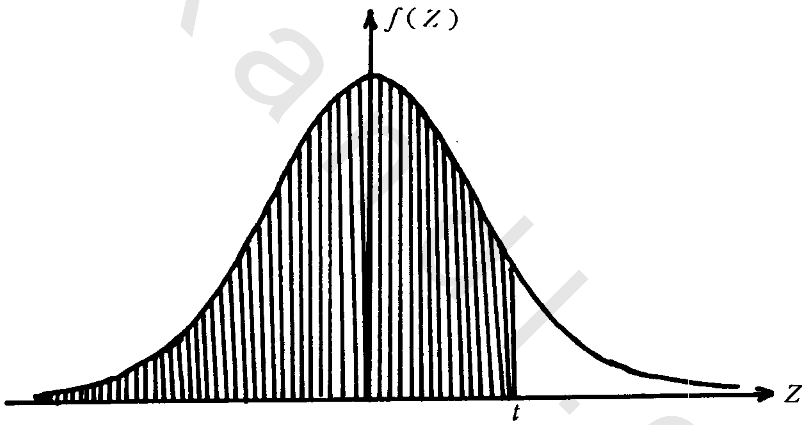
أي أن للمتغير Z متوسطا يساوي الصفر وانحرافا معياريا يساوي الواحد ، ويمكن البرهان على أن التوزيع الاحتمالي لـ Z هو التوزيع الطبيعي . وبذلك يكون

منحنى الكثـافة الموافق لـ Z عضـوا في أسرة المنحنيـات الطبيعيـة، وبـالـذات ذلك العضو المقابل لـ $\mu = 0$ و $\sigma = 1$. وهو بالضبط منحنى الكثافـة المذكور في ختام الفقرة (٥ ـ ٢):

$$f(Z) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{1}{2} Z^2} \quad ; \quad -\infty < Z < +\infty .$$

وربما أصبح واضحا الآن سبب تسمية هذا المنحنى بالمنحنى الطبيعي المعياري.

ويقـدم جدول التـوزيع الطبيعي في الملحق، المسـاحات تحت هـذا المنحنى إلى اليسار من نقطة معينة $Z = t$. ونقصد المساحة المظللة في الشكل (٥ ـ ٣).



شكل (٥ ـ ٣) دالة التوزيع المتجمع للمتغير الطبيعي المعياري Z

وبالعودة إلى جدول التـوزيع الطبيعي المشار إليه في الملحق، نلاحظ أن قيم Z في الجدول تبدأ مـن الصفر بفاصـل قدره 0.01 بين كل قيمـة والقيمة التي تليهـا. وهكذا تكون كل قيمـة لـ Z معطاة بـرقمين عشريين. ويتضمن العمـود الأول قيما لـ Z بفاصل يساوي 0.1 من قيمة إلى القيمة التي تليها. وتشكل هذه القيم عناوين لسطور الجدول، إذ نبدأ بالسطر 0 يليه السطر 0.1 ، فالسطر 0.2 ، وهكذا حتى نصل إلى السطر 3.4 . أما المنزلـة العشرية الثانيـة من قيمة Z فهي معطـاة في السطـر الأفقي الأول من الجدول، وتشكل عناوين لأعمـدة الجدول، بدءا من العمود الثاني حتى العمـود الأخير، وهكذا

نجد العمود 0.00 يليه العمود 0.01 ، يليه العمود 0.02 ، وهكذا حتى نصل إلى العمود 0.09 وهو العمود الأخير. وكل عدد في صلب الجدول، وهو ملتقى سطر مع عمود، يمثل المساحة تحت منحنى الكثافة المعياري وإلى اليسار من قيمة Z التي يحددها عنوان السطر حتى الرقم العشري الأول ويستكمل عنوان العمود رقمها العشري الثاني. وهكذا فإن العدد 0.8212 الواقع في ملتقى السطر 0.9 مع العمود 0.02 ، يمثل المساحة تحت منحنى الكثافة الواقعة إلى اليسار من $Z = 0.92$ ، أي المساحة تحت المنحنى وفوق الفترة الممتدة بين $-\infty$ والنقطة 0.92 من المحور Z . وعلى العكس، إذا أردنا المساحة الواقعة إلى اليسار من $Z = 1.96$ ، مثلا، ندخل الجدول وفق السطر 1.9 والعمود 0.06 فنجد عند ملتقاهما العدد 0.9750 وهو المساحة المطلوبة. وإذا كانت قيمة Z معطاة بأكثر من رقمين عشريين فإننا نحصرها بين قيمتين مذكورتين في الجدول ثم نقوم بعملية تناسب طردي، (عملية استيفاء).

والأسئلة الوجيهة التي تطرح نفسها هنا هي:

١ ـ إذ يقتصر الجدول على القيم الموجبة لـ Z ، ما العمل لو كانت القيمة المعطاة لـ Z سالبة؟

٢ ـ ما العمل لو كان المطلوب هو المساحة إلى اليمين من قيمة لـ Z سالبة أو موجبة؟

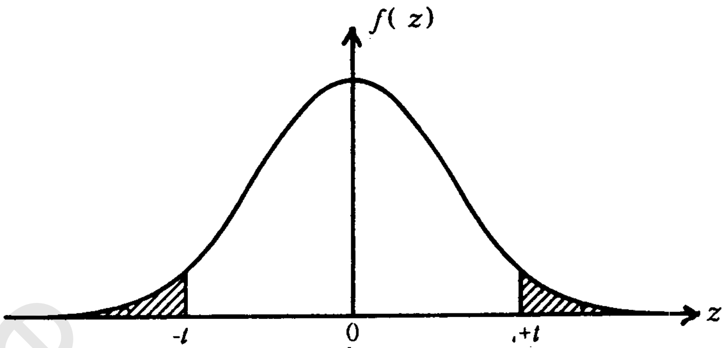
٣ ـ ما العمل لو كان المطلوب هو المساحة بين أي قيمتين لـ Z ؟

وللإجابة عن هذه التساؤلات نعود إلى التعريف في (٣ ـ ٦ ـ ٢) لدالة التوزيع المتجمع، ونكتب: المساحة تحت المنحنى الطبيعي المعياري الواقعة إلى اليسار من النقطة t = $P(Z \leq t) = F(t)$

وتتمتع هذه الدالة $F(t)$ بالخاصة المهمة التالية:

$$F(-t) = 1 - F(t)$$

وهي نتيجة مباشرة لتناظر المنحنى الطبيعي المعياري بالنسبة إلى المحور الرأسي. إذ لو نظرنا إلى الشكل (٥ ـ ٤) لوجدنا أن $F(-t)$ يساوي المنطقة I المظللة على اليسار من



شكل (٥ ـ ٤)

$Z = -t$. وأن $F(t)$ هي مجموع المنطقـة المظللة في أقصى اليسار والمنطقـة غير المظللة في الوسط . و $1 - F(t)$ يساوي بوضـوح المنطقة المظللة في أقصى اليمين، وبما أن المنطقتين المظللتين متساويتان بحكم التناظر فإن $F(-t) = 1 - F(t)$. ولإيجاد $F(-t)$ يكفي إذن حساب $F(t)$ من الجدول الموصوف أعلاه، حيث t موجبة، ثم نطرح القيمة الناتجة من 1 وهذا يجيب عن السؤال الأول .

ومن خاصة الحادثتين المتتامتين، $P(A) = 1 - p(\bar{A})$ ، نجد مباشرة أن :

$$P(Z > t) = 1 - P(Z \leq t) = 1 - F(t)$$

وهذا يجيب عن السؤال الثاني .

وللإجابة عن السؤال الثالث، لنفرض أن المطلوب هو حساب

$$P(a < Z \leq b)$$

فمن الواضح أنـه يمكن التعبير عن الحادثة $(Z \leq b)$ كإتحاد حـادثتين منفصلتين على الشكل

$$(Z \leq b) = (Z \leq a) \cup (a < Z \leq b)$$

ومنه :

$$P(Z \leq b) = P(Z \leq a) + P(a \leq Z \leq b)$$

أي :

$$F(b) = F(a) + P(a < Z \leq b)$$

أو

$$P(a < Z \leq b) = F(b) - F(a)$$

وبما أن الاحتمال الموافق لنقطة في التوزيعات المستمرة يساوي الصفر فيمكن كتابة

$$P(a \leq Z < b) = P(a < Z \leq b) = P(a \leq Z \leq b) = P(a < Z < b)$$

مثال (٥ ـ ١)

احسب $P(Z \leq 1.35)$ ، $P(Z > 0.5)$ ، $P(Z < -1.79)$ ، $P(Z \geq -0.68)$ ، $P(1 < Z < 3.27)$ ، $P(-0.1 < Z < 2.5)$ ، $P(-1.85 < Z < -0.16)$.

الحل

$$P(Z \leq 1.35) = F(1.35) = 0.09115$$

(ندخل الجدول وفق السطر 1.3 والعمود (0.05) .

$$P(Z > 0.5) = 1 - F(0.5) = 1 - 0.6915 = 0.3085$$

(ندخل الجدول وفق السطر 0.5 والعمود 0.00) .

$$P(Z < -1.79) = F(-1.79) = 1 - F(1.79) = 1 - 0.9633 = 0.0367$$

(ندخل الجدول وفق السطر 1.7 والعمود 0.09) .

$$P(Z \geq -0.68) = 1 - P(Z < -0.68) = 1 - F(-0.68)$$
$$= 1 - [1 - F(0.68)] = F(0.68) = 0.7517$$

(ندخل الجدول وفق السطر 0.6 والعمود 0.08) .

$$P(1 < Z < 3.27) = F(3.27) - F(1)$$
$$= 0.9995 - 0.8413 = 0.1582$$

$$P(-0.1 < Z < 2.5) = F(2.5) - F(-0.1)$$
$$= F(2.5) - [1 - F(0.1)]$$
$$= F(2.5) + F(0.1) - 1$$
$$= 0.9938 + 0.5398 - 1 = 1.5336 - 1 = 0.5336$$

$$P(-1.85 < Z < -0.16) = F(-0.16) - F(-1.85)$$
$$= [1 - F(0.16)] + [1 - F(1.85)]$$
$$= 2 - F(0.16) - F(1.85)$$
$$= 2 - 0.5636 - 0.9678 = 2 - 1.5314 = 0.4686$$

لاحظ أننا نعود إلى الجدول عندما يكون المطلوب $F(t)$ حيث t عدد موجب .

مثال (٥ ـ ٢)

احسب c بحيث يكون

$P(Z \le c) = 0.8264$ ، $P(Z < c) = 0.2981$ ، $P(Z > c) = 0.9292$ ، $P(-c < Z < c) = 0.9500$ ، $P(-c < Z < c) = 0.90$.

الحل

$$P(Z \le c) = F(c) = 0.8264$$

والعدد c هو قيمة Z في جدول التوزيع الطبيعي المعياري التي يقع إلى اليسار منها مساحة تساوي 0.8264 . ونبحث في صلب الجدول عن هذه القيمة لنجدها بالذات وعندئذ نحدد قيمة Z المطلوبة من السطر والعمود الموافقين ، أو نحصرها بين عددين في الجدول ثم نستنتج قيمة Z المطلوبة بعملية تناسب طردي (استيفاء). وفي حالتنا هنا نجد أن 0.8264 واقع في السطر 0.9 والعمود 0.04 وتكون القيمة c المطلوبة 0.94 .

$$P(Z < c) = 0.2981 \Leftrightarrow F(c) = 0.2981$$

وإذا كانت قيمة $F(c)$ أصغر من 0.5 فمن الواضح أن c ستكون سالبة . ولكن الجدول لا يحوي القيم السالبة لـ Z . وفي مثل هذه الحالة نأخذ :

$$F(-c) = 1 - F(c) = 1 - 0.2981 = 0.7019$$

ونبحث في صلب الجدول عن 0.7019 فنجده في السطر 0.5 والعمود 0.03 وتكون $-c = 0.53$ أو $c = -0.53$.

$$P(Z > c) = 0.9292 \Leftrightarrow 1 - F(c) = 0.9292$$

أي

$$F(-c) = 0.9292 \Leftrightarrow -c = 1.47 \Leftrightarrow c = -1.47$$
$$P(-c < Z < c) = 0.95 \Leftrightarrow F(c) - F(-c) = 0.95$$

ومنه

$$F(c) - [1 - F(c)] = 0.95$$

$$2F(c) = 1.95,\ F(c) = 0.975,\ c = 1.96.$$

$$P(-c < Z < c) = 0.90 \Leftrightarrow 2F(c) = 1.90$$

أي

$$F(c) = 0.95$$

ولدينا من الجدول

$Z = 1.64$ تقابل 0.9495

$Z = 1.65$ تقابل 0.9505

ومنه

تزايد المساحة	تزايد Z
0.001	0.01
0.0005	?

التزايد المطلوب في Z $= \dfrac{0.0005 \times 0.01}{0.001} = 0.005$

وتكون قيمة Z المطلوبة هي

$$1.64 + 0.005 = 1.645$$

وسنصطلح على كتابة Z_α لتعني قيمة المتغير الطبيعي المعياري التي يقع إلى اليمين منها مساحة تساوي α. أي أن $F(Z_\alpha) = 1 - \alpha$. وبهذا المعنى يكون :

$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = F(Z_{\alpha/2}) - F(-Z_{\alpha/2})$$

$$= 2F(Z_{\alpha/2}) - 1 = 2\left(1 - \frac{\alpha}{2}\right) - 1 = 1 - \alpha$$

وعلى سبيل المثال :

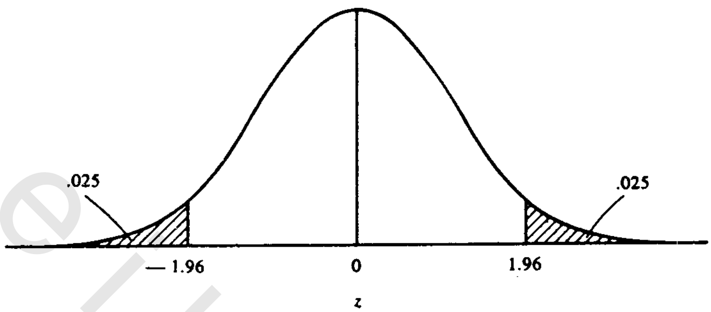
$Z_{.025}$ هي قيمة Z التي تحصر إلى اليمين منها مساحة تساوي 0.025. ويكون

$$P(-Z_{0.025} < Z < Z_{0.025}) = 0.95$$

وقد رأينا في المثال السابق أن

$$Z_{0.025} = 1.96$$

(أنظر الشكل ٥ ـ ٥).



شكل (٥ ـ ٥)

لقد تعلمنـا حتى الآن كيف نحسب احتمالات حوادث معبرا عنها بـدلالة المتغير المعياري Z ، وذلك بـالاستفـادة من جـدول التـوزيع الطبيعي المعيـاري. ولكن كيف نستفيد من هذا الجدول نفسه لحساب احتمالات حوادث معبر عنها بدلالة متغير طبيعي غير معياري، X ، مثلا؟ بالطبع لا يمكننا حسـاب مثل هذه الاحتمالات إلا إذا حددنا منحنى الكثافة للمتغير X تحديدا تاما. أي علمنا متوسطه μ وانحرافه المعياري σ . وعند معرفة قيمة μ وقيمة σ يصبح الأمر في غاية السهولة، إذ نقوم بمعايرة X ، أي نكتب:

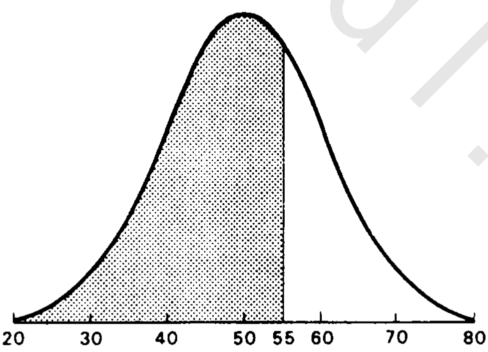
$$Z = \frac{X - \mu}{\sigma}$$

ونحول العبارة الاحتمالية بدلالة X إلى عبارة احتمالية مكافئة بدلالة Z ، ثم نعود إلى جدول التوزيـع الطبيعي المعياري، الذي تدرَّبنـا لتونا على كيفية استخـدامه، لحساب المطلوب وفيما يلي توضيح عملي للفكرة.

يقدم اختصاصي في علم النفـس نصائح حول أفضل المهن أو الوظـائف المناسبة لفتى. ولهذه الغاية يقـدم للفتى عددا من الاختبارات. أحدهـا ، مثلا، اختبار يهدف إلى قيـاس مهـارات التحـدث أو المهارات الشفهيـة. لنفـرض أن درجـة الفتى في هـذا

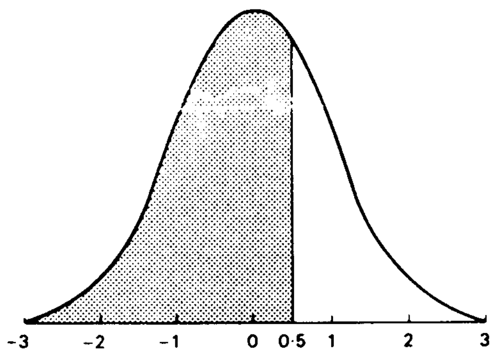
الاختبار كانت 55 . فهذا الرقم لذاته ليس له أي مدلول بالنسبة إلينا . إلا أن الاختصاصي النفسي يعلم توزيع درجات هذا الاختبار بالنسبة للرجال في المجتمع بصورة عامة . فمثل هذه الاختبارات قد استخدمت في الماضي على نطاق واسع وقدمت لعينة تمثيلية كبيرة من الرجال والنساء . وبالنسبة إلى الرجال تتوزع درجات هذا الاختبار، بصورة تقريبية، وفق التوزيع $N(50, 10^2)$. (في الواقع يتعمد مصممو هذه الاختبارات وضعها بحيث تتوزع الدرجات الناتجة عنها طبيعيا، على وجه التقريب) . وتوزيع هذه الدرجات مع درجة الفتى مبينة في الشكل (٥ ـ ٦) . وما يهم الاختصاصي النفسي حقا هو كيف يمكن مقارنة هذا الرجل مع بقية الرجال في المجتمع . ويمكن تلخيص هذه المقارنة بسهولة من خلال النسبة المئوية للرجال الذين يتوقع حصولهم على درجات في هذا الاختبار أسوأ من 55 وللحصول على هذه النسبة نحسب المساحة تحت منحنى الكثافة للتوزيع $N(50, 100)$ الواقعة إلى اليسار من النقطة 55 وبمعايرة الدرجة 55 تأخذ القيمة :

$$Z = \frac{55 - 50}{10} = 0.5$$



شكل (٥ ـ ٦) : التوزيع N (50, 100) لدرجات اختبار المهارة الشفهية، والمساحة المظللة هي احتمال الحصول على درجة أقل من 55 .

والمساحة المطلوبة هي إذا المساحة الواقعة إلى اليسار من النقطة 0.5 تحت منحنى الكثافة الطبيعي المعياري والمبينة في الشكل (٥ ـ ٧) . وهي تساوي من الجدول ١ في



شكل (٥ ـ ٧) درجة الاختبار بعد معايرتها.

الملحق 0.6915 . وهكذا نستنتج أن 69% من المجتمع يتوقع حصولهم على درجة أسوأ، و 31% من المجتمع يتوقع حصولهم على درجة أفضل وهذا يحدد بوضوح موقعه النسبي من الآخرين.

ولو فرضنا أن درجة هذا الشاب كانت 40 في اختبار لقياس المهارات الحسابية. وهذا الاختبار مصمم بدوره بحيث يكون توزيع الدرجات الناتجة عنه $N(50, 100)$. وبمعايرة هذه الدرجة نجد أنها تصبح في سلم القياس المعياري:

$$Z = \frac{40 - 50}{10} = -1$$

وبالعودة إلى جدول التوزيع الطبيعي المعياري نجد أن

$$F(Z) = F(-1) = 1 - F(1) = 1 - 0.8413 = 0.1587$$

وهكذا نتوقع أن ينال 16% فقط من المجتمع درجات أسوأ، وأن ينال 84% درجات أفضل.

وبصورة عامة، تسمى معايرة متغير طبيعي X توزيعه $N(\mu, \sigma^2)$ ، أي التحويل من X إلى المتغير الطبيعي Z وفق العلاقة

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

تعبيرا عن قيمة المتغير X وفق سلم القياس المعياري. وهو سلم قياس يعتبر μ مبدأ للقياسات، ويعتبر الانحراف المعياري σ وحدة قياس. وعندما لا نهتم بقيمة X لذاتها بل بموقع X النسبي من المتوسط μ، فإن القيمة Z توضح لنا بالضبط هذا الموقع النسبي ومنطوق العبارة الجبرية $X - \mu = Z\sigma$، هو أن موقع X يحيد عن النقطة μ بمقدار Z مرة الانحراف المعياري.

مثال (٥ ـ ٣)

إذا كانت درجات حاصل الذكاء تتوزع طبيعيا بمتوسط يساوي 100 وانحراف معياري يساوي 15، فما نسبة الناس ذوي درجة ذكاء:

أ ـ فوق 125، تحت 80، بين 70 و 130؟

الحل

لنرمز لدرجة حاصل الذكاء بـ X، فلدينا بالفرض أن توزيع X هو $N(100, 152)$. والمطلوب

أ ـ

$$P(X > 125) = 1 - F\left(\frac{125-\mu}{\sigma}\right) = 1 - F\left(\frac{125-100}{15}\right)$$
$$= 1 - F(1.67) = 1 - 0.9525 = 0.0475$$

والنسبة المطلوبة هي 4.75% .

ب ـ

$$P(X < 80) = F\left(\frac{80-\mu}{\sigma}\right) = F\left(\frac{80-100}{15}\right)$$
$$= F(-1.33) = 1 - F(1.33) = 1 - 0.9082 = 0.0918$$

والنسبة المطلوبة هي 9.18% .

جـ ـ

$$P(70 < X < 130) = F\left(\frac{130-\mu}{\sigma}\right) - F\left(\frac{70-\mu}{\sigma}\right)$$
$$= F\left(\frac{130-100}{15}\right) - F\left(\frac{70-100}{15}\right) = F(2) - F(-2)$$
$$= 2F(2) - 1 = 2(0.9772) - 1 = 1.9544 - 1 = 0.9544$$

والنسبة المطلوبة هي 95.44%.

وكثيرا مـا نستخدم عـلاقة المعـايرة، $Z=\dfrac{X-\mu}{\sigma}$ ، بطـريقة عكسيـة . فنحن نعرف أو نحدد سلفا قيمة Z ، أي القياس المطلوب على السلم المعياري، ونريد القياس المقابـل له على السلم الأصلي (قبل المعـايرة). فلنفـرض، مثلا، أن لـدى مديـر شركة وظيفة شاغرة، وهو لا يقبل مرشحين لهذه الوظيفة إلا إذا كانوا في مهاراتهم الحسابية من الربع الأعلى في المجتمع. ولترجمة رغبتـه هذه بدلالة الـدرجة الدنيـا التي ينبغي أن ينالها المرشح في اختبار المهـارات الحسابية، نقوم بما يلي، مفترضين أن درجـات الاختبار تتبع التوزيع $N(50, 100)$. نحدد من عبـارة «المرشح من الربع الأعلى في المجتمع في مهـاراته الحسابية» قيمة Z ، وذلك لأن هذه العبارة مكافئة للمعادلة $F(Z) = 0.75$ ، ومن جدول التوزيع الطبيعي المعياري نجد، باستخدام الاستيفاء، أن

$$Z = 0.67 + \left(\frac{14}{31}\right)(0.01) = 0.6745$$

وبالتالي

$$X = \mu + Z\sigma = 50 + 10\,(0.6745) = 56.745$$

وبالتدوير إلى أقرب عدد صحيح، نستنتج أن الدرجة المطلوبة هي 57 وهكذا لا يقبل طلب متقدم لهذه الوظيفـة إلا إذا كانت درجته في اختبار المهـارات الحسابية 57 أو أكثر.

مثال (٥ ـ ٤)

بالإشارة إلى المثال (٥ ـ ٣) وتوزيع درجات حـاصل الذكاء. لنفرض أن الحكومة تقدم تعليما خـاصا للخمسة في المائة الأدنى في حـاصل ذكائهم. وتقدم تعليما جـامعيا للسبعة في المائة الأعلى في حاصل ذكائهـم. أوجد القيم المعيارية Z المقابلة لهذه النسب ثم استنتج الحدود الفـاصلة في درجـات حاصـل الذكـاء لأولئك الذين يتطلبـون تعليما خاصا، ولأولئك الذين يدخلون الجامعات.

الحل

لنفرض أن القيمة المعيارية المقابلة لنسبة جماعة التعليم الخاص هي a ، والمقابلة لنسبة جماعة التعليم الجامعي هي b فعندئذ :

$P(Z \le a) = 0.05,\ F(a) = 0.05,\ F(-a) = 0.95,\ -a = 1.645,\ a = -1.645.$

$P(Z > b) = 0.07;\ 1 - F(b) = 0.07,\ F(b) = 0.93$

$b = 1.47 + 8(0.01)/14 = 1.47 + 0.0057 = 1.4757$

ويكون الحد الأعلى لدرجة حاصل الذكـاء لأولئك الذين يتطلبون تعليما خاصا ، مقربا إلى أقرب عدد صحيح هو:

$$X = \mu + a\sigma = 100 + 15(-1.645) = 75$$

والحد الأدنى لدرجة حاصل الذكاء لأولئك الذين يدخلون الجامعات، مقربا إلى أقرب عدد صحيح، هو:

$$X = \mu + b\sigma = 100 + 15(1.4757) = 122$$

مثال (٥ ـ ٥)

إذا كـــان X متغيرا طبيعيـــا متوسطـــه $\mu = 56$ وانحرافـه المعيـاري $\sigma = 3$ ، فاحسب $P(X \le 60.5)$ ، $P(X > 65)$ ، $P(53 < X < 59)$.

الحل

$$P(X \le 60.5) = P\left(\frac{X-\mu}{\sigma} \le \frac{60.5-\mu}{\sigma}\right) = P\left(Z \le \frac{60.5-56}{3}\right)$$
$$= F\left(\frac{60.5-56}{3}\right) = F(1.5) = 0.9332\ .$$

و

$$P(X > 65) = P\left(\frac{X-\mu}{\sigma} > \frac{65-\mu}{\sigma}\right) = P\left(Z > \frac{65-56}{3}\right)$$
$$= 1 - F\left(\frac{65-56}{3}\right) = 1 - F(3) = 1 - 0.9987 = 0.0013$$

و

$$P(53 < X < 59) = F\left(\frac{59-56}{3}\right) - F\left(\frac{53-56}{3}\right)$$
$$= F(1) - F(-1) = 2F(1) - 1$$
$$= 2(0.8413) - 1 = 1.6826 - 1 = 0.6826\ .$$

مثال (٥ ـ ٦)

ليكن X متغيرا عشوائيـا يتوزع وفق التـوزيع الطبيعي بمتوسـط يساوي 2 وتباين يساوي 16 . والمطلوب حساب احتمالات الحوادث العددية التالية :

$P(X < 3)$ ، $P(X > 1)$ ، $P(-1 < X < 35)$.

الحل

$$P(X<3) = F\left(\frac{3-\mu}{\sigma}\right) = F\left(\frac{3-2}{4}\right)$$
$$= F(0.25) = 0.5987 .$$
$$P(X>1) = 1 - F\left(\frac{1-\mu}{\sigma}\right) = 1 - F\left(\frac{1-2}{4}\right)$$
$$= 1 - F(-0.25) = F(0.25) = 0.5987$$
$$P(-1 < X < 3.5) = F\left(\frac{3.5-\mu}{\sigma}\right) - F\left(\frac{-1-\mu}{\sigma}\right)$$
$$= F\left(\frac{3.5-2}{4}\right) - F\left(\frac{-1-2}{4}\right)$$
$$= F(0.375) - F(-0.75)$$
$$= F(0.375) - [1 - F(0.75)]$$
$$= F(0.375) + F(0.75) - 1$$

ولحساب $F(0.375)$ نأخذ منتصف الطريق بين $F(0.37)$ و $F(0.38)$ ، أي منتصف الطريق بين 0.6443 و 0.648 وهو إلى أربعة أرقام عشرية 0.6462 . وهكذا يكون

$$P(-1 < X < 3.5) = 0.6462 + 0.7734 - 1 = 0.4196$$

مثال (٥ ـ ٧)

في عملية تعبئة آلية لعبوات السكر، من المفترض أن تضع الآلة في كل عبوة 2 كغ من السكر. وبالطبع يتغير ما تضعه الآلة من عبوة إلى أخرى بشكل عشوائي. إذا افترضنا أن ما تضعه الآلة بالفعل هو متغير $N(\mu, \sigma^2)$.

أ ـ تشير السجلات السابقة للإنتاج إلى أن $\sigma = 0.2$ ، وإلى أن احتمال أن تتضمن عبوة أقل من 2 كغ هو 0.01 . أوجد قيمة μ التي تعمل الآلة وفقا لها. (أي القيمة المتوسطة لما تضعه هذه الآلة في العبوة الواحدة على المدى الطويل.)

ب ـ إذا قمنا بعملية تحسين لعمل الآلة تتوخى تخفيض σ (أي انتاج عبوات أكثر تجانسا من حيث الوزن) مع بقاء μ كما هو . كم يجب أن تكون قيمة σ بحيث نطمئن إلى أن احتمال عبوة بأقل مما ينبغي من السكر هو 0.001 ؟

الحل

أ ـ لنرمز بـ X لوزن السكر الفعلي في العبوة. والمطلوب هو حساب μ علما أن

$$P(X < 2) = 0.01$$

و $\sigma = 0.02$. ولكن

$$P(X < 2) = F\left(\frac{2-\mu}{0.02}\right) = 0.01$$

أو

$$F\left(\frac{\mu-2}{0.02}\right) = 0.99$$

ومن الجدول نجد أن :

$$\frac{\mu-2}{0.02} = 2.33$$

أو

$$\mu = 0.02\,(2.33) + 2 = 2.047$$

ب ـ إذا اشتغلت الآلة وفقا لـ $\mu = 2.047$ فعندئذ يكون X متغيرا $N(2.047, \sigma^2)$.

ونريد قيمة σ بحيث يكون :

$$P(X < 2) = 0.001$$

ولكن الآن :

$$P(X < 2) = F\left(\frac{2-2.047}{\sigma}\right)$$

إذا نريد σ بحيث يكون

$$F\left(\frac{-0.047}{\sigma}\right) = 0.001$$

أو

$$F\left(\frac{0.047}{\sigma}\right) = 0.999$$

ومن الجدول نجد :

$$\frac{0.047}{\sigma} = 3.09$$

أي أن

$$\sigma = \frac{0.047}{3.09} = 0.015$$

مثال (٥ ـ ٨)

مفترضا أن طول الذكر البالغ X، مقاسا بالسنتمتر، هو متغير $N(175, 56.25)$. كيف يحدد مهندس ارتفاع أبواب الغرف في فيلا يقوم بتصميمها بحيث لا يضطر أكثر من 2% من الرجال إلى طأطأة رؤوسهم عند الدخول أو الخروج؟

الحل

لنفرض أن ارتفاع الباب a سم فيكون المطلوب تحديد قيمة a بحيث يكون :

$$P(X > a) \leq 0.02$$

ولكن

$$P(X > a) = 1 - P(X < a) = 1 - F\left(\frac{a - 175}{7.5}\right) \leq 0.02$$

وبالتالي يكون

$$F\left(\frac{a - 175}{7.5}\right) \geq 0.98$$

ومن الجدول نجد أن :

$$\frac{a - 175}{7.5} \geq 2.057$$

وهكذا يكون :

$$a \geq 175 + 2.057\,(7.5) = 190.43$$

أي أن ارتفاع الباب ينبغي أن يكون 190.5 سم على الأقل .

تمارين (٥ ـ ٢)

١) باستخدام جدول التوزيع الطبيعي أحسب الاحتمالات التالية، حيث Z المتغير الطبيعي المعياري $N(0, 1)$.

$P(Z \leq 1.2)$، $P(-0.9 < Z < 0)$، $P(0.3 < Z < 1.56)$، $P(|Z| < 0.2)$، $P(-1.3 < Z < 1.74)$، $P(Z > -0.75)$، $P(Z \leq -0.32)$.

٢) أوجد المساحة تحت منحنى كثافة التوزيع الطبيبي المعياري الواقعة :

أ ـ إلى اليسار من 1 ،

ب ـ إلى اليسار من 2 ،
جـ ـ بين 1 و 2 ،
د ـ إلى اليمين من 0.5 – ،
هـ ـ إلى اليسار من 1– ،
و ـ بين 1– و 1+ .

٣) أوجد العدد c بحيث يكون:
أ ـ $P(Z < c) = 0.8643$ ،
ب ـ $P(Z < c) = 0.2266$ ،
جـ ـ $P(Z \geq -c) = 0.6554$ ،
د ـ $P(Z < c) = 0.05$
هـ ـ $P(-c < Z < c) = 0.90$ ،
و ـ $P(-c < Z < c) = 0.95$ ،
ز ـ $P(-c < Z < c) = 0.99$.

٤) إذا رمزنا بـ Z_α لقيمة المتغير الطبيعي المعياري Z التي تحصر إلى اليمين منها مساحة تساوي α ، فاحسب $Z_{0.005}$ ، $Z_{0.025}$، $Z_{0.05}$، $Z_{0.02}$، $Z_{0.01}$، $Z_{0.10}$.

٥) متغير عشوائي X يتبع التوزيع الطبيعي $N(16, 7)$ ، احسب $P(|X-16| > 3)$.

٦) متغير عشوائي X يتبع التوزيع $N(50, 25)$ ، احسب: $P(X > 62)$، $P(|X-50| < 8)$، $P(X = 60)$، $P(|X-40| > 5)$.

٧) تتوزع معدلات مجتمع كبير من طلبة الكليات تقريبا وفق التوزيع $N(2.4, 0.64)$. ما نسبة الطلاب الذين تتجاوز معدلاتهم 3.0 ؟ (المعدل التام هو 4).

٨) بالإشارة إلى المسألة السابقة إذا شطبت أسماء الطلاب الذين تقل معدلاتهم عن 1.9 فكم ستبلغ نسبة الأسماء المشطوبة؟

٩) متغير عشوائي X يتبع التوزيع الطبيعي . إذا كان $E(X^2) = 68$ و $P(X < 10) = 0.8413$ فاحسب μ و σ^2 .

١٠) يتوزع عمر نوع من الغسالات مقدرا بالسنوات وفق التوزيع الطبيعي $N(3.1, 1.2)$. إذا كانت الغسالات مكفولة لمدة سنة ، فما هي نسبة الغسالات المباعة التي سيضطر المصنع إلى استبدالها بغسالة جديدة؟

١١) وجدنا أن الفترة الزمنية الضرورية لإتمام اختبار ذكاء مخصص لطلبة الكليات تتوزع وفق التوزيع الطبيعي بمتوسط 70 دقيقة وانحراف معياري يساوي 12 دقيقة . كيف يجب تحديد زمن الاختبار إذا أردنا إتاحة وقت كاف لإتمام الاختبار لـ 90% من الطلاب المتقدمين؟

١٢) نظمت آلة لتقديم شراب مرطب بحيث تضع ، في المتوسط ، μ أونزة في الكأس الواحدة . إذا كان ما تضعه بالفعل في الكأس الواحدة متغيرا طبيعيا بانحراف معياري $\sigma = 0.3$ أونزة . فما القيمة التي ينبغي تحديدها لـ μ بحيث تفيض الكؤوس ذات السعة 8 أونزة بنسبة 1% فقط ؟

١٣) وزن بيضة الدجاج بالغرام يتوزع وفق التوزيع الطبيعي $N(60, 225)$. وتصنّف البيضة «صغيرة» إذا قل وزنها عن 45 غراما ، إذا رغبت أن يصنف باقي البيض بالتساوي بين عادي وكبير، إقترح الوزن الذي يفصل بين هذين الصنفين مقربا إلى أقرب غرام .

١٤) تتوزع أوزان قوالب الصابون في مصنع طبيعيا . وفي الأسبوع الماضي كان وزن $6\frac{2}{3}\%$ من القوالب المصنوعة أقل من 90.5 غراما بينما زاد وزن 4% من القوالب على 100.25 غراما . والمطلوب :

أ ـ أوجد متوسط وتباين توزيع وزن القالب ، والنسبة المئوية للقوالب التي يتوقع أن تزن أقل من 88 غراما .

ب ـ إذا خفضنا تباين الوزن بنسبة الثلث فما هي النسبة المئوية من إنتاج الأسبوع القادم التي تتوقع أن يقل وزنها عن 88 غراما. مفترضا أن المتوسط لم يتغير؟

١٥) يقدر أن 1400 راكبا ممن يبدلون قطارهم في محطة معينة يهدفون بصورة منتظمة إلى اللحاق بقطار الخامسة والنصف مساء، وأن 50 راكبا يصلون قبل الساعة الخامسة وعشرين دقيقة مساء، موعد فتح البوابة الخاصة بهذا القطار، وأن 70 راكبا يفوتهم القطار عند التزامه التام بموعد المغادرة. مفترضا أن زمن وصول الراكب إلى المحطة متغير عشوائي يتبع التوزيع الطبيعي، أحسب متوسط هذا التوزيع وتباينه. ومن ثم قدر:

ا ـ موعد فتح البوابة بحيث لا يزيد عدد المنتظرين أمامها على عشرين راكبا.

ب ـ عدد المستبدلين الذين سيفوتهم القطار في يوم يغادر فيه (على غير المتوقع) قبل الوقت المحدد بدقيقتين.

١٦) يغادر رجل منزله كل صباح الساعة السابعة كي يصل إلى عمله في الساعة الثامنة. وقد وجد خلال فترة طويلة أنه يتأخر عن عمله بنسبة مرة في كل أربعين مرة. وبدأ يغادر المنزل في الساعة السادسة وخمس وخمسين دقيقة فوجد خلال فترة مماثلة أنه يتأخر مرة في كل مائة مرة. بفرض أن الزمن الذي تستغرقه الرحلة يتوزع طبيعيا كيف ينبغي أن يحدد موعد المغادرة بحيث لا يتأخر أكثر من مرة كل ما ئتي مرة؟

١٧) في كتاب معين يمكن اعتبار عدد الكلمات في الصفحة الواحدة متغيرا طبيعيا، على وجه التقريب، بمتوسط 800 كلمة وانحراف معياري 50 كلمة. إذا اخترت عشوائيا ثلاث صفحات فما احتمال ألا تتضمن أي منها ما بين 830 إلى 845 كلمة؟

١٨) في بلد معين، متوسط طول الذكر البالغ 170 سم بانحراف معياري 10 سم، ومتوسط طول الأنثى البالغة 160 سم بانحراف معياري 8 سم، وبالنسبة لكل

مـن الجنسين يعتبر التوزيـع الطبيعي نموذجا منـاسبا لوصف تغير الطول . بفرض أن الطول ليس من العـوامل التي تؤخـذ في الاعتبـار عنـد اختيار الزوجة أو الزوج . أحسب احتمـال أن زوجـا وزوجته اخترناهما عشـوائيا سيكـون كل منهما أطول من 164 سم .

١٩) في بستـان للبرتقال متـوسط وزن الثمـرة 19.3 أونزة بـانحراف معيـاري 2.3 أونزة . مفترضا أن وزن الثمرة متغير يتبع التوزيع الطبيعي ، أحسب :

ا ـ نسبة الثمار التي يقل وزنها عن 18 أونزة .

ب ـ نسبة الثمار التي لا يقل وزنها عن 20 أونزة .

جـ ـ نسبة الثمار التي يتراوح وزنها بين 18.5 و 20.5 أونزة .

د ـ الوزن الذي سيقل عنه 15% من الثمار.

هـ ـ الوزن الذي سيزيد عليه 25% من الثمار.

٢٠) ملاحظة عدد كبير من السيارات عند نقطة محددة من طريق عام بينت لنا أن السرع تتوزع طبيعيا . إذا علمت أن سرعة 90% من السيارات تقل عن 124.3 كم/ سا ، وأن سرعة 5% فقط من السيـارات تقل عن 101 كم/ س . حـدد السرعة المتـوسطة μ والانحراف المعياري σ .

٢١) من المفترض أن يكون قطـر كريات معدنيـة تنتجها شركة صناعيـة مساويا 2 مم . ولكن الكريات ستكون مقبولـة إذا تراوحت أقطارها بين 1.90 مم و 2.10 مم . وقد لوحظ في دفعـة إنتاج كبيرة أن 2.5% منها مرفوض لأنـه أكبر مما يمكن التساهل فيه وأن 2.5% منها مرفوض لأنه أصغر مما يمكن التساهل فيه . حدد، بصورة تقريبية ، ما ستصبحه نسبتا الرفض إذا غيرنا حدود التساهل إلى 1.95 مم و 2.15 مم .

٢٢) تتوزع درجات امتحان وفق التوزيع الطبيعي $N(50, 100)$ ، ونرغب في إعادة النظر في سلـم الدرجات بحيث تكـون درجة النجاح 40 ونسبة الناجحين 70% ، ودرجة

التفوق 70 ونسبة المتفوقين 20% . أحسب الدرجة الجديدة لمتقدم للامتحان كانت درجته الأصلية 60 .

٢٣) يمكن تصنيف البيض إلى عادي إذا كان الوزن أقل من 46 غراما، ومتوسط إذا كان الوزن بين 46 و 56 غراما، وكبير إذا كان الوزن أكبر من 56 غراما. لنفرض أن البيض الذي تضعه سلالة معينة من الدجاج يتوزع، من حيث وزن البيضة، وفق التوزيع الطبيعي $N(50, 25)$. أحسب نسبة كل صنف من الأصناف الثلاثة. وإذا كانت أسعار البيع للبيضة الواحدة من الأصناف الثلاثة هي، على الترتيب، 4 هللة، 5 هللة،6 هللة. وكانت كلفة الإنتاج 4 هللة لكل بيضة، فما الربح المتوقع للبيضة الواحدة؟

وبالنسبة لسلالة أخرى من الدجاج فإنها تضع بيضا يتبع، من حيث الوزن، التوزيع الطبيعي $N(52, 25)$. إلا أنه يستهلك أكثر من الطعام مما يرفع كلفة البيضة إلى 4.5 هللة. ما الربح المتوقع للبيضة الواحدة في هذه السلالة؟

٢٤) لنفرض أن مقاس الحذاء لذكر بالغ هو عدد صحيح k يرتبط بطول القدم، y ، مقاسا بالبوصة بالعبارة التالية: «حذاء مقاسه k سيكون مناسبا لقدم طولها يتراوح بين $0.5k + 5.5$ و $0.5k + 6$ ؛ حيث $k = 5, 6, \ldots, 14$ ». ويمكن اعتبار y ، طول قدم ذكر بالغ، متغيرا يتبع التوزيع الطبيعي $N(10.2, 1.21)$.

ا ـ ما النسبة من مجتمع الذكور البالغين التي تتطلب حذاء مقاسه أكبر من 14؟
ب ـ ما المقاس الأكثر تواترا وما نسبة أولئك الذين يطلبون هذا المقاس؟

٢٥) حدود التساهل في طول قطعة مصنعة هي 10.00 ± 0.05 مم. وتُفحص كل قطعة يجري إنتاجها لرؤية ما إذا كانت تحقق هذه الحدود أم لا. والتوزيع الاحتمالي لطول القطعة هو التوزيع الطبيعي بمتوسط يساوي 10.01 مم وانحراف معياري 0.04 مم. وكلفة إنتاج القطعة 10 ريالات. وجميع القطع التي لا يقع طولها ضمن حدود

التساهل تهمل وتعتبر خسارة للشركة المصنعة. ولتخفيض حجم الخسارة يمكن:

ا ـ إزالة الانحياز في عمل الآلة وجعل متوسط التوزيع $\mu = 10$ وذلك بكلفة إضافية قدرها 4 ريالات لكل قطعة.

ب ـ تخفيض الانحراف المعياري إلى 0.03 وذلك بكلفة إضافية قدرها ريالان لكل قطعة.

جـ ـ القيام بالإجرائين (ا) و (ب) معا لقاء كلفة إضافية 6 ريالات للقطعة الواحدة. إذا كنت تعمل في قسم الإحصاء في هذه الشركة فبأي الإجراءات الثلاثة المذكورة تنصح؟

٢٦) تقضي مواصفات الإنتاج لعبوات نوع معين من الحلويات أن وزن كل عبوة يجب أن يقع بين 140 غ و 160 غ. إذا كان وزن العبوة يتوزع طبيعيا بتباين يساوي 4غ^{2}. كيف تحدد متوسط التوزيع الذي ينبغي أن تهدف إليه الشركة المنتجة ولماذا؟

٢٧) يستخدم أحد المصانع 2000 مصباح كهربائي للإضاءة. وعمر المصباح الكهربائي مقاسا بالساعات يتبع التوزيع الطبيعي $N(550, 2500)$. وحرصا على وجود عدد قليل من المصابيح المحترقة خلال أوقات الإنتاج يستبدل المصنع المصابيح جميعها كل فترة وبصورة دورية. كيف ينبغي تحديد طول فترة الاستبدال لكي لا يوجد في المصنع في أي وقت أكثر من 20 مصباحا محروقا؟

ومع نوع أفضل من المصابيح حيث يتوزع عمر المصباح وفق التوزيع الطبيعي $N(600, 1600)$ تتغير فترة الاستبدال إلى 500 ساعة، بين أن عدد المصابيح المحترقة في المصنع في أي وقت سينخفض عندئذ إلى حوالي 12 مصباحا.

٢٨) مبيعات بقال من سلعة معينة كل أسبوعين هي متغير يتبع التوزيع الطبيعي بمتوسط يساوي 200 كغ وتباين يساوي 225 كغ. أوجد احتمال أن تكون مبيعاته من هذه السلعة خلال أسبوعين أقل من 185 كغ، وعندما يطلب مزيدا من هذه

السلعة تأخذ عملية تسليم البضاعة المطلوبة فترة أسبوعين. حدد إلى أقرب كيلوغرام المخزون الذي ينبغي تأمينه من هذه السلعة عند إعادة طلبها بحيث يكون البقال مطمئنا باحتمال 0.95 إلى أن هذه السلعة لن تنفذ قبل وصول الطلب.

(٥ ـ ٤) خواص التوزيع الطبيعي وبعض التطبيقات*

اصطلحنا على كتابة $N(\mu, \sigma^2)$ لتعني توزيعا طبيعيا بمتوسط يساوي μ وتباين يساوي σ^2 . وهكذا نكتب، على سبيل المثال: X متغير $N(8, 4)$ لتعني أن X متغير يتبع التوزيع الطبيعي بمتوسط يساوي 8 وتباين يساوي 4 . وفيما يلي بعض خواص التوزيع الطبيعي:

١ ـ ليكن X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ و $N(\mu_2, \sigma_2^2)$ ، على الترتيب. فعندئذ يكون مجموعهما $X+Y$ ، ولنرمز له بـ U ، متغيرا $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. أي متغيرا طبيعيا أيضا بمتوسط يساوي مجموع المتوسطين وتباين يساوي مجموع التباينين.

٢ ـ وبصورة أعم إذا كان X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ و $N(\mu_2, \sigma_2^2)$ ، على الترتيب فإن المتغير $U = aX + bY + c$ ، حيث a ، b ، c أية أعداد حقيقية، هو بدوره متغير طبيعي متوسطه، حسب خواص التوقع:

$$
\begin{aligned}
E(U) = E(aX+bY+c) &= a\,E(X) + b\,E(Y) + c \\
&= a\,\mu_1 + b\,\mu_2 + c
\end{aligned}
$$

وتباينه حسب خواص التباين:

$$
\begin{aligned}
V(U) = V(aX+bY+c) = V(aX+bY) &= V(aX) + V(bY) \\
&= a^2 V(X) + b^2 V(Y) \\
&= a^2 \sigma_1^2 + b^2 \sigma_2^2
\end{aligned}
$$

* للقراءة فقط.

ونكتب باختصار:

إذا كان X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ ، $N(\mu_2, \sigma_2^2)$ وكانت a ، b ، c أية أعداد ثابتة فإن $U = aX + bY + c$ يكون متغيرا

$$N(a\mu_1 + b\mu_2 + c, \; a^2\sigma_1^2 + b\sigma_2^2)$$

وعلى سبيل المثال إذا كان X متغيرا $N(15, 2)$ و Y متغيرا $N(-7, 4)$ فإن $U = 2X - 3Y + 1$ وهو متغير طبيعي متوسطه يساوي

$$2(15) - 3(-7) + 1 = 52$$

وتباينه

$$2^2(2) + (-3)^2(4) = 44$$

أي أن U متغير $N(52, 44)$.

٣ ـ ويمكن بوضوح تعميم الخاصة ٢ إلى أكثر من متغيرين، لتصبح في الحالة الخاصة التالية، وهي في حد ذاتها بالغة الأهمية، كما يلي:

إذا كانت $X_1, X_2, \ldots, X_n$ متغيرات مستقلة وكل منها $N(\mu, \sigma^2)$ ، [أي إذا كانت $X_1, X_2, \ldots, X_n$ عينة عشوائية من $N(\mu, \sigma^2)$] فإن:

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$$ يكون متغيرا $N(n\mu, n\sigma^2)$ ،

ويكون

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$ متغيرا $N\left(\mu, \frac{\sigma^2}{n}\right)$.

وتجدر ملاحظة أنه بالرغم من أن المتغير الطبيعي يتحول بين $-\infty$ و $+\infty$ ، إلا أنه يمكن استخدامه استخداما مقبولا تماما لوصف متغير، X ، موجب بطبيعته. وذلك

شريطة أن يكـون $P(X \le 0)$ عـددا صغيرا جدا يمكن إهمالـه. أي أننا نتجـاوز المقولـة الدقيقة بأن $P(X \le 0) = 0$ ، وتعني استحالة أن يكون X سالبا إلى مقولة ، تقريبية وعملية في آن واحد، تكتفـي بالتأكيـد على أن احتمال أن يكون X سالبا هـو احتمال قريب من الصفر. وبما أن

$$p(X \le 0) = F\left(\frac{0-\mu}{\sigma}\right) = F\left(\frac{-\mu}{\sigma}\right)$$

وأن σ موجب ، فإن $P(X \le 0)$ سيكون مهملا إذا كان μ كبيرا بالمقارنة مع σ .

وعلى سبيل المثال، إذا كان $\mu = 4.5\sigma$ فإن $P(X \le 0)$ يكون أقل من 0.000005 ، وهو صغير إلى الحد الذي يجعله غير ذي بال في التطبيقات العملية.

مثال (٥ ـ ٩)

إذا كانت X ، Y ، T متغيرات مستقلة $N(2, 1)$ ، $N(3, 2)$ ، $N(4, 3)$ ، على الترتيب ، فاحسب:

أ ـ $P(1 < X < 3)$ ،

ب ـ $P(X \le Y)$ ،

جـ ـ $P(3X - 2Y > 1)$ ،

د ـ $P(X + Y < 2T - 4)$ ،

الحـــــل

أ ـ
$$P(1 < X < 3) = F(3-2) - F(1-2) = F(1) - F(-1)$$
$$= 2F(1) - 1 = 0.6826$$

ب ـ
$$P(X \le Y) = P(X - Y \le 0)$$

ولكن $X - Y$ متغير $N(-1, 3)$ وفق الخاصة ٢ . وبالتالي يكون

$$P(X \le Y) = P(X - Y \le 0) = F\left(\frac{0-(-1)}{\sqrt{3}}\right) = F(0.577) = 0.718$$

جـ ـ وفق الخاصة ٢ نجد أن $3X - 2Y$ متغير $N(0, 17)$ وهكذا نجد :

$$P(3X - 2Y > 1) = 1 - P(3X - 2Y \le 1) = 1 - F\left(\frac{1 - 0}{\sqrt{17}}\right)$$
$$= 1 - F(0.243) = 0.404$$

د ـ ووفق الخاصة ٣ يكون $X + 2Y - 2T$ متغيرا $N(-3, 15)$ ، وبالتالي :

$$P(X + Y \le 2T - 4) = P(X + Y - 2T \le -4) = F\left(\frac{-4 - (-3)}{\sqrt{15}}\right)$$
$$= 1 - F\left(\frac{1}{\sqrt{15}}\right) = 1 - F(0.258) = 0.398 .$$

مثال (٥ ـ ١٠)

يتم إنتاج مسامير البرشام التي تستخدم لبرشمة صفيحة معدنية بطريقة تسمح لنا بوصف قطر المسمار X كمتغير $N(3; 0.04)$. وبطريقة مستقلة يجري إنتاج صفائح معدنية ذات ثقوب دائرية يمكن اعتبار قطر الثقب Y متغيرا $N(3.2, 0.01)$. (القياس في الحالتين بالسنتمتر) .

ا ـ ما هو احتمال أن يناسب المسمار ثقب الصفيحة؟

ب ـ إذا اخترنا أربعة أزواج (مسمار ـ صفيحة) فما هو احتمال أن يكون زوجان منهما، على الأقل، متناسبين؟

الحل

ا ـ X و Y متغيران طبيعيان مستقلان . واحتمال تناسب المسمار مع الثقب هو:

$$P(X < Y) = P(X - Y < 0)$$

ولكن $X - Y$ متغير $N(-0.2, 0.05)$ ، وبالتالي :

$$P(X - Y < 0) = F\left(\frac{0 - (-0.2)}{\sqrt{0.05}}\right) = F(0.894) = 0.814$$

ب ـ يمكننا اعتبار إنتاج مسمار وصفيحة تكرارا لتجربة ثنائية احتمال النجاح فيها $p = 0.814$ ، $n = 4$ ، وإذا رمزنا بـ U لعدد الأزواج المتناسبة ، يصبح المطلوب :

$$P(U\geq 2) = 1 - P(U=0) - P(U=1)$$
$$= 1 - (0.186)^4 - 4\,(0.814)\,(0.186)^3 = 0.978\,.$$

مثال (٥ ـ ١١)

$$P(|\bar{X} - \mu| > 2) = P(\bar{X} - \mu > 2) + P(\bar{X} - \mu < -2) \leq 0.025$$

أخذنا عينة عشوائية حجمها n من مجتمع طبيعي فيه $\mu = 10$ و $\sigma = 20$ ، ما هي أصغر قيمة ممكنة لـ n بحيث لا يزيد عن 0.025 احتمال أن يتجاوز الفرق بين متوسطي العينة والمجتمع المقدار 2 ؟

الحل

ليكن $\bar{X}$ متوسط العينة. نعلم من الخاصة ٣ أن $\bar{X}$ متغير $N\left(10, \frac{400}{n}\right)$. والمطلوب تحديد حجم العينة n بحيث يكون، $p(|\bar{X} - \mu| > 2) \leq 0.025$ ولكن الحادثة $|\bar{X} - \mu| > 2$ تعني إما $\bar{X} - \mu > 2$ أو $X - \mu < -2$ ، وبالتالي:

$$P(|\bar{X} - \mu| > 2) = P(\bar{X} - \mu > 2) + P(\bar{X} - \mu < -2) \leq 0.025$$

أو

$$P\left(\frac{\bar{X} - \mu}{20/\sqrt{n}} > \frac{2}{20/\sqrt{n}}\right) + P\left(\frac{\bar{X} - \mu}{20/\sqrt{n}} < \frac{-2}{20/\sqrt{n}}\right) \leq 0.025$$

أو

$$1 - F\left(\frac{2\sqrt{n}}{20}\right) + F\left(\frac{-2\sqrt{n}}{20}\right) \leq 0.025$$

أو

$$2 - 2F\left(\frac{\sqrt{n}}{10}\right) \leq 0.025$$

$$F\left(\frac{\sqrt{n}}{10}\right) \geq 0.9875$$

ونجد من الجدول أن

$$\frac{\sqrt{n}}{10} \geq 2.24$$

$$\sqrt{n} \geq 22.4 \Leftrightarrow n \geq 501.76$$

أي أن حجم العينة ينبغي ألا يقل عن 502 .

مثال (٥ ـ ١٢)

ليكن X متغيرا عشوائيا يتبع التوزيع الطبيعي $N(100, 25)$. أحسب

$P(|\bar{X} - 100| > 1)$ ، إذا كان :

ا ـ $\bar{X}$ متوسط عينة حجمها $n = 25$ ،

ب ـ $\bar{X}$ متوسط عينة حجمها $n = 100$.

الحل

ا ـ بالاستناد إلى الخاصة ٣ نعلم أن $\bar{X}$ متغير يتبع التوزيع الطبيعي $N(100, 1)$

ويكون

$$\begin{aligned} P(|\bar{X} - 100| > 1) &= P(\bar{X} - 100 > 1) + P(\bar{X} - 100 < -1) \\ &= 1 - F(1) + F(-1) \\ &= 1 - F(1) + [1 - F(1)] = 2 - 2F(1) \\ &= 2 - 2 \times 0.8413 = 0.3174 \end{aligned}$$

ب ـ $\bar{X}$ يتبع الآن التوزيع الطبيعي $N(100, 0.25)$ ومنه :

$$\begin{aligned} P(|\bar{X} - 100| > 1) &= P\left(\frac{\bar{X} - 100}{0.5} > \frac{1}{0.5}\right) + P\left(\frac{\bar{X} - 100}{0.5} < \frac{-1}{0.5}\right) \\ &= 1 - F(2) + F(-2) \\ &= 2[1 - F(2)] = 0.0456 \end{aligned}$$

تمارين (٥ ـ ٣)

١) في المثال (٥ ـ ١٢) ، كم يجب أن يكون حجم العينة n ليصبح :

ا ـ $P(|\bar{X} - 100| > 0.5) \leq 0.01$ ،

ب ـ $P(|\bar{X} - 100| > 0.5) \leq 0.001$.

٢) إذا افترضنا أن الدرجات في امتحان عام تتوزع ، على وجه التقريب ، وفق التوزيع الطبيعي $N\,(72,\ 100)$ ففي مجموعة عشوائية تتضمن مائة طالب ممن أدوا هذا الامتحان ، ما احتمال أن يختلف متوسط درجاتهم عن 72 بأكثر من 3 درجات؟

٣) ما أصغر حجم عينة ينبغي أخذها من مجتمع طبيعي فيه $\mu = 10$ و $\sigma = 20$ ، كي لا يزيد احتمال تجاوز متوسط العينة لضعف متوسط المجتمع عن 0.025 ؟

٤) إذا كان X ، Y و Z ثلاثة متغيرات مستقلة وتوزيعاتها ، على الترتيب ، $N(2, 2)$ ، $N(3, 3)$ ، $N(4, 4)$ ، فاحسب :

أ ـ $P(1 \le X \le 4)$ ،

ب ـ $P(X - 2 \le 4)$ ،

جـ ـ $P(2X + Y \ge 5)$ ،

د ـ $P(Z + 2 \le 4X - Y \le +3)$ ،

هـ ـ $P(X \ge Y,\ Z - 3 > 0)$.

٥) يتوزع المتغيران المستقلان X و Y وفق $N(\mu, \sigma^2)$ و $N(2\mu, 2\sigma^2)$ ، على الترتيب .

ا ـ إذا كان $\sigma = 3$ و $P(X + 2Y \le 10)$ فاحسب μ .

ب ـ إذا كان $\mu = 0$ و $P(4X - Y < 3) = 0.4$ فاحسب σ .

جـ ـ إذا كان $P(|2X - Y| > 10) = 0.05$ و $P(Y \le s) = 0.9$ فاحسب μ و σ .

٦) يتوزع طول نصف قطر دولاب صغير ينتجه مصنع معين وفق التوزيع الطبيعي $N(1, 0.0001)$ (القياس بالسنتيمتر) . ويتم إنتاج الدواليب بصورة مستقلة ثم تجمع عقب ظهورها في خط الإنتاج أزواجا . ونعتبر أن الزوج من الدواليب مُرض إذا اختلف نصفا القطرين للدولابين بأقل من 0.03 سم .

ا ـ ما نسبة الأزواج المرضية من الدواليب ؟

ب ـ من بين خمسة أزواج ما احتمال أن يكون أحدها على الأقل غير مُرض ؟

جـ ـ إلى أي حد ينبغي تخفيض الانحراف المعياري لطريقة الإنتاج كي تصبح نسبة الأزواج المرضية 99% ؟

٧) وجد طبيب يعمل في عيادة أن الأوقات التي تستغرقها استشارات المرضى مستقلة بعضها عن بعض، وتتبع التوزيع الطبيعي بمتوسط يساوي 5 دقائق وانحراف معياري 1.5 دقيقة. ويقابل مرضاه، على التتالي، بدون فواصل زمنية بين مريضين، مبتدئا عمله الساعة العاشرة صباحا. ما الموعد الذي ينبغي للمريض العاشر أن يرتبه مع سيارة أجرة بحيث يطمئن باحتمال 99% أن السيارة سوف لا تنتظره ؟ وإذا كان الطبيب سيقابل 22 مريضا قبل انصرافه، فما احتمال مغادرته للعيادة قبل الساعة 12 ظهرا ؟

٨) عمر قطعة إلكترونية مقاسا بالساعات يتوزع وفق التوزيع الطبيعي، لنفرض أن 92.5% من هذه القطع يتجاوز عمرها 2160 ساعة و 3.92% يتجاوز عمرها 17040 ساعة.

ا ـ أحسب متوسط التوزيع وانحرافه المعياري.

ب ـ إذا أخذنا عينة من 100 قطعة فاحسب احتمال أن يكون متوسط العمر في العينة:

(i) أكبر من 10000 ساعة،

(ii) أقل من 8000 ساعة.

(iii) واقعا بين 8000 و 10000 ساعة.

٩) الأجر الأسبوعي بالريال الذي تدفعه شركة إلى عمالها يتوزع تقريبا وفق التوزيع الطبيعي $N(200, 324)$.

أ ـ أحسب احتمال ألا يختلف متوسط الأجر الأسبوعي لعينة عشوائية من 9 عمال عن متوسط المجتمع 200 بأكثر من 12 ريالا.

ب ـ كم يجب أن يكون حجم العينة حتى لا يختلف متوسطها عن متوسط المجتمع بأكثر من ستة ريالات إلا بنسبة بسيطة لا تتجاوز 10% ؟

١٠) على مدير شركة أن يقابل 20 مرشحا لوظيفة. ويعلم من تجربته السابقة أن وقت المقابلة مقاسا بالدقيقة يتوزع وفق التوزيع الطبيعي $N\ (10, 9)$. ويبدأ مقابلاته الساعة التاسعة صباحا. في أي وقت ينبغي له أن يطلب فنجان القهوة ويرتاح لمدة ربع ساعة إذا أراد أن يكون مطمئنا باحتمال 99% إلى أنه قد انتهى في ذلك الوقت من مقابلة 50% من المرشحين؟ وما احتمال أن ينتهي من كل المقابلات عند الساعة الواحدة بعد الظهر؟

١١) يتوزع وزن أمتعة المسافر جوا وفق التوزيع الطبيعي بمتوسط 20 كغ وانحراف معياري 5 كغ . ويتسع نوع معين من الطائرات لـ 100 راكب. ما هو احتمال أن يتجاوز الوزن الكلي لأمتعة المسافرين 2150 كغ؟

١٢) عمر صمام كهربائي مقاس بالساعة يتبع التوزيع الطبيعي $N\,(200, \sigma^2)$. إذا اشترى شخص عشر صمامات وأراد باحتمال 0.95 ألا يقل متوسط عمر الصمامات العشرة عن 190 ساعة، فما هي أكبر قيمة يمكن أن يأخذها الانحراف المعياري σ؟

١٣) بالإشارة إلى التمرين رقم ٢٨ من مجموعة التمارين (٥ ـ ٢)، لنفرض أن خمس بقالات متجاورة متضامنة بالنسبة إلى توفير تلك السلعة للزبائن. وأن مبيعاتها خلال أسبوعين من تلك السلعة مستقلة بعضها عن بعض وأن كلا منها تتبع

التوزيع الطبيعي بمتوسطات هي 200 ، 240 ، 180 ، 260 ، و 320 كغ ، وتباينات هي ، على الترتيب ، 225 ، 240 ، 225 ، 265 ، 270 كغ . اكتب متوسط وتباين الطلب على السلعة خلال أسبوعين ، وحدد إلى ثلاثة أرقام معنوية المستوى الإجمالي لمخزونها من تلك السلعة الذي ينبغي توفره عند طلب بضاعة جديدة بحيث يكون احتمال عدم نفاذها 0.99 .

احسب احتمال أن يتجاوز مجموع مبيعات البقالات الخمس من تلك السلعة خلال عشرة أسابيع 6200 كغ .

١٤) مصنع مربيات يضع في كل عبوة ثماني علب من ثمانية أنواع مختلفة . والمفروض أن تزن كل علبة 50 غراما . ولكن عمليا يتبع وزن كل علبة التوزيع الطبيعي $N(52, 1.21)$ ، وبصورة مستقلة من نوع إلى آخر.

أ ـ ما نسبة العلب التي تزن أقل من 50 غراما؟

ب ـ ما نسبة العبوات التي تقل عن 400 غراما؟

جـ ـ ما احتمال أن تزن واحدة أو أكثر من العلب ضمن عبوة أقل من 50 غراما ؟

د ـ كم ينبغي أن يكون الانحراف المعياري لوزن العلبة إذا أردنا لـ 99% من العبوات أن تزن أكثر من 400 غراما؟

١٥) أوزان الأشخاص الذين يستخدمون مصعدا معينا تتوزع وفق التوزيع الطبيعي بمتوسط يساوي 150 ليبرة وانحراف معياري 20 ليبرة . والحد الأعلى المسموح لحمولة المصعد هو 650 ليبرة .

ا ـ بصورة عشوائية ، يجتمع أربعة أشخاص في المصعد . ما هو احتمال تجاوز الحمولة القصوى؟

ب ـ بصورة عشوائية يوجد شخص واحد في المصعد ومعه أمتعة تزن ثلاثة أمثال

وزنه، ما هو احتمال تجاوز الحمولة القصوى؟
فسر أي اختلاف بين جوابيك في (ا) و (ب).

(٥ ـ ٥) نظرية النهاية المركزية

تعرض نظرية النهاية المركزية، وتحت شروط عامة جدا، أن كلا من مجموع ومتوسط عينة عشوائية، مسحوبة من مجتمع ما، يمتلك عند تكرار هذه العينات عددا كبيرا من المرات، توزيعا له، على وجه التقريب، شكل الجرس. وربما كان من الأفضل إيضاح هذه العبارة بمثال.

لنعتبر المجتمع المتولد عن قذف حجر نرد عددا كبيرا جدا من المرات. وقد رأينا توزيعه في المثال (٣ ـ ٦). لنسحب عينة من خمسة قياسات، $n = 5$، من المجتمع وذلك بقذف حجر النرد خمس مرات وتسجيل الملاحظات الخمس الناتجة. ثم نحسب مجموع هذه الملاحظات الخمس Σx_i ومتوسطها $\bar{x}$، ويبين الجدول (٥ ـ ١) نتائج تكرار هذه العملية مائتي مرة. كما يبين الشكل (٥ ـ ٨) المدرج التكراري للقيم المائتين لـ $\bar{x}$ (أو Σx_i). وتنبغي ملاحظة النتيجة المهمة التالية:

بالرغم من أن التوزيع الاحتمالي لـ X له شكل أفقي تماما، إلا أن المدرج التكراري لمائتين من قيم $\bar{X}$ (وهو يقدم صورة أولية عن شكل التوزيع الاحتمالي للمتغير $\bar{X}$ أو للمتغير ΣX_i) يتخذ شكلا مقببا قريبا من شكل الجرس، وكلما زاد حجم العينات المسحوبة عن خمسة اعتدل شكل المضلع التكراري ليقترب أكثر فأكثر من شكل التوزيع الطبيعي. وبعبارة أخرى، لو أننا أخذنا $n = 10$ في مثالنا، أي لو أننا قذفنا حجر النرد عشر مرات بدلا من خمس، ثم سجلنا نتائج مائتي عينة من هذا الحجم، ورسمنا المدرج التكراري للقيم المائتين لـ $\bar{X}$، فمن المتوقع الحصول على

شكل أكثر قربا من شكل الجرس. ولا بد من ملاحظة أنه للحصول على فكرة أدق عن شكل التوزيع الإحتمالي لـ $\bar{x}$ نحتاج، نظريا، إلى عدد لا نهائي من العينات، أو لنقل، بصورة عملية، إننا نحتاج إلى عدد من العينات أكبر بكثير من المائتين التي تضمنتها التجربة هنا. ومع ذلك فإن الشكل الذي تقدمه العينات المائتان كاف

جدول (٥ ـ ١): مئتا عينة من مجتمع قذف حجر نرد

رقم العينة	قياسات العينة	$\sum x_i$	$\bar{x}$	رقم العينة	قياسات العينة	$\sum x_i$	$\bar{x}$
1	3,1,6,4,1	15	3.0	33	6,3,5,4,5	23	4.6
2	4,6,6,5,2	23	4.6	34	6,5,3,3,3	20	4.0
3	5,5,2,5,2	19	3.8	35	2,6,2,6,3	19	3.8
4	4,4,5,2,2	17	3.4	36	2,2,1,6,6	17	3.4
5	2,3,6,3,3	17	3.4	37	4,3,2,5,4	18	3.6
6	6,6,2,5,4	23	4.6	38	5,1,2,5,6	19	3.8
7	6,3,3,2,6	20	4.0	39	5,5,2,5,6	23	4.6
8	3,1,5,1,5	15	3.0	40	5,6,6,5,2	24	4.8
9	6,2,5,5,4	22	4.4	41	3,1,6,3,6	19	3.8
10	6,5,6,6,6	29	5.8	42	1,6,2,6,1	17	3.4
11	6,6,1,1,2	16	3.2	43	3,2,3,4,6	18	3.6
12	1,4,1,4,6	16	3.2	44	3,2,5,1,6	17	3.4
13	4,6,3,5,5	23	4.6	45	4,6,5,3,2	20	4.0
14	4,3,3,4,5	19	3.8	46	6,2,5,4,5	22	4.4
15	4,6,2,3,1	16	3.2	47	6,1,1,2,5	15	3.0
16	1,4,3,4,5	17	3.4	48	1,1,5,5,2	14	2.8
17	3,4,3,1,4	15	3.0	49	2,2,3,3,4	14	2.8
18	3,3,3,6,4	19	3.8	50	5,4,2,2,1	14	2.8
19	6,3,4,4,6	21	4.2	51	3,5,1,5,3	17	3.4
20	5,4,2,2,6	19	3.8	52	5,2,3,3,2	15	3.0
21	4,5,5,2,2	18	3.6	53	4,1,5,2,6	18	3.6
22	1,5,2,3,1	12	2.4	54	5,4,4,2,4	19	3.8
23	3,5,6,5,3	22	4.4	55	4,5,2,1,4	16	3.2
24	5,3,6,4,3	21	4.2	56	4,5,6,3,1	19	3.8
25	6,2,3,2,5	18	3.6	57	3,5,5,1,4	18	3.6
26	5,4,5,1,6	21	4.2	58	6,6,5,3,4	24	4.8
27	4,1,6,2,6	19	3.8	59	6,3,2,5,4	20	4.0
28	6,6,6,2,2	22	4.4	60	4,6,5,1,1	17	3.4
29	3,4,2,1,5	15	3.0	61	5,1,1,2,2	11	2.2
30	1,2,2,3,3	11	2.2	62	2,6,2,2,3	15	3.0
31	6,5,1,6,2	20	4.0	63	2,4,4,1,1	12	2.4
32	6,3,1,2,5	17	3.4	64	3,1,2,2,2	10	2.0

تابع جدول (٥ ـ ١)

رقم العينة	قياسات العينة	Σx_i	$\bar{x}$	رقم العينة	قياسات العينة	Σx_i	$\bar{x}$
65	3,4,1,1,6	15	3.0	107	5,2,5,1,1	14	2.8
66	6,2,5,5,6	24	4.8	108	3,3,4,1,2	13	2.6
67	3,1,1,4,6	15	3.0	109	3,1,4,3,3	14	2.8
68	3,2,6,5,4	20	4.0	110	5,2,6,1,2	16	3.2
69	6,4,1,5,3	19	3.8	111	1,2,6,3,1	13	2.6
70	3,2,2,6,4	17	3.4	112	4,6,2,2,1	15	3.0
71	5,4,1,2,2	14	2.8	113	4,4,4,1,4	17	3.4
72	1,4,2,4,5	16	3.2	114	3,3,6,3,2	17	3.4
73	1,6,1,5,2	15	3.0	115	2,1,5,4,6	18	3.6
74	3,1,1,4,4	13	2.6	116	6,6,4,2,4	22	4.4
75	1,5,6,5,4	21	4.2	117	3,2,2,1,4	12	2.4
76	4,1,6,6,5	22	4.4	118	3,2,2,4,3	14	2.8
77	2,4,6,4,5	21	4.2	119	5,3,1,1,4	14	2.8
78	6,2,2,6,1	17	3.4	120	6,1,3,3,4	17	3.4
79	5,1,2,4,1	13	2.6	121	3,3,6,3,1	16	3.2
80	6,1,6,1,6	20	4.0	122	5,2,2,2,3	14	2.8
81	6,5,5,5,1	22	4.4	123	3,2,6,1,1	13	2.6
82	5,3,3,1,6	18	3.6	124	5,1,6,5,5	22	4.4
83	3,6,4,5,4	22	4.4	125	5,1,2,6,5	19	3.8
84	3,4,4,2,3	16	3.2	126	2,3,6,3,3	17	3.4
85	2,5,6,1,4	18	3.6	127	4,3,2,1,5	15	3.0
86	2,1,2,2,1	8	1.6	128	4,5,5,1,3	18	3.6
87	2,4,3,3,5	17	3.4	129	6,3,4,5,1	19	3.8
88	1,2,2,6,5	16	3.2	130	1,6,2,2,1	12	2.4
89	4,3,5,3,3	18	3.6	131	3,1,1,2,5	12	2.4
90	4,6,1,1,2	14	2.8	132	5,4,1,2,5	17	3.4
91	4,2,1,1,2	10	2.0	133	3,2,6,6,2	19	3.8
92	3,3,4,4,2	16	3.2	134	3,4,5,5,3	20	4.0
93	4,1,4,5,4	18	3.6	135	3,5,5,5,4	22	4.4
94	4,1,2,6,3	16	3.2	136	6,2,5,5,1	19	3.8
95	1,1,6,1,5	14	2.8	137	2,3,2,4,2	13	2.6
96	3,2,5,1,5	16	3.2	138	6,1,4,1,5	17	3.4
97	5,2,4,6,6	23	4.6	139	5,6,1,6,5	23	4.6
98	3,3,6,5,1	18	3.6	140	2,2,6,2,6	18	3.6
99	4,4,5,2,6	21	4.2	141	1,3,2,4,3	13	2.6
100	4,2,4,4,2	16	3.2	142	6,4,4,5,5	24	4.8
101	4,5,5,2,1	17	3.4	143	3,1,6,2,4	16	3.2
102	2,5,5,3,2	17	3.4	144	2,1,1,6,2	12	2.4
103	2,3,3,1,5	14	2.8	145	4,4,1,5,5	19	3.8
104	1,5,2,3,2	13	2.6	146	2,4,5,1,2	14	2.8
105	3,4,2,2,3	14	2.8	147	5,1,3,2,3	14	2.8
106	5,3,2,3,4	17	3.4	148	3,2,2,5,6	18	3.6

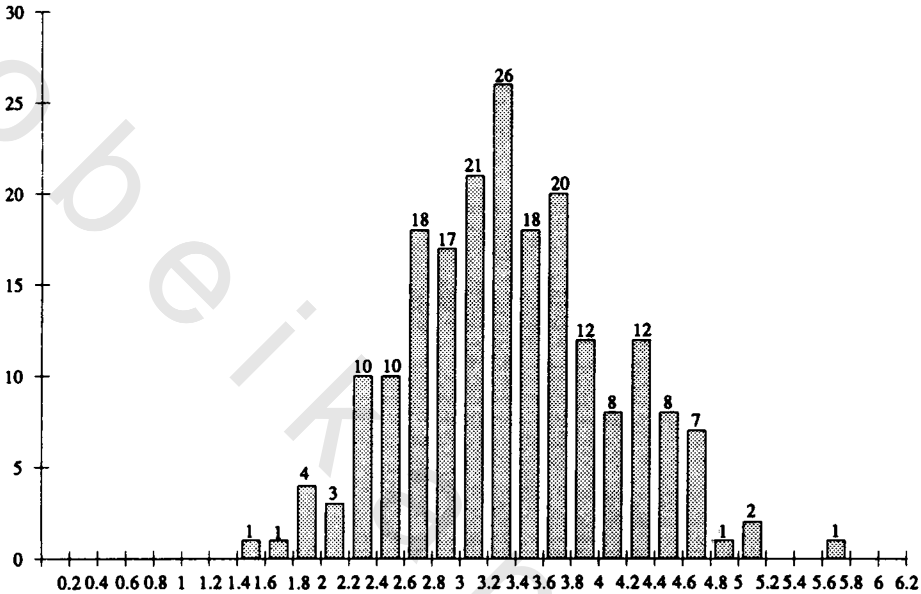
تابع جدول (٥ ـ ١)

رقم العينة	قياسات العينة	Σx_i	$\bar{x}$	رقم العينة	قياسات العينة	Σx_i	$\bar{x}$
149	1,3,6,1,3	14	2.8	175	2,4,2,2,2	12	2.4
150	6,3,1,4,6	20	3.8	176	4,6,6,6,2	24	4.8
151	3,6,6,1,3	19	3.8	177	3,6,5,4,4	22	4.4
152	3,5,2,6,2	18	3.6	178	2,3,4,4,3	16	3.2
153	3,1,2,2,5	13	2.6	179	2,6,5,3,5	21	4.2
154	4,6,4,3,3	20	4.0	180	6,3,5,2,1	17	3.4
155	1,4,2,4,3	14	2.8	181	4,3,2,2,1	12	2.4
156	5,5,4,6,4	24	4.8	182	3,5,2,2,3	15	3.0
157	4,1,4,4,3	16	3.2	183	4,3,6,1,2	16	3.2
158	3,2,1,5,5	16	3.2	184	5,5,1,6,2	19	3.8
159	5,6,1,3,5	20	4.0	185	6,2,3,3,2	16	3.2
160	2,5,6,3,3	19	3.8	186	1,4,4,4,2	15	3.0
161	1,4,2,5,3	15	3.0	187	5,6,3,6,4	24	4.8
162	4,2,4,3,5	18	3.6	188	5,1,3,5,3	17	3.4
163	1,2,5,2,6	16	3.2	189	4,4,1,3,5	17	3.4
164	1,1,3,5,2	12	2.4	190	5,3,1,2,4	15	3.0
165	3,5,3,4,5	20	4.0	191	1,1,1,6,1	10	2.0
166	3,1,2,2,4	12	2.4	192	4,5,4,4,6	23	4.6
167	2,4,3,5,2	16	3.2	193	5,2,6,6,6	25	5.0
168	2,6,3,5,3	19	3.8	194	5,6,5,5,5	26	5.2
169	5,4,3,1,1	14	2.8	195	6,5,1,6,4	22	4.4
170	6,2,6,6,6	26	5.2	196	4,2,3,4,6	21	4.2
171	1,5,5,1,1	13	2.6	197	5,2,4,2,2	15	3.0
172	3,5,5,3,1	17	3.4	198	2,3,3,3,6	18	3.6
173	1,2,2,3,1	9	1.8	199	6,1,4,5,2	18	3.6
174	2,1,4,1,2	10	2.0	200	2,3,1,1,4	11	2.2

لتوضيح الفكرة الأساسية التي تتضمنها نظرية النهاية المركزية، والتي نعرضها في العبارة المبسطة التالية:

(٥ ـ ٥ ـ ١) الفكرة الأساسية لنظرية النهاية المركزية

إذا سحبنا عينات عشوائيـة حجم كل منها n ، من مجتمع متـوسطه μ وانحرافه المعياري σ محدودان، فإن توزيع متـوسط العينة $\bar{x}$ يتطابق تقريبـا مع التوزيع الطبيعي بمتوسط يساوي μ وانحراف معياري $\frac{\sigma}{\sqrt{n}}$. وستزداد دقة التقريب كلما ازداد n .



شكل (٥ ـ ٨) مدرج التكرار لمتوسطات العينات المائتين المسحوبة من مجتمع قذف حجر النرد

ويمكن إعـادة صيـاغة النظريـة لتتفق مع $\sum_{1}^{n} X_i$ بـدلا مـن $\bar{X}$. فنقـول إن تـوزيع $\sum_{1}^{n} X_i$ يسعى أيضـا إلى أن يصبـح طبيعيـا بمتـوسط يسـاوي $n\,\mu$ وانحراف معياري $\sigma\sqrt{n}$ ، وذلك عندما يصبح n كبيرا جدا .

وتبدو أهمية نظرية النهاية المركزية من زاويتين، فهي توضح أولا نزوع العديد من المتغيرات العشوائيـة إلى أن يكون تـوزيعها، بصورة تقـريبية، هـو التوزيع الطبيعي. إذ يمكن، مثلا، أن نتصور طول الإنسان حصيلـة عدد كبير من المؤثرات العشوائية، مثل طـول الأب، وطول الأم، والمورثـات (وعـددهـا كبير)، ونشاط الغـدة أو الغـدد ذات العلاقـة بالطـول، والبيئة أو المحيط بأنواعـه، والتغذيـة، إلخ. وإذا كانت آثـار هذه العوامل، تضاف بعضها إلى بعض، لتنتج واقعا معينا بالنسبة إلى طول الإنسان فعندئذ

يمكن اعتبار الطول كحصيلة لعدد كبير من المتغيرات العشوائية. وهكذا تنطبق نظرية النهاية المركزية، ويكون توزيع متغير الطول هو، على وجه التقريب، التوزيع الطبيعي، وذلك بصرف النظر عن توزيع أي من المتغيرات العشوائية التي تؤثر في تحديد الطول. وهذه بالطبع محاولة للتعليل، ليس أكثر، إذ أن ما يجري في الواقع غير معروف لنا بصورة دقيقة، ولكن ما يمكن قوله، على كل حال، هو إن نظرية النهاية المركزية توضح سبب وجود العديد من المتغيرات العشوائية التي نصادفها في حياتنا العامة، والتي نعتبر أن توزيعها الاحتمالي هو التوزيع الطبيعي.

ومن زاوية أخرى نجد أن العطاء الأكثر أهمية لنظرية النهاية المركزية، يتعلق بمسألة الاستقراء الاحصائي. فالعديد من الإحصاءات التي تستخدم للقيام باستقراءات حول معلمات توزيع (وهي تمثل خصائص مهمة لمجتمع القياسات) مثل p، احتمال النجاح في التوزيع الثنائي، أو μ متوسط التوزيع الطبيعي إلخ. هذه الإحصاءات تأخذ شكل مجموع لقياسات العينة أو متوسط هذه القياسات. وإذا كان الحال كذلك، وكانت n كبيرة بكفاية، فيمكننا اعتبار التوزيع الطبيعي تقريبا جيدا للتوزيع الاحتمالي لذلك الإحصاء. وهو ما تمس الحاجة إليه عند القيام بأي استقراء إحصائي. وسنجد في الفقرات القادمة العديد من الاستخدامات المفيدة للغاية لنظرية النهاية المركزية.

والسؤال الذي يفرض نفسه هنا، هو: كم يجب أن يبلغ حجم العينة n حتى يصبح التقريب الناشيء عن تطبيق نظرية نهاية المركزية تقريبا جيدا من وجهة النظر العملية؟

ولسوء الحظ لا يوجد جواب عام ومحدد تماما لهذا السؤال. ويتعلق الأمر بالتوزيع الاحتمالي الموافق للمجتمع الذي جاءت منه العينة، وبالغاية من استخدام التقريب،

وهكذا. وغالبا ما يكون لكل حالة حكمها، معتمدين، بصورة رئيسة، على الخبرة السابقة والتجربة. ونشعر بكثير من الراحة عند النظر إلى مثال قذف حجر النرد المذكور أعلاه، فقد لاحظنا أن المدرج التكراري للقيم الـ 200 لـ $\bar{X}$ قريب من شكل الجرس بالرغم من أن حجم العينة الذي استخدمناه لم يتعد الخمس، وبالرغم من أن التوزيع الذي تأتي منه العينات هو خط أفقي (انظر الشكل (٣ ـ ٣)) وبعيد جدا عن شكل الجرس. وبصورة عامة، يمكن القول إنه كلما كانت درجة التناظر في التوزيع الذي نعاينه عالية كان التقريب جيدا حتى في عينات صغيرة الحجم.

تمارين (٥ ـ ٤)

١) بالإشارة إلى التمرين ١١ من مجموعة التمارين (٣ ـ ١)، لنفرض أن الشخص يقوم بـ 250 رحلة في السنة إلى عمله. وليكن $\bar{Y}$ متوسط عدد الإشارات الحمر التي يواجهها في الرحلة الواحدة، احسب $E(\bar{Y})$ ، $V(\bar{Y})$، ثم احسب $P(\bar{Y} \geq 1.5)$

٢) في مدينة معينة 1/3 الأسر ليس لديها سيارة، و 1/3 الأسر لديها سيارة واحدة، و 1/6 الأسر لديها سيارتان، و 1/12 من الأسر لديها ثلاث سيارات، و 1/12 من الأسر لديها أربع سيارات، ليكن X عدد السيارات للأسرة الواحدة:

أ ـ احسب $E(X)$ ، $V(X)$.

ب ـ احسب $E(\bar{X})$ ، $V(\bar{X})$ حيث $\bar{X}$ متوسط عينة عشوائية من 100 أسرة.

جـ ـ إذا كان لكل سيارة خمس عجلات فما المتوسط والانحراف المعياري لعدد العجلات للأسرة الواحدة.

د ـ احسب بصورة تقريبية $P(\bar{X} < 1)$.

٣) تذبح مضافة عربية كل يوم 1 ، 2 ، 3 ، أو 4 خراف باحتمالات هي، على الترتيب، 0.4 ، 0.3 ، 0.2 ، 0.1. ما هو الحد الأدنى لعدد الخراف التي ستلبي باحتمال لا يقل

عن 0.99 حاجة المضافـة من الذبائح لفترة 120 يوما؟ (نفترض أن حاجـة المضافة في يوم مستقلة عن حاجتها في يوم آخر).

٤) متوسط الوزن في قطيع ضخم من الخراف هو 8.2 كغ بتباين يساوي 4.84 كغ٢ . ما احتمال أن يقع متوسط الوزن في عينة عشوائية من 80 خروفا بين 8.3 و 8.4 كغ؟

(٥ ـ ٦) تقريب التوزيع الثنائي بالتوزيع الطبيعي

رأينا في الفصل السـابق عدة تطبيقات للتـوزيع الثنائي إقتضت جميعهـا حساب احتمال أن يأخذ X ، وهو عدد النجـاحات من بين n تكرارا، قيمـة معينة أو يقع ضمن فترة معينة، وقـد اقتصرنا هنـاك على أمثلة تكـون n صغيرة فيها، وذلك بسبـب مشقة الحسابات عندما تكون n كبيرة. ولنفرض، مثلا، أننا في حاجة لحساب احتمال وقوع X ضمن فترة معينة ، حيث $n = 1000$ ، فمع أن مثل هذا العمل ليس مستحيـلا، إلا أنه ممتنع إلى الحد الذي نريد معه تجنب الغوص في الحسابات. وتقدم نظرية النهاية المركزية حلا لهذه المشكلة. ذلك لأنه يمكن النظر إلى عدد النجاحات X كمجموع يحقق شروط نظريـة النهاية المركـزية. فإذا اصطلحنا على أن يـوافق النتيجة S (أو النجاح) العـدد 1 ويوافق النتيجة F (أو الفشل) العدد صفر. فعندئذ تكـون نتائج التكرارات المستقلة لـ n عبارة عن متتالية من المتغيرات المستقلة $X_1, X_2, \ldots, X_n$ ، حيث يأخذ كل X_i إما القيمة 1 أو القيمة صفر. ويكون عدد النجاحات X هو بالضبط عدد مرات ورود الـ 1 في تلك المتتالية أو مجموعها. أي أن

$$X = \sum_{i=1}^{n} X_i$$

وبما أن كل X_i يتوزع وفق التوزيع الثنائي النقطي أو توزيع بيرنوللي، [انظر مطلع الفقرة (٤ ـ ٢) ونهاية الفقرة (٤ ـ ٧)] فتصبح نتائج التكرارات المستقلة الـ n وهي

X_1، $X_2, ..., X_n$ عينة عشوائية من مجتمع بيرنوللي، ويصبح X مجموع هذه العينة. ووفقا لنظرية النهاية المركزية يكون التوزيع التقريبي لـX ، في حالة n كبيرة بكفاية، هو التوزيع الطبيعي بمتوسط يساوي $n\,p$ وتباين يساوي $n\,p\,q$. وبالتالي يمكن استخدام جدول التوزيع الطبيعي المعياري لحساب احتمالات تتعلق بالمتغير X ، ولكن بصورة تقريبية.

مثال (٥ ـ ١٣)

لنأخذ التوزيع الثنائي في حالة $n = 10$ ، $p = 1/2$. وعندئذ يكون $\mu = np = 5$ و$\sigma = \sqrt{npq} = 1.58$ احسب $p(2 \le X \le 4)$ باستخدام التوزيع الثنائي أولا ثم باستخدام التوزيع الطبيعي لحساب قيمة تقريبية.

الحل

$$P(2 \le X \le 4) = \binom{10}{2}\left(\frac{1}{2}\right)^{10} + \binom{10}{3}\left(\frac{1}{2}\right)^{10} + \binom{10}{4}\left(\frac{1}{2}\right)^{10} = 0.3662$$

وهذا الاحتمال هو مجموع مساحات المستطيلات المقامة فوق 2 و 3 و 4 في المدرج الاحتمالي (انظر الشكل (٥ ـ ٩)) وإذا اعتبرنا X كأنه، على وجه التقريب، متغير $N(5, 2.5)$ ، فإن نظرة سريعة إلى الشكل (٥ ـ ٩) ستوضح أن المساحة تحت منحنى الكثافة الطبيعي من 2 إلى 4 تهمل النصف الأيسر من مساحة المستطيل المقام فوق 2 ، والنصف الأيمن من مساحة المستطيل المقام فوق 4 ، وأن التقريب سيكون أفضل لو أخذنا بدلا من $P(2 \le X \le 4)$ ، العبارة $P(2 - 1/2 \le X \le 4 + 1/2)$. ولكن:

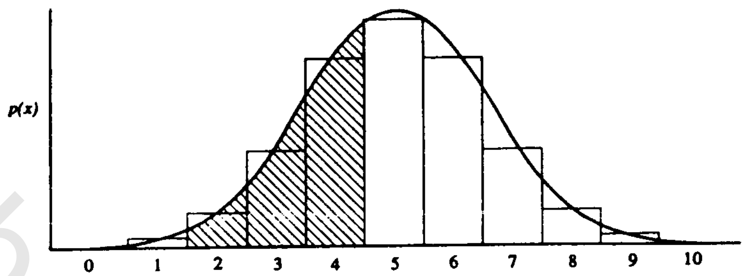
$$P(1.5 \le X \le 4.5) = F\left(\frac{4.5-5}{1.58}\right) - F\left(\frac{1.5-5}{1.58}\right)$$

$$= F(-0.316) - F(-2.215)$$

$$= 1 - F(0.316) - [1 - F(2.215)]$$

$$= F(2.215) - F(0.316) = 0.9866 - 0.6240$$

$$= 0.3626$$



شكل (٥ ـ ٩) تقريب التوزيع الثنائي بالتوزيع الطبيعي

والقيمة الناتجة صحيحة إلى رقمين عشريين بالرغم من أن n لا تتجاوز العشرة. ويعود الفضل في جودة التقريب هنا إلى تناظر التوزيع الثنائي في حالة $p = 0.5$ ، وإلى تعديل فترة تغير X ، مأخوذا كمتغير طبيعي مستمر، بحيث تغطي تماما المستطيلات الموافقة للحادثة التي نحسب احتمالها. وتسمى إضافة أو طرح 1/2 ، عملية تصحيح من أجل الاستمرار.

وعندما يكون n صغيرا و p قريبا من الصفر، أو قريبا من الواحد، فإن شكل المدرج الاحتمالي سيكون ملتويا بشدة (أي تتجمع معظم المساحة إلى جانب $X = 0$ أو إلى جانب $X = n$ ، على الترتيب) وبالتالي سيكون بعيدا جدا عن وضع التناظر. وفي مثل هذه الحالات سيكون التقريب سيئا ما لم تكن n كبيرة بكفاية.

مثال (٥ ـ ١٤)

موثوقية قطعة إلكترونية هي احتمال أن نختار واحدة من كومة إنتاج فنجدها تؤدي المهمة التي صممت من أجلها. أحسب احتمال أن نجد ما لا يقل عن 27 قطعة لا تعمل من بين عينة عشوائية تتضمن 1000 قطعة وذلك تحت الفرض بأن الموثوقية هي 0.98.

الحـــل

المسألـة هي مسألة تـوزيع ثنائي فكل قطعـة تختارها إمـا أن تعمل أو لا تعمل. وإذا اعتبرنـا نتيجـة «القطعـة لا تعمل» نجـاحـا، يكـون $p = 0.02$ ويكون المطلـوب حساب:

$$P(X \geq 27) = \sum_{x=27}^{1000} \binom{1000}{27} (0.02)^{x} (0.98)^{1000-x}$$

والحساب الدقيق لهذه النتيجة يتطلب جهدا كبيرا. وبـاستخدام تقريب التوزيع الطبيعي نحسب المسـاحـة تحت المنحنى الطبيعي إلى اليمين من $X = 26.5$ (لاحظ أنه ينبغي استخدام $X = 26.5$ بدلا من $X = 27$ بحيث تشمل المستطيل الاحتمالي المقـام فوق النقطة $X = 27$). وذلك بـاعتبـار أن X يتبع على وجـه التقـريب، التـوزيع الطبيعي بمتوسط يساوي

$$\mu = np = 1000 \times 0.02 = 20$$

وانحراف معياري

$$\sigma = \sqrt{n\ p\ q} = \sqrt{1000 \times 0.02 \times 0.98} = 4.43$$

وهكذا نجد قيمة تقريبية للاحتمال المطلوب:

$$\begin{aligned} P(X \geq 26.5) &= P\left(\frac{X-\mu}{\sigma} \geq \frac{26.5-\mu}{\sigma}\right) \\ &= P\left(Z \geq \frac{26.5-20}{4.43}\right) = P(Z > 1.4) \\ &= 1 - F(1.4) = 1 - 0.9292 = 0.0708 \end{aligned}$$

مثال (٥ ـ ١٥)

اختبرنا لقاحـا جديدا ضد الزكام. وقـد أعطي اللقاح لمائة شخص، وروقبوا من حيث إصابتهم بالزكـام لمدة سنة. وقد نجا 68 منهم من الإصابة بالـزكام. ولنفرض أننا

نعلم من معلومات سابقة أن احتمال عدم الإصابة بالزكام هي بصورة طبيعية وبدون استخدام اللقاح 0.5 . أية نتائج يمكنك استخلاصها من هذه التجربة حول فعالية اللقاح؟

الحـــل

لنحسب احتمال نجاة 68 أو أكثر من الإصابة بالزكام تحت الفرض بأن $p = 0.5$ ، أي أن اللقاح لم يكن له أي تأثير، فنجد باستخدام التقريب الطبيعي :

$$\mu = n\,p = 100\,(0.5) = 50\ ;\ \sigma = \sqrt{50 \times 0.5} = 5\ ,$$
$$P(X \geq 68) = P\left(Z \geq \frac{67.5 - 50}{5}\right) = 1 - F\,(3.5) = 0.0002$$

لقد قمنا بالحسابات مفترضين أن اللقاح غير فعال، وأن العدد 68 الذي حصلنا عليه، وهو أكبر من المتوقع تحت هذا الفرض، كان محض مصادفة. ولكن الاحتمال الناتج صغير جدا، وهو يعني، عمليا، أنه لو كان ما افترضناه صحيحا وكررنا التجربة نفسها عددا كبيرا جدا من المرات فإننا سنجد نتيجة كالنتيجة التي حصلنا عليها، أو أفضل، في تجربتين من كل عشرة آلاف تجربة، وهذا يثير الكثير من الريبة في صحة ما افترضناه، ويدعو إلى الاعتراف بفعالية اللقاح في الوقاية من الزكام.

مثال (٥ ـ ١٦)

يتضمن امتحان خمسين سؤالا من النوع متعدد الاختيارات، ولكل سؤال ثلاثة أجوبة مقترحة، واحد منها فقط هو الجواب الصحيح؛ ولكي ينجح الطالب لا بد له من الإجابة بصورة صحيحة على عشرين سؤالا على الأقل.

ا ـ احسب احتمال نجاح طالب غير مؤهل يختار جوابه عن كل سؤال عشوائيا.

ب ـ مع بقاء عدد الأسئلة ودرجة النجاح كما هي، كم يجب أن يكون عدد

الاختيارات المطروحة أمـام كل سؤال ليصبح احتمال نجاح طالب يختار جـوابه عشوائيا أقل من 0.01؟

جـ ـ في حال وجود اختيـارين فقط، كم يجب أن تكون درجة النجاح بحيث لا يزيد احتمال نجاح طالب يختار جوابه عشوائيا على الواحد في المائة؟

الحل

١ ـ ليكـن X عـدد الأجوبـة الصحيحـة، فلدينـا $p = 1/3$ ، $n = 50$ ، والمطلوب $P(X \geq 20)$. وباستخدام التقريب الطبيعي نجد:

$$\mu = np = 50 \times \frac{1}{3} = \frac{50}{3} \;,\; \sigma = \sqrt{npq} = \sqrt{\frac{100}{9}} = \frac{10}{3}$$

$$P(X \geq 20) \approx P\left(Z \geq \frac{19.5 - \frac{50}{3}}{\frac{10}{3}}\right) = 1 - F(0.85) = 1 - 0.8023$$
$$= 0.198$$

ب ـ

$$P(X \geq 20) = P\left(Z \geq \frac{19.5 - 50p}{\sqrt{50p(1-p)}}\right) \leq 0.01$$

$$F\left(\frac{19.5 - 50p}{\sqrt{50P(1-p)}}\right) \geq 0.99$$

$$\frac{19.5 - 50p}{\sqrt{50p(1-p)}} \geq 2.33$$

والمطلوب قيمة p التي تحقق هـذه المتباينة وتجعـل المقدار $19.5 - 50p$ موجبا كما ينبغي أن يكون. وبتربيع الطرفين والإصلاح نجد:

$$2771.45\, p2 - 2221.45\, p + 380.25 \geq 0$$

وهذه تتحقـق إذا كان $p > 0.55$ أو $p < 0.248$. ولكن قيـم p الأكبر مـن 0.55 مرفـوضة لأنها تجعل $19.5 - 50p$ سـالبا. وبما أن عـدد الاختيارات هـو بالضرورة عـدد

صحيح فعلينا أخذ أول نسبة تقل عن 0.248 ويكون جداؤها بعدد صحيح مساويا للواحد تماما. والنسبة المطلوبة هي إذا 0.2 ، وهذا يعني أن عدد الاختيارات المطروحة أمام كل سؤال ينبغي أن تكون خمسة.

جـ ـ المطلوب تحديد عدد صحيح a يحقق المتباينة التالية:

$$P(X \geq a) \leq 0.01$$

حيث $n = 50$ ، $p = 1/2$ ، وبالتالي $\mu = 25$ ، $\sigma = \sqrt{12.5} = 3.54$.

$$P(X \geq a) \approx P\left(Z \geq \frac{a - \frac{1}{2} - 25}{3.54}\right) \leq 0.01$$

أي

$$F\left(\frac{a - 25.5}{3.54}\right) \geq 0.99$$

$$\frac{a - 25.5}{3.54} \geq 2.33 \Leftrightarrow a = 34 .$$

تمارين (٥ ـ ٥)

١) عند تصالب حبتي بازيلاء لكل منها زوج من المورثات (أحمر، أبيض) يُتوقَّع أن تكون زهور ربع النسل بيضاء. إذا فحصنا 64 نبتة ناتجة عن مثل هذا التصالب فما احتمال أن نجد 16 منها بالضبط ذات زهور بيض؟

٢) نسبة القطع غير الصالحة التي تنتجها آلة هي 20% . أحسب بصورة تقريبية احتمال أن تتضمن عينة عشوائية من 400 قطعة من إنتاج هذه الآلة أكثر من 96 قطعة غير صالحة؟

٣) نقذف حجر نرد 300 مرة، ونعتبر الحصول على 1 أو 2 «نجاحا». استخدم تقريب التوزيع الثنائي بالتوزيع الطبيعي لحساب احتمال ألا يحيد عدد النجاحات عن 100 بأكثر من 15 .

٤) في بلدة كبيرة يعطي نصف الناخبين، عادة، أصواتهم للمرشح A . ويسأل كل من 20 باحثا إحصائيا عينة عشوائية من 16 من الناخبين عن المرشح المفضل. استخدم جدول التوزيع الطبيعي لحساب تقريبي لعدد الباحثين الذين تتوقع أن يفيدوا بأن أقل من 6 من عينتهم فضلوا المرشح A .

٥) يزرع رجل في حديقة منزله بذور زهور يقال أن 60% منها ينبت. إذا زرع 60 بذرة فما احتمال أن ينبت منها 15 بذرة أو أقل؟

٦) لتعيين مشرف على آلة حاسبة الكترونية تتطلب إحدى الشركات من المرشحين اجتياز اختبار كتابي. وتتألف ورقة الامتحان من 100 سؤال متعدد الاختيارات، ولكل سؤال ثلاثة أجوبة مقترحة أحدها فقط صحيح. والنجاح في الاختبار يقتضي الإجابة الصحيحة على 40 سؤالا، على الأقل. والمطلوب

ا ـ احتمال نجاح متقدم يختار الجواب على كل سؤال عشوائيا؟

ب ـ أكبر عدد من الأسئلة ينبغي أن تتضمنها ورقة الامتحان إذا أردنا لاحتمال نجاح متقدم يختار أجوبته عشوائيا أن لا يتجاوز الـ 1% ؟

٧) 25% من تلاميذ مدرسة لم يكن في سجلهم خلال عام دراسي بأكمله أي يوم غياب بسبب المرض. وفي الصف السادس من هذه المدرسة يوجد 120 تلميذا. أوجد عددا r بحيث يكون احتمال أقل من r تلميذ صف سادس بدون أي يوم غياب مرضي يساوي 0.01 أعرض الفرضيات التي اعتمدت عليها؟

٨) إذا كان 55% من الناخبين في مدينة كبيرة يؤيدون قضية فما احتمال أن تظهر عينة عشوائية من 100 ناخب من هذه المدينة أغلبية لصالح القضية؟

٩) احتمال أن نستكمل بنجاح سلسلة من العمليات في تجربة معينة هو 0.44 . إذا بدأنا 65 من مثل هـذه التجارب بصورة تضمن استقلال كل تجربة عن غيرها من التجارب، فما احتمال أن نستكمل بنجاح أقل من 25 منها؟ بين أنه إذا كان احتمال النجاح 0.04 فقط فإن احتمال أربع نجاحات على الأقل هو حوالي 1/4 ؟

١٠) بالإشارة إلى التمرين ١٥ من مجموعة التمارين (٣ ـ ١) هل يمكنك الآن إعطاء جواب تقريبي؟

١١) بالإشارة إلى التمرين ١٦ من مجموعة التمارين (٣ ـ ١)، هل يمكن إعطاء جواب تقريبي؟

(٥ ـ ٧) فترة ثقة لمتوسط مجتمع طبيعي تباينه معروف

ذكرنا عبر هذا الكتاب أن الإحصاء يهدف إلى التنبؤ أو اتخاذ قرار حول خاصة من خصائص مجتمع اعتمادا على المعلومات المتيسرة من عينة نأخذها من هذا المجتمع. وكما يوحي عنوان الفقرة فإن المجتمع الذي ينبغي دراسته هو مجتمع يخضع للتوزيع الطبيعي، أو مجتمع موصوف رياضيا بنموذج هو النموذج الطبيعي. وأن الخاصة التي تهمنا من خصائص هذا المجتمع هي متوسطه μ ، مثلا، مع افتراض لأن تباينه معروف ويساوي σ^2 . وما نريده هنا هو تحديد فترة، أي تحديد عددين حقيقيين، نستطيع أن نقول، بثقة عالية، إن المتوسط يقع بينهما.

لنأخذ عينة عشوائية حجمها n من هذا المجتمع، ولنرمز لمقاديرها بـ $X_1, X_2, \ldots, X_n$ ولمتوسطها بـ $\bar{X}$. فكما رأينا في الفقرة [٥ ـ ٤ (الخاصة ٣)]، يتوزع $\bar{X}$ وفق التوزيع الطبيعي $N\left(\mu, \frac{\sigma^2}{n}\right)$. ومعرفة التوزيع الاحتمالي لـ $\bar{X}$ يعني بالنسبة لنا

الشيء الكثير، إذ نستطيع تقديم وصف رياضي لمجتمع القياسات الموافق لـ $\bar{X}$ ، أي للقيم كافة التي يمكن أن يأخذها المتوسط $\bar{X}$ لو أننا قمنا بأخذ عدد هائل من العينات المختلفة ذات الحجم n من هذا المجتمع. وسيسمح لنا هذا التوزيع بالإجابة بيسر وسهولة على أسئلة هامة من النوع: ما نسبة العينات التي يتجاوز متوسطها قيمة محددة؟ أو يقل عن قيمة محددة؟ أو يقع بين عددين محددين؟ الخ. وبصورة عامة، يمكننا اعتمادا على معلوماتنا من الفقرة (٦ ـ ٣) أن نكتب:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

و

$$P\left(-Z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1 - \alpha$$

وهذا يكافىء:

$$P\left(-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

أو

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\frac{\sigma}{n}\right) = 1 - \alpha$$

وفي هذه العبارة يمكن معرفة $Z_{\alpha/2}$ من جدول التوزيع الطبيعي المعياري حالما نحدد قيمة α ، و σ معروف، و n حجم العينة محدد سلفا. وإذا أمعنا النظر، سنجد منطوق هذه العبارة قبل أخذ العينة كالتالي:

إن نسبة $100\ (1 - \alpha)\%$ من العينات ذات الحجم n التي يمكن أخذها من هذا المجتمع ستؤدي إلى قيمة لـ $\bar{X}$ بحيث تتضمن الفترة التي تبدأ بالعدد $\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ وتنتهي بالعدد $\bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ القيمة الصحيحة لمتوسط المجتمع μ . ويسمى $\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ حد الثقة الأدنى و $\bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ حد الثقة الأعلى.

وتيسيرا للفهم، ولتشكيل تصور محسوس للفكرة التي نطرحها هنا، دعنا نحدد قيمة لـ α، ولتكن 0.05 ، وعندئذ $Z_{\alpha/2} = Z_{0.025} = 1.96$. ويكون $1 - \alpha = 0.95$. وتصبح المقولة التي تشكل منطوق العبارة الاحتمالية أعلاه كالتالي:

إن نسبة 95% من العينات ذات الحجم n التي يمكن أخذها من هذا المجتمع ستؤدي إلى قيمة لـ $\bar{X}$ بحيث تتضمن الفترة

$$\left(\bar{X} - 1.96\,\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + 1.96\,\frac{\sigma}{\sqrt{n}}\right)$$

القيمة الصحيحة لمتوسط المجتمع μ .

وعندما نأخذ العينة سنحصل على قيمة محددة $\bar{x}$ للمتغير العشوائي $\bar{X}$ ، وسنجد فترة معرفة تماما هي الفترة الممتدة بين العدد $\bar{x} - 1.96\,\frac{\sigma}{\sqrt{n}}$ والعدد $\bar{x} + 1.96\,\frac{\sigma}{\sqrt{n}}$ ، وهذه الفترة إما أن تتضمن القيمة الصحيحة μ أو لا تتضمنها وليس هناك خيار ثالث. لم يعد هناك احتمالات للموقف، فقد أطلقنا طلقة على الهدف (أخذنا عينة وحددنا فترة ثقة) والنتيجة هي حتما واحدة من اثنتين فإما أننا أصبنا الهدف (الفترة تغطي μ) أو أننا لم نصبه (الفترة لا تغطي μ). وبما أننا نعلم قبل أخذ العينة أن نسبة عالية من العينات، (95% منها) تصيب الهدف، فستتولد عندنا ثقة عالية بأن العينة التي حصلنا عليها قد أصابت فعلا، مما يقترح تسمية النسبة العالية تلك «معامل ثقة». فنقول إن الفترة الناتجة هي فترة ثقة تتضمن μ بمعامل ثقة يبلغ 95%. والله سبحانه وتعالى وحده يعلم ما إذا كانت العينة التي حصلنا عليها حسنة الطالع (من بين الـ 95% التي تغطي القيمة الصحيحة للمتوسط μ) أم أنها سيئة الطالع (من بين الـ 5% التي يجانبها الصواب، إذ لا تغطي الفترة الناشئة عنها القيمة الصحيحة للمتوسط μ).

وبالطبع يمكن أن نكون أشد تحفظا فنأخذ $\alpha = 0.01$ ويكون معامل الثقة $100\,(1 - \alpha) = 99\%$. ومن الطبيعي أن تكون الفترة التي نحصل عليها في هذه الحالة

أطول من سابقتها المقابلة لمعامل ثقة 95% . كما يمكن، على الوجه الآخر، أن نكون أقل تحفظا فنأخذ $\alpha = 0.10$ ، ويكون معامل الثقة 90% لفترة تمتاز بأنها أقصر من سابقتيها .

مثال (٥ ـ ١٧)

يمثل البيان الإحصائي التالي إنتاج عشر شجيرات من الطماطم مقاسا بالكيلوغرام .

2.3, 2.6, 2.2, 3.1, 4.0, 1.9, 2.7, 1.9, 3.3, 3.0

ونعلم أن قياسات الإنتاج في مجتمع شجيرات الطماطم يوصف بتوزيع طبيعي تباينه $\sigma^2 = 0.36$. أحسب 90% ، 95% ، و 99% فترة ثقة لمتوسط الإنتاج μ .

الحل

متوسط العينة $\bar{x}$ هو:

$$\bar{x} = \frac{2.3 + 2.6 + \dots + 3.0}{10} = 2.7$$

90% فترة ثقة :

$$Z_{\alpha/2} = Z_{0.05} = 1.645 \text{ ،} \alpha/2 = 0.05 \text{ ،} \alpha = 0.10 \text{ ،} 1 - \alpha = 0.9$$

وتكون الفترة المطلوبة :

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 1.645 \frac{\sqrt{0.36}}{\sqrt{10}} = 2.7 \pm 0.31 = (2.39, 3.01)$$

95% فترة ثقة :

$$Z_{\alpha/2} = Z_{0.025} = 1.96 \text{ ،} \alpha/2 = 0.025 \text{ ،} \alpha = 0.05 \text{ ،} 1 - \alpha = 0.95$$

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 1.96 \frac{0.6}{3.16} = 2.7 \pm 0.37 = (2.33, 3.07)$$

99% فترة ثقة :

$Z_{\alpha/2} = Z_{0.005} = 2.58$ ، $\alpha/2 = 0.05$ ، $\alpha = 0.10$ ، $1 - \alpha = 0.99$

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 2.58 \frac{0.6}{3.16} = 2.7 \pm 0.49 = (2.21 , 3.19)$$

لاحظ أن فترة الثقة تتسع مع ازدياد معامل الثقة .

في هذا المثـال نخرج ، مثـلا ، بالتقديـر التالي : «بمعـامل ثقة 95% يقع متـوسط الإنتاج μ بين 2.33 كغ و 3.07 كغ . ولكن هب أننا اتفقنا على اعتبار منتصف الفترة قيمة تقديرية أو تقديرا لـ μ ، فهذا شيء منطقي تماما إذ نقول إن متوسط العينة $\bar{X} = 2.7$ هو تقديرنـا لمتوسط المجتمع الذي جاءت منه العينـة ، ولكن الاكتفاء بذلك لا يضيف أي جديد إلى ما هـو معروف تاريخيا ، إذ يلجأ كل خبير يريـد القيام بعملية تخمين إلى أخذ عينة تمثل المجتمع ، في رأيـه ، تمثيلا جيدا ، ثم يأخـذ معلومات العينة ليعممهـا بصورة مباشرة على المجتمع . وكأن العينة هي صـورة مصغرة للمجتمع ، وليس علينا إلا تكبير هذه الصورة حتى نحصل على صورة المجتمع . ولكن ماذا عن الخطأ في هذا التقدير؟ لو رجعنا إلى العبارة الاحتمالية في مطلع الفقرة وهي :

$$P\left(- Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

لوجدنا أنها مكافئة للعبارة :

$$P\left(\left| \bar{X} - \mu \right| < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

و $|\bar{X} - \mu|$ يمثل الخطأ المطلق للتقدير، فهو القيمـة المطلقة لحيدان التقدير عن الشيء المراد تقديره . والعبارة الإحتمالية تقول إنه باحتمال يبلغ $(1 - \alpha)$ لا يتجاوز الخطأ في هذا التقديـر المقدار $\pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ ، وفي المثال السابق يمكن القـول ، مثلا ، إنه في 95% من العينات الممكنـة سوف لا يحيد متـوسط العينة عن متـوسط المجتمع بأكثر من 0.37 كغ زيادة أو نقصـانا.وبعبـارة أخرى ، سوف لا يتعـدى الخطأ في تقديـرنا إلا فيما

ندر، القيمـــة 0.37 كغ زيـادة أو نقصـانـا. وسنطلق على المقـدار $Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ ، مع شيء من التجـاوز، اسم «الحد الأعلى لخطأ التقـديـر»، فهـو في حقيقة أمـره حـد أعلى تقريبي لخطأ التقدير. وسنرمز له بالرمز e ، ونكتب:

$$e = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

ولو سألنا الخبير التقليدي عن الخطأ في تقديره لعجز عن الإجابة إذ ليس لديه أية وسيلـة تسمح لـه بـذلك. وبينما يعتمد الخبير التقليـدي إعتمادا كليـا على العينـة التي أخذها فإن الإحصـائي اليوم لا يعتمد على العينة إلا كجزء من صـورة متكاملة تتضمن إلى جانب العينة المأخوذة العينات كافة التي كان يمكن الحصول عليها لو أنه كرر تجربة أخذ العينـة عددا هائلا من المرات. وهـو ما يسمى بتوزيع المعـاينة، ممثلا هنـا بتوزيع المتوسط $\bar{x}$. وهذه هي الإضافة الجديدة لعلـم الإحصاء في مسألة كهذه. (انظر الفقرة (٥ ـ ٧) والفقرة (٥ ـ ٩)).

ويتضح من عبارة e أنـه يمكننـا التحكم في حجم الخطأ من خـلال التحكم في حجم العينة n ، فالمقدار e يتناسب عكسا مع الجذر التربيعي لحجم العينـة. ولو أردنا تخفيض e إلى نصف مـا هو عليـه لاحتجنا إلى زيادة حجـم العينة إلى أربعـة أضعاف. وبـالطبع يمكننا قبل تنفيـذ البحث الإحصـائي، أي قبل أخذ العينـة، تصميم حجم العينة n بصورة تتناسب مـع مقدار الخطأ الذي يمكن التساهل فيـه. وسنوضح الفكرة بمثال.

مثال (٥ ـ ١٨)

بالإشـارة إلى المثال السـابق (٥ ـ ١٧)، لنفرض أننـا نريد تقـدير متـوسط إنتاج شجيرة الطماطم μ بحيث لا يزيد الخطأ عن 0.2 كغ إلا باحتمال زهيد لا يتجاوز الواحد في المائة. فكم يجب أن يكون حجم العينة؟

الحل

الحد الأعلى للخطأ e يساوي 0.2 ،

$$Z_{\alpha/2} = Z_{0.005} = 2.58 \qquad \alpha = 0.01$$

$$e = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2.58\,\frac{0.6}{\sqrt{n}} \leq 0.2$$

ومنه :

$$0.2\sqrt{n} \geq 1.548$$

$$\sqrt{n} \geq 7.74 \quad , \quad n \geq 59.9$$

أي أن حجم العينة يجب أن لا يقل عن 60 .

تمارين (٥ـ٦)

١) بفرض عينات عشوائية من مجتمعات طبيعية تباينها معروف، أوجد فترات ثقة للقيمة الحقيقية μ لمتوسط المجتمع بمعامل الثقة المبين في كل حالة :

ا ـ $n = 9$ ، $\bar{X} = 4$ ، $\sigma^2 = 16$ ، معامل الثقة 90% .

ب ـ $n = 100$ ، $\bar{X} = 29$ ، $\sigma^2 = 49$ ، معامل الثقة 95% .

جـ ـ $n = 64$ ، $\bar{X} = 4$ ، $\sigma^2 = 100$ ، معامل الثقة 99% .

٢) مجتمع طبيعي انحرافه المعياري $\sigma = 0.75$. كم يجب أن يكون حجم عينة مأخوذة من هذا المجتمع كي لا يزيد الحد الأعلى للخطأ عن 0.4 ، وذلك باحتمال 0.95 ؟

٣) نعلم أن الخطأ المرتكب في قياس طول، عند استخدام جهاز لقياس الأطوال، يتوزع وفق التوزيع الطبيعي، بمتوسط يساوي الصفر، وانحراف معياري 1 مم .

ا ـ أحسب احتمال أن يقل الخطأ عند استخدام الجهاز لمرة واحدة عن 0.5مم .

ب ـ إذا استخدم الجهاز بصورة مستقلة 9 مرات لقياس طول معين، فاحسب احتمال أن يقع متوسط القياسات التسعة في حدود 0.5 مم من القيمة الحقيقية للطول .

٤) مجتمع طبيعي انحرافه المعياري $\sigma = 0.75$ ، كم يجب أن يكون حجم عينة مأخوذة من هذا المجتمع كي لا يزيد الحد الأعلى للخطأ عن 0.4 ، وذلك باحتمال 0.95 ؟

٥) يريدإحصائي تحديد متوسط الأجر اليومي لمستخدمي مهنة معينة. ويريد باحتمال 0.95 حدا أقصى للخطأ قدره 9 ريالات. ومن دراسات مماثلة أخرى يعلم أن بإمكانه افتراض مجتمع طبيعي انحرافه المعياري $\sqrt{650}$ ريالا. ما هو حجم العينة التي ينبغي أن يخطط للحصول عليها؟

٦) إذا كان الانحراف المعياري لمجتمع طبيعي هو $\sigma = 5$. ما هو حجم العينة التي ينبغي أخذها حتى نطمئن باحتمال قدره 0.95 إلى عدم اختلاف متوسط العينة عن متوسط المجتمع بأكثر من الواحد؟

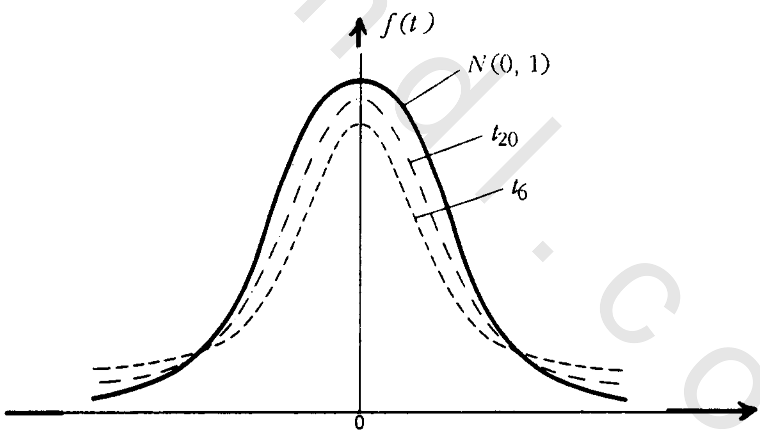
(٥ ـ ٨) فترة ثقة لمتوسط مجتمع طبيعي تباينه غير معروف وحجم العينة صغير

لو تتبعنا المناقشة في الفقرة السابقة لوجدنا أنه لا بد من تعويض σ ، التي افترضناها معروفة هناك، بتقدير لها من العينة. والتقدير الذي تمليه البداهة هو اعتماد s، الانحراف المعياري للعينة، كتقدير لـ σ ، الانحراف المعياري للمجتمع الذي جاءت منه العينة. وهكذا يأخذ المقدار $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ ، الذي يتبع تماما التوزيع الطبيعي المعياري، الصيغة:

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

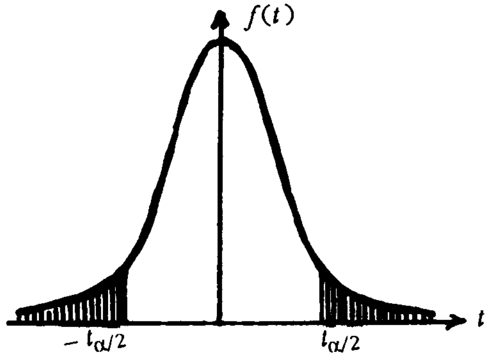
والمتغير الجديد الذي رمزنا له بـ t يحوي مركبة عشوائية في البسط هي $\bar{X}$ ومركبه عشوائية في المقام هي S. ولم يعد توزيعه هو التوزيع الطبيعي المعياري. وقد تمكن «ستيودنت» ، وهو لقب لكاتب إحصائي كان ينشر أبحاثه بتوقيع «ستيودنت»، أن

يشتق العبارة المضبوطة لتوزيع t ويسمى هذا التوزيع في كتب الإحصاء المختلفة «التوزيع t»، أو توزيع «ستيودنت». وفي الشكل (٥ ــ ١٠) نجد أمثلة من منحنيات الكثافة لهذا التوزيع. فهو متناظر حول المحور الرأسي، شأنه في ذلك شأن منحنى الكثافة الطبيعي المعياري. ويعتمد المنحنى على حجم العينة n ونصطلح على تسمية المقدار $n-1$، «عدد درجات الحرية» ونرمز له بـ ν (حرف يوناني ينطق نو). والجدول ٢ في الملحق يعطي القيمة الموجبة لـ t التي يقع إلى اليسار منها $(1-\alpha)$ من المساحة الكلية تحت المنحنى وذلك من أجل قيم مختلفة لـ α و ν. وسنرمز بـ $t_\alpha(n-1)$ للدلالة على قيمة المتغير t في صلب الجدول الواقعة في ملتقى السطر $n-1$ والعمود الذي عنوانه $1-\alpha$. وعلى سبيل المثال، لإيجاد $t_{0.025}(14)$، ندخل الجدول وفق السطر 14 ونتحرك حتى نصل إلى العمود الذي عنوانه 0.975 لنجد القيمة 2.145.



شكل (٥ ـ ١٠) أشكال مقارنة للتوزيعين t_6، t_{20} والتوزيع $N(0, 1)$.

وإذا أخذنا عينة عشوائية حجمها n من مجتمع طبيعي $N(\mu, \sigma^2)$ حيث μ و σ^2 غير معروفين، وكالعادة رمزنا بـ $\bar{X}$ و S لمتوسط العينة وانحرافها المعياري فيمكننا إستنادا إلى تناظر التوزيع t (انظر الشكل (٥ ـ ١١)) كتابة:



شكل (٥ ـ ١١)

$$P\left(-t_{\alpha/2}(n-1) < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right) = 1-\alpha$$

أو

$$P\left(\frac{-S}{\sqrt{n}}\,t_{\alpha/2}(n-1) < \bar{X}-\mu < \frac{S}{\sqrt{n}}\,t_{\alpha/2}(n-1)\right) = 1-\alpha$$

أو

$$P\left(\bar{X}-\frac{S}{\sqrt{n}}\,t_{\alpha/2}(n-1) < \mu < \bar{X}+\frac{S}{\sqrt{n}}\,t_{\alpha/2}(n-1)\right) = 1-\alpha$$

وتكون فترة الثقة للمتوسط μ بمعامل ثقة $100(1-\alpha)\%$ هي $\left(\bar{X}-\frac{S}{\sqrt{n}}\,t_{\alpha/2}(n-1)\,,\ \bar{X}+\frac{S}{\sqrt{n}}\,t_{\alpha/2}(n-1)\right)$. ولوضع فترة ثقة نحسب إذا $\bar{X}$ و S من العينة وقيمة $t_{\alpha/2}(n-1)$ من جدول التوزيع t ثم نعوض.

مثال (٥ ـ ١٩)

قسنا ارتفاع خمس عشرة شجيرة باذنجان بعد فترة من زرعها فكان متوسط الإرتفاع 83 سم بانحراف معياري 5.8 سم. ضع 95% فترة ثقة لمتوسط الارتفاع في المجتمع الذي اخترنا منه الشجيرات الخمس عشرة في العينة، مفترضا أن ارتفاع الشجيرة في المجتمع يخضع للتوزيع الطبيعي.

الحل

حـــدا الثقـــة همـــا $\bar{x} \pm \frac{S}{\sqrt{n}}\, t_{0.025}\,(14)$. ومـن جـدول التوزيع t نجـد $t_{0.025}\,(14) = 2.145$ ، لاحظ أنه لـو كان σ معروفا لكان هذا العدد 1.96 فقط)، وتكون فترة الثقة بمعامل 95% هي :

$$83 \pm 2.145 \times 5.8 \,/\, \sqrt{15}$$

أي من 79.8 سم إلى 86.2 سم .

وكلما ازداد حجم العينة أصبح S تقديرا أفضل لـσ وبالتالي اقتربت قيم t من قيم المتغير الطبيعي المعياري Z الموافقة لها . ولو نظرنـا في السطور الأخيرة في جدول التوزيع t (السطور التـي تلي السطر 30) لوجـدنا أن الفـروق بين قيم t وقيم Z المقابلـة لها تصبح صغيرة، وفي السطر الأخير من الجدول حيث كتب حـذاءه الرمز "∞" تتطابق قيم t مع قيم Z المقابلة .

مثال (٥ ـ ٢٠)

وضعت عينـة من 12 فأرا تجريبيـا على نظام تغـذية معين خـلال الأشهر الثـلاثة الأولى من حياتها وقيست الزيادة في وزن كل فأر بالغرام فكانت كما يلي :

55, 62, 54, 58, 65, 64, 60, 62, 59, 67, 62, 61

والمطلوب وضع فترة ثقة بمعامل ثقة 90% لمتوسط الزيادة في الوزن خلال الأشهر الثلاثة الأولى من حياة مجتمع الفئـران الذي جاءت منـه العينة، علما أنه يمكن اعتبـار التوزيع الطبيعي توزيعا مناسبا لمتغير زيادة الوزن خلال الأشهر الثلاثة الأولى .

الحل

نحسب متوسط العينة وانحرافها المعياري فنجد $\bar{X} = 60.75$ ، $S = 3.84$ ، ولدينا $1 - \alpha = 0.90$ ، $\alpha = 0.10$ ، $\alpha/2 = 0.05$ ، ومن جدول التوزيع t نجد :

$$t_{\alpha/2}(n-1) = t_{0.05}(11) = 1.796$$

بالتعويض نجد فترة الثقة المطلوبة :

$$60.75 \pm \frac{3.84}{\sqrt{12}} \times 1.796 = 60.75 \pm 1.99$$

أي من 58.76غ إلى 62.74غ .

تمارين (٥ ـ ٧)

١) في ستة اختبارات لتجميع وتركيب قطع آلية معينة ، استغرق وقت التجميع والتركيب 13 ، 14 ، 12 ، 16 ، 13 ، و 11 دقيقة. مفترضا أن زمن التجميع والتركيب يتبع التوزيع الطبيعي ، ضع فترة ثقة لمتوسط الزمن الحقيقي للتجميع والتركيب بمعامل ثقة 99% .

٢) عينة عشوائية من 30 درجة من درجات اختبار للذكاء أعطي لطلاب المرحلة الثانوية ، أنتجت متوسطا قدره 423 وإنحرافا معياريا $S = 68$. أوجد فترة ثقة لمتوسط المجتمع بمعامل ثقة 95% ، مفترضا أن درجات الاختبار في المجتمع تتبع التوزيع الطبيعي .

٣) وجد طبيب أسنان في فحصه الدوري لستة طلاب ابتدائي أنهم احتاجوا إلى 2 ، 3 ، 6 ، 0 ، 4 ، و 3 عمليات حشوة .

ا ـ إذا استخدم الطبيب متوسط هذه العينة كتقدير لمتوسط المجتمع الذي جاءت منه العينة فماذا يمكنه أن يقول باحتمال 0.95 عن الحد الأعلى للخطأ الذي ارتكبه؟

ب ـ ضع فترة ثقة لمتوسط المجتمع بمعامل ثقة 95% .

جـ ـ ما الفرض الذي استندت إليه في حساباتك؟

٤) أحد الثوابت الفيزيائية المهمة هو e/m نسبة شحنة الكهروب (الإلكترون) إلى كتلته. وفي تجربة فيزيائية لقياس هذا الثابت أعيدت، بصورة مستقلة، 12 مرة، كانت النتائج التالية:

$$1.7604 \times 10^7 , 1.7638 \times 10^7 , 1.7609 \times 10^7$$

$$1.7563 \times 10^7 , 1.7556 \times 10^7 , 1.7582 \times 10^7$$

$$1.7526 \times 10^7 , 1.7663 \times 10^7 , 1.7624 \times 10^7$$

$$1.7620 \times 10^7 , 1.7605 \times 10^7 , 1.7621 \times 10^7$$

ا ـ ما تقديرك للقيمة الحقيقة لـ e/m ؟ وما هو الحد الأعلى لخطأ هذا التقدير باحتمال 0.95 ؟

ب ـ ضع فترة ثقة لقيمة e/m بمعامل ثقة 99% .

٥) تأتي مادة غذائية من مصنع معين في علب مكتوب عليها «الوزن الصافي 38 أونزة». وقد وجد أن الوزن الذي تحتويه كل من عينة عشوائية من 6 علب كان كما يلي:

34.06, 39.65, 34.75, 40.00, 39.50, 34.25

ضع فترة ثقة لمتوسط محتوى العلبة من إنتاج المصنع من تلك المادة الغذائية، وذلك بمعامل ثقة 98% .

(٥ ـ ٩) فترة الثقة لمتوسط مجتمع في حالة عينات كبيرة الحجم

لا نفترض هنا أن المجتمع الذي نأخذ منه العينة مجتمع طبيعي، ولكننا نفترض أن حجم العينة n كبير إلى الحد الذي يسمح بالاستفادة من نظرية النهاية المركزية، واعتبار توزيع $\bar{X}$ ، متوسط العينة، مطابقا تقريبا للتوزيع الطبيعي. وبالتالي تطبيق ما جاء في الفقرة (٥ ـ ٧) بحذافيره. فإذا كان تباين المجتمع σ^2 معروفا كانت فترة الثقة لمتوسط المجتمع μ، بمعامل ثقة $100\ (1 - \mu)\%$ هي

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

حيث تؤخذ قيمة $Z_{\alpha/2}$ من جدول التوزيع الطبيعي المعياري. وإذا كان التباين σ^2 غير معروف، وغالبا ما يكون الأمر كذلك في التطبيقات العملية، فإن تباين العينة S^2 يشكل تقديرا جيدا لـ σ^2، نظرا لكبر حجم العينة، مما يسمح بتعويض S بدلا من σ في فترة الثقة لتصبح:

$$\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

وكما رأينا في الفقرة (٥ ـ ٧) نعتبر متوسط العينة $\bar{X}$ تقديرا لمتوسط المجتمع μ ويكون الحد الأعلى التقريبي لخطأ التقدير:

$$e= Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

وفي حالة σ غير معروف نعوض عن σ بـ S لنجد:

$$e= Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

وبما أن معامل الثقة الأكثر استخداما في التطبيقات العملية هو المعامل 0.95 أو 95% . فقد جرت العادة على كتابة الحد الأعلى التقريبي لخطأ التقدير على الشكل:

$$e= 1.96 \ \sigma_{\bar{x}} = 1.96 \frac{\sigma}{\sqrt{n}}$$

حيث نعني بـ $\sigma_{\bar{x}}$ الانحراف المعياري لمتوسط العينة $\bar{x}$، وهو وفقا لنظرية النهاية المركزية $\frac{\sigma}{\sqrt{n}}$. وبما أن النتائج تقريبية، علي أي حال، فقد جرت العادة أيضا على استخدام 2 بدلا من 1.96 ، تسهيلا للحسابات، وهكذا نكتب:

$$e= 2\sigma_{\bar{x}} = 2\frac{\sigma}{\sqrt{n}}$$

أو في حالة σ غير معروف:

$$e= 2\frac{S}{\sqrt{n}}$$

ويجدر التنويه بخطأ شائع بالنسبة إلى المبتدئين، ينبغي الانتباه إليه، وهو استخدام 2σ كحد أعلى تقريبي للخطأ بدلا من $2\sigma_{\bar{x}}$.

مثال (٥ ـ ٢١)

لنفرض أننا نرغب في تقدير متوسط الإنتاج اليومي في شركة للصناعات الكيميائية. وقد سجلنا الإنتاج اليومي لفترة $n = 60$ يوما فكان متوسط هذه العينة وانحرافها المعياري بالأطنان:

$$S = 23 \quad , \bar{X} = 941$$

والمطلوب تقدير μ متوسط الإنتاج اليومي في هذه الشركة.

الحل

التقدير الأفضل هو $\mu = 941$ طنا في اليوم. وحدود الخطأ بالزيادة أو النقصان في هذا التقدير هي:

$$\pm 2\frac{S}{\sqrt{n}} = \pm \frac{2(23)}{\sqrt{60}} = 5.94$$

(تذكر أننا عندما نستخدم العدد 2 يكون معامل الثقة 95%). وهكذا نقول، بمعامل ثقة 95%، إن التقدير 941 هو في حدود 5.94 طنا، زيادة أو نقصانا، من القيمة الحقيقية لمتوسط الإنتاج.

مثال (٥ ـ ٢٢)

نعلم أن عمر مركبة معينة من دائرة كهربائية يتبع توزيعا احتماليا ملتويا. أخذنا عينة عشوائية من 250 من هذه المركبات فكان متوسط العمر فيها 840 ساعة بانحراف معياري $S = 21.98$ ساعة. أوجد فترة ثقة تقريبية لمتوسط العمر في مجتمع إنتاج هذه المركبات، مستخدما معامل ثقة 95%.

الحل

فترة الثقة المطلوبة هي:

$$\bar{X} \pm 2\frac{S}{\sqrt{n}}$$

أي

$$840 \pm 2 \frac{21.98}{\sqrt{250}} = 840 \pm 2.78$$

وهكذا نقدر بمعامل ثقة 95% أن متوسط العمر في مجتمع إنتاج هذه المركبات واقع بين 837.2 و 842.8 ساعة .

مثال (٥ ـ ٢٣)

نريد تقدير μ متوسط الطول في إنتاج مصنع للبراغي في حدود خطأ لا يزيد عن 1/2 مم إلا باحتمال لا يتجاوز الخمسة في المائة. فكم يجب أن يكون حجم العينة علما بأننا نعرف من سجلات الإنتاج السابقة أن الانحراف المعياري للطول يساوي 1.2 مم؟

الحـــل

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.96 \frac{1.2}{\sqrt{n}} \leq \frac{1}{2}$$

ومنه

$$\sqrt{n} \geq 2\,(1.96)(1.2) = 4.704$$

$$n \geq 22.1$$

أي أن حجم العينة يجب ألا يقل عن 23 .

تمارين (٥ ـ ٨)

١) من المعروف أن عمر أحد عناصر دائرة كهربائية يتبع توزيعا احتماليا ملتويا. وقد وجدنا أن متوسط العمر في عينة عشوائية من 300 عنصرا يساوي 920 ساعة بتباين يساوي 483 ساعة². ضع فترة ثقة تقريبية لمتوسط العمر الحقيقي لهذا العنصر مستخدما معامل ثقة 95% .

٢) في عينة عشوائية من 100 كيس تفاح كتب عليه «الوزن الصافي 1 كغ» وجدنا أن متوسط الوزن 1002 غراما بتباين يساوي 144 غ^{٢} . ضع فترة ثقة تقريبية لمتوسط وزن التفاح الحقيقي ضمن الكيس الواحد وذلك بمعامل ثقة 90% .

٣) عينة عشوائية من 60 مخزنا أظهرت أن متوسط سعر الحليب $\bar{X} = 77.3$ سنتا للكيلوغرام، بانحراف معياري 4.2 سنتا . أوجد فترة ثقة لمتوسط سعر الحليب بمعامل ثقة 95% .

٤) مالك سيارة يريد أن يعرف المتوسط الأسبوعي للمسافة التي يسيرها مقاسة بالميل. وقد سجل المسافات التي قطعها في 52 أسبوعا متتاليا ووجد متوسطها 176 ميلا في الأسبوع، بانحراف معياري 96 ميلا . ضع فترة ثقة لمتوسط ما يقطعه في الأسبوع بمعامل 95% .

٥) يرغب مستشفى في تقدير عدد الأيام التي يحتاجها علاج مرضى يقع سنهم بين 25 و 34 سنة . وقد وجدت إدارة المستشفى أن متوسط عدد أيام الإقامة لعينة عشوائية من 500 مريض من هذه الفترة من العمر، يساوي 5.4 يوما بانحراف معياري 1.3 يوما . ضع فترة ثقة لمتوسط الإقامة في المستشفى لمجتمع المرضى الذي جاءت منه العينة، وذلك بمعامل ثقة 99% .

٦) لنفرض أن تباين مجتمع $\sigma^2 = 100$ ، وبمعامل ثقة 95% نريد أن يكون تقديرنا $\bar{x}$ في حدود 2.5 وحدة قياس من μ المتوسط الحقيقي للمجتمع . كم يجب أن يكون حجم العينة؟

٧) فيما يلي جدول التوزيع التكراري للعمر عند الزواج، لأقرب سنة، لـ 175 رجلا :

مركز الفئة	17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5	62.5
التكرار	28	28	43	18	9	4	2	1	0	2

ا ـ أحسب متوسط العمر عند الزواج وانحرافه المعياري .

ب ـ إذا افترضنا أن هذه الأعمار عينة عشوائية من مجتمع كبير. فاحسب بمعامل ثقة 95% فترة ثقة لمتوسط العمر عند الزواج في ذلك المجتمع .

٨) في تجربة لبنج جديد ، أعطي لمائة فأر وقيس زمن الانتعاش لكل منها إلى أقرب عشر من الدقيقة ، فكانت النتائج كما يلي :

45.0	58.2	55.1	52.5	61.7	52.9	70.4	62.5	71.3	50.1
84.9	60.9	35.4	64.3	75.7	48.5	41.3	53.8	66.8	37.4
32.4	50.7	82.3	71.8	66.4	49.7	51.7	56.0	88.8	64.7
77.9	41.4	52.7	53.4	57.9	51.7	55.6	44.1	85.4	67.3
87.3	52.5	46.7	48.3	60.1	66.0	77.3	46.5	54.3	52.6
53.1	67.9	55.9	64.2	68.0	48.2	41.2	56.3	79.4	80.9
58.7	49.0	51.2	70.2	54.0	74.6	51.9	42.6	95.4	51.9
83.5	70.4	76.7	47.0	55.9	43.8	49.1	60.0	38.3	44.3
63.5	45.4	57.3	54.5	73.9	64.1	80.6	68.8	73.5	84.0
65.9	58.3	59.6	59.1	46.7	51.3	44.5	54.2	63.8	56.9

والمطلوب تقدير متوسط زمن الانتعاش في المجتمع الذي جاءت منه العينة وإعطاء حد أعلى لخطأ التقدير باحتمال 0.99 .

كم فأرا تحتاج كي يكون خطأ التقدير 1 تقريبا؟

(٥ ـ ١٠) فترة الثقة لنسبة

لنفرض أن صنفا معينا A يوجد في مجتمع كبير بنسبة تساوي π . إذا سحبنا عينة عشوائية حجمها n من هذا المجتمع ، وعرفنا النجاح بأنه الحصول على عنصر من الصنف A ، فإن احتمال النجاح عند كل سحب هو، عمليا ، π . ونسبة النجاح في

العينة هي عدد عناصر الصنف A ولنرمز لها بـ X (أي عدد النجاحات) مقسوما على حجم العينة n . وإذا رمزنا لنسبة الصنف A في العينة بـ p فإن :

$$p = \frac{X}{n}$$

وعندما ناقشنا تقريب التوزيع الثنائي بالتوزيع الطبيعي في الفقرة (٥ ـ ٦) وجدنا أنه يمكن اعتبار عدد النجاحات X مجموع عينة، وبالتالي تكون النسبة p هي متوسط العينة. وكما أن تطبيق نظرية النهاية المركزية على X يسمح لنا بالقول إن X يتوزع تقريبا وفق التوزيع الطبيعي $N(n\pi\,,\,n\pi(1-\pi))$ ، فإن تطبيق نظرية النهاية المركزية على المتوسط p يسمح لنا بالقول إن للنسبة p (نسبة النجاح في العينة) توزيعا مطابقا تقريبا للتوزيع الطبيعي $N\left(\pi\,,\,\frac{\pi\,(1-\pi)}{n}\right)$ ، حيث :

$$E(p) = E\left(\frac{X}{n}\right) = \frac{1}{n}\,E(X) = \frac{n\pi}{n} = \pi$$

$$V(p) = V\left(\frac{X}{n}\right) = \frac{1}{n^2}\,V(X) = \frac{n\pi(1-\pi)}{n^2} = \frac{\pi(1-\pi)}{n}$$

وذلك كله شريطة أن يكون n كبيرا (مثلا أكبر من 30 في حالة قيمة لـ π ليست قريبة من الصفر أو قريبة من الواحد). وهذه النتيجة تسمح لنا بوضع فترة ثقة للنسبة π على الشكل التالي بمعامل ثقة $100\,(1-\mu)\%$:

$$p \pm Z_{\alpha/2}\,\sigma_p = p \pm Z_{\alpha/2}\sqrt{\frac{\pi\,(1-\pi)}{n}}$$

حيث σ_p تعني الانحراف المعياري للنسبة p . ووجود π المجهولة في هذه العبارة يمنع من تطبيقها. ويمكننا الاستعاضة عن π ، نسبة النجاح في المجتمع، بتقدير لها هو p ، نسبة النجاح في العينة. (تماما كما عوضنا عن σ بـ s في الفقرة السابقة) . وتصبح فترة الثقة لـ π كما يلي :

$$p \pm Z_{\alpha/2}\sqrt{\frac{p\,(1-p)}{n}}$$

ومن أجل معـامل ثقة 95% يمكن اعتبـار $Z_{\alpha/2}$ مساويـا لـ 2 بـدلا من 1.96 ، تسهيـلا للحساب . وهكذا نكتب فترة الثقة لـ π ، بمعامل ثقة 95% كما يلي :

$$p-2\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p+2\sqrt{\frac{p(1-p)}{n}}$$

مثال (٥ ـ ٢٤)

مـن بين 300 أسـرة اخترنـاها من بلدة كبيرة وجدنا 123 أسرة تمتلك تلفازا ملونا . ضـع فترة ثقـة لنسبـة الأسر التي تمتلك تلفـازا ملونـا في مجمل البلـدة . وذلك بمعامل ثقة 95% .

الحل

لدينا $n = 300$ ، $p=\frac{123}{300}=0.41$

$$\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{0.41\times 0.59}{300}}=\sqrt{0.00080633}=0.0284$$

وهكذا تكون فترة الثقة المطلوبة كما يلي :

$$0.41 \pm 2\,(0.0284) = 0.41 \pm 0.057$$

أي أن π واقعة بين 0.353 و 0.467 أو أن ما بين 35.3% إلى 46.7% من الأسر في هذه البلدة تمتلك تلفازا ملونا .

مثال (٥ ـ ٢٥)

تضمنت عينة من 250 من طلبة الجامعة 30 طالبا أعسر (يستخدم اليد اليسرى) . أعط فترة ثقة تقريبية بمعامل ثقة 95% لنسبة الطلاب العسر في الجامعة؟

الحل

$n = 250$ ، $p=\frac{30}{250}=0.12$

$$\sqrt{\frac{p\,(1-p)}{n}} = \sqrt{\frac{0.12 \times 0.88}{250}} = \sqrt{0.0004224} = 0.021$$

وتكون فترة الثقة المطلوبة :

$$0.12 \pm 2\,(0.021) = 0.12 \pm 0.042$$

أي أن ما بين 7.8% إلى 16.2% من طلبة الجامعة يستخدمون اليد اليسري .

تمارين (٥ ـ ٩)

١) اختبرنا نوعا جديدا من مصابيح آلات التصوير لتقدير p ، احتمال أن ينتج المصباح الإضاءة اللازمة وفي الوقت المناسب . ووجدنا من بين 1000 مصباح أن 920 قد عملت وفقا للمواصفات المطلوبة . والمطلوب :

ا ـ تقدير p ووضع حد لخطأ التقدير (باحتمال 0.95) .

ب ـ وضع فترة ثقة للقيمة الحقيقية p بمعامل ثقة 99% .

٢) اختبرنا عينة عشوائية من 400 صماما خاصا بأجهزة الراديو، فوجدنا من بينها 40 صماما عاطلا عن العمل . ضع فترة ثقة للنسبة الحقيقية للصمامات العاطلة في مجتمع الصمامات المنتجة من هذا النوع، مستخدما معامل ثقة 90% .

٣) حضر كيميائي مبيدا يهدف عند تطبيقه على نوع معين من الحشرات، إلى قتل 60% منها فكم يجب أن يكون حجم العينة المستخدمة ، إذا رغب في أن يطمئن باحتمال 0.95 إلى أنه في حدود 0.02 من النسبة الحقيقية التي يهدف إليها من الحشرات الهالكة؟

٤) كم ناخبا يجب أن تضم عينة جمعناها لتقدير نسبة الناخبين الذين يفضلون مرشحا معينا ، إذا رغبنا في أن يكون التقدير صحيحا في حدود 0.005 ؟ ولنفرض أن النسبة الحقيقية ينبغي أن تقع في جوار الـ 0.5 .