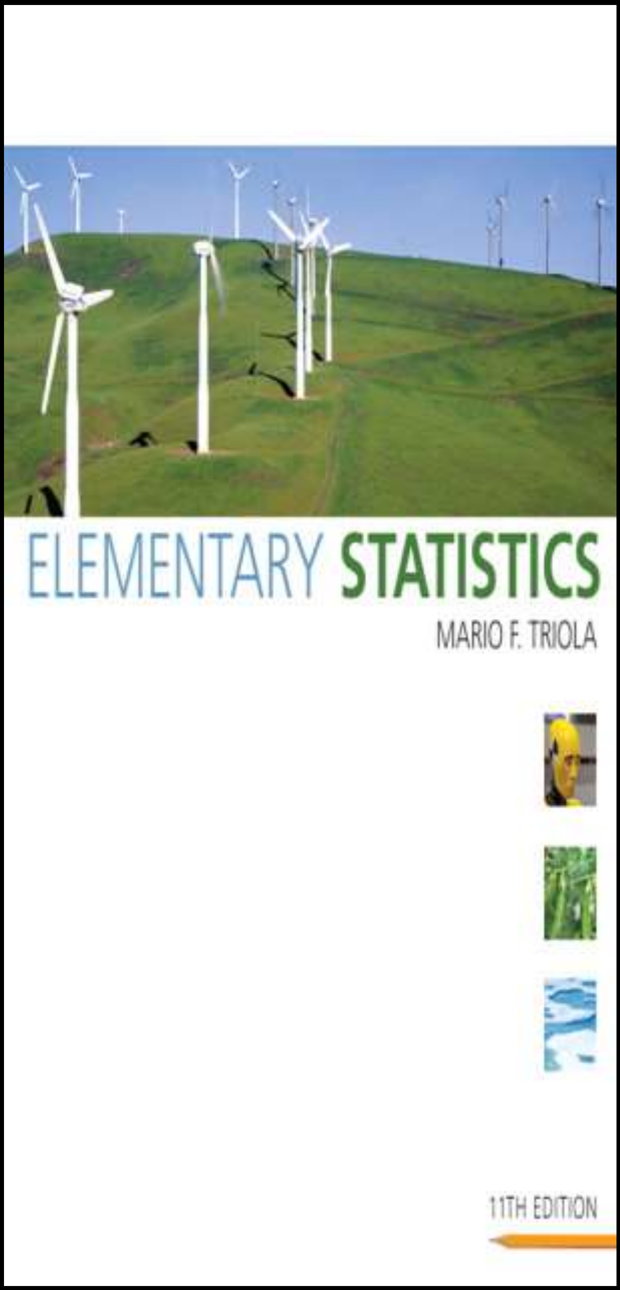
Lecture Slides



Elementary Statistics
Eleventh Edition

and the Triola Statistics Series

by Mario F. Triola

PEARSON

Chapter 6
Normal Probability Distributions
الفصل 6 التوزيعات الاحتمالية العادية

Key Concept
المفهوم الرئيسي

This section presents the *standard normal distribution* which has three properties:

1. It's graph is bell-shaped.

2. It's mean is equal to 0 ($\mu = 0$).

3. It's standard deviation is equal to 1 ($\sigma = 1$).

Develop the skill to find areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution. Find *z*-scores that correspond to area under the graph.

يعرض هذا القسم التوزيع الطبيعي القياسي الذي يحتوي على ثلاث خصائص:

1إنه رسم بياني على شكل جرس.

2هذا يعني تساوي 0 (0= ميو).

3الانحراف المعياري يساوي 1 (1= كيو).

طور المهارة للعثور على المناطق (أو الاحتمالات أو الترددات النسبية) المقابلة لمختلف المناطق تحت التي تتوافق مع مساحة تحت الرسم البياني zالرسم البياني للتوزيع الطبيعي القياسي. العثور على عشرات

التوزيع الموحد

A continuous random variable has a uniform distribution if its values are spread evenly over the range of probabilities. The graph of a uniform distribution results in a rectangular shape.

يحتوي المتغير العشوائي المستمر على توزيع موحد إذا كانت قيمه موزعة بالتساوي على مدى الاحتمالات. يؤدي الرسم البياني للتوزيع الموحد إلى شكل مستطيل.

Density Curve منحنى الكثافة

A density curve is the graph of a continuous probability distribution. It must satisfy the following properties:

منحنى الكثافة هو الرسم البياني لتوزيع الاحتمال المستمر. يجب أن تفي بالخصائص التالية:

1. The total area under the curve must equal 1.
2. Every point on the curve must have a vertical height that is 0 or greater. (That is, the curve cannot fall below the *x*-axis.)

1. يجب أن تساوي المساحة الإجمالية تحت المنحنى 1.

2. يجب أن يكون لكل نقطة على المنحنى ارتفاع رأسي يساوي 0 أو أكبر. (وهذا يعني أن المنحنى لا يمكن أن يقع أسفل المحور السيني.)

Area and Probability المنطقة والاحتمال

Because the total area under the density curve is equal to 1, there is a correspondence between area and probability.

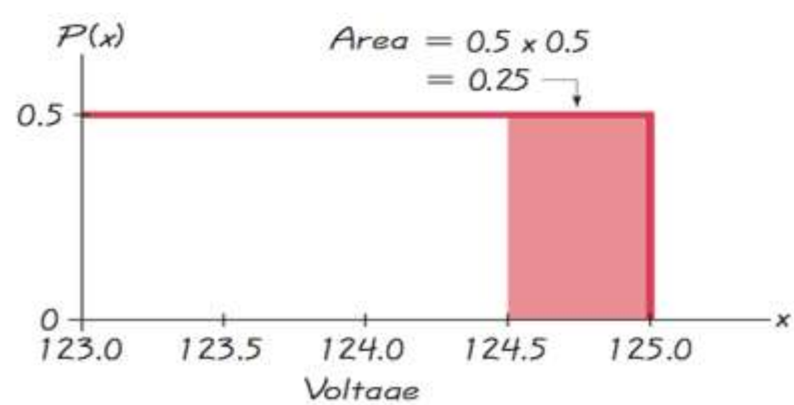
نظرًا لأن المساحة الكلية تحت منحنى الكثافة تساوي 1 ، فهناك علاقة بين المنطقة والاحتمال.

Using Area to Find Probability استخدام المنطقة لإيجاد الاحتمالات

Given the uniform distribution illustrated, find the probability that a randomly selected voltage level is greater than 124.5 volts.

إذا وجدت توزيعًا موحدًا مصورًا ، ابحث عن احتمال أن يكون مستوى الجهد المحدد عشوائيًا أكبر من 124.5 فولت.

تمثل المنطقة المظللة مستويات الجهد أكبر من 124.5 فولت. المراسلات بين المنطقة والاحتمال: 0.25.

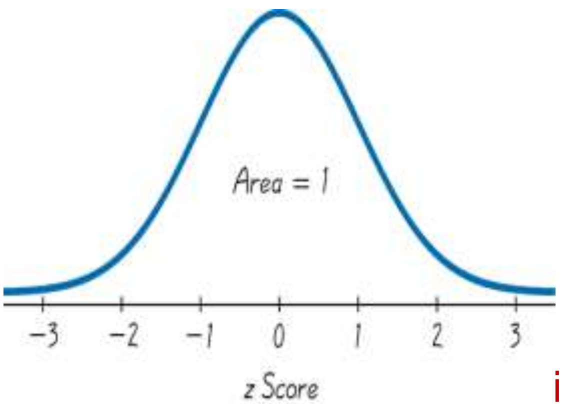


Standard Normal Distribution
التوزيع القياسي

The **standard normal distribution** is a normal probability distribution with $\mu = 0$ and $\sigma = 1$. The total area under its density curve is equal to 1.

التوزيع الطبيعي القياسي هو توزيع احتمالي طبيعي مع ميو= 0 و كيو=1

المساحة الكلية تحت منحنى الكثافة تساوي 1



ities When Given z-scores

العثور على الاحتمالات عند إعطاء زي درجات .

Table A-2 (in Appendix A)

Formulas and Tables insert card

Find areas for many different regions

الجدول أ -2 (في الملحق أ)

الصيغ والجداول إدراج بطاقة

البحث عن مناطق للعديد من المناطق المختلفة

Finding Probabilities – Other Methods
إيجاد الاحتمالات - طرق أخرى

- STATDISK
- Minitab
- Excel
- TI-83/84 Plus

Methods for Finding Normal Distribution Areas
طرق لإيجاد مناطق التوزيع العادية

Table A-2, STATDISK, Minitab, Excel

Gives the cumulative area from the left up to a vertical line above a specific value of z.

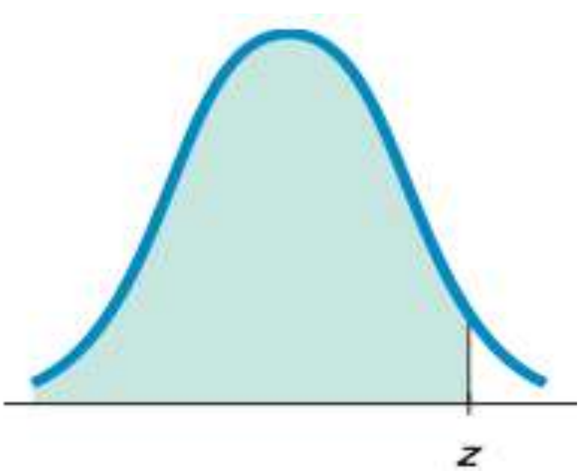


Table A-2 The procedure for using Table A-2 is described in the text.

STATDISK Select **Analysis, Probability Distributions, Normal Distribution.** Enter the *z* value, then click on **Evaluate.**

MINITAB Select **Calc, Probability Distributions, Normal.** In the dialog box, select **Cumulative Probability, Input Constant.**

EXCEL Select **fx, Statistical, NORMDIST.** In the dialog box, enter the value and mean, the standard deviation, and "true."

Methods for Finding Normal Distribution Areas

طرق لإيجاد مناطق التوزيع العادية

TI-83/84 Plus Calculator

Gives area bounded on the left and bounded on the right by vertical lines above any specific values.

TI-83/84 Press **2ND** **VARS**

[2: normal cdf ( ], then enter the two *z* scores separated by a comma, as in (left *z* score, right *z* score).

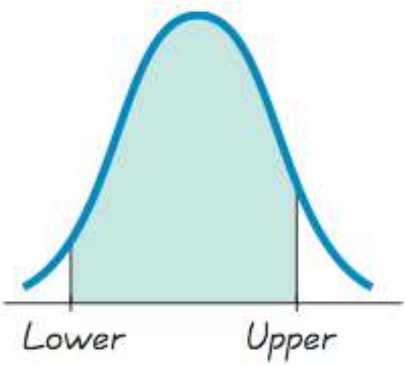


Table A-2

TABLE A-2 Standard Normal (*z*) Distribution: Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 and lower	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	* .0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	* .0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559

Using Table A-2 باستخدام الجدول A-2

1. It is designed only for the *standard* normal distribution, which has a mean of 0 and a standard deviation of 1.

إنه مصمم فقط للتوزيع العادي القياسي ، والذي يكون له متوسط 0 وانحراف معياري 1.

1. It is on two pages, with one page for *negative z*-scores and the other page for *positive* *z*-scores.

إنه موجود على صفحتين ، مع صفحة واحدة للنتائج السلبية للإشارة السلبية والصفحة الأخرى للنتائج الموجبة للإيجابية.

1. Each value in the body of the table is a *cumulative area from the left* up to a vertical boundary above a specific *z*-score.

كل قيمة في نص الجدول هي منطقة تراكمية من اليسار إلى الحد الرأسي أعلى درجة محددة زي

Using Table A-2 باستخدام الجدول A-2

4.When working with a graph, avoid confusion between z-scores and areas. z Score

Distance along horizontal scale of the standard normal distribution; refer to the leftmost column and top row of Table A-2

4. عند العمل مع الرسم البياني ، وتجنب الخلط بين عشرات و النتيجه زي

المسافة على طول المقياس الأفقي للتوزيع الطبيعي القياسي ؛ الرجوع إلى العمود في أقصى اليسار والصف العلوي من الجدول A-2.

5.The part of the z-score denoting hundredths is found across the top.

5. تم العثور على جزء من علامة زي تشير إلى المئات عبر الجزء العلوي.

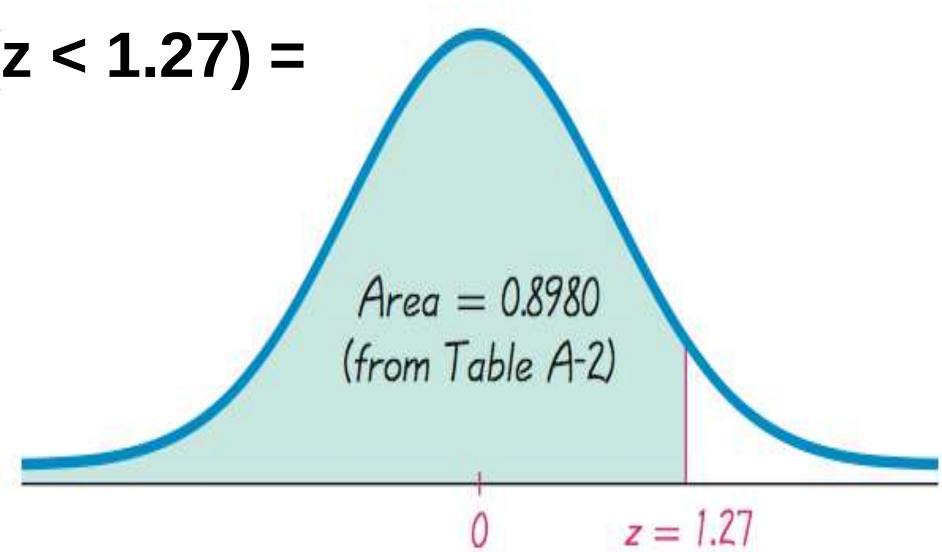
Example – Thermometers
مثال - موازين الحرارة

The Precision Scientific Instrument Company manufactures thermometers that are supposed to give readings of 0ºC at the freezing point of water. Tests on a large sample of these instruments reveal that at the freezing point of water, some thermometers give readings below 0º (denoted by negative numbers) and some give readings above 0º (denoted by positive numbers). Assume that the mean reading is 0ºC and the standard deviation of the readings is 1.00ºC. Also assume that the readings are normally distributed. If one thermometer is randomly selected, find the probability that, at the freezing point of water, the reading is less than 1.27º.

تقوم شركة Precision Scientific Instrument بتصنيع موازين الحرارة التي من المفترض أن تعطي قراءات 0 درجة مئوية عند نقطة التجمد المائية. تكشف الاختبارات التي أجريت على عينة كبيرة من هذه الأدوات أنه عند نقطة التجمد المائية تعطي بعض موازين الحرارة قراءات أقل من 0 º (يرمز إليها بالأرقام السالبة) ويعطي البعض قراءات أعلى من 0 º (يرمز إليها بالأرقام الموجبة). افترض أن متوسط القراءة هو 0 درجة مئوية وأن الانحراف المعياري للقراءات هو 1.00 درجة مئوية. افترض أيضًا أن القراءات توزع عادةً. إذا تم اختيار مقياس حرارة واحد بشكل عشوائي ، ابحث عن احتمال أن تكون القراءة عند نقطة التجمد المائية أقل من 1.27 درجة مئوية.

Example - (Continued)

$P(z < 1.27) =$



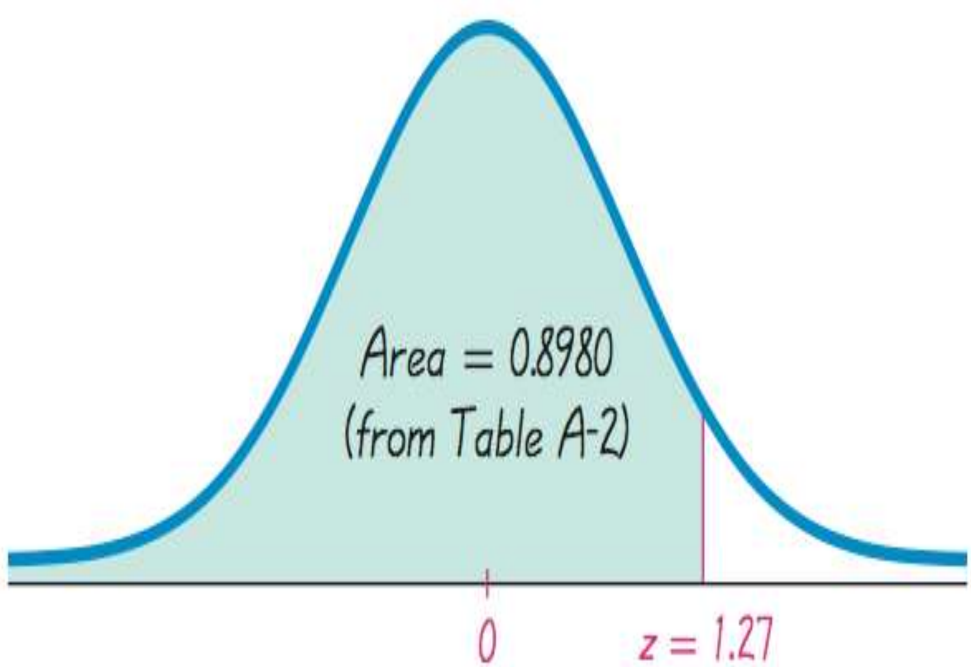
Look at Table A-2 انظر إلى الجدولA-2

TABLE A-2 (*continued*) Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292

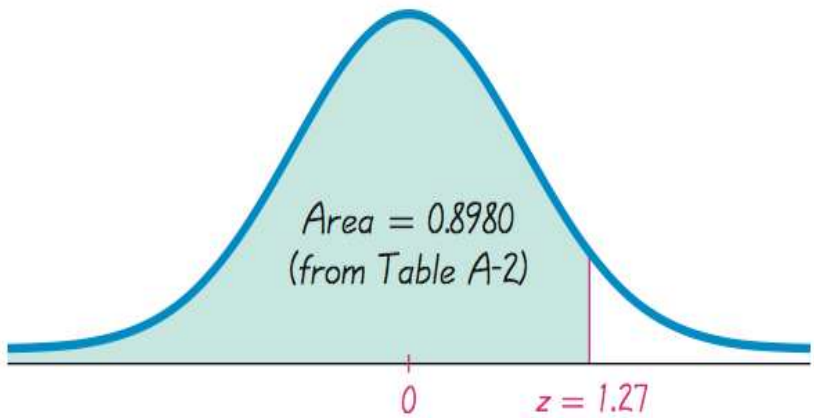
Example - cont

$$P(z < 1.27) = 0.8980$$



Example – cont مثال - تابع

$P(z < 1.27) = 0.8980$

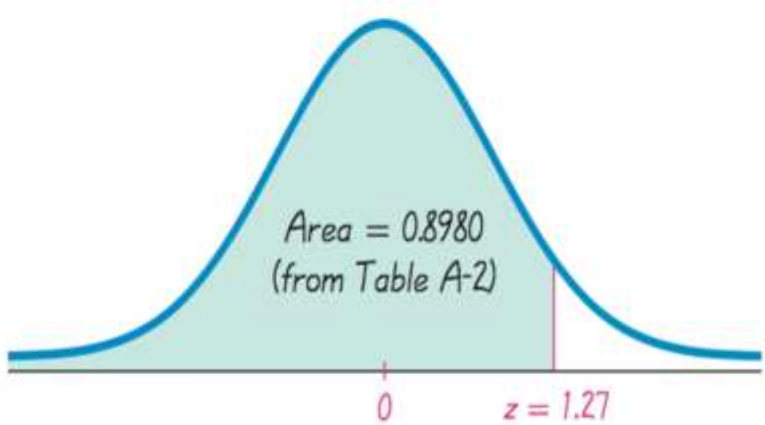


The *probability* of randomly selecting a thermometer with a reading less than 1.27º is 0.8980.

هو ºاحتمال اختيار عشوائي لمقياس حرارة مع قراءة أقل من 1.27 0.8980.

Example - cont مثال – تابع

$P(z < 1.27) = 0.8980$



Or 89.80% will have readings below 1.27º.

أو 89.80 ٪ سيكون لها قراءات أقل من 1.27º.

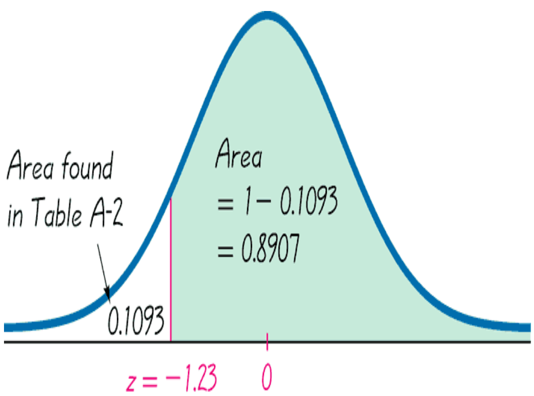
Example - Thermometers Again

مثال - موازين الحرارة مرة أخرى

If thermometers have an average (mean) reading of 0 degrees and a standard deviation of 1 degree for freezing water, and if one thermometer is randomly selected, find the probability that it reads (at the freezing point of water) above –1.23 degrees.

إذا كانت موازين الحرارة لها متوسط قراءة (متوسط) قدره 0 درجة وانحراف معياري قدره 1 درجة لمياه التجميد ، وإذا تم اختيار مقياس حرارة واحد بشكل عشوائي ، فابحث عن احتمال قراءته (عند نقطة التجمد للماء) أعلى من -1.23 درجة

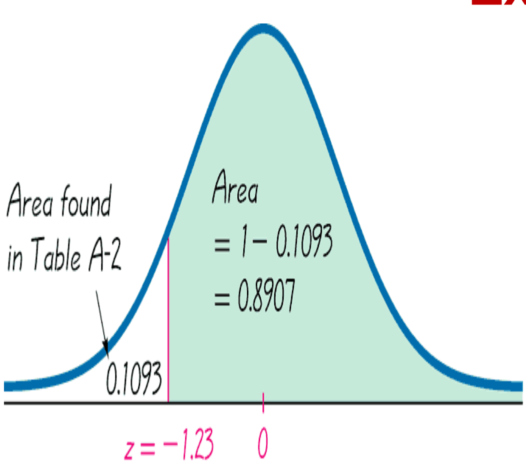
$$P\ (z > -1.23) = 0.8907$$



Probability of randomly selecting a thermometer with a reading above –1.23º is 0.8907.

احتمال اختيار عشوائي لمقياس الحرارة مع قراءة أعلاه –1.23º هو 0.8907.

Example - cont



89.07% of the thermometers have readings above –1.23 degrees.

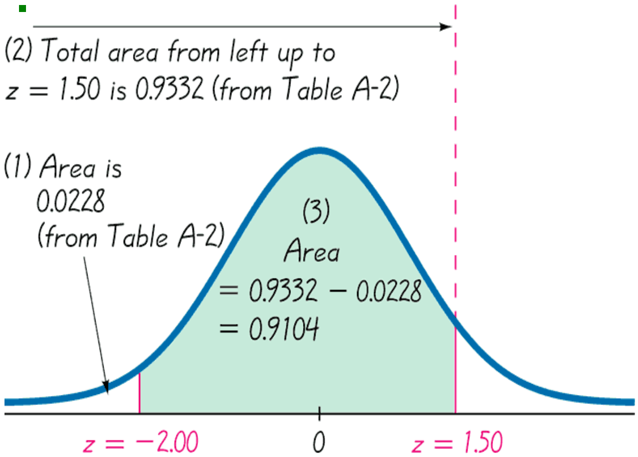
89.07 ٪ من موازين الحرارة لها قراءات أعلى من -1.23 درجة

Example - Thermometers III

مثال - موازين الحرارة III

A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.

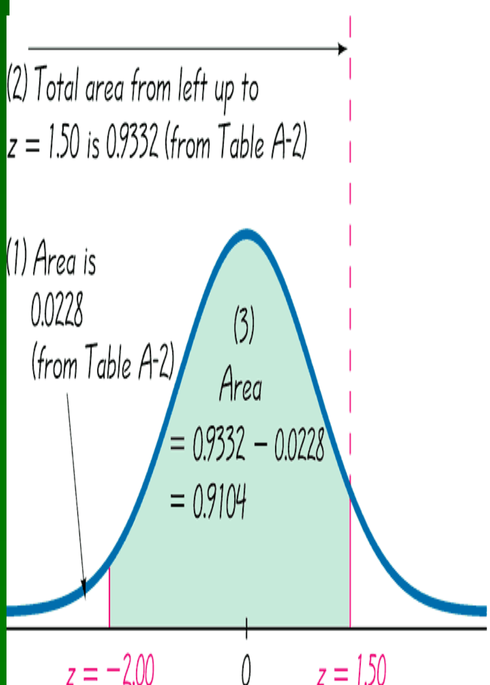
يتم اختيار مقياس الحرارة بشكل عشوائي. أوجد احتمال أن يقرأ (عند نقطة التجمد للماء) بين - 2.00 و 1.50 درجة



$P(z < -2.00) = 0.0228$
$P(z < 1.50) = 0.9332$
$P(-2.00 < z < 1.50) =$
$0.9332 - 0.0228 = 0.9104$

The probability that the chosen thermometer has a reading between – 2.00 and 1.50 degrees is 0.9104.

احتمال أن يكون لدى مقياس الحرارة المختار قراءة بين 2.00 و 1.50 درجة هو 0.9104.



Example - cont

A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.

يتم اختيار مقياس الحرارة بشكل عشوائي. أوجد احتمال أن يقرأ (عند نقطة التجمد للماء) بين -2.00 و 1.50 درجة.

$P(z < -2.00) = 0.0228$
$P(z < 1.50) = 0.9332$
$P(-2.00 < z < 1.50) =$
$0.9332 - 0.0228 = 0.9104$

If many thermometers are selected and tested at the freezing point of water, then 91.04% of them will read between –2.00 and 1.50 degrees.

إذا تم اختيار العديد من موازين الحرارة واختبارها عند نقطة تجمد الماء ، فسيتم قراءة 91.04٪ منها بين -2.00 و 1.50 درجة.

Notation الرموز

P(*a* < z < *b*)

denotes the probability that the *z* score is between *a* and *b.*

يدل على احتمال أن تكون النتيجة z بين a و b.

P(z > *a*)

denotes the probability that the *z* score is greater than *a.*

يدل على احتمال أن تكون النتيجة z أكبر من a.

P(z < *a*)

denotes the probability that the z score is less than *a.*

يدل على احتمال أن تكون النتيجة z أقل من a.

Finding a z Score When Given a Probability Using Table A-2

العثور على نقاط z عند إعطاء احتمال باستخدام جدول A-2

1. **Draw a bell-shaped curve and identify the region under the curve that corresponds to the given probability. If that region is not a cumulative region from the left, work instead with a known region that is a cumulative region from the left.**

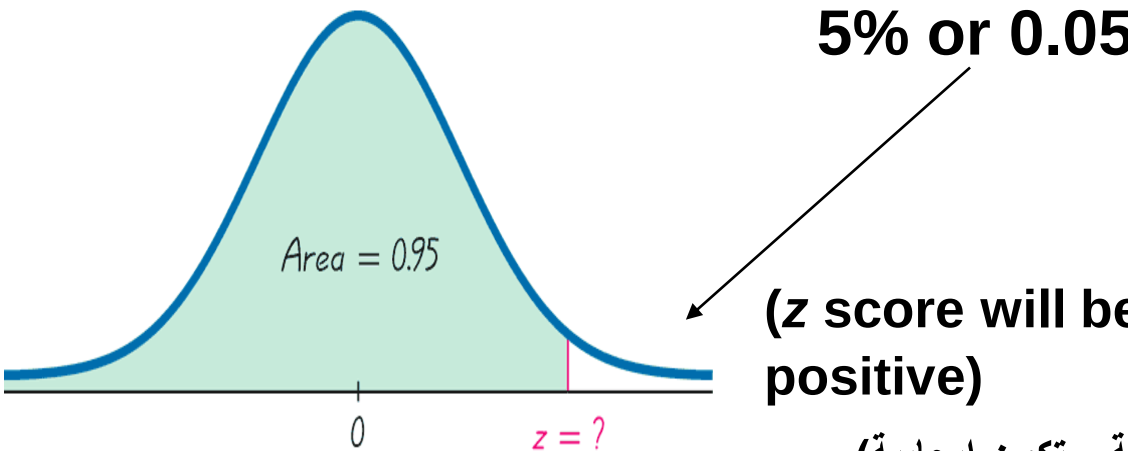
1. رسم منحنى على شكل جرس وتحديد المنطقة تحت المنحنى الذي يتوافق مع احتمال معين. إذا لم تكن تلك المنطقة منطقة تراكمية من اليسار ، فعمل بدلاً من ذلك مع منطقة معروفة تمثل منطقة تراكمية من اليسار.

2. Using the cumulative area from the left, locate the closest probability in the body of Table A-2 and identify the corresponding z score.

2. باستخدام المنطقة التراكمية من اليسار ، حدد أقرب الاحتمالات في نص الجدول A-2 المقابلة. وحدد النتيجة z.

Finding z Scores When Given Probabilities

العثور على عشرات ض
عندما تعطى الاحتمالات



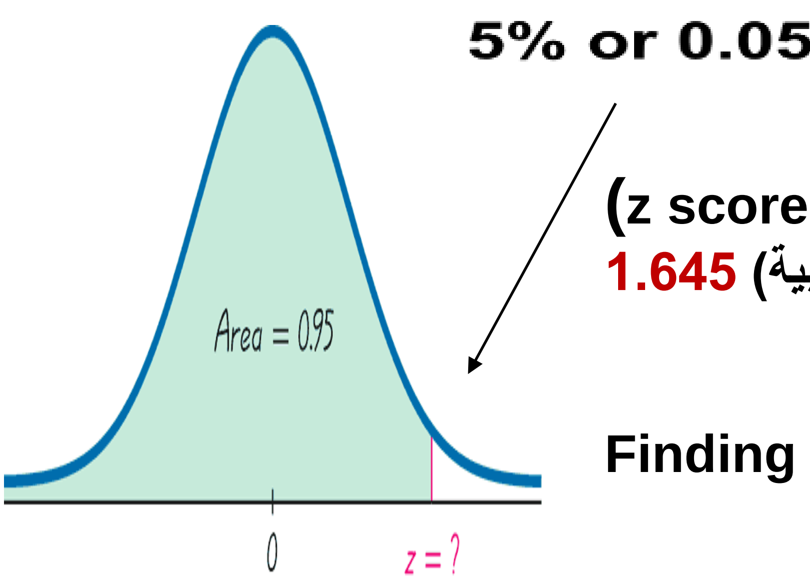
(*z* score will be positive)

(النتيجة ستكون إيجابية)

Finding the 95th Percentile
العثور على 95 في المئة

Finding z Scores When Given Probabilities – cont

العثور على عشرات ض
عندما تعطى الاحتمالات - تابع

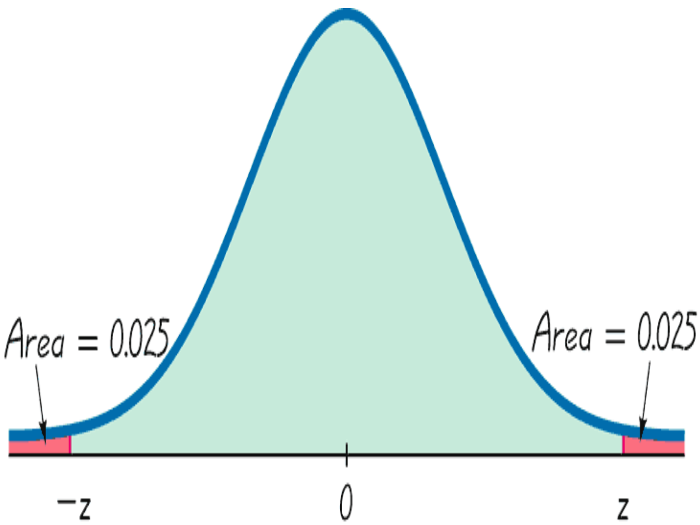


(z score will be positive)
1.645 (النتيجة ستكون إيجابية)

Finding the 95th Percentile
العثور على 95 في المئة

Finding *z* Scores When Given Probabilities – cont

العثور على عشرات ض
عندما تعطى الاحتمالات - تابع



(One *z* score will be negative and the other positive)

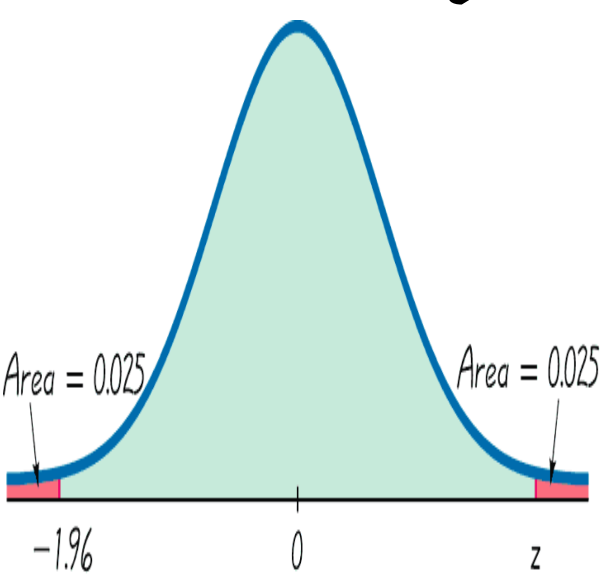
واحدة ستكون z(درجة سالبة والأخرى موجبة)

Finding the Bottom 2.5% and Upper 2.5%

العثور على القاع 2.5 ٪ والعليا 2.5 ٪

Finding z Scores When Given Probabilities – cont

العثور على عشرات ض
عندما تعطى الاحتمالات - تابع



(One z score will be negative and the other positive)

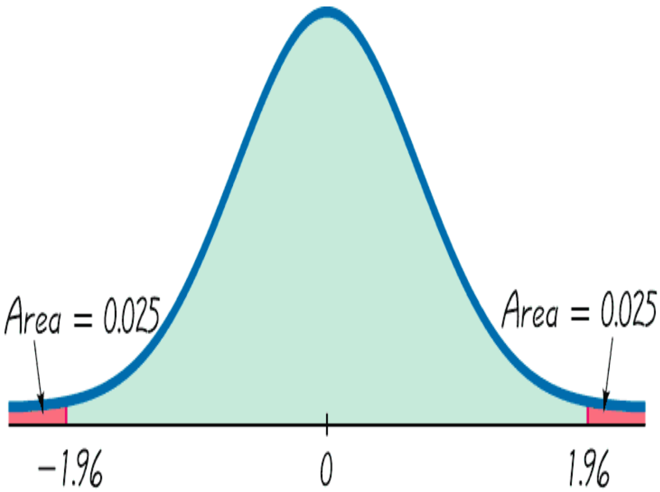
واحدة ستكون z(درجة سالبة والأخرى موجبة)

Finding the Bottom 2.5% and Upper 2.5%

العثور على القاع 2.5 ٪ والعليا 2.5 ٪

Finding *z* Scores When Given Probabilities – cont

العثور على عشرات ض عندما تعطى الاحتمالات - تابع



(One *z* score will be negative and the other positive)

واحدة ستكون سالبة والأخرى z)درجة موجبة)

Finding the Bottom 2.5% and Upper 2.5%

العثور على القاع 2.5 ٪ والعليا 2.5 ٪

Recap

In this section we have discussed:

Density curves.

Relationship between area and probability.

Standard normal distribution.

Using Table A-2.

في هذا القسم ناقشنا:

منحنيات الكثافة.

العلاقة بين المنطقة والاحتمال.

التوزيع القياسي.

A-2. باستخدام الجدول

Key Concept
المفهوم الرئيسي

This section presents methods for working with normal distributions that are not standard. That is, the mean is not 0 or the standard deviation is not 1, or both.

The key concept is that we can use a simple conversion that allows us to standardize any normal distribution so that the same methods of the previous section can be used.

يقدم هذا القسم طرقًا للعمل مع توزيعات عادية غير قياسية. وهذا يعني أن الوسط ليس 0 أو الانحراف المعياري ليس 1 أو كلاهما.

المفهوم الرئيسي هو أنه يمكننا استخدام تحويل بسيط يسمح لنا بتوحيد أي توزيع عادي بحيث يمكن استخدام نفس الأساليب في القسم السابق.

Conversion Formula صيغة التحويل

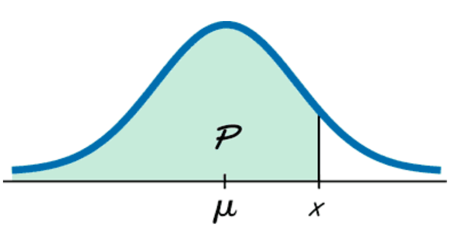
$$z = \frac{x - \mu}{\sigma}$$

Round z scores to 2 decimal places

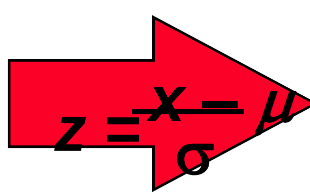
جولة عشرات عشرات إلى 2 المنازل العشرية

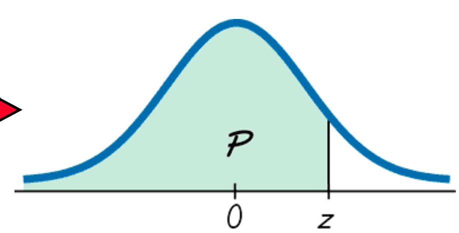
Converting to a Standard Normal Distribution

تحويل إلى توزيع عادي قياسي



(a) Nonstandard Normal Distribution





(b) Standard Normal Distribution

Example – Weights of Water Taxi Passengers
مثال - أوزان ركاب التاكسي المائي

In the Chapter Problem, we noted that the safe load for a water taxi was found to be 3500 pounds. We also noted that the mean weight of a passenger was assumed to be 140 pounds. Assume the worst case that all passengers are men. Assume also that the weights of the men are normally distributed with a mean of 172 pounds and standard deviation of 29 pounds. If one man is randomly selected, what is the probability he weighs less than 174 pounds?

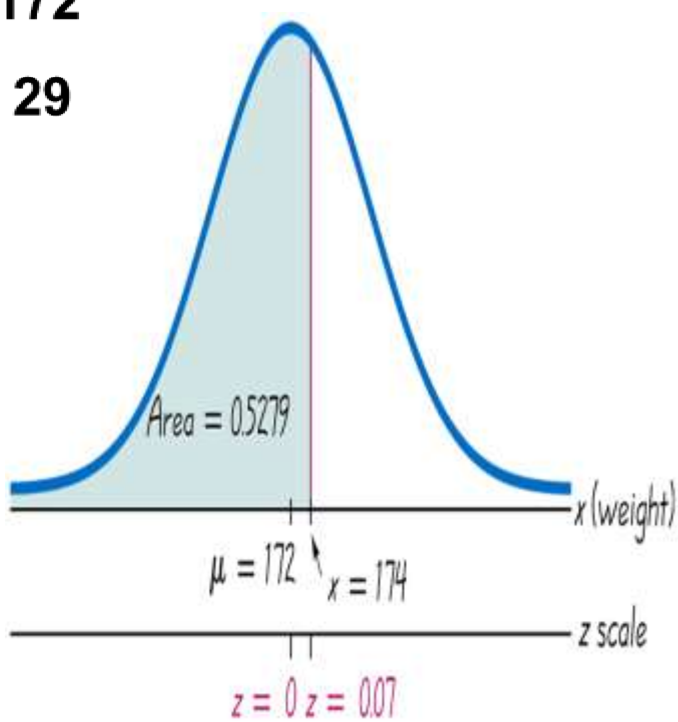
في مشكلة الفصل ، لاحظنا أن الحمولة الآمنة لسيارة أجرة مائية تم العثور عليها بمبلغ 3500 جنيه. لاحظنا أيضًا أن متوسط وزن الراكب كان يفترض أن يكون 140 رطلاً. افترض أسوأ حالة أن جميع الركاب رجال. افترض أيضًا أن أوزان الرجال يتم توزيعها عادة بمتوسط 172 رطل وانحراف معياري قدره 29 رطلاً. إذا تم اختيار رجل واحد بشكل عشوائي ، فما هو احتمال أن يزن أقل من 174 رطلاً؟

Example – cont مثال - تابع

$\mu = 172$

$\sigma = 29$

$$z = \frac{174 - 172}{29} = 0.07$$

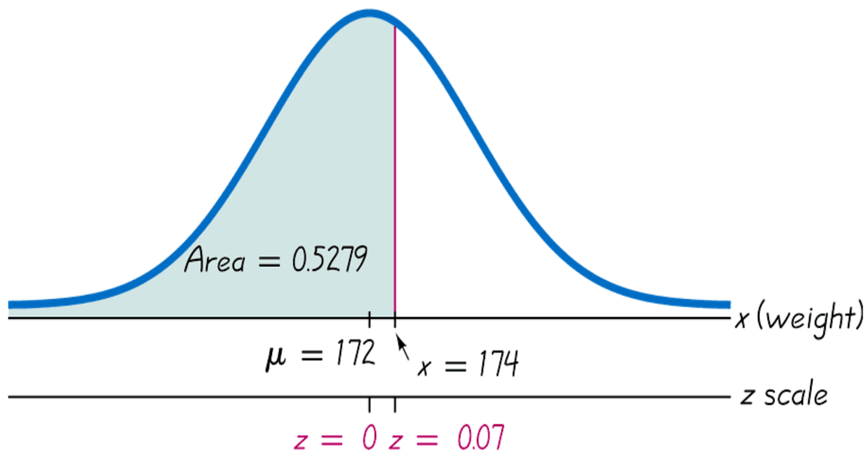


Example – cont مثال - تابع

$\mu = 172$

$\sigma = 29$

P (x < 174 lb.) = P(z < 0.07)
= 0.5279



Helpful Hints
تلميحات مفيدة

1. Don't confuse z scores and areas. z scores are distances along the horizontal scale, but areas are regions under the normal curve. Table A-2 lists z scores in the left column and across the top row, but areas are found in the body of the table.
2. Choose the correct (right/left) side of the graph.
3. A z score must be negative whenever it is located in the left half of the normal distribution.
4. Areas (or probabilities) are positive or zero values, but they are never negative.

هي مسافات على طول المقياس الأفقي ، zوالمناطق. عشرات z. 1 لا تخلط بين عشرات
في العمود zدرجات A-2ولكن المناطق هي مناطق تحت المنحنى العادي. يسرد الجدول
الأيسر وعبر الصف العلوي ، ولكن توجد مساحات في نص الجدول.
.2 اختر الجانب الصحيح (الأيمن / الأيسر) من الرسم البياني.
سالبة كلما وجدت في النصف الأيسر من التوزيع الطبيعي. 3z. يجب أن تكون النتيجة
.4 المناطق (أو الاحتمالات) هي قيم موجبة أو صفرية ، لكنها ليست سلبية أبدًا.

Procedure for Finding Values Using Table A-2 and Formula 6-2
والصيغة 6-2 A-2الإجراء الخاص بإيجاد القيم باستخدام الجدول

1. Sketch a normal distribution curve, enter the given probability or percentage in the appropriate region of the graph, and identify the *x* value(s) being sought.

2. Use Table A-2 to find the *z* score corresponding to the cumulative left area bounded by *x*. Refer to the body of Table A-2 to find the closest area, then identify the corresponding *z* score.

3. Using Formula 6-2, enter the values for μ, σ, and the *z* score found in step 2, then solve for *x*.

   $x = \mu + (z \cdot \sigma)$ (Another form of Formula 6-2)

   (If *z* is located to the left of the mean, be sure that it is a negative number.)

4. Refer to the sketch of the curve to verify that the solution makes sense in the context of the graph and the context of the problem.

1. ارسم منحنى التوزيع الطبيعي ، وأدخل الاحتمال أو النسبة المئوية المعطاة في المنطقة المناسبة من الرسم البياني ، وحدد القيمة (القيم) المطلوبة.

المقابلة للمنطقة اليسرى التراكمية التي تحدها zللعثور على الدرجة 2A-2. استخدم الجدول المقابلة. zللعثور على أقرب منطقة ، ثم حدد النتيجة A-2ارجع إلى نص الجدول x.

xالموجودة في الخطوة 2 ، ثم حل لـ 3z. باستخدام الصيغة 6-2 ، أدخل قيم μ ، والنتيجة

نموذج آخر من الصيغة 6-2)) x = μ + (z • q)

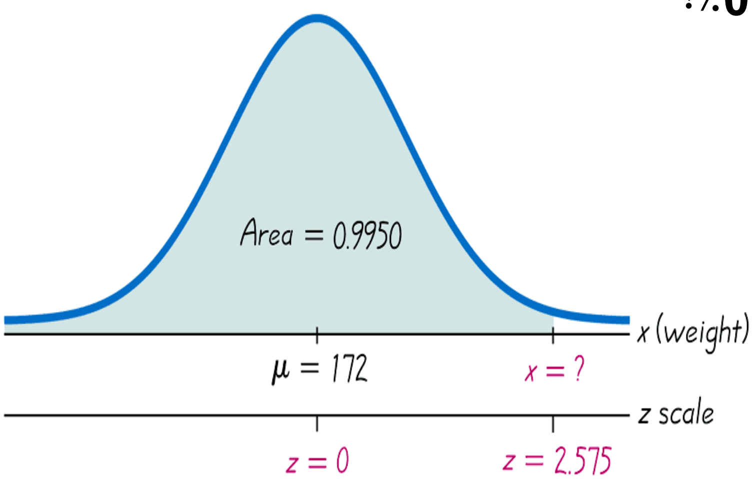
موجودة على يسار الوسط ، فتأكد من أنه رقم سالب.) z(إذا كانت

4. ارجع إلى مخطط المنحنى للتحقق من أن الحل منطقي في سياق الرسم البياني وسياق المشكلة.

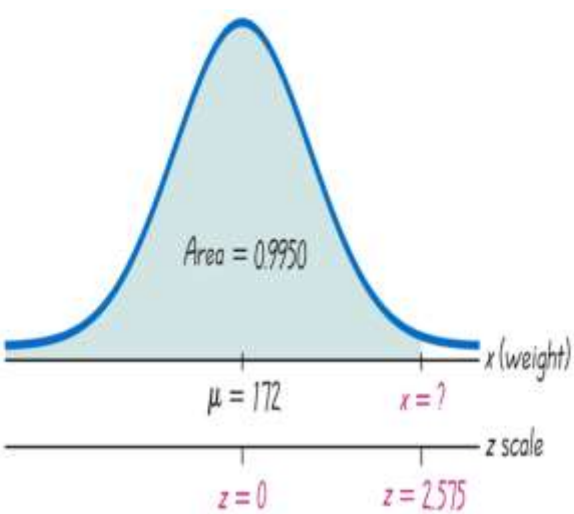
Example – Lightest and Heaviest
مثال - الأخف وزنا والأثقل

Use the data from the previous example to determine what weight separates the lightest 99.5% from the heaviest 0.5%?

استخدم البيانات الواردة في المثال السابق لتحديد الوزن الذي يفصل الأخف وزناً بنسبة 99.5٪ عن أثقل 0.5٪؟



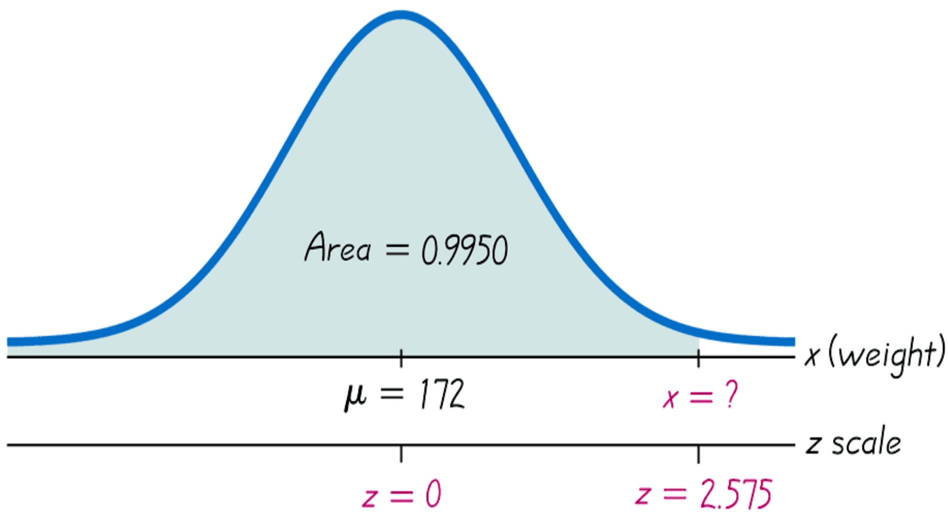
Example – Lightest and Heaviest – cont - مثال الأخف وزنا - تابع

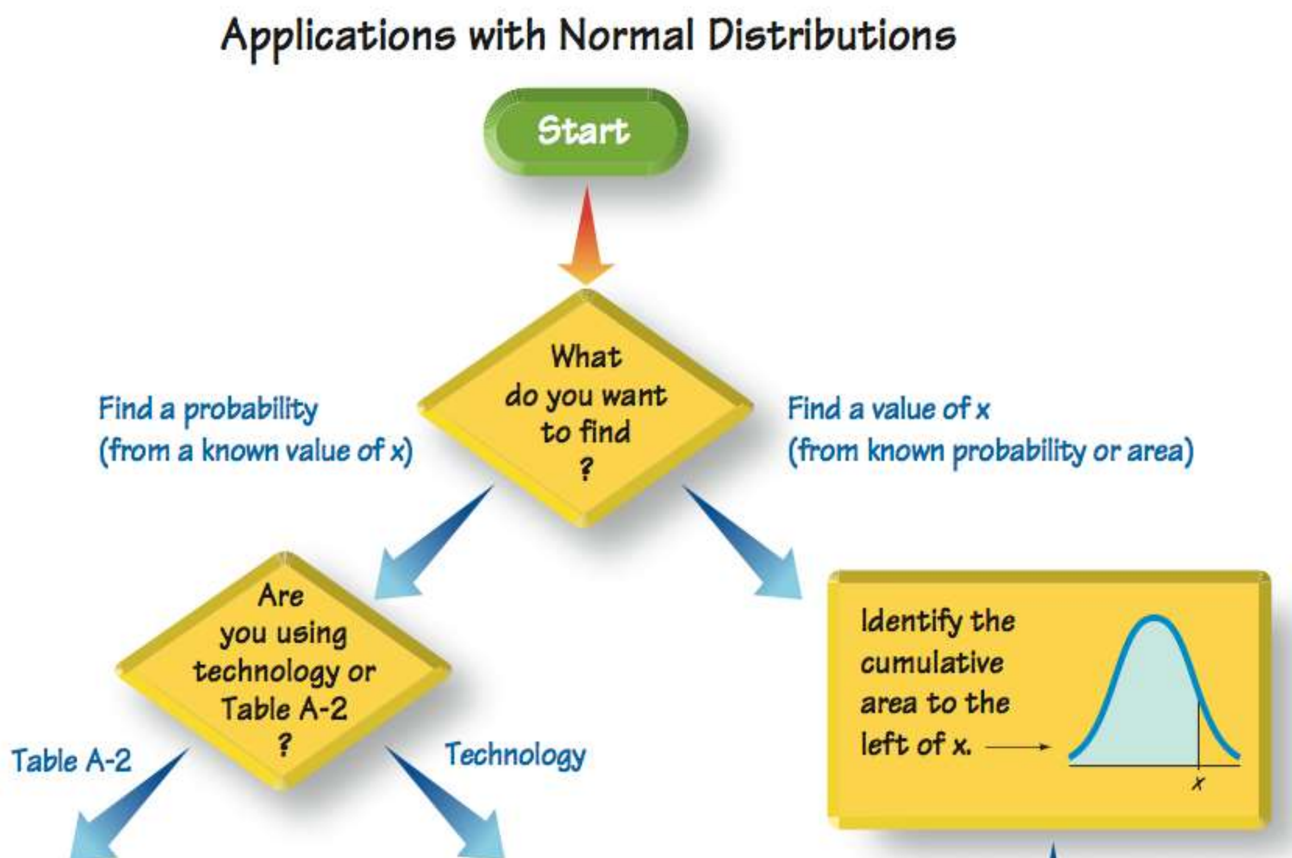


x = m + (z • q)
x = 172 + (2.575 . 29)
x = 246.675 (247 rounded)

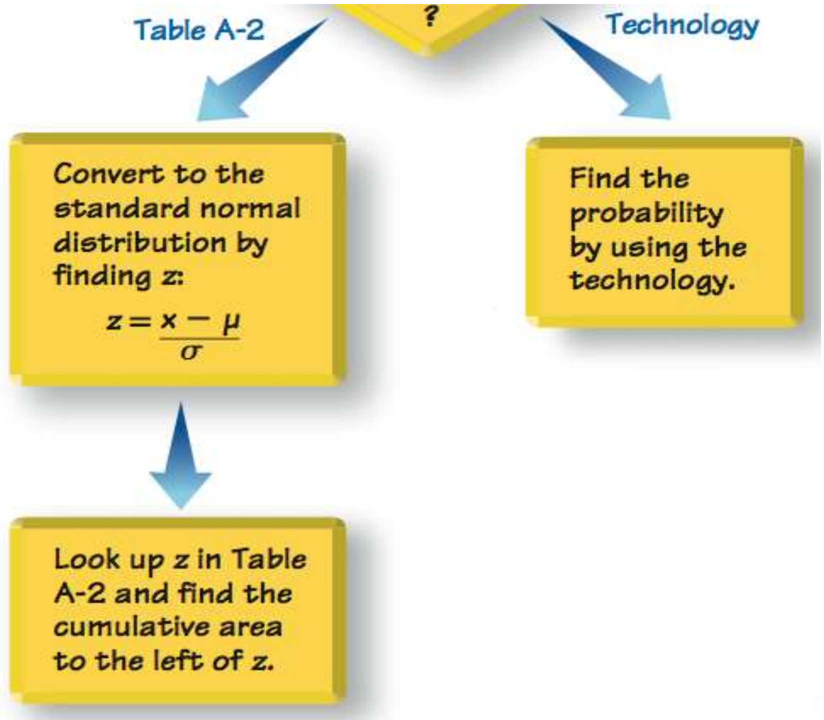
Example – Lightest and Heaviest – cont
مثال - الأخف وزنا - تابع

The weight of 247 pounds separates the lightest 99.5% from the heaviest 0.5%
وزن 247 رطل يفصل الأخف وزناً بنسبة 99.5٪ عن 0.5٪

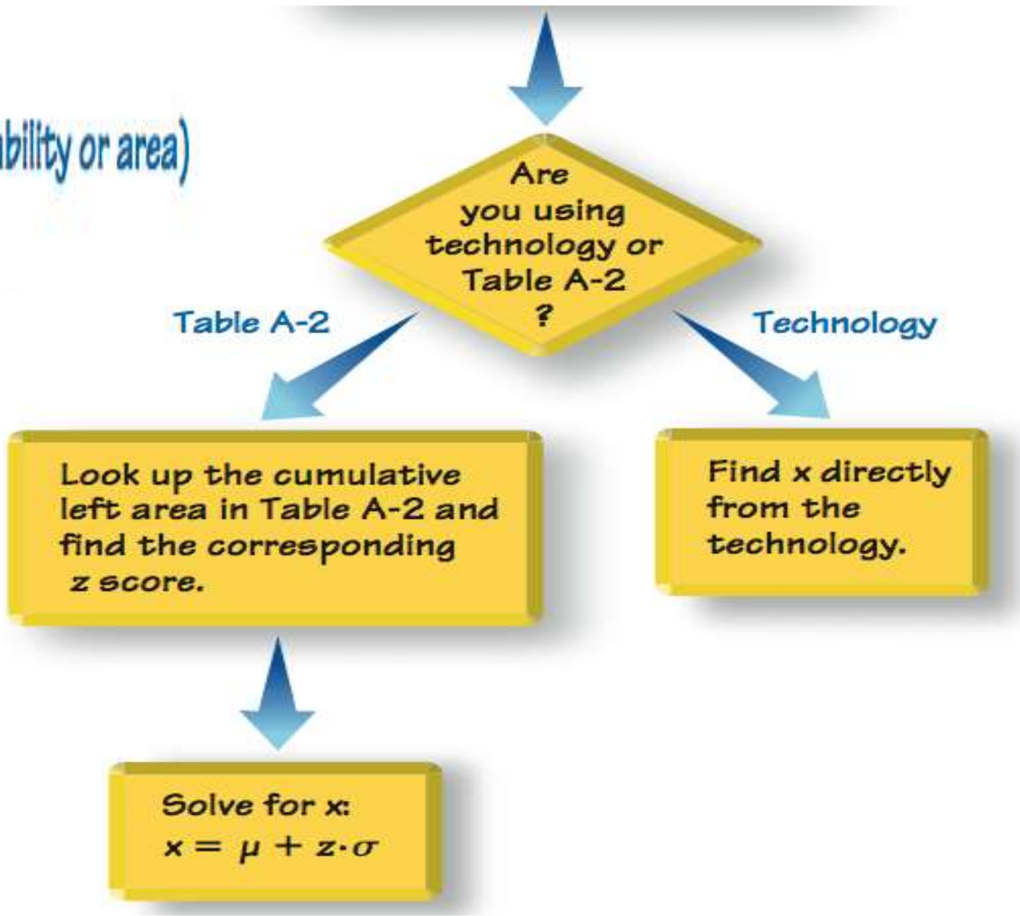




Find a probability (from a known value of x)



Find a value of x (from known probability or area)



Recap خلاصة

In this section we have discussed:

- ❖ **Non-standard normal distribution.**
- ❖ **Converting to a standard normal distribution.**
- ❖ **Procedures for finding values using Table A-2 and Formula 6-2.**

في هذا القسم ناقشنا:

توزيع عادي غير قياسي. تحويل إلى التوزيع الطبيعي القياسي.

والصيغة 6-2. A-2إجراءات البحث عن القيم باستخدام الجدول

Key Concept المفهوم الرئيسي

The main objective of this section is to understand the concept of a sampling distribution of a statistic, which is the distribution of all values of that statistic when all possible samples of the same size are taken from the same population.

We will also see that some statistics are better than others for estimating population parameters.

الهدف الرئيسي من هذا القسم هو فهم مفهوم توزيع أخذ العينات للإحصاء ، وهو توزيع جميع قيم تلك الإحصائية عندما تؤخذ جميع العينات الممكنة من نفس الحجم من نفس السكان.

سنرى أيضًا أن بعض الإحصاءات أفضل من غيرها لتقدير المعلمات السكانية.

Definition تعريف

The sampling distribution of a statistic (such as the sample mean or sample proportion) is the distribution of all values of the statistic when all possible samples of the same size *n* are taken from the same population. (The sampling distribution of a statistic is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

توزيع أخذ العينات للإحصاء (مثل متوسط العينة أو نسبة العينة) هو توزيع جميع قيم الإحصاء عندما تؤخذ جميع العينات الممكنة من نفس الحجم n من نفس المجموعة السكانية. (يتم تمثيل توزيع أخذ العينات للإحصاء عادةً كتوزيع احتمالي بتنسيق جدول أو رسم بياني لاحتمال أو صيغة.)

Definition تعريف

The sampling distribution of the mean is the distribution of sample means, with all samples having the same sample size n taken from the same population. (The sampling distribution of the mean is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

توزيع أخذ العينات للمتوسط هو توزيع وسيلة العينة ، مع كل العينات لها نفس حجم العينة n مأخوذة من نفس المجموعة. (يتم تمثيل توزيع العينة للمتوسط عادةً على أنه توزيع احتمالي بتنسيق جدول أو رسم بياني للاحتمال أو صيغة.)

Properties الخصائص

- **Sample means target the value of the population mean. (That is, the mean of the sample means is the population mean. The expected value of the sample mean is equal to the population mean.)**
- **The distribution of the sample means tends to be a normal distribution.**

العينة تعني استهداف قيمة متوسط السكان. (وهذا يعني أن متوسط عينة يعني هو متوسط عدد السكان. القيمة المتوقعة لمتوسط العينة مساوية لمتوسط عدد السكان).

توزيع العينة يعني تميل إلى أن يكون التوزيع الطبيعي.

Definition تعريف

The sampling distribution of the variance is the distribution of sample variances, with all samples having the same sample size n taken from the same population. (The sampling distribution of the variance is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

توزيع العينات من التباين هو توزيع تباينات العينة ، مع جميع العينات لها نفس nحجم العينة مأخوذة من نفس المجموعة. (يتم تمثيل توزيع العينات من التباين عادةً على أنه توزيع احتمالي بتنسيق جدول أو رسم بياني للاحتمال أو صيغة.)

Properties الخصائص

- **Sample variances target the value of the population variance. (That is, the mean of the sample variances is the population variance. The expected value of the sample variance is equal to the population variance.)**
- **The distribution of the sample variances tends to be a distribution skewed to the right.**

تستهدف الفروق عينة قيمة التباين السكان. (وهذا يعني أن متوسط تباينات العينة هو التباين السكاني. القيمة المتوقعة لتباين العينة تساوي التباين السكاني.)

يميل توزيع تباينات العينة إلى أن يكون توزيعًا منحرفًا إلى اليمين.

Definition تعريف

The sampling distribution of the proportion is the distribution of sample proportions, with all samples having the same sample size n taken from the same population.

توزيع أخذ العينات من النسبة هو توزيع نسب العينة ، مع كل العينات لها نفس حجم مأخوذة من نفس المجموعة. nالعينة

Definition تعريف

We need to distinguish between a population proportion p and some sample proportion:

وبعض نسبة العينة: pنحتاج إلى التمييز بين نسبة السكان

***p* = population proportion**

$\hat{p}$ **= sample proportion**

Properties الخصائص

- Sample proportions target the value of the population proportion. (That is, the mean of the sample proportions is the population proportion. The expected value of the sample proportion is equal to the population proportion.)

نسب العينة تستهدف قيمة نسبة السكان. (وهذا يعني أن نسبة نسب العينة هي نسبة السكان. القيمة المتوقعة لنسبة العينة تساوي نسبة السكان.)

- The distribution of the sample proportion tends to be a normal distribution.

توزيع نسبة العينة يميل إلى أن يكون التوزيع الطبيعي.

Unbiased Estimators مقدرات غير متحيزة

Sample means, variances and proportions are unbiased estimators.
That is they target the population parameter.
These statistics are better in estimating the population parameter.

عينة يعني ، الفروق والنسب هي مقدرات غير متحيزة.
هذا هو أنها تستهدف المعلمة السكان.
هذه الإحصاءات هي أفضل في تقدير المعلمة السكان.

Biased Estimators مقدرون متحيزون

Sample medians, ranges and standard deviations are biased estimators.
That is they do NOT target the population parameter.
Note: the bias with the standard deviation is relatively small in large samples so s is often used to estimate.

عينة المتوسطات ، النطاقات والانحرافات المعيارية هي مقدرات متحيزة.
هذا هو أنها لا تستهدف المعلمة السكان.
ملاحظة: يكون التحيز مع الانحراف المعياري صغيرًا نسبيًا في العينات الكبيرة ، لذا غالبًا ما يتم استخدام s لتقدير.

Example - Sampling Distributions

مثال - توزيعات أخذ العينات

Consider repeating this process: Roll a die 5 times, find the mean , variance s^2, and the proportion of *odd* numbers of the results. What do we know about the behavior of all sample means X that are generated as this process continues indefinitely?

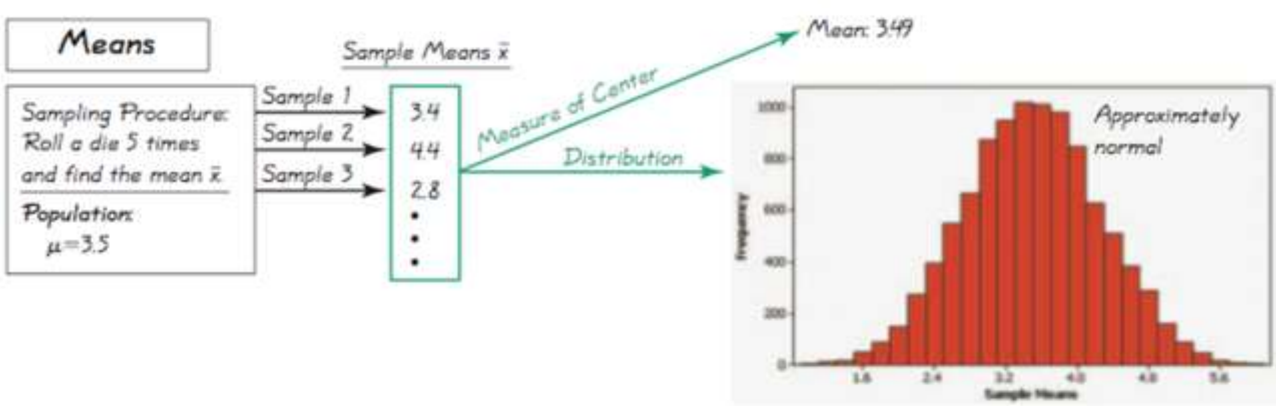
، s2فكر في تكرار هذه العملية: مرر خمس مرات ، ابحث عن المتوسط ، التباين ونسبة الأعداد الفردية للنتائج. ما الذي نعرفه عن سلوك كل عينة يعني أنه يتم توليدها مع استمرار هذه العملية إلى أجل غير مسمى؟

Example - Sampling Distributions

مثال - توزيعات أخذ العينات

Specific results from 10,000 trials

نتائج محددة من 10000 تجربة

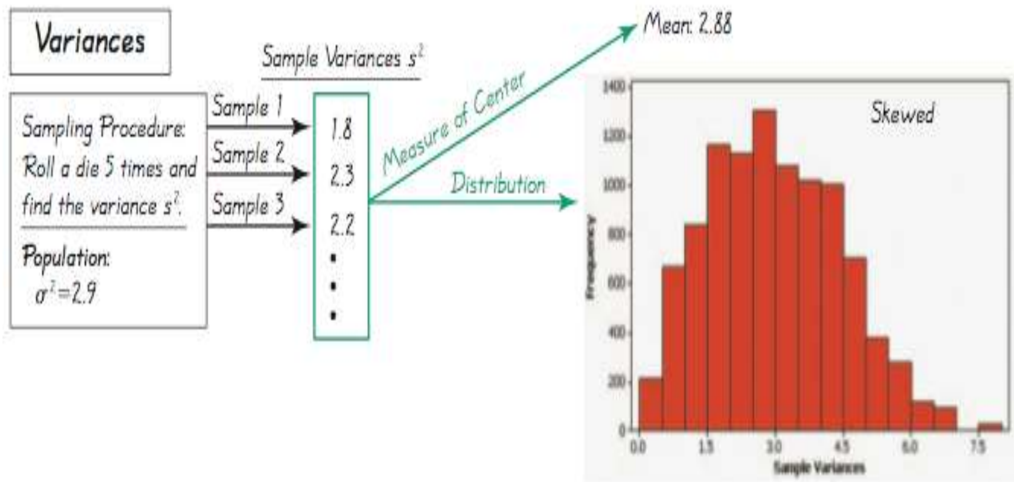


All outcomes are equally likely so the population mean is 3.5; the mean of the 10,000 trials is 3.49. If continued indefinitely, the sample mean will be 3.5. Also, notice the distribution is "normal."

جميع النتائج محتملة على قدم المساواة وبالتالي فإن متوسط عدد السكان هو 3.5 ؛ متوسط 10000 تجربة هو 3.49. إذا استمرت إلى أجل غير مسمى ، فإن متوسط العينة سيكون 3.5. لاحظ أيضًا أن التوزيع "طبيعي".

Example - Sampling Distributions
مثال - توزيعات أخذ العينات

Specific results from 10,000 trials
نتائج محددة من 10000 تجربة

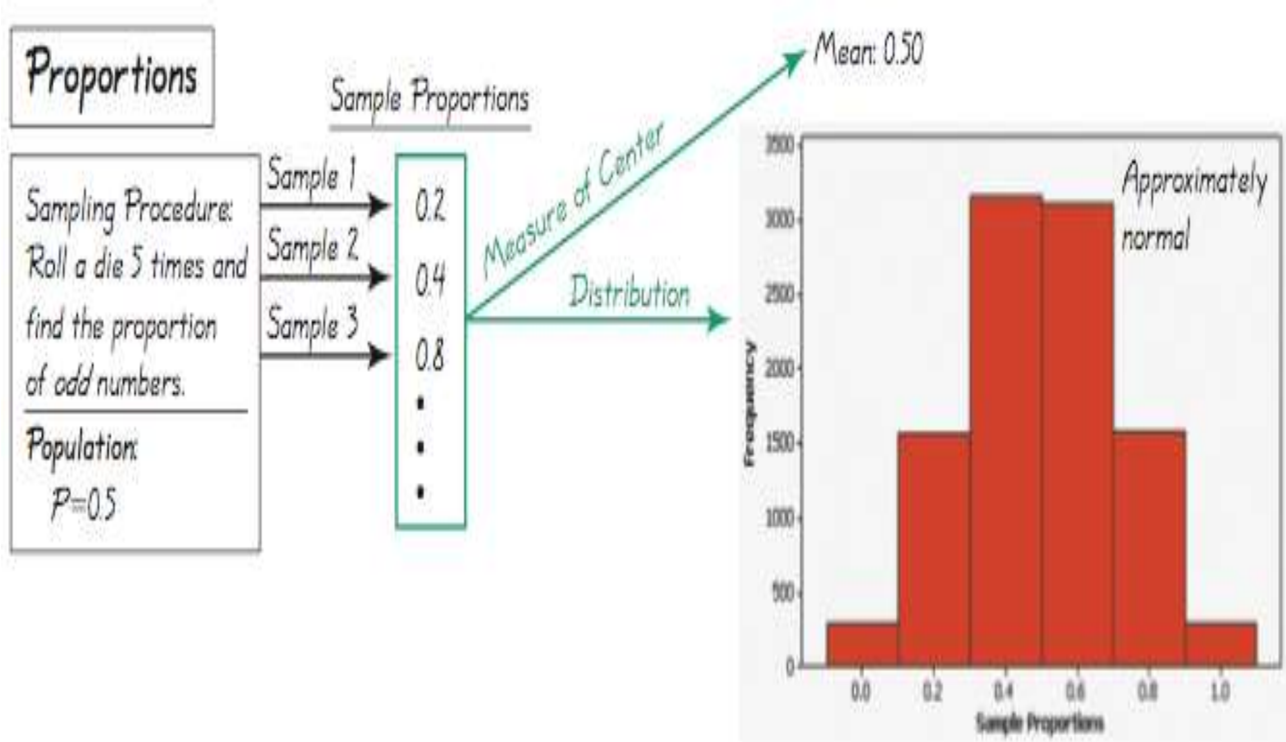


All outcomes are equally likely so the population variance is 2.9; the mean of the 10,000 trials is 2.88. If continued indefinitely, the sample variance will be 2.9. Also, notice the distribution is "skewed to the right."

جميع النتائج محتملة على قدم المساواة لذلك التباين السكاني هو 2.9 ؛ متوسط 10000 تجربة هو 2.88. إذا استمرت إلى أجل غير مسمى ، فإن تباين العينة سيكون 2.9. لاحظ أيضًا أن التوزيع "منحرف إلى اليمين".

Example - Sampling Distributions
مثال - توزيعات أخذ العينات

Specific results from 10,000 trials
نتائج محددة من 10000 تجربة



All outcomes are equally likely so the population proportion of odd numbers is 0.50; the proportion of the 10,000 trials is 0.50. If continued indefinitely, the mean of sample proportions will be 0.50. Also, notice the distribution is "approximately normal."

جميع النتائج محتملة على قدم المساواة ، وبالتالي فإن نسبة السكان من الأرقام الفردية هي 0.50 ؛ نسبة 10000 تجربة هي 0.50. إذا استمرت إلى أجل غير مسمى ، فإن متوسط نسب العينة سيكون 0.50. لاحظ أيضًا أن التوزيع "طبيعي تقريبًا".

Why Sample with Replacement?
لماذا عينة مع استبدال؟

Sampling *without replacement* would have the very practical advantage of avoiding wasteful duplication whenever the same item is selected more than once.
However, we are interested in sampling *with replacement* for these two reasons:

سيكون لأخذ العينات دون الاستبدال ميزة عملية للغاية تتمثل في تجنب الازدواجية المهدرة كلما تم تحديد العنصر نفسه أكثر من مرة.
ومع ذلك ، نحن مهتمون بأخذ العينات مع الاستبدال لهذين السببين:

1. When selecting a relatively small sample form a large population, it makes no significant difference whether we sample with replacement or without replacement.

عند اختيار عينة صغيرة نسبيًا تشكل عددًا كبيرًا من السكان ، لن يحدث فرق كبير سواء أخذنا عينة بديلة أو بدون بديل.

2. Sampling with replacement results in independent events that are unaffected by previous outcomes, and independent events are easier to analyze and result in simpler calculations and formulas.

يؤدي أخذ العينات مع الاستبدال إلى أحداث مستقلة لا تتأثر بالنتائج السابقة ، وتكون الأحداث المستقلة أسهل في التحليل وتؤدي إلى عمليات حساب وصيغ أبسط.

Cautionالحذر

Many methods of statistics require a simple random sample. Some samples, such as voluntary response samples or convenience samples, could easily result in very wrong results.

تتطلب العديد من طرق الإحصاء عينة عشوائية بسيطة. بعض العينات ، مثل عينات الاستجابة الطوعية أو عينات الراحة ، يمكن أن تؤدي بسهولة إلى نتائج خاطئة.

Recap خلاصة

In this section we have discussed:

- **Sampling distribution of a statistic.**
- **Sampling distribution of the mean.**
- **Sampling distribution of the variance.**
- **Sampling distribution of the proportion.**
- **Estimators.**

في هذا القسم ناقشنا:

توزيع العينات من الإحصائية.

توزيع العينات من الوسط.

توزيع العينات من الفرق.

توزيع العينات من النسبة.

المقدرات.

Key Concept المفهوم الرئيسي

The Central Limit Theorem tells us that for a population with any distribution, the distribution of the sample means approaches a normal distribution as the sample size increases.

The procedure in this section form the foundation for estimating population parameters and hypothesis testing.

تخبرنا نظرية الحد المركزي بأن توزيع العينة يعني بالنسبة للسكان الذين لديهم أي توزيع أن التوزيع يقترب من التوزيع الطبيعي كلما زاد حجم العينة.

يشكل الإجراء في هذا القسم الأساس لتقدير المعلمات السكانية واختبار الفرضيات.

Central Limit Theorem
نظرية الحد المركزي

Given:معطى

1. The random variable *x* has a distribution (which may or may not be normal) with mean *μ* and standard deviation *σ*.
2. Simple random samples all of size *n* are selected from the population. (The samples are selected so that all possible samples of the same size *n* have the same chance of being selected.)

1. المتغير العشوائي اكس له توزيع (والذي قد يكون أو لا يكون طبيعياً) بمتوسط ميو والانحراف المعياري كيو .

2. يتم اختيار عينات عشوائية بسيطة من كل حجم ن من السكان. (يتم اختيار العينات بحيث نفس فرصة التحديد.) nيكون لكل العينات الممكنة من نفس الحجم

Central Limit Theorem – cont.تابع - نظرية الحد المركزي

Conclusions:الاستنتاجات

1. The distribution of sample x will, as the sample size increases, approach a normal distribution.
2. The mean of the sample means is the population mean µ.

1. توزيع العينة، اكس مع زيادة حجم العينة ، يقترب من التوزيع الطبيعي.

2. يعني متوسط العينة هو متوسط عدد السكان ميو

3. The standard deviation of all sample means is
3. الانحراف المعياري لجميع وسائل العينة هو

$$\sigma/\sqrt{n}$$

Practical Rules Commonly Used
القواعد العملية شائعة الاستخدام

1. **For samples of size *n* larger than 30, the distribution of the sample means can be approximated reasonably well by a normal distribution. The approximation gets closer to a normal distribution as the sample size *n* becomes larger.**

أكبر من 30 ، يمكن تقريب توزيع العينة جيدًا n.1 بالنسبة للعينات ذات الحجم بشكل معقول عن طريق التوزيع الطبيعي. يقترب التقريب من التوزيع الطبيعي أكبر. nعندما يصبح حجم العينة

2. **If the original population is *normally distributed*, then for any sample size *n*, the sample means will be normally distributed (not just the values of *n* larger than 30).**

2. إذا تم توزيع السكان الأصليين بشكل طبيعي ، فسيتم توزيع وسيلة العينة عادةً أكبر من 30). nوليس فقط قيم) nلأي حجم عينة

Notation الرموز

the mean of the sample means

يعني متوسط العينة

$$\mu_{\bar{x}} = \mu$$

the standard deviation of sample mean

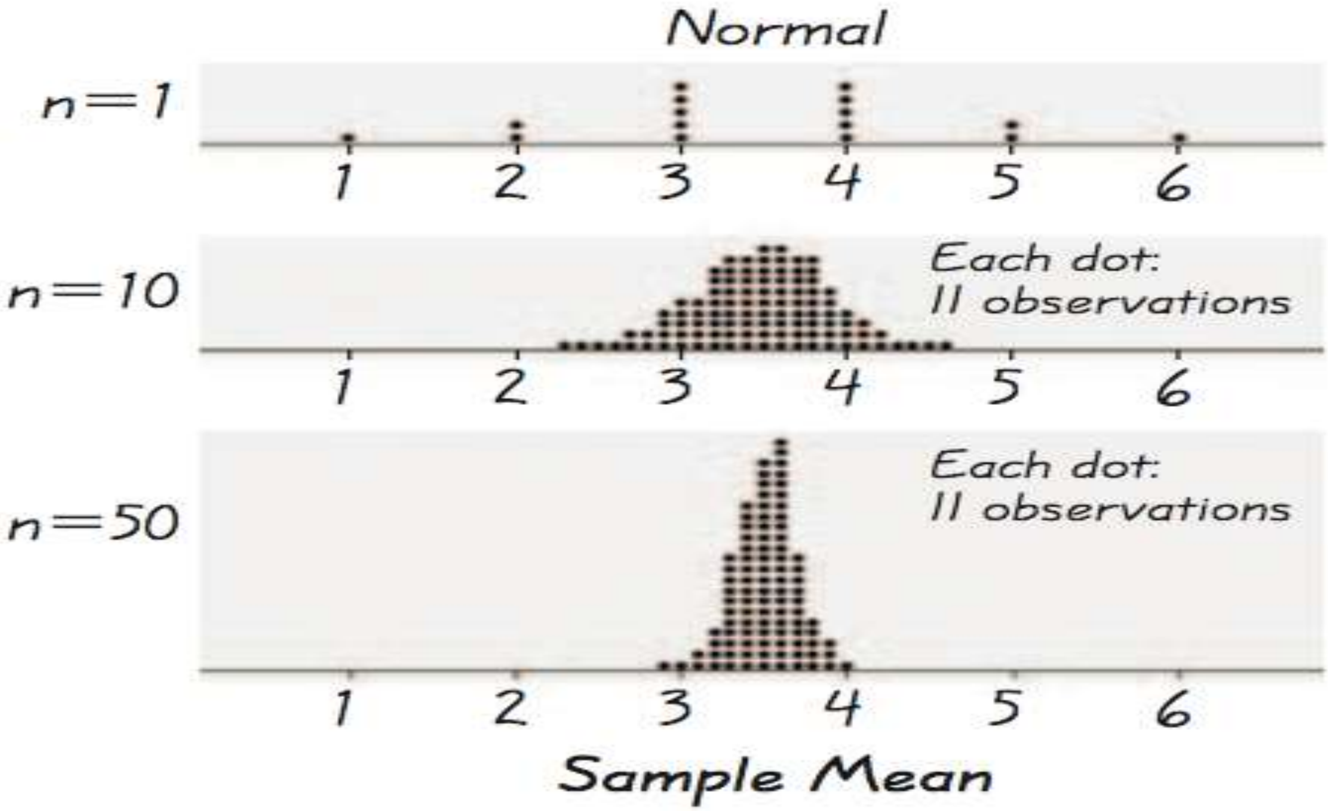
الانحراف المعياري للعينة المتوسطة

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

(often called the standard error of the mean)

(غالبًا ما تسمى الخطأ القياسي للمتوسط)

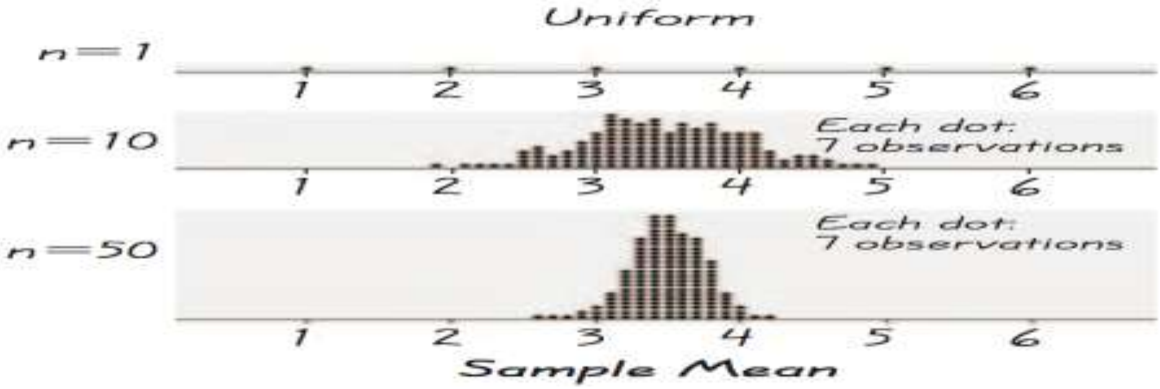
Example - Normal Distribution
مثال - التوزيع الطبيعي



As we proceed from *n* = 1 to *n* = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.

أثناء تقدمنا من n = 1 إلى n = 50 ، نرى أن توزيع العينة يعني اقتراب شكل التوزيع الطبيعي.

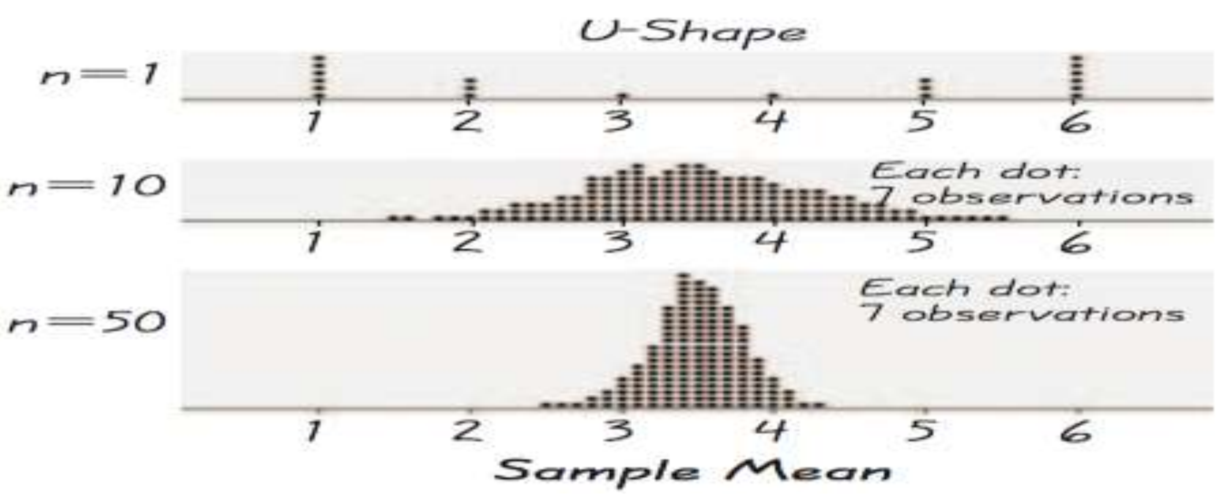
Example - Uniform Distribution مثال - توزيع موحد



As we proceed from *n* = 1 to *n* = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.

، نرى أن توزيع العينة يعني اقتراب شكل التوزيع 50 = nإلى 1 = nأثناء تقدمنا من الطبيعي.

Example - U-Shaped Distribution Uمثال - توزيع على شكل حرف



As we proceed from n = 1 to n = 50, we see that the distribution of sample means is approaching the shape of a normal distribution.

، نرى أن توزيع العينة يعني اقتراب شكل التوزيع 50 = nإلى 1 = nأثناء تقدمنا من الطبيعي.

Important Point نقطة مهمة

As the sample size increases, the sampling distribution of sample means approaches a normal distribution.

كلما زاد حجم العينة ، فإن توزيع العينة يعني اقتراب التوزيع الطبيعي.

Example – Water Taxi Safetyمثال - سلامة تاكسي المياه

Use the Chapter Problem. Assume the population of weights of men is normally distributed with a mean of 172 lb and a standard deviation of 29 lb.

استخدم مشكلة الفصل. افترض أن عدد أوزان الرجال يتم توزيعه عادة بمتوسط 172 رطل وانحراف معياري قدره 29 رطلاً.

a) Find the probability that if an individual man is randomly selected, his weight is greater than 175 lb.
b) Find the probability that 20 randomly selected men will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).

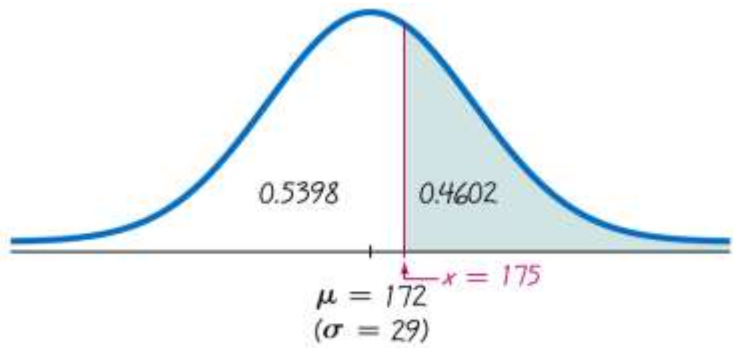
أ) ابحث عن احتمال أنه إذا تم اختيار رجل فردي بشكل عشوائي ، يكون وزنه أكبر من 175 رطلاً.
ب) أوجد احتمال أن يكون لدى 20 رجلاً تم اختيارهم عشوائياً متوسط وزن أكبر من 175 رطلاً (بحيث يتجاوز وزنهم الإجمالي القدرة الآمنة 3500 رطل).

Example – cont مثال - تابع

a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

أ) ابحث عن احتمال أنه إذا تم اختيار رجل فردي بشكل عشوائي ، يكون وزنه أكبر من 175 رطلاً.

$$z = \frac{175 - 172}{29} = 0.10$$



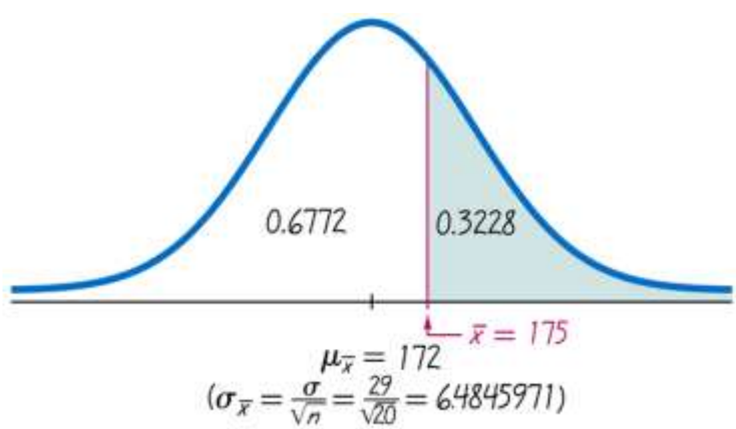
(a)

Example – cont مثال - تابع

b) Find the probability that 20 randomly selected men will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).

ب) أوجد احتمال أن يكون لدى 20 رجلاً تم اختيارهم عشوائياً متوسط وزن أكبر من 175 رطلاً (بحيث يتجاوز وزنهم الإجمالي القدرة الآمنة 3500 رطل).

$$z = \frac{175 - 172}{\frac{29}{\sqrt{20}}} = 0.46$$



(b)

Example – cont مثال - تابع

a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

أ) ابحث عن احتمال أنه إذا تم اختيار رجل فردي بشكل عشوائي ، يكون وزنه أكبر من 175 رطلاً.

$$P(x > 175) = 0.4602$$

b) Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).

ب) أوجد احتمال أن يكون لدى 20 رجلاً تم اختيارهم عشوائياً متوسط وزن أكبر من 175 رطلاً (بحيث يتجاوز وزنهم الإجمالي القدرة الآمنة 3500 رطل).

$$P(\overline{x} > 175) = 0.3228$$

It is much easier for an individual to deviate from the mean than it is for a group of 20 to deviate from the mean.

من الأسهل على الفرد أن يحيد عن
يعني مما كانت عليه لمجموعة من 20 للانحراف عن الوسط.

Interpretation of Results تفسير النتائج

Given that the safe capacity of the water taxi is 3500 pounds, there is a fairly good chance (with probability 0.3228) that it will be overloaded with 20 randomly selected men.

بالنظر إلى أن السعة الآمنة لسيارات الأجرة المائية تبلغ 3500 رطل ، فهناك فرصة جيدة إلى حد ما (مع احتمال 0.3228) بأن يتم تحميلها بشكل زائد مع 20 رجلاً تم اختيارهم بشكل عشوائي.

Correction for a Finite Population
تصحيح لعدد محدود من السكان

When sampling without replacement and the sample size *n* is greater than 5% of the finite population of size *N* (that is, *n* > 0.05*N*), adjust the standard deviation of sample means by multiplying it by the *finite population correction factor*:

أكبر من 5٪ من السكان n عندما تكون العينة دون الاستبدال وحجم العينة
، اضبط الانحراف المعياري (N أي ، <> 0.05) N المحددين من الحجم
للعينة عن طريق ضربها بعامل تصحيح السكان المحدود:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \underbrace{\sqrt{\frac{N-n}{N-1}}}$$

finite population correction factor
عدد محدود من السكان
معامل التصحيح

Recap خلاصة

In this section we have discussed:

- **Central limit theorem.**
- **Practical rules.**
- **Effects of sample sizes.**
- **Correction for a finite population.**

في هذا القسم ناقشنا:

نظرية الحد المركزي.

القواعد العملية.

آثار أحجام العينات.

تصحيح لعدد محدود من السكان.

Key Concept المفهوم الرئيسي

This section presents criteria for determining whether the requirement of a normal distribution is satisfied. The criteria involve visual inspection of a histogram to see if it is roughly bell shaped, identifying any outliers, and constructing a graph called a normal quantile plot.

يقدم هذا القسم معايير لتحديد ما إذا كانت متطلبات التوزيع الطبيعي مستوفاة أم لا. تتضمن المعايير الفحص البصري للرسم البياني لمعرفة ما إذا كان شكل الجرس تقريبًا ، وتحديد أي القيم المتطرفة ، وإنشاء رسم بياني يُدعى مؤامرة كمية طبيعية.

Definition تعريف

A normal quantile plot (or normal probability plot) is a graph of points (*x*,*y*), where each *x* value is from the original set of sample data, and each *y* value is the corresponding *z* score that is a quantile value expected from the standard normal distribution.

المؤامرة الكمية المعتادة (أو المؤامرة الاحتمالية الطبيعية) هي رسم بياني للنقاط (س ، ص) ، حيث تكون كل قيمة س من المجموعة الأصلية لبيانات العينة ، وكل قيمة ص هي النتيجة المقابلة التي هي قيمة كمية متوقعة من التوزيع الطبيعي القياسي. z

Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

إجراء لتحديد ما إذا كان من المعقول افتراض أن بيانات العينة من السكان الموزعين بشكل طبيعي

1. Histogram: Construct a histogram. Reject normality if the histogram departs dramatically from a bell shape.

1. المدرج التكراري: بناء المدرج التكراري. رفض الحالة الطبيعية إذا كان الرسم البياني يغادر بشكل كبير من شكل الجرس.

2. Outliers: Identify outliers. Reject normality if there is more than one outlier present.

2. القيم المتطرفة: تحديد القيم المتطرفة. رفض الحالة الطبيعية إذا كان هناك أكثر من واحد الحاضر الحاضر .

3. Normal Quantile Plot: If the histogram is basically symmetric and there is at most one outlier, use technology to generate a normal quantile plot.

3. مؤامرة كمية طبيعية: إذا كان الرسم البياني متماثلًا في الأساس وكان هناك حد أقصى واحد ، استخدم التكنولوجيا لإنشاء مؤامرة كمية طبيعية.

Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

إجراء لتحديد ما إذا كان من المعقول افتراض أن بيانات العينة من السكان الموزعين بشكل طبيعي

3. Continued

Use the following criteria to determine whether or not the distribution is normal.

Normal Distribution: The population distribution is normal if the pattern of the points is reasonably close to a straight line and the points do not show some systematic pattern that is not a straight-line pattern.

3. تابع

استخدم المعايير التالية لتحديد ما إذا كان التوزيع طبيعيًا أم لا.

التوزيع الطبيعي: يكون التوزيع السكاني طبيعيًا إذا كان نمط النقاط قريبًا بشكل معقول من خط مستقيم ولا تظهر النقاط بعض النمط المنهجي الذي لا يمثل نمط خط مستقيم.

Procedure for Determining Whether It Is Reasonable to Assume that Sample Data are From a Normally Distributed Population

إجراء لتحديد ما إذا كان من المعقول افتراض أن بيانات العينة من السكان الموزعين بشكل طبيعي

3. Continued

Not a Normal Distribution: The population distribution is not normal if either or both of these two conditions applies:

The points do not lie reasonably close to a straight line.

The points show some systematic pattern that is not a straight-line pattern.

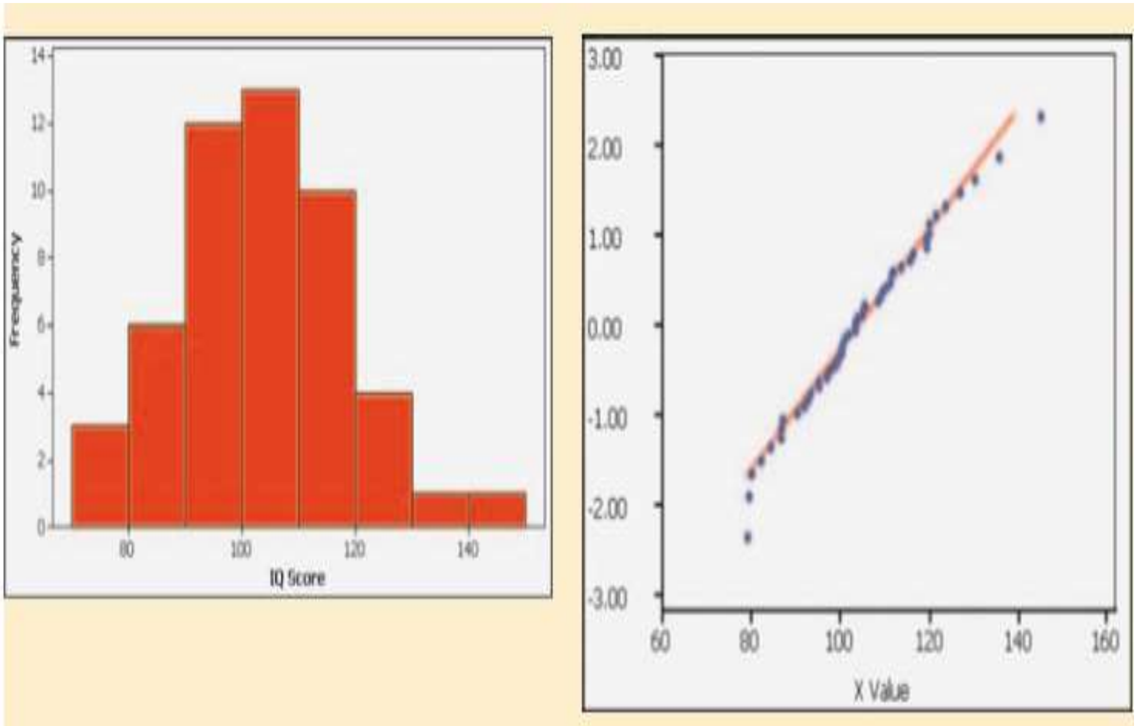
3. تابع

ليس التوزيع الطبيعي: توزيع السكان غير طبيعي إذا كان أي من هذين الشرطين أو كليهما ينطبق:

النقاط لا تقع بالقرب من خط مستقيم.

تُظهر النقاط نمطًا منهجيًا ليس نمط خط مستقيم.

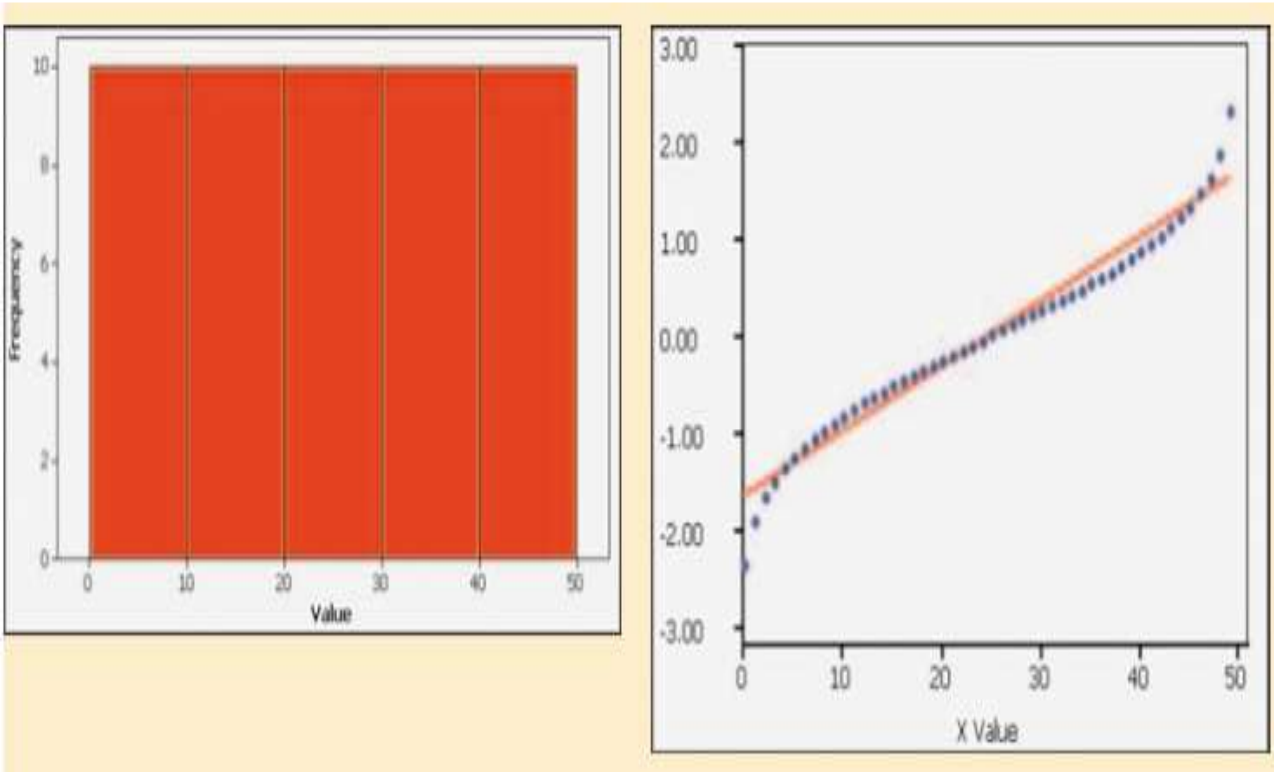
Example مثال



Normal: Histogram of IQ scores is close to being bell-shaped, suggests that the IQ scores are from a normal distribution. The normal quantile plot shows points that are reasonably close to a straight-line pattern. It is safe to assume that these IQ scores are from a normally distributed population.

عادي: الرسم البياني لدرجات الذكاء قريب من كونه على شكل جرس ، يشير إلى أن درجات معدل الذكاء هي من التوزيع الطبيعي. تظهر المؤامرة الكمية العادية نقاط قريبة بشكل معقول من نمط خط مستقيم. من الآمن افتراض أن درجات حاصل الذكاء هذه هي من مجموعة موزعة بشكل طبيعي.

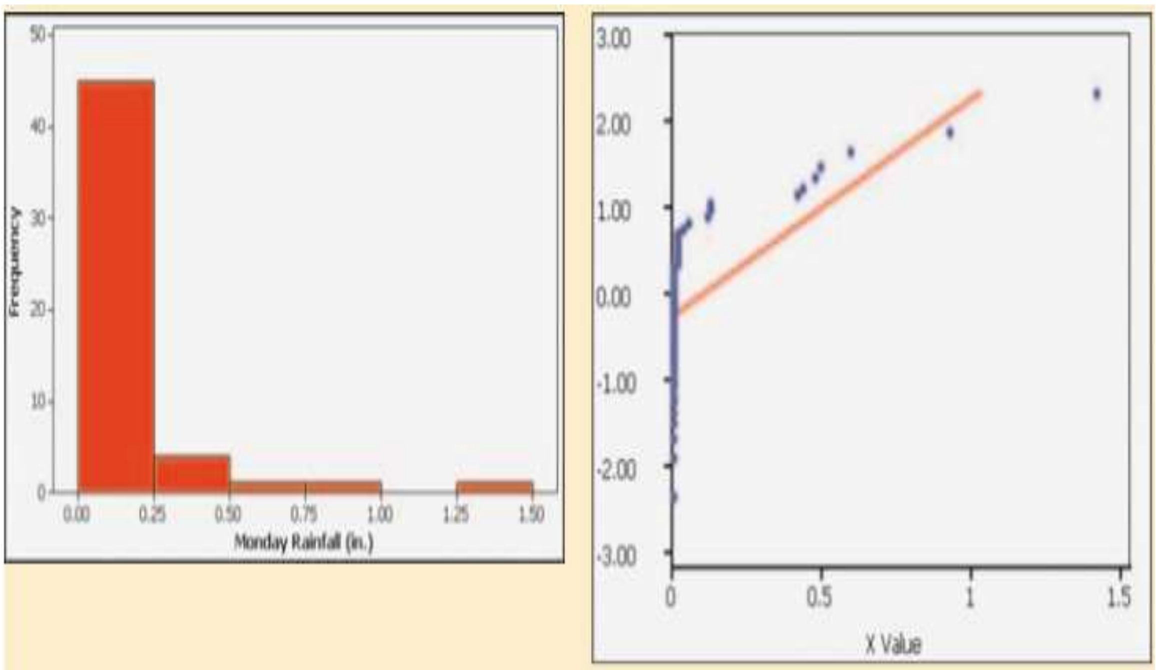
Example مثال



Uniform: **Histogram of data having a uniform distribution. The corresponding normal quantile plot suggests that the points are not normally distributed because the points show a systematic pattern that is not a straight-line pattern. These sample values are not from a population having a normal distribution.**

منتظم: رسم بياني للبيانات ذات توزيع موحد. تشير المؤامرة الكمية الطبيعية المقابلة إلى أن النقاط لا يتم توزيعها بشكل طبيعي لأن النقاط تظهر نمطًا منهجيًا ليس نمط خط مستقيم. هذه القيم عينة ليست من السكان الذين لديهم التوزيع الطبيعي.

Example مثال



Skewed: Histogram of the amounts of rainfall in Boston for every Monday during one year. The shape of the histogram is skewed, not bell-shaped. The corresponding normal quantile plot shows points that are not at all close to a straight-line pattern. These rainfall amounts are not from a population having a normal distribution.

منحرف: رسم بياني لكميات هطول الأمطار في بوسطن عن كل يوم اثنين خلال عام واحد. شكل الرسم البياني منحرف ، وليس على شكل جرس. تظهر المؤامرة الكمية الطبيعية المقابلة النقاط التي ليست قريبة على الإطلاق من نمط خط مستقيم. كميات الأمطار هذه ليست من السكان الذين لديهم توزيع طبيعي.

Manual Construction of a Normal Quantile Plot

البناء اليدوي لمؤامرة كمية عادية

Step 1. First sort the data by arranging the values in order from lowest to highest.

Step 2. With a sample of size *n*, each value represents a proportion of 1/*n* of the sample. Using the known sample size *n*, identify the areas of 1/2*n*, 3/2*n*, and so on. These are the cumulative areas to the left of the corresponding sample values.

Step 3. Use the standard normal distribution (Table A-2 or software or a calculator) to find the *z* scores corresponding to the cumulative left areas found in Step 2. (These are the *z* scores that are expected from a normally distributed sample.)

الخطوة 1. أولاً فرز البيانات عن طريق ترتيب القيم بالترتيب من الأدنى إلى الأعلى.

الخطوة 2. مع عينة من حجم ن ، تمثل كل قيمة نسبة 1 / ن من العينة. باستخدام ، n و3 / 2 n، حدد المساحات من 1 / 2 nحجم العينة المعروف وما إلى ذلك. هذه هي المناطق التراكمية على يسار القيم عينة المقابلة.

الخطوة 3. استخدم التوزيع العادي القياسي (الجدول تايبل ا أو البرنامج أو الحاسبة) المقابلة للمناطق اليسرى التراكمية zللعثور على الدرجات المتوقعة من عينة zالموجودة في الخطوة 2. (هذه هي الدرجات موزعة بشكل طبيعي.)

Manual Construction of a Normal Quantile Plot
البناء اليدوي لمؤامرة كمية عادية

Step 4. Match the original sorted data values with their corresponding *z* scores found in Step 3, then plot the points (*x*, *y*), where each *x* is an original sample value and *y* is the corresponding *z* score.

الخطوة 4. قم بمطابقة قيم البيانات الأصلية التي تم فرزها مع النتائج z المقابلة لها الموجودة في الخطوة 3 ، ثم ارسم النقاط (x، y) ، حيث أن كل x هي قيمة عينة أصلية و y هي النتيجة z المقابلة.

Step 5. Examine the normal quantile plot and determine whether or not the distribution is normal.

الخطوة 5. فحص المؤامرة الكمية المعتادة وتحديد ما إذا كان التوزيع طبيعيًا أم لا.

Ryan-Joiner Test اختبار ريان النجار

The Ryan-Joiner test is one of several formal tests of normality, each having their own advantages and disadvantages. STATDISK has a feature of Normality Assessment that displays a histogram, normal quantile plot, the number of potential outliers, and results from the Ryan-Joiner test. Information about the Ryan-Joiner test is readily available on the Internet.

يعد اختبار Ryan-Joiner أحد الاختبارات الرسمية العديدة للحالة الطبيعية ، ولكل منها مزاياها وعيوبها. يتميز موقع STATDISK بميزة تقييم الحالة التي تعرض رسم بياني ، ومؤامرة كمية طبيعية ، وعدد القيم المتطرفة المحتملة ، ونتائج اختبار ريان جوينر. معلومات حول اختبار Ryan-Joiner متاحة بسهولة على الإنترنت.

Data Transformations تحويلات البيانات

Many data sets have a distribution that is not normal, but we can transform the data so that the modified values have a normal distribution. One common transformation is to replace each value of x with log (*x* + 1). If the distribution of the log (*x* + 1) values is a normal distribution, the distribution of the x values is referred to as a lognormal distribution.

تحتوي العديد من مجموعات البيانات على توزيع غير طبيعي ، لكن يمكننا تحويل البيانات بحيث يكون للقيم المعدلة توزيع عادي. أحد التحويلات الشائعة هي استبدال كل قيمة x (بالسجل x + 1). إذا كان توزيع القيم log (x + 1) ، توزيعًا طبيعيًا ، فسيتم الإشارة إلى توزيع القيم x كتوزيع غير طبيعي.

Other Data Transformations تحويلات البيانات الأخرى

In addition to replacing each x value with the log (x + 1), there are other transformations, such as replacing each x value with $\sqrt{x}$, or 1/x, or x2. In addition to getting a required normal distribution when the original data values are not normally distributed, such transformations can be used to correct other deficiencies, such as a requirement (found in later chapters) that different data sets have the same variance.

، هناك تحويلات أخرى ، مثل (x + 1 بالسجل) x بالإضافة إلى استبدال كل قيمة بالإضافة إلى الحصول على توزيع طبيعي .x2، أو x بـ ، أو 1 / x استبدال كل قيمة مطلوب عندما لا يتم توزيع قيم البيانات الأصلية بشكل طبيعي ، يمكن استخدام هذه التحويلات لتصحيح أوجه القصور الأخرى ، مثل المتطلب (الموجود في الفصول اللاحقة) بأن مجموعات البيانات المختلفة لها نفس التباين.

Recap خلاصة

In this section we have discussed:

- **Normal quantile plot.**
- **Procedure to determine if data have a normal distribution.**

في هذا القسم ناقشنا:

مؤامرة كمية طبيعية.

الإجراء لتحديد ما إذا كانت البيانات لها توزيع عادي.