Solutions

1. What is the z-score for a trip that takes 13.8 min?

A −3.6

B −1.5

C 1.5

D 3.6

C

Roberta rides her bicycle to school. She records her times for one week, and determines a mean of 10.2 min and a standard deviation of 2.4 min. She believes the time follows a normal distribution.

$z = ?$

$x = 13.8$

$\bar{x} = 10.2$

$s = 2.4$

$$z = \frac{x - \bar{x}}{s}$$

$z = (13.8 - 10.2) / 2.4$

$z = 1.5$

2. A trip has a z-score of -0.5. How long did the trip take?

- A 9.0 min
- B 9.7 min
- C 10.7 min
- D 11.4 min

A

Roberta rides her bicycle to school. She records her times for one week, and determines a mean of 10.2 min and a standard deviation of 2.4 min. She believes the time follows a normal distribution.

$z = -0.5$

$x = ?$

$\overline{x} = 10.2$

$s = 2.4$

$$z = \frac{x - \overline{x}}{s}$$

$-0.5 = (x - 10.2) / 2.4$

$-0.5(2.4) = x - 10.2$

$-1.2 = x - 10.2$

$9 = x$

3. What is the probability that a trip will take twice as long as the mean?

Roberta rides her bicycle to school. She records her times for one week, and determines a mean of 10.2 min and a standard deviation of 2.4 min. She believes the time follows a normal distribution.

The mean = 10.2 minutes

A trip that is twice as long as the mean = 2(10.2) = 20.4 minutes.

Recall that the probability of a specific value is ZERO (the area under the graph).

→ **The probability of the trip taking 20.4 minutes is zero.**

Roberta rides her bicycle to school. She records her times for one week, and determines a mean of 10.2 min and a standard deviation of 2.4 min. She believes the time follows a normal distribution.

4. Communication Roberta makes similar measurements the following week, and calculates a different mean and standard deviation. Why did this happen? How can she obtain more reliable values for the mean and standard deviation?

The data collected could be different for a number of reasons:

- Traffic
- Different route
- Weather conditions

To obtain more reliable values, Roberta should combine the data recorded from the two weeks. Like with probability, the more trials (or in this case, data) the more reliable the results.

5. The corner of Bloor and Bathurst streets in Toronto is the home of Honest Ed's, a huge discount store famous all over the world. The sign contains about 23 000 light bulbs.

The filament of an incandescent light bulb slowly evaporates, limiting the life of the bulb. Turning the bulb on and off also shortens its lifetime. Technicians tested a sample of 500 bulbs. The table shows the frequencies for the lifetime of the bulbs.

a) Sketch a frequency histogram and a frequency polygon for these data.

b) Estimate the mean life of these light bulbs.

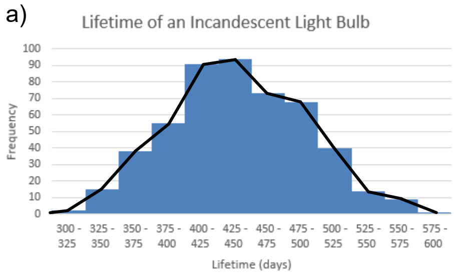
c) Add a relative frequency column to the table.

d) What is the probability that a given light bulb will fail in 400 days or fewer?

e) If you want to be reasonably sure that there would never be a burned out bulb on the sign, how often should you replace all of the bulbs? Explain.

Lifetime (days)	Frequency
300–325	2
325–350	15
350–375	38
375–400	55
400–425	91
425–450	94
450–475	73
475–500	68
500–525	40
525–550	14
550–575	9
575–600	1

a)



b)

Lifetime (days)	Midpoint	Frequency	Mid x Freq	Relative Frequency
300 - 325	312.5	2	625	0.004
325 - 350	337.5	15	5062.5	0.030
350 - 375	362.5	38	13775	0.076
375 - 400	387.5	55	21312.5	0.110
400 - 425	412.5	91	37537.5	0.182
425 - 450	437.5	94	41125	0.188
450 - 475	462.5	73	33762.5	0.146
475 - 500	487.5	68	33150	0.136
500 - 525	512.5	40	20500	0.080
525 - 550	537.5	14	7525	0.028
550 - 575	562.5	9	5062.5	0.018
575 - 600	587.5	1	587.5	0.002

Mean = Total of (Mid x Freq) / Total Freq

= 220,025 / 500

= 400.05 days

d) $P(0 \le x \le 400) = 0.004 + 0.030 + 0.076 + 0.110$

Probability of bulb failing in 400 days or fewer = 0.220

e) To be **reasonably** sure that there will never be a burned out bulb, you should replace them every 440 days (the estimated mean).

9. Application Terry's Tree Farm planted Christmas tree seedlings last year. The table shows the current heights of a sample of 20 trees.

a) Determine the mean height.

b) Determine the standard deviation of the heights.

c) Are there enough data to predict whether the distribution of the heights follows a normal distribution? Explain. Include a graph or table as part of your answer.

Tree Height (cm)			
37.4	35.0	34.8	35.7
32.4	35.2	38.0	36.1
32.8	38.0	38.5	37.6
32.0	36.3	31.1	35.5
30.4	34.1	36.2	35.8

a) Mean = Total of heights / # of Trees

= 702.9 / 20

= 35.145 cm

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

b) The data is a **SAMPLE** of 20 trees, so use the sample formula.

$$s = \sqrt{\frac{\sum x^2 - n(\bar{x})^2}{n-1}}$$

Tree Height (cm)	x^2
37.4	1398.8
32.4	1049.8
32.8	1075.8
32.0	1024.0
30.4	924.2
35.0	1225.0
35.2	1239.0
38.0	1444.0
36.3	1317.7
34.1	1162.8
34.8	1211.0
38.0	1444.0
38.5	1482.3
31.1	967.2
36.2	1310.4
35.7	1274.5
36.1	1303.2
37.6	1413.8
35.5	1260.3
35.8	1281.6

$\sum x^2$ = 24809.35

$n(\bar{x})^2$ = 20(35.145)^2

= 24703.4205

$$s = \sqrt{\frac{24809.35 - 24703.4205}{20 - 1}}$$

$$s = \sqrt{\frac{105.9295}{19}}$$

s = √5.575236842

s = 2.361193944

The standard deviation of the heights is 2.36 cm

c) If the distribution is said to be normal then a representative sample can be taken from it.

The problem we have here is that our sample is relatively small (only 20 trees), so there is not enough data to predict whether the distribution is normal or not. The data may be centred (based on the table) but it isn't conclusive whether it will drop off symmetrically to the left and right sides. There may be some skewness.

Tree Height (cm)	Frequency
30 - 32	3
32 - 34	2
34 - 36	7
36 - 38	7
38 - 40	1

Possibly centred around these?

Not enough data to see if these will be equal or not

There may be trees not in the sample that may be greater than 40 cm

10. A student from region A scored 400 on a standardized math test with a mean of 350 and a standard deviation of 35. A student from region B scored 67 on a standardized math test with a mean of 62 and a standard deviation of 5. Both are being considered for a scholarship at the same university.

a) How can the university decide which candidate is the better student?

b) What assumptions must the university make?

c) Apply the method in part a), and determine which candidate is the better student. Give a reason for your answer.

a) The university can use z-scores for each student to see how many standard deviations above the mean each student is performing.

b) If doing this the university must assume that the tests are of comparable difficulty and cover the same material.

c) Region A: x = 400, $\overline{x}$ = 350, s = 35

$$z = \frac{x - \overline{x}}{s}$$

z = (400 - 350) / 35

z = 1.42857...

Region B: x = 67, $\overline{x}$ = 62, s = 5

$$z = \frac{x - \overline{x}}{s}$$

z = (67 - 62) / 5

z = 1.00

The student from Region A performed at a level 1.43 standard deviations above the mean compared to the student from Region B who performed only 1 standard deviation above the mean.

The student from Region A is the better candidate.

13. **Application** A coal-fired power plant releases some radioactive substances into the air. For example, a 1000-MW coal plant releases a mean of 5200 kg of uranium per year, with a standard deviation of 1300 kg. The release of uranium follows a normal distribution.

a) What is the probability that the plant will release less than 4000 kg of uranium in a given year?

b) What is the probability that the plant will release more than 6000 kg of uranium in a given year?

c) What is the probability that the release will be within 10% of the mean?

Calculate the z-scores and then use the table on Pages 480-481.

a) x = 4000, $\overline{x}$ = 5200, s = 1300

$$z = \frac{x - \overline{x}}{s}$$

z = (4000 - 5200) / 1300

z = -0.9230 ⟶ P(x < 4000) = 0.1788

z	0.00	0.01	0.02	0.03
−2.9	0.0019	0.0018	0.0018	0.0017
−2.8	0.0026	0.0025	0.0024	0.0023
−2.7	0.0035	0.0034	0.0033	0.0032
−2.6	0.0047	0.0045	0.0044	0.0043
−2.5	0.0062	0.0060	0.0059	0.0057
−2.4	0.0082	0.0080	0.0078	0.0075
−2.3	0.0107	0.0104	0.0102	0.0099
−2.2	0.0139	0.0136	0.0132	0.0129
−2.1	0.0179	0.0174	0.0170	0.0166
−2.0	0.0228	0.0222	0.0217	0.0212
−1.9	0.0287	0.0281	0.0274	0.0268
−1.8	0.0359	0.0351	0.0344	0.0336
−1.7	0.0446	0.0436	0.0427	0.0418
−1.6	0.0548	0.0537	0.0526	0.0516
−1.5	0.0668	0.0655	0.0643	0.0630
−1.4	0.0808	0.0793	0.0778	0.0764
−1.3	0.0968	0.0951	0.0934	0.0918
−1.2	0.1151	0.1131	0.1112	0.1093
−1.1	0.1357	0.1335	0.1314	0.1292
−1.0	0.1587	0.1562	0.1539	0.1515
−0.9	0.1841	0.1814	0.1788	0.1762
−0.8	0.2119	0.2090	0.2061	0.2033

b) x = 6000, $\overline{x}$ = 5200, s = 1300

$$z = \frac{x - \overline{x}}{s}$$

z = (6000 - 5200) / 1300

z = 0.6154 ⟶ P(x < 6000) = 0.7324

z	0.00	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985
0.6	0.7257	0.7291	0.7324

However we want P(x > 6000) = 1 - 0.7324

= 0.2676

c) Find the lower and upper limits

Lower = Mean - 10%
= 5200 - 520
= 4680

Upper = Mean + 10%
= 5200 + 520
= 5720

To be within 10% of the mean we need
P(x > 4680) and P(x < 5720)

x = 4680, $\overline{x}$ = 5200, s = 1300

$$z = \frac{x - \overline{x}}{s}$$

z = (4680 - 5200) / 1300

z = -0.4   P(x < 4680) = 0.3446

z	0.00	0.01
–2.9	0.0019	0.0018
–2.8	0.0026	0.0025
–2.7	0.0035	0.0034
–2.6	0.0047	0.0045
–2.5	0.0062	0.0060
–2.4	0.0082	0.0080
–2.3	0.0107	0.0104
–2.2	0.0139	0.0136
–2.1	0.0179	0.0174
–2.0	0.0228	0.0222
–1.9	0.0287	0.0281
–1.8	0.0359	0.0351
–1.7	0.0446	0.0436
–1.6	0.0548	0.0537
–1.5	0.0668	0.0655
–1.4	0.0808	0.0793
–1.3	0.0968	0.0951
–1.2	0.1151	0.1131
–1.1	0.1357	0.1335
–1.0	0.1587	0.1562
–0.9	0.1841	0.1814
–0.8	0.2119	0.2090
–0.7	0.2420	0.2389
–0.6	0.2743	0.2709
–0.5	0.3085	0.3050
–0.4	0.3446	0.3409

x = 5720, $\overline{x}$ = 5200, s = 1300

$$z = \frac{x - \overline{x}}{s}$$

z = (5720 - 5200) / 1300

z = 0.4   P(x < 5720) = 0.6554

z	0.00	0.01
0.0	0.5000	0.5040
0.1	0.5398	0.5438
0.2	0.5793	0.5832
0.3	0.6179	0.6217
0.4	0.6554	0.6591
0.5	0.6915	0.6950

P(4680 < x < 5720) = P(x < 5720) - P(x < 4680)
= 0.6554 - 0.3446
= 0.3108

The probability of being within 10% of the mean is 0.3108