M1 Maths

# P6-6 Normal Distributions

- finding probabilities from values
- finding values from probabilities
- $z$-scores

Summary Learn Solve Revise Answers

## Summary

Many attributes which are continuous and which cluster around a mean value are normally distributed. The distribution curve is bell shaped.

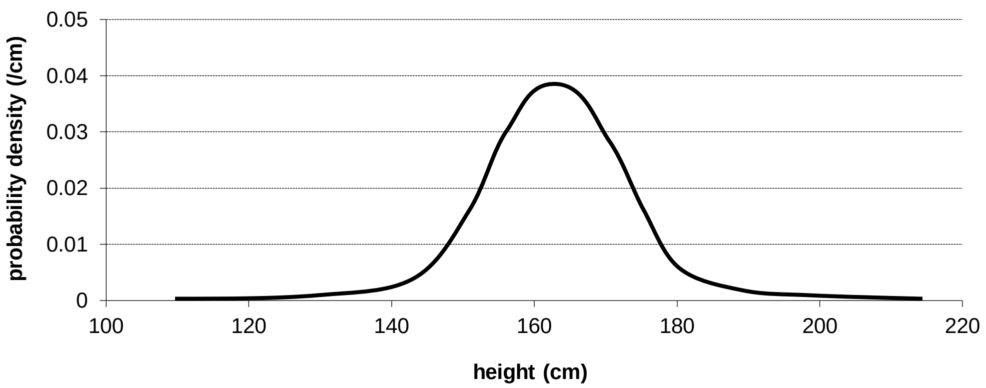
Calculators can be used to find the probability that the value of an attribute will lie within a given range if the mean and standard deviation of the distribution are known. Conversely, they can be used to find values given the probability.

A $z$-score of a value is the number of standard deviations above the mean. If the value is $x$, then $z = \frac{x-\mu}{\sigma}$ .

It is possible to use this to find the mean and/or standard deviation from other information about the distribution.

## Learn

Height is a continuous random variable. The probability density function for the heights of Australian women is shown below. The mean is 162 cm and the standard deviation is 10 cm.




This shape of distribution is very common in nature – so common in fact that it is called a *normal distribution*. Attributes which are continuous and which cluster around a mean are often normally distributed.

The curve is sometimes called the *bell curve* because it is shaped a bit like a bell. The normal distribution is symmetrical and unimodal, so the mean, median and mode always coincide. Normal distributions can differ in their mean (a higher mean shifts the curve to the right on the axes, a lower mean shifts it to the left) and they can differ in their spread or standard deviation (a higher spread causes the curve to increase in width and decrease in height, and vice versa). The mean of the distribution above is 164 cm and the standard deviation is 10 cm.

The probability distribution function for a normal distribution with mean $\mu$ and standard deviation $\sigma$ is $P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ where $x$ is the random variable (in this case height).

Fortunately you don't have to use this function or even remember it. Your calculator has functions for normal distributions and will do all the necessary calculations for you in response to a few simple inputs.

## Finding the probability that a woman's height will lie within a given range

Suppose we want to know the probability that an Australian woman's height is between 160 cm and 170 cm. This is the same as wanting to know what fraction (or percentage) of Australian women are between 160 cm and 170 cm tall.

Our calculator will do the calculations. On a Casio, select MENU 2 – STAT, then DIST (distribution), then NORM (normal), then Ncd (normal cumulative distribution). Then we enter numbers as follows, pressing EXE after each entry:

| | |
|---|---|
| Lower | :160 |
| Upper | :170 |
| $\sigma$ | :10 |
| $\mu$ | :162 |

Then we Execute (or just press EXE again) and we get $p$ = 0.36740431.

This tells us that the probability that an Australian woman selected at random will be between 160 and 170 cm is about 0.37 or 37%. The other way of reading it is that 37% of Australian women are between 160 and 170 cm.

If we need to find the percentage of Australian women who are more than 175 cm tall, we make the same entries except that we choose an upper limit which is higher than any woman is likely to be, e.g. 500 cm.

Lower :175

Upper :500

$\sigma$ :10

$\mu$ :162

This will give us the result $p = 0.097$, so 9.7% of women are more than 175 cm tall.

Similarly for 'less than 168 cm', choose a ridiculously low lower limit

Lower :0

Upper :168

$\sigma$ :10

$\mu$ :162

## Practice

Q1 Assuming that the heights of women are normally distributed with a mean of 162 cm and a standard deviation of 10 cm, find:

(a) the probability than a woman selected at random will be between 150 and 160 cm tall

(b) the percentage of women between 170 cm and 180 cm tall

(c) the probability that a randomly selected woman will be taller than 158 cm

(d) the percentage of women shorter than 155 cm

Q2 Dutch men have an average height of 182.5 cm with a standard deviation of 10.5 cm. Find:

(a) the probability than a Dutch man selected at random will be between 170 and 190 cm tall

(b) the percentage of Dutch men taller than 2 m.

(c) the percentage of Dutch men shorter than 170 cm

(d) the percentage of Dutch men who would be an acceptable height for a movie role that requires them to be taller than 190 cm, but less than 205 cm.

Q3 IQs (intelligence quotients – a largely obsolete measure of intelligence) of Americans are normally distributed with a mean of 110 and a standard deviation of 15. Roughly what fraction of Americans would have IQs between:

(a) 104 and 105

(b) 104 and 106

(c) 122 and 126

(d) above 148

(e) 148 or above

Q4 The masses of the 1729 sheep slaughtered at the Williams Brothers Meat Works in the past month were roughly normally distributed with a mean of

28.4 kg and a standard deviation of 2.1 kg. How many are likely to have weighed:

(a) between 24 and 25 kg

(b) between 31.0 kg and 31.6 kg

(c) less than 27 kg

(d) more than 28.6 kg

Q5 Babe aimed 1000 darts at a vertical line 2 m from the left end of a wall. When measured, the distances from the left end were roughly normally distributed with a mean of 2.05 m (Babe has a slight systematic bias to the right) and a standard deviation of 73 cm (Babe is not a very good shot anyway). How many darts are likely to have landed:

(a) between 1 and 3 metres from the end of the wall

(b) more than 3 m from the end of the wall

(c) less than 10 cm from the line he was aiming at

(d) completely to the left of the wall

## Finding the range, given the probability

We can also do the reverse calculations. Suppose we wanted to know the height above which the tallest 2% of Australian women lie. Here we are given the probability (2%) and need to find the lower limit. [Note that we can only find one of the limits, the other limit being assumed to be completely above or below the population spread.]

What we do is go to MENU 2, DIST, NORM, InvN (inverse normal).

The calculator asks us if we want the left, right or centre tail. In this example, we are interested in the 2% of women at the right-hand end of the distribution, so we choose RIGHT. Had the question been about the shortest women, we would have chosen LEFT. We don't need to use CENTRE: you can play with this and work out what it does if you like.

We enter these values:

| | |
|---|---|
| Tail | :Right |
| Area | :0.02 |
| $\sigma$ | :10 |
| $\mu$ | :162 |

[Note than the area under the curve between two $x$ values represents the probability that $x$ lies between those values. The formula for the curve is set so that the total area under it is 1.]

When we execute, we should get 182.5. This shows that 2% of Australian women are taller than 182.5 cm.

## Practice

Q6 British men's heights have $\mu$ = 175.5 cm, $\sigma$ = 11.0 cm.

(a) Find the lower cut-off height of the tallest 5% of British men.

(b) If 2.53% of British men are considered dwarves, what is the maximum height for a dwarf?

(c) How tall would a doorway have to be for only 1% of British men to have to duck to walk through it?

(d) What is the median height for British men?

(e) What is the upper quartile height for British men?

(f) What if the 90$^{th}$ percentile height?

Q7 The actual contents of 375 mL Sucky-Cola cans are normally distributed with mean 369 mL and variance 81 mL.

(a) If the lowest 3% are considered seriously under-filled, what is the minimum amount needed not to be considered seriously under-filled?

(b) The fullest 20% are considered 'prize cans'. How many mL must a can contain to be considered a prize can?

(c) What percentage of the cans will have the correct amount or more?

(d) What percentage will have exactly the right amount?

Q8 As mentioned, IQs of American are normally distributed with a mean of 110 and a standard deviation of 15.

(a) What is the minimum IQ required to be in the top 10%?

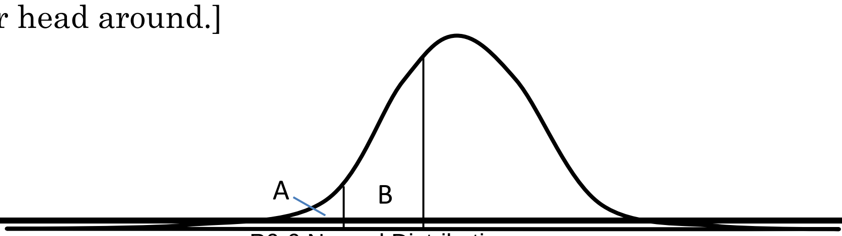
(b) A nincompoop is someone in the lowest 3% of the population IQ-wise. What is the IQ range for nincompoops?

(c) Mensa, a club for people with very high IQs, accepts those in the top half a percent of the population. What IQ would be required to join?

## More Complicated Calculations

A machine makes dog biscuits called Grunters. The masses are normally distributed with mean 6.5 g and standard deviation 0.8 g. Those under 5.5 g are fed to the dogs in the factory. The next 30% of those produced are sold as Mini-Grunters and the rest as Maxi-Grunters. What mass do the biscuits have to be to be a Maxi-Grunter?

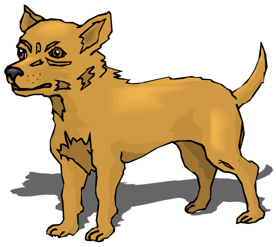
Solving problems like this can be much easier if we draw a sketch of the distribution. [Actually, drawing the sketch can be helpful in any normal distribution problem which is hard to get your head around.]



Those under 5.5 g which are fed to the factory dogs are represented by area A. The calculator will tell us that this area is 10.6% of the biscuits.

The area B is the Mini-Grunters. Areas A plus B must be 10.6% + 30%, i.e. 40.6%.

An inverse normal calculation will then tell us that $x$ is 6.31 g. So to be a Maxi-Grunter, a biscuit has to be more than 6.31 g.



Here are some more to practise on. Draw a sketch for each.

## Practice

Q9 They adjusted the Grunter machine so that the mean mass is now 7 g with a standard deviation of 0.6 g. So that they don't go hungry, they now feed all biscuits under 6 g to the factory dogs. The next 25% are sold as Mini-Grunters and the rest as Maxi-Grunters. What mass do the biscuits have to be now to be a Maxi-Grunter?

Q10 1230 Australian women (mean height = 162 cm, sd = 10 cm) auditioned for a role in a crowd scene of a movie. Those under 160 cm were rejected. The shortest 444 of those remaining got the part. How tall was the tallest person in the crowd in the movie?

Q11 The energy consumption of light bulbs which are nominally 23 Watts has a mean of 24.1 W and a standard deviation of 1.7 W. Those above 26 W burn out during testing. What is the median energy consumption of those which survive the test?

Q12 Find $c$ in the following situations
(a) $c$ if $P(38 < x < c) = 0.01$, $\mu = 30$ and $\sigma = 9$
(b) $c$ if $P(c < x < 2.2) = 0.61$, $\mu = 1.8$ and $\sigma = 0.2$

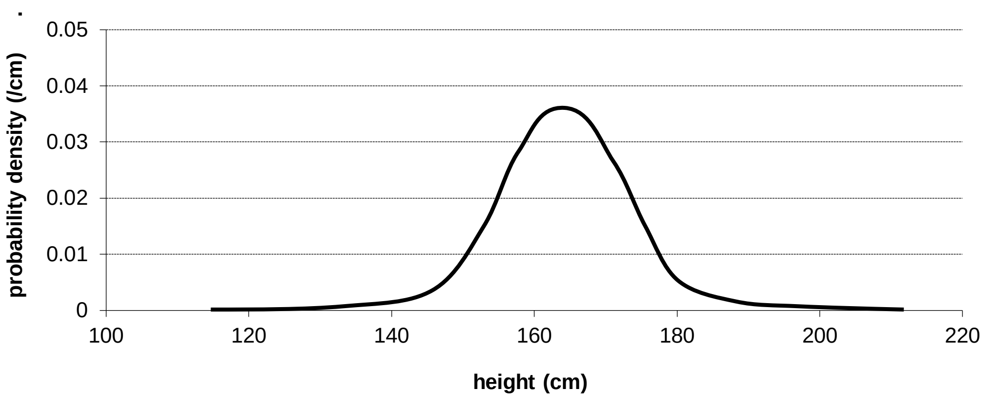
Q13 X ~ N(140, 4). [Note that this means that X is a variable or quantity that is normally distributed with mean 140 and standard deviation 4.]

If the highest 30% and the lowest 55% are removed from the population, what is the lower quartile of those left?

Q14 X ~ N(30, 2.5). If those above 34 and those below 28 are removed from the population, what is the 60$^{th}$ percentile of those left?

## z-Scores

Consider the distribution of heights of Australian women.




Height is a random variable we will call $X$. The values of $X$ we will call $x$. So, for example, if we take a 177 cm woman, for her, $x = 177$.

We often use another variable $Z$ with values $z$. $z$ is the number of standard deviations above the mean for a given $x$ value. As the mean height is 162 and the standard deviation is 10, our 177 cm woman is 15 cm above the mean, this is 1.5 lots of 10 or 1.5 standard deviations, so for her, $z = 1.5$ or we say her $z$-score is 1.5.

We can calculate the $z$ score from the $x$ score quite intuitively: subtract the mean from $x$ to find how many centimetres $x$ is above the mean; then divide this by the standard deviation to find how many standard deviations it is above the mean.

Or we can use the formula $z = \frac{x-\mu}{\sigma}$, which comes to the same thing.

### Practice

Q15 Heights of Australian women are normally distributed with mean 162 and sd 10. Find the $z$-scores for the following heights

(a) 172 cm
(b) 167 cm
(c) 162 cm
(d) 152 cm [Note we give this as –1.]
(e) 157 cm
(f) 164 cm
(g) 148 cm
(h) 193 cm
(i) 213 cm
(j) 172.7 cm
(k) 156.3 cm

Q16 Masses of new-born humans are normally distributed with mean 3260 g and standard deviation 485 g. Find the $z$-scores for the following masses:

(a) 3600 g
(b) 3160 g
(c) 2730 g
(d) 4020 g
(e) 3000 g
(f) 3760 g

We can calculate the $x$-score from the $z$-score by the reverse process: multiply the $z$-score by $\sigma$ to find how many units that is above the mean, then add that to the mean. Or use $x = \mu + z\sigma$.

## Practice

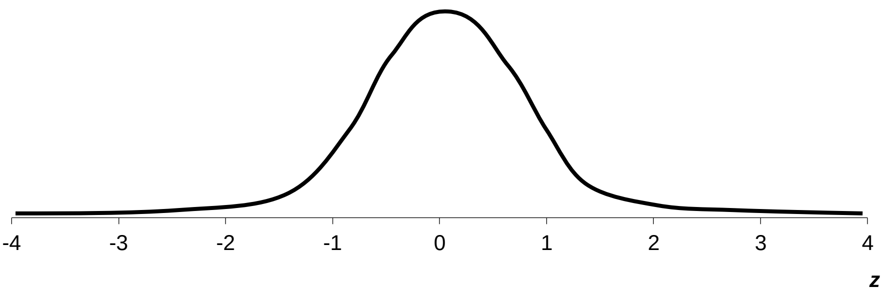
Q17 The masses of new-born wart hogs are normally distributed with mean 820 g and standard deviatin 65 g. Find the $x$-scores (masses) corresponding to the following $z$-scores.

(a) 1
(b) –2
(c) 0.65
(d) –0.71
(e) 3.11
(f) 0
(g) –4.5

Q18 $X \sim N(55.3, 8.7)$ means X is normally distributed with mean 55.3 and standard deviation 8.7. If $X \sim N(55.3, 8.7)$, find the $x$ scores for the following $z$-scores.

(a) 0.8
(b) 2.25
(c) –1.39
(d) 1.71
(e) – 0.37
(f) 11.6

$z$-scores are convenient because the distribution of $z$-scores is exactly the same for every normal distribution.




As you would expect, the mean of the distribution is 0 and the standard deviation is 1.

What's more, in every case, 68% of the values lie between –1 and 1, 95% lie between –2 and 2 and 99.7% lie between –3 and 3.

This distribution is called the standard normal distribution.

To do calculations with $z$-scores on the calculator, we use the same functions as before, but set $\mu = 0$ and $\sigma = 1$.

## Practice

Q19 For a normal distribution, find:

(a) $P(0.4 < z < 1.2)$

(b) $P(-0.4 < z < 0.4)$

(c) $P(-2 < z < 2)$

(d) $P(-1.1 < z < -0.3)$

Q20 What percentage of a normally distributed population would have a $z$-score

(a) $> 1.4$

(b) $< -1$ [Make your lower limit a large negative number, like –99.]

(c) $< 2.25$

(d) $> -0.28$

Q21 (a) Above what $z$-score would you find 12% of a normally distributed population?

(b) Above what $z$-score would you find 61% of a normally distributed population?

(c) Below what $z$-score would you find 5% of a normally distributed population?

(d) Below what $z$-score would you find 90% of a normally distributed population?

## Finding the mean and standard deviation of a normal distribution

Up to now, we have used the mean and standard deviation to work out $p$ from $x$ or vice versa. There are 4 quantities in the relation between $p$, $x$, $\mu$ and $\sigma$. If we know any three of them, we can find the other one. Firstly, we work out the value of $z$ from the value of $p$ using InvN on the calculator with $\mu = 0$ and $\sigma = 1$. Then we sub the three known values into $z = \frac{x-\mu}{\sigma}$ to get the unknown value.

### *Finding σ*

20% of a normally distributed population with mean 77 is over 82. What is the standard deviation?

We use $p = 0.2$ to get $z = 0.842$,

then we sub into $z = \frac{x-\mu}{\sigma}$ to get $0.842 = \frac{82-77}{\sigma}$, which we can solve to get $\sigma = 5.94$.

### *Finding μ*

20% of a normally distributed population with standard deviation 5.94 is over 82. What is the mean?

Again, we use $p = 0.2$ to get $z = 0.842$,

then we sub into $z = \frac{x-\mu}{\sigma}$ to get $0.842 = \frac{82-\mu}{5.94}$, which we can solve to get $\mu = 77$.

### *Finding both*

We can find $\mu$ and $\sigma$, but, of course, we will need to produce two equations and solve these simultaneously.

20% of a normally distributed population is over 82 and 60% of it is over 75.5. What are the mean and standard deviation?

We use $p = 0.2$ to get $z = 0.842$, and $p = 0.6$ to get $z = 0.253$

then we sub into $z = \frac{x-\mu}{\sigma}$ to get $0.842 = \frac{82-\mu}{\sigma}$ and $0.253 = \frac{75.5-\mu}{\sigma}$,

then we solve these two equations simultaneously to get $\mu = 77$ and $\sigma = 5.94$.

## Practice

Q22 The amount of milk in Happy Cow 1 L milk cartons is normally distributed with mean 1002 mL. If 10% have more than 1010 L, what is the standard deviation?

Q23 The amount of milk in Sad Cow 1L milk cartons is normally distributed with standard deviation 60 mL. If 74% have less than 990 mL, what is the mean?

Q24 The amount of milk in Bipolar Cow 1 L milk cartons is normally distributed. 80% have more than 985 mL and 5% have more than 1050 mL. Find the mean and standard deviation.

Q25 The lengths of Fire Bug matches are normally distributed. 10% are shorter than 5 cm; 20% are longer than 5.3 cm. Find the mean length and the standard deviation.

Q26 The number of people on the 6:35 a.m. train from Beenleigh to Brisbane on weekdays (excluding public holidays) is normally distributed. On 60% of the days there are less than 354 people. On 60% of the days there are more than 328 people. What are the mean number and the standard deviation?

## Solve

Q51 The men on Fargo Island have heights that are distributed normally with mean 172 cm and standard deviation 9 cm. A film company wants to offer one third of them a role in the mob scene of a pirate movie. But they want their range of heights to be as small as possible. What is the smallest possible height range?

Q52 In a maths competition, the awards rule is that those scoring more than 2 standard deviations above the mean get a medal and those scoring more than 2.5 standard deviations below the mean get publicly humiliated. The minister for fair treatment of maths competition participants decides that the percentage of competitors getting prizes should be doubled and the percentage getting publicly humiliated should be halved. What should the awards rule be changed to?

Q53 25 kg bags of cement contain an average of 25.6 kg of cement and the standard deviation is 0.8 kg. On a palette of 30 bags, what is the probability that there are 5 or more under-weight bags?

## Revise

### Revision Set 1

Q71 The masses of Yum Yum fruit cakes are normally distributed with mean 511 g and standard deviation 8 g.

(a) Find the probability that a given fruitcake will be between 500 and 510 g

(b) What percentage of the cakes are under 500 g?

(c) What is the $90^{th}$ percentile mass?

(d) If the lightest 8% are sold cheap and the heaviest 15% are taken home by the factory workers, what is the range of masses for those sold at the normal price?

(e) What is the median mass of those sold at the normal price?

Q72 (a) Find the $z$ score if $x = 24.9$, $\mu = 26$, $\sigma = 0.8$
(b) Find $x$ if the $z$ score is –1.31, $\mu = 120$, $\sigma = 4$
(c) What percentage of a normally distributed population would have a $z$-score above 2?
(d) What percentage of a normally distributed population would have a $z$-score below −0.033?
(e) Below what $z$-score would you find 18% of a normally distributed population?

Q73 The heights of Angolan men are normally distributed with mean 177 cm. If 27% of them are more than 186 cm, what is the standard deviation?

Q74 The masses of Chocolate Groggos are normally distributed. 10% are more than 55 g; 6% are less than 46 g. What are the mean and standard deviation?

## Answers

| | | | | | | |
|---|---|---|---|---|---|---|
| Q1 | (a) 0.306 | (b) 17.6% | (c) 65.5% | (d) 24.2% | | |
| Q2 | (a) 0.646 | (b) 4.8% | (c) 11.7% | (d) 22.1% | | |
| Q3 | (a) 0.025 | (b) 0.050 | (c) 0.069 | (d) 0.0056 | (e) 0.0056 | |
| Q4 | (a) 60 | (b) 76 | (c) 437 | (d) 799 | | |
| Q5 | (a) 828 | (b) 97 | (c) 109 | (d) 2 | | |
| Q6 | (a) 193.6 cm | (b) 154 cm | (c) 201.1 cm | (d) 175.5 cm | (e) 182.9 cm | (f) 189.6 cm |
| Q7 | (a) 352 mL | (b) 377 mL | (c) 25.2% | (d) 0% | | |
| Q8 | (a) 129 | (b) 0-82 | (c) 148 | | | |
| Q9 | 6.68 g | | | | | |
| Q10 | 169.8 cm | | | | | |
| Q11 | 23.81 W | | | | | |
| Q12 | (a) 38.34 | (b) 1.732 | | | | |
| Q13 | 140.88 | | | | | |
| Q14 | 30.98 | | | | | |
| Q15 | (a) 1 | (b) 0.5 | (c) 0 | (d) –1 | (e) –0.5 | (f) 0.2 |
| | (g) –1.4 | (h) 3.1 | (i) 5.1 | (j) 1.07 | (k) –0.57 | |
| Q16 | (a) 0.701 | (b) –0.206 | (c) –1.093 | (d) 1.567 | (e) –0.536 | (f) 1.031 |
| Q17 | (a) 885 g | (b) 690 g | (c) 862 g | (d) 774 g | (e) 1022 g | (f) 820 g |
| | (g) 528 g | | | | | |
| Q18 | (a) 62.3 | (b) 74.9 | (c) 43.2 | (d) 70.2 | (e) 52.1 | (f) 156.2 |
| Q19 | (a) 0.230 | (b) 0.311 | (c) 0.954 | (d) 0.246 | | |
| Q20 | (a) 8.1% | (b) 15.9% | (c) 98.8% | (d) 0.610 | | |
| Q21 | (a) 1.17 | (b) -0.28 | (c) −1.64 | (d) 1.28 | | |
| Q22 | 6.24 mL | | | | | |
| Q23 | 951.4 mL | | | | | |
| Q24 | 1007 mL, 26.1 mL | | | | | |
| Q25 | 0.141 cm | | | | | |
| Q26 | 341, 51.3 | | | | | |

Q51 7.75 cm (168.12 cm – 175.88 cm)

Q52 2.74 below and 1.69 above
Q53 84.2%

| | | | | | |
|---|---|---|---|---|---|
| Q71 | (a) 36.6% | (b) 8.5% | (c) 521.3 g | (d) 499.8 g – 519.3 g | (e) 510.3 g |
| Q72 | (a) –1.375 | (b) 114.76 | (c) 2.28% | (d) 48.68% | (e) –0.915 |

Q73 14.7 cm
Q74 $\mu = 50.93$ g $\sigma = 3.173$