Normal distribution (part 2)

We know how to determine the probability $P(Z \le z)$ where Z is a random variable following the STANDARD NORMAL DISRIBUTION ($Z^N(0,1)$)

The objective of this lesson is

to work out the probability $P(Z \le z)$ for any normal distribution to understand and work out "percentage points" of the normal distribution

Reminder:

Consider a random variable $Z^N(0,1)$, find out

a)
$$P(Z \le 2.43) =$$

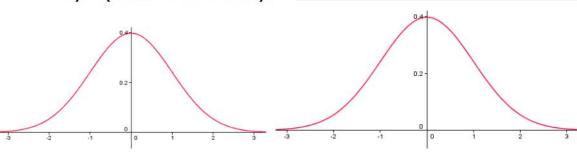
b)
$$P(Z \ge 0.34) =$$

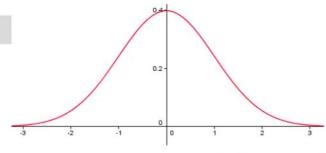
c)
$$P(Z \le -1.29) =$$

d)
$$P(Z > -3.5) =$$

e)
$$P(0 \le Z \le 0.97) =$$

f)
$$P(-2.5 \le Z \le 0.8) =$$

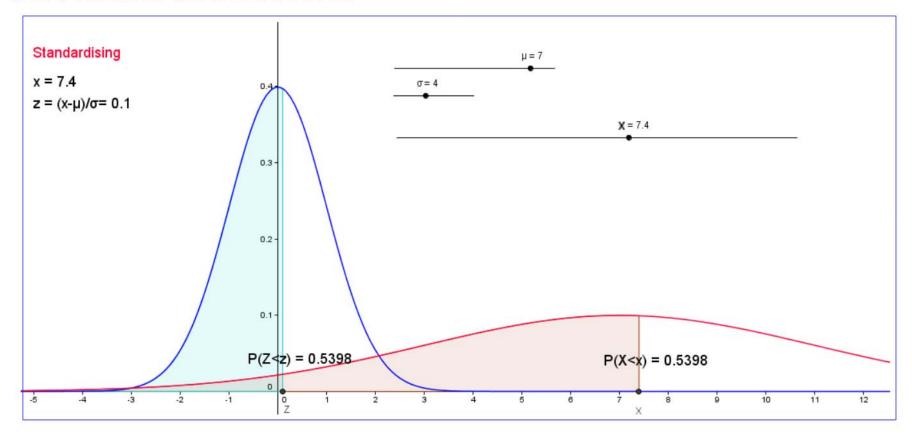




24262.0 26392.0=70553.0-1 28693.0=74102.0-1 77692.0 89555.0=2.0-89558.0 89555.0=2.0-89558.0

Standardising a normal variable

The table of values provides the probabilities of $P(Z \le z)$ for $Z \sim N(\mu=0, \sigma^2=1)$. What can we do when the normal distribution has a mean μ different from 0 and a variance different from 1?



In Blue: Z~N(0,1) In red: X~N(7,16)

Example:

If we want to work out P(X<x), we need to work out the value of z, so that the "blue" area is equal to the "red" one

This value is
$$z = \frac{x - \mu}{\sigma}$$

Worked example

The chest measurements of adult male customers for T-shirts may be modelled by a normal distribution with mean 101 cm and standard deviation 5 cm. Find the probability that a randomly selected customer will have a chest measurement which is:

- (a) less than 103 cm,
- 1b) 98 cm or more,
- (c) between 95 cm and 100 cm,
- d) between 90 cm and 110 cm.

z	0.00	0.01
0.0	0.50000	0.50399
0.1	0.53983	0.54380
0.2	0.57926	0.58317
0.3	0.61791	0.62172
0.4	0.65542	0.65910
0.5	0.69146	0.69497
0.6	0.72575	0.72907
0.7	0.75804	0.76115
8.0	0.78814	0.79103
0.9	0.81594	0.81859
1.0	0.84134	0.84375
1.1	0.86433	0.86650
1.2	0.88493	0.88686
1.3	0.90320	0.90490
1.4	0.91924	0.92073

$$a)z = \frac{103 - 101}{5} = 0.4$$

$$P(X < 103) = P(Z < 0.4) = 0.65542$$

$$b)z = \frac{98 - 101}{5} = -0.6$$

$$P(X \ge 98) = P(Z \ge -0.6) = P(Z \le 0.6) = 0.72575$$

c)
$$z = \frac{95-101}{5} = -1.2$$
 $z = \frac{100-101}{5} = -0.2$

$$P(95 \le X \le 100) = P(-1.2 \le Z \le -0.2) = 0.88493 - 0.57926 = 0.306$$

Standardising: exercises

Give your answer to question 1 to 5 correct to 3 decimal places

- **1** Given that $X \sim N(12, 9)$, find: a) P(X > 15), b) P(X < 16.8),
 - c) P(X < 8.31), d) P(X > 9.39).
- **2** Given that $X \sim N(50, 100)$, find:

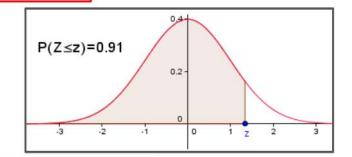
 - a) P(36 < X < 62) b) P(40 < X < 50)
 - c) P(55.8 < X < 72.3) d) P(38 < X < 42)
- **3** Given that $X \sim N(1.6, 4)$, find:
 - a) P(X > 0)
- b) P(X < -1.8)
- c) P(-2 < X < 2) d) P(0 < X < 2)
- **4** Given that $X \sim N(-4, 25)$, find:
 - a) P(X>0)
- b) P(-5.2 < X < -2.3)
- c) P(-2 < X < 1) d) P(X > 1 or X < -1)
- **5** Given that $Y \sim N(3.7, 2.4)$, find:
 - a) P(Y > 4)
- b) P(Y < 4.5)
- c) P(3.1 < Y < 4.2) d) P(2.8 < Y < 3.5)

Answers

- **1** a) 0.159 b) 0.945 c) 0.109 d) 0.808 **2** a) 0.804 b) 0.341 c) 0.268 d) 0.097 **3** a) 0.788 b) 0.045
- **3** c) 0.543 d) 0.367 **4** a) 0.212 b) 0.228 c) 0.186 d) 0.884 **5** a) 0.425 b) 0.698 c) 0.277 d) 0.167

Percentage points of the normal distribution

The purpose of this chapter is to "solve the equation" $P(Z \le z) = p$ Where p is a given number and z is the unknown

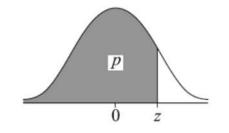


Note:

The value of z we look for is called the z-score

TABLE 4 PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

The table gives the values of z satisfying $P(Z \le z) = p$, where Z is the normally distributed random variable with mean = 0 and variance = 1.



p	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	p
0.5	0.0000	0.0251	0.0502	0.0753	0.1004	0.1257	0.1510	0.1764	0.2019	0.2275	0.5
0.6	0.2533	0.2793	0.3055	0.3319	0.3585	0.3853	0.4125	0.4399	0.4677	0.4958	0.6
0.7	0.5244	0.5534	0.5828	0.6128	0.6433	0.6745	0.7063	0.7388	0.7722	0.8064	0.7
0.8	0.8416	0.8779	0.9154	0.9542	0.9945	1.0364	1.0803	1.1264	1.1750	1.2265	0.8
0.9	1.2816	1.3408	1.4051	1.4758	1.5548	1.6449	1.7507	1.8808	2.0537	2.3263	0.9
p	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	p
0.95	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392	0.95
0.96	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663	0.96
0.97	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335	0.97
0.98	2.0537	2.0749	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904	0.98
0.99	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902	0.99

Some examples: $P(Z \le z) = 0.85 = 85\%$ for $z = P(Z \le z) = 0.972 = 97.2\%$ for z = 0.972 = 97.2%

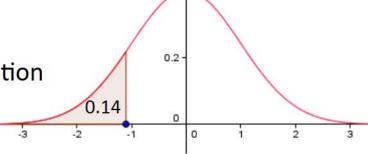
Note: in the percentage point table $p \geq 0.5$, $p \geq 50\%$

if you want to find the z-score for p < 0.5, you need to consider the symmetry of

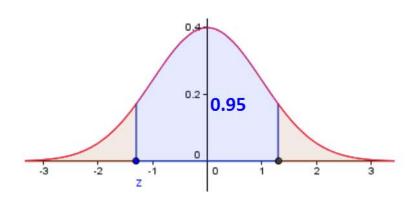
the distribution

Example:

Find the z-score which is greater than 14% of the population

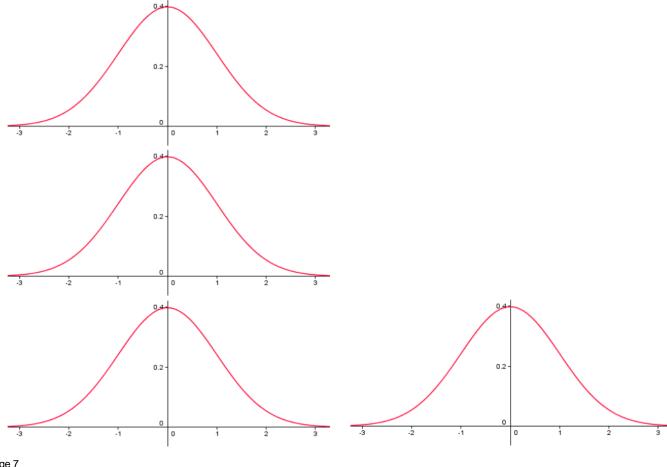


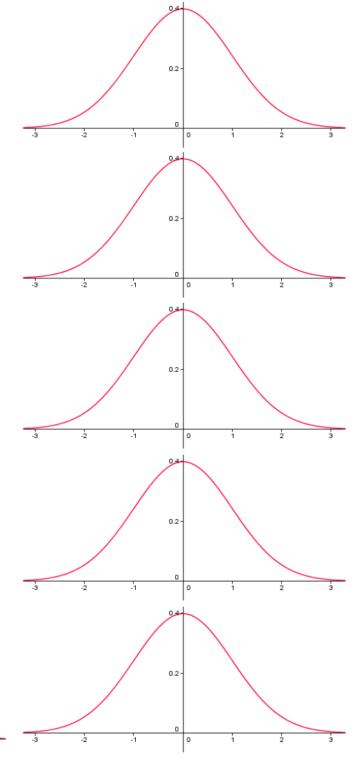
Find the limits of the CENTRAL 95%



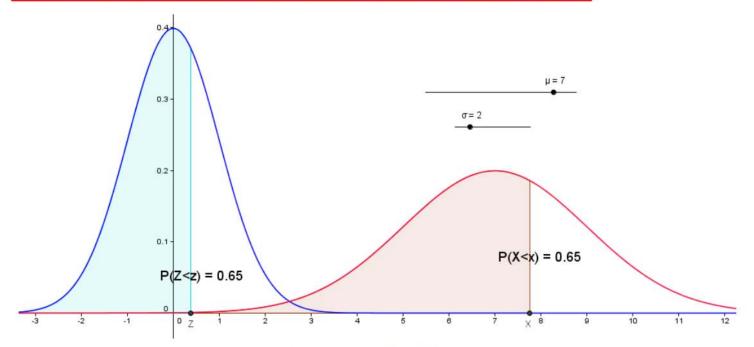
Exercises

- 1 Find the *z*-score which:
 - (a) is greater than 97.5% of the population,
 - **(b)** is less than 90% of the population,
 - (c) exceeds 5% of the population,
 - (d) is exceeded by 7.5% of the population,
 - (e) is greater than 2.5% of the distribution,
 - **(f)** is less than 15% of the population,
 - (g) exceeds 20% of the distribution,
 - (h) is greater than 90% of the distribution,
 - (i) is less than 1% of the population.





Percentage points of any normal distribution $N(\mu, \sigma^2)$



Consider the variable X^N(μ =7 , σ^2 =4) and Z^N(0,1) We want to find the x-score so that P(X \leq x) = 65% = 0.65

The standardisation transforms x into z so that $P(X \le x) = P(Z \le z)$ Therefore, solving $P(X \le x) = p$ where p is a given number is equivalent to solve $P(Z \le z) = p$.

Once we find z, we "reverse" the standardisation in order to find x.

if
$$z = \frac{x - \mu}{\sigma}$$
 then $x = \mu + \sigma z$

Example:

The wingspans of a population of birds are normally distributed with mean 14.1 cm and standard deviation 1.7 cm. Find:

- (a) the wingspan which will exceed 90% of the population,
- (b) the wingspan which will exceed 20% of the population,
- (c) the limits of the central 95% of the wingspans.

Solution

(a) The *z*-score which exceeds 90% of the population is 1.2816. The value required is therefore 1.2816 standard deviations above the mean, i.e.

$$14.1 + 1.2816 \times 1.7 = 16.3$$
 cm.

(b) The *z*-score which will exceed 20% of the population will be exceeded by 80% of the population.

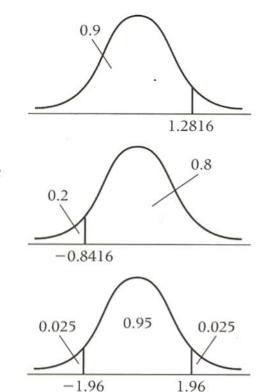
$$z = -0.8416$$

 $x = 14.1 - 0.8416 \times 1.7 = 12.7 \text{ cm}.$

(c) $z = \pm 1.96$

The central 95% of wingspans are 14.1 \pm 1.96 \times 1.7, i.e.

 14.1 ± 3.33 or 10.8 cm to 17.4 cm.



EXERCISE

- 1 A large shoal of fish have lengths which are normally distributed with mean 74 cm and standard deviation 9 cm.
 - (a) What length will be exceeded by 10% of the shoal?
 - (b) What length will be exceeded by 25% of the shoal?
 - (c) What length will be exceeded by 70% of the shoal?
 - (d) What length will exceed 95% of the shoal?
 - (e) Find the limits of the central 90% of lengths.
 - (f) Find the limits of the central 60% of lengths.
- 2 Hamburger meat is sold in 1 kg packages. The fat content of the packages is found to be normally distributed with mean 355 g and standard deviation 40 g.

Find the fat content which will be exceeded by:

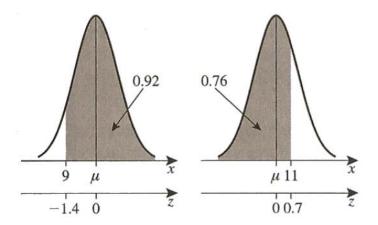
- (a) 5% of the packages,
- (b) 35% of the packages,
- (c) 50% of the packages,
- (d) 80% of the packages,
- (e) 99.9% of the packages.
- (f) Find the limits of the central 95% of the contents.

	-			12E	CKC	FY
;5.69	(c)	;1.08	(q)	; 5. 58	(9)	I
;6.18-₽.66	(1)	;8.88-1.9 č	(e)	;8.88	(p)	
322;	(c)	370;	(q)	:124	(a)	7
.884-772	(1)	:152	(ə)	:178	(p)	
	·ccb-//7	.EE1-772 (1)	'cc+-//7 (1) :1c7	(e) 231; (f) 277–433.	'55t-//7 (I) (157 (a) (176	(d) 321; (e) 231; (f) 277–433.

Mean, variance and standard deviation of normal distribution

The random variable X has a normal distribution. It is known that P(X > 9) = 0.92 and that $P(X \le 11) = 0.76$.

- \blacksquare) Determine the mean and the standard deviation of X, giving your answers to two decimal places.
- b) Determine P(X > 10), to two decimal places.



Have a go!

A vending machine discharges soft drinks. A total of 5% of the discharges have a volume of more than 475 cm³, while 1% have a volume less than 460 cm³. The discharges may be assumed to be normally distributed. Find the mean and standard deviation of the discharges.

```
3% of 2-scores exceed 1.6449. Hence, if the mean is \mu and the standard deviation is \sigma, \mu + 1.6449\sigma = 475. 1% of 2-scores are below -2.3263. Hence, \mu - 2.3263\sigma = 460. Subtracting equation [2] from equation [1] gives 3.9712\sigma = 15 o = 3.7772 o = 3.7772 Substituting in equation [1] = 468.787 Substituting in equation [1] = 468.787
```

Give your answers to Questions 1 to 4 to three significant figures.

- **1** Given that $X \sim N(\mu, 2.5)$ and that P(X > 3.5) = 0.970, find μ .
- **2** Given that $X \sim N(\mu, 0.5)$ and that P(X < -1.2) = 0.050, find μ .
- **3** Given that $X \sim N(32.4, \sigma^2)$ and that P(X > 45.2) = 0.300, find σ .
- **4** Given that $X \sim N(-7.21, \sigma^2)$ and that P(X < 0) = 0.900, find σ .
- 5 Given that $X \sim N(\mu, \sigma^2)$, that P(X > 0) = 0.800, and that P(X < 5) = 0.700, find μ and σ .
- 7 The quantity of milk in a bottle is normally distributed with a mean of 1000 ml. Find the standard deviation given that the probability that there is less than 995 ml in a randomly chosen bottle is 5%. What can you say about the standard deviation if the probability is to be less than 5%?
- 8 A variety of hollyhock grows to great heights. Assuming a normal distribution of heights, find the mean and standard deviation, given that the 30th and 70th percentiles are 1.83 m and 2.31 m respectively.
- 9 Due to manufacturing variations, the length of string in a randomly chosen ball of string can be modelled by a normal distribution. Find the mean and the standard deviation, given that 95% of balls of string have lengths exceeding 495 m, and that 99% have lengths exceeding 490 m. Give your answers in metres to two decimal places.
- The length of a brass cylinder has a normal distribution with a mean of μ and a variance of σ^2 , both being unknown. A large sample reveals that 10% of the cylinders are longer than 3.68 cm and that 3% are shorter than 3.52 cm. Find the values of μ and σ^2 .

```
1) \mu = 6.47

2) \mu = -0.0369

3) \sigma = 24.4

4) \sigma = 5.63

5) \mu = 3.08 and \sigma = 3.66

7) \sigma = 6.08. If p < 5\%, \sigma < 6.08

8) \mu = 2.07m and \sigma = 0.458m

9) \mu = 507.07m and \sigma = 7.34m

10) \mu = 3.62cm and \sigma^2 = 0.00256 cm<sup>2</sup>
```

Exam-Style Questions

1	The exam marks for 1000 candidates can be modelled by a normal distribution						
	with	with mean 50 marks and standard deviation 15 marks.					
	a)	One candidate is selected at random.	Find the probability that they scored				

than 30 marks on this exam.

(3 marks)

The pass mark is 41.
 Estimate the number of candidates who passed the exam.

(3 marks)

 Find the mark needed for a distinction if the top 10% of the candidates achieved a distinction.

(3 marks)

- 2 The random variable X follows a normal distribution with mean μ and standard deviation 6. The probability that X takes a value of less than 50 is 0.12302.
 - a) Find the mean of this distribution.

(4 marks)

b) Find P(X > 71).

(3 marks)

c) Find the value of a such that $P(\mu - a < X < \mu + a) = 0.8$.

(4 marks)

- 3 The lifetimes of a particular type of battery are normally distributed with mean μ and standard deviation σ . A student using these batteries finds that 40% last less than 20 hours and 80% last less than 26 hours.
 - a) Find μ and σ .

(7 marks)

b) Find the probability that a randomly selected battery of this type has a lifetime of at least 15 hours.

(3 marks)

Page 14

- 4 The random variable X has a normal distribution with mean 120 and standard deviation 25.
 - a) Find P(X > 145).

(3 marks)

b) Find the value of j such that $P(120 \le X \le j) = 0.46407$

(4 marks)

- 5 The diameters of the pizza bases made at a restaurant are normally distributed.
 The mean diameter is 12 inches, and 5% of the bases measure more than 13 inches.
 - a) Write down the median diameter of the pizza bases.

(1 mark)

b) Find the standard deviation of the diameters of the pizza bases.

(4 marks)

Any pizza base with a diameter of less than 10.8 inches is considered too small and is discarded.

c) If 100 pizza bases are made in an evening, approximately how many would you expect to be discarded due to being too small?

(3 marks)

Three pizza bases are selected at random.

d) Find the probability that at least one of these bases is too small.

(3 marks)

- 6 A garden centre sells bags of compost. The volume of compost in the bags is normally distributed with a mean of 50 litres.
 - a) If the standard deviation of the volume is 0.4 litres, find the probability that a randomly selected bag will contain less than 49 litres of compost.
 - b) If 1000 of these bags of compost are bought, find the expected number of bags containing more than 50.5 litres of compost.

A different garden centre sells bags of similar compost. The volume of compost, in litres, in these bags is described by the random variable Y, where $Y \sim N(75, \sigma^2)$. It is found that 10% of the bags from this garden centre contain less than 74 litres of compost.

c) Find σ.

(3 marks)

Exam-Style Questions

- 1 Let X represent the exam marks. Then $X \sim N(50, 15^2)$.
 - a) $P(X < 30) = P(Z < \frac{30 50}{15}) = P(Z < -1.33)$ [1 mark] $= P(Z > 1.33) = 1 - P(Z \le 1.33)$ [1 mark] = 1 - 0.90824 = 0.09176 [1 mark]
 - b) $P(X \ge 41) = P(Z \ge \frac{41 50}{15}) = P(Z \ge -0.6)$ [1 mark] $= P(Z \le 0.6) = 0.72575$ [1 mark] So $0.72575 \times 1000 = 726$ is the expected number who passed the exam [1 mark].
 - c) If a is the mark needed for a distinction, then: $P(X \ge a) = 0.1 \Rightarrow P\left(Z \ge \frac{a - 50}{15}\right) = 0.1 \text{ [1 mark]}$ $\Rightarrow P\left(Z < \frac{a - 50}{15}\right) = 0.9$ p = 0.9 for z = 1.2816 $\Rightarrow \frac{a - 50}{15} = 1.2816 \text{ [1 mark]}$ $\Rightarrow a = 69 \text{ (to the nearest whole mark) [1 mark]}$
- 2 a) P(X < 50) = 0.12302 $\Rightarrow P\left(Z < \frac{50 - \mu}{6}\right) = 0.12302$ [1 mark] $\Rightarrow P\left(Z > \frac{\mu - 50}{6}\right) = 0.12302$ $\Rightarrow P\left(Z \le \frac{\mu - 50}{6}\right) = 1 - 0.12302 = 0.87698$ [1 mark]

$$\Phi(z) = 0.87698 \text{ for } z = 1.16$$

 $\Rightarrow \frac{\mu - 50}{6} = 1.16 \text{ [1 mark]}$
 $\Rightarrow \mu = 1.16 \times 6 + 50 = 56.96 \text{ [1 mark]}$

- **b)** $P(X > 71) = P(Z > \frac{71 56.96}{6})$ = P(Z > 2.34) [1 mark] = $1 - P(Z \le 2.34)$ [1 mark] = 1 - 0.99036 = 0.00964 [1 mark]
- c) Since the distribution is symmetrical about μ , $P(X < \mu + a) = 0.9$ [1 mark] $\Rightarrow P\left(Z < \frac{\mu + a - \mu}{6}\right) = P\left(Z < \frac{a}{6}\right) = 0.9$ [1 mark] p = 0.9 for z = 1.2816 $\Rightarrow \frac{a}{6} = 1.2816$ [1 mark] $\Rightarrow a = 7.69$ (to 3 s.f.) [1 mark]

a) Let X represent the lifetime of a battery in hours. Then $X \sim N(\mu, \sigma^2)$.

$$P(X < 20) = 0.4 \Rightarrow P\left(Z < \frac{20 - \mu}{\sigma}\right) = 0.4 \text{ [1 mark]}$$

$$\Rightarrow P\left(Z > \frac{\mu - 20}{\sigma}\right) = 0.4$$

$$\Rightarrow P\left(Z \le \frac{\mu - 20}{\sigma}\right) = 0.6$$

Using the percentage-points table, p = 0.6 for z = 0.2533

$$\Rightarrow \frac{\mu - 20}{\sigma} = 0.2533$$
 [1 mark]

$$\Rightarrow \mu - 20 = 0.2533\sigma$$
 (equation 1)

$$P(X < 26) = 0.8 \Rightarrow P(Z < \frac{26 - \mu}{\sigma}) = 0.8$$
 [1 mark]

Using the percentage-points table,

$$p = 0.8$$
 for $z = 0.8416$

$$\Rightarrow \frac{26 - \mu}{\sigma} = 0.8416$$
 [1 mark]

$$\Rightarrow$$
 26 - μ = 0.8416 σ (equation 2)

Adding equations 1 and 2 gives:

 $\mu - \mu - 20 + 26 = 0.2533\sigma + 0.8416\sigma$ [1 mark]

$$\Rightarrow$$
 6 = 1.0949 σ \Rightarrow σ = 5.47995... = 5.48 (to 3 s.f.)

Putting $\sigma = 5.47...$ into equation 1 gives:

$$\mu = 0.2533 \times 5.47... + 20 = 21.4$$
 (to 3 s.f.)

So
$$\mu = 21.4$$
 hours [1 mark]

and $\sigma = 5.48$ hours [1 mark] (to 3 s.f.)

b) From part a), $X \sim N(21.4, 5.48^2)$

So
$$P(X \ge 15) = P(Z \ge \frac{15 - 21.4}{5.48})$$
 [1 mark]

$$= P(Z \ge -1.17) = P(Z \le 1.17)$$
 [1 mark]
= 0.8790 [1 mark]

4 a)
$$P(X > 145) = P(Z > \frac{145 - 120}{25})$$

$$= P(Z > 1) [1 mark]$$

= 1 = P(Z < 1) [1 mark]

$$= 1 - P(Z \le 1)$$
 [1 mark]

= 1 - 0.84134 = 0.15866 [1 mark]

b) P(120 < X < j) = 0.46407

$$\Rightarrow P\left(\frac{120 - 120}{25} < Z < \frac{j - 120}{25}\right) = 0.46407$$

$$\Rightarrow P(0 < Z < \frac{j-120}{25}) = 0.46407 [1 mark]$$

$$\Rightarrow P\left(Z < \frac{j - 120}{25}\right) - P(Z \le 0) = 0.46407$$

$$\Rightarrow P(Z < \frac{j-120}{25}) = 0.46407 + 0.5$$

$$\Rightarrow P(Z < \frac{j-120}{25}) = 0.96407 [1 mark]$$

$$\Phi(z) = 0.96407$$
 for $z = 1.80$

$$\Rightarrow \frac{j-120}{25} = 1.8 [1 \text{ mark}]$$

$$\Rightarrow$$
 j = 1.8 × 25 + 120 = 165 [1 mark]

- a) For a normal distribution, mean = median, so median = 12 inches [1 mark].
- **b)** Let X represent the base diameters. Then $X \sim N(12, \sigma^2)$. P(X > 13) = 0.05 [1 mark], so $P\left(Z > \frac{13}{\sigma} = 12\right) = P\left(Z > \frac{1}{\sigma}\right) = 0.05$ [1 mark] and $P\left(Z \le \frac{1}{\sigma}\right) = 0.95$

Using the percentage-points table,

$$p = 0.95$$
 for $z = 1.6449$

$$\Rightarrow \frac{1}{\sigma} = 1.6449$$
 [1 mark]

$$\Rightarrow \sigma = 1 \div 1.6449 = 0.608$$
 (to 3 s.f.) [1 mark]

c)
$$P(X < 10.8) = P(Z < \frac{10.8 - 12}{0.608})$$

$$= P(Z < -1.97) [1 mark]$$

$$= P(Z > 1.97) = 1 - (Z \le 1.97)$$

$$= 1 - 0.97558 = 0.02442$$
 [1 mark]

So you would expect $0.02442 \times 100 \approx 2$ pizza bases to be discarded [1 mark].

d) P(at least 1 base too small)

= 1 - P(no bases too small).

P(base not too small) = 1 - 0.02442 = 0.97558.

 $P(\text{no bases too small}) = 0.97558^3 = 0.9285$

[1 mark]

P(at least 1 base too small) = 1 - 0.9285 [1 mark] = 0.0715 [1 mark].

6 a) Let X represent the volume of compost in a bag. Then $X \sim N(50, 0.4^2)$.

$$P(X < 49) = P\left(Z < \frac{49 - 50}{0.4}\right)$$

$$= P(Z < -2.5) [1 mark]$$

$$= P(Z > 2.5) = 1 - P(Z \le 2.5)$$
 [1 mark]

$$= 1 - 0.99379 = 0.00621$$
 [1 mark]

b)
$$P(X > 50.5) = P(Z > \frac{50.5 - 50}{0.4})$$

$$= P(Z > 1.25) [1 mark]$$

$$= 1 - P(Z \le 1.25)$$
 [1 mark]

$$= 1 - 0.89435 = 0.10565$$
 [1 mark]

So in 1000 bags, 0.10565 × 1000 [1 mark] ≈ 106 bags [1 mark] (approximately) would be expected to contain more than 50.5 litres of compost.

c) P(Y < 74) = 0.1

$$\Rightarrow P\left(Z < \frac{74 - 75}{\sigma}\right) = P\left(Z < -\frac{1}{\sigma}\right) = 0.1$$

 $\Rightarrow P(Z > \frac{1}{\sigma}) = 0.1 \Rightarrow P(Z \le \frac{1}{\sigma}) = 0.9$ [1 mark]

Using the percentage-points table,

$$p = 0.9$$
 for $z = 1.2816$

$$\Rightarrow \frac{1}{\sigma} = 1.2816$$
 [1 mark]

$$\Rightarrow \sigma = 1 \div 1.2816$$

= 0.780 litres (to 3 s.f.) [1 mark]