Cambridge AS & A Level · Mathematics 9709 · Notes

The normal distribution AS

Read online at **pastopic.com/cambridge/mathematics-9709/notes/the-normal-distribution**

Last checked 29 September 2026

Read first: **Discrete random variables**

What you need to know

By the end of this topic you should be able to:

1. **Use a normal distribution as a model for a continuous random variable** (a length, a mass, a time), and read the **normal distribution tables** in MF19: the table of $\Phi(z)$ and the table of critical values. Draw a quick sketch of the bell-shaped curve to show the distribution or the probability you want.
2. **Find a probability** when $X \sim N(\mu, \sigma^2)$ and you know μ and σ: for example $P(X > x_1)$, $P(X < x_1)$ or $P(x_1 < X < x_2)$. You do this by **standardising**, $z = \frac{x-\mu}{\sigma}$, and the working must be shown in full.
3. **Work backwards from a probability**: when you are told something like $P(X > x_1) = 0.2$, find the z-value and use $\frac{x_1-\mu}{\sigma} = z$ to find whichever of x_1, μ or σ is unknown. With two such facts you can find both μ and σ.
4. **Approximate a binomial distribution by a normal one.** Remember the conditions: n large enough that $np > 5$ and $nq > 5$ (where $q = 1 - p$). Then use $N(np,\ npq)$ with a **continuity correction** to find probabilities.

Anything about adding or combining normal variables, or the sample mean, belongs to Probability & Statistics 2.

Understand it

A model for measurements

Many measured quantities, such as heights, masses and journey times, cluster around an average, with values far above or far below it becoming rarer and rarer, equally on both sides. The **normal distribution** is the model for this. It is a **continuous** distribution: X can take any value in a range, not just whole numbers.

We write $X \sim N(\mu, \sigma^2)$, where μ is the mean and σ^2 is the **variance**. So $N(20, 9)$ has mean 20 and standard deviation $\sqrt{9} = 3$. Watch for this: the second number is the variance, not the standard deviation.

The graph of a normal distribution is a symmetrical **bell-shaped curve**, highest at μ. Probabilities are **areas under the curve**, and the total area is 1:

- half the area is on each side of μ, so $P(X < \mu) = P(X > \mu) = 0.5$;

- about 68% of values lie within one standard deviation of the mean, and about 95% within two;
- a larger σ gives a wider, flatter curve; changing μ slides the curve along without changing its shape.

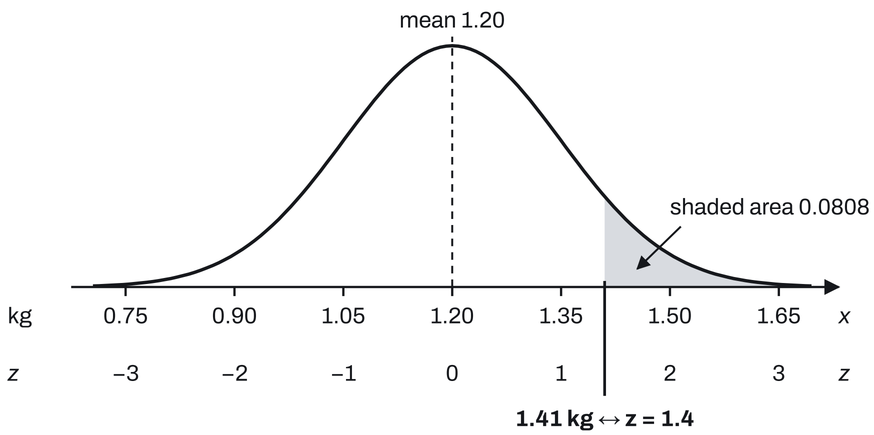
Because X is continuous, the probability of one exact value is 0, so $P(X < a)$ and $P(X \leqslant a)$ are the same. This is why **no continuity correction** is used for a normal variable such as a length or a time.

Standardising

There is only one table, for the **standard normal distribution** $Z \sim N(0, 1)$. Any normal variable is turned into Z by asking "how many standard deviations above the mean is this value?":

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

For melons with masses $X \sim N(1.20,\ 0.15^2)$ kg, a melon of 1.41 kg has $z = \frac{1.41-1.20}{0.15} = 1.4$: it is 1.4 standard deviations above the mean. So $P(X > 1.41) = P(Z > 1.4)$.



Reading the tables

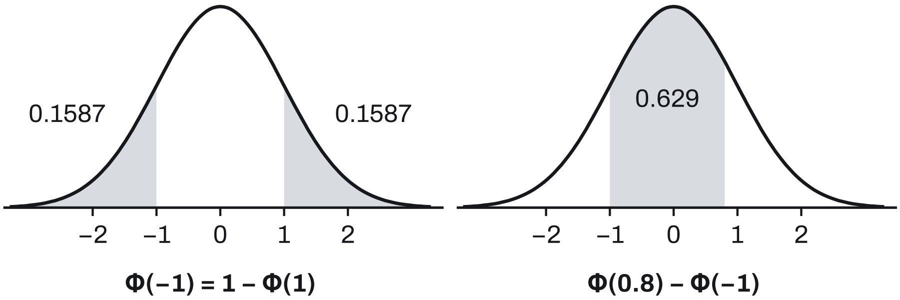
The MF19 table gives $\Phi(z) = P(Z \leqslant z)$, the area **to the left** of z, for $z \geqslant 0$. Look up the first two digits of z in the rows and columns and use the "ADD" columns for the third decimal place, so $\Phi(1.375) = 0.9147 + 0.0008 = 0.9155$ (the ADD entry is in units of the fourth decimal place). Always use at least 3 decimal places of z.

Everything else comes from the symmetry of the curve and the total area of 1:

You want	Use
$P(Z < a), a > 0$	$\Phi(a)$
$P(Z > a), a > 0$	$1 - \Phi(a)$
$P(Z < -a)$	$1 - \Phi(a)$
$P(Z > -a)$	$\Phi(a)$
$P(-a < Z < b)$	$\Phi(b) - \Phi(-a) = \Phi(b) + \Phi(a) - 1$

You want	Use
$P(-a < Z < a)$	$2\Phi(a) - 1$

For example, $P(Z > 1.4) = 1 - \Phi(1.4) = 1 - 0.9192 = 0.0808$, the shaded area above.



The **critical values** table works the other way round: for $p = 0.75, 0.9, 0.95, 0.975, 0.99, \ldots$ it gives the z with $P(Z \leqslant z) = p$. The ones you meet most are:

p	0.75	0.9	0.95	0.975	0.99
z	0.674	1.282	1.645	1.960	2.326

For any other probability, find it inside the Φ table and read off z (for example $\Phi(z) = 0.8$ gives $z = 0.842$, and $\Phi(z) = 0.85$ gives $z = 1.036$).

Working backwards

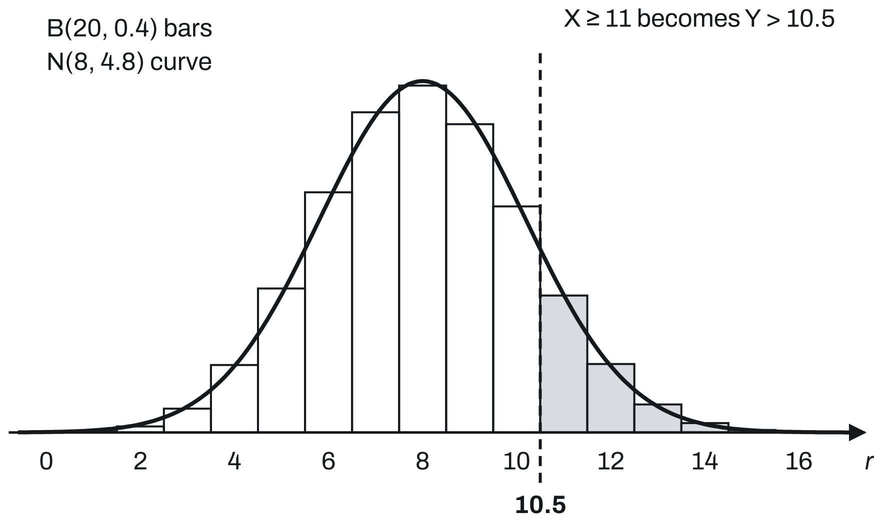
If 5% of parcels weigh more than 12.4 kg, the z-value that cuts off the top 5% is the one with $\Phi(z) = 0.95$, which is 1.645. So $\frac{12.4-\mu}{\sigma} = 1.645$: one equation linking x_1, μ and σ. Know two of them and you can find the third. A value **below** the mean always has a **negative** z; a sketch tells you which sign you need.

Approximating a binomial distribution

When $X \sim B(n, p)$ with n large, working out many binomial terms is slow, but the bar chart of the distribution looks very like a normal curve. It works well when $np > 5$ and $nq > 5$, so that the distribution is not squashed against 0 or n. Then use the normal distribution with the **same mean and variance**:

$$X \approx Y \sim N(np,\ npq).$$

X only takes whole numbers, but Y is continuous. Each whole number r is the bar from $r - 0.5$ to $r + 0.5$, so we change each inequality to include or exclude whole bars. This is the **continuity correction**.



For $X \sim B(20, 0.4)$: $np = 8$ and $nq = 12$, both more than 5, so use $Y \sim N(8, 4.8)$. Then $P(X \geqslant 11) \approx P(Y > 10.5) = 1 - \Phi\left(\frac{2.5}{\sqrt{4.8}}\right) = 1 - \Phi(1.141) = 0.127$, very close to the exact binomial value 0.1275.

Binomial	Normal with continuity correction
$P(X \leqslant 30)$	$P(Y < 30.5)$
$P(X < 30)$, that is $X \leqslant 29$	$P(Y < 29.5)$
$P(X \geqslant 30)$	$P(Y > 29.5)$
$P(X > 30)$, that is $X \geqslant 31$	$P(Y > 30.5)$
$P(20 \leqslant X \leqslant 25)$	$P(19.5 < Y < 25.5)$

The rule: first write the whole numbers you want (for example "fewer than 30" is $0, 1, \ldots, 29$), then go half a unit beyond the last one you include.

Key facts and formulas

"Given" means it is in the MF19 list of formulae and tables in the exam. "Learn it" means you must remember it.

Formula	Status
$\Phi(z) = P(Z \leqslant z)$ for $Z \sim N(0,1)$, in the table	Given
$\Phi(-z) = 1 - \Phi(z)$	Given
Critical values: $P(Z \leqslant z) = p$, e.g. $p = 0.9 \Rightarrow z = 1.282$	Given
$X \sim N(\mu, \sigma^2)$: mean μ, **variance** σ^2	Learn it
Standardising: $Z = \dfrac{X - \mu}{\sigma}$	Learn it

Formula	Status
$P(Z > a) = 1 - \Phi(a)$; $P(-a < Z < a) = 2\Phi(a) - 1$	Learn it
$P(X = a) = 0$ for a continuous variable, so $<$ and $\leqslant$ give the same probability	Learn it
Binomial $B(n, p)$: $\mu = np$, $\sigma^2 = np(1-p)$	Given
Normal approximation valid when $np > 5$ and $nq > 5$	Learn it
$B(n, p) \approx N(np,\ npq)$ with a continuity correction (± 0.5)	Learn it
Expected number $= n \times$ probability	Learn it

How to do it

In the Paper 5 questions we have tagged for this topic (59 questions, 2021–2025), the types came up this often. 34 of the 59 questions mix types, so the counts add up to more than 59.

Question type	Questions
Probabilities and expected numbers from $N(\mu, \sigma^2)$	36
Finding an unknown value, mean or standard deviation	26
Normal approximation to the binomial	26
Finding both the mean and the standard deviation	9

1. Probabilities and expected numbers from $N(\mu, \sigma^2)$ (36 questions)

1. **Sketch** the curve, mark μ and the value(s) given, and shade the area you want. Say whether you expect the answer to be more or less than 0.5.
2. **Standardise** each value, showing the numbers substituted: $P(X > 55) = P\left(Z > \frac{55-48}{5}\right) = P(Z > 1.4)$. Use σ, not σ^2, and no continuity correction.
3. **Turn the area into** Φ with the table in "Understand it", for example $1 - \Phi(1.4)$.
4. **Look up** and give the answer to 3 significant figures.
5. **Expected number:** multiply the probability (at least 4 figures) by the number in the group and give **one whole number** without "about" or "≈".

For example, $X \sim N(48, 25)$: $P(X > 55) = 1 - \Phi(1.4) = 0.0808$. In a group of 400, expect $400 \times 0.0808 = 32.3$, so 32 of them.

Variations you will meet:

- **Between two values:** standardise **both** ends. Don't assume the two areas are equal unless the values are the same distance from the mean. For example, $P(-1 < Z < 0.8) = \Phi(0.8) + \Phi(1) - 1$.
- **"Within d of the mean"** means $\mu - d < X < \mu + d$, which is symmetrical: $2\Phi\left(\frac{d}{\sigma}\right) - 1$.

- **"Within k standard deviations of the mean"** is $P(-k < Z < k) = 2\Phi(k) - 1$; you don't need μ or σ. **"More than k standard deviations above the mean"** is $P(Z > k)$, and **"more than k standard deviations from the mean"** is **both tails**, $2(1 - \Phi(k))$.
- **Notation:** "$X \sim N(3.5, 0.2^2)$" has $\sigma = 0.2$; "$N(150, 12^2)$" has $\sigma = 12$; "$N(30, 16)$" has $\sigma = 4$.
- **The mean in terms of** σ, such as $\sigma = \mu$: $P(X > 0) = P\left(Z > \frac{0-\mu}{\mu}\right) = P(Z > -1) = \Phi(1) = 0.8413$.
- **Several independent items:** find one item's probability p first. "All four" is p^4; "each of three" is p^3; one item from each of two different distributions, both below a value, is the product of the two probabilities; "fewer than 3 out of 12 days" uses the binomial $B(12, p)$.
- **Classes of items** (small, medium, large): use the percentages to find the boundaries, and use a money value per item to estimate total income, number of items × probability × price for each class.
- **Given as counts**, such as "210 of 3000 bags weighed more than 1.5 kg": the probability is $\frac{210}{3000} = 0.07$.

2. Finding an unknown value, mean or standard deviation (26 questions)

1. **Sketch** and mark the given percentage as an area. Decide which side of the mean the value lies: below the mean means $z < 0$.
2. **Convert to "area to the left"**: "20% are more than x" means $\Phi(z) = 0.8$.
3. **Find** z: from the critical values table if the probability is there ($0.75 \to 0.674$, $0.9 \to 1.282$, $0.95 \to 1.645$, $0.975 \to 1.960$); otherwise from inside the Φ table, to 3 decimal places. Put the minus sign in if the value is below the mean.
4. **Form the equation** $\frac{x-\mu}{\sigma} = z$. The right-hand side must be a z-value, never a probability.
5. **Solve** for the unknown and round as asked.

For example, with $\mu = 42$ and 10% of values below t: $\frac{t-42}{\sigma} = -1.282$. If instead σ is unknown and you know that 15% are above 50: $\Phi(z) = 0.85$, $z = 1.036$, and $\frac{50-42}{\sigma} = 1.036$ gives $\sigma = 7.72$.

Variations you will meet:

- **Find the value** t such that 90% are more than t: t is below the mean, $z = -1.282$, so $t = \mu - 1.282\sigma$.
- **Find** μ given σ and a probability, e.g. $P(X < 3.5) = 0.3$: $z = -0.524$, so $\mu = 3.5 + 0.524\sigma$.
- **A percentage between two values**, such as "65% lie between 48 and k" (with $k > 48$): first find $P(X < 48)$ with part 1's method, add 0.65 to get $P(X < k)$, then find z and solve for k.
- **Boundaries of classes**: "30% of eggs are small" gives the greatest mass of a small one from $\Phi(z) = 0.3$, $z = -0.524$. "45% are medium" with the small percentage known: add the two to get the percentage below the large boundary.
- **Unusual wording**: "on 85% of days more than t" means t is well below the mean, $z = -1.036$.

3. Normal approximation to the binomial (26 questions)

1. **Recognise it**: a count out of n with a fixed probability, n large, and usually the words "use an approximation".

2. **Mean and variance**: $\mu = np$ and $\sigma^2 = npq$, written out, for example $200 \times 0.2 = 40$ and $200 \times 0.2 \times 0.8 = 32$. Keep $\sigma = \sqrt{32}$ exact or to at least 4 figures.
3. **Continuity correction**: write the whole numbers you want, then go 0.5 beyond, e.g. "fewer than 35" is $X \leqslant 34$, so $P(Y < 34.5)$.
4. **Standardise** with $\sqrt{npq}$: $\frac{34.5-40}{\sqrt{32}} = -0.972$.
5. **Find the area** and give the answer to 3 significant figures.

Variations you will meet:

- **"Justify the use of your approximation"**: show np and nq with their values and say both are more than 5, e.g. "$np = 54 > 5$ and $nq = 6 > 5$". It is not enough to say "n is large".
- **The probability comes from somewhere else**: from a percentage in the question, a three-way split (three categories: add the right percentages first), or a normal probability found in an earlier part.
- **Between two numbers**: "more than 12 but fewer than 18" is $13 \leqslant X \leqslant 17$, so $P(12.5 < Y < 17.5)$. "Between 60 and 75 inclusive" is $P(59.5 < Y < 75.5)$.
- **The count of a complicated event**, such as "at least two sixes when three dice are thrown", with p found from the binomial first. Keep p as an exact fraction.
- **Same question, different** n: a small n (like 10) uses the binomial formula exactly; a large n asks for the approximation.

4. Finding both the mean and the standard deviation (9 questions)

1. **Sketch** with both values marked and both percentages shaded, so you can see which side of μ each value is on.
2. **Find the two z-values**, with their signs.
3. **Write two equations**: $x_1 - \mu = z_1\sigma$ and $x_2 - \mu = z_2\sigma$.
4. **Subtract** to remove μ: $x_2 - x_1 = (z_2 - z_1)\sigma$. Solve for σ, then substitute back for μ.
5. **Check**: σ must be positive, and μ must lie on the side of each value that the sketch shows.

For example, if 10% are below 20 and 25% are above 31: $20 - \mu = -1.282\sigma$ and $31 - \mu = 0.674\sigma$. Subtracting, $11 = 1.956\sigma$, so $\sigma = 5.62$ and $\mu = 27.2$.

Variations you will meet:

- **Both values on the same side of the mean**, such as "70% more than 8.2 and 40% less than 9.1": both are below the mean, so both z-values are negative (-0.524 and -0.253).
- **Counts instead of percentages**: "57 of 600 are shorter than 4 cm" gives $P = 0.095$, so $\Phi(z) = 0.095$ and $z = -1.311$.
- **Then a follow-on part** using the μ and σ you found, such as an expected number or a binomial probability.

Worked examples

The bold codes show where marks are earned, as in Cambridge mark schemes. The number is how many marks.

- **M1** method mark: for a correct method, even if the numbers slip.
- **A1** accuracy mark: for a correct answer; it needs the M mark before it.
- **B1** mark for a correct result on its own, with no method needed.

Example 1

The masses of melons sold by a shop are normally distributed with mean 1.20 kg and standard deviation 0.15 kg.

(a) Find the probability that a randomly chosen melon has a mass greater than 1.41 kg. **[3]**

(b) The shop has 300 of these melons. How many of them would you expect to have masses between 1.05 kg and 1.32 kg? **[4]**

(a)

$$P(X > 1.41) = P\left(Z > \frac{1.41 - 1.20}{0.15}\right) = P(Z > 1.4)$$

M1 for standardising with the right numbers.

$= 1 - \Phi(1.4) = 1 - 0.9192$ **M1** for the correct area, $= 0.0808$ **A1**

(b)

$$P(1.05 < X < 1.32) = P\left(\frac{1.05 - 1.20}{0.15} < Z < \frac{1.32 - 1.20}{0.15}\right) = P(-1 < Z < 0.8)$$

M1

$= \Phi(0.8) + \Phi(1) - 1 = 0.7881 + 0.8413 - 1 = 0.6294$ **M1 A1** (the second diagram in "Understand it" shows this area).

Expected number $= 300 \times 0.6294 = 188.8$, so 189 melons. **B1** (any single whole number from a correct probability, 188 or 189, is accepted).

Example 2

The times, in minutes, taken by members of a club to run a course are normally distributed with mean 42.0 and standard deviation σ. 15% of the members take more than 50.0 minutes.

(a) Find the value of σ. **[3]**

(b) 90% of members take more than t minutes. Find the value of t. **[3]**

(a) $P(X > 50) = 0.15$, so $P\left(Z < \frac{50-42}{\sigma}\right) = 0.85$ and $z = 1.036$. **B1**

$$\frac{50-42}{\sigma} = 1.036$$

M1 for the standardised expression equal to a z-value.

$\sigma = \frac{8}{1.036} = 7.72$ **A1**

(b) t is below the mean, so $z = -1.282$ **B1** (critical value).

$$\frac{t-42}{7.722} = -1.282$$

M1, so $t = 42 - 1.282 \times 7.722 = 32.1$ **A1**

Example 3

The lengths of fish in a lake are normally distributed. 10% of the fish are shorter than 20.0 cm and 25% are longer than 31.0 cm. Find the mean and standard deviation of the lengths. **[5]**

z-values: $P(Z < z_1) = 0.1$ gives $z_1 = -1.282$ **B1**; $P(Z < z_2) = 0.75$ gives $z_2 = 0.674$ **B1**.

$$\frac{20-\mu}{\sigma} = -1.282, \qquad \frac{31-\mu}{\sigma} = 0.674$$

M1 for one standardised equation with a z-value.

So $20 - \mu = -1.282\sigma$ and $31 - \mu = 0.674\sigma$. Subtracting: $11 = 1.956\sigma$, so $\sigma = 5.624$. **M1** for solving the pair.

$\mu = 20 + 1.282 \times 5.624 = 27.21$. So $\mu = 27.2, \sigma = 5.62$. **A1**

Example 4

35% of the customers at a café pay with cash. A random sample of 120 customers is chosen.

(a) Use a suitable approximation to find the probability that at least 50 of them pay with cash. **[5]**

(b) Justify the use of your approximation. **[1]**

(a) $X \sim B(120, 0.35)$. Mean $= 120 \times 0.35 = 42$, variance $= 120 \times 0.35 \times 0.65 = 27.3$. **B1**

"At least 50" is $X \geqslant 50$, so with $Y \sim N(42, 27.3)$ use $P(Y > 49.5)$.

$$P(Y > 49.5) = P\left(Z > \frac{49.5-42}{\sqrt{27.3}}\right) = P(Z > 1.435)$$

M1 for standardising with the mean and $\sqrt{\text{variance}}$, **M1** for the continuity correction 49.5.

$= 1 - \Phi(1.435) = 1 - 0.9243$ **M1**, $= 0.0757$ **A1**

(b) $np = 42 > 5$ and $nq = 120 \times 0.65 = 78 > 5$. **B1**

Common mistakes

- **Putting a probability where a z-value belongs.** Writing $\frac{x-\mu}{\sigma} = 0.9$ (or 0.8159, which is $\Phi(0.9)$) earns nothing. Change the probability into a z-value first. Examiners report this every year.
- **Getting the sign of z wrong.** A value below the mean has a negative z. Use a sketch; then check that your answer is on the correct side of μ.
- **Using the variance instead of the standard deviation**, or $\sqrt{\sigma}$. In $N(20, 9)$, $\sigma = 3$. In the approximation, divide by $\sqrt{npq}$, not by npq.
- **A continuity correction on a continuous variable.** Lengths, masses and times need none. Only the binomial approximation uses one.
- **No or wrong continuity correction on a binomial approximation**, or the wrong boundary: "more than 60 " is $P(Y > 60.5)$, not 59.5. Write the whole numbers you want first.
- **Assuming a "between" area is symmetrical** and standardising only one end. Only "within d of the mean" is symmetrical.
- **Not using the critical values.** When the probability is 0.9, 0.95 and so on, use 1.282, 1.645 from the critical values table, not a rougher value found another way.
- **Rounding too early.** Keep z to at least 3 decimal places and probabilities to 4 significant figures until the end; many lose the final mark in approximation questions this way.
- **Stopping at the probability** when the question asks "how many": multiply by the group size and give one whole number.
- **Giving σ to 3 decimal places** instead of 3 significant figures, for example 0.041 instead of 0.0413.
- **Forgetting both tails** in "more than one standard deviation from the mean".

How to get the marks

- **Show the standardisation with numbers in**, such as $\frac{1.41-1.20}{0.15}$. The instructions say unsupported calculator answers earn nothing, and here "full details of the working" are required by the syllabus.
- **Show the area**: write $1 - \Phi(1.4)$ or $\Phi(0.8) + \Phi(1) - 1$ before the value. A sketch is not required but it catches wrong areas.
- **Critical values**: state the z-value you use (1.282, 0.674, ...). It usually earns its own **B1**.
- **Approximations**: write the mean and variance as calculations, show the continuity-corrected value in the standardisation, then the area. That is **B1 M1 M1 M1 A1**.
- **"Justify"**: np and nq with their values, each compared with 5.
- **Accuracy**: 3 significant figures unless told otherwise, from at least 4-figure working. "Correct to 1 decimal place" means exactly that.
- **"How many would you expect"**: one whole number, found from the probability times the total; no "approximately".

- **Two unknowns**: two equations with z-values, a clear method of solving (elimination or substitution), and both values stated.
- **Method marks follow through.** A wrong probability in part (a) used correctly in part (b) still earns the later method marks.

Quick check

1. $X \sim N(30, 16)$. Find $P(26 < X < 36)$.
2. $X \sim N(\mu, 5^2)$ and $P(X < 40) = 0.2$. Find μ.
3. $X \sim B(n, p)$ is approximated by the normal variable Y. Write the continuity-corrected probability for (a) $P(X < 30)$, (b) $P(X \leqslant 30)$, (c) $P(20 \leqslant X < 25)$.
4. Can $B(50, 0.08)$ be approximated by a normal distribution? Give a reason.
5. Heights of adults are normally distributed with mean 170 cm and standard deviation 6 cm. The tallest 5% are taller than h cm. Find h, correct to 1 decimal place.

Answers

1. $\sigma = 4$. $P\left(\frac{26-30}{4} < Z < \frac{36-30}{4}\right) = P(-1 < Z < 1.5) = \Phi(1.5) + \Phi(1) - 1 = 0.9332 + 0.8413 - 1 = 0.775$.
2. $\Phi(z) = 0.2$ gives $z = -0.842$. $\frac{40-\mu}{5} = -0.842$, so $\mu = 40 + 4.21 = 44.2$.
3. (a) $P(Y < 29.5)$; (b) $P(Y < 30.5)$; (c) $20 \leqslant X \leqslant 24$, so $P(19.5 < Y < 24.5)$.
4. No: $np = 50 \times 0.08 = 4$, which is not more than 5 (although $nq = 46 > 5$).
5. $\Phi(z) = 0.95$ gives $z = 1.645$. $\frac{h-170}{6} = 1.645$, so $h = 170 + 9.87 = 179.9$.

Past questions to try

- **9709/53/O/N/24 Q3**: a probability, then a value from a percentage lying between two values.
- **9709/51/M/J/23 Q4**: an expected number, then the mean and standard deviation from two percentages.
- **9709/52/O/N/22 Q6**: a binomial approximation, and justifying it.
- **9709/51/O/N/22 Q4**: finding σ, then several independent people each within a range.