The normal distribution

S4.2 · Statistics · revision notes

The normal distribution $N(\mu, \sigma^2)$ as a model for continuous data: its shape and properties, probabilities and inverse values on your calculator, standardising with Z, and finding an unknown mean or standard deviation. Normal distribution questions are common on Paper 3 and often link to other topics.

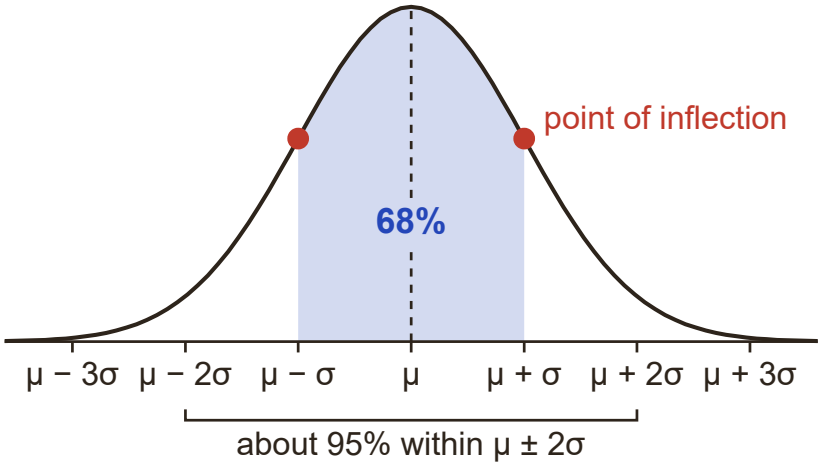
Key facts

- $X \sim N(\mu, \sigma^2)$: σ^2 is the variance
- Points of inflection at $x = \mu \pm \sigma$
- $P(X = a) = 0$, so $P(X < a) = P(X \leq a)$
- $P(X > a) = 1 - P(X < a)$
- $Z = \dfrac{X - \mu}{\sigma}$, $Z \sim N(0, 1^2)$
- About 68% within $\mu \pm \sigma$; 95% within $\mu \pm 2\sigma$; 99.7% within $\mu \pm 3\sigma$
- $P(Z < 1.6449) = 0.95$; $P(Z < 1.9600) = 0.975$; $P(Z < 2.3263) = 0.99$

Notes

Properties

- $X \sim N(\mu, \sigma^2)$ is a continuous distribution with mean μ and variance σ^2 (standard deviation σ).
- The curve is bell-shaped and symmetrical about μ, so mean = median = mode. The total area under it is 1, and area represents probability.
- $P(X = a) = 0$ for a continuous variable, so $P(X < a) = P(X \leq a)$
- The points of inflection are at $x = \mu \pm \sigma$
- About 68% of values lie within $\mu \pm \sigma$, about 95% within $\mu \pm 2\sigma$ and almost all (99.7%) within $\mu \pm 3\sigma$



Normal curve: symmetrical about μ, inflection points at μ ± σ

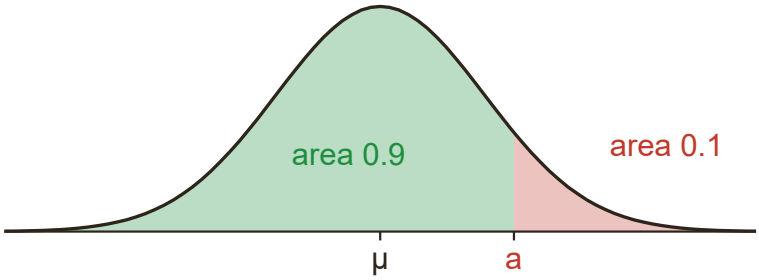
- A normal model suits continuous data that are roughly symmetrical, with most values near the mean, e.g. masses, heights or lengths.

Finding probabilities

- Use your calculator's normal CD with a lower bound, an upper bound, μ and σ. For ‘greater than’, use a very large upper bound.
- The standard normal distribution is $Z \sim N(0, 1^2)$. Standardise with $Z = \dfrac{X - \mu}{\sigma}$
- Use symmetry, e.g. $P(X < \mu - a) = P(X > \mu + a)$

Inverse normal

- Inverse normal gives the value a with $P(X < a) = p$: most calculators use the area to the left, so for $P(X > a) = 0.1$ enter an area of 0.9.



P(X > a) = 0.1 means an area of 0.9 to the left of a

- The percentage points table in the formula booklet gives z values, e.g. $P(Z > 1.6449) = 0.05$

Finding an unknown mean or standard deviation

- Standardise the known value: $\frac{a - \mu}{\sigma} = z$, where z comes from the inverse normal for Z
- One unknown: solve directly. Both unknown: form two equations and solve them simultaneously.
- Use z values to 4 decimal places, negative for values below the mean.

How to answer each type of question

M = method mark, A = answer mark, B = a mark on its own.

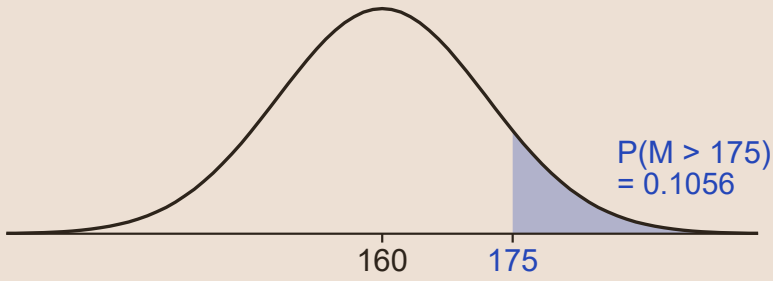
Find normal probabilities

2 to 4 marks | Grade E

1. Write the probability statement, e.g. $P(M > 175)$
2. Use normal CD with μ and σ, or standardise first.
3. Sketch the curve to check whether the answer should be above or below 0.5.

Example. The masses of apples, M grams, are modelled by $M \sim N(160, 12^2)$. Find
(a) $P(M > 175)$
(b) $P(150 < M < 170)$

Model answer.



(a) $P\left(Z > \dfrac{175 - 160}{12}\right) = P(Z > 1.25)$ (M1) $= 0.1056$ (A1)

(b) $P(-0.8333 < Z < 0.8333)$ (M1) $= 0.5953$ (A1)

Use the points of inflection

2 to 3 marks | Grade C

1. The mean is halfway between the points of inflection.
2. The standard deviation is the distance from the mean to either point.
3. Then use normal CD as usual.

Example. The curve of the distribution $X \sim N(\mu, \sigma^2)$ has points of inflection at $x = 42$ and $x = 58$. Find μ and σ, and hence find $P(X < 45)$

Model answer.

$\mu = \dfrac{42 + 58}{2} = 50$ and $\sigma = 58 - 50 = 8$ (B1)

$P(X < 45) = P(Z < -0.625)$ (M1) $= 0.2660$ (A1)

Inverse normal: find a value 3 marks **Grade C**

1. Write the probability statement, e.g. $P(M > m) = 0.05$
2. Change it to an area to the left if your calculator needs one.
3. Use inverse normal directly, or $m = \mu + z\sigma$

Example. Using $M \sim N(160, 12^2)$, the heaviest 5% of apples are graded ‘extra large’. Find the least mass of an ‘extra large’ apple.

Model answer.

$P(M > m) = 0.05$, so $\dfrac{m - 160}{12} = 1.6449$ (M1 B1)

$m = 160 + 1.6449 \times 12$

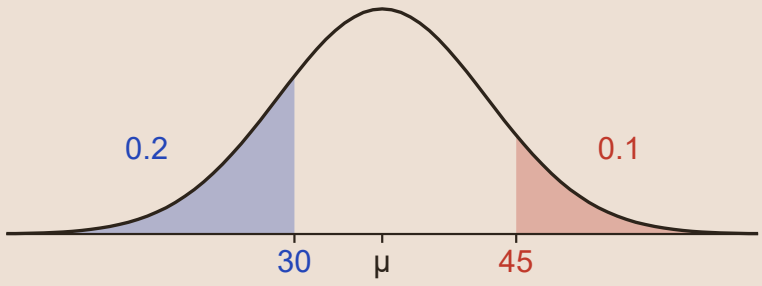
$= 179.7$ g (A1)

Find an unknown mean and standard deviation 5 to 6 marks **Grade A**

1. Write each condition as a probability about Z
2. Find the z values to 4 decimal places (negative below the mean).
3. Form two equations $\dfrac{x - \mu}{\sigma} = z$
4. Solve them simultaneously.

Example. Journey times, T minutes, are modelled by $T \sim N(\mu, \sigma^2)$. Given that $P(T < 30) = 0.2$ and $P(T > 45) = 0.1$, find μ and σ

Model answer.



$\dfrac{30 - \mu}{\sigma} = -0.8416$ (M1 B1) and $\dfrac{45 - \mu}{\sigma} = 1.2816$ (B1)

So $\mu - 0.8416\sigma = 30$ and $\mu + 1.2816\sigma = 45$; subtracting gives $2.1232\sigma = 15$ (M1)

$\sigma = 7.06$ (A1) and $\mu = 35.9$ (A1)

Normal probability used in a binomial model

4 to 5 marks | Grade A*

1. Find the probability for one item from the normal model.
2. Define a binomial random variable for the number of items, using that probability.
3. Find the binomial probability asked for.

Example. Using $M \sim N(160, 12^2)$, apples are packed in bags of 8, chosen at random. Find the probability that more than 2 of the apples in a bag have a mass greater than 175 g.

Model answer.

$p = P(M > 175) = 0.10565$ (B1)

$Y \sim B(8, 0.10565)$, where Y is the number of apples in the bag over 175 g (M1)

$P(Y > 2) = 1 - P(Y \leq 2)$ (M1) $= 0.0439$ (A1)

Don't lose marks

- Using the variance instead of the standard deviation.
- Using the wrong tail with inverse normal (most calculators give the value with the stated area to its left).
- Rounding z values too early in simultaneous equations.
- Using a positive z for a value below the mean.

More tips

Memory tricks

- Sketch and shade every time: it shows whether your answer should be more or less than 0.5, and which tail to use.
- Enter σ, not σ^2: $N(160, 144)$ means $\sigma = 12$
- A value above the mean has a positive z; a value below the mean has a negative z
- Symmetry: $P(\mu - a < X < \mu + a) = 1 - 2P(X > \mu + a)$

Exam technique

- When finding unknowns, show the standardised equation, e.g. $\frac{30 - \mu}{\sigma} = -0.8416$: the method marks depend on it.
- Quote z values to 4 decimal places, from the percentage points table or your calculator.
- To judge whether a normal model is suitable, check that the data are roughly symmetrical (e.g. the mean is close to the median) and that about 95% of values lie within 2 standard deviations of the mean.

What each grade needs

What you need to be able to do, from the first marks up to the top grade.

Grade E

Describe the shape and properties of the normal distribution

Bell-shaped and symmetrical about the mean μ, with mean = median = mode, and total area 1.

Grade E — **Find normal probabilities using a calculator**

Use normal CD with μ and σ (not σ^2).

Grade D — **Use the approximate 68%, 95% and 99.7% rules**

About 68% (roughly two-thirds) of values lie within $\mu \pm \sigma$, 95% within $\mu \pm 2\sigma$ and almost all within $\mu \pm 3\sigma$

Grade C — **Find a value from a given probability**

Use inverse normal, remembering that most calculators work with the area to the left.

Grade C — **Use the points of inflection at $\mu \pm \sigma$**

The mean is halfway between the points of inflection and σ is the distance from the mean to either one.

Grade B — **Find an unknown mean or standard deviation**

Standardise: $\frac{x - \mu}{\sigma} = z$, with z from the inverse normal for $Z \sim N(0, 1)$

Grade A — **Find μ and σ using simultaneous equations**

Form two standardised equations and solve them together.

Grade A* — **Combine normal and binomial models in multi-stage problems**

E.g. find $p = P(X > a)$ from a normal model, then use $B(n, p)$

Free notes, Full packs, practice papers and mock exams at chhetriacademy.com. Written for this site; not taken from past papers. Stuck? 1-to-1 tutoring with a free 30-minute trial lesson.