Mathematical Methods Units 3 & 4

Chapter 14 — The normal distribution

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Attribution: Classbasics Pty Ltd, vwx.au

Aboriginal and Torres Strait Islander content has been excluded as accuracy cannot be guaranteed, and it is better to be respectful than cover the entire curriculum inaccurately.

Significant effort has been made to ensure this content is legal. Please send any areas of concern to admin@vwx.au with appropriate document/legal references so errors can be verified and changes be made.

Contents

Theory

The **normal distribution** is the most important continuous distribution in statistics. A normally distributed random variable is written $X \sim N(\mu, \sigma^2)$, where μ is the **mean** and σ is the **standard deviation** (so σ^2 is the variance).

The bell-shaped curve. The probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

a smooth "bell" curve. You are **not** asked to integrate it by hand — its key features come from μ and σ.

Key features.

- The curve is **symmetric about the vertical line** $x = \mu$. Because of this symmetry the **mean, median and mode are all equal to** μ.
- The total area under the curve is 1, and by symmetry the area on each side of μ is $\frac{1}{2}$; that is,

  $Pr(X > \mu) = Pr(X < \mu) = \frac{1}{2}$.
- The curve has **points of inflection at** $x = \mu - \sigma$ **and** $x = \mu + \sigma$ — one standard deviation either side of the mean. (Reading the inflection points off a graph is a quick way to find σ.)
- The curve extends indefinitely in both directions, approaching the x-axis as an asymptote but never touching it.

The effect of μ and σ.

- μ fixes the **location** (centre) of the curve. Changing μ slides the whole curve left or right without changing its shape.
- σ fixes the **spread**. A **larger** σ gives a **wider, flatter** curve; a **smaller** σ gives a **taller, narrower** curve. (The total area stays 1, so a narrower curve must be taller.)

Using symmetry to find probabilities. Even before computing exact probabilities with technology (Section 14C), symmetry gives many results. If $a > 0$, then by the symmetry about μ:

$$Pr(X < \mu - a) = Pr(X > \mu + a), \quad Pr(\mu - a < X < \mu + a) = 1 - 2\,Pr(X < \mu - a).$$

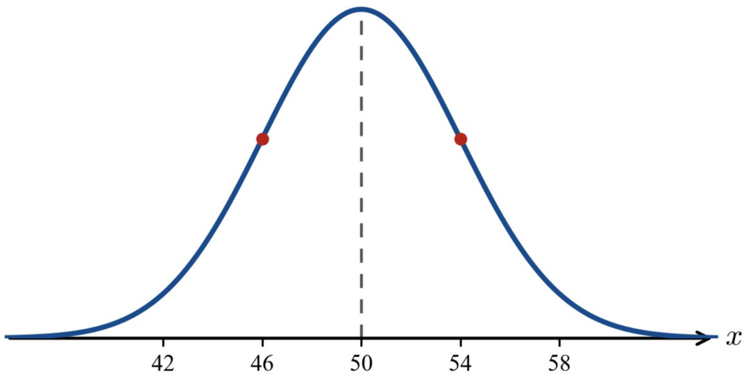
Worked examples

Example 1 — scaffolded

The random variable $X \sim N(50, 4^2)$. State the mean, median and mode, the coordinates (on the x-axis) of the points of inflection, and the value of $Pr(X > 50)$.

Working	Comment
Here $\mu = 50$ and $\sigma = 4$.	Read μ and σ from $N(\mu, \sigma^2)$; the second entry is $\sigma^2 = 16$.
Mean = median = mode = 50**.**	For a normal distribution all three central measures equal μ.
Points of inflection at $x = \mu \pm \sigma = 50 \pm 4$, i.e. $x = 46$ and $x = 54$.	The inflection points are one standard deviation either side of the mean.

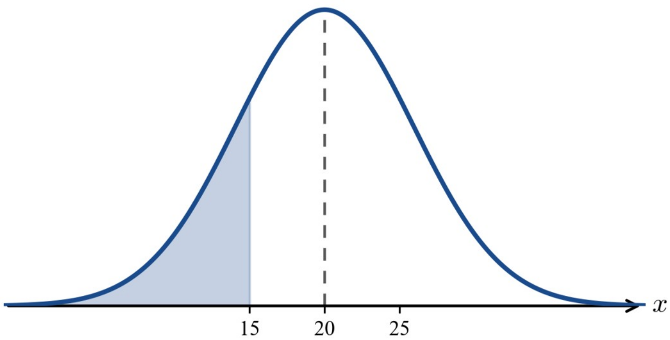
Working	Comment
$Pr(X>50)=\frac{1}{2}$.	The curve is symmetric about $x=\mu=50$, so half the area lies to the right.



Example 2 — scaffolded

X is normally distributed with mean $\mu=20$, and $Pr(X<15)=0.2$. Using symmetry, find (a) $Pr(X>25)$, (b) $Pr(15<X<25)$, and (c) $Pr(X>15)$.

Working	Comment
15 and 25 are equidistant from $\mu=20$ (each 5 units away).	Set up the symmetry: $15=\mu-5$ and $25=\mu+5$.
(a) $Pr(X>25)=Pr(X<15)=0.2$.	By symmetry, the right tail beyond $\mu+5$ equals the left tail below $\mu-5$.
(b) $Pr(15<X<25)=1-Pr(X<15)-Pr(X>25)=1-0$.	The middle area is 1 minus the two equal tails.
(c) $Pr(X>15)=1-Pr(X<15)=1-0.2=0.8$.	Complementary events.



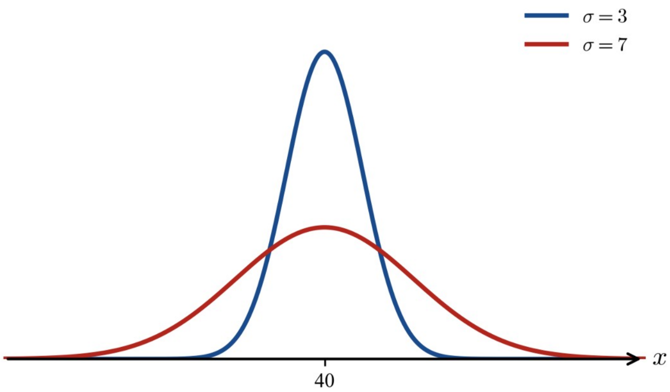
Example 3 — unscaffolded

Two classes sit the same test. The marks of class A are modelled by $A\sim N(65,5^2)$ and those of class B by $B\sim N(65,10^2)$. Compare the two distributions.

Both distributions have the **same mean**, $\mu=65$, so both curves are centred on 65 and have mean = median = mode = 65.

Class B has the larger standard deviation ($\sigma_B=10$ versus $\sigma_A=5$), so the marks of class B are **more spread out**: its bell curve is **wider and flatter**, while class A's is **taller and narrower** (its marks cluster more tightly around 65).

$\therefore$ same centre, but class B shows greater variability.



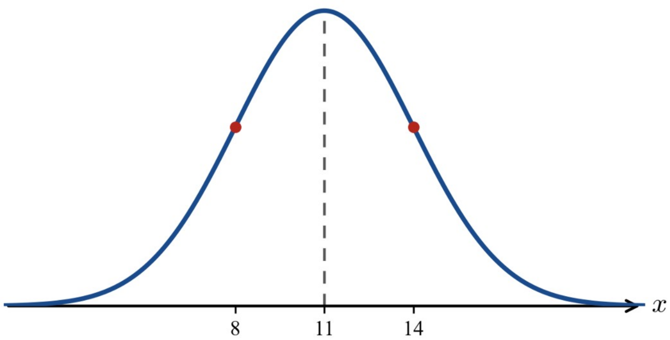
Example 4 — unscaffolded

The graph of a normal distribution has points of inflection at $x=8$ and $x=14$. Find the mean μ and the standard deviation σ.

The inflection points are at $\mu-\sigma$ and $\mu+\sigma$, so μ is midway between them and σ is half the distance between them:

$$\mu=\frac{8+14}{2}=11, \sigma=\frac{14-8}{2}=3.$$

$\therefore X\sim N(11,3^2)$.



Practice questions

Scaffolded

Q1. The heights (cm) of a large group are modelled by $X\sim N(100,15^2)$. (a) State the mean, median and mode. (b) Write the x-coordinates of the points of inflection. (c) State $Pr(X<100)$. (d) Between which two values does the "middle" one standard deviation either side of the mean lie?

Q2. $X\sim N(30,2^2)$. (a) State μ and σ. (b) Give the points of inflection. (c) State $Pr(X>30)$.

Q3. X is normal with mean 40 and $Pr(X<34)=0.15$. (a) By symmetry, state $Pr(X>46)$. (b) Find $Pr(34<X<46)$. (c) Find $Pr(X>34)$.

Q4. A normal curve has points of inflection at $x=27$ and $x=33$. (a) Find the mean. (b) Find the standard deviation. (c) Write the distribution in the form $N(\mu,\sigma^2)$.

Q5. $X\sim N(12,\sigma^2)$ and $Y\sim N(20,\sigma^2)$ have the same standard deviation. (a) How are the two curves related in shape? (b) How are they related in position?

Q6. For $X \sim N(50, 5^2)$, use symmetry to find: (a) $Pr(X<50)$; (b) $Pr(X>55 \text{ or } X<45)$ given $Pr(X>55)=0.16$; (c) $Pr(45<X<55)$.

Unscaffolded

Q7. For $X \sim N(25, 2^2)$, state the mean, the points of inflection, and $Pr(X>25)$.

Q8. A normal distribution has mean 70 and standard deviation 6. Between which two x-values are its points of inflection?

Q9. The points of inflection of a normal curve are at $x=4.5$ and $x=7.5$. Find μ and σ.

Q10. $X \sim N(60, 8^2)$. Given $Pr(60<X<68)=0.34$, use symmetry to find $Pr(X>68)$ and $Pr(52<X<68)$.

Q11. X is normal with mean 15 and $Pr(X<11)=0.25$. Find $Pr(X>19)$, $Pr(11<X<19)$ and $Pr(X>11)$.

Q12. Two normal curves have the same standard deviation but means $\mu_1=30$ and $\mu_2=45$. Describe, with reasons, how the graphs differ.

Q13. Two normal curves have the same mean 40 but standard deviations $\sigma_1=3$ and $\sigma_2=7$. Which curve is taller at its centre, and why?

Q14. A normal distribution is symmetric about $x=100$ and $Pr(X>118)=0.1$. Find $Pr(X<82)$ and $Pr(82<X<118)$.

Q15. For $X \sim N(\mu, \sigma^2)$, explain why $Pr(X=\mu)=0$ even though $x=\mu$ is the most likely region.

Q16. A machine fills bottles with a mean of 500 mL. The filling volume is normal with points of inflection at 498 mL and 502 mL. State the standard deviation and $Pr(X<500)$.

Q17. Sketch, on one set of axes, the curves $N(50, 4^2)$ and $N(50, 8^2)$, and describe how they differ.

Q18. X is normal with mean μ. Given $Pr(X<\mu-3)=0.12$, write down $Pr(X>\mu+3)$ and $Pr(\mu-3<X<\mu+3)$.

Q19. The marks in an exam are modelled by $N(62, 11^2)$. State the mean mark and the two marks at which the distribution has its points of inflection.

Q20. A normal curve has its maximum at $x=8$ and a point of inflection at $x=11$. Find μ and σ.

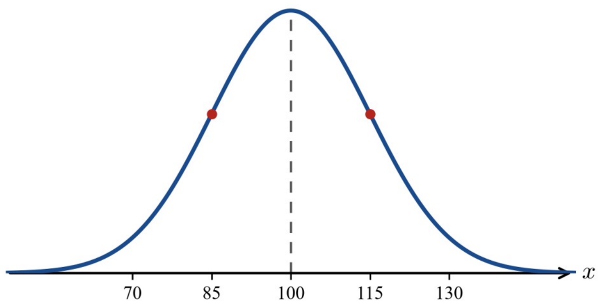
Q21. $X \sim N(20, \sigma^2)$ satisfies $Pr(X<14)=Pr(X>26)$. Explain why this must be true for **any** value of σ, and state the common value in terms of $p=Pr(X<14)$.

Answers

Q1. (a) 100; (b) 85 and 115; (c) $\frac{1}{2}$; (d) 85 to 115. **Q2.** (a) $\mu=30$, $\sigma=2$; (b) 28 and 32; (c) $\frac{1}{2}$. **Q3.** (a) 0.15; (b) 0.7; (c) 0.85. **Q4.** (a) 30; (b) 3; (c) $N(30,3^2)$. **Q5.** (a) identical shape (same σ); (b) Y's curve is Y translated 8 units right of X's. **Q6.** (a) 0.5; (b) 0.32; (c) 0.68. **Q7.** mean 25; inflection 23 and 27; $Pr(X>25)=\frac{1}{2}$. **Q8.** 64 and 76. **Q9.** $\mu=6$, $\sigma=1.5$. **Q10.** $Pr(X>68)=0.16$; $Pr(52<X<68)=0.68$. **Q11.** $Pr(X>19)=0.25$; $Pr(11<X<19)=0.5$; $Pr(X>11)=0.75$. **Q12.** identical shape; the second curve is the first translated 15 units to the right. **Q13.** the $\sigma_1=3$ curve is taller (smaller spread ⇒ narrower and, since total area $=1$, taller). **Q14.** $Pr(X<82)=0.1$; $Pr(82<X<118)=0.8$. **Q15.** for a continuous variable the probability of any single exact value is the area of a line, which is 0. **Q16.** $\sigma=2$ mL; $Pr(X<500)=\frac{1}{2}$. **Q17.** same centre 50; $N(50,8^2)$ is wider and flatter (twice the spread). **Q18.** $Pr(X>\mu+3)=0.12$; $Pr(\mu-3<X<\mu+3)=0.76$. **Q19.** mean 62; inflection at 51 and 73. **Q20.** $\mu=8$, $\sigma=3$. **Q21.** 14 and 26 are equidistant from $\mu=20$, so symmetry forces the tails equal for every σ; common value $=p$.

Fully-worked solutions

Q1. $X\sim N(100,15^2)$, so $\mu=100$, $\sigma=15$. (a) Mean = median = mode = 100. (b) Inflection at $\mu\pm\sigma=100\pm15$, i.e. 85 and 115. (c) By symmetry $Pr(X<100)=\frac{1}{2}$. (d) One standard deviation either side of the mean is from 85 to 115.



Q2. $\mu=30$, $\sigma=2$. (a) as stated. (b) Inflection at 30 ± 2, i.e. 28 and 32. (c) $Pr(X>30)=\frac{1}{2}$ by symmetry about $\mu=30$.

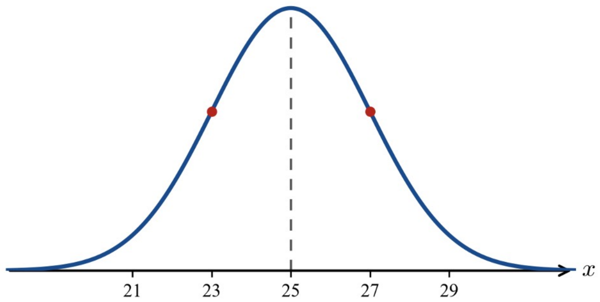
Q3. Mean 40; $34=\mu-6$ and $46=\mu+6$. (a) $Pr(X>46)=Pr(X<34)=0.15$ by symmetry. (b) $Pr(34<X<46)=1-0.15-0.15=0.7$. (c) $Pr(X>34)=1-0.15=0.85$.

Q4. Inflection at $\mu\pm\sigma$. (a) $\mu=\frac{27+33}{2}=30$. (b) $\sigma=\frac{33-27}{2}=3$. (c) $N(30,3^2)$.

Q5. Same σ, so (a) the curves have **identical shape**. (b) The means differ by $20-12=8$, so Y's curve is X's curve translated 8 units to the right.

Q6. $X\sim N(50,5^2)$. (a) $Pr(X<50)=\frac{1}{2}$. (b) $45=\mu-5$ and $55=\mu+5$; by symmetry $Pr(X<45)=Pr(X>55)=0.16$, so $Pr(X>55 \text{ or } X<45)=0.16+0.16=0.32$. (c) $Pr(45<X<55)=1-0.32=0.68$.

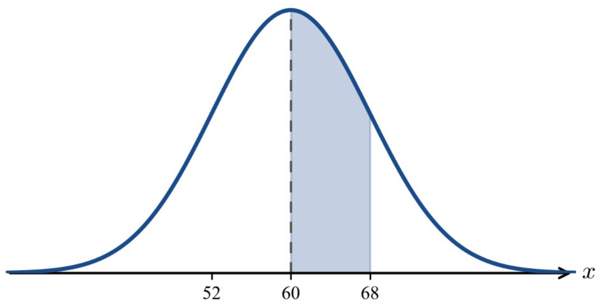
Q7. $\mu=25$, $\sigma=2$: mean 25; inflection $25\pm2=23$ and 27; $Pr(X>25)=\frac{1}{2}$.



Q8. Inflection at $\mu \pm \sigma = 70 \pm 6$, i.e. 64 and 76.

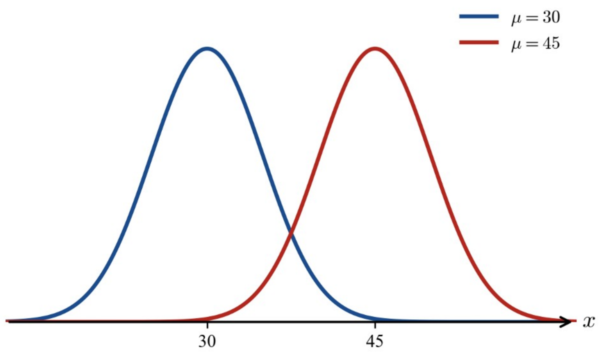
Q9. $\mu = \dfrac{4.5 + 7.5}{2} = 6$; $\sigma = \dfrac{7.5 - 4.5}{2} = 1.5$.

Q10. $X \sim N(60, 8^2)$, symmetric about 60. Since $Pr(X > 60) = 0.5$ and $Pr(60 < X < 68) = 0.34$, $Pr(X > 68) = 0.5 - 0.34 = 0.16$. By symmetry $Pr(52 < X < 60) = 0.34$, so $Pr(52 < X < 68) = 0.34 + 0.34 = 0.68$.



Q11. Mean 15; $11 = \mu - 4$, $19 = \mu + 4$. $Pr(X > 19) = Pr(X < 11) = 0.25$; $Pr(11 < X < 19) = 1 - 0.25 - 0.25 = 0.5$; $Pr(X > 11) = 1 - 0.25 = 0.75$.

Q12. Same $\sigma \Rightarrow$ identical shape; the means differ by $45 - 30 = 15$, so the second curve is the first translated 15 units to the right (same height and width).



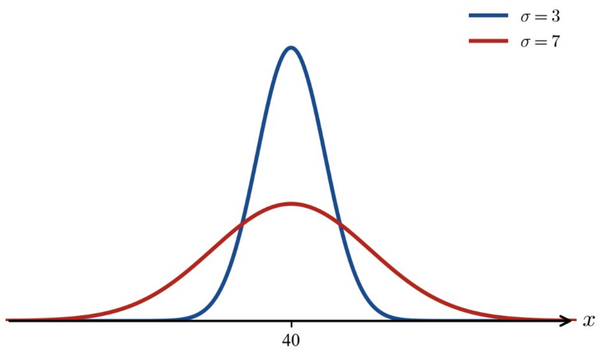
Q13. Both are centred on 40. The curve with $\sigma_1 = 3$ has the smaller spread, so it is narrower; because the area under each curve is 1, a narrower curve must be **taller**. Hence the $\sigma = 3$ curve is taller at its centre.

Q14. Symmetric about 100; $118 = \mu + 18$ and $82 = \mu - 18$. By symmetry $Pr(X < 82) = Pr(X > 118) = 0.1$; then $Pr(82 < X < 118) = 1 - 0.1 - 0.1 = 0.8$.

Q15. X is continuous, so probability is area under the density curve. The “area” above a single point $x = \mu$ is a line segment of zero width, hence zero area, so $Pr(X = \mu) = 0$; probability accrues only over an interval.

Q16. Inflection at $\mu \pm \sigma$: $\mu = \frac{498+502}{2} = 500$ and $\sigma = \frac{502-498}{2} = 2$ mL. $Pr(X < 500) = \frac{1}{2}$ by symmetry.

Q17. Both curves are centred on $x = 50$ (same mean), with inflection points at $46, 54$ for $N(50, 4^2)$ and at $42, 58$ for $N(50, 8^2)$. The $\sigma = 8$ curve is **wider and flatter** (twice the spread) while the $\sigma = 4$ curve is taller and narrower.



Q18. $\mu - 3$ and $\mu + 3$ are equidistant from μ, so by symmetry $Pr(X > \mu + 3) = Pr(X < \mu - 3) = 0.12$; then $Pr(\mu - 3 < X < \mu + 3) = 1 - 0.12 - 0.12 = 0.76$.

Q19. $\mu = 62$, $\sigma = 11$: mean 62; inflection at $62 \pm 11 = 51$ and 73.

Q20. The maximum of a normal curve is at the mean, so $\mu = 8$. A point of inflection is at $\mu + \sigma$ (or $\mu - \sigma$): $11 = 8 + \sigma \Rightarrow \sigma = 3$.

Q21. $14 = 20 - 6$ and $26 = 20 + 6$ are the same distance (6) from the mean $\mu = 20$. The normal curve is symmetric about $x = \mu$ for **every** σ, so the left tail $Pr(X < 20 - 6)$ always equals the right tail $Pr(X > 20 + 6)$, regardless of σ. The common value is $p = Pr(X < 14)$.

Theory

Every normal distribution is symmetric and bell-shaped, centred on its mean μ, with its spread governed by the standard deviation σ. We write $X \sim N(\mu, \sigma^2)$.

The standard normal. The normal distribution with mean 0 and standard deviation 1 is called the **standard normal distribution**, and its variable is written $Z \sim N(0, 1)$.

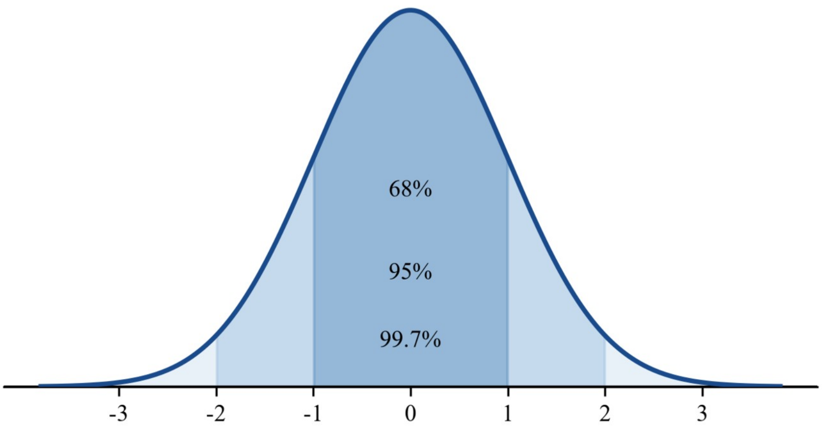
Standardisation. Any normal variable $X \sim N(\mu, \sigma^2)$ can be converted to the standard normal by

$$Z = \frac{X - \mu}{\sigma}.$$

The value z is the z**-score** (or standard score): it measures how many standard deviations x lies from the mean. A positive z is above the mean, a negative z below it. To go back, $x = \mu + z\sigma$.

The 68–95–99.7% rule. For any normal distribution, approximately:

- **68%** of values lie within **one** standard deviation of the mean ($\mu - \sigma$ to $\mu + \sigma$);
- **95%** lie within **two** standard deviations ($\mu - 2\sigma$ to $\mu + 2\sigma$);
- **99.7%** lie within **three** standard deviations ($\mu - 3\sigma$ to $\mu + 3\sigma$).



The 68–95–99.7% rule for the standard normal

Consequences (using symmetry). Because the curve is symmetric, the central percentages split evenly on each side. The most useful values are:

Region	Probability
$Pr(X < \mu)$	0.5
$Pr(\mu - \sigma < X < \mu + \sigma)$	0.68
$Pr(\mu < X < \mu + \sigma)$	0.34
$Pr(X > \mu + \sigma) = Pr(X < \mu - \sigma)$	0.16
$Pr(X < \mu + \sigma)$	0.84
$Pr(\mu + \sigma < X < \mu + 2\sigma)$	0.135
$Pr(X > \mu + 2\sigma) = Pr(X < \mu - 2\sigma)$	0.025
$Pr(X < \mu + 2\sigma)$	0.975

Region	Probability
$Pr(X > \mu + 3\sigma)$	0.0015

Combining these gives any probability whose boundaries fall on whole numbers of standard deviations — for example $Pr(\mu - \sigma < X < \mu + 2\sigma) = 0.34 + 0.475 = 0.815$.

Worked examples

Example 1 — scaffolded

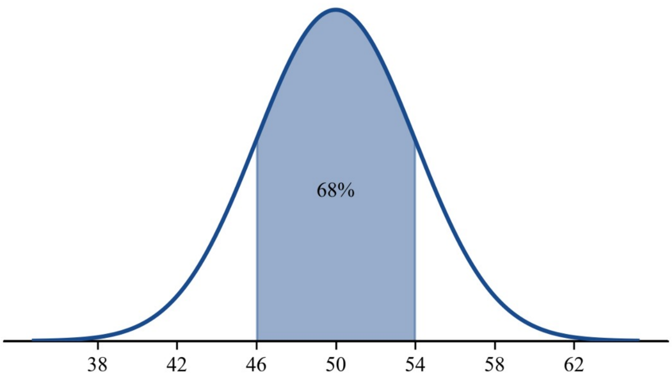
The heights of a large group have distribution $X \sim N(100, 15^2)$. Find the z-score of each value and interpret it: (a) $x = 115$; (b) $x = 70$; (c) $x = 130$.

Working	Comment
Here $\mu = 100$, $\sigma = 15$, and $z = \dfrac{x - \mu}{\sigma}$.	Identify the mean and standard deviation.
(a) $z = \dfrac{115 - 100}{15} = 1$.	115 is **1 standard deviation above** the mean.
(b) $z = \dfrac{70 - 100}{15} = -2$.	70 is **2 standard deviations below** the mean.
(c) $z = \dfrac{130 - 100}{15} = 2$.	130 is **2 standard deviations above** the mean.

Example 2 — scaffolded

For $X \sim N(50, 4^2)$, use the 68–95–99.7% rule to find: (a) $Pr(46 < X < 54)$; (b) $Pr(X > 58)$; (c) $Pr(42 < X < 58)$.

Working	Comment
$\mu = 50$, $\sigma = 4$.	Note the boundaries in terms of σ.
(a) $46 = \mu - \sigma$ and $54 = \mu + \sigma$, so $Pr(46 < X < 54) = 0.68$.	Within one standard deviation.
(b) $58 = \mu + 2\sigma$, so $Pr(X > 58) = Pr(X > \mu + 2\sigma) = 0.025$.	Upper tail beyond two standard deviations.
(c) $42 = \mu - 2\sigma$, $58 = \mu + 2\sigma$, so $Pr(42 < X < 58) = 0.95$.	Within two standard deviations.



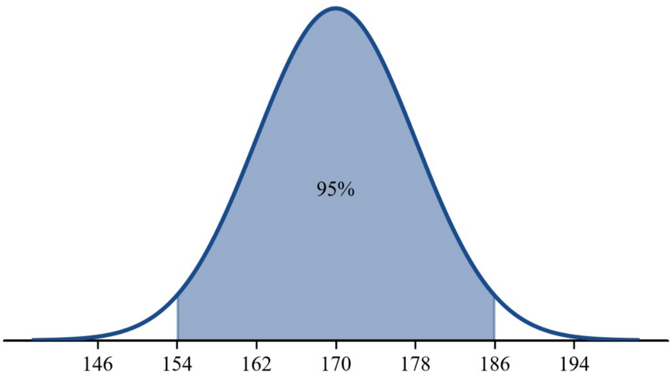
Example 3 — unscaffolded

The lengths (mm) of a component are $X \sim N(170, 8^2)$. Find $Pr(154 < X < 186)$ and $Pr(X < 162)$.

$\mu = 170$, $\sigma = 8$. Now $154 = 170 - 2(8) = \mu - 2\sigma$ and $186 = 170 + 2(8) = \mu + 2\sigma$, so $Pr(154 < X < 186) = 0.95$.

Also $162 = 170 - 8 = \mu - \sigma$, so $Pr(X < 162) = Pr(X < \mu - \sigma) = 0.16$.

$\therefore Pr(154 < X < 186) = 0.95$ and $Pr(X < 162) = 0.16$.

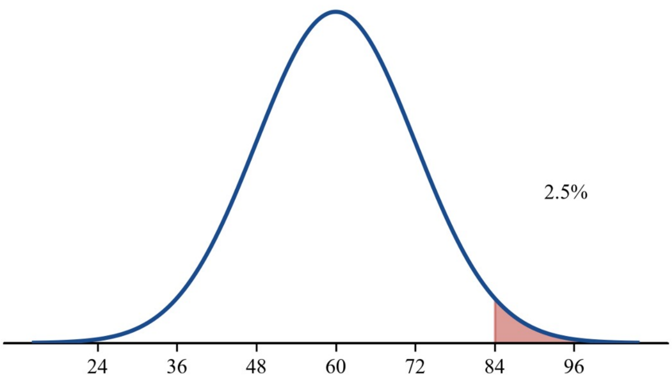


Example 4 — unscaffolded

Exam marks are $X \sim N(60, 12^2)$. The top 2.5% of students receive a distinction. Find the lowest mark that earns a distinction.

The top 2.5% satisfies $Pr(X > x) = 0.025$. Since $Pr(X > \mu + 2\sigma) = 0.025$, the cut-off is at $x = \mu + 2\sigma = 60 + 2(12) = 84$.

$\therefore$ a mark of 84 (or above) earns a distinction.



Practice questions

Scaffolded

Q1. Scores are $X \sim N(500, 100^2)$. (a) Find the z-score of $x = 600$. (b) Find the z-score of $x = 300$. (c) Find the score with a z-score of 1.5. (d) Interpret the z-score in part (a).

Q2. For $X \sim N(25, 3^2)$, use the 68–95–99.7% rule to find: (a) $Pr(22 < X < 28)$; (b) $Pr(19 < X < 31)$; (c) $Pr(X > 34)$.

Q3. For $X \sim N(1000, 50^2)$, find: (a) $Pr(950 < X < 1050)$; (b) $Pr(X < 900)$; (c) $Pr(850 < X < 1150)$.

Q4. For $X \sim N(20, 5^2)$, find the z-score of: (a) $x = 30$; (b) $x = 15$; (c) $x = 35$.

Q5. For $X \sim N(80, 10^2)$, find: (a) $Pr(80 < X < 90)$; (b) $Pr(90 < X < 100)$; (c) $Pr(X > 100)$.

Q6. For $X \sim N(200, 20^2)$: (a) find the value exceeded by 16% of the distribution; (b) find the value exceeded by 2.5% of the distribution; (c) between which two values does the middle 68% lie?

Unscaffolded

Q7. For $X \sim N(70, 6^2)$, find the z-scores of $x = 82$ and $x = 64$.

Q8. For $X \sim N(150, 25^2)$, find $Pr(125 < X < 175)$ and $Pr(X > 200)$.

Q9. For $X \sim N(40, 5^2)$, find $Pr(30 < X < 50)$ and $Pr(X < 35)$.

Q10. For $X \sim N(12, 2^2)$, find $Pr(X > 18)$ and $Pr(8 < X < 16)$.

Q11. Student A scored 75 on a test with $X \sim N(60, 10^2)$; student B scored 82 on a different test with $Y \sim N(70, 8^2)$. By comparing z-scores, determine who performed better relative to their group.

Q12. For $X \sim N(500, 80^2)$: (a) find the value with a z-score of -1.5; (b) find the value exceeded by 84% of the distribution.

Q13. For $X \sim N(100, 20^2)$, find $Pr(100 < X < 120)$, $Pr(60 < X < 80)$ and $Pr(X < 40)$.

Q14. For $X \sim N(25, 4^2)$, between which two values does the middle 95% lie, and what is $Pr(X > 29)$?

Q15. A normal variable X has $Pr(X < 50) = 0.16$ and $Pr(X > 80) = 0.16$. Find its mean μ and standard deviation σ.

Q16. For the standard normal $Z \sim N(0, 1)$, find $Pr(Z < 1)$, $Pr(Z > 2)$ and $Pr(-1 < Z < 2)$.

Q17. For $X \sim N(60, 15^2)$, find $Pr(45 < X < 90)$ and $Pr(X > 75)$.

Q18. Bags of sugar have mass $X \sim N(500, 10^2)$ grams. What percentage of bags have a mass between 490 g and 520 g, and what percentage are underweight (less than 480 g)?

Q19. A normal variable X has $Pr(X < 20) = 0.025$ and $Pr(X < 50) = 0.975$. Find μ and σ.

Q20. A normal variable X has mean 70 and $Pr(X > 85) = 0.0015$. Show that $\sigma = 5$.

Q21. The lifetimes of a brand of globe are $X \sim N(1000, 50^2)$ hours. (a) Between which two values do the middle 99.7% of lifetimes lie? (b) Find the proportion of globes that last more than 1100 hours. (c) Hence estimate how many of a batch of 2000 globes last less than 900 hours.

Answers

Q1. (a) 1 (b) -2 (c) 650 (d) 600 is one standard deviation above the mean. **Q2.** (a) 0.68 (b) 0.95 (c) 0.0015. **Q3.** (a) 0.68 (b) 0.025 (c) 0.997. **Q4.** (a) 2 (b) -1 (c) 3. **Q5.** (a) 0.34 (b) 0.135 (c) 0.025. **Q6.** (a) 220 (b) 240 (c) 180 to 220. **Q7.** 2 and -1. **Q8.** 0.68 and 0.025. **Q9.** 0.95 and 0.16. **Q10.** 0.0015 and 0.95. **Q11.** Both have $z=1.5$, so they performed **equally well** relative to their groups. **Q12.** (a) 380 (b) 420. **Q13.** 0.34, 0.135 and 0.0015. **Q14.** 17 to 33; $Pr(X>29)=0.16$. **Q15.** $\mu=65$, $\sigma=15$. **Q16.** 0.84, 0.025 and 0.815. **Q17.** 0.815 and 0.16. **Q18.** 81.5% between 490 and 520 g; 2.5% underweight. **Q19.** $\mu=35$, $\sigma=7.5$. **Q20.** $\sigma=5$. **Q21.** (a) 850 to 1150 hours (b) 0.025 (c) about 50 globes.

Fully-worked solutions

Q1. $\mu=500$, $\sigma=100$. (a) $z=\dfrac{600-500}{100}=1$. (b) $z=\dfrac{300-500}{100}=-2$. (c) $x=\mu+z\sigma=500+1.5(100)=650$. (d) A mark of 600 is one standard deviation above the mean.

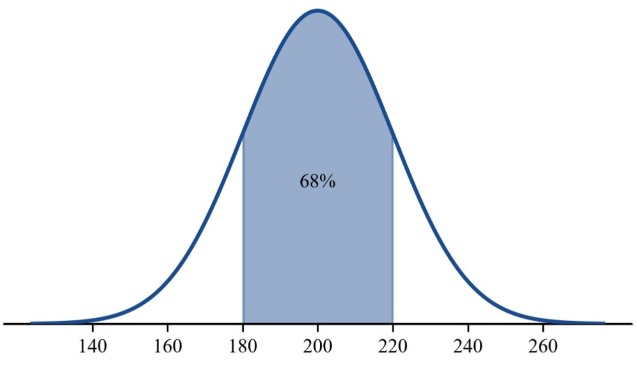
Q2. $\mu=25$, $\sigma=3$. (a) $22=\mu-\sigma$, $28=\mu+\sigma$, so $Pr(22<X<28)=0.68$. (b) $19=\mu-2\sigma$, $31=\mu+2\sigma$, so $Pr(19<X<31)=0.95$. (c) $34=\mu+3\sigma$, so $Pr(X>34)=Pr(X>\mu+3\sigma)=0.0015$.

Q3. $\mu=1000$, $\sigma=50$. (a) $950=\mu-\sigma$, $1050=\mu+\sigma$: $Pr=0.68$. (b) $900=\mu-2\sigma$: $Pr(X<900)=0.025$. (c) $850=\mu-3\sigma$, $1150=\mu+3\sigma$: $Pr=0.997$.

Q4. $\mu=20$, $\sigma=5$. (a) $z=\dfrac{30-20}{5}=2$. (b) $z=\dfrac{15-20}{5}=-1$. (c) $z=\dfrac{35-20}{5}=3$.

Q5. $\mu=80$, $\sigma=10$. (a) $80=\mu$, $90=\mu+\sigma$, so $Pr(80<X<90)=\dfrac{0.68}{2}=0.34$. (b) $90=\mu+\sigma$, $100=\mu+2\sigma$, so $Pr=\dfrac{0.95-0.68}{2}=0.135$. (c) $100=\mu+2\sigma$, so $Pr(X>100)=0.025$.

Q6. $\mu=200$, $\sigma=20$. (a) $Pr(X>x)=0.16$ at $x=\mu+\sigma=220$. (b) $Pr(X>x)=0.025$ at $x=\mu+2\sigma=240$. (c) The middle 68% lies between $\mu-\sigma=180$ and $\mu+\sigma=220$.



Q7. $\mu=70$, $\sigma=6$. $z=\dfrac{82-70}{6}=2$; $z=\dfrac{64-70}{6}=-1$.

Q8. $\mu=150$, $\sigma=25$. $125=\mu-\sigma$, $175=\mu+\sigma$: $Pr(125<X<175)=0.68$. $200=\mu+2\sigma$: $Pr(X>200)=0.025$.

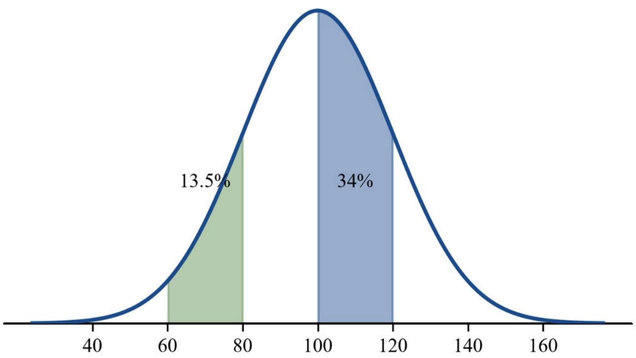
Q9. $\mu=40$, $\sigma=5$. $30=\mu-2\sigma$, $50=\mu+2\sigma$: $Pr=0.95$. $35=\mu-\sigma$: $Pr(X<35)=0.16$.

Q10. $\mu=12$, $\sigma=2$. $18=\mu+3\sigma$: $Pr(X>18)=0.0015$. $8=\mu-2\sigma$, $16=\mu+2\sigma$: $Pr(8<X<16)=0.95$.

Q11. Student A: $z=\frac{75-60}{10}=1.5$. Student B: $z=\frac{82-70}{8}=1.5$. The z-scores are equal, so relative to their respective groups the two students performed **equally well**.

Q12. $\mu=500$, $\sigma=80$. (a) $x=\mu+z\sigma=500+(-1.5)(80)=380$. (b) $Pr(X>x)=0.84$ means x is below the mean at $Pr(X>\mu-\sigma)=0.84$, so $x=\mu-\sigma=420$.

Q13. $\mu=100$, $\sigma=20$. $Pr(100<X<120)=Pr(\mu<X<\mu+\sigma)=0.34$.
$Pr(60<X<80)=Pr(\mu-2\sigma<X<\mu-\sigma)=0.135$. $40=\mu-3\sigma$, so $Pr(X<40)=0.0015$.



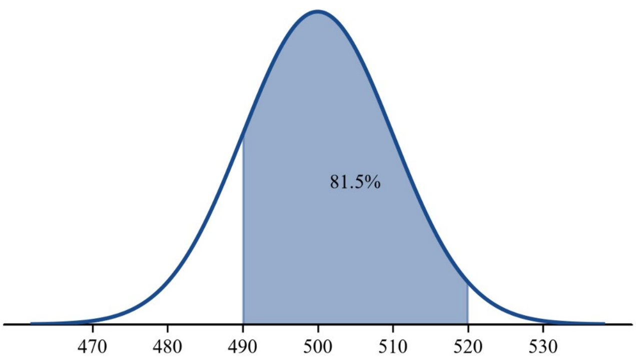
Q14. $\mu=25$, $\sigma=4$. The middle 95% lies within 2σ: from $\mu-2\sigma=17$ to $\mu+2\sigma=33$. Also $29=\mu+\sigma$, so $Pr(X>29)=0.16$.

Q15. By symmetry, $Pr(X<50)=Pr(X>80)=0.16$ means $50=\mu-\sigma$ and $80=\mu+\sigma$. Adding: $2\mu=130$, so $\mu=65$; subtracting: $2\sigma=30$, so $\sigma=15$.

Q16. For $Z\sim N(0,1)$: $Pr(Z<1)=0.84$; $Pr(Z>2)=0.025$;
$Pr(-1<Z<2)=Pr(-1<Z<0)+Pr(0<Z<2)=0.34+0.475=0.815$.

Q17. $\mu=60$, $\sigma=15$. $45=\mu-\sigma$, $90=\mu+2\sigma$, so $Pr(45<X<90)=0.34+0.475=0.815$. Also $75=\mu+\sigma$, so $Pr(X>75)=0.16$.

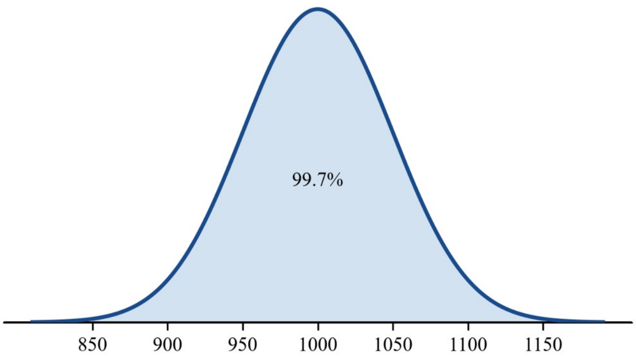
Q18. $\mu=500$, $\sigma=10$. $490=\mu-\sigma$, $520=\mu+2\sigma$, so $Pr(490<X<520)=0.34+0.475=0.815=81.5\%$.
$480=\mu-2\sigma$, so $Pr(X<480)=0.025=2.5\%$ underweight.



Q19. $Pr(X<20)=0.025=Pr(X<\mu-2\sigma)$, so $20=\mu-2\sigma$. $Pr(X<50)=0.975=Pr(X<\mu+2\sigma)$, so $50=\mu+2\sigma$. Adding: $2\mu=70$, $\mu=35$; subtracting: $4\sigma=30$, $\sigma=7.5$.

Q20. $Pr(X>85)=0.0015=Pr(X>\mu+3\sigma)$, so $85=\mu+3\sigma=70+3\sigma$. Then $3\sigma=15$, so $\sigma=5$.

Q21. $\mu=1000$, $\sigma=50$. (a) The middle 99.7% lies within 3σ: from $1000-150=850$ to $1000+150=1150$ hours. (b) $1100=\mu+2\sigma$, so $Pr(X>1100)=0.025$. (c) $900=\mu-2\sigma$, so $Pr(X<900)=0.025$; of 2000 globes, about $2000\times0.025=50$ last less than 900 hours.


99.7%
850
900
950
1000
1050
1100
1150

Theory

If a continuous random variable X is **normally distributed** with mean μ and standard deviation σ, we write $X \sim N(\mu, \sigma^2)$. Its graph is the symmetric bell-shaped curve centred on $x = \mu$, and a probability $Pr(a \leq X \leq b)$ is the **area** under the curve between $x = a$ and $x = b$.

Because there is no elementary antiderivative for the normal density, these areas are found with **technology** (a CAS calculator) — this is a technology-active skill. Using the CAS command $normalcdf(\text{lower}, \text{upper}, \mu, \sigma)$:

- $Pr(X < a) = normalcdf(-\infty, a, \mu, \sigma)$;
- $Pr(X > a) = normalcdf(a, \infty, \mu, \sigma)$;
- $Pr(a < X < b) = normalcdf(a, b, \mu, \sigma)$.

Three facts often let you find a probability more quickly, or check a CAS answer:

- **Symmetry.** The mean splits the area in half: $Pr(X < \mu) = Pr(X > \mu) = 0.5$. Also $Pr(\mu - k < X < \mu) = Pr(\mu < X < \mu + k)$.
- **Complement.** $Pr(X > a) = 1 - Pr(X < a)$, since the total area is 1.
- **Interval.** $Pr(a < X < b) = Pr(X < b) - Pr(X < a)$.

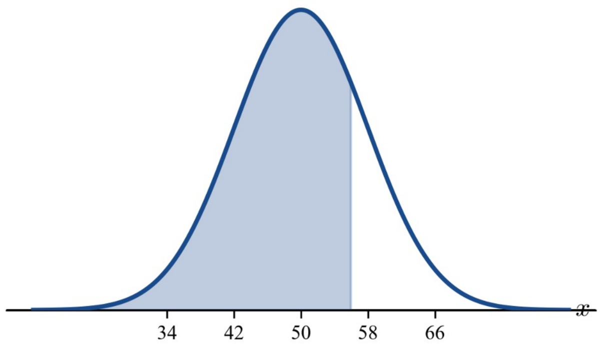
For a continuous variable $Pr(X = a) = 0$, so $<$ and $\leq$ give the same probability. Unless an exact form is available, state normal probabilities **correct to four decimal places**.

Worked examples

Example 1 — scaffolded

The lengths X mm of a machine part are normally distributed with $\mu = 50$ and $\sigma = 8$. Find $Pr(X < 56)$, correct to four decimal places.

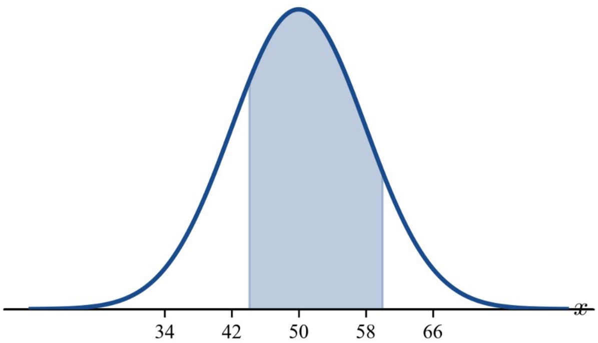
Working	Comment
$X \sim N(50, 8^2)$.	Identify $\mu = 50$, $\sigma = 8$.
Required area: $Pr(X < 56)$ — the area to the **left** of 56.	56 is above the mean, so expect a probability above 0.5.
$Pr(X < 56) = normalcdf(-\infty, 56, 50, 8)$	Set the lower limit to $-\infty$ (a very large negative value on the CAS).
$= 0.7734$.	Read the CAS value to four decimal places.



Example 2 — scaffolded

For the same variable $X \sim N(50, 8^2)$, find (a) $Pr(X > 60)$ and (b) $Pr(44 < X < 60)$.

Working	Comment
(a) $Pr(X > 60) = normalcdf(60, \infty, 50, 8) = 0.1056.$	The area to the **right** of 60; equivalently $1 - Pr(X < 60)$.
(b) $Pr(44 < X < 60) = normalcdf(44, 60, 50, 8)$ $= 0.6677.$	The area **between** 44 and 60. Equivalently $Pr(X < 60) - Pr(X < 44)$.



Example 3 — unscaffolded

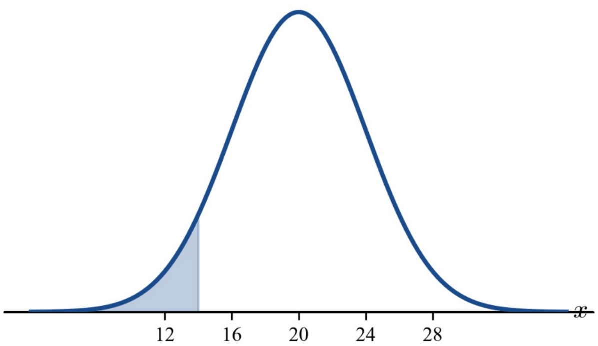
The mass X kg of newborn lambs is $X \sim N(20, 4^2)$. Find $Pr(X < 14)$ and $Pr(X > 26)$, and comment.

$Pr(X < 14) = normalcdf(-\infty, 14, 20, 4) = 0.0668.$

$Pr(X > 26) = normalcdf(26, \infty, 20, 4) = 0.0668.$

The values 14 and 26 are each $6 = 3/2\,\sigma$ from the mean 20, on opposite sides, so by the **symmetry** of the normal curve the two tail areas are equal.

$\therefore Pr(X < 14) = Pr(X > 26) = 0.0668.$



Example 4 — unscaffolded

IQ scores are modelled by $X \sim N(100, 15^2)$. Given that a person has an IQ above 100, find the probability that it is above 115.

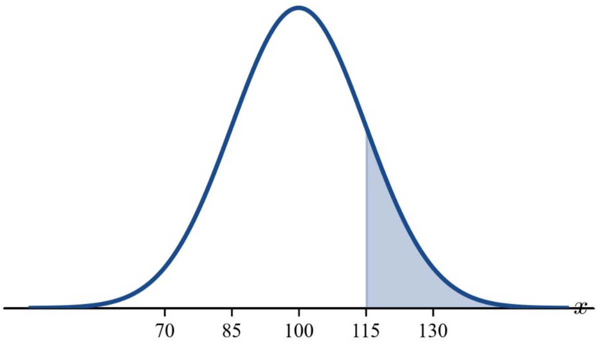
This is the conditional probability $Pr(X > 115 / X > 100)$. Since $\{X > 115\} \subseteq \{X > 100\}$,

$$Pr(X > 115 / X > 100) = \frac{Pr(X > 115)}{Pr(X > 100)}.$$

$Pr(X>115)=normalcdf(115,\infty,100,15)=0.158655$, and $Pr(X>100)=0.5$ (symmetry).

$$Pr(X>115 / X>100)=\frac{0.158655}{0.5}=0.3173.$$

$\therefore$ the conditional probability is 0.3173.



Practice questions

Give all probabilities correct to four decimal places.

Scaffolded

Q1. $X\sim N(70,10^2)$. Find: (a) $Pr(X<85)$; (b) $Pr(X>60)$; (c) $Pr(65<X<80)$.

Q2. $X\sim N(25,5^2)$. Find: (a) $Pr(X<25)$ (use symmetry — no CAS needed); (b) $Pr(X>35)$; (c) $Pr(20<X<30)$.

Q3. $Z\sim N(0,1^2)$ (the standard normal). Find: (a) $Pr(Z<1.5)$; (b) $Pr(Z>-1)$; (c) $Pr(-1<Z<2)$.

Q4. The contents X mL of a drink bottle are $X\sim N(500,50^2)$. Find: (a) $Pr(X>600)$; (b) $Pr(450<X<550)$; (c) hence write down $Pr(X<450 \text{ or } X>550)$.

Q5. $X\sim N(12,2^2)$. Find: (a) $Pr(X>12)$; (b) $Pr(10<X<14)$; (c) what does the answer to (b) tell you about the proportion of values within one standard deviation of the mean?

Q6. The diameters X cm of ball bearings are $X\sim N(1.6,0.1^2)$. Given that a bearing has diameter greater than 1.6 cm, find the probability that it is greater than 1.75 cm.

Unscaffolded

State each probability correct to four decimal places.

Q7. $X\sim N(60,12^2)$. Find $Pr(X<75)$.

Q8. $X\sim N(60,12^2)$. Find $Pr(X>80)$.

Q9. $X\sim N(60,12^2)$. Find $Pr(50<X<78)$.

Q10. The heights X cm of adult males are $X\sim N(170,8^2)$. Find the probability that a randomly chosen man is taller than 180 cm.

Q11. For the heights in **Q10**, find $Pr(160<X<175)$.

Q12. The masses X kg of bags of sugar are $X \sim N(3.2, 0.4^2)$. Find the probability that a bag has mass less than 3 kg.

Q13. Battery lifetimes X hours are $X \sim N(500, 75^2)$. Find $Pr(X > 620)$.

Q14. Exam marks X are $X \sim N(72, 9^2)$. A mark below 50 is a fail. Find the probability that a randomly chosen student fails.

Q15. For the marks in **Q14**, find $Pr(65 < X < 90)$.

Q16. $X \sim N(250, 20^2)$. Find $Pr(X > 300)$.

Q17. $Z \sim N(0, 1^2)$. Find $Pr(Z < -0.5)$.

Q18. $X \sim N(20, 3^2)$. Find $Pr(17 < X < 26)$.

Q19. Waiting times X minutes are $X \sim N(15, 2.5^2)$. Given that a wait is longer than 15 minutes, find the probability that it is longer than 18 minutes.

Q20. $X \sim N(100, 15^2)$. Find $Pr(85 < X < 130)$.

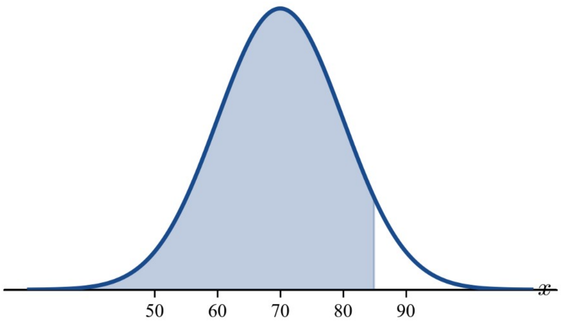
Q21. $X \sim N(50, 10^2)$. Given that $X > 60$, find the probability that $X > 70$, correct to four decimal places.

Answers

Q1. (a) 0.9332 (b) 0.8413 (c) 0.5328. **Q2.** (a) 0.5 (b) 0.0228 (c) 0.6827. **Q3.** (a) 0.9332 (b) 0.8413 (c) 0.8186. **Q4.** (a) 0.0228 (b) 0.6827 (c) 0.3173. **Q5.** (a) 0.5 (b) 0.6827 (c) about 68.27 % of values lie within one standard deviation of the mean. **Q6.** 0.1336. **Q7.** 0.8944. **Q8.** 0.0478. **Q9.** 0.7309. **Q10.** 0.1056. **Q11.** 0.6284. **Q12.** 0.3085 . **Q13.** 0.0548. **Q14.** 0.0073. **Q15.** 0.7589. **Q16.** 0.0062. **Q17.** 0.3085. **Q18.** 0.8186. **Q19.** 0.2301. **Q20.** 0.7377. **Q21.** 0.1434.

Fully-worked solutions

Q1. $X \sim N(70, 10^2)$. (a) $Pr(X<85) = normalcdf(-\infty, 85, 70, 10) = 0.9332$. (b) $Pr(X>60) = normalcdf(60, \infty, 70, 10) = 0.8413$. (c) $Pr(65<X<80) = normalcdf(65, 80, 70, 10) = 0.5328$.



Q2. $X \sim N(25, 5^2)$. (a) By symmetry about the mean, $Pr(X<25) = 0.5$. (b) $Pr(X>35) = normalcdf(35, \infty, 25, 5) = 0.0228$. (c) $Pr(20<X<30) = normalcdf(20, 30, 25, 5) = 0.6827$ (20 and 30 are each one standard deviation from the mean).

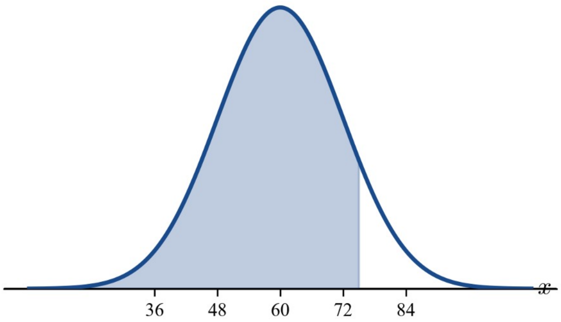
Q3. $Z \sim N(0, 1^2)$. (a) $Pr(Z<1.5) = 0.9332$. (b) $Pr(Z>-1) = 0.8413$. (c) $Pr(-1<Z<2) = normalcdf(-1, 2, 0, 1) = 0.8186$ (equivalently $Pr(Z<2) - Pr(Z<-1)$).

Q4. $X \sim N(500, 50^2)$. (a) $Pr(X>600) = normalcdf(600, \infty, 500, 50) = 0.0228$. (b) $Pr(450<X<550) = normalcdf(450, 550, 500, 50) = 0.6827$. (c) The remaining area is $Pr(X<450 \text{ or } X>550) = 1 - 0.6827 = 0.3173$.

Q5. $X \sim N(12, 2^2)$. (a) $Pr(X>12) = 0.5$ (symmetry). (b) $Pr(10<X<14) = normalcdf(10, 14, 12, 2) = 0.6827$. (c) 10 and 14 are exactly one standard deviation either side of the mean, so about 68.27 % of the values lie within one standard deviation of the mean.

Q6. $X \sim N(1.6, 0.1^2)$. $Pr(X>1.75 / X>1.6) = \dfrac{Pr(X>1.75)}{Pr(X>1.6)}$. Now $Pr(X>1.75) = normalcdf(1.75, \infty, 1.6, 0.1) = 0.066807$ and $Pr(X>1.6) = 0.5$, so the probability is $\dfrac{0.066807}{0.5} = 0.1336$.

Q7. $Pr(X<75) = normalcdf(-\infty, 75, 60, 12) = 0.8944$.



Q8. $Pr(X>80)=normalcdf(80,\infty,60,12)=0.0478$.

Q9. $Pr(50<X<78)=normalcdf(50,78,60,12)=0.7309$.

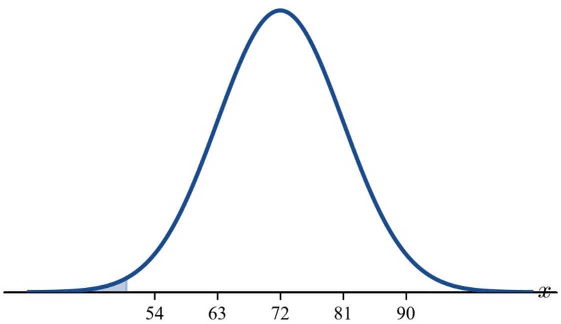
Q10. $X\sim N(170,8^2)$. $Pr(X>180)=normalcdf(180,\infty,170,8)=0.1056$.

Q11. $Pr(160<X<175)=normalcdf(160,175,170,8)=0.6284$.

Q12. $X\sim N(3.2,0.4^2)$. $Pr(X<3)=normalcdf(-\infty,3,3.2,0.4)=0.3085$.

Q13. $Pr(X>620)=normalcdf(620,\infty,500,75)=0.0548$.

Q14. $X\sim N(72,9^2)$. $Pr(X<50)=normalcdf(-\infty,50,72,9)=0.0073$.



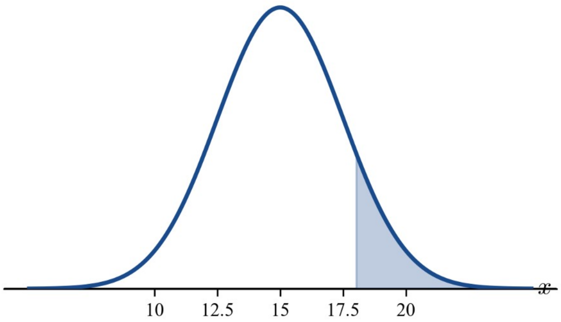
Q15. $Pr(65<X<90)=normalcdf(65,90,72,9)=0.7589$.

Q16. $Pr(X>300)=normalcdf(300,\infty,250,20)=0.0062$.

Q17. $Pr(Z<-0.5)=normalcdf(-\infty,-0.5,0,1)=0.3085$.

Q18. $Pr(17<X<26)=normalcdf(17,26,20,3)=0.8186$.

Q19. $X\sim N(15,2.5^2)$. $Pr(X>18/X>15)=\dfrac{Pr(X>18)}{Pr(X>15)}$. Here $Pr(X>18)=normalcdf(18,\infty,15,2.5)=0.115070$ and $Pr(X>15)=0.5$, so the probability is $\dfrac{0.115070}{0.5}=0.2301$.



Q20. $Pr(85 < X < 130) = normalcdf(85, 130, 100, 15) = 0.7377$.

Q21. $X \sim N(50, 10^2)$. $Pr(X > 70 / X > 60) = \dfrac{Pr(X > 70)}{Pr(X > 60)}$. Now $Pr(X > 70) = normalcdf(70, \infty, 50, 10) = 0.022750$ and $Pr(X > 60) = normalcdf(60, \infty, 50, 10) = 0.158655$, so the probability is $\dfrac{0.022750}{0.158655} = 0.1434$.

Theory

In 14C you found probabilities for a normal random variable $X \sim N(\mu, \sigma^2)$. This section solves the **inverse** problems: given a probability, find a value; or find an unknown **mean** μ or **standard deviation** σ. These are technology-active — the CAS **inverse normal** command does the numerical work — but you must set up the equation by hand.

Standardisation. If $X \sim N(\mu, \sigma^2)$ then $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$. A probability statement about X becomes one about Z, and the boundary value x corresponds to $z = \dfrac{x - \mu}{\sigma}$.

The inverse normal. For a given probability p, the value z with $Pr(Z < z) = p$ is written $z = \Phi^{-1}(p)$ and read from the CAS *inverse normal*. Useful values (know these):

p	0.10	0.20	0.80	0.90	0.95	0.975	0.99
z	-1.2816	-0.8416	0.8416	1.2816	1.6449	1.9600	2.3263

The three inverse problems.

- **Find a value** a with $Pr(X < a) = p$: compute $a = \mu + \sigma\Phi^{-1}(p)$ (directly, with CAS inverse normal on $N(\mu, \sigma^2)$).
- **Find** μ (given σ): from $\dfrac{a - \mu}{\sigma} = z$ where $z = \Phi^{-1}(p)$, so $\mu = a - \sigma z$.
- **Find** σ (given μ): $\sigma = \dfrac{a - \mu}{z}$.
- **Find both** μ **and** σ: two probability statements give two equations $\dfrac{a_1 - \mu}{\sigma} = z_1$ and $\dfrac{a_2 - \mu}{\sigma} = z_2$; solve simultaneously.

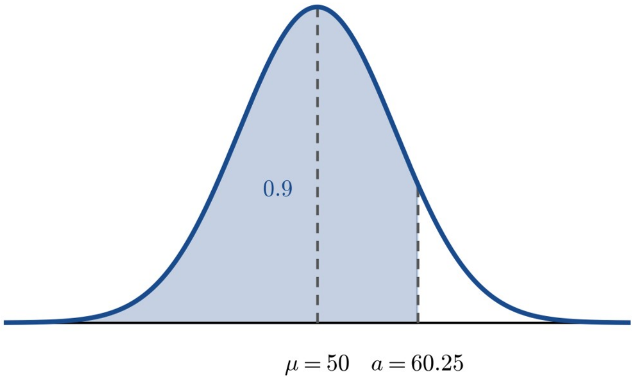
Watch the tail. For $Pr(X > a) = p$, convert to $Pr(X < a) = 1 - p$ first. Always sketch a bell curve and shade the given area — it tells you whether the required z is positive or negative and prevents sign errors.

Worked examples

Example 1 — scaffolded

$X \sim N(50, 8^2)$. Find the value a such that $Pr(X < a) = 0.9$, correct to two decimal places.

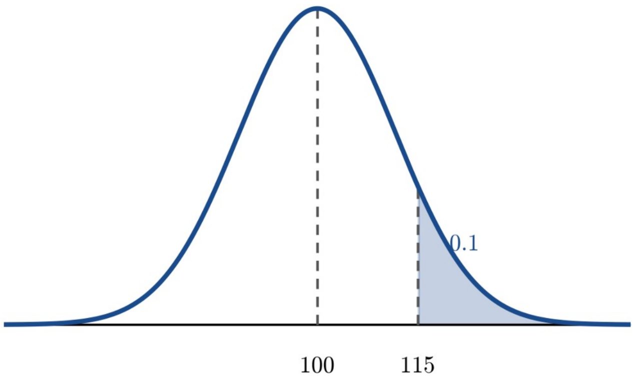
Working	Comment
$Z = \dfrac{X - 50}{8}$; require $Pr(Z < z) = 0.9$.	Standardise: the left area is 0.9.
Inverse normal: $z = \Phi^{-1}(0.9) = 1.2816$.	Read from CAS (or the table).
$\dfrac{a - 50}{8} = 1.2816 \Rightarrow a = 50 + 8(1.2816)$.	Un-standardise: $a = \mu + \sigma z$.
$a = 60.25$.	To two decimal places.



Example 2 — scaffolded

The lifetime X (hours) of a component is $X \sim N(100, \sigma^2)$. Given $Pr(X > 115) = 0.10$, find σ.

Working	Comment
$Pr(X > 115) = 0.10 \Rightarrow Pr(X < 115) = 0.90$.	Convert the right tail to a left area.
$z = \Phi^{-1}(0.90) = 1.2816$; $\dfrac{115 - 100}{\sigma} = 1.2816$.	Standardise the boundary $x = 115$.
$\sigma = \dfrac{15}{1.2816} = 11.70$.	Solve for σ.



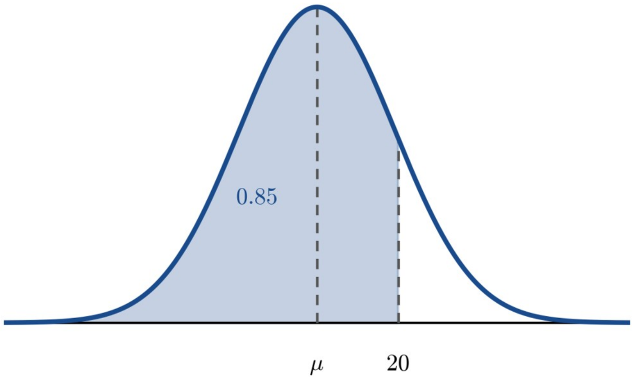
Example 3 — unscaffolded

$X \sim N(\mu, 5^2)$ and $Pr(X < 20) = 0.85$. Find μ.

Standardise: $\dfrac{20 - \mu}{5} = \Phi^{-1}(0.85) = 1.0364$.

So $20 - \mu = 5(1.0364) = 5.182$, giving $\mu = 20 - 5.182 = 14.82$.

$\therefore \mu = 14.82$.



Example 4 — unscaffolded

$X \sim N(\mu, \sigma^2)$ with $Pr(X<60)=0.2$ and $Pr(X<80)=0.9$. Find μ and σ.

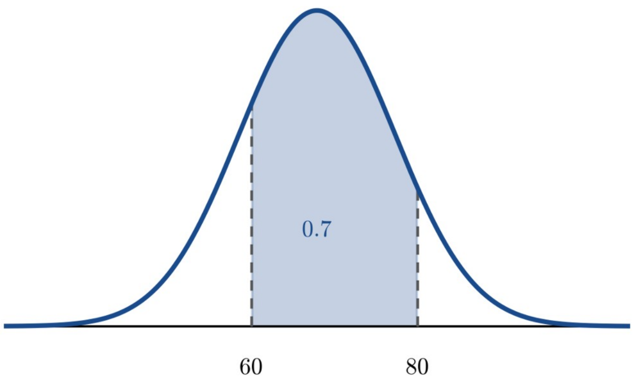
Two standardised equations, using $\Phi^{-1}(0.2)=-0.8416$ and $\Phi^{-1}(0.9)=1.2816$:

$$\frac{60-\mu}{\sigma}=-0.8416, \frac{80-\mu}{\sigma}=1.2816.$$

Subtracting: $\frac{80-60}{\sigma}=1.2816-(-0.8416)=2.1232$, so $\sigma=\frac{20}{2.1232}=9.42$.

Then $\mu=60+0.8416\sigma=60+0.8416(9.42)=67.93$.

$\therefore \mu=67.93, \sigma=9.42$.



Practice questions

Scaffolded

Q1. $X \sim N(70, 10^2)$. (a) State the z-value with $Pr(Z<z)=0.95$. (b) Hence find a such that $Pr(X<a)=0.95$, to two decimal places.

Q2. $X \sim N(25, \sigma^2)$ with $Pr(X<30)=0.8$. (a) Write the standardised equation for σ. (b) Find σ.

Q3. $X \sim N(\mu, 4^2)$ with $Pr(X>50)=0.05$. (a) Convert to a left-tail probability and state the z-value. (b) Find μ.

Q4. $X \sim N(500, 50^2)$. Find the value a exceeded with probability 0.9 (i.e. $Pr(X>a)=0.9$).

Q5. $X \sim N(\mu, \sigma^2)$ with $Pr(X<10)=0.1$ and $Pr(X<25)=0.7$. (a) Write the two standardised equations. (b) Solve for μ and σ.

Q6. Exam scores are $X \sim N(65, 12^2)$. The top 15% of students receive an A. Find the minimum score (to two decimal places) needed for an A.

Unscaffolded

Q7. $X \sim N(40, 6^2)$. Find a with $Pr(X < a) = 0.975$.

Q8. $X \sim N(200, \sigma^2)$ with $Pr(X > 230) = 0.025$. Find σ.

Q9. $X \sim N(\mu, 8^2)$ with $Pr(X < 100) = 0.6$. Find μ.

Q10. $X \sim N(15, \sigma^2)$ with $Pr(X < 12) = 0.2$. Find σ.

Q11. $X \sim N(\mu, \sigma^2)$ with $Pr(X < 5) = 0.15$ and $Pr(X < 9) = 0.75$. Find μ and σ.

Q12. The charge time X (minutes) of a battery is $X \sim N(50, \sigma^2)$ and $Pr(X > 60) = 0.1$. (a) Find σ. (b) Hence find $Pr(X > 45)$, to four decimal places.

Q13. For the standard normal $Z \sim N(0, 1)$, find the value a such that $Pr(-a < Z < a) = 0.9$.

Q14. $X \sim N(500, 20^2)$. Find the first and third quartiles Q_1 and Q_3, and hence the interquartile range.

Q15. The heights (cm) of a population are $X \sim N(170, 7^2)$. Find the height below which the shortest 10% lie.

Q16. $X \sim N(\mu, 5^2)$ with $Pr(X > 18) = 0.2$. Find μ.

Q17. $X \sim N(80, \sigma^2)$ with $Pr(74 < X < 86) = 0.6$. Find σ. *(Use the symmetry of the interval about the mean.)*

Q18. $X \sim N(\mu, \sigma^2)$ with $Pr(X < 30) = 0.3$ and $Pr(X > 50) = 0.1$. Find μ and σ.

Q19. A machine fills bottles with $X \sim N(\mu, 3^2)$ mL. Only 1% of bottles may contain less than 500 mL. Find the smallest mean μ (to two decimal places) that achieves this.

Q20. $X \sim N(100, 15^2)$. Find the 60th percentile.

Q21. $X \sim N(\mu, \sigma^2)$ with $Pr(X < 45) = 0.2$ and $Pr(X < 65) = 0.9$. (a) Find μ and σ. (b) Hence find $Pr(X > 70)$, to four decimal places.

Answers

Q1. (a) 1.6449 (b) 86.45 **Q2.** $\sigma=5.94$ **Q3.** $\mu=43.42$ **Q4.** $a=435.92$ **Q5.** $\mu=20.64$, $\sigma=8.31$ **Q6.** 77.44 **Q7.** 51.76 **Q8.** $\sigma=15.31$ **Q9.** $\mu=97.97$ **Q10.** $\sigma=3.56$ **Q11.** $\mu=7.42$, $\sigma=2.34$ **Q12.** (a) $\sigma=7.80$ (b) 0.7392 **Q13.** $a=1.64$ **Q14.** $Q_1=486.51$, $Q_3=513.49$, IQR $=26.98$ **Q15.** 161.03 cm **Q16.** $\mu=13.79$ **Q17.** $\sigma=7.13$ **Q18.** $\mu=35.81$, $\sigma=11.07$ **Q19.** $\mu=506.98$ mL **Q20.** 103.80 **Q21.** (a) $\mu=52.93$, $\sigma=9.42$ (b) 0.0350

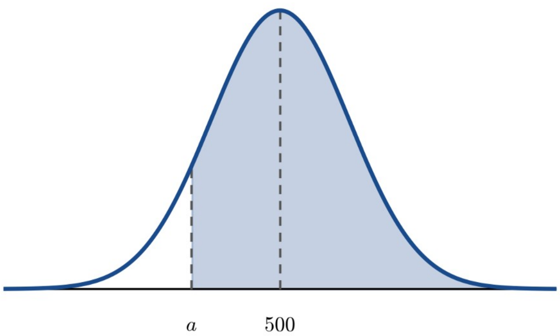
Fully-worked solutions

Q1. (a) $\Phi^{-1}(0.95)=1.6449$. (b) $a=70+10(1.6449)=86.45$.

Q2. $Pr(X<30)=0.8 \Rightarrow \dfrac{30-25}{\sigma}=\Phi^{-1}(0.8)=0.8416$, so $\sigma=\dfrac{5}{0.8416}=5.94$.

Q3. $Pr(X>50)=0.05 \Rightarrow Pr(X<50)=0.95$, $z=\Phi^{-1}(0.95)=1.6449$. Then $\dfrac{50-\mu}{4}=1.6449 \Rightarrow \mu=50-4(1.6449)=43.42$.

Q4. $Pr(X>a)=0.9 \Rightarrow Pr(X<a)=0.1$, $z=\Phi^{-1}(0.1)=-1.2816$. So $a=500+50(-1.2816)=435.92$.



Q5. $\dfrac{10-\mu}{\sigma}=\Phi^{-1}(0.1)=-1.2816$ and $\dfrac{25-\mu}{\sigma}=\Phi^{-1}(0.7)=0.5244$. Subtracting: $\dfrac{15}{\sigma}=0.5244-(-1.2816)=1.8060$, so $\sigma=8.31$; then $\mu=10+1.2816(8.31)=20.64$.

Q6. Top 15% means $Pr(X<a)=0.85$, $z=\Phi^{-1}(0.85)=1.0364$. $a=65+12(1.0364)=77.44$.

Q7. $a=40+6\Phi^{-1}(0.975)=40+6(1.9600)=51.76$.

Q8. $Pr(X>230)=0.025 \Rightarrow Pr(X<230)=0.975$, $z=1.9600$. $\dfrac{230-200}{\sigma}=1.96 \Rightarrow \sigma=\dfrac{30}{1.96}=15.31$.

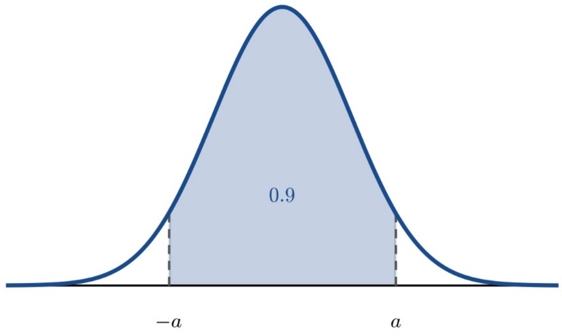
Q9. $\dfrac{100-\mu}{8}=\Phi^{-1}(0.6)=0.2533 \Rightarrow \mu=100-8(0.2533)=97.97$.

Q10. $\dfrac{12-15}{\sigma}=\Phi^{-1}(0.2)=-0.8416 \Rightarrow \sigma=\dfrac{-3}{-0.8416}=3.56$.

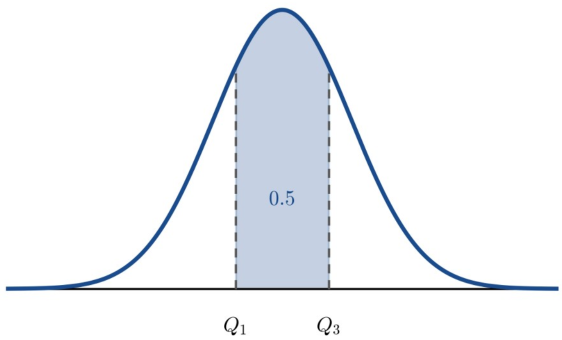
Q11. $\dfrac{5-\mu}{\sigma}=\Phi^{-1}(0.15)=-1.0364$ and $\dfrac{9-\mu}{\sigma}=\Phi^{-1}(0.75)=0.6745$. Subtracting: $\dfrac{4}{\sigma}=1.7109 \Rightarrow \sigma=2.34$; $\mu=5+1.0364(2.34)=7.42$.

Q12. (a) $Pr(X>60)=0.1 \Rightarrow Pr(X<60)=0.9$, $\dfrac{60-50}{\sigma}=1.2816 \Rightarrow \sigma=7.80$. (b) $Pr(X>45)=Pr\left(Z>\dfrac{45-50}{7.80}\right)=Pr(Z>-0.641)=0.7392$.

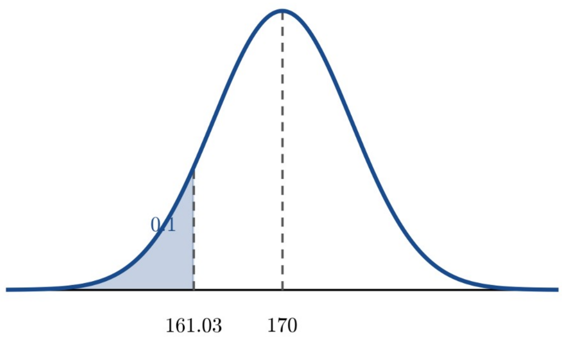
Q13. By symmetry $Pr(Z<a)=0.5+\dfrac{0.9}{2}=0.95$, so $a=\Phi^{-1}(0.95)=1.64$.



Q14. Q_1: $Pr(X<Q_1)=0.25$, $z=-0.6745$, $Q_1=500+20(-0.6745)=486.51$. Q_3: $z=0.6745$, $Q_3=513.49$. IQR $=513.49-486.51=26.98$.



Q15. $Pr(X<a)=0.1$, $z=-1.2816$, $a=170+7(-1.2816)=161.03$ cm.



Q16. $Pr(X>18)=0.2 \Rightarrow Pr(X<18)=0.8$, $z=0.8416$. $\dfrac{18-\mu}{5}=0.8416 \Rightarrow \mu=18-5(0.8416)=13.79$.

Q17. The interval $74<X<86$ is symmetric about $\mu=80$, so $Pr(X<86)=0.5+\dfrac{0.6}{2}=0.8$, $z=0.8416$. $\dfrac{86-80}{\sigma}=0.8416 \Rightarrow \sigma=\dfrac{6}{0.8416}=7.13$.

Q18. $\frac{30-\mu}{\sigma}=\Phi^{-1}(0.3)=-0.5244$; $Pr(X>50)=0.1 \Rightarrow Pr(X<50)=0.9$, $\frac{50-\mu}{\sigma}=1.2816$. Subtracting: $\frac{20}{\sigma}=1.8060 \Rightarrow \sigma=11.07$; $\mu=30+0.5244(11.07)=35.81$.

Q19. $Pr(X<500)=0.01$, $z=\Phi^{-1}(0.01)=-2.3263$. $\frac{500-\mu}{3}=-2.3263 \Rightarrow \mu=500+3(2.3263)=506.98$ mL.

Q20. The 60th percentile is a with $Pr(X<a)=0.6$: $a=100+15(0.2533)=103.80$.

Q21. (a) $\frac{45-\mu}{\sigma}=\Phi^{-1}(0.2)=-0.8416$ and $\frac{65-\mu}{\sigma}=\Phi^{-1}(0.9)=1.2816$. Subtracting: $\frac{20}{\sigma}=2.1232 \Rightarrow \sigma=9.42$; $\mu=45+0.8416(9.42)=52.93$. (b) $Pr(X>70)=Pr\left(Z>\frac{70-52.93}{9.42}\right)=Pr(Z>1.812)=0.0350$.

Chapter 14 The normal distribution — 14E The normal approximation to the binomial distribution

Theory

When the number of trials n is large, the binomial distribution $X \sim \mathrm{Bi}(n, p)$ has a graph that is close to a bell shape, and it can be **approximated by a normal distribution**.

The approximating normal. If $X \sim \mathrm{Bi}(n, p)$ then X is approximately

$$X \approx N(\mu, \sigma^2), \mu = np, \sigma^2 = np(1-p),$$

that is, the approximating normal has the **same mean and the same variance** as the binomial.

When is the approximation reasonable? The rule of thumb is that **both** np and $n(1-p)$ should be reasonably large — a common guide is that each is at least 5 (better, at least 10). The approximation is best when p is close to $\frac{1}{2}$ (so the binomial is nearly symmetric) and improves as n increases.

Using the approximation. To estimate a binomial probability such as $Pr(X \le k)$:

1. Find $\mu = np$ and $\sigma^2 = np(1-p)$ (so $\sigma = \sqrt{np(1-p)}$).
2. Treat X as $N(\mu, \sigma^2)$ and find the required probability — either directly with technology, or by **standardising** $Z = \dfrac{X-\mu}{\sigma}$ and using the standard normal.

For example, $Pr(X \le k) \approx Pr\left(Z \le \dfrac{k-\mu}{\sigma}\right)$.

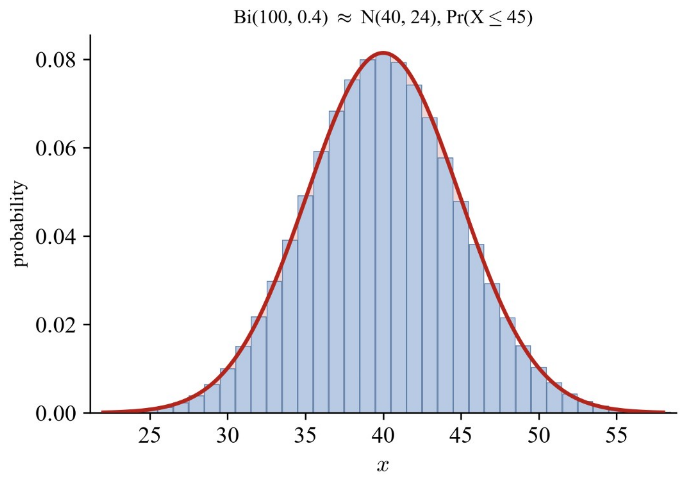
A note on accuracy. The approximate value is *close to* — but not equal to — the exact binomial value. In Mathematical Methods no continuity correction is applied; the normal approximation is used as a quick estimate, and the exact answer can always be found directly from the binomial.

Worked examples

Example 1 — scaffolded

The number of successes in 100 independent trials, each with probability 0.4 of success, is $X \sim \mathrm{Bi}(100, 0.4)$. Use a normal approximation to estimate $Pr(X \le 45)$.

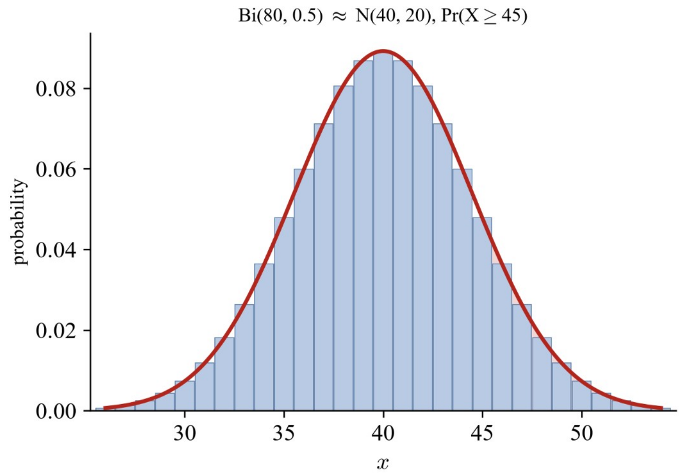
Working	Comment
$\mu = np = 100 \times 0.4 = 40$.	The approximating normal has mean np.
$\sigma^2 = np(1-p) = 100 \times 0.4 \times 0.6 = 24$, so $\sigma = 2\sqrt{6} \approx 4.899$.	…and variance $np(1-p)$.
Check: $np = 40$ and $n(1-p) = 60$ are both large, so the approximation is reasonable.	Both ≥ 10.
Approximate X by $N(40, 24)$. Standardise: $z = \dfrac{45-40}{2\sqrt{6}} \approx 1.0206$.	Convert to the standard normal.
$Pr(X \le 45) \approx Pr(Z \le 1.0206) \approx 0.8463$.	Read the standard-normal probability.
(The exact binomial value is 0.8689, so the estimate is close.)	Sense-check against the exact value.



Example 2 — scaffolded

For $X \sim \mathrm{Bi}(80, 0.5)$, use a normal approximation to estimate $Pr(X \geq 45)$.

Working	Comment
$\mu = 80 \times 0.5 = 40$; $\sigma^2 = 80 \times 0.5 \times 0.5 = 20$, so $\sigma = 2\sqrt{5} \approx 4.472$.	Same mean and variance as the binomial.
$np = n(1-p) = 40$ — both large, and $p = 0.5$, so the approximation is good.	Symmetric case.
$z = \dfrac{45-40}{2\sqrt{5}} \approx 1.1180$.	Standardise the value 45.
$Pr(X \geq 45) \approx Pr(Z \geq 1.1180) \approx 0.1318$.	Upper-tail probability.
(Exact binomial: 0.1572.)	Sense-check.



Example 3 — unscaffolded

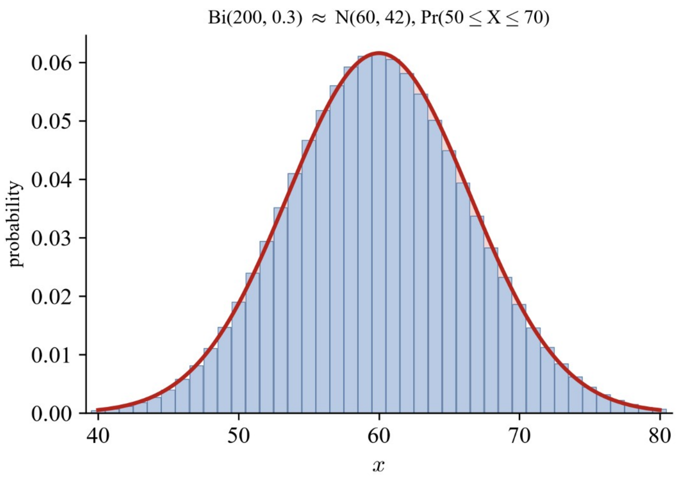
For $X \sim \mathrm{Bi}(200, 0.3)$, estimate $Pr(50 \leq X \leq 70)$ using a normal approximation.

$\mu = 200 \times 0.3 = 60$ and $\sigma^2 = 200 \times 0.3 \times 0.7 = 42$, so $\sigma = \sqrt{42} \approx 6.481$. Both $np = 60$ and $n(1-p) = 140$ are large, so $X \approx N(60, 42)$.

Standardising the endpoints: $z_1 = \dfrac{50-60}{\sqrt{42}} \approx -1.5430$ and $z_2 = \dfrac{70-60}{\sqrt{42}} \approx 1.5430$.

$Pr(50 \leq X \leq 70) \approx Pr(-1.5430 \leq Z \leq 1.5430) \approx 0.8772$.

$\therefore$ the required probability is about 0.8772 (exact binomial 0.8952).



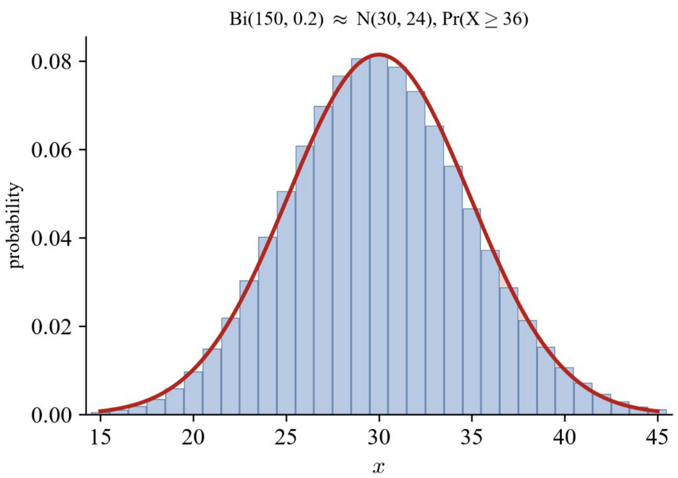
Example 4 — unscaffolded

A fair spinner shows a particular colour with probability 0.2. In 150 spins, let X be the number of times that colour appears. Estimate $Pr(X \geq 36)$ using a normal approximation.

$X \sim \text{Bi}(150, 0.2)$, so $\mu = 150 \times 0.2 = 30$ and $\sigma^2 = 150 \times 0.2 \times 0.8 = 24$, giving $\sigma = 2\sqrt{6} \approx 4.899$. Here $np = 30$ and $n(1-p) = 120$ are both large, so $X \approx N(30, 24)$.

$z = \dfrac{36-30}{2\sqrt{6}} \approx 1.2247$, so $Pr(X \geq 36) \approx Pr(Z \geq 1.2247) \approx 0.1103$.

$\therefore$ the colour appears at least 36 times with probability about 0.1103 (exact 0.1317).



Practice questions

Scaffolded

For each, give probabilities correct to four decimal places.

Q1. For $X \sim \text{Bi}(100, 0.5)$: (a) state μ and σ^2 for the approximating normal; (b) confirm the approximation is reasonable; (c) estimate $Pr(X \leq 55)$.

Q2. For $X \sim \text{Bi}(64, 0.25)$: (a) find μ and σ; (b) estimate $Pr(X \geq 20)$.

Q3. For $X \sim \text{Bi}(400, 0.5)$: (a) find μ and σ; (b) estimate $Pr(190 \leq X \leq 210)$; (c) which two standardised values are the endpoints?

Q4. For $X \sim \text{Bi}(50, 0.4)$: (a) find μ and σ^2; (b) estimate $Pr(X \leq 18)$.

Q5. For $X \sim \text{Bi}(300, 1/3)$: (a) find μ and σ^2; (b) estimate $Pr(X \geq 110)$.

Q6. For $X \sim \text{Bi}(144, 0.5)$: (a) find μ and σ; (b) estimate $Pr(66 \leq X \leq 81)$.

Unscaffolded

Use a normal approximation to estimate each probability, correct to four decimal places, and state the mean and variance of the approximating normal.

Q7. $X \sim \text{Bi}(100, 0.3)$; find $Pr(X \leq 35)$.

Q8. $X \sim \text{Bi}(250, 0.4)$; find $Pr(X \geq 110)$.

Q9. In a large batch, 50% of components pass a test. For a sample of 324 components, let X be the number that pass. Estimate $Pr(X \leq 174)$, and comment on whether the normal approximation is suitable here.

Answers

Q1. (a) $\mu = 50$, $\sigma^2 = 25$; (c) 0.8413. **Q2.** (a) $\mu = 16$, $\sigma = 2\sqrt{3} \approx 3.4641$; (b) 0.1241. **Q3.** (a) $\mu = 200$, $\sigma = 10$; (b) 0.6827; (c) $z = \pm 1$. **Q4.** (a) $\mu = 20$, $\sigma^2 = 12$; (b) 0.2819. **Q5.** (a) $\mu = 100$, $\sigma^2 = \frac{200}{3}$; (b) 0.1103. **Q6.** (a) $\mu = 72$, $\sigma = 6$; (b) 0.7745. **Q7.** $N(30, 21)$; 0.8624. **Q8.** $N(100, 60)$; 0.0984. **Q9.** $N(162, 81)$; 0.9088; suitable ($np = n(1-p) = 162$, both large, $p = 0.5$).

Fully-worked solutions

Q1. $\mu = 100(0.5) = 50$, $\sigma^2 = 100(0.5)(0.5) = 25$, $\sigma = 5$. Both np and $n(1-p)$ equal 50, so the approximation is reasonable. $z = \frac{55-50}{5} = 1$, so $Pr(X \le 55) \approx Pr(Z \le 1) \approx 0.8413$. (Exact 0.8644.)

Q2. $\mu = 64(0.25) = 16$, $\sigma^2 = 64(0.25)(0.75) = 12$, $\sigma = 2\sqrt{3} \approx 3.4641$. $z = \frac{20-16}{2\sqrt{3}} \approx 1.1547$, so $Pr(X \ge 20) \approx Pr(Z \ge 1.1547) \approx 0.1241$. (Exact 0.1561.)

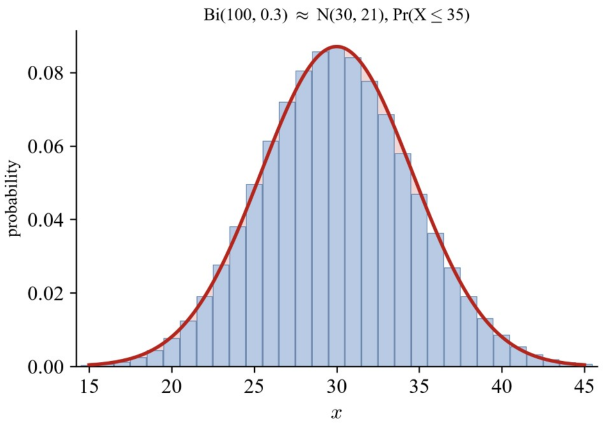
Q3. $\mu = 200$, $\sigma^2 = 400(0.5)(0.5) = 100$, $\sigma = 10$. Endpoints standardise to $z = \frac{190-200}{10} = -1$ and $z = \frac{210-200}{10} = 1$, so $Pr(190 \le X \le 210) \approx Pr(-1 \le Z \le 1) \approx 0.6827$. (Exact 0.7063.)

Q4. $\mu = 20$, $\sigma^2 = 50(0.4)(0.6) = 12$, $\sigma = 2\sqrt{3} \approx 3.4641$. $z = \frac{18-20}{2\sqrt{3}} \approx -0.5774$, so $Pr(X \le 18) \approx Pr(Z \le -0.5774) \approx 0.2819$. (Exact 0.3356.)

Q5. $\mu = 100$, $\sigma^2 = 300 \cdot 1/3 \cdot 2/3 = \frac{200}{3}$, $\sigma \approx 8.165$. $z = \frac{110-100}{8.165} \approx 1.2247$, so $Pr(X \ge 110) \approx Pr(Z \ge 1.2247) \approx 0.1103$. (Exact 0.1228.)

Q6. $\mu = 72$, $\sigma^2 = 144(0.5)(0.5) = 36$, $\sigma = 6$. Endpoints: $z_1 = \frac{66-72}{6} = -1$, $z_2 = \frac{81-72}{6} = 1.5$, so $Pr(66 \le X \le 81) \approx Pr(-1 \le Z \le 1.5) \approx 0.7745$. (Exact 0.8042.)

Q7. $\mu = 30$, $\sigma^2 = 100(0.3)(0.7) = 21$, $\sigma \approx 4.5826$. $z = \frac{35-30}{\sqrt{21}} \approx 1.0911$, so $Pr(X \le 35) \approx 0.8624$. (Exact 0.8839.)



Q8. $\mu = 100$, $\sigma^2 = 250(0.4)(0.6) = 60$, $\sigma \approx 7.746$. $z = \frac{110-100}{\sqrt{60}} \approx 1.2910$, so $Pr(X \ge 110) \approx 0.0984$. (Exact 0.1104.)

Q9. $\mu = 324(0.5) = 162$, $\sigma^2 = 324(0.5)(0.5) = 81$, $\sigma = 9$. Here $np = n(1-p) = 162$ are both large and $p = 0.5$, so the normal approximation is very suitable. $z = \dfrac{174 - 162}{9} \approx 1.3333$, so $Pr(X \leq 174) \approx 0.9088$. (Exact 0.9176.)