NCEA Level 2 · AS 91267: Probability Methods

Name: _____ Date: _____

Hi! This standard is all about working out chances using tables, trees, risk and the normal distribution. Work through each section in order, check your answers against the workings, and don't skip the Excellence justifications, that's where the top marks live.

Revision summary

- Probability of an event: P(A) = (favourable outcomes) ÷ (total outcomes), always between 0 and 1.
- Two-way tables: read off joint totals; row/column totals give marginal probabilities.
- Conditional probability: P(A given B) = P(A and B) ÷ P(B). On a tree, follow the relevant branch.
- Tree diagrams: multiply along branches, add across branches. Branch probabilities from each node sum to 1.
- Risk (relative frequency) = (number with outcome) ÷ (total in group). Relative risk = risk of group A ÷ risk of group B.
- Relative risk > 1 means higher risk than the baseline group; = 1 means equal risk; < 1 means lower risk.
- Normal distribution: symmetric, bell-shaped about the mean μ. Standardise with $z = (x - \mu) \div \sigma$, then use tables/calculator for the area (probability).
- Roughly 68% of data lies within 1 σ of the mean, 95% within 2 σ, 99.7% within 3 σ.

Key skills

1. A bag holds 5 red, 3 blue and 2 green marbles. One marble is drawn at random. Find P(blue).

2. In a survey of 200 people, 120 own a bike and of those 45 also own a scooter. Find P(owns a scooter given owns a bike).

3. A spinner lands on 'win' with probability 0.4. It is spun twice. Draw/consider a tree and find P(win then win).

4. Heights are normally distributed with mean $\mu = 170$ cm and standard deviation $\sigma = 8$ cm. Find the z-score for a height of 182 cm.

Achieved

5. At Avondale College, 300 students were asked if they walk to school. The two-way table shows: Junior students – 90 walk, 60 do not; Senior students – 50 walk, 100 do not. Find the probability a randomly chosen student walks to school.

6. A café finds 70% of customers order coffee. Of those who order coffee, 30% also order a muffin. Draw a tree for the first branch and find P(orders coffee and a muffin).

7. The time to fill a petrol tank is normally distributed with mean 4.5 minutes and standard deviation 0.8 minutes. Find the probability a fill takes less than 5 minutes.

8. In a town, 40 of 500 dog owners reported their dog had fleas, while 12 of 300 cat owners reported fleas. Calculate the risk of fleas for dog owners and for cat owners.

Merit

9. A factory has two machines. Machine A makes 60% of items with a 3% defect rate; Machine B makes 40% with a 5% defect rate. An item is chosen at random and found to be defective. Find the probability it came from Machine A.

10. Using the flea data (dogs: risk 0.08, cats: risk 0.04), calculate the relative risk of fleas for dogs compared with cats and interpret it in context.

11. Bags of flour are filled to a mean of 1500 g with standard deviation 12 g, normally distributed. Bags under 1485 g are rejected. In a batch of 800 bags, estimate how many are rejected.

Excellence

12. A medical test for a condition is 96% accurate for people who have it (true positive) and gives a false positive 5% of the time. In the population, 2% of people have the condition. A person tests positive. Find the probability they actually have the condition, and comment on why the result may seem surprising.

13. Two events A and B satisfy P(A) = 0.5, P(B) = 0.4 and P(A and B) = 0.2. Determine whether A and B are independent, and prove your conclusion using probability rules.

14. The lifetimes of two brands of LED bulb are normally distributed. Brand X: mean 20000 h, σ = 2000 h. Brand Y: mean 22000 h, σ = 3500 h. A retailer wants bulbs lasting at least 18000 h. Which brand gives a higher chance of meeting this, and by how much? Justify whether a larger mean always guarantees a better result here.

Answer key (for parents and tutors)

1. A bag holds 5 red, 3 blue and 2 green marbles. One marble is drawn at random. Find P(blue).
   Total marbles = 5 + 3 + 2 = 10. P(blue) = 3 ÷ 10 = 0.3.
   Achieved: applies basic probability as favourable ÷ total.

2. In a survey of 200 people, 120 own a bike and of those 45 also own a scooter. Find P(owns a scooter given owns a bike).
   P(scooter given bike) = 45 ÷ 120 = 0.375.
   Achieved: applies the conditional probability method from a subgroup.

3. A spinner lands on 'win' with probability 0.4. It is spun twice. Draw/consider a tree and find P(win then win).
   Branches are independent, so multiply along the path: P(win, win) = 0.4 × 0.4 = 0.16.
   Achieved: multiplies along tree branches for independent events.

4. Heights are normally distributed with mean μ = 170 cm and standard deviation σ = 8 cm. Find the z-score for a height of 182 cm.
   $z = (x - \mu) \div \sigma = (182 - 170) \div 8 = 12 \div 8 = 1.5$.
   Achieved: standardises a value using the z-score formula.

5. At Avondale College, 300 students were asked if they walk to school. The two-way table shows: Junior students – 90 walk, 60 do not; Senior students – 50 walk, 100 do not. Find the probability a randomly chosen student walks to school.
   Total who walk = 90 + 50 = 140. Total students = 300. P(walks) = 140 ÷ 300 = 0.467 (3 s.f.).
   Achieved: reads a two-way table and computes a marginal probability.

6. A café finds 70% of customers order coffee. Of those who order coffee, 30% also order a muffin. Draw a tree for the first branch and find P(orders coffee and a muffin).
   P(coffee) = 0.7, P(muffin given coffee) = 0.3. Multiply along branches: P(coffee and muffin) = 0.7 × 0.3 = 0.21.
   Achieved: applies the multiplication rule along a tree path.

7. The time to fill a petrol tank is normally distributed with mean 4.5 minutes and standard deviation 0.8 minutes. Find the probability a fill takes less than 5 minutes.
   $z = (5 - 4.5) \div 0.8 = 0.625$. $P(z < 0.625) \approx 0.734$. So about 0.734 (73.4%).
   Achieved: standardises and reads a normal-distribution probability.

8. In a town, 40 of 500 dog owners reported their dog had fleas, while 12 of 300 cat owners reported fleas. Calculate the risk of fleas for dog owners and for cat owners.
   Risk (dogs) = 40 ÷ 500 = 0.08. Risk (cats) = 12 ÷ 300 = 0.04.
   Achieved: calculates risk as a relative frequency for two groups.

9. A factory has two machines. Machine A makes 60% of items with a 3% defect rate; Machine B makes 40% with a 5% defect rate. An item is chosen at random and found to be defective. Find the probability it came from Machine A.
   P(A and defective) = 0.60 × 0.03 = 0.018. P(B and defective) = 0.40 × 0.05 = 0.020. P(defective) = 0.018 + 0.020 = 0.038. P(A given defective) = 0.018 ÷ 0.038 = 0.474 (3 s.f.).
   Merit: links a tree diagram to conditional probability (reverse conditioning).

10. Using the flea data (dogs: risk 0.08, cats: risk 0.04), calculate the relative risk of fleas for dogs compared with cats and interpret it in context.
    Relative risk = risk(dogs) ÷ risk(cats) = 0.08 ÷ 0.04 = 2. Interpretation: dog owners in this town are twice as likely to report their pet having fleas as cat owners.
    Merit: computes relative risk and interprets it in context.

11. Bags of flour are filled to a mean of 1500 g with standard deviation 12 g, normally distributed. Bags under 1485 g are rejected. In a batch of 800 bags, estimate how many are rejected.
    $z = (1485 - 1500) \div 12 = -1.25$. $P(x < 1485) = P(z < -1.25) \approx 0.1056$. Expected rejects = 0.1056 × 800 ≈ 84 bags.
    Merit: links a normal probability to an expected count for a real batch.

12. A medical test for a condition is 96% accurate for people who have it (true positive) and gives a false positive 5% of the time. In the population, 2% of people have the condition. A person tests positive. Find the probability they actually have the condition, and comment on why the result may seem surprising.
P(have) = 0.02, P(positive given have) = 0.96, so P(have and positive) = 0.02 × 0.96 = 0.0192. P(not have) = 0.98, P(positive given not have) = 0.05, so P(not have and positive) = 0.98 × 0.05 = 0.049. P(positive) = 0.0192 + 0.049 = 0.0682. P(have given positive) = 0.0192 ÷ 0.0682 = 0.282 (3 s.f.). Comment: even with a fairly accurate test, only about 28% of positives truly have the condition, because the condition is rare, so the many false positives from the large healthy group outnumber the true positives.
Excellence: justifies a counter-intuitive result by reasoning about base rates.

13. Two events A and B satisfy P(A) = 0.5, P(B) = 0.4 and P(A and B) = 0.2. Determine whether A and B are independent, and prove your conclusion using probability rules.
For independence we require P(A and B) = P(A) × P(B). Compute P(A) × P(B) = 0.5 × 0.4 = 0.2, which equals P(A and B) = 0.2. Since they are equal, A and B are independent. Equivalently, P(A given B) = P(A and B) ÷ P(B) = 0.2 ÷ 0.4 = 0.5 = P(A), confirming B does not change the probability of A.
Excellence: justifies independence formally using two equivalent criteria.

14. The lifetimes of two brands of LED bulb are normally distributed. Brand X: mean 20000 h, σ = 2000 h. Brand Y: mean 22000 h, σ = 3500 h. A retailer wants bulbs lasting at least 18000 h. Which brand gives a higher chance of meeting this, and by how much? Justify whether a larger mean always guarantees a better result here.
Brand X: z = (18000 − 20000) ÷ 2000 = −1.0, P(x ≥ 18000) = P(z ≥ −1.0) ≈ 0.8413. Brand Y: z = (18000 − 22000) ÷ 3500 = −1.143, P(x ≥ 18000) = P(z ≥ −1.143) ≈ 0.8735. Brand Y is better here, by about 0.8735 − 0.8413 = 0.032 (about 3 percentage points). Justification: a larger mean does not automatically guarantee a better chance at a threshold, because the larger spread (σ) of Brand Y works against it. In this case Y's higher mean outweighs its larger σ at the 18000 h cut-off, but at a higher threshold the greater spread could reverse the advantage, the outcome depends on both mean and standard deviation relative to the threshold.
Excellence: generalises about the combined effect of mean and spread with justification.