NCEA Level 3 · AS 91585: Probability Concepts in Action

Name: _____ Date: _____

Hey! This standard is all about digging into probability, conditional probability, independence, risk, and reading two-way tables and trees. Work through each section in order, check your working against the answers, and pay attention to the 'reasoning' style questions in Merit and Excellence, because that's where the top grades live.

Revision summary

- Conditional probability: $P(A|B) = P(A \text{ and } B) \div P(B)$. It answers 'given B has happened, what's the chance of A?'
- Independence: A and B are independent if $P(A|B) = P(A)$, or equivalently $P(A \text{ and } B) = P(A) \times P(B)$. If these don't hold, the events are dependent.
- Two-way tables: read counts off the table, use row/column totals for conditional probabilities, and the grand total for unconditional ones.
- Tree diagrams: multiply probabilities ALONG branches; add across separate branches that give the same outcome. Second-layer branches are conditional probabilities.
- Relative risk = (risk for exposed group) ÷ (risk for baseline/unexposed group). Risk = proportion with the outcome. RR > 1 means increased risk, RR = 1 means no difference, RR < 1 means reduced risk.
- Absolute risk is the actual proportion (e.g. 0.02); relative risk is a ratio comparing two groups, always state which group is the baseline.
- Probability distribution: probabilities sum to 1. Expected value $E(X) = \Sigma(x \times P(X = x))$.
- Always define events clearly and check whether a probability is conditional (restricted to a subgroup) or unconditional (out of the whole population).

Key skills

1. A survey of 200 Avondale residents recorded whether they own a bike and whether they cycle to work. 60 own a bike; of those, 45 cycle to work. Find the probability that a randomly chosen bike owner cycles to work.

2. For two events, $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \text{ and } B) = 0.2$. Are A and B independent?

3. In a trial, 8% of people taking a new supplement caught a cold, compared with 20% of people not taking it. Calculate the relative risk of catching a cold for supplement-takers compared with non-takers.

4. A tree diagram shows $P(\text{rain}) = 0.3$, and $P(\text{bus is late} \mid \text{rain}) = 0.6$, $P(\text{bus is late} \mid \text{no rain}) = 0.1$. Find the probability the bus is late.

Achieved

5. At Vertex Academy, 150 students were asked about their transport. The two-way table: Drive, Year 12: 20, Year 13: 40; Take bus, Year 12: 50, Year 13: 40. Find the probability that a randomly selected student is a Year 13 student who drives.

6. A café's coffee machine fails on 5% of mornings. If it works, 90% of customers are served on time; if it fails, only 30% are served on time. Draw the tree in your head and find the probability a customer is served on time.

7. In a health study, 15 of 300 people who exercise regularly developed high blood pressure, compared with 48 of 400 people who don't. Calculate the relative risk for non-exercisers compared with exercisers.

8. The number of texts X a student receives in a lunch break has this distribution: P(X=0)=0.2, P(X=1)=0.3, P(X=2)=0.4, P(X=3)=k. Find k, then find the expected number of texts.

Merit

9. A factory in Penrose has two machines producing USB drives. Machine A makes 70% of drives with a 3% defect rate; Machine B makes 30% with a 6% defect rate. A drive is found defective. What is the probability it came from Machine B? Show your reasoning.

10. A study reports the relative risk of a skin condition for people with high sun exposure is 3.0 compared with low exposure. The risk for the low-exposure group is 0.02. A news headline says 'high sun exposure triples your chance of the condition, huge danger!'. Using the numbers, discuss whether the headline is misleading.

11. Of 500 Auckland commuters surveyed, 180 use a bike-share app. Of app users, 63 had a minor accident last year; of non-users, 64 had one. Determine whether having an accident is independent of using the app, showing the comparison.

Excellence

12. A medical test for a rare condition affecting 1 in 1000 people has sensitivity 99% (correctly positive if you have it) and specificity 95% (correctly negative if you don't). A person tests positive. Find the probability they actually have the condition, and explain why the answer seems surprisingly low.

13. Events A and B satisfy P(A) = a and P(B) = b, and are independent. Show that A and the complement of B (B') are also independent, giving a general algebraic argument.

14. A game charges $2 to play. You roll a fair die: rolling a 6 wins $10, rolling a 4 or 5 wins $3, otherwise nothing. Determine whether the game is worth playing in the long run, then find the winning payout the $10 prize would need to become for the game to be exactly fair (expected profit zero), keeping all else the same.

Answer key (for parents and tutors)

1. A survey of 200 Avondale residents recorded whether they own a bike and whether they cycle to work. 60 own a bike; of those, 45 cycle to work. Find the probability that a randomly chosen bike owner cycles to work.
   This is conditional, restricted to bike owners. P(cycles | owns bike) = 45 ÷ 60 = 0.75.
   Achieved-level fluency: applies the conditional probability formula from counts.

2. For two events, P(A) = 0.4, P(B) = 0.5, and P(A and B) = 0.2. Are A and B independent?
   Check P(A) × P(B) = 0.4 × 0.5 = 0.2. This equals P(A and B) = 0.2, so A and B are independent.
   Key skill: tests independence using the multiplication condition.

3. In a trial, 8% of people taking a new supplement caught a cold, compared with 20% of people not taking it. Calculate the relative risk of catching a cold for supplement-takers compared with non-takers.
   Relative risk = risk (exposed) ÷ risk (baseline) = 0.08 ÷ 0.20 = 0.4. Supplement-takers had 0.4 times the risk (a 60% reduction).
   Key skill: computes relative risk with a stated baseline group.

4. A tree diagram shows P(rain) = 0.3, and P(bus is late | rain) = 0.6, P(bus is late | no rain) = 0.1. Find the probability the bus is late.
   Multiply along branches and add: P(late) = P(rain)×P(late|rain) + P(no rain)×P(late|no rain) = 0.3×0.6 + 0.7×0.1 = 0.18 + 0.07 = 0.25.
   Key skill: combines tree branches using the law of total probability.

5. At Vertex Academy, 150 students were asked about their transport. The two-way table: Drive, Year 12: 20, Year 13: 40; Take bus, Year 12: 50, Year 13: 40. Find the probability that a randomly selected student is a Year 13 student who drives.
   Total = 20 + 40 + 50 + 40 = 150. Year 13 drivers = 40. P(Year 13 and drives) = 40 ÷ 150 = 4/15 ≈ 0.267.
   Achieved: selects and applies a method to read a joint probability from a two-way table.

6. A café's coffee machine fails on 5% of mornings. If it works, 90% of customers are served on time; if it fails, only 30% are served on time. Draw the tree in your head and find the probability a customer is served on time.
   P(on time) = P(works)×0.90 + P(fails)×0.30 = 0.95×0.90 + 0.05×0.30 = 0.855 + 0.015 = 0.87.
   Achieved: applies tree-diagram method with the law of total probability.

7. In a health study, 15 of 300 people who exercise regularly developed high blood pressure, compared with 48 of 400 people who don't. Calculate the relative risk for non-exercisers compared with exercisers.
   Risk (non-exercisers) = 48 ÷ 400 = 0.12. Risk (exercisers) = 15 ÷ 300 = 0.05. Relative risk = 0.12 ÷ 0.05 = 2.4. Non-exercisers were 2.4 times as likely to develop high blood pressure.
   Achieved: calculates relative risk with a clearly identified baseline.

8. The number of texts X a student receives in a lunch break has this distribution: P(X=0)=0.2, P(X=1)=0.3, P(X=2)=0.4, P(X=3)=k. Find k, then find the expected number of texts.
   Probabilities sum to 1: 0.2+0.3+0.4+k = 1, so k = 0.1. E(X) = 0×0.2 + 1×0.3 + 2×0.4 + 3×0.1 = 0 + 0.3 + 0.8 + 0.3 = 1.4 texts.
   Achieved: completes a probability distribution and computes expected value.

9. A factory in Penrose has two machines producing USB drives. Machine A makes 70% of drives with a 3% defect rate; Machine B makes 30% with a 6% defect rate. A drive is found defective. What is the probability it came from Machine B? Show your reasoning.
   P(defect) = 0.70×0.03 + 0.30×0.06 = 0.021 + 0.018 = 0.039. P(B | defect) = P(B and defect) ÷ P(defect) = 0.018 ÷ 0.039 ≈ 0.462. So about 46.2% of defective drives come from Machine B, even though B makes less of the total output, its higher defect rate lifts its share of defects.
   Merit: links tree probabilities to a reverse conditional (Bayes-style) and interprets the result.

10. A study reports the relative risk of a skin condition for people with high sun exposure is 3.0 compared with low exposure. The risk for the low-exposure group is 0.02. A news headline says 'high sun exposure triples your chance of the condition, huge danger!'. Using the numbers, discuss whether the headline is misleading.
    Risk (high exposure) = relative risk × baseline risk = 3.0 × 0.02 = 0.06. So the absolute risk rises from 2% to 6%, a 4 percentage-point increase. While 'triples' (relative risk 3) is technically correct, the actual chance is still only 6%,

so the headline exaggerates by quoting the relative risk without the low absolute risk. A more honest statement links both: the risk tripled but remains fairly small (6 in 100).
Merit: connects relative and absolute risk to critique a claim.

11. Of 500 Auckland commuters surveyed, 180 use a bike-share app. Of app users, 63 had a minor accident last year; of non-users, 64 had one. Determine whether having an accident is independent of using the app, showing the comparison.
P(accident) = (63+64) ÷ 500 = 127 ÷ 500 = 0.254. P(accident | app user) = 63 ÷ 180 = 0.35. Since P(accident | app user) = 0.35 ≠ P(accident) = 0.254, the events are NOT independent, app users have a higher accident rate, so accident and app use are dependent.
Merit: links a two-way table to a formal independence check and interprets it.

12. A medical test for a rare condition affecting 1 in 1000 people has sensitivity 99% (correctly positive if you have it) and specificity 95% (correctly negative if you don't). A person tests positive. Find the probability they actually have the condition, and explain why the answer seems surprisingly low.
Let D = has condition. P(D) = 0.001, P(not D) = 0.999. P(positive) = P(D)×0.99 + P(not D)×0.05 = 0.001×0.99 + 0.999×0.05 = 0.00099 + 0.04995 = 0.05094. P(D | positive) = 0.00099 ÷ 0.05094 ≈ 0.0194, about 1.9%. It is low because the condition is very rare: even a small 5% false-positive rate applied to the huge healthy majority (999 people) produces far more false positives than the true positives from the tiny affected group. The base rate dominates.
Excellence: justifies a counter-intuitive conditional result by reasoning about base rates.

13. Events A and B satisfy P(A) = a and P(B) = b, and are independent. Show that A and the complement of B (B') are also independent, giving a general algebraic argument.
Independence gives P(A and B) = ab. Now P(A and B') = P(A) − P(A and B) = a − ab = a(1 − b). Also P(A)×P(B') = a×(1 − b) = a(1 − b). Since P(A and B') = P(A)×P(B'), A and B' are independent for any a, b. This generalises: if A is independent of B, it is automatically independent of B'.
Excellence: constructs a general algebraic proof of a probability property.

14. A game charges $2 to play. You roll a fair die: rolling a 6 wins $10, rolling a 4 or 5 wins $3, otherwise nothing. Determine whether the game is worth playing in the long run, then find the winning payout the $10 prize would need to become for the game to be exactly fair (expected profit zero), keeping all else the same.
Let W = winnings before the $2 fee. P(W=10)=1/6, P(W=3)=2/6, P(W=0)=3/6. E(W) = 10×1/6 + 3×2/6 + 0 = 10/6 + 6/6 = 16/6 ≈ $2.67. Expected profit = E(W) − 2 = 2.67 − 2 = $0.67 per play, so it is worth playing (positive in the long run). For a fair game, expected profit = 0, so E(W) must equal $2. Let the '6' prize be x: x×1/6 + 3×2/6 = 2, so x/6 + 1 = 2, x/6 = 1, x = $6. Lowering the top prize from $10 to $6 makes the game exactly fair.
Excellence: applies expected value to evaluate a game and solves an inverse problem to generalise fairness.