2.7 Statistics: Probability Depth: 30-minute Pack

EDEXCEL 9MA0 · Calculator allowed · about 30 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

A-Level Statistics: Probability Depth Practice

Original material written for Revision Library, styled on Edexcel AS/A-level Mathematics (9MA0) Statistics content.

This pack extends core probability skills to full A-level depth: Venn diagrams and set notation for two and three events, tree diagrams for dependent and independent trials, conditional probability, testing and proving independence algebraically, and probability models involving an unknown to be found from a given equation, including quadratics arising from probability statements. All contexts and figures are original.

1 For two events C and D, P(C) = 0.4, P(D) = 0.3 and P(C or D) = 0.58.

(a) Determine, showing your working, whether C and D are independent. **(3)**

(b) State the value of P(D | C). **(1)**

(Total for Question 1 is 4 marks)

---

2 At a college with 200 students in total, E is the event that a randomly chosen student studies Economics and B is the event that a randomly chosen student studies Business. It is known that 84 students study Economics, 90 students study Business, and 37 students study both Economics and Business.

(a) Find the number of students who study neither Economics nor Business. **(2)**

(b) Find P(not E and B), the probability that a randomly chosen student studies Business but not Economics. **(2)**

(c) Find P(E | B). **(2)**

(Total for Question 2 is 6 marks)

© Revision Library | revisionlibrary.co.uk

3 A gym in Cardiff has 120 members. Y is the event that a randomly chosen member attends yoga classes and P is the event that a randomly chosen member attends Pilates classes. Of the 120 members: 28 attend yoga only, 34 attend Pilates only, 22 attend both yoga and Pilates, and 36 attend neither.

(a) Find P(Y). **(1)**

(b) Find P(Y | P). **(2)**

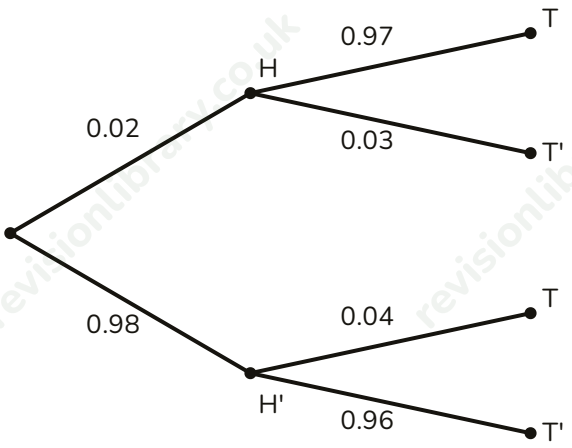
(c) Find P(P | not Y). **(2)**

(d) Determine, showing your working, whether Y and P are independent. **(2)**

(Total for Question 3 is 7 marks)

---

4 A screening test is used to detect a medical condition in a large population, in which 2% of people have the condition. Let H be the event that a randomly chosen person has the condition, and T be the event that the test gives a positive result. If a person has the condition, the test gives a positive result with probability 0.97. If a person does not have the condition, the test still incorrectly gives a positive result with probability 0.04.



(a) State one assumption needed for the tree diagram model above to be valid. **(1)**

(b) Find P(T). **(3)**

(c) Find the probability that a randomly selected person does not have the condition, given that their test result is positive. **(4)**

(d) Comment on the usefulness of this test as a diagnostic tool, with reference to your answer to part (c). **(3)**

(Total for Question 4 is 11 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.7 Statistics: Probability Depth: 30-minute Pack

Question 1

(a) P(C and D) = P(C) + P(D) - P(C or D) **M1**

(a) P(C and D) = 0.12, and P(C) x P(D) = 0.4 x 0.3 = 0.12 **A1**

(a) cso: since P(C and D) = P(C) x P(D) (both equal 0.12), C and D are independent **A1**

(a) Answer: Independent, since P(C and D) = 0.12 = P(C) x P(D)

(b) cao 0.3, using independence from part (a) so P(D | C) = P(D) **B1ft**

(b) Answer: P(D | C) = 0.3

Question 2

(a) n(E or B) = 84 + 90 - 37 **M1**

(a) cao 63, from 200 - 137 **A1**

(a) Answer: 63 students study neither Economics nor Business

(b) n(B) - n(E and B) = 90 - 37 **M1**

(b) cao 53/200 (= 0.265) **A1**

(b) Answer: P(not E and B) = 53/200 = 0.265

(c) n(E and B) / n(B), oe **M1**

(c) cao 37/90 (awrt 0.411) **A1**

(c) Answer: P(E | B) = 37/90 (awrt 0.411)

Question 3

(a) cao 50/120 (= 5/12, awrt 0.417) **B1**

(a) Answer: P(Y) = 5/12 (awrt 0.417)

(b) 22/(34+22), oe **M1**

(b) cao 11/28 (awrt 0.393) **A1**

(b) Answer: P(Y | P) = 11/28 (awrt 0.393)

(c) identify n(not Y) = 120 - 50 = 70 and n(P and not Y) = 34 **M1**

(c) cao 17/35 (awrt 0.486) **A1**

(c) Answer: P(P | not Y) = 17/35 (awrt 0.486)

(d) compute P(Y) x P(P) = (5/12) x (7/15), using P(P) = 56/120 = 7/15 **M1**

(d) cso: P(Y) x P(P) = 7/36 (awrt 0.194), which does not equal P(Y and P) = 11/60 (awrt 0.183), so Y and P are not independent **A1**

(d) Answer: Not independent, since P(Y)P(P) = 7/36 (awrt 0.194) not equal to P(Y and P) = 11/60 (awrt 0.183)

Question 4

(a) e.g. whether one person has the condition (or tests positive) is independent of whether any other person has the condition (or tests positive), oe; or the given sensitivity and false-positive rate apply equally to every person tested, oe **B1**

(a) Answer: For example: the test result for one person is independent of the test result for any other person

(b) P(H) x P(T|H) = 0.02 x 0.97 **M1**

(b) P(not H) x P(T|not H) = 0.98 x 0.04, and add the two products **M1**

(b) cao 0.0586 **A1**

(b) Answer: P(T) = 0.0586

(c) identify P(not H and T) = 0.0392 **M1**

(c) divide by P(T), following through from part (b) **M1ft**

(c) 0.0392/0.0586 or exact fraction 196/293 **A1**

(c) awrt 0.669 **A1**

(c) Answer: P(not H | T) = 196/293 (awrt 0.669)

(d) despite a positive result, there is still a probability of about 0.669 (roughly two out of three) that the person does not have the condition, oe **B1**

(d) this is because the condition is rare (only 2% prevalence), so false positives from the large healthy population outnumber true positives from the small group who have the condition, oe **B1**

(d) conclusion: a single positive result alone is not sufficient to diagnose the condition; further/confirmatory testing would be needed, oe **B1**

(d) Answer: The test is not very reliable on its own for a rare condition: a positive result still leaves about a 0.669 chance the person is unaffected, because far more healthy people are tested than affected people, so their false positives dominate; confirmatory testing is needed

© Revision Library | revisionlibrary.co.uk