

Assignment 1 – Part 2 – Solution

For Part1:

- Must be done on MyStatLab
- Due Date & Time: Tuesday, May 20, 2014 by 23:59 hrs.

For Part 2:

- Due Date & Time: Tuesday, May 20, 2014, in class
- Can be done in a group of at most 2 students, or individually.
- You MUST also attach the Printed/ Signed “Integrity Statement”.
- Use MiniTab or Excel if you have to. However you should cut and paste your result and sufficient explanation.
- You should provide detailed explanation for each of your solution. You must show your complete manual calculation if the question involves using formula or any kind of analysis.

Total Points: 40 points**Question 1. (12 marks)**

The mission of a research firm is to "explore the impact of the Internet on families, communities, work, home, and daily life." In August — September 2010 they conducted telephone interviews with a sample of Canadian adults aged 18 and older about online shopping. The following table displays some of the data collected by this research firm in their study.

a) List the variables in the data set. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, give the units.

Age	Gender	Household Income	Participated in online auction?	Monthly amount spent online	Time (hours) per week browsing online retailers
25	Male	\$55,000	Yes	\$250	6
47	Female	\$60,000	No	\$50	4
38	Female	\$110,000	No	\$120	10
30	Male	\$62,000	No	\$75	4

Answer: (6 marks)

Categorical: Gender, Participated in online auction?

Quantitative: Age (years), Household Income (\$), Monthly amount spent online (\$), Time per week browsing online retailers (hours).

b) In addition to the variables listed above, data were also collected on the variables listed below. Indicate whether each is nominal or ordinal.

- Region (Urban, Suburban, Rural)
- Education (Less than High School, High School, College, University)
- Marital Status (Single, Widowed, Divorced, Married)

Answer: (3 marks)

- Region (Urban, Suburban, Rural)
Nominal
- Education (Less than High School, High School, College, University)
Ordinal
- Marital Status (Single, Widowed, Divorced, Married)
Nominal

c) For each of the following, indicate whether the data are cross-sectional or time series:

- Weekly receipts at a clothing boutique
- Monthly demand for an automotive part
- Percentage of adults who bank online

Answer: (3 marks)

- Weekly receipts at a clothing boutique
Time Series
- Monthly demand for an automotive part
Time Series
- Percentage of adults who bank online
Cross-Sectional

Question 2. (5 marks)

For each of the following, indicate whether the data are nominal, ordinal, interval or ratio data.

a) A respondent of a survey indicates that he is a resident in the state of Ohio.

Answer : nominal data. (1 mark)

b) The Graduate Management Admission Test (GMAT) is a standardized test used by schools to determine the aptitude of individuals who are applying for MBA programs. The range of the GMAT score is 200-800. Brian has recently taken the exam and scored 720.

Answer : interval data. (1 mark)

c) A respondent of a survey indicates that she is currently in her junior year at the University of Texas.

Answer : ordinal data. (1 mark)

d) A respondent of a survey indicates that he owns the home that he currently resides.

Answer : nominal data. (1 mark)

e) The number of iPhones sold today at an Apple store

Answer : ratio data. (1 mark)

Question 3. (5 marks)

The following contingency table shows the number of customers who bought various brands of inkjet printers and their gender.

Printer Brand	Men	Women
HP	30	21
Canon	18	6
Epson	9	6
Brother	3	3
Other	24	30

a) What is the probability that a randomly selected customer from this sample bought a Canon printer or a Brother printer? (1 mark)

$$\text{Answer: } P(\text{Canon or Brother}) = \frac{24}{150} + \frac{6}{150} = \frac{30}{150} = 0.20$$

b) What is the probability that a randomly selected customer from this sample bought an HP printer or is a female? (1 mark)

$$\text{Answer: } P(\text{HP or Female}) = \frac{51}{150} + \frac{66}{150} - \frac{21}{150} = \frac{96}{150} = 0.64$$

c) The probability that a randomly selected customer from this sample bought an Epson printer and is a male ? (1 mark)

$$\text{Answer: } P(\text{Epson and Male}) = \frac{9}{150} = 0.06$$

d) Given that a randomly selected printer from this sample was a Canon, determine the probability that it was purchased by a man. (2 mark)

Answer:

$$P(\text{Man} | \text{Canon}) = \frac{P(\text{Man and Canon})}{P(\text{Canon})}$$

$$P(\text{Man and Canon}) = \frac{18}{150} = 0.12$$

$$P(\text{Canon}) = \frac{24}{150} = 0.16$$

$$P(\text{Man} | \text{Canon}) = \frac{0.12}{0.16} = 0.75$$

e) Given that a randomly selected printer from this sample was purchased by a female customer, determine the probability that the printer was an HP. (2 mark)

Answer:

$$P(\text{HP} | \text{Woman}) = \frac{P(\text{HP and Woman})}{P(\text{Woman})}$$

$$P(\text{HP and Woman}) = \frac{21}{150} = 0.14$$

$$P(\text{Woman}) = \frac{66}{150} = 0.44$$

$$P(\text{HP} | \text{Woman}) = \frac{0.14}{0.44} = 0.318$$

Question 4. (5 marks)

The following contingency table shows the number of adults with credit scores 725 or above and below 725 as well as their marital status.

Credit Score	Single (C)	Married (D)
Below 725 (A)	39	33
725 or Above (B)	15	63

a) Identify all the marginal probabilities for this table. (P(A), P(B), P(C), P(D)) (1 mark)

Answer:

Credit Score	Single (C)	Married (D)	Total
Below 725 (A)	0.26	0.22	0.48
725 or Above (B)	0.10	0.42	0.52
Total	0.36	0.64	1.00

$$P(A) = 0.48$$

$$P(B) = 0.52$$

$$P(C) = 0.36$$

$$P(D) = 0.64$$

b) Identify all the joint probabilities for this table. ($P(A \text{ and } C)$, $P(A \text{ and } D)$, $P(B \text{ and } C)$, $P(B \text{ and } D)$)
(1 mark)

Answer:

Credit Score	Single (C)	Married (D)	Total
Below 725 (A)	0.26	0.22	0.48
725 or above (B)	0.10	0.42	0.52
Total	0.36	0.64	1.00

$$P(A \text{ and } C) = 0.26$$

$$P(A \text{ and } D) = 0.22$$

$$P(B \text{ and } C) = 0.10$$

$$P(B \text{ and } D) = 0.42$$

c) Identify all the conditional probabilities for this table. ($P(A | C)$, $P(A | D)$, $P(B | C)$, $P(B | D)$)
(1 mark)

Answer:

Credit Score	Single (C)	Married (D)	Total
Below 725 (A)	0.26	0.22	0.48
725 or above (B)	0.10	0.42	0.52
Total	0.36	0.64	1.00

$$P(A | C) = \frac{P(A \text{ and } C)}{P(C)} = \frac{0.26}{0.36} = 0.722$$

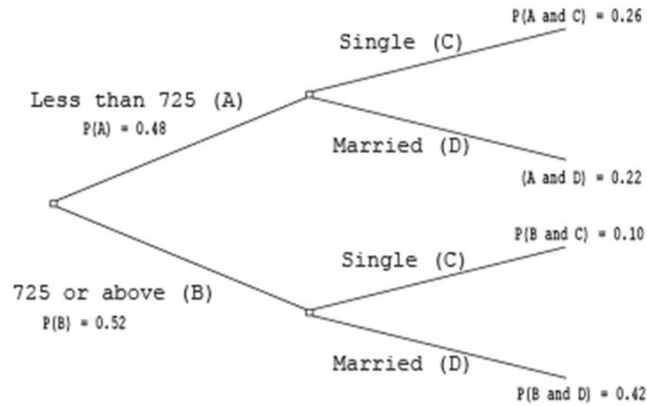
$$P(A | D) = \frac{P(A \text{ and } D)}{P(D)} = \frac{0.22}{0.64} = 0.344$$

$$P(B | C) = \frac{P(B \text{ and } C)}{P(C)} = \frac{0.10}{0.36} = 0.278$$

$$P(B | D) = \frac{P(B \text{ and } D)}{P(D)} = \frac{0.42}{0.64} = 0.656$$

d) Construct a decision tree for this table. (1 mark)

Answer:



e) Determine whether Events A and C are independent or dependent. What conclusions can be made? (1 mark)

Answer:

$$P(A) = 0.48$$

$$P(A | C) = \frac{P(A \text{ and } C)}{P(C)}$$

$$P(A \text{ and } C) = 0.26$$

$$P(C) = 0.36$$

$$P(A | C) = \frac{0.26}{0.36} = 0.722$$

$$P(A | C) \neq P(A)$$

Events A and C are dependent. Married adults appear to have higher credit scores than single adults.

Question 5. (1 mark)

According to a Rasmussen survey, 23% of adults responded that vanilla was their favorite ice cream flavor. A random sample of three adults was selected. Determine the probability that vanilla ice cream is the favorite flavor for all three adults.

Answer:

$$P(\text{Vanilla is favorite for all 3}) = 0.23^3 = 0.012$$

Question 6. (5 marks)

A college student population of freshman and seniors was surveyed to determine if they prefer traditional textbooks or digital textbooks. According to this sample, 60% of freshman students preferred digital textbooks. Forty percent of the sample consisted of freshman students. Also, 20% of senior students preferred digital textbooks. Given that a student prefers digital textbooks, use Bayes' Theorem to determine the probability that the student is a freshman.

Answer:

Define A_1 = student is a freshman, A_2 = student is a senior, B = student prefers a digital textbook.

$$P(A_1) = 0.40 \text{ (1 mark)}$$

$$P(A_2) = 1 - 0.40 = 0.60 \text{ (1 mark)}$$

$$P(B | A_1) = 0.60 \text{ (1 mark)}$$

$$P(B | A_2) = 0.20 \text{ (1 mark)}$$

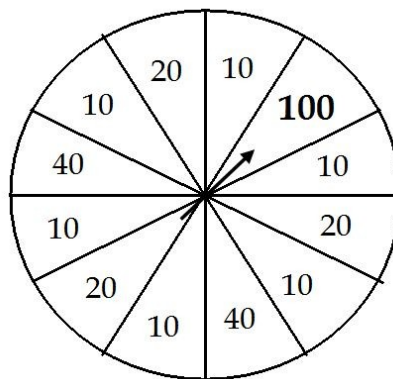
$$P(A_1 | B) = \frac{P(A_1) P(B|A_1)}{P(A_1) P(B|A_1) + P(A_2) P(B|A_2)} \text{ (1 mark)}$$

$$P(A_1 | B) = \frac{(0.40)(0.60)}{(0.40)(0.60) + (0.60)(0.20)}$$

$$P(A_1 | B) = \frac{0.24}{0.24 + 0.12} = \frac{0.24}{0.36} = 0.67$$

Question 7. (5 marks)

During its grand opening week, Stickler's bicycle shop offers a "wheel of discount savings". After customers select the items they wish to purchase, they spin the wheel to determine the discount they will receive. The wheel is divided into 12 slices. Six slices are red and award a 10% discount, three slices are white and award a 20% discount, and two slices are blue and award a 40% discount. The remaining slice is gold and awards a 100% discount!



a) What is the probability that a customer gets at least a 40% discount? (1 mark)

b) What is the probability that a customer does not get at least a 40% discount? (1 mark)

c) What is the probability that a customer gets a 10% or 20% discount? (1 mark)

d) What is the probability that two customers in a row get a 20% discount? (1 mark)

Answer:

a. 0.25 ($3/12$)

b. 0.75 ($1 - 0.25$) complement rule

c. 0.75 ($6/12 + 3/12 = 9/12$) addition rule with mutually exclusive events

d. 0.0625 ($3/12 \times 3/12$) multiplication rule

e) Suppose you visit Stickler's bicycle shop in the hopes of getting a 100% discount on your purchases. As you wait your turn in line, there are three gold winners in a row. The two customers in line behind you begin to discuss what's happened. One believes that the streak of three gold winners has killed anyone else's chances of getting a 100% discount, while the other says just the opposite... that the wheel's "hot streak" increases their chances of getting a 100% discount. Comment on these opinions. (1 mark)

Answer:

The spins are independent, so if the wheel is fair, the three gold winners in a row have no effect on the next person's chances.