SOLUTION

STAT 113 – Midterm 1 VERSION A

Note: Version B had slightly different numbers. But the basic problems were the same. You can recognize Version B by "Name" (instead of "Name:") on the top line, i.e., a missing ":" after "Name".

1. On questions 1a-f: [2pts] for the correct choice; [0pts] for no choice; [-1pt] for a wrong choice. If more than one choice is correct, any correct choice is fine.

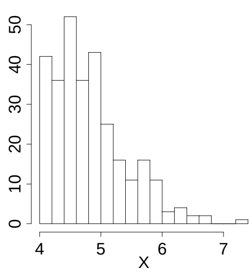
1a.

$$P(A \cap \underbrace{B \cap C}_{D}) = P(A|D)P(D) = P(A|B \cap C)\ \underbrace{P(B|C)P(C)}_{=P(D)=P(B\cap C)}$$

1b. By Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)\underbrace{[1 - P(A)]}_{=P(A^c)}}$$

1c & d. Consider the following histogram of $n = 300$ measurements: Denote with $\bar{x}$ the sample mean, with Md the sample median, and with s^2 the sample variance. Which of the following statements is correct?



$\bar{x} > Md$ and $s^2 < 4$.

1e. Peter is preparing for the final exam in his history of France course. The exam will consist of 5 essay questions selected at random from a list of 10 the professor has handed out in advance. Not exactly a Napoleon buff, Peter has only prepared eight of the questions. Let y denote the number of questions on the exam which Peter has prepared.

y is a hypergeometric r.v. with $N = 10$, $r = 8$, $n = 5$.

1f. $p(2) = \frac{1}{5}, p(-1) = \frac{4}{5}$ and hence

$$E(y) = 2\frac{1}{5} + (-1)\frac{4}{5} = -\frac{2}{5} = -0.4$$

2. Businesses commonly project revenues under alternative economic scenarios. For a stylized example, inflation could be high or low and unemployment could be high or low. There are four possible scenarios, with the assumed probabilities:

Scenario	Inflation	Unemployment	Probability
1	high	high	0.20
2	high	low	0.20
3	low	high	0.36
4	low	low	0.24

Hint: Denote with A the event "high inflation", and with B the event "high unemployment".

2a [3pts] $P(A) = P(A \cap B) + P(A \cap B^c) = 0.20 + 0.20 = 0.4$.

2b [4pts] $P(B) = P(A \cap B) + P(A^c \cap B) = 0.20 + 0.36 = 0.56$ and hence $P(A|B) = P(A \cap B)/P(B) = 0.20/0.56 = 0.36$.

2c [3pts] Are inflation and unemployment independent? No, because $P(A) = 0.4 \neq P(A|B) = 0.36$.

3. A family has two dogs (Rex and Rover) and a little boy (Russ). None of them is fond of the mailman. Given that they are outside, Rex and Rover have a 30% and a 40% chance, respectively, of biting the mailman. Russ, if he is outside, has a 15% chance of doing the same thing. Suppose only one of the three is outside when the mailman comes. Rex is outside 50% of the time, Rover 20% of the time and Russ 30% of the time.

3a [5pts] What is the probability the mailman will be bitten? Denote with A_1, A_2, A_3 the events that Rex, Rover and Russ, respectively are outside. Denote with B the event that the mailman is bitten.

$$P(B) = \sum_{i=1}^{3} P(B|A_i)P(A_i) = 0.3 \cdot 0.5 + 0.4 \cdot 0.2 + 0.15 \cdot 0.3 = 0.275$$

3b [5pts] *If* the mailman is bitten, what are the chances that Russ did it?

$$P(A_3|B) = \frac{P(B|A_3)P(A_3)}{\sum_{i=1}^{3} P(B|A_i)P(A_i)} = \frac{0.15 \cdot 0.3}{0.275} = 0.16$$

4. James Bond insists that his martinis be shaken, not stirred. A skeptical bartender tests Bond with 6 martinis (using six coin flips to determine which drinks to shake and which to stir). Bond errs on one and correctly identifies the other 5 before passing out. Denote with p the probability that Bond can tell the difference between shaken and stirred Martinis.

4a [3pts] If $p = 0.5$, what is the probability of guessing 5 or more Martinis correctly?

Let $y = \#$ correctly guesses. y is a binomial r.v. with $n = 6$, $p = 0.5$.
Using the tables in the appendix we find $P(y \geq 5) = 1 - \sum_{k=0}^{4} p(k) = 0.11$.
(or, use the Minitab command: `pdf; binomial n=6, p=0.5`)

4b [3pts] Find the value of p such that guessing 5 out of 6 Martinis correctly is highest

To find $\hat{p}$ such that $p(5)$ is maximized we consider $p(5)$ as a function of p and maximize it in p. It is easier to maximize $f(p) = \log p(5)$ instead of $p(5)$[1]:

$$f(p) = \log\left[c \cdot p^5(1-p)^1\right] = \log(c) + 5\log(p) + \log(1-p),$$

where c is a factor which does not depend on p (and is hence irrelevant for the maximization). To find the maximum of $f(p)$, get the first derivative, and set it equal to zero:

$$\frac{df}{dp} = \frac{5}{p} - \frac{1}{1-p} = 0 \qquad \Rightarrow p = \frac{5}{6}$$

4c [2pts] Assume that instead of deciding initially to serve 6 Martinis, the bartender was serving Martinis until Bond guessed 5 correctly. Assuming $p = 0.5$, find the probability of erring once.

Denote with x the number of trials until Bond correctly guesses five times. x is a negative binomial r.v. with $r = 5$ and $p = 0.5$. Erring once implies $x = 6$, and thus:

$$p_x(6) = 5 \cdot p^5(1-p)^1 = 5\frac{1}{2^6} = 5/64 = 0.08.$$

4d [2pts] Under the assumptions of 4c, find the value of p such that guessing 5 out of 6 Martinis correctly is hightest

The answer is the same as for 4b because $p_x(6) = 1 \cdot p^5(1-p)^1$ and $p_y(5) = c \cdot p^5(1-p)^1$, i.e., as a function of p the probability mass functions for x and y are proportional. Proportional functions have the same maximum.

[1] Since the logarithm is a monotone transformation the value p which maximizes the logarithm also maximizes the original function.