Homework #7, Mei Maddox.

Mei Maddox

Submit the solution on Canvas into the corresponding assignment (e.g. "Homework #1") in the form of R Markdown report, knitted into either of the available formats (HTML, pdf or Word). Provide only code and output. NO NEED TO COPY THE PROBLEM FORMULATION (!)

Problem 1

5.48 Consider your birthday is Jan 1st. Suppose the number of students in the class is 25.

a) The probability of finding at least one student with a birthday that matches yours is less than the probability of a match for at least any two students in the class (0.5686997) because you are specifying the date which must be matched.

b) The probability of finding at least one student with a birthday that matches yours in a class size of 25 is $1 - (\frac{364}{365})^{24} \approx 0.0637228$.

5.50

```
##    P(hole in one)
## 4          0.0005
## 6          0.0015
## 12         0.0005
## 16         0.0025
```

a) Because the probability is independent from one hole to the next, the probability of entire round is simply the product of all the probabilities for each hole. The listed probabilities is the complement of the probability of NO hole in one. Therefore, the probability is $1-P(4)\times 1-P(6)\times 1-P(12)\times 1-P(16) \approx 0.995008$

b) Because the probability is indepenent from game to game, the probability of no holes in one across 20 games is simply the product of the probability of no holes in one for one round 20 times. $P(\{NoHoleInOne|20Rounds\}) = P(\{NoHoleInOne|SingleRound\})^{20} \approx 0.9047559$.

c) The probability of making at least one hole in one across 20 rounds is the complement to the answer of part b. $P(\{AtLeastOneHoleInOne|20Rounds\}) = 1 - P(\{NoHoleInOne|20Rounds\}) \approx 0.0952441$.

5.57 Prevelence of breast cancer in women who get mammograms: $P(S) = 0.01$
Sensitivity: $P(POS|S) = 0.86$
Specificity: $P(NEG|S^c) = 0.88$

a) The prevalence of no breast cancer is $P(S^c) = 1 - P(S) = 1 - 0.01 = 0.99$

- $P(S \cap POS) = P(POS|S) \times P(S) = 0.86 \times 0.01 \approx 0.0086$

- $P(S \cap NEG) = P(NEG|S) \times P(S) = (1 - P(POS|S)) \times P(S) = 0.14 \times 0.01 \approx 0.0014$
- $P(S^c \cap POS) = P(POS|S^c) \times P(S^c) = (1 - P(NEG|S^c)) \times P(S^c) = 0.12 \times 0.99 \approx 0.1188$
- $P(S^c \cap NEG) = P(NEG|S^c) \times P(S^c) = 0.88 \times 0.99 \approx 0.8712$

b) $P(POS) = P(POS \cap S) + P(POS \cap S^c) \approx 0.1274$

c) The proportion of women who receive a positive mammogram and actually have breast cancer is $P(S|POS) = \frac{P(S \cap POS)}{P(POS)} \approx 0.0675039$.

d) You get the frequencies by multiplying the probability by the number of people. Although $P(S|POS) \approx 0.0675039$, the sample size of 100 women would suggest that $P(S|POS) = \frac{1}{13} \approx 0.0769231 \approx 0.08$.

5.62

```
##      POS  NEG
## Yes   48    6
## No  1307 3921
```

a) $P(Yes) = 54/5282 \approx 0.0102234$

b)
- i: Sensitivity $= P(POS|Yes) = \frac{P(POS \cap Yes)}{P(Yes)} = \frac{48}{54} \approx 0.8888889$
- ii: Specificity $= P(NEG|No) = \frac{P(NEG \cap No)}{P(No)} = \frac{3921}{5228} \approx 0.75$

c)
- i: $P(Yes|POS) = \frac{P(Yes \cap POS)}{P(POS)} = \frac{48}{1355} \approx 0.0354244$
- ii:$P(No|NEG) = \frac{P(No \cap NEG)}{P(NEG)} = \frac{3921}{3927} \approx 0.9984721$

d)
- $P(POS|Yes)$: Proportion of children who receive a positive diagnosis of down syndrome out of all sampled children who *ACTUALLY* have it.
- $P(NEG|No)$: Proportion of children who receive a negative diagnosis of down syndrome out of all sampled children who do *NOT* have it.
- $P(Yes|POS)$: Proportion of children who *ACTUALLY* have down syndrome out of all sampled children which receive a positive diagnosis.
- $P(No|NEG)$: Proportion of children who do *NOT* have down syndrome out of all sampled children which receive a negative diagnosis.

6.5

```
##   grade points   P
## 1     A      4 0.2
## 2     B      3 0.4
## 3     C      2 0.3
## 4     D      1 0.1
```

a) The probability distribution is for the grade point score of a randomly selected student is $\{P(X = 4) = 0.2, P(X = 3) = 0.4, P(X = 2) = 0.3, P(X = 1) = 0.1\}$

b) A student of this particular instructor receives $\mu = \sum_{i=1}^{4} x_i P(X = x_i) = 4 \times 0.2 + 3 \times 0.4 + 2 \times 0.3 + 1 \times 0.1 \approx 2.7$ points on an assignment, on average.

6.12 You bid \$30 for first item and \$20 for second item on eBay. The probability of winning the bid is 0.1 and 0.2, respectively. The two bids are independent events. Let X denote the random variable representing the total amount of money you will spend on the two items.

a) The sample space of winning or loosing the two bids is {WW,WL,LW,LL}.

b)

- $P(WW) = 0.1 \times 0.2 = 0.02$
- $P(WL) = 0.1 \times 0.8 = 0.08$
- $P(LW) = 0.9 \times 0.2 = 0.18$
- $P(LL) = 0.9 \times 0.8 = 0.72$

c) $P(X = 50) = 0.02, P(X = 30) = 0.08, P(X = 20) = 0.18, P(X = 0) = 0.72$

d) $\mu = 50 \times 0.02 + 30 \times 0.08 + 20 \times 0.18 + 0 \times 0.72 = 1 + 2.4 + 3.6 + 0 = 7$

Problem 2

1 Using formula $E[X] = \sum_{i=1}^{k} x_i \times P(X = x_i)$, derive the form of E[aX + b] as a function of a, b and E[X].

$E[aX + b]$
$= \sum_{i=1}^{k}(aX + b)_i \times P((aX + b)_i)$
$= \sum_{i=1}^{k}(ax_i + b) \times P(ax_i + b)$
$= \sum_{i=1}^{k} ax_i \times P(X = x_i) + b \times P(ax_i + b)$
$= \sum_{i=1}^{k} ax_i \times P(X = x_i) + \sum_{i=1}^{k} b \times P(ax_i + b)$
$= a\sum_{i=1}^{k} x_i \times P(X = x_i) + b\sum_{i=1}^{k} P(ax_i + b)$
$= a \times E[X] + b \times 1$
$= aE[X] + b$

2 Using formula $V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = \sum_i (x_i - E[X])^2 \times P(X = x_i)$ and result from part 1, derive the form of V [aX +b] as a function of a, b and V [X].

$V[aX + b]$
$= E[(aX + b)^2] - (E[(aX + b)])^2$
$= E[(aX + b)^2] - (E[(aX + b)])^2$
$= E[a^2X^2 + 2abX + b^2] - (aE[X] + b)(aE[X] + b)$
$= E[a^2X^2] + E[2abX] + b^2 - (a^2(E[X])^2 + 2abE[X] + b^2)$
$= a^2E[X^2] + 2abE[X] + b^2 - a^2(E[X])^2 - 2abE[X] - b^2$
$= a^2E[X^2] - a^2(E[X])^2$
$= a^2(E[X^2] - (E[X])^2)$
$= a^2 \times V[X]$