Week 3 Assignment Math - Puneet Auluck

Short Answers:

- **2.34**

  (a)

	X1	X2	X3	X4
x	0	5	10	30
p_x	0.500	0.250	0.231	0.019

  E(x) = 4.13, SD = 3.41

  (b) $4 maximum
- **2.40**

  (a)

	X1	X2	X3	X4
x	0	5	10	30
p_x	0.500	0.250	0.231	0.019

  E(x) = 12.7, SD = 10.09

  (b) E(x) = 407.28, SD = 324.42
- **2.42**

  (a) $72, SD = 6.40

  (b) $11.00, SD = 0.40
- **2.46**

  (a) Symmetrical

  (b) 0.62

  (c) 0.26

  (d) No, assumption not valid

Detailed Answers:

Below are some functions that will be used for first 2 questions.

```
getExpectedValue <- function(x,y) {
  return(sum(x * y))
}

getVar <- function (x, e, p){
  diff_xe_sq <- ((x*p)-e)^2
  return(sum(diff_xe_sq * p))
}

getSD <- function(x){
  return(sqrt(x))
}

buildPModel <- function(x,p){
  return(data.frame(rbind(ceiling(x), format((p),digits=2)),row.names =
c("x","p_x")))
}
```

2.34 Card game. Consider the following card game with a well-shuffled deck of cards. If you draw a red card, you win nothing. If you get a spade,you win $5. For any club, you win $10 plus an extra $20 for the ace of clubs.

(a) Create a probability model for the amount you win at this game. Also, find the expected winnings for a single game and the standard deviation of the winnings.

```
x <- c(0,5,10,30)
p_x <- c(26/52,13/52,12/52,1/52)
e <- getExpectedValue(x,p_x)
v <- getVar(x,e,p_x)
s <- getSD(v)

print(buildPModel(x,p_x))

##        X1    X2    X3    X4
## x       0     5    10    30
## p_x 0.500 0.250 0.231 0.019

sprintf("Average revenue %1.2f, variance: %1.2f, standard deviation: %1.2f",
e, v,s)

## [1] "Average revenue 4.13, variance: 11.64, standard deviation: 3.41"
```

(b) What is the maximum amount you would be willing to pay to play this game? Explain.

The maximum amount I am willing to pay is $4 which is near to the expected value. If I play this game 100 times, with the expected value of $4.13, I would profit $13.
If I play this game 100 times @ $5/game, I should expect a loss of $87.

2.40 Baggage Fees. An airline charges the following baggage fees: $25 for the first bag and $35 for the second. Suppose 54% of passengers have no checked luggage, 34% have one piece of checked luggage and 12% have two pieces. We suppose a negligible portion of people check more than two bags.

(a) Build a probability model, compute the average revenue per passenger, and compute the corresponding standard deviation.

```
x <- c(0,25,35)
p_x <- c(0.54,0.34, 0.12)

e <- getExpectedValue(x,p_x)
v <- getVar(x,e,p_x)
s <- getSD(v)


print(buildPModel(x,p_x))

##       X1   X2   X3
## x      0   25   35
## p_x 0.54 0.34 0.12
```

```r
sprintf("Average revenue per passenger %1.2f, variance: %1.2f, standard deviation: %1.2f", e, v,s)
```

```
## [1] "Average revenue per passenger 12.70, variance: 101.76, standard deviation: 10.09"
```

(b) About how much revenue should the airline expect for a flight of 120 passengers? With what standard deviation? Note any assumptions you make and if you think they are justified.

```r
x <- c(120*.54*0, 120*.34*25,120*.12*35)

e <- getExpectedValue(x,p_x)
v <- getVar(x,e,p_x)
s <- getSD(v)
sprintf("The expected revenue should be: $%1.2f, with standard deviation of: %1.2f", e, s)
```

```
## [1] "The expected revenue should be: $407.28, with standard deviation of: 324.42"
```

We make the assumption that the 120 passengers are random variables and none of the passengers are associated with each other.

2.42 Selling on Ebay. Marcie has been tracking the following two items on Ebay: A textbook that sells for an average of $110 with a standard deviation of $4 and Mario Kart for the Nintendo Wii, which sells for an average of $38 with a standard deviation of $5.

(a) Marcie wants to sell the video game and buy the textbook. How much net money (profits - losses) would she expect to make or spend? Also compute the standard deviation of how much she would make or spend.

```r
book_e <- 110
book_s <- 4
game_e <- 38
game_s <- 5
# Expected value of expenditure is: E(book - game) = E(book)-E(game)
sprintf("Marcie should expect to spend: $%1.2f", abs(110-38))
```

```
## [1] "Marcie should expect to spend: $72.00"
```

```r
# Standard deviation: (-1)^*var(book)+(1)^2*var(game)
sprintf("With standard deviation is %1.2f. ", sqrt((book_s^2)+(game_s^2)))
```

```
## [1] "With standard deviation is 6.40. "
```

(b) Lucy is selling the textbook on Ebay for a friend, and her friend is giving her a 10% commission (Lucy keeps 10% of the revenue). How much money should she expect to make? With what standard deviation?

```r
sprintf("Lucy should expect to make: $%1.2f, with standard deviation of: %1.2f", 0.10*book_e, sqrt(0.10^2*book_s^2))
```

```
## [1] "Lucy should expect to make: $11.00, with standard deviation of: 0.40"
```

2.46 Income and gender. The relative frequency table below displays the distribution of annual total personal income (in 2009 inflation-adjusted dollars) for a representative sample of 96,420,486 Americans. These data come from the American Community Survey for 2005-2009. This sample is comprised of 59% males and 41% females.

Income	*Total*
\$1 to \$9,999 or loss	2.2%
\$10,000 to \$14,999	4.7%
\$15,000 to \$24,999	15.8%
\$25,000 to \$34,999	18.3%
\$35,000 to \$49,999	21.2%
\$50,000 to \$64,999	13.9%
\$65,000 to \$74,999	5.8%
\$75,000 to \$99,999	8.4%
\$100,000 or more	9.7%

(a) Describe the distribution of total personal income.

The distribution appears to be symmetircal. It is heavily concentrated in the middle, meaning majority of the population's income is between 25K-50K.

(b) What is the probability that a randomly chosen US resident makes less than \$50,000 per year?

```
income <- c(2.2,4.7,15.8,18.3,21.2,13.9,5.8,8.4,9.7)*96420486/100
p_50less <-  sum(income[1:5])/(96420486)
sprintf("P(x<$50,000) is %1.2f. ", p_50less)

## [1] "P(x<$50,000) is 0.62. "
```

(c) What is the probability that a randomly chosen US resident makes less than \$50,000 per year and is female? Note any assumptions you make.

Assumption: propbability of person making less than 50K is independent of probability of being a female.

```
p_female <- 0.41
sprintf("P(x<$50,000 and female) is %1.2f:", p_50less * p_female)

## [1] "P(x<$50,000 and female) is 0.26:"
```

(d) The same data source indicates that 71.8% of females make less than \$50,000 per year. Use this value to determine whether or not the assumption you made in part (c) is valid.

No, because `P(f) = 0.41` should be same as `p(f) = P(female & less than 50K)*P(50K) = (0.78)*(0.622)=0.49`.