Joseph Blubaugh
jblubau1@tamu.edu
STAT 630-720
HW 01

Book Questions

1.2.3) Suppose $S = (1, 2, 3)$, with $P(1) = \frac{1}{2}$ and $P(1, 2) = \frac{2}{3}$. What must $P(2)$ be?

$P(2) = \frac{4}{6} - \frac{3}{6} = \frac{1}{6}$
$P(3) = 1 - P(1, 2) = \frac{1}{3}$

1.2.6) Label the subregions in the Venn diagram in Figure 1.2.4 using the sets A, B and C and their complements. Page 9.

  a) $A \cap B^c \cap C^c$
  b) $A \cap B \cap C^c$
  c) $A^c \cap B \cap C^c$
  d) $A \cap C \cap B^c$
  e) $A \cap B \cap C$
  f) $B \cap C \cap A^c$
  g) $C \cap B^c \cap A^c$

1.2.12) Suppose that $S = (1, 2, 3, 4)$ and $P(1) - \frac{1}{8} = P(2) = 3P(3) = 4P(4)$. Compute $P(1)$, $P(2)$, $P(3)$, and $P(4)$.

   $P(1) + P(2) + P(3) + P(4) = 1$, If we know the relationship between the probabilities, we can rewrite the equation and solve for x.

   $x - \frac{1}{8} + x + \frac{x}{3} + \frac{x}{4} = 1$
   $2x + \frac{x}{3} + \frac{x}{4} = 1 + \frac{1}{8}$
   $12(2x + \frac{x}{3} + \frac{x}{4}) = (\frac{9}{8})12$
   $24x + 4x + 3x = \frac{27}{2}$
   $31x = \frac{27}{2}$
   $x = \frac{27/2}{31} = \frac{27}{62}$

   $P(1) = \frac{27}{62} - \frac{1}{8} = .310$
   $P(2) = \frac{27}{62} = .435$
   $P(3) = \frac{27}{62}/3 = .145$
   $P(4) = \frac{27}{62}/4 = .109$

1.3.2) Suppose that Al watches the six o'clock news 2/3 of the time, watches the eleven o'clock news 1/2 of the time, and watches both the six o'clock and

the eleven o'clock news 1/3 of the time. a) For a randomly selected day, what is the probability that Al watches only the six o'clock news? b) For a randomly selected day, what is the probability that Al watches neither news?

$P(NewsAtSix) = \frac{2}{3}$, $P(NewsAtEleven) = \frac{1}{2}$, $P(Both) = \frac{1}{3}$

a) $P(NewsAtSixOnly) = P(NewsAtSix) - P(NewsAtSix \cap NewsAtEleven) = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$

b) $P(NoNews) = P(NewsAtSix^c \cap NewsAtEleven^c)$
$P(NewsAtSixOnly) = \frac{1}{3}$
$P(NewsAtElevenOnly) = P(NewsAtEleven) - P(NewsAtSix \cap NewsAtEleven) = \frac{1}{2} - \frac{1}{3} = \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$
$P(NoNews) = P(NewsAtSix^c \cap NewsAtEleven^c) = 1 - (\frac{1}{3} + \frac{1}{3} + \frac{1}{6}) = 1 - \frac{5}{6}$
$P(NoNews) = .166$

1.3.8) Suppose 55 percent of students are female, of which 4/5 (44 percent of all students) have long hair, 45 percent are male, of which 1/3 (15 percent of all students) have long hair. What is the probability that a student chosen at random will either be female or have long hair (or both)?

1) Probability of a student being female or having long hair: $.55 + .15 = .7$
2) Probability of a student being female with long hair: $.44$

1.3.10a) Generalize the principle of inclusion-exclusion, as follows. Suppose there are 3 events A, B, C. Prove that

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

The principle of inclusion shows how you calculate the total probability of all events in a set without double counting. In this case, $P(A \cap B), P(A \cap C), P(B \cap C), P(A \cap B \cap C)$ are all disjoint. Dealing with odd and even number events are different. With two event, you subtract the intersection, with 3 events, you subtract to double intersections but add back the triple intersection.

Example:
$P(A) = .08$, $P(B) = .08$, $P(C) = .08$
The intersections are all $= .01$
$P(A \cap B) = .01$, $P(A \cap C) = .01$, $P(B \cap C) = .01$, $P(A \cap B \cap C) = .01$

Using a venn diagram, summing up each section individually (where the non overlapping sections of A, B, C are now $= .05$), $P(A \cup B \cup C) = .19$. This is the same as using $P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$ where $.08 + .08 + .08 - .02 - .02 - .02 + .01 = .19$

1.4.1) Suppose that we roll eight fair six-sided dice.

a) what is the probability that all eight dice show a 6?

To get the probability of this combination, we need to know the total combinations that are possible. We can get this by multiplying the possible combinations from each die. A dice has 6 possible values $S = (1, 2, 3, 4, 5, 6)$, to get the total number of combinations, we have to multiply 6 times the number of dice, 8. 6^8 = 1,679,616, so there is a $\frac{1}{1,679,616}$ chance of this happening.

b) what is the probability that all eight dice show the same number?

Since there are 8 dice, there are only 6 possible outcomes where all 8 dice will have the same number. $\frac{8}{1,679,616}$ simplifies to $\frac{1}{209,952}$

c) what is the probability that the sum of the 8 dice is equal to 9?

In order for all dice to sum up to 9, 7 dice will have to roll a 1 and one die will have to roll a 2. There will be only 8 times when this can occur:
$S = (2, 1, 1, 1, 1, 1, 1, 1)$
$S = (1, 2, 1, 1, 1, 1, 1, 1)$
$S = (1, 1, 2, 1, 1, 1, 1, 1)$
. . .
The probability is the same as in b) $\frac{8}{1,679,616}$ simplifies to $\frac{1}{209,952}$

1.4.4) Suppose we are dealt five cards from an ordinary 52-card deck. What is the probability that:

a) We get all four aces, plus the king of spades?

Since the order does not matter, we use the binomial coefficient formula to get our combinations of unordered pairs: $\binom{N}{n} = \frac{N!}{n!(N-n)!}$, in this case $\frac{52!}{n!(52-5)!} = 2,598,960$. Since we are looking for just one combination the probability of drawing all 4 aces and the king of spades is $\frac{1}{2,598,960}$

b) All five cards are spades?

Since there are 13 cards in a suit, and we are looking for a specific suit: $\frac{13!}{5!(13-5)!} = 1,287$, so $\frac{1,287}{2,598,960} = 0.0004$

c) We get no pairs

To get the number of hands with no pairs, it may be easier to calculate the number of hands with pairs and subtract that number from the total number of possible hands. To get a pair you have to have one card that matches another card of a different suit, so $\binom{13}{1}\binom{4}{2} = 78$ so there are 78 possible pairs of 2 cards. To account for the remaining 3 cards for which the value doesn't matter we take $\binom{50}{3}$ so we get $\binom{13}{1}\binom{4}{2}\binom{50}{3} = 1,528,800$ hands with a pair, subtract that from the total number of unique hands and you get $2,598,960 - 1,528,800 = 1,070,160$, $\frac{1,070,160}{2,598,960} = .411$

d) We get a full house

To get a full house you need a 3 of a kind and a pair. To calculate the probability of a full house we can use $\binom{12}{1}\binom{4}{3} = 52$ for the 3 of a kind and $\binom{12}{1}\binom{4}{2} = 72$ for the pair. $52 * 72 = 3,744$, so the probability of a full house is $\frac{3,744}{2,598,960} = 0.0014$

1.4.11) Consider two urns, labeled urn #1 and urn #2. Suppose urn #1 has 5 red and 7 blue balls. Suppose that urn #2 has 6 red and 12 blue balls. Suppose we pick three balls uniformly at random for each urn. What is the probability that 6 chosen balls are the same color?

a) For Urn #1 Possible total combinations: $\binom{12}{3} = 220$
Possible combinations of red: $\binom{5}{3} = 10$, Probability of all red: $\frac{10}{220} = .045$
Possible combinations of blue: $\binom{7}{3} = 35$, Probability of all blue: $\frac{35}{220} = .159$

b) For Urn #2 Possible total combinations: $\binom{18}{3} = 816$
Possible combinations of red: $\binom{6}{3} = 20$, Probability of all red: $\frac{20}{816} = .0245$
Possible combinations of blue: $\binom{12}{3} = 220$, Probability of all blue: $\frac{220}{816} = .2696$

c) The total number of combinations from both urns: $\binom{12}{3}\binom{18}{3} = 179,520$
The total number of combinations from both urns being red: $\binom{5}{3}\binom{6}{3} = 200$, so $\frac{200}{179,520} = .001$
The total number of combinations from both urns being blue: $\binom{7}{3}\binom{12}{3} = 7,700$, so $\frac{7,700}{179,520} = .0428$

The probability of all 6 balls being the same color is $.001 + .0428 = .0438$

1.4.12) Suppose we roll a fair six-sided die and flip three fair coins. What is the probability that the total number of heads is equal to the number showing on the die?

There are $6 * 2 * 2 * 2 = 48$ possible combinations. The number of combinations where the number of heads is equal to the number on the die is 3, 1 for each possible head on any given role. So $\frac{3}{48} = .0625$

1.5.3) Suppose we flip 3 fair coins.

a) what is the probability that all three coins are heads?
$\frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8} = .125$

b) what is the conditional probability that all three coins are heads, conditional on knowing that the number of heads is odd?
$\frac{1}{2} * \frac{1}{2} = \frac{1}{4} = .25$

c) what is the conditional probability that all three coins are heads, given that the number of heads is even?

$\frac{1}{2} = .5$

1.5.7) Suppose that a baseball pitcher throws fastballs 80 percent of the time and curveballs 20 percent of the time. Suppose a batter hits a home run on 8 percent of all fastballs, and on 5 percent on all curveballs. What is the probability that this batter will hit a homerun on the pitchers next pitch?

Probability of hitting a homerun P(A) on a fastball P(B): $.08 * .9 = .072$ Probability of hitting a homerun P(A) on a curveball P(B): $.05 * .2 = .01$ The probability of hitting a homerun off the next pictch is $.082$

1.5.10) Consider two urns, urn #1 and urn #2. Urn #1 has 5 red and 7 blue balls. Urn #2 has 6 red and 12 blue balls. We pick 3 balls randomly free each urn. Conditional that all 6 chosen balls are red, what is the conditional probability that this color is red?

Urn1: $\frac{5}{12} * \frac{4}{11} * \frac{3}{10} = \frac{60}{1320} = .045$
Urn2: $\frac{6}{18} * \frac{5}{17} * \frac{4}{16} = \frac{120}{4896} = .024$

The conditional probability that all balls are red: $.045 * .024 = .001$

Additional Questions

A) Two 6-sided dice (red and green) are thrown and the values they show are recorded.

a) List the sample space. There are 36 outcomes.

S = {R1G1, R1G2, R1G3, R1G4, R1G5, R1G6, R2G1, R2G2, R2G3, R2G4, R2G5, R2G6, R3G1, R3G2, R3G3, R3G4, R3G5, R3G6, R4G1, R4G2, R4G3, R4G4, R4G5, R4G6, R5G1, R5G2, R5G3, R5G4, R5G5, R5G6, R6G1, R6G2, R6G3, R6G4, R6G5, R6G6}

b) List the outcomes that make up the following events: A = "the sum of the two values is at least 9", B = "the value of the red die is higher than the value of the green die" and C = "the green die has value 4".

s = {R3G6, R6G3, R4G5, R5G4}

c) List the elements of the following events: $A \cap C$, $B \cup C$, $A \cap (B \cup C)$.

$A \cap C$: Both A and C

$B \cup C$: B but not C, C but not B, or both B and C

$A \cap (B \cup C)$: A and B, A and C, B but not C, C but not B, or both B and C

d) Assume the outcomes are equally likely (they each have the same probability) and find the probabilities of the events in part (c).
   If all events are equally likely then the probability of an event happening is 1/36.
   $A \cap C = \frac{1}{36} * \frac{1}{36} = \frac{1}{72}$
   $B \cup C = \frac{1}{36}$ or $\frac{1}{72}$
   $A \cap (B \cup C) = \frac{1}{36}$ or $\frac{1}{72}$ or $\frac{1}{46656}$

e) Can P(A|C) be computed by multiplying the probabilities of A and C?
   yes, because they have the same probability

f) Imagine this experiment being repeated many times. What would be the longterm proportion of all the experiments for which the sum of the two dice is 7?
   List out the possible values: s = {R1G6, R2G5, R3G4, R4G3, R5G2, R6G1}. There are 6 outcomes where the dice could add to 7 so the expected probability is 1/6

B) Assume that grades are posted according to the last 4 digits of one's social security number ranging from 0000 to 9999 (I did this many years ago, but now this would be illegal)

a) Use R to estimate the probability that at least two students in a class of 100 share the same ID
   The estimated probability is .3919

b) Find the actual probability and compare it to your estimate (recall the birthday problem discussed in lecture).
   The actual probability is .3914

c) What is the smallest class enrollment for which the probability that at least two students have the same ID numbers is at least 0.5?
   The smallest class enrollment for a probability of at least .5 is 117

```
ss.match = function(sims, N, n) {

  ## create an empty object
  results = NULL

  for (i in 1:sims){

    ## create a random sample of 10000 possibilities
    random.event = sample(N, size = n, replace = TRUE)
```

```r
    ## sort and take the difference of each number
    difference = diff(sort(random.event))

    ## search for zero to find a match
    match = any(difference == 0)

    ## bind results to tabulate the probabily
    results = rbind(results, match)
  }

  ## devide the number of matches by the total simulations
  output = length(results[results == TRUE])/sims

  ## return the estimated probability
  return(output)

}

## function for estimating the probability of matching
## the last 4 digits of a social security number in a
## group of 100 people
ss.match(sims = 20000, N = 10000, n = 100)

## actual probability
1 - exp(lfactorial(10000) - lfactorial(10000 - 100) - 100 * log(10000))

table = NULL

for (i in 100:125) {

  ## estimate the probability for 100 students through 125 students
  results = data.frame("obs" = i, `probability' = ss.match(5000, 10000, i))
  table = rbind(table, results)

}

print(table)
```