Math 451 Test #3

Maxwell Levin

November 30, 2018

Problem 1

A deck consists of 52 cards labeled 1 to 52. You deal 5 cards without replacement at random from a well-shuffled deck of 52 cards. Let *X*, *Y*, and *Z* denote the minimum, median, and maximum, respectively, of the 5 observed numbers. The marginal pmf $\mathbf{f_X(x)}$ can be shown to be

$$\mathbf{f_X(x) = Pr\{X = x\} = \frac{\binom{52-x}{4}}{\binom{52}{5}}, x = 1, 2, \ldots, 48.}$$

To evaluate $\mathbf{E(X^2)}$, you may use *Mathematica* or some other numerical tool.

(a) Find E[X] and Var[X].

We calculate $E[X]$ by

$$E[X] = \sum_{x=1}^{48} x f_X(x),$$
$$= \sum_{x=1}^{48} x \frac{\binom{52-x}{4}}{\binom{52}{5}}.$$

This is difficult to compute manually, so instead I wrote the following python script to evaluate the expected value of X for me:

```
from scipy.misc import comb as ncr
from fractions import Fraction as frac
# The pmf of X
def f_X(x):
    return ncr(52 - x, 4) / ncr(52, 5)
expect = 0
for x in range(1, 49):
    expect += x * f_X(x)
print("E[X] = ", frac(expect).limit_denominator(), ", approx = {0:.3f}".format(expect))
```

```
E[X] =  53/6 , approx = 8.833
```

Thus we see that

$$E[X] = \frac{53}{6} \approx 8.833.$$

To calculate variance we take

$$Var[X] = E[X^2] - (E[X])^2,$$
$$= \sum_{x=1}^{48} x^2 \frac{\binom{52-x}{4}}{\binom{52}{5}} - \left(\sum_{x=1}^{48} x \frac{\binom{52-x}{4}}{\binom{52}{5}} \right)^2.$$

We again compute this by writing some python code:

```python
# This is E[X^2]
expect_2 = 0
for x in range(1, 49):
    expect_2 += x * x * f_X(x)
var_x = expect_2 - expect**2
print("Var[X] = ", frac(var_x).limit_denominator(), ", approx = {0:.3f}".format(var_x))
```

```
Var[X] =  12455/252 , approx = 49.425
```

Thus we see that the variance of X is

$$Var[X] = \frac{12455}{252} \approx 49.425.$$

(b) Find the probability mass function $\mathbf{f_Z(z)}$ of Z, $\mathbf{E[Z]}$, and $\mathbf{Var[Z]}$.

It took me a while to understand how to setup the pmf $f_Z(z)$ of Z for this problem. After talking to Yung-Pin I learned that we can think of X and Z in terms of combinatorics. We have $\binom{1}{1}$ ways of picking the z called for by $f_Z(z)$. Once the z is chosen, we need to pick four cards that are (strictly) smaller than Z. We have $\binom{z-1}{4}$ ways of doing this. Finally we have $\binom{52}{5}$ ways of picking 5 cards without replacement from our deck. Thus our pmf is given by

$$f_Z(z) = \frac{\binom{z-1}{4}}{\binom{52}{5}}, z = 5, 6, \ldots, 52.$$

Note the support for Z starts at 5 instead of 1 because we require a hand of five cards and four of those must be smaller than z. Thus z must be at least 5.

We can now calculate $E[Z]$ by

$$E[Z] = \sum_{z=5}^{52} z \frac{\binom{z-1}{4}}{\binom{52}{5}},$$

which we again calculate using python:

```python
# The pmf of Z
def f_Z(z):
    return ncr(z - 1, 4) / ncr(52, 5)

expect = 0
for z in range(5, 53):
    expect += z * f_Z(z)
print("E[Z] = ", frac(expect).limit_denominator(), ", approx = {0:.3f}".format(expect))
```

```
E[Z] =  265/6 , approx = 44.167
```

Thus we see that

$$E[Z] = \frac{265}{6} \approx 44.167.$$

To calculate variance we take

$$Var[Z] = E[Z^2] - (E[Z])^2.$$

We follow the same drill and compute this in python:

```python
expect_2 = 0
for z in range(5, 53):
    expect_2 += z * z * f_Z(z)
var_z = expect_2 - expect**2
print("Var[Z] = ", frac(var_z).limit_denominator(), ", approx = {0:.3f}".format(var_z))
```

```
Var[Z] =  12455/252 , approx = 49.425
```

Thus we see that the variance of Z is

$$Var[Z] = \frac{12455}{252} \approx 49.425.$$

Note that this is exactly te same as the variance for X.

(c) Find the joint prbability mass function $\mathbf{f_{X,Z}(x,z)}$ of $\boldsymbol{X}$ and $\boldsymbol{Z}$. Clearly state the support of the joint pmf and find $\mathbf{E[XZ]}$.

To solve this problem I again thought it from a perspective of combinatorics. We have $\binom{1}{1}$ ways to choose the x specified by $f_{X,Z}(x,z)$ and likewise we have $\binom{1}{1}$ ways to choose z, as it is also fixed from our inputs. Now we need to pick three cards in between x and z. Since there are $z-x-1$ numbers in between x and z the number of ways to pick these cards is $\binom{z-x-1}{3}$. Thus we have the pmf

$$f_{X,Z}(x,z) = \frac{\binom{z-x-1}{3}}{\binom{52}{5}}.$$

Note that we still have the requirements

$$x = 1, 2, \ldots, 48.$$

$$z = 5, 6, \ldots, 52.$$

And that we also require $z - x - 1 \geq 3$, ie

$$z \geq x + 4.$$

To find the expected value $E[XZ]$ we take

$$E[XZ] = \sum_{x=1}^{48} \sum_{z=5, z\geq x+4}^{52} xz \frac{\binom{z-x-1}{3}}{\binom{52}{5}}.$$

We compute this in python with the following script:

```python
# The joint pmf of X and Z
def f_XZ(x,z):
  return ncr(z-x-1, 3) / ncr(52, 5)
expect = 0
for x in range(1, 49):
  for z in range(5, 53):
    if z < x + 4:
      continue
    expect += x * z * f_XZ(x, z)
print("E[XZ] = ", frac(expect).limit_denominator(), ", approx = {0:.3f}".format(expect))
```

```
E[XZ] =  16801/42 , approx = 400.024
```

Thus we see that the expected value is approximately given by

$$E[XZ] \approx 400.024.$$

(d) The quantity $\mathbf{Z - X}$ is called the *range*. Use the results in (a), (b), and (c) to find $\mathbf{Var[Z - X]}$.

Recall that we can write

$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y] = 2abCov(X, Y),$$

and that

$$Cov(X, Y) = E[XY] - E[X]E[Y].$$

For $Var[Z - X]$ this translates to

$$Var[Z - X] = Var[Z] + Var[X] - 2Cov(Z, X),$$
$$= Var[Z] + Var[X] - 2(E[XZ] - E[X]E[Z]).$$

We can use our results from (a), (b), and (c) to plug in for $E[X], E[Z], Var[X], Var[Z],$ and $E[XZ]$ to find that

$$Var[Z - X] = \frac{12455}{252} + \frac{12455}{252} - 2\left(\frac{16801}{42} - \frac{265}{6}\frac{53}{6}\right) \approx 79.079.$$

(e) Find the conditional pmf $\mathbf{f_{X|Y}(x|Y = y)}$ of $\boldsymbol{X}$ given $\boldsymbol{Y}$. Clearly state the support of this conditional pmf and use it to find $\mathbf{E[X|Y = y]}$ and $\mathbf{Var[X|Y = y]}$.

Note that we can find this conditional probability in parts by using the fact that

$$f_{X|Y}(x|Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

We first calculate $f_{X,Y}(x, y)$ by

$$f_{X,Y}(x, y) = \frac{\binom{1}{1}\binom{1}{1}\binom{x-1}{0}\binom{y-x-1}{1}\binom{52-y}{2},}{\binom{52}{5}}$$
$$= \frac{(y - x - 1)\binom{52-y}{2}}{\binom{52}{5}}.$$

Note that $x = 1, 2, \ldots, 48, y = 3, 4, \ldots, 50,$ and $y \geq x + 2$ are the three requirements we must satisfy for this joint pmf to hold. We now find $f_Y(y)$:

$$f_Y(y) = \frac{\binom{1}{1}\binom{y-1}{2}\binom{52-y}{2}}{\binom{52}{5}},$$
$$= \frac{\binom{y-1}{2}\binom{52-y}{2}}{\binom{52}{5}}.$$

Here the support for Y remains the same: $y = 3, 4, \ldots, 50$. We can now compute $f_{X|Y}(x|Y = y)$ by

$$f_{X|Y}(x|Y = y) = \frac{\frac{(y-x-1)\binom{52-y}{2}}{\binom{52}{5}}}{\frac{\binom{y-1}{2}\binom{52-y}{2}}{\binom{52}{5}}},$$
$$= \frac{y - x - 1}{\binom{y-1}{2}},$$

$$= \frac{2(y-x-1)}{(y-1)(y-2)}.$$

Thus we've found our conditional pmf of X given Y. We now calculate the expected value of X given Y using the following from classnotes 31:

$$E[X|Y=y] = \sum_{\text{all } x} x f_{X|Y}(x|Y=y).$$

That is,

$$E[X|Y=y] = \sum_{x=1}^{y-2} \frac{2x(y-x-1)}{(y-1)(y-2)},$$

$$= \frac{2}{(y-1)(y-2)} \sum_{x=1}^{y-2} xy - x^2 - x,$$

$$= \frac{2}{(y-1)(y-2)} \left((y-1) \sum_{x=1}^{y-2} x - \sum_{x=1}^{y-2} x^2 \right)$$

$$= \frac{2}{(y-1)(y-2)} \left((y-1) \left(\frac{(y-2)(y-1)}{2} \right) - \frac{(y-2)(y-1)(2(y-2)+1)}{6} \right),$$

$$= y - 1 - \frac{2y-3}{3},$$

$$E[X|Y=y] = \frac{y}{3}$$

To calculate the variance we take

$$Var[X|Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2,$$

$$E[X^2|Y=y] = \sum_{x=1}^{y-2} \frac{2x^2(y-x-1)}{(y-1)(y-2)},$$

$$= \frac{2}{(y-1)(y-2)} \left((y-1) \sum_{x=1}^{y-2} x^2 - \sum_{x=1}^{y-2} x^3 \right),$$

$$= \frac{2}{(y-1)(y-2)} \left((y-1) \left(\frac{(y-2)(y-1)(2y-3)}{6} \right) - \left(\frac{(y-2)(y-1)}{2} \right)^2 \right)$$

$$= \frac{(y-1)(2y-3)}{3} - \frac{(y-1)(y-2)}{2},$$

$$= \frac{4y^2-10y+6}{6} - \frac{3y^2-9y+6}{6},$$

$$= \frac{y^2-y}{6},$$

$$E[X^2|Y=y] = \frac{y(y-1)}{6}.$$

Thus the variance is given by

$$Var[X|Y=y] = \frac{y(y-1)}{6} - \left(\frac{1}{3}y\right)^2 = \frac{y(y-3)}{18}.$$

(f) If you deal the entire well-shuffled deck without replacement and record the observed values by $\mathbf{V_1, V_2, \ldots, V_{52}}$, find the correlation coefficient $\rho(\mathbf{V_1 + V_2 + \cdots + V_{26}, V_{27}, V_{28}, \ldots V_{52}})$.

We know from class 28 that ρ represents the strength of the linear relationship between two random variables and that $\rho = 1$ indicates a perfectly positive linear relationship while $\rho = -1$ indicates a perfectly negative linear relationship.

Let $A = V_1 + V_2 + \cdots + V_{26}$ and let $B = V_{27} + V_{28} + \cdots + V_{52}$. Since $A + B = \sum_{i=1}^{52} i$ we know that $A = -B + \sum_{i=1}^{52} i$. Thus A and B have a perfectly negative linear relationship, i.e. $\rho(A, B) = -1$. That is,

$$\rho(V_1 + V_2 + \cdots + V_{26}, V_{27}, V_{28}, \ldots V_{52}) = -1.$$

Problem 2

Alice and Bob go fishing. On a typical fishing trip the time X in hours until Alice catches her first fish can be modeled by an exponential distribution with mean $\beta_{\mathbf{A}}$ hours/fish, and the time Y in hours until Bob catches his first fish can be modeled by an exponential distribution with mean $\beta_{\mathbf{B}}$ hours/fish. Assume their fishing times X and Y are independent.

(a) Find the probability that Alice will catch her first fish before Bob does. That is, find $\mathbf{Pr\{X < Y\}}$.

Note that since X and Y are exponentially distributed we know that they have the following pdf's:

$$X \sim f_X(x) = \frac{1}{\beta_A} e^{-x/\beta_A},\ 0 < x < \infty,$$

$$Y \sim f_Y(y) = \frac{1}{\beta_B} e^{-y/\beta_B},\ 0 < y < \infty.$$

Note that we can express $Pr\{X < Y\}$ as

$$\int_0^\infty \int_0^y f_X(x) f_Y(y) dx dy,$$

i.e.

$$\begin{aligned}
Pr\{X < Y\} &= \int_0^\infty \int_0^y \frac{1}{\beta_A} e^{-x/\beta_A} \cdot \frac{1}{\beta_B} e^{-y/\beta_B} dx dy, \\
&= \int_0^\infty \frac{1}{\beta_B} e^{-y/\beta_B} \left(-e^{-x/\beta_A} \Big|_0^y \right) dy, \\
&= \int_0^\infty \frac{1}{\beta_B} e^{-y/\beta_B} - \frac{1}{\beta_B} e^{-y(1/\beta_A + 1/\beta_B)} dy \\
&= -e^{-y/\beta_B} \Big|_0^\infty + \left(\frac{1}{\beta_B} \left(\frac{1}{\beta_A} + \frac{1}{\beta_B} \right)^{-1} e^{-y(1/\beta_A + 1/\beta_B)} \right) \Big|_0^\infty, \\
&= 1 - \frac{1}{\beta_B} \cdot \frac{1}{\frac{1}{\beta_A} + \frac{1}{\beta_B}}, \\
&= 1 - \frac{1}{\beta_B} \cdot \frac{\beta_A \beta_B}{\beta_A + \beta_B}, \\
&= 1 - \frac{\beta_A}{\beta_A + \beta_B}, \\
&= \frac{\beta_B}{\beta_A + \beta_B}.
\end{aligned}$$

(b) Let $\mathbf{U} = \mathbf{min(X, Y)}$ denote the time when the first fish is caught by either Alice or Bob, whichever occurs first. Find the pdf of U.

Note that $F_U(u) = Pr\{\text{A fish is caught by time } u\}$ and that $f_U(u) = \frac{d}{du}[F_U(u)]$. We use these notions to find $f_U(u)$, but we first search for $F_U(u)$:

$$
\begin{aligned}
F_U(u) &= 1 - Pr\{\text{A fish has not been caught by time } u\}, \\
&= 1 - Pr\{X > u \cap Y > u\}.
\end{aligned}
$$

Note that since X and Y are independent we can write this as

$$
\begin{aligned}
F_U(u) &= 1 - Pr\{X > u\}Pr\{Y > u\}, \\
&= 1 - \left(\int_u^\infty f_X(x)dx\right)\left(\int_u^\infty f_Y(y)dy\right), \\
&= 1 - \left(\int_u^\infty \frac{1}{\beta_A} e^{-x/\beta_A} dx\right)\left(\int_u^\infty \frac{1}{\beta_B} e^{-y/\beta_B} dy\right), \\
&= 1 - \left(-e^{x/\beta_A}\Big|_u^\infty\right)\left(-e^{y/\beta_B}\Big|_u^\infty\right), \\
&= 1 - e^{-u/\beta_A} e^{-u/\beta_B}, \\
F_U(u) &= 1 - e^{-u(1/\beta_A + 1/\beta_B)}.
\end{aligned}
$$

Note that this makes sense for $0 < u < \infty$. We now take the derivative with respect to u to get the pdf $f_U(u)$:

$$
\begin{aligned}
f_U(u) &= \frac{d}{du}\left[1 - e^{-u(1/\beta_A + 1/\beta_B)}\right], \\
f_U(u) &= \left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right) e^{-u(1/\beta_A + 1/\beta_B)}, \ 0 < u < \infty.
\end{aligned}
$$

(c) Does the random variable U defined in part (b) have the memoryless property? Justify your answer.

Note that U would satisfy the memoryless property if for some times a and b greater than zero the equality below holds:

$$Pr\{U > a + b | U > a\} = Pr\{U > b\}.$$

This can also be written as

$$\frac{Pr\{U > a + b\}}{Pr\{U > a\}} = Pr\{U > b\}.$$

In part (b) we relied on the fact that

$$F_U(u) = 1 - Pr\{U > u\}.$$

We now express this fact as

$$Pr\{U > u\} = 1 - F_U(u).$$

We now find $Pr\{U > a + b\}$, $Pr\{U > a\}$, and $Pr\{U > b\}$:

$$Pr\{U > a + b\} = 1 - F_U(a + b) = exp\left(-(a + b)\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right),$$

$$Pr\{U > a\} = 1 - F_U(a) = exp\left(-a\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right),$$

$$Pr\{U > b\} = 1 - F_U(b) = exp\left(-b\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right).$$

Now note that our memoryless condition becomes

$$\frac{Pr\{U > a+b\}}{Pr\{U > a\}} = \frac{exp\left(-(a+b)\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right)}{exp\left(-a\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right)},$$

$$= exp\left(-b\left(\frac{1}{\beta_A} + \frac{1}{\beta_B}\right)\right) = Pr\{U > b\}.$$

Thus our memoryless condition is satisfied, so U has the memoryless property.

(d) Let $\lceil \mathbf{X} \rceil$ denote the smallest integer greater than or equal to *x*. Let $\mathbf{V} = \lceil \mathbf{x} \rceil$. Find the pmf of *V*.

Note that the pmf of V is given by

$$f_V(v) = Pr\{V = v\} = Pr\{\lceil X \rceil = v\} = Pr\{v - 1 < X \leq v\},$$

which we can obtain by integrating the pdf of X:

$$Pr\{v - 1 < X \leq v\} = \int_{v-1)}^{v} \frac{1}{\beta_A} e^{-x/\beta_A} dx,$$

$$= -e^{-x/\beta_A}\Big|_{v-1}^{v} = e^{-(v-1)/\beta_A} - e^{-v/\beta_A},$$

$$f_V(v) = e^{-v/\beta_A}(e^{1/\beta_A} - 1),\ v = 1, 2, \ldots$$

Thus we've found the pmf of V.

(e) Does the random variable *V* defined in part (d) have the memoryless property? Justify your answer.

Note that the memoryless condition here would be:

$$\frac{Pr\{V > a+b\}}{Pr\{V > a\}} = Pr\{V > b\},$$

where a and b are two times in hours. To compute these probabilities we will need to find the cumulative distribution function of V, which we obtain by taking the summation (V is a discrete random variable):

$$F_V(v) = \sum_{k=1}^{v} f_V(k),$$

$$= \sum_{k=1}^{v} (e^{1/\beta_A} - 1)e^{-k/\beta_A}.$$

I asked Wolfram Alpha to compute this sum for me and found that

$$F_V(v) = 1 - e^{-v/\beta_A},\ v = 1, 2, \ldots$$

We now compute $Pr\{V > a+b\}, Pr\{V > a\},$ and $Pr\{V > b\}$:

$$Pr\{V > a+b\} = 1 - F_V(a+b) = e^{-(a+b)/\beta_A},$$

$$Pr\{V > a\} = 1 - F_V(a) = e^{-a/\beta_A},$$
$$Pr\{V > b\} = 1 - F_V(b) = e^{-b/\beta_A}.$$

Now note that our memoryless condition becomes

$$\frac{Pr\{V > a+b\}}{Pr\{V > a\}} = \frac{e^{-(a+b)/\beta_A}}{e^{-a/\beta_A}} = e^{-b/\beta_A} = Pr\{V > b\}.$$

Thus our memoryless condition is satisfied.

Problem 3

Suppose 7 males and 9 females enter an elevator at the ground floor (level 0) of a dorm. There are 10 floors (levels 1 through 10) above the ground floor. Each male passenger would stop at an odd-numbered floor twice as likely as an even-numbered floor. Each female passenger would stop at an even-numbered floor twice as likely as an odd-numbered floor. Suppose all passsengers stop at a floor/leave the elevator independently.

(a) Let W denote the number of stops that the elevator would make until all passengers leave the elevator. Find E[W] and Var[W].

Let $W = W_1 + W_2 + \cdots + W_{10}$, where

$$W_i = \begin{cases} 1, & \text{if someone got off on floor } i; \\ 0, & \text{otherwise.} \end{cases}$$

Let m be the probability that male j (just some individual male) gets off on floor i if i is even, and let $2m$ be the probability that male j gets off on floor i if i is odd. Since $5m + 5(2m) = 1, m = \frac{1}{15}$. Thus the probability male j gets off on floor i is given by

$$f_{M_j}(i) = \begin{cases} 1/15, & \text{if } i \text{ is even;} \\ 2/15, & \text{if } i \text{ is odd.} \end{cases}$$

By symmetry we have the following probability mass function for female j:

$$f_{F_j}(i) = \begin{cases} 2/15, & \text{if } i \text{ is even;} \\ 1/15, & \text{if } i \text{ is odd.} \end{cases}$$

We now make note the following probabilities, as we'll use these throughout this problem:

$$Pr\{W_i = 0 | i \text{ is even}\} = \left(1 - \frac{1}{15}\right)^7 \left(1 - \frac{2}{15}\right)^9 = \left(\frac{14}{15}\right)^7 \left(\frac{13}{15}\right)^9$$

$$Pr\{W_i = 1 | i \text{ is even}\} = 1 - \left(\frac{14}{15}\right)^7 \left(\frac{13}{15}\right)^9 \approx 0.82981$$

$$Pr\{W_i = 0 | i \text{ is odd}\} = \left(1 - \frac{2}{15}\right)^7 \left(1 - \frac{1}{15}\right)^9 = \left(\frac{13}{15}\right)^7 \left(\frac{14}{15}\right)^9$$

$$Pr\{W_i = 1 | i \text{ is odd}\} = 1 - \left(\frac{13}{15}\right)^7 \left(\frac{14}{15}\right)^9 \approx 0.80262$$

We can now begin our calculation of $E[W]$:

$$E[W] = E[W_1 + W_2 + \cdots + W_{10}],$$
$$= E[W_1 + W_3 + \cdots + W_9] + E[W_2 + W_4 + \ldots W_{10}],$$
$$= 5E[W_i | i \text{ is odd}] + 5E[W_i | i \text{ is even}],$$
$$= 5Pr\{W_i = 1 | i \text{ is odd}\} + 5Pr\{W_i = 1 | i \text{ is even}\},$$
$$\approx 8.16218.$$

Thus we've found $E[W]$. We now seek to calculate $Var[W]$. We do this by

$$Var[W] = \sum_{i=1}^{10} Var[W_i] + \sum_{i=1}^{10} \sum_{j=1, i \neq j}^{10} Cov(W_i, W_j).$$

This is obviously too complicated to calculate outright in any manner that would be nice to read, so we break it into parts. We start with $\sum_{i=1}^{10} Var[W_i]$:

$$Var[W_i] = E[W_i^2] - (E[W_i])^2,$$
$$= Pr\{W_i = 1\} - (Pr\{W_i = 1\})^2.$$

Note that this depends on i, so we do this in two parts:

$$Var[W_i | i \text{ is even}] = Pr\{W_i = 1 | i \text{ is even}\} - (Pr\{W_i = 1 | i \text{ is even}\})^2 \approx 0.14122,$$
$$Var[W_i | i \text{ is odd}] = Pr\{W_i = 1 | i \text{ is odd}\} - (Pr\{W_i = 1 | i \text{ is odd}\})^2 \approx 0.15842.$$

Thus our individual variance sum is

$$\sum_{i=1}^{10} Var[W_i] = 5Var[W_i | i\text{i is even}] + 5Var[W_i | i\text{i is odd}] \approx 1.49821.$$

We now begin our calculations of the covariance terms:

$$Cov(W_i, W_j) = E[W_i W_j] - E[W_i]E[W_j].$$

Note that this calculation will require us to consider three cases for combinations of i and j: the first case we'll consider is when *i,j* are both even, then when *i,j* are both odd, and then when one of *i,j* is odd and the other is even. We examine the case where *i,j* are even:

$$E[W_i W_j | i, j \text{ are even}] = Pr\{W_i W_j = 1 | i, j \text{ are even}\} = 1 - Pr\{W_i W_j = 0 | i, j \text{ are even}\},$$
$$= 1 - (2Pr\{W_i = 0 | i \text{ is even}\} - Pr\{W_i = 0 \cap W_j = 0 | i, j \text{ are even}\}),$$
$$= 1 - \left(2Pr\{W_i = 0 | i \text{ is even}\} - \left(\frac{13}{15}\right)^7 \left(\frac{11}{15}\right)^9\right),$$
$$\approx 0.68215.$$
$$E[W_i | i \text{ is even}] = Pr\{W_i = 1 | i \text{ is even}\}.$$

Thus the covariance in the case the i,j are both even is

$$Cov(W_i, W_j|i, j \text{ are even}) = E[W_iW_j|i, j \text{ are even}] - (E[W_i|i \text{ is even}])^2 \approx -0.00644.$$

Note that there are $\binom{5}{2} = 20$ such covariance terms.

We now consider the case where i,j are both odd:

$$E[W_iW_j|i, j \text{ are odd}] = Pr\{W_iW_j = 1|i, j \text{ are odd}\} = 1 - Pr\{W_iW_j = 0|i, j \text{ are odd}\},$$
$$= 1 - (2Pr\{W_i = 0|i \text{ is odd}\} - Pr\{W_i = 0 \cap W_j = 0|i, j \text{ are odd}\}),$$
$$= 1 - \left(2Pr\{W_i = 0|i \text{ is odd}\} - \left(\frac{11}{15}\right)^7 \left(\frac{13}{15}\right)^9\right),$$
$$\approx 0.63671.$$
$$E[W_i|i \text{ is odd}] = Pr\{W_i = 1|i \text{ is odd}\}.$$

Thus the covariance in the case the i,j are both odd is

$$Cov(W_i, W_j|i, j \text{ are odd}) = E[W_iW_j|i, j \text{ are odd}] - (E[W_i|i \text{ is odd}])^2 \approx -0.00750.$$

Note that there are $\binom{5}{2} = 20$ such covariance terms.

We finally consider the case where one of i,j is odd and the other is even. For convenience we say i is even and j is odd:

$$E[W_iW_j|i \text{ is even, } j \text{ is odd}] = Pr\{W_iW_j = 1|i, j \text{ are odd}\} = 1 - Pr\{W_iW_j = 0|i, j \text{ are odd}\}$$
$$= 1 - (Pr\{W_i = 0|i \text{ is even}\} + Pr\{W_j = 0|j \text{ is odd}\} - Pr\{W_i = 0 \cap W_j = 0|i \text{ is even, } j \text{ is odd}\}),$$
$$= 1 - \left(Pr\{W_i = 0|i \text{ is even}\} + Pr\{W_j = 0|j \text{ is odd}\} - \left(\frac{12}{15}\right)^7 \left(\frac{12}{15}\right)^9\right),$$
$$\approx 0.66058.$$
$$E[W_i|i \text{ is even}]E[W_j|j \text{ is odd}] = Pr\{W_i = 1|i \text{ is even}\}Pr\{W_j = 1|j \text{ is odd}\}.$$

Thus the covariance in the case that i is even and j is odd is

$$Cov(W_i, W_j|i \text{ is even, } j \text{ is odd}) = E[W_iW_j|i \text{ is even, } j \text{ is odd}] - E[W_i|i \text{ is even}]E[W_j|j \text{ is odd}] \approx 0.62699.$$

Note that there are $2\binom{5}{1}\binom{5}{1} = 50$ such covariance terms.

We can now compute $Var[W]$ by

$$Var[W] = \sum_{i=1}^{10} Var[W_i] + \sum_{i=1}^{10} \sum_{j=1, j\neq i}^{10} Cov(W_i, W_j),$$
$$= 5Var[W_i|i \text{ is even}] + 5Var[W_i|i \text{ is odd}] + 20Cov(W_i, W_j|i, j \text{ are even}) + 20Cov(W_i, W_j|i, j \text{ are even})$$
$$+50Cov(W_i, W_j|i \text{ is even, } j \text{ is odd}),$$
$$\approx 32.56891.$$

Thus we've found our variance.

(b) Let $\mathbf{X_i}$ be the number of people leaving the elevator at the ith floor, $\mathbf{i = 1, 2, \ldots, 10}$. Find $\mathbf{Var[X_1 + X_2 + X_3 + X_4]}$.

We first define $M = M_1 + M_2 + \cdots + M_7$ and $F = F_1 + F_2 + \cdots + F_9$, where

$$M_i = \begin{cases} 1, & \text{if male } i \text{ gets off on one of the first 4 floors;} \\ 0, & \text{otherwise.} \end{cases}$$

$$F_i = \begin{cases} 1, & \text{if female } i \text{ gets off on one of the first 4 floors;} \\ 0, & \text{otherwise.} \end{cases}$$

We can now say that $X_1 + X_2 + X_3 + X_4 = M + F$, and therefore that

$$Var[X_1 + X_2 + X_3 + X_4] = Var[M + F].$$

Note that since individuals leave the elevator independently we can say that

$$\begin{aligned} Var[M+F] &= Var[M] + Var[F], \\ &= \sum_{i=1}^{7} Var[M_i] + \sum_{i=1}^{9} Var[F_i], \\ &= 7Var[M_i] + 9Var[F_i]. \end{aligned}$$

We now compute the variance for males:

$$\begin{aligned} Var[M_i] &= E[M_i^2] - (E[M_i])^2, \\ &= Pr\{M_i = 1\} - (Pr\{M_i = 1\})^2 \end{aligned}$$

To compute $Pr\{M_i = 1\}$ it's easier to compute $1 - Pr\{M_i = 0\}$:

$$\begin{aligned} Pr\{M_i = 1\} &= 1 - Pr\{M_i = 0\}, \\ &= 1 - \left(\frac{13}{15}\right)\left(\frac{14-1}{15}\right)\left(\frac{13-2}{15}\right)\left(\frac{14-3}{15}\right), \\ &= 1 - \left(\frac{11}{15}\right)^2\left(\frac{13}{15}\right)^2, \\ &\approx 0.59607. \end{aligned}$$

Thus the variance for males is

$$\begin{aligned} Var[M_i] &= Pr\{M_i = 1\} - (Pr\{M_i = 1\})^2, \\ &\approx 0.24077. \end{aligned}$$

We now calculate the variance for females:

$$\begin{aligned} Var[F_i] &= E[F_i^2] - (E[F_i])^2, \\ &= Pr\{F_i = 1\} - (Pr\{F_i = 1\})^2 \end{aligned}$$

To compute $Pr\{F_i = 1\}$ it's easier to compute $1 - Pr\{F_i = 0\}$:

$$Pr\{F_i = 1\} = 1 - Pr\{F_i = 0\},$$

$$= 1 - \left(\frac{14}{15}\right)\left(\frac{13-1}{15}\right)\left(\frac{14-2}{15}\right)\left(\frac{13-3}{15}\right),$$

$$= 1 - \left(\frac{10}{15}\right)\left(\frac{12}{15}\right)^2\left(\frac{14}{15}\right),$$

$$\approx 0.60178.$$

Thus the variance for females is

$$Var[F_i] = Pr\{F_i = 1\} - (Pr\{F_i = 1\})^2,$$

$$\approx 0.23964.$$

Thus the variance overall is

$$Var[M + F] = 7Var[M_i] + 9Var[F_i],$$

$$\approx 3.84217.$$

This concludes my third test in Math 451.