ECS 256 - Problem set 2

Olga Prilepova, Christopher Patton, Alexander Rumbaugh, John Chen, Thomas Provan

Problem 1.a

A coin is flipped k times with p probability of heads. For each head, the coin is flipped one additional time (a bonus flip). The number of bonus flips is referred to as Y and the total number of heads X

Var(X) can be found using the Law of Total Variance, and properties of binomial distributions. We will also need to use part of the derivation of EX:

$$\begin{aligned}
E(X|Y) &= E(X - Y + Y|Y) \\
&= E((X - Y)|Y) + E(Y|Y) && \text{(by 3.13)} \\
&= pY + Y && \text{(by 3.110)} \\
&= (1 + p)Y
\end{aligned}$$

$$\begin{aligned}
Var(X) &= E[Var(X|Y)] + Var[E(X|Y)] && \text{(by 9.8)} \\
&= E[Var(X|Y)] + Var[(1 + p)Y] && \text{(from above)} \\
&= E[Var(X|Y)] + (1 + p)^2 kp(1 - p) && \text{(by 3.34 and 3.109)} \\
&= E[Yp(l - p)] + (1 + p)^2 kp(1 - p) && \text{(by 3.111)} \\
&= kp^2(1 - p) + (1 + p)^2 kp(1 - p) && \text{(by 3.103)} \\
&= kp(1 - p)\left(p + (1 + p)^2\right) \\
\text{Using p=0.5} & \\
&= k(0.25)(0.5 + (1.5^2)) \\
&= 0.6875k
\end{aligned}$$

Problem 1.b

In the trapped miner example, a miner chooses between three doors with only one leading to safety after 2 hours. The other two doors lead back to the door room after 3 and 5 hours respectively.

We are interesting the variance of Y, the time it takes to escape the mine. We will build upon Ahmed Ahmedin's solution to EY, where N refers to the total attempts needed to escape and U_i refers to the time spent traveling on the i^{th} attempt.

$$
\begin{aligned}
Var(Y) &= E[Var(Y|N)] + Var[E(Y|N)] && \text{(by 9.8)} \\
&= E[Var(Y|N)] + Var[4N-2] && \text{(by 9.16)} \\
&= E[Var(Y|N)] + 16Var[N] && \text{(by 3.34 and 3.41)} \\
&= E[Var(Y|N)] + 16 \cdot \frac{1-1/3}{(1/3)^2} && \text{(by 3.93)} \\
&= E[Var(U_1 + U_2 + ... + U_n|N)] + 96 \\
&= E[Var(U_1|N) + ... + Var(U_{N-1}|N + Var(U_N|N)] + 96 && \text{(by 3.51)} \\
&= E[1 + 1 + ...1 + 0] + 96 \\
&= E[N-1] + 96 \\
&= E[N] - 1 + 96 && \text{(by 3.17)} \\
&= 3 - 1 + 96 && \text{(by 3.92)} \\
&= 98
\end{aligned}
$$

We know that Var(U_i—N) is independent because the miner's choice of door does not depend of a previous choice. Since we are conditioning this event on there being N attempts, the values of the first N-1 attempts will either be 3 or 5. So the variance of an individual attempt in this case is 1. The variance of the N^{th} attempt is 0 because that attempt always is the same tunnel.

Problem 2.a
For a vector Q of random variables $(Q_1, ..Q_n)$ we have:

$$Cov(Q) = E(QQ') - E(Q)E(Q') \quad \text{(by 13.53)}$$

Let $Q = Y|X$, where Y is vector valued. Then:

$$Cov(Y|X) = E\big((Y|X)(Y|X)'\big) - E(Y|X)E(Y|X)' \quad \text{(by 13.53)}$$

Taking expected value of both sides we have:

$$\begin{aligned}
E\big(Cov(Y|X)\big) &= E\Big(E\big((Y|X)(Y|X)'\big) - E(Y|X)E(Y|X)'\Big) \\
&= E\Big(E\big((Y|X)(Y|X)'\big)\Big) - E\Big(E(Y|X)E(Y|X)'\Big) \\
&= E(YY') - E\Big(E(Y|X)E(Y|X)'\Big) && \text{(by Law of Tot. Expect.)}
\end{aligned}$$

Now let $Q = E(Y|X)$, where Y is vector valued. Then:

$$\begin{aligned}
Cov(E(Y|X)) &= E\big(E(Y|X)E(Y|X)'\big) - E\big(E(Y|X)\big)E\big(E(Y|X)\big)' && \text{(by 13.53)} \\
&= E\big(E(Y|X)E(Y|X)'\big) - E(Y)E(Y)' && \text{(by Law of Tot. Expect.)}
\end{aligned}$$

Summing up the left sides and the right sides of these 2 equations we get:

$$\begin{aligned}
E\big(Cov(Y|X)\big) + Cov(E(Y|X)) &= E(YY') - E\Big(E(Y|X)E(Y|X)'\Big) \\
&\quad + E\big(E(Y|X)E(Y|X)'\big) - E(Y)E(Y)' \\
E\big(Cov(Y|X)\big) + Cov(E(Y|X)) &= E(YY') - E(Y)E(Y)' \\
E\big(Cov(Y|X)\big) + Cov(E(Y|X)) &= Cov(Y) && \text{(by 13.53)}
\end{aligned}$$

Problem 2.b

TODO.

First, just an equation to remind us of what we're actually trying to find here, the correlation between X and Y.

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

But X is too complicated and annoying to work with, so we're going to change the random vector to find the correlation matrix of from $(X,Y)'$ to $(B,Y)'$, where B is the random variable that expresses the number of bonus heads flipped, via eq. 13.54. Then we'll expand that forumla out so we know what the final equation we need to solve for is.

$$\begin{aligned}
\begin{pmatrix} X \\ Y \end{pmatrix} &= \begin{pmatrix} B+Y \\ Y \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} B \\ Y \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
Cov\big[(X,Y)'\big] &= \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Cov\big[(B,Y)'\big] \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} && \text{(By 13.54)} \\
&= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Var(B) & Cov(B,Y) \\ Cov(B,Y) & Var(Y) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} Var(B)+Cov(B,Y) & Cov(B,Y)+Var(Y) \\ Cov(B,Y) & Var(Y) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} Var(B)+Var(Y)+2Cov(B,Y) & Cov(B,Y)+Var(Y) \\ Cov(B,Y)+Var(Y) & Var(Y) \end{pmatrix}
\end{aligned}$$

Now that we know the final form, we'll go back and calculate $Cov[(B,Y)']$. From the problem **2.a**, we have the following.

$$Cov(W) = E\big(Cov(W|Z)\big) + Cov\big(E(W|Z)\big)$$

Now we'll allow $W = (B,Y)'$ and $Z = Y$ in order to calculate for $Cov\big[(B,Y)'\big]$.

$$Cov\big((B,Y)'\big) = E\big[Cov((B,Y)'|Y)\big] + Cov\big[E((B,Y)'|Y)\big]$$

Now we'll consider each half of that individually, then add them together. First, let's consider the random variable $Q_c = Cov\big((B,Y)'|Y\big)$

$$Cov\big((B,Y)'|Y\big) = \begin{cases} \text{for } i \text{ from } 0..k \\ Cov((B,Y)'|Y=i)\ w.p. \binom{k}{i} 0.5^k \end{cases}$$

$$Cov\big((B,Y)'|Y=i\big) = \begin{pmatrix} Var(B|Y=i) & Cov(B,Y|Y=i) \\ Cov(B,Y|Y=i) & Var(Y|Y=i) \end{pmatrix}$$
$$= \begin{pmatrix} i/4 & Cov(B,Y|Y=i) \\ Cov(B,Y|Y=i) & 0 \end{pmatrix}$$

$$\begin{aligned} Cov\big(B,Y|Y=i\big) &= E(BY|Y=i) - E(B|Y=i) * E(Y|Y=i) \\ &= E(iB|Y=i) - i * E(B|Y=1) \\ &= i * E(B|Y=i) - i * E(B|Y=1) \\ &= 0 \end{aligned}$$

$$Cov\big((B,Y)'|Y=i\big) = \begin{pmatrix} i/4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$E\Big[Cov\big((B,Y)'|Y=i\big)\Big] = \sum_{i=0}^{k} \begin{pmatrix} i/4 & 0 \\ 0 & 0 \end{pmatrix} * \binom{k}{i} 0.5^k$$

((Should be able to reduce that, don't remember how. Do double check my work on those covariances. I am reasonably certain they all zero out, which would be greatand makes intuitive sense to me (the covariance of anything with a scalar, which (Y—y=i) is, is 0) but I'm not 100% confident about it))

Now then, onto the other half, $Cov\big[E((B,Y)'|Y)\big]$ First, that expected value part.

$$E\big((B,Y)'|Y\big) = \begin{cases} \text{for } i \text{ from } 0..k \\ E((B,Y)'|Y=i)\ w.p. \binom{k}{i} 0.5^k \end{cases}$$

$$E\big((B,Y)'|Y=i\big) = \begin{pmatrix} E(B|Y=i) \\ E(Y|Y=i) \end{pmatrix}$$
$$= \begin{pmatrix} i/2 \\ i \end{pmatrix}$$

((Okay, here's the point at which I get stuck. We now need to get the covariance matrix of this. Hmm... Covariance is the hardest one by far. I think the two variances are pretty easy, fairly sure they fall out from the variance of the binomial. I'll write up the matrix form of what we need, hopefully Olga has made progress by the time I wake up.))

$$Cov\big[E((B,Y)'|Y)\big] = \left(\begin{array}{cc} Var(E(B|Y)) & Cov\big(E(B|Y), E(Y|Y)\big) \\ Cov\big(E(B|Y), E(Y|Y)\big) & Var(E(Y|Y)) \end{array} \right)$$

Problem 3 Let X_i denote the state of the machine at time index i, where $X_0 \sim \Pi$, the stationary distributions of the states. Finally, suppose there are n states, named numerically 1 to n. We want to compute

$$\rho(X_{i+j}, X_i) = \frac{\operatorname{Cov}(X_{i+j}, X_i)}{\sqrt{\operatorname{Var}(X_{i+j})} \cdot \sqrt{\operatorname{Var}(X_i)}}$$

for all $1 \leq j \leq k$ and some k. We can simplify the denomonator by realizing that, since X_i and X_{i+j} are drawn from the same distribution, they have the same variance. The equation becomes:

$$\rho(X_{i+j}, X_i) = \frac{\operatorname{Cov}(X_{i+j}, X_i)}{\operatorname{Var}(X_i)}.$$

By definition, the expected value of X_i is

$$E(X_i) = \sum_{l=1}^{n} l \cdot \pi_l.$$

Thus,

$$\operatorname{Var}(X_i) = E(X_i^2) - (EX_i)^2 = \sum_{l=1}^{n} (l^2 \pi_l) - \Big[\sum_{l=1}^{n} l \pi_l \Big]^2.$$

Now, let's derive $\operatorname{Cov}(X_{i+j}, X_i)$. For this, we also need the expected value of X_{i+j}:

$$E(X_{i+j} | X_i) = \sum_{k=1}^{n} k \cdot m_{l,k}^{j}$$

where M^j is the transition matrix $M \cdot M \cdots M$ (j times), and $X_i = l$.

$$\operatorname{Cov}(X_{i+j}, X_i) = E(X_{i+j} X_i) - E(X_{i+j}) E(X_i)$$

Suppose $Q = X_{i,j} X_i$. By the law of total expectations, $E(Q) = E(E(Q|X_i))$. Thus,

$$\begin{aligned} E(X_{i+j} X_i) &= \sum_{l=1}^{n} \pi_l l \cdot E(X_{i+j} | X_i) \\ &= \sum_{l=1}^{n} \Bigg[\pi_l l \cdot \Big[\sum_{k=1}^{n} k \cdot m_{l,k}^{j} \Big] \Bigg]. \end{aligned}$$

Since X_{i+j} and X_i have the same distribution (Π), the covariance becomes

$$\mathrm{Cov}(X_{i+j}, X_i) = \sum_{l=1}^{n} \left[\pi_l l \cdot \left[\sum_{k=1}^{n} k \cdot m_{l,k}^{j} \right] \right] - (EX_i)^2.$$

We can now write code to calculate the correlation of X_i and X_{i+j}. The following function `mccor(tm, k)` returns a vector corresponding to $\rho(X_i, X_{i+1}) \ldots \rho(X_i, X_k)$.

```
# Calculate Pi distribution from a transition matrix.
findpis <- function(p) {
  n <- nrow(p)
  imp <- diag(n) - t(p)
  imp[n,] <- rep(1,n)
  rhs <- c(rep(0,n-1),1)
  solve(imp, rhs)
}

# Calculate correlations between the current state and
# the next K states, given transition matrix tm.
mccor <- function(tm, K) {

  Pi <- findpis(tm)
  n <- nrow(p)

  # mu = E[Xi] for all i.
  mu <- 0
  for (i in 1 : n) {
    mu <- mu + (i * Pi[i])
  }

  # denom = Var(Xi).
  denom <- 0
  for (i in 1 : n) {
    denom <- denom + (i^2 * Pi[i])
  }
  denom <- denom - mu^2

  # corr = Correlations.
  corr <- c()

  i <- 1
```

```
  for(j in 1 : K) {

    outer_sum <- 0
    for (l in 1 : n) {
      inner_sum <- 0
      for (k in 1 : n) {
        inner_sum <- inner_sum + (k * tm[l,k])
      }
      outer_sum <- outer_sum + (l * Pi[l] * inner_sum)
    }

    # numer = Cov(Xi, Xi+j).
    numer <- outer_sum - mu^2

    corr <- c(corr, c(numer / denom))
    tm <- tm %*% tm # Next M^j.
  }

  # Result is c(rho(Xi, Xi+1), rho(Xi, Xi+2) ... rho(Xi, Xi+k)).
  return (corr)
}
```