

## 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  $X$

$\text{Var}(X)$  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  $E(X)$ :

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

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

Using  $p=0.5$

$$\begin{aligned} &= k(0.25)(0.5 + (1.5^2)) \\ &= 0.6875k \end{aligned}$$

### 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 interested in 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.

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

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:

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

Taking expected value of both sides we have:

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

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

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

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

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

**Problem 2.b**

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

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

From the problem **2.a**, we have the following.

Let  $B = X - Y$  and use that random variable in the following computations.

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

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

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

$$Var(B + Y) = Var(B) + Var(Y) + 2Cov(B, Y)$$

$$Cov(B, Y) = (Var(B + Y) - Var(B) - Var(Y))/2$$

$$Cov(B, Y|Y) = (Var(B + Y|Y) - Var(B|Y) - Var(Y|Y))/2$$

$$Var(B + Y|Y) = Var(B|Y)$$

$$Var(Y|Y) = 0$$

$$Cov(B, Y|Y) = 0$$

$$\begin{aligned} Cov(E(B, Y|Y)) &= (Var(EB + EY|Y) - Var(EB|Y) - Var(EY|Y))/2 \\ &= (Var(1.5Y) - Var(0.5Y) - Var(Y))/2 \\ &= (2.25Var(Y) - 0.25Var(Y) - Var(Y))/2 \\ &= Var(Y) \end{aligned}$$

Hence:

$$\begin{aligned} Cov(B, Y) &= 0 + Var(Y) \\ &= Var(Y) \\ &= kp(1 - p) \\ &= 0.25k \end{aligned}$$

Now, let's see how we can get  $Cov(X, Y)$  using that fact that we now know  $Cov(B, Y)$ . Remember,  $X = B + Y$

$$\begin{aligned} Cov(X, Y) &= Cov(B + Y, Y) \\ Cov(B + Y, Y) &= Cov(B, Y) + Cov(Y, Y) \quad (13.2) \\ &= Cov(B, Y) + Var(Y) \end{aligned}$$

Now we have all the ingredients to find the correlation:

$$\begin{aligned}
 \rho(X, Y) &= \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \\
 &= \frac{Cov(B, Y) + Var(Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \\
 &= \frac{0.25k + 0.25k}{\sqrt{0.6875k}\sqrt{0.25k}} \\
 &= \frac{0.5k}{k\sqrt{0.6875}\sqrt{0.25}} \\
 &= \frac{0.5}{\sqrt{0.6875 * 0.25}} \\
 &= 1.2
 \end{aligned}$$

**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

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

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

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

By definition, the expected value of  $X_i$  is

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

Thus,

$$\text{Var}(X_i) = E(X_i^2) - (E X_i)^2 = \sum_{l=1}^n (l^2 \pi_l) - \left[ \sum_{l=1}^n l \pi_l \right]^2.$$

Now, let's derive  $\text{Cov}(X_{i+j}, X_i)$ . 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  $X_i = l$ .

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

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

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

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

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

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  $\rho(X_i, X_{i+1}) \dots \rho(X_i, X_k)$ .

```

1 # Calculate Pi distribution from a transition matrix.
2 findpis <- function(p) {
3   n <- nrow(p)
4   imp <- diag(n) - t(p)
5   imp[n,] <- rep(1, n)
6   rhs <- c(rep(0, n-1), 1)
7   solve(imp, rhs)
8 }
9
10 # Calculate correlations between the current state and
11 # the next K states, given transition matrix tm.
12 mccor <- function(tm, K) {
13
14   Pi <- findpis(tm)
15   n <- nrow(p)
16
17   # mu = E[Xi] for all i.
18   mu <- 0
19   for (i in 1 : n) {
20     mu <- mu + (i * Pi[i])
21   }
22
23   # denom = Var(Xi).
24   denom <- 0
25   for (i in 1 : n) {
26     denom <- denom + (i^2 * Pi[i])
27   }
28   denom <- denom - mu^2
29
30   # corr = Correlations.
31   corr <- c()
32
33   # Mj = transition matrix at time i+j.

```

```

34  Mj <- tm
35
36  i <- 1
37  for(j in 1 : K) {
38
39      outer_sum <- 0
40      for (l in 1 : n) {
41          inner_sum <- 0
42          for (k in 1 : n) {
43              inner_sum <- inner_sum + (k * Mj[l,k])
44          }
45          outer_sum <- outer_sum + (l * Pi[l] * inner_sum)
46      }
47
48      # numer = Cov(Xi, Xi+j).
49      numer <- outer_sum - mu^2
50
51      corr <- c(corr, c(numer / denom))
52      Mj <- Mj %*% tm # Next Mj.
53  }
54
55  # Result is c(rho(Xi, Xi+1), rho(Xi, Xi+2) ... rho(Xi, Xi+k)).
56  return (corr)
57 }

```