Homework 6
36-325/725
due Friday Oct 12

(1) Suppose we generate a random variable X in the following way. First we flip a fair coin. If the coin is heads, take X to have a Unif(0,1) distribution. If the coin is tails, take X to have a Unif(3,4) distribution.

(a) Find the mean of X.

(b) Find the standard deviation of X.

(2) Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be random variables and let $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ be constants. Show that

$$Cov\left(\sum_{i=1}^{m} a_i X_i,\ \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j).$$

(3) Let

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{3}(x+y) & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\mathrm{Var}(2X - 3Y + 8)$.

(4) Let r(x) be a function of x and let s(y) be a function of y. Show that

$$E(r(X)s(Y)|X) = r(X)E(s(Y)|X).$$

Also, show that $E(r(X)|X) = r(X)$.

(5) Prove that

$$\mathrm{Var}(Y) = \mathrm{E}\,\mathrm{Var}(Y \mid X) + \mathrm{Var}\,\mathrm{E}(Y \mid X).$$

Hint: Let $m = E(Y)$ and let $b(x) = E(Y|X = x)$. Note that $E(b(X)) = EE(Y|X) = E(Y) = m$. Bear in mind that b is a function of x. Now write

$Var(Y) = E(Y-m)^2 = E((Y-b(X)) + (b(X)-m))^2$. Expand the square and take the expectation. You then have to take the expectation of three terms. In each case, use the rule of the iterated expectation: i.e. $E(\text{stuff}) = E(E(\text{stuff}|X))$. You may want to use the result of problem (4) to evaluate one of the terms.

(6) Show that if $E(X|Y=y) = c$ for some constant c then X and Y are uncorrelated.

(7) This question is to help you understand the idea of a *sampling distribution.* Let $X_1, \ldots, X_n$ be iid with mean μ and variance σ^2. Let $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$. Then $\overline{X}_n$ is a *statistic*, that is, a function of the data. Since $\overline{X}_n$ is a random variable, it has a distribution. This distribution is called the *sampling distribution of the statistic.*

(7a) Find the mean and variance of the sampling distribution, i.e. find $E(\overline{X}_n)$ and $\text{Var}(\overline{X}_n)$.

(7b) Don't confuse the distribution of the data f_X and the distribution of the statistic $f_{\overline{X}_n}$. To make this clear, here is an example. Let $X_1, \ldots, X_n \sim \text{Uniform}(0,1)$. Let f_X be the density of the $\text{Uniform}(0,1)$. Plot f_X. Now let $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$. Find $E(\overline{X}_n)$ and $\text{Var}(\overline{X}_n)$. Plot them as a function of n. Comment. Now we will simulate the sampling distribution.

```
nsim <- 10000   ### think of nsim as essentially being infinity
n <- 25
xbar <- rep(0,nsim)
for(i in 1:nsim){
     x       <- runif(n,0,1)
     xbar[i] <- mean(x)
     }
hist(xbar)
expected.value <- mean(xbar)
variance       <- var(xbar)
print(expected.value)
```

```
print(variance)
```

Repeat this experiment for $n = 1, 2, 5, 10, 25, 50, 100$. Check that the simulated values of $E(\overline{X}_n)$ and $\mathrm{Var}(\overline{X}_n)$ agree with your theoretical calculations. What do you notice about the sampling distribution of $\overline{X}_n$ as n increases?