Homework 8, DATA 556: Due Tuesday, 11/20/2018

Alexander Van Roijen

November 20, 2018

Please complete the following:

1. Problem 1 Let Z,W be bivariate Normal defined as

$$Z = X \tag{1}$$

$$W = \rho X + \sqrt{1 - \rho^2} Y \tag{2}$$

with X,Y i.i.d N(0,1) and $-1 < \rho < 1$. Find $E[W|Z]$ & $\text{Var}(W|Z)$

$$E[W|Z] = E[\rho X + \sqrt{1 - \rho^2} Y | Z = z] = E[\rho X + \sqrt{1 - \rho^2} Y | X = z] \tag{3}$$

$$= \rho z + \sqrt{1 - \rho^2} E[Y] \text{ by linearity of expectation} \tag{4}$$

$$= Z\rho \tag{5}$$

$$Var(W|Z) = E[W^2|Z] - E[W|Z]^2 = E[(\rho X + \sqrt{1 - \rho^2} Y)^2 | Z = z] - z^2 \rho^2 \tag{6}$$

$$= E[\rho^2 z^2 + (1 - \rho^2) Y^2 + 2\rho z \sqrt{1 - \rho^2} Y] - z^2 \rho^2 \tag{7}$$

$$= \rho^2 z^2 + (1 - \rho^2) E[Y^2] + 2\rho z \sqrt{1 - \rho^2} E[Y] - z^2 \rho^2 \text{ by linearity of expectation} \tag{8}$$

$$\text{We also know } E[Y^2] = Var(Y) + E[Y]^2 = 1 \tag{9}$$

$$=> Var(W|Z) = \rho^2 z^2 + 1 - \rho^2 - z^2 \rho^2 = 1 - \rho^2 \tag{10}$$

2. Let $X = (X_1, X_2, X_3, X_4, X_5) \sim Mult_5(n, p)$ with $p = (p_1, p_2, p_3, p_4, p_5)$

(a) Find $E[X_1|X_2]$ and $\text{Var}(X_1|X_2)$

$$X_1, X_3, X_4, X_5 | X_2 \sim Mult_4(n - X_2, p') \text{ with} \tag{11}$$

$$p' = (p_1', p_3', p_4', p_5') \text{ and } p_i' = \frac{p_i}{1 - p_2} \tag{12}$$

Now we recognize that $P(X_i = x_i)$ in either distribution is $\binom{N'}{x_i} p_1^{x_i} (1 - p_i)^{N' - x_i}$

$$\tag{13}$$

$$\text{Where we know that } N' - x_1 = \text{ remaining possible successes} \tag{14}$$

$$\text{Thus, we see a binomial distribution for } X_i \tag{15}$$

$$=> E[X_1|X_2] = N' * p_i' = (n - X_2)p_1' \tag{16}$$

$$=> Var[X_1|X_2] = N'p_1'(1 - p_1') = (n - X_2)p_1'(1 - p_1') \tag{17}$$

This further makes sense as we expect both values to be functions of our RV X_2

$$\tag{18}$$

(b) Find $E[X_1|X_2 + X_3]$

Similar to above, we can determine a new multinomial

$$X_1, X_4, X_5 | X_2 + X_3 \sim Mult_3(n - (X_2 + X_3), p') \text{ with} \tag{19}$$

$$p' = (p'_1, p'_4, p'_5) \text{ and } p'_i = \frac{p_i}{1 - (p_2 + p_3)} \tag{20}$$

$$\text{By the same logic as above, we determine that} X_1 \sim Binom(N', p'_1) \tag{21}$$

$$=> E[X_1|X_2 + X_3] = N' * p'_1 = (n - (X_2 + X_3))p'_1 \tag{22}$$

3. Show that the following version of LOTP follows from Adam's law: for any event A and continuous random variable X with PDF f_X:

$$P(A) = \int_{-\infty}^{\infty} P(A|X = x) f_X(x) dx \tag{23}$$

$$\text{Let's use indicator variables to prove this} \tag{24}$$

$$I_A = \text{Inidicator for when event A occurs} \tag{25}$$

$$E[I_A] = P(A) \text{ and } E[I_A|X = x] = P(A|X = x) \text{By the fundamental bridge} \tag{26}$$

$$E[I_A] = E[E[I_A|X]] \text{ Adam's Law} \tag{27}$$

$$\text{all together we get} \tag{28}$$

$$P(A) = E[E[I_A|X]] = E[P(A|X = x)] = \int_{-\infty}^{\infty} P(A|X = x) f_X(x) dx \text{ by lotus} \square \tag{29}$$

4. Let $N \sim Pois(\lambda_1)$ be the number of movies that will be released next year. Suppose that for each movie the number of tickets sold is $Pois(\lambda_2)$, independently

(a) Find the mean and the variance of the number of movie tickets that will be sold next year.

We can describe the number of tickets sold overall as the conditional probability on tickets sold per movie given the number of movies released that year. We get the

3

following

$$\text{tickets sold} = E[M_1 + M_2 + ... + M_n | N = n] \text{ with} M_i \sim Pois(\lambda_2) \tag{30}$$

$$= E[M_1 | N = n]... + E[M_n | N = n] \text{ by linearity of expectation} \tag{31}$$

$$= n * \lambda_2 => S = N * \lambda_2 \text{ where S is a random variable expressing tickets sold next year} \tag{32}$$

$$=> E[S] = E[N] * \lambda_2 = \lambda_1 * \lambda_2 \tag{33}$$

$$\text{Var(Tickets Sold)} = E[\text{Var}(M_1 + ...M_n | N = n)] + \text{Var}(E[M_1 + ...M_n | N = n]) \tag{34}$$

$$= E[\lambda_2 + \lambda_2... + \lambda_2] + \text{Var}(\lambda_2 N) = E[N\lambda_2] + \text{Var}(N\lambda_2) = \lambda_1\lambda_2 + \lambda_2^2\lambda_1 \tag{35}$$

(b) Use simulations in R (the statistical programming language) to numerically estimate mean and the variance of the number of movie tickets that will be sold next year assuming that the mean number of movies released each year in the US is 700, and that, on average, 800000 tickets were sold for each movie.

```
> set.seed(123)
> numReleased=700
> numSold=800000
> numTrials=100000
> numMovies = rpois(numTrials,numReleased)
> counter = 0
> oneYearTix=0
> total=numeric()
> while(counter<=numTrials)
+ {
+    oneYearTix = rpois(numMovies[counter],numSold)
+    total[counter] = sum(oneYearTix)
+    counter = counter+1
+ }
> print(mean(total))
[1] 560023366
> print(numSold*numReleased)
```

```
[1] 5.6e+08

> print(var(total))

[1] 4.47973e+14

> print(numSold^2 * numReleased)

[1] 4.48e+14
```

Look at that. How wonderful! It matches our expectations!

5. Show that if $E(Y|X) = c$ is a constant, then X and Y are uncorrelated.

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \qquad (36)$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] \qquad (37)$$

$$\text{By Adam's law we get } E[E[Y|X]] = E[Y] => E[Y] = E[c] = c \qquad (38)$$

$$E[XY] = \sum_{x \in X} E[XY|X = x]P(X = x) = \sum_{x \in X} xE[Y]P(X = x) = c\sum_{x \in X} xP(X = x) = cE[X]$$
$$(39)$$

$$=> Cov(X, Y) = E[XY] - E[X]E[Y] = cE[X] - cE[X] = 0 \qquad (40)$$

$$\text{Thus X and Y are uncorrelated } \square \qquad (41)$$

6. Show that for any random variables X and Y,

$$E[Y|E[Y|X]] = E[Y|X] \qquad (42)$$

$$\text{We know } E[Y|X] \text{ is a function of X, call it } h(X) \qquad (43)$$

$$=> E[Y|E[Y|X]] = E[Y|h(X)] \text{ Inuitively, it makes sense that this equals } E[Y|X] \qquad (44)$$

$$\text{As knowing a function of X details us the same information as knowing X} \qquad (45)$$

$$\text{However, we will be more rigorous using Adams law with extra conditioning} \qquad (46)$$

$$E[Y|h(X)] = E[E[Y|h(X), M]|h(X)] \text{ Now we let } M = X \qquad (47)$$

$$= E[E[Y|h(X), X]|h(X)] = E[E[Y|X]|h(X)] \text{ as the event g(X) is known with X} \qquad (48)$$

$$= E[h(X)|h(X)] \text{ by substituting from our earlier definition} = h(X) = E[Y|X]\square \qquad (49)$$

7. Let Y denote the number of heads in n flips of a coin, whose probability of heads is θ.

(a) Suppose θ follows a distribution $P(\theta) = Beta(a, b)$, and then you observe y heads out of n flips. Show algebraically that the mean $E(\theta|Y = y)$ always lies between the mean $E(\theta)$ and the observed relative frequency of heads

$$min\left\{E[\theta], \frac{y}{n}\right\} \leq E(\theta|Y = y) \leq max\left\{E[\theta], \frac{y}{n}\right\} \tag{50}$$

Here $E(\theta|Y = y)$ is the mean of the distribution $P(\theta|Y = y)$, and $E(\theta)$ is the mean of the distribution $P(\theta) = Beta(a, b)$.

$$E[\theta|Y = y] = \int_{-\infty}^{\infty} \theta f(\theta|Y = y) d\theta \tag{51}$$

$$f(\theta|Y = y) = \frac{P(Y|\theta) * P(\theta)}{P(Y)} \text{By bayes rule} \tag{52}$$

We recognize that the numerator is a new beta distribution with new $\alpha\beta$ $\quad$ (53)

Resultingly, depending on which one is large, y and n or a and b, the exepected alue will lie between

$$\tag{54}$$

(b) Show that if $P(\theta) = Unif(0, 1)$

We have $\text{Var}(\theta|Y = y) \leq \text{Var}(\theta)$

Here $\text{Var}(\theta|Y = y)$ is the variance of the distribution of $P(\theta|Y = y)$ and $\text{Var}(\theta)$ is the

variance of the distribution $P(\theta) = Unif(0,1)$

$$\text{Var}(\theta) = \frac{1}{12} \text{ Known about the distribution} \tag{55}$$

$$\text{Similarly to above, we use bayes rule to find} f(\theta|Y=y) = \frac{P(Y|\theta) * P(\theta)}{P(Y)} \tag{56}$$

$$=> f(\theta)|Y = y \propto \binom{n}{y} \theta^y (1-\theta)^{n-y} * 1 \sim Beta(Y, n-Y) \tag{57}$$

$$\text{Var}(Beta(Y, n-Y)) = \frac{Y(n-Y)}{(Y+n-Y)^2(Y+n-Y+1)} = \frac{Y(n-Y)}{(n)^2(n+1)} \tag{58}$$

$$\text{Now lets analyze the derivate of this variance with respect to Y} \tag{59}$$

$$\text{This makes sense as our variance should be some function of Y} \tag{60}$$

$$\frac{d}{dy}\frac{Y(n-Y)}{(n)^2(n+1)} \propto n - 2y \tag{61}$$

$$\text{This function is negative with Y and thus decreasing for all y.} \tag{62}$$

$$\text{We can easily see that there is a point of maximum return at } y = \frac{n}{2} \tag{63}$$

$$\text{Plug this into our original equation we get} \tag{64}$$

$$\text{Var}(Beta(Y, n-Y)) = \frac{Y(n-Y)}{(Y+n-Y)^2(Y+n-Y+1)} = \frac{1}{4(n+1)} \tag{65}$$

Since n is strictly positive and $y = \frac{n}{2}$, this value is highest with lowest n possible, $n = 2$

$$\tag{66}$$

$$=> max(\text{Var}(Beta(Y, n-Y))) = \frac{1}{4 * (2+1)} = \frac{1}{12} \leq \text{Var}(\theta)\square \tag{67}$$

8. Let A, B and C be independent random variables with the following distributions:

$$P(A = 1) = 0.4; P(A = 2) = 0.6 \tag{68}$$

$$P(B = -3) = 0.25; P(B = -2) = 0.25; P(B = -1) = 0.25; P(B = 1) = 0.25 \tag{69}$$

$$P(C = 1) = 0.5; P(C = 2) = 0.4; P(C = 3) = 0.1 \tag{70}$$

(a) What is the probability of the quadratic equation $Ax^2 + Bx + C = 0$ has two real

roots that are different?

$$\text{we are solving} \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \tag{71}$$

For our roots to be real and different, we require $B^2 - 4AC > 0 \ \& B^2 - 4AC \neq 0$

$$\tag{72}$$

With some thinking, we can see that this is only satisfied in the following cases

$$\tag{73}$$

$$A = 1, C = 1, B = -3 \tag{74}$$

$$A = 1, C = 2, B = -3 \tag{75}$$

$$A = 2, C = 1, B = -3 \tag{76}$$

$$\tag{77}$$

If A or C were both 2 or either was great than two, we would only have negative roots as B^2 can only take the maximum value of 9 at any time.

Thus we simply add the probabilities of these three joint events

$$P(A = 1, C = 1, B = -3) + P(A = 2, C = 1, B = -3) + P(A = 1, C = 2, B = -3)$$

$$\tag{78}$$

$$= .4 * .5 * .25 + .6 * .5 * .25 + .4 * .4 * .25 = 0.165 \tag{79}$$

(b) What is the probability that the quadratic equation $Ax^2 + Bx + C = 0$ has two real

roots and are both strictly positive.

$$\text{we are solving} \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \tag{80}$$

For our roots to be real positive, we require $B^2 - 4AC \geq 0 \ \& -B - \sqrt{B^2 - 4AC} > 0$

$$\tag{81}$$

With some thinking, we can see that this is only satisfied in the following cases

$$\tag{82}$$

$$A = 1, C = 1, B = -2 \tag{83}$$

$$A = 1, C = 1, B = -3 \tag{84}$$

$$A = 1, C = 2, B = -3 \tag{85}$$

$$A = 2, C = 1, B = -3 \tag{86}$$

$$\tag{87}$$

We can see that unlike the last problem, we can allow our roots to be one in the same, so we add another case where $-B - \sqrt{B^2 - 4AC} = 0$

we get the following

$$P(A = 1, C = 1, B = -3) + P(A = 2, C = 1, B = -3) + P(A = 1, C = 2, B = -3) + P(A = 1, C = 1,$$

$$\tag{88}$$

$$= .4 * .5 * .25 + .6 * .5 * .25 + .4 * .4 * .25 + .4 * .5 * .25 = 0.215 \tag{89}$$

Have a great thanksgiving!