STA360: Homework 1

Samuel Eure

9/3/2018

Problem 2.1

The following data is provided on fathers and sons with each row corresponding to the father's occupation and each column corresponding to the son's occupation.

```
X = c(0.018, 0.035, 0.031, 0.008, 0.018,
      0.002, 0.112, 0.064, 0.032, 0.069,
      0.001, 0.066, 0.094, 0.032, 0.084,
      0.001, 0.018, 0.019, 0.010, 0.051,
      0.001, 0.029, 0.032, 0.043, 0.130)
data = t(array(X, dim = c(5,5)))
data = data.frame(data)
attach(data)
jobs = c('farm', 'operatives', 'craftsmen', 'sales', 'professional')
names(data) = jobs
row.names(data) = c('farm', 'operatives', 'craftsmen', 'sales', 'professional')
print(data)
```

```
##               farm operatives craftsmen sales professional
## farm         0.018      0.035     0.031 0.008        0.018
## operatives   0.002      0.112     0.064 0.032        0.069
## craftsmen    0.001      0.066     0.094 0.032        0.084
## sales        0.001      0.018     0.019 0.010        0.051
## professional 0.001      0.029     0.032 0.043        0.130
```

Let Y_1 represent the random variable which is the father's occupation and Y_2 represent random variable which is the son's occupation.

(2.1a): Find the marginal probaility distribution of a father's occupation.

Let $P(Y_1, Y_2)$ represent the joint probability distribution for the occupation of a father and son, $P(Y_i)$ represent the marginal probability distribution of either the father(if $i = 1$) or son (if $i = 2$), and Φ represent the space of all occupation for the son or the father. i.e.

$$\Phi = \{farm, operatives, craftsmen, sales, professional\}$$

Using this notation, the marginal probability distribution of the father's occupation can be found as follows:

$$\forall i \in \Phi, P(Y_1 = i) = \sum_{k \in \Phi} P(Y_1 = i, Y_2 = k)$$

Thus, using the table presented above, the marginal probability distribution for the occupation of the father can be found by summing along the rows of the table, resulting in:

```
for(a in jobs){
  prob = sum(data[a,])
  start = paste("P( Y1 =", a,") = ", prob)
```

```
  print(start)
}
```

```
## [1] "P( Y1 = farm ) =  0.11"
## [1] "P( Y1 = operatives ) =  0.279"
## [1] "P( Y1 = craftsmen ) =  0.277"
## [1] "P( Y1 = sales ) =  0.099"
## [1] "P( Y1 = professional ) =  0.235"
```

which is indeed a probability distribution since $0.11 + 0.270 + 0.277 + 0.099 + 0.235 = 1.00$, as required.

(2.1b): Find the marginal probaility distribution of a son's occupation

This is achived in almost an identical way to problem 2.1a, however now:

$$\forall k \in \Phi, P(Y_2 = k) = \sum_{i \in \Phi} P(Y_1 = i, Y_2 = k)$$

resulting in the marginal probability distribution of:

```
for(a in jobs){
  prob = sum(data[,a])
  start = paste("P( Y2 =", a,") = ", prob)
  print(start)
}
```

```
## [1] "P( Y2 = farm ) =  0.023"
## [1] "P( Y2 = operatives ) =  0.26"
## [1] "P( Y2 = craftsmen ) =  0.24"
## [1] "P( Y2 = sales ) =  0.125"
## [1] "P( Y2 = professional ) =  0.352"
```

(2.1c): Find the conditional distribution of a son's occupation, given that the father is a farmer.

$P(Y_2 \mid Y_1 = farm)$ can be found through the following way: Since

$$P(Y_1, Y_2) = P(Y_1)P(Y_2 \mid Y_1)$$

$\Rightarrow$

$$P(Y_2 \mid Y_1) = \frac{P(Y_1, Y_2)}{P(Y_1)} = \frac{P(Y_1, Y_2)}{\sum_{k \in \Phi} P(Y_1, Y_2 = k)}$$

And thus,

$$P(Y_2 \mid Y_1 = farm) = \frac{P(Y_1 = farm, Y_2)}{\sum_{k \in \Phi} P(Y_1 = farm, Y_2 = k)} = \frac{P(Y_1 = farm, Y_2)}{0.11}$$

Where $P(Y_1 = farm) = 0.11$ was taken from part a. Thus, the conditional distribution is as follows:

```
for(a in jobs){
  probY1Y2 = data['farm',a]/sum(data['farm',])
  probY1Y2 = round(probY1Y2,7)
  print(paste("P( Y2 =", a, '| Y1 = farm) = ', probY1Y2))
}
```

```
## [1] "P( Y2 = farm | Y1 = farm) =  0.1636364"
## [1] "P( Y2 = operatives | Y1 = farm) =  0.3181818"
## [1] "P( Y2 = craftsmen | Y1 = farm) =  0.2818182"
## [1] "P( Y2 = sales | Y1 = farm) =  0.0727273"
## [1] "P( Y2 = professional | Y1 = farm) =  0.1636364"
```

(2.1d): Find the conditional distribution of a father's occupation, given that the son is a farmer.

Similar to what was shown in 2.1c,

$$P(Y_1 \mid Y_2 = farm) = \frac{P(Y_1, Y_2 = farm)}{P(Y_2 = farm)} = \frac{P(Y_1, Y_2 = farm)}{0.023}$$

Where $P(Y_2 = farm) = 0.023$ was taken from part b.

Thus, the conditional distribution of the father's occupation given the son is a farmer is:

```
for(a in jobs){
  probY1Y2 = data[a, 'farm']/sum(data[,'farm'])
  probY1Y2 = round(probY1Y2,7)
  print(paste("P( Y1 =", a, '| Y2 = farm) = ', probY1Y2))
}
```

```
## [1] "P( Y1 = farm | Y2 = farm) =  0.7826087"
## [1] "P( Y1 = operatives | Y2 = farm) =  0.0869565"
## [1] "P( Y1 = craftsmen | Y2 = farm) =  0.0434783"
## [1] "P( Y1 = sales | Y2 = farm) =  0.0434783"
## [1] "P( Y1 = professional | Y2 = farm) =  0.0434783"
```

Problem 2.2

Let Y_1 and Y_2 be two independent random variables s.t. $\mathbb{E}[Y_i] = \mu_i$ and $Var[Y_i] = \sigma_i^2$. Compute the following quantiles, where a_1 and a_2 are given constants:

(2.2a): $\mathbb{E}[a_1Y_1 + a_2Y_2], Var[a_1Y_1 + a_2Y_2]$

Let the sample space of Y_i be Φ_i. Then, if $a_i \in \mathbb{R}$

$$\mathbb{E}[a_iY_i] = \sum_{y_{i,k} \in \Phi_i} a_i y_{i,k} P(Y_i = y_{i,k}) = a_i \sum_{y_{i,k} \in \Phi_i} y_{i,k} P(Y_i = y_{i,k}) = a_i \mathbb{E}[Y_i] = a_i \mu_i$$

Also, note that

$$\mathbb{E}[Y_1 + Y_2] = \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} (y_1 + y_2) P(Y_1 = y_1, Y_2 = y_2)$$

and since Y_1 & Y_2 are independent

$$\mathbb{E}[Y_1 + Y_2] = \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} (y_1 + y_2) P(Y_1 = y_1) P(Y_2 = y_2)$$

$$= \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} [y_1 P(Y_1 = y_1) P(Y_2 = y_2) + y_2 P(Y_1 = y_1) P(Y_2 = y_2)]$$

$$= \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} y_1 P(Y_1 = y_1) P(Y_2 = y_2) + \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} y_2 P(Y_1 = y_1) P(Y_2 = y_2)$$

$$= \sum_{y_1 \in \Phi_1} y_1 P(Y_1 = y_1) \sum_{y_2 \in \Phi_2} P(Y_2 = y_2) + \sum_{y_2 \in \Phi_2} y_2 P(Y_1 = y_1) \sum_{y_1 \in \Phi_1} P(Y_2 = y_2)$$

by the law of total probability

$$= \sum_{y_1 \in \Phi_1} y_1 P(Y_1 = y_1) + \sum_{y_2 \in \Phi_2} y_2 P(Y_2 = y_2) = \mathbb{E}[Y_1] + \mathbb{E}[Y_2] = \mu_1 + \mu_2$$

Thus

$$\mathbb{E}[a_1 Y_1 + a_2 Y_2] = \mathbb{E}[a_1 Y_1] + \mathbb{E}[a_2 Y_2] = a_1 \mathbb{E}[Y_1] + a_2 \mathbb{E}[Y_2] = a_1 \mu_1 + a_2 \mu_2$$

Now, for any a_i, Y_i,

$$Var[a_i Y_i] = \mathbb{E}[(a_i Y_i - \mathbb{E}[a_i Y_i])^2] = \mathbb{E}[(a_i Y_i)^2] - (\mathbb{E}[a_i Y_i])^2$$
$$= a_i^2 \mathbb{E}[Y_i^2] - (a_i \mathbb{E}[Y_i])^2 = a_i^2 [\mathbb{E}[Y_i^2] - (\mathbb{E}[Y_i])^2] = a_i^2 Var[Y_i] = a_i^2 \sigma_i^2$$

additionally

$$Var[Y_1 + Y_2] = \mathbb{E}[(Y_1 + Y_2)^2] - (\mathbb{E}[Y_1 + Y_2])^2 = \mathbb{E}[Y_1^2 + 2Y_1Y_2 + Y_2^2] - (\mathbb{E}[Y_1] + \mathbb{E}[Y_2])^2$$
$$= \mathbb{E}[Y_1^2] + \mathbb{E}[2Y_1Y_2] + \mathbb{E}[Y_2^2] - \mathbb{E}[Y_1]^2 - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2] - \mathbb{E}[Y_2]^2$$
$$= \mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]^2 + (2\mathbb{E}[Y_1Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2]) + \mathbb{E}[Y_2^2] - \mathbb{E}[Y_2]^2$$
$$= Var[Y_1] + (2\mathbb{E}[Y_1Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2]) + Var[Y_2]$$
$$= \sigma_1^2 + (2\mathbb{E}[Y_1Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2]) + \sigma_2^2$$

Now, this term is quite interesting. But I shall now show that $(2\mathbb{E}[Y_1Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2]) = 0$ due to the fact that $P(Y_1, Y_2) = P(Y_1)P(Y_2)$. This is shown below.

$$\mathbb{E}[Y_1Y_2] = \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} y_1 y_2 P(Y1, Y2) = \sum_{y_1 \in \Phi_1} \sum_{y_2 \in \Phi_2} y_1 y_2 P(Y1) P(Y2)$$
$$= \sum_{y_1 \in \Phi_1} y_1 P(Y1) \sum_{y_2 \in \Phi_2} y_2 P(Y2) = \mathbb{E}[Y_1]\mathbb{E}[Y_2]$$

Thus, $2\mathbb{E}[Y_1Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2] = 2\mathbb{E}[Y_1]\mathbb{E}[Y_2] - 2\mathbb{E}[Y_1]\mathbb{E}[Y_2] = 0$. Using these facts proved above, the problem is quite simple, with

$$Var[a_1 Y_1 + a_2 Y_2] = Var[a_1 Y_1] + Var[a_2 Y_2] = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2$$

(2.2b): $\mathbb{E}[a_1Y_1 - a_2Y_2], Var[a_1Y_1 - a_2Y_2]$

For this problem, let $b_2 = (-1)a_2$. Then, the answer to $\mathbb{E}[a_1Y_1 + b_2Y_2]$ is simply

$$\mathbb{E}[a_1Y_1 + b_2Y_2] = \mathbb{E}[a_1Y_1] + \mathbb{E}[b_2Y_2] = a_1\mathbb{E}[Y_1] + b_2\mathbb{E}[Y_2] = a_1\mu_1 + b_2\mu_2 = a_1\mu_1 - a_2\mu_2$$

And similarly, the answer to $Var[a_1Y_1 - a_2Y_2]$ is

$$\begin{aligned} Var[a_1Y_1 + b_2Y_2] = Var[a_1Y_1] + Var[b_2Y_2] = a_1^2\sigma_1^2 + b_2^2\sigma_2^2 = a_1^2\sigma_1^2 + (-a_2)^2\sigma_2^2 \\ = a_1^2\sigma_1^2 + (-1)^2a_2^2\sigma_2^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 \end{aligned}$$

Problem 2.3

Let X, Y, Z be random variables with join density $p(x, y, z) \propto f(x, z)g(y, z)h(z)$.

Let Φ_X and Φ_Y represent the sample space of X and Y, respectively.

(2.3a): Show $p(x \mid y, z) \propto f(x, z)$, i.e. $p(x \mid y, z)$ is a function of x and z

By definition, $[p(x, y, z) \propto f(x, z)g(y, z)h(z)] \Rightarrow \exists\beta \in \mathbb{R}$ s.t. $p(x, y, z) = \beta f(x, z)g(y, z)h(z)]$ Using this fact, it can be shown that

$$p(x \mid y, z) = \frac{p(x, y, z)}{p(y, z)} = \frac{p(x, y, z)}{\int_{x\in\Phi_X} p(x, y, z)dx} = \frac{\beta f(x, z)g(y, z)h(z)}{\int_{x\in\Phi_X} \beta f(x, z)g(y, z)h(z)dx}$$

$$p(x \mid y, z) = \frac{\beta g(y, z)h(z)f(x, z)}{\beta g(y, z)h(z)\int_{x\in\Phi_X} f(x, z)dx} = \frac{f(x, z)}{\int_{x\in\Phi_X} f(x, z)dx}$$

Thus, $p(x \mid y, z)$ is a function of x and z.

(2.3b): Show $p(y \mid x, z) \propto g(y, z)$, i.e. $p(y \mid x, z)$ is a function of y and z

This can be found in a similar fashion to part a. Let Φ_Y be the sample space of Y. Then

$$p(y \mid x, z) = \frac{p(x, y, z)}{p(x, z)} = \frac{p(x, y, z)}{\int_{y\in\Phi_Y} p(x, y, z)dy} = \frac{\beta f(x, z)g(y, z)h(z)}{\int_{y\in\Phi_Y} \beta f(x, z)g(y, z)h(z)dy}$$

$$p(y \mid x, z) = \frac{\beta f(x, z)h(z)g(y, z)}{\beta f(x, z)h(z)\int_{y\in\Phi_Y} g(y, z)dy} = \frac{g(y, z)}{\int_{y\in\Phi_Y} g(y, z)dy}$$

and thus, $p(y \mid x, z)$ is a function of only y and z.

(2.3c): X and Y are conditionally independent given Z.

By definition, $X, Y \perp Z$ if $p(x \mid y, z) = p(x \mid z)$ & $p(y \mid x, z) = p(y \mid z)$. Thus, I will solve for $p(x \mid z)$ & $p(y \mid z)$, as follows:

$$p(x \mid z) = \frac{p(x,z)}{p(z)} = \frac{\int_{y\in\Phi_Y} p(x,y,z)dy}{\int_{x\in\Phi_X}\int_{y\in\Phi_Y} p(x,y,z)dydx} = \frac{\int_{y\in\Phi_Y} \beta f(x,z)g(y,z)h(z)dy}{\int_{x\in\Phi_X}\int_{y\in\Phi_Y} \beta f(x,z)g(y,z)h(z)dydx}$$

$$= \frac{\beta f(x,z)h(z)\int_{y\in\Phi_Y} g(y,z)dy}{\beta h(z)\int_{x\in\Phi_X} f(x,z)dx\int_{y\in\Phi_Y} g(y,z)dy} = \frac{\beta h(z)\int_{y\in\Phi_Y} g(y,z)dy}{\beta h(z)\int_{y\in\Phi_Y} g(y,z)dy}\frac{f(x,z)}{\int_{x\in\Phi_X} f(x,z)dx} = \frac{f(x,z)}{\int_{x\in\Phi_X} f(x,z)dx} = p(x \mid y,z)$$

$$\Longrightarrow$$

$$p(x \mid z) = p(x \mid y, z)$$

Similarly for $p(y \mid z)$

$$p(y \mid z) = \frac{p(y,z)}{p(z)} = \frac{\int_{x\in\Phi_X} p(x,y,z)dx}{\int_{y\in\Phi_Y}\int_{x\in\Phi_X} p(x,y,z)dxdy} = \frac{\int_{x\in\Phi_X} \beta g(y,z)f(x,z)h(z)dx}{\int_{y\in\Phi_Y}\int_{x\in\Phi_X} \beta f(x,z)g(y,z)h(z)dxdy}$$

$$= \frac{\beta g(y,z)h(z)\int_{x\in\Phi_X} f(x,z)dx}{\beta h(z)\int_{y\in\Phi_Y} g(y,z)dy\int_{x\in\Phi_X} f(x,z)dx} = \frac{\beta h(z)\int_{x\in\Phi_X} f(x,z)dx}{\beta h(z)\int_{x\in\Phi_X} f(x,z)dx}\frac{g(y,z)}{\int_{y\in\Phi_Y} g(y,z)dy} = \frac{g(y,z)}{\int_{y\in\Phi_Y} g(y,z)dy} = p(y \mid x,z)$$

$$\Longrightarrow$$

$$p(y \mid z) = p(y \mid x, z)$$

Thus

$$(p(y \mid z) = p(y \mid x, z)) \wedge (p(x \mid z) = p(x \mid y, z)) \Longrightarrow (x \perp y \mid z)$$

Problem 2.6

Suppose $A \perp B \mid C$. Show $(A \perp B \mid C) \Rightarrow (A^c \perp B \mid C) \wedge (A \perp B^c \mid C) \wedge (A^c \perp B^c \mid C)$

By definition, $(A \perp B \mid C) \Rightarrow P(A \mid B, C) = P(A \mid C)$. Thus

$$(A \perp B \mid C) \Rightarrow [P(A \mid B, C) = P(A \mid C)] \Rightarrow [1 - P(A^c \mid B, C) = P(A \mid C)]$$

$$\Longrightarrow [P(A^c \mid B, C) = 1 - P(A \mid C)] \Rightarrow [P(A^c \mid B, C) = P(A^c \mid C)] \Rightarrow (A^c \perp B \mid C)$$

And thus, $(A \perp B \mid C) \Rightarrow (A^c \perp B \mid C)$. Similarly, by definition, $(A \perp B \mid C) \Rightarrow P(B \mid A, C) = P(B \mid C)$. Thus

$$(A \perp B \mid C) \Rightarrow [P(B \mid A, C) = P(B \mid C)] \Rightarrow [1 - P(B^c \mid A, C) = P(B \mid C)]$$

$$\Longrightarrow [P(B^c \mid A, C) = 1 - P(B \mid C)] \Rightarrow [P(B^c \mid A, C) = P(B^c \mid C)] \Rightarrow (B^c \perp A \mid C)$$

And thus, $(A \perp B \mid C) \Rightarrow (A \perp B^c \mid C)$. Using this fact, I show that

$$(A \perp B \mid C) \Rightarrow (A \perp B^c \mid C) \Rightarrow [P(A \mid B^c, C) = P(A \mid C)] \Rightarrow [1 - P(A^c \mid B^c, C) = P(A \mid C)]$$

$$\Longrightarrow [P(A^c \mid B^c, C) = 1 - P(A \mid C)] = P(A^c \mid C)$$
$$\Longrightarrow A^c \perp B^c \mid C$$

and thus

$$(A \perp B \mid C) \Longrightarrow (A^c \perp B^c \mid C)$$

Find an example where $A \perp B \mid C$ holds but $(A \perp B \mid C^c)$ does not.

Suppose A is the event that a lamp sitting in a room far away from here turns on. Suppose B is the event that Haley pushes a button which turns on the lamp (assume that Haley is far away from the lamp and does not know if it is turned on or off). Finally, suppose event C is the event that Sam pushes a button which turns on the lamp (assume Sam is far away from the lamp and does not know if it is on or off). Suppose that $P(B) = P(C) = 1/2$. Assume futher that the buttons Sam and Haley hold are the only means of turning on this lamp. (i.e. $P(A \mid B^c, C^c) = 0$). Take note that

$$B \Rightarrow A$$

and

$$C \Rightarrow A$$

however $P(B \mid A, C) = P(B \mid C) = P(B) = 1/2$. Moreover, if C is to occur, A must also occur, so $P(A \mid B, C) = P(A \mid C) = 1$. Given this set up, $A \perp B \mid C$ since $P(A \mid B, C) = P(A \mid C)$ & $P(B \mid A, C) = P(B \mid C)$. However, if C^c occurs (i.e. Sam does not push his button), then

$$P(A \mid C^c) = P(B) = 1/2$$

Since $(B \mid C^c) \Leftrightarrow (A \mid C^c)$ due to the fact that the only way for A to occur or have occured is for Haley to have pushed her button. Thus

$$P(A \mid B, C^c) = 1 \neq P(A \mid C^c) = P(B) = 1/2$$

Thus, in this situation, $A \perp B \mid C$, however $\sim(A \perp B \mid C^c)$.