# STA 601 homework 1

*Lingyun Shao, MS stats*

*Sep. 2, 2018*

## 2.1

### a)

This distribution can be obtained by calculating the row sums of the original chart. The marginal probability distribution of a father's occupation is:

| Father's occupation ($y_1$) | farm | operatives | craftsmen | sale | professional |
|---|---|---|---|---|---|
| Marginal distribution ($p_{Y_1}(y_1)$) | 0.11 | 0.279 | 0.277 | 0.099 | 0.235 |

### b)

This distribution can be obtained by calculating the column sums of the original chart. The marginal probability distribution of a son's occupation is:

| Son's occupation ($y_2$) | farm | operatives | craftsmen | sale | professional |
|---|---|---|---|---|---|
| Marginal distribution ($p_{Y_2}(y_2)$) | 0.023 | 0.26 | 0.24 | 0.125 | 0.352 |

### c)

This distribution can be obtained by normalizing (each probability divided by the sum of them) the row where the father's occupation is 'farm'. The conditional distribution of a son's occupation, given that the father is a farmer, is:

| Son's occupation ($y_2$) | farm | operatives | craftsmen | sale | professional |
|---|---|---|---|---|---|
| Conditional distribution ($p_{Y_2\|Y_1}(y_2\|y_1 = farm)$) | 0.1635 | 0.318 | 0.282 | 0.073 | 0.1635 |

*Note*: The round-ups of last digits might vary because I try to make the sum equal to 1.

### d)

This distribution can be obtained by normalizing the column where the son's occupation is 'farm'. The conditional distribution of a father's occupation, given that the son is a farmer, is:

| Father's occupation ($y_1$) | farm | operatives | craftsmen | sale | professional |
|---|---|---|---|---|---|
| Conditional distribution ($p_{Y_1\|Y_2}(y_1\|y_2 = farm)$) | 0.7826 | 0.0869 | 0.0435 | 0.0435 | 0.0435 |

*Note*: The round-ups of last digits might vary because I try to make the sum equal to 1.

## 2.2

### a)

Suppose $Y_1, Y_2$ are continuous (similarly for discrete situations, just substitute integral with sum and its corresponding notation) and respectively follow the pdf of $p_{Y_1}(y_1), p_{Y_2}(y_2)$.

By finition, $E[Y_i] = \int_{\mathcal{Y}_i} y_i p(y_i) dy_i$, where $p(y_i)$ refers to $p_{Y_i}(y_i)$. Because of independence, $p(y_1, y_2) = p(y_1)p(y_2)$. Therefore,

$$
\begin{aligned}
E[a_1Y_1 + a_2Y_2] &= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} (a_1y_1 + a_2y_2)p(y_1, y_2)dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} (a_1y_1 + a_2y_2)p(y_1)p(y_2)dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} [a_1y_1p(y_1)p(y_2) + a_2y_2p(y_1)p(y_2)]dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} a_1y_1p(y_1)p(y_2)dy_1dy_2 + \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} a_2y_2p(y_1)p(y_2)dy_1dy_2 \\
&= a_1\int_{\mathcal{Y}_1} y_1p(y_1)dy_1 + a_2\int_{\mathcal{Y}_2} y_2p(y_2)dy_2 \qquad (\because \int_{\mathcal{Y}_i} p(y_i)dy_i = 1) \\
&= a_1\mu_1 + a_2\mu_2
\end{aligned}
$$

By the definition of variance, $Var[Y_i] = EY_i^2 - [EY_i]^2$. Therefore, $EY_i^2 = Var[Y_i] + E[Y_i]^2 = \mu_i^2 + \sigma_i^2$

$$
\begin{aligned}
Var[a_1Y_1 + a_2Y_2] &= E[(a_1Y_1 + a_2Y_2)^2] - (E[a_1Y_1 + a_2Y_2])^2 \\
&= E[a_1^2Y_1^2 + a_2^2Y_2^2 + 2a_1a_2Y_1Y_2] - (a_1\mu_1 + a_2\mu_2)^2 \\
&= a_1^2E[Y_1]^2 + a_2^2E[Y_2]^2 + 2a_1a_2E[Y_1Y_2] - (a_1\mu_1 + a_2\mu_2)^2 \qquad (\because \textit{property of } E) \\
&= a_1^2(\mu_1^2 + \sigma_1^2) + a_2^2(\mu_2^2 + \sigma_2^2) + 2a_1a_2E[Y_1]E[Y_2] - (a_1\mu_1 + a_2\mu_2)^2 \qquad (\because \textit{independence}) \\
&= a_1^2(\mu_1^2 + \sigma_1^2) + a_2^2(\mu_2^2 + \sigma_2^2) + 2a_1a_2\mu_1\mu_2 - (a_1\mu_1 + a_2\mu_2)^2 \\
&= a_1^2\sigma_1^2 + a_2^2\sigma_2^2
\end{aligned}
$$

**b)**

$$
\begin{aligned}
E[a_1Y_1 - a_2Y_2] &= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} (a_1y_1 - a_2y_2)p(y_1, y_2)dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} (a_1y_1 - a_2y_2)p(y_1)p(y_2)dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} [a_1y_1p(y_1)p(y_2) - a_2y_2p(y_1)p(y_2)]dy_1dy_2 \\
&= \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} a_1y_1p(y_1)p(y_2)dy_1dy_2 - \int_{\mathcal{Y}_2}\int_{\mathcal{Y}_1} a_2y_2p(y_1)p(y_2)dy_1dy_2 \\
&= a_1\int_{\mathcal{Y}_1} y_1p(y_1)dy_1 - a_2\int_{\mathcal{Y}_2} y_2p(y_2)dy_2 \qquad (\because \int_{\mathcal{Y}_i} p(y_i)dy_i = 1) \\
&= a_1\mu_1 - a_2\mu_2
\end{aligned}
$$

$$
\begin{aligned}
Var[a_1Y_1 - a_2Y_2] &= E[(a_1Y_1 - a_2Y_2)^2] - (E[a_1Y_1 - a_2Y_2])^2 \\
&= E[a_1^2Y_1^2 + a_2^2Y_2^2 - 2a_1a_2Y_1Y_2] - (a_1\mu_1 - a_2\mu_2)^2 \\
&= a_1^2E[Y_1]^2 + a_2^2E[Y_2]^2 - 2a_1a_2E[Y_1Y_2] - (a_1\mu_1 - a_2\mu_2)^2 \qquad (\because \textit{property of } E) \\
&= a_1^2(\mu_1^2 + \sigma_1^2) + a_2^2(\mu_2^2 + \sigma_2^2) - 2a_1a_2E[Y_1]E[Y_2] - (a_1\mu_1 - a_2\mu_2)^2 \qquad (\because \textit{independence}) \\
&= a_1^2(\mu_1^2 + \sigma_1^2) + a_2^2(\mu_2^2 + \sigma_2^2) - 2a_1a_2\mu_1\mu_2 - (a_1\mu_1 - a_2\mu_2)^2 \\
&= a_1^2\sigma_1^2 + a_2^2\sigma_2^2
\end{aligned}
$$

## 2.3

The supports of $X, Y, Z$ are denoted by $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$. We suppose these random variables are continuous (results can be similarly proved in discrete situations by substituting integral with sum and its corresponding notation) $p(x, y, z) \propto f(x, z)g(y, z)h(z)$, let $p(x, y, z) = Cf(x, z)g(y, z)h(z)$, where $C \neq 0$ is a constant number.

### a)

$$
\begin{aligned}
p(x|y,z) &= \frac{p(x,y,z)}{p(y,z)} \\
&= \frac{p(x,y,z)}{\int_{\mathcal{X}} p(x,y,z)dx} \\
&= \frac{Cf(x,z)g(y,z)h(z)}{\int_{\mathcal{X}} Cf(x,z)g(y,z)h(z)dx} \\
&= \frac{Cf(x,z)g(y,z)h(z)}{Cg(y,z)h(z)\int_{\mathcal{X}} f(x,z)dx} \\
&= \frac{f(x,z)}{\int_{\mathcal{X}} f(x,z)dx} \\
&\propto f(x,z)
\end{aligned}
$$

Obviously, $p(x|y,z)$ is a function of $x$ and $z$.

### b)

$$
\begin{aligned}
p(y|x,z) &= \frac{p(x,y,z)}{p(x,z)} \\
&= \frac{p(x,y,z)}{\int_{\mathcal{Y}} p(x,y,z)dy} \\
&= \frac{Cf(x,z)g(y,z)h(z)}{\int_{\mathcal{Y}} Cf(x,z)g(y,z)h(z)dy} \\
&= \frac{Cf(x,z)g(y,z)h(z)}{Cf(x,z)h(z)\int_{\mathcal{Y}} g(y,z)dy} \\
&= \frac{g(y,z)}{\int_{\mathcal{Y}} g(y,z)dy} \\
&\propto g(y,z)
\end{aligned}
$$

Obviously, $p(y|x,z)$ is a function of $y$ and $z$.

### c)

$$
\begin{aligned}
p(z) &= \int_{\mathcal{Y}}\int_{\mathcal{X}} p(x,y,z)dxdy \\
&= \int_{\mathcal{Y}}\int_{\mathcal{X}} Cf(x,z)g(y,z)h(z)dxdy \\
&= C\ h(z)\int_{\mathcal{Y}} g(y,z)dy\int_{\mathcal{X}} f(x,z)dx
\end{aligned}
$$

According to part a), we already have $p(x|y,z) = \frac{f(x,z)}{\int_{\mathcal{X}} f(x,z)dx}$

$$\begin{aligned} p(x|z) &= \frac{p(x,z)}{p(z)} \\ &= \frac{\int_{\mathcal{Y}} p(x,y,z)dy}{p(z)} \\ &= \frac{C\ f(x,z)h(z)\int_{\mathcal{Y}} g(y,z)dy}{C\ h(z)\int_{\mathcal{Y}} g(y,z)dy \int_{\mathcal{X}} f(x,z)dx} \\ &= \frac{f(x,z)}{\int_{\mathcal{Y}} f(x,z)dx} \\ &= p(x|y,z) \end{aligned}$$

Similarly, we can prove from the other side and get $p(y|z) = p(y|x,z)$. It's no difference.

$\frac{p(x,z)}{p(z)} = p(x|z) = p(x|y,z) = \frac{p(x,y,z)}{p(y,z)}$, so $\frac{p(x,z)}{p(z)}\frac{p(y,z)}{p(z)} = \frac{p(x,y,z)}{p(z)}$. Therefore, we have $p(x|z)p(y|z) = p(x,y|z)$.

Actually, $p(x|z) = p(x|y,z)$ is equivalent to conditional independence and it means that once we know the information of $Z$, $Y$ provides no extra information about $X$. Therefore, we have proved that $X$ and $Y$ are conditionally independent given $Z$.

## 2.6

### a)

Given $A \perp B|C$, $Pr(A \bigcap B|C) = Pr(A|C)Pr(B|C)$.

$$\begin{aligned} Pr(A^c \cap B|C) &= Pr(B|C) - Pr(A \cap B|C) \\ &= Pr(B|C) - Pr(A|C)Pr(B|C) \\ &= [1 - Pr(A|C)]Pr(B|C) \\ &= Pr(A^c|C)Pr(B|C) \end{aligned}$$

$$\begin{aligned} Pr(A \cap B^c|C) &= Pr(A|C) - Pr(A \cap B|C) \\ &= Pr(A|C) - Pr(A|C)Pr(B|C) \\ &= [1 - Pr(B|C)]Pr(A|C) \\ &= Pr(B^c|C)Pr(A|C) \end{aligned}$$

$$\begin{aligned} Pr(A^c \cap B^c|C) &= Pr(A^c|C) - Pr(A^c \cap B|C) \\ &= Pr(A^c|C) - [Pr(B|C) - Pr(A \cap B|C)] \\ &= [1 - Pr(A|C)] - Pr(B|C) + Pr(A|C)(B|C) \\ &= [1 - Pr(A|C)][1 - Pr(B|C)] \\ &= P(A^c|C)P(B^c|C) \end{aligned}$$

As shown above, $A^c \perp B|C, A \perp B^c|C, A^c \perp B^c|C$ all hold if $A \perp B|C$ holds.

**b)**

One example where $A \perp B | C$ holds but $A \perp B | C^c$ does not is:

Suppose $X_1, X_2$ are i.i.d random variables from $N(0,1)$. We choose twice from $\{X_1, X_2\}$ with replacement to get two random variables $T_1, T_2$.

$$A = \{T_1 > 0\}, B = \{T_2 > 0\}, C = \{We\ did\ not\ choose\ the\ same\ random\ variable\}$$

In this case, $Pr(A \cap B | C) = 1/4$ and $Pr(A|C) = Pr(B|C) = 1/2$. However, $Pr(A \cap B | C^c) = 1/2$ and $Pr(A|C^c) = Pr(B|C^c) = 1/2$. So this is a example where $A \perp B | C$ holds but $A \perp B | C^c$ does not.