Bayesian theory and data analysis - HW - 1

Anirudh K M - anikamal

8th February, 2016

1 Problem 1

Entry level test checked.

2 Problem 2

Table 1: Probability of joint distribution based on father and son occupations

Father \ Son	Farm	Operatives	Craftmen	Sales	Professional
Farm	0.018	0.035	0.031	0.008	0.018
Operatives	0.002	0.112	0.064	0.032	0.069
Craftmen	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 P1 and P2 be random variables denoting fathers and sons occupation respectively.

Rule of marginal probability

$$P_r(P1) = \sum_{k=1}^{K} P_r(P1 \cap P2_k)$$

a) The marginal probability distribution of father's occupation.

Based on the rule of marginal probability, we could find the probability distribution for father's occupation.

P$_r(P1 = farm) = 0.018 + 0.035 + 0.031 + 0.008 + 0.018 = 0.11$

$P_r(P1 = operatives) = 0.002 + 0.112 + 0.064 + 0.032 + 0.069 = 0.279$

$P_r(P1 = craftsmen) = 0.001 + 0.066 + 0.094 + 0.032 + 0.084 = 0.277$

$P_r(P1 = sales) = 0.001 + 0.018 + 0.019 + 0.010 + 0.051 = 0.099$

$P_r(P1 = professional) = 0.001 + 0.029 + 0.032 + 0.043 + 0.130 = 0.235$

b) The marginal probability distribution of son's occupation.

We could compute this too based on marginal probability again.

$\mathrm{P}_r(P2 = farm) = 0.018 + 0.002 + 0.001 + 0.001 + 0.001 = 0.023$

$P_r(P2 = operatives) = 0.035 + 0.112 + 0.066 + 0.018 + 0.029 = 0.26$

$P_r(P2 = craftsmen) = 0.031 + 0.064 + 0.094 + 0.019 + 0.032 = 0.24$

$P_r(P2 = sales) = 0.008 + 0.032 + 0.032 + 0.010 + 0.043 = 0.125$

$P_r(P2 = professional) = 0.018 + 0.069 + 0.084 + 0.051 + 0.130 = 0.352$

c) The conditional distribution of sons occupation given that father is a farmer

$$\mathrm{P}_r(P2 = farmer|P1 = farmer) = \frac{P_r(P2 = farmer \cap P1 = farmer)}{(P1 = farmer)}$$

$$\equiv \frac{0.018}{0.11} = 0.163$$

$$\mathrm{P}_r(P2 = operatives|P1 = farmer) = \frac{P_r(P2 = operatives \cap P1 = farmer)}{(P1 = farmer)}$$

$$\equiv \frac{0.035}{0.11} = 0.318$$

$$\mathrm{P}_r(P2 = craftsmen|P1 = farmer) = \frac{P_r(P2 = craftsmen \cap P1 = farmer)}{(P1 = farmer)}$$

$$\equiv \frac{0.031}{0.11} = 0.281$$

$$\mathrm{P}_r(P2 = sales|P1 = farmer) = \frac{P_r(P2 = sales \cap P1 = farmer)}{(P1 = farmer)}$$

$$\equiv \frac{0.008}{0.11} = 0.072$$

$$\mathrm{P}_r(P2 = professional|P1 = farmer) = \frac{P_r(P2 = professional \cap P1 = farmer)}{(P1 = farmer)}$$

$$\equiv \frac{0.018}{0.11} = 0.163$$

d) The conditional distribution of fathers occupation given that son is a farmer

$$\mathrm{P}_r(P1 = farmer|P2 = farmer) = \frac{P_r(P1 = farmer \cap P2 = farmer)}{(P2 = farmer)}$$

$$\equiv \frac{0.018}{0.023} = 0.78$$

$$\mathrm{P}_r(P1 = operatives|P2 = farmer) = \frac{P_r(P1 = operatives \cap P2 = farmer)}{(P2 = farmer)}$$

$$\equiv \frac{0.002}{0.023} = 0.086$$

$$\mathrm{P}_r(P1 = craftsmen|P2 = farmer) = \frac{P_r(P1 = craftsmen \cap P2 = farmer)}{(P2 = farmer)}$$

$$\equiv \frac{0.001}{0.023} = 0.0435$$

$$\mathrm{P}_r(P1 = sales|P2 = farmer) = \frac{P_r(P1 = sales \cap P2 = farmer)}{(P2 = farmer)}$$

$$\equiv \frac{0.001}{0.023} = 0.0435$$

$$\mathrm{P}_r(P1 = professional|P2 = farmer) = \frac{P_r(P1 = professional \cap P2 = farmer)}{(P2 = farmer)}$$

$$\equiv \frac{0.001}{0.023} = 0.0435$$

3 Problem 3

Given:

$$p(x, y, z) \propto f(x, z)g(y, z)h(z)$$

Assumption

We will assume this to be continuous joint density.

a

To Prove that

$$p(x|y, z) \propto f(x, z)$$

We know that by conditional probability, we have

$$p(x|y, z) = \frac{p(x, y, z)}{p(y, z)}$$

But from the given we can say that,

$$p(x, y, z) = kf(x, z)g(y, z)h(z)$$

where k is a constant.

We can represent

$$p(y, z) = \int p(x, y, z)dx$$

This can be further simplified as follows

$$p(y, z) = \int f(x, z)g(y, z)h(z)dx$$

$$\equiv p(y, z) = g(y, z)h(z) \int f(x, z)dx$$

Now g(y,z) and h(z) acts like a constant.

Substituting this in the first equation we get

$$p(x|y, z) = \frac{kf(x, z)g(y, z)h(z)}{g(y, z)h(z) \int f(x, z)dx}$$

From this we can easily say that

$$p(x|y, z) \propto \frac{f(x, z)}{\int f(x, z)dx}$$

This implies that $p(x|y, z)$ is a function of x and z alone.

b

To prove that

$$p(y|x, z) \propto g(y, z)$$

We know that by conditional probability, we have

$$p(y|x, z) = \frac{p(x, y, z)}{p(x, z)}$$

But from the given we can say that,

$$p(x, y, z) = kf(x, z)g(y, z)h(z)$$

where k is a constant.

We can represent

$$p(x, z) = \int p(x, y, z)dy$$

This can be further simplified as follows

$$p(x, z) = \int f(x, z)g(y, z)h(z)dy$$

$$\equiv p(x, z) = f(x, z)h(z) \int g(y, z)dy$$

Now f(x,z) and h(z) acts like a constant.

Substituting this in the first equation we get

$$p(y|x,z) = \frac{kf(x,z)g(y,z)h(z)}{f(x,z)h(z)\int g(y,z)dy}$$

From this we can easily say that

$$p(y|x,z) \propto \frac{g(y,z)}{\int g(y,z)dy}$$

This implies that $p(y|x,z)$ is a function of y and z alone.

c

To show that X and Y are conditionally independent given Z.

We know

$$p(x,y|z) = \frac{p(x,y,z)}{p(z)}$$

But from the given we can say that,

$$p(x,y,z) = kf(x,z)g(y,z)h(z)$$

where k is a constant.

We can represent

$$p(z) = \iint p(x,y,z)dxdy$$

$$\equiv h(z) \iint f(x,z)g(y,z)dxdy$$

From these equations we can say that

$$p(x,y|z) \propto \frac{f(x,z)g(y,z)}{\iint f(x,z)g(y,z)dxdy}$$

From this we can conclude that X and Y are conditionally independent given Z.

4 Problem 4

a

In this case, we see that A has already observed the outcome. So he knows whether the die showed up a six or not.

Based on we can say that

$$Pr(E|I_A) = 0$$

When the dice doesn't show up six.

$$Pr(E|I_A) = 1$$

When the dice turned out to be six.

On the other hand, B doesn't observe the outcome, so we can say that

$$Pr(E|I_B) = \frac{1}{Total\ number\ of\ sides\ of\ fair\ die}$$
$$= \frac{1}{6}$$

Where 6 is the number of sides of a die.

b

Since A has no knowledge on soccer, he gives equal chance to all teams

$$(Pr(E|I_A)) = \frac{1}{Total\ number\ of\ teams}$$

But it is given that B is a knowledgable sports fan, this can be divided in two cases.

He has no knowledge in soccer but on other sports

Then,

$$(Pr(E|I_B)) = \frac{1}{Total\ number\ of\ teams}$$

He has knowledge in soccer

Then,

$$Pr(E|I_B) > Pr(E|I_A)$$

or

$$Pr(E|I_B) < Pr(E|I_A)$$

Then the probability given by B will become higher than that of A or lesser than depending on the performance measure of team Brazil.

References

1. A First Course in Bayesian Statistical Methods(Peter D Hoff)
2. Think Bayesian (Allen B. Downey)