Economics 257 – Midterm Exam

1. The textbook provides two seemingly different definitions of the independence of two events, A and B. They are:

(a) Two events A and B are independent if and only if $P(A|B) = P(A)$ (for $P(B) > 0$); equivalently $P(B|A) = P(B)$ (for $P(A) > 0$).

(b) Two events A and B are independent if and only if $P(A \cap B) = P(A)P(B)$.

Prove that the two definitions are equivalent, that is, that (a) implies (b) and that (b) implies (a).

The definition of the conditional probability $P(A|B)$ is

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

and similarly for $P(B|A)$. Based on this definition, Bayes' Theorem tells us that

$$P(A|B)P(B) = P(B|A)P(A) = P(A \cap B).$$

From this, if (a) holds, so that $P(A|B) = P(A)$, we see that $P(A \cap B) = P(A)P(B)$, and if $P(B|A) = P(B)$ we see similarly that $P(A \cap B) = P(B)P(A)$. This shows that (a) implies (b).

If (b) holds, then, since $P(B|A) = P(A \cap B)/P(A)$ if $P(A) > 0$, it follows that $P(B|A) = P(A)P(B)/P(A) = P(B)$. In exactly the same way, one shows that, if (b) holds, then $P(A|B) = P(A)$. This shows that (b) implies (a).

Are the events Ω (the entire outcome space) and $\emptyset$ (the null set) independent? If so, why; if not why not?

We recall that $P(\Omega) = 1$ and $P(\emptyset) = 0$. Thus $P(\Omega)P(\emptyset) = 1 \cdot 0 = 0$. But $\Omega \cap \emptyset = \emptyset$, and so $P(\Omega \cap \emptyset) = P(\emptyset) = 0$. Thus, by (b) above, Ω and $\emptyset$ are independent. For (a), if we let $A = \Omega$ and $B = \emptyset$, we cannot use the definition that requires $P(B) > 0$. The other way round, there is no problem, since $P(\Omega) > 0$, and then independence would mean that $P(\emptyset|\Omega) = P(\emptyset)$ We know at once that $P(\emptyset) = 0$, and by definition $P(\emptyset|\Omega) = P(\emptyset \cap \Omega)/P(\Omega) = P(\emptyset)/P(\Omega) = 0/1 = 0$, and so, by this definition as well, we can conclude that Ω and $\emptyset$ are independent.

2. A large industrial firm uses three local motels to provide overnight accommodations for its clients. From past experience it is known that 20% of the clients are assigned rooms at the Ramada Inn, 50% at the Sheraton, and 30% at the Lakeview Motor Lodge. If the plumbing is faulty in 5% of the rooms at the Ramada Inn, in 4% of the rooms at the Sheraton, and in 8% of the rooms at the Lakeview Motor Lodge, what is the probability that

(a) a client will be assigned a room with faulty plumbing?

(b) a person with a room having faulty plumbing was assigned accommodations at the Lakeview Motor Lodge?

(a) Start by defining the events of interest: A_1 a client assigned to the Ramada Inn, A_2 a client assigned to the Sheraton, A_3 a client assigned to the Lakeview. We have:

$$P(A_1) = 0.2, \quad P(A_2) = 0.5, \quad P(A_3) = 0.3.$$

These three events are mutually exclusive and exhaustive. Then more events: B_1 faulty plumbing at the Ramada, B_2 ditto at the Sheraton, B_3 ditto at the Lakeview. We have

$$P(B_1) = 0.05, \quad P(B_2) = 0.04, \quad P(B_3) = 0.08.$$

The event that a client is assigned a room at the Ramada with faulty plumbing is $A_1 \cap B_1$, and similarly for the other two motels. We must assume that clients are assigned rooms randomly, which implies that A_i is independent of B_i, $i = 1, 2, 3$. Thus the probability of a random client getting a room at the Ramada with faulty plumbing is $P(A_1 \cap B_1) = P(A_1)P(B_1) = 0.2 \times 0.05 = 0.01$. Then $P(A_2 \cap B_2) = 0.5 \times 0.04 = 0.02$, and $P(A_3 \cap B_3) = 0.3 \times 0.08 = 0.024$. Now the event where a client is assigned a room with faulty plumbing is the disjoint union of the events, of the form $A_i \cap B_i$, for which a client gets a faulty room in a given motel. We may write this event as $\bigcup_{i=1}^{3}(A_i \cap B_i)$. The probability of a disjoint union of events is the sum of the probabilities of the events, and so here we have

$$P\big((A_1 \cap B_1) \cup (A_2 \cap B_2) \cup (A_3 \cap B_3)\big) = 0.01 + 0.02 + 0.024 = 0.054.$$

(b) Here, we want a conditional probability. The conditioning event is the client getting a room with faulty plumbing. That's the event $\bigcup_{i=1}^{3}(A_i \cap B_i)$, with probability 0.054. The event of which we seek the probability is A_3, where the client is assigned to the Lakeview. Thus we want

$$P\Big(A_3 \mid \bigcup_{i=1}^{3}(A_i \cap B_i)\Big) = \frac{P\big(A_3 \cap \bigcup_{i=1}^{3}(A_i \cap B_i)\big)}{P\big(\bigcup_{i=1}^{3}(A_i \cap B_i)\big)}$$

We saw that the denominator in the last expression is 0.054. By the distributive property, we have, for any three events A, B, and C, that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

If we apply this to the event in the numerator, we see that

$$A_3 \cap \bigcup_{i=1}^{3} (A_i \cap B_i) = \bigcup_{i=1}^{3} \big(A_3 \cap (A_i \cap B_i)\big).$$

Since the A_i are mutually exclusive, we see that only the last event on the right-hand side, with $i = 3$, is non-empty. It is $A_3 \cap (A_3 \cap B_3) = A_3 \cap B_3$, and we saw that the probability of this is 0.024. So, the probability we are seeking, of the event of getting a room at the Lakeview, conditional on being assigned a room with faulty plumbing, is $0.024/0.054 = 4/9$.

3. From a sack of fruit containing 3 oranges, 2 apples, and 3 bananas, a random sample of 4 pieces of fruit is selected. If X is the number of oranges and Y is the number of apples in the sample, find

(a) the joint probability distribution of X and Y;

(b) $P[(X,Y) \in A]$, where A is the region that is given by $\{(x,y) \mid x + y \le 2\}$.

The sack contains 3+2+3 = 8 pieces of fruit. Of these, 4 are selected. Let Z be the number of bananas in the sample. Then the possible values of the triple of random variables (X,Y,Z) are, for $Z = 3$, (1,0,3) and (0,1,3); for $Z = 2$, (2,0,2), (1,1,2), and (0,2,2); for $Z = 1$, (3,0,1), (2,1,1), (1,2,1); for $Z = 0$, (3,1,0) and (2,2,0). Note that Y, the number of apples, cannot exceed 2, since there are only 2 apples in the sack. Note also that there must be at least one orange or one apple in the set of four, since there are only three bananas.

Consider first the event $(X,Y,Z) = (1,0,3)$. There are 4 ways of arriving at this, as the orange could be the first, second, third, or fourth piece of fruit to come out of the sack. These four ways are all equally probable, and the probability of the first is

$$\frac{3}{8} \cdot \frac{3}{7} \cdot \frac{2}{6} \cdot \frac{1}{5} = \frac{3}{280}.$$

Explanation: the probability of an orange first out is 3/8, since, at the beginning, there are 3 oranges among the 8 pieces of fruit. This is multiplied by the probability of a banana next, which is 3/7, since there are now 3 bananas among 7 pieces of fruit. For a banana after that, we get 2/6, and, for the last banana, 1/5. Since this outcome, $(1,0,3)$, can be arrived at in 4 ways, its probability is $4 \cdot 3/280 = 3/70$.

We can argue similarly for all the other possible events. Altogether we have:

Event	no of ways	probability of each way	total probability
(1,0,3)	4	$3/8 \cdot 3/7 \cdot 2/6 \cdot 1/5$	3/70
(0,1,3)	4	$2/8 \cdot 3/7 \cdot 2/6 \cdot 1/5$	2/70
(2,0,2)	6	$3/8 \cdot 2/7 \cdot 3/6 \cdot 2/5$	9/70
(1,1,2)	12	$3/8 \cdot 2/7 \cdot 3/6 \cdot 2/5$	18/70
(0,2,2)	6	$2/8 \cdot 1/7 \cdot 3/6 \cdot 2/5$	3/70
(3,0,1)	4	$3/8 \cdot 2/7 \cdot 1/6 \cdot 2/5$	3/70
(2,1,1)	12	$3/8 \cdot 2/7 \cdot 2/6 \cdot 3/5$	18/70
(1,2,1)	12	$3/8 \cdot 2/7 \cdot 1/6 \cdot 3/5$	9/70
(3,1,0)	4	$3/8 \cdot 2/7 \cdot 1/6 \cdot 2/5$	2/70
(2,2,0)	6	$3/8 \cdot 2/7 \cdot 2/6 \cdot 1/5$	3/70

We can check that the sum of the probabilities of all the events is 1.

The joint distribution of X and Y can be read off from the table above. The distribution would usually be presented as in the following table, in which the probabilities are given for values of X by column, and values of Y by rows.

$Y \downarrow \backslash\ X \rightarrow$	0	1	2	3
0	0	3/70	9/70	3/70
1	2/70	18/70	18/70	2/70
2	3/70	9/70	3/70	0

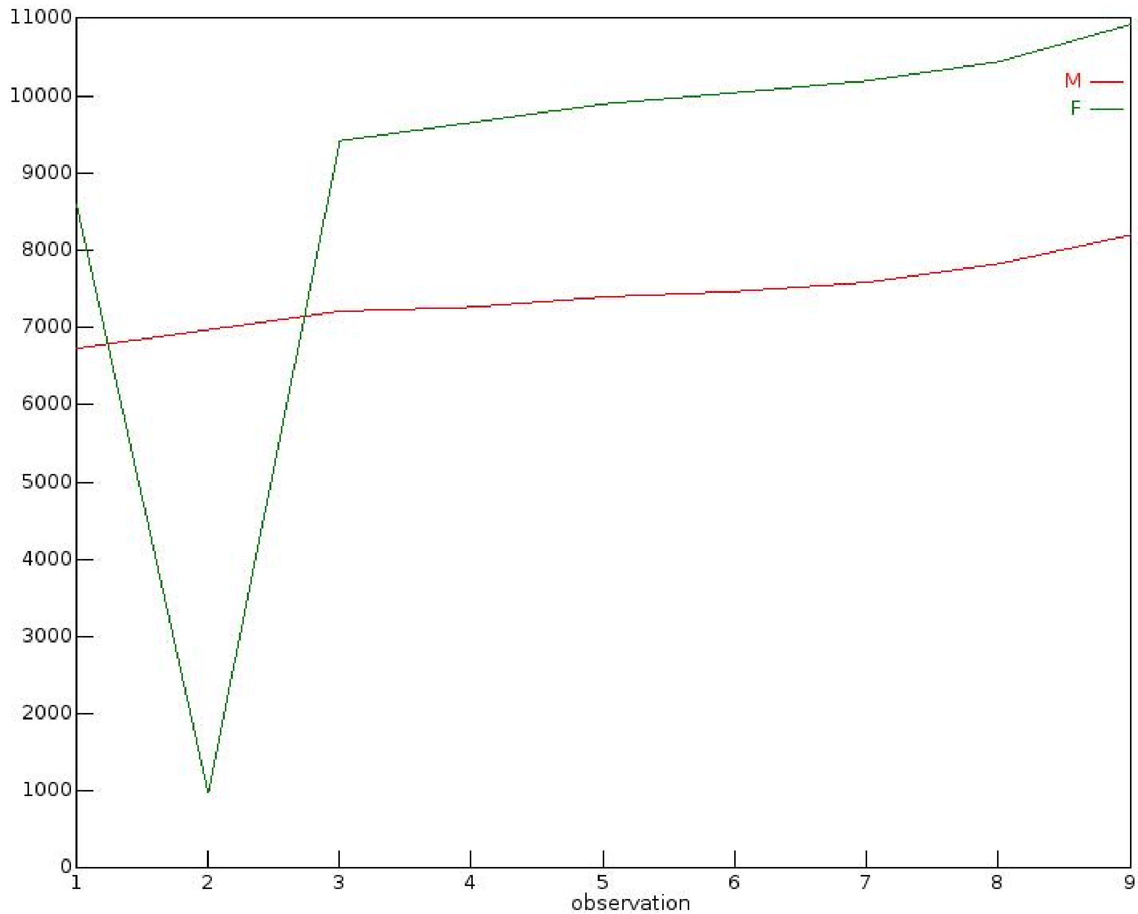
The event A corresponds to the following events: $(X, Y) = (0,0), (1,0), (2,0), (0,1), (1,1),$ $(0,2)$, and, since these events are mutually exclusive, the probability of the event A is the sum of the probabilities of these events:

$$(0 + 3 + 9 + 2 + 18 + 3)/70 = 35/70 = 1/2.$$

4. The number of males and females enrolled in colleges (undergraduate and postbaccalaureate) in the United States from 2000 through 2008 is given here. Graphically present these data with a time-series plot.

College Enrollment (in thousands)	Males	Females
2000	6,721.8	8,590.5
2001	6,960.8	967.2
2002	7,202.1	9,409.6
2003	7,255.6	9,644.9
2004	7,387.3	9,884.4
2005	7,455.9	10,031.6
2006	7,574.8	10,184.1
2007	7,815.9	10,432.2
2008	8,188.9	10,913.9

Here is the plot:



(i) Compute the mean college enrollment over the period 2000–2008 for males, females, and for both sexes together.

(ii) Compute the variance over the period in the enrollment of males, females, and both sexes.

(iii) Compute the covariance of the two series, male and female enrollment, and then compute the correlation.

(iv) Consider two series, the first being male enrollment from 2001 to 2008 (8 observations), the second being male enrollment from 2000 to 2007 (also 8 observations). Consider the joint distribution of these two series: for instance, the first pair would be (6960.8, 6721.8). Compute the correlation of the two series.

(v) Repeat the exercise for female enrollment.

(vi) Interpret the results of (iv) and (v).

(i) The mean for males is 7395.9, for females 8895.38, and for both sexes together 8145.639, which, not surprisingly, is the mean of the male and female means.

(ii) The variances are: for men 192,736.69, for women 9,267,673.345, and for both 5,720,082.1.

(iii) The covariance is 748,328.45, and the correlation is 0.560.

(iv and v) The correlation between male enrollment and its own lag is 0.9999; that for female enrollment is 0.906.

(vi) It can be seen that the graph of the time series of male enrollment is almost a straight line. If it were an exactly straight line, the correlation between the series and its lag would be exactly 1. The actual value of the correlation is very close to 1, but not quite, ans the graph is not exactly a straight line.

Female enrollment, on the other hand, underwent a huge dip in 2001 – I have no idea why – and this interrupted what would otherwise have been nearly a straight line. The dip accounts for a correlation that is much lower than that for men.