DS-GA 1002 - Homework 1

Eric Niblock

September 8th, 2020

1. (True or False) Prove the following statements or provide a counterexample. Let A, B, and C be events in a probability space.

(a) If A and B are independent, then so are A^C and B.

True. We know that it is the case that,

$$B = B \cap (A \cup A^C) \tag{1}$$

$$B = (B \cap A) \cup (B \cap A^C) \tag{2}$$

Then we can apply the probability function to both sides, noting that $(B \cap A)$ and $(B \cap A^C)$ are disjoint sets. Thus,

$$P(B \cap A^C) = P(B) - P(B \cap A) \tag{3}$$

Using the property of independence on A and B, we know that $P(A \cap B) = P(A)P(B)$, so,

$$P(B \cap A^C) = P(B) - P(B)P(A) \tag{4}$$

$$P(B \cap A^C) = P(B)(1 - P(A)) \tag{5}$$

$$P(B \cap A^C) = P(B)P(A^C) \tag{6}$$

The final line shows that the original statement is true; if A and B are independent, then so are A^C and B, since $P(B \cap A^C) = P(B)P(A^C)$ implies the independence of B and A^C.

(b) If A and B are conditionally independent given C, then they are also conditionally independent given C^C.

False. Let $A, B, C \in \Omega$, with each being defined below,

$$\begin{aligned} A &= \{1, 2, 3, 10\} \\ B &= \{3, 4, 5, 10, 11\} \\ C &= \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \\ \Omega &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \end{aligned} \tag{7}$$

We also define our probability measure so that,

$$P(S) = \frac{n(S)}{n(\Omega)} \tag{8}$$

Where the function n counts the number of elements in a set. First we find that A, B are conditionally independent given C,

$$P(A \cap B \,|\, C) = \frac{P(A \cap B \cap C)}{P(C)} = \frac{n(A \cap B \cap C)}{n(C)} = \frac{1}{9} \tag{9}$$

$$P(A \,|\, C) = \frac{P(A \cap C)}{P(C)} = \frac{n(A \cap C)}{n(C)} = \frac{1}{3} \tag{10}$$

$$P(B \,|\, C) = \frac{P(B \cap C)}{P(C)} = \frac{n(B \cap C)}{n(C)} = \frac{1}{3} \tag{11}$$

So,

$$P(A \cap B \,|\, C) = P(A \,|\, C)P(B \,|\, C) = \frac{1}{9} \tag{12}$$

Now, we show that A, B are not conditionally independent given C^C,

$$P(A \cap B \,|\, C^C) = \frac{P(A \cap B \cap C^C)}{P(C^C)} = \frac{n(A \cap B \cap C^C)}{n(C^C)} = \frac{1}{3} \tag{13}$$

$$P(A \,|\, C^C) = \frac{P(A \cap C^C)}{P(C^C)} = \frac{n(A \cap C^C)}{n(C^C)} = \frac{1}{3} \tag{14}$$

$$P(B \,|\, C^C) = \frac{P(B \cap C^C)}{P(C^C)} = \frac{n(B \cap C^C)}{n(C^C)} = \frac{2}{3} \tag{15}$$

So,

$$P(A \cap B \,|\, C^C) \neq P(A \,|\, C^C)P(B \,|\, C^C) \tag{16}$$

Therefore, if A and B are conditionally independent given C, they are not necessarily conditionally independent given C^C.

(c) Events in a partition cannot be independent (assume that every event in the partition has nonzero probability).

True. Proof by contradiction. Take any two events in a partition, S_i and S_j, and assume that they are independent with some non-zero probability. Then, it follows from the definition of partition that,

$$P(S_i \cap S_j) = 0 \tag{17}$$

However, assuming independence, we see that,

$$P(S_i \cap S_j) = P(S_i)P(S_j) = 0 \tag{18}$$

This implies that either $P(S_i) = 0$ or $P(S_j) = 0$. Contradiction. Events in a partition cannot be independent.

(d) If $P(A|B) = 1$ then $P(B^C|A^C) = 1$.

True. We note that,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 1 \tag{19}$$

$$P(A \cap B) = P(B) \tag{20}$$

$$A \cap B = B \tag{21}$$

This implies that,

$$B \subseteq A \tag{22}$$

Using this, we can now tackle the fact that $P(B^C|A^C) = 1$, which similarly implies

$$P(B^C|A^C) = \frac{P(B^C \cap A^C)}{P(A^C)} = 1 \tag{23}$$

By application of De Morgan's law,

$$\frac{P(B^C \cap A^C)}{P(A^C)} = \frac{P((B \cup A)^C)}{P(A^C)} = 1 \tag{24}$$

And because of the fact that $B \subset A$, we note that,

$$B \cup A = A \tag{25}$$

$$\frac{P((B \cup A)^C)}{P(A^C)} = \frac{P(A^C)}{P(A^C)} = 1 \tag{26}$$

So, it is true that if $P(A|B) = 1$ then $P(B^C|A^C) = 1$.

(e) $P(B|A \cup B) \geq P(B|A)$**.**

True. The inequality $P(B|A \cup B) \geq P(B|A)$ can be expressed as,

$$\frac{P(B \cap (A \cup B)}{P(A \cup B)} \geq \frac{P(B \cap A)}{P(A)} \tag{27}$$

$$\frac{P((B \cap A) \cup B)}{P(A \cup B)} \geq \frac{P(B \cap A)}{P(A)} \tag{28}$$

$$\frac{P(B)}{P(A \cup B)} \geq \frac{P(B \cap A)}{P(A)} \tag{29}$$

$$P(A)P(B) \geq P(A \cap B)P(A \cup B) \tag{30}$$

Now since $P(A) \leq 1$ and $P(B) \leq 1$, the most that the left hand side could be is $P(A)P(B) = 1$. Then we have,

$$1 \geq P(A \cap B)P(A \cup B) \tag{31}$$

But we know this must be true, because $P(A \cap B) \leq 1$ and $P(A \cup B) \leq 1$. So we have,

$$1 \geq 1 \tag{32}$$

Thus, the statement has been shown.

2. **(Probability spaces)**

(a) **Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let A be an event in the σ-algebra $\mathcal{F}$, such that $P(A) \neq 0$, on which we want to condition. We define a collection of events $\mathcal{F}_A$ as the collection of the intersection of A with all the events in F:**

$$\mathcal{F}_A = \{A \cap F : F \in \mathcal{F}\} \tag{33}$$

If we consider a new sample space $\Omega_A = A$, prove that $\mathcal{F}_A$ is a valid σ-algebra, and also that the conditional probability measure

$$P_A(S \cap A) = \frac{P(S \cap A)}{P(A)} \tag{34}$$

In order to prove that prove that $\mathcal{F}_A$ is a valid σ-algebra, we need to satisfy the following conditions as noted in the definition:

> *A σ-algebra $\mathcal{F}$ is a collection of subsets of Ω such that: (1) If a set $S \in \mathcal{F}$ then $S^C \in \mathcal{F}$. (2) If the sets $S_1, S_2 \in \mathcal{F}$, then $S_1 \cup S_2 \in \mathcal{F}$. This also holds for infinite sequences; if $S_1, S_2, ... \in \mathcal{F}$ then $\cup_{i=1}^{\infty} S_i \in \mathcal{F}$. (3) $\Omega \in \mathcal{F}$ [Def. 1]*

We begin with condition (3), which is satisfied by the problem statement. Given that $A \in \mathcal{F}$, we have

$$(A \cap A) = A = \Omega_A \in \mathcal{F}_A \tag{35}$$

Now we prove (1). Take some event $F \in \mathcal{F}$, this implies that $A \cap F$ and its compliment are in $\mathcal{F}_A$,

$$(A \cap F) \in \mathcal{F}_A \implies (A \cap F)^C = \emptyset \cup (\Omega_A - F) = (\Omega_A - F) \in \mathcal{F}_A \tag{36}$$

Choosing the right event from $\mathcal{F}$ will produce this compliment. We take $A^C \cup F^C$, then

$$(A \cap (A^C \cup F^C)) = \emptyset \cup (A \cap (\Omega - F)) = (\Omega_A - F) \in \mathcal{F}_A \tag{37}$$

So, to find the compliment of any event in $\mathcal{F}_A$ which was the result from some $F \in \mathcal{F}$, we use event $A^C \cup F^C \in \mathcal{F}$.

Finally, we show (2). Take events $F_1, F_2 \in \mathcal{F}$. Then,

$$(A \cap F_1),\, (A \cap F_2) \in \mathcal{F}_A \implies (A \cap F_1) \cup (A \cap F_2) \in \mathcal{F}_A \tag{38}$$

$$(A \cap F_1) \cup (A \cap F_2) = A \cap (F_1 \cup F_2) \in \mathcal{F}_A \tag{39}$$

And we know that $(F_1 \cup F_2) \in \mathcal{F}$ because it is a valid σ-algebra. So, to find the union of any two events in $\mathcal{F}_A$ which resulted from some $F_1, F_2 \in \mathcal{F}$, we use event $F_1 \cup F_2 \in \mathcal{F}$.

Having satisfied all three conditions from *Definition 1*, we confirm that $\mathcal{F}_A$ is a valid σ-algebra. We now show that $P_A(S \cap A)$ is a valid probability measure. We must satisfy the conditions given in the definition:

> *A probability measure is a function defined over the sets in a σ-algebra $\mathcal{F}$ such that: (1) $P(S) \geq 0$ for any event $S \in \mathcal{F}$. (2) If the sets $S_1, S_2, ..., S_n \in \mathcal{F}$ are disjoint (i.e. $S_i \cap S_j = \emptyset$ for $i \neq j$) then*
>
> $$P(\cup_{i=1}^{n} S_i) = \sum_{i=1}^{n} P(S_i) \tag{40}$$
>
> *Similarly, for a countably infinite sequence of disjoint sets $S_1, S_2, ... \in \mathcal{F}$*
>
> $$P(\lim_{n\to\infty} \cup_{i=1}^{n} S_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(S_i) \tag{41}$$
>
> *(3) $P(\Omega) = 1$. [Def. 2]*

We begin by proving (3),

$$P_A(A \cap A) = P_A(\Omega_A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1 \tag{42}$$

We have satisfied (3), so now we show (1),

$$P_A(S \cap A) = \frac{P(S \cap A)}{P(A)} \geq 0 \tag{43}$$

But this is evident, because we know $P(A) > 0$ and $P(S \cap A) \geq 0$ since they are results from a valid probability space. So (1) is satisfied. Therefore, we are left with showing (2). Take $F_1, F_2 \in \mathcal{F}$ such than F_1 and F_2 are disjoint. Then, $A \cap F_1$ and $A \cap F_2$ are disjoint. So we must show,

$$P_A(F_1 \cap A) + P_A(F_2 \cap A) = P_A((F_1 \cup F_2) \cap A) \tag{44}$$

So,

$$\frac{P(F_1 \cap A)}{P(A)} + \frac{P(F_2 \cap A)}{P(A)} = \frac{P((F_1 \cup F_2) \cap A)}{P(A)} \tag{45}$$

$$P(F_1 \cap A) + P(F_2 \cap A) = P((F_1 \cup F_2) \cap A) \tag{46}$$

$$P(F_1 \cap A) + P(F_2 \cap A) = P((F_1 \cap A) \cup (F_2 \cap A)) \tag{47}$$

And since $A \cap F_1$ and $A \cap F_2$ are disjoint, we have,

$$P(F_1 \cap A) + P(F_2 \cap A) = P(F_1 \cap A) + P(F_2 \cap A) \tag{48}$$

So, we have shown (2). Therefore, we have shown that $P_A(S \cap A)$ is a valid probability measure.

(b) Suppose we have a sample space $\Omega = \{1, ..., M\}$ with σ-algebra $\mathcal{F} = 2^{\Omega}$, the power set of Ω. To determine P, the probability measure, we employ the following empirical procedure:

i. Collect N data points taking values in Ω (e.g., N rolls of an M-sided die). Call these observations $x_1, ..., x_N$.

ii. For each $S \subseteq \Omega$,

$$P(S) = \frac{\text{number of } i\text{-values such that } x_i \in S}{N} \tag{49}$$

As an example, suppose $M = 2$ and we flip a coin $N = 10$ times getting 6 heads and 4 tails, where 1 denotes head and 2 denotes tail. Then

$$P(\emptyset) = 0, \quad P(\{1\}) = 0.6, \quad P(\{2\}) = 0.4, \quad P(\{1,2\}) = 1$$

If P is defined using the above procedure, will it always result in a valid probability measure? Either prove that it will, or give a counterexample.

In order to show that this is a valid probability measure, we must show that it satisfies the three conditions given in *Definition 2*.

Condition (1) is easily satisfied. We know that the count of i-values in $x_1, ..., x_N$ will always be positive or zero concerning every $S \in \Omega$. Similarly, N will always be positive. So, $P(S)$ will always be positive or zero.

Condition (3) is the second easiest to satisfy, and we wish to show $P(\Omega) = 1$. We have,

$$P(\Omega) = \frac{\text{number of } i\text{-values such that } x_i \in \Omega}{N} \tag{50}$$

But since every $x_i \in \Omega$, the numerator just becomes the number of values in $x_1, ..., x_N$. So then we have,

$$P(\Omega) = \frac{N}{N} = 1 \tag{51}$$

The final condition to satisfy is condition (2). Take any two disjoint sets S_1 and S_2. Then,

$$P(S_1) + P(S_2) = P(S_1 \cup S_2) \tag{52}$$

$$\frac{n(x_i \in S_1)}{N} + \frac{n(x_i \in S_2)}{N} = \frac{n(x_i \in (S_1 \cup S_2))}{N} \tag{53}$$

Where function n simply implies the count of x_i values in the list $x_1, ..., x_N$ such that the condition of the argument is met. Now, since S_1 and S_2 are disjoint,

$x_i \in S_1 \Rightarrow x_i \notin S_2$, and $x_i \in S_2 \Rightarrow x_i \notin S_1$. Because there is no x_i that exists in both sets, the count of x_i from each set must equal the count of x_i from the union of both sets. Therefore (2) is satisfied.

With all of the conditions satisfied, we have shown that empirical procedure is a valid probability measure.

3. (Testing) A company with 10 employees decides to test them for COVID-19 before they go back to work in person. From available data, they determine that the probability of each employee being ill is 0.01. The employees have not been in contact with each other for a while, so the events *Employee i is ill*, for $1 \leq i \leq 10$, are modeled as independent. If an employee is ill, the test is positive with probability 0.98. If they are not ill, the test is positive with probability 0.05.

(a) Is it reasonable to model the events Test i is positive, for $1 \leq i \leq 10$, as independent? From now on model them as independent whether you think it is reasonable or not.

Given the information we have, it is somewhat reasonable to model the events $P_i =$ *Test i is positive*, as independent, which is equivalent to saying $P(P_i|P_j) = P(P_i)$. In words, "knowing employee j is positive has no bearing on whether employee i is positive." This assumption is reasonable, because the employees have not been in contact, and therefore employee j could not have spread the disease to employee i. However, there could be more individuals outside of the company which are mutual companions of i and j that could have given them the disease. Knowing this information would ruin the notion of independence. However, given the information we have currently, it seems fairly reasonable to model events P_i as independent, due to the lack of contact. There is also a possibility that the tests themselves are all part of a faulty batch, though this is also unlikely.

(b) The company tests all employees. What is the probability that there is at least one positive test?

We first define the following events: $P_i =$ *Test i is positive*, $P_i^C =$ *Test i is negative*, $S_i =$ *i is sick*, and $S_i^C =$ *i is healthy*. We also know the following information: $P(S_i) = 0.01$, $P(P_i|S_i) = 0.98$, $P(P_i|S_i^C) = 0.05$. Then, the probability of one individual testing positive is given by,

$$\begin{aligned} P(P_i) &= P(S_i)P(P_i|S_i) + P(S_i^C)P(P_i|S_i^C) \\ &= (0.01)(0.98) + (0.99)(0.05) = 0.0593 \end{aligned} \tag{54}$$

So, we also know the probability of i testing negative,

$$P(P_i^C) = 1 - P(P_i) = 0.9407 \tag{55}$$

We will also define the event $\cup_{i=1}^{10} P_i =$ *At least one i is positive*. It follows then, due to the independence of positive tests, that,

$$P(\cup_{i=1}^{10} P_i) = 1 - P(\cap_{i=1}^{10} P_i^C)$$
$$= 1 - (P(P_1^C)P(P_2^C)...P(P_{10}^C)) \quad (56)$$
$$= 1 - (0.9407^{10}) = 0.4574$$

(c) If there is at least one positive test, what is the probability that nobody is ill? If you make any independence or conditional independence assumptions, please justify them.

The probability of interest is $P(\cap_{i=1}^{10} S_i^C \mid \cup_{i=1}^{10} P_i)$. We can employ Bayes Theorem to find an equality that is easier to work with,

$$P(\cap_{i=1}^{10} S_i^C \mid \cup_{i=1}^{10} P_i) = \frac{P(\cap_{i=1}^{10} S_i^C) P(\cup_{i=1}^{10} P_i \mid \cap_{i=1}^{10} S_i^C)}{P(\cup_{i=1}^{10} P_i)} \quad (57)$$

We examine the most difficult of the term within the expression, which is,

$$P(\cup_{i=1}^{10} P_i \mid \cap_{i=1}^{10} S_i^C) = 1 - P(\cap_{i=1}^{10} P_i^C \mid \cap_{i=1}^{10} S_i^C) \quad (58)$$

Now, we assume conditional independence between $\{P_1^C, ..., P_{10}^C\}$ given that no one is ill. In other words, person i testing negative has no bearing on person j testing negative given that everyone is healthy. Then,

$$P(\cup_{i=1}^{10} P_i \mid \cap_{i=1}^{10} S_i^C) = 1 - (P(P_1^C \mid \cap_{i=1}^{10} S_i^C)...P(P_{10}^C \mid \cap_{i=1}^{10} S_i^C)) \quad (59)$$

We also make the assumption that $P(P_i^C \mid \cap_{i=1}^{10} S_i^C) = P(P_i^C \mid S_i^C)$. Then,

$$P(\cup_{i=1}^{10} P_i \mid \cap_{i=1}^{10} S_i^C) = 1 - (P(P_i^C \mid S_i^C))^{10} = 1 - (0.95^{10}) = 0.4013 \quad (60)$$

Furthermore, from the independence of events concerning employees being ill, we have,

$$P(\cap_{i=1}^{10} S_i^C) = (P(S_i^C))^{10} = (0.99)^{10} = 0.9044 \tag{61}$$

So then, our final answer becomes,

$$P(\cap_{i=1}^{10} S_i^C \mid \cup_{i=1}^{10} P_i) = 0.7934 \tag{62}$$

This result is the probability that everyone is healthy given that one person tested positive.

4. **(Student Performance) A group of researchers are interested in predicting students performance based on different factors. To achieve this goal, they collect some real world data. The file *student.csv* contains the data of 395 students' performance in a math class and the file *student.txt* contains a description of the labels. Please show your work when answering the following questions.**

   (a) **Based on the data, are Internet access, taking an extra paid class and school support independent of having good grades (that is, having a final grade above 11)?**

   Observe the information obtained through analysis with Python below.

```
import pandas as pd
f = pd.read_csv( r'C:\Users\Eric\Downloads\student.csv')
```

```
print('Total # of Students: ', len(f))
print('I: # Students with Internet: ',len(f[f['internet']=='yes']))
print('G: # Students with Good Grades: ',len(f[f['G3'] > 11]))
print('P: # Students with Extra Paid Class: ',len(f[f['paid']=='yes']))
print('S: # Students with Extra School Support: ',len(f[f['schoolsup']=='yes']))
print('# Students with Internet and Good Grades: ',len(f[(f['G3'] > 11) & (f['internet']=='yes')]))
print('# Students with Internet and Extra Paid Class: ',len(f[(f['G3'] > 11) & (f['paid']=='yes')]))
print('# Students with Internet and Extra School Support: ',len(f[(f['G3'] > 11) & (f['schoolsup']=='yes')]))
```

```
Total # of Students:  395
I: # Students with Internet:  329
G: # Students with Good Grades:  162
P: # Students with Extra Paid Class:  181
S: # Students with Extra School Support:  51
# Students with Internet and Good Grades:  140
# Students with Internet and Extra Paid Class:  74
# Students with Internet and Extra School Support:  9
```

Given this information, we can test if variables are independent. We define the probability of some event, as the number of times an event occurs divided by the total number of events. In this case, it is always the number of students possessing some characteristic(s), divided by the total number of students. We will say that two events, A, B, are independent if,

$$P(A \cap B) \approx P(A)P(B) \tag{63}$$

Where we have defined approximately equal to mean within 0.02. We can begin by testing if having access to internet and having good grades are independent.

$$P(I \cap G) = \frac{140}{395} = 0.3544 \tag{64}$$

$$P(I) = \frac{329}{395} = 0.8329 \tag{65}$$

$$P(G) = \frac{162}{395} = 0.4101 \tag{66}$$

So,

$$P(I \cap G) = 0.3544 \approx P(I)P(G) = 0.3416 \tag{67}$$

So access to internet and having good grades are approximately independent. Next we test if taking an extra paid class and having good grades are independent.

$$P(P \cap G) = \frac{74}{395} = 0.1873 \tag{68}$$

$$P(P) = \frac{181}{395} = 0.4582 \tag{69}$$

$$P(G) = \frac{162}{395} = 0.4101 \tag{70}$$

So,

$$P(I \cap G) = 0.1873 \approx P(I)P(G) = 0.1879 \tag{71}$$

So taking a paid class and having good grades are approximately independent. Next we test if having extra school support and having good grades are independent.

$$P(S \cap G) = \frac{9}{395} = 0.0228 \tag{72}$$

$$P(S) = \frac{51}{395} = 0.1291 \tag{73}$$

$$P(G) = \frac{162}{395} = 0.4101 \tag{74}$$

So,

$$P(S \cap G) = 0.0228 \neq P(S)P(G) = 0.0530 \tag{75}$$

So having extra school support and having good grades are dependent.

(b) If we know the family size of the student, is having good grades independent of taking an extra paid class?

Observe the information obtained through analysis with Python below.

```
print('Total # Students Family Size > 3: ',len(f[(f['famsize'] == 'GT3')]))
print('# Students Family Size > 3, with Good Grades: ', len(f[(f['famsize'] == 'GT3') & (f['G3'] > 11)]))
print('# Students Family Size > 3, with Extra Paid Class: ', len(f[(f['famsize'] == 'GT3') & (f['paid']=='yes')]))
print('# Students Family Size > 3, with Good Grades & Extra Paid Class: ', \
      len(f[(f['famsize'] == 'GT3') & (f['paid']=='yes') & (f['G3'] > 11)]))
```

```
Total # Students Family Size > 3:  281
# Students Family Size > 3, with Good Grades:  109
# Students Family Size > 3, with Extra Paid Class:  130
# Students Family Size > 3, with Good Grades & Extra Paid Class:  48
```

```
print('Total # Students Family Size <= 3: ',len(f[(f['famsize'] == 'LE3')]))
print('# Students Family Size <= 3, with Good Grades: ', len(f[(f['famsize'] == 'LE3') & (f['G3'] > 11)]))
print('# Students Family Size <= 3, with Extra Paid Class: ', len(f[(f['famsize'] == 'LE3') & (f['paid']=='yes')]))
print('# Students Family Size <= 3, with Good Grades & Extra Paid Class: ', \
      len(f[(f['famsize'] == 'LE3') & (f['paid']=='yes') & (f['G3'] > 11)]))
```

```
Total # Students Family Size <= 3:  114
# Students Family Size <= 3, with Good Grades:  53
# Students Family Size <= 3, with Extra Paid Class:  51
# Students Family Size <= 3, with Good Grades & Extra Paid Class:  26
```

Here, we will need to test for conditional independence on both types of family size (this is because of our previous result: If A and B are conditionally independent given C, then they are not necessarily also conditionally independent given C^C). Conditional independence is given by,

$$P(A \cap B \,|\, C) \approx P(A \,|\, C) P(B \,|\, C) \tag{76}$$

Where we have defined approximately equal to mean within 0.02. Given a family size greater than 3 (noted at F^+), we define probability as before, and we calculate the following values in order to determine if taking an extra paid class is conditionally independent of having good grades:

$$P(P \cap G \,|\, F^+) = \frac{48}{281} = 0.1708 \tag{77}$$

$$P(P \,|\, F^+) = \frac{130}{281} = 0.4626 \tag{78}$$

$$P(G \,|\, F^+) = \frac{109}{281} = 0.3879 \tag{79}$$

So,

$$P(P \cap G \,|\, F^{+}) = 0.1708 \approx P(P \,|\, F^{+})P(G \,|\, F^{+}) = 0.1795 \tag{80}$$

So having good grades and taking an extra paid class are approximately conditionally independent given a family size greater than 3. Now we consider their conditional independence given a family size less than or equal to 3.

$$P(P \cap G \,|\, F^{-}) = \frac{26}{114} = 0.2281 \tag{81}$$

$$P(P \,|\, F^{-}) = \frac{51}{114} = 0.4474 \tag{82}$$

$$P(G \,|\, F^{-}) = \frac{53}{114} = 0.4649 \tag{83}$$

$$P(P \cap G \,|\, F^{-}) = 0.2281 \neq P(P \,|\, F^{-})P(G \,|\, F^{-}) = 0.2080 \tag{84}$$

So having good grades and taking an extra paid class are conditionally dependent given a family size less than or equal to 3.