Tutorial 3 Suggested Solutions

Some Theoretical and Conditional Probability Problems

Sept. 30

Chapter 2: Some more Theoretical Exercises

TE 19

As my hint mentioned in the outline, let's take a look inside the front cover at the Negative Hypergeometric Distribution.

There are different interpretations, so I'll share mine. Similar to the hypergeometric distribution discussed earlier, we have a product. In this case, a product of probabilities $\rightarrow$ Probability we have $r-1$ red balls in our first $k-1$ draws TIMES the final kth draw, gives us our rth desired red ball.

Have a look at the binomial and negative binomial distributions, and see if there are any similarities. (There are.)

First we find: P(drawing r-1 red balls in k-1 draws):

This can be regarded as a hypergeometric. Recall, there are n red balls, and m blue. If we draw $k-1$ balls, and we state that $r-1$ are red, then $k-1-(r-1)$ must be blue:

$$\frac{\binom{n}{r-1}\binom{m}{k-r}}{\binom{n+m}{k-1}}$$

Second: P(kth draw completes our r red ball count requirement), we have stated that the last ball will be chosen as red. From the n red balls, we subtract $r-1$ from them. From the $n+m$ total balls, we subtract the $k-1$ we removed:

$$\frac{n-r+1}{n+m-k+1}$$

Thus we multiply these together to get our answer:

$$P(k \text{ balls drawn}) = \frac{\binom{n}{r-1}\binom{m}{k-r}}{\binom{n+m}{k-1}} \frac{n-r+1}{n+m-k+1}$$

which is valid for $r \leq k$.

Thus we have demystified the most "complicated" of the distributions presented there. There are other ways of deriving this, such has finding the sample space and counting the outcomes separately, but this is how I found it easiest.

TE 20

Our sample space consists of countably infinite points (see question sheet for some definitions). Let's answer the two questions:

All equal probability?

Proof by contradiction: Assume all points, e_i are equally likely, where i indexes the countably infinite points. That is, $\mathcal{S} = \{e_1,\ e_2,\ \ldots\}$ is some arbitrary space.

Assign each point $P(e_i) = p$. By axioms of probability

$$\sum_{i=1}^{\infty} P(e_i) \stackrel{?}{=} 1$$

However this is a contradiction because if $p > 0$:

$$\sum_{i=1}^{\infty} p = \infty$$

The other case is if $p = 0$, but then

$$\sum_{i=1}^{\infty} 0 = 0$$

Thus by contradiction in all possible cases of equally likely scenarios, we state it is impossible.

What if we drop "Equally Likely" ...

Yes, in fact, is the answer to the second part. If we drop the equally likely condition, and define $P(e_i)$ cleverly, we can make it work. Let's think of an example for this.

Let

$$P(e_i) = \frac{1}{2^i}$$

Then check

$$\sum_{i=1}^{\infty} P(e_i) \stackrel{?}{=} 1$$

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{2^i} &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots \\ &= \frac{1}{2}\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) \\ &= \frac{1}{2}\left(\frac{1}{1-\frac{1}{2}}\right) \\ &= 1\end{aligned}$$

as required.

This is the geometric series, a classical result, and will be seen again later when we study the Geometric Distribution.

Chapter 3: Conditional Probabilities & Independence

Example 2f

Please read chapter 2's *Example 5m.* There is a very elegant proof using: inclusion/exclusion, probability and counting, and the Maclaurin series for the exponential function.

The final solution that there are NO matches among N men is

$$P_N = \sum_{i=0}^{N} \frac{(-1)^i}{i!}$$

Hence, as $n \to \infty, P(NoMatches) = \frac{1}{e}$, and NOT 0, as some might think initially.

This question asks us to use Conditional Probability (the *multipication rule* here specifically) to figure out what is the probability that exactly k of N men get their hat back.

We focus on a specific set of k people, and will calculate the probability that exactly they, and no others, get their own hat.
Let E be the event that this set of k men get their hat.
Let G be the event that the other $N-k$ men did not get theirs.

Let's write out something to build upon:

$$P(EG) = P(E)P(G|E)$$

Let the event $F_i,\ i = [1, k]$ refer to man "i" in that "k" group, getting his hat. Then

$$P(E) = P(F_1F_2\dots F_k)$$

$$= P(F_1)P(F_2|F_1)P(F_3|F_2F_1)\dots P(F_k|F_{k-1}\dots F_2F_1)$$

Intuitively, each of these probabilities should be clear now, if we start with the simplest and work across:

$$= \frac{1}{N}\frac{1}{N-1}\frac{1}{N-2}\cdots\frac{1}{N-k+1}$$

This should look familiar if you have done your homework from chapter 1. Thus

$$P(E) = \frac{(N-k)!}{N!}$$

Let's use what we know about NO one getting their hat to get $P(G|E)$. Since k men have their hat, we can just reduce the size of N to $N-k$ and apply the formula. Thus:

$$P(G|E) = P_{N-k} = \sum_{i=0}^{N-k}\frac{(-1)^i}{i!}$$

Now we can just multiply these and simplify:

$$P(EG) = P(E)P(G|E) = \frac{(N-k)!}{N!}\cdot\left(\sum_{i=0}^{N-k}\frac{(-1)^i}{i!}\right)$$

Recall, what we just solved, is for a *specific* set of k men, however, if there are N total, there are $\binom{N}{k}$ ways to do this:

$$P(\text{exactly k matches}) = \binom{N}{k}\cdot P(EG)$$

$$= \frac{N!}{k!(N-k)!}\frac{(N-k)!}{N!}P_{N-k}$$

$$= \frac{\sum_{i=0}^{N-k}\frac{(-1)^i}{i!}}{k!}$$

As N grows very large, the answer converges to $\frac{1}{e\cdot k!}$

Example 3c

If you know the answer, K, your chances of getting it correct, C, are $P(C|K) = 1$. *Using "C" is an abuse of notation here, as C is currently being used as the symbol for complement, however we will just stick to the book's notation. Complement symbols can also look like "primes" which are also used for derivatives, so it's all about context.* This question is interested in, what is the probability that a student knows the answer GIVEN they answered it correctly. Using Bayes':

$$P(K|C) = \frac{P(KC)}{P(C)} = \frac{P(CK)}{P(C \cap \mathcal{S})} = \frac{P(C|K)P(K)}{P(CK) + P(CK^C)}$$

$$= \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^C)P(K^C)}$$

which is now in terms we understand, and can put in values for to solve. We are told the probability that a student knows the answer $P(K) = p$ (think of this as the proportion of the course material the understand), hence $P(K^C) = 1 - p$ which is traditionally denoted as $q = 1 - p \Rightarrow p + q = 1$. If they have to guess, naturally they have a 1 out of something chance. If there are m choices, then $P(C|K^C) = \frac{1}{m}$. Thus we can plug that all in to get our answer:

$$P(K|C) = \frac{p}{p + \frac{q}{m}}$$

Here's another form of this question if you want to practice last week's material: If the test contains N questions, what's the probability the student will get n correct just by purely guessing?

Example 3e

This question is very wordy: Jones is given some test, that ALWAYS correctly determines if someone has a disease, but if they have diabetes, it can give a false positive 30% of the time.

We want to know the probability that Jones has the disease, GIVEN that the result shows up positive. Let us call these events, D and E, respectively:

$$P(D|E) = \frac{P(ED)}{P(E)}$$

A lot of students make the mistake here by assuming that we know what $P(ED)$ is. It may seem counter-intuitive, but we need to rearrange the above formula – which we will do using Bayes' rules. All we know is the physician's belief P(D), and the probability of having the disease GIVEN some information from the test.

$$\frac{P(ED)}{P(E \cap \mathcal{S})} = \frac{P(E|D)P(D)}{P(E \cap D) + P(E \cap D^C)}$$

$$= \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^C)P(D^C)}$$

Now the problem is in terms that we know given the question. According to the physician's professional opinion, $P(D) = 0.6$ thus $P(D^C) = 0.4$. There is a certainty that if someone is sick, the test will pick it up $\rightarrow P(E|D) = 1$. However, if they have diabetes a false positive occurs according to $P(E|D^C) = 0.3$. Hence we can plug in our numbers and solve for an answer of 0.833. Thus the physician should recommend the surgery as this is $> 80\%$ (their confidence threshold).

TE 2

Let's use the definition of conditional probability (presented in this text). Let B be a proper subset ($B \subset \mathcal{S}$) (stated since we can't divide by 0), and let $A \subset B$,

$P(A|B)$

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

and since

$$A \subset B \Rightarrow P(AB) = P(A)$$

which leads to the answer immediately

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

$P(A|B^C)$

Recall, $B^C := \mathcal{S} \setminus B \Rightarrow B \cap B^C = \emptyset$ since any set intersected with a disjoint set, in this case, its complement, is empty. Hence,

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

Since $A \subset B \Rightarrow A \cap B^C = \emptyset$ also. Which leads us to our final answer:

$$P(A|B^C) = \frac{P(\emptyset)}{P(B^C)} = \frac{0}{P(B^C)} = 0$$

$P(B|A)$

This should be intuitively clear. If A has occurred, then B must have also occurred, since A is in B:

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

$P(B|A^C)$

Firstly, $B \cap A^C = B \setminus A$, think of it as cutting out A from B, which doesn't have any special way to simplify any more in the general case. Thus we can only get as far as:

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

which is fine. *Simplest form* is a relative term of course ...

TE 3

n_i families, means there are n_1 families with 1 child, n_4 families with 4 children, etc. Thus there are $\sum_{i=1}^{k} n_i = m$ total families. In theory, $k \to \infty$, however for large i, $n_i = 0$ (let's keep sociology out of this), and the model is appropriate.

It should follow that there are then $\sum_{i=1}^{k} i \cdot n_i$ total kids in the school.

There are two methods the book wants us to compare:

1. Choose 1 of the m families at random, then choose a random child from that family.
2. Choose 1 child from the whole school at random.

We want to show that method 1 is more likely in selecting a *firstborn child*.

Now that we've digested that wordy question, let's show the condition. The book suggests we try to show:

$$\sum_{i=1}^{k} in_i \sum_{j=1}^{k} \frac{n_j}{j} \geq \sum_{i=1}^{k} n_i \sum_{j=1}^{k} n_j$$

Notice, we these sums are split up into their respective indices, but can also be written as:

$$\sum_{i=1}^{k}\sum_{j=1}^{k} in_i \frac{n_j}{j} \geq \sum_{i=1}^{k}\sum_{j=1}^{k} n_i n_j$$

$$\sum_{i=1}^{k}\sum_{j=1}^{k} \frac{i}{j} n_i n_j \geq \sum_{i=1}^{k}\sum_{j=1}^{k} n_i n_j$$

Let's start from the beginning and see how to get back to this point using conditional probability (since there is a reason this question is found in chapter 3).
Let's define A as the event that a firstborn is chosen.
Let F_i be the event that we choose a family with i children.
Let's express method 1 using Bayes' theorem:

$$P(A) = \sum_{i=1}^{k} P(A|F_i)P(F_i) = \sum_{i=1}^{k} \left(\frac{1}{i}\right)\left(\frac{n_i}{m}\right) = \frac{1}{m}\sum_{i=1}^{k}\frac{n_i}{i}$$

The second method just has us pick a random child:

$$P(A) = \frac{m}{\sum_i in_i}$$

We need to show:

$$\frac{1}{m}\sum_{i=1}^{k}\frac{n_i}{i} \geq \frac{m}{\sum_i in_i}$$

Rearranging terms,

$$\sum_i \frac{n_i}{i} \sum_i in_i \geq m^2$$

We temporarily complicate our expression to show it in terms of i and j as is required for us:

$$\sum_j \frac{n_j}{j} \sum_i in_i \geq \sum_j n_j \sum_i n_i$$

$$\sum_{i=1}^{k}\sum_{j=1}^{k} \frac{i}{j} n_i n_j \geq \sum_{i=1}^{k}\sum_{j=1}^{k} n_i n_j$$

We are now back to where the hint told us we should be, so we're on the right track, and should be almost done. The only thing that remains is to show that our coefficient on the $n_i n_j$ term on the left makes it larger or equal.

Of course, if

$$i = j \Rightarrow \frac{i}{i} n_i n_i = n_i^2$$

, but he double sum will also leave terms in the form of

$$\frac{a}{b} n_a n_b + \frac{b}{a} n_b n_a$$

which we can factor into

$$= \left(\frac{a}{b} + \frac{b}{a}\right) n_a n_b$$

Hence we must show that

$$\left(\frac{i}{j} + \frac{j}{i}\right) \geq 1 \ , \forall i, j \in \mathbb{N}$$

But we can do better by showing,

$$\left(\frac{i}{j} + \frac{j}{i}\right) \geq 2 \ , \forall i, j \in \mathbb{N}$$

(Think about why this is true, it should be immediate, otherwise try and plug in some numbers)
If we multiply out our denominator, we get:

$$i^2 + j^2 \geq 2ij$$

$$i^2 + j^2 - 2ij \geq 0$$

$$(i - j)^2 \geq 0$$

Which holds true, thus we have completed our proof.

■