ESA Statistics Tutorial

Session 3, 6 October 2021

Tiffany Yong

1 Basic Probability

1.1 Set Theory

Definition 1.1. (Set) A **set** is a collection of items, with a rule on how to pick them.

Example 1.1. A set of integers from 1 to 5 would be $\{1, 2, 3, 4, 5\}$. We can assign this set to any letter or symbol we like, as it is cumbersome to rewrite all elements of a set. Let's name our set of integers $S = \{1, 2, 3, 4, 5\}$. But what if the set is extremely large, that you can't easily write out every element? Then you would need to define a set in a different way, which gives rise to the following set notation:

$$\{x \text{ (insert rule here) : (insert rule here)}\}$$

For example, if you want to make a set A comprised of every even number strictly above 100, you would write it as such: Define E as all even numbers. Then $\{x \in E : x > 100\}$.

Definition 1.2. (Element) An **element** of a set is one of the items in that set.

Sometimes, we want to state whether or not something is in a set. From here, we can see that $102 \in$ (in) A, but $1 \notin$ (not in) A. These are the symbols we use to make writing easier.

Definition 1.3. (Subsets) Given sets A, B, we say that A is a **subset** of B if every single element in A is also in B. This is denoted as $A \subseteq B$.

Example 1.2. An example would be that all even numbers greater than 100 is a subset of all even numbers, because if you are an even number greater than 100, you are by definition an even number.

Definition 1.4. (Union) Given sets A, B, we say that x belongs to the **union** of A and B if x is in either A or B. This is written as $x \in A \cup B$ if $x \in A$ **or** $x \in B$.

To belong in a set union, you can belong in either category, or both categories. For example, let A be the set of all people who take economic statistics. Let B be the set of all people who take microeconomics. If you are in $A \cup B$, this means that you be only taking one of the classes, or you can be taking both.

Definition 1.5. (Intersection) Given sets A, B, we say that x belongs to the **intersection** of A and B if x is in both A and B. This is written as $x \in A \cap B$ if $x \in A$ **and** $x \in B$.

To belong in a set intersection, you must belong in both categories. Taking the example from above, let A be the set of all people who take economic statistics, and B be the set of all people who take microeconomics. If you are in $A \cap B$, this means you must be in both classes. If you are in one or none of them, you are not in $A \cap B$.

1.2 Notation

Since mathematicians hate writing out actual words, there are some important bits of notation that you will probably need to know for your classes.
First, let's say you want to take the intersection of k sets. But then, lets say $k = 10,000$. You wouldn't really want to write down $A_1 \cap A_2 \cap \cdots \cap A_{10,000}$ over and over again, so the following notation was created.

$$\bigcap_{i=1}^{k} A_i = A_1 \cap A_2 \cap \cdots \cap A_k$$
$$\bigcup_{i=1}^{k} A_i = A_1 \cup A_2 \cup \cdots \cup A_k$$

In mathematics, you often talk about something being true "for all" elements in the set. For example, "for all" numbers greater than 1, they are greater than 0. Then, in order to shorten the "for all" notation, we use the $\forall$ symbol instead. Thus, the statement above becomes $\forall x > 1, x > 0$.
Another thing we like to talk about is "there exists", to talk about there being an element in a set that fulfils a specific condition. For example, "there exists" a number that is greater than 1. We use the $\exists$ symbol, and the statement above becomes $\exists x$ such that $x > 1$.
Next, we have implication arrows. $P \Rightarrow Q$ means that if P is true, then Q must also be true. Notice that this does not include the reverse implication.

Example 1.3. Let P be the event that I broke my leg. Let Q be the event that I can't play a soccer game. Then, P implies Q, but not the other way round. Breaking a leg will most definitely make me incapable of playing soccer, but I can miss a soccer match for many reasons other than me breaking my leg.

This gives rise to a need for a stronger implication, which results in the "if and only if" condition. $P \iff Q$ means that if P is true then Q must be true, but also if Q is true then P must be true. You can see that this means that both of them must always be either true or false together if they are linked with this relation.

Example 1.4. Let's say that the only way I can quit my soccer club is by filling in a form, and as soon as I fill in the form, I'm out of the club. Let P be the event that I quit my soccer club, and Q be the event that I filled in the form. I had to fill in the form to quit the club, so if P is true, Q is true, and we have that $P \Rightarrow Q$. But if I filled in the form, then I left the soccer club, so if Q is true, then P is true, and we have that $Q \Rightarrow P$. The combination of both of these implications gives me $P \iff Q$.

1.3 Probabilities

Definition 1.6. (Space) A **space** is the collection of all possible elements that sets can be defined, in a particular context.

Example 1.5. The space of a coin flip is $\{H, T\}$, as a coin flip can take only these two values. The space of a roll of a dice is $\{1, 2, 3, 4, 5, 6\}$, since it has six sides. Each element in the space represents all the possible states that the system can take.

Definition 1.7. (Complement) Given a set A in the space ($A \in \Omega$), then the **complement** of A is the set of all the elements in the space that are not in the set A. We denote this complement as $\overline{A}$.

Example 1.6. With the dice roll example above, let's say that $A = \{1, 3, 5\}$. Then, $\overline{A} = \{2, 4, 6\}$, or whatever is left in the space that was not in A.

Definition 1.8. (Properties of Unions and Intersections)

1. (Commutativity) $A \cap B = B \cap A$.
2. (Associativity) $A \cap B \cap C = (A \cap B) \cap C$.
3. (Distributivity) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
4. (De Morgan's Laws) $\overline{(A \cup B \cup C)} = \overline{A} \cap \overline{B} \cap \overline{C}$, and $\overline{(A \cap B \cap C)} = \overline{A} \cup \overline{B} \cup \overline{C}$.

Example 1.7. (Proof of De Morgan's Law) I will write a proof in plain English, set theoretic notation, and then have a Venn Diagram proof on the board.
The word version of $\overline{(A \cup B \cup C)}$ is "not in either A, B or C". This is the same as saying "not in A and not in B and not in C", which is $\overline{A} \cap \overline{B} \cap \overline{C}$.
The word version $\overline{(A \cap B \cap C)}$ is "not in all three of A, B and C". This is the same as saying "not in A, or not in B, or not in C".

Proof. The set theoretic proof of one of the laws is as follows:

$$\begin{aligned}
\overline{(A \cup B \cup C)} &= \{x : x \notin (A \text{ or } B \text{ or } C)\} \\
&= \{x : (x \notin A) \text{ and } (x \notin B) \text{ and } (x \notin C)\} \\
&= \{x : x \in \overline{A} \text{ and } x \in \overline{B} \text{ and } x \in \overline{C}\} \\
&= \{x : x \in \overline{A} \cap \overline{B} \cap \overline{C}\}
\end{aligned}$$

Try to prove the other law by yourself. □

Definition 1.9. (Random experiment) A **random experiment** is an experiment for which you don't know the results beforehand.

Why do we care about random experiments? Well, the notion of probability is based on the idea of randomness. If we knew the outcome of an experiment before even doing it, we can intuitively say that there is no "randomness" involved. If you're even in the slightest uncertain about something, you can assign a probability to it. I can assume that the sun will rise tomorrow, because it always has. But there is still a vanishingly slim probability that it might not, who knows? Not me, because I can't see into the future.

Definition 1.10. (Sample space) A **sample space** is the collection of all possible outcomes of a random experiment, where the space is defined in the context of the experiment. This sample space is denoted as Ω in probability.

Definition 1.11. (Event) An **event** is a subset of the sample space.

Definition 1.12. (Mutually exclusive) Two events A and B are **mutually exclusive (or disjoint)** if the intersection of these subsets is the empty set. This is written as $A \cap B = \emptyset$.

Example 1.8. A very simple example of two events that are mutually exclusive are events which are the complement of each other. Event A is that it is raining in Montreal, and event $\overline{A}$ is that it is not raining in Montreal. We can intuitively see that there is no way it can be raining and not raining in Montreal at the same time, so these are mutually exclusive events.

Definition 1.13. (Exhaustive) A set of events is **exhaustive** if taking the union across all of these events gives you the entire sample space. This means that you can reach every single outcome with this set of events. This is written as $\bigcup_{n=1}^{k} A_n = \Omega$.

Example 1.9. Let the random experiment be a dice roll, such that $\Omega = \{1, 2, 3, 4, 5, 6\}$. Then, let O be the event that the outcome of the dice roll is odd, and E be the event that the outcome of the dice roll is even. We can see that $O = \{1, 3, 5\}$ and $E = \{2, 4, 6\}$. The set of events $\{O, E\}$ is exhaustive, as it covers every single element of Ω.

Definition 1.14. (Algebra) The **algebra** is the set of all subsets of the sample space. As we have defined above, a subset of a sample space is known as an event. Thus, an algebra is the set of all events in the sample space. Here are the properties of an algebra:

1. $\Omega \in \mathcal{F}$
2. If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
3. If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$

In finite cases, algebras are often taken to be 2^{Ω}, which is the set of all the possible things that can happen (we will be able to see this in the future after combinatorics is introduced). The intuition is that you can make a combination of events by choosing at each stage to include or exclude an outcome.

This definition is only here for completion purposes, you can do well in this course without knowing what an algebra is at all.

Example 1.10. The algebra of a coin flip is the set of subsets of the sample space, or the set of events. For a coin flip, the sample space is $\Omega = \{H, T\}$. Thus, we can see that an algebra is $\mathcal{F} = \{\{H, T\}, \{H\}, \{T\}, \{\emptyset\}\}$. Some events can never happen, just like you can't flip one coin and get both heads and tails. But it is there for completion purposes, to fulfil the criteria above.

If you like, you can try proving that $\mathcal{F}$ is an algebra.

Definition 1.15. (Probability function) A **probability function** $\mathbb{P}(.)$ is a function defined on an algebra, $\mathcal{F}$, with range $[0, 1]$, with the following properties:

1. $\mathbb{P}(A) \geq 0, \forall A \in \mathcal{F}$
2. $\mathbb{P}(\Omega) = 1$
3. Let $A_1, A_2, \ldots A_l$ be a set of mutually exclusive events in $\mathcal{F}$, and define $B = A_1 \cup A_2 \cup \cdots \cup A_k$. Then, we have that $\mathbb{P}(B) = \sum_{n=1}^{k} P(A_n)$.

These rules are not arbitrary. They follow the idea of probability that makes intuitive sense to us. For property (1), if someone told you that there was a probability of -1 that an event would happen, this statement would make no sense to you under the framework of probability. For property (2), if someone told you that there was a probability of 2 that something would happen, that would also make no sense. We want all possible outcomes to correspond with 100% probability, also known as $\mathbb{P}(\Omega) = 1$. For property (3), it seems difficult at first glance, but all it's saying is that if you add up probabilities of disjoint events, it adds up to the probability of the union of all those events happening. If this confuses you, I can draw a Venn Diagram proof.

Example 1.11. With the following properties of probability functions that make intuitive sense, we can work through simple examples.

1. What is the probability that one coin flip returns heads? $\mathbb{P}(H) = \frac{1}{2}$.
2. What is the probability that a dice rolls the number 1?
   Define A as the event that the dice rolls a 1. $\mathbb{P}(A) = \mathbb{P}(\{1\}) = \frac{1}{6}$.

3. What is the probability that a dice rolls not the number 1?
   $\mathbb{P}(\overline{A}) = \mathbb{P}(\Omega) - \mathbb{P}(A) = 1 - \frac{1}{6} = \frac{5}{6}$.

4. What is the probability that a dice rolls a number? $\mathbb{P}(\Omega) = 1$.

5. What is the probability that a dice rolls no number? $\mathbb{P}(\emptyset) = 0$.

6. What is the probability that a dice rolls either 1 or 2?
   Define B as the probability that the dice rolls a 2. Since a dice can't roll 1 and 2 at the same time, A and B are mutually exclusive.
   $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

7. What is the probability that a dice rolls either 1 or an odd number? Define C as the probability that the dice rolls an odd number. Since there is some overlap between the dice rolling 1 and an odd number, we use the following formula:
   $\mathbb{P}(A \cup C) = \mathbb{P}(A) + \mathbb{P}(C) - \mathbb{P}(A \cap C) = \frac{1}{2} + \frac{1}{6} - \frac{1}{6}$.

   The intuition of this formula is as follows. You add the probabilities of events that belong to A, and then add the probabilities of events that belong to C. But what if an event belongs to both A and C? Through this process, it was double counted. So we subtract the probability of all the events that were double counted, which is $\mathbb{P}(A \cap B)$.

2 Conditional Probability

Definition 2.1. (Conditional probability) The **conditional probability** of A given that B holds is written as $\mathbb{P}(A|B)$, and follows this formula:

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

The intuition for this formula is as follows. Since we already know that B holds, we move into the space that B holds. There, we take the probability of the intersection of B with A, and divide it by the probability of B, to see how likely it is to happen under the new sample space B. The Venn Diagram proof will be on the board. From this, you can also see that the following holds:

1. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B|A)$.

2. $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$.

Definition 2.2. (Bayes' Theorem) From the definition of conditional probability above, we can see **Bayes' theorem**, which is as follows:

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

Definition 2.3. (Independence) Two events A and B are **independent** if one event having occurred does not affect the probability of the other event occurring. This is written as $\mathbb{P}(A|B) = \mathbb{P}(A)$, or equivalently, $\mathbb{P}(B|A) = \mathbb{P}(B)$.

An easy way to check this identity is borne of the conditional probabilities we have proven above.

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

We sometimes confuse independence with being mutually exclusive, but mutually exclusive events are very rarely independent. Let there be two events, C and D that are mutually exclusive. We know that if C happens, $\mathbb{P}(D) = 0$. Unless $\mathbb{P}(D) = 0$ to begin with, this proves that the probability of D occurring has changed because of C.

Example 2.1. Let A be the event that I win the lottery, and let B be the event that I tripped and fell today. We can see no obvious causal link between the both. If I trip and fall, I'm not more or less likely to win the lottery. If I win the lottery, I'm not more or less likely to trip and fall. Thus, A and B are independent.

3 Combinatorics

Definition 3.1. (Factorial) The factorial operator is such that $n!$ equals the product sum $n \cdot n - 1 \cdot \dots \cdot 1$, and where 0! is defined to be 1.

The factorial is a way to represent an ordering of a set of items with size n. A few examples of when we might want to order things are:

1. What's the number of ways in which you can line a given group of people up?
2. What is the number of ways that you can draw all the balls out of a bag?

Example 3.1. I will give an example. Let's say we have 3 balls in a bag. We want to take out the balls one by one, and take note of the order in which we take them out. How many unique orderings can we get?

Well, for the first ball, we have 3 balls that we can possibly take out, so we have 3 possibilities for the first ball. Then, now that we have removed the first ball, we only have 2 balls left in the bag. So, there are only 2 options left to put in the second slot for the ordering. Finally, we only have 1 ball left in the bag, so we have to pick that last ball. The number of ways we can write the order is thus $3(2) = 6$.

Definition 3.2. (Permutations) The number of possible **permutations** of k items chosen from a set of n items is the number of distinct sets of k that can be assembled when order is considered. We write this as ${}^{n}P_{k}$, and its formula is given by:

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

Examples of situations where order might matter when you're choosing things is:

1. What's the number of ways you can pick the first place, second place and third place winner of the race? Well, if you put person A in first place, it's a different result from putting person A in third place.
2. What's the number of words you can pick out from a set of letters? The word "tea" and the word "eat" contain the exact same letters, but they mean very different things, so the order matters.
3. If you give someone your phone number, you won't tell them the number in any random order right? It has to be that specific number in a specific sequence. So if you choose a sequence of numbers, it doesn't only matter which numbers you choose, but also the order in which you choose them.

Example 3.2. The idea of a permutation is the same as a factorial, but you stop taking combinations after a certain point. Let's say we have 10 balls in a bag, and we want to find unique orderings of 3 balls in the bag. How many unique orderings can we get?

For the first ball, we have 10 balls that we can possibly take out, so we have 3 possibilities for the first ball. Then, now that we have removed the first ball, we have 9 balls left in the bag. So, there are 9 options left to pick as the second ball. We remove one of these 9 options, leaving us 8 balls in the bag. There are now 8 balls left in the bag, and we pick one of them. Thus, the number of ways we can write the order is 10(9)(8), or $\frac{10!}{7!}$.

Definition 3.3. (Combinations) The number of possible **combinations** of k items chosen from a set of n items is the number of distinct sets of k that can be assembled when order is not considered. We write this as ${}^{n}C_k$ or $\binom{n}{k}$, and its formula is given by:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

A few examples where we use combinations because order does not matter are:

1. If you are drawing 10 winners for a grand prize, it does not matter which order in which you draw these winners, as long as the prize for all of them is the same. So the number of possible winners is a combination, not a permutation

2. If you are trying to decide which of your 10 friends go on a retreat with you, it doesn't matter in which order you choose them, because they all go on the same retreat with you, whether they're picked first or last.

Example 3.3. The idea of a combination is that you take the number of permutations for k items chosen from a set of n items, and then you remove the part where order matters. Since we have proven that $k!$ gives you the unique ordering of a given set of items, then what we do is we take the permutations formula, and divide by $k!$ to remove repeat cases. Like the example above, if you have 10 balls in a bag, and you want to remove 3 balls. It does not matter the order in which you remove them, you just want to get 3 balls out of the bag. What's the number of ways that you can do so?

Well, you take the permutation of the number of ways you can remove 3 balls from a bag of 10 balls, ${}^{10}P_3$. However, this permutation takes removing $\{A, B, C\}$ as a different result as removing $\{A, C, B\}$, even though we get the same final set of balls. So what we want to do is ignore this. How do we ignore this? To answer this question, we must know how many times is each set of 3 counted extra. We have answered this question with our definition of the factorial, that there are 3! ways to order each set of 3 balls.

Thus, we divide ${}^{10}P_3$ by the number of ways each unique set of 3 balls can be ordered, 3!. This removes all the extra counts, and leaves only the unique sets of items where order is not considered. This gives us the derivation of $\binom{n}{k} = \frac{n!}{(n-k)!n!}$.

4 Probability Exercises

Now we move on to some practice exercises, from past midterms.

Exercise 4.1. (2020 Midterm Question 2)

2. [40 points] Government economists have constructed an indicator (red or amber or green) of stress on the financial system which they intend as a signal of risk that the economy may enter a recession (that is, a period of negative growth—policymakers need to find out as soon as possible if the economy is at risk of moving into recession, so that they can react early). The indicator is found to perform as follows: when the economy does enter recession, the previous month's indicator was red 60% of the time and amber 35% of the time, and was green only 5% of the time, whereas overall, the indicator is red 20% of the time, amber 30% of the time, and green 50% of the time. [The indicator is observable one month beforehand, so it's intended as a `leading indicator'. The time difference between indicator and recession is one month throughout this problem.] The economy is either in recession or it isn't; there is no other state.

Over long spans of time, the economy is in recession 10% of the time.

a) Find the probabilities that the economy will enter recession next month if the indicator is red, amber or green respectively.
b) Find the probabilities at the economy will **not** enter recession next month if the indicator is red or amber.

(a) Let D be the event that the economy is in recession. Let R, A and G be the event that the indicator is red, amber or green respectively. Since the indicator and the recession are independent, we know that $\mathbb{P}(D \cap X) = \mathbb{P}(D)\mathbb{P}(X)$, where X is either R, A or G.

$$\mathbb{P}(D|R) = \frac{\mathbb{P}(D \cap R)}{\mathbb{P}(R)} = \frac{0.6(0.1)}{0.2} = 0.3$$
$$\mathbb{P}(D|A) = \frac{\mathbb{P}(D \cap A)}{\mathbb{P}(A)} = \frac{0.35(0.1)}{0.3} = \frac{7}{60}$$
$$\mathbb{P}(D|G) = \frac{\mathbb{P}(D \cap G)}{\mathbb{P}(G)} = \frac{0.05(0.1)}{0.5} = 0.01$$

(b)

$$\mathbb{P}(\overline{D}|R) = 1 - \mathbb{P}(D|R) = 0.3$$
$$\mathbb{P}(\overline{D}|A) = 1 - \mathbb{P}(D|A) = \frac{53}{60}$$

Exercise 4.2. (2020 Midterm Question 1)

(c) Next, consider two events C and D. P(C) = 0.2 and P(D)=0.3. The events C and D (and therefore C-complement and D-complement) are independent. Find the probability that C and D will not both happen (ie the probability that one of them may happen, or neither, but not both).

We want to find $\mathbb{P}(\overline{C \cap D})$.

$$P(C \cap D) = \mathbb{P}(C) \cdot \mathbb{P}(D) = 0.2(0.3) = 0.06$$
$$\mathbb{P}(\overline{C \cap D}) = 1 - \mathbb{P}(C \cap D) = 1 - 0.09 = 0.94$$

Exercise 4.3. (2019 Midterm Question 2) After several years of bad harvests, Zak, a farmer, is heavily in debt and in danger of losing his farm. He has one month to raise the necessary funds. Fortunately, he has three children who work in the financial industry, and who are going to make some speculative bets with their savings just before earnings season, in the hop of making enough to pay off the debts and save the farm. There is no other way that the farm can be saved.

Aline bets on a mining stock; she has a 1% chance of making a huge profit and saving the farm. Bernadette bets on a healthcare startup; she has a 2% chance (of making a huge profit and saving the farm). Finally, Carmen bets on a tech company; she has a 3% chance. With any of these bets, either it pays off so richly that it's enough to save the farm, or the person's savings will be lost. The bets are statistically independent.

(a) What's the probability that Zak loses the farm?

Let A, B, C be the events that Aline, Bernadette and Carmen make a huge profit respectively. Zak loses the farm if none of them make a huge profit, which is $\mathbb{P}(\overline{A \cup B \cup C})$.

$$\begin{aligned}\mathbb{P}(\overline{A \cup B \cup C}) &= \mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) \\ &= \mathbb{P}(\overline{A}) \cdot \mathbb{P}(\overline{B}) \cdot \mathbb{P}(\overline{C}) \\ &= 0.99(0.98)(0.97) = 0.941\end{aligned}$$

(b) A month passes, and against the odds, the farm is saved. Knowing this, what's the probability that only one (that is, exactly one) of the children lost her savings?

Let S be the event that the farm is saved, and let L be the event that only one of the children lost her savings.

$$\begin{aligned}\mathbb{P}(L) &= \mathbb{P}(\overline{A} \cap B \cap C) + \mathbb{P}(\overline{B} \cap A \cap C) + \mathbb{P}(\overline{C} \cap A \cap B) \\ &= 0.99(0.02)(0.03) + 0.98(0.01)(0.03) + 0.97(0.01)(0.02) \\ &= 0.00108 \\ \mathbb{P}(L|S) &= \frac{\mathbb{P}(S \cap L)}{\mathbb{P}(S)} = \frac{\mathbb{P}(S|L)\mathbb{P}(L)}{\mathbb{P}(S)} \\ &= \frac{1(0.00108)}{1 - 0.941} = 0.0183\end{aligned}$$

Exercise 4.4. (2019 Midterm Question 3) Securities regulators are monitoring a stockbroker whom they suspect of passing inside information on a company to her clients. Because she tells clients to buy or sell the company's shares frequently, it's hard to distinguish legitimate recommendations, based on her interpretation of publicly-available information, from possible illegal recommendations based on tips from insiders.

Regulators have followed her recommendations for three years, or 750 trading days, during which there have been 75 days on which she has called clients to recommend selling the stock, and 25 days on which she has called to recommend buying. Regulators know the days on which negative and positive earning developments were known to insiders, but were not yet announced to the public. Of the days on which the broker said 'sell', 30% were days in which such negative information was known to insiders only. Over the three-year period, there were 12 cases in which insiders had information earlier, of which five were related to negative earnings news. In these cases, insiders got the information 7 days ahead of the public.

(a) What is the probability that the broker says 'sell' when it's a day on which insiders have negative information not available to the public?

Let N be the event that there is negative information available to insiders only on a given day, and S be the event where the stockbroker gives a 'sell' recommendation on a given day. Based on information from the question, $\mathbb{P}(S) = 0.1$, and $\mathbb{P}(N) = \frac{5 \cdot 7}{750} = 0.0467$.

$$\mathbb{P}(S|N) = \frac{\mathbb{P}(S \cap N)}{\mathbb{P}(N)} = \frac{\mathbb{P}(N|S) \cdot \mathbb{P}(S)}{\mathbb{P}(N)} = \frac{0.3(0.1)}{0.467} = 0.643$$

(b) What is the probability that we have both a 'sell' recommendation and negative information available to insiders only, on any given days?

$$\mathbb{P}(S \cap N) = \mathbb{P}(N|S) \cdot \mathbb{P}(S) = 0.3(0.1) = 0.03$$

5 Combinatorics Exercises

Exercise 5.1. (2020 Midterm Question 3)

3. [30 points] You are at an amusement park and are considering taking a ride on a Ferris wheel (this is an amusement park ride in which people are raised high in the air on a revolving wheel with seats). However, you have just read of a major study of equipment maintenance standards which suggests that some of these rides are unsafe. The study, which you believe to be accurate, found that 20% of major rides are poorly maintained and unsafe. For such rides, there is a 10% probability that a passenger will be injured on any given ride (that is, riding once). Customers can't see whether the machine has been well maintained or not, so they don't ride the unsafe ones proportionally less. For safe rides, the probability of injury will be taken to be zero.

a)One of your friends comes back from a ride on the ferris wheel, uninjured. What is the probability that the ferris wheel is *un*safe? Interpret your calculation.

b)A government agency monitoring safety is comparing results from two major manufacturers of ferris wheels, Consolidated Rides (CR) and Blue Sky Amusements (BSA). It compares accident rates from 6 CR rides (i.e. ferris wheels) and 8 BSA rides. If the two manufacturers in fact produce equally safe products, so that any differences are purely the result of random chance, what is the probability that the top three highest-accident-rate rides are all from BSA?

Exercise 5.2. (a) Let U be the event where a major ride is poorly maintained and unsafe. Let I be the event that a passenger is injured on any given ride.
From the question, we have that $\mathbb{P}(U) = 0.2$, $\mathbb{P}(I|U) = 0.1$ and $\mathbb{P}(I|\overline{U}) = 0$ (probability of injury on safe rides is 0).

$$\begin{aligned}
\mathbb{P}(\overline{I}|U) &= 1 - \mathbb{P}(I|U) = 0.9 \\
\mathbb{P}(I) &= \mathbb{P}(I \cap U) + \mathbb{P}(I \cap \overline{U}), \text{ since } U \text{ and } \overline{U} \text{ exhaustive} \\
&= \mathbb{P}(I|\overline{U})\mathbb{P}(\overline{U}) + \mathbb{P}(I|U)\mathbb{P}(U) \\
&= 0 + 0.1(0.2) = 0.02 \\
\mathbb{P}(\overline{I}) &= 1 - \mathbb{P}(I) = 0.98 \\
\mathbb{P}(U|\overline{I}) &= \frac{\mathbb{P}(\overline{I}|U)\mathbb{P}(U)}{\mathbb{P}(\overline{I})}, \text{ by Bayes' Theorem} \\
&= \frac{0.9(0.2)}{0.98} = 0.184
\end{aligned}$$

(b) 6 CR rides, 8 BSA rides

$$\begin{aligned}
\mathbb{P}(\text{top 3 from BSA}) &= \frac{\text{number of ways that top 3 are from BSA}}{\text{total number of ways to choose top 3}} \\
&= \frac{{}^{8}C_3}{{}^{14}C_3} = \frac{56}{364} = 0.154
\end{aligned}$$

Exercise 5.3. (2020 Assignment 2 Q2)

2. An international development organization has eight microeconomists and six regional economic development specialists, among others. A team of four regional specialists and two microeconomists is to be chosen for an extended mission to a country undergoing debt restructuring.
   a) How may different combinations are possible?
   b) One of the microeconomists detests one of the regional specialists (and the sentiment is mutual). A team with both will have difficulty functioning. If the people picked for the project are chosen randomly, what is the probability that these two are both chosen?
   c) What is the probability that neither of them is chosen?

(a)

$$\begin{aligned}\text{number of combinations} &= {}^6C_4({}^8C_2)\\ &= 15(28) = 420\end{aligned}$$

(b)

$$\begin{aligned}\mathbb{P}(\text{both are chosen}) &= \frac{\text{number of ways both are chosen}}{\text{total number of ways}}\\ &= \frac{\text{number of ways to pick everyone else}}{420}\\ &= \frac{{}^5C_3({}^7C_1)}{420} = \frac{1}{6}\end{aligned}$$

(c)

$$\begin{aligned}\mathbb{P}(\text{neither are chosen}) &= \mathbb{P}(\text{microeconomist not chosen})\mathbb{P}(\text{regional specialist not chosen})\\ &= \left(1-\frac{1}{4}\right)\left(1-\frac{2}{3}\right)\\ &= \frac{3}{4}\cdot\frac{1}{3} = \frac{1}{4}\end{aligned}$$