14.30 Introduction to Statistical Methods in Economics
Spring 2009

Problem Set #1 - Solutions

14.30 - Intro. to Statistical Methods in Economics

Instructor: Konrad Menzel

Due: Tuesday, February 17, 2009

Instructions

You may work together to solve the problems but must each hand in independently-written solutions, so make sure to show all of your work. Each part of each question is worth 1 point, although partial credit may be rewarded for incorrect answers.

Question 1

A delegation of three is to be chosen from the untenured faculty of the MIT Economics Department (numbering ten) to represent the department in an Institute-wide committee. In how many ways

a) can the delegation be chosen?

- Solution to (a): Since there are 10 untenured faculty members, there are "10 choose 3" $\binom{10}{3} = \frac{10!}{3!7!} = \frac{10\cdot9\cdot8}{3\cdot2\cdot1} = 120$ different ways to choose the three faculty members. Further, if the three faculty members are selected to unique positions such as the President, Vice-President, and Secretary of the delegation, then there would be $\frac{10!}{7!} = 10\cdot9\cdot8 = 720$ total ways since each of the 120 different combinations of faculty members can be mixed into the 3 positions in 6 different ways.

b) can it be chosen, if two people refuse to go together?

- Solution to (b): Since from part (a) we know that there are 120 different ways to selcted among the 10 untenured faculty members, we can approach this problem from at least two directions. We can either subtract out the combinations where the two odious faculty members have been included, or we can directly count the number of ways that we can have at most one of the two who refuse.
  The first way is simple. We compute 120 - #(Both are chosen to go together). In order to get them both, we first choose the two of them (this is a constraint on the problem–we should try to satisfy these either first or last to avoid more complex counting issues). There is only one way to choose both of them in a combination sense (although with permutations there are definitely 2 ways to pick them: you can pick one or the other first); this leaves us with 8 different ways to pick the remainder member of the committee. Thus, we arrive at our answer of $120 - 8 = 112$.
  In order to demonstrate that the other method would give the same answer, let's first count the number of ways to pick a group without having to deal with the two odious individuals: $\binom{8}{3} = \frac{8!}{3!5!} = \frac{8\cdot7\cdot6}{3\cdot2\cdot1} = 56$. Now, recalling that there are two ways to pick an odious member, let's count the number of ways we can pick two additional faculty members: $\binom{8}{2} = \frac{8!}{2!6!} = \frac{8\cdot7}{2\cdot1} = 28$. So, we arrive at our final count:

$$\begin{aligned}\#(at\,most\,1\,odious\,member) &= \#(0\,odious\,members) + 2\cdot\#(1\,odious\,member)\\ &= 56 + 2\cdot28 = 112.\end{aligned}$$

  As we hoped, both approaches gave the same answer.

c) can it be chosen, if two particular members insist on either both going or neither going?

- Solution to (c): If two members insist on either both going or neither going, we have a similar problem as before, but we can compute it very simply from the numbers from part (b). The event A that we're interested in is $A = \{Both\, go \cup Neither\, go\}$. In part (b) we discovered the following:

$$\begin{aligned} \#\{3\, Person\, Committees\} &= 120 \\ \#\{Both\, go\} &= 8 \\ \#\{Only\, 1\, goes\} &= 28 \times 2 = 56 \\ \#\{Neither\, goes\} &= 56 \end{aligned}$$

So, in order to use the addition counting rule, we need to have mutually exclusive events. But, obviously $\{Both\, go\} \cap \{Neither\, goes\} = \emptyset$ so we simply add the two counts together: $8 + 56 = 64$.

d) can it be chosen, if two people must be chosen from MIT assistant faculty (6 professors) and one person must be chosen from visiting assistant faculty (4 professors)?

- Solution to (d): This should be easy if you've already figured out parts (a)-(c), not because you've computed the answer yet, but because the methods are similar. We just break the problem into its two parts: choosing MIT assistant faculty and choosing visiting assistant faculty. This give us $\underbrace{\frac{6 \cdot 5}{2}}_{MIT} \cdot \underbrace{4}_{Visiting} = 60$. Does this answer make sense? It is less than if there were just two people who hated each other, as this is equivalent to the problem where four people refuse to go together, but we have to have one of them. So, we would expect the count to be less than what we found in (b).

Question 2

In the seventeenth century, Italian gamblers used to bet on the total number of spots rolled with three dice. They believed that the chance of rolling a total of 9 ought to equal the chance of rolling a total of 10. They noted that altogether there are six combinations to make 9: (1,2,6), (1,3,5), (1,4,4), (2,3,4), (2,2,5), and (3,3,3). Similarly, there are six combinations for 10: (1,4,5), (1,3,6), (2,2,6), (2,3,5), (2,4,4), (3,3,4). Thus, argued the gamblers, 9 and 10 should have the same chance. Empirically, they found this not to be true, however. Galileo solved the gambers' problem. How?

a) How many permutations of three dice are there that sum to 9?

- Solution to (a): Here is the list of how many permutations for each 3-tuple that sum to 9:

$$\begin{aligned} \#\{1,2,6\} &= 3! = 6 \\ \#\{1,3,5\} &= 3! = 6 \\ \#\{1,4,4\} &= \tfrac{3!}{2!} = 3 \\ \#\{2,3,4\} &= 3! = 6 \\ \#\{2,2,5\} &= \tfrac{3!}{2!} = 3 \\ \#\{3,3,3\} &= \tfrac{3!}{3!} = 1 \end{aligned}$$

So, we find that there are a total of $6 + 6 + 3 + 6 + 3 + 1 = 25$ possible ways to obtain 9 as the sum of the three dice.

b) How many permutations of three dice are there that sum to 10?

- Solution to (b): Here is the list of how many permutations for each 3-tuple that sum to 10:(1,4,5), (1,3,6), (2,2,6), (2,3,5), (2,4,4), (3,3,4)

$$\begin{aligned}
\#\{1,4,5\} &= 3! = & 6 \\
\#\{1,3,6\} &= 3! = & 6 \\
\#\{2,2,6\} &= \tfrac{3!}{2!} = & 3 \\
\#\{2,3,5\} &= 3! = & 6 \\
\#\{2,4,4\} &= \tfrac{3!}{2!} = & 3 \\
\#\{3,3,4\} &= \tfrac{3!}{2!} = & 3
\end{aligned}$$

So, we find that there are a total of $6+6+3+6+3+3=27$ possible ways to obtain 9 as the sum of the three dice.

c) How many total permutations of three dice are there? What was Galileo's solution? Explain.

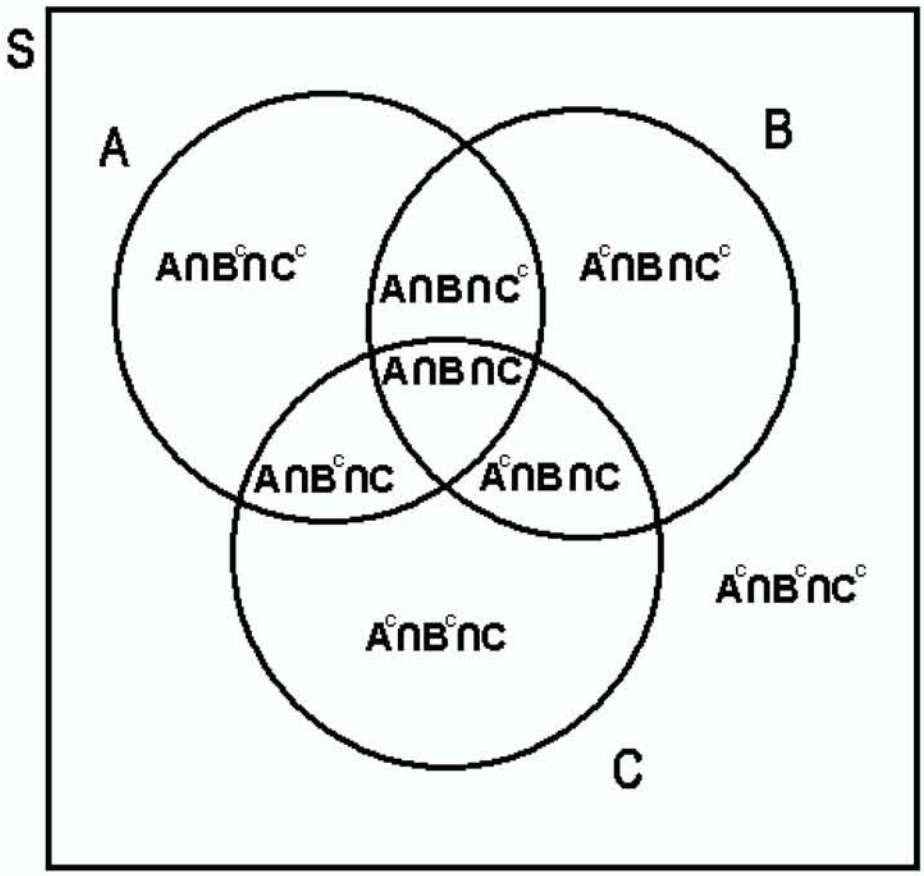
- Solution to (c): Since there are a total of $2^3 = 216$ possible permutations of the three dice, Galileo's solution must have been the application of the multiplication rule. Since each die gives us an independent outcome, we know that each permutation of the 3-tuples is equally likely. So, what that means is that rather than thinking about the combinations of the 3 dice, we must consider the permutations which are uniformly or evenly likely. The Italian gamblers then would have adjusted their bets to account for the fact that 10's are more likely than 9's: $\frac{25}{216} \leq \frac{27}{216}$.
Note: a quick (albeit incomplete) answer to the problem would have observed that since we expect each dice to have an average contribution of $3.5 = (1+2+3+4+5+6)/6 = \frac{21}{6}$, we will expect 10 and 11 to have the same probabilities due to symmetry and a Central Limit Theorem will tell us that the probability of getting events away from the center of the possible outcomes (when summing random variables) are less probable. So, since 9 is away from the average outcome of 10.5, we know that the probabilities should decline monotonically. You'll learn more about this later in the course.

Question 3

Venn diagrams or set diagrams are diagrams that show all hypothetically possible logical relations between a finite collection of sets (groups of things). Venn diagrams were invented around 1880 by John Venn. They are used in many fields, including set theory, probability, logic, statistics, and computer science (Wikipedia: http://en.wikipedia.org/wiki/Venn_Diagram).

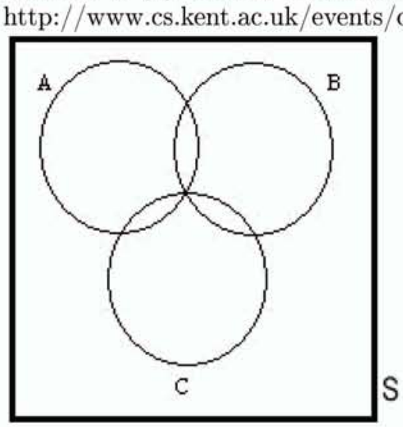
a) Draw a Venn diagram for the three events A, B, and C contained in the sample space S and properly label all possible union and intersections of events.

- Solution to (a): Here is an example:



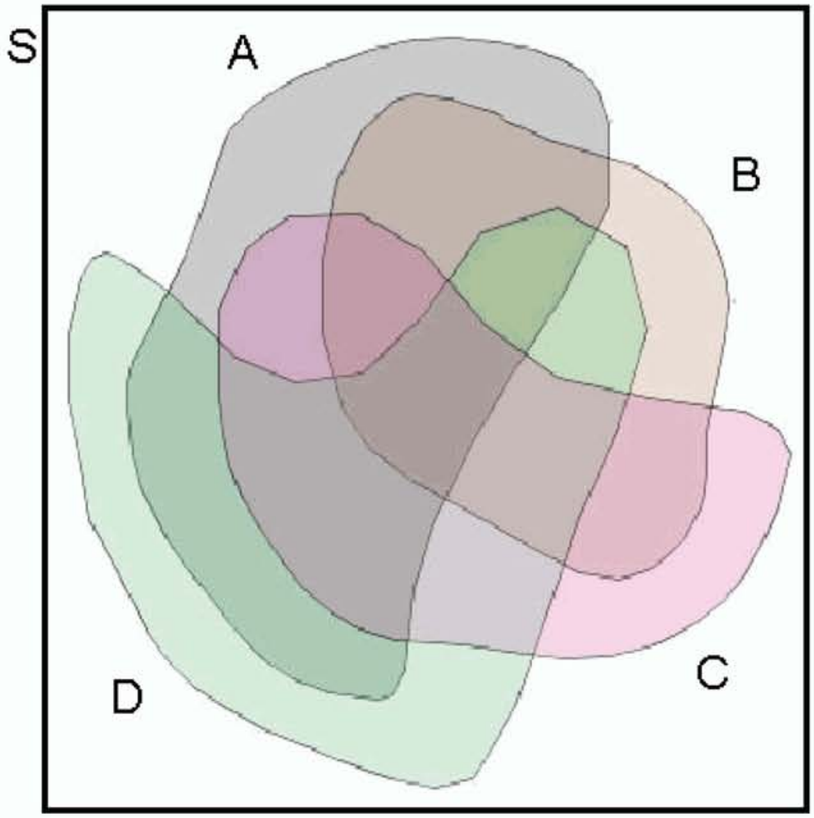
b) Draw a Venn diagram for the three events A, B, and C contained in the sample space S and properly label all possible combination of events where $A \cap B \cap C = \emptyset$.

- Solution to (b): Here is an adapted example from the following website: http://www.cs.kent.ac.uk/events/conf/2004/euler/eulerdiagrams.html.



c) Try (but don't spend too much time–it's just for fun) to draw a complete Venn diagram for the four events A, B, C, and D contained in the sample space S where you include all possible unions and intersections of events. How many mutually exclusive regions should such a diagram include?

- Solution to (c): Here is an example from the website cited in part (b).



d) How many mutually exclusive regions should such a diagram with $k \in \mathbb{N}$ events include?

- Solution to (d): We will try this problem two different ways. A proof by induction (show true for 1 and then true for k and $k+1$) would be preferred as it provides the logical justifications, but we'll start by looking at small examples.
  We first note that if you have just one event, there are two possible regions: A and A^c. With two events, we know that there are $A \cap B$, $A^c \cap B$, $A \cap B^c$, *and* $A^c \cap B^c$. That's four. Perhaps we're seeing a pattern. If you look at your picture for part (a) and count all of the separate regions, you'll find that there are eight for 3 events. We hypothesize that the formula for the number of is 2^k.
  A proof by induction would go like this: With one event, you're either in or out, which gives us 2. Assume the formula holds for k:2^k is the number of mutually exclusive events in such a diagram. For $k+1$, we recognize, by the same logic that for the $k+1$ event, both it and its complement will intersect the 2^k events, generating 2^{k+1} mututally exclusive events.
  Note: I suspect that some of you may have interpreted this question differently. I will try to take into account any confusion for which I am responsible in the grading process.

Question 4

Does a monkey have a better chance of rearranging "ACCLLUUS" to spell "CALCULUS," or of rearranging "AABEGLR" to spell "ALGEBRA?" (2 points.)

- Solution: How many letters is "CALCULUS"? Eight. How many letters is "ALGEBRA"? Seven. How many unique permutations of the letters "ACCLLUUS" are there? We have $\frac{8!}{2!2!2!} = 7!$ different permutations since there are two C's, L's, and U's. How many unique permutations of the letters "AABEGLR" are there? We start out with at most 7!, but we have duplicate A's, so we know that there must be 7!/2! different permutations, which is less than 7!. Thus, we know that "ALGEBRA" is twice as likely to come up randomly on a monkey's Shakespearean typewriter.

Question 5

In Lecture 1, you learned about event partitions. Give three different examples of partitions of a single draw from a deck of playing cards.

- Solution: There are many different partitions that could be made. I will provide just three examples here.
  - Example 1: $\{\{All\,Hearts\}, \{All\,Clubs\}, \{All\,Spades\}, \{All\,Diamonds\}\}$
  - Example 2: $\{\{All\,Red\,Cards\}, \{All\,Black\,Cards\}\}$
  - Example 3: $\{\{All\,Even\,Numbers\}, \{All\,Odd\,Numbers\}, \{All\,Aces\}, \{All\,Royalty\}\}$

Question 6

The MIT football team plays 12 games in a season. In each game they have $\frac{1}{3}$ probability of winning, $\frac{1}{2}$ probability of losing, and $\frac{1}{6}$ probability of tying. Games are independent. What is the probability that the team has 8-3-1 record? (8 wins, 3 losses, and 1 ties)

- Solution: Sorry, there was a typo on the problem set. If you solved it correctly for either an 8-3-1 or 6-4-2 record, you'll receive full credit.
  You should have received the following solution for the 8-3-1 record. Since each game is an independent event, we can compute the probability of getting that particular record in a particular order: $\left(\frac{1}{3}\right)^8 \left(\frac{1}{2}\right)^3 \left(\frac{1}{6}\right)^1 = 3.175e-06$. Now, that's pretty small! However, we still need to compute the number of ways that we could obtain that record. If there are twelve games and we have 8 wins, 3 losses, and 1 tie, we can order those in $\frac{12!}{8!3!1!} = 1980$ ways. That's a lot of ways, though! This means that the total probability of the Beavers getting a decent record like that is
  $$\frac{12!}{8!3!1!}\left(\frac{1}{3}\right)^8 \left(\frac{1}{2}\right)^3 \left(\frac{1}{6}\right)^1 = .00628715$$
  which is still pretty small, but we have to remember that this is just the probability of getting exactly this record. If we want at least that good of a record, we have to add in all better possible seasons as well.
  Note: These probabilities and records weren't actually taken from the Beaver's true performance. ;)

Question 7

You and your friends just rented a car from Enterprise for an 8,000 mile cross-country road trip to see all of the sights from from Boston Harbor to the Golden Gate Bridge. Your rental car may be of three different types: brand new (and not a lemon), nearly 1 year old, or a lemon (bound to break down). That many miles can be demanding on a rental car. If the car you receive is brand new (New), it will break down with probability 0.05. If it is one year old (One), it will break down with probability 0.1. If it is just a lemon (Lemon), it will break down with probability 0.9. The probability that the car Enterprise gives you a car that is New, One, or Lemon is 0.8, 0.1, and 0.1, respectively. Compute the probability that your car is going to break down on your road trip.

- Solution: Since getting each car is a mutually exclusive event, we can simply add up the probabilities of getting each car after multiplying by the probabilities of break down for each car. In particular, we get that the probability of getting a new car and it breaking down is jointly $0.8 \times 0.05 = 0.04$. Similarly, we find for one year old cars and lemons the joint probability to be $0.1 \times 0.1 = 0.01$ and $0.1 \times 0.9 = 0.09$. Thus, we obtain the total probability of the car breaking down is simply

$$\underbrace{0.04}_{New} + \underbrace{0.01}_{One} + \underbrace{0.09}_{Lemon} = 0.13$$

Make sure to take your cell phones with you so you can call for help!

Question 8

a) Bayes' formula is really important. Write down Bayes' formula and describe it in words.

- Solution to (a): Bayes' formula is easily derived from the conditional expectation formula:

$$\begin{aligned} P(A|B) &= \frac{P(A,B)}{P(B)} \\ P(A|B)\,P(B) &= P(A,B) \end{aligned}$$

which by symmetry gives us Bayes' formula: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$. A more general version makes use of the Law of Total Probability:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}.$$

What Bayes' formula tells us is that if we know the marginals of two events and we know even one of the conditionals, we can back out what the other conditional distribution is. This is really helpful, as illustrated by your work on the problems below.

Further, here are a couple of common applications.

b) Suppose that five percent of men and 0.25 percent of women are color blind. A colorblind person is chosen at random. What is the probability of this person being male? Assume that there are an equal number of males and females. What if there were twice as many males as females?

- Solution to (b): The events A and B are being male and colorblind. We write the formula:

$$\begin{aligned} P(Male|Colorblind) &= \frac{P(Colorblind|Male)\,P(Male)}{P(Colorblind)} \\ &= \frac{(0.05)\,(0.5)}{(0.5 \times 0.05 + 0.5 \times 0.0025)} \\ &= 0.95238095 = 95.2\% \end{aligned}$$

which we arrive at by making the assumption that 50% of the population is male and 50% is female and have computed the marginal probability of being colorblind by multiplying the probability (independent draws) of being male and colorblind and adding that to the probability of female and colorblind (two mututally exclusive events): $0.5 \times 0.05 + 0.5 \times 0.0025$. If there are twice as many males as females we just have to change the formula slightly to account for the different marginal of A. After doing this we obtain:

$$\begin{aligned} P(Male|Colorblind) &= \frac{P(Colorblind|Male)\,P(Male)}{P(Colorblind)} \\ &= \frac{(0.05)\left(\frac{2}{3}\right)}{\left(\frac{2}{3} \times 0.05 + \frac{1}{3} \times 0.0025\right)} \\ &= 0.97560976 = 97.6\% \end{aligned}$$

c) Suppose that there exists an imperfect test for Tuberculosis (TB). If someone has TB, there is a ninety-five percent chance that the test will come up "red." If someone does not have TB, there is only a two percent chance that the test will come up red. Finally, the chance that anyone has TB is, say, five percent (in the United States; in other countries Tuberculosis is endemic). Once someone takes the test and it comes up red, what is the probability that they have TB?

- Solution to (c): We do the same thing as in part (b), just with different marginal and conditional probabilities and some algebra. In particular, we have $P(Positive|TB) = 0.95$, $P(Positive|\sim TB) = 0.02$, and $P(TB) = 0.05$. What we want to know is $P(TB|Positive)$. We can compute the marginal $P(Positive)$ by using the information that we know from the marginal and conditional distributions:

$$\begin{aligned} P(Positive) &= P(TB)\,P(Positive|TB) + P(\sim TB)\,P(Positive|\sim TB) \\ &= (0.05) \times (0.95) + (1 - 0.05) \times (0.02) \end{aligned}$$

Now that we know that marginal, we can use Bayes' formula:

$$\begin{aligned} P(TB|Positive) &= \frac{P(Positive|TB)\,P(TB)}{P(Positive)} \\ &= \frac{0.95 \times 0.05}{(0.05) \times (0.95) + (1 - 0.05) \times (0.02)} \\ &= 0.71428571 = 71.4\% \end{aligned}$$

This is why rare diseases can be so hard to accurately diagnose, because even small false positive rates get amplified by the large population that is going to test positive that does not have the disease.