Astronomy 505: Homework #2 (Probability)

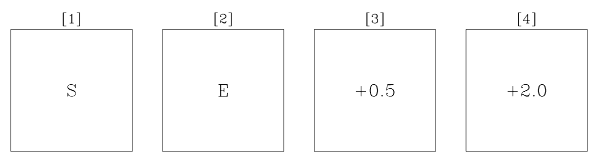
1. (a) An Astronomy 110 student was extremely keen on astronomy. She often went to night sky observing parties at White Sands Missile Range, attended the open houses at the campus observatories every month, and became the secretary of the Astronomical Society of Las Cruces. After graduating she takes a job in a local company in order to keep living in a region with such clear skies. Which is greater: the probability that her occupation is that of electrical engineer, or the probability that she is an electrical engineer who is active in local amateur astronomy circles? (2 pts)

   (b) A fair coin is flipped three times. (A coin is said to be fair if the possible outcomes, heads and tails, are equally likely. The sample description space is defined to be $\Omega = \{H, T\}$.) What is the probability of obtaining two heads and one tail? (2 pts)

   (c) A fair die is tossed twice. Given that a 3 appears on the first toss, what is the probability obtaining the sum 7 in two tosses? (2 pts)

2. Consider various samples of data drawn from the Sloan Digital Sky Survey (SDSS).

   (a) You are given a set of cards, one for each of four galaxies, that list morphological type on one side and $V - I$ colour on the other. You are asked to test the rule *A spiral galaxy has a* $V - I$ *colour between -0.5 and 1.0.* Which of the following cards should you turn over to test directly the veracity of the rule? (3 pts)



   (b) A subsample contains five galaxies numbered 1, 2, 3, 4, 5. A galaxy can be selected at random, and is then replaced before the next galaxy is selected. (This is called sampling with replacement.) What is the probability of drawing the same galaxy (any one) twice in two tries? (3 pts)

   (c) Describe the sample description space Ω for two galaxies selected (1) with replacement, and (2) without replacement. (3 pts)

   (d) Two galaxies are drawn at random, with replacement. Three students were asked to compute the probability p that the sum of their numbers was equal to 5. The first computed $p = \frac{2}{15}$, arguing that there are 15 distinguishable pairs and only 2 are favourable. The second computed $p = \frac{1}{9}$, arguing that there are 9 distinguishable sums, of which only 1 was favourable. The third computed $p = \frac{4}{25}$, arguing that there were 25 distinguishable ordered outcomes of which 4 were favourable. What is the correct answer? Explain what is wrong with the reasoning of the two incorrect answers. (5 pts)

   (e) For a subsample of 800 stellar objects, 200 are identified as *variable*, 200 are identified as *M dwarfs*, and the remaining are identified as *variable M dwarfs*. One element is drawn at random from the sample. Let A and B represent the events of observing states *variable* and *M dwarf*, respectively. Find $P[A]$, $P[B]$, and $P[AB]$. Are A and B independent? (5 pts)

   (f) Assume that 0.1% of the objects within a subsample are quasars. You examine 100 spectra from the subsample in detail. What is the probability that at least one of the spectra will be that of a quasar? (5 pts)

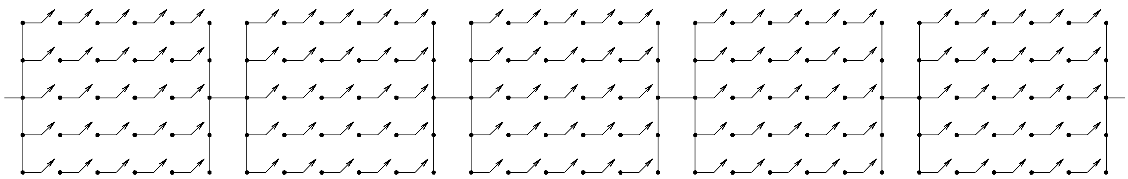
   (g) You then write a spectral analysis routine that fits absorption lines to each spectrum, and compares it to a set of template quasar spectra. The program successfully detects all quasars, but also misidentifies 0.5% of other spectral candidates as quasars. If the first object which you analyze is identified by your program as a quasar, what is the probability that it actually is one? (5 pts)

3. The automatic guider on the Keck I telescope fails with probability P_g, and an independent monitor system fails to detect and counteract guider failure with probability P_m. Professor Pompous argues that if $P_m > P_g$, installation of the monitor is useless. Show that he needs to retake Astronomy 505 by computing the probability that the telescope loses guiding with and without the monitor system in place. Assume that $P_m = 10\% = 2 \times P_g$. (5 pts)

4. A group of Astronomy 110 students is taking a final exam. For a particular multiple-choice question on the test, the fraction of students who know the answer is p. The probability of answering the question correctly is unity for those who know the answer, and $\frac{1}{m}$ for those who guess, given m possible multiple-choice answers. Compute the probability that a student knew the answer to the question, given that they answered correctly. (5 pts)

5. Consider the hardware which controls the rotator angle (the orientation of the spectroscope slit on the sky) for the Keck I telescope. For simplicity, we will model a portion of it as a series of electronic switches. In order for the rotator angle to be correctly updated, there must exist a continuous path for current to flow through these switches. (You don't actually need to know anything about the telescope or about spectroscopy in order to study the behavior of the switches; this information is merely provided to give you a framework for the problem.)

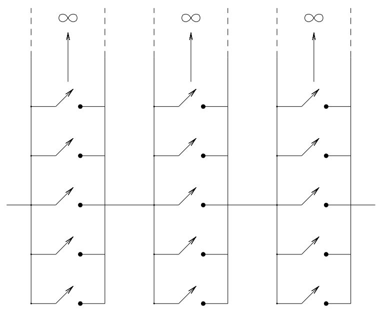
(a) Visualize n switches in series which form a single row of switches. Next, place n such rows in parallel to form a block of switches. Finally, place n such blocks of n^2 switches in a series. Imagine that each of the n^3 switches is,



independently, either closed with probability p or open (as shown in the figure, for $n = 5$) with probability $1 - p$. What is the probability $\mathrm{P}(n, p)$ that this circuit can conduct current from on end to the other? (10 pts)

(b) Now imagine that the n blocks are connected in parallel rather than in series. What is the probability $\mathrm{P}(n, p)$ that this circuit can conduct current? (5 pts)

(c) *For extra credit (and extra respect)*: We can adapt the hardware slightly to model a random process, one in which time is a factor. Consider the case of three blocks placed in series, where each block is composed of an arbitrarily large number of single parallel switches (see figure). At time $t = 0$, all of the switches are open. Thereafter, once each microsecond, there is a 3% chance that one switch (somewhere) will close (and stay closed). At what time will the probability that this circuit can conduct current rise above 50%? (10 pts)



(d) Write a computer program to calculate and plot $\mathrm{P}(n, p)$ for case (a) and case (b), for $0 \leq p \leq 1$ and $n = 5$. Attach a printout of your well-commented code and an annotated copy of your plot to your assignment. (If you are not sure what programming environment to select, consider that both the numerical calculations and the plotting can be done fairly simply with supermongo, which has a low threshold for competency.) Compare the behavior of $\mathrm{P}(n, p)$ for case (a) versus case (b). (10 pts)

(e) Complete the same exercise, for the case $n = 10$. Discuss any trends in $\mathrm{P}(n, p)$ with n. Is there a characteristic threshold of significance, in either case? (5 pts)

(f) *For extra credit*: Calculate and plot the probability that zero blocks, or any one, any two, or all three blocks in case (c) will conduct, as a function of time. (*Note that the simplest way to complete (c) is by completing (f).*) (5 pts)

Note: Problems 5(c) through 5(f) will require a bit more effort than the others, because of the need to write, compile, and execute a program. Because of this, you may work through them in the assigned groups of two. You may also take a second week if necessary; be sure to allow sufficient time to work through the necessary steps.