33-765 — Statistical Physics

Department of Physics, Carnegie Mellon University, Spring Term 2019, Deserno

Problem sheet #1

1. Playing games with non-standard dice *(5 points)*

The picture on the right shows a curious quadruplet of dice: their faces are not labeled consecutively with the numbers one to six, but rather have the following numbers printed on them:

D_1:	{3,3,3,3,3,3}	(yellow)
D_2:	{0,0,4,4,4,4}	(red)
D_3:	{1,1,1,5,5,5}	(green)
D_4:	{2,2,2,2,6,6}	(blue)

We will assume that, like for an ordinary die, the likelihood of any of the six sides of any of the four dice coming up when we roll the die is $1/6$.
Assume now that two people are playing a game: each one picks one of the dice and then they roll them. Whoever gets the higher number wins the game. Since the dice are different, it is conceivable that the probability of winning depends on which die one chooses. There is no reason to believe it is still 50%, as it would be for normal dice!
Hence, let us introduce the following terminology: We call die A "stronger" than die B (or die B "weaker" than die A), when the probability of a roll of A beating a roll of B is larger than 50%. Let us formally write this as "$A > B$", or "$B < A$".

1. Show that $D_1 < D_2$, $D_2 < D_3$, and $D_3 < D_4$!
   Hint: Of course you can do this by carefully calculating the probabilities of winning, but you don't have to do it that way. In all cases it is possible to give a simple (but stringent!) argument to show why one die is stronger than the other.
2. Based on the sequence you have just shown, do you expect $D_4 > D_1$? Check explicitly whether this holds!
3. How does D_1 compare to D_3, and how does D_2 compare to D_4?

2. Pairwise and mutual independence *(5 points)*

We flip a fair coin twice and see whether it falls head ("H") or tail ("T"). Let us look at the following three events: A: "There was a head in the first throw." B: "There was a head in the second throw." C: "The result was the same in the first and the second throw." Show that the three events A, B and C are *pairwise independent* but not *mutually independent*!

3. Simpson's paradox *(5 points)*

A college offers two majors (A and B), to which both male and female students apply. The fraction of male and female students interested in major A is μ_A and ϕ_A, respectively. Show that the following can happen: *in both majors* the acceptance probabilities f_A and f_B for women are *larger* than those for men (m_A and m_B), and yet the *overall* acceptance rate for women is *lower* than that for men. Give a *complete and precise characterization* of the circumstances under which this situation occurs!

4. Sick Bayes *(5 points)*

Consider a disease that exists with some small probability p in the general population. Assume that people can be checked for the disease with a test that correctly picks it up with a (large) probability α (which is often called the "sensitivity" of the test). Of course, any test also has a (hopefully small) false positive rate β. (Incidentally, $1-\beta$ is often called the "specificity" of the test). If a random person gets tested positive, what is the probability of them having the disease? How does one have to design such a test so that test-takers are not unnecessarily scared? Give an illustrative numerical example!