

Math 4600: Homework 7

Due: March 8th

1. A box has 4 balls: 2 blue and 2 white. You pick two at random with replacement. What is the probability that both will be blue? Solve this in two ways:
 - (a) Compute probability analytically, similar to what we did in class using “n choose k”.
 - (b) **(computing)** Do a “Monte-Carlo” simulation of this experiment, namely, pretend the balls are numbered from 1 to 4 (with blues being, say, 1 and 2), and ask the MATLAB to draw 2 of them at random, say, 10 times, and each time record whether you got the desired result (blue-blue) or not. From this, make an estimate of the probability that both drawn balls are blue. Do this with 10, 100 and 1000 trials and compare results with your theoretical result from (a). There is a code on the homework page that you can use as a part of your code.

2. **(computing)** When the number of trials becomes very large ($n > 100$) and the probability of success becomes very small ($p < 0.01$), it is hard to compute binomial probabilities even using a computer. In a situation like this, we can find the binomial probabilities approximately using the following fact:

$$b(k; n, p) \approx p(k; np) = \frac{e^{-np}(np)^k}{k!}.$$

The function $p(k; \Lambda)$ with parameter $\Lambda = np$ is the probability distribution function for a Poisson distribution with parameter np .

Look at the probability distribution function for the binomial distribution and its poisson approximation (use the provided code if you like). Start with $n = 50$, $p = 0.1$. What happens when p gets smaller? larger? is the approximation working? What happens if n gets larger? smaller?

3. In the lab there are animals of 2 different phenotypes (A and B). There are 10 and 7 animals of first and second phenotype, respectively. A researcher picks up 3 animals at random (without replacement).
 - (a) What is the probability to get (AAA) combination?
 - (b) What is the probability that the third animal has phenotype B?
 - (c) **(computing)** In general, if there are K objects of one kind and M objects of the other kind, and n objects are picked without replacement, then X - the number of objects of the first kind in the sample - has hypergeometric distribution:

$$\Pr(X = k) = \frac{\binom{K}{k} \binom{M}{n-k}}{\binom{K+M}{n}}$$

Plot in MATLAB the hypergeometric probability distribution function with given parameters as a bar graph (similar to what you did in problem 2 for the binomial distribution). Use the graph to look up the probability to have all three animals to be of phenotype B.

4. (From F. Adler “Modeling the Dynamics of Life”). An important application of Bayes’ theorem is the planning of rare disease screening. Suppose we have a disease that affects 1% of people. A diagnostic test always picks up the disease but also generates 5% false positives (a positive test when there is no disease).
 - (a) A patient tests positive. What is the probability that she has the disease? (You may find it useful to consider the following events: D - the disease is present, N - there is no disease, P - the test is positive).
 - (b) If you solved the part a) correctly, you find that even with a positive test an individual is still unlikely to have the disease. For this reason, it is often more useful to test only the risk groups, where the disease is more common. Assume that in a risk group the occurrence of the disease is 40%. Check that in this case the probability for a positively-tested individual to have the disease is about 93%, i.e. now the test is a nearly certain indicator of the disease.