Homework 8

April 9th
Due April 16th

Your assignment may be handwritten or typeset, but in any case it should be neat and readable. Your name and assignment number should be clearly visible (like on this document for example). It is due on Gradescope.

You are encouraged to work in groups for this assignment. However, the redaction should be done on one's own: do not copy some other student's work, or give your assignment to some other student. To consult textbooks or online resources is fair game; on the other hand, to look up the exact exercise and its solution is not. I will be available at my office hours[1] if you need help.

Exercise 1

One Summer night, you decide to stay up to try and see shooting stars. Let us call N the number of shooting stars you see during your watch.

1. What type of distribution do you think would be a good model for N?
2. There is a 0.6% chance of you not seeing any shooting star during your watch.[2] What is the expected number of shooting stars you will see?

Exercise 2

We say that X and Y are uncorrelated if $\mathrm{Cov}(X, Y) = 0$.

Suppose that X and Y are uncorrelated Bernoulli variables. Show that

$$\mathbb{P}(X = 1 \text{ and } Y = 0) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 0).$$

Hint: First, show the same relation with $Y = 1$ rather than $Y = 0$ everywhere.

We can actually show that X and Y are independent. Even though it is not true that uncorrelated variables are independent, see previous homework, it is the case when we consider only Bernoulli variables.

[1] Monday 2:30-3:30, Thursday 10:30-11:30, or by appointment.
[2] Somewhat accurate for one hour of a Summer night.

Exercise 3

A random variable X takes values -2, -1, 1 and 2 with respective probabilities p, $1/2 - p$, $1/2 - p$ and p. A random variable Y takes values -1 and 1 with probabilities q and $1 - q$.

1. What are the possible values for p and q?
2. For which value(s) of p is the variance of X minimal?
3. For which value(s) of q is the variance of Y maximal?

Exercise 4

Let N be a random variable of distribution $\mathcal{G}eom(1/3)$.

1. Give an upper bound for $\mathbb{P}(N \geq 20)$ using Markov's inequality.
2. Give an upper bound for $\mathbb{P}(N \geq 20)$ using Chebyshev's inequality.
3. Recall that $\sum_{k=0}^{\infty} x^k = 1/(1-x)$ whenever $|x| < 1$. For $0 < \alpha < 3/2$, show that
$$\mathbb{E}[\alpha^N] = \frac{\alpha}{3 - 2\alpha}.$$
4. Show that we have
$$\mathbb{P}(N \geq 20) \leq 0.8\%.$$
Hint: $\alpha^{-19}/(3 - 2\alpha) \leq 0.8\%$ for $\alpha = 1.425$.
5. What is the actual value of $\mathbb{P}(N \geq 20)$?

Exercise 5

The annual cost for repairs on a car, in dollars, is a random variable X with expectation 2000 and variance 500,000.

1. Suppose you have a yearly budget of \$3000 for repairs. Using Chebyshev's inequality, give an upper bound for the probability that the repairs will go over budget.
2. This probability seems too high for you. Instead, you decide to gather 100 people to chip in \$2500 each. Assuming the cost for repairs is independent for each person, use Chebyshev's inequality to give an upper bound for the probability that the collective repairs will go over the shared budget.
3. There is no way to know the actual probabilities if we do not know the distribution of X. Nevertheless, based on the estimates we have, which of the above strategies seems to be the safest?