Homework 8
Solution

April 16th

Exercise 1

1. We are doing an experiment continuously and independently as time goes on, and counting the number of successes. This is a Poisson experiment.
2. Let N be the number of shooting stars we see during the (say one hour) watch. According to the previous question, it has distribution $\mathcal{P}oi(\lambda)$. The probability to not see any shooting stars is
$$\mathbb{P}(N = 0) = \frac{\mathsf{e}^{-\lambda}\lambda^0}{0!} = \mathsf{e}^{-\lambda}.$$
According to the exercise, this probability is 0.6%, so $\lambda = -\ln(0.006) \approx 5.1$.

   Now the expected number of shooting stars you will see is
$$\mathbb{E}[N] = \lambda \approx 5.1.$$

Exercise 2

Suppose the covariance is zero. This means exactly that
$$0 = \mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y].$$
Since X, Y and XY can only take the value 0 or 1, we know that
$$\mathbb{E}[X] = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) = \mathbb{P}(X = 1),$$
and this holds for Y and XY as well. In particular,
$$0 = \mathrm{Cov}(X, Y) = \mathbb{P}(XY = 1) - \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1).$$
But $XY = 1$ if and only if X and Y are both 1. This means
$$\mathbb{P}(X = 1 \text{ and } Y = 1) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1).$$
At this point, we have the result suggested by the hint. Now we need to replace $Y = 1$ by $Y = 0$.

The idea is that we want to express $\mathbb{P}(X = 1 \text{ and } Y = 0)$ using $\mathbb{P}(X = 1 \text{ and } Y = 1)$, then it flows fairly naturally.
$$\begin{aligned}\mathbb{P}(X = 1 \text{ and } Y = 0) &= \mathbb{P}(X = 1) - \mathbb{P}(X = 1 \text{ and } Y = 1) = \mathbb{P}(X = 1) - \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1)\\ &= \mathbb{P}(X = 1) \cdot \big(1 - \mathbb{P}(Y = 1)\big) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 0)\end{aligned}$$

Exercise 3

1. The probabilities all sum to one, so there is no constraint there. Since we need every probability to be non-negative, this means $0 \le p \le 1/2$ and $0 \le q \le 1$.

2. The variance of X is $\mathbb{E}[X^2] - \mathbb{E}[X]^2$.

$$\mathbb{E}[X] = -2 \cdot p + (-1) \cdot (1/2 - p) + 1 \cdot (1/2 - p) + 2 \cdot p = 0,$$
$$\mathbb{E}[X^2] = (-2)^2 \cdot p + (-1)^2 \cdot (1/2 - p) + 1^2 \cdot (1/2 - p) + 2^2 \cdot p = 8p + (1 - 2p) = 1 + 6p,$$
$$\mathrm{Var}(X) = (1 + 6p) - 0^2 = 1 + 6p.$$

   The variance is minimal for p minimal, i.e. $p = 0$.

3. Similarly,

$$\mathbb{E}[Y] = (-1) \cdot q + 1 \cdot (1 - q) = 1 - 2q,$$
$$\mathbb{E}[Y^2] = (-1)^2 \cdot q + 1^2 \cdot (1 - q) = 1,$$
$$\mathrm{Var}(Y) = 1 - (1 - 2q)^2.$$

   This is a quadratic term in "completed square" form; its maximum is 1, attained for $1-2q = 0$, i.e. $q = 1/2$. Since this is a possible value for q, it is the maximal possible variance.

   Just for fun, there is another way to find the variance. Since Y takes only two values, we can tweak it a bit to make it Bernoulli: $Z = (Y + 1)/2$ takes the value 1 with probability $1 - q$, and 0 with probability q. Hence

$$\mathrm{Var}(Y) = \mathrm{Var}(2Z - 1) = \mathrm{Var}(2Z) = 2^2\mathrm{Var}(Z) = 4 \cdot (1 - q)\big(1 - (1 - q)\big) = 4q(1 - q).$$

   One can check that it is the same result as the one we found before.

Exercise 4

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

1.

$$\mathbb{P}(N \ge 20) \le \frac{\mathbb{E}[N]}{20} = \frac{3}{20} = 15\%.$$

2. We know that $\mathbb{E}[N] = 5$ and $\mathrm{Var}(N) = 6$. If $N \ge 20$, then $|N - 5| \ge 15$, so

$$\mathbb{P}(N \ge 20) \le \mathbb{P}(|N - 5| \ge 15) = \mathbb{P}(|N - 5| \ge \sqrt{75/2} \cdot \sqrt{6}) \le \frac{1}{75/2} \approx 2.7\%.$$

3.

$$\mathbb{E}\left[\alpha^N\right] = \sum_{k=1}^{\infty} \alpha^k \cdot \mathbb{P}(N = k) = \sum_{k=1}^{\infty} \alpha^k (1 - 1/3)^{k-1}(1/3) = \frac{1}{3} \cdot \alpha \sum_{k=0}^{\infty} \underbrace{\alpha^k \left(\frac{2}{3}\right)^k}_{=\left(\frac{2\alpha}{3}\right)^k}$$
$$= \frac{\alpha}{3} \cdot \frac{1}{1 - 2\alpha/3} = \frac{\alpha}{3 - 2\alpha}.$$

4. Using Markov's inequality, we get

$$\mathbb{P}(N \geq 20) = \mathbb{P}\big(\alpha^N \geq \alpha^{20}\big) \leq \frac{\mathbb{E}\big[\alpha^N\big]}{\alpha^{20}} = \frac{\alpha^{-19}}{3-2\alpha}.$$

   This is true for any α between 0 and 3/2, so we can use $\alpha = 1.425$ and get $\mathbb{P}(N \geq 20) \leq 0.8\%$. The example $\alpha = 1.425$ was chosen becase it corresponds to the smallest value of $\alpha^{-19}/(3-2\alpha)$.

5. N is a geometric random variable, so it corresponds to the first success in a series of independent experiments. The fact that $N \geq 20$ corresponds to the first 19 experiments being failures, which has probability $(2/3)^{19} \approx 0.054\%$ because the experiments are independent.

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. Call X the cost for repairs, so that $\mathbb{E}[X] = 2,000$ and $\mathrm{Var}(X) = 500,000$.

$$\mathbb{P}(X \geq 3000) \leq \mathbb{P}(|N-2000| \geq 1000) \leq \frac{1}{(1000/\sqrt{\mathrm{Var}(X)})^2} = \frac{500,000}{1,000,000} = 50\%.$$

2. Call X_i the cost for the ith person, so the total cost is $X = X_1 + \cdots + X_{100}$. Because the expectation is linear, we have $\mathbb{E}[X] = 100 \cdot 2,000$, and because the variables are independent, we have $\mathrm{Var}(X) = 100 \cdot 500,000$ (otherwise we would have to deal with covariances). Now the total budget is $250,000$ and

$$\mathbb{P}(X \geq 250,000) \leq \mathbb{P}(|N-200,000| \geq 50,000) \leq \frac{1}{(50,000/\sqrt{\mathrm{Var}(X)})^2} = \frac{50,000,000}{2,500,000,000} = 2\%.$$

3. Strategy 2 seems the safest, even though the individual cost is lower. Of course this is why insurance companies exist, and why it is always easier to be rich.