Week 7

Exercise 67 Suppose we have a coin which lands on heads with probability p. Suppose that we toss the coin a lot of times and that each toss is independent.

(1) What is the probability that the fifth head appears on the 12th toss?

(2) What is the probability that the same number of heads appear in the first eight tosses as in the next five tosses?

Answer. (1) This is asking that we have four heads show up on the first eleven coin tosses and then a head on the 12th. The chance of getting four heads in the first eleven coin tosses is the binomial distribution

$$\binom{11}{4} p^4 (1-p)^7$$

and the chance of getting a heads on the 12th flip is p. Since these are independent, we can just multiply them giving:

$$\binom{11}{4} p^4 (1-p)^7 \cdot p$$

(2) Since these are independent events (the eight tosses don't influence the five tosses), we can multiply their probabilities. We're asking for the same number of heads to appear and since five is the lesser of the two numbers, we can have at most 5 heads. Otherwise, the probability is 0 (We can't get six heads in five tosses). In order to calculate the chance that they're the same, we need to look at the chance that they both have 0 heads, 1 heads, 2 heads, etc. So if we let X be the number of heads in eight tosses and Y be the number of heads in five tosses we have:

$$\sum_{k=0}^{5} P(X=k, Y=k) = \sum_{k=0}^{5} \binom{8}{k} p^k (1-p)^{8-k} \cdot \binom{5}{k} p^k (1-p)^{5-k} = \sum_{k=0}^{5} \binom{8}{k}\binom{5}{k} p^{2k} (1-p)^{13-2k}$$

□

Exercise 68 Suppose that we have two independent geometric random variables X_1 and X_2. Let X_1 have parameter p_1 and X_2 have parameter p_2. Find the following:

(1) $P(X_1 = X_2)$

(2) $P(X_1 < X_2)$

(For the second one, you can leave it in a summation.)

Answer. (1)

$$\begin{aligned} P(X_1 = X_2) &= \sum_{k=1}^{\infty} P(X_1 = k, X_2 = k) \\ &= \sum_{k=1}^{\infty} P(X_1 = k) P(X_2 = k) \qquad \text{(independence)} \\ &= \sum_{k=1}^{\infty} (1-p_1)^{k-1} p_1 (1-p_2)^{k-1} p_2 \\ &= \frac{p_1 p_2}{1 - (1-p_1)(1-p_2)} \qquad \text{(geometric series)} \\ &= \frac{p_1 p_2}{p_1 + p_2 - p_1 p_2} \end{aligned}$$

(2)

$$\begin{aligned} P(X_1 < X_2) &= \sum_{k=1}^{\infty} P(X_1 = k, X_1 < X_2) \\ &= \sum_{k=1}^{\infty} P(X_1 = k, k < X_2) \\ &= \sum_{k=1}^{\infty} P(X_1 = k) \cdot P(k < X_2) \qquad \text{(independence)} \\ &= \sum_{k=1}^{\infty} (1-p_1)^{k-1} p_1 (1-p_2)^k \end{aligned}$$

We can't go much further than here, so it's an ok spot to leave this.

□

Exercise 69 Suppose you're playing volleyball where the first team to get two points more than their opponents is declared the winner. Let G be the total number of games you played where each game is independent of all other games. Assume that your team ends up winning each game with probability p.

(1) What is the probability that the total number of games played is n?

(2) Find the expected value of G.

(3) Find the variance of G.

Answer. We can break down the games played into pairs. If at any point any team wins a pair of games, they are declared the winners. If we play n games, that means there are $\frac{n}{2}$ total pairs. Therefore, in total we know that n must be even.

(1) Since n must be even we know $P(G = n) = 0$ when n is odd. For the first $n-2$ games we know that our team won half the time, and the other team won half the time. In fact for every two games played, we must have won once and they must have won once. Therefore, if we look at pairs, there are $\frac{n-2}{2}$ pairs and the order between us winning and the other team winning doesn't matter so we get:

$$(2p(1-p))^{\frac{n-2}{2}}$$

for the probability.

The final two games must have been won by either our team or the other team. Since we're using "or", we know that we want to sum these up. And we get

$$p^2 + (1-p)^2$$

Since our games are independent, we multiply these together to get:

$$P(G = n) = (p^2 + (1-p)^2) \cdot (2p(1-p))^{\frac{n-2}{2}}$$

(When n is even.)

(2) Let $G = 2X$ where X is the random variable for the number of pairs of games played. The chance of a team winning has the geometric distribution $\left(p^2 + (1-p)^2\right)$. (You can see this by: $p^2+(1-p)^2+(2p(1-p)) = 1$) Then

$$E(G) = 2E(X) = \frac{2}{p^2 + (1-p)^2}$$

(See Example 36.1 for why this is true)

(3) The variance is then given by

$$\operatorname{Var}(G) = 4\operatorname{Var}(X) = \frac{8p(1-p)}{(p^2+(1-p)^2)^2}$$

(See Example 3.61 for why this is true)

□

Exercise 70 Suppose we're making chocolate chip cookies and we want the probability of a cookie containing at least 1 chocolate chip to be at least 99%. How many chocolate chips must a cookie contain on average?

Answer. Let μ be the average number of chocolate chips in a cookie. As we're looking for chocolate chips in a unit area, we're using Poisson distribution here. Therefore, "probability of X chocolate chip" translates to $P(X)$ and so we are looking for $P(X \geq 1) \geq 0.99$. So we have:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - e^{-\mu}\frac{\mu^0}{\mu!} \\ &= 1 - e^{-\mu} \\ &\geq 0.99 \end{aligned}$$

Solving for μ we get $\mu \geq -\ln(0.01) \approx 4.6$. So we want an average of 5 chocolate chips per cookie. □

Exercise 71 You've been hired by Aram to make sure his exercises don't have errors. On average there is 1 error per page. What is the probability that after Aram has written all 300 pages of his exercise sheets there will be at least one page that has at least 5 errors on it? (Assume Poisson distribution and that each page having errors is independent.)

Answer. We're told we have Poisson distribution, so that makes life easier. We know the average is 1 so $\mu = 1$. We want to know what the probability that any one page has at least five errors. Let X be the number of errors on any given page. To calculate this we use Poisson, which gives us:

$$\begin{aligned} P(X \geq 5) &= 1 - \sum_{i=0}^{4} e^{-\mu}\frac{\mu^i}{i!} \\ &= 1 - e^{-1}\sum_{i=0}^{4}\frac{1}{i!} \\ &= 1 - e^{-1}\left(\frac{1}{1}+\frac{1}{1}+\frac{1}{2}+\frac{1}{6}+\frac{1}{24}\right) \\ &= 1 - e^{-1}\cdot 2.708\overline{3} \\ &= 1 - 0.99634 \\ &= 0.00366 \end{aligned}$$

Let Y be the number of pages which contain at least five errors. To calculate when this happens on at least one page, we need to look at:

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - e^{-\mu'}$$

Here μ' is the average number for Y which is given by np. In this case, since there's 300 pages, $n = 300$, and since the probability for any page having at least 5 errors is 0.00366, we have $p = 0.00366$. Therefore $\mu' = np = 0.00366 \cdot 300 = 1.097954$.

$$P(Y \geq 1) = 1 - e^{-1.097954} = 1 - 0.33355282 = 0.6664472$$

(Note that $P(Y)$ is technically a binomial distribution. The above is giving a Poisson approximation.) □

Exercise 72 Suppose it's a snowy day and roughly 30 snowflakes are falling per square inch per minute. What is the chance that any particular square inch is *not* hit by any snowflakes during a given 10 second period? (Use Poisson distribution)

Answer. We know the average is 30 snowflakes per minute. Since we want to know within a 10 second period, we need to calculate the average in a 10 second period. Since expectation is linear (we can just add things) we can do $30/6 = 5$ in order to find the average in a 10 second period. Let X be the number of snowflakes that hit any square inch. We want to calculate:

$$P(X = 0) = e^{-\mu}\frac{\mu^0}{0!} = e^{-5} = 0.00674$$

□

Exercise 73 We're on a class field trip to Chernobyl! In order to be safe, we brought a Geiger counter to make sure we're not going to die from radiation. We notice that we receive roughly 10 pulses per minute. What is the probability that we receive 3 pulses in any given half-minute period? (Use Poisson distribution)

Answer. Let X be the Poisson random variable which counts the number of pulses in a given half-minute period. To calculate this, we need to find the parameter (aka the average). We know there's 10 pulses per minute, which means 5 pulses for every half-minute period. Plugging in, we get:

$$P(X = 3) = e^{-\mu}\frac{\mu^3}{3!} = e^{-5}\frac{5^3}{3!} = 0.00674 \cdot \frac{125}{6} = 0.1404$$

□

Exercise 74 Instead of Chernobyl, pretend we're in a laboratory where we're looking into radioactive substances (safely). We know that radioactive substances release particles known as α-particles. We set-up a counter (like a Geiger counter) to see how many α-particles are given off in a time period. Suppose that we have two different substances and they are emitting α-particles independently of one another. The first substance has a Poisson distribution with parameter 3.87 and the second substance has a Poisson distribution with parameter 5.41 (in a given time frame). What is the probability that the counter is hit by at most 4 α-particles (based on the time frames given by the distributions)?

Answer. Since the distributions are over the same time frame, we can just sum up their averages. In other words, the global average is $3.87 + 5.41 = 9.28$. Let X be the number of α-particles we count. Then

$$\begin{aligned}
P(X \leq 4) &= P(0) + P(1) + P(2) + P(3) + P(4) \\
&= e^{-\mu}\frac{\mu^0}{0!} + e^{-\mu}\frac{\mu^1}{1!} + e^{-\mu}\frac{\mu^2}{2!} + e^{-\mu}\frac{\mu^3}{3!} + e^{-\mu}\frac{\mu^4}{4!} \\
&= e^{-9.28} + e^{-9.28} \cdot 9.28 + e^{-9.28} \cdot \frac{9.28^2}{2} + e^{-9.28}\frac{9.28^3}{6} + e^{-9.28} \cdot \frac{9.28^4}{24} \\
&= e^{-9.28}(1 + 9.28 + 43.0592 + 133.19646 + 309.01578) \\
&= e^{-9.28} \cdot 495.5514 \\
&= 0.04622
\end{aligned}$$

□

Exercise 75 Suppose that X has density function $f(x) = cx^2(1-x)^2$ for $0 < x < 1$ and is equal to 0 everywhere else.

(1) What is the value of c?

(2) What is the expected value of X?

(3) What is the variance of X?

Answer. (1) We know that the are under the curve must be equal to 1. So we have

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,\mathrm{dx} &= \int_0^1 cx^2(1-x)^2\,\mathrm{dx}\\ &= c\int_0^1 x^2 - 2x^3 + x^4\,\mathrm{dx}\\ &= c\,\frac{x^3}{3} - \frac{2x^4}{4} + \frac{x^5}{5}\Big]_0^1\\ &= c\left(\frac{1}{3} - \frac{2}{4} + \frac{1}{5}\right)\\ &= c\left(\frac{1}{30}\right)\\ &\Rightarrow\\ 1 &= \frac{c}{30}\\ c &= 30\end{aligned}$$

(2) We want to calculate the expected value:

$$\begin{aligned}E(X) &= \int_0^1 x\cdot 30x^2(1-x)^2\,\mathrm{dx}\\ &= 30\int_0^1 x^3 - 2x^4 + x^5\,\mathrm{dx}\\ &= 30\,\frac{x^4}{4} - \frac{2x^5}{5} + \frac{x^6}{6}\Big]_0^1\\ &= 30\left(\frac{1}{4} - \frac{2}{5} + \frac{1}{6}\right)\\ &= 30\left(\frac{1}{60}\right)\\ &= \frac{1}{2}\end{aligned}$$

(3) We want to calculate $\mathrm{Var}(X) = E(X^2) - E(X)^2$. So we need to find $E(X^2)$ first.

$$\begin{aligned}E(X^2) &= \int_0^1 x^2\cdot 30x^2(1-x)^2\,\mathrm{dx}\\ &= 30\int_0^1 x^4 - 2x^5 + x^6\,\mathrm{dx}\\ &= 30\,\frac{x^5}{5} - \frac{2x^6}{6} + \frac{x^7}{7}\Big]_0^1\\ &= 30\left(\frac{1}{5} - \frac{2}{6} + \frac{1}{7}\right)\\ &= 30\left(\frac{1}{105}\right)\\ &= \frac{2}{7}\end{aligned}$$

So we get

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \frac{2}{7} - \left(\frac{1}{2}\right)^2 = \frac{2}{7} - \frac{1}{4} = \frac{1}{28} = 0.0357142857$$