
Note: You **must** show all your work. Numerical answers without a proper explanation or a clearly written down path to the solution will be assigned zero points.

Problem 7.1. (5 points) The number of cars one sees passing by the local playground in an afternoon is modeled using a Poisson distribution with mean 25. The proportion of black cars in the stream is 1/5. The color of the cars is independent of the number of cars that drive by. What is the probability that exactly 5 black cars and exactly 10 non-black cars drive by in a particular afternoon?

Solution: Let N_1 denote the number of black cars and let N_2 denote the number of non-black cars. By the thinning theorem,

$$N_1 \sim Poisson(\lambda_1 = 5) \quad \text{and} \quad N_2 \sim Poisson(\lambda_2 = 20)$$

with N_1 and N_2 independent. Hence, the probability we are looking for is

$$\mathbb{P}[N_1 = 5, N_2 = 10] = \mathbb{P}[N_1 = 5]\mathbb{P}[N_2 = 10] = \left(e^{-5} \cdot \frac{5^5}{5!}\right)\left(e^{-20} \cdot \frac{20^{10}}{10!}\right) = 0.00102.$$

Problem 7.2. (5 points) Let the number of customers N who walk into Hooper's store on a given day be Poisson with variance 30. The probability that a particular customer is a monster is 2/3. The number of customers is independent from whether they are monsters or not.

What is the probability that the total number of customers in a particular day is 25, **given** that the number of monster-customers equals 12?

Solution: Due to thinning, with N_1 denoting the number of customers who are monsters, and N_2 denoting the number of customers who are not monsters, we have that

$$N_1 \sim Poisson(\lambda_1 = 30(2/3) = 20) \quad \text{and} \quad N_2 \sim Poisson(\lambda_2 = 30(1/3) = 10)$$

with N_1 and N_2 independent. So, the conditional probability we are looking for is

$$\begin{aligned}\mathbb{P}[N = 25 \mid N_1 = 12] &= \frac{\mathbb{P}[N = 25, N_1 = 12]}{\mathbb{P}[N_1 = 12]} = \frac{\mathbb{P}[N_1 + N_2 = 25, N_1 = 12]}{\mathbb{P}[N_1 = 12]} \\ &= \frac{\mathbb{P}[12 + N_2 = 25, N_1 = 12]}{\mathbb{P}[N_1 = 12]} = \frac{\mathbb{P}[N_2 = 13, N_1 = 12]}{\mathbb{P}[N_1 = 12]}.\end{aligned}$$

Recall that N_1 and N_2 are independent. So, the above probability equals

$$\frac{\mathbb{P}[N_2 = 13, N_1 = 12]}{\mathbb{P}[N_1 = 12]} = \frac{\mathbb{P}[N_2 = 13]\,\mathbb{P}[N_1 = 12]}{\mathbb{P}[N_1 = 12]} = \mathbb{P}[N_2 = 13] = e^{-10} \cdot \frac{10^{13}}{13!}.$$

Problem 7.3. (5 points) In a large population, the purple party and the mauve party are facing off in a two-party election. You are surveying people exiting from a polling booth and asking them if they voted purple. The probability that a randomly chosen person voted purple is 20%. What is the probability that exactly 15 people must be asked before you can find exactly 5 people who voted purple?

Solution: This is a case where the negative binomial distribution is appropriate. The parameters r and β are chosen so that r counts the required number of successes 5 (i.e., the people voting for the purple party) and β satisfies that $\frac{1}{1+\beta}$ is the probability of success $20\% = 1/5$ (i.e., the probability that a person chosen at random votes purple). So, the total number of failures before the 5^{th} success in encountered is

$$N \sim NegBinomial(r = 5, \beta = 4).$$

Since we are looking for the probability that the total number of surveyed is equal to 15, this means that we are looking for the probability that the number of failures (i.e., the ones voting for mauve) is 10. Using our STAM tables, we get

$$\mathbb{P}[N = 10] = \binom{14}{4}\left(\frac{1}{5}\right)^5\left(\frac{4}{5}\right)^{10} = \frac{14 \cdot 13 \cdot 12 \cdot 11}{4 \cdot 3 \cdot 2} \cdot \frac{4^{10}}{5^{15}} = 0.034394.$$

Problem 7.4. (20 points) A (6-sided) die is thrown and the number shown is written down (call it X). After that, a biased coin with the probability of *heads* equal to $1/(X + 1)$ is tossed until the first *heads* appears.

i. (10 points) Compute the probability mass function for, as well as the expected of, the number of tosses.
ii. (10 points) Suppose that the number of tosses it took to get *heads* was observed, and it turned out to be equal to 5. The number on the die, on the other hand, is not known. What is the most likely number on the die?

Solution: Let Y be the number of tosses until the first head. Conditionally on X, Y is geometrically distributed with success probability $p(X) = \frac{1}{1+X}$.

(1) The set of the values Y can take is $\mathbb{N}_0 = \{0, 1, \dots\}$, so the pmf is given by the sequence $p_k = \mathbb{P}[Y = k]$, $k \in \mathbb{N}_0$. By the formula of total probability

$$\mathbb{P}[A] = \sum_{i=1}^{n} \mathbb{P}[A|B_i]\mathbb{P}[B_i],$$

where $\{B_i\}_{i=1,\dots,n}$ is a partition of Ω, we have for $k \in \mathbb{N}_0$,

$$\begin{aligned} p_k &= \sum_{i=1}^{6} \mathbb{P}[Y = k|X = i]\mathbb{P}[X = i] = \sum_{i=1}^{6}(1 - p(i))^{k-1}p(i)\frac{1}{6} \\ &= \frac{1}{6}\sum_{i=1}^{6} \frac{1}{1+i}\left(\frac{i}{i+1}\right)^{k-1} \end{aligned}$$

The expectation is given by $\mathbb{E}[Y] = \sum_{k=0}^{\infty} kp_k$, but this is not a nice computational exercise. For your your information, you could do the following (note the exchange of the sums):

$$\mathbb{E}[Y] = \sum_{k=0}^{\infty} k\frac{1}{6}\sum_{i=1}^{6}\frac{1}{1+i}(\frac{i}{i+1})^{k-1} = \sum_{i=1}^{6}\frac{1}{6}\sum_{k=0}^{\infty} k\frac{1}{1+i}(\frac{i}{i+1})^{k-1}$$

Each inner sum represents the expectation of a geometric random variable (with success probability $1/(1+i)$) plus 1 (for the extra toss for the final success). Therefore,

$$\mathbb{E}[Y] = \sum_{i=1}^{6}\frac{1}{6}(i+1) = \frac{9}{2}.$$

There is a significant chance that folks will be off by one when counting tosses (since we usually count the number of **failures** before the final success). So, both 9/2 and 7/2 are graded as correct as they both exhibit a correct thought process.

(2) For this part, we need to compute the conditional probabilities $\mathbb{P}[X = i|Y = 5]$, for $i = 1, \dots, 6$. We use Bayes' formula: let $\{B_i\}_{i\in 1,\dots,n}$ be a partition of Ω. Then

$$\mathbb{P}[B_i|A] = \frac{\mathbb{P}[B_j \cap A]}{\mathbb{P}[A]} = \frac{\mathbb{P}[B_j]\mathbb{P}[A|B_j]}{\mathbb{P}[A]} = \frac{\mathbb{P}[B_j]\mathbb{P}[A|B_j]}{\sum_{k=1}^{n}\mathbb{P}[A|B_k]\mathbb{P}[B_k]}.$$

In our case $B_i = \{X = i\}$ and $A = \{Y = 5\}$, so

$$r_i = \mathbb{P}[X = i|Y = 5] = \mathbb{P}[Y = 5|X = i]\frac{\mathbb{P}[X = i]}{\mathbb{P}[Y = 5]} = \frac{1}{1+i}(\frac{i}{i+1})^4\frac{\frac{1}{6}}{p_5}.$$

In order to find the most likely i, we need to find the one which maximizes r_i. Since r_i and $\frac{i^4}{(1+i)^5}$ differ only by a multiplicative constant, it suffices to maximize the expression

$$f(i) = \frac{i^4}{(1+i)^5}, \text{ over } i = \{1, 2, \dots, 6 \,:\, .\}$$

You can use 'R' to get the six numbers. I used "Mathematica, for example, to get the following numerical values:

i	1	2	3	4	5	6
$f(i)$	0.031	0.065	0.079	0.082	0.080	0.077

so the most likely die outcome is $i = 4$.

Problem 7.5. (10 points) You repeatedly and independently spin a uniform spinner with three colored regions: yellow, blue, and red. After 1800 spins, you tally the number of times your spinner landed on red.

Independently from the spinner, you randomly extract balls from an urn containing 7 teal and 21 beige balls with replacement. After 1000 extractions, you tally the number of times a teal ball was extracted.

What is the variance of the difference between the number of times your spinner landed on red and the number of times a teal ball was extracted?

Solution: The two tallies are independent binomial random variables. Let the number of times the spinner landed on red be denoted by N_1. Then,

$$N_1 \sim Binomial(m_1 = 1800, q_1 = 1/3).$$

Let the number of times a teal ball was extracted be denoted by N_2. Then,

$$N_2 \sim Binomial(m_2 = 1000, q_2 = 1/4).$$

We are looking for

$$Var[N_1 - N_2].$$

Since the two random variables are independent, we have

$$\begin{aligned} Var[N_1 - N_2] &= Var[N_1] + Var[N_2] = m_1 q_1(1 - q_1) + m_2 q_2(1 - q_2) \\ &= 1800\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + 1000\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = 587.5 \end{aligned}$$

Problem 7.6. (5 points) *Source: Prof. Jim Daniel (personal communication).*
Let the random variable N be in the $(a, b, 0)$ class with $a = b = 3/4$. Find $Var[N]$.

Solution: According to the tables we use in our exams, the only r.v. from the $(a, b, 0)$ class that can satisfy $a = b = 3/4$ is the negative binomial with parameters $r = 2$ and $\beta = 3$. Again, using our tables,

$$Var[N] = r\beta(1 + \beta) = 2 \cdot 3 \cdot 4 = 24.$$