Probability and Statistics: Homework 2

Joel Christoph
Guido Gongioanni
Marius Grünewald
David McCarthy
Patricia Paskov

Natalia Polakova

08.09.2021

1. Exercise 1

   (a) Find the probability mass function (PMF) of X

We calculate the PMF for each event individually and obtain:

$$
\begin{aligned}
P(X=1) &= 1-p_1 \\
P(X=2) &= p_1(1-p_2) \\
P(X=3) &= p_1 * p_2(1-p_3) \\
P(X=4) &= p_1 * p_2 * p_3(1-p_4) \\
P(X=5) &= p_1 * p_2 * p_3 * p_4(1-p_5) \\
P(X=6) &= p_1 * p_2 * p_3 * p_4 * p_5(1-p_6) \\
P(X=7) &= p_1 * p_2 * p_3 * p_4 * p_5 * p_6(1-p_7) \\
P(X=8) &= p_1 * p_2 * p_3 * p_4 * p_5 * p_6 * p_7
\end{aligned}
$$

   (b) Show that it is a valid PMF

A valid PMF must satisfy two conditions. Its sum must equal one and each element $P(X=x)$ needs to be strictly non-negative.
We know that each probability for reaching any following level is strictly positive in a sense that $1 \geq (P(X=x) \geq 0$. This implies that $1 \geq 1 - P(X=x) \geq 0$ due to the law of total probability. Further, $1 \geq P(X=x_1)P(X=x_2) \geq 0$ has to hold. Therefore, each instance in the PMF is strictly non-negative and bounded by 0 and 1.
To show that the PMF integrates to 1, we add the individual elements of the PMF and show that this equals to 1.

$$(1-p_1) + (p_1(1-p_2)) + p_1p_2(1-p_3) + p_1p_2p_3(1-p_4) + p_1p_2p_3p_4(1-p_5) +$$

$$p_1p_2p_3p_4p_5(1-p_6)+p_1p_2p_3p_4p_5p_6(1-p_7)+p_1p_2p_3p_4p_5p_6p_7$$
$$=1-p_1+p_1-p_1p_2+p_1p_2-p_1p_2p_3+p_1p_2p_3-p_1p_2p_3p_4+p_1p_2p_3p_4$$
$$-p_1p_2p_3p_4p_5+p_1p_2p_3p_4p_5-p_1p_2p_3p_4p_5p_6+p_1p_2p_3p_4p_5p_6-p_1p_2p_3p_4p_5p_6p_7+p_1p_2p_3p_4p_5p_6p_7$$
$$=1$$

(c) Find E(X) and the standard deviation of X

To simplify later notation, we adjust the $P(X=8)$ notation slightly. We define p_8 as being the chance of reaching level 9, which doesn't exist, and therefore has to be equal to zero. This yields
$P(X=8)=p_1p_2p_3p_4p_5p_6p_7(1-p_8)=p_1p_2p_3p_4p_5p_6p_7(1-0)$
With this neutral addition to $P(X=8)$. Further, we amend the vector with $1-P(X=0)=p_0=1$, the probability of entering level 1, which is equal to 1. We do not change the relevant properties of the PMF because the probability of not entering level 1 is 0, thus not distorting the properties. The sum of all events in the PMF is still 1, since $P(X=0)=0$. Further, 0 is non-negative, making it a valid probability. It allows us to rewrite $P(X=1)=(1-p_1)*p_0=(1-p_1)*1$
We can generalize the PMF to

$$(1-p_i)\prod_{j=0}^{i-1}p_j$$
$$\text{where } i=[1,8] \text{ and } j=[0,7].$$

The expected value is defined as the probability of an event multiplied with its payoff. Subsequently, the expectation is defined in terms of p and X as

$$E(X)=\sum_{i=1}^{8}(X=i)(1-p_i)\prod_{j=0}^{i-1}p_j$$

with $(X=i)$ being equal to the highest level reach (e.g. $(X=1)=1$, ..., $(X=8)=8$. The standard deviation is generally stated as

$$SD(X)=\sqrt{VAR(X)}=\sqrt{E(X^2)-(E(X))^2}$$

In our application this means

$$SD(X)=\sqrt{\sum_{i=1}^{8}(X=i)^2(1-p_i)\prod_{j=0}^{i-1}p_j-\left(\sum_{i=1}^{8}(X=i)(1-p_i)\prod_{j=0}^{i-1}p_j\right)^2}$$

(d) Calculate the $E(X)$ and $SD(X)$ for increasingly harder levels.

By running the formulas with the suggested values we obtain an expected value of 3.656 with a standard deviation of 2.903.

(e) Now assume there is a secret passage somewhere in level 2 that takes you to level 8, which you can find with probability x. Without re-doing all of your calculations, find $E(X)$ for any vector p and then for the particular vector given above.

Generally, we assume the following set up. In level 2, there are three distinct option conditional on surviving the first level: regular success p_2, and finding the tunnel x and failure of either $(1-p_2)(1-x)$. By this we assume, that the tunnel is somewhere in the game and strictly prior to the end of the game.
As a result, the probabilities of reaching X as the highest level change accordingly (except for level 1),

$$P(X=1) = 1-p_1,$$
$$P(X=x) = (1-p_i)(1-x)\prod_{j=1}^{x-1} p_j \; \forall x \in (2,3,4,5,6,7)$$
$$\text{and } P(X=8) = xp_1 + (1-x)\prod_{j=1}^{7} p_j$$

The expected value can therefore be written as

$$E(X) = (1-p_1)(X=1) + \sum_{i=2}^{7}(X=i)(1-p_i)(1-x)\prod_{j=1}^{x-1} p_j + (X=8)(xp_1 + (1-x)\prod_{j=1}^{7} p_j)$$

(f) Does the variance of X go up with the introduction of the secret tunnel? No need for calculations, just explain intuitively why that is the case.

The variance will increase. The chances of reaching level increase, putting more weight on the, so far, least likely outcome. Since level 8 is (likely) very far from the mean, it implies that the difference to mean is great and now multiplied by a larger probability. This will increase the variance.

2. Exercise 2

(a) What is the distribution of X, number of eggs hatching?

It is a binomial distribution. More precisely, we specify the distribution of X as

$$X \sim \text{Binom}(n, p_h) = \binom{n}{k} p_h^k (1-p_h)^{n-k}$$

with k being the number of dragons hatched from eggs. This is possible due to the fact that they are i.i.d.

(b) What is the distribution of Y, number of full grown dragons you get in the end?

Each dragon has a probability to grow up of $p_h p_m$. Since dragon are i.i.d. (for example, they do not eat each other), we have a repeated game of survival for each dragon. Therefore, it follows a binomial distribution as follows

$$Y \sim \text{Binom}(n, p_h p_m) = \binom{n}{g} (p_h p_m)^g (1-p_h p_m)^{n-g}$$

where g represents the number of dragons grown up.

(c) Find the probability that you get at least 3 adult dragons.

We calculate the probability of not getting at least 3 adult dragons and subtract it from 1.

$$P(n, p_h p_m) = \binom{n}{2}(p_h p_m)^2(1-p_h p_m)^{n-2}$$
$$P(n, p_h p_m) = \binom{n}{1}(p_h p_m)^1(1-p_h p_m)^{n-1}$$
$$P(n, p_h p_m) = \binom{n}{0}(p_h p_m)^0(1-p_h p_m)^{n-0}$$

This results in the following expression

$$P(\text{at least 3 adult dragons}) =$$
$$1 - \binom{n}{2}(p_h p_m)^2(1-p_h p_m)^{n-2} - n(p_h p_m)(1-p_h p_m)^{n-1} - (1-p_h p_m)^n$$

3. Exercise 3

- **mean**: what is the average wealth of Tuscan households? **median**: what is the level of wealth such that half of the population holds more of it and the other half holds less? **mode**: what is the salary that is earned by the highest percentage of workers?
- If we assume that the idea of an egalitarian society is that at birth each individual has the same chances of extracting the highest or the lowest values of the wealth distribution, this means that the distribution would be a $Unif \sim (0, W)$ where W represents the highest realization of the wealth distribution.
- The Lorenz curve is a plot comparing the empirical CDF of a variable against a hypothetical uniform distribution of that variable. The mean of the distribution appears in the denominator of the Lorenz Curve's formula to rescale the

4. Exercise 4

$$F(x) = \int_{-\infty}^{x} \frac{a}{x^{a+1}} dx = x^{-a} - (-1^{-a}) = 1 - x^{-a} \tag{1}$$

(a) To check that the CDF we obtained is a valid CDF we take the limit with respect to x.

$$\lim_{x \to +1} F(x) = 1 - 1^{-a} = 0 \tag{2}$$

$$\lim_{x \to +\infty} F(x) = 1 - \infty^{-a} = 1 \tag{3}$$

(b) The expected value is

$$\begin{aligned}
\mathbb{E}(x) &= \int_{-\infty}^{x} \frac{a}{x^{a+1}} dx \\
&= a \int_{-\infty}^{x} a x^{-a} dx \\
&= x^{1-a}\Big|_1^{\infty} \\
&= \begin{cases} \dfrac{a}{a-1} & \text{if } a > 1. \\ \infty & \text{if } 0 < a \leq 1. \end{cases}
\end{aligned} \tag{4}$$

and the variance is

$$
\begin{aligned}
Var(x) &= \int_1^\infty (x-\mu)^2 \frac{a}{x^{2+1}} dx \\
&= \int_1^\infty \left(x^2 - 2x\mu + \mu^2\right) f(x) dx \\
&= \int_1^\infty x^2 f(x) dx + \int_1^\infty 2x\mu f(x) dx + \int_1^\infty \mu^2 f(x) dx \\
&= \int_1^\infty x^2 f(x) dx - 2\mu \cdot \mu + \mu^2 \cdot 1 \\
&= \int_1^\infty x^2 \frac{a}{x^{a+1}} dx - \mu^2 \\
&= a \int_1^\infty x^{1-a} dx - \mu^2 \\
&= \frac{x^{2-a}}{2-a}\Big|_1^\infty \\
&= \begin{cases} \dfrac{a}{(a-1)^2(a-2)} & \text{if } a > 2. \\ \infty & \text{if } 0 < a \leq 2. \end{cases}
\end{aligned} \tag{5}
$$

(c) The ffatnessöf Pareto distribution is determined by the parameter α; the kth moment of the distribution will exist only for $k < \alpha$. Thus the smaller is α the ffatterïs the distribution. Hence, Pareto distribution has some excess of kurtosis. In the real world scenario, this means that the wealthiest 1% of population possesses a substantially larger portion of the wealth than would be predicted by extrapolating the distribution of middle income population.

(d) To find 10th, 50th and 90th percentile of the wealth distribution, we express a quantile function as:

$$F^{-1}(p) = \frac{1}{(1-p)^{\frac{1}{a}}} \tag{6}$$

where p is a percentile and $0 \leq p < 1$ and $a = 2$.

$$
\begin{aligned}
F^{-1}(0.1) &= \frac{1}{(1-0.1)^{\frac{1}{2}}} = (\frac{1}{0.9})^{\frac{1}{2}} = 1.05 \\
F^{-1}(0.5) &= \frac{1}{(1-0.1)^{\frac{1}{2}}} = (\frac{1}{0.5})^{\frac{1}{2}} = 1.41 \\
F^{-1}(0.9) &= \frac{1}{(1-0.1)^{\frac{1}{2}}} = (\frac{1}{0.1})^{\frac{1}{2}} = 3.16
\end{aligned}
$$

The 10-90 ratio equals:

$$\frac{F^{-1}(0.1)}{F^{-1}(0.9)} = \frac{1.05}{3.16} = 0.33$$

(e) This is called a "80-20 law", according to which 20% of all people receive 80% of all income, and 20% of the most affluent 20% receive 80% of that 80%, holds when the Pareto index α equals:

$$\alpha = log_4 5 = \frac{log_{10} 5}{log_{10} 4} \approx 1.161$$

(f) Let us denote CDF of Pareto distribution as:

$$y = 1 - x^{-a}$$

We express x as:

$$x = (1 - y)^{-\frac{1}{a}}$$

In order to generate to generate Pareto RVs, we take $y \sim Unif(0, 1)$ and from there make a large number of draws. In this way we can plug the values

5. Exercise 5: The falafel roulette

(a)

$$\begin{aligned} P(\text{at least 1 teaspoon}) =& 1 - P(\text{at most 1 teaspoon}) \\ =& 1 - \text{normalCDF}(\mu - 2\sigma) \\ =& 1 - \text{normalCDF}(3 - 2) \qquad (\mu = 3, \sigma = 1) \\ =& 1 - 0.0228 \\ =& 0.9772 \end{aligned}$$

(b) The spiciness of each falafel is independent and identically distributed (i.i.d.) and there are $3! = 6$ possible permutations of the ranking of spiciness among the three friends, two rankings of which have your falafel as the spiciest. Therefore,

$$P(\text{Your falafel is spiciest}) = \frac{2}{3!} = \frac{2}{6} = \frac{1}{3}$$

Equivalently, let X be the spiciness of your falafel, and Y and Z be the spiciness of your friend 1 and friend 2's falafels, respectively. As the spiciness is i.i.d. distributed, your changes of having a spicier falafel than a friend can be represented by integrating across a uniform from 0 to 1:

$$\begin{aligned} P(\text{Your probability is spiciest}) =& P(X > Y \cap X > Z) \\ =& \int_0^1 \int_0^x \int_0^x \text{dz dy dx} \\ =& \int_0^1 \int_0^x x \text{dy dx} \\ =& \int_0^1 \frac{x^2}{2} \text{dx} \\ =& \frac{1}{3} \end{aligned}$$

(c) The falafel has more than 5 teaspoons in her falafel if the realization of a normally distributed random variable is more than two standard deviations above the mean for a normally distributed variable. We know that roughly 95% of the normal distribution lies within two standard deviations of the mean, leaving roughly 5% at least two standard deviations from the mean. As the normal distribution is symmetric, this exactly half of the roughly 5% of the distribution that is at least two standard deviations of the mean is on the right tail, that is over two standard deviations from the mean. Therefore, the probability that the falafel has spiciness of five teaspoons or more is roughly 2.5%.

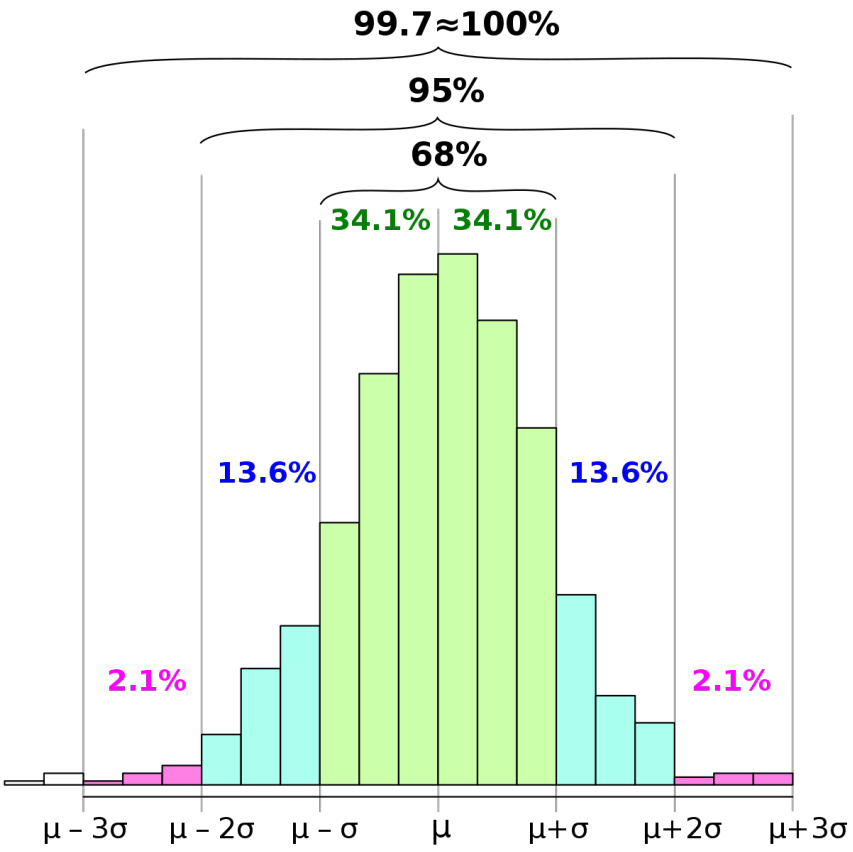

Fig. 1. The 68–95–99.7 empirical rule for the percentage values that lie within 1, 2, or 3 standard deviations of the mean in a normal distribution. Source.

(d) The spiciness for each cook is independently distributed.

$$\begin{aligned}\mathbb{E}[S] =& p\mathbb{E}[S|\text{Cook 1}] + (1-p)\mathbb{E}[S|\text{Cook 2}] \\ =& \frac{1}{2}\mathbb{E}[S|\text{Cook 1}] + \frac{1}{2}\mathbb{E}[S|\text{Cook 2}] \\ =& \frac{1}{2}(3+5) \\ =& 4 \text{ tablespoons}\end{aligned}$$

(e) Your friend can order falafel safely as long as the amount of spice is no more than 5 teaspoons in her falafel. One can expect the spiciness to not exceed five teaspoons as long as:

$$\begin{aligned}\mathbb{E}[S] =& p\mathbb{E}[S|\text{Cook 1}] + (1-p)\mathbb{E}[S|\text{Cook 2}] \\ =& 3p + 5(1-p) \\ =& 5 - 2p\end{aligned}$$

Moreover

$$
\begin{aligned}
5 - 2p \leq & 5 \\
\iff 0 \leq & p
\end{aligned}
$$

Therefore, the expected level of spiciness is below five for any p. For the probability that the falafel has five teaspoons or more to be strictly less than 5%,

$$
\begin{aligned}
P(\text{five teaspoons or more}) =& p \cdot (1 - \Phi(\text{max. 5 tablespoons})) + (1-p) \cdot \frac{1}{2} \\
=& 0.9772p + \frac{1-p}{2} = \frac{1}{2} - 1.0228p
\end{aligned}
$$

6. Exercise 6: Elevator problem

We assume that people choose which of floors 2 3,...., 47 to go to.

(a) Assume for this part only that the probabilities for floors 1, 2, 3,..., 47 are equal. Find the expected number of stops the elevator makes on floors 1, 2, 3,..., 47.

k people decide on which floor they will get off. The probability of person j not getting off on floor i, $i \in \{2..., 47\}$, is equal to $\big(1-\text{P(getting off on floor i)}\big) = \Big(1 - \frac{1}{46}\Big) = \Big(\frac{45}{46}\Big)$.

The probability of k people not getting off on floor i. is equal to $\big(1-\text{P(getting off on floor i)}\big)^k = \Big(1 - \frac{1}{46}\Big)^k = \Big(\frac{45}{46}\Big)^k$ because they independently decide which floor to go to.

So that means that the P(at least someone getting off on floor i) = 1 - P(no one getting off on floor i) = 1 - $\Big(\frac{45}{46}\Big)^k$.

There are 46 floors where the elevator can possibly stop, and for every floor there is the same probability of at least 1 person getting off so the expected number of stops is equal to $46\left(1 - \Big(\frac{45}{46}\Big)^k\right)$

(b) Generalize (a) to the case that floors 1, 2, 3,..., 47 have probabilities $p_1, p_2, p_3, ..., p_{47}$ (respectively); you can leave your answer as a finite sum.

p_1 is the probability that person j gets of at floor 2. So, the probability of person j not getting of on floor 2 is (1-p_2) or more generally for floor i, $i \in \{2..., 47\}$, is equal to (1-p_i). Prob that among k people at least. 1 person gets off the elevator is: 1-P(no one gets off) = $\left(1 - \Big(1 - p_i\Big)^k\right)$. So, the expected number of stops in this case is: $\sum_{i=1}^{46}\left(1 - \Big(1 - p_i\Big)^k\right)$

(c) Back to the case where the probabilities for all floors are equal. Assume you are in a hurry because it is the economist job market and you have an interview in floor 20. If there are no stops, you are there in 10 seconds. However, every time the elevator stops you lose 5 seconds. If your interview is in a minute, and the probability you are late for it as a function of k, the number of people that get in the elevator with you.

If there are no stops, you only lose 10 seconds, and you lose 5 seconds every single time the elevator stops. That means the elevator should not stop more than 10 times before floor

20 in order to be on time at the interview. The probability of not being late is equal to P (max. 10 stops before floor 20). So, the probability of being late is equal to 1- P (max. 10 stops from floor 2 to floor 19).

X $\sim$ Bin(γ,18) where γ is equal to the probability that at least someone gets off at a certain floor. γ, as calculated in (a) is equal to $\left(1 - \left(\frac{45}{46}\right)^k \right)$.

1- P(X $\leqslant$10) = 1 - (P(X=1) + P(X=2) +...+ P(X=10))

$= 1 - \sum_{i=1}^{10} P(X = i) = 1 - \sum_{i=1}^{10} \binom{18}{i} \left(1 - \left(\frac{45}{46}\right)^k\right)^i \left(\left(\frac{45}{46}\right)^k\right)^{18-i}$

7. Exercise 7

Anyway, assume that crossing the tunnel takes 2 minutes walking and that the probability that a car arrives at side i of the tunnel any given second is λ_i for i = 1, 2. The traffic lights work at regular intervals, changing from red to green every 5 minutes.

(a) Suppose you start walking from side 1 of the tunnel when the traffic light changes to red at your end. What is the probability you encounter a car in the tunnel?

It is 2 minute walk (120 seconds) and cars have been accumulating on the other side for the last 5 minutes during the red light (300 seconds). You can only meet cars coming from side 2. P(Encountering a car from side 2) $= 1 - e^{(120+300)\lambda_2}$

(b) Assuming no car comes in your way and you arrive safely at the other side, how many cars would have accumulated at side 1 of the tunnel, on average, as a function of λ_1 when you reach the other end?

Accumulation of cars $\sim Expo(\lambda_1)$. A Poisson distribution has a following PDF: $f(X) = \lambda_1 e^{-\lambda_1 X}$
Therefore, the expected amount of cars adding up in the two minutes is given as $E(X) = \lambda_1$

(c) Suppose the traffic lights are broken. What is the probability you encounter a car in the tunnel? What is the probability that you encounter a car in both directions if you start walking through the tunnel at a random point in time?

The walk takes two minutes with no accumulated cars prior to the start. This means that cars are freely entering the tunnel from either side at any given point in time. Since we don't want to die, we start walking when the tunnel is empty. We assume that if a car enters the tunnel in the 120 seconds during which we are walking, we will encounter it.
We add the distinct probabilities of meeting a car from side 1 and side 2. Since we do not know whether the events are independent, we need to account for the union. We arrive at the following expression:

$$\begin{aligned} &P(\text{Encountering a car from side 1, side 2, or both sides}) = \\ &(1 - e^{-120\lambda_1}) + (1 - e^{-120\lambda_2}) - (1 - e^{-120\lambda_1})(1 - e^{-120\lambda_2}) \end{aligned} \tag{7}$$

To calculate the probability of meeting a car from both sides, we consider the joint probability, which is accounted for in the final part of Equation 6:

$$P(\text{Encountering a car from both sides}) = (1 - e^{-120\lambda_1})(1 - e^{-120\lambda_2})$$

8. Exercise 8

(a) h is called the *hazard function* because is represents the rate of death of an object of given age t, such as the loss of the status of being unemployed. Equivalently, it gives the probability of the unemployment status surviving the job-finding process until time t.
Let the event be the fact that the worker has left unemployment, for instance through finding a job. $h(t)$ is the probability density for leaving unemployment at time t, given that the worker has been unemployed until then, because $h(t)$ is the rate of occurrence of the event after a duration t and equal to

$$\frac{\text{Density of events at } t}{\text{P(surviving up to that duration without experiencing the event}}$$

(b)

(c)