Probability and Statistics
Autumn 2025

MATH-235
Mats J. Stensrud

Problem Sheet 5 [1]

Optional Revision Problems

Exercise 1. People are arriving at a party one at a time. While waiting for more people to arrive they entertain themselves by comparing their birthdays. Let X be the number of people needed to obtain a birthday match, i.e., before the X-th person arrives no two people have the same birthday, but when person X arrives there is a match. Find the PMF of X.

Solution 1. As X denotes the number of people, we only consider positive integers, therefore $P(X = k) = 0$ for $k \notin \mathbb{N}^+$. If only one person arrives clearly there can be no birthday match, so $P(X = 1) = 0$. In contrast, if we have more than 367 people at the party (accounting for birthdays in leap years), then for sure there was a birthday match before the last person arrived, so $P(X = k) = 0$ for $k > 367$, $k \in \mathbb{N}^+$.

Consider $P(X = k)$ for $1 < k \leq 367$. The first person can have their birthday on any day out of the 366, and there will be no birthdays shared. The second one to arrive must have it some other day, so the probability that their birthdays are not shared is 365/366. The third one must not share their birthday either with person 1 or person 2, so they can have it any day out of the remaining 364, meaning that the probability that they have their birthday in one of the "non-taken" days is 364/366. In general, if the j-th person arrives, $j < k$, the probability that they have their birthday on a day that no one else from the previous $j - 1$ people, is $(366 - (j - 1))/366$. Therefore, using the multiplication rule, the probability that all the people up until the $(k-1)$-th had their birthday on a different day is

$$\prod_{i=1}^{k-1} \frac{366 - (i - 1)}{366}.$$

Finally, the k-th arrival must have their birthday shared with one of the previously arrived guests. They have $k - 1$ options for that, so

$$P(X = k) = \prod_{i=1}^{k-1} \frac{366 - (i - 1)}{366} \cdot \frac{k - 1}{366} \text{ for } 2 \leq k \leq 367.$$

See Figure 1 for the PMF

Exercise 2. Let X be an r.v. with CDF F, and $Y = \mu + \sigma X$, where μ and σ are real numbers with $\sigma > 0$. (Then Y is called a location-scale transformation of X.) Find the CDF of Y, in terms of F.

---

[1] Exercises are based on the coursebook Statistics 110: Probability by Joe Blitzstein

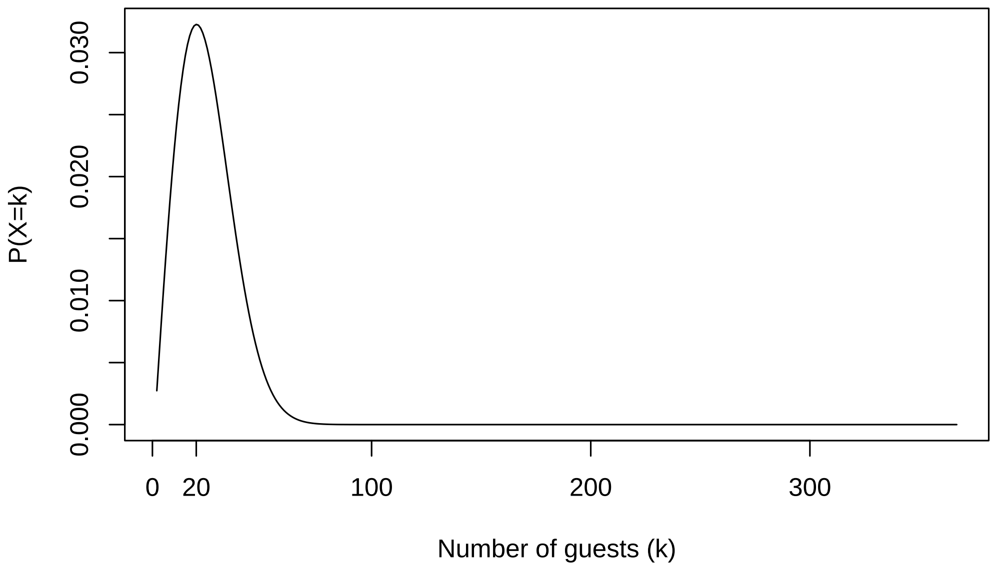

Figure 1: **Disclaimer:** The PMF here is represented with a continuous line, but it is still a discrete random variable (i.e. $P(X = 20.1) = 0$).

Solution 2. The CDF of X is $F(x) = P(X \leq x)$, while the CDF of Y can be written as

$$F_Y(y) = P(Y \leq y) = P(\mu + \sigma X \leq y),$$

by the definition of Y. From simple algebra and from the fact that $\sigma > 0$, it follows that

$$P(\mu + \sigma X \leq y) = P\left(X \leq \frac{y-\mu}{\sigma}\right) = F\left(\frac{y-\mu}{\sigma}\right).$$

Exercise 3.

1. Show that $p(n) = \left(\frac{1}{2}\right)^{n+1}$ for $n = 0, 1, 2, \ldots$ is a valid PMF for a discrete r.v.
2. Find the CDF of a random variable with a PMF from part 1.

Solution 3.

1. The candidate PMF has to satisfy two conditions to be valid: It has to be non-negative and it has to sum to one over the support of the random variable. The former is clearly satisfied, so let's focus on the latter.

   $$\sum_{n=0}^{\infty} p(n) = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^{n+1} = \frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^{n} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}} = 1.$$

   (If the second to last equation is not clear, please consult the Math Appendix of the course book.)

2. For integer-valued random variables, the following equality holds for the CDF:

   $$F_X(x) = P(X \leq x) = P(X \leq \lfloor x \rfloor),$$

hence it is sufficient to find the CDF for $k \in \mathbb{N}$.

$$F_X(k) = \sum_{n=0}^{k} P(X = n) = \sum_{n=0}^{k} \left(\frac{1}{2}\right)^{n+1} = \frac{1}{2}\sum_{n=0}^{k} \left(\frac{1}{2}\right)^{n} = \frac{1}{2} \cdot \frac{1 - \left(\frac{1}{2}\right)^{k+1}}{1 - \frac{1}{2}} = 1 - \left(\frac{1}{2}\right)^{k+1}.$$

(Again, please consult the Math Appendix of the course notes, if a refresher is needed on the sum of geometric series.)

Exercise 4. In a chess tournament, n games are being played, independently. Each game ends in a win for one player with probability 0.4 and ends in a draw (tie) with probability 0.6. Find the PMFs of the number of games ending in a draw, and of the number of players whose games end in draws.

Solution 4. Denote the number of games ending in a draw with D. Since the games are conducted independently, D is a sum of n many independent Bernoulli trials, therefore $D \sim Bin(n, p)$, so

$$P(D = k) = \binom{n}{k} p^k (1 - p)^{n-k} \text{ for } k \in \{0, 1, \ldots, n\},$$

and $P(D = k) = 0$ otherwise.

Denote the number of players whose games resulted in a draw with R. Note that if a game ended in a draw, then exactly two players finished their games with a draw, thus $R = 2D$. Thinking hastily, one could conclude that $R \sim Bin(2n, p)$, but **it is not!** A random variable with the distribution $Bin(2n, p)$ could take any of the values $\{0, 1, \ldots, 2n\}$ with positive probability, however, as the players whose games ended in a draw should be always multiples of 2, R cannot take odd values. Rather $R \sim 2 \cdot Bin(n, p)$, so

$$P(R = k) = \binom{n}{k/2} p^{k/2} (1 - p)^{n-k/2} \text{ for } k \in \{0, 2, 4, \ldots, 2n\},$$

and $P(R = k) = 0$ otherwise.

Week 5 exercises

Exercise 5. 1. A fair die is rolled. Find the expected value of the roll.

2. Four fair dice are rolled. Find the expected total of the rolls.

Solution 5. 1. Denote the outcome of the roll with X, then we are interested in the quantity $E(X)$. Since this is a 6-faced fair die, each of the faces have the same probability, 1/6, i.e. $P(X = x) = \frac{1}{6}$ for $x \in \{1, \ldots, 6\}$ and 0 otherwise. Then by the definition of expectation,

$$\begin{aligned} E(X) &= \sum_{x=1}^{6} P(X = x) \cdot x \\ &= \frac{1}{6}\sum_{x=1}^{6} x = \frac{21}{6} = \frac{7}{2}. \end{aligned}$$

2. Denote the total of the rolls (i.e. the sum of the rolls) with T, and the roll of each of the four dice with X_1, X_2, X_3, and X_4, respectively. Then by the linearity of expectation and by Part 1.

$$\begin{aligned} E(T) &= E(X_1 + X_2 + X_3 + X_4) \\ &= E(X_1) + E(X_2) + E(X_3) + E(X_4) \\ &= 4 \cdot \frac{7}{2} = 14. \end{aligned}$$

Exercise 6. Consider the St. Petersburg paradox (Example 4.3.14 in the course book), except that you receive \$n rather than $\$2^n$ if the game lasts for n rounds. What is the fair value (i.e. the amount a rational player would be willing to offer to play this and the amount that a rational bettor would be willing to accept) of this game? What if the payoff is $\$n^2$?

Hint: Try to use the trick from Example 4.3.6

Solution 6. Let X be the total winnings from the game, then the fair value corresponds to $E(X)$, as if the player offers $E(X)$ to the bettor, then both of their expected gains (and losses, as it is a zero-sum game) are 0.

Suppose that the payout is $\$n$ if the game lasted n rounds. Then $E(X) = \sum_{n=1}^{\infty} n \cdot \left(\frac{1}{2}\right)^n$, as we have a fair coin, thus a sequence of $n-1$ tails and one head has probability $\frac{1}{2^n}$.

We know (see the Math Appendix) that for $|p| < 1$ the sum is $\sum_{n=1}^{\infty} p^n = \frac{1}{1-p}$. Taking the derivative in p on both sides, we get

$$\sum_{i=1}^{\infty} n \cdot p^{n-1} = \frac{1}{(1-p)^2}.$$

Multiplying both sides with p

$$\sum_{i=1}^{\infty} n \cdot p^{n} = \frac{p}{(1-p)^2},$$

where the left-hand side is exactly the formula for $E(X)$, if we substitute with $p = \frac{1}{2}$ (note that crucially $|1/2| < 1$). Evaluating the right-hand side at $p = 1/2$, we get

$$E(X) = \frac{1/2}{1/4} = 2,$$

that is a steep decrease from the infinite fair value of the original problem.

Now suppose that the payout is modified to $\$n^2$, if the game lasted for n rounds. Thus, the expected payout is $E(X) = \sum_{n=1}^{\infty} n^2 \cdot \left(\frac{1}{2}\right)^n$. Following the previous approach, let us take the derivative of $\sum_{n=1}^{\infty} p^n = \frac{1}{1-p}$ twice in p:

$$\sum_{n=1}^{\infty} n(n-1)p^{n-2} = \frac{2}{(1-p)^3}.$$

By multiplying both sides by p^2, we get

$$\sum_{n=1}^{\infty} (n^2 - n)p^{n} = \frac{2 \cdot p^2}{(1-p)^3},$$

so it follows that

$$
\begin{aligned}
E(X) &= \sum_{n=1}^{\infty} n^2 \cdot p^n = \sum_{n=1}^{\infty} (n^2 - n) \cdot p^n + \sum_{n=1}^{\infty} n \cdot p^n \\
&= \frac{2 \cdot p^2}{(1-p)^3} + \frac{p}{(1-p)^2} = \frac{2 \cdot p^2 + p(1-p)}{(1-p)^3},
\end{aligned}
$$

where in the second line we used our findings from the first part. Using the substitution $p = 1/2$, it follows that the game has the fair value of $\frac{1/2+1/4}{1/8} = 4 + 2 = 6$.

Exercise 7. Are there discrete random variables X and Y such that $E(X) > 100E(Y)$ but Y is greater than X with probability at least 0.99?

Solution 7. Yes. Consider what happens if we make X usually 0, but on rare occasions, X is extremely large (like the outcome of a lottery); Y, on the other hand, can be more moderate. For a simple example, let X be 10^6 with probability 1/100 and 0 with probability 99/100, and let Y be the constant 1 (which is a degenerate r.v.). There are literally infinite examples for a pair of random variables satisfying these properties, so if you have a different answer, that is completely fine.

Exercise 8. Let $X \sim Bin(n, p)$ and $Y \sim NBin(r, p)$. Using a story about a sequence of Bernoulli trials, prove that $P(X < r) = P(Y > n - r)$.

Solution 8. Interpret X as the number of successes out of n many independent Bernoulli(p) trials, and Y as the number of failures in a sequence of independent Bernoulli(p) trials, before the r-th success is reached.

Consider an infinite sequence of independent Bernoulli(p) trials. X just focuses on the first n and counts the number of successes among them, while Y counts the number of failures among all the trials that were performed before the r-th success is reached. A crucial difference is that $P(X \leq n) = 1$, but Y is unbounded ($\nexists M$ such that $P(Y < M) = 1$). The event $\{X < r\}$ corresponds to less than r successes in the first n trials, which means that there were more than $n - r$ failures. On the other hand $\{Y > n - r\}$ describes the event that more than $n - r$ failures occurred before the r-th success was reached, so in the first $(n - r) + r = n$ trials, they must have been more than $n - r$ failures. Therefore, $\{X < r\}$ and $\{Y > n - r\}$ describe the same event, hence they have the same probability.

Exercise 9. A discrete random variable X has the *memoryless property* if it has the distribution $P(X \geq j + k | X \geq j) = P(X \geq k)$ for all nonnegative integers j, k.

1. If X has a memoryless distribution with CDF F and PMF $p_i = P(X = i)$, find an expression for $P(X \geq j + k)$ in terms of $F(j), F(k), p_j, p_k$.
2. Name a discrete distribution which has the memoryless property. Justify your answer with a clear interpretation in words or with a computation.

Solution 9. 1. By the memoryless property,

$$
P(X \geq k) = P(X \geq j + k | X \geq j) = \frac{P(X \geq j + k, X \geq j)}{P(X \geq j)} = \frac{P(X \geq j + k)}{P(X \geq j)},
$$

so

$$\begin{aligned}
&P(X \geq j+k) = P(X \geq j)P(X \geq k) = P((X > j) \cup (X = j))P((X > k) \cup (X = k)) \\
&= (P(X > j) + P(X = j)) \cdot (P(X > k) + P(X = k)) \\
&= (1 - P(X \leq j) + P(X = j)) \cdot (1 - P(X \leq k) + P(X = k)) \\
&= (1 - F(j) + p_j)(1 - F(k) + p_k),
\end{aligned}$$

where we used the axiom of probability for disjoint events and the definition of CDF-s.

2. The *Geometric distribution* is memoryless (in fact, it turns out to be essentially the *only* discrete memoryless distribution!). This follows from the story of the Geometric: consider Bernoulli trials, waiting for the first success (and defining waiting time to be the number of failures before the first success). Say we have already had j failures without a success. Then the additional waiting time from that point forward has the same distribution as the original waiting time (the Bernoulli trials neither are conspiring against the experimenter nor act as if he or she is due for a success: the trials are independent). A calculation agrees: for $X \sim \text{Geom}(p)$,

$$P(X \geq j+k | X \geq j) = \frac{P(X \geq j+k)}{P(X \geq j)} = \frac{q^{j+k}}{q^j} = q^k = P(X \geq k).$$

Exercise 10. A coin with probability p of Heads is flipped n times. The sequence of outcomes can be divided into *runs* (blocks of H's or blocks of T's), e.g., HHHHTTHTTTH becomes $\boxed{\text{HHH}}\ \boxed{\text{TT}}\ \boxed{\text{H}}\ \boxed{\text{TTT}}\ \boxed{\text{H}}$, which has 5 runs. Find the expected number of runs.

Hint: Start by finding the expected number of tosses (other than the first) where the outcome is different from the previous one.

Solution 10. If a toss has a different outcome from the previous toss, that means that a new *run* has started. Hence, the expected number of tosses with a different outcome than the previous one is almost equal to the expected number of runs. The only difference is that, this way, we have not accounted for the first run (since the first toss cannot be different from the nonexistent "zeroth"), therefore, by adding 1 to this expectation, we get the expected number of runs.

Denote the number of runs with R and the event that the toss i is different from the $(i-1)$-th toss with D_i for $i = 2, \ldots, n$. The D_i-s are independent and identically distributed indicator/Bernoulli random variables. Using the observations of the first paragraph, we can obtain the expected number of runs as

$$E(R) = E\left(1 + \sum_{i=2}^{n} D_i\right).$$

Using the linearity of expectation (and that the expectation of constant is equal to the constant) we have

$$E(R) = 1 + \sum_{i=2}^{n} E(D_i),$$

therefore we only have to determine $E(D_i)$. There are two possibilities for the i-th toss to be different from the preceding toss: Either the $(i-1)$-th toss was heads and the current toss is tails, or the $(i-1)$-th toss was tails and the current one is heads. These two disjoint events can occur with probabilities $p(1-p)$ and $(1-p)p$, respectively, and we know that if either of these two occurs, $D_i = 1$. Hence $D_i \sim Bern(2 \cdot (1-p)p)$. We have seen in class that a $Bern(p^*)$ random variable has expectation p^* (but you can show it from the definition of expectation too), hence $E(D_i) = 2 \cdot (1-p)p$. Finally, plugging back into the formula for $E(R)$ we get

$$E(R) = 1 + \sum_{i=2}^{n} 2p(1-p) = 1 + 2p(1-p)\sum_{i=2}^{n} 1 = 1 + 2(n-1)p(1-p).$$

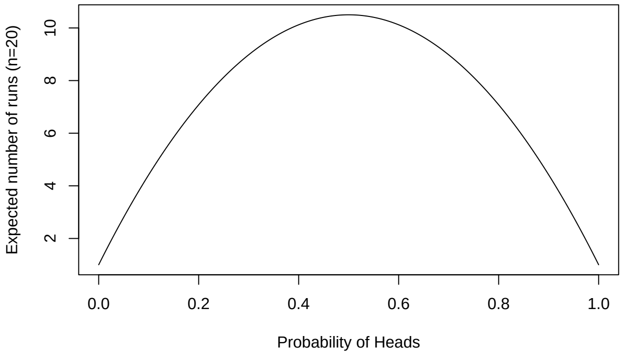

Figure 2: Expected number of runs in $n = 20$ coin tosses.