Problem Sheet 8 [1]

Based on Chapters 5.4, 5.5, part of 5.7, and Chapter 6.1 of the course book.

Optional Revision Problems

Exercise 1. The Rayleigh distribution has PDF $f(x) = xe^{-x^2/2}$, $x > 0$, and zero otherwise. Let X have the Rayleigh distribution.

1. Find $P(1 < X < 3)$.
2. Find the first quartile, median, and third quartile of X; these are defined to be the values q_1,q_2,q_3 (respectively) such that $P(X \leq q_j) = j/4$ for $j = 1, 2, 3$.

Hint: For calculating probabilities from the PDF, you can use the information in the table in Section 5.8 in the book.

Solution 1. 1. By definition

$$P(1 < X < 3) = \int_1^3 f(x)\mathrm{d}x = \int_1^3 xe^{-x^2/2}\mathrm{d}x = \left[-e^{-x^2/2}\right]_1^3 = e^{-1/2} - e^{-9/2} \approx 0.595$$

2. From the PDF it follows that the support of X is $(0, \infty)$, i.e. that X is positive with probability 1, therefore $P(X \leq q_j) = P(0 < X \leq q_j)$. Then by the same derivation as in part 1.

$$P(0 < X < q_j) = \left[-e^{-x^2/2}\right]_0^{q_j} = 1 - e^{-q_j^2/2}.$$

We want to find the value q_j for which $P(0 < X < q_j) = 1 - e^{-q_j^2/2} = \frac{j}{4}$ holds. Then,

$$\begin{aligned} & 1 - e^{-q_j^2/2} = \frac{j}{4} \\ \iff & \frac{4-j}{4} = e^{-q_j^2/2} \\ \iff & log\left(\frac{4-j}{4}\right) = -q_j^2/2 \\ \iff & -2 \cdot log\left(\frac{4-j}{4}\right) = q_j^2 \\ \iff & \sqrt{-2 \cdot log\left(\frac{4-j}{4}\right)} = q_j. \end{aligned}$$

Substituting in for $j = 1, 2, 3$; $q_1 \approx 0.759$, $q_2 \approx 1.177$, and $q_3 \approx 1.665$.

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

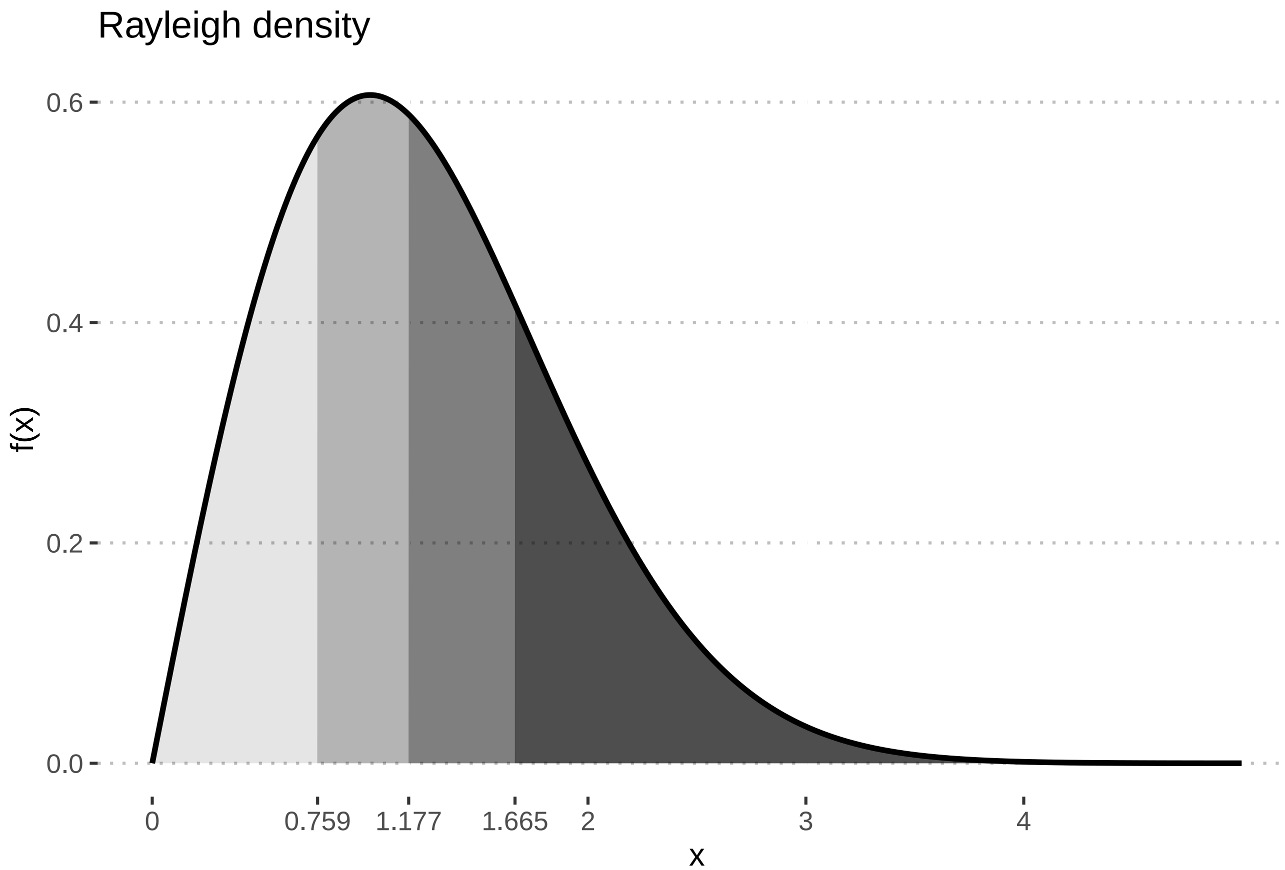

Figure 1: PDF of the Rayleigh distribution. All four shaded regions have the same area, i.e. the first, second, and third quartiles are at the "color-change".

Exercise 2. A stick is broken into two pieces, at a uniformly random breakpoint. Find the CDF and average of the length of the longer piece.

Solution 2. We can assume the units are chosen so that the stick has length 1. Let L be the length of the longer piece, and let the break point be $U \sim Unif(0,1)$. For any $l \in [1/2, 1]$, observe that $L < l$ is equivalent to $\{U < l \text{ and } 1 - U < l\}$, i.e. that the piece before the breakpoint (U) has at most length l, and the piece after the breakpoint ($1 - U$) has at most length l. For $l < 1/2$, this event happens with probability 0, as it is impossible that both pieces are shorter than 1/2. The event $\{U < l \text{ and } 1 - U < l\}$ can be written as $1 - l < U < l$. Therefore we can obtain L's CDF as

$$F_L(l) = P(L < l) = P(1 - l < U < l) = 2l - 1,$$

so $L \sim Unif(1/2, 1)$. In particular, $E(L) = 3/4$.

Week 8 Exercises

Exercise 3. Let $U \sim Unif(0,1)$. Using U, construct $X \sim Expo(\lambda)$

Solution 3. We know from the universality of the uniform (Thm 5.3.1 in the course book), that if there exists CDF F, that is continuous and strictly increasing, then we can construct X from the $Uniform(0,1)$ random variable U, as $X = F^{-1}(U)$, such that it will have the CDF F. For the current problem this implies, that we have to find the inverse of the CDF of the Exponential distribution. If $X \sim Expo(\lambda)$, then $F_X(x) = 1 - e^{-\lambda x}, \quad x > 0$, that satisfies the continuity and the strictly increasing criteria.

$$\begin{aligned} u = F_X(x) = 1 - e^{-\lambda x} & \\ &\iff e^{-\lambda x} = 1 - u \\ &\iff \lambda x = -log(1-u) \\ &\iff x = \frac{-log(1-u)}{\lambda}, \end{aligned}$$

therefore $F_X^{-1}(u) = \frac{-log(1-u)}{\lambda}$, so $X := \frac{-log(1-U)}{\lambda} \sim Expo(\lambda)$. By the remark from S7E6 (if $U \sim Unif(0,1)$, then $(1-U) \sim Unif(0,1)$, as well), we can define X as an even simpler function of U, as $X := \frac{-log(U)}{\lambda}$.

Figure 2 visualizes the procedure above: I simulated 500 samples from a $Unif(0,1)$ distribution and then transformed each of them according to the function $g(u) = \frac{-log(1-U)}{\lambda}$, with $\lambda = 3$. Figure 2 is the histogram of these transformed values, with the theoretical PDF of $Expo(3)$ as the red line.

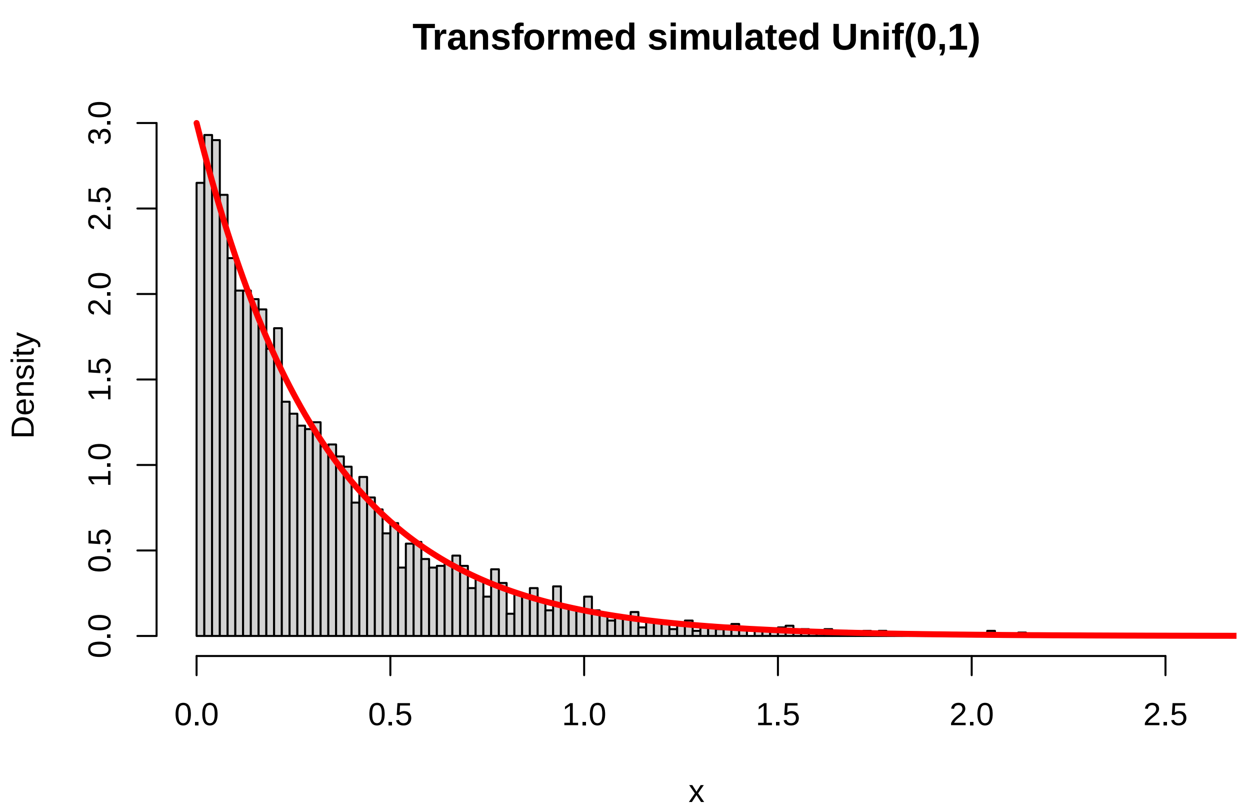

Figure 2: Histogram of simulated and then transformed Uniform random variables, and the PDF of $Expo(3)$.

Exercise 4. Let $Z \sim \mathcal{N}(0,1)$. Create an r.v. $Y \sim \mathcal{N}(1,4)$, as a simple-looking function of Z. Make sure to check that your Y has the correct mean and variance.

Hint: If the transformation is not immediate, check Definition 5.4.3

Solution 4. From Definition 5.4.3 or from the standardization of normal random variables, it follows that if $Y := 1 + 2Z$, then $Y \sim \mathcal{N}(1, 4)$. To check that it has the correct mean, we can use the linearity of expectation

$$E(Y) = E(1 + 2Z) = 1 + 2E(Z) = 1 + 2 \cdot 0 = 1,$$

and for the variance, we can use the observations made after Thm 4.6.3, that $Var(X+c) = Var(X)$ and $Var(cX) = c^2Var(X)$. (If in doubt that these equalities hold, try to show them using the definition of the variance, and the linearity of expectation.) Then,

$$Var(Y) = Var(1 + 2Z) = Var(2Z) = 4Var(Z) = 4 \cdot 1,$$

so Y has the correct mean and variance.

Exercise 5. Let $Z \sim N(0, 1)$. We know from the 68-95-99.7% rule that there is a 68% chance of Z being in the interval $(-1, 1)$. Give a visual explanation of whether or not there is an interval (a, b) that is shorter than the interval $(-1, 1)$, yet which has at least as large a chance as $(-1, 1)$ of containing Z.

Solution 5. The probability that Z is contained in the interval $(-1, 1)$, is simply the area under the PDF of Z (as $P(a < Z < b) = \int_a^b f(y)\mathrm{d}z$). For the Standard normal and $a = -1$, $b = 1$, this is visualized in Figure 3.

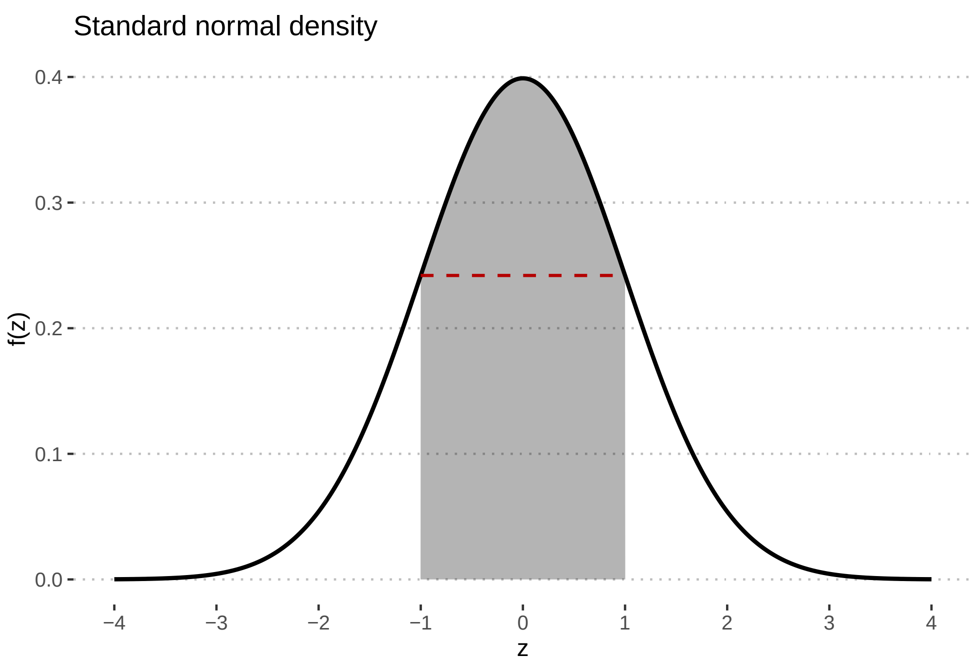

Figure 3: PDF of a Standard normal. The shaded area represents the probability that Z is between -1 and 1.

Since the PDF takes the highest values within this interval, i.e. $f(x) > f(y)$ for any $x \in [-1, 1]$, $y \notin [-1, 1]$, if we were to choose any other interval, that has the same area under it, we would find it to be greater in length. For example, if we fixed the left end of the interval at -0.5, to ensure that we have the same chance of having Z in this new interval, we would need to select the right end to be approximately 2.27, so the interval length has increased from 2 to 2.77.

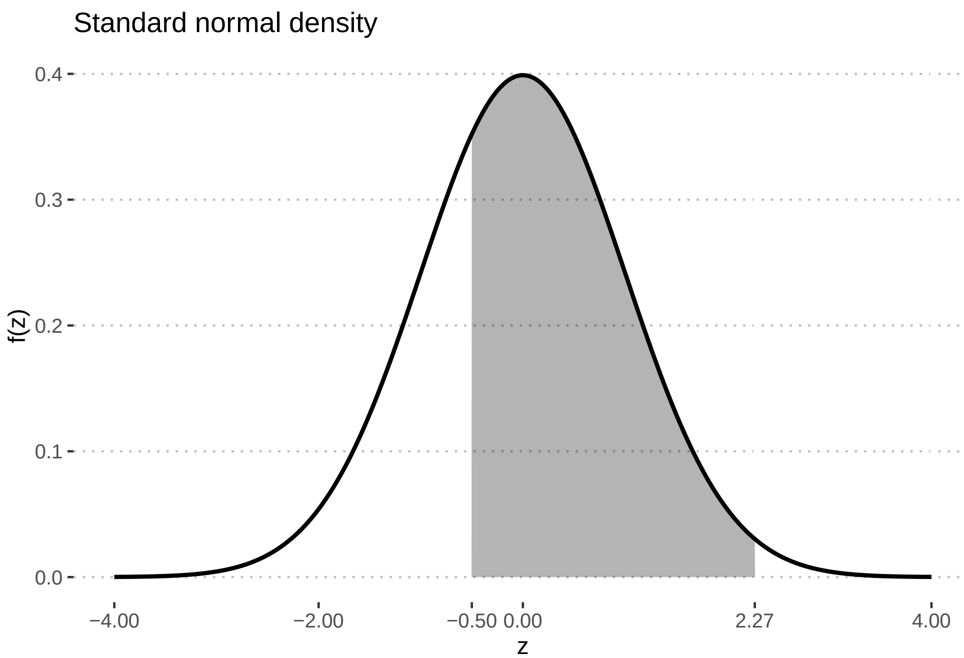

Figure 4: PDF of a Standard normal. The shaded area represents the probability that Z is between -0.5 and 2.27.

So in summary, there is no other interval $[a, b]$ that is shorter, but still ensures that $P(Z \in [a, b]) = 0.68$

This is vaguely related to a concept in Bayesian statistics, where people try to find the shortest credible interval of the posterior density, that guarantees that the unobserved parameter falls within that interval with a certain probability.

Exercise 6. A post office has 2 clerks. Alice enters the post office while 2 other customers, Bob and Claire, are being served by the 2 clerks. She is next in line. Assume that the time a clerk spends serving a customer has an $Expo(\lambda)$ distribution.

1. What is the probability that Alice is the last of the 3 customers to be done being served?

   Hint: No integrals are needed.

2. What is the expected total time that Alice needs to spend at the post office?

   Hint: Example 5.6.3 can be useful.

Solution 6.

1. Alice begins to be served when either Bob or Claire leaves. By the memoryless property, by conditioning on the waiting time needed until Bob or Claire has been served, the additional time needed to serve whichever of Bob or Claire is still there is $Expo(\lambda)$. The time it takes to serve Alice is also $Expo(\lambda)$, so by symmetry the probability is 1/2 that Alice is the last to be done being served.

2. The expectation of random variable with $Expo(\lambda)$ distribution is $\frac{1}{\lambda}$. Therefore, the expected time spent waiting in line is $\frac{1}{2\lambda}$, since the minimum of two independent $Expo(\lambda)$ r.v.s is $Expo(2\lambda)$ (by Example 5.6.3). The expected time spent being served is $\frac{1}{\lambda}$. So the expected total time is $\frac{1}{2\lambda} + \frac{1}{\lambda} = \frac{3}{2\lambda}$.

Exercise 7. Let T be the time until a radioactive particle decays, and suppose (as is often done in physics and chemistry) that $T \sim Expo(\lambda)$.

1. The half-life of the particle is the time at which there is a 50% chance that the particle has decayed (in statistical terminology, this is the median of the distribution of T). Find the half-life of the particle.

2. Show that for ϵ a small, positive constant, the probability that the particle decays in the time interval $[t, t+\epsilon]$, given that it has survived until time t, does not depend on t and is approximately proportional to ϵ.

   Hint: $e^x \approx 1 + x$ if $x \approx 0$.

Solution 7. 1. The median for continuous random variables is defined as the value t, for which $P(T \leq t) = 1/2$ holds. Thus we have to find t, such that $F_T(t) = 1 - e^{-\lambda t} = 0.5$. Then,

$$\begin{aligned} 1 - e^{-\lambda t} &= 0.5 \\ \iff e^{-\lambda t} &= 0.5 \\ \iff \lambda t &= -log(0.5) \\ \iff t &= \frac{log(2)}{\lambda}. \end{aligned}$$

2. We want to find the probability that the particle decays in the given interval, conditional on the fact, that it survived until the beginning of the interval. In probabilistic notation we want to find $P(t \leq T \leq t+\epsilon | T > t)$. From the memoryless property of the exponential it follows that $P(t \leq T \leq t+\epsilon | T > t) = P(T \leq \epsilon)$. You can see this from the observation from the lecture that the memoryless property means that "even if you've waited for a long time without success, the success is not more likely to occur soon", where success is the decay. Alternatively, you can show this by rewriting $P(t \leq T \leq t+\epsilon | T > t) = 1 - P((t > T) \cup (T > t+\epsilon) | T > t)$ and by the axioms of probability this is equal to $1 - P(t > T | T > t) - P(T > t+\epsilon | T > t) = 1 - 0 - P(T > \epsilon) = P(T \leq \epsilon)$, where in the second to last step we used the memoryless property per definition. (Or the last option is writing out the probabilities as integrals, solving them and getting the same result). By the definition of the CDF of an exponential r.v.

$$P(T \leq \epsilon) = F(\epsilon) = 1 - e^{-\lambda\epsilon},$$

so using the hint, this is

$$1 - (1 - \lambda\epsilon) = \lambda\epsilon,$$

that does not depend on t and proportional to ϵ, as requested.

Exercise 8. Let $U \sim Unif(a, b)$. Find the median and mode of U.

Solution 8. For continuous r.v.-s the median is just the value u, for which $F(u) = 0.5$ holds. The CDF of $Unif(a, b)$ is $\frac{u-a}{b-a}$ for $u \in [a, b]$, so we have to find the solution in u for $\frac{u-a}{b-a} = 0.5$ After some simple algebra we get that the median is at $u = \frac{a+b}{2}$.

A quicker solution is noting that symmetric distributions have the same mean and median (this can be a practice exercise), and remembering that the mean of $Unif(a, b)$ is $\frac{a+b}{2}$.

The mode is defined as the value that maximizes the PDF of the random variable. Since the PDF of U is zero outside $[a, b]$, and constant $\frac{1}{b-a}$ on $[a, b]$, it follows that any value between a and b is the mode.

Exercise 9. Let Y be Log-Normal with parameters μ and σ^2. So $Y = e^X$ with $X \sim N(\mu, \sigma^2)$. Three students are discussing the median and the mode of Y . Evaluate and explain whether or not each of the following arguments is correct.

(a) Student A: The median of Y is e^μ because the median of X is μ and the exponential function is continuous and strictly increasing, so the event $Y \leq e^\mu$ is the same as the event $X \leq \mu$.

(b) Student B: The mode of Y is e^μ because the mode of X is μ, which corresponds to e^μ for Y since $Y = e^X$.

(c) Student C: The mode of Y is μ because the mode of X is μ and the exponential function is continuous and strictly increasing, so maximizing the PDF of X is equivalent to maximizing the PDF of $Y = e^X$.

Solution 9. Student A is correct because the exponential function is strictly increasing, therefore

$$0.5 = P(X < \mu) = P(e^X < e^\mu) = P(Y < e^\mu).$$

Student B is incorrect because the exponential function introduces skewness to the PDF of the new random variable, "moving it to the left" so the maximiser of the new PDF will be less than e^μ (in fact the mode of Y is $e^{\mu-\sigma^2}$).

Student C is incorrect as well. Consider $\mu < 0$, then C would imply that the mode of Y is negative, that is impossible, as Y can only take positive values.