Full Name: ______________________

Student No: ______________________

Instructor: Siamak Taati

Grade:

Read before you start:

- Please make sure you write your full name and student number.
- The exam consists of 9 questions, most with multiple parts, and a total score of 155 points. All answers require justifications.
- During the exam, you are allowed to use your phone to access the app for probability distributions (and possibly a calculator app), but the internet access of your phone must be switched off.
- The duration of the exam is 2 hours.

You can use the remainder of this page as scratch paper.

1. (15 points) Determine which of the following statements is True and which is False. In each case, give a short justification.

FALSE If A and B are two independent events, then necessarily $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

> **Solution:** For arbitrary events A and B we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ by inclusion-exclusion. When A and B are independent, it is not necessarily the case that $\mathbb{P}(A \cap B) = 0$. Namely, in that case, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$, so $\mathbb{P}(A \cap B) = 0$ only if either A or B has probability 0.

TRUE If A and B are two events with non-zero probabilities, then $\mathbb{P}(A \cap B \mid A \cup B) \leq \mathbb{P}(A \mid B)$.

> **Solution:** Since $B \subseteq A \cup B$, we have $\mathbb{P}(B) \leq \mathbb{P}(A \cup B)$. Therefore,
>
> $$\mathbb{P}(A \cap B \mid A \cup B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cup B)} \leq \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A \mid B)\ .$$

FALSE If X and Y are discrete random variables with probability mass functions $p_X(a)$ and $p_Y(a)$ respectively, then $X + Y$ is also discrete and has probability mass function $p_X(a) + p_Y(a)$.

> **Solution:** The function $q(a) \coloneqq p_X(a) + p_Y(a)$ is not even a probability mass function because its values add up to 2 instead of 1.

2. (15 points) A jar contains 10 blue balls and 5 red balls. We draw the balls one after another at random without replacement until there is no ball left in the jar.

 (a) What is the probability that the last ball is red?

> **Solution:** We can follow at least 3 different approaches:
>
> - Approach 1: (short and intuitive)
>   Any of the 15 balls in the jar can be the last one to be drawn, and by symmetry, they are all equally likely. Hence,
>
> $$\mathbb{P}(\text{the last ball is red}) = \frac{\#\text{ of red balls}}{\text{total }\#\text{ of balls}} = \frac{5}{15} = \boxed{\frac{1}{3}}\ .$$
>
> - Approach 2: (using model 1)
>   If we only take note of the color of the ball drawn at each turn, then each possible outcome can be represented by a string of 15 characters, 10 of which are `B`s and 5 of which are `R`s. For instance, `BRRBRBBBRBBBBRB` represents the outcome that the first ball we draw is blue, the second and the third are red, the fourth is blue, and so on. The sample space is thus the set Ω of all such strings.
>
>   Let E be the event the last ball is red. As a subset of Ω, the event E consists of all the strings that end with `R`.
>
>   By symmetry, all the possible outcomes in Ω are equally likely. Note that $|\Omega| = \binom{15}{5}$ (why?) and $|E| = \binom{14}{4}$ (why?). Hence,
>
> $$\mathbb{P}(E) = \frac{|E|}{|\Omega|} = \frac{\binom{14}{4}}{\binom{15}{5}} = \frac{\frac{14!}{4!10!}}{\frac{15!}{5!10!}} = \boxed{\frac{1}{3}}\ .$$

- Approach 3: (using model 2)
  If we distinguish between the balls of the same color, then the sample space can be represented by the set Ω' of all $15!$ orderings of the 15 balls.

  Let E' be the event that the last ball is red. As a subset of Ω', the event E' consists of all those orderings that end with one of the 5 red balls.

  Again, by symmetry, all the possible outcomes in Ω' are equally likely. We have $|\Omega'| = 15!$ and $|E'| = 5 \times 14!$ (why?). Hence,

$$\mathbb{P}(E') = \frac{|E'|}{|\Omega'|} = \frac{5 \times 14!}{15!} = \boxed{\frac{1}{3}}.$$

(b) What is the probability that no two consecutive balls we draw are red?

Solution: We use the first model used in Part (a). Let F denote the event that no two consecutive balls we draw are red. The elements of F are represented by the strings in which no two `R`s sit next to each other. There are $\binom{11}{5}$ such strings. Indeed, to ensure that between every two consecutive `R`s there is at least one `B`, glue a `B` to the right of each of the first four `R`s. We are left with 6 blues and 5 reds. To count the number of possibilities now we can count the number of possible positions of the reds among the $5 + 6 = 11$ different positions.
Since all the possible outcomes in Ω are equally likely, we have

$$\mathbb{P}(F) = \frac{|F|}{|\Omega|} = \frac{\binom{11}{5}}{\binom{15}{5}} = \frac{\frac{11!}{5!6!}}{\frac{15!}{5!10!}} = \frac{10 \times 9 \times 8 \times 7}{15 \times 14 \times 13 \times 12} = \boxed{\frac{2}{13}} \approx 0.15385 .$$

A solution using model 2 from Part (a) would be equally valid, and would (naturally) lead to the same answer.

(c) If the 2nd ball is red, what is the probability that the 1st ball has also been red?

Solution: For each k, let B_k denote the event that the k-th ball is blue and R_k denote the event that the k-th ball is red.

- Approach 1: (using Bayes's rule)
  Using Bayes' rule, we have

$$\begin{aligned}\mathbb{P}(R_1 \mid R_2) &= \frac{\mathbb{P}(R_1 \cap R_2)}{\mathbb{P}(R_2)} \\ &= \frac{\mathbb{P}(R_1)\,\mathbb{P}(R_2 \mid R_1)}{\mathbb{P}(R_2)} \\ &= \frac{\frac{5}{15} \times \frac{4}{14}}{\frac{1}{3}} \\ &= \boxed{\frac{2}{7}} \approx 0.28571 .\end{aligned}$$

  For the third equality, we have used the fact that $\mathbb{P}(R_2) = 1/3$. This can be seen using a similar reasoning as in Part (a). Alternatively, we can use the principle of total probability and the chain rule to write

$$\mathbb{P}(R_2) = \mathbb{P}(R_1)\,\mathbb{P}(R_2 \mid R_1) + \mathbb{P}(B_1)\,\mathbb{P}(R_2 \mid B_1) = \frac{5}{15} \times \frac{4}{14} + \frac{10}{15} \times \frac{5}{14} = \frac{1}{3} .$$

- Approach 2: (using symmetry)
  By symmetry, the desired probability is the same as the probability that the 2nd balls is red given that the 1st ball has been red. Namely,

$$\mathbb{P}(R_1 \mid R_2) = \mathbb{P}(R_2 \mid R_1) = \frac{4}{14} = \boxed{\frac{2}{7}} \approx 0.28571 \, .$$

3. (25 points) Let U be a random variable uniformly distributed over the interval $[0, 1]$, and let $X = \frac{1}{1+U}$.

(a) Find the expected value of X.

Solution: Let f_U denote the pdf of U. Note that $f_U(u) = 1$ for $u \in (0, 1)$ and $f_U(u) = 0$ for $u \notin [0, 1]$. Using the distribution of U, we can compute

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \frac{1}{1+u} f_U(u)\, \mathrm{d}u = \int_0^1 \frac{1}{1+u}\, \mathrm{d}u = \ln(1+u)\Big|_0^1 = \boxed{\ln 2} \, .$$

(b) What are the possible values of X?

Solution: As we vary u over $[0, 1]$, the function $\frac{1}{1+u}$ varies continuously from 1 (when $u = 0$) to $1/2$ (when $u = 1$). Hence, the possible values of X are the numbers in the interval $[1/2, 1]$.

(c) Find the cdf of X.

Solution: Let F_X denote the cdf of X. For $x \in [1/2, 1]$, we have

$$\begin{aligned} F_X(x) = \mathbb{P}(X \le x) &= \mathbb{P}\left(\frac{1}{1+U} \le x\right) \\ &= \mathbb{P}\left(U \ge \frac{1}{x} - 1\right) \\ &= 1 - \mathbb{P}\left(U < \frac{1}{x} - 1\right) \\ &= 1 - \left(\frac{1}{x} - 1\right) \qquad \text{(because } 0 \le \frac{1}{x} - 1 \le 1\text{)} \\ &= 2 - \frac{1}{x} \, . \end{aligned}$$

Hence, in general,

$$F_X(x) = \begin{cases} 0 & \text{if } x < 1/2, \\ 2 - \dfrac{1}{x} & \text{if } 1/2 \le x \le 1, \\ 1 & \text{if } x > 1. \end{cases}$$

(d) Find the pdf of X.

Solution: Differentiating the cdf of X, we obtain that its pdf:

$$f_X(x) = \begin{cases} \dfrac{1}{x^2} & \text{if } x \in (1/2, 1), \\ 0 & \text{if } x \notin [1/2, 1]. \end{cases}$$

(e) What is the probability that $1/4 < X < 3/4$?

Solution: We have

$$\begin{aligned} \mathbb{P}(1/4 < X < 3/4) &= \mathbb{P}(X < 3/4) - \mathbb{P}(X \leq 1/4) \\ &= \lim_{x \nearrow \frac{3}{4}} F_X(x) - F_X(1/4) \\ &= F_X(3/4) - F_X(1/4) && (F_X \text{ is continuous at } x = 3/4) \\ &= \left(2 - \frac{1}{3/4}\right) - 0 \\ &= \boxed{2/3}. \end{aligned}$$

4. (15 points) Let $X \sim \mathsf{N}(\mu = 2, \sigma^2 = 1)$ and $Y \sim \mathsf{N}(\mu = -1, \sigma^2 = 2^2)$ be two independent normal random variables. What is the distribution of $-2X + Y + 1$?

Solution: By the stability of the normal distribution, $W := -2X + Y + 1$ is normally distributed. Its parameters are

$$\begin{aligned} \mathbb{E}[W] &= \mathbb{E}[-2X + Y + 1] \\ &= -2\,\mathbb{E}[X] + \mathbb{E}[Y] + 1 && \text{(linearity of expectation)} \\ &= -2 \times 2 + (-1) + 1 \\ &= \boxed{-4}. \\ \mathbb{V}\mathrm{ar}[W] &= \mathbb{V}\mathrm{ar}[-2X + Y + 1] \\ &= 4\,\mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y] && (\text{independence of } X \text{ and } Y) \\ &= 4 \times 1 + 2^2 \\ &= \boxed{8}. \end{aligned}$$

5. (15 points) Two random variables X and Y have joint probability density

$$f(x, y) = \begin{cases} x + y & \text{if } 0 < x < 1 \text{ and } 0 < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal distributions of X and Y.

Solution: Note that the possible values of X and Y are in the interval $(0, 1)$.

For $x \in (0,1)$, the marginal pdf of X is

$$f_X(x) = \int_{y=-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int_{y=0}^{1} x + y\,\mathrm{d}y = \Big[xy + \frac{1}{2}y^2\Big]_{y=0}^{1} = x + 1/2\ .$$

For $x \notin [0,1]$, we have $f_X(x) = 0$.
By symmetry, the marginal distribution of Y is the same as the marginal distribution of X.

(b) Are X and Y independent?

Solution: They are not independent. This can be seen for instance by noting that, for $0 < x < 1$ and $0 < y < 1$,

$$f(x,y) = x + y \neq (x + 1/2)(y + 1/2) = f_X(x)f_Y(y)\ .$$

(c) Compute $\mathbb{P}(X + Y > 1)$.

Solution: To obtain $\mathbb{P}(X + Y > 1)$, we integrate the joint probability density of X and Y over the region $\{(x,y) : x + y > 1\}$ of the plane. Namely,

$$\begin{aligned}
\mathbb{P}(X + Y > 1) &= \iint_{(x,y):x+y>1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{x=0}^{1}\int_{y=1-x}^{1} (x+y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{1} \Big[xy + \frac{1}{2}y^2\Big]_{y=1-x}^{1}\,\mathrm{d}x \\
&= \int_{x=0}^{1} \Big(x + \frac{1}{2}x^2\Big)\,\mathrm{d}x \\
&= \Big[\frac{1}{2}x^2 + \frac{1}{6}x^3\Big]_{x=0}^{1} \\
&= \frac{1}{2} + \frac{1}{6} \\
&= \boxed{2/3}\ .
\end{aligned}$$

6. (20 points) In a digital communication channel, messages are sent and received as strings of `0`s and `1`s. Due to noise, the transmission of each bit of information is subject to error (i.e., receiving a 0 instead of a 1, or vice versa). We assume that every bit is transmitted incorrectly with probability 0.01, independently of the other bits.

(a) What is the probability that a message with 100 bits is transmitted without error?

Solution: For $k = 1, 2, \ldots$, let E_k denote the event that the k-th bit is transmitted with an error. By assumption, $E_1, E_2, \ldots$ are independent, each having probability 0.01. The event that a message with 100 bits is transmitted without error is the event $E_1^c \cap E_2^c \cap \cdots \cap E_{100}^c$. We have

$$\begin{aligned}\mathbb{P}(E_1^c \cap E_2^c \cap \cdots \cap E_{100}^c) &= \mathbb{P}(E_1^c)\,\mathbb{P}(E_2^c)\cdots\mathbb{P}(E_{100}^c) && \text{(independence)} \\ &= (1-0.01)^{100} \\ &\approx \boxed{0.36603}\,.\end{aligned}$$

(b) What is the expected number of errors in a message with 10000 bits?

Solution: Let N_{10000} denote the number of errors in a message with 10000 bits. We can write

$$N_{10000} = X_1 + X_2 + \cdots X_{10000}$$

where, for each k, X_k is a Bernoulli random variable indicating whether the k-th bit is transmitted with an error or not. By assumption, $X_1, X_2, \ldots$ are independent random variables, each with Bernoulli parameter $p = 0.01$. By the linearity of expectation,

$$\mathbb{E}[N_{10000}] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_{10000}] = 10000 \times 0.01 = \boxed{100}\,.$$

(Alternatively, you could note that N is a binomial random variable with parameters $n = 10000$ and $p = 0.01$, and remember that the expected value of a binomial random variable with parameters n and p is np.)

(c) What is the variance of the number of errors in a message with 10000 bits?

Solution: Using the notation of the previous part, and using the independence of $X_1, X_2, \ldots$, we have

$$\mathbb{V}\mathrm{ar}[N] = \mathbb{V}\mathrm{ar}[X_1] + \mathbb{V}\mathrm{ar}[X_2] + \cdots + \mathbb{V}\mathrm{ar}[X_{10000}] = 10000 \times 0.01 \times (1 - 0.01) = \boxed{99}\,.$$

(d) Find the (approximate) probability that the number of errors in a message with 10000 bits is less than 85?

Solution:

- Approach 1: (exact solution)
  As noted in Part (b), N_{10000} is a binomial random variable with parameters $n = 10000$ and $p = 0.01$. Computing the desired probability by hand would be out of the question, as it would require too much time and space and is very sensitive to rounding errors. Using the app for the binomial distribution (or any other computer software), we can find

  $$\mathbb{P}(N_{10000} < 85) = \mathbb{P}(N_{10000} \leq 84) \approx \boxed{0.05665}\,.$$

  It is interesting to remark that the browser version of the app cannot handle this tricky computation.

- Approach 2: (approximation via the CLT)
  By the central limit theorem, the random variable

  $$N_{10000} = X_1 + X_2 + \cdots X_{10000}$$

is approximately normally distributed with parameters $\mu = 100$ and $\sigma^2 = 99$ as calculated in Parts (b) and (c). Using the app for the normal distribution (or any other computer software), we can find

$$\mathbb{P}(N_{10000} < 85) \approx \boxed{0.06583}\,.$$

A better approximation is obtained if we use the "histogram correction":

$$\mathbb{P}(N_{10000} < 85) = \mathbb{P}(N_{10000} < 84.5) \approx \boxed{0.05964}\,.$$

7. (25 points) A random variable X has a uniform distribution over the interval $[0, \beta]$, where β is an unknown parameter. Consider the following point estimators:

$$\widehat{\beta} := 2\overline{X}_n \qquad\qquad \widetilde{\beta} := \max\{X_1, X_2, \ldots, X_n\}$$

for β based on an independent sample $X_1, X_2, \ldots, X_n$. (Here, $\overline{X}_n := (X_1 + X_2 + \cdots + X_n)/n$ stands for the sample mean.)

(a) Argue that $\widehat{\beta}$ is consistent.

Solution: By the law of large numbers $\overline{X}_n \to \mathbb{E}[X]$ as $n \to \infty$. Since X is uniformly distributed over $[0, \beta]$, we have $\mathbb{E}[X] = \frac{1}{2}\beta$. This can be seen either by symmetry, or by direct computation:

$$\mathbb{E}[X] = \int_0^\beta \frac{1}{\beta} x \,\mathrm{d}x = \frac{1}{2\beta} x^2 \Big|_{x=0}^{\beta} = \frac{1}{2}\beta\,.$$

It follows that,

$$\widehat{\beta} = 2\overline{X}_n \to 2 \times \frac{1}{2}\beta = \beta$$

as $n \to \infty$, which means $\widehat{\beta}$ is a consistent estimator for β.

(b) Argue that $\widetilde{\beta}$ is consistent.

Solution:

- Justification 1: (non-rigorous, but sufficient for this course)
  Recall that saying that X is uniformly distributed over $[0, \beta]$ roughly means that, if we repeat the experiment involving X many many times, then the values of X in these repetitions will be distributed roughly uniformly over $[0, \beta]$.

  This means that, when n is large, the values of $X_1, X_2, \ldots, X_n$ are distributed roughly uniformly over $[0, \beta]$. In particular, all the values will be smaller than or equal to β, yet there will be values very close to β. Hence, $\widetilde{\beta} = \max\{X_1, X_2, \ldots, X_n\}$ will be very close to β.

- Justification 2: (more rigorous)
  Clearly, $\widetilde{\beta} \leq \beta$ for every n (why?). Consider an arbitrarily small $\varepsilon > 0$. Let us show that when n is large, $\widehat{\beta} > \beta - \varepsilon$ with a probability close to 1.

Indeed,

$$\begin{aligned}\mathbb{P}(\widehat{\beta} \leq \beta - \varepsilon) &= \mathbb{P}(\max\{X_1, X_2, \ldots, X_n\} \leq \beta - \varepsilon) \\ &= \mathbb{P}(X_k \leq \beta - \varepsilon \text{ for every } k \in \{1, 2, \ldots, n\}) \\ &= \mathbb{P}(X_1 \leq \beta - \varepsilon)\, \mathbb{P}(X_1 \leq \beta - \varepsilon) \cdots \mathbb{P}(X_1 \leq \beta - \varepsilon) \qquad \text{(independence)} \\ &= \left(\frac{\beta - \varepsilon}{\beta}\right)^n .\end{aligned}$$

Since $0 < \frac{\beta-\varepsilon}{\beta} < 1$ (as long as ε is not too large), we have $\left(\frac{\beta-\varepsilon}{\beta}\right)^n \to 0$ as $n \to \infty$. In conclusion, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(\beta - \varepsilon < \widetilde{\beta} < \beta) = 1 ,$$

which means that $\widetilde{\beta}$ is a consistent estimator for β.

(c) Is $\widehat{\beta}$ biased?

Solution: No. By the linearity of expectation,

$$\mathbb{E}[\widehat{\beta}] = \mathbb{E}[2\overline{X}_n] = \frac{2}{n}(\mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]) = \frac{2}{n} \times n\, \mathbb{E}[X] = 2\, \mathbb{E}[X] .$$

As we calculated in Part (a), we have $\mathbb{E}[X] = \frac{1}{2}\beta$. It follows that $\mathbb{E}[\widehat{\beta}] = \beta$, which means $\widehat{\beta}$ is unbiased.

Let $n = 5$, and suppose we observe the values

$$X_1 = 0.2312 , \qquad X_2 = 3.8882 , \qquad X_3 = 2.0016 , \qquad X_4 = 2.1474 , \qquad X_5 = 0.1431 .$$

(d) Compute the estimates provided by $\widehat{\beta}$ and $\widetilde{\beta}$.

Solution: The above observed values lead to the estimates

$$\begin{aligned}\widehat{\beta} &= 2 \times \frac{1}{5}(0.2312 + 3.8882 + 2.0016 + 2.1474 + 0.1431) = \boxed{3.3646} \\ \widetilde{\beta} &= \max\{0.2312, \quad 3.8882, \quad 2.0016, \quad 2.1474, \quad 0.1431\} = \boxed{3.8882}\end{aligned}$$

for β.

(e) Given the above observed values, is the estimate provided by $\widehat{\beta}$ reasonable?

Solution: Since $X_1, X_2, \ldots, X_5$ are chosen from $[0, \beta]$, their values cannot be larger than β. However, the estimate $\widehat{\beta} = 3.3646$ is not consistent with the observed values, in particular with $X_2 = 3.8882$ which is larger than the value of $\widehat{\beta}$.

8. (5 points) At AUBMC, on average, 3 babies are born per day. The arrival times of the babies can be modeled by a Poisson process. What is the probability that, on Sunday, at least 2 babies arrive before noon? (We take the midnight as the start of the day.)

Solution:

- Approach 1:
  Let N denote the number of babies arriving on Sunday before noon (i.e., between 12 am and 12 pm). By the properties of the Poisson process, N is a Poisson random variable with parameter $\mu = \langle\text{rate}\rangle \times \langle\text{interval length}\rangle = 3 \times 0.5 = 1.5$. Therefore,

$$\mathbb{P}(N \geq 2) = 1 - \mathbb{P}(N = 0) - \mathbb{P}(N = 1) = 1 - \mathrm{e}^{-1.5} - \mathrm{e}^{-1.5} \times 1.5 \approx \boxed{0.44217}\,.$$

- Approach 2:
  Let S_2 denote the arrival time of the second baby on Sunday. By the properties of the Poisson process, S_2 is a gamma random variable with shape $\alpha = 2$ and rate $\lambda = 3$ per day. Using the app for the gamma distribution,

$$\mathbb{P}(\text{the second baby arrives before noon}) = \mathbb{P}(S_2 < 0.5\,\text{day}) \approx \boxed{0.44217}\,.$$

9. (20 points) An ear thermometer measures the body temperature (more specifically, the temperature inside your ear canal) using an infrared sensor. Like any other measurement device, an ear thermometer is subject to errors. The reading on the thermometer can be modeled as

$$T = \theta + R\,,$$

where θ is the actual (non-random, but unknown) body temperature and R is the random error. We assume that R is normally distributed with mean 0 and an unknown standard deviation σ.

(a) Explain why repeating the measurement a few times and taking the average of the readings gives a better estimate for θ than just a single measurement.

Solution: Let $T_1, T_2, \ldots, T_n$ be the results of n independent measurements and $\overline{T}_n = (T_1 + T_2 + \cdots + T_n)/n$ their average. Then, $\mathbb{E}[\overline{T}_n] = \mathbb{E}[T] = \theta$ and $\mathbb{V}\mathrm{ar}[\overline{T}_n] = \mathbb{V}\mathrm{ar}[T]/n = \sigma^2/n$, (see Part (b)). Therefore, the standard error of the measurement is reduced by a factor of $\sqrt{n}$. For instance, repeating the measurement 4 times will half the standard error.

Intuitively, in some measurements, the error is positive, while in others the error is negative. When we take the average, these errors partially compensate for one another, resulting in a smaller typical error.

Let $T_1, T_2, \ldots, T_5$ be the results of 5 separate (independent) measurements, and

$$\overline{T} = \frac{T_1 + T_2 + \cdots + T_5}{5} \qquad \text{and} \qquad S^2 = \frac{1}{4}\sum_{i=1}^{5}(T_i - \overline{T})^2$$

their sample mean and sample variance.

(b) Compute $\mathbb{E}[\overline{T}]$ and $\mathbb{V}\mathrm{ar}[\overline{T}]$ in terms of the parameters of the model.

Solution: We have

$$\begin{aligned}
\mathbb{E}[\overline{T}] &= \mathbb{E}\left[\frac{1}{5}(T_1 + T_2 + \cdots + T_5)\right] \\
&= \frac{1}{5}\big(\mathbb{E}[T_1] + \mathbb{E}[T_2] + \cdots \mathbb{E}[T_5]\big) && \text{(linearity of expectation)} \\
&= \frac{1}{5} \times 5\,\mathbb{E}[T] = \mathbb{E}[T] \\
&= \mathbb{E}[\theta + R] = \theta + \mathbb{E}[R] \\
&= \boxed{\theta}\,. \\
\mathbb{V}\mathrm{ar}[\overline{T}] &= \mathbb{V}\mathrm{ar}\left[\frac{1}{5}(T_1 + T_2 + \cdots + T_5)\right] \\
&= \frac{1}{25}\big(\mathbb{V}\mathrm{ar}[T_1] + \mathbb{V}\mathrm{ar}[T_2] + \cdots \mathbb{V}\mathrm{ar}[T_5]\big) && \text{(independence of } T_1, T_2, \ldots, T_5) \\
&= \frac{1}{25} \times 5\,\mathbb{V}\mathrm{ar}[T] = \frac{1}{5}\,\mathbb{V}\mathrm{ar}[T] \\
&= \frac{1}{5}\,\mathbb{V}\mathrm{ar}[\theta + T] = \frac{1}{5}\,\mathbb{V}\mathrm{ar}[R] \\
&= \boxed{\frac{1}{5}\sigma^2}\,.
\end{aligned}$$

(c) Based on these 5 measurements, construct an interval estimator for θ with confidence level 98%.

Solution: Since R is assumed to be normally distributed, so is T. Furthermore, by the stability of the normal distribution, the sample mean $\overline{T}$ is also normally distributed. The parameters of this normal distribution are the ones derived in Part (b). Since σ is unknown, we will use its estimate provided by S.

We know that $Z := \frac{\overline{T}-\theta}{\sigma/\sqrt{5}}$ (the standardized form of $\overline{T}$) has the standard normal distribution, whereas $W := \frac{\overline{T}-\theta}{S/\sqrt{5}}$ has the t-distribution with $5 - 1 = 4$ degrees of freedom. Therefore, for every $a > 0$,

$$\begin{aligned}
\mathbb{P}\left(\overline{T} \pm a\frac{S}{\sqrt{5}} \text{ includes } \theta\right) &= \mathbb{P}\left(\theta \pm a\frac{S}{\sqrt{5}} \text{ includes } \overline{T}\right) \\
&= \mathbb{P}\left(\theta - a\frac{S}{\sqrt{5}} < \overline{T} < \theta + a\frac{S}{\sqrt{5}}\right) \\
&= \mathbb{P}\left(-a < \frac{\overline{T} - \theta}{S/\sqrt{5}} < a\right) \\
&= F_{\mathsf{T}(4)}(a) - F_{\mathsf{T}(4)}(-a) \\
&= 2F_{\mathsf{T}(4)}(a) - 1\,.
\end{aligned}$$

where $F_{\mathsf{T}(4)}$ is the cdf of the t-distribution with 4 degrees of freedom. Thus, for every $a > 0$, we obtain an interval estimator $\overline{T} \pm a\frac{S}{\sqrt{5}}$ for θ that has confidence level $2F_{\mathsf{T}(4)}(a) - 1$.

In order to have a confidence level of 98%, we must find $a > 0$ such that $2F_{\mathsf{T}(4)}(a) - 1 = 0.98$, or equivalently, $F_{\mathsf{T}(4)}(a) = 0.99$. Using the app for the t-distribution (with 4 degrees of freedom), we find $a \approx 3.74695$.

In conclusion, we find the interval estimator

$$\boxed{\overline{T} \pm 1.675687S \quad \text{(with 98\% confidence)}}$$

for the true temperature θ. (Here, $1.675687 = 3.74695/\sqrt{5}$.)

Suppose you measure your temperature 5 times in a row and find values

$$36.9\ ^\circ\mathrm{C} \qquad 36.8\ ^\circ\mathrm{C} \qquad 37.0\ ^\circ\mathrm{C} \qquad 36.8\ ^\circ\mathrm{C} \qquad 36.8\ ^\circ\mathrm{C}$$

which have mean 36.86 °C and standard deviation 0.08944 °C.

(d) Use the latter interval estimator to give a 98% confidence interval for your body temperature.

Solution: Substituting $\overline{T} = 36.85$ °C and $S = 0.08944$ °C in the interval estimator found in Part (c), we obtain the confidence interval

$$\boxed{36.85 \pm 0.15\ ^\circ\mathrm{C} \quad \text{(with 98\% confidence)}}$$

or equivalently,

$$\boxed{(36.70\ ^\circ\mathrm{C}, 37.00\ ^\circ\mathrm{C}) \quad \text{(with 98\% confidence)}}$$

for your body temperature.