Body temperature investigation

DAPR1 staff

Note: this is not a report, but a live programming lecture transcript!

1 Body temperature data

We will be using data comprising measurements on body temperature and pulse rate for a random sample of $n = 50$ healthy subjects. The data are stored at the following address: https://uoepsy.github.io/data/BodyTemp.csv

We are interested in estimating what is the average body temperature for healthy individuals. Next week, we will investigate if this is significantly different from the commonly though value of 37°C.

It would be extremely time-consuming and costly to take everyone's body temperature. Instead, the data were collected for a simple random sample of healthy humans and we will use the average body temperature of this sample to estimate the true population mean.

1.1 Simple random sampling (SRS)

Simple random sampling is a type of sampling technique. Sampling techniques are used by companies, researchers and individuals for a variety of reasons. Sampling strategies are useful when conducting surveys and answering questions about populations. There are many different methods researchers can use to obtain individuals to be in a sample. These are known as sampling methods.

Simple random sampling is, unsurprisingly, the simplest form of probability sampling: every member in the population has an equal chance of being selected in the sample. Individuals are usually selected by a random number generator or some other mean of random sampling.

The biggest benefit of SRS is it removes bias, as everyone has an equal chance of being selected. Furthermore, the sample is representative of the population.

1.2 Questions of interest

We are going to investigate the average body temperature for all healthy humans.

- This week: What is the average body temperature for healthy humans?
- Next week: Has the average body temperature for healthy humans changed from the long-thought 37 °C?

2 Introduction

The data available at https://uoepsy.github.io/data/BodyTemp.csv comprise measurements of the body temperature (`BodyTemp`, in Celsius) and pulse rate (`Pulse`) for a sample of 50 healthy individuals. We are interested in estimating the mean body temperature for all healthy humans and testing whether this is different from the commonly thought value of 37 °C. As such, for the purpose of this investigation we will only focus on the variable `BodyTemp`.

Table 1: Descriptive statistics of body temperatures (°C)

n	M	SD	Min	Max
50	36.81	0.43	35.78	38.22

3 Analysis

Figure 1 shows that the body temperatures of the sample of 50 healthy individuals follow roughly a bell-shaped distribution, with most values between 36.5 and 37.5 °C and fewer in the tails of the distribution. No values were lower than 35.75 or larger than 38.22. The average body temperature in the sample was 36.81 °C, see Table 1, with a SE of 0.06 and 95% CI [36.69, 36.93]. Hence, we are 95% confident that the average body temperature for a healthy individual is between 36.69 °C and 36.93 °C.

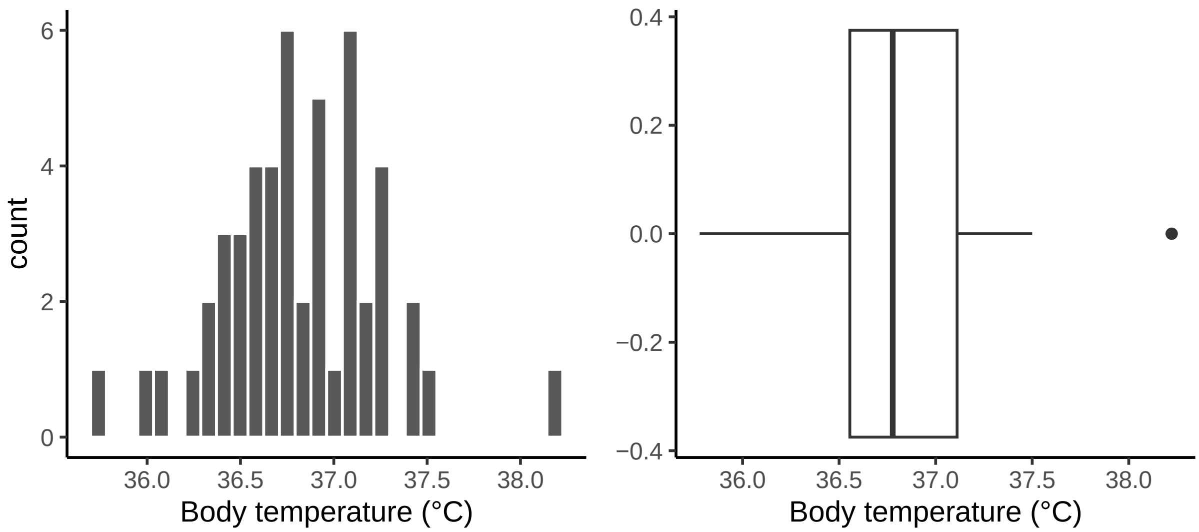

Figure 1: Distribution of body temperatures (°C)

At the 5% significance level, we tested whether the average body temperature of healthy individuals was significantly different from the commonly thought value of 37 °C. The sample data provide very strong evidence against the null hypothesis and in favour of the alternative one that the mean body temperature is different from the hypothesised value of 37 °C, $t(49) = -3.14, p = .003$, two-sided.

4 Discussion

This report estimated the mean body temperature of all healthy humans and tested whether, at the 5% significance level, that value is different from the long-thought 37°C.

The sample data provided very strong evidence that the mean body temperature of all healthy humans is different from 37 °C. Furthermore, we are 95% confident that the value is between 36.69 °C and 36.93 °C, i.e. between 0.07 and 0.31 lower than the commonly thought value of 37 °C.

4.1 Appendix

```
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE)
###### week 1 code
```

```r
library(tidyverse)
library(patchwork)
library(kableExtra)

temp_data <- read_csv('https://uoepsy.github.io/data/BodyTemp.csv')

temp_data %>%
    select(BodyTemp)

temp_data <- temp_data %>%
    select(BodyTemp)

dim(temp_data)
head(temp_data)
glimpse(temp_data) # or str()
summary(temp_data)

plt.h <- ggplot(temp_data, aes(x = BodyTemp)) +
    geom_histogram(color = 'white') +
    labs(x="Body temperature (°C)") +
    theme_classic()

plt.d <- ggplot(temp_data, aes(x = BodyTemp)) +
    geom_density() +
    labs(x="Body temperature (°C)") +
    theme_classic()

plt.b <- ggplot(temp_data, aes(x = BodyTemp)) +
    geom_boxplot() +
    labs(x="Body temperature (°C)") +
    theme_classic()

plt.h / plt.b

# Option 1: no descriptives table
xbar <- mean(temp_data$BodyTemp)
s <- sd(temp_data$BodyTemp)
n <- nrow(temp_data)
se <- s / sqrt(n)
tstar <- qt(c(0.025, 0.975), df = n - 1)
xbar + tstar * se

# Option 2: descriptives table
stats <- temp_data %>%
    summarise(
        n = n(),
        M = mean(BodyTemp),
        SD = sd(BodyTemp),
        Min = min(BodyTemp),
        Max = max(BodyTemp)
        )
stats
```

```r
tstar <- qt(c(0.025, 0.975), df = stats$n - 1)
stats$M + tstar * (stats$SD / sqrt(stats$n))

kbl(stats, booktabs = TRUE, digits = 2,
    caption = "Descriptive statistics for Body Temperature (°C)")


# week 2 code (draft)

tobs <- (stats$M - 37) / (stats$SD / sqrt(stats$n))
tobs

pvalue <- 2 * pt(abs(tobs), df = stats$n - 1, lower.tail = FALSE)
pvalue

t.test(temp_data$BodyTemp, mu = 37)


###### week 2 code

# Step 1. Specify null and alternative hypotheses
# H0 : mu = 37
# H1 : mu not equal to 37

# Step 2. Compute t-statistic
# t_obs = (xbar - mu0) / SE     where SE = s / sqrt(n)

t_obs <- (stats$M - 37) / (stats$SD / sqrt(stats$n))
t_obs

# Step 3. Compute the p-value
pvalue <- 2 * pt(abs(t_obs), df = stats$n - 1, lower.tail = FALSE)
pvalue

# Step 4. Make a decision by comparing the p-value to alpha (significance level)
# - Reject H0 if pvalue <= alpha
# - Do not reject H0 if pvalue > alpha

pvalue <= 0.05

# Step 5. Writing up


###### week 3 code (draft)
t_obs <- (stats$M - 37) / (stats$SD / sqrt(stats$n))
t_obs

qt(c(0.025, 0.975), df = stats$n - 1)

t.test(temp_data$BodyTemp, mu = 37)


###### week 3 code
```

```r
# Step 1. Specify null and alternative hypotheses
# H0 : mu = 37
# H1 : mu not equal to 37

# Step 2. Compute t-statistic
# t_obs = (xbar - mu0) / SE     where SE = s / sqrt(n)

t_obs <- (stats$M - 37) / (stats$SD / sqrt(stats$n))
t_obs

# Step 3. Compute the critical values from the null distribution t(n-1)
tstar <- qt(c(0.025, 0.975), df = stats$n - 1)
tstar

# Step 4. Make a decision by comparing the observe t-statistic to the critical values -t^* and +t^*

t_obs
tstar

# Step 5. Writing up


plt.h | plt.b
kbl(stats, booktabs = TRUE, digits = 2,
    caption = "Descriptive statistics of body temperatures (°C)")
```