

Quiz 1 — 1/18/2023

Instructions

This take-home quiz is due on **Wednesday, January 18 at 23:59**.

You may use your own course materials, as well as any materials directly linked from the course website. **No collaboration allowed.**

Type your answers **directly in this Jupyter notebook**, and submit this notebook (just the `ipynb` file) using the submission form on the [course website](#).

Problem 1

Suppose X has a t -distribution with 10 degrees of freedom.

Write R code to compute the following.

(a) $P(X \leq 0.75)$

In []:

(b) $P(X > -0.25)$

In []:

(c) $P(-0.4 < X \leq 0.1)$

In []:

Feedback. Most of you had the right idea with this problem.

Some things to remember about `pt()` vs `qt()` in R. Assume X has a t -distribution with df degrees of freedom:

1. `pt(x, df)` = $P(X \leq x)$. See Problem 4 in the Lesson 2 Exercises.
2. `qt(p, df)` = the value of x such that $P(X \leq x) = p$, or in other words, the p -quantile. See Problem 1 in the Lesson 2 Exercises.
3. For part (b), note that $P(X > -0.25) = 1 - P(X \leq -0.25)$. You can compute $P(X \leq -0.25)$ using `pt()`; see item 1 above.

Problem 2

In the same folder as this notebook, there is a CSV file `data/HumanTemp.csv`, with three columns/variables: `Temp`, `Sex`, and `Pulse`.

Write R code to do the following.

(a) Find the median of the values in `Pulse` . (Read the data first!)

In []:

(b) Compute the 3rd quantile of the values in `Pulse` . Do **not** use `summary()` .

In []:

(c) Create a normal QQ-plot of the values in `Pulse` .

In []:

(d) Based on your answer to (c), do the values in `Pulse` approximately follow a Normal distribution? Briefly explain.

Write your answer here. Double-click to edit.

Feedback. Most of you had the right idea with parts (a)-(c). For a similar problem, see Problem 5 in the Lesson 2 Exercises.

For part (d): see the **Normal Q-Q plot** section in Lesson 2. Also, make sure your explanation is *precise*: use the correct words. In addition, keep your language simple. Some tips:

- Be specific. For example, consider the following sentence:

It follows a Normal distribution because it roughly is a straight line.

What does "it" refer to here? The first "it" refers to the **values in `Pulse`** , and the second "it" refers to the Normal Q-Q plot.

- The Normal Q-Q plot consists of *points*, not values. You can say:

The points in the Normal Q-Q plot do not approximately follow a straight line.

You can also say:

The Normal Q-Q plot is not an approximately straight line.

But you should **not** say:

The values do not follow a straight line.

- You should be able to answer this problem using the following template:

The values in **`Pulse`** (do or do not) approximately follow a Normal distribution because the points in the Normal Q-Q plot (fill in the blank here).

Grading rubric

Problem	Weight
1a	0.5
1b	0.5
1c	0.5
2a	0.5
2b	0.5

Problem	Weight
2c	0.5
2d	1
Max Score	40