STAT 551 Case 06 Report

One Size Does Not Fit All: Assuming the Same Normal Temperature for Everyone is Not Justified

Consultants: Harper Cheng, Lily Xia
Client: David Abbott

Abstract

To investigate the variation of body temperature at the population level (between-individual variation) as well as at the individual level (within-individual variation), we propose several statistical techniques to address these two questions separately. We recommend one sample t-test and Levene's statistic for examining between-individual variation, and the use of boxplot for analyzing within-individual variation. One important advice is to always inspect the data with exploratory analyses and support hypothesis testing results with data visualization and summary statistics.

1 Introduction

The human body temperature at 37°C has long been used as a universally accepted approximation for one's normal temperature. This standard is recently questioned by researchers as they concern that this approximation might not be an accurate estimation for some people due to between-individual temperature variation. The goal of this research is to corroborate that assuming 37°C as the normal temperature for everyone is unjustified. Instead of attempting to estimate the normal temperature for the entire population, researchers hope to implement new protocols on routine temperature-taking at an individual level such that physicians can diagnose fever based on patients' own normal temperature ranges.

2 Statistical Questions

We believe that you are seeking advice on testing temperature variation, both at the population and the individual level. For the sake of convenience, we use terms "between-individual variation" and "within-individual variation" to differentiate these two perspectives, though the two terms might less accurate for addressing the corresponding statistical questions. Initial questions formed for the study are modified based on our discussion during the meeting to make the statistical questions more relevant. We have formulated the hypotheses as follows:

1. **Between-individual Variation:** The average standard deviation of normal body temperature is within 0.325°C. The standard deviation of body temperature is homogeneous over individuals.

2. **Within-individual Variation:** A given individual's temperature is fairly stable. In other words, the inter-day standard deviation within an individual is relatively small (compared with 0.325°C).

Standard deviation is used to quantify the amount of variability. 0.325°C is chosen as the cutoff for the following reasons. After some literature review, we have decided to use 36.2 to 37.5° as the range for normal body temperature [2]. A rule of thumb for estimating standard deviation from range states that standard deviation is about one fourth of the range [1]. Therefore, $\frac{37.5-36.2}{4} = 0.325$°C was selected as the cutoff for testing the scale of standard deviation. You are welcome to use other values to reflect the normal body temperature range, but the approximation on standard deviation from range still applies.

3 Data Description and Collection

96 adults (42 men and 54 women) between the age of 18 and 67 were recruited. They were instructed to take their oral temperatures using digital thermometers of the same brand with accuracy of 0.1°C over the range of 34°C to 42°C. Each participant took seven pairs of temperature readings in the morning and evening over a two-week period, resulting in roughly 14 readings per person and a total of 1333 readings (excluding missing values). Covariates that are believed to be significantly related to body temperature are age, gender, and time of day. Considering the strict recruitment protocol and detailed instructions on when and how temperature should be taken, we believe that study subjects were random and independent and that the data collected can be considered accurate.

4 Proposed Statistical Methods

Descriptive analyses and statistical tests have been performed with statistical software R on real data you sent to us. A few measurements have been found to fall outside the detection limit of the thermometer. This is likely to have been found to caused by human errors and we suggest to remove these erroneous observations before conducting analyses. For the purpose of demonstration, we have performed the following analyses with 0.325°C as the hypothesized value (or cutoff) and with outliers removed.

4.1 Between-individual Variation

This section is on testing variation of temperature at the population level. We propose one sample one-sided t-test for testing if the average standard deviation of normal body temperature is statistically larger than the cutoff. Additionally, we confirm the validity of using Levene's statistic (as proposed by you) for testing homogeneity.

One Sample One-sided t-test

First of all, summary statistics are explored prior to hypothesis testing. To understand the characteristics of the data and avoid relying solely on the statistical test or its p-value. Table 1 summarizes the mean and range of standard deviation of individual's body temperature as well as the proportion of people with standard deviation higher than 0.325°C. For both male and female, the mean value is smaller than the hypothesized value and less than half of the participants have experienced large temperature fluctuation. A descriptive summary on the data suggests that variation in body temperature is likely to be small, regardless of gender.

Table 1: Summary Statistics for Standard Deviation of Individual's Body Temperature

	Mean	**Range**	**Proportion of people with SD Higher than 0.325°C**
Female	0.30	(0.074, 0.66)	42%
Male	0.32	(0.14, 0.74)	48%

To use a one sample t-test, normality of the response needs to be checked to ensure the assumption for t-test is met. We can visually assess normality by using a density plot (or a histogram) or a quantile-quantile plot of the response as shown in Figure 1. In this case, the response, which is the standard deviation of one's body temperature over 14 readings, is slightly skewed to the right which means that some individuals experienced relatively large fluctuations in body temperature compared to other participants. However, the response can still be considered as normally distributed as the skewness is not severe. The one-sample t-test gives reliable results with a large sample size. With 96 participants and more than a thousand observations, the normality assumption required of the response can be relaxed. Since the assumption is satisfied, a one sample t-test can be applied to test the hypothesis.

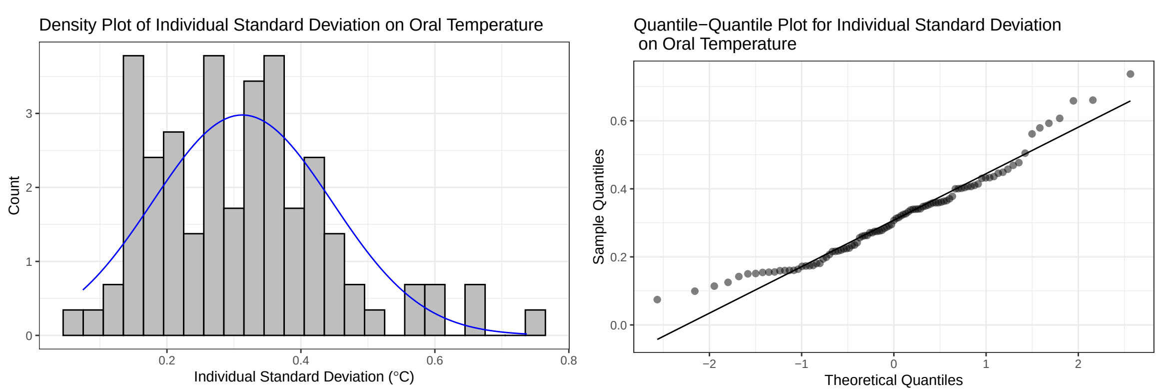


Figure 1: Checking t-test assumptions by visualizing the normality of response. To see if the response variable follows a normal distribution, a density plot (Left panel) or a quantile-quantile plot (Right panel) can be used. A normal distribution would result in a bell-shaped and symmetric curve as represented by the blue line, and points on the quantile-quantile plot would fall on the straight line.

The null hypothesis states that the average standard deviation of one's body temperature is equal to 0.325°C. A p-value of 0.82 is obtained after performing the t-test under significance level of 5%. There is no evidence against the null hypothesis due to the fact that the average standard deviation is smaller than the hypothesized value and that p-value is considerably large than 0.05.

Based on summary statistics and results from one sample t-test, the average standard deviation of normal body temperature is found to be within 0.325°C.

Levene's Statistic

To test whether different individuals have comparable variation in body temperature, we believe Levene's statistic proposed by you is suitable, but it has to be used with caution. Levene's statistic is often used to evaluate homogeneity of variances for two or more groups and it is preferred to alternative tests such as Bartlett's test in that it is less sensitive to skewed data [3]. When performing the test, each individual would be treated as a group such that variance of body temperature among different persons can be assessed. However, Levene's statistic requires the sample within each group (i.e., each individual) to be independent. It is likely that body temperature on different days for an individual is correlated, thus violating the independence assumption. Even so, Levene's test seems to be the most relevant test in addressing homoscedasticity. Since the test assumption might be violated, we suggest not to draw conclusions based solely on the p-value. Levene's test produces a p-value much smaller than 0.05 suggesting that it is possible that variance of body temperature is not to be the same for different individuals. In other words, fluctuation in body temperature might not be of the same magnitude for everyone.

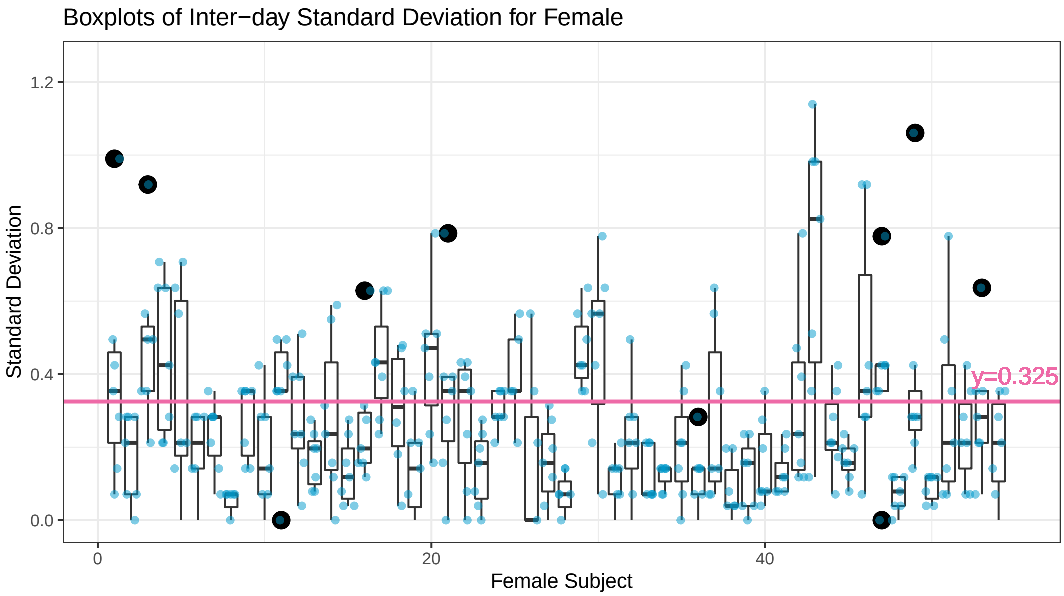
As a sanity check, result obtained from Levene's test can be verified by comparing range of standard deviation as recorded in Table 1. We can measure the difference between value V_1 and V_2 by percentage difference which is calculated as $\frac{|V_1-V_2|}{(V_1+V_2)/2} \times 100\%$. Percentage differences for standard deviation of one's body temperature are 159.7% and 136.4% for female and male, respectively. Together with results given by Levene's test, variance of body temperature is likely to be heterogeneous.

4.2 Within-individual Variation

This section is on testing variation of temperature at the individual level, or investigating inter-day standard deviation within an individual. Since most statistical tests require independent samples, finding a test that is feasible for analyzing correlated data is quite challenging. Instead of hypothesis testing, we suggest inter-day standard deviation with individual boxplot be visualized. Then, the proportion of people with mean standard deviation higher than the cutoff can be calculated to corroborate our results observed from boxplot.

Visualization with Side-by-side Boxplot

We can visually examine the distribution of inter-day standard deviation for each individual's body temperature with a boxplot. As Figure 2 suggests, the majority of participants have standard deviation below 0.325°C, albeit a few people with very high variability in day-to-day body temperature. Daily fluctuations for most people are within a narrow range as the box and whiskers are short in length. The proportion of individuals with mean standard deviation higher than 0.325°C is found to be 12.4%. Numerical calculation on the proportion as well as illustration on the distribution of daily standard deviation for one's body temperature provide evidence to support our hypothesis that a given individual's body temperature is fairly stable and daily variation is relatively small.



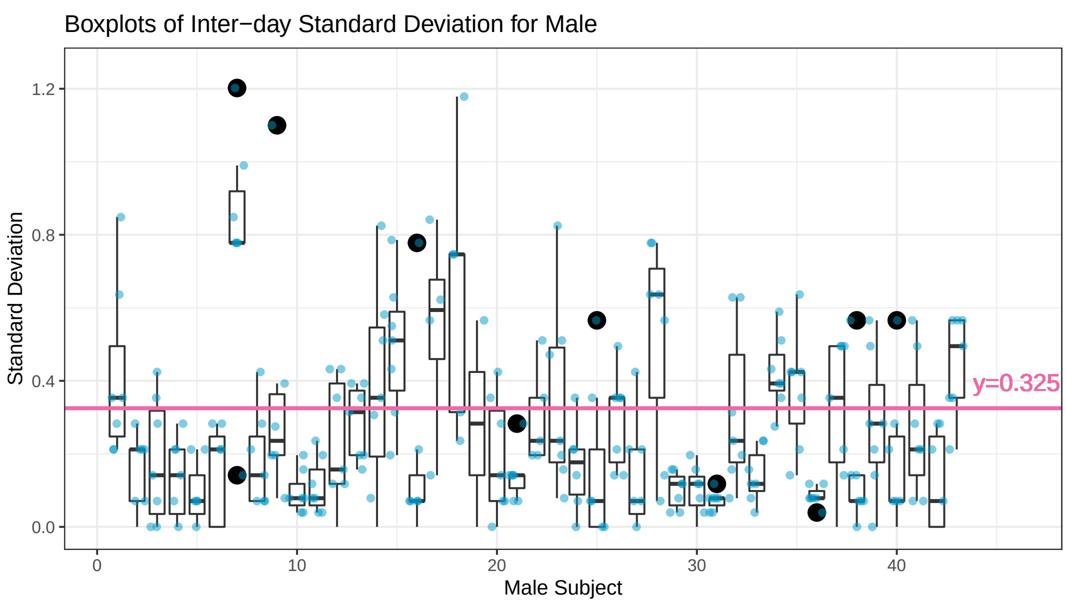


Figure 2: Visually examine the distribution of inter-day standard deviation of body temperature by gender. Actual standard deviations are depicted by blue dots. Unusual observations for each individual are represented by black dots of slightly larger size. The cutoff value 0.325°C is plotted with a pink line. Standard deviation above this cutoff is considered large.

We have noticed a few individuals with unusually high temperature variation. To elucidate what factors contribute to such large standard deviation (i.e., whether the variation comes from inter-day or intra-day variability), we have selected 31 subjects with abnormally high inter-day temperature variation (subjects who have standard deviation above 0.325°C) and plotted their temperatures in a Spaghetti plot. Figure 3 shows that body temperature within the same day is not stable for these individuals. To see if individual temperature has any trend, we have included a Spaghetti plot for the 31 subjects with each individual's temperature trajectory plotted separately. Figure 4 suggests that there is no obvious trend in temperature (i.e., slope is approximately zero) during the two-week period, that is inter-day variation does not contribute much to temperature variation. It is likely that the large within-individual variation for selected individuals is mainly attributable to large intra-day variation.

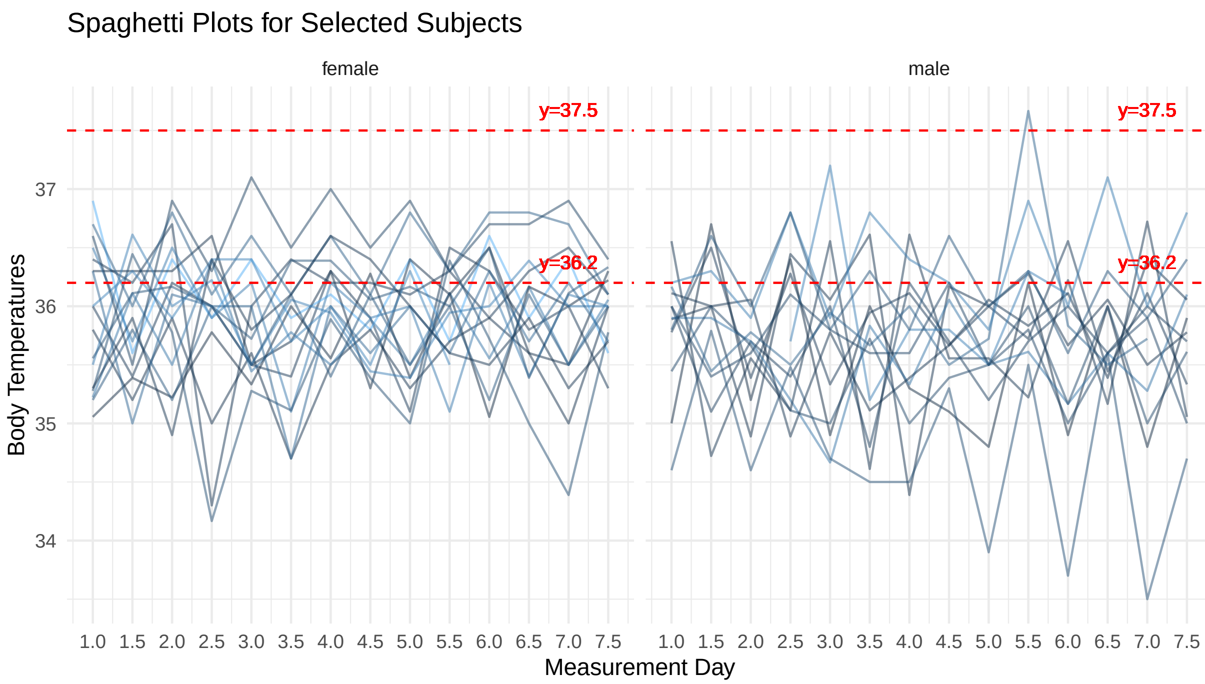


Figure 3: A Spaghetti plot on temperatures of selected individuals with abnormally high variation, that is those with inter-day average standard deviation above $0.325°$C. Males seem to experience larger fluctuation in body temperature compared to females. Body temperature within the same day is not stable for these individuals as suggested by the sawtooth pattern.

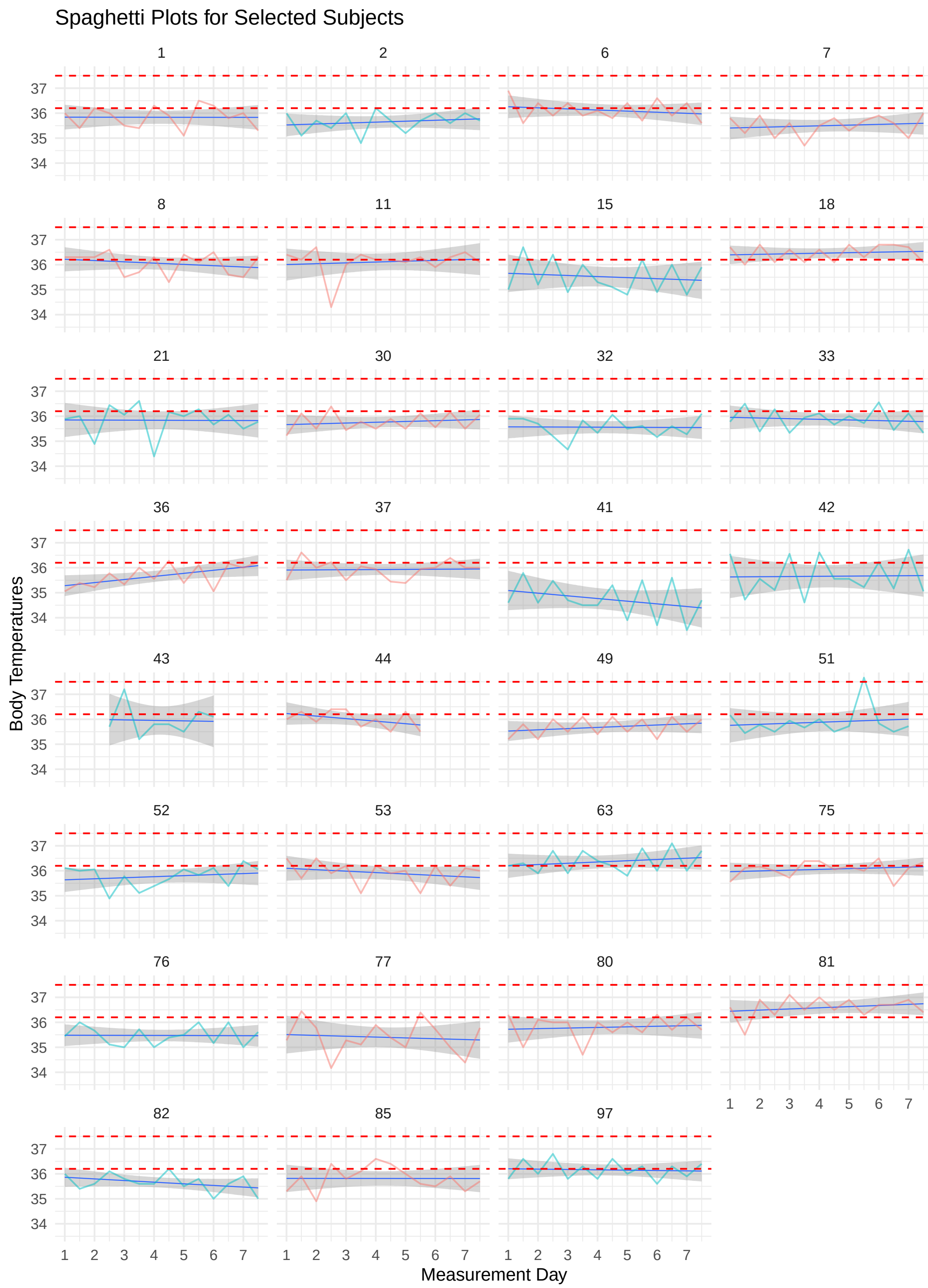


Figure 4: Spaghetti plots for selected individuals. Each plot records 14 individual body temperatures during the two-week period with the red line representing female and the blue line representing male. The linear regression line is plotted in blue with 95% confidence interval bands in gray. The regression lines have slope close to zero indicating that there is no obvious trend in body temperature over time.

5 Conclusion

We first chose standard deviation to quantify the amount of variation within an individual. The cutoff for defining how large the variation should be considered large is based on literature review, but the cutoff value can be altered according to your own reference. For between-individual variation, we propose a one sample one-sided t-test provided that assumptions for t-test be satisfied. We confirm the use of Levene's statistic for testing homogeneity. Notice that it should be used with caution because measurements within each person might be correlated. For within-individual variation, boxplot can be used to visually examine the overall distribution of temperature variation in lieu of hypothesis testing considering correlated observations within an individual. We have included some analyses to demonstrate our ideas and the use of the above-mentioned statistical tools to help with decision-making. No definitive conclusions are drawn based on the analyses conducted. As a final suggestion, it is prudent to support your results by visualization and/or summary statistics to alleviate over-reliance on hypothesis testing.

References

[1] Alfredo, R. and Charles, C. (2012). Improving on the range rule of thumb. *Rose Hulman Undergraduate Mathematics Journal*, 13(2).

[2] Chen, W. (2019). Thermometry and interpretation of body temperature. *Biomedical Engineering Letters*, 9(1):3–17.

[3] Levene, H. (1960). *Robust tests for equality of variances.* Stanford, Calif., Stanford University Press.