Springboard Project 8.3 (Human Body Temperature)

Using Human Body Temperature data

1. *Is the distribution of body temperature normal?*

   To answer this question, I first generated a histogram. The distribution was not perfectly normal, as it showed a distribution with more observations below the mode than above it. However, since the shape was approximately normal and the number of observations exceeded 29, I was comfortable treating the data as normal.

   For another glimpse at normality, I generated a Q-Q plot, which plots the ordered observations against their theoretical location in a standard normal distribution. If the plot forms a straight line increasing from left to right, the distribution is normal. Like this histogram, this plot showed an approximately normal distribution with a few aberrations in the tails, particularly the right one.

2. *Is the sample size large? Are the observations independent?*

   With respect to the criterion for using the Central Limit Theorem, I considered the sample size is large since it is greater than 29.

   Concerning independence of observations, I also treated this assumption as a reasonable one since I know it is a random sample. I also found no evidence of skew in the data set's other variables: there is an equal number of male and female observations, and a histogram for heart rates showed an approximately normal distribution.

   Where independence is concerned, though, it's worth mentioning the data was constructed based on an actual experiment, and little information is provided about data-collection methods in that experiment. Ideally, I'd have more information about those methods to answer this question with certainty.

3. *Is the true population mean really 98.6 degrees F?*

   Put briefly, I don't have the resources to answer this question. I would need to collect body temperatures from every member of the population of interest. I don't know what that population is, and if I did, it's unlikely that I would be able to readily collect all body temperatures – unless the population is a very small one (an unlikely proposition since the results would not be very generalizable).

   - *Would you use a one-sample or two-sample test? Why?*

     I chose a one-sample test because we are interested in the mean body temperature for a single population, not the mean difference between two populations.

- *In this situation, is it appropriate to use the t or z statistic?*

  I chose the t-statistic because we don't know the population's standard deviation.

- *Now try using the other test. How is the result different? Why?*

  The p-value is lower because *z* and *t* follow different distributions. This means that observations in these distributions follow different patterns for mean, shape, and spread. So when hypothesis tests are conducted, there are different values produced for measuring the statistical likelihood of observed outcomes.

4. At what temperature should we consider someone's temperature to be "abnormal"?

To answer this question, I calculated a 95% confidence interval for the mean human body temperature. I first found the standard error and then used it to scale the interval, based on a *t* distribution. A temperature outside this interval would be below 98.12 or above 98.38 degrees Fahrenheit and thus considered abnormal here.

Out of curiosity, I also looked at results according to the empirical rule, which holds that 95% of observations in a normal distribution are within 2 standard deviations of the mean. (Note: this means I've decided, fairly arbitrarily, that the most extreme 5% of temperatures would be abnormal, which is not an entirely defensible assumption for a few reasons. Again, I looked at the empirical rule largely out of curiosity.) Using this as a guideline, an abnormal temperature is below 96.79 or above 99.71 – a wider, and perhaps less useful, pair of limits.

5. Is there a significant difference between males and females in normal temperature?

- *Which test did you use and why?*

  I chose the unpaired, two-sample t-test because I am looking for a significant difference between the means of two independent groups. I performed this test assuming equal variances between the two groups based on the results of Bartlett's Test, which checks for equal variances.

- *Write a story with your conclusion in the context of the original problem.*

  Since the p-value (.024) was less than .05, I rejected the null hypothesis and concluded that there is a significant difference between the average body temperatures for males and females.