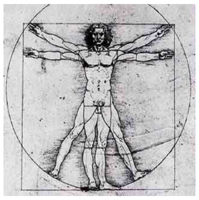

Human Proportions

1 The shape of body part measurements

The human body comes in various shapes and sizes. However, as daVinci knew, there are certain proportions that are consistent throughout. For this project two data sets are used which contain various measurements of human bodies.

To download the data sets issue these commands:

```
> source("http://www.math.csi.cuny.edu/st/R/normtemp.R")
> source("http://www.math.csi.cuny.edu/st/R/fat.R")
```

The `normtemp` data set[1] contains measurements of normal body temperature for 300 healthy adults in the variable `temperature`. The variable `gender` records the gender of the subject, and `hr` the heart rate in beats per minute.

The `fat` data set[2] contains many measurements of human bodies that can be done with a tape measure (circumference measurements), for instance the variable `wrist` contains measurements of wrist size in centimeters. Additionally, the variable `body.fat` contains body fat measurements.

After downloading the data sets, they may be attached so that the variable names are visible from the command line.

```
> attach(normtemp)
> attach(fat)
```

2 *t*-based statistical Inferences

What is normal body temperature? Common wisdom is that it is 98.6 °F. The `normtemp` data set may be used to test this assumption using a test of significance.

2.1 Basics of significance tests

We will test the *null hypothesis* that the distribution of normal body temperatures is normal with mean $\mu = 98.6$ and unspecified variance, against the *alternative hypothesis* that the mean is less than this. That is

$$H_0 : \mu = 98.6, \qquad H_A : \mu < 98.6.$$

[1]This data set was contributed to the *Journal of Statistical Education* by Allen L. Shoemaker, `http://www.amstat.org/publications/jse/v4n2/datasets.shoemaker.html`

[2]This data set was contributed to the *Journal of Statistical Education* by Roger W. Johnson, `http://www.amstat.org/publications/jse/v4n1/datasets.johnson.html`.

As we assume the data is normally distributed, we will use the T-statistic as a test statistic:

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\text{observed} - \text{expected}}{\textbf{SE}}.$$

Finally, we'll use $\alpha = 0.05$ for the *level of significance*.

Question 1: An assumption for using the t-statistic is that the data be normally distributed, or n be large enough. In this case, n is quite large, but nonetheless, verify graphically that the data appears to come from a normal distribution. Which graph did you use, and how did it show normality.

We can find the p-value with the following commands:

```
> xbar = mean(temperature)
> s = sd(temperature)
> n = length(temperature)
> T.obs = (xbar - 98.6)/(s/sqrt(n))
> pt(T.obs, df = n - 1)

[1] 1.205316e-07
```

We see that the p-value is tiny, about 10^{-7}, and is much less than the significance level α. The p-value is computed using the alternative hypothesis to specify what values of the test statistic are more extreme than the observed value, `T.obs`. In this case, as smaller values of $\bar{x}$ support H_A, we see that smaller values are more extreme.

Alternatively, one can find the critical value(s) for this problem using H_A and α. For this problem, the lone critical value corresponds to having $\alpha \cdot 100\%$ of the area to the left of the critical value. That is, the critical value is the 0.05 quantile of the t-distribution and is found with

```
> qt(0.05, df = n - 1)

[1] -1.656752
```

```
> T.obs

[1] -5.454823
```

One can see that `T.obs` is much less than the critical value.

For this problem, the computer can do all of the work with the `t.test()` function. To use this function, one must specify the data (`temperature`), the null hypothesis (`mu=98.6`), and the alternative hypothesis (`alt="less"`). The function returns the observed value of T, the degrees of freedom, the p-value and more:

```
> t.test(temperature, mu = 98.6, alt = "less")

        One Sample t-test

data:  temperature
t = -5.4548, df = 129, p-value = 1.205e-07
alternative hypothesis: true mean is less than 98.6
95 percent confidence interval:
     -Inf 98.35577
sample estimates:
mean of x
 98.24923
```

You are left to compare the p-value to α if so desired.

The alternative hypothesis may be specified three ways to the `t.test` function: $H_A : \mu < \mu_0$ is specified with `alt="less"`, $H_A : \mu > \mu_0$ is specified with `alt="greater"`, and $H_A : \mu \neq \mu_0$ is specified with `alt="two.sided"`.

Question 2: Some have suggested that 98.2°F is a better value for the mean body temperature. Do a test of significance of

$$H_0 : \mu = 98.2, \qquad H_A : \mu \neq 98.2.$$

Is the difference statistically significant at the $\alpha = 0.05$ level?

Question 3: Do females have a higher body temperature? We will ask this by testing if the mean body temperature for females is 98.2°F or is it more. That is

$$H_0 : \mu = 98.2, \qquad H_A : \mu > 98.2.$$

You can extract the temperatures for just the females in the study with this command:

```
> females = temperature[gender == "female"]
```

1. First check that using the t-statistic is appropriate for this data. That is, either the data appears to come from a normal population or the data set (n) is large enough.

   Is the statistic appropriate and why?

2. Now apply `t.test()` to find the p-value. Are the differences statistically significant at a level of $\alpha = 0.05$.

Question 4: A watch band maker assumes that the average wrist size is seven inches. We can use the `wrist` variable to check this assumption using a significance test of

$$H_0 : \mu = 7 \text{ inches}, \qquad H_A : \mu \neq 7\text{inches}.$$

1. Check that a t-test is appropriate for this data, by verifying normality of the data, or a large n. Why is the t-test appropriate?

2. As the measurements in `wrist` are in centimeters, convert the null and alternative hypotheses into centimeters (1 inch = 2.54 cms).

3. Test to see if the difference between the mean wrist size of the data is statistically significant at the $\alpha = 0.05$ level.

Question 5: The body mass index, `BMI`, is a measurement of weight divided by height squared in metric units. "Normal" values are in the range $[18, 5, 25]$. Do a significance test at the $\alpha = 0.05$ level to see if the mean BMI is 25 versus an alternative that it is more.

1. First plot the data in `BMI` and argue if the t-test is appropriate for this data and why.

2. Write down the null and alternative hypotheses

$$H_0: \qquad\qquad H_A:$$

3. What is the p-value?

4. Is the difference statistically significant?

Question 6: A manufacturer of one-size-fits-all sweatpants wishes to know whether the mean waist size (the data is stored in the variable `abdomen`) has changed from the last time it designed a pair of sweatpants. Before, the manufacturer assumed the mean waist size was 34 inches. Do a significant test at the $\alpha = 0.05$ level, testing

$$H_0: \mu = 34, \qquad H_A: \mu > 34.$$

1. Why is a t-test approporiate for the data in `abdomen`?

2. Rewrite the null and alternative hypotheses using cms (1 inch = 2.54 centimeters).

3. What is the p-value?

4. Is the difference statistically significant?

Question 7: The mean height of 17-year old Finnish boys is estimated to be 180cms. Suppose the data collected in the `height` variable is a random sample from adult males in the United States. Is the mean height statistically different from that of full grown 17-year olds from Finland at the $\alpha = 0.05$ level of significance?

1. Specify the null and alternative hypotheses using inches.

2. Comment on the appropriateness of the t-test?

3. What is the p-value?

2.2 two-sample tests

Is there a statistically significant difference between the body temperatures for males and females? Let μ_m be the population mean for males, and μ_f be the population mean for females. If we assume that the two populations are normally distributed, then a significance test of

$$H_0 : \mu_1 = \mu_2, \qquad H_A : \mu_1 \neq \mu_2$$

(or $H_A : \mu_1 < \mu_2$, or $H_A : \mu_1 > \mu_2$) can also be done using a t-distributed statistic.

The `t.test()` will still perform this test. In the two-sample case, the null hypothesis is not specified as it is always the same.

For example, the two-sample test of equivalence of means for the temperature data can be performed as follows

```
> males = temperature[gender == "Male"]
> females = temperature[gender == "female"]
> t.test(males, females, alt = "two.sided")


        Welch Two Sample t-test

data:  males and females
t = -2.2854, df = 127.51, p-value = 0.02394
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.53964856 -0.03881298
sample estimates:
mean of x mean of y
 98.10462  98.39385
```

The small p-value indicates that there is a statistically significant difference.

Question 8: The normality of the two populations (sampled in `males` and `females`) was not verified. Comment on the validity of this assumption. (A large sample size is also sufficient to use the t-test.)

Question 9: If the population variances can be assumed to be equal, then the extra argument `var.equal=TRUE` to `t.test()` will generally give smaller p-values, as the sampling distribution of the test statistic has smaller degrees of freedom in most cases. First check if this is a reasonable assumption about our data, and then check if it makes a difference in the t-test.

An alternate syntax (the model formula syntax) for performing a two-sample test when there is one variable (`gender`) which is used to indicate which level of some factor the subject has may be used. (These variables are referred to as *indicator variables.*) The t-test above, may have been carried out with the syntax:

```
> t.test(temperature ~ gender)
```

Question 10: Verify that using the model formula gives the same output and also that the default value for `alt=` is `"two.sided"`.

Question 11: Perform a two-sample t-test of the resting heart rate (`hr`) for males and females. Is the difference statistically significant at the $\alpha = 0.10$ level?

One can define indicator variables to break up a data set. For instance, a test to see if age has an effect on body fat percentage could be done with

```
> over40 = age > 40
> t.test(body.fat ~ over40)
```

Question 12: Perform the t-test described above.

1. What is the p-value reported?
2. What is the null and alternative hypotheses tested?
3. You can readily check the assumptions on the population by creating side-by-side boxplots. These are created with the same model formula syntax:

   ```
   > boxplot(body.fat ~ over40)
   ```

Question 13: The age 40 splits the data on weights (in `weight`) into two groups. Perform a significance test of

$$H_0 : \mu_{\text{40 or under}} = \mu_{\text{over 40}}, \qquad H_A : \mu_{\text{40 or under}} < \mu_{\text{over 40}}$$

where each μ is, as usual, the population mean for the population the sample comes from.

1. What is the p-value?
2. Justify you choice of test statistic.

Question 14: Are taller people less susceptible to high body fat? Define an indicator variable as follows:

```
> tall = height > 72
```

Using this, perform a significance test of

$$H_0 : \mu_{\text{6 feet or under}} = \mu_{\text{over 6 feet}}, \qquad H_A : \mu_{\text{6 feet or under}} < \mu_{\text{over 6 feet}}$$

where μ is the population mean of the body fat for the respective populations. (The data is in the variable `body.fat`.)

1. What is the p-value of the significance test?

2. Justify you choice of test statistic.

The function `wilcox.test()` for two samples performs a Wilcoxon rank-sum test of the equivalence of medians. The assumption on the populations is that they have the same shape, although perhaps different medians. They need not be normally distributed. As the population of weights is typically long-tailed, it is often best to compare sub-populations using this test instead of the t-test.

Question 15: Repeat the significance test investigating if age and weight are somehow related using the Wilcoxon rank-sum test. Compare your p-value to that returned by a t-test for the same data.

Question 16: Does a person's density (recorded in `density`) decrease with age? Use the `old` indicator variable to investigate this significance test

$$H_0 : M_{40\text{ or under}} = M_{\text{over }40}, \qquad H_A : M_{40\text{ or under}} < M_{\text{over }40},$$

where M is the subpopulation median of densities.

1. Decide if the Wilcoxon rank-sum test is appropriate for this data set. If so, compute the p-value.

2. Decide if the t-test is appropriate for this data set (where for a symmetric population the population mean and median are identical). If so, compute the p-value.

3. If both tests are appropriate, compare the p-values. Are they similar? Different?