Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Calculating a t-test by hand**

This document includes information that is important for you to know about t-tests, the process of calculating a t-test statistic, and a problem that you will calculate by hand. The goals of this are to have you practice thinking about the components that go into calculating a t-test to help you better understand what they are and how they work, and to practice interpreting the results of a t-test. Note that there [are many different types of t-test](https://en.wikipedia.org/wiki/Student%27s_t-test#Unpaired_and_paired_two-sample_t-tests) and in this case you are calculating a form of 2-sample t-statistic.

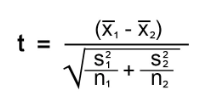
***SUBMIT THIS COMPLETED DOCUMENT BY MOODLE TUES ~~MON~~ OCT. 25 BEFORE MIDNIGHT***

Conceptually, a t-test compares the size of the difference between two means with the standard error of that difference. The numerator is the difference between the means, and the denominator is the difference in “standard error”. “Standard error” is a measure of variance in descriptive statistics that combines the standard deviations (*s*) of the groups with the sample sizes of each group. (Note the

t =

So, t =

*\* Note that “pop 1” and “pop 2” will be two groups or categories in a categorical variable in a spreadsheet (data frame).*



To calculate a t-statistic :

1. Calculate the mean of both groups (x̅1, x̅2)
2. Calculate the standard deviation for each group (s1, s2)
3. Find the population size for each group (n1 , n2)
4. Add these values to the equation and calculate the t-statistic.

To translate that t-statistic to a p-value

1. How many degrees of freedom in a two-sample t-test (df)?
   1. Df = (n1 + n2 ) - 2
2. Use the [t-statistic table](https://www.statology.org/here-is-how-to-find-the-p-value-from-the-t-distribution-table/) on the last page to find the p-value that is associated with your t-statistic and degrees of freedom. Use the 2-tailed option with a significance level of 0.05.

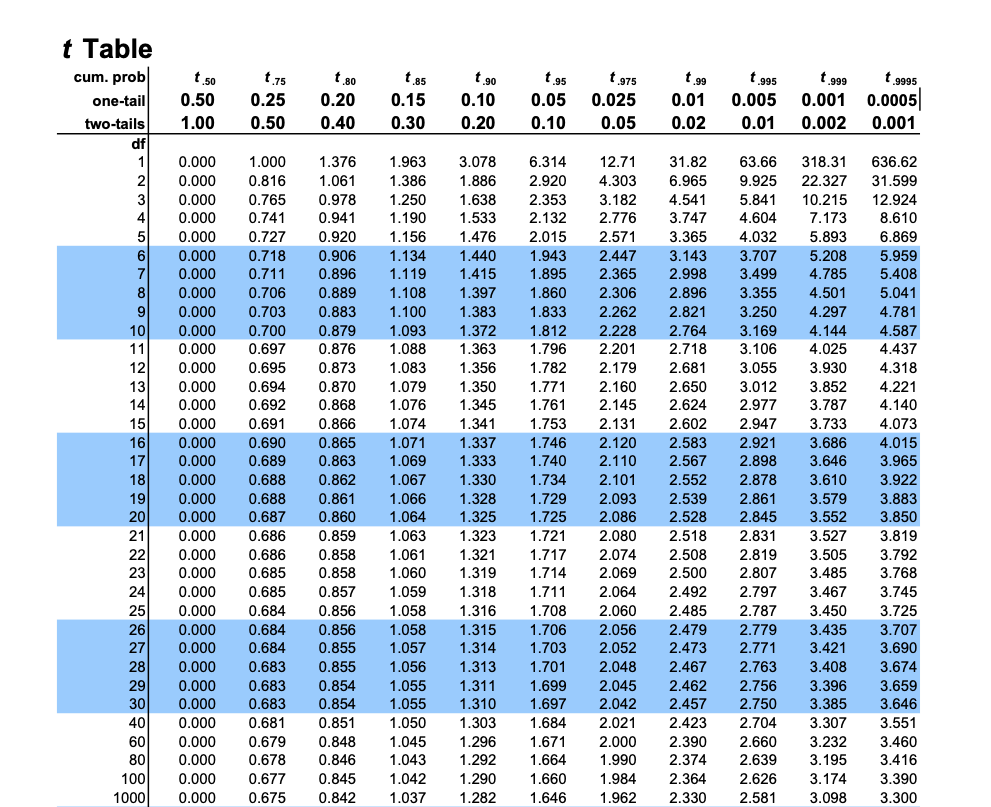


We are going to use a randomly-drawn subset from the penguins data set to compare the body mass of Adelie and Gentoo penguins.

|  |  |  |  |
| --- | --- | --- | --- |
| **species** | **island** | **body\_masss\_g** | **sex** |
| Adelie | Biscoe | 3950 | male |
| Adelie | Dream | 3425 | male |
| Adelie | Torgersen | 3350 | female |
| Adelie | Biscoe | 4300 | male |
| Adelie | Biscoe | 3600 | female |
| Adelie | Biscoe | 3825 | female |
| Adelie | Biscoe | 3950 | male |
| Adelie | Torgersen | 3450 | female |
| Adelie | Torgersen | 4300 | male |
| Adelie | Biscoe | 2925 | female |
| Gentoo | Biscoe | 4875 | female |
| Gentoo | Biscoe | 5600 | male |
| Gentoo | Biscoe | 5950 | male |
| Gentoo | Biscoe | 5050 | male |
| Gentoo | Biscoe | 4450 | female |
| Gentoo | Biscoe | 4400 | female |
| Gentoo | Biscoe | 4150 | female |
| Gentoo | Biscoe | 5000 | female |
| Gentoo | Biscoe | 4750 | female |
| Gentoo | Biscoe | 4400 | female |

|  |  |  |
| --- | --- | --- |
|  | Adelie | Gentoo |
| Mean body mass |  |  |
| Standard deviation of body mass |  |  |
| Sample size |  |  |

How many degrees of freedom?



What is the critical value for your t-statistic given your degrees of freedom and a significance level of 0.05? What does this value mean?

Write a sentence or two that accurately reports your t-statistic, the p-value associated with it the critical value, and what that means in terms of whether or not the two populations you are comparing are statistically significantly different.