t Procedures

Worksheet 22*

1. Genetic evolution is based on mutation. Consequently, one fundamental question in evolutionary biology is the rate of *de novo* mutations, i. e., those mutations that neither parent possessed. For example, these mutations can occur in the germs cells, the egg or the sperm of the parent. To investigate this question in humans, Kong et al, sequenced the entire genomes of 78 Icelandic trios (mother, father, offspring) and recorded the number of *de novo* mutations in the offspring. The mean $\bar{x} = 63.24$ and the standard deviation $s = 14.18$.

   (a) Give a 95% confidence interval for the mean number of mutations.

   (b) A mutation rate of one mutation in one-hundred million base pairs would lead to a mean number of *de novo* mutations of 60. State a hypothesis test for this situation.

   (c) What does the confidence interval in (a) say about a test with a 5% significance level?

   (d) Give the t statistic for this situation and state whether or not the test can be rejected at the 1% level. Explain your answer.

   (e) Using the standard deviation above, fill in the following table for the power for the given alternatives.

alternative hypothesis	significance level	power	number of observations
63	0.05	0.80	
63	0.02	0.80	
63	0.05		80
63	0.05		240
62	0.05	0.80	
62	0.02	0.80	
62	0.05		80
62	0.05		240

2. July and August are the strongest months for monsoon rains in Tucson. Your friend suspects that more rain takes place in July.

   (a) State a hypothesis appropriate to you friend's suspicion.

   (b) For the years 2007 to 2016, here are the data for rainfall total in inches.

	2007	2008	2009	2010	2011	2012	2013	2104	2015	2016
July	4.38	4.03	2.15	1.44	1.69	0.71	1.44	4.58	2.06	3.32
August	4.27	1.69	0.86	2.89	1.03	0.00	1.85	3.32	1.80	1.09

(c) What test procedure will you use in this case? Explain your answer.

(d) Compute the appropriate t statistics for the hypothesis in part (a).

(e) What is the conclusion concerning your hypothesis? Use you answer in part (c) to explain your conclusion.

(f) If you incorrectly used the two-sample t procedure, how does the p-value change?