

Inferential Statistics Exercises

8.1

The mean for diet A is higher than diet B, implying that when people did lose weight, losses were higher on diet A than B. Standard deviation to mean ratio was lower for diet B which I am interpreting as weight loss results were less consistent. 28 people lost more than the mean on diet A, 25 lost more than the mean on diet B, a 6% difference yet the mean to stdev ratio was 2.1 to 1.34, considerably different.

8.2

Those on diet A did lose about 50% more weight than those on diet B but both diets did work for those that experienced weight loss. The median and mean values for both diet A and diet B are close, 5.341/5.642 and 3.710/3.745 implying there are not likely a large number of outliers in the data, or weight loss is fairly consistent.

8.3

60% of people in area 1 preferred a cereal other than brands A or B, in area 2 Brand B has 33% market share and 55% of people in area two preferred brands B or A over any other choice. Brand B is consistently more popular in both areas.

8.4

Presuming the dataset was to be taken from the Data Annexe doc since there was no excel file with the name in the exercise.

Batch	Agent1	Agent2
1	7.7	8.5
2	9.2	9.6
3	6.8	6.4
4	9.5	9.8
5	8.7	9.3
6	6.9	7.6
7	7.5	8.2
8	7.1	7.7
9	8.7	9.4
10	9.4	8.9
11	9.4	9.7
12	8.1	9.1

t-Test: Paired Two Sample for Means

	Agent1	Agent2
Mean	8.25	8.683333
Variance	1.059090909	1.077879
Observations	12	12
Pearson Correlation	0.901055812	
Hypothesized Mean Difference	0	
df	11	
	-	
t Stat	3.263938591	
P(T<=t) one-tail	0.003772997	

t Critical one-tail	1.795884819
P(T<=t) two-tail	0.007545995
t Critical two-tail	2.20098516

Difference in Means	0.433333333
---------------------	-------------

Running the formulas was straightforward, but after reading the first portion of chapter 9 in business stats text and a great deal of additional stats tutorials for a few hours I have no confidence in my interpretation what-so-ever, such is the reality of asynchronous learning and very short deadlines. Referencing the text from example 8.4 "The associated two-tailed p-value is $p = 0.018$, so the observed t is significant at the 5% level (two-tailed)." there is no indication of why 0.018 is significant at the 5% level, the assumption is the student actually understands what they are interpreting which isn't necessarily the case.

- Is the two tail P value less than 0.05, yes
- Are the filtration agents different, no idea because the difference in means is minimal, whereas the example worked through with the container displays was almost 10% of the sales volume

8.5

Similar explanation as in 8.4,

- Is the one tail P value less than 0.01, yes it is 0.38%
- Unsure how this relates to the difference in means which is very low so I am still assuming this implies the two agents are roughly the same.

Also noticed the excel formula for standard deviation, STDEV(), presumes this is only a sample, calculating the sum of individual variance values as $n-1$.

<https://www.youtube.com/watch?v=HvDqbzu0i0E>

8.6

The Null hypothesis if the mean income for male cardholders exceeds the mean income for female cardholders.

Averaging the mean incomes of each sample groups shows arithmetically that the male sample group had a higher income, 52,900 pounds VS. 44,200 pounds.

There was considerably more variance between high and low incomes in the male sample group which could be skewing the mean higher. As a result, the median was calculated for each sample identifying that both sample groups were positively skewed so the sample sets have similar shape. With similar sample shapes the mean value can be considered a strong indicator of average annual income.

			F-Test Two-Sample for Variances		
Male Cardholders	n	60		<i>Variable 1</i>	<i>Variable 2</i>
	Mean	52.9			
	SD	15.26856	Mean	52.91333	44.23333
	Median	52.1	Variance	233.129	190.1758
Female Cardholders	n	60	Observations	60	60
	Mean	44.2	df	59	59
	SD	13.79042	F	1.22586	
	Median	38.2	P(F<=f) one-tail	0.218246	
			F Critical one-tail	1.539957	
			P2	0.436492	

A second test was performed, t-Test Two-Sample Assuming Equal Variances. According to the course textbook, "Basic Business Statistics", if the t Stat value exceeds the t Critical value the null hypothesis can not be rejected. A review of the tested output below confirms this to be the case, further evidence that the mean value assessment in the first test is correct.

t-Test: Two-Sample Assuming Equal Variances

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	52.91333333	44.23333333
Variance	233.1289718	190.1758192
Observations	60	60
Pooled Variance	211.6523955	
Hypothesized Mean Difference	0	
df	118	
t Stat	3.267900001	
P(T<=t) one-tail	0.000709735	
t Critical one-tail	1.657869522	
P(T<=t) two-tail	0.00141947	
t Critical two-tail	1.980272249	
P2	0.002838941	