Unit 7: Hypothesis Testing Worksheet

Exercise 7.1 The Related Samples T Test

Q: Recall that in the previous unit exercises, a two-tailed test was undertaken whether the population mean impurity differed between the two filtration agents in Data Set G. Suppose instead a one-tailed test had been conducted to determine whether Filter Agent 1 was the more effective. What would your conclusions have been?

A: The results of the Hypothesis testing are shown in the picture below , and the Formulation of the Null Hypothesis for the one-tailed test would be that Agent 1 is not more effective than Agent 2 (i.e. (H_0 is $\mu_1 >= \mu_2$), while the Alternative Hypothesis H_1 is that Agent 1 is more effective (i.e. $\mu_1 < \mu_2$).

When comparing the test data *t Stat* and *t Critical one-tail*, we find that *t Stat* < - *t Critical one-tail,* and we also see that *P(T<=t) one-tail t* < 0.05 therefore we can reject the null hypothesis.

According to the testing, Agent 1 is significantly more effective than Filter Agent 2 at reducing impurities, and this conclusion is the same as with the one from the two-tailed test that Agent 1 is the recommended choice for the filtration process.

	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

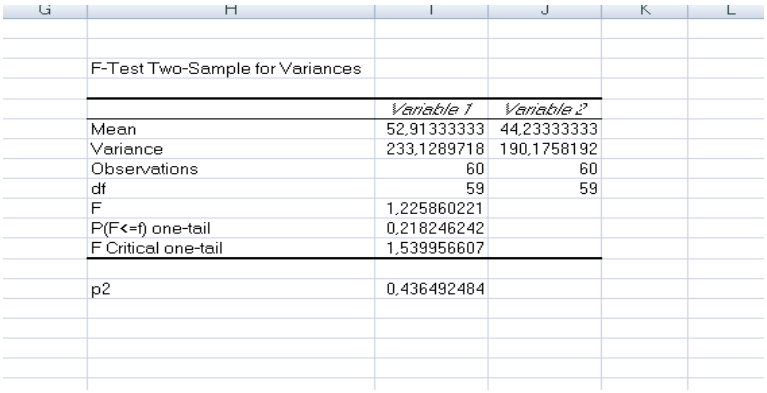
	Agent1	*Agent2*
Mean	8,25	8,683333333
Variance	1,059091	1,077878788
Observations	12	12
Pearson Correlation	0,901056	
Hypothesized Mean Difference	0	
df	11	
t Stat	-3,26394	
P(T<=t) one-tail	0,003773	
t Critical one-tail	1,795885	
P(T<=t) two-tail	0,007546	
t Critical two-tail	2,200985	

Difference in Means	-0,43333

Exercise 7.2 The Independent Samples T Test

Q: Consider the bank cardholder data of Data Set C. Open the Excel workbook Exa8.6C.xlsx which contains this data from the Exercises folder. Assuming the data to be suitably distributed, complete an appropriate test of whether the population mean income for males exceeds that of females and interpret your findings. What assumptions underpin the validity of your analysis, and how could you validate them?

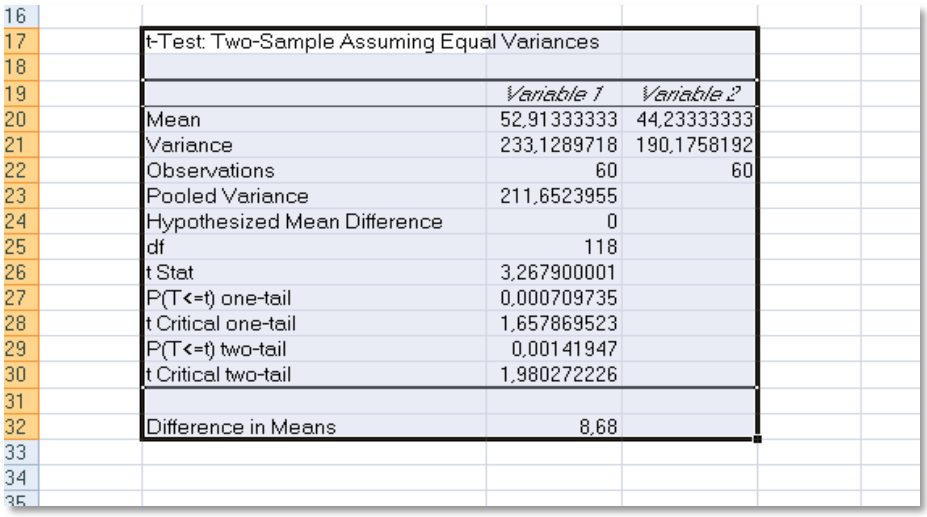
A": We assume that the population variances underlying the incomes of men and women do not differ, and we use the F-test Two-sample for Variances with the null hypothesis being that the variances are significantly different and the alternate hypothesis is that the variances are equal. The results of the test are in the image below:



G	H	I	J	K	L
	F-Test Two-Sample for Variances				
		Variable 1	*Variable 2*		
	Mean	52,91333333	44,23333333		
	Variance	233,1289718	190,1758192		
	Observations	60	60		
	df	59	59		
	F	1,225860221			
	P(F<=f) one-tail	0,218246242			
	F Critical one-tail	1,539956607			
	p2	0,436492484			

From the F-Test in Excel we find that p2 = 0, 43649 (reject the null Hypothesis) and that the observed F test statistic = 1,225860221. The observed F ratio is therefore not significant and we can use the Two-Sample t-test Assuming Equal Variances to test if the population mean income for males exceeds that of females.

From this test we find that t = 3.268 with 118 degrees of freedom. The two-tailed p-value is p = 0.00142, so the observed t is significant. The results strongly suggest that the income for men is greater than that of women, with a difference in Means of 8, 68.



t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	*Variable 2*
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	1,657869523	
P(T<=t) two-tail	0,00141947	
t Critical two-tail	1,980272226	
Difference in Means	8,68	