Hypothesis Testing

8.4 Two-tailed t-test

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	8.683333333
Variance	1.059090909	1.077878788
Observations	12	12
Pearson Correlation	0.901055812	
Hypothesized Mean Difference	0	
df	11	
t Stat	-3.26393859	
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	-0.43333333	

*Note: I needed to add Analysis ToolPak in Excel to get the Data Analysis option.

The obtained related samples t = -3.264 with 11 degrees of freedom.

The associated two-tailed p-value is p = 0.0075, so the observed t is significant at the 1% level (two-tailed).

The sample mean impurity for Agent1 was 8.25 and for Agent2 was 8.68.

The data therefore constitute strong evidence that the underlying mean impurity was lower for Agent1, by an estimated 8.68-8.25 = 0.43. The results suggest that Agent1 should be preferred.

8.5 One-tailed t-test

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	8.683333333
Variance	1.059090909	1.077878788
Observations	12	12
Pearson Correlation	0.901055812	
Hypothesized Mean Difference	0	
df	11	
t Stat	-3.26393859	
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	-0.43333333	

Since impurity a lower impurity level is preferred, the hypothesis is that Agent2 gives lower impurity levels than Agent 1, making the null hypothesis that Agent2 gives the same or higher impurity levels than Agent1:

$$H_0: \mu_1 \geq \mu_2 \qquad \text{against} \qquad H_1: \mu_1 < \mu_2$$

The sample mean impurity for Agent1 was 8.25 and for Agent2 was 8.68, so the data are NOT consistent with H1.

We can stop there, however, completing the exercise:

The associated one-tailed p-value is p = 0.0038, so the observed t is significant at the 1% level (one-tailed).

The data therefore constitute strong evidence (on a one-tailed test) that the underlying mean impurity was lower for Agent1, by an estimated 8.68-8.25 = 0.43. The results continue to suggest that Agent1 should be preferred.

8.6 Independent samples t-test

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	0.21824624	
F Critical one-tail	1.539956607	
p2	0.43649248	

The sample variances for the samples are, respectively

$s_1^2 = 233.129$ and $s_2^2 = 190.176$

The observed F test statistic is F = 1.226 with 59 and 59 associated degrees of freedom, giving a two tailed p-value of p = 0.436.

The observed F ratio is thus *not significant*. The data are consistent with the assumption that the population variances underlying the incomes between the two samples, and we therefore proceed to use the equal variances form of the unrelated samples t test.

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	1.657869522	
P(T<=t) two-tail	0.00141947	
t Critical two-tail	1.980272249	
Difference in means	8.68	

The obtained independent samples t = 3.628 with 118 degrees of freedom.

The associated two-tailed p-value is p = 0.0014, so the observed t is significant at the 1% level (two-tailed).

The sample mean incomes for Males and Females were, respectively, 52.91 and 44.23.

The data therefore constitute strong evidence that the underlying mean income was greater for Males, by an estimated 52.91 – 44.23 = 8.68. The results strongly suggest that Males earn a higher income than Females.

Since it is generally thought that Males earn more than Females, the hypothesis could have explicitly stated that rather than just stating that there would be a difference. This would then have been a one-tailed t-test instead of two-tailed. In that case the associated one-tailed p-value is $p = 0.0007$, so the observed t is significant at the 1% level (one-tailed).

The result in the same, it’s just that the method would have been slightly different if the hypotheses was stated positively in the first instance.

The results are based upon the assumptions:

- The sample mean is an unbiased estimator of the population mean. This would be based upon the sample being a random sample, and big enough to be representative of the population.
- The sample distributions are normal. This could be tested with a simple histogram.