Hypothesis Testing Worksheet

**Exercise 7.1:**

Test type: Paired t-test (same samples under two agents)

Hypotheses (one-tailed):

* H₀: μ₁ ≤ μ₂
* H₁: μ₁ > μ₂

(where μ₁ = mean impurity for Filter Agent 1, μ₂ = mean impurity for Filter Agent 2)

We assume that:

* Differences are approximately normally distributed
* Pairs are correctly matched
* No extreme outliers in the differences

Excel methods:

* Data Analysis ToolPak:

|  |
| --- |
| Data → Data Analysis → t-Test: Paired Two Sample for Means   * Hypothesized Mean Difference = 0 * Read the p-value under P(T<=t) one-tail * Formula function:     =T.TEST(range\_agent1, range\_agent2, 1, 1)   * third argument 1 = one-tailed test * fourth argument 1 = paired test |

Decision rule:

* If the sample mean for Agent 1 is better (lower impurity), check the one-tailed p-value.
* One-tailed p-value = half the two-tailed p-value (when difference is in predicted direction).
* If p < 0.05, reject H₀.

Conclusion:

Since the earlier two-tailed test was significant, the one-tailed test will give an even smaller p-value. Therefore, we reject H₀ and conclude there is strong evidence that Filter Agent 1 is more effective than Filter Agent 2.

**Exercise 7.2**

Test type: Independent samples t-test (two groups: male vs female incomes)

Hypotheses (one-tailed):

* H₀: μ\_{male} ≤ μ\_{female}
* H₁: μ\_{male} > μ\_{female}

(μ is population mean)

Descriptive statistics:

* Males: x̄ = 52.91, s = 15.27, n = 60
* Females: x̄ = 44.23, s = 13.79, n = 60
* Difference in means: 52.91 - 44.23 = 8.68

Step 1 - Equality of variances (F-test):

* F = 1.226, p(two-tailed) = 0.436
* Conclusion: not significant → assume equal variances are plausible.

Step 2 - Independent samples t-test:

* Test statistic: t = 3.268
* Degrees of freedom ≈ 113 (Welch’s method)
* p(two-tailed) = 0.00142
* p(one-tailed) = 0.00071

Decision rule:

At α = 0.05, since p\_{one-tailed} = 0.00071 < 0.05, reject H₀.

Conclusion:

There is very strong evidence that the population mean income for males exceeds that of females. The difference in sample means is about 8.68 units. We reject H₀. The data provide strong evidence that male mean income is higher than female mean income.

We assume that:

1. Random, independent samples of males and females
2. Approximately normal distribution of incomes in each group (can be checked with histograms/Q-Q plots)
3. Homogeneity of variances (F-test supports this)

Validation:

* Inspect histograms/boxplots for normality and outliers
* Confirm F-test result (equal variances assumption reasonable)