## The INDEPENDENT Samples T Test (Exercise 7.2)

1. **Understanding the test performed**

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances |  |  |
|  | Male | Female |
|  | *Income* | *33.1* |
| Mean | 52.91333333 | 44.42203 |
| Variance | 233.1289718 | 191.2814 |
| Observations | 60 | 59 |
| df | 59 | 58 |
| F | 1.218774896 |  |
| P(F<=f) one-tail | 0.226087561 |  |
| F Critical one-tail | 1.5435835 |  |

From the output:

* An **F-test for variances** was done first:
  + ,
  + ,
  + .

Since , we **do not reject**  that the variances are equal.  
But the subsequent **t-test** used **“Unequal Variances”** (Welch’s t-test) anyway.

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | | |
|  |  |  |
|  | *Male* | *Female* |
| Mean | 52.91333333 | 44.23333 |
| Variance | 233.1289718 | 190.1758 |
| Observations | 60 | 60 |
| Hypothesised Mean Difference | 0 |  |
| df | 117 |  |
| t Stat | 3.267900001 |  |
| P(T<=t) one-tail | 0.000711286 |  |
| t Critical one-tail | 1.657981659 |  |
| P(T<=t) two-tail | 0.001422572 |  |
| t Critical two-tail | 1.980447599 |  |

**2. Hypotheses and conclusion**

The problem says:

test of whether the population mean income for males exceeds that of females

That’s a **one-tailed test**:

From the output:  
**One-tailed p-value** =

Since  (and even ), we **reject**.

**Conclusion:** There is strong evidence that the population mean income for males is greater than that for females.

**3. Assumptions for the t-test (unequal variances)**

1. **Independence**: The male and female samples are independent of each other (not paired).
2. **Normality**: The distribution of income in each group should be approximately normal — especially important because sample sizes are moderate (), but the t-test is fairly robust to non-normality with such sample sizes.
3. **Random sampling**: The data should come from a random sample from the population.

**4. How to validate assumptions**

* **Independence**: Known from study design (different individuals, no pairing).
* **Normality**: Check using:
  + Histograms / Q-Q plots of incomes for each group.
  + Shapiro–Wilk or Kolmogorov–Smirnov tests (though with , CLT helps).
* **Equal variance not assumed**: The Welch test doesn’t assume equal variances, so the F-test’s non-significance is just extra info; we already used the safer Welch df.

**5. Interpretation of practical significance**

The mean difference is about  (in thousands? the units are not given, but likely same as data).  
We can also compute Cohen’s d for effect size:

But since we used unequal variances, maybe use:

Roughly:

That’s a **medium to large** effect size.

**Final summary:**  
The analysis provides strong statistical evidence () that males have a higher mean income than females in the population. The Welch’s t-test was appropriate here given possible unequal variances, though the F-test suggested no significant difference in variances. Assumptions of normality and independence should be checked for rigor. The effect size is substantial.

## Appendix A

Data Analysis:

****