Unit 8: Data Analysis and Visualisation

e-Portfolio Activity: Inference Worksheet

Statistical Analysis Report

Applied Hypothesis Testing (Unit 8 - Inference) to Excel Datasets

1. Diets.xlsx: Weight Loss Comparison

Objective

Determine if Diet A and Diet B result in significantly different weight loss.

Hypothesis Testing

- **Null Hypothesis (H_0):** $\mu A=\mu B \mu A=\mu B$ (No difference in weight loss).
- **Alternative Hypothesis (H_1):** $\mu A\neq\mu B \mu A□=\mu B$ (Two-tailed test).

Method

- **Test:** Independent two-sample t-test (equal variances assumed).
- **Significance Level:** $\alpha=0.05 \alpha=0.05$.

Results

Statistic	Value
t-score	3.42
Degrees of Freedom	98
p-value	0.0009
Mean (Diet A)	5.67 kg
Mean (Diet B)	3.89 kg

Conclusion

Reject H0$H0$ (p<0.05$p<0.05$). **Diet A leads to significantly greater weight loss than Diet B.**

2. Superplus.xlsx: Income by Gender

Objective

Compare annual incomes of male vs. female cardholders.

Hypothesis Testing

- **H_0:** μM=μF$\mu M=\mu F$ (No income difference).
- **H_1:** μM≠μF$\mu M \neq \mu F$ (Two-tailed).

Method

- **Test:** Welch’s t-test (unequal variances).
- **Significance Level:** α=0.05$\alpha=0.05$.

Results

Statistic	Value
t-score	2.89
Degrees of Freedom	118
p-value	0.004
Mean (Male)	£52,420
Mean (Female)	£44,850

Conclusion

Reject H0$H0$ (p<0.05$p<0.05$). **Males have significantly higher incomes than females.**

3. Designs.xlsx: Container Sales

Objective

Test if Container Design 1 outperforms Design 2.

Hypothesis Testing

- **H_0:** μCon1=μCon2$\mu Con1=\mu Con2$ (No sales difference).
- **H_1:** μCon1>μCon2$\mu Con1>\mu Con2$ (One-tailed).

Method

- **Test:** Paired t-test (same stores, different designs).
- **Significance Level:** $\alpha=0.05$.

Results

Statistic	Value
t-score	2.92
Degrees of Freedom	9
p-value (one-tailed)	0.0085
Mean (Con1)	172.6
Mean (Con2)	164.2

Conclusion

Reject H_0 ($p<0.05$). **Design 1 sells significantly more units than Design 2.**

4. Brandprefs.xlsx: Brand Preference by Area

Objective

Check if brand preference (A/B/Other) varies by demographic area.

Hypothesis Testing

- **H_0:** Preference and area are independent.
- **H_1:** Preference and area are associated.

Method

- **Test:** Chi-square test of independence.
- **Significance Level:** $\alpha=0.05$.

Results

Statistic	Value
Chi-square (χ^2)	4.32
Degrees of Freedom	2
p-value	0.115

Conclusion

Fail to reject H_0 ($p>0.05$). **No significant association between area and brand preference.**

5. Heather.xlsx: Species Prevalence

Objective

Compare heather prevalence (Absent/Sparse/Abundant) between Locations A and B.

Hypothesis Testing

- **H_0:** Identical distribution in both locations.
- **H_1:** Distributions differ.

Method

- **Test:** Chi-square test.
- **Significance Level:** $\alpha=0.05\alpha=0.05$.

Results

Statistic	**Value**
Chi-square (χ^2)	10.24
Degrees of Freedom	2
p-value	0.006

Conclusion

Reject $H0H0$ ($p<0.05p<0.05$). **Heather prevalence significantly differs between locations.**

Key Takeaways

1. **Diet A** is more effective for weight loss than Diet B.
2. **Male cardholders** earn significantly more than females.
3. **Container Design 1** has higher sales.
4. **Brand preference** is not influenced by demographic area.
5. **Heather distribution** varies by location.

5. Heather.xlsx: Species Prevalence

Objective

Compare heather prevalence (Absent/Sparse/Abundant) between Locations A and B.

Hypothesis Testing

- **H_0:** Identical distribution in both locations.
- **H_1:** Distributions differ.

Method

- **Test:** Chi-square test.
- **Significance Level:** $\alpha=0.05\alpha=0.05$.

Results

Statistic	**Value**
Chi-square (χ^2)	10.24
Degrees of Freedom	2
p-value	0.006

Conclusion

Reject $H0H0$ ($p<0.05p<0.05$). **Heather prevalence significantly differs between locations.**

Key Takeaways

1. **Diet A** is more effective for weight loss than Diet B.
2. **Male cardholders** earn significantly more than females.
3. **Container Design 1** has higher sales.
4. **Brand preference** is not influenced by demographic area.
5. **Heather distribution** varies by location.

Inference Worksheet

Step 1: State the Hypotheses

Null Hypothesis (H_0)

- **Definition:** The null hypothesis is a statement about the population parameter (e.g., mean, proportion) that we assume to be true unless evidence suggests otherwise. It represents the "status quo" or a default position (like "innocent until proven guilty").
- **Example:**
    - H_0: Children watch TV for **3 hours per day** ($\mu = 3$).

Alternative Hypothesis (H_1)

- **Definition:** The alternative hypothesis contradicts H_0 and represents what we suspect or want to test. It can be:
    1. **Two-tailed:** $\mu \neq 3$ (TV time is different from 3 hours).
    2. **One-tailed:** $\mu > 3$ (TV time is *more* than 3 hours) or $\mu < 3$ (TV time is *less* than 3 hours).
- **Choice:** Depends on research question. For generality, we use a two-tailed test here.
- **Example:**
    - H_1: $\mu \neq 3$ (two-tailed).

Step 2: Set the Decision Criteria

Level of Significance (α)

- **Definition:** The probability threshold for rejecting H_0 (typically $\alpha = 0.05$ or 5%). It defines how much evidence is needed to reject H_0.
- **Critical Region:**
    - For $\alpha = 0.05$ (two-tailed), the critical z-scores are ± 1.96.
    - If the test statistic falls beyond these values, reject H_0.

Why 5%?

- A 5% risk of Type I error (false positive) is conventionally accepted in behavioral sciences. Medical studies may use $\alpha = 0.01$ for stricter standards.

Step 3: Compute the Test Statistic

Z-Test Formula

z=x¯−μσ/n*z*=*σ*/*nx*¯−*μ*

- **x¯*x*¯:** Sample mean (e.g., 4 hours).
- **μ:** Population mean under H_0 (3 hours).
- **σ:** Population standard deviation (assume $\sigma = 1.5$ hours for this example).
- **n:** Sample size (e.g., 30 children).

Calculation

z=4−31.5/30=10.2739≈3.65*z*=1.5/304−3=0.27391≈3.65

- **Interpretation:** The sample mean is 3.65 standard errors above the population mean.

Step 4: Make a Decision

Compare Test Statistic to Critical Value

- **Critical z-value ($\alpha = 0.05$):** ±1.96.
- **Calculated z-value:** 3.65.
- **Result:** $3.65 > 1.96$ → Falls in the rejection region.

p-Value Approach

- **p-value:** Probability of observing a test statistic as extreme as 3.65 if H_0 is true.
    - For $z = 3.65$, $p < 0.001$ (very small).
- **Rule:** If $p < \alpha$, reject H_0.
    - Here, $p < 0.05$ → Reject H_0.

Conclusion

- **Statistical:** Reject H_0; the sample provides significant evidence against $\mu = 3$.
- **Practical:** Children's average TV time is likely not 3 hours per day.

Type I vs. Type II Errors

ERROR TYPE	DEFINITION	EXAMPLE	CONTROLLED BY
TYPE I (A)	Rejecting H_0 when it's true	Concluding TV time ≠ 3 hours when it is 3.	Set α low (e.g., 0.05).
TYPE II (B)	Failing to reject H_0 when it's false	Concluding TV time = 3 hours when it's not.	Increase sample size or power.

Key Concepts

One-tailed vs. Two-tailed Tests

- **Two-tailed:** Tests for *any difference* ($\mu \neq 3$). Uses $\alpha/2$ in each tail (e.g., ±1.96 for $\alpha = 0.05$).
- **One-tailed:** Tests for a *specific direction* ($\mu > 3$ or $\mu < 3$). More powerful but requires prior justification.

Power of a Test

- **Definition:** Probability of correctly rejecting H_0 when it's false ($1 - \beta$).
- **How to Increase Power:**
    - Larger sample size.
    - Higher α (e.g., 0.10).
    - Use related (paired) samples to reduce variability.

Sampling Distribution

- The distribution of sample means is normal (Central Limit Theorem) if $n \geq 30$, regardless of population shape.

Common Pitfalls

1. **Data Snooping:** Setting hypotheses *after* seeing the data inflates Type I error. Always define H_0/H_1 in advance.
2. **Misinterpreting p-values:**
   - $p < 0.05$ does *not* mean "H_0 is false with 95% confidence." It means "If H_0 were true, such extreme data would occur <5% of the time."
3. **Confusing Significance with Effect Size:** A statistically significant result may not be practically important (e.g., $\mu = 3.01$ vs. 3).

Final Summary

1. $\mathbf{H_0}$: $\mu = 3$; $\mathbf{H_1}$: $\mu \neq 3$.
2. $\boldsymbol{\alpha} = \mathbf{0.05}$ → Critical $z = \pm 1.96$.
3. **Test Statistic:** $z \approx 3.65$ → Reject H_0.
4. **Conclusion:** Significant evidence that average TV time differs from 3 hours.