Statistical Tests for Use Case 1: Equality, Diversity & Inclusion

Understanding how to properly interpret statistical tests is essential for conducting fair and accurate EDI audits. This guide covers three key tests used in People Analytics to examine diversity patterns and their correct interpretation.

CHAPTER 1

Chi-Square Test: Testing Association

What This Test Does

The chi-square test examines whether two categorical variables are statistically associated. In EDI contexts, it helps answer questions like: Are men and women distributed across job grades in the same way?

Critical point: This test reveals association, not discrimination or causality. It's a descriptive fairness audit tool.

Chi-square = 10.82

df = 3

p-value = 0.013

Understanding Chi-Square Output

Chi-square statistic (χ^2)

Measures the size of the difference between observed and expected frequencies. Larger values indicate stronger associations between variables.

Degrees of freedom (df)

Based on the number of categories in your variables. Used to determine the critical value for significance testing.

p-value

Probability that results occurred by chance. Values below 0.05 indicate statistical significance and evidence of association.

Expected frequencies

What the distribution would look like if no association existed. Compare with observed values to identify inequality patterns.

Interpreting Chi-Square Results Correctly

The Right Interpretation

"The chi-square test indicates a statistically significant association between gender and job grade ($\chi^2 = 10.82$, $p = 0.013$). This suggests that men and women are not evenly distributed across job grades, beyond what would be expected by chance."

What This Does NOT Say

- That discrimination has occurred
- That gender causes job grade differences
- That the pattern is intentional or systematic

Common Student Mistakes

- 🚫 Interpreting chi-square as proof of bias
- 🚫 Ignoring which grades drive the result
- 🚫 Forgetting this is a descriptive fairness audit, not a causal model

CHAPTER 2

Independent Samples T-Test: Comparing Groups

Purpose of the Test

The independent samples t-test compares mean values between two independent groups. In EDI analysis, it answers questions like: Do two functions differ in average diversity representation?

This test examines whether observed differences between groups are statistically meaningful or likely due to chance variation.

T-Test Parameters and Interpretation

01

Mean Difference

Shows the size and direction of difference between groups. Indicates which function has higher diversity representation.

02

t-statistic

Measures how many standard errors the means are apart. Example: t = 2.45 indicates a moderate difference.

03

Degrees of Freedom

Based on sample sizes. Example: df = 98 for two groups totaling 100 observations.

04

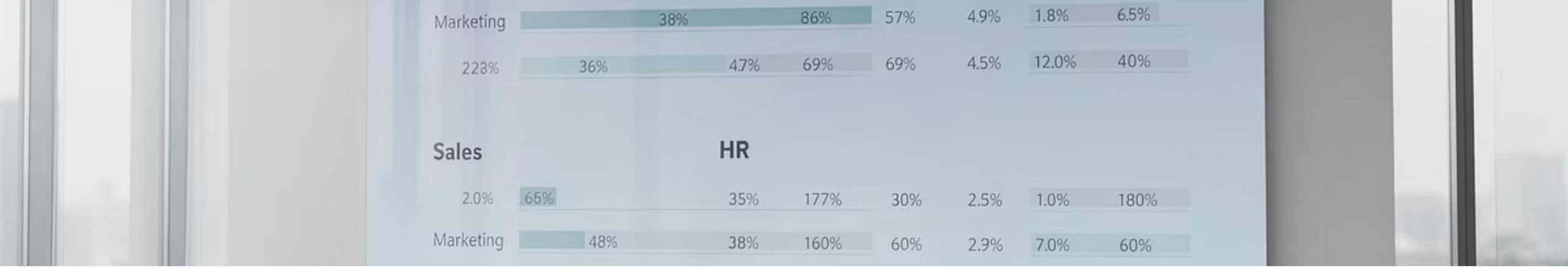
p-value

Determines statistical significance. p = 0.016 means less than 2% chance results occurred randomly.

05

Levene's Test

Checks whether equal variances can be assumed between groups, affecting which t-test formula to use.



Correct T-Test Interpretation

> "The results show a statistically significant difference in ethnic diversity between the two functions (t = 2.45, p = 0.016), with Function A displaying higher average diversity than Function B."

✓ What This Tells Us

- Group-level averages differ significantly
- The difference is unlikely due to chance
- One function shows higher diversity metrics

✗ Common Mistakes

- 🚫 Talking about individuals instead of group averages
- 🚫 Ignoring variance assumptions
- 🚫 Overinterpreting small but significant differences

CHAPTER 3

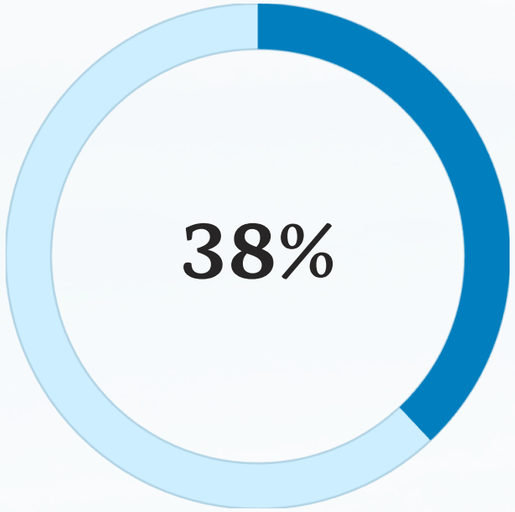
Multiple Linear Regression: Predicting Diversity

Multiple regression explains variation in a continuous outcome using multiple predictors. In EDI contexts, it helps identify what factors are associated with higher or lower ethnic diversity at team level.

This powerful technique allows us to examine several factors simultaneously while controlling for their interrelationships, providing a more nuanced understanding of diversity patterns.

Regression Model Output Example

Predictor	Coefficient	p-value
Team size	0.12	0.03
Team lead gender	0.25	0.01
Location	-0.18	0.04



R^2 Value

Variance explained by the model

Significance Level

Threshold for statistical significance

Understanding Regression Parameters

Coefficient (β)

Direction and size of association. Positive values indicate increases; negative values indicate decreases in the outcome variable.

p-value

Whether predictor contributes beyond chance. Values below 0.05 indicate the predictor is statistically significant.

R^2 / Adjusted R^2

How much variation the model explains. Higher values indicate better model fit and predictive power.

Interpreting Regression Results

Correct Interpretation

"The regression model explains 38% of the variance in ethnic diversity across teams. Team size and team leader gender are statistically significant predictors, suggesting that structural team characteristics are associated with diversity outcomes."

Critical Limitations

Remember: coefficients show association, not causation. The model identifies patterns but cannot prove that changing one variable will directly cause changes in diversity outcomes.

Common Mistakes to Avoid

- Interpreting coefficients as causal effects and ignoring multicollinearity when interpreting individual predictors
- Ignoring model fit (R^2) when evaluating results
- Overloading policy recommendations based solely on statistical associations

Key Takeaways for EDI Analysis

Test for Association

Chi-square reveals whether variables are related, but doesn't prove discrimination or causality. Use it as a starting point for deeper investigation.

Compare Groups Carefully

T-tests show differences between group averages. Always check variance assumptions and avoid overinterpreting small differences.

Model Complexity Wisely

Regression identifies patterns across multiple factors. Remember that association is not causation, and model fit matters for interpretation.

Proper statistical interpretation is essential for fair, evidence-based EDI decision-making. Always consider context, limitations, and the difference between statistical significance and practical importance.