Statistics and Data Analysis
Unit 04 – Lecture 02 Notes

Tofik Ali

February 14, 2026

Topic

Simple linear regression model, interpretation, residuals and R-squared.

Learning Outcomes

- Write the simple linear regression model
- Interpret slope and intercept in context
- Compute a prediction and a residual
- Explain R-squared (intuition)

Detailed Notes

These notes are designed to be read alongside the slides. They expand each slide bullet into plain-language explanations, small worked examples, and common pitfalls. When a formula appears, emphasize (1) what each symbol means, (2) the assumptions needed to use it, and (3) how to interpret the final number in the problem context.

Model

- y = b0 + b1 x + error
- Slope: expected change in y for 1-unit increase in x
- Intercept: predicted y at x=0 (interpret carefully)

Fit and Diagnostics

- Look at residual plots for patterns
- Outliers can dominate the fitted line
- High R^2 does not guarantee a good model

Exercises (with Solutions)

Exercise 1: Prediction

Model: yhat = 10 + 2x. Predict y when x=7.

Solution

- yhat = 24

Exercise 2: Residual

If actual y=20 at x=7, compute residual.

Solution

- 20-24 = -4

Exercise 3: Interpret slope

Slope is 5 thousand INR per extra room. Interpret.

Solution

- Each extra room increases predicted price by 5k INR (on average).

Exit Question

Why do we check residual plots even if R^2 is high?

Demo (Python)

Run from the lecture folder:

```
python demo/demo.py
```

Output files:

- `images/demo.png`
- `data/results.txt`

References

- Montgomery, D. C., & Runger, G. C. *Applied Statistics and Probability for Engineers*, Wiley.
- Devore, J. L. *Probability and Statistics for Engineering and the Sciences*, Cengage.
- McKinney, W. *Python for Data Analysis*, O'Reilly.