Logistic Regression Intuition

Classification Methods

Daniel E. Acuna
Associate Professor, University of Colorado Boulder

Contents of This Video

In this video, we will cover:

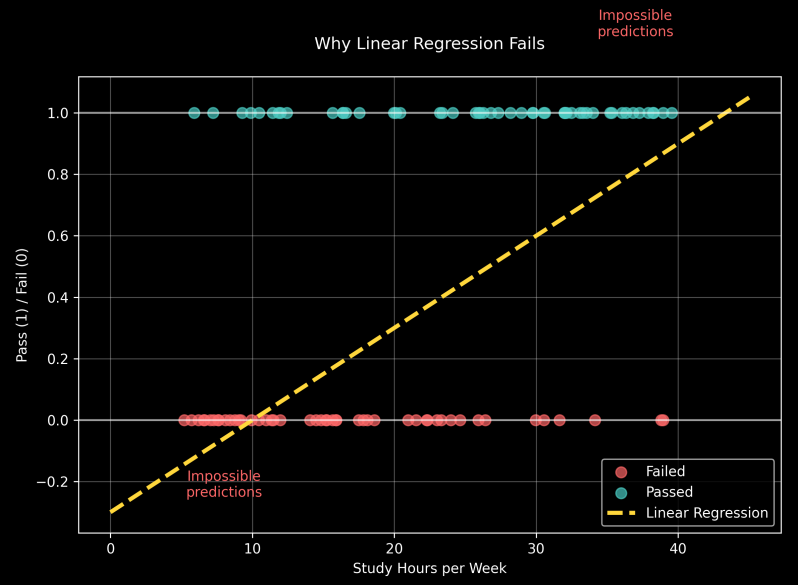
- Simple logistic regression with one feature
- The logistic function and S-shaped curve
- Why linear regression doesn’t work for classification
- Extending to multiple features
- Decision threshold and predictions
- Real-world applications in education

The Problem with Linear Regression

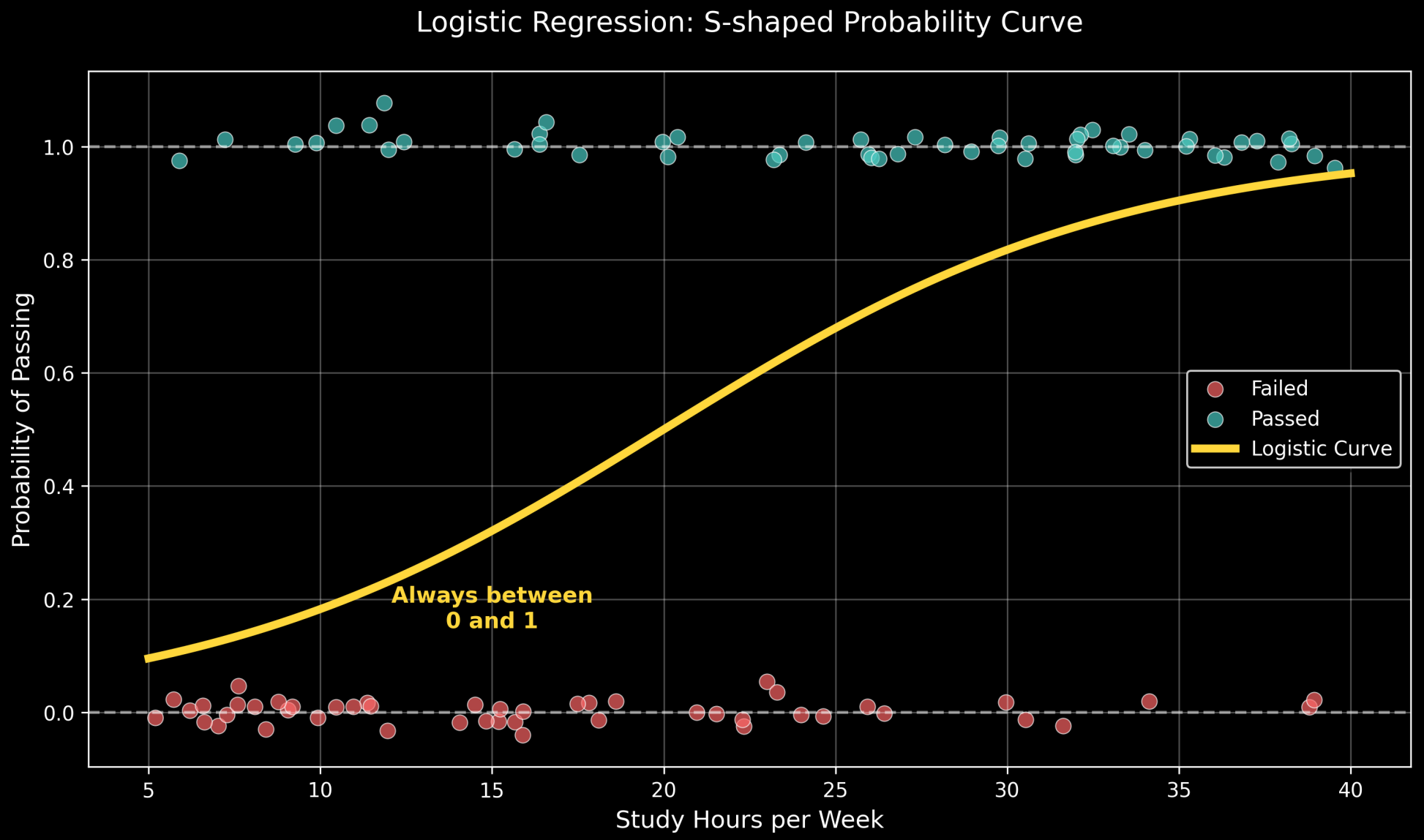
Student Example: Predict pass/fail from study hours

Linear Regression Issues:

- Can predict values < 0 or > 1
- Doesn’t represent probabilities well
- Straight line doesn’t capture the behavior
- Poor classification decisions



The Logistic Function Solution



The Logistic Function Formula

From Linear Score to Probability

$$z = b_0 + b_1 \cdot X_1$$

$$P(\text{Pass}) = \frac{1}{1 + e^{-z}}$$

Where:

- X_1 = study hours per week
- b_0 = intercept (bias term)
- b_1 = coefficient for study hours
- z = linear combination (log-odds)

Adding More Features

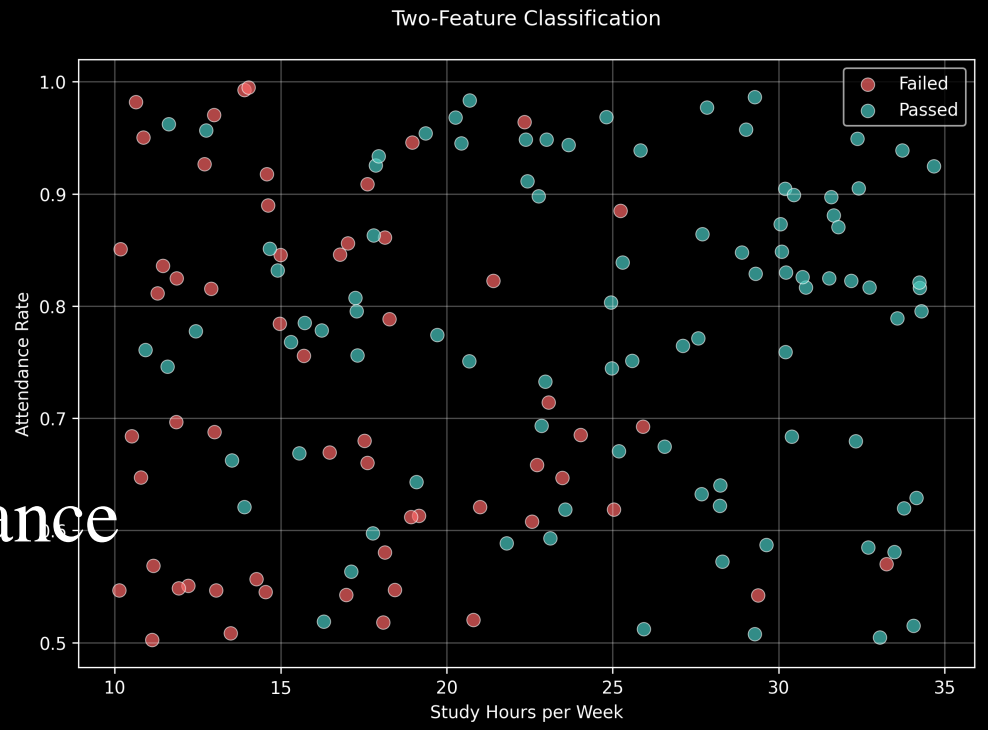
Two Features:

- Study hours per week
- Attendance rate

New Formula:

$$z = b_0 + b_1 \cdot \text{Hours} + b_2 \cdot \text{Attendance}$$

$$P(\text{Pass}) = \frac{1}{1 + e^{-z}}$$



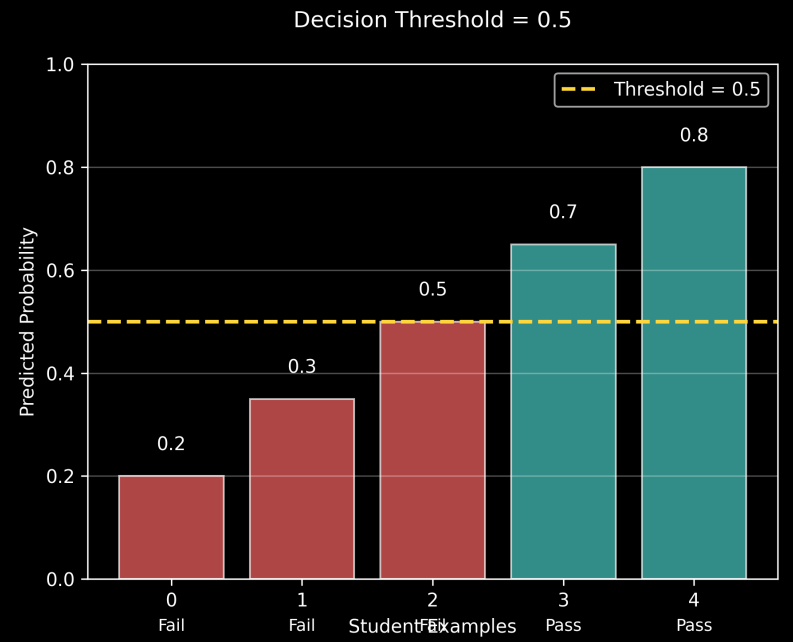
Making Predictions with Thresholds

Decision Rule:

- If $P(\text{Pass}) > 0.5 \rightarrow$ Predict “Pass”
- If $P(\text{Pass}) \leq 0.5 \rightarrow$ Predict “Fail”

Threshold can be adjusted:

- Lower threshold (0.3): More students predicted to pass
- Higher threshold (0.7): Fewer students predicted to pass
- Choice depends on cost of different errors



Expanding to Multiple Features

Real-world student data includes:

- Study hours per week
- Attendance rate
- Assignments submitted on time
- Previous GPA
- Average sleep hours

Formula: $z = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$

Real-World Applications

Academic Settings:

- Early warning systems for at-risk students
- Admission decision support
- Course recommendation systems
- Graduation probability prediction

Key Advantages:

- Fast computation
- Interpretable coefficients
- Probabilistic outputs
- Works well with many features

What We've Covered

In this video, we've explored:

- Why linear regression fails for classification
- The logistic function and S-shaped probability curve
- How logistic regression calculates probabilities
- Extending from one to multiple features
- Decision thresholds and predictions
- Real-world applications in academic settings