

# False Friends Between Statistics and Machine Learning

By Radoslav Harman (language editing and formatting by ChatGPT 5.4)

## Why this handout exists

Statistics and machine learning often use the **same words for different ideas**. This can create unnecessary confusion. When you read a paper, always ask:

1. Which community is speaking?
2. What mathematical object is being discussed?
3. Is the word naming a goal, a model, or a computational procedure?

## Common false friends

| Term       | In classical statistics   | In much of ML  | What to watch for  |
|------------|---|--|--|
| Inference  | Drawing conclusions about unknown quantities from data; estimation, confidence intervals, tests | Often: running a trained model on new inputs   | “Inference time” in ML usually does <b>not</b> mean statistical inference  |
| Hypothesis | A statement to be tested, such as $H_0 : \beta = 0$   | A candidate model or function from a hypothesis space  | A “hypothesis” in ML is often <b>not</b> a null hypothesis   |
| Regression | A family of models relating a response to predictors; the response need not be continuous       | Often means predicting a continuous target   | Logistic regression is called regression because it is a regression model, even though it is used for classification |
| Model      | A probabilistic description of data generation  | Often any predictive function or architecture  | In ML, “model” may mean a black-box predictor with little probabilistic interpretation                               |
| Parameter  | Usually a finite-dimensional unknown quantity in a statistical model                            | May mean trainable weights, often millions of them   | In deep learning, “number of parameters” usually means learned weights, not interpretable coefficients               |
| Bias       | Systematic deviation of an estimator from the truth   | Often bias in the bias–variance tradeoff, but also dataset bias, societal bias, inductive bias | Same word, several distinct meanings even within ML  |

| Term         | In classical statistics   | In much of ML  | What to watch for   |
|--------------|---|--|---|
| Consistency  | An estimator converges to the true quantity as sample size grows                                  | Sometimes used loosely for stable performance or agreement across datasets | In statistics, it has a precise asymptotic meaning  |
| Significance | Usually statistical significance  | Often informal importance or practical effect                              | “Significant improvement” in ML may mean practically noticeable, not hypothesis-test significance |
| Learning     | Traditionally less central term; one speaks more of parameter estimation, sometimes model fitting | Core term: the algorithm learns from data                                  | In ML, “learning” often includes optimization and generalization, not human-like understanding    |

## Three especially important distinctions

### 1. Inference

- **Statistics:** What can we conclude about unknown truths from data?
- **ML engineering:** What happens when a trained model is applied to a new input?

### 2. Hypothesis

- **Statistics:** A claim to test.
- **ML:** A candidate function or predictor.

### 3. Regression

- **Statistics:** A modeling framework.
- **ML:** Often a task label meaning continuous prediction.

## Rule of thumb

In data science, do not trust the word alone. Trust the **mathematics, the context, and the object being defined.**