CS 480/680 Tutorial 2 - Stats and Probability Review
Jan 23, 2026
Gavin Deane

Exercise 1: Conditional Probability and Bayes' Rule

We review conditional probability and Bayes' rule through a medical testing example.

A person may or may not have a certain disease, and a diagnostic test is used to detect it. The test is imperfect.

We define the following random variables:

- $x \in \{0, 1\}$ represents the disease status:
  - $x = 0$: the person does *not* have the disease
  - $x = 1$: the person *has* the disease
- $y \in \{0, 1\}$ represents the test outcome:
  - $y = 0$: the test result is *negative*
  - $y = 1$: the test result is *positive*

The test has the following known properties:

- $P(y = 0 \mid x = 0) = 0.90$
- $P(y = 1 \mid x = 1) = 0.99$

For each of the following cases, compute the posterior probability

$$P(x = 1 \mid y = 1),$$

that is, the probability that the person has the disease given that the test result is positive.

(a) $P(x = 1) = 10^{-3}$

(b) $P(x = 1) = 10^{-2}$

(c) $P(x = 1) = 10^{-1}$

Exercise 2: Likelihood and Log-Likelihood

In this exercise, we review the concepts of likelihood and log-likelihood for common probability distributions used in machine learning.

Exercise 2(a): Coin Flip Model

Let $x_1, x_2, \ldots, x_n$ be independent and identically distributed (i.i.d.) random variables such that

$$x_i \sim \text{Bernoulli}(p),$$

where $x_i \in \{0, 1\}$ and $p \in (0, 1)$ is an unknown parameter.

(i) Write down the likelihood function $P(x_{1:n} \mid p)$.

(ii) Derive the log-likelihood function $\log P(x_{1:n} \mid p)$.

(iii) Express the log-likelihood in terms of the total number of observed successes ($x_i = 1$ means success) $\sum_{i=1}^{n} x_i$.

Exercise 2(b): Gaussian Model

Let $x_1, x_2, \ldots, x_n$ be independent and identically distributed random variables such that

$$x_i \sim \mathcal{N}(\mu, \sigma^2),$$

where $\mu \in \mathbb{R}$ is the mean and $\sigma^2 > 0$ is the variance.

(i) Write down the likelihood function $P(x_{1:n} \mid \mu, \sigma^2)$.

(ii) Derive the log-likelihood function $\log P(x_{1:n} \mid \mu, \sigma^2)$.

(iii) Identify which terms of the log-likelihood depend on μ.

Exercise 3: Squared Error Minimization

Consider the following optimization problem:

$$\min_{a \in \mathbb{R}} \sum_{i=1}^{n} (x^i - a)^2,$$

where $x^1, x^2, \ldots, x^n$ are real-valued data points.

(i) Compute the value of a that minimizes the objective.

(ii) Interpret the solution in terms of the given data.

(iii) Briefly explain how this optimization problem relates to parameter estimation in probabilistic models.

Optional Exercise: Central Limit Theorem

Let $x_1, x_2, \ldots, x_n$ be independent and identically distributed random variables with finite mean and variance.

(i) State the Central Limit Theorem in your own words.

(ii) Explain why the Central Limit Theorem helps justify the use of Gaussian models in machine learning.