STAT 131 – Lecture Review Week 4

Review for intuition, setup, and problem reading

not a second lecture

Antonio Aguirre

Spring 2026

Based on the Lecture 3.1 and 3.2 notes.

How to use this review

What this hour is for

We will review the **main formulas**, but the focus is on **how to read the problem**, **what is given**, and **which tool fits the question**.

The four recurring *moves*

1. **Go forward:** use conditional probability or the product rule.
2. **Use independence:** turn an intersection into a product.
3. **Average over cases:** use a partition and the law of total probability.
4. **Reverse the conditioning:** use Bayes' theorem.

Main habit

Before calculating anything, ask:
What is the event? What is the conditioning? What is the right tool?

Quick recap: conditional probability and the product rule

Conditional probability

For events A and B with $\Pr(B) > 0$,

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

Intuition

Conditioning on B means we **shrink the sample space** to the outcomes where B occurred.

Product rule

Rearranging gives

$$\Pr(A \cap B) = \Pr(A \mid B)\Pr(B).$$

More generally,

$$\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1, A_2).$$

Reading cue

If a problem says "given", you are already in conditional-probability land.

Independence: what changes after you learn new information?

Definition

Two events A and B are **independent** if

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

This is equivalent to

$$\Pr(A \mid B) = \Pr(A) \qquad \text{and} \qquad \Pr(B \mid A) = \Pr(B).$$

Interpretation

Knowing one event occurred does **not change** the probability of the other.

Important contrast

Independent means *not informative*. **Disjoint** means *cannot happen together*.

Independent is not the same as disjoint

Disjoint events

If $A \cap B = \varnothing$, then

$$\Pr(A \cap B) = 0.$$

So ,

$$\Pr(A \mid B) = 0 \neq \Pr(A).$$

Are they **independent**?

Independent events

If A and B are independent, then

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

The events **can occur together**?.

A good midterm habit: do not swap these two ideas.

Example: two independent machines

Suppose

$$A = \{\text{Machine 1 breaks during a shift}\}, \qquad \Pr(A) = \frac{1}{3},$$

$$B = \{\text{Machine 2 breaks during the same shift}\}, \qquad \Pr(B) = \frac{1}{4},$$

and the two machines operate independently.

Question

What is the probability that **at least one** machine breaks during the shift?

Setup and calculation

$$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \frac{1}{3} + \frac{1}{4} - \left(\frac{1}{3} \cdot \frac{1}{4}\right) = \frac{1}{2}. \end{aligned}$$

Independent repeated trials: exactly k vs. at least one

Suppose each item is defective with probability p and non-defective with probability $q = 1 - p$. Assume 6 items are inspected independently.

Exactly 2 defective

Any one arrangement with 2 defective and 4 non-defective has probability

$$p^2 q^4.$$

There are

$$\binom{6}{2}$$

ways to place the 2 defective items, so

$$\Pr(\text{exactly 2 defective}) = \binom{6}{2} p^2 q^4.$$

At least 1 defective

The complement is easier:
Pr(at least 1 defective)
=1-Pr(all 6 non-defective)
=1-q^6.

Reading tip

Exactly usually means **count arrangements**.
At least one often means **use the complement**.

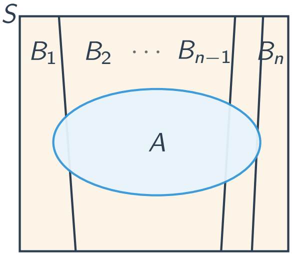
Partitions and the law of total probability

Partition

A collection $B_1, \dots, B_n$ is a **partition** of the sample space if

$$B_i \cap B_j = \varnothing \ (i \neq j), \qquad \bigcup_{i=1}^{n} B_i = S.$$

So the events are **disjoint** and **exhaustive**.



Law of total probability

If $B_1, \dots, B_n$ form a partition, then for any event A,

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A \mid B_i) \Pr(B_i).$$

Interpretation

Add the different paths through which A can happen.

Example: overall defect rate from three machines

Three machines produce a batch of items. Let D be the event "sampled item is defective."

Machine	$\Pr(M_i)$	$\Pr(D \mid M_i)$	Contribution to $\Pr(D)$
M_1	0.20	0.01	$0.20 \times 0.01 = 0.002$
M_2	0.30	0.02	$0.30 \times 0.02 = 0.006$
M_3	0.50	0.03	$0.50 \times 0.03 = 0.015$
Total			$\Pr(D) = 0.023$

Takeaway

The overall defect rate is a **weighted average** of the machine-specific defect rates.

Bayes' theorem: reversing the conditioning

Short form

$$\Pr(B \mid A) = \frac{\Pr(A \mid B)\Pr(B)}{\Pr(A)}.$$

Full two-case version

$$\Pr(B \mid A) = \frac{\Pr(A \mid B)\Pr(B)}{\Pr(A \mid B)\Pr(B) + \Pr(A \mid B^c)\Pr(B^c)}.$$

Vocabulary

$$\underbrace{\Pr(B)}_{\text{prior}}, \qquad \underbrace{\Pr(A \mid B)}_{\text{likelihood}}, \qquad \underbrace{\Pr(B \mid A)}_{\text{posterior}}$$

Most important warning

In general,

$$\Pr(A \mid B) \neq \Pr(B \mid A).$$

Bayes is the tool for moving from one to the other.

Bayes example: a positive screening test

Suppose a screening test is used in a population where the prevalence is

$$\Pr(D) = \frac{630}{100000} = 0.0063.$$

The test characteristics are

$$\Pr(T^+ \mid D) = 0.90, \qquad \Pr(T^+ \mid D^c) = 0.05.$$

Question

If a person tests positive, what is

$$\Pr(D \mid T^+)?$$

Why this is Bayes

We know

$$\Pr(T^+ \mid D)$$

and want

$$\Pr(D \mid T^+).$$

So we need to **reverse the conditioning**.

Bayes example: posterior probability after a positive test

Compute the posterior

$$\Pr(D \mid T^+) = \frac{\Pr(T^+ \mid D)\Pr(D)}{\Pr(T^+ \mid D)\Pr(D) + \Pr(T^+ \mid D^c)\Pr(D^c)}$$

$$= \frac{0.9 \cdot 0.0063}{0.9 \cdot 0.0063 + 0.05 \cdot 0.9937} \approx 0.1024.$$

Interpretation

A positive test raises the probability a lot:

$$0.0063 \longrightarrow 0.1024.$$

But not to 90%

The test is fairly sensitive, but the disease is still rare in the population. The **base rate** matters.

Random variables: outcomes vs. numerical summaries

Definition

A random variable X is a **function** that assigns a real number to each outcome in the sample space:

$$X : S \to D, \qquad D \subset \mathbb{R}.$$

Example. Toss a coin 3 times. Gain \$1 for each head and lose \$1 for each tail. Let $X =$ total gain/loss.

Outcome(s)	X
HHH	3
HHT, HTH, THH	1
HTT, THT, TTH	-1
TTT	-3

Key point

The random variable is **not the outcome**. It is the **number attached to the outcome**.

Important observation

Different outcomes can produce the same value of X.

Discrete random variables and the p.m.f.

A random variable is **discrete** if it takes at most a countable number of values.

Probability mass function

For a discrete random variable X,

$$f_X(x) = \Pr(X = x), \qquad x \in D.$$

Its main properties are

$$0 \leq f_X(x) \leq 1, \qquad \sum_{x \in D} f_X(x) = 1, \qquad \Pr(X \in A) = \sum_{x \in A} f_X(x).$$

Example: number of heads in 10 fair tosses

If $X =$ total number of heads, then

$$D = \{0, 1, 2, \ldots, 10\}, \qquad \Pr(X = x) = \binom{10}{x} \left(\frac{1}{2}\right)^{10}.$$

Two basic examples to remember

Uniform on $\{1, 2, \ldots, k\}$

All values are equally likely:

$$f_X(x) = \begin{cases} \dfrac{1}{k}, & x = 1, 2, \ldots, k, \\ 0, & \text{otherwise.} \end{cases}$$

Bernoulli(p)

If $X = 1$ when an event happens and $X = 0$ otherwise, then

$$f_X(x) = \begin{cases} 1 - p, & x = 0, \\ p, & x = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Mental model

Uniform = **all values equally likely.** Bernoulli = **success/failure encoded as 1/0.**

Final checklist for problem solving

Before calculating

- If the problem says given, write a conditional probability.
- If events or trials are independent, use products for intersections.
- If an event happens through several disjoint cases, use total probability.
- If you know $\Pr(A \mid B)$ but need $\Pr(B \mid A)$, use Bayes.
- If the problem turns outcomes into a number, define the random variable and then its distribution.