Decision and Risk Analysis

Lecture 1: **Bayesian Thinking for Risk and Decision Analysis**

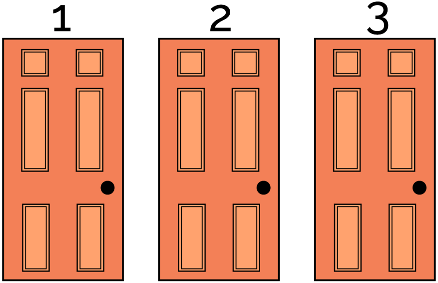
DAEN 427/ISEN 427/ISEN 627

Alexander Abuabara

Fall 2026

There are **three doors** for you to choose between:

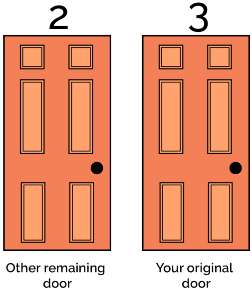
two contained **nothing** of value and **one** contained a **prize**.



You **pick one** of three closed doors,

making your guess as which door contained the prize.

Then, a door you **did not** choose is revealed,
and it **does not** contain the prize.



With two doors still closed (including the one you originally picked), you are given the choice:

- do you want to **swap doors**,
- or **stay** with the one you picked?

Your Decision: Should you stay or switch?

A decision problem begins with a **choice to be made** under certain **objectives** and **constraints**.

Risk and Uncertainty

- **Choice under risk**: Outcome probabilities are known (or can be estimated reliably).
- **Choice under uncertainty**: Outcome probabilities are unknown (or difficult to estimate).

Decision and Risk Analysis

1. **Modeling Decisions**: Define the available choices or actions.
2. **Modeling Risk**: Analyze situations where outcome probabilities are known.
3. **Modeling Uncertainty**: Analyze situations where outcome probabilities are unknown (or difficult to estimate).
4. **Modeling Preferences**: Represent how decision makers value outcomes and trade-offs.

Decision and Risk Analysis, Examples

1. **Modeling Decisions**
   A student chooses between studying for an exam tonight or attending a social event.
2. **Modeling Risk**
   A company launches a new product knowing there is a 70% chance of high demand and a 30% chance of low demand.
3. **Modeling Uncertainty**
   A farmer decides what crop to plant without knowing how much rain will fall this season.
4. **Modeling Preferences**
   A traveler chooses between a cheaper flight with a long layover and a more expensive direct flight.

Review

Exploratory data analysis

1. Exploratory data analysis, or EDA, is an initial data analysis that summarizes main characteristics. It often is done through visual means or in basic summary statistics.
2. Data visualization can be done through many ways, and helps us see patterns and identify potential issues or interesting patterns.

Categorical data

Nominal data

1. Named categories with no inherent order or numeric meaning.
2. When there are only two categories, the variable is often called binary or dichotomous
3. Examples: Material type (steel, aluminum, concrete), Model outcome (approved, rejected), Class label (spam, not spam)

Ordinal data

1. Categories with a natural order, but where the spacing between levels is not necessarily measurable or equal.
2. Examples: Risk level (low, medium, high), Damage severity (minor, moderate, severe), Rating (poor, fair, good, excellent)

Numeric data

Count or rank data

1. (hopefully self-explanatory)
2. Examples: number of defects, number of clicks, number of missing values.

Continuous data

1. Measurable quantities where difference between possible values can be arbitrarily small.
2. Data might lie within a range or be unbounded (in either direction)
3. Examples: temperature, pressure, time spent, latency, model error.

Some complications

Suppose you're designing a study and want to measure smoking exposure. We might treat it as:

1. Nominal: yes/no
2. Ordinal: current vs. former vs. never smoker
3. Count: number of cigarettes smoked in past week
4. Continuous: calculation of lifetime pack-years smoked

In the real world, decisions are made based on statistical considerations such as modeling interpretations, likelihood of measurement error, or simply convenience and real-world burden of collecting such data.

The population vs. a sample

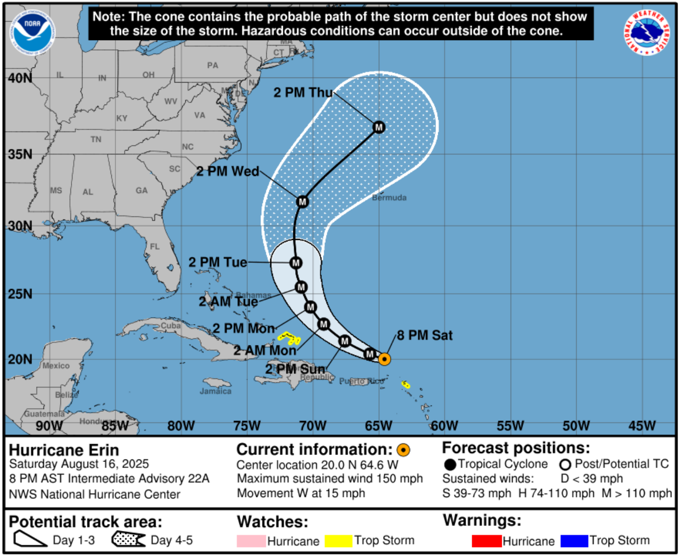
1. A population is the entire group that you want to draw conclusions about.
2. A sample is the specific group that you will collect data from. The size of the sample is always less than the total size of the population.
3. In research, a population doesn't always refer to people. It can mean a group containing elements of anything you want to study, such as objects, events, organizations, countries, etc.

Parameters

1. Attribute of the population of interest.
2. Not computable directly.
   (unless entire population is perfectly measured)
3. Usually written in Greek letters.

Statistics

1. Attribute of a sample.
2. Function of the observed values at hand.
3. Usually written in Roman letters.



Are point estimates of location enough?

Interpretations of probability

There is a 1 in 3 chance of selecting a white ball

The surgery has a 50% probability of success

Interpretations of probability

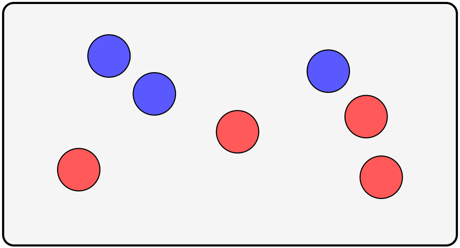
Long-run frequency

- A probability describes what would happen over many repetitions of the same random process.

Degree of belief

- A probability describes uncertainty about a single event, given available information.

A simple random experiment



If one ball is selected at random, what is the probability it is **red**?

Probability spaces are objects that model random experiments - real-world processes involving states that occur "randomly".

A probability space consists of three components:

1. A **sample space**, containing all possible outcomes.
2. Subsets of the sample space, called **events**, which comprise any number of possible outcomes (including none of them!).
3. A function that assigns **probabilities** to events.

An event is said to **occur** if the outcome of the random experiment is contained in that event.

Sample space

Depends on what we observe.

Random experiment	Sample space
Toss one coin	$\{H, T\}$
Toss two coins	$\{HH, HT, TH, TT\}$
Roll two dice and sum	$\{2, 3, 4, \dots, 12\}$
Time until component failure	$[0, \infty)$
Classify an email	{spam, not spam}

Events

Are subsets of the sample space.

They are the statements to which we assign probabilities.

Sample space	Example event
$\{H, T\}$	The coin lands heads
$\{HH, HT, TH, TT\}$	At least one head occurs
$\{2, 3, 4, \ldots, 12\}$	The sum is odd or equal to 8
$[0, \infty)$	Failure occurs after one year
Defect categories	The part is classified as defective

Probability

Is a number between 0 and 1 assigned to an event.

- $P(\text{heads in one fair coin toss}) = 0.5$
- $P(\text{at least one head in two fair coin tosses}) = 0.75$
- $P(\text{sum is 8 when rolling two dice}) = 5/36$
- $P(\text{machine fails before inspection})$ may be estimated from data or a model

Events as subsets

Remember, events are subsets of the entire sample space. Let's take for now the example of tossing a single fair coin and recording the outcome. For one fair coin toss, the sample space is $S = \{H, T\}$.

We can define the simple events A and B in the outcome space:

- A: getting a head, $A = \{H\}$.
- B: getting a tail, $B = \{T\}$.

Other possible events include: $\emptyset, \{H, T\}$. *What do the empty event and the full sample space mean in words?*

Set operations for events

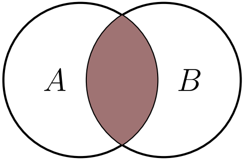
For two events A and B:

- **Intersection**: $A \cap B$ means A and B both occur.
- **Union**: $A \cup B$ means A or B occurs, including both.
- **Complement**: A^c means A does not occur.
- **Set difference**: $A \setminus B$ means A occurs but B does not.

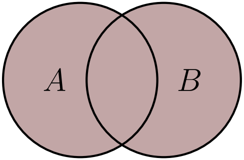
Events A and B are **disjoint** or **mutually exclusive** if

$$A \cap B = \emptyset.$$

Intersection: $A \cap B$



Union: $A \cup B$



Kolmogorov axioms

A probability rule $P(\cdot)$ must satisfy three basic axioms.

1. For every event A, $P(A) \geq 0$.
2. The probability of the entire sample space is $P(S) = 1$.
3. If A and B are disjoint, then

$$P(A \cup B) = P(A) + P(B).$$

These axioms are enough to derive many useful probability rules.

https://en.wikipedia.org/wiki/Probability_axioms

Two useful rules ...

From the axioms, we obtain two rules used constantly.

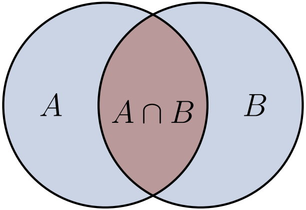
Complement rule

$$P(A^c) = 1 - P(A).$$

Inclusion-exclusion

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Why subtract the intersection?



When we add $P(A) + P(B)$, the overlap is counted twice.
Inclusion-exclusion subtracts one copy of the overlap.

De Morgan's laws

Complements, unions, and intersections can be combined:

Complement of a union

$$(A \cup B)^c = A^c \cap B^c$$

Complement of an intersection

$$(A \cap B)^c = A^c \cup B^c$$

The rules can be expressed in English as:

- The negation of "A or B" is the same as "not A and not B".
- The negation of "A and B" is the same as "not A or not B".

Example: engineering reliability

A manufacturing line classifies parts using two screening tests.

- A: a part fails the visual inspection.
- B: a part fails the sensor-based inspection.

Suppose

$$P(A) = 0.08, \qquad P(B) = 0.12, \qquad P(A \cap B) = 0.03.$$

Then the probability that a part fails at least one inspection is

$$P(A \cup B) = 0.08 + 0.12 - 0.03 = 0.17.$$

Example: data science classification

A model flags observations for review.

- A: the observation is truly anomalous.
- B: the model flags the observation.

Important event statements include:

- $A \cap B$: true positive.
- $A^c \cap B$: false positive.
- $A \cap B^c$: false negative.
- $A^c \cap B^c$: true negative.

Probability gives a formal language for **uncertainty**.

- A probability space consists of outcomes, events, and probabilities.
- Events are sets, so set operations matter.
- The complement rule and inclusion-exclusion are basic tools for probability calculations.

These ideas support inference, modeling, and risk analysis.

Quiz 1

When flipping a fair coin, we say that "the probability of flipping Heads is 0.5." How do you interpret this probability?

a. If I flip this coin over and over, roughly 50% will be Heads.

b. Heads and Tails are equally plausible.

c. Both (a) and (b) make sense.

Quiz 2

An election is coming up and a pollster claims that candidate A has a 0.9 probability of winning. How do you interpret this probability?

a. If we observe the election over and over, candidate A will win roughly 90% of the time.

b. Candidate A is much more likely to win than to lose.

c. The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

Quiz 3

Consider two claims:

1. Zuofu claims that he can predict the outcome of a coin flip. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10.
2. Kavya claims that she can distinguish natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies each.

In light of these experiments, what do you conclude?

a. You're more confident in Kavya's claim than Zuofu's claim.

b. The evidence supporting Zuofu's claim is just as strong as the evidence supporting Kavya's claim.

Quiz 4

Suppose that during a recent doctor's visit, you tested positive for a rare disease. If you only get to ask the doctor one question, which would it be?

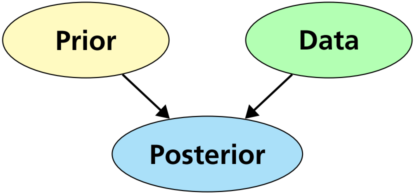
a. What's the chance that I actually have the disease?

b. If in fact I don't have the disease, what's the chance that I would've gotten this positive test result?

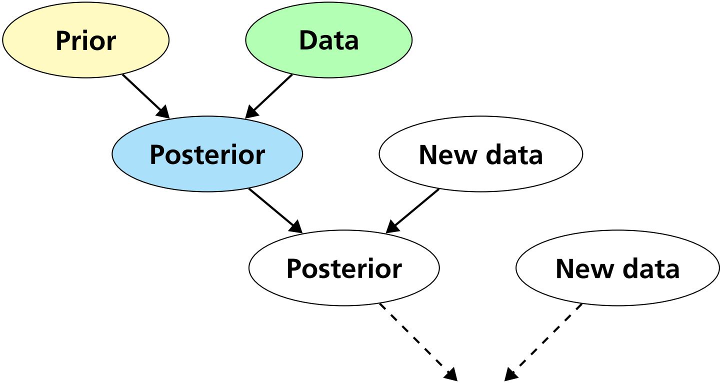
Bayesian Modeling of Decisions, Uncertainty, Risk, and Preferences

- Bayesian modeling provides a way to update beliefs as new information becomes available.
- It combines:
  1. **Prior beliefs**: what we know / assume before seeing new data.
  2. **Evidence/data**: new information relevant to the decision.
  3. **Posterior beliefs**: updated probabilities after incorporating the evidence.
- Bayesian modeling helps decision makers revise risk estimates and make better choices under uncertainty.
- Supporting learning from evidence and improving decisions over time.

A Bayesian Knowledge-Building Diagram



A Bayesian Knowledge-Building Diagram



Interpreting Probability

- The frequentist philosophy is so named for its interpretation of probability as the long-run relative frequency of a repeatable event.
- In the Bayesian philosophy, a probability measures the relative plausibility of an event.

Is this coin fair?

- **Frequentist:** "If the coin were fair, how surprising would these flips be?"
- **Bayesian:** "Given these flips, how much should I believe the coin is fair?"

One-Time Events

- Many important events are unrepeatable:

  Big Bang, elections, weather and climate change, ...

- **Frequentist** probability can be awkward for one-time events because there is no long-run repetition.
- **Bayesian** probability is more flexible:
  - It expresses degree of belief.
  - It incorporates model, evidence, and uncertainty.
- A 0.9 probability of winning means the candidate is considered 9 times more likely to win than lose.

The Bayesian Balancing Act

- Bayesian and frequentist approaches interpret the same evidence differently:
  - **Frequentist:** Treats "10 out of 10" identically in all cases.
  - **Bayesian:** Plausibility on observed evidence given prior beliefs.

  Both achieve 10 out of 10 correct in testing:
- Zuofu claims he can predict coin flips.

  ... is doubtful, predicting coin flips is highly implausible.
- Kavya claims she can distinguish natural / artificial sweeteners.

  ... is more believable.

Ask Questions!

- A **frequentist** analysis assesses the uncertainty of the observed data in light of an assumed hypothesis.
- A **Bayesian** analysis assesses the uncertainty of the hypothesis in light of the observed data.

Scenario: Population 100 people, 4 have the disease, 96 do not.

	Test +	**Test –**	**Total**
Disease	3	1	4
No disease	9	87	96
Total	12	88	100

Frequentist question: $P(\text{Positive Test} \mid \text{No Disease})$

- Among the 96 healthy people, 9 tested positive
  - False positive rate: $\frac{9}{96} \approx 0.10$
    If a person does *not* have the disease,
    there is about a 10% chance of testing positive.
- Frequentist analysis focuses on: $P(\text{Data} \mid \text{Null Hypothesis})$
- A p-value is *not* the probability that the hypothesis is true. (common misunderstanding)

Bayesian question: $P(\text{Disease} \mid \text{Positive Test})$

- Only 3 of the 12 positive tests are true positives.
  - Probability of actually having the disease: $\frac{3}{12} = 0.25$
    Even with a positive test, disease probability is only 25%.
- Sensitivity : $P(+ \mid \text{Disease}) = \frac{3}{4} = 75\%$
- Specificity : $P(- \mid \text{No Disease}) = \frac{87}{96} \approx 90.6\%$

The Meaning of Probability

Probability describes uncertainty.

Frequentist:

- Probability is a **long-run frequency**.
- Describes what happens over many repeated trials.

Bayesian:

- Probability is a **degree of belief**.
- Measures how plausible something is given current information.

Bayesian thinking updates beliefs using evidence:

$$\textbf{Posterior} \propto \textbf{Prior} \times \textbf{Likelihood}$$

Probability is not only about randomness;
it is a mathematical language for learning under uncertainty.

Homework

- Submit your answers before the next class.
- You may discuss with peers but submissions must be individual.
- Responses may be shared anonymously during class discussion!

Homework

Question 1

Think of a recent situation in which you changed your mind. Make a Bayesian knowledge-building diagram that includes your prior information, explain how new data helped you change your mind and your posterior conclusion.

Homework

Question 2

There are several data scientist openings at a heavily hyped company. Having read the job description, you know for a fact that you are qualified for the position (this is your data). Your goal is to ascertain whether you will actually be offered a position (this is your hypothesis).

a. From the perspective of someone using frequentist thinking, what question is answered in testing the hypothesis that you'll be offered the position?
b. Repeat part (a) within a Bayesian framework
c. Which question would you rather have the answer to: the frequentist or the Bayesian? Explain.

Homework

Question 3

a. Identify a topic that you know about (e.g., a sport, a school subject, music).
b. Identify a hypothesis about this subject.
c. How would your current expertise inform your conclusion about this hypothesis?
d. Which framework are you employing here, Bayesian or frequentist?

Homework

Question 4

A university health center is considering whether to recommend a costly new screening test for a rare disease among students. Historical data suggest that only 2% of students have the disease. The test correctly identifies diseased students 95% of the time. The test falsely identifies healthy students as diseased 8% of the time. A student tests positive.

a. What is the probability the student actually has the disease?
b. Should the health center automatically recommend treatment after one positive test?
c. What are the risks of a false positive? a false negative?
d. How would the decision change if the disease were more common?

Homework

Question 5

Suppose two analysts disagree about the probability of a cyberattack before seeing evidence:

- Analyst A believes the probability is 1%.
- Analyst B believes the probability is 20%.

Then, both observe the same security alert.

a. How can Bayesian reasoning explain why the analysts may still reach different conclusions?
b. What role do priors play in decision-making under uncertainty?
c. When is it reasonable for two people to maintain different posterior beliefs after observing the same data?
d. How do costs and risks influence the final decision, beyond just the posterior probability?

Homework

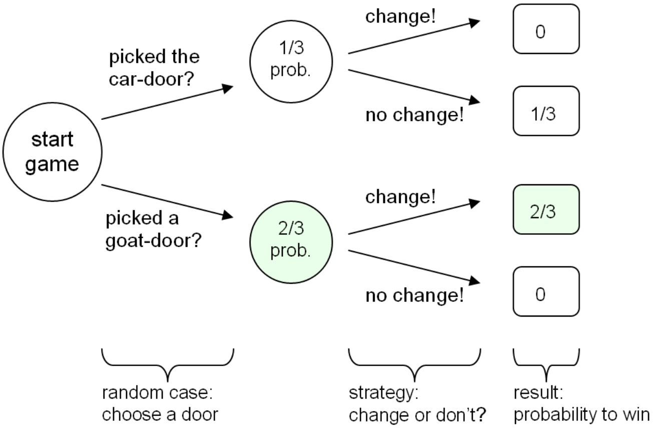
Question 6

Your friend just became interested in Bayesian statistics.
Explain the following to them:

a. Why is Bayesian statistics useful?
b. What are the similarities in Bayesian and frequentist statistics?

The Monty Hall Problem

Switching **Doubles** Your Chance of Winning!



Refs.: Wikipedia and University of Illinois