DSC 40A

Theoretical Foundations of Data Science I

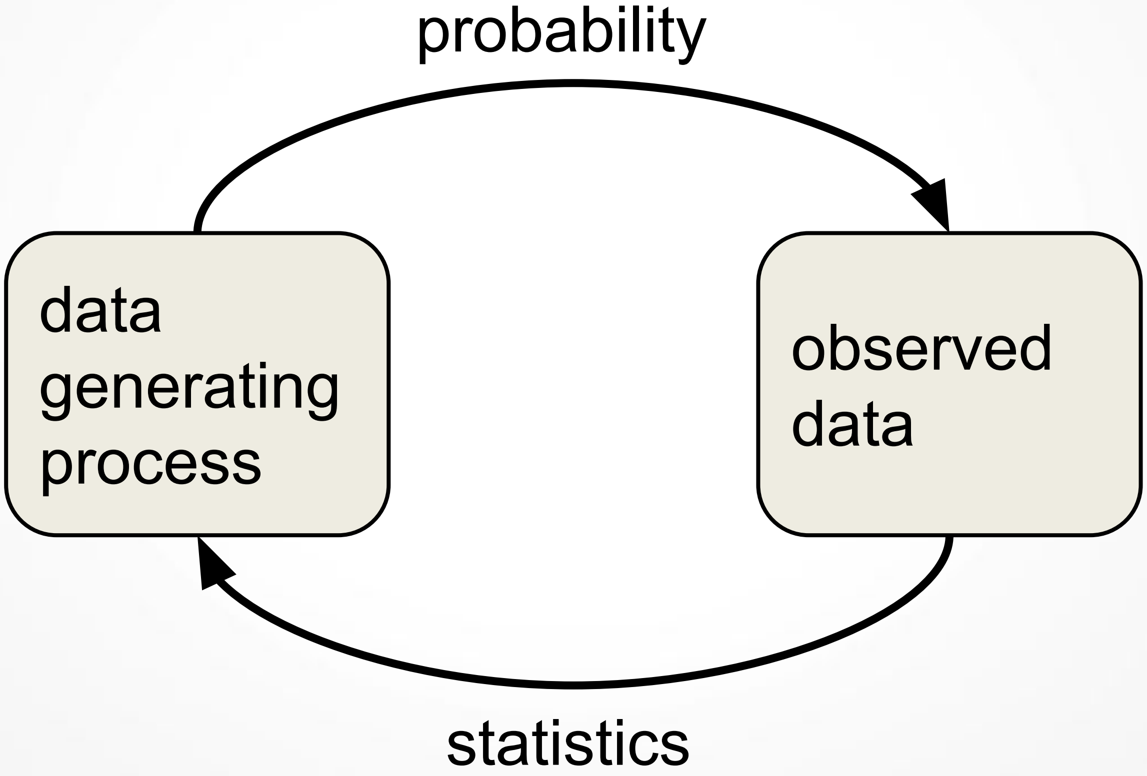
Predicting from Samples

- So far in this class, we have made predictions based on a data set, or sample.
- For our predictions to be any good, we need to know that other samples generated in the same way will be similar.

In This Video

- We'll study the basic definitions and rules of discrete probability.

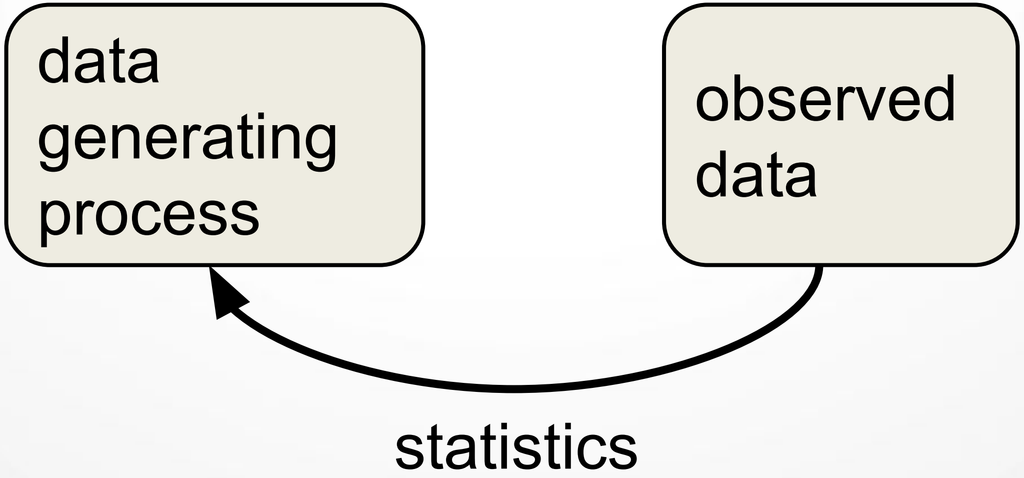
Probability and Statistics



Statistical Inference

Given observed data, we want to know how it was generated or where it came from. Maybe we want to

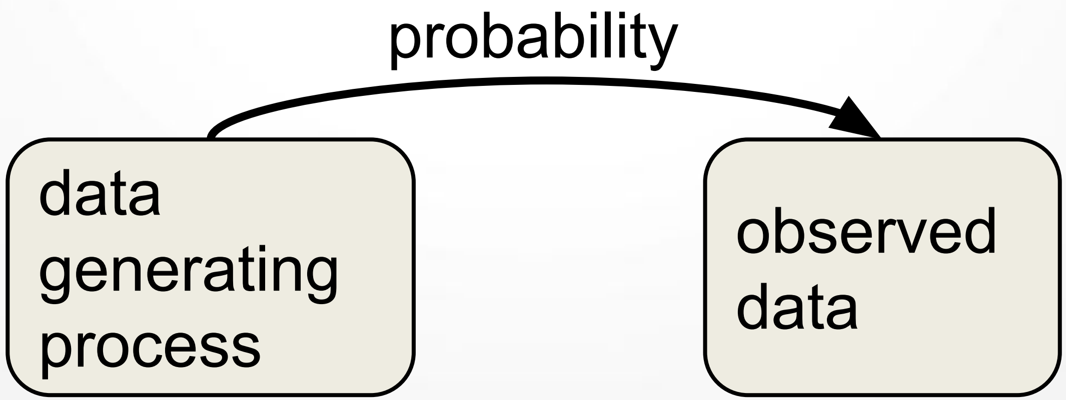
- predict other data generated from the same source
- know how different our sample could have been
- draw conclusions about whole population and not just observed sample - generalize



Probability

Given a certain model for data generation, what kind of data do you expect the model to produce? How similar is it to the data you have? Probability is the tool to answer these questions.

- expected value versus sample mean
- variance versus sample variance
- likelihood of producing exact observed data



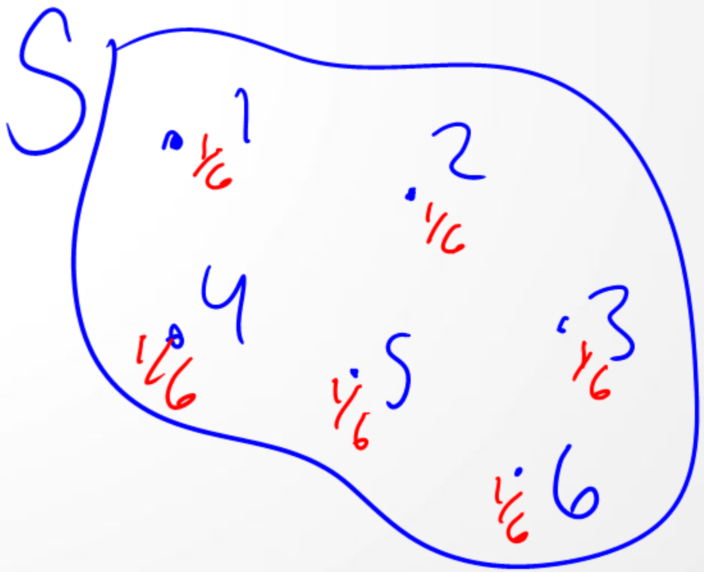
Probability

Sample space, S: (finite or countable) set of possible outcomes.

Probability distribution, p: assignment of probabilities to outcomes in S so that

$$0 \le p(s) \le 1$$

$$\sum_{s \in S} p(s) = 1$$



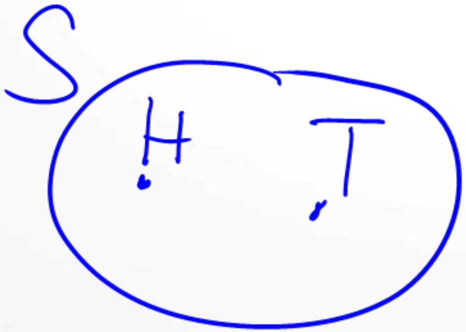
Probability

Sample space, S: (finite or countable) set of possible outcomes.

Probability distribution, p: assignment of probabilities to outcomes in S so that

- 0<= p(s) <=1   for each s in S

- Sum of probabilities is 1, $\sum_{s \in S} p(s) = 1$



Compare flipping a **fair coin** and **biased coin**:

A. Different sample spaces, different probability distributions.

B. Different sample spaces, same probability distributions.

C. Same sample spaces, different probability distributions.

D. Same sample spaces, same probability distributions.

Probability

Sample space, S: (finite or countable) set of possible outcomes.

Probability distribution, p: assignment of probabilities to outcomes s in S so that

- $0 <= p(s) <= 1$ for each s in S
- Sum of probabilities is 1, $\sum_{s \in S} p(s) = 1$

Event, E: an event E is a subset of the sample space

$$p(E) = \sum_{s \in E} p(s) \Rightarrow 0 \leq p(E) \leq 1$$

Uniform distribution

For sample space S with n elements, **uniform distribution** assigns the probability 1/n to each element of S.

- flipping fair coin 3 times in a row
- rolling a die

When flipping a fair coin successively three times:

A. The sample space is {H, T}
B. The event {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT} has probability less than 1.
C. The uniform distribution assigns probability 1/8 to each outcome.
D. None of the above.

Uniform distribution

For sample space S with n elements, **uniform distribution** assigns the probability 1/n to each element of S.

- flipping fair coin 3 times in a row
- rolling a die

For uniform distribution, the probability of an event E is:

$$p(E) = \sum_{s \in E} p(s) = \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{\#\text{ terms } = \text{ size of } E} = \boxed{\frac{\#\text{ outcomes in } E}{\#\text{ outcomes in } S}}$$

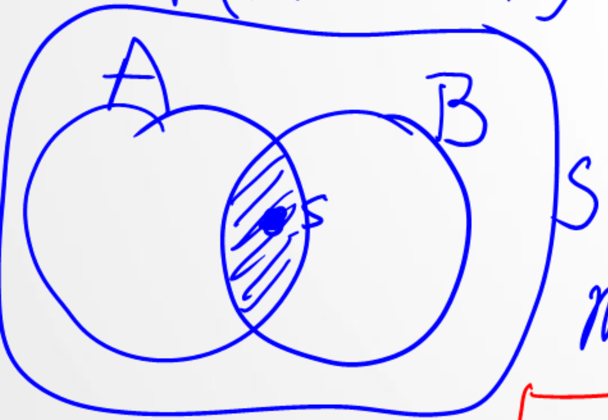
if uniform, $p(s) = \frac{1}{n}$

Addition Rule

If A and B are mutually exclusive events (cannot happen simultaneously), then

$$P(A \text{ or } B) = \sum_{s \in A \text{ or } B} p(s) = \sum_{s \in A} p(s) + \sum_{s \in B} p(s)$$

$$= P(A) + P(B)$$



more general rule:

$$\boxed{P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)}$$

Addition Rule

If A and B are mutually exclusive events (cannot happen simultaneously), then

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$$

union/OR

In general

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

intersection/AND

Multiplication Rule

$$P(A \text{ and } B) = \underbrace{P(A)}_{\text{prob of } A} * \underbrace{P(B|A)}_{\text{prob of } B \text{ given that } A \text{ happened}}$$

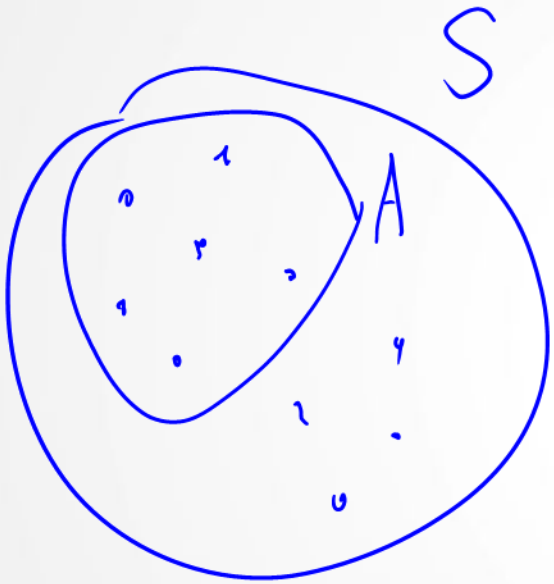
bar means "given that" or "assuming that"

Multiplication Rule

P(A and B) = P(A ∩ B)
= P(A)*P(B given that A has happened)
= P(A)*P(B|A)

$P(A \text{ and } B) = P(B \text{ and } A) = P(B) * P(A|B)$

Complement Rule



$$P(\bar{A}) = 1 - P(A)$$

bar means
"not A"
or complement of A

Complement Rule

$$P(\bar{A}) = 1 - P(A)$$

Practice Problems

Example 1. Rolling A Die. A fair 6-sided die has numbers from 1 to 6. Each time it is rolled, the outcome will be a number from 1 to 6. The probability of getting any of the six numbers is the same, which is 1/6. No roll affects the outcome of any other roll.

(i) Suppose the die is rolled once. What is the probability of rolling a 1 and a 2?

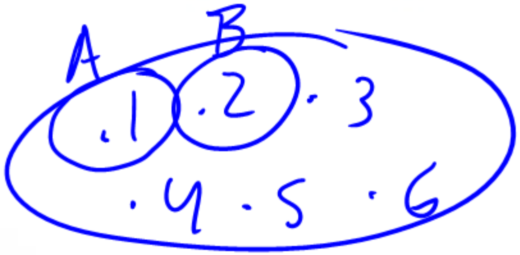
(ii) If the die is rolled once, what is the probability of rolling a 1 or a 2?

(iii) If the die is rolled twice, what is the probability of rolling a 1 on the first roll and a 2 on the second roll?

(i) 0

(ii) $2/6 = 1/3$

(iii) $\frac{1}{6} * \frac{1}{6} = \frac{1}{36}$



$P(A \text{ or } B) = P(A) + P(B)$

$P(A \text{ and } B) = \underbrace{P(A)}_{1/6} * \underbrace{P(B|A)}_{0}$

from Theory Meets Data by Ani Adhikari, Chapter 4

Practice Problems

P(A and B and C) = P(A) * P(B | A) * P(C | (A and B))

Example 2. A die is rolled 3 times. What is the probability that the face 1 never appears in any of the rolls?

P(not 1 on first roll AND not 1 on 2nd roll AND not 1 on 3rd roll) = $\underbrace{\text{P(not 1 on 1st roll)}}_{5/6}$ *

P(not 1 on 2nd roll assuming not 1 on 1st roll) * P(not 1 on 3rd roll assuming not 1 on 1st roll and not 1 on 2nd roll) $= \left(\frac{5}{6}\right)^3$

5/6 5/6

from Theory Meets Data by Ani Adhikari, Chapter 4

Practice Problems

Example 3. A die is rolled $\underline{n \text{ times}}$. What is the chance that only faces 2, 4 or 6 appear?

$$\frac{3}{6} = \frac{1}{2}$$

$$\underbrace{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2}}_{n} = \boxed{\left(\frac{1}{2}\right)^n}$$

Practice Problems

Example 4. A die is rolled two times. What is the probability that the two rolls had different faces?

S = set of pairs for first + second rolls

second roll

1 2 3 4 5 6

first roll

1
2
3
4
5
6

36 outcomes

$30/36 = \boxed{5/6}$

first roll gave one outcome

$S = \{1, 2, 3, 4, 5, 6\}$

outcomes for second roll

$\boxed{5/6}$

from Theory Meets Data by Ani Adhikari, Chapter 4

Summary

- We saw the basic definitions and rules in probability:
  - addition rule
  - multiplication rule
  - complement rule

- **Next time:** We'll learn about conditional probability, the probability of one event given that another has occurred.