Probability Theory

Rosen Section 7.2

Tom Michoel

MNF130V2020 – Week 13

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is $p(E) = \frac{|E|}{|S|}$.*

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is $p(E) = \frac{|E|}{|S|}$.*

- What if not all outcomes are equally likely?

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is $p(E) = \frac{|E|}{|S|}$.*

- What if not all outcomes are equally likely?

- Examples:

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is* $p(E) = \frac{|E|}{|S|}$.

- What if not all outcomes are equally likely?

- Examples:
  - Biased coin flip.

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is $p(E) = \frac{|E|}{|S|}$.*

- What if not all outcomes are equally likely?

- Examples:
  - Biased coin flip.
  - Rolling two dice when the order does not matter.

Introduction

- Recall:

  If S is a finite nonempty sample space of **equally likely outcomes**, *and E is an event (a subset of S), then the probability of E is $p(E) = \frac{|E|}{|S|}$.*

- What if not all outcomes are equally likely?
- Examples:
  - Biased coin flip.
  - Rolling two dice when the order does not matter.
  - ...

Assigning probabilities

Definition

Let S be the sample space of an experiment with a finite or countable number of outcomes. A **probability distribution** is a function p that assigns a probability $p(s)$ to each outcome $s \in S$, such that:

Assigning probabilities

Definition

Let S be the sample space of an experiment with a finite or countable number of outcomes. A **probability distribution** is a function p that assigns a probability $p(s)$ to each outcome $s \in S$, such that:

1. The probability of each outcome is nonnegative and no greater than one:

$$0 \leq p(s) \leq 1 \ \text{ for each } \ s \in S,$$

Assigning probabilities

Definition

Let S be the sample space of an experiment with a finite or countable number of outcomes. A **probability distribution** is a function p that assigns a probability $p(s)$ to each outcome $s \in S$, such that:

1. The probability of each outcome is nonnegative and no greater than one:

$$0 \leq p(s) \leq 1 \;\text{ for each }\; s \in S,$$

2. The probabilities of all outcomes sum to one, that is, when we do the experiment, one of the outcomes is certain to occur:

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

Example

Suppose that a die is biased (or loaded) so that 3 appears twice as often as each other number, but that the 5 other outcomes are equally likely. What is the probability distribution to describe random rolls of this die?

Example

Suppose that a die is biased (or loaded) so that 3 appears twice as often as each other number, but that the 5 other outcomes are equally likely. What is the probability distribution to describe random rolls of this die?

- The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.

Example

Suppose that a die is biased (or loaded) so that 3 appears twice as often as each other number, but that the 5 other outcomes are equally likely. What is the probability distribution to describe random rolls of this die?

- The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.
- The probability distribution must satisfy

$$p(s) = \begin{cases} q & \text{if } s \in \{1, 2, 4, 5, 6\} \\ 2q & \text{if } s = 3 \end{cases}$$

Example

Suppose that a die is biased (or loaded) so that 3 appears twice as often as each other number, but that the 5 other outcomes are equally likely. What is the probability distribution to describe random rolls of this die?

- The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.
- The probability distribution must satisfy

$$p(s) = \begin{cases} q & \text{if } s \in \{1, 2, 4, 5, 6\} \\ 2q & \text{if } s = 3 \end{cases}$$

- Plugging this into the condition that $\sum_{s \in S} p(s) = 1$ gives $1 = \sum_{s \in S} p(s) = 5q + 2q$, or $q = \frac{1}{7}$.

Example

Suppose that a die is biased (or loaded) so that 3 appears twice as often as each other number, but that the 5 other outcomes are equally likely. What is the probability distribution to describe random rolls of this die?

- The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.
- The probability distribution must satisfy

$$p(s) = \begin{cases} q & \text{if } s \in \{1, 2, 4, 5, 6\} \\ 2q & \text{if } s = 3 \end{cases}$$

- Plugging this into the condition that $\sum_{s \in S} p(s) = 1$ gives $1 = \sum_{s \in S} p(s) = 5q + 2q$, or $q = \frac{1}{7}$.
- Hence the probability distribution for our loaded die is:

$$p(s) = \begin{cases} \frac{1}{7} & \text{if } s \in \{1, 2, 4, 5, 6\} \\ \frac{2}{7} & \text{if } s = 3 \end{cases}$$

Definition

The **probability** of an event $E \subseteq S$ is the sum of the probabilities of the outcomes in E, that is,

$$p(E) = \sum_{s \in E} p(s)$$

Example

Suppose that S is a set with n elements. The **uniform distribution** is the probability distribution that assigns the probability $p(s) = \frac{1}{n}$ to each element $s \in S$.

Example

Suppose that S is a set with n elements. The **uniform distribution** is the probability distribution that assigns the probability $p(s) = \frac{1}{n}$ to each element $s \in S$.

- The uniform disrtibution is a probability distribution because

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

Example

Suppose that S is a set with n elements. The **uniform distribution** is the probability distribution that assigns the probability $p(s) = \frac{1}{n}$ to each element $s \in S$.

- The uniform disrtibution is a probability distribution because

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

- The uniform distribution assigns the same probability to an event as our previous definition of discrete probability (see Section 7.1):

$$p(E) = \sum_{s \in E} p(s) = \sum_{s \in E} \frac{1}{n} = \frac{|E|}{n} = \frac{|E|}{|S|}$$

Example

- A **Bernoulli trial** is an experiment with only two possible outcomes, generally called **success** or **failure** ($S = \{\text{success}, \text{failure}\}$, or $S = \{1, 0\}$). If p is the probability of success and q the probability of failure, it follows that $p + q = 1$.

Example

- A **Bernoulli trial** is an experiment with only two possible outcomes, generally called **success** or **failure** ($S = \{\text{success}, \text{failure}\}$, or $S = \{1, 0\}$). If p is the probability of success and q the probability of failure, it follows that $p + q = 1$.
- The **binomial distribution** is the probability ditribution on the set $S = \{0, 1, 2, \ldots, n\}$ that assigns to each element $k \in S$ the probability $b(k; n, p)$ of k successes in n Bernoulli trials with probability of success p.

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

$\square$

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.

$\square$

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.
- The probability $p(s) = p^k q^{n-k}$ where $k = \sum_{i=1}^{n} s_i$.

□

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.
- The probability $p(s) = p^k q^{n-k}$ where $k = \sum_{i=1}^{n} s_i$.
- Let $E = \{s \mid \sum_{i=1}^{n} s_i = k\}$ be the subset of all n-tuples with k successes.

□

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.
- The probability $p(s) = p^k q^{n-k}$ where $k = \sum_{i=1}^{n} s_i$.
- Let $E = \{s \mid \sum_{i=1}^{n} s_i = k\}$ be the subset of all n-tuples with k successes.
- All elements $s \in E$ have the same probability $p(s) = p^k q^{n-k}$.

$\square$

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.
- The probability $p(s) = p^k q^{n-k}$ where $k = \sum_{i=1}^{n} s_i$.
- Let $E = \{s \mid \sum_{i=1}^{n} s_i = k\}$ be the subset of all n-tuples with k successes.
- All elements $s \in E$ have the same probability $p(s) = p^k q^{n-k}$.
- The number of elements in E is the number of *unordered* ways to select k elements of a set with n elements, that is, $|E| = \binom{n}{k}$ (see Section 6.3).

$\square$

The binomial distribution

Theorem

The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Proof.

- The outcome of n Bernoulli trials is an n-tuple $s = (s_1, s_2, \ldots, s_n)$ where $s_i = 1$ (for success) or $s_i = 0$ (for failure), for $i = 1, 2, \ldots, n$.
- The probability $p(s) = p^k q^{n-k}$ where $k = \sum_{i=1}^{n} s_i$.
- Let $E = \{s \mid \sum_{i=1}^{n} s_i = k\}$ be the subset of all n-tuples with k successes.
- All elements $s \in E$ have the same probability $p(s) = p^k q^{n-k}$.
- The number of elements in E is the number of *unordered* ways to select k elements of a set with n elements, that is, $|E| = \binom{n}{k}$ (see Section 6.3).
- Hence $p(E) = \binom{n}{k} p^k q^{n-k}$.

$\square$

The binomial distribution

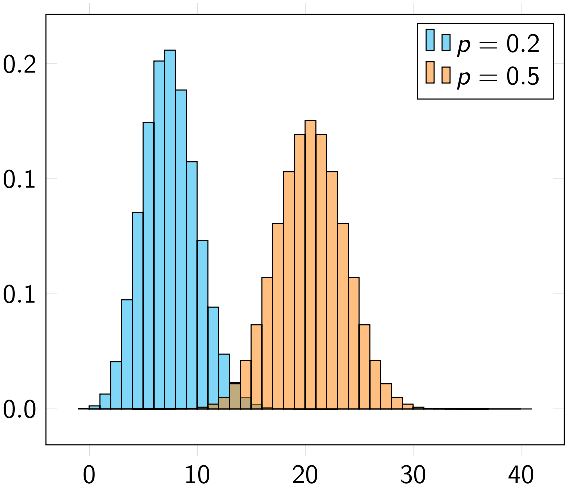

Figure 1: The binomial distributions for $n = 40$ and $p = 0.2$ and $p = 0.5$

Probabilities of complements of events

Theorem

Let E be an event in a sample space S. The probability of the event $\bar{E} = S - E$, the **complementary event** *of E, is*

$$p(\bar{E}) = 1 - p(E)$$

Proof.

$\square$

Probabilities of complements of events

Theorem

Let E be an event in a sample space S. The probability of the event $\bar{E} = S - E$, the **complementary event** *of E, is*

$$p(\bar{E}) = 1 - p(E)$$

Proof.

- Because $S = E \cup \bar{E}$ and $E \cap \bar{E} = \emptyset$,

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

$\square$

Probabilities of complements of events

Theorem

Let E be an event in a sample space S. The probability of the event $\bar{E} = S - E$, the **complementary event** *of E, is*

$$p(\bar{E}) = 1 - p(E)$$

Proof.

- Because $S = E \cup \bar{E}$ and $E \cap \bar{E} = \emptyset$,

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

- Hence

$$p(\bar{E}) = 1 - p(E)$$

□

Probabilities of unions of events

Theorem

Let E_1 and E_2 be events in a sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) \tag{1}$$

Proof.

$\square$

Probabilities of unions of events

Theorem

Let E_1 and E_2 be events in a sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) \tag{1}$$

Proof.

- We can write $E_1 \cup E_2 = E_1 \cup [E_2 - (E_1 \cap E_2)]$.

$\square$

Probabilities of unions of events

Theorem

Let E_1 and E_2 be events in a sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) \tag{1}$$

Proof.

- We can write $E_1 \cup E_2 = E_1 \cup [E_2 - (E_1 \cap E_2)]$.
- Since $E_1 \cap [E_2 - (E_1 \cap E_2)] = \emptyset$, it follows that

$$\begin{aligned} p(E_1 \cup E_2) &= \sum_{s \in E_1} p(s) + \sum_{s \in [E_2 - (E_1 \cap E_2)]} p(s) \\ &= \sum_{s \in E_1} p(s) + \sum_{s \in E_2} p(s) - \sum_{s \in E_1 \cap E_2} p(s) \\ &= p(E_1) + p(E_2) - p(E_1 \cap E_2) \end{aligned}$$

$\square$

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?
   - There are $6^3 = 216$ equally possible (ordered) outcomes when 3 dice are rolled.

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?
   - There are $6^3 = 216$ equally possible (ordered) outcomes when 3 dice are rolled.
   - Hence the probability to obtain $p(⚅⚅⚅) = \frac{1}{216}$.

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?
    - There are $6^3 = 216$ equally possible (ordered) outcomes when 3 dice are rolled.
    - Hence the probability to obtain $p(⚅⚅⚅) = \frac{1}{216}$.
2. What is the probability to obtain ⚅ ⚅ ⚅ when rolling the first two dice gave ⚅ ⚅?

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?
   - There are $6^3 = 216$ equally possible (ordered) outcomes when 3 dice are rolled.
   - Hence the probability to obtain $p(⚅⚅⚅) = \frac{1}{216}$.
2. What is the probability to obtain ⚅ ⚅ ⚅ when rolling the first two dice gave ⚅ ⚅?
   - There are 6 equally possible outcomes for the 3rd dice.

Conditional probability

Example

1. What is the probability to obtain ⚅ ⚅ ⚅ when 3 dice are rolled?
   - There are $6^3 = 216$ equally possible (ordered) outcomes when 3 dice are rolled.
   - Hence the probability to obtain $p(⚅⚅⚅) = \frac{1}{216}$.
2. What is the probability to obtain ⚅ ⚅ ⚅ when rolling the first two dice gave ⚅ ⚅?
   - There are 6 equally possible outcomes for the 3rd dice.
   - Hence the probability to obtain $p(⚅⚅⚅ \mid ⚅⚅) = \frac{1}{6}$.

Conditional probability and independence

Definition

Let E and F be events with $p(F) > 0$. The **conditional probability** of E given F, denoted by $p(E \mid F)$, is defined as

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Definition

The events E and F are **independent** if and only if $p(E \cap F) = p(E)p(F)$.

Remark

Conditional probability and independence

Definition

Let E and F be events with $p(F) > 0$. The **conditional probability** of E given F, denoted by $p(E \mid F)$, is defined as

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Definition

The events E and F are **independent** if and only if $p(E \cap F) = p(E)p(F)$.

Remark

- Conditional probability expresses the information we gain about E from observing F.

Conditional probability and independence

Definition

Let E and F be events with $p(F) > 0$. The **conditional probability** of E given F, denoted by $p(E \mid F)$, is defined as

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Definition

The events E and F are **independent** if and only if $p(E \cap F) = p(E)p(F)$.

Remark

- Conditional probability expresses the information we gain about E from observing F.
- If E and F are independent, observing F tells us nothing new about E, because

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} = \frac{p(E)p(F)}{p(F)} = p(E)$$

What to do next?

What to do next?

- Read Section 7.2, especially all the extra examples. Only the topics covered in these slides are part of the syllabus!

What to do next?

- Read Section 7.2, especially all the extra examples. Only the topics covered in these slides are part of the syllabus!
- Solve exercises. Some recommended exercises are in the assignment for week 14:
  `https://mitt.uib.no/courses/21678/assignments/26392`

What to do next?

- Read Section 7.2, especially all the extra examples. Only the topics covered in these slides are part of the syllabus!
- Solve exercises. Some recommended exercises are in the assignment for week 14:
  `https://mitt.uib.no/courses/21678/assignments/26392`
- Post questions on the discussion forum and participate in the discussion:
  `https://mitt.uib.no/courses/21678/discussion_topics/157826`