Probability Theory

Section 3.4 and 3.5

of

Grimaldi's book

Acknowledgement

- I have used materials from the following sources:
  - Textbook
  - Lecture notes slides prepared by Prof. Bulatov.

Sample space

- Experiment: tossing a coin, rolling a die, selecting subjects from a group at random
- The set of all possible outcomes of an experiment is called a sample space.

Experiments and Outcomes

- Experiment: Tossing a coin
  Outcomes: {heads, tails}

- Experiment: Rolling a dice
  Outcomes: {1,2,3,4,5,6}

- Experiment: Rolling two dice
  Outcomes: $\{1,\ldots,6\} \times \{1,\ldots,6\}$
  or $\{A \subseteq \{1,\ldots,6\} : |A| \leq 2\}$

- Experiment: Buying 3 lottery tickets (out of 100,000)
  Outcomes: 3-element subsets of $\{1,\ldots,100000\}$

Sample Space and Events

- The set of all outcomes of an experiment is called the sample space

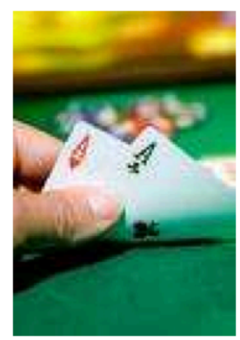
- Sometimes we are interested not in a single outcome, but an event that happens in several outcomes

Examples:

- Get heads at least 3 times when tossing 5 coins

- Win a prize in lottery

- Get 2 aces in a poker hand



Events

- Let S be the sample space of a certain experiment. An event is any subset of S

Examples:

- Experiment: Tossing 2 coins

  Sample space: $S = \{heads,tails\} \times \{heads,tails\}$

  Event: Get exactly 1 heads

  $A = \{(heads,tails),(tails,heads)\}$

- Experiment: Rolling 2 dice

  Sample space: $S = \{1,\ldots,6\} \times \{1,\ldots,6\}$

  Event: The sum of the dice is 6

  $A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$

Examples

- The probability of getting heads in the coin tossing experiment

  Sample space: S = {heads,tails} , Event: A = {heads},

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{1}{2}$$

- The probability to get even number in the dice rolling experiment

  Sample space: S = {1,2,3,4,5,6}, Event: A = {2,4,6}

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{3}{6} = \frac{1}{2}$$

- 100 tickets, numbered 1,2,3,…, 100, are sold to 100 different people for a drawing. Four different prizes are awarded, including a grand prize (a trip to Tahiti). Find the probability that ticket 47 wins a prize while ticket 73 does not.

Examples

- The probability of getting heads in the coin tossing experiment

  Sample space: S = {heads,tails} , Event: A = {heads},

  $$\Pr(A) = \frac{|A|}{|S|}$$ 1

- The probabili ment

  Sample spac

  $$\Pr(A) = \frac{|A|}{|S|}$$

- A: event of subset of size 4 where 47 is included and 73 is not included.
- |A|= 1 x 98 x 97 x 96.
- Sample space size = C(100,4).

- 100 tickets, numbered 1,2,3,…, 100, are sold to 100 different people for a drawing. Four different prizes are awarded, including a grand prize (a trip to Tahiti). Find the probability that ticket 47 wins a prize while ticket 73 does not.

Sample space

- Under the assumption of equal likelihood, let S be the sample space for an experiment E. Each subset of S, including empty set, is called an event. Each element of S determines an outcome, so if |S| =n and a ε S and A ⊆ S, then
  - Pr({a}) = Pr(a) = |{a}|/|S| = 1/n
  - Pr(A) = |A|/n

Axioms of Probability

- Let S be the sample space for an experiment E. Let A be any event (i.e. $A \subseteq S$).
  - $Pr(A) \geq 0$
  - $Pr(S) = 1$
  - Rule of Complement: $Pr(A)=1-Pr(A^c)$ where A^c is the complement of event A.

More General Probability: Geometric Probability

- How to measure the area of an island?

- Draw a rectangle around the island and drop many random points
- Then $$\frac{\text{area of the island}}{\text{area of the rectangle}} \approx \frac{\text{\# of points within the island}}{\text{total \# of points}}$$
- Sample space: Points in the rectangle

  Events: Measurable sets of points

  Probability: The area of an event

Examples of sample space

1. **Tossing a fair (unbiased) coin:** Sample space $S = \{H, T\}$, and $Pr(\{H\}) = Pr(\{T\}) = \frac{1}{2}$

2. **Tossing a fair coin three times:** Here $S = \{(t_1, t_2, t_3) | t_i \in \{H, T\}\}$. Here t_i is the outcome of the i^{th} toss. The probability of each outcome, such as (H, T, T), is $\frac{1}{8}$. If we toss the coin n times, the size of the sample space is 2^n, and each point having probability $\frac{1}{2^n}$.

3. **Rolling two distinguishable dice (one red, one blue):** The sample space is $S = \{(i, j) : 1 \leq i, j \leq 6\}$. There are 36 elements in the sample space. Each of the outcomes has equal probability, $\frac{1}{36}$.

Examples of sample space

4. **Rolling two indistinguishable dice (both red):** The sample space here is $S = \{\{i,j\} : 1 \leq i \leq j \leq 6\}$. An outcome of one die 3 and the other 5 is written as $\{3,5\}$ with the smaller one first. There are 21 elements in the sample space. Now the probability of each sample point is not the same. The probability of an outcome of the form $\{i,i\}$ is $\frac{1}{36}$. However the probability of an outcome $\{i,j\}$, $i \neq j$, is $\frac{2}{36}$ (why?).

5. **Card shuffling:** The deck of cards has 52 cards. Shuffle a deck of cards. The sample space consists of 52! permutations of the deck, each with equal probability $\frac{1}{52!}$.

6. **Poker hands:** The sample space consists of all possibe five-card hands. The sample space has $\binom{52}{5}$ elements, each with probability $\frac{1}{\binom{52}{5}}$.

Examples of sample space

7. **Balls and bins with distinguishable balls:** Bins (there are k such bins) are distinguishable and the balls (there are n such balls) are distinguishable. An outcome is $(b_1, b_2, \ldots b_n)$ where b_i is the bin number ball i lies. The sample space has k^n n-tuples, each with equal probability $\frac{1}{k^n}$.

8. **Balls and bins with indistinguishable balls:** Bins (there are k such bins) are distinguishable and the balls (there are n such balls) are indistinguishable. After throwing the balls, we see only the number of balls that landed in each bin. Each outcome is a k-tuple $(m_1, m_2, \ldots, m_k)$ where m_i denotes the number of balls in bin i. The number of sample points is therefore $\binom{n+k-1}{k-1}$. The probalities of sample points are not the same. Why?

Examples of sample space

7. **Balls and bins with distinguishable balls:** Bins (there are k such bins) are disting[illegible] nguishable. An outcom[illegible] The sample space has [illegible]

$$Pr(\{m_1, m_2, \ldots, m_k\}) = \frac{n!}{m_1! \times m_2! \times \ldots \times m_k!}$$

8. **Balls and bins with indistinguishable balls:** Bins (there are k such bins) are distinguishable and the balls (there are n such balls) are indistinguishable. After throwing the balls, we see only the number of balls that landed in each bin. Each outcome is a k-tuple $(m_1, m_2, \ldots, m_k)$ where m_i denotes the number of balls in bin i. The number of sample points is therefore $\binom{n+k-1}{k-1}$. The probalities of sample points are not the same. Why?

Probabilities do not depend on the order

- Among 100 lottery tickets there is 1 winning ticket. I buy 2 tickets. Find the probability I win. (Suppose ticket #1 wins.)
- Method 1. Experiment: I buy 2 tickets (unordered)

  Sample space: S = {2-element subsets of {1,…,100} }

  Event: A = {1 belongs to my set}

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{99}{C(100,2)} = \frac{1}{50}$$

- Method 2. Experiment: I buy a ticket, and after hesitating one more

  Sample space: S = {permutations of size 2 }

  Event: A = {permutations of size 2 containing 1}

  $$\Pr(A) = \frac{|A|}{|S|} = \frac{99+99}{P(100,2)} = \frac{2\cdot 99}{100\cdot 99} = \frac{1}{50}$$

More General Probability: Crooked Dice

- Suppose we made a loaded dice

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

  $Pr(1) = 1/16$,

  $Pr(2) = Pr(3) = Pr(4) = Pr(5) = 1/8$

  $Pr(6) = 7/16$

  $Pr(\{i,j,\ldots,m\}) = Pr(i) + Pr(j) + \ldots + Pr(m)$

- Find $Pr(\{1,3,5\})$

Monty Hall Game (1970s)

- A contestant is shown three doors; behind one of the doors was a prize (car), and behind the other two were goats.
- The contestant picks a door (but doesn't open it)
- Monty's assistant opens one of the other two doors, revealing a goat. (This is easy since the assistant knows where the prize is.)
- The contestant is then given the option of sticking with his current door, or switching to the other unopened one.
- **Question is**: **Does the contestant have a better chance of winning if he/she switches it?**

The Monty Hall Game (1970s) (contd.)

- Intuitively, it seems obvious that there are only two remaining doors; they must have equal probability.
- You are better off switching. Why?

The Monty Hall Game (1970s) (contd.)

- Intuitively, it seems obvious that there are only two remaining doors; they must have equal probability.
- You are better off switching. Why?
- First write the sample space, and then assign probability to the sample space.
- The sample space: S = {(car, goat, goat), (goat, car, goat), (goat, goat, car)}.
- The probability of each sample point is 1/3. (equal likelihood, uniform probability)

The Monty Hall Game (1970s) (contd.)

- Intuitively, it seems obvious that there are only two remaining doors; they must have equal probability.
- You are better off switching. Why?
- First write the sample space, and then assign probability to the sample space.
- The sample space: S = {(car, goat, goat), (goat, car, goat), (goat, goat, car)}.
- The probability of each sample point is 1/3. (equal likelihood, uniform probability)
- If you don't switch, Pr(winning) = 1/3.
- If you switch, you always win if your initial selection was wrong.
- If you switch, Pr(winning) = 2/3.

Birthday Paradox

- Examine the chances that two people in a group have the same birthday.
- Suppose U={1,2, …, 365}, is the number of days in a year.
- The experiment consists of drawing a sample of n elements from U.
- The size of the sample space : $|S| = 365^n$.

Birthday Paradox

- Examine the chances that two people in a group have the same birthday.
- Suppose U={1,2, …, 365}, is the number of days in a year.
- The experiment consists of drawing a sample of n elements from U.
- The size of the sample space : $|S| = 365^n$.
- Let A be the event that at least two people have the same birthday.
- Let A^c be the complement of A, i.e. it is the event of no two people have the same birthday.
- $|A^c| = 365 \times 364 \times \ldots \times (365 – n + 1)$.
- $Pr(A^c) = (365 \times 364 \times \ldots \times (365 – n + 1))/(365)^n$
- $Pr(A) = 1 – P(A^c)$

Birthday Paradox

- Examine the chances that two people in a group have the same birthday.
- Suppose U={1,2, …, 365}, is the number of days in a year.
- The experiment consists of drawing a sample of n elements from U.
- The size of the sample space : $|S| = 365^n$.
- Let A be the event that at least two people have the same birthday.
- Let A^c be the complement of A, i.e. it is the event of no two people have the same birthday.
- $|A^c| = 365 \times 364 \times \ldots \times (365 - n + 1)$.
- $Pr(A^c) = (365 \times 364 \times \ldots \times (365 - n + 1))/(365)^n$
- $Pr(A) = 1 - P(A^c)$
- **As n increase, Pr(A) increases.**
- **When n = 23, Pr(A) is larger than 50%.**
- **When n= 60 people, Pr(A) is over 99%.**

Tossing with a biased coin

- A coin is biased if Pr({head}) ≠ Pr({tail}).
- Is it possible to toss a coin whose outcome is unbiased, even when you are tossing with a biased coin?
- Let Pr({head}) = p, p is unknown.
- Pr({tail}) = 1-p.
- A simple experiment:
  - flip the coin twice : events A = {head}; B = {tail}
  - if event A is followed by event B, output unbiased-head
  - if event B is followed by event A, output unbiased-tail
  - Otherwise, repeat the experiment.

Example

Example: (6, Exercise 3.4) If two integers are selected, at random and without replacement, from {1, 2, 3, ..., 100}, what is the probability the integers are consecutive.

Summary

- It is important to do probability calculations systematically.
- The key steps are:
  - what is the sample space?
  - what is the probability of each outcome?
  - what is the event we are interested in?
  - Finally, compute the probability of the event by adding up the probabilities of the sample points inside it
- Be careful, there are many pitfalls if you are not careful.

Practice Problems from Grimaldi's book (fifth edition)

- Example 3.31 (page 151), 3.33 (page 152), 3.34 (page 153), 3.37 (page 155).
- Example 3.39 (page 157), 3.40 (page 159).
- Exercises 3.4: 4, 7, 9
- Exercises 3.5: 3, 4, 7