Week 6 6.2 Note

Lecture 19: Chance

---

1. The Prosecutor's Fallacy（检察官谬误）

> The prosecutor's fallacy is a statistical mistake that assumes the **chance of a random match** *equal* to the **chance of defendant's innocence**.

- **Low chance of matching the evidence provided, high chance that the person is guilty.**
- Example:
  - Fact:
    - Suppose there are 5 million people in Sydney.
    - A murder occurs with *DNA left*, and *a person matching* that is arrested.
  - Faulty Arguments:
    - The chance of DNA match is *1 in 500000* (very **small**).
    - So, the *chance* that the arrested person is *guilty* is very **high**.
  - Error in Thinking:

Type	DNA Match	DNA not Match
Guilty	1	0
Innocent	9	4,999,990

  - Note:
    - Only 1 person is guilty and has a DNA match.
    - No one (0) is guilty and has not matched the DNA.
    - If the chance of matching is 1/500000, then 10 people are expected to match (1 guilty, 9 innocent)
    - This leaves 4,999,990 people not matching.
  - Hence:
    - The chance that innocent person **has a DNA match** is tiny.
      - P(DNA Match | Innocent) = 9/4,999,999
    - The chance that a DNA matched person is **innocent** is high.
      - P(Innocent | DNA Match) = 9/10
    - P(DNA Match | Innocent) is **not equal** to P(Innocent | DNA Match)
    - **So we can't say P(Guilty|DNA Match) is high for a DNA matched person.**
      - P(Guilty | DNA Match) = 1 - P(DNA Match | Innocent) = 1 - 9/4,999,999

2. Properties of Chance

> Chance (probability): the **percentage of time** a certain event is expected to *happen*, if the same process is *repeated* long-term.

- Basic Properties:
  1. Chances are between **0%** (impossible) and **100%** (certain).

2. The chance of something equal to **100% minus its opposite (complement)**. —P(Event) = 1 - P(Complement Event)
3. Drawing at **random** means that a collection of objects have **the same chance of being picked**.

3. Conditional Probability

> Conditional Probability（条件概率）: the chance that **a certain event occurs**, given **another event has occurred**. —P(Event 1 | Event 2)

- Multiplication Principle:
  - The probability of 2 events occur is **the chance of 1st event multiplied by the chance of 2nd event**, _given the 1st event has occurred.
  - **P(Event 1 and Event 2 Occur) = P(Event 1) x P(Event 2 | Event 1)**

> Independence（独立性）: 2 events are *independent* if **the chance of 2nd given the 1st is the same as the 2nd**. **—P(2nd Event | 1st Event) = P(2nd Event)**

- Drawing randomly with *replacement* ensures independence.
- P(Event 1 and Event 2 occur) = P(Event 1) x P(Event 2)
- Example: the rain yesterday is irrelevant to the rain today.

> Dependence（依靠性）: 2 events are *dependent* if **the chance of 2nd given the 1st is not the same as the 2nd**. —P(2nd Event | 1st Event) not= P(2nd Event)__

- Drawing randomly without *replacement* ensures dependence.

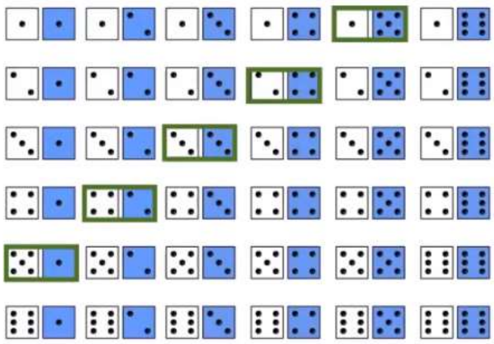
Lecture 20: More Chance

---

1. Making Lists

- For simple chance problems, a good way to start:
  - Write a list of all outcomes
  - Count which outcomes belong to the event of interest.
    - This can lead to a simple way to summarize the outcomes, or use a probability rule.
    - This can also generalizing complicate problems.
- Example: Throw 2 dice. What is the chance of getting a total of 6 spots?
- Write a full list and count the outcomes:

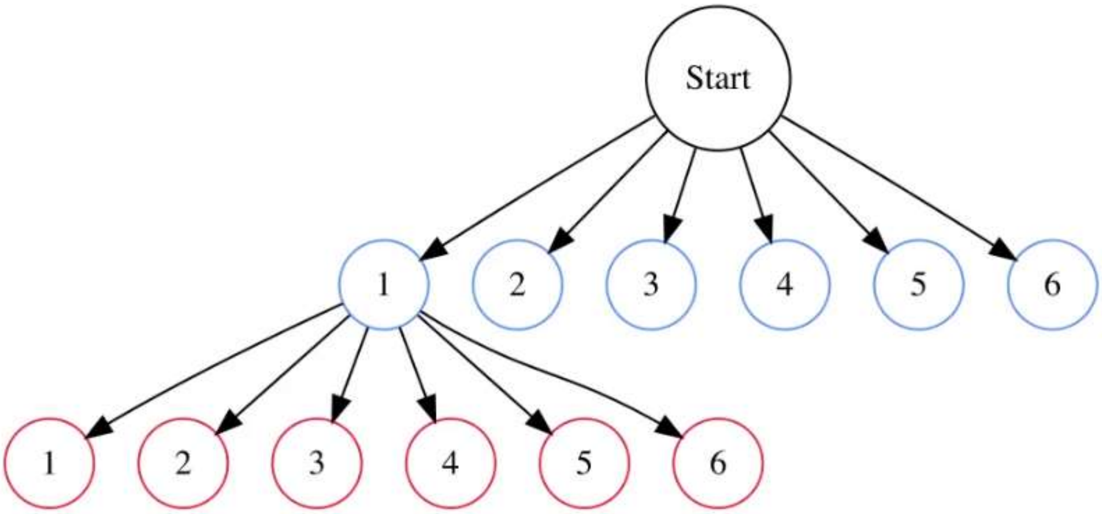
Method1: Write a full list of outcomes and count the outcomes of interest.



So the chance is 5/36 (approx 0.14).

- Summarize in a tree diagram:

Method2: Summarise in a tree diagram



The totals of 6 are (1,5), (2,4), (3,3), (4,2), (5,1) giving 5/36.

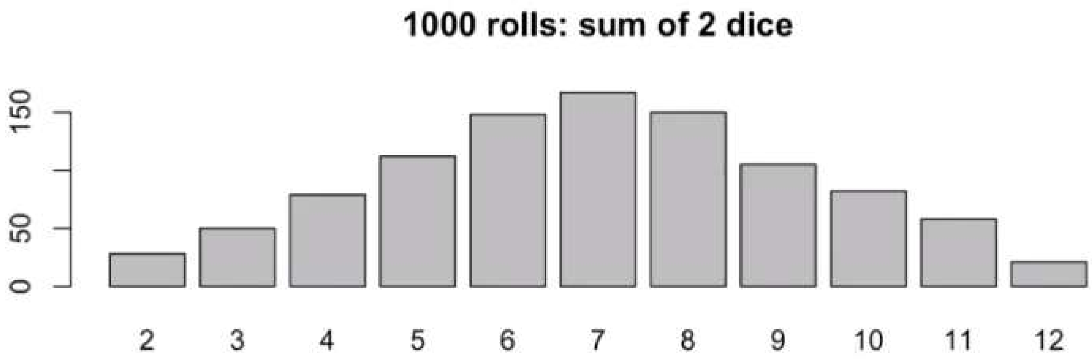
- Simulate using R:

```r
# Start the simulation:
set.seed(1)
totals = sample(1:6, 1000, rep = T) + sample(1:6, 1000, rep = T)
# Introducing commands:
# sample(): create the sample for simulation, each one represents a dice
# 1:6: the sample is from 1 to 6 (numbers on the dice)
# 1000: repeat the simulation for 1000 times
# rep = T: use replacement to make each simulation independent

# Create a table:
table(totals)
```

```
## totals
##   2   3   4   5   6   7   8   9  10  11  12
##  28  50  79 112 148 167 150 105  82  58  21
```

```
barplot(table(totals), main="1000 rolls: sum of 2 dice")
```



- So the chance of getting a total of 6 is 148/1000 = 0.148, very close to 5/36.

2. Addition Rule

> Mutually Exclusive（相互排斥）: 2 things are mutually exclusive if **one event prevents the other**.

- If 2 things are ME, then the chance of at least 1 occurring is **the sum of individual chances**.
- **P(At least 1 event occurs) = P(Event 1) + P(Event 2)**

Lecture 21: Binomial Formula

---

1. Binomial Coefficients

- We can work out the exact chance of P by using the binomial model.
- Suppose we have n objects in a row, made up 2 types: x(type 1) and n - x(type 2):
  - **The number of ways of rearranging the n objects** is given by the **binomial coefficient**: (**n!**)

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

- Example:
  Show (6 2) = 15.

```
choose(6, 2)
# Output:
15
```

2. Binomial Model

> Binary Trial: the case where only 2 things can occur (binary outcomes). So, **P(Event) = p and P(not Event) = 1 - p**.

- **If each trial is dependent (without replacement), then we cannot use binomial model.**
- Binomial Theorem:

  Binomial theorem

  - Suppose we have n independent, binary trials, with P(event)=p at every trial, and n is fixed.
  - The chance that exactly x events occur is

  $$\binom{n}{x} p^x (1-p)^{n-x}$$

  - Example: A fair coin is tossed 5 times. What is the probability of getting 3 heads?

```
# dbinom(x, y, z):
# x - event: we want 3 heads (the event)
# y - time: we toss it 5 times
# z - the probability of getting a head is 50% (0.5)
dbinom(3, 5, 0.5)
# Output:
0.3125
```