Summary---Topic 2: Basic Probability

Terminology

- **Outcome**: result of an experiment (An experiment is a situation which an outcome can be observed)
- **Event**: a collection of outcomes
- **Sample space**: set of all possible outcomes

Simple event	Event which contains only a single outcome in the sample space
Joint event **(A and B)←→($\mathbf{A \cap B}$)**	Intersection of events A and B: An event that both event A and event B occur
Compound event **(A or B) ←→($\mathbf{A \cup B}$)**	Union of events A and B: An event that either event A or event B or both occur
Complement of an event	A' is the complement of the event A, where A' means the event that A does not occur
Mutually exclusive events	Events cannot occur together
Collectively exhaustive events	At least one of the events must occur; the set of events covers the whole sample space

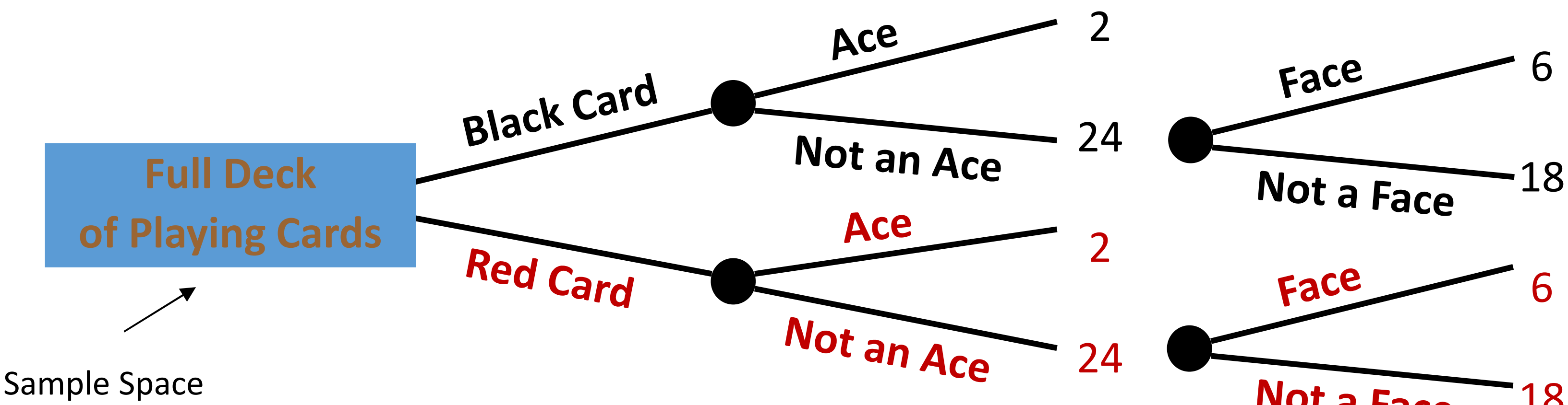
Visualizing Events

- Contingency table

	Ace	Not Ace	Total
Black	2	24	**26**
Red	2	24	**26**
Total	**4**	**48**	**52**

Sample Space

- Decision tree



Probability: Measures of likelihood of an event occurrence

- $0 \leq P(\text{Event}) \leq 1$
- Probability of simple (single characteristic) event A :

$$\mathrm{P(A)} = \frac{Number\ of\ ways/times\ A\ can\ occur}{Total\ number\ of\ possible\ outcomes\ in\ sample\ space}$$

- Probability of joint events A and B (A∩B) :

$$\mathrm{P(A\ and\ B)} = \frac{Number\ of\ outcomes\ from\ both\ A\ and\ B}{Total\ number\ of\ possible\ outcomes\ in\ sample\ space}$$

- Probability of compound events A or B (A∪B) :

$$\mathrm{P}(\mathrm{A}\ or\ B) = \frac{Number\ of\ outcomes\ from\ either\ A\ or\ B\ or\ both}{Total\ number\ of\ possible\ outcomes\ in\ sample\ space}$$

*Addition rule for compound events A or B (A ∪ B) :

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

- Probability of complement of an event A :

$$\mathrm{P(A')} = 1 - P(A)$$

- If event A and event B are said to be **mutually exclusive**, then $\mathrm{P}(\mathrm{A}\ and\ B) = 0$.
- If event A and event B are said to be **collectively exhaustive**, then $\mathrm{P}(\mathrm{A}\ or\ B) = 1$.

Joint Probability

- Joint probability refers to the probability of an occurrence involving two or more events, denoted as P(events A and B and C and D)

Marginal Probability

- Marginal probability is the probability of the occurrence of the single event (simple probability)
- It is often derived from summing up all possible joint events that involves the concerned event
- Formal definition

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \ldots\ldots + P(A \text{ and } B_k),$$

where $B_1, B_2, \ldots\ldots, B_k$ are **mutually exclusive** and **collectively exhaustive** events.

$P(B_i \text{ and } B_j) = 0,\ i \neq j,\ i, j = 1, \ldots, k$

$P(B_1 \text{ or } B_2 \ldots \text{ or } B_k) = 1$

Conditional Probability

- Conditional probability of event A given that event B has occurred,

$$\mathrm{P(A|B)} = \frac{P(A\ and\ B)}{P(B)}$$

- Multiplication rule for joint events:

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

- To determine whether Event A is **statistically independent** on Event B, test whether

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

or

$$\mathrm{P(A|B) = P(A)}$$

or

$$\mathrm{P(B|A) = P(B)}$$

Counting Rules

- Rule 1:

If any one of **k** different mutually exclusive and collectively exhaustive events can occur on each of **n** trials, the number of possible outcomes is equal to k^n

- Rule 2:

If there are k_1 events on the first trial, k_2 events on the second trial, … and k_n events on the nth trial, the number of possible outcomes is

$$k_1 * k_2 * \cdots k_{n-1} * k_n$$

- Rule 3:

The number of ways that n items can be arranged **in order** is

$$\mathrm{n}! = n * (n-1) * \cdots * 2 * 1$$

Counting Rules

- Rule 4: Permutations

The number of ways of arranging X objects selected from n objects in order is

$$ {}_{n}P_{X} = \frac{n!}{(n-X)!} $$

- Rule 5: Combinations

The number of ways of selecting X objects from n objects, irrespective of order, is

$$ {}_{n}C_{X} = \frac{n!}{X! * (n-X)!} $$

Exercises and Solutions

Q1. Each year, ratings are compiled concerning the performance of new cars during the first 90 days of use. Suppose that the cars have been categorized according to whether the car needs warranty-related repair (yes or no) and the country in which the company manufacturing the car is based (United States or not United States). Based on the data collected, the probability that the new car needs warranty repair is 0.04, the probability that the car was manufactured by a U.S.-based company is 0.60, and the probability that the new car needs a warranty repair and was manufactured by a U.S.-based company is 0.025.

Construct a **contingency table** to evaluate the probabilities of a warranty-related repair. What is the probability that a new car selected at random

a)needs a warranty repair?

b)needs a warranty repair and was manufactured by a U.S.-based company?

c)needs a warranty repair or was manufactured by a U.S.-based company?

d)needs a warranty repair or was not manufactured by a U.S.-based company?

Manufacture Country / Needs warranty-related repair	U.S.	Non-U.S.	Total
Yes	0.025	0.04-0.025=0.015	**0.04**
No	0.6-0.025=0.575	0.96-0.575=0.385	**1-0.04=0.96**
Total	**0.600**	**1-0.6=0.4**	**1.00**

Q1. a)needs a warranty repair?

b)needs a warranty repair and was manufactured by a U.S.-based company?

c)needs a warranty repair or was manufactured by a U.S.-based company?

d)needs a warranty repair or was not manufactured by a U.S.-based company?

Manufacture Country / Needs warranty-related repair	U.S.	Non-U.S.	Total
Yes	0.025	0.015	**0.04**
No	0.575	0.385	**0.96**
Total	**0.600**	**0.400**	**1.00**

Solution:

a) P(needs warranty repair) = 0.04

*Addition rule for compound events A or B :

$$P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$$

b) P(needs warranty repair and manufacturer based in U.S.) = 0.025

c) P(needs warranty repair or manufacturer based in U.S.)
=P(needs warranty repair) + P(manufacturer based in U.S.)
- P(needs warranty repair and manufacturer based in U.S.)
= 0.04 + 0.6 – 0.025 = 0.615

d) P(needs warranty repair or manufacturer not based in U.S.)
=P(needs warranty repair) + P(manufacturer not based in U.S.)
- P(needs warranty repair and manufacturer not based in U.S.)
= 0.04 + 0.4 - 0.015 = 0.425

Q2. A sample of 500 respondents was selected in a large metropolitan area to study consumer behavior with the following results:

Enjoys Shopping for Clothing	Gender		
	Male	Female	**Total**
Yes	126	234	**360**
No	104	36	**140**
Total	**230**	**270**	**500**

a)Suppose the respondent chosen is a female. What is the probability that she does not enjoy shopping for clothing?

b)Suppose the respondent chosen enjoys shopping for clothing. What is the probability that the individual is a male?

c)Are enjoying shopping for clothing and the gender of the individual independent? Explain.

Conditional probability :

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

Solution:

a) $\mathrm{P(not\ enjoy|female)} = \frac{P(not\ enjoy\ and\ female)}{P(female)} = \frac{36/500}{270/500} = 0.1333$

b) $\mathrm{P(male|enjoy)} = \frac{P(male\ and\ enjoy)}{P(enjoy)} = \frac{126/500}{360/500} = 0.35$

Q2. A sample of 500 respondents was selected in a large metropolitan area to study consumer behavior with the following results:

Enjoys Shopping for Clothing	Gender		
	Male	Female	**Total**
Yes	126	234	**360**
No	104	36	**140**
Total	**230**	**270**	**500**

c)Are enjoying shopping for clothing and the gender of the individual independent? Explain.

To determine whether Event A is **statistically independent** on Event B, test whether

$$1. P(A \text{ and } B) = P(A) * P(B)$$

or

$$2. \mathrm{P}(\mathrm{A}|\mathrm{B}) = P(A)$$

or

$$3. \mathrm{P}(B|\mathrm{A}) = P(B)$$

Solution:

1. $\mathrm{P}(enjoy\ and\ male) = 126/500 = 0.252$,
   $\mathrm{P}(enjoy) = 360/500 = 0.72$
   $\mathrm{P}(male) = 230/500 = 0.46$

→ $0.72 * 0.46 = 0.3312 \neq 0.252$ → not independent

2. $\mathrm{P}(male|enjoy) = 0.35$, $\mathrm{P}(male) = 230/500 = 0.46$

→ $0.35 \neq 0.46$ → not independent

3. $\mathrm{P}(enjoy|male) = \cdots$, $\mathrm{P}(enjoy) \ldots$, → ?

Q3. a) An advertising executive is studying television viewing habits of married men and women during prime time hours. On the basis of past viewing records, the executive has determined that during prime time, husbands are watching television 60% of the time. It has also been determined that when the husband is watching television, 40% of the time the wife is also watching. When the husband is not watching television, 30% of the time the wife is watching television.

(i) Find the probability that both the wife and the husband are watching television in prime time. ***P(H and W) = ?***

(ii) Find the probability that the wife is watching television in prime time. ***P(W) = ?***

Solution:

- Let H: husband watch TV, H': husband does not watch TV
- Let W: wife watch TV, W': wife does not watch TV

✓ $P(H)=0.6$, ➔ P(H') = 1-0.6=0.4

✓ $P(W|H) = 0.4$,

✓ $P(W|H') = 0.3$

*Multiplication rule for joint events:

$P(A\ and\ B) = P(A|B) * P(B) = P(B|A) * P(A)$

*Marginal probability:

$P(A) = P(A \text{ and } B_1)+P(A \text{ and } B_2)+\ldots\ldots+P(A \text{ and } B_k)$,

i) $P(H \text{ and } W) = P(W|H)*P(H) = 0.4*0.6 = 0.24$

ii) Since $P(W) = P(W \text{ and } H) + P(W \text{ and } H') = 0.24 + P(W \text{ and } H')$,

and $P(W \text{ and } H') = P(W|H') * P(H') = 0.3*0.4 = 0.12$,

we obtain that $P(W) = 0.24+0.12 = 0.36$.

Q3. b) A sample of 500 households was selected in a large metropolitan area to determine various information concerning consumer behavior. The following information was obtained:

Enjoys shopping for clothing	Gender		
	Male	Female	**Total**
Yes	136	224	**360**
No	104	36	**140**
Total	**240**	**260**	**500**

(i) Suppose the respondent chosen is a female. What, then, is the probability that she does not enjoy shopping for clothing?

(ii) Suppose the respondent chosen enjoys shopping for clothing. What, then, is the probability that the individual is a male?

(iii) Are enjoy shopping for clothing and the gender of the individual statistically independent? Explain.

Solution: i) $\mathrm{P}(\text{not enjoy}|\text{female}) = \frac{P(not\ enjoy\ and\ female)}{P(female)} = \frac{36/500}{260/500} = 0.1384$

ii) $\mathrm{P}(male|enjoy) = \frac{P(enjoy\ and\ male)}{P(enjoy)} = \frac{136/500}{360/500} = 0.3778$

iii) $\mathrm{P}(male|enjoy) = 0.3778, \quad \mathrm{P}(male) = \frac{240}{500} = 0.48$

$\rightarrow$ Not equal $\rightarrow$ not independent.

Q4. If there are 10 multiple-choice questions on an exam, each having three possible answers, how many different sequences of answers are there?

Rule 1:
If any one of k different mutually exclusive and collectively exhaustive events can occur on each of n trials, the number of possible outcomes is equal to k^n

Solution:

- $\mathrm{n} = 10,\ \mathrm{k} = 3$

$$\text{the number of sequences of answers} = 3^{10} = 59049.$$

Q5. You would like to "build-your-own-burger" at a fast-food restaurant. <u>There are five different breads, seven different chesses, four different cold toppings, and five different sauces on the menu.</u> If you want to include one choice from each of these ingredient categories, how many different burgers can you build?

Rule 2:
If there are k_1 events on the first trial, k_2 events on the second trial, … and k_n events on the n^{th} trial, the number of possible outcomes is

$$k_1 * k_2 * \cdots k_{n-1} * k_n$$

Solution:

- There are 4 steps to build the burger → $n=4$
- $k_1 = 5,\ k_2 = 7,\ k_3 = 4,\ k_4 = 5$

$$\text{the number of possible burgers} = k_1 * k_2 * k_3 * k_4 = 700.$$

Q6. Four members of a group of 10 people are to be selected to a team. How many ways are there to select these four members?

Rule 5: Combinations

The number of ways of selecting X objects from n objects, irrespective of order, is

$$_nC_X = \frac{n!}{X! * (n-X)!}$$

Solution:

- $\mathrm{X} = 4$
- $n = 10$

$$_{10}C_4 = \frac{10!}{4! * (10-4)!} = \frac{10!}{4! * 6!} = \frac{10 * 9 * \cdots * 2 * 1}{(4 * \cdots * 1) * (6 * 5 * \cdots * 1)} = 210\ ways$$