Econ 325: Introduction to Empirical Economics

Lecture 2

Probability

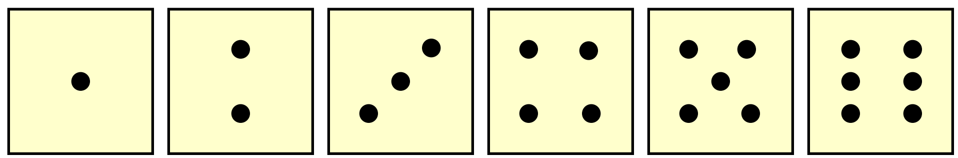
3.1

Definition

- Random Experiment – a process leading to an uncertain outcome
- Basic Outcome – a possible outcome of a random experiment
- Sample Space – the collection of all possible outcomes of a random experiment
- Event – any subset of basic outcomes from the sample space

Example

Let the Sample Space be the collection of all possible outcomes of rolling one die:



S = [1, 2, 3, 4, 5, 6]

Let A be the event “Number rolled is even”

Let B be the event “Number rolled is at least 4”

Then

A = [2, 4, 6] and B = [4, 5, 6]

Clicker Question 2.1

What is Sample Space of rolling two dice?

A). S = [(1,1),(1,2),…,(1,6), (2,1),(2,2), …,(2,6), (3,1),…,(3,6), (4,1),…, (4,6), (5,1),…, (5,6), (6,1),…,(6.6)]

B). S = [(1,1),(1,2),…,(1,6), (2,2), …,(2,6), (3,3),…,(3,6), (4,4),(4,5),(4,6), (5,5),(5,6), (6.6)]

C). It depends on how you define sample space.

Clicker Question 2.1

- For A, the order is taken into account to define basic outcomes, and there are 36 basic outcomes.

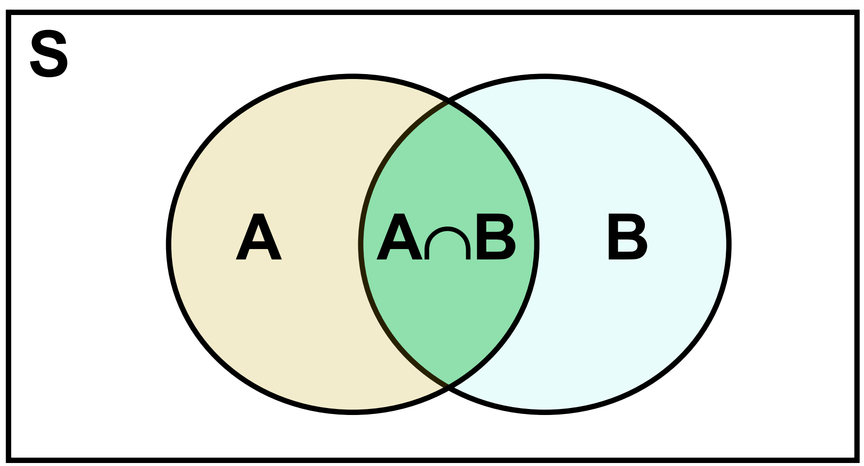
- For B, the order does not matter. For example, (2,1) is the same outcome as (1,2). In this case, there are 21 basic outcomes.

Sample Space for A

	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Definition

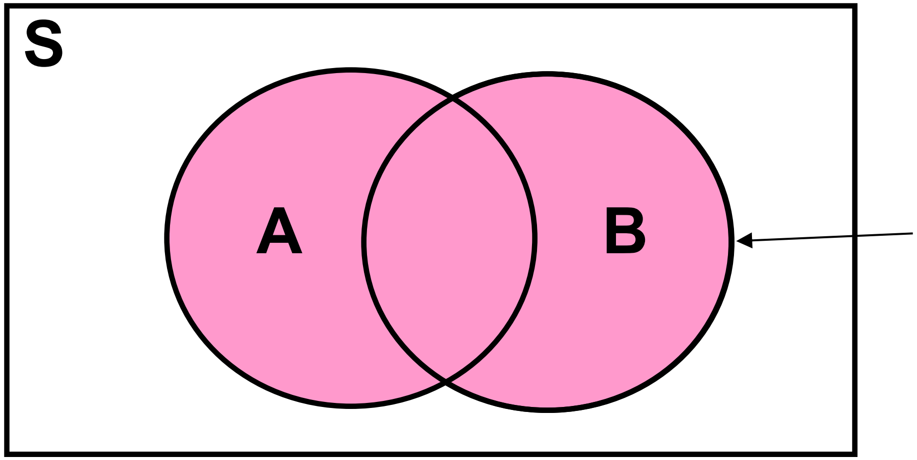
- Intersection of Events – If A and B are two events in a sample space S, then the intersection, A ∩ B, is the set of all outcomes in S that belong to both A and B



Definition

- Union of Events – If A and B are two events in a sample space S, then the union, A ∪ B, is the set of all outcomes in S that belong to either

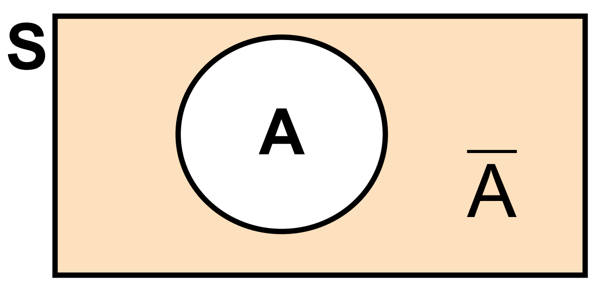
  A or B



The entire shaded area represents A ∪ B

Definition

- The Complement of an event A is the set of all basic outcomes in the sample space that do not belong to A. The complement is denoted $\overline{A}$



Properties of Set Operations

- Commutative: $A \cup B = B \cup A$, $A \cap B = B \cap A$
- Associative: $(A \cup B) \cup C = A \cup (B \cup C)$
- Distributive Law:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

- De Morgan's Law:

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B}$$

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B}$$

Examples

S = [1, 2, 3, 4, 5, 6] | **A = [2, 4, 6]** | **B = [4, 5, 6]**

Complements:

$$\overline{A} = [1, 3, 5] \qquad \overline{B} = [1, 2, 3]$$

Intersections:

$$A \cap B = [4, 6] \qquad \overline{A} \cap B = [5]$$

Unions:

$$A \cup B = [2, 4, 5, 6] \qquad A \cup \overline{B} = [1,2,3,4,6]$$

$$A \cup \overline{A} = [1, 2, 3, 4, 5, 6] = S$$

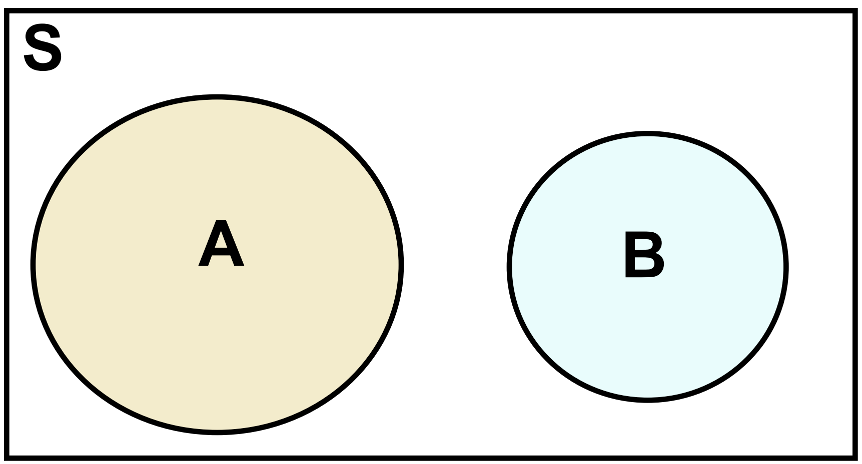
Definition

- Events $E_1, E_2, \ldots E_k$ are Collectively Exhaustive events if $E_1 \cup E_2 \cup \ldots \cup E_k = S$

- i.e., the collection of events are collectively exhaustive if they completely cover the sample space.

Definition

- A and B are Mutually Exclusive Events if they have no basic outcomes in common
  - i.e., the set A ∩ B is empty



True or False

S = [1, 2, 3, 4, 5, 6]

A = [2, 4, 6] ,   B = [4, 5, 6]

A and B are mutually exclusive.

A). True

B). False

True or False

S = [1, 2, 3, 4, 5, 6]

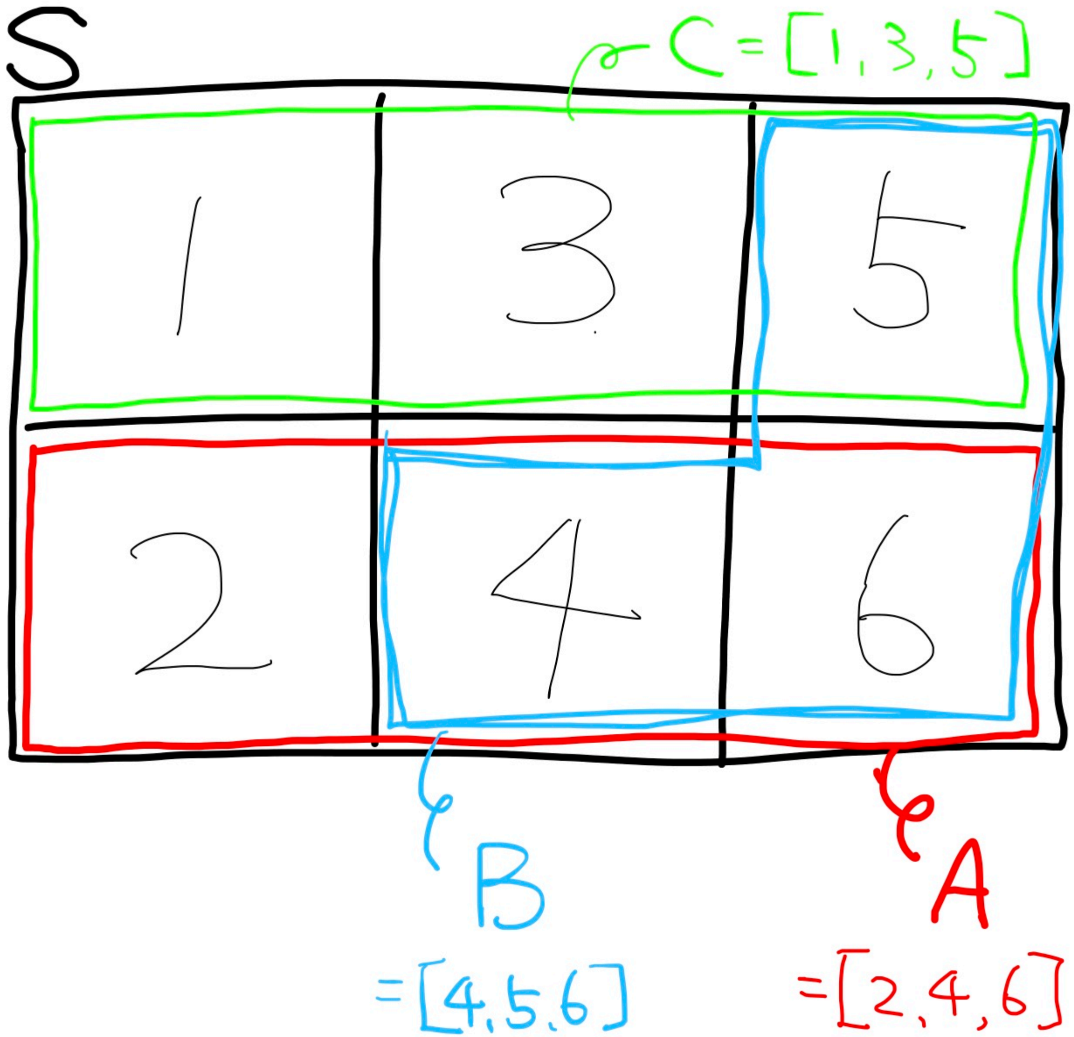
A = [2, 4, 6] ,   B = [4, 5, 6],   C = [1, 3, 5]

A, B, and C are collectively exhaustive.

A). True

B). False

True or False



Example

Consider a set A and its complement $\overline{A}$.

- A and $\overline{A}$ is mutually exclusive because they share no common element.
- A and $\overline{A}$ is collectively exhaustive because their union covers all elements in the sample space.

Clicker Question 2.2

- For any two events A and B,

1. $(A \cap B) \cap (A \cap \overline{B})$ is empty.
2. $A = (A \cap B) \cup (A \cap \overline{B})$

A). Both 1 and 2 are true.

B). Both 1 and 2 are false.

C). 1 is true and 2 is false.

D). 1 is false and 2 is true.

Probability

Rolling a fair die, what is the probability of A = [2,4,6]?

P(A) = 1/2

Are you sure?

How can we define the probability of event A?

Probability as relative frequency

- Consider repeating the experiment n times.
- Count the number of times that event A occurred: n_A
- The relative frequency is n_A / n
- Take the limit of $n \to \infty$

$$P(A)$$

$$= \lim_{n \to \infty} \frac{n_A}{n}$$

$$= \frac{\text{number of events in the population that satisfy event A}}{\text{total number of events in the population}}$$

Clicker Question 2.3

If two dice are rolled, what is the probability that the sum of two numbers is less than or equal to 3?

A). 1/3

B). 1/6

C). 1/12

Clicker Question 2.3

What is the probability that the sum of the numbers is less than or equal to 3?

	2	3	4	5	6	7
	3	4	5	6	7	8
	4	5	6	7	8	9
	5	6	7	8	9	10
	6	7	8	9	10	11
	7	8	9	10	11	12

Factorial & Permutation formula

- Marathon race by 8 runners:

$$\{A, B, C, D, E, F, G, H\}$$

- How many ways to order 8 runners in a sequence?

Answer: $8 \cdot 7 \cdot 6 \cdot \quad \cdot 2 \cdot 1 = 8!$

- How many ways to pick the 1st, 2nd, and 3rd places (with ordering)?

Answer: $8 \cdot 7 \cdot 6 = \frac{8!}{5!}$

Combination formula

- How many ways to pick 3 out of 8 runners?

Answer: $\frac{\left(\frac{8!}{5!}\right)}{3!}$

- This is the combination formula:

$$C_k^n = \frac{\left(\frac{n!}{(n-k)!}\right)}{k!} = \frac{n!}{k!\,(n-k)!}$$

Counting the Possible Outcomes

- Use the Combination formula to determine the number of unordered ways in which k objects can be selected from n objects

$$C_k^n = \frac{n!}{k!(n-k)!}$$

- where
  - $n! = n(n-1)(n-2)\ldots(1)$
  - $0! = 1$ by definition

Example

- A personnel officer has 5 candidates to fill 2 positions.
- 3 candidates are men and 2 candidates are women.
- If every candidate is equally likely to be chosen, what is the probability that no women will be hired?

Example

- The total number of possible combinations:

$$C_2^5 = \frac{5!}{2!(5-2)!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{2\cdot1\cdot(3\cdot2\cdot1)} = \frac{20}{2} = 10$$

- The number of possible combinations that both hired persons are men:

$$C_2^3 = \frac{3!}{2!(3-2)!} = \frac{3\cdot2\cdot1}{2\cdot1\cdot(1)} = \frac{6}{2} = 3$$

- The probability that no women is hired:

3/10=30%

Example

- 5 candidates

  W_1, W_2, M_1, M_2, M_3

- Basic Outcomes

  (W_i, M_j) for $i = 1, 2$, $j = 1, 2, 3$

  (W_1, W_2)

  (M_1, M_2), (M_2, M_3), (M_1, M_3)

$C^5_2 = 10$ possible combinations

$C^3_2 = 3$ possible combinations

Answer: $\frac{C^3_2}{C^5_2} = \frac{3}{10}$

Another Example

- A personnel officer has 10 candidates to fill 3 positions.
- 4 candidates are men and 6 candidates are women.
- If every candidate is equally likely to be chosen, what is the probability that no women will be hired?
- Answer: C_3^4 / C_3^{10}

Probability as a set function

Probability is a real-valued set function P that assigns, to each event A in the sample space S, a number P(A) that satisfies the following three properties:

1. $P(A) \geq 0$
2. $P(S) = 1$
3. If $A_1, A_2, \ldots, A_k$ are mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = P(A_1) + P(A_2) + \ldots + P(A_k)$$

for any positive integer k.

3.3

Probability Rules

- The Complement rule:

$$P(\overline{A}) = 1 - P(A) \qquad \text{i.e., } P(A) + P(\overline{A}) = 1$$

- The Addition rule:
  - The probability of the union of two events is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

A Probability Table

Probabilities and joint probabilities for two events A and B are summarized in this table:

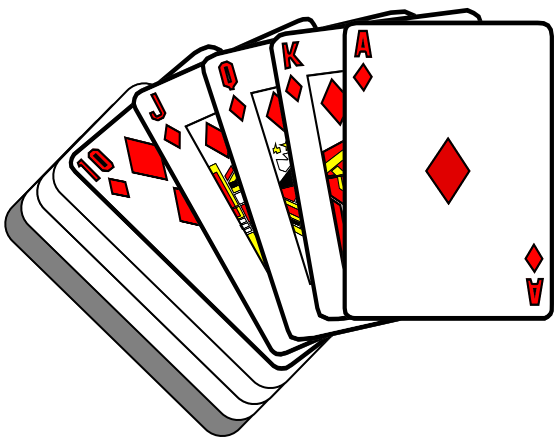
	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	$P(S) = 1.0$

Addition Rule Example

Consider a standard deck of 52 cards, with four suits: ♥ ♣ ♦ ♠

Let event A = card is an Ace

Let event B = card is from a red suit



Addition Rule Example

(continued)

$$P(\text{Red} \cup \text{Ace}) = P(\text{Red}) + P(\text{Ace}) - P(\text{Red} \cap \text{Ace})$$

$$= 26/52 + 4/52 - 2/52 = 28/52$$

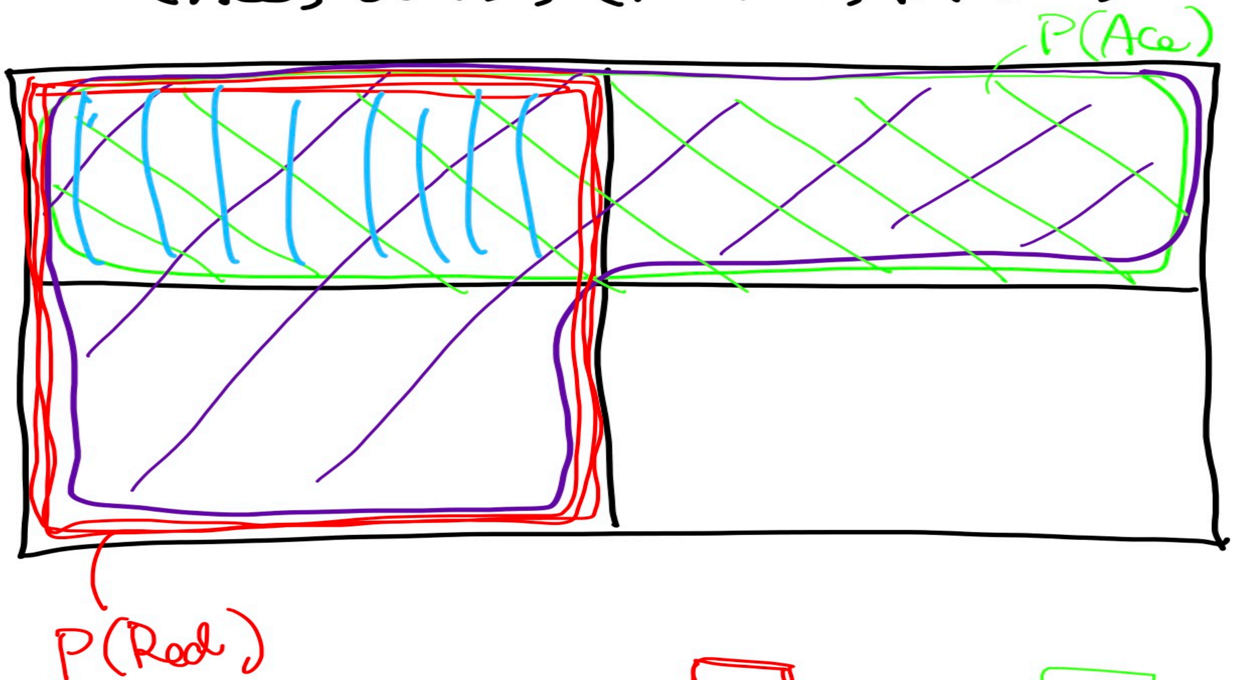
Don't count the two red aces twice!

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Addition Rule

4 Basic Outcomes

(Ace, Red), (Non-Ace, Red),
(Ace, Black), (Non-Ace, Black)



$$P(\text{Red} \cup \text{Ace}) = P(\text{Red}) + P(\text{Ace}) - P(\text{Red} \cap \text{Ace})$$

Clicker Question 2.4

- **True or false?**

$$P(A) = P(A \cap B) + P(A \cap \overline{B})$$

A). True
B). False
C). Depends on case by case

Conditional Probability

- A conditional probability is the probability of one event, given that another event has occurred:

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

→ **The conditional probability of A given that B has occurred**

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

→ **The conditional probability of B given that A has occurred**

Clicker Question 2.5

Throw a fair die. I tell you that the outcome is an even number. What is the probability of having rolled a ``6'' given the information that it is an ``even number''?

A). 1/2
B). 1/3
C). 1/6

Clicker Question 2.5

A = [6]

B = [2, 4, 6]

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$

Clicker Question 2.6

- Roll two dice. What is the probability that at least one die is equal to 2 when the sum of two numbers is less than or equal to 3?

A). 1/2

B). 1/3

C). 1/4

D). 2/3

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

Multiplication Rule

- Multiplication rule for two events A and B:

$$P(A \cap B) = P(A|B)P(B)$$

- also

$$P(A \cap B) = P(B|A)P(A)$$

Multiplication Rule Example

P(Red ∩ Ace) = P(Red| Ace)P(Ace)

$$= \left(\frac{2}{4}\right)\left(\frac{4}{52}\right) = \frac{2}{52}$$

$$= \frac{\text{number of cards that are red and ace}}{\text{total number of cards}} = \frac{2}{52}$$

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Statistical Independence

- Two events are statistically independent if and only if:

$$P(A \cap B) = P(A)P(B)$$

  - Events A and B are independent when the probability of one event is not affected by the other event

- If A and B are independent, then

$$P(A|B) = P(A)$$ **if P(B)>0**

$$P(B|A) = P(B)$$ **if P(A)>0**

Statistical Independence

- To check if two events are independent or not, we need to check the formal condition:

$$P(A \cap B) = P(A)P(B)$$

- In some cases, we have a good intuition for independence.
  - ``Red'' and ``Ace'' are independent.
  - Tossing two coins, ``Head in the first coin'' is independent of ``Head in the second coin.''

Example

P(Red ∩ Ace) = 2 / 52 = 1/26

P(Red) x P(Ace) = (1/2) x (1/13) = 1/26

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Clicker Question 2.7

Roll a fair die once, define

A = [2, 4, 6],

B = [1, 2, 3, 4].

Are the events A and B statistically independent?

A). Yes

B). No

Clicker Question 2.7

- Check if $P(A \cap B) = P(A)P(B)$ holds or not.
- $P(A \cap B) = P([2,4]) = \frac{2}{6} = \frac{1}{3}$
- $P(A)\ P(B) = \frac{1}{3}$

  because

$$P(A) = P([2,4,6]) = \frac{3}{6} = \frac{1}{2}$$

$$P(B) = P([1,2,3,4]) = \frac{4}{6} = \frac{2}{3}$$

Bivariate Probabilities

Outcomes for bivariate events:

	B_1	B_2	. . .	B_k
A_1	$P(A_1 \cap B_1)$	$P(A_1 \cap B_2)$	. . .	$P(A_1 \cap B_k)$
A_2	$P(A_2 \cap B_1)$	$P(A_2 \cap B_2)$	. . .	$P(A_2 \cap B_k)$
. . .	. . .	. . .	. . .	. . .
A_h	$P(A_h \cap B_1)$	$P(A_h \cap B_2)$	. . .	$P(A_h \cap B_k)$

Marginal Probabilities

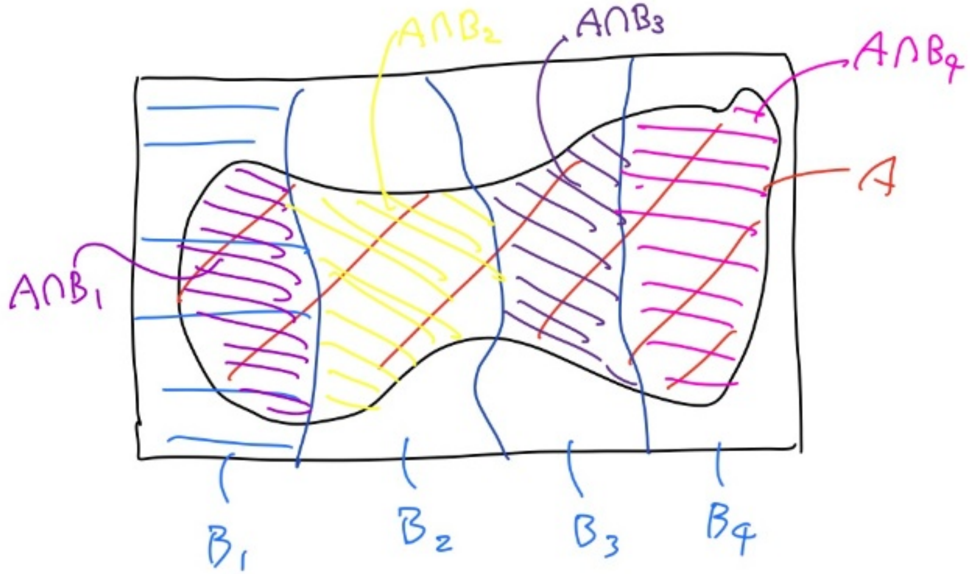
$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)$$

$$B_i \cap B_j = \phi \quad \text{for } i \neq j$$

$$\Rightarrow (A \cap B_i) \cap (A \cap B_j) = \phi \quad \text{for } i \neq j.$$

Then,

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$$



Marginal Probabilities

- The Marginal Probability of P(A):

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$$

when B_1, B_2, …, B_k are k mutually exclusive and collectively exhaustive events

Marginal Probability Example

P(Ace)

$$= P(\text{Ace} \cap \text{Red}) + P(\text{Ace} \cap \text{Black}) = \frac{2}{52} + \frac{2}{52} = \frac{4}{52}$$

Type	Color		Total
	Red	Black	
Ace	2	2	4
Non-Ace	24	24	48
Total	26	26	52

Joint Distribution of X and Y

- Consider two random variables: X and Y
- X takes n possible values:

$$\{x_1, x_2, \ldots, x_n\}$$

- Y takes m possible values:

$$\{y_1, y_2, \ldots, y_m\}$$

- Joint Distribution of X and Y can be described by Bivariate probabilities.

3.4

Distribution of (X,Y)

	X=x_1	X=x_2	. . .	X=x_n
Y=y_1	P$(X = x_1, \mathrm{Y} = y_1)$	P$(X = x_2, \mathrm{Y} = y_1)$	. . .	P$(X = x_n, \mathrm{Y} = y_1)$
Y=y_2	P$(X = x_1, \mathrm{Y} = y_m)$	P$(X = x_2, \mathrm{Y} = y_2)$	. . .	P$(X = x_n, \mathrm{Y} = y_2)$
. . .	. . .	. . .	. . .	. . .
Y=y_m	P$(X = x_1, \mathrm{Y} = y_m)$	P$(X = x_2, \mathrm{Y} = y_m)$	. . .	P$(X = x_n, \mathrm{Y} = y_m)$

Example

- Joint Distribution of University Degree (X) and Annual Income (Y)

	Y=30	Y=60	Y=100
X=0	0.24	0.12	0.04
X=1	0.12	0.36	0.12

Marginal Probabilities

- The Marginal Probability of P$(X = x_1)$

$$P(X = x_1) = P(X = x_1, Y = y_1) + P(X = x_1, Y = y_2) + \cdots + P(X = x_1, Y = y_m)$$

Clicker Question 2.8

- What is P(X=1)?

	Y=30	Y=60	Y=100	Marginal Dist. of X
X=0	0.24	0.12	0.04	0.40
X=1	0.12	0.36	0.12	??
Marginal Dist. of Y	0.36	0.48	0.16	1.00

A). 0.6 B). 0.24 C). 0.2

Bayes' theorem example

- You took a blood test for cancer diagnosis.
- Your blood test was positive.
- This blood test is positive with probability 95 percent if you indeed have a cancer.
- This blood test is positive with probability 10 percent if you don't have a cancer.
- How do you compute the probability of your having a cancer if 1 percent of people in population have cancer?

Bayes' theorem example

- Event B = {Cancer}
- Event A = {Positive Blood test}
- How can we learn about the probability of B from observing that A happened?
- We would like to compute P(B|A).

Bayes' Theorem (simple version)

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

$$= \frac{P(A \cap B)}{P(A \cap B) + P(A \cap \overline{B})}$$

$$= \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})}$$

Bayes' Theorem Example

- If a person has cancer (B), a blood test is positive (A) with 95% probability. If a person is free of cancer ($\bar{B}$), the test comes back positive (A) with 10% probability.

  $$P(A \mid B) = 0.95 \quad \text{and} \quad P(A \mid \bar{B}) = 0.10$$

- 1% people have cancer: $P(B) = 0.01$.
- What is the probability that you have cancer when your blood test is positive?

Bayes' Theorem Example

- $P(B) = .01$, $P(\bar{B}) = 1\text{-}P(B) = .99$
- $P(A \mid B) = .95$, $P(A \mid \bar{B}) = .10$
- Join distribution of A and B

	$\bar{B}$	B
$\bar{A}$	8910 (89.1%)	5 (0.05%)
A	990 (9.9%)	95 (0.95%)
Total	9900 (99%)	100 (1%)

Bayes' Theorem Example

(continued)

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid \bar{B})P(\bar{B})}$$

$$= \frac{(.95)(.01)}{(.95)(.01) + (0.10)(.99)}$$

$$= \frac{.0095}{.0095. + .099} = .0876$$

The revised probability of having cancer is 8.76 percent!

Total Law of Probability

$B_1 \cup B_2 \cup \cdots \cup B_k = S$ (Collectively exhaustive)

$B_i \cap B_j = \phi$ for $i \neq j$ (Mutually exclusive)

Partition A by $\{B_i\}_{i=1}^{k}$ as

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)$$

Note: $(A \cap B_i) \cap (A \cap B_j) = \phi$ for $i \neq j$.

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cup B_k)$$

$$= \sum_{i=1}^{k} P(A \cap B_i)$$

$$= \sum_{i=1}^{k} P(A \mid B_i) P(B_i)$$

Bayes' Theorem (general version)

3.5

- Let B_1, B_2 , …, B_k be mutually exclusive and collectively exhaustive events. Then,

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + \ldots + P(A \mid B_k)P(B_k)}$$