Probability and Counting Rules

A researcher claims that 10% of a large population have disease H.

A random sample of 100 people is taken from this population and examined.

If 20 people in this random sample have the disease, what does it mean? **How likely** would this happen if the researcher is right?

1

A Simple Example

- What's the probability of getting a **head** on the toss of a single fair coin? Use a scale from **0** (**no way**) to **1** (**sure thing**).
- **So toss a coin twice.** Do it! Did you get one head & one tail? What's it all mean?

2

Sample Space and Probability

- **Random Experiment**: (Probability Experiment) an experiment whose outcomes depend on chance.
- **Sample Space (S)**: collection of all possible outcomes in random experiment.
- **Event**: a collection of outcomes

3

Sample Space and Event

- Sample Space:

S = {Head, Tail}
S = {0, 1, 2, 3, 4} = {Grade points for a course}
S = {Life span of a human} = $\{x \mid x \geq 0, x \in R\}$

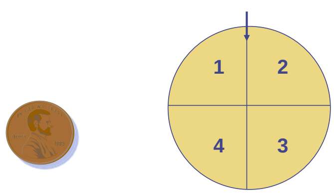
- Event:

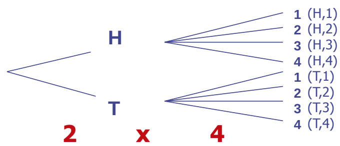
E = {Head}
E = {2, 3, 4} = $\{x \mid 2 \leq x \leq 4, x \in I\}$
E = {Life span of a human is less than 3 years}

4

Sample Space = {(H,1),(H,2),(H,3),(H,4),(T,1),(T,2),(T,3),(T,4)}

Tree Diagram

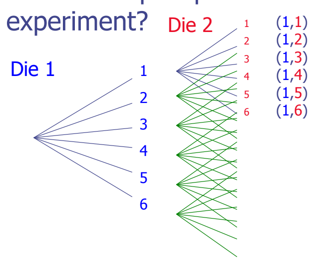




5

Tree Diagram

What is the sample space for casting two dice experiment?



6 x 6 = 36 outcomes

6

Sample Space (Two Dice)

Sample Space:

S = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),
(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),
(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),
(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),
(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),
(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}

Compound event

Simple event

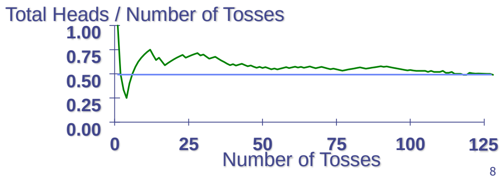
How many outcomes are "**Sum is 7**"?

7

Definition of Probability

A rough definition: (frequentist definition)

Probability of a certain outcome to occur in a *random experiment* is the **proportion of times** that the this outcome would occur in a very long series of repetitions of the random experiment.



8

Determining Probability

How to determine probability?

- Empirical Probability
- Theoretical Probability
- (Subjective approach)

9

Empirical Probability Assignment

Empirical study: (Don't know if it is a balanced Coin?)

Outcome	Frequency
Head	512
Tail	488
Total	1000

10

Empirical Probability Assignment

Empirical probability assignment:

$$P(E) = \frac{\text{Number of times event E occurs}}{\text{Number of times experiment is repeated}}$$

$$= \frac{m}{n}$$

Probability of Head:

$$\textbf{P(Head)} = \frac{512}{1000} = \underline{.512} = 51.2\%$$

11

Empirical Probability Distribution

Empirical study:

Outcome	Frequency	Probability
Head	512	.512
Tail	488	.488
Total	1000	1.0

Empirical Probability Distribution

12

Theoretical Probability Assignment

Make a reasonable assumption:

What is the probability distribution in tossing a coin?

Assumption:

We have a balanced coin!

Theoretical Probability Assignment

Theoretical probability assignment:

$$P(E) = \frac{\text{Number of equally likely outcomes in event E}}{\text{Size of the sample space}}$$

$$= \frac{n(E)}{n(S)}$$

Probability of Head:

$$P(\text{Head}) = \frac{1}{2} = .5 = 50\%$$

Theoretical Probability Distribution

Empirical study:

Outcome	Probability
Head	.50
Tail	.50
Total	1.0

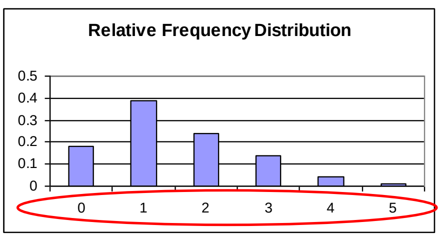
Empirical Probability Distribution

Relative Frequency and Probability Distributions

Number of times visited a doctor from a random sample of 300 individuals from a community

Class	Frequency	Relative Frequency	
0	54	.18	P(0) = .18
1	117	.39	P(1) = .39
2	72	.24	P(2) = .24
3	42	.14	P(3) = .14
4	12	.04	P(4) = .04
5	3	.01	P(5) = .01
Total	300	1.00	

Discrete Distribution



Relative Frequency and Probability

When selecting one individual at random from a population, the probability distribution and the relative frequency distribution are the same.

Discrete Distribution

If an individual is randomly selected from this group 300, what is the probability that this person visited doctor 4 or 5 times?

P(4 or 5 times) = P(4) + P(5)
= .04 + .01
= .05

Class	Frequency	Relative Frequency
0	54	.18
1	117	.39
2	72	.24
3	42	.14
4	12	.04
5	3	.01
Total	300	1.00

It would be an empirical probability distribution, if the sample of 300 individuals is utilized for understanding a large population.

19

Properties of Probability

- **Probability is always a value between 0 and 1.**
- **Total probability (sum of the probabilities of all outcomes in the sample space) equals 1.**
- **Probability of an event that can never occur is 0.**

20

Probability Distribution (Probability Model)

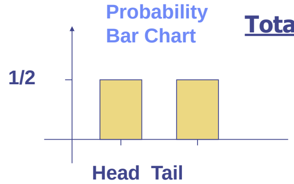
If a balanced coin is tossed, *Head* (H) *and Tail* (T) *are equally likely to occur,*

P(H) =.5 and **P(T) =.5**

P(all possible outcomes) = P(H) + P(T) = .5 + .5 = 1

Probability Bar Chart

Total probability is 1



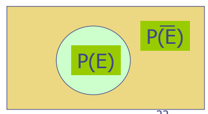
21

Complementation Rule

For any event E,

P(E does not occur) = 1 – P(E)

Complement of E = $\bar{E}$



* Some places use E^c or E′

22

Complementation Rule

If an unbalanced coin has a probability of 0.7 to turn up Head each time tossing this coin. What is the probability of not getting a Head for a random toss?

P(not getting Head) = 1 – 0.7
= 0.3

23

Complementation Rule

If the chance of a randomly selected individual living in community A to have disease H is .001, what is the probability that this person does not have disease H?

P(having disease H) = .001

P(not having disease H)
= 1 – P(having disease H)
= 1 – 0.001
= **0.999**

24

Rules of Probability

Topics

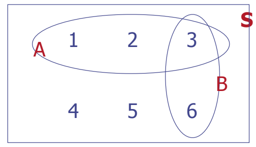
- Notations ∪ and ∩
- Venn Diagram
- Addition Rules
- Conditional Probability
- Multiplication Rules
- Odds Ratio & Relative Risk

25

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

A = {1, 2, 3}

B = {3, 6}



Venn Diagram

(with elements listed)

Intersection of events:

A ∩ B <=> A and B

Example: A ∩ B = {3}

Union of events:

A ∪ B <=> A or B

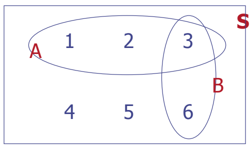
Example: A ∪ B = {1, 2, 3, 6}

26

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

A = {1, 2, 3}

B = {3, 6}



Venn Diagram

(with elements listed)

If a balanced die is used,

P(A ∩ B) = **1/6**

P(A ∪ B) = 4/6 = **2/3**

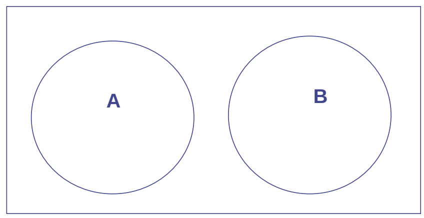
27

Addition Rules for Probability

Additive Rule 1

If A and B are mutually exclusive (disjointed) events,

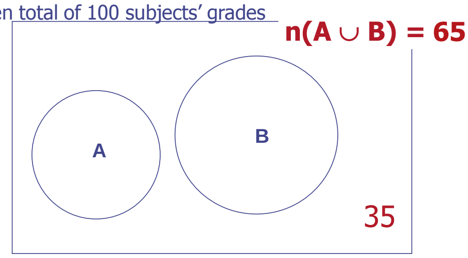
P(A ∪ B) = P(A) + P(B)



28

Venn Diagram (with counts)

Given total of 100 subjects' grades



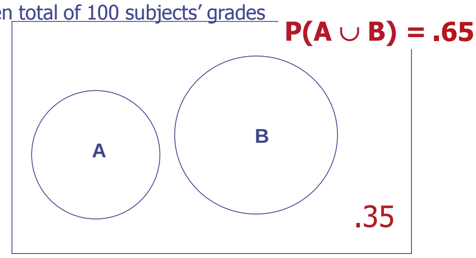
n(A) = 20

n(B) = 45

29

Venn Diagram (with relative frequencies)

Given total of 100 subjects' grades



P(A) = .20

P(B) = .45

Additive Rule 1: P(A ∪ B) = P(A) + P(B) = .2 + .45 = **.65**

30

Addition Rule 1
(Special Addition Rule)

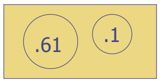
In an experiment of casting an unbalanced die, the event A and B are defined as the following:

A = an odds number

B = 6

and given P(A) = .61, P(B) = 0.1. What is the probability of getting an odds number or a six?

A and B are mutually exclusive.



$P(A \cup B) = P(A) + P(B) = .61 + .1 = .71$

31

Addition Rule 1 - generalized
(for multiple events)

In $A_1, A_2, \ldots, A_k$ are defined k mutually exclusive (m.e.) then

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

Example: When casting a balanced die, the probability of getting 1 or 2 or 3 is

$P(1 \cup 2 \cup 3) = P(1) + P(2) + P(3)$
$= 1/6 + 1/6 + 1/6 = 1/2$

32

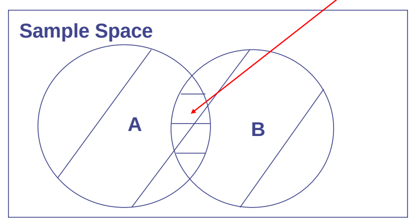
Example: When casting a unbalanced die, such that P(1) = .2, P(2) = .1, P(3) =.3, the probability of getting 1 or 2 or 3 is

$P(1 \cup 2 \cup 3) = P(1) + P(2) + P(3)$
$= .2 + .1 + .3 = \mathbf{.6}$

33

General Addition Rule

Not M.E.!
$A \cap B \neq \varnothing$



Addition Rule 2

If A and B are any two events, then

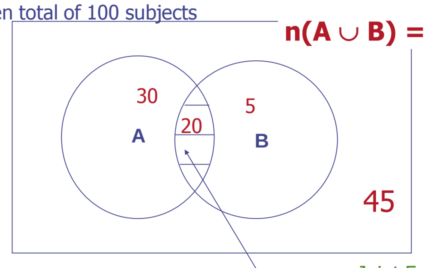
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

34

Venn Diagram (with counts)

Given total of 100 subjects

$n(A \cup B) = 55$



A=Smokers, n(A) = 50

B=Lung Cancer, n(B) = 25

$A \cap B$ ← Joint Event

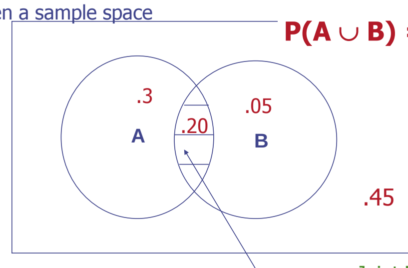
$n(A \cap B) = 20$

35

Venn Diagram (with relative frequencies)

Given a sample space

$P(A \cup B) = .55$



A=Smokers, P(A) = .50

B=Lung Cancer, P(B) = .25

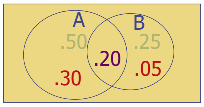
$A \cap B$ ← Joint Event

$P(A \cap B) = .20$

36

Probability of Joint Events

In a study, an individual is randomly selected from a population, and the event A and B are defined as the followings:



A = the individual is a smoker.
B = the individual has lung cancer.

P(A) = .50, P(B) = .25
P(A and B) = P(A $\cap$ B)
= P(smoker and has lung cancer) = .20

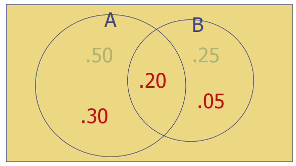
What is the probability P(A or B) = P(A $\cup$ B) = ?

37

Addition Rule 2
(General Addition Rule)

What is the probability that this randomly selected person is a smoker or has lung cancer? P(A or B) = ?

P(A or B) = P(A) + P(B) - P(A and B)
= .50 + .25 - .20
= **.55**

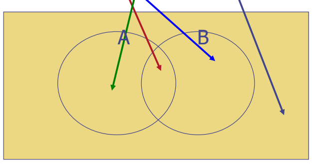


38

Contingency Table

	Cancer, B	No Cancer, B^c	Total
Smoke, A	20	30	50
Not Smoke, A^c	5	45	50
	25	75	100

Venn Diagram

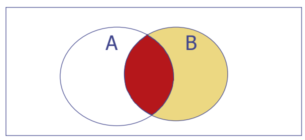


39

Conditional Probability

The conditional probability of event A to occur given event B has occurred (or given the condition B) is denoted as $P(A|B)$ and is, if $P(B)$ is not zero, $n(E)$ = # of equally likely outcomes in E,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{or} \quad P(A|B) = \frac{n(A \cap B)}{n(B)}$$



40

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

	Cancer C	No Cancer C′	Total
Smoke S	20	30	50
Not Smoke S′	5	45	50
	25	75	100

P(C|S) = 20/50 = .4
P(C|S′) = 5/50 = .1

41

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

	Cancer C	No Cancer C′	Total
Smoke S	20 (.2)	30 (.3)	50 P(S) =(.5)
Not Smoke S′	5 (.05)	45 (.45)	50 P(S′) =(.5)
	25 P(C) =(.25)	75 P(C)=(.75)	100 (1.0)

P(C|S) = .2/.5 = .4
P(C|S′) = .05/.5 = .1

What is $\frac{P(C|S)}{P(C|S')}$ = **4**
(*Relative Risk*)

42

Independent Events

Events A and B are independent if

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

or

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

or

$P(A \text{ and } B) = P(A) \cdot P(B)$

43

Multiplication Rule 1
(Special Multiplication Rule)

If event A and B are independent,

$P(A \text{ and } B) = P(A) \cdot P(B)$

44

Example

If a balanced die is rolled twice, what is the probability of having two 6's?

6_1 = the event of getting a 6 on the 1st trial

6_2 = the event of getting a 6 on the 2nd trial

$P(6_1) = 1/6$,

$P(6_2) = 1/6$, 6_1 and 6_2 are independent events

$P(6_1 \text{ and } 6_2) = P(6_1)\,P(6_2) = (1/6)(1/6) = 1/36$

45

Independent Events

"10%" of the people in a large population has disease H. If a random sample of two subjects was selected from this population, what is the probability that both subjects have disease H?

H_i : Event that the i-th randomly selected subject has disease H.

$P(H_2|H_1) = P(H_2)$ [Events are almost independent]

$P(H_1 \cap H_2) =$? $P(H_1)\,P(H_2) = .1 \times .1 = .01$

46

Independent Events

If events $A_1, A_2, \ldots, A_k$ are independent, then

$P(A_1 \text{ and } A_2 \text{ and } \ldots \text{ and } A_k)$
$= P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_k)$

What is the probability of getting all heads in tossing a balanced coin four times experiment?

$P(H_1) \cdot P(H_2) \cdot P(H_3) \cdot P(H_4) =$ **$(.5)^4 = .0625$**

47

Independent Events

"10%" of the people in a large population has disease H. If a random sample of **3** subjects was selected from this population, what is the probability that all subjects have disease H?

H_i : Event that the i-th randomly selected subject has disease H.

[Events are almost independent]

$P(H_1 \cap H_2 \cap H_3) = P(H_1)\,P(H_2)\,P(H_3)$
$= .1 \times .1 \times .1 =$ **.001**

48

Multiplication Rule 2
(General Multiplication Rule)

For any two events A and B,

P(A and B) = P(A|B) P(B) = P(B|A) P(A)

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

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

49

Multiplication Rule 2

If in the population, **50% of the people smoked**, and **40% of the smokers have lung cancer**, what percentage of the population that are smoker and have lung cancer?

P(S) = 50% of the subjects smoked
P(C|S) = 40% of the smokers have cancer
P(C and S) = P(C|S) P(S) = .4 x .5 = .2

50

Multiplication Rule 2

Box 1 contains 2 red balls and 1 blue ball

Box 2 contains 1 red ball and 3 blue balls

A coin is tossed. If "Head" turns up a ball is drawn from Box 1, and if "Tail" turns up then a ball is drawn from Box 2. Find the probability of selecting a red ball.

P(H and Red) + P(T and Red)
= P(H)P(Red|H) + P(T)P(Red|T)
= (1/2) x (2/3) + (1/2) x (1/4)
= 1/3 + 1/8 = **11/24**

51

Counting Rules

Fundamental Principle of Counting

Counting Rule: (use of the notation k_n)
In a sequence of n events in which the first one has k_1 possibilities and the second event has k_2 and the third has k_3, and so forth, the total possibilities of the sequence will be $k_1 \times k_2 \times k_3 \times \ldots \times k_n$

2 x 4 = 8 possible outcomes (coin & wheel)
6 x 6 = 36 possible outcomes (2 dice)
$k_1 \times k_2$

52

Example:

There are three shirts of different colors, two jackets of different styles and five pairs of pants in the closet. How many ways can you dress yourself with one shirt, one jacket and one pair of pants selected from the closet?

Sol: 3 x 2 x 5 = **30** ways
$k_1 \times k_2 \times k_3$

53

Permutations

B A S E

How many different four-letter code words can be formed by using the four letters in the word "BASE" without repeating use of the same letter?

$k_1 = 4$ $k_2 = 3$ $k_3 = 2$ $k_4 = 1$

4·3·2·1 = **24**

54

Factorial Formula

For any counting number n

n! = 1 x 2 x … x (n − 1) x n
0! = 1

Example: 3! = 1 x 2 x 3 = **6**

B A S E

How many different four-letter code words can be formed by using the four letters in the word "BASE" and the same letter can be used repeatedly?

$k_1 = 4$ $k_2 = 4$ $k_3 = 4$ $k_4 = 4$

$4 \cdot 4 \cdot 4 \cdot 4 =$ **256**

B A S E

How many different four-letter code words can be formed by using the four letters selected from letters A through Z and the same letter can **not** be used repeatedly?

$k_1 = 26$ $k_2 = 25$ $k_3 = 24$ $k_4 = 23$

$26 \cdot 25 \cdot 24 \cdot 23 =$ **358,800**

B A S E

How many different four-letter code words can be formed by using the four letters selected from letters A through Z and the same letter can be used repeatedly?

$k_1 = 26$ $k_2 = 26$ $k_3 = 26$ $k_4 = 26$

$26 \cdot 26 \cdot 26 \cdot 26 =$ **456,976**

Permutation Rule

Permutation Rule:

The number of possible permutations of r objects from a collection of n distinct objects is

$${}_nP_r = \frac{n!}{(n-r)!}$$

Order does count!

Permutations

$${}_nP_r = \frac{n!}{(n-r)!}$$

How many ways can a four-digit code be formed by selecting 4 distinct digits from nine digits, 1 through 9, without repeating use of the same digit?

$${}_9P_4 = \frac{9!}{(9-4)!} = \frac{9!}{5!} = \frac{9 \cdot 8 \cdot 7 \cdot 6 \cdot \cancel{5!}}{\cancel{5!}}$$

$= 9 \cdot 8 \cdot 7 \cdot 6 =$ **3024**

Combination Rule

Combination Rule:

The number of possible combinations of r objects from a collection of n distinct objects is

$$_nC_r = \frac{n!}{r!\cdot(n-r)!}$$

Order does not count!

61

Combinations

$$_nC_r = \frac{n!}{r!\cdot(n-r)!}$$

How many ways can a committee be formed by selecting 3 people from a group 10 candidates?

n = 10, r = 3

$$_{10}C_3 = \frac{10!}{3!\cdot(10-3)!} = \frac{10!}{3!\cdot 7!} = \frac{10\cdot 9\cdot 8\cdot \cancel{7!}}{3!\cdot \cancel{7!}}$$

$$= \frac{10\cdot 9\cdot 8}{3!} = \mathbf{120}$$

ABC, ACB, BAC, BCA, CAB, CBA are the same combination.

62

Combinations

$$_nC_r = \frac{n!}{r!\cdot(n-r)!}$$

How many ways can a combination of 4 distinct digits be selected from nine digits, 1 through 9?

$$_9C_4 = \frac{9!}{4!\cdot(9-4)!} = \frac{9!}{4!\cdot 5!} = \frac{9\cdot 8\cdot 7\cdot 6\cdot \cancel{5!}}{4!\cdot \cancel{5!}}$$

$$= \frac{9\cdot 8\cdot 7\cdot 6}{4\cdot 3\cdot 2\cdot 1}$$

$= 126$

63

Probability Using Counting (Theoretical Approach)

$$_nC_r = \frac{n!}{r!\cdot(n-r)!}$$

In a lottery game, one selects 6 distinct numbers from a set of 47 distinct numbers, what is the probability of winning the jackpot?

n = 47, r = 6

$$= \frac{47!}{6!\cdot 41!} = \frac{47\cdot 46\cdot 45\cdot 44\cdot 43\cdot 42\cdot 41!}{6!\cdot 41!}$$

$$= \frac{47\cdot 46\cdot 45\cdot 44\cdot 43\cdot 42}{6!} = \mathbf{10{,}737{,}573}$$

$$P(\text{Jackpot}) = \frac{1}{10{,}737{,}573}$$ (Theoretical Prob.)

64

Birthday Problem

In a group of randomly select 23 people, what is the probability that at least two people have the same birth date? (Assume there are 365 days in a year.)

P(at least two people have the same birth date)

Too hard !!!

= 1 – P(everybody has different birth date)

$= 1 - [365 \times 364 \times \ldots \times (365-23+1)] / 365^{23}$

65

Binomial Probability

What is the probability of getting two 6's in casting a balanced die 5 times experiment?

$P(S\cap S\cap S'\cap S'\cap S') = (1/6)^2 \times (5/6)^3 = \mathbf{0.016}$

$P(S\cap S'\cap S\cap S'\cap S') = (1/6)^2 \times (5/6)^3$

$P(S\cap S'\cap S'\cap S\cap S') = (1/6)^2 \times (5/6)^3$

...

How many of them? $\binom{5}{2} = \frac{5!}{2!\cdot 3!} = 10$

Probability (two 6's) = 0.016 x 10 = **0.16**

66