Defining probability

Random processes

- A *random process* is a situation in which we know what outcomes could happen, but we don't know which particular outcome will happen.

Random processes

- A *random process* is a situation in which we know what outcomes could happen, but we don't know which particular outcome will happen.
- A *random process* is a situation in which we know what outcomes could have happened, but we don't know which particular outcome happened.

Random processes

- A *random process* is a situation in which we know what outcomes could happen, but we don't know which particular outcome will happen.
- A *random process* is a situation in which we know what outcomes could have happened, but we don't know which particular outcome happened.
- Examples: coin tosses, die rolls, iTunes shuffle, whether the stock market goes up or down tomorrow, etc.
- It can be helpful to model a process as random even if it is not truly random. For example, technically a coin follows Newtonian mechanics, which are deterministic...

Some history

- Race board games (backgammon, senet) have been around for at least 5000 years.
- Al-Kindi, around 850 AD, developed frequency analysis for decrypting messages. (This work stemmed from deep numerical study of the Quran.)
- Cardano, around 1560, determined that when 3 dice are rolled, they are more likely to sum to 10 than to 9.

Senet board from ancient Egypt, about 1400 BC

Historical example: Problem of points

The problem of points was a math problem discussed at least as early as 1497, but went unsolved until 1654, when Pascal and Fermat discussed it extensively through written correspondence.

- Anne and Bert have decided that $100 will go to the first to win 3 rounds, where each round is a fair coin flip.
- Anne has won 2 rounds and Bert has won 1 round.
- The game is interrupted.
- How should the gamblers fairly split the $100?

Historical example: Problem of points

The problem of points was a math problem discussed at least as early as 1497, but went unsolved until 1654, when Pascal and Fermat discussed it extensively through written correspondence.

- Anne and Bert have decided that \$100 will go to the first to win 3 rounds, where each round is a fair coin flip.
- Anne has won 2 rounds and Bert has won 1 round.
- The game is interrupted.
- How should the gamblers fairly split the \$100?

The full problem generalizes this example to any number of total rounds and rounds won before interruption. Notice it took a long time for humans to find the solution. Quantitative probability is not in our genes.

Probability

- *Classical definition of probability:*
  - “The probability of an event is the ratio of the number of cases favorable to it, to the number of all cases possible when nothing leads us to expect that any one of these cases should occur more than any other, which renders them, for us, equally possible.”
    - Laplace 1812

Probability

- There are several possible interpretations of probability but they (almost) completely agree on the mathematical rules probability must follow.
  - $P(A)$ = Probability of event A
  - $0 \leq P(A) \leq 1$

Probability

- There are several possible interpretations of probability but they (almost) completely agree on the mathematical rules probability must follow.
    - $P(A)$ = Probability of event A
    - $0 \leq P(A) \leq 1$
- *Frequentist interpretation:*
    - The probability of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times.

Probability

- There are several possible interpretations of probability but they (almost) completely agree on the mathematical rules probability must follow.
  - $P(A)$ = Probability of event A
  - $0 \leq P(A) \leq 1$
- *Frequentist interpretation:*
  - The probability of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times.
- *Bayesian interpretation:*
  - A Bayesian interprets probability as a subjective degree of belief: for the same event, two separate people could have different viewpoints and so assign different probabilities.
  - Largely popularized by revolutionary advance in computational technology and methods during the last twenty years.

Intuition?

Which of the following events would you be most surprised by?

(a) 3 heads in 10 coin flips

(b) 30 heads in 100 coin flips

(c) 300 heads in 1000 coin flips

Intuition?

Which of the following events would you be most surprised by?

(a) 3 heads in 10 coin flips

(b) 30 heads in 100 coin flips

(c) *300 heads in 1000 coin flips*

Law of large numbers

Law of large numbers states that as more observations are collected, the proportion of occurrences with a particular outcome, $\hat{p}_n$, converges to the probability of that outcome, p.

$\hat{p}$ is the **sample** proportion.

$\hat{p}_n$ is the sample proportion when sample size is n.

p, in this case, is the probability (proportion of infinite **population**).

Law of large numbers (cont.)

When tossing a *fair* coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

Law of large numbers (cont.)

When tossing a *fair* coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

Law of large numbers (cont.)

When tossing a *fair* coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

- The coin is not "due" for a tail.

Law of large numbers (cont.)

When tossing a *fair* coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

$$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$$

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$$

- The coin is not "due" for a tail.
- The common misunderstanding of the LLN is that random processes are supposed to compensate for whatever happened in the past; this is just not true and is also called *gambler's fallacy* (or *law of averages*).

Disjoint and non-disjoint outcomes

Disjoint (mutually exclusive) outcomes: Cannot happen at the same time.

- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Disjoint and non-disjoint outcomes

Disjoint (mutually exclusive) outcomes: Cannot happen at the same time.

- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Non-disjoint outcomes: Can happen at the same time.

- A student can get an A in Stats and A in Econ in the same semester.

Union of non-disjoint events

What is the probability of drawing a jack or a red card from a well shuffled full deck?

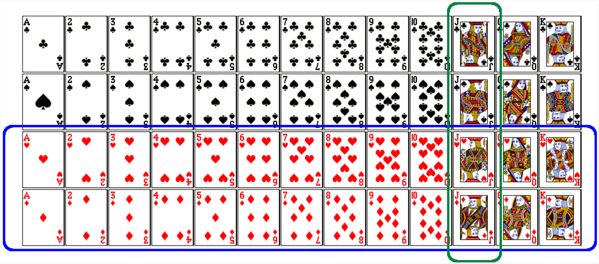
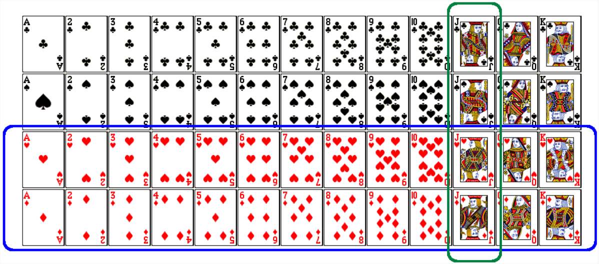


Figure from *http://www.milefoot.com/math/discrete/counting/cardfreq.htm*.

Union of non-disjoint events

What is the probability of drawing a jack or a red card from a well shuffled full deck?



$$P(jack\ or\ red) = P(jack) + P(red) - P(jack\ and\ red)$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

Figure from *http://www.milefoot.com/math/discrete/counting/cardfreq.htm*.

Practice

What is the probability that a randomly sampled student thinks marijuana should be legalized <u>or</u> they agree with their parents' political views?

	Share Parents' Politics		
Legalize MJ	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(a) $\frac{40+36-78}{165}$

(b) $\frac{114+118-78}{165}$

(c) $\frac{78}{165}$

(d) $\frac{78}{188}$

(e) $\frac{11}{47}$

Practice

What is the probability that a randomly sampled student thinks marijuana should be legalized <u>or</u> they agree with their parents' political views?

	Share Parents' Politics		
Legalize MJ	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(a) $\frac{40+36-78}{165}$

(b) $\frac{114+118-78}{165}$

(c) $\frac{78}{165}$

(d) $\frac{78}{188}$

(e) $\frac{11}{47}$

Recap

General addition rule

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

Note: *For disjoint events* $P(A \; and \; B) = 0$*, so the above formula simplifies to* $P(A \; or \; B) = P(A) + P(B)$.

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

- Rules for probability distributions:
  1. The events listed must be disjoint
  2. Each probability must be between 0 and 1
  3. The probabilities must total 1

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

- Rules for probability distributions:
  1. The events listed must be disjoint
  2. Each probability must be between 0 and 1
  3. The probabilities must total 1
- The probability distribution for the genders of two kids:

Event	*MM*	*FF*	*MF*	*FM*
Probability	*0.25*	*0.25*	*0.25*	*0.25*

Practice

In a survey, 52% of respondents said they are Democrats. What is the probability that a randomly selected respondent from this sample is a Republican?

(a) 0.48

(b) more than 0.48

(c) less than 0.48

(d) cannot calculate using only the information given

Practice

In a survey, 52% of respondents said they are Democrats. What is the probability that a randomly selected respondent from this sample is a Republican?

(a) 0.48
(b) more than 0.48
(c) less than 0.48
(d) *cannot calculate using only the information given*

If the only two political parties are Republican and Democrat, then (a) is possible. However it is also possible that some people do not affiliate with a political party or affiliate with a party other than these two. Then (c) is also possible. However (b) is definitely not possible since it would result in the total probability for the sample space being above 1.

Sample space and complements

Sample space is the collection of all possible outcomes of a trial.

- A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$
- A couple has two kids, what is the sample space for the gender of these kids?

Sample space and complements

Sample space is the collection of all possible outcomes of a trial.

- A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$
- A couple has two kids, what is the sample space for the gender of these kids? $S = \{MM, FF, FM, MF\}$

Sample space and complements

Sample space is the collection of all possible outcomes of a trial.

- A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$
- A couple has two kids, what is the sample space for the gender of these kids? $S = \{MM, FF, FM, MF\}$

Complementary events are two mutually exclusive events whose probabilities that add up to 1.

- A couple has one kid. If we know that the kid is not a boy, what is gender of this kid? { ~~M~~, *F* } → Boy and girl are *complementary* outcomes.
- A couple has two kids, if we know that they are not both girls, what are the possible gender combinations for these kids?

Sample space and complements

Sample space is the collection of all possible outcomes of a trial.

- A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$
- A couple has two kids, what is the sample space for the gender of these kids? $S = \{MM, FF, FM, MF\}$

Complementary events are two mutually exclusive events whose probabilities that add up to 1.

- A couple has one kid. If we know that the kid is not a boy, what is gender of this kid? { ~~M~~, *F* } → Boy and girl are *complementary* outcomes.
- A couple has two kids, if we know that they are not both girls, what are the possible gender combinations for these kids? { *MM*, ~~*FF*~~, *FM*, *MF* }

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

- Knowing that the coin landed on a head on the first toss does not provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.

Independence

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

- Knowing that the coin landed on a head on the first toss does not provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.
- Knowing that the first card drawn from a deck is an ace does provide useful information for determining the probability of drawing an ace in the second draw. → Outcomes of two draws from a deck of cards (without replacement) are dependent.

Practice

Between January 9-12, 2013, SurveyUSA interviewed a random sample of 500 NC residents asking them whether they think widespread gun ownership protects law abiding citizens from crime, or makes society more dangerous. 58% of all respondents said it protects citizens. 67% of White respondents, 28% of Black respondents, and 64% of Hispanic respondents shared this view. Which of the below is true?

Opinion on gun ownership and race ethnicity are most likely

(a) complementary
(b) mutually exclusive
(c) independent
(d) dependent
(e) disjoint

http://www.surveyusa.com/client/PollReport.aspx?g=a5f460ef-bba9-484b-8579-1101ea26421b

Practice

Between January 9-12, 2013, SurveyUSA interviewed a random sample of 500 NC residents asking them whether they think widespread gun ownership protects law abiding citizens from crime, or makes society more dangerous. 58% of all respondents said it protects citizens. 67% of White respondents, 28% of Black respondents, and 64% of Hispanic respondents shared this view. Which of the below is true?

Opinion on gun ownership and race ethnicity are most likely

(a) complementary
(b) mutually exclusive
(c) independent
(d) *dependent*
(e) disjoint

http://www.surveyusa.com/client/PollReport.aspx?g=a5f460ef-bba9-484b-8579-1101ea26421b

Checking for independence

If P(A occurs, given that B is true) = $P(A \mid B) = P(A)$, then A and B are independent.

Checking for independence

If P(A occurs, given that B is true) = $P(A \mid B) = P(A)$, then A and B are independent.

P(protects citizens) = 0.58

Checking for independence

If P(A occurs, given that B is true) = $P(A \mid B) = P(A)$, then A and B are independent.

P(protects citizens) = 0.58

P(randomly selected NC resident says gun ownership protects citizens, given that the resident is white) =
P(protects citizens | White) = 0.67

P(protects citizens | Black) = 0.28

P(protects citizens | Hispanic) = 0.64

Checking for independence

If P(A occurs, given that B is true) = $P(A \mid B) = P(A)$, then A and B are independent.

P(protects citizens) = 0.58

P(randomly selected NC resident says gun ownership protects citizens, given that the resident is white) =
P(protects citizens | White) = 0.67

P(protects citizens | Black) = 0.28

P(protects citizens | Hispanic) = 0.64

P(protects citizens) varies by race/ethnicity, therefore opinion on gun ownership and race ethnicity are most likely dependent.

Determining dependence based on sample data

- If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.
- If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.
- If a sample is large, then even a small difference can provide strong evidence of a real difference.

Determining dependence based on sample data

- If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.
- If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.
- If a sample is large, then even a small difference can provide strong evidence of a real difference.

We saw that P(protects citizens | White) = 0.67 and P(protects citizens | Hispanic) = 0.64. Under which condition would you be more convinced of a real difference between the proportions of Whites and Hispanics who think gun widespread gun ownership protects citizens? $n = 500$ or $n = 50,000$

Determining dependence based on sample data

- If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.
- If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.
- If a sample is large, then even a small difference can provide strong evidence of a real difference.

We saw that P(protects citizens | White) = 0.67 and P(protects citizens | Hispanic) = 0.64. Under which condition would you be more convinced of a real difference between the proportions of Whites and Hispanics who think gun widespread gun ownership protects citizens? $n = 500$ or $n = 50,000$

$n = 50,000$

Product rule for independent events

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

Or more generally, $P(A_1 \text{ and } \cdots \text{ and } A_k) = P(A_1) \times \cdots \times P(A_k)$

Product rule for independent events

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

Or more generally, $P(A_1 \text{ } and \text{ } \cdots \text{ } and \text{ } A_k) = P(A_1) \times \cdots \times P(A_k)$

You toss a coin twice, what is the probability of getting two tails in a row?

Product rule for independent events

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

Or more generally, $P(A_1 \text{ and } \cdots \text{ and } A_k) = P(A_1) \times \cdots \times P(A_k)$

You toss a coin twice, what is the probability of getting two tails in a row?

$$P(\text{T on the first toss}) \times P(\text{T on the second toss}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Practice

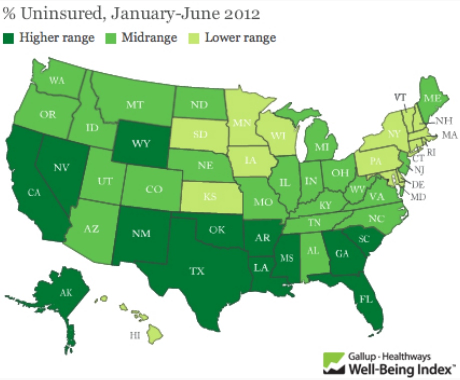
A recent Gallup poll suggests that 25.5% of Texans do not have health insurance as of June 2012. Assuming that the uninsured rate stayed constant, what is the probability that two randomly selected Texans are both uninsured?

(a) 25.5^2

(b) 0.255^2

(c) 0.255×2

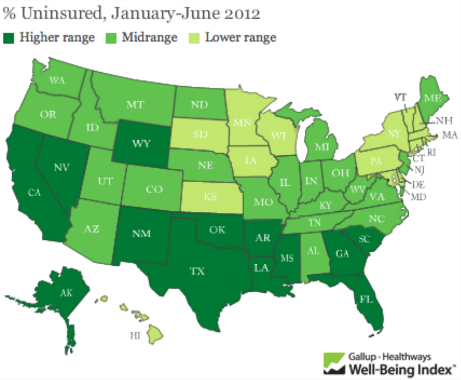
(d) $(1 - 0.255)^2$



http://www.gallup.com/poll/156851/uninsured-rate-stable-across-states-far-2012.aspx

Practice

A recent Gallup poll suggests that 25.5% of Texans do not have health insurance as of June 2012. Assuming that the uninsured rate stayed constant, what is the probability that two randomly selected Texans are both uninsured?



(a) 25.5^2

(b) 0.255^2

(c) 0.255×2

(d) $(1 - 0.255)^2$

http://www.gallup.com/poll/156851/uninsured-rate-stable-across-states-far-2012.aspx

Disjoint vs. complementary

Do the sum of probabilities of two disjoint events always add up to 1?

Disjoint vs. complementary

Do the sum of probabilities of two disjoint events always add up to 1?

Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.

Disjoint vs. complementary

Do the sum of probabilities of two disjoint events always add up to 1?

Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.

Do the sum of probabilities of two complementary events always add up to 1?

Disjoint vs. complementary

Do the sum of probabilities of two disjoint events always add up to 1?

Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.

Do the sum of probabilities of two complementary events always add up to 1?

Yes, that's the definition of complementary, e.g. heads and tails.

Putting everything together...

If we were to randomly select 5 Texans, what is the probability that at least one is uninsured?

- If we were to randomly select 5 Texans, the sample space for the number of Texans who are uninsured would be:

$$S = \{0, 1, 2, 3, 4, 5\}$$

- We are interested in instances where at least one person is uninsured:

$$S = \{0, 1, 2, 3, 4, 5\}$$

- So we can divide up the sample space into two categories:

$$S = \{0, \textit{at least one}\}$$

Putting everything together...

Since the probability of the sample space must add up to 1:

$$Prob(at\ least\ 1\ uninsured) \quad = \quad 1 - Prob(none\ uninsured)$$

Putting everything together...

Since the probability of the sample space must add up to 1:

$$
\begin{aligned}
Prob(at\ least\ 1\ uninsured) &= 1 - Prob(none\ uninsured) \\
&= 1 - [(1 - 0.255)^5]
\end{aligned}
$$

Putting everything together...

Since the probability of the sample space must add up to 1:

$$\begin{aligned} Prob(at\ least\ 1\ uninsured) &= 1 - Prob(none\ uninsured) \\ &= 1 - [(1 - 0.255)^5] \\ &= 1 - 0.745^5 \end{aligned}$$

Putting everything together...

Since the probability of the sample space must add up to 1:

$$\begin{aligned} Prob(at\ least\ 1\ uninsured) &= 1 - Prob(none\ uninsured) \\ &= 1 - [(1 - 0.255)^5] \\ &= 1 - 0.745^5 \\ &= 1 - 0.23 \end{aligned}$$

Putting everything together...

Since the probability of the sample space must add up to 1:

$$\begin{aligned} Prob(at\ least\ 1\ uninsured) &= 1 - Prob(none\ uninsured) \\ &= 1 - [(1 - 0.255)^5] \\ &= 1 - 0.745^5 \\ &= 1 - 0.23 \\ &= 0.77 \end{aligned}$$

At least 1

$$P(at\ least\ one) = 1 - P(none)$$

Practice

Roughly 20% of undergraduates at a university are vegetarian or vegan. What is the probability that, among a random sample of 3 undergraduates, at least one is vegetarian or vegan?

(a) $1 - 0.2 \times 3$

(b) $1 - 0.2^3$

(c) 0.8^3

(d) $1 - 0.8 \times 3$

(e) $1 - 0.8^3$

Practice

Roughly 20% of undergraduates at a university are vegetarian or vegan. What is the probability that, among a random sample of 3 undergraduates, at least one is vegetarian or vegan?

(a) $1 - 0.2 \times 3$

(b) $1 - 0.2^3$

(c) 0.8^3

(d) $1 - 0.8 \times 3$

(e) $1 - 0.8^3$

$$P(\text{at least 1 from veg}) = 1 - P(\text{none veg})$$
$$= 1 - (1 - 0.2)^3$$
$$= 1 - 0.8^3$$
$$= 1 - 0.512 = 0.488$$

Overview

Let A and B be two events.

$0 \leq P(A) \leq 1$

$0 \leq P(B) \leq 1$

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

A and B are mutually exclusive $\iff$ $P(A \text{ and B}) = 0$

A and B are independent $\iff$ $P(A \text{ and B}) = \text{P(A)} \times \text{P(B)}$

A and B are exhaustive $\iff$ $P(A \text{ or B}) = 1$

A and B are complements $\iff$ A and B are exhaustive and mutually exclusive

The notation A^c is used to denote the complement of A

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