William Shakespeare Probability

If you rearrange the letters in William Shakespeare you can spell the phrase

“here was I like a psalm”

William Shakespeare was 46 years old when the Kings James version of the Bible was completed. In the Kings James Bible in Psalm 46 the 46th word from the top is Shake and the 46th word from the bottom is spear. The two words that are 46th from the top and bottom are the words Shake Spear. There is a missing letter e at the end of spear but even with that issue it is still an neat discovery. Is it proof the author knew there would be a famous writer named Shakespeare or is it a coincidence?

Is this a coincidence? Yes.

Given enough words, enough last names and enough searching, you should expect to find some kind of neat connection between the words in the Bible and a last name of some person. There are about 780,000 words in bible. The number of possible combinations between these almost 8 hundred thousand words in the bible and the billions of possible peoples last names is an extremely large number. You cannot determine in advance if a specific connection will happen but there are guaranteed to be many interesting connections.

The fact that the relationship between the name Shakespeare and the 46th psalm exists just seems so unlikely that many people say it just cannot happen by chance alone. Many people have a misunderstanding of how rare it is for an an event to occur because they misinterpret the order of the events. If you create 1 million of these possible combinations between any name and the words in bible and find that 1 of these outcomes has an interesting meaning then that outcome be 1 in a million.

There is a small probability of that you can select one specific outcome in advance and then find it in a search of the words in the bible but the probability that SOME outcome with SOME meaning can be found in a search is not as small as you may think. If you just start with showing the one in a million outcome and you do not mention that you looked at a million possible outcomes to find it then it does seem a bit more surprising

Another interesting Bible Fact

Psalm 118 is the middle chapter of the entire Bible. Psalm 118 verse 8 is the middle verse of the entire Bible. Psalm 118:8 “It is better to take refuge in the Lord than to trust man”.

There are 594 chapters before and after Psalm 118. Add 594 + 594 =1188.

Before Psalm 118, Psalm 117 is the shortest chapter in the Bible. “ Jesus Wept”

After Psalm 118, Psalm 119 is the longest chapter of the Bible.

The Birthday Problem

A famous problem, called the **Birthday Problem,** involves finding the probability that 2 people in a group have the same birthday. There are 365 days in a year if we do not count leap years. A persons birthday can be any of the 365 dates. If there are 366 people in a group then at least 2 of them must have the same birthday. The first 365 people can all have a different birthday but the 366th person must cannot have a different birthday from all the others as all the days in the year are taken. So we can say that if more than 365 people are in a group than at least 2 people in the group will have the same birthday. If there are less that 365 people in a group you cannot be guaranteed that 2 people in the group will have the same birthday.

If there is 1 person in the room then there is no chance that at least 2 of them must have the same birthday. If we let N be the number of people in the room then if N > 365 we can say that P(at least 2 matching birthdays) is 1 and if N = 1 then we can say that P(at least 2 matching birthdays) is 0

The Birthday Problem asks what is the probability that in a group of any size less than 365 at least 2 of them must have the same birthday. It is clear that the more people are in the group the larger the probability that 2 people in a group will have the same birthday. If N = 364 then it is almost a sure thing and if N = 2 it is not very probable. What may seem hard to believe is that if N = 70 there is a 99.9% probability of 2 matching birthdays and If N = 23 there is a 50% probability. The calculations that produced there values are based on the assumption that each day of the year (except February 29) is equally probable for a birthday. Real-life birthday distributions are not equally probable.

NOTE: This problem is often misstated. It does not ask what is the probability that 2 people in a group have May 3rd as a common birthday or any other specific date. It only says that 2 have some common birthday. That allows for all of the 365 days to be the common day.

Let Event A represent the case where at least two people in the room have the same birthday. The complement of Event A is the probability that NO two people in the room have the same birthday. The complement of Event A is written as $\overline{A}$ and is read NOT A. A and $\overline{A}$ are the only two possible outcomes and only one of them can happen in a group. Either there are two people in the room have the same birthday or there are NOT two people in the room with the same birthday so $P(A)+P(\overline{A})=1$ and P(A) = 1 – $P(\overline{A})$. To compute $P(A)$ the probability that at least two people in the room have the same birthday we will compute 1 – $P(\overline{A})$ which is 1 minus the probability that no two people in the room have the same birthday.

Calculating the probability that in a group of size N at least 2 of them must have the same birthday. Lets start by calling the 1st persons birthday Event 1, the 2nd persons birthday Event 2, the 3rd persons birthday Event 3, and so on.

If N = 6 then

Event 1 is the 1st person's birthday is any day of the year. The probability of that is $\frac{365}{365}$

Event 2 is the 2nd persons birthday is not the same as the 1st person's birthday. There are 355 –1 days that do not match the birthdays already used. year. The probability of that is $\frac{364}{365}$

Event 3 is the 3rd persons birthday is not the same as the first 2 birthdays. There are 355 –2 days that do not match the birthdays already used. year. The probability of that is $\frac{363}{365}$

Event 4 is the 4rd persons birthday is not the same as the first 3 birthdays. There are 355 –3 days that do not match the birthdays already used. year. The probability of that is $\frac{362}{365}$

Event 5 is the 5th persons birthday is not the same as the first 5 birthdays. There are 355 –4 days that do not match the birthdays already used. year. The probability of that is $\frac{361}{365}$

Event 6 is the 6th persons birthday is not the same as the first 5 birthdays. There are 355 –5 days that do not match the birthdays already used. year. The probability of that is $\frac{360}{365}$

Event n is the nth persons birthday is not the same as the first n–1 birthdays. There are 355 – (n–1) days that do not match the birthdays already used. year. The probability of that is $\frac{365-(n-1)}{365}$

If N = 6 then there are 6 events and $P(\bar{A})$ will have 6 products

$$P(\bar{A})=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet\frac{360}{365} \qquad P(\bar{A})=\left(\frac{1}{365}\right)^6\bullet 365\bullet 364\bullet 363\bullet 362\bullet 361\bullet 360\simeq .96$$

1 – $1-P(\bar{A})\simeq 1-.96\simeq .04$

If N = 6 then there is a 4% chance that at least 2 people out of the 6 will have the same birthday.

If N = 10 then

$$P(\bar{A})=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet\frac{360}{365}\bullet\frac{359}{365}\bullet\frac{358}{365}\bullet\frac{357}{365}\bullet\frac{356}{365}$$

$$P(\bar{A})=\left(\frac{1}{365}\right)^{10}\bullet 365\bullet 364\bullet 363\bullet 362\bullet 361\bullet 360\bullet 359\bullet 358\bullet 357\bullet 356\simeq .88$$

$$1-P(\bar{A})\simeq 1-.88\simeq .12$$

If N =10 then there is a 12% chance that at least 2 people out of 6 people will have the same birthday.

If N = 23 then

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet..........\bullet\frac{365-(23-1)}{365}$$

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet..........\bullet\frac{343}{365}$$

$$P\left(\bar{A}\right)=\left(\frac{1}{365}\right)^{23}\bullet 365\bullet 364\bullet 363\bullet........\bullet 343\simeq .49$$

$$1-P\left(\bar{A}\right)\simeq 1-.49\simeq .51$$

If N = 23 then there is a 51% chance that at least 2 people out of 23 people will have the same birthday.

If N = 40 then

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet..........\bullet\frac{365-(40-1)}{365}$$

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet\frac{362}{365}\bullet\frac{361}{365}\bullet..........\bullet\frac{326}{365}$$

$$P\left(\bar{A}\right)=\left(\frac{1}{365}\right)^{60}\bullet 365\bullet 364\bullet 363\bullet........\bullet 326\simeq .11$$

$$1-P\left(\bar{A}\right)\simeq 1-.11\simeq .89$$

If N = 40 then there is a 89% chance that at least 2 people out of 60 people will have the same birthday.

If N = 70 then

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet..........\bullet\frac{365-(70-1)}{365}$$

$$P\left(\bar{A}\right)=\frac{365}{365}\bullet\frac{364}{365}\bullet\frac{363}{365}\bullet..........\bullet\frac{296}{365}$$

$$P\left(\bar{A}\right)=\left(\frac{1}{365}\right)^{70}\bullet 365\bullet 364\bullet 363\bullet........\bullet 296\simeq .001$$

$$1-P\left(\bar{A}\right)\simeq 1-\left(\frac{1}{365}\right)^{70}\bullet 365\bullet 364\bullet 363\bullet........\bullet 296\simeq .001$$

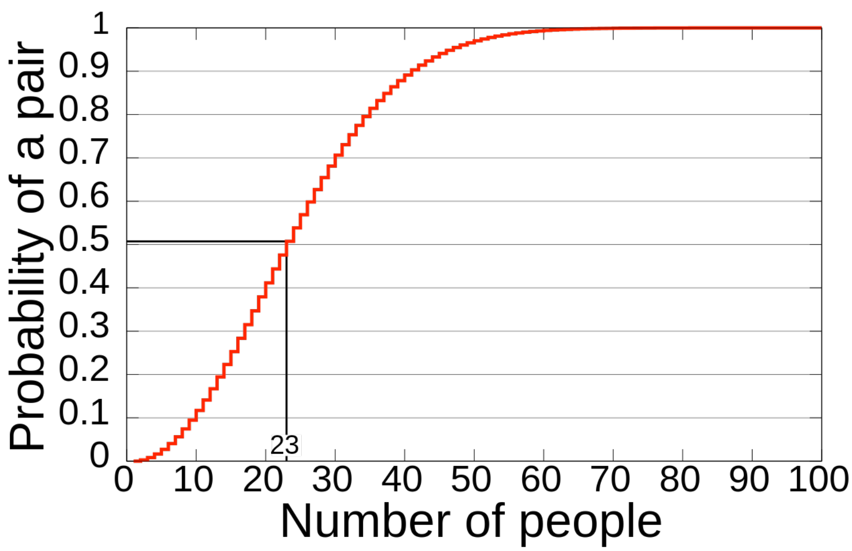
Given n people where n < 366 the probability that at least 2 people out of the n people will have the same birthday can be found as follows.

$$P(\bar{A}) = \frac{365}{365} \bullet \frac{364}{365} \bullet \frac{363}{365} \bullet .... \bullet \frac{365-(n-1)}{365}$$

$$P(\bar{A}) = \left(\frac{1}{365}\right)^n \bullet 365 \bullet 364 \bullet 363 \bullet .... \bullet (365-(n-1))$$

$$1-P(\bar{A}) \simeq 1-\left(\frac{1}{365}\right)^n \bullet 365 \bullet 364 \bullet 363 \bullet .... \bullet (365-(n-1))$$

The graph below shows the probabilities of a common birthday for different values of n. Note that the values for P(n) approach 1 at values of n close to 70



Approximations for P(N)

Many formulas have been developed to approximate the probability that at least 2 people out of the n people will have the same birthday. 2 of the most common are listed below.

$$1-P(\bar{A}) \simeq 1-\left(\frac{364}{365}\right)^{{}_nC_2} \quad \text{and} \quad 1-P(\bar{A}) \simeq 1-e^{-n(n-1)/2(365)}$$

Note: There is a Taylor Series Expansion that is developed in Calculus that also can be used to approximate P(A)

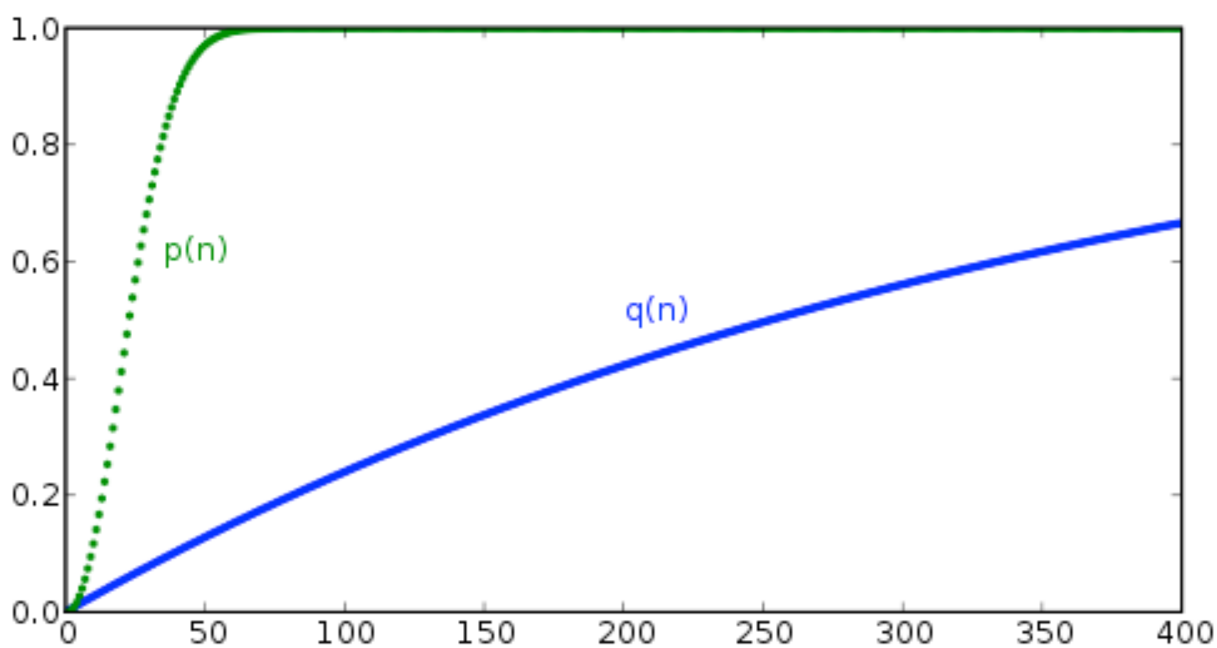
Same birthday as you problem

There is a difference between finding the probability that at least 2 people out of the n people will have the same birthday and the probability that someone in the group of size n has the same birthday as you. 2 people in the room can have the same birthday by matching birthdays on any of the 365 days in the year. For a person in the room to match **your birthday** they must have exactly 1 of the 365 day in the year. Just having 2 matches of any date allows for a match on any of the 365 days in the year, That is less probable to occur.

We can call the probability that someone in the group of size n have the same birthday p(n) and the probability that someone in the group of size n has the same birthday as you q(n). The difference in the formulas for each probability are shown below.

$$p(n) \simeq 1-\left(\frac{1}{365}\right)^{n} \bullet 365 \bullet 364 \bullet 363 \bullet \ldots\ldots\bullet\left(365-(n-1)\right) \quad \text{and} \quad q(n) \simeq 1-\left(\frac{365-1}{365}\right)^{n}$$

The graph below compares probabilities of p(n) and q(n) for different values of n. Note that the values for p(n) approach 1 at values of n close to 70 and q(n) the values for p(n) approach 1 at a much slower rate.



Card Guess

I am going to try and impress you with my clairvoyant ability. I will ask you to think of any card in a deck of 52 playing cards. I will then ask you to think of a second card in the deck that is **differen**t from the first card you thought of. Would you be impressed if I could name the 2 cards?

The probability of guessing Event A and then Event B is written P (A and B). If the occurrence of Event A does not affect the the probability of Event B happening than the events are **independent** and the formula for finding this probability is $\mathrm{P(A\ and\ B) = P(A) \bullet P(B)}$. If the occurrence of Event A does affect the the probability of Even B happening the events are **dependent** and the formula for finding this probability and the formula for finding this probability is $\mathrm{P(A\ and\ B) = P(A) \bullet P(B\ |A)}$ In the case of guessing 2 different cards from the deck events are **dependent**. When I ask you to think of any of the 52 cards in the deck I have a 1/ 52 probability of guessing your card correctly so P(A) = 1/ 52. I will be correct about 1.923% of the time.. I then ask you to think of a **different** second card.. There are only 51 cards that are different from the one you selected so P(B I A) = 1/ 51. . I have a 1/ 51 probability of guessing your second card correctly. I will be correct is about 1.961% of the time. The events are **dependent**. so $\mathrm{P(A\ and\ B) = P(A) \bullet P(B\ |A)} \cong \frac{1}{52} \bullet \frac{1}{51} \cong .000377$. I will be correct is about .0377% of the time. It would seem that I would have a low probability of impressing you with my clairvoyant ability with this approach but If I did this live on the web with 1 million viewers then about $\frac{1}{52} \bullet \frac{1}{51} \bullet 1{,}000{,}000 \simeq 377$ of the viewers will be shocked that I was able to predict their 2 different cards in order. What these 377 people do not see is that the other 999,623 people saw me fail.

How can I improve the probability that I am correct?

A study by Jay Olson 's titled "Revealing the Psychology of Playing Card Magic " found that the most common cards people named when asked to name a card were the Ace of Spades and the Queen of Hearts . When a person was asked to name a card he found that the 8 most common cards named and the probability that the card was named was as follows: AS – 24.5% , QH – 13.7 %, AH – 6.2%, KH – 5.9% , JS – 4.3% , AD – 3.8% , JH – 2.8%

This means that if I predict the first card named will be the Ace of Spades (AS) I will be correct 24.5 % of the time rather than the $\frac{1}{52}$ or 1.923% if all cards are equally likely to be chosen. If I predict the second card named will be the Queen of Hearts (QH) than I will be correct 13.7 % of the time rather than the $\frac{1}{51}$ or 1.961% if all cards are equally likely to be chosen. The probability of the Ace of Spades (AS) and then the Queen of Hearts (QH) being selected by a person is written as $\mathrm{P(AS\ and\ QH) = P(AS) \bullet P(AS\ |\ QH)} \cong .245 \bullet .137 \cong .0336$. I will be correct is about 03.36% of the time.

If I did this live on the web with 1 million viewers then about $.245 \bullet .137 \bullet 1{,}000{,}000 = 33{,}600$ of the viewers will be shocked that I was able to predict their 2 different cards in order. That is a vast improvement over the 377 people that saw me correctly guess their cards in the first case.

Can I improve the probability that I am correct even more?

The calculations for $\mathrm{P}(\mathrm{AS\ and\ QH})$ are based on the fact that you predict the first card correctly and then name the second card in the correct order. If you ask a person to name any card in a deck of 52 playing cards and then ask them to think of a second card in the deck that is **differen**t from the first card and then predict what the the 2 cards were without stating the order they were selected in you can improve the probability that you are correct.

The probability of guessing Event A **or** Event B is written P (A or B). If the occurrence of Event A does not affect the the probability of Even B happening than the events are **independent** and the formula for finding this probability is $\mathrm{P}(\mathrm{A\ or\ B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B})$. If the occurrence of Event A does affect the the probability of Event B happening the events are **dependent** and the formula for finding this probability and the formula for finding this probability is $\mathrm{P}(\mathrm{A\ or\ B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B}) - \mathrm{P}(\mathrm{A\ and\ B})$

You are correct if the order the cards were selected in was (AS and then QH) OR if the the order was (QH and then AS). If event A is (AS and then QH) and Event B is (QH and then AS) then we can see that $\mathrm{P}(\mathrm{A\ and\ B})$ would mane that the AS was named first and the QH was named second AND the QH was named first and the AS was named second. This cannot happen so $\mathrm{P}(\mathrm{A\ and\ B}) = 0$ and $\mathrm{P}(\ (\mathrm{AS\ or\ QH})\ \mathrm{or}(\mathrm{QH\ or\ AS})\ ) = \mathrm{P}(\mathrm{AS\ and\ QH}) + \mathrm{P}(\mathrm{QH\ and\ AS}) = .0336 + .336 = .0672$. If you predict that the 2 cards were the Ace of Spades and the Queen of Hearts without naming the order you will be correct about 6.72% of the time. That is a vast improvement over the .0377% value for just guessing 2 random cards.