DSC 40A

Theoretical Foundations of Data Science I

Today

- More examples of using combinatorics to solve probability questions.

Question

Answer at q.dsc40a.com

Remember, you can always ask questions at q.dsc40a.com!

If the direct link doesn't work, click the "Lecture Questions" link in the top right corner of dsc40a.com.

Counting as a Tool for Probability

Example 8. What is the probability that a randomly generated bitstring of length 10 is the string 0011001101?

$\left(\frac{1}{2}\right)^{10}$

_ _ _ _ _ _ _ _ _ _

↑ bit 0 or 1 → 2 options

assuming uniform prob $= \frac{1}{2}$

possible bit strings $= 2^{10}$

event = 1 string : 0011001101

prob $= \frac{1}{2^{10}} = \left(\frac{1}{2}\right)^{10}$

$$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting as a Tool for Probability

Example 7. What is the probability that a randomly generated bitstring of length 10 contains an equal number of zeros and ones?

_ 0 0 _ 0 _ 0 _ 0 _

5 0's 5 1's

one to one mapping between positions and the bit string

5 positions for 0's

{3, 2, 7, 9, 5}

sets

|E| = selecting 5 numbers (= position of 0's) out of 10 possible numbers

n = 10

k = 5

$$\text{prob} = \frac{C(10,5)}{2^{10}} = \frac{|E|}{|S|} = \frac{\#\text{ sets of 5 numbers out of 10}}{\#\text{ bit strings of length 10}}$$

$$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Why is the positions of 5 0's a set, not a sequence?

The following sets are all the same:

$\{1, 3, 5, 7, 10\} = \{5, 3, 1, 7, 10\} = \{10, 7, 5, 3, 1\} = \ldots$

The order in which the positions are specified doesn't matter, therefore it's a set and it's equivalent to:

bitstring: 0 1 0 1 0 1 0 1 1 0

Positions: 1 (0), 3 (0), 5 (0), 7 (0), 10 (0)

How to solve using permutations?

(Recall:

$$\#\text{sequences} = \#\text{sets} \cdot \#\text{orderings})$$

10 options, 9 options, 8 options, ..., 1 option

we have to place: 5 0's 5 1's in 10 places _ _ _ _ _ _ _ _ _ _

so we could calculate all possible permutations $= 10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

However since the objects we are permuting (bits) aren't unique note the following:

Let's start out with 0 0 0 0 0 1 1 1 1 1

if we swap first and last bits we get

1 0 0 0 0 1 1 1 1 0

which is a different string from the previous one we had

if we shuffle the first two bits we get

0 0 0 0 0 1 1 1 1 1 which is the same

So all permutations of the 0's amongst themselves are equivalent $= 5!$

and all permutations of the 1's amongst themselves are equivalent $= 5!$

Therefore # of unique bitstrings $= \frac{10!}{5!\,5!} = C(10,5)$

which is the same answer as using sets

Different version of this question

1) We have a string of length 10 composed of characters from a, b, c, d, e, f, g, h, i, j
What is the number of strings of length 10 such that all characters in the string are unique.

Ex: abcdefghij, jihgfedcba but not bbccadefji

Answer _ _ _ _ _ _ _ _ _ _ $\Rightarrow 10 \cdot 9 \cdot 8 \cdot \_ \ldots = 10!$

2) We have a string of length 10 composed of characters from a, b, c, d, e, f
What is the number of strings that have 5 a's and the rest of the characters are all different?

ex: aaaaabcdef   abacadaeaf

Solution 1: How many permutations? 10! (In first slot we select from 5 a's, 1 b, 1 c, 1 d, 1 e 1 f)

How many of these are unique strings?

There are 5! orderings of the 5 a's

a a a a a
a a a a a
a a a a a
a a a a a
⋮

} = a a a a a

Therefore: $\frac{10!}{5!} = P(10,5)$

Solution 2: If we select the positions for b, c, d, e, f then the remaining are all a's

To assign positions we need to select at random a <u>sequence</u> of 5 numbers from 1-10 without replacement

$(1, 2, 3, 4, 5) \Rightarrow$ b c d e f a a a a a
$(5, 4, 3, 2, 1) \Rightarrow$ f d e c b a a a a a
$(2, 4, 6, 8, 10) \Rightarrow$ a b a c a d a e a f

$\Rightarrow P(10,5)$

Counting as a Tool for Probability

Example 10. What is the probability that a fair coin flipped 10 times turns up HHTTHHTTHT?

$p(H) = p(T) = \frac{1}{2}$ (uniform dist.)

H H T T _ _ _ _ _ _

$\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ ...

10 times $= \left(\frac{1}{2}\right)^{10}$

prob $= \frac{1}{2^{10}}$

$$P(n,k) = (n)(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting as a Tool for Probability

Example 9. What is the probability that a fair coin flipped 10 times turns up an equal number of heads and tails?

_ _ _ _ _ _ _ _ _ _ with equal num of H / T

$C(n,k) = C(n, n-k) = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!\,k!}$

set of positions where tails are:

n = 10 possible positions

k = 5 positions to place 'tails' (then the rest are 'heads')

$\text{Prob} = \frac{|E|}{|S|} = \frac{C(10,5)}{2^{10}} = C(10,5)\cdot\left(\frac{1}{2}\right)^{10} = C(10,5)\cdot\left(\frac{1}{2}\right)^{5}\left(\frac{1}{2}\right)^{5}$

Example 2: 6H, 4T $\quad C(10,4) = C(10,6)$

$$P(n,k) = (n)(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting as a Tool for Probability

Example 12. What is the probability that a biased coin with $Prob(H) = \frac{1}{3}$ flipped 10 times turns up HHTTHHTTHT?

$Prob(T) = \frac{2}{3}$

$\frac{1}{\# \text{ total number of seq. of length } 10}$?

no longer uniform dist

H H T T H H T T H T =
$\frac{1}{3}$ $\frac{1}{3}$ $\frac{2}{3}$ $\frac{2}{3}$ $\frac{1}{3}$ $\frac{1}{3}$ $\frac{2}{3}$ $\frac{2}{3}$ $\frac{1}{3}$ $\frac{2}{3}$

$\left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^5$

$prob(HH\ldots H) = \left(\frac{1}{3}\right)^{10}$

$prob(TT\ldots T) = \left(\frac{2}{3}\right)^{10}$

$$P(n,k) = (n)(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting as a Tool for Probability

Example 11. What is the probability that a biased coin with $Prob(H) = \frac{1}{3}$ flipped 10 times turns up an equal number of heads and tails?

$Prob(T) = \frac{2}{3}$

S = coin toss of length 10

E = coin toss with 5 H 5 T

$prob(s \in E) = \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^5$

$$p(E) = \sum_{s \in E} p(s) = C(10,5)\left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^5 \neq \frac{\#\text{ outcomes in } E}{\#\text{ outcomes in } S}$$

not uniform probability

$p(s) \neq \frac{1}{|S|}$

$$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Example: 4H 6T

S = num of seq of length 10

E = coin toss seq with 4H 6T

$$|E| = C(10,6) = C(10,4) \qquad \text{prob}(s \in E) = \left(\tfrac{1}{3}\right)^4 \left(\tfrac{2}{3}\right)^6$$

$$\text{prob} = \sum_{s \in E} p(s) = |E|\, p(s \in E) = C(10,4) \left(\tfrac{1}{3}\right)^4 \left(\tfrac{2}{3}\right)^6$$

$$C(10,4) = C(10,6)$$

Example: 6H 4T

what changes

$$p(s \in E) = \left(\tfrac{1}{3}\right)^6 \left(\tfrac{2}{3}\right)^4$$

$$\text{prob} = C(10,4) \left(\tfrac{1}{3}\right)^6 \left(\tfrac{2}{3}\right)^4$$

The Easy Way

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Solved using sequences x2
sets x2

from Theory Meets Data by Ani Adhikari, Chapter 4

The Easy Way

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Another way to think of sampling without replacement:

1. randomly shuffle all 20 students
2. take the first 5

S = all possibles positions to place a student in a sequence of length 20

E = student 17 is in one of first 5 positions

$$\frac{|E|}{|S|} = \frac{5}{20} = \frac{1}{4}$$

Practice Problems

Example 13. You were one of 238 individuals who reported for jury duty. If 54 of these people will be assigned to a courtroom, what is the probability that you get assigned to a courtroom?

sampling without replacement

sample space = 238 positions

shuffle everyone and select the first 54 $\Rightarrow$ prob = $\frac{54}{238}$

Can also solve with sets / seq.

S is all sets of 54 out of 238

E is all sets of 54 out of 238 including you

$|S| = C(238, 54)$

Practice Problems

Example 13. You were one of 238 individuals who reported for jury duty. If 54 of these people will be assigned to a courtroom, what is the probability that you get assigned to a courtroom?

How many sets of 54 individuals include you?

A. C(238, 54)

B. C(237, 54)

C. C(238, 53)

D. C(237, 53)

remaining people after you were selected

$C(237, 53)$

spots left after your were chosen

$$P = \frac{C(237, 53)}{C(238, 54)}$$

Practice Problems

Example 14. You were one of 238 individuals who reported for jury duty. Suppose 28 of these individuals are doctors. If 54 of these people will be assigned to a courtroom, what is the probability that exactly 5 doctors get assigned to a courtroom?

S = sets of 54 out of 238

E = sets of 54 with 5 doctors

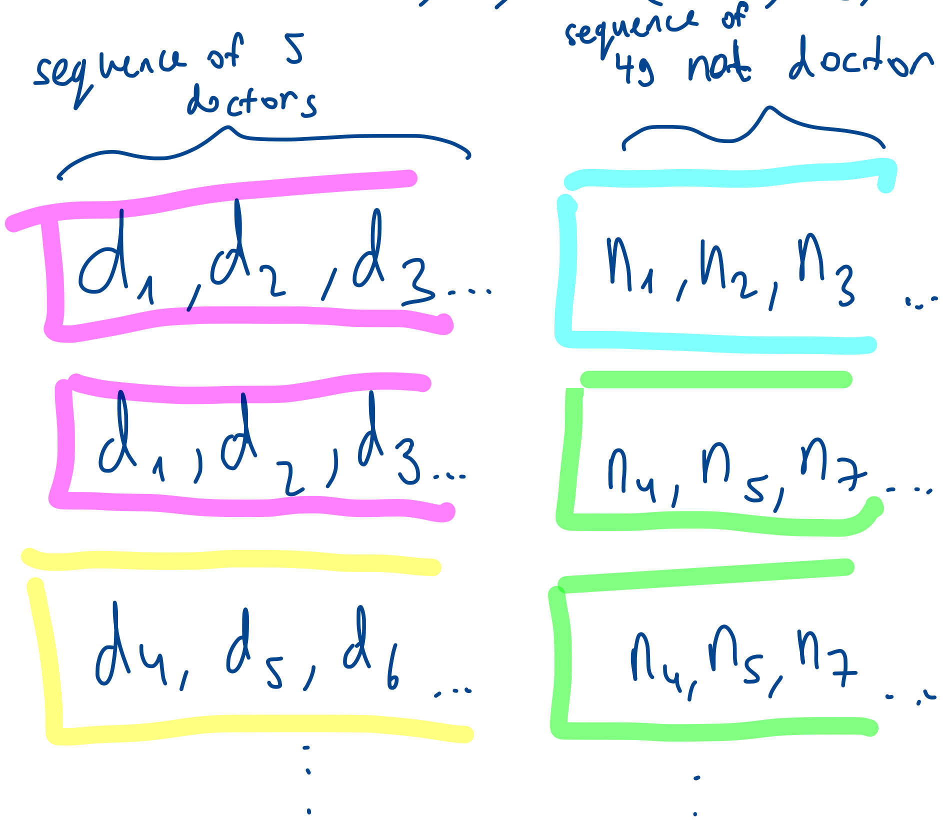
$$\text{Prob} = \frac{\#\text{ sets in } S \text{ with 5 doctors}}{\#\text{ sets in } S} = \frac{C(28,5)\,C(210,49)}{C(238,54)}$$

C(28,5): choosing 5 doctors

C(210,49): choosing remaining non-doctors to have 54 total

238 = 28 + 210, 54 = 5 + 49

Why product $C(238,5) \cdot C(210,49)$?



Practice Problems

Example 15. What is the probability that your five-card poker hand is a straight?

Practice Problems

Example 16. Suppose you look at your first card as it is dealt, and you see that it is a Queen. What is the probability that your five-card hand is a straight?

Summary

- Counting is a useful tool for probability.
- Sometimes there’s an easy way!
- **Next time:** Bayes Theorem