Ma 3 KC Border
Introduction to Probability and Statistics Winter 2015

Chapter 3: Learning to count

Relevant textbook passages:

Pitman [4]: Sections 1.5–1.6, pp. 47–77; Appendix 1, pp. 507–514.

Larsen–Marx [3]: Sections 2.6, 2.7, pp. 67–101.

3.1 The great coin-flipping experiment

There were 214 submissions of 128 flips, for a grand total of 27,392! You can soon find the data at http://www.math.caltech.edu/%7E2014-15/2term/ma003/Data/FlipsMaster.txt

Treatment	Sample size	Heads	Tails
Start Heads	17,536	50.8326%	49.1674%
Start Tails	9,856	48.8839%	51.1161%
Pooled	27,392	50.1314%	49.8686%

The data do indeed support the hypothesis that there is a bias toward coming up the same way as starting out. Recall that I put predictions into a sealed envelope. Here are the predictions of the average number of runs, by length, compared to the experimental results.

Run length	Theoretical average[1]	Predicted range[2]	Total runs	Average runs	How well did I do?
1	32.5	31.648769 –33.383511	6819	31.864486	Nailed it.
2	16.125	15.608301 –16.616951	3329	15.556075	Off by 0.04
3	8	7.657826 – 8.350408	1808	8.448598	Off by 0.09
4	3.96875	3.725101 – 4.214796	826	3.859813	Nailed it.
5	1.96875	1.787162 – 2.147520	423	1.976636	Nailed it.
6	0.976563	0.848447 – 1.101768	205	0.957944	Nailed it.
7	0.484375	0.389673 – 0.581934	103	0.481308	Nailed it.
8	0.240234	0.171202 – 0.307798	65	0.303738	Nailed it.
9	0.119141	0.072501 – 0.168266	24	0.112150	Nailed it.
10	0.059082	0.025071 – 0.090817	20	0.093458	Off by 0.0026
11	0.0292969	N/A	6	0.028037	
12	0.0145264	N/A	5	0.023364	
13	0.00720215	N/A	1	0.004673	

[a] The formula for the theoretical average is the object of the Bonus Question.
[b] This is based on a Monte Carlo simulation of the 95% confidence interval.

Yes! There are Laws of Chance.

Bonus Question: A bonus question carries no points, but the first two worthy answers entitle the submitters to be my guest for lunch at the Athenaeum. Limit one lunch per student per term.

Question: What is the average number of runs of length k in n tosses. (A run is a *maximal* consecutive string.) Of course you must prove you are correct.

Robert Ash [1], section 1.4 has a really good discussion of counting principles. Also see Pitman [4, Section 1.6] and Larsen–Marx [3, Section 2.6].

3.2 Generally accepted counting principles

3.2.1 Lists versus sets

If you've had CS 1, you've studied Python which has both **lists** and **sets**. Both are collections of n objects, but two lists are different unless the same object appears in the same *position* in both lists.

For instance,

123 and 213 are distinct lists of three elements, but the same set.

A list is sometimes referred to as a **permutation**.

3.2.2 Number of lists of length *n*

If I have n distinct objects, how many distinct ways can I arrange them into a list (without repetition)? Think of the objects being numbered and starting out in a bag and having to be distributed among n numbered boxes.

$$\underset{1}{\sqcup}\quad \underset{2}{\sqcup}\quad \underset{3}{\sqcup}\quad \underset{4}{\sqcup}\quad \underset{5}{\sqcup}\quad \cdots\quad \cdots\quad \cdots\quad \underset{n-1}{\sqcup}\quad \underset{n}{\sqcup}$$

There are n choices for box 1, and for each such choice, there are $n-1$ for position 2, etc., so all together

> there are $n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$ distinct lists of n objects.

The number $n!$ is read as ***n* factorial**.

> By definition,
> $$0! = 1.$$

3.2.3 Number of lists of length k of n objects

How many distinct lists of length k can I make with n objects? As before, there are n choices of the first position on the lists, and then $n-1$ choices for the second position, etc., down to $n-(k-1) = n-k+1$ choices for the k^{th} position on the list. Thus there are

$$\underbrace{n \times (n-1) \times \cdots \times (n-k+1)}_{k \text{ terms}}$$

distinct lists of k items chosen from n items. There is a more compact way to write this. Observe that

$$\begin{aligned}
& n \times (n-1) \times \cdots \times (n-k+1) \\
=& \frac{n \times (n-1) \times \cdots \times (n-k+1) \times (n-k) \times (n-k-1) \times \cdots \times 2 \times 1}{(n-k) \times (n-k-1) \times \cdots \times 2 \times 1} \\
=& \frac{n!}{(n-k)!}
\end{aligned}$$

Thus

$$\boxed{\text{there are } \frac{n!}{(n-k)!} \text{ distinct lists of length } k \text{ chosen from } n \text{ objects.}}$$

Note that when $k = n$ this reduces to $n!$ (since $0! = 1$), which agrees with the result in the previous section.

3.2.4 Number of subsets of size k of n objects

How many distinct subsets of size k can I make with n objects? (A subset is sometimes referred to as a **combination** of elements.) Well there are $\frac{n!}{(n-k)!}$ distinct lists of length k chosen from n objects. But when I have a set of k objects, I can write it $k!$ different ways as a list. Thus each set appears $k!$ times in my listing of lists. So I have to take the number above and divide it by $k!$ to get the number of. Thus

$$\boxed{\text{there are } \frac{n!}{(n-k)! \cdot k!} \text{ distinct subsets of size } k \text{ chosen from } n \text{ objects.}}$$

3.2.1 Definition *For natural numbers* $0 \leqslant k \leqslant n$

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

is read as

"n choose k"

It is the number of distinct subsets of size k chosen from a set with n elements. It is also known as the **binomial coefficient**.

Other notations you may encounter include $C(n,k)$, ${}^nC_k$, and ${}_nC_k$. (These notations are easier to typeset in lines of text.)

3.2.5 Some useful identities

$$\binom{n}{n} = 1$$

$$\binom{n}{1} = n$$

$$\binom{n}{k} = \binom{n}{n-k}$$

$$\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k} \tag{1}$$

Here is a mechanical proof of (1).

$$\binom{n}{k+1} = \frac{n!}{(k+1)!\big(n-(k+1)\big)!} = \frac{n!}{(k+1)k!\,\frac{(n-k)!}{n-k}} = \frac{n-k}{k+1}\binom{n}{k}$$

So

$$\begin{aligned}\binom{n}{k} + \binom{n}{k+1} &= \frac{k+1}{k+1}\binom{n}{k} + \frac{n-k}{k+1}\binom{n}{k} = \frac{n+1}{k+1}\binom{n}{k} \\ &= \frac{(n+1)n!}{(k+1)k!(n-k)!} = \frac{(n+1)!}{(k+1)!\big((n+1)-(k+1)\big)!} \\ &= \binom{n+1}{k+1}\end{aligned}$$

But a more transparent argument is this: $\binom{n}{k+1}$ is the number of subsets of size $k+1$ of a set A with $n+1$ elements. So let B be the first n elements of A.

There are two kinds of subsets of size $k+1$ of A. In the first kind they all are from B, and there are $\binom{n}{k+1}$ subsets of this kind. Or, k could be from B and the the last element would be the $n+1^{st}$ element of B. There are $\binom{n}{k}$ of this kind.

It gives rise to **Pascal's Triangle**, which gives $\binom{n}{k}$ as the k^{th} entry of the n^{th} row (where the numbering starts with $n=0$ and $k=0$). Each number is the sum of the two above it:

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
 & & & & & & \text{etc.} & & & & & &
\end{array}
$$

Equation (1) also implies (by the telescoping method) that

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^k \binom{n}{k} = (-1)^k \binom{n-1}{k}.$$

3.2.6 Number of all subsets of a set

Given a subset A of a set X, its **indicator function** is defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

There is a one-to-one correspondence between sets and indicator functions. How many different indicator functions are there? For each element the value can be either 0 or 1, and there are n elements so

there are 2^n distinct subsets of a set of n objects.

3.2.7 And so ...

If we sum the number of sets of size k from 0 to n, we get the total number of subsets, so

$$\boxed{\sum_{k=0}^{n} \binom{n}{k} = 2^n.}$$

This is a special case of the following result, which you may remember from high school or Ma 1a.

3.2.2 Binomial Theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

3.3 Examples of counting and probability

To calculate the probability of the event E, when the experimental outcomes are all equally likely, simply count the number of outcomes that belong to E and divide by the total number of outcomes in the outcome space S.

3.3.1 How many different outcomes are there for the experiment of tossing a coin n times?

$$2^n$$

3.3.2 Binomial probabilities

What is the probability of getting k heads in n independent tosses of a fair coin?

$$\frac{\binom{n}{k}}{2^n} = \frac{n!}{k!(n-k)!2^n}.$$

We can use Pascal's Triangle to write down these probabilities.

$$\begin{array}{ccccccccc}
 & & & & 1 & & & & \text{(Prob of 0 Heads in 0 tosses)} \\
 & & & \frac{1}{2} & & \frac{1}{2} & & & \text{(Prob of 0, 1 Heads in 1 toss)} \\
 & & \frac{1}{4} & & \frac{2}{4} & & \frac{1}{4} & & \text{(Prob of 0, 1, 2 Heads in 2 tosses)} \\
 & \frac{1}{8} & & \frac{3}{8} & & \frac{3}{8} & & \frac{1}{8} & \text{etc.} \\
\frac{1}{16} & & \frac{4}{16} & & \frac{6}{16} & & \frac{4}{16} & & \frac{1}{16} \\
 & & & & \text{etc.} & & & &
\end{array}$$

3.3.3 How many ways can a standard deck of 52 cards be arranged?

(order matters)

$$52! \approx 8.06582 \times 10^{67}$$

or more precisely:

80,658,175,170,943,878,571,660,636,856,403,766,975,289,505,440,883,277,824,000,000,000,000.

3.3.4 How many different 5-card poker hands are there?

(order does not matter)

$$\binom{52}{5} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2,598,960.$$

3.3.5 How many different deals?

How many different *deals* of 5-card poker hands for 7 players are there? (order of hands matters, but order of cards within hands does not),

$$\underbrace{\binom{52}{5}\binom{47}{5}\binom{42}{5}\binom{37}{5}\binom{32}{5}\binom{27}{5}\binom{22}{5}}_{\text{7 terms}} \approx 6.3 \times 10^{38}.$$

Each succeeding hand has 5 fewer cards to choose from, the others being used by the earlier hands.

3.3.6 How many 5-card poker hands are flushes?

To get a *flush* all five cards must be of the same suit. There are thirteen ranks in each suit, so there $\binom{13}{5}$ distinct flushes from a given suit. There are four suits, so there

$$4\binom{13}{5} = 5148 \text{ possible flushes.}$$

(This includes straight flushes.)

So what is the **probability** of a flush?

$$\frac{4\binom{13}{5}}{\binom{52}{5}} = \frac{5148}{2,598,960} \approx .00197$$

3.3.7 Deals in bridge

In Contract Bridge, all fifty-two cards are dealt out to four players, so each has thirteen. The first player can have any one of $\binom{52}{13}$ hands, so the second may have any of $\binom{39}{13}$ hands, the third may have any of $\binom{26}{13}$ hands, and the last player is stuck with the $\binom{13}{13} = 1$ hand left over.

Thus there are

$$\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13} \approx 5.36447 \times 10^{28}.$$

distinct *deals* in bridge.

3.3.8 Splits in bridge

Suppose your opponents have n Clubs between them. What is the probability that they are split k–$(n-k)$ between West and East?

This is the probability that West (the player on your left) has k of the n. East will have the remaining $n-k$.

There are $\binom{26}{13} = 10,400,600$ possible hands for West. In order for West's hand to have k Clubs, they[1] must have one of the $\binom{n}{k}$ subsets of size k from the n Clubs. The remaining $13-k$ must be made up from the $26-n$ non-Clubs. There are $\binom{26-n}{13-k}$ possibilities. Thus there are

$$\binom{n}{k}\binom{26-n}{13-k}$$

hands in which West has k clubs, so the probability is

$$\frac{\binom{n}{k}\binom{26-n}{13-k}}{\binom{26}{13}}$$

that West has k clubs.

For the case $n = 3$ this is $11/100$ for $k = 0, 3$, and $39/100$ for $k = 1, 2$.

3.3.9 Sampling with and without replacement

Suppose you have an urn U with N balls, of which n are red and the remaining $N - n$ are green.

[1] Some pedants will claim that the use of *they* or *their* as an ungendered singular pronoun is a grammatical error. There is a convincing argument that those pedants are wrong. See, for instance, Huddleston and Pullum [2, pp. 103–105]. Moreover there is a great need for an ungendered singular pronoun, so I will use *they* in that role.

• If $k \leqslant n$ balls are drawn *without* replacement, what is the probability that all are red?

Here the random experiment is to choose a subset of size k.

1. First argument: There are $\binom{N}{k}$ distinct subsets of size k. But in order for all of them to be red, they must in fact be drawn form the set of red balls. There are only $\binom{n}{k}$ distinct size-k subsets of red balls. Thus the probability is

$$\frac{\binom{n}{k}}{\binom{N}{k}}.$$

2. Second argument: (This argument is a little trickier, and tacitly relies on some intuitive properties of conditional probability that I'll gloss over until later.) In a order to draw a size-k set of red balls, the first time a ball is drawn there are n red balls out of N total, so the probability of the first being red is n/N. If the first is red, the probability of the second is red is $(n-1)/(N-1)$, as there are $n-1$ remaining red balls out $N-1$ remaining. We shall see in a bit that we should multiply these probabilities, etc., so the probability is

$$\underbrace{\frac{n}{N}\frac{n-1}{N-1}\cdots\frac{n-k+1}{N-k+1}}_{k \text{ terms}}.$$

Fortunately, these two arguments give the same answer, since

$$\frac{\binom{n}{k}}{\binom{N}{k}} = \frac{\frac{n!}{k!(n-k)!}}{\frac{N!}{k!(N-k)!}} = \frac{\frac{n!}{(n-k)!}}{\frac{N!}{(N-k)!}} = \frac{n(n-1)\cdots(n-k+1)}{N(N-1)\cdots(N-k+1)}$$

• For $k > n$ what is the probability that all k are red? Zero.

• If k balls are drawn *with* replacement, and the draws are independent, what is the probability that all are red?

This is easier. Let E_i be the event that the i^{th} ball drawn is red. Each E_i has probability n/N and they are independent. The event that all k are red is the intersection $E_1E_2\cdots E_k$, which has probability

$$\left(\frac{n}{N}\right)^k.$$

• With replacement, for $k > n$ what is the probability that all k are red? $(n/N)^k$, same as before.

It is intuitive that with replacement the probability of all red balls should be greater than with out replacement. This is certainly true for $k > n$, where we are comparing $(n/N)^k$ to 0. But even for $k \leqslant n$, think of it this way:

The first ball has an n/N chance of being red, and if its, then the second ball has a smaller chance, $(n-1)/(N-1)$, of being red, and the next one has an even smaller chance, etc., But with replacement each ball has an n/N chance of being red. Algebraically, $(n-i)/(N-i) < n/N$, so

$$\underbrace{\frac{n}{N}\frac{n-1}{N-1}\cdots\frac{n-k+1}{N-k+1}}_{k\text{ terms}} < \left(\frac{n}{N}\right)^k.$$

3.3.10 Matching

There are n consecutively numbered balls and n consecutively numbered bins. The balls are arranged in the bins (one ball per bin) at random (all arrangements are equally likely). What is the probability that at least one ball matches its bin? (See Exercise 28 on page 135 of Pitman [4].)

Intuition is not a lot of help here for understanding what happens for large n. When n is large, there is only a small chance that any given ball matches, but there are a lot of them, so one could imagine that the probability could converge to zero, or to one, or perhaps something in between.

Let A_i denote the event that Ball i is placed in Bin i. We want to compute the probability of $\bigcup_{i=1}^{n} A_i$. This looks like it might be a job for the the Inclusion–Exclusion Principle, since these events are not disjoint. Recall that it asserts that

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) = & \sum_i p(A_i) \\ & - \sum_{i<j} P(A_iA_j) \\ & + \sum_{i<j<k} P(A_iA_jA_k) \\ & \ \vdots \\ & + (-1)^k \sum_{i_1<i_2<\cdots<i_k} P(A_{i_1}A_{i_2}\cdots A_{i_k}) \\ & \ \vdots \\ & + (-1)^{n+1} P(A_1A_2\dots A_n). \end{aligned}$$

Consider the intersection $A_{i_1}A_{i_2}\cdots A_{i_k}$, where $i_1 < i_2 < \cdots < i_k$. In order for this event to occur, ball i_j must be in bin i_j for $j = 1, \dots, k$. This leaves $n-k$ balls unrestricted, so there are $(n-k)!$ arrangements in this event. And there are $n!$ total arrangements. Thus

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = \frac{(n-k)!}{n!}.$$

Note that this depends only on k. Now there are $\binom{n}{k}$ size-k sets of balls. Thus the k term in the formula above satisfies

$$\sum_{i_1<i_2<\cdots<i_k} P(A_{i_1}A_{i_2}\cdots A_{i_k}) = \binom{n}{k}\frac{(n-k)!}{n!}.$$

Therefore the Inclusion–Exclusion Principle reduces to

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}\frac{(n-k)!}{n!} = \sum_{k=1}^{n}(-1)^{k+1}\frac{1}{k!}.$$

Here are the values for $n = 1, \dots, 10$:

n:	Prob(match)
1:	1
2:	$\frac{1}{2} = 0.5$
3:	$\frac{2}{3} \approx 0.666667$
4:	$\frac{5}{8} = 0.625$
5:	$\frac{19}{30} \approx 0.633333$
6:	$\frac{91}{144} \approx 0.631944$
7:	$\frac{177}{280} \approx 0.632143$
8:	$\frac{3641}{5760} \approx 0.632118$
9:	$\frac{28673}{45360} \approx 0.632121$
10:	$\frac{28319}{44800} \approx 0.632121$

Notice that the results converge fairly rapidly, but to what? The answer is $\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k!}$, which you may recognize as $1 - (1/e)$. (See the supplementary notes on series.)

Bibliography

[1] R. B. Ash. 2008. *Basic probability theory*. Mineola, New York: Dover. Reprint of the 1970 edition published by John Wiley and Sons.

[2] R. Huddleston and G. K. Pullum. 2005. *A student's introduction to English grammar*. Cambridge: Cambridge University Press.

[3] R. J. Larsen and M. L. Marx. 2012. *An introduction to mathematical statistics and its applications*, fifth ed. Boston: Prentice Hall.

[4] J. Pitman. 1993. *Probability*. Springer Texts in Statistics. New York, Berlin, and Heidelberg: Springer.