Garrett Ordner

Probability and Statistics: A Primer for Beginners and Pre-Beginners

The Journey Begins: Probability Theory Part Seven: Finishing Up

Primary reference: Casella-Berger 2nd Edition

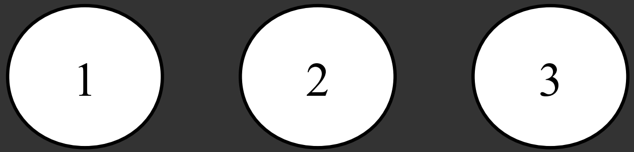
From Counts to Probabilities

Last time, we learned how to count all the possible outcomes of an experiment. In our examples, like picking a lottery number, all the possible outcomes had the same probability of occurring. This property turns out to be pretty handy.

Let's say we're picking one of three balls. Our sample space is

$$\Omega = \{1,2,3\}$$

We know that the probability of picking any ball is 1/3. Each ball has the same probability as any other.



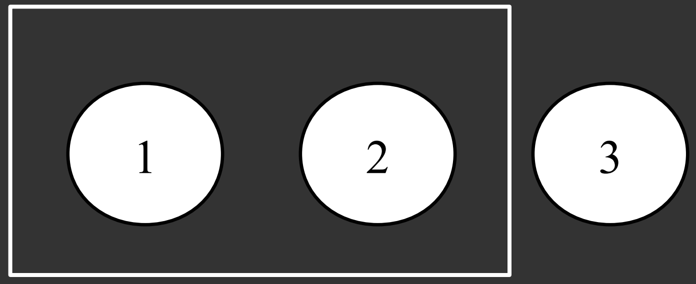
From Counts to Probabilities

Now, though, let's examine the probability that we would pick, say, either ball 1 or ball 2. Let that event be defined as

$$A = \{1,2\}$$

Since picking ball 1 is *disjoint* with picking ball 2, we can use the third axiom of probability:

$$P(A) = P(1)+P(2) = 1/3 + 1/3 = 2/3$$



A

Generalization

Ok, so now you probably see that we get the probability of A by dividing the number of outcomes in A (2) by the number of outcomes in Ω(3).

Of course, we can calculate this for any number of outcomes. If an experiment has N outcomes that are equally likely, then

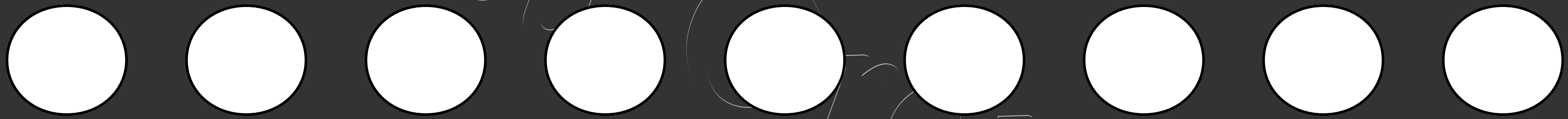
$\Omega = \{a_1, \ldots, a_n\}$ and $P(a_i) = 1/N$ for every outcome in Ω

$$P(A) = \sum_{a_i \in A} P(a_i) = \sum_{a_i \in A} \frac{1}{N}$$

So you just divide the number of outcomes in A by the total possible outcomes!

Counting ever higher

When we last left off, we were considering a scenario where a lottery player picks four from nine possible numbers:



We found that if the *order of the picks was important*, and they were picked *without replacing them*, then the number of possible lottery tickets was

$$n_1 * n_2 * n_3 * n_4 = 9 * 8 * 7 * 6 = 3{,}024 \text{ possible tickets}$$

Counting ever higher

A quick note here on an operation called “factorial”(!):

$$n! = n * (n-1) * (n-2) * \cdots * 2 * 1$$

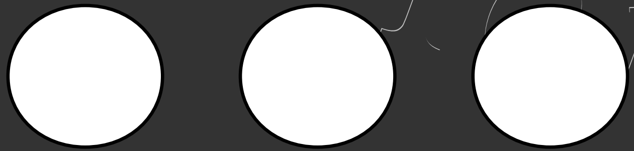
$$\text{So } 9 * 8 * 7 * 6 = \frac{9*8*7*6*5*4*3*2*1}{5*4*3*2*1} = \frac{9!}{5!}$$

So if we’re picking r items from n possibilities *without replacement*, and *order is important*, the job can be done in

$$\frac{n * (n-1) * (n-2) * \cdots * 2 * 1}{(n-r) * (n-r-1) * \cdots * 2 * 1} = \frac{n!}{(n-r)!}$$

But what if the order doesn’t matter?

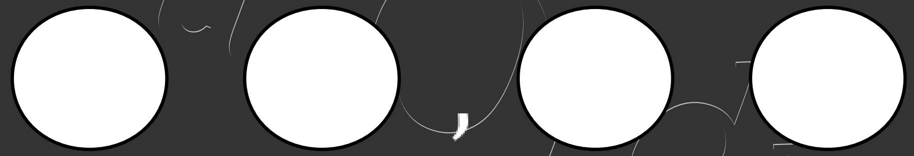
Good question, hypothetical reader! If the order doesn’t matter, then counting using the previous method won’t work. Let’s take a smaller example using three balls instead of nine where we need to choose two *without replacement*:



If order mattered, we could pick (1,2), (1,3), (2,1), (2,3), (3,1), or (3,2), for six possibilities. But if order doesn’t matter, then (1,2) and (2,1) are equivalent. Thus, three of these six possibilities are redundant, and we only have three possible picks.

Sounds like it'll get complicated.

Don't worry, there's an intuitive way to think about this: For any 2 balls you pick, you can pick them in two possible orders. For example:

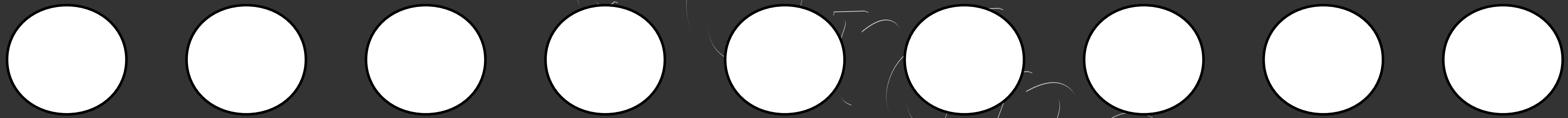


So just take the total number of ordered *permutations* for your initial pick (in this example, six), and divide by the total number of ordered permutations for the items you've picked (in this case, 2):

$$\frac{3!}{2!} = \frac{6}{2} = 3$$

Can we expand this example?

Definitely: Let’s go back to the nine-ball lottery. Now, you must pick four numbers *without* replacing them, but the order in which you pick them doesn’t matter (e.g. 1-2-3-4 is the same as 2-4-3-1)



We’ve established the number of picks when *order is important* to be:

$$\frac{n!}{(n-r)!} = \frac{9!}{(9-4)!} = \frac{9!}{5!}$$

Can we expand this?

But now, those four numbers we pick can be *in any order* and still be a winner. How many ways can we pick a given group of four numbers? Using the fundamental theorem of counting:

$$n_1 * n_2 * n_3 * n_4 = 4 * 3 * 2 * 1 = 4!$$

So we need to divide the number of ordered possibilities in the initial pick (4 from 9) by the number of ways we can pick them (4!):

$$\frac{\left(\frac{9!}{5!}\right)}{4!} = \frac{9!}{5!\,4!}$$

Looks like there might be a formula here.

Right you are! If we say we're picking r=4 numbers from n=9 possibilities *without replacement*, we know that if *order is important*, the number of possible picks is:

Right you are! If we say we're picking r=4 numbers from n=9 possibilities *without replacement*, we know that if *order is important*, the number of possible picks is:

$$\frac{n!}{(n-r)!} = \frac{9!}{(9-4)!} = \frac{9!}{5!}$$

So if *order is not important*, we simply divide this formula by the number of possible ordered sets of a given pick, r!=4!:

So if *order is not important*, we simply divide this formula by the number of possible ordered sets of a given pick, r!=4!:

126 possible picks

$$\frac{n!}{(n-r)!r!} = \frac{9!}{(9-4)!4!} = \frac{9!}{5!4!} = 126 \text{ possible picks}$$

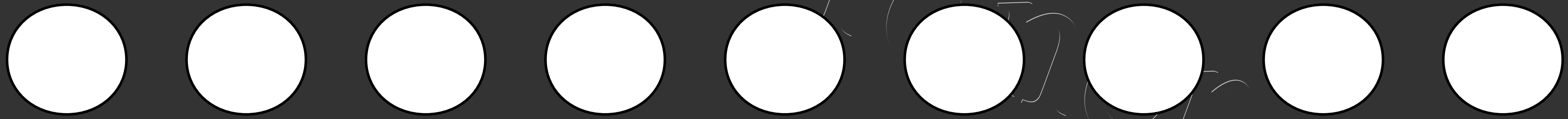
LEVEL UP!

You're ready to learn some new notation! The formula on the previous slide calculates the number of possible *combinations* and can be represented like this:

$$\frac{n!}{(n-r)!\,r!} = \binom{n}{r} = \binom{9}{4} = \frac{9!}{(9-4)!\,4!}$$

Hey, this all seems pretty easy!

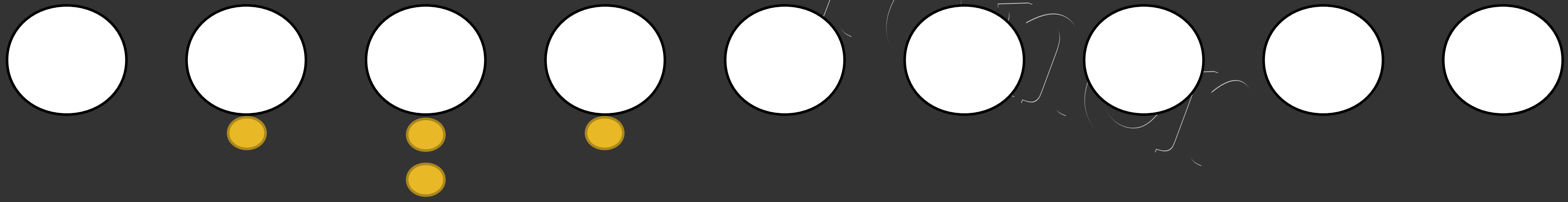
Well, I’ve got some bad news for you: Now we’re picking numbers *with replacement* but where *order is not important* (e.g. we could pick 1-1-1-2, but it would be the same as 1-2-1-1), and that, my friend, is not so fun. ☹ Once again, we’re picking r = 4 from n=9:



If we start trying to count the balls themselves, this’ll get complicated in a hurry. Instead, let’s try a different way of looking at the problem.

How should we approach this, then?

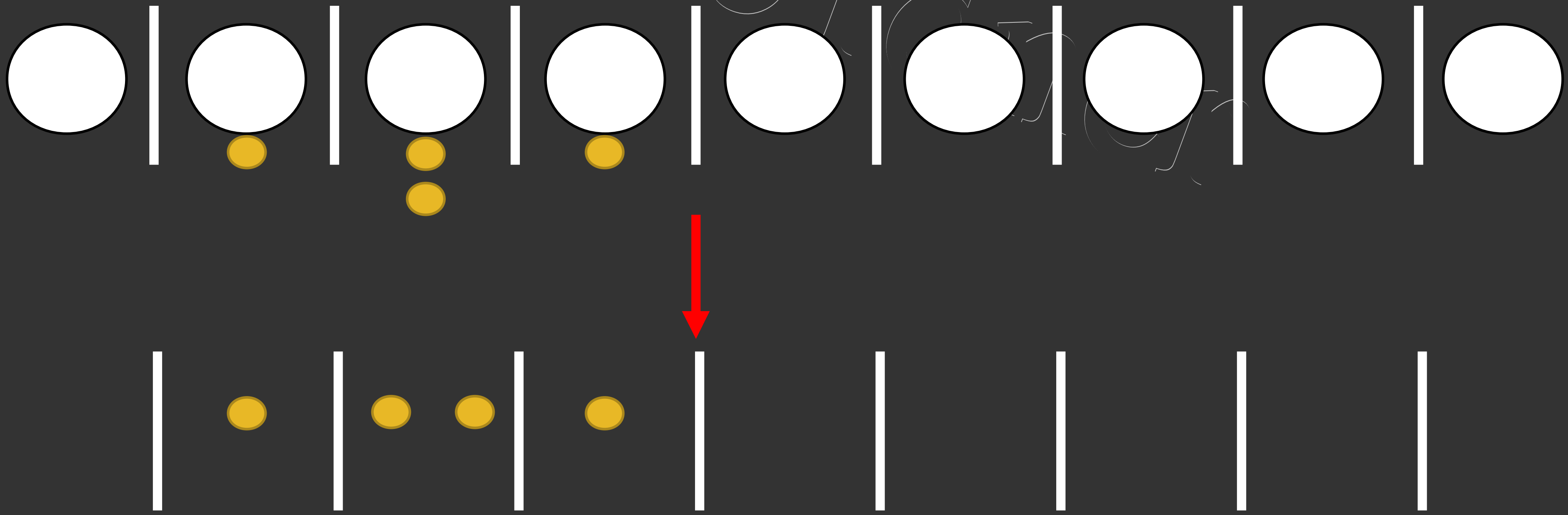
Instead of picking the balls, let's pretend we have four identical tokens, and each of the balls is a bin in which we place a token. For example, if we pick 2-3-4-3 (or equivalently, 3-3-2-4, for example), we mark it like this:



Still, this doesn't quite help us as much as we might like. Let's add one more element to this graphic to help us gain a new perspective.

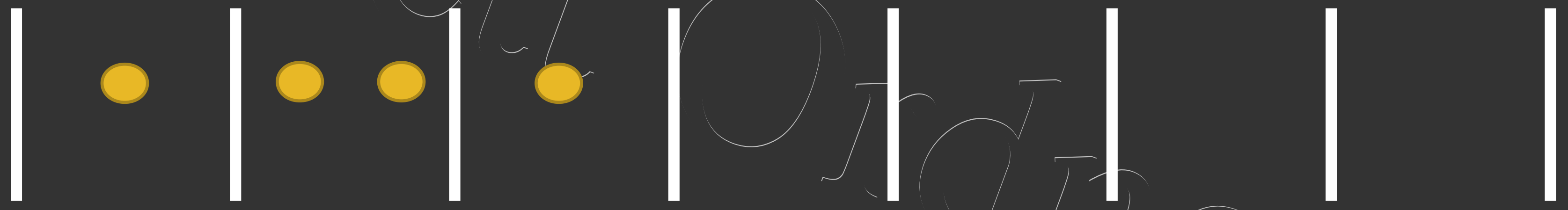
How should we approach this, then?

Now that we’re treating the balls as “bins” into which we can place our tokens, we don’t really need to see the balls at all. Rather, we can simply look at the separations between them:



What good does that do?

We’ve transformed this into a collection of 12 objects: four tokens, and the eight separations between our nine “bins”. (So n+r-1 objects)



Now we just need to order these n+r-1 objects. If every object was unique, that would give us 12! orders. But remember two things:

- The order of the tokens doesn’t matter.
- The order of the separators | doesn’t matter.

How do we handle the redundant token and separator orders?

Same as before: We divide by the number of equivalent orders! Remember, the eight identical separators could be arranged in 8! ways, and the four identical tokens could be arranged in 4! ways. So the number of possible picks *with replacement* where *order doesn't matter* is:

$$\frac{12!}{8!\,4!} = 495$$

Are we gonna generalize this one too?

Of course! Let's remember that we picked r=4 from n=9 and then work backwards from our final result:

Of course! Let's remember that we picked r=4 from n=9 and then work backwards from our final result:

$$495 = \frac{12!}{8!\,4!} = \frac{(9+4-1)!}{(12-4)!\,4!} = \frac{(n+r-1)!}{(n+r-1-r)!\,r!} = \binom{n+r-1}{r}$$

Now we've learned the final counting scenario: Choosing r from n *with replacement* and where *order is not important*.

Now we've learned the final counting scenario: Choosing r from n *with replacement* and where *order is not important*.

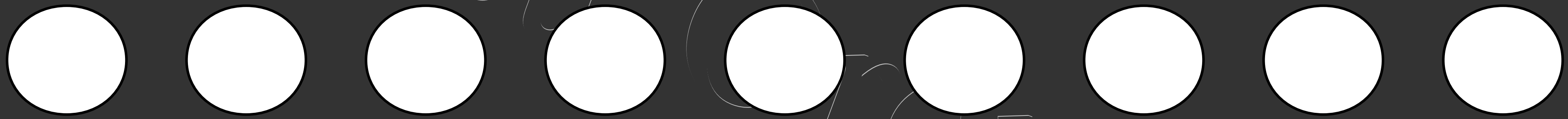
Counting Reference Table

For picking r items from a collection of n items

	Without Replacement			With Replacement		
		Without Replacement	**With Replacement**		**Without Replacement**	**With Replacement**
Ordered	**Ordered**	$\frac{n!}{r!}$	n^r	**Ordered**	$\frac{n!}{r!}$	n^r
	Unordered	$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$	$\frac{(n+r-1)!}{(n+r-1)!\,(r)!} = \binom{n+r-1}{r}$	**Unordered**	$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$	$\frac{(n+r-1)!}{(n+r-1)!\,(r)!} = \binom{n+r-1}{r}$
		Without Replacement	**With Replacement**		**Without Replacement**	**With Replacement**
Unordered	**Ordered**	$\frac{n!}{r!}$	n^r	**Ordered**	$\frac{n!}{r!}$	n^r
	Unordered	$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$	$\frac{(n+r-1)!}{(n+r-1)!\,(r)!} = \binom{n+r-1}{r}$	**Unordered**	$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$	$\frac{(n+r-1)!}{(n+r-1)!\,(r)!} = \binom{n+r-1}{r}$

Counting ever higher

When we last left off, we were considering a scenario where a lottery player picks four from nine possible numbers:



We found that if the *order of the picks was important*, and they were picked *without replacing them*, then the number of possible lottery tickets was

$$n_1 * n_2 * n_3 * n_4 = 9 * 8 * 7 * 6 = 3{,}024 \text{ possible tickets}$$