Multinomial Coefficients

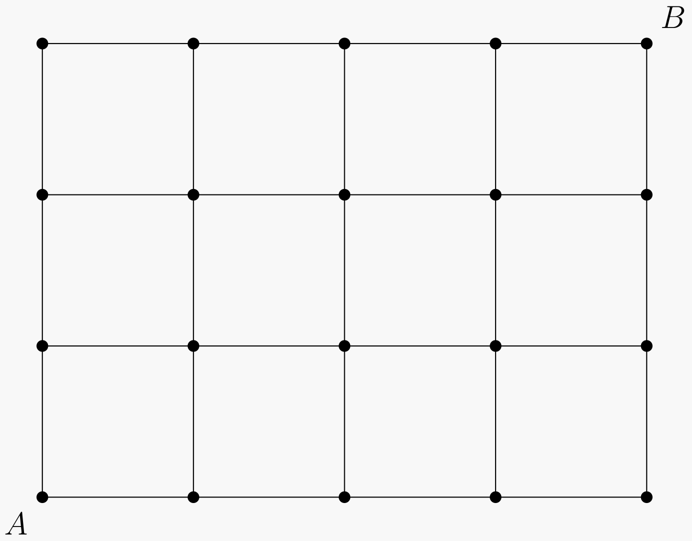
Multinomial coefficients count the number of divisions of n objects into r distinct and non-overlapping subsets of sizes $n_1, n_2, \ldots, n_r$.

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!n_2!\ldots n_r!}$$

Note that $\sum_i^n n_i = n$; the sizes of the subsets must sum to the size of the set of objects.

Example

Consider the grid below. Starting at the point labeled A, you can go one step up or one step up each move. You continue this until you reach point B. How many paths are there from A to B?



To reach B from A, 4 steps are to be taken to the right, and 3 steps are to be taken up. So, the total number of steps to reach B from A is 7.

Let the total number of steps be $n = 7$. Let the steps to the right be $n_1 = 4$. Let the steps upward be $n_2 = 3$.

So, to find the possible number of paths from A to B use the multinomial rule. The possible number of paths from A to B is:

$$\begin{aligned}\frac{n!}{n_1!n_2!} &= \frac{7!}{4!3!} \\ &= \frac{7 \cdot 6 \cdot 5 \cdot 4!}{4! \cdot 3 \cdot 2 \cdot 1} \\ &= 35\end{aligned}$$

So, the number of paths from A to B is 35.

Motivation for Probability Theory

Are the events happening around us certain?

- You leave home for the office, is it certain that you arrive on time?
- You take an exam, is it certain you pass it, or how much you score?
- Is it certain that tomorrow would be rainy?

Like these examples, we have countless other events whose realization is not certain. Why is the fact that events are not happening for certain relevant to us? If the event is relevant, they affect our decisions. Therefore, understanding how uncertainty is important in our decision making process. Which of the following statements is more effective and helpful in making a decision in the face of this uncertainty?

- The chance that it rains tomorrow is high
- The chance that it rains tomorrow is 90%

Obviously, the second is more helpful. This is where probability and statistics come into play.

Definition : Probability Theory

Probability theory is a set of methods used to quantify the degree of uncertainty attached the the realization of an event.

Basic Terms

Definition : Random Experiment

A random experiment is a procedure that:

1. Can be repeated as many times as we want
2. Has well-defined outcomes
3. Which outcome would be realized is revealed only after the experiment is executed

Definition : Sample Space

The sample space is the set of all possible outcomes of an experiment, denoted by Ω.

Definition : Event

Any subset of the sample space Ω is an event.

Definition : Sigma Algebra

A sigma algebra is a set containing all the possible events related to an experiment.

Definition : Set Definitions

Suppose that E and F are events in a sample space Ω.

- Complement: The complement E^c of E is the event containing everything in Ω not contained in E. E^c is the event "not E".
- Union: The union $E \cup F$ of E and F is the event containing everything in E or F (or both).
- Intersection: The intersection $E \cap F$ of E and F is the event containing everything in both E and F. This is sometimes written as EF.
- DeMorgan's Laws: These laws allow us to compute the complements of unions and intersections.

$$(E \cup F)^c = E^c \cap F^c \quad \text{and} \quad (E \cap F)^c = E^c \cup F^c$$

The Three Axioms of Probability

Consider an experiment with sample space Ω. For each event $E \subseteq \Omega$, we assume that a number $\Pr(E)$, called the probability of the event $E \subseteq \Omega$ is defined. Probabilities encode likelihood, and satisfy three axioms.

1. $0 \leq \Pr(E) \leq 1$
2. $\Pr(\Omega) = 1$
3. For any sequence of mutually exclusive events $E_1, E_2, \ldots$

$$\Pr\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \Pr(E_i)$$

An immediate consequence of this is that the empty set has a probability of 0.

Proof. Use Axiom 3 with $E_1 = \Omega, E_2 = \emptyset, E_3 = \emptyset, \ldots$ with $E_i = \emptyset$ for all $i > 1$. Then,

$$\begin{aligned}
\Pr(\Omega) &= \Pr\left(\bigcup_{i=1}^{\infty} E_i\right) \\
&= \sum_{i=1}^{\infty} \Pr(E_i) \\
&= \Pr(E_1) + \Pr(E_2) + \Pr(E_3) + \ldots \\
&= \Pr(\Omega) + \Pr(\emptyset) + \Pr(\emptyset) + \ldots
\end{aligned}$$

Subtracting $\Pr(\Omega)$ from both sides, we have:

$$0 = \Pr(\emptyset) + \Pr(\emptyset) + \ldots$$

Since $\Pr(\emptyset)$ is never negative, the only way to add an infinite number of copies of $\Pr(\emptyset)$ and have it equal 0 is if $\Pr(\emptyset) = 0$. $\square$

Example

The U.S. Census of 2010 gives the proportions / fractions of households in the United States with certain numbers of habitants:

Size of Household	1	2	3	4	5	6	7 or more
Probability	0.267	0.336	0.158	0.137	0.063	0.024	0.015

(a) Check that is a valid probability distribution.

(b) Let A be the event "the number of people in a household is between 3 and 5, inclusive." What is $\Pr(A)$?

(a) All of the probabilities are non-negative.

$$\begin{aligned}\sum_{i=1}^{7} \Pr(i) &= \Pr(1) + \Pr(2) + \Pr(3) + \Pr(4) + \Pr(5) + \Pr(6) + \Pr(\geq 7) \\ &= 0.267 + 0.336 + 0.158 + 0.137 + 0.063 + 0.024 + 0.015 \\ &= 1\end{aligned}$$

Therefore, this is a valid probability distribution.

(b)

$$\begin{aligned}\Pr(A) &= \Pr(3) + \Pr(4) + \Pr(5) \\ &= 0.158 + 0.137 + 0.063 \\ &= 0.358\end{aligned}$$