AGAINST ALL ODDS
EPISODE 21 – "BINOMIAL DISTRIBUTIONS"
TRANSCRIPT

FUNDER CREDITS

Funding for this program is provided by Annenberg Learner.

INTRO

Pardis Sabeti
Hello, I'm Pardis Sabeti and this is *Against All Odds*, where we make statistics count.

We've already learned that in the world of random phenomena, probability models provide us with a list of all possible outcomes and proportions for how often they would each occur over the very long term.

I could use a probability model to find out how many times I can expect to get heads on a coin toss…

Or how many daffodil blossoms I can expect to see in the spring based on the number of bulbs I plant in the fall…

Or even how many children in a family inherit a genetic disease.

You may notice a common thread throughout all of these examples. They're all concerned with things that have only *two* possible outcomes…
Heads vs. Tails
Blooms vs. None
Sick vs. Healthy

Traditionally we think of one possible outcome as a success and the other as a failure. Like when dealing with random variables, in this case what we're interested in is the overall count of successes. The count forms a particular kind of discrete probability model…the binomial distribution.

A device called a Quincunx illustrates the idea of a binomial setting. The ball has only two choices when it hits a peg – to go left or to go right. Those are the success or failure options in this case.

In addition to a situation having only two possible outcomes, there are three more conditions that must be true in order for it to be in the binomial setting. I could do yet another coin toss to demonstrate these, but let's take a look at something a bit more exciting – basketball free throws.

Free throws are mostly clear of any defensive pressure or other external factors during a basketball game. We only need to be concerned with whether or not the ball goes into the net…or not.

The first trait of a binomial distribution is that there is a repeated, fixed number of trials, or observations. In this case, *n* equals the number of shots taken. A basketball player could make hundreds of free-throws during a single season.

Second, all of these trials are independent; meaning that the outcome of one trial does not change the probabilities of other trials.

Now, conventional basketball wisdom might tell you that this isn't true if a player has a "Hot Hand." Some people believe that a player on a shooting streak is more likely to make subsequent baskets because of the success of previous shots. But a 1985 study proved that this wasn't the case.

> **Thomas Gilovich**
> There is no positive correlation between successive shots, in fact, there's a slight negative correlation, and what that means, translated into a more intuitive statistic is that, the probability of making a shot, after having just made your shot, isn't greater after just having missed your shot. And we find players are not any more likely to make their second shot after making or missing their first. The outcome of the second shot is independent of the first shot.

Pardis Sabeti
Therefore, free throws do fit this trait of a binomial distribution.

The third trait that we must look at is whether or not each of the trials end in one of two outcomes: Success, *S*, or Failure, *F*. Our basketball example fits: either the ball goes into the net, a success, or it misses, a failure.

Lastly, the probability of success, or *p*, must be the same for all trials. Free throws are always shot from the same distance, and there's no defensive pressure that would be present during play. Each time a player lines up at the free throw line he has the same probability of making his shot – it's based on his particular shooting skills.

So there you have it. Basketball free throws fit the binomial distribution!

Now that you've learned how to identify a binomial distribution, you'll see that pattern start to pop up everywhere. For example, binomial distributions can help determine the probability of how many children in a particular family will inherit a genetic disease.

> **Matthew Heeney**
> Hey Mikey.
>
> **Michael Jacob**
> 'Sup
>
> **Matthew Heeney**
> How ya doing?

Michael Jacob
Good.

Matthew Heeney
Hi mom.

Deborah Francis-Jacob
Hi Dr. Heeney.

Matthew Heeney
Come on in. Hop up on the table.

Where did you get the pain when you had it?

Pardis Sabeti
Sickle cell disease is a genetic disorder of the red blood cells, estimated to affect around 100,000 people in the United States. It can cause a great deal of pain, life-threatening infections, strokes, and chronic organ damage.

Matthew Heeney
Inside the red blood cells is a protein called hemoglobin. And that hemoglobin, there are many different types around the world. In the womb we all have fetal hemoglobin, or Hemoglobin F, and when we're born we change over to Hemoglobin A, or hemoglobin adult. But many people around the world have inherited other types of hemoglobins, and one of them is the sickle hemoglobin or Hemoglobin S.

Pardis Sabeti
In people with Sickle Cell disease, the Hemoglobin S molecules cause the normally round red blood cells to distort into a sickle shape, which causes blockages in the blood vessels. Tissues downstream are starved of oxygen, causing damage and much pain.

Like all genetic traits, the genes that determine an individual's hemoglobin type are inherited, one version from each parent. Since it's a recessive disease, a child needs to receive two bad versions of the gene, one from each parent, to have the disease.

Matthew Heeney
So in the typical situation, two parents who are carriers, or who have trait, one bad copy of the gene, each time the mother makes an egg, there's a 50% chance that she'll produce an egg that has the sickle mutation, and a 50% chance she'll produce a normal egg. On the other side, the father, 50% chance his sperm will have a sickle mutation, and 50% normal. And so when the sperm and the egg come together, statistically there's a 25%

chance of having a so-called normal AA genotype, about a 50% chance of having the AS trait, just like the parents themselves, and about a 25% chance of having the disease SS.

Pardis Sabeti
Even though there are four possible genetic combinations, only one of them results in a child with sickle cell. The other three combinations would result in a child without the disease. So inheritance of this trait fits the binomial distribution because there are only two possible outcomes—a child with sickle cell disease, or a child without. And, because the parents' genetic makeup never changes, these odds are constant for every child they conceive.

Matthew Heeney
Some of the common misperceptions or complications are that families feel that once they've had one child with sickle cell disease, that their next three are going to be ok. But in fact no it's with each pregnancy that there's that one in four or 25% risk.

Pardis Sabeti
Sickle cell disease fits the binomial distribution.
There are only two possible outcomes— sick or healthy child.

The outcome for each child is independent.

n is the number of children in a particular family.

And the parents' genetic makeup never changes, making the probability of having a child with sickle cell the same for each pregnancy.

We know that the probability of a child having sickle cell disease is .25 if both parents are carriers. Having sickle cell is labeled as a "Success," for our statistical purposes.

Success is the trait that we are counting, and of course has nothing to do with the successful health outcome for the child.

You can calculate binomial probabilities by hand or using software or graphing calculator. Once we have them, we can use the numbers to visualize the data on a probability histogram.

Let's look at the probability histogram for sickle cell disease in families with six children.

The *y*-axis of our histogram shows probability, while the *x*-axis shows the number of children with sickle cell disease in a family with two carrier parents. The *x*-axis runs from no children having sickle cell, to all six having the disease.

We can see from this histogram that the probability of having three children with sickle cell, for example, is 13%.

Public health officials might want to find the mean of this distribution. Luckily, we can calculate this pretty easily. We just multiply the number of trials, *n*, times the probability of success in an individual trial, *p*. In this case, we would multiply 6 by .25 to get 1.5. So the mean number of children with sickle cell disease, in families of 6 children where both parents are carriers, is 1.5. Of course no family has one and a half children! This number is the statistical average.

Until recently, most sickle cell care treated the disease's symptoms, trying to break down those blood flow blockages and alleviate pain. Now, many patients also take a preventative drug called hydroxyurea that was found to cause patients to start making fetal blood again.

> **Matthew Heeney**
> Even though they were well outside the womb, they started making that Hemoglobin F. And what's interesting is the Hemoglobin F stops the hemoglobins in the sickle cell from snapping together and changing the shape of the cell. So now we have large groups of patients who are being treated with this medication. And although it's not a cure it does decrease a lot of the complications that we see, particularly the pain, admissions to hospital, so it's really made incredible inroads into this disease for those who actually have it available to them and take it.

Pardis Sabeti
Dr. Heeney and his colleagues are working on a gene therapy protocol that would get sickle cell patients' bodies to make more of that fetal hemoglobin on their own.

> **Matthew Heeney**
> Not so much to fix the sickle gene but to sort of fool the blood stem cells into thinking they're in the womb again so they stop making sickle blood and they make fetal blood. So we hope that in a couple years we'll be starting a trial for that.

Pardis Sabeti
Stem cell transplants are also possible for patients who are interested and have a good donor match. So as the medical science advances, researchers and physicians like Dr. Heeney hope that children born with Sickle Cell Disease can look forward to a day in the not so distant future when there is a real cure for every patient.

For *Against All Odds*, I'm Pardis Sabeti. See you next time!

PRODUCTION CREDITS

Host – Dr. Pardis Sabeti

Writer/Producer/Director – Maggie Villiger

Associate Producer – Katharine Duffy

Editor – Seth Bender

Director of Photography – Dan Lyons

Additional Camera – Kristopher Cain

Sound Mix – Richard Bock

Animation – Jason Tierney

Title Animation – Jeremy Angier

Web + Interactive Developer – Matt Denault / Azility, Inc.

Website Designer – Dana Busch

Production Assistant – Kristopher Cain

Teleprompter – Sue Willard-Kiess

Hair/Makeup – Amber Voner

Additional Footage & Stills
- Pond5/James Cawley
- Vimeo/Phil Fried
- Pond5/Pavel Losevsky
- Centers for Disease Control
- Boston Children's Hospital

Additional Sound
- Freesound.org/CGEffex

Music
DeWolfe Music Library

Based on the original Annenberg/CPB series *Against All Odds*,
Executive Producer Joe Blatt

Annenberg Learner Program Officer – Michele McLeod

Project Manager – Dr. Sol Garfunkel

Chief Content Advisor – Dr. Marsha Davis

Executive Producer – Graham Chedd

FUNDER CREDITS

Funding for this program is provided by Annenberg Learner.

For information about this, and other Annenberg Learner programs, call 1-800-LEARNER, and visit us at www.learner.org.