AGAINST ALL ODDS
EPISODE 24 – "CONFIDENCE INTERVALS"
TRANSCRIPT

FUNDER CREDITS

INTRO

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

You're probably an expert on sampling by now. You know the point isn't to report about the individuals in your sample but to extrapolate from the information you have about them to something about the whole population. That's the basis of statistical inference – estimating a population parameter from your sample statistic.

For a quick refresher, let's use the example of a patient's blood pressure. A nurse takes one blood pressure reading at one moment in time. But of course blood pressure actually varies throughout the day, or the week, depending on things like if you're rushing around, super-stressed, or just relaxing on the couch. We'd get a better picture of one person's overall mean blood pressure – the population parameter – if we take a bunch of readings and calculate their mean – the sample statistic.

OK, so let's say twice a day for a week we measured our patient's systolic blood pressure. That's the force of blood in the arteries as the heart beats. The mean of those readings was 130. But how well does that number – based on our sample – really match the true value of his mean blood pressure? In other words, how trustworthy is our conclusion? After all, we might have come up with a number other than 130 for the mean if the readings that happened to be in our sample were higher or lower. Well, statisticians address this issue by calculating confidence intervals. Rather than a single number, like 130, we can compute a range of values along with a confidence level for that range.

The ability to state a precise confidence interval is extremely useful, particularly when millions of dollars are riding on your answer. That's the situation in the highly competitive business of batteries.

> **Duracell Commercial**
> The Duracell batteries we make now live longer than the ones from a few years back.

Pardis Sabeti
Battery companies have always trumpeted their products' long lives in their commercials…

Because the companies promise specific improvements and battery lifetimes, they need proof before the ads go on the air. Back on the factory floor – like here at Kodak's Ultra Technologies – technicians use rigorous testing and construct confidence intervals to back up the marketers' claims.

Richard Cataldi
When we're creating a confidence interval, there are many things that we have to take into account. One, we have to see that we have a Normal distribution, we have to have representative samples of the product.

Pardis Sabeti
Random samples of batteries are pulled from the warehouse. Then it's time for The Rack, a device from which no battery emerges alive! The Rack is more subdued in appearance than the toys, clocks, and flashlights these AA's might have been headed for, but it's mimicking the load of real products. It might seem more fun to test the batteries in the actual products themselves, but The Rack provides a much more controlled environment.

Larry Morgan
You'd have the variable of the life of the battery, but also the variable of the performance of the toy.

Pardis Sabeti
As the batteries continue to discharge in the standard test, a computer turns them on and off in cycles that represent the consumers' typical use pattern. Battery voltage is constantly monitored, keeping careful track of when each battery runs out of juice.

Richard Cataldi
The computer will compile for you the mean value, the standard deviation of that data and you can go ahead to the final step of creating the confidence interval with that data. The parameters you need of course to calculate that are the mean, the standard deviation, and the sample size, and your choice of what confidence interval you want. And what we can report in our engineering report is that this population of AA batteries, when used in a toy, can be considered to last seven and a half hours, plus or minus 20 minutes. And our confidence in that range is 95%. Our management can go ahead and make decisions based on that confidence level.

Pardis Sabeti
Let's retrace Kodak's steps and figure out how they came up with this confidence interval. Before we even get started, we need to make a few assumptions:

One—That we're working with independent observations. The life of one battery doesn't affect the lives of any of the others.
Two—That our data are from a Normal population or the sample size is sufficiently large. That's reasonable.

And three—That we know the population standard deviation. That's actually not a reasonable assumption in the real world, and there are ways to get around the problem. But for now, we'll assume we do know the standard deviation sigma.

Our sample mean "*x*-bar" is what statisticians call a "Point Estimate." It's a single number statistic that can be used to estimate a single number for the population parameter mu. One problem with using sample mean "*x*-bar" to infer the population mean mu is that "*x*-bar" can vary depending on the sample we take. If we randomly included a different selection of batteries in our test sample, we likely would have wound up with a different value for "*x*-bar." A single number like "*x*-bar" is not a very helpful estimate of mu without some indication of how accurate it is. Including a margin of error allows us to move beyond just a point estimate.

Luckily we know that since the sampling distribution of "*x*-bar" is normal, the mean of the sampling distribution, the mu of "*x*-bar," is the same as the unknown mean of the population, mu. We also know that the variability of "*x*-bar" is described by its standard deviation, which we can calculate with this familiar formula. In this case the population standard deviation sigma is 63.5 minutes and *n*, the number of batteries in our sample, is 40 so sigma of "*x*-bar" is about 10 minutes.

Think back to our 68-95-99.7 Rule. In any normal distribution, 95% of the observations lie within plus or minus two standard deviations of the mean. So 95% of all possible battery life samples give data so that mu is within plus or minus 20 minutes of that sample's "*x*-bar."

The population mean mu is unknown. But we do know that 95% of the population samples we would ever take would have an "*x*-bar" that is within plus or minus 20. Here "*x*-bar" was 450 minutes. So we can say with 95% confidence that mu lies within 20 of "*x*-bar." This interval, "*x*-bar" ± 20, is called the confidence interval. To say that we are 95% confident in our calculated range means that we got the numbers using a method that gives correct results 95% of the time. In other words, that's how likely it is that the method we used to compute the range actually comes up with an interval that contains the unknown population mean.

Think of the interval as a net. Each time we take a new sample we throw the net again. Sometimes we catch mu and sometimes we miss it. Here's the interval from one sample. It contains mu. Here's a second: another catch. And a third: that one missed. We can never know whether a particular confidence interval, like the one Kodak came up with for battery life, is a hit or a miss. But we do know that as we continue to take samples, we will catch the true value of mu 95% of the time over many, many samples. In other words, the probability that the method works is .95. So when we say that we are 95% sure that Kodak's batteries last between 430 and 470 minutes, we mean we got this interval by a method that works 95% of the time.

What if Kodak were willing to settle for only 90% confidence? Or what if they insist on 99% confidence? In that case, we need to find the z-value that leaves just 1% of the area under the curve in these tails. Statisticians have a general equation that will allow us to compute any confidence level we'd like. Here's the formula. Using the z-table or software to figure out the value for any critical values – called z^* – on the Normal curve you can create a confidence interval for any confidence level your heart desires. You always wind up with something in the form of a point estimate plus or minus a margin of error.

Though we can never achieve a 100% confidence level in statistics, I am completely sure that these concepts will come in handy as we move forward. Stick with us to find out how.

For *Against All Odds*, I'm Pardis Sabeti.

PRODUCTION CREDITS

Host – Dr. Pardis Sabeti

Writer/Producer/Director – Maggie Villiger

Associate Producer – Katharine Duffy

Editor – Seth Bender

Director of Photography – Noah Brookoff

Audio – Dan Casey

Sound Mix – Richard Bock

Animation – Jason Tierney

Title Animation – Jeremy Angier

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

Website Designer – Dana Busch

Teleprompter – Kelly Cronin

Hair/Makeup - Amber Voner

Additional Footage:
- Pond5/Jake Hellbach
- Always HD
- Pond5/Dmitriy Shironosov
- Pond5/TrueAutonomy
- Pond5/styrofilm
- Pond5/NSWE
- Duracell
- Energizer Commercials Used with Permission ©Eveready Battery Company, Inc.
- Pond5/alisbalb
- Pond5/Pavel Losevsky
- Pond5/P2HD

Additional Sound:
- Freesound.org/Timbre

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