AGAINST ALL ODDS
EPISODE 26 – "SMALL SAMPLE INFERENCE FOR ONE MEAN"
TRANSCRIPT

FUNDER CREDITS

Funding for this program is provided by Annenberg Learner.

INTRO

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

Inference is the name of the game when we want to make conclusions about a whole population based on our sample data. In our module about Confidence Intervals, we learned how to use *z*-procedures to calculate a confidence interval for a population mean. The *z*-procedure works in cases where we know the population's standard deviation. But that's hardly ever the case in real life! Now we'll learn a way around that limitation. For times when we don't know the population standard deviation but want to figure out a confidence interval and do significance tests, statisticians turn to *t*-procedures.

These *t*-inference procedures were invented in 1908 by William S. Gosset. If you're picturing a stuffy university professor, think again! Gosset was actually a chemist at the Guinness Brewery in Ireland, striving to find the best ingredients to use in their beer.

Making ale involves frequent sampling of everything from barley to yeast to the beer itself. Gosset wanted to save time and money by using small samples and their standard deviation as an estimate of the unknown population standard deviation. But using the standard deviation derived from only a few data values doesn't give a sufficiently good estimate of the entire population's standard deviation; and so he couldn't proceed with a *z*-procedure. After much laborious work, Gosset figured out a way around the issue, and finally created a new class of distribution called the *t*-distributions.

Let's pay a visit to Pretty Things Beer and Ale Project, a modern brewer where Gosset's century-old innovations can still help with quality control.

Brewing days start early for this crew. A small operation with no brewery of their own, Pretty Things had its beginnings in one connoisseur couple's love of all things beer.

> **Dann Paquette**
> Pretty Things is a company that my wife Martha and I started in 2008. Basically at the time, basically to employ me. I was in between brewing jobs and we had just moved back to the U.S. after living in England for three years and we needed to get me employed!

Pardis Sabeti

That modest goal met, they've grown and prospered to the point where Pretty Things is producing the equivalent of 8,000 kegs of beer every year. Their brewing process is a cross between chemistry and cookery. First Dann Paquette tweaks pH and mineral levels in the water to mimic what today's recipe calls for – a water profile from the northwest of Belgium.

Dann Paquette
The beer we are brewing today is Baby Tree. And that is one of our year-round beers. And it's based on the old monastic tradition of very strong dark beer that would be served in monasteries to visiting monks, or abbots more than likely. It was their best beer that they would make, so they would try to impress people with it, that was the whole idea, and this is our take on that.

Pardis Sabeti
The next big job falls to Martha Holley-Paquette, who's responsible for milling the malted barley that's the base of today's beer.

Martha Holley-Paquette
A lot of the grain comes from a silo outside and that's what we call the base malt and it's just sort of a gentle flavor pale color, and then we add flavors by using different types of grain.

Pardis Sabeti
Milling cuts each grain in half so the starch and enzymes within these toasted seeds are accessible.

Dann Paquette
It's going to come down the shoot and this, we're hitting it with water … it's going to hydrate the grain which allows the enzymes to break the starch down to sugar.

Pardis Sabeti
This hot porridgey mix is called the mash, and it cooks for about an hour, with Dann monitoring the temperature closely.

Dann Paquette
So I'm just raising the temperature of the mash to get a different blend of enzymes. I was at 152 and I'm going for 158.

Pardis Sabeti

Then it's pumped next door into the lauter tun, where the spent grains are separated from the clear liquid. At this point, the proto-beer liquid is called wort. Next it heads to the kettle, where the hops get added.

Dann Paquette
So inside the kettle you're boiling for an hour, you're sanitizing the liquid, you're creating bitterness, you're separating out the protein. Then after the boil you're shipping to a whirlpool which is essentially like a bit of a centrifuge, where you're allowing all of that material, hops and protein to come out of solution, and form a cone, so that we can draw the sweet wort off of that.

Pardis Sabeti
Dann cools the wort and sends it to the fermenting tanks where over the course of several weeks the yeast will work their magic and transform the sugars in the wort into alcohol and the finished beer. But before the yeast can get down to it, Dann takes a wort sample to check the liquid's density.

Dann Paquette
Should I be hopeful or pessimistic 'cause I can go two different ranges…

Pardis Sabeti
Dann and Martha aim for a density reading of at least 19.5 degrees Plato for this pre-fermentation wort. Every degree Plato means the liquid is a certain amount denser than water.

Dann Paquette
I'd call that 20.3.
That's ridiculous…

Martha Holley-Paquette
20.3 is good

Dann Paquette
It's really good, I mean I wouldn't mind making the beer like that.

Pardis Sabeti
The density tells Dann how much sugar is in the wort. It's kind of like measuring the potential for fermentation in this batch, and it looks like they won't have any trouble hitting the 9% alcohol content listed on the Baby Tree label.

There are several points in the brewing process where Dann and Martha take samples to make sure everything's running smoothly.

Dann Paquette
This is our last chance to dump the beers if they've turned out rubbish.

Pardis Sabeti
So far they've never had to dump a batch, thanks to their careful craftsmanship and some judicious sampling!

Let's imagine that Pretty Things wants to see how closely their production of Baby Tree beer is hitting their pre-fermentation density goal. We can go back in their records to collect data from a sample of 10 batches. Since our sample size is small, its standard deviation could be quite far from the actual population standard deviation of all the Baby Tree beer they've ever brewed. So the *z*-procedure won't work. It's time to call on the *t*-procedure that brewer William Gosset invented.

Compare a *t*-distribution to the standard Normal curve. You can immediately see they share certain features. They're both roughly bell-shaped, and symmetrical, with the mean at zero. But the *t*-curve is broader with a shorter peak, and its tails are higher. These fatter tails mean there's more probability of getting results far from zero. That's because the sample standard deviation varies from sample to sample, particularly when the sample size is small, adding uncertainty.

Another difference is that although there is a single standard normal distribution, there's a different *t*-distribution for every sample size. As our sample size increases, the sample standard deviation *s* gets closer and closer to the population standard deviation sigma. The curve narrows and gets closer to the standard normal curve.

Remember the different *t*-distributions are specified by sample size. Of course statisticians have their own particular way of saying that… instead of saying, "sample size," they use "degrees of freedom." The degrees of freedom for this *t*-statistic are always one less than the sample size. In our beer example, the sample size is ten, so our degrees of freedom are nine.

Here are ten density measurements Pretty Things got with their hydrometer. For these calculations to work, our data must be from a Normal distribution, which is a safe assumption in this case. Now we can calculate our confidence interval for the population mean for the density of all Baby Tree beer. Here's our equation: "*x*-bar" is the mean of our ten density samples. *s* is the standard deviation of our sample. Our sample mean is 19.72 and our calculation for *s* gives us .85. *n*, of course, is ten. We choose whatever confidence level we like… let's go with 95%.

We can look up in a *t*-table or use software to find the appropriate *t*-distribution critical value. We chose 95% and need to look for nine degrees of freedom… that gives us 2.262 for *t**, the critical value for our 95% confidence interval. Once we've plugged all the values into the equation, we calculate the confidence interval for the mean density of Baby Tree beer to be $19.72 \pm .61$. Like all confidence intervals, we're looking at a point estimate plus or minus the margin

of error. The interval estimate, 19.11 to 20.33 degrees Plato, gives us a range of plausible values for mu.

You might remember that the *z**, the critical value, for a 95% *z*-confidence interval was 1.96. Here the *t*-critical value of 2.262 gives us a wider confidence interval. That's the price we pay for having a small sample and not knowing sigma.

Pretty Things' goal for their Baby Tree beer is a density of 19.5 degrees Plato, which tells them they've got the right amount of fermentable sugars in the wort. Using the confidence interval we've calculated, we can say with 95% confidence that a 19.5 population mean is safely within our range. So based on our sample, it looks like Dann and Martha are consistently hitting the mark.

The personalized care and attention the Pretty Things brewers put into every batch is certainly paying off with its growing legion of fans.

Martha Holley-Paquette
So we're at Bukowski's, which is a bar in Cambridge in Inman Square, and we're doing a pint night.

Dann Paquette
So basically a pint night is people come in, they buy your beer, you give them a free pint glass. It's…if we were a young company it would be a desperate measure to get people to drink our beer, but we're just trying to give nice, cool things to our customers, basically.

Martha Holley-Paquette
And drink beer.

Pardis Sabeti
Pretty Things enthusiasts turned out to share in the beer love… even one pregnant fan due to give birth that day!

Pregnant Woman
I'm nine months pregnant so I can't have any, but I love Baby Tree, it's one of my favorites.

Husband
Gotta take one for the team, I'm drinking for three here.

Two Male Customers
Cheers!

Male Customer
It's a great beer. Great beer.

Table of Customers
Cheers! Hey hey!

Pretty Things Fan
Thank you for the glass…add it to the collection. Cheers.

Dann Paquette & Martha Holley-Paquette
Thank you! Cheers!

Pardis Sabeti
See, statistics really does have some very practical applications! For *Against All Odds*, I'm Pardis Sabeti. Cheers!

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 – Noah Brookoff

Audio – Dave Graceffa

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 Stills
- Getty Images/Picture Post

Special Thanks
Bukowski Tavern, Cambridge

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.