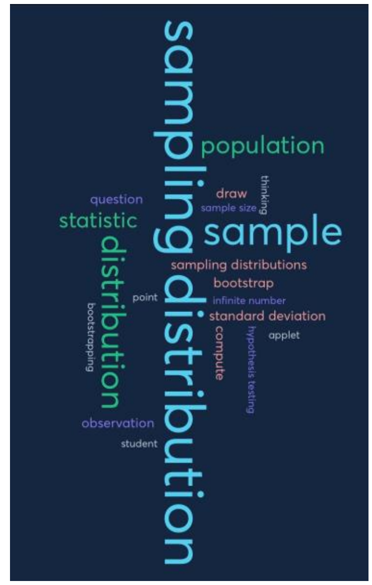
The Podcast *Quantitude*

with Greg Hancock & Patrick Curran

Season 3, Episode 21:

A Low-Resolution Discussion of Sampling Distributions

Published February 22, 2022 • 51:25



SUMMARY KEYWORDS
sampling distribution, sample, distribution, population, statistic, bootstrap, sampling distributions, compute, standard deviation, draw, question, observation, hypothesis testing, sample size, infinite number, bootstrapping, applet, student, thinking, point

Patrick 00:01
Welcome, my name is Patrick Curran and along with my bootstrapped friend Greg Hancock, we make up quantitative food. We are a podcast dedicated to all things quantitative ranging from the relevant to the completely irrelevant. In this week's episode, Greg and I discussed the critical distinction between sample distributions in sampling distributions, and we explore all the different ways in which sampling distributions are foundational to how we conduct research. Along the way. We also discuss Starbucks jazz, one item tests, Hot Pockets, delusions of grander Tetris, and Pong, drawing inappropriate distributions, magical properties, texting pictures of Kindle pages, Roman arches 1970s graphics, never say never mumbling, Green Day, ignoring Roy Lavie, real life bootstraps. And good night, Gracie. We hope you enjoy this week's episode. So we both have two kids in high school. My kids are juniors and Quinn is a junior and Tate is in ninth grade. Is that right? That's exactly right. Even though Tate is playing saxophone in the senior band,

Greg 01:11
they had a concert just the other night, if you close your eyes, you could be in Starbucks, where wherever jazz,

Patrick 01:19
that's your appreciation of jazz is Starbucks, I'm sorry, but I'm going to post edit this. And I'm going to put in the clip out of whiplash where they're talking in the bar, and he makes that comment about modern jazz. So that's going in right here.

MovieClip 01:35
Wonder why jazz is dying? Every Starbucks jazz album just proves my point really,

Patrick 01:43
Anyway...are your kids having trouble at their school, getting teachers?

Greg 01:48
There's a difference between getting teachers and getting qualified teachers.

Patrick 01:52
My kids are twins, and they share a Spanish class. They do not have a Spanish teacher, the gym coach comes in and just sits, wow, we're having dinner last night. We're talking about the teacher. And one of the kids said, Well, maybe they should just lower the criteria for teacher just to get somebody in the conversation went around to if you could hire a teacher by asking a single question, what would that be? And the kids distilled down to? What is your comfort level in spending a lifetime in poverty? That was one item question. Now it got me thinking about two things. There's a long con and a short con, we'll get to the long con in a minute. The short con is I was going to give you some professions that required some degree of qualification. And you have to tell me what the one question you would ask would be oh, gosh, okay, he just made a rude gesture at me. You and I have both hired babysitters, you are interviewing a babysitter and you have one question, what do you ask?

Greg 02:59
If I'm interviewing with my wife, then I won't get a chance to ask a question. I don't mean that in a bad way. I mean that because she will have 1000 questions that will all be oriented toward the safety and security of our children, all of those kinds of things. For me, it kind of will boil down to something like how many minutes? Do you put a hot pocket in the microwave? Right? If you can. If you can figure that out. I'm good.

Patrick 03:29
Okay, so you're one item question is how many minutes does a Hot Pocket take? That is not bad. And it reflects how you have raised your children when Goldie is not home. I have seen boxes of hot pockets and Alright, let's get a little bit closer to home because I see that you were in your cinderblock and asbestos recording room. So you must be on the campus of the University of Maryland. You brought your own paint to huff though I appreciate that. All of us in one way or the other have been involved in department chairs searches. If you had one item that would qualify someone to be a department chair, what would that be?

Greg 04:08
Are you applying for this job because you want it or because you don't want someone else to get it?

Patrick 04:16
That is exactly right. Anyway, it was a really fun conversation, because then it expanded and one that I liked was what would you ask a politician? And he said, Are you dedicated to the health and well being of all of the people who you represent? And the answer has to be no. For podcaster Oh, Christie's was Do you have delusions of grandeur. But it got me thinking last night if you were to ask a grad student one question upon which you would waive the first year grad stat sequence, what would that question be? Wow. So you've been head to your area. What would you ask them?

Greg 04:57

Oof. They're totally different directions. You could go with this right? You could go with a question about the coursework that they've had previously, no password be content has to be content. So I get one question.

Patrick 05:08
One question that if they know the answer to that question, they don't have to take the first year stat sequence, oh,

Greg 05:15
I don't know if I could pick one thing or not. There are a lot of really important things. But I don't know if I could narrow it down to one.

Patrick 05:19
What got me thinking in this way, when I came into Carolina, it was years and years ago, 23 years ago, now, something like that. Yeah, the program director to the time as a friend of mine, senior colleague, and occasionally students would transfer into the program in psychology more generally in developmental or clinical or social or they'd have a master's. And they would petition to not have to take the first year grad stats sequence, the policy at the time was that you would do an evaluation of the student to test them out of the first year sequence. And he did do a formal evaluation of the student, he would stop them in the hallway and ask them, Can you please differentiate a sample distribution from a sampling distribution? And if the student could answer that in just one or two sentences, he would waive the first year stats sequence? Hmm. Can you differentiate a sample distribution from a sampling distribution? And I thought it was brilliant, because if you think about it, if there is one concept that is maybe the keystone of everything we do, right, those old Roman arcs where the Keystone is that one that holds it all together, that is understanding what a sampling distribution is.

Greg 06:34
So I will say I'm having a bit of a little post traumatic reaction here, because that question almost verbatim was something that you asked me during our very first pop quiz episode during season one.

Patrick 06:46
No kidding. I will start the clock after I asked the question. In 90 seconds, please differentiate a sample distribution from a sampling distribution. To add insult to injury, I want to start by describing on a podcast a graphical animated applet, would you allow me to try to do that

Greg 07:14
I happen to know exactly what you're talking about. I would love to hear you distract.

Patrick 07:20
Okay, you know, what I love and I can see it in your eyes right now is you would not love for me to do that as a pedagogical mode of dissemination. You want to watch me screw it up big time. Okay, everybody. So what we're gonna do is we're going to talk about sampling distribution. This is a fundamental concept. A whole lot of people absolutely understand this and understand this deeply. Others are new to this. And we can puzzle through what is it that this thing is and why this is maybe

one of the most important concepts in all of quantitative methods I got to tell you is when my colleague told me this story is I grimaced a little bit. And I was like, I don't know if I would pass that. So a sample distribution we can think about you have a population. All right, let's come up with a construct. I always focus on these negative things of drug and alcohol use and anxiety and depression. We're going to talk about, I don't know, gratitude, it's really interesting area of research, gratitude in children, predictors of developmental trajectories, that serving potentially as a protective factor against other things. So how grateful Are you for the good things that happened in your life, picture in your mind's eye, a population distribution of gratitude, and let's make it normal, but we're going to come back and revisit that, okay. And we draw sample from that. But let's say we have 100 kids, we have some sample distribution of one sample, right? Yep. And we look at the distribution of gratitude. All right, that's what we're going to refer to is just a sample distribution, right. So the goal here is to differentiate the sample distribution from the sampling distributes. That's the sample distribution. All of us do this every day. Now, we want to use that as a mechanism to get to the underlying parameters. So we compute the mean of that. So there's the average of gratitude, and we have a standard deviation of gratitude. All right, what can possibly go wrong? Well, let's go to that applet. Everybody picture three plots that are in a column so there's an XY plot at the top in the middle and below now to find this will post it on the show notes but it is online stat book.com, which is a really wonderful resource. What I love about this Greg and I assume this is the one that you were knowing that I was gonna raise. It is like Tetris level graphics like it's just these little rectangles. Yeah, that animate but oh my gosh, yeah. If you're interested in this stuff, you have to go and play with this.

Greg 09:44
Whatever resolution your mind's eye has turned that way down to as Patrick said, Tetris level, I would have gone with Pong.

Patrick 09:56
I want you to picture in the upper one. That's the population two distribution of let's say gratitude, and like I said, is imagine that that's normal. For now, the middle one, we're going to draw a sample from that parent population. All right, so what I described was 100. And imagine your little Pong things, it literally rains down your little observations, and that creates the sample distribution that you have. That's the middle one. And then the lower one is it computes the mean of that. And like the world's simplest game of Tetris, the mean goes bla, bla, bla, bla, and drops on to the lower one. Alright, so Has everybody got that you got the parents on the top, you got the sample in the middle, and then you have what that mean is at the bottom, I will let's play around with some different sample sizes. Let's say that we were to do the world's dumbest experiment, and we draw a sample of one. Alright, so we have the parent population, the single little rectangle goes down to our sample, we're going to compute the mean of that. What is the mean of one? It's the observation? Now we're going to let that mean, go do to do and drop to the bottom one, I will do the sound effects throughout, would you especially because it doesn't do the sound effects on the applet,

Patrick 11:06
But when I teach I do!

Patrick 11:13

I do too! how have the one now we're going to do that, again, we're going to sample another single observation, we're going to compute the mean, which is the observation, and then we're going to let that drop to that bottom distribution. We're going to do that again. And again, and again and again, and an infinite number of times and that bottom distribution, think about it yourself. As you're listening, what is that going to look like? Well, it's going to be exactly what the top distribution is, because we only have one observation at a time. But now let's do the world's second dumbest experiment. And we're going to sample two observations. Alright, now Greg, you're gonna have to make the sound as I do as well. So you have the parrot in two boxes drop, do tu tu, tu tu tu. Alright, that was horrible on your part, you're gonna have to work on that I have two observations. And we're going to compute the mean, well, now it's going to split the difference between those two observations. So it's going to pull the lower one up, it's going to pull the upper one down to get the mean. And then that mean is going to go do to bloop on to the bottom distribution. Now we're going to do it again. And again, and again, and again, an infinite number of times. Now we have a distribution of the means for samples of size two, for a sample size of two. That is a sampling distribution. Alright, we're going to talk about this a lot. So this is just our first introduction, that is a distribution of a statistic at a given sample size, we are going to find that sample size, of course plays a critical role in this. All right, what is that sampling distribution center done? Well, the population mean, that's really cool, whatever the mean of gratitude is in the population, is that in the limit? That is if we do this an infinite number of times, the mean of the sampling distribution is the mean of the population. All right, how cool is that? What is the standard deviation of the sampling distribution? Is that the same spread of the parent population? No, it's a little less dispersed. Why? Because we took those two observations, and we combine them into the mean. So it reduced that variability when it dropped that down, right, so it's a little bit less disperse.

Greg 13:25

So whatever variance those two observations have, it reduced that down to a single observation.

Patrick 13:32

Now let's do it with three, we compute the mean, drop it down to the lower one, do that an infinite number of times, you already know where this is going, folks, we have a normal distribution centered at the population mean, but it's even a little bit less disperse, we've taken those three measures, and have reduced them to a single summary stat. So we've kind of squashed them together to use a technical term. And now that is less dispersed, still, now go to five, now go to 10. Go to 20. There are two things that we observed that sampling distribution of sample means remains normal. But as your sample size goes up, it becomes less and less disperse, right? There's less variability in it, and indeed, smarter people than you. And I showed that that gets less dispersed by a factor of the inverse of the square root of n, which is potentially one of the most magical properties in all of statistics,

Greg 14:37

even without mathematically proving that that has a lot of intuitive value for me, because if I have samples of size 10, those samples of size 10 have a certain amount of variance, but you're taking those 10 pieces of information, and you're reducing them down to a single representative value. So whereas you had variants of 10 observations, you are reducing the amount of them information down to 1/10 of what it was. So the variance of that sampling distribution, it makes intuitive sense that it would become

1/10 of what the variance was of the original population distribution. So when you tell me the square root of n, that makes total sense, too, because when I take the square root of one nth of the variance, what do I get, I get the standard deviation over the square root of n,

Patrick 15:22
go back to the applet, because you're going to picture that parent distribution, it's been normal, here's the fun thing is in this applet, grab your mouse and drag over it, you can make it skewed, you can make it by modal you can draw the finger I actually did that in class once is you can draw an extended finger, you can draw the St. Louis Arch, you can draw the McDonald sign, draw anything you want, in that parrot population, draw your sample, compute the mean and drop it down to the sampling distribution, and do that hundreds of 1000s of times, whatever you drew in that upper distribution, no matter how wildly inappropriate, it would be in front of a class of undergrads to draw the finger. The event is sufficiently large sample size, that sampling distribution of sample means is going to resolve to a normal distribution, it is freakin magic,

Greg 16:21
complete magic, it is so beautiful to watch, and what I do in my class. In addition to me drawing inappropriate shapes in the population distribution, you do it for samples of size two, and you look at the sampling distribution, and it still retains some of the characteristics of the original population, although it's sort of muted, it doesn't quite look like your finger, although a little bit, but then you do it with samples of size five, and 10, and 25. And all of that. And exactly as you said, the whole idea is that as your sample size goes up in the limit, it becomes this gorgeous normal curve. I love it. Magic is just the right word for this,

Patrick 17:00
someone would almost be tempted to develop a theorem that talked about the limit of the central distribution. Wow, would you call that? This is the central limit theorem, the sum of independent random variables tends toward a normal distribution, regardless of the parent distribution. Let me read a fun quote to you. We don't write like sending more, I will occasionally text you a photo taken on my cell phone, have a page on my Kindle of the Winston Churchill book I'm reading was it two nights ago that I did that?

Greg 17:36
I believe it was I think you sent me three different passages and you read that and you go, nobody writes like this anymore. It's so sad.

Patrick 17:43
I'm just like, oh, this is so beautiful. And so I literally take pictures of the Kindle and text them to Greg, which is a reflection that we need to get out more often. But I love beautifully written things mostly because I'm not capable of doing it myself. But golden has a little clip on this that I will read that I love when he's talking about a sampling distribution. He says I know of scarcely anything so apt to impress the imagination is the wonderful form of cosmic order expressed by the law of frequency of error. The law would have been personified by the Greeks and deified if they had known of it, it rains with serenity and incomplete self effacement amidst the wildest confusion, the huge are the mob and the greater the

apparent anarchy, the more perfect is it sway. It is the supreme law of unreason? Whenever a large sample of chaotic elements are taken in hand and marshaled in the order of their magnitude, and unsuspected and most beautiful form of regularity proves to have been latent all along.

Greg 18:48
Wow, I love that gorgeous.

Patrick 18:51
This is what my colleague was asking. If you can differentiate a sample distribution from a sampling distribution, then you have that Keystone in the Roman arch, it is

Greg 19:04
exactly the Keystone that you described. Because our whole life, our whole statistical life is not just about the behavior of a mean, it's about the behavior of statistics relative to that backdrop of what could happen just randomly. And so even though we're describing this in terms of a 1970s style graphic display

Greg 19:30
this applies to everything that we do. Even though I'm not a huge fan of one item tests. I totally get why you're calling chose this because it pervades everything.

Patrick 19:40
This is the point of the podcast where I might be tempted to say, so what, by

MovieClip 19:46
the way, when you're telling these little stories, here's a good idea. Have a point. It makes it so much more interesting for the listener.

Patrick 19:53
If it's so important, well, what role does this play in what we do? Okay, so we have have a distribution of a sample statistic. Now one thing that drives my students crazy, I don't know, if you have to wrestle with this with your own, we have leather bound books that catalog, the things that we teach you, and then criticize you for telling us back the way we taught you. It really is just a theme of academia. But we say if you were to sample an infinite number of times of sample size n from a given population, compute the mean gathered together the means and form the distribution, that would be the sampling distribution of a sample statistic at a given sample size. But you will never do that in practice. That's always the hard right turn. people much smarter than us have said, Look, if you have a sample of some given size, let's say we've got a sample go back to our 100 Kids measured on gratitude. If you compute the mean, and you compute the sample standard deviation, and you divide that sample standard deviation by the square root of n, that is an unbiased estimate of the standard deviation of that sampling distribution. Had you done that an infinite number of times. So I will ask you Hancock, you get a sample mean, you get a sample standard deviation, you divide the standard deviation by the square root of n, and you get the estimate of the standard deviation of the sampling distribution of means even though you only have one sample drawn from that finger distributed parent population. So what what do we do with that?

Greg 21:42
Well, it depends what year it is. If it was before 1908, you would say, oh, that's normal. But after 1908, thanks to Gossett right student, we realized that that follows a distribution that isn't quite normal, although it gravitates toward normal, the sample size gets better. So we understand that there is some distribution that is represented by this particular test statistic, in this case, that is the t distribution family. And for the sample mean, as you described, it's going to have n minus one degrees of freedom. But to your point, we don't have to do the infinite number of samplings for this, because we know we can derive or smarter people can derive exactly what the asymptotic distribution is going to be for this, whether it's for sample means or for differences between two sample means, or for a whole host of other statistics, we don't have to do this sampling process, even though it is completely fun to watch. Because we can figure out analytically what that shape is ultimately going to be. And that's what gives us the key family of distributions that we tend to rely on so heavily whether it's the z distribution, the t distribution, family, chi square distributions, or the F distributions,

Patrick 22:57
when you're talking about the two samples, I love thinking about that as well. Because let's go back to the mind's eye, what I described was a single population, a single sample a single mean, and then you gather those together, not just put a second column where you have two parent distributions, you have a sample, you get a mean, in each sample, you take the difference between those means. And that's what goes doo doo, doo, doo doo, boop, and drops down, and we build a distribution of the difference between two means. Now why on earth? Would anybody bother to do that? Well, you have a treatment and a control group, we have a distribution in the difference in the means. Now going back to why would we do that we teach our students that if you have a distribution of a random variable, one thing that we can do is if you have the mean, the standard deviation, we can actually compute probabilities that are randomly drawn individual would be above a point or below a point or between two points. So all of us who have taught our undergrad stat, you have had students curse at you, because we say assume a population value of mean of 20. When a standard deviation of to what is the probability we would draw a random observation that was greater than 12 for less than five, and then if you really want to piss them off, between eight and 11, right, and this is where my 17 year olds are out at the kitchen counter, grousing up a storm about When will I ever find the area under a curve in real life? And it's like, Well, funny, you should ask. That's what paid for the headphones that you're wearing right now. That's with a random variable. Well, what we just described then gives us the ability to do exactly the same thing, but with a sample statistic. So now it's not what is the probability that we would draw an observation that had a score? or 12 or higher, we can say, well, what is the probability that we would have observed a sample mean? Given some might I say no hypothesis, if you're doing a difference between two means, and we get a sampling distribution of the differences between means, we can look at the area under that curve to compute what is the probability we would have observed a difference in our sample means if there was really no difference in the population, this is how we do hypothesis testing, we reference sampling distributions.

Greg 25:40
And this really helps us to understand the consonance between the confidence interval world of thinking about things and the null hypothesis testing world of things. Sometimes people talk about how

confidence intervals are the new statistics. And we really need to think about things like that. I think confidence intervals, if interpreted properly, are very useful things. But in this world, you know, what we will often say, when you build a confidence interval, whether it's for a mean, or in this case, a difference between two means will often say, does that competence interval contain zero. And so we are doing our hypothesis testing with that. And the reason that we can do that is because the shape of the sampling distribution, whether it's for sample means, or differences between sample means the shape of that sampling distribution is the same whether it's centered at zero, or 48.7, or 96. That sampling distribution only differs in terms of where it's centered. And that's an incredibly useful property for hypothesis testing associated with means.

Patrick 26:43
Now, just to confuse all of you at home, because it would make so much sense to call the standard deviation of the sampling distribution, the standard deviation of the sampling distribution, but heaven forbid, we do that, we're going to call that standard deviation, the standard error. So all of you have looked at output, you have seen a point estimate, you have seen a standard error. And what do you do you divide the point estimate by the standard error, and you get the critical ratio, right? That is exactly the same thing as forming a z score of taking the individual observation deviated from the mean, and dividing by the standard deviation? Well, now we're taking the sample statistic, dividing by the standard deviation of the distribution of that sample statistic, when we take the mean and divide by the standard error and say we look for values that are greater or less than 1.965. What what are we doing? We have a standardized distribution of sample means. And we're saying Does your sample statistic fall more than two standard deviations away from the mean? Because if it does, then I'm going to air quote, call that significant. And say that it would be unlikely you would have observed that sample mean, if that mean, were equal to zero in the population? Now, here's the poke in the eye, oh, can I introduce a poke in the eye? Would you everything we've been doing this analytic. And what I mean is, we have done a thought experiment with this infinite number of samples gathered together the means compute the standard deviation, we never do that in practice, what we do is we have one sample one mean one standard deviation, we divide the standard deviation by the square root of n. And we say this is what the standard deviation of the sampling distribution would be had we done an infinite number of samples. One thing we know about statistics is you don't get something for nothing. And you got to pay the report, right? Those are two Cardinal mile markers in statistics, we have to invoke certain assumptions to allow that inverse root and to take hold. And we can't always use that there's some statistics that we don't know that sampling distribution, and there's some features where that doesn't work. Enlighten us, Dr. Hancock. Wow, that's really broad. That's why I pitched it to you because I kind of figure Yeah, run with it. You can

Greg 29:18
go to the bathroom now. I.

Patrick 29:21
It's okay. Thank you, me now.

Greg 29:25
I mean, go now I didn't go now.

Patrick 29:27
All right. I'll remember that for next time. Well, with regard

Greg 29:30
to the analytical properties, certainly those hold under the assumptions of random sampling, and that there isn't dependence among the observations, some sort of more complex samples. So that square root of n does correspond to a thought experiment with a lot of assumptions baked in about the nature of those samples. But the more interesting question is, what happens when you have a statistic whose behavior we can't just sort of arm chair like, I know, for example, that sample variance says ultimately are going to follow some version of a chi square distribution and ratios of sample variances are going to follow some version of an F distribution, under particular characteristics. Someone has done the heavy lifting on those already. But what if someone says, I want to do hypothesis testing about the median, where I want to do hypothesis testing about the 75th percentile point, then I go, Oh, my gosh, you know, I might make a first stab at the math, but that's gonna die hard. And so I have to figure out if I can arm chair this, how can I get the sampling distribution for something that I can't just figure out mathematically,

Patrick 30:44
and this is really fun, because it kind of goes back to the PoI applet, which is we started with a parent distribution, a sample, some sample statistic, and you do that a very large number of times, and you build the sampling distribution. And then we say, but you never do that in practice. And then just to jerk your chain, we say, but actually, sometimes we do do that in practice.

Greg 31:12
Never say never, ever say

Patrick 31:14
no. Okay, well add that to the list, you don't get something for nothing. You got to pay the Reaper and never say never.

Greg 31:26
So that primitive applet not only describes a bunch of things that we don't need to do, it describes a bunch of things that we actually do need to do. So if someone wants to do hypothesis testing about a median, or some other statistic that's analytically challenging, then go back to that thing and say, let me specify what I think the population looks like. And run this an infinite number of times. If I can't derive what the ultimate sampling distribution of that statistic like medians gonna look like, then turn the computer loose, specify the population, have it gather samples of whatever size 100 Do to compute the median, and then drop the median, boop, there it goes. And do that again, and again, and again. And I will get a sampling distribution of medians. And I can use that to characterize the randomness of those medians. I can use that to do hypothesis testing of that particular statistic that is just resampling. From the population,

Patrick 32:22

one might then refer to that as the parametric Bootstrap. Parametric because somehow the regression fairy came and gave us the parameters and not bootstrap bill out of Pirates of the Caribbean.

MovieClip 32:38
And you would encounter a bootstrap bill. We knew never sat well with Bootstrap what we did to Jack Sparrow.

Greg 32:46
Let me give you an example where we do a parametric bootstrap that is used fairly often. And it comes in principal component analysis. For example, when people are trying to decide how many components do I extract from my data? One strategy that has been used historically with different flavors on it is something called Horne's parallel analysis. And the idea here is that when you do a principal components analysis, you get out these Eigen values that represent how much variance each of those components accounts for in your original set of variables. When you're trying to decide how many components to extract, one thing that you could do is say, Well, how big would I expect those eigenvalues to be? By chance, if you can grind through the math of random Eigen values, then God bless you. But here's another way to do it. Imagine if we have, let's say, 20 variables that we're interested in doing principal components analysis on, let's create a world where we have 20 variables that don't correlate at all in the population, nothing going on. If you reach into the population and drought samples of size, let's say 200, you're not going to get zero correlations among your 20 variables, you're going to get little Irish correlations, oh 20304, those are going to yield Eigen values that will reflect the fact that there is some shared variability among those variables. So how big can that eigenvalue get just by chance? Well, this thing called horns parallel analysis says go create a population where all of the variables are uncorrelated reach into that drought a sample, let's say samples of size 200, compute the first eigen value, second eigenvalue, third eigenvalue, fourth eigenvalue, do that for another sample there another sample and another sample. And ultimately, what you will get is a sampling distribution for how big first Eigen values get by chance, how big second Eigen values get by chance, how big third Eigen values get by chance, etc. And then you go to your data, and you say, wow, is my first eigen value really big? Or is it in the realm of what I would expect to happen just by chance, and you figure that out by consulting these parametrically resampled or bootstrap sampling distributions for each of those, and whether you use the middle of those distributions as a cut off or 95th percentile or some other point, you know, doesn't really matter as much as the idea that you built the sampling distribution, you need to make this decision. That's an example of a parametric bootstrap for something that's otherwise kind of hard to get your head around analytically. And the only problem with that was having that population in the first place. And that's kind of the hitch for so many things that we do in hypothesis testing. We say, Well, if you know this about the population, and then we kind of mumble that out, and then

Patrick 35:32
we're very good at mumbling. Yeah.

Greg 35:37
That whole process only required was you to specify what the population look like, other than that piece of cake.

Patrick 35:44
Okay, Wiseguy, what, if the regression fairy has not left the parameters under our pillow upon which to build a sampling distribution? What do you do then?

Greg 35:55
Yeah, that's a great question. And that, frankly, describes the circumstance that we're in very, very often. So you know, we're going to have to do

Patrick 36:02
DIY, we DIY, I'm going to go out in the garage, I'm going to put on Green Day, which Frank does not appreciate a while ago, he came over, I was lifting in the garage, and I had Green Day on, he came over and pointed out that his wife who was charming and I really, really liked her is a cello player, and they have strings quartet that practices on their back deck, and I had Green Day, ranked while I was lifting. And he came over and asked me to run the garage door down. So we're gonna run the garage door down, we're gonna crank Green Day, don't want to be in a Mac. And we're gonna build literally build this is not figuratively, this is literally Yeah. Alright, so we're gonna go, we're down now.

Patrick 37:03
It's not a figured of sampling distribution, we are literally going to build our own sampling distribution,

Greg 37:10
if I have no idea what the shape of the population might be, which is like all the time.

Patrick 37:18
Hey, here's an example of what it looks like.

Greg 37:22
Patrick just showed me a distribution using his hand that had a very pronounced mode. Let's just say, Where do I have information that might give me some insights into what the distribution of the population might look like? And the answer is right in front of me, it's in the sample that I hold in my hand, just like Patrick said, we have a population, we draw a sample from that population, that sample has a distribution, a sample distribution, and if Oh, this is a big F, capital I capital F asterisks on either side, bold, underlined shadow.

Patrick 38:00
So how I write a grant. That's how I put the unique contribution is you just described how I write that in a grant application.

Greg 38:11
If your sample is representative in shape of the population from which it came, then we can actually use your sample and pretend that it's a population because we can't draw samples from the actual population, you know, we're going to do we're going to draw samples from your sample

Patrick 38:29
Didn't you have a gag once for a sponsor?

Greg 38:33
Today's episode is sponsored by bootstrapping, solving the problem of crappy small samples by making other crappy small samples out of your crappy small sample,

Patrick 38:42
was this what you're talking about? Exactly. But here, it's a good thing. And it

Greg 38:49
all hinges on the big F, if I have a sample that isn't representative of the shape of the population, then there's no amount of pretending nothing that comes out the other side is going to be particularly good. But if I have a sample that is sufficiently representative of the population in terms of shape, what I'm going to do to get a sense of what the sampling variability could be, is I'm going to reach into my sample and draw out a sample. But the key is, if my sample, let's say is a sample of size 300, I'm going to reach into that sample and draw a sample of size 300. But I'm going to do it with replacement. So I'm going to pretend the population I reach in there, I draw out an observation, I write it down, I throw it back in the mix, I stirred up, I draw out another observation. And I do that 300 times. But with replacement, so there's a very good chance, almost a perfect chance that I'm going to get at least some duplication of scores. But that's okay. When I draw scores from a population, we often get scores that are the same as well, I compute the test statistic that I'm interested in from that non parametrically bootstrap sample. I'm not assuming that I have the population and all its characteristics. I only have an estimate of that. So I get my sample statistic from that bootstrap sample. And then do it again. And I do it again and again and again. And when I'm done exactly like what Patrick described, I have a sampling distribution, but it is an empirical sampling distribution, based on the distribution of statistics that were from samples that were bootstrapped from my original parents sample,

Patrick 40:17
where does this stuff come in, in a really important way, while certain things that we estimate don't resolve to a normal sampling distribution, a mean does, the difference between two means does a whole variety of things does. But let's think about something like a product. What if you're doing something crazy that nobody would ever ever think about doing, say, looking at mediation, that you want to have X predicting y that in turn predict Z? And you want to get an estimate of a mediated effect? Well, what do we do, we take the product of the regression coefficient of x, predicting y and the regression coefficient of y predicting z, we multiply those two together, that is the mediator, the fact well, that actually does not have a symmetric sampling distribution, just to jerk Chris preachers chain, I am a vocal advocate of what's called delta method, and you get a sample distribution of this product that is symmetric. And I say, Oh, that's good enough, you can just use delta method, it's good enough, what's actually not, because what we do is divide by the standard error, and we look for values have wait for it, plus or minus 1.965. Okay, but that assumes a symmetric distribution of that particular sample estimate that is not governed by a symmetric distribution. So we do a bootstrap, which is now state of the art and every program will do this, where you do what Greg just described, you fit the model, you get the two regression coefficients, you take the product, and it goes do tu Tu, Tu, Tu,

bloop, on to that bottom? Do that 1000 times or 2000 times or given processors, now people will do it 10,000 times. And what you're doing is putting on Green Day, do you have that time to listen to me wine in building the sampling distribution for your mediated effect, and then you don't look for plus or minus some value, you get a confidence interval that is a symmetric, right is it's not centered around what that point estimate is. It's magic, right? We're back to Golden's I know of scarcely anything so apt to impress the imagination, you don't like a normal distribution for your sampling distribution. Okay, cool. Go out in the garage and build one, bring it back in. And you know, what works really, really well.

Greg 42:53
Absolutely. Not just really well for testing indirect effects, for example, but for a variety of things, you get this empirical sampling distribution, and you go out to whatever percentile points you want, you want a 95% confidence interval, great, then go to the two and a half percent out point, the 97 and a half percentile point. And as Patrick said, they're not symmetrically balanced around the center of that empirical distribution. Who cares? You don't need them to be we're not doing any plus or minus anything. Right? We go out there. And now we have our empirical sampling distribution. It is incredibly handy and has wide applicability. I mean, honestly, you could say, why not just do this for everything? Right? Yeah, the idea why not just go in your garage every time and build what you need? Well, there's some Reaper to be paid in this. Because when you do know the characteristics of the population, or there are assumptions that you can make reasonably, you will probably do better using those parametric results. But when you have statistics that don't follow those kinds of distributions, or you don't have conditions in the population that meet standard assumptions, this is very clever.

Patrick 44:01
Did you just feel a wrinkle in the force like somebody is screaming obscenities at us from a distance?

Greg 44:09
Like Arizona distance?

Patrick 44:12
I have this sense that Roy levy exactly is screaming obscenities at his Alexa, this is exactly a foundational concept in Bayesian statistics is we have some prior beliefs about the distribution of parameters. We bring data to bear. We update those prior distributions with the characteristics of the data and we get an updated posterior distribution. We can form confidence intervals, we can look at cut offs, we can integrate distributions. We can take modes of that we can do all sorts of things. This stuff scales up to the very highest level of statistics. are being conducted right now.

Greg 45:01
I love that you brought Roy up and the whole Bayesian thing. And what I'd like to do with that is ignore it. So the usual, the usual, the idea of sampling distributions that you DIY, whether you do it in a Bayesian form, or you do it in this kind of nonparametric resampling form that we're talking about has broad applicability, right? If you think about the world that you and I tend to work in, which is this ever expanding structural modeling kind of universe, bootstrapping has played a role in a variety of places. Sometimes people talk about it, not just for the indirect effects, as you've described, but also just generally for ways of characterizing the empirical sampling distribution of any statistic that you are

interested in cooking up in that world. For example, if I use a structural model to generate different types of reliability coefficients, whether they are Omega reliability or maximal reliability, there are a variety of statistics that you can generate, you go, what the heck is the distribution of that going to look like? Well, I want to know because I would like to be able to put a confidence interval around that thing that I care about, and maybe do so under whatever distributional characteristics my sample is embodying from the population, hey, Bootstrap is your buddy bootstraps your friend, we can even think about bootstrapping from the perspective of data model fit. Now, that's a little bit tricky, because if I draw samples from my sample, and compute, let's say, a comparative fit index, or a standardized root mean square residual, if I drew samples from my sample, and did it over and over and over and over and over, guess what, I'm going to get fit indices that tend to bounce around the fantasies that I have for my own samples, that doesn't feel so useful. But what gets done in this particular world is that your data are often transformed such that your model is assumed to be perfect, which it isn't. And then you bootstrap what fit indices might look like for a perfectly fitting model with your structure of the individual data. And then you looked how deviant your original models fit indices are from that, do your fit indices deviate? statistically significantly? I'll just use that term, right? Are they pretty out of whack relative to what you would expect? Well, then you have an empirical distribution that helps you to attach some sampling variability, some unlikelihood to what you observed with respect to your model, there's tremendous potential for bootstrapping. And I know I poo pooed it in a gag a long time ago. But that really was just because if your sample is crap, then your bootstraps are crap. But if you have a representative sample, and good thought has gone into gathering that particular sample, then bootstrapping can take you incredibly far, for just a wide variety of statistics. I really, really like it, you should

Patrick 47:58
think about any estimate that you get, is it a mean? Is it a correlation? Is it a regression coefficient? Is it a multiple r squared? Is that a test statistic for your SEM? Right? Anything that you estimate in that way, you should immediately think, what is the probability distribution that governs this random variable? Everything we do is sampling distributions. Yeah. When you ask a student, or when you ask yourself, can you articulate the difference between a sample distribution and a sampling distribution, this is what we're talking about.

Greg 48:36
So now, you know, what we really need is bootstrapping, in real life. So it's like the multiverse. If I could play stuff out over and over and over an infinite number of times randomly, and see what the distribution of outcomes is going to be, then I would know what the probability is that something bad is going to happen, or that I wind up sleeping on the couch.

Patrick 49:00
So you want to build a sampling distribution of arguments with your partner. And we can compute the inter quartile range of an infinitely large number of partner arguments, and that you can compute the probability that you are going to be above a certain level. I like that. All right, we often say here's a research question that needs to be pursued if you're looking for one. So we need somebody out there to do a real life Bootstrap. Yeah, well, maybe Quinn can program something for you. Because this is the same kid who wrote a batch script to log in to zoom for school, so he didn't have to get up in the

morning. So put him on that is we want a real life bootstrapping app. Alright, so that we can make better decisions in our lives. I'll see what he can do. Thanks, everybody, for hanging out with us. We always appreciate it. We hope you have found this remotely entertaining or informative or if on a good day, maybe both

Greg 49:59
and that may It helps you view things more comprehensively with regard to sampling distributions and variability. And I didn't need to say any of that.

Patrick 50:09
Say good night, Gracie.

Greg
Goodnight Gracie

Greg 50:12
Thanks so much for joining us. Don't forget to tell your friends to subscribe to us on Apple podcasts, Spotify, or wherever they go to resample stuff to avoid grading that huge stack of assignments. Wow, what were they thinking when they assign those. You can also follow us on Twitter where we are at quietude pod and visit our totally redone website quantitative pod.org where you can leave us a message, find organized playlists and show notes. Listen to past episodes and other fun stuff. And finally, you can get amazing quantitative merch like shirts mugs and notepads from Red bubble.com Where All proceeds go to Donors Choose that org to help support low income schools. You've been listening to quantitate the podcast based on the principle of most squares, well, two squares Anyway, today's episode has been sponsored by the generalized linear model. It's like the general linear model, but way like even general are and by one tailed hypothesis tests like our social lives, there's a good chance all the fun stuff is happening exactly where we are not. And finally, by the variance inflation factor, to be honest, it kind of does make your standard errors but look big in those pants. I mean in that regression model, this is most definitely not NPR.

MovieClip 51:19
Come out to the coast, we'll get together, have a few laughs