

The Podcast *Quantitude*

with Greg Hancock & Patrick Curran

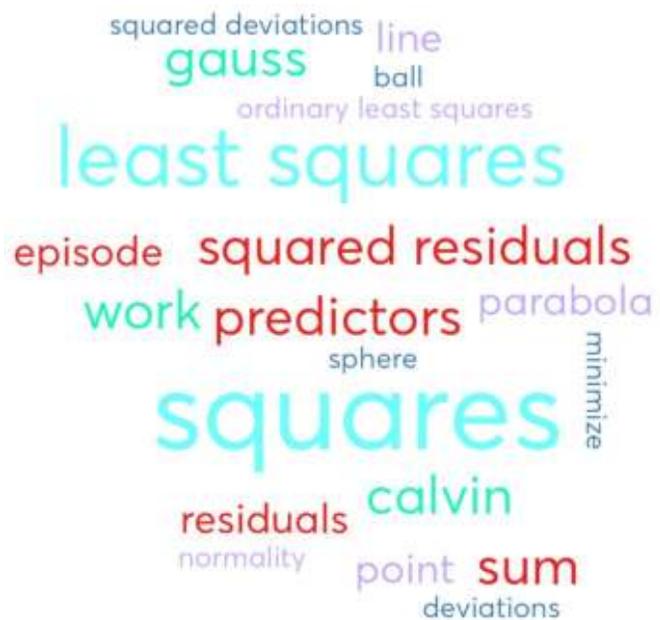
Season 4, Episode 1:

*Ordinary Least Squares:
Back Where It All Began*

Published Tuesday September 13, 2022 • 51:26

SUMMARY KEYWORDS

squares, least squares, Calvinball, work, Gauss, squared residuals, predictors, sum, episode, residuals, line, point, parabola, normality, ordinary least squares, deviations, sphere, squared deviations, minimize



Greg 00:00

He was thinking about trying something for the season for opening.

Patrick 00:02

I'm glad you have an idea because I have no freaking idea what we're doing here.

Greg 00:07

Well, you know, when the baseball players are walking up to the plate, you're a big baseball fan, right? So yeah, so they walk up and sometimes they'll play some song that the batter has chosen.

Patrick 00:15

Oh, dude, it's a walk up song. Absolutely.

Greg 00:18

Okay, excellent. walk up song. I thought that we could do that to introduce us for the season. What do you think?

Patrick 00:23

I love it. Okay,

Greg 00:24

what song would you pick for your walk up song?

Patrick 00:27

I gotta tell you I'm tempted to pick something by Olivia Newton John. Just because I was in love with her for so long.

Greg 00:34

You're Hopelessly Devoted to her.

Patrick 00:36

Okay, I'm gonna cut back. Do I have to pick just one because they're like a ton I can pick I'm going to go old school like AC DC. Thunderstruck excellent AC DC is from Australia and Olivia Newton John's from Australia. See we got a theme for the site.

Greg 00:53

Okay, nice. And I'm sure he collaborated on

Patrick 00:57

the wiggles. The Wiggles for my walk up sock to jagged, jagged Big Red. Travel Marin will travel? No, no, stay on task. Thunderstruck.

Greg 01:12

Alright, so this would be your walk up.

Baseball Announcer 01:16

Friends and fans standing of six feet zero inches 180 pounds of raw muscle. Haley from the Mile High City of Denver, Colorado. As a dedicated runner and trumpet player and longtime faculty member at the University of North Carolina Chapel Hill. Please welcome Quantitude co-host Patrick. What do you think?

Patrick 01:40

Oh, dude, I made a career error in not going into baseball. Okay, what's your walkup song?

Greg 01:47

I probably go with something like one of my favorite songs is take the power back but Rage Against the Machine. I think in the end, I'd probably go with that one.

Patrick 01:54

Okay, do you not know I this is a judge free zone. But when you could have chosen from Metallica or Led Zeppelin

Greg 02:03

or Olivia Newton John, pay back off

Patrick 02:12

Okay, all right. Let's give that one a go.

Baseball Announcer 02:17

Next, formerly five feet nine inches, weighing in at 7938 grams. Born and raised in the Emerald City, Seattle, Washington. He's an Aries. He's half Transylvanian, raised by a single dad and let him watch five hours of TV on school days. He's a longtime faculty member at the University of Maryland. Please welcome your other Quannah to co host Craig Han ha

Greg 02:47

ah, what do you think?

Patrick 02:48

I think that had you chosen big red car by the way those you wouldn't have had to bleep out a word. Two minutes into our first episode.

Greg 02:57

I think the wiggles go blue and a couple of episodes.

Patrick 03:01

Dark Side of the wiggles.

Greg 03:04

Maybe we should just have taped to the theme music again.

Patrick 03:07

Oh Can Tate help us out?

Greg 03:09

Let's ask him. Tate What do you think can you help us out with the music

Patrick 03:12

you got a gentleman Holy Cow I guess he's on the other side of puberty Welcome, my name is Patrick Curran and along with my far from ordinary least square friend Greg Hancock, we make up quantities. We are a podcast dedicated to all things quantitative ranging from the irrelevant to the completely irrelevant. In our opening episode two season four, Greg and I delve into ordinary least squares estimation, where it came from what it attempts to achieve and where it can take us from here along the way. We also discuss Golden Retrievers who are neither gold nor retrieve Olivia Newton John some calling out our own work. The meat sweats. You should know being intolerably self righteous, losing your camel saying ham and French Calvin bull Frank finding Gauss corpse crudely describing regression and sexy Hulk. We hope you enjoy this week's episode.

Greg 04:15

Welcome back season four.

Patrick 04:18

Do to take 30 in the morning. We don't need that. No. Whoo. Not now. Not ever.

Greg 04:24

All right. Well, can you quietly tell us Did you have a good summer

Patrick 04:27

I had a great summer, took some trips hung out on my brother's back deck and Denver ice went to LA with teenagers to visit colleges who a low of 68 and a high of 72 and I stepped off the aircraft and Raleigh and was hit in the face with humidity and I turned and tried to fight my way back when we took the plane and they just shoved me out I don't think so. took a road trip up to see you You did and that was great fun. Greg has a dog named Gus. What kind of dog is Gus

Greg 05:07

is called an English cream golden retriever, which means he's not golden at all. He's white. And for the record, he actually doesn't retrieve either.

Patrick 05:15

His Latin name is Leggus Lickus, because evidently his reason for living is to lick my legs, specifically out of a room full of bare legs. Mine get licked. Gus is adorable. And he has an entire basket of these little stuffed animals one by one throughout the day, he brings them outside in the front yard and leaves them and I pulled up in like some crop circle out front door near drove off. I mean, it's about the best security system you can have. Because clearly a psychotic family lives in the house with the crop circle of stuffies. How about you? How was your summer?

Greg 06:01

Well, I think you hit the nail on the head with the trips. I didn't go out to California, but the college trips are starting to be a thing, right? You have two kids who are seniors this year in high school, I have someone who's a senior in high school. So life is consumed with all things college. And frankly, I don't remember any of that for myself. I think I just remember submitting one or two applications. And that was pretty much it. But doing it we are as are you obviously

Patrick 06:27

and your son Quinn helped me with some entertainment value because he was working on his essay at the dining room table. I took a picture on my phone and texted it to my wife and said Quinn is almost done with his college essays. When my two kids don't even know where they're going to apply yet. 10 minutes later, one of my kids texted me back and said, What's the deal? So I want to thank Quinton for that.

Greg 06:57

I'll pass that on. What I didn't do over the summer is make any episodes with

Patrick 07:02

you. You know, this time we didn't even pretend now you and I both studied longitudinal data and we have some passing respect for trends in the first summer. How many episodes did we do eight. Okay, so that was like some kind of summer campers. Yep. Then the second summer to how many?

Greg 07:22

Three historical episodes? Yeah,

Patrick 07:25

this past summer? None. None episodes. And not only that, but we sub Condor work to everyone else by making them make the memes. There were some funny ass memes submitted

Greg 07:42

hilarious and clearly people have too much free time. Because we were getting like hundreds of submissions.

Patrick 07:51

You sent out what some kind of spiral binder or something?

Greg 07:54

Yeah, right. each week's winner got a quantitative spiral binder. Now we have a number of winners throughout and so we decided to pick a grand prize winner from among those

Patrick 08:05

Do you still have a printer at home? I do have a print day you're going to print out each of the memes. Okay, I want you to spread them out on a grid in the front lawn. I was traveling in Central America, I went into a bar there was a big crowd in the back there was a grid painted on the floor. People were frantically betting and exchanging monies they blew a whistle and let a chicken go. And everybody had bought a square. The square that the chicken pooped in is the person who won. And I lost \$20 You were going to put these out on a grid in your front lawn. Okay, and the one that Gus drops his stuffy on is going to be the winner. We're going to pause recording and I want you to go do this right now. Okay, we will come back when we have a winner

Greg 08:57

All right deal you know, my friend that was no small task. I put each of the images in an envelope so that I did not know what was what right this is that make single blind, double blind. So I put them around the yard. I didn't do a grid, but I spread them around the yard and then I let Gus out. And what did he do? He just flopped over on his side. He did not care at all about what was going on. But eventually I sort of got him up and around and he did in fact drop his stuffie outstanding.

Patrick 09:31

Which one was it was his moose stuffy? Okay, when I said Which one was it? I was a little less interested in which I was more interested.

Greg 09:42

It's the moose. The problem was he dropped it right between two of the envelopes.

Patrick 09:48

Okay, so fine. So you have two envelopes. I mean

Greg 09:51

to open the first one, okay. This is so exciting. It's dramatic.

Patrick 09:56

Okay, hold it up so I can see it do is that a picture of By leg.

Greg 10:01

Yep, I put a picture of your leg in there. Okay, obviously Gus even though he couldn't see it through the envelope, he clearly gravitated toward it. Do you

Patrick 10:08

mind if I ask just more broadly? How do you have a picture of my leg?

Greg 10:14

And that's a longer story. Okay. All right. So let's take a look at the real winner in the other envelope. Are you ready?

Patrick 10:20

Wait, I don't when you've got one.

Greg 10:24

Congratulations. It's a major award.

Movie Clip 10:28

Well of course it's a lamp you nincompoop. But it's a major award. I won it. Damn hell, how you say you won it? I have mind powers, mind powers.

Greg 10:43

The real winner is

Patrick 10:45

I can't see what it is. Describe it to me. I see the name Michael Matta.

Greg 10:49

Yep. Michael. Matta. He had one of the first memes of the summer the one about when you should or should not dichotomize a variable.

Patrick 10:55

And what was the image? I forget?

Greg 10:56

It was a pie chart.

Patrick 10:58

Oh, that's right. Right. Well, congratulations, Michael, does he want a picture of my leg bend is that

Greg 11:06

I will give him the option. Okay, I'll give him the option of either a picture of your leg or maybe the quantity of swag of his choosing. How about that?

Patrick 11:13

Nice. Yeah. Well, congratulations, Michael. And thank you to everyone for squandering your valuable time on this stuff as we had a ball with

Greg 11:22

that. Yeah, it was so much fun. Thanks, everybody. All right.

Patrick 11:25

Now, let's get back to work. So dude, I drove up. I did not have my red Camaro. No, that's right. What was it two or three years ago, I rented a car for reasons I won't bore you with the car that I selected was not there. And they gave me a red Camaro. And I was never going to come back. Although I drove the red Camaro. And then what five miles an hour on i 95. For like 100 No red Camaro, but I went up. We hung out in your backyard. We hung out with your family. You took me to meet fest? Yes. In downtown Baltimore. At the Brazilian steakhouse. Yep. And I got the meat sweats. And I was so proud of your son. Oh my god. If anybody has been to this restaurant, what's it called FOMO. To show these men and women walk around with grilled meat on spikes and cut it for you at your table. You have a little card in front of you that's green or red. I never turned mine over from green. I just love green. Greg's son Quinn sat down, and we all went to the salad bar table and came back and Quinn hadn't moved. And he sat there with a blank plate and over 90 minutes. All he ate was meat.

Greg 12:39

It was amazing. Unbelievable. I mean, this is a kid who sits in his pajamas all day coding, eating honey nut cheerios, that's pretty much it. We have to force food groups on him. And he shows up at this restaurant and cleans them out. I have never seen him eat that much. It's crazy.

Patrick 12:58

When we were downtown on the Inner Harbor, you and I took a walk down memory lane and sat on the bench where this entire boondoggle started. Now we did not have to scale a barbed wire top chain link fence like the first time. But it was really fun to sit there and say, Wow, that was a horrible decision.

Greg 13:24

The bench it was a lot of years ago. Now

Patrick 13:27

we're moving into season four, we got Episode One, what do we want to do? We got some ideas and knowing you and me, we won't do half of them. One of the things that we talked about was doing some more fundamentals kind of stuff. People are foolish enough to use a handful of our episodes in their teaching. I have an amazing graduate student named Chris Strauss. Chris is in their last year they're working on their dissertation and deed, they're going to propose it in a couple of weeks, they have tea aid for me for a number of classes, they are a gifted teacher, in something they use in their own class is what they call you should know. And at the end of each week, they send an email to their class of s YSK. You should know I have stolen this from my own teaching. And out of the lectures out of the readings out of the problem set. In the last seven days of everything that we've talked about. This is the that you should know, Greg and I thought we could weave some of that stuff in here. If you're navigating these waters, whether you're quantity or non quantity, whether you're doing developmental working quantitative methods, or you're a user or a teacher or whatever it might be. There are a box full of things that you should just have at your fingertips. We talked about this a season or two ago, my daughter plays piano and she has a professional teacher who talks about having songs at your fingertips, right and it's a similar kinda thing, but we thought that we would start season one with an s y s k on ordinary least squares.

Greg 15:08

And it's really just symbolic of how we have devolved, because in season one, it was, Oh, you just songs at your fingertips. And now it's like, yeah, it's Believe it or not, we actually get a ton of requests, right. Like last year, we had partial correlation. And we did sampling distributions and statistical degrees of freedom and information theory.

Patrick 15:28

Okay, information. You made me

Greg 15:32

do that. That's fine. All right. So today, we are going to talk about ordinary least squares. Because this is something that is foundational. This is like, well, it's like the bench that we sat on three and a half hours ago, Patrick gets wet. Okay, just stop I was having a moment was lead to you.

Patrick 16:01

My kids are out at the dining room table. And every time I come out from recording one of these, they look at me and say, What is your job? This is something you need in your back pocket because it is a foundation to so much of what we do and whether it's not exactly what we ended up using as an estimator. It is the motivation for moving to something else. So this could be you should know episode. This could be a how to be intolerably righteous episode.

Greg 16:35

That's you every. Sorry,

Patrick 16:39

okay, I see that. Somebody says something and you say, you know, that's not quite right, actually, and then go into how to be intolerably self righteous. So let's talk about that. Because we're a logical starting point is the orbits of the planets, I can always work in Ptolemaic and Copernican models of the universe. So

Greg 17:02

take us there get in our way back machine,

Patrick 17:04

this stuff goes back 2000 years, a lot of where what we can think about as curve fitting comes from is back where there were stars, there was the moon, there was the sun. People had figured out north, south, east west, but there were absolute fundamentals that were unknown. And a big one was what is the shape of the earth? And some people argued it was a disk. Some argued it was an infinite plane. That kind of made sense, right? Because if your camel wandered off and never came back, well, it just Dang, I got to get a new camel now. Some argued what was a sphere, you had empirical observations, you have the positions of the stars, you knew what city you were in, you knew how far the cities were apart, you knew where the stars were relative to the city. And with some very complicated geometry, they figured out was the sphere. How did they do that? Well, they fit a function to observations, that the only shape that would make sense is a sphere. And it's like, Alright, guys got it. Great. It's a sphere. If we wait long enough, the camera will probably come back. If it doesn't stop and start to live somewhere else. Except somebody said, Okay, I'm cool. It's a sphere. How big is it? Ooh, good question. We got to figure out how big the sphere is. Right. So then we got now we got a pretty good estimate how big it is cool. Hey, notice sun comes up. Sun comes down. We seem to be a sphere. It seems to be a sphere. We seem to be rotating around stuff. The Orbit gets a circle. Kepler says, you're all morons. It's an ellipse. Well, how do you know it's an ellipse? Look at the data dude, I fit this trajectory to the data. It's all curve fitting. Damned if Laplace back in the day right? We did. I'm sure we talked about Laplacian or summer, did we, you're better at remembering these things than I am. The French were involved. Let's just say that's always the French and dumb often

Greg 19:02

laugh. I just like glass, okay, you're

Patrick 19:06

just making stuff up. I traveled in France for several weeks, and all I ate were ham sandwiches, because that was the only French word I knew. sandwiches on. sandwiches on. We have observations, they are imperfect. It was trying to do some kind of curve fitting to figure out these massively fundamental things about whether you should wait for your camel to come back or get a new camel. And that brings up another issue that arises in everyday life. And it sounds silly to say but it's really important to think about in these settings is what is best, right? That's fundamentally when I think about estimation is what does it mean best? How do you best get an estimate of the curvature of the earth? And it's kind of funny because you hear people talk about well, what's the best kind of car to get What's the best restaurant to go to? And heavens? You and I are dealing with this with the kids. What's the

best college? Every time you think about it? It's with respect to what right? What do you mean the best college, you could go to Harvard and have a horrible experience, you could go to Cape Fear Community College and flourish? It all depends on how we define best.

Greg 20:23

So if you have a set of these observations, how would Kelvin bowl?

Patrick 20:27

No, it's all Calvin Ball estimation is all Calvin Ball. Okay, I am a die hard Calvin and Hobbes fan in Calvin and Hobbes is a thing called Calvin Ball, he plays it with himself and his stuffed tiger. And Calvin Ball is defined by rules that you make up as you go. And so in the middle of Calvin Ball, you make a new rule, and then that becomes rule for Calvin Ball until the next rule is made. I see estimation is Calvin Ball, which is we have a goal in which we have a belief about something that exists. The circumference of the Earth, we have a set of observations that we've made over time in different places where we want to estimate that circumference. So we have some population value, we have some sample data. And we want to get that and here's the best estimate of what that is based on the sample data. We can set up whatever criteria we would like. That's the Calvin ballroom.

Greg 21:28

It absolutely is, right. So if you wanted to stand somewhere in the middle of a set of scores, and talk about how much variability there is, first of all, Where do I stand in the middle of this set of scores? It's not entirely clear. And in part, maybe it depends on how you wanted to find the rules for variability. So you might want to say, where could I stand in the middle, and then judge what the biggest residual is what the biggest deviation is from where I'm standing, maybe that will help you to decide where a good place to stand is. So what we might say is, variability is the place where the average residual, the average deviation, I don't care whether it's above me or below me, or the average deviation is the smallest that simultaneously would accomplish two things define a middle place to stand. And it would also define a measure of variability, which would be like an average deviation from that middle that did not win. That version of Calvin Ball was not the ultimate winner

Patrick 22:25

think about two points, one that has a residual of plus one and one that has a residual of minus one. The mean of those two residuals is zero. Well done. Now you have a point that's positive 10 and negative 10. Well done. You got a mean Z row crap. Yeah, we need a new Calvin Ball Roll.

Greg 22:49

Yeah, I'll give you some got some positive deviations and some negative deviations, we could ignore the sign associated with that, which sometimes we do, but I will tell you mathematically dealing with absolute values is a complete pain. And so the other option, which we're all familiar with by now is, drumroll please.

Patrick 23:10

We want to avoid the plus one minus one plus 10 minus 10. And we don't want to strip off a negative by looking at absolute we can square so now we represent the plus one minus one, we have one and one

the plus 10 minus 10. We have 100 and 100, who now it represents that distance. But as you alluded to working with absolute values are an absolute punch in the face for deriving fit functions. Squares are not because squares take us to freshman math, they start defining really beautiful parabolas. And we can work with really beautiful parabolas.

Greg 23:54

So this idea of squaring deviations and using that as a criterion to me, I remember when I first learned this, it sounded kind of lame. It's like, oh, that's why we square it to get rid of the sign. And the answer is, well, yeah, partly, but dealing with the square has a lot of mathematically beautiful properties. But it also ties to other things that we care about statistically. Now I'm describing it as one came before the other. But if you think about other stuff that we do, we take moments of distributions right first moment, second moment, third moment, fourth moment, which has to do with powers associated with the deviation so it aligns really nicely with those and also squared deviations tied to distributions that we know and love. The F distribution is a way to compare variances that are based on squared deviations. The chi square distribution is a way to compare not a variance but a sum of squares to some known population value. So the idea of squaring these deviations is something that just brings this tremendous harmony of the spheres.

Patrick 24:55

I saw them open for Olivia Newton

Greg 25:00

You better shape up? No.

Patrick 25:09

So Laplace anticipated Calvin Ball like 250 years before, because what Laplace said is, hey, I've got all these observations, I've got a thing I'm trying to do. This is not just Hey, guys, look what I found. These are very specific goals. What is the shape? What is the size? What is the orbit? Can we predict relative positions of one planet against the face of the other planet? These are very targeted questions. And he says, Well, we can minimize the biggest residual, we could minimize the sum of the signed residual, we could minimize the average absolute residual each of these as a Calvin Ball criterium. And each of these as a form of best where he landed for the reasons that you just described is we can minimize the sum of the squared residuals. Okay, so let's pick that apart, minimize the sum of the squared residuals, least squares, why is an ordinary? Well, we'll talk about that because it turns out there are different variations of that. But that is where the term least squares comes from, is our criterion that we're going to use to define what is best is they are the values that we calculate that results in the smallest possible sum of the squared residuals.

Greg 26:37

So that means when we have a univariate case, the sample mean for a set of scores is the least squares estimator because the squared deviations from that the sum of squared deviations is a minimum. That's the principle of least squares. But now this extends well beyond the case of well, we've got 10 measurements of how Saturn transits across the heavens to where we start to have systems of multiple variables. But this principle of least squares carries forward I know

Patrick 27:03

it is just the first episode and I should have gotten my shovel before we started recording. Because I should probably go dig up Gauss.

Greg 27:14

You know, why do we even rebury Gauss anymore?

Patrick 27:17

Frank?

Greg 27:20

That's Patrick's neighbor, Frank, by the way,

Patrick 27:22

who is retired now the number of times he said, Hey, Patrick, I saw you got a new leaf blower and I can't help but notice Gauss corpse. I'm like, geez, let's drag in Gauss corpse because Gauss said, dude, Laplace, that is amazing, I got some things I can add to that. And that turns into the classical linear prediction model, which

Greg 27:47

for those of you who haven't done the calculus, it's such a beautiful thing, like you will weep tears of joy. If you go through and derive it, the line of best fit has a slope and intercept that minimizes the sum of squared deviations of the points from that line, where we define the deviations or residuals from that line as being in the vertical direction in the y direction. So we're not literally minimizing the perpendicular distances from the line. So if you were given a scatterplot, and someone said, Hey, throw a line in there, your eyeball might gravitate toward a line that minimizes how far the points are in terms of their most immediate direct distance to the line. But that's not what we actually do, we don't set a criterion as a distance that is a combination of x and y information. Because the goal of regression is often to be able to make predictions about or understand an outcome variable y, we focus on minimizing those residuals that in this case, are in the vertical dimension. And so if you go through the calculus and say what intercept and slope would minimize the sum of squared deviations of points in the vertical direction from a line, it just falls out the formula for the intercept of the formula for the slope and you really get all choked up after you do the calculus

Patrick 29:03

sexy hook so one of the blank memes was like mean hook and then evidently what's called sexy hold that's lost on me pothole occur whatever. But one of the submissions that was made and it didn't make it to the Final Four, but I have to admit, it was one of my favorite I still love the winner. But it was one of my favorite the raid gene Hulk was simulation and sexy Hulk was derivation and I know who did it and if you're listening, well done that was one of my favorite because I've told stories in the past where Bower and I've worked together for a lot of years and he actually understands math and I don't I do massive simulation of just hit it as hard as you can. And Dan is sexy whole can does the derivations we had an episode on are you tell me buddy is it was the regression In a friend you had in high school, we talked

about this a little bit before, I don't remember what episode that was. We talked about that a little bit in there. And just to briefly repeat it, of what Greg just described is it really is a beautiful thing. Because imagine that you have a scatterplot between x and y , and you want to go back to fifth or sixth grade. And we learn this here of the line of best fit a plus Bx , it's a line, the scatterplot around it, and so on. We want to get what is the best line? All right, well think about what is the best college what is the best car? What is the best estimate? Well, we're going to define best as that estimate, that results in the smallest possible sum of the squared residuals. So we could rage Holcomb, right, so we have an XY scatterplot. And we just pick point five for the slope, we calculate the residuals, we square them, we add them up, and we go to another plot where the x axis is that value of your SLO. So it's point five, and the y axis is the sum of the squared residuals. So we're going to take point five, and we're going to go up and we're going to put a dot right where that sum of the squared residuals now I'm going to rage HK point seven. I'm going to compute the residuals I'm going to square I'm going to add them up, I'm going to go to my second plot, I'm going to go to point seven. And I'm going to go up to what that sum of the squared residuals is. Point 3.2, negative point 1.5. I'm going to range the hell out of this, you're going to start to see the outline of a parabola. Sexy Hulk is going to say, Okay, do your Hulk voice of like no rage Hulk, I don't know, whatever you're doing. That was it. She's hot hook says I can write an equation for that parabola. And not only can I write an equation for that, I can write an equation for the derivative of that, meaning it is the slope of the tangent line, at any point, I can take that derivative equation and set it to zero, which means that is the point when it's horizontal, which means that's the point at the bottom of that parabola, which means that is the minimum sum of the squared residuals, we're going to call that a normal equation, the point at which that tangent line is zero. We're going to solve for that and damned if that isn't the least squares best estimate of the slope of that line. Given the Calvin bowl criteria, you said defined best, which was minimizing the sum of the squared residuals,

Greg 32:39

you rage with your rage hog, you stimulate the heck out of it, and you draw that parabola and Bower next door hot Hulk says, Yeah, I could derive that and does and ultimately, you guys come to the same answer. I will say that there's some value in rage Hulk for other things where it's much harder to derive, then you simulate the heck out of it, you get a sense of things. But for this, this is something that has a closed form. This is something that's drivable. And so we have exact equations for the intercept and the slope, there's so much that's beautiful in this. First of all, if I may say the idea of the sum of squares is not just something that solved a problem of distances being positive or negative. It's not just tied to other distributions, we use squares and sums of squares all the time in our daily lives. Think about the first time you took a geometry class, and you learned about a right triangle. And you were given the length of one side, the length of the other side, and you had to figure out the length of the hypotenuse, and you squared the length of each of those legs and you sum them up, you summed some squares, and the length of the hypotenuse was the square root of those, we are doing the exact same thing. So this whole sum of squares thing is really nothing more than an extension of Euclidean geometry and extension of pathetic Iris. It's got high pot noose written all over it. So this aligns so beautifully with everything that we have done, things that you have done since you were a little geometry student. And then it also extends beautifully to having more variables. Rather than talking about points vertically above or below a line. This extends into the multiple regression world.

Patrick 34:22

And a line becomes a plain picture facing the corner of a room and the floor that comes out to the left is x one. The floor that comes out to the right is x two y is the wall that goes up to the ceiling. You have a three dimensional space, you have a sphere of points. There's a relation, it's more like a football of points, right? You're taking a sheet of plywood and trying to at the same time, get that x one slope with y and that x two slope with y that jointly minimize the sum of the squared residuals of the points above and below the plane. Now we can rage whole can sexy for that as well. But instead of a two dimensional where there's a beta one on the x axis and a sum of the squared residuals on the y , now, there's a beta one with x one, beta two with x two, and the parabola becomes an egg. And now instead of computing a derivative of the nonlinear function, that gives us that tangent line that we set to zero, that gives us the bottom of the parabola. Now we have partial derivatives, where we're going to look at those tangent lines of each, we get a little cross at the bottom of the egg. And that is the value of b one and B two that jointly give the smallest sum of the squared residuals.

Greg 35:53

And that's about the last opportunity, we have to be able to visualize it the way you describe, but it obviously extends to more x 's as well bring as many x 's as you want. One way that I sometimes describe multiple regression, which is a very crude way, is that you have a bunch of X 's right and your goal, someone says, I'm going to give you these x 's, and you have to build something that looks as much like y as possible. And you come up with a recipe Oh, okay, so I need three tablespoons of x one and I need two teaspoons of x two, and I need three milliliters of x three, and then you come up with this particular thing that creates things that look as much like y as possible. So when we talk about sums of squares with respect to lines, and planes, and all of that, at the end of the day, the sums of squares that we're minimizing are the differences between the observed y 's and the predicted Y 's where the predicted y 's are the best concoction you can make where you only have the x 's to play with, this extends to as many x 's as you have. And it is a beautiful thing. When you

Patrick 36:57

said that it was a crude way of describing regression, I thought I was gonna have to beat out a bunch of stuff. You have x and y , and you take some damn combination of the mother. And that's your model, are you happy? That's how I crudely describe regression. Thank

Greg 37:17

you, Rachel.

Patrick 37:20

I'm telling you, the analogy works. Your point is exceedingly well taken, the squared deviations start coming up all over the place of sums of squares, and things that we do with those, whether that be a variance, whether that be taking a square root to get a standard deviation, whether it means powering them to third and fourth powers to get things like skewness and kurtosis. Whether that means we take the length of vectors and areas of a hyperplane to get a determinant. They are all variations on sums of squares. Alright, so to review so far, we have rage Hulk sexy, wandering camels, Calvin Ball, Calvin Ball, what about the Reaper?

Greg 38:11

The Reaper has to make an appearance in episode one, season four.

Patrick 38:13

Okay, buddy, where do we pay the Reaper,

Greg 38:16

there are a variety of places where we pay the Reaper, or we make assumptions that then make us pay the Reaper. One of the problems when you square things is that bigger things when you square them get way way bigger. And smaller things don't that means that whenever we square residuals or deviations, things that are very far away, get way way far away in a squared metric. And what that does, for better or worse is it starts to give certain data points more leverage, potentially more say in that measure of variability. So when we have a score, that is an outlier, you square that guy, and it really contributes and maybe even over contributes. So one of the downsides of least squares is that it's extremely sensitive to outliers.

Patrick 39:03

Okay, so I'm going to throw a little pop quiz atcha, then, I mean, we might as well work all sorts of things into this first episode. Oh, boy. All right, what what that takes us back to is the best. What is the best estimator? Well, there's not a single best, we have to have a set of Calvin Ball rules that if they apply least squares is the best method available. But if they do not apply, it may not be so here's your pop quiz. I made the effort to dig up Gauss and drag in his corpse. I think I've already done my part in this discussion. Tell me what are some of Gauss conditions that would lead ordinary least squares to be the best Linear Unbiased Estimator?

Greg 39:50

Okay, I'll throw one out and then you throw one out, but no, I Calvin Ball rules state that I can make up a new rule right now and it becomes a new rule. Did I understand that correct?

Patrick 40:04

You understood that correctly? Okay. How about normality

Greg 40:07

associated with the residuals.

Patrick 40:11

Gauss didn't say one word about normality of the residuals. This is a little bit of an urban myth of least squares does least squares estimation require normality. Gauss Markov never invokes the assumption of normality. That was a trick question. You're gonna come out swinging, aren't you?

Greg 40:32

I am. So when we want to make inferences later about predictions, normality starts to rear its head for the significance test. But specifically for the minimization for the least squares estimation. You're right normality doesn't even creep in there.

Patrick 40:45

That's another how to lose friends and be intolerably self righteous. It's a book Greg and I are working on. That's something that you can drop at a cocktail party, which is to obtain the best point estimates using least squares estimation invokes nothing about the shape of the distribution. Now, the shape of the distribution becomes fundamentally important. If we want to start making assumptions about the shape of sampling distributions. Why do we do that? Because we need to find areas under the curve. Why do we need to do that? So we can assign probabilities? Well, how do you get area under the curve if you don't know what the curve is? All right, all of that comes in. But Gauss, and then later Markov, who set out these Kelvin ball rules for least squares estimation, never invoked normality. So I saw your little intellectual Judo you just did to me, because not only did you bounce it back to me, but you didn't even do one to begin with. You started number one on me. So what did Gauss say?

Greg 41:52

Wait, no. And now it's your turn? How does that work? That's what we already said that you would do that. So okay, go ahead. Now you do the next one

Patrick 42:01

linear in the parameters?

Greg 42:02

What the heck does that mean?

Patrick 42:04

Dude, are you reflecting the entire thing on me, this is another urban myth, in least squares, as I have heard people say, you should never use linear regression. Because everything we study is nonlinear. That's not correct to say, the linearity is in the relation between x and y picture and x y relation, that is a big St. Louis Arch parabola have occurred, we can do that in the linear regression model, we have $\beta_0 + \beta_1 x + \beta_2 x^2$, there you go. We can have nonlinear predictors. But notice β_0 , β_1 and β_2 are themselves entering the model linearly, we do not have $\beta_0 + \beta_1 e^{\beta_2 x}$. That is nonlinear in the parameters, because β_2 lives up in the cockpit of the exponential, okay, that we can't do. But Gauss said, as an opening gambit, it has to be linear in the parameters. So then I guess I throw another one at you. Well, I've done like three at this point. And so I can run the table if you want. Let's rip through these. Alright, Gauss nor Markov ever said anything about measurement error. All right, this turned into an urban myths episode. It is an urban myth, that Gauss Markov says that the predictors are error free. What Gauss Markov states is the distributions are fixed and known. There you go, it means that we don't make an assumption about the shape of the predictors, and we assume the predictors are error free, but Gauss, even at the bottom of my grave that I'm trying to hide from Frank, is you can still hear Gauss mumble out there. I'm just gonna use like a British accent, but first, it'll come out as a pirate. And second, I don't know what his country of origin is. He's German. Okay, so I'm not even going there. At the bottom of his grave, Gauss keeps grousing about he never said anything about measurement error, it is that they are fixed and known. And then a logical outcome of that is they are error free.

Greg 44:26

That's number three. Actually. We're listening all the things it's not. That's great. What else do you have up your list speed through?

Patrick 44:35

No, you have not given me one. I know what's going on here. I see this.

Greg 44:40

How about homoscedasticity? Yes, Gauss did say that the variance across the length of the line across the plane across whatever surface it is that you have, the variability is constant across that.

Patrick 44:54

So that is the three dimensional baguette, you have an x y you have an ellipse, it comes away from the piece of paper from you because it's a joint distribution. And no matter where you cut the bag, get at different values of x and turn and look at the sides. That is the same distribution. Now, Gauss said that doesn't have to be normal, but it has to be homoscedastic. Later, when we want to do inferences, then that does have to be normal. When you cut that back get, it has to have a mean of zero. And the same variance no matter where you

Greg 45:30

cut. How about independence of errors? No two

Patrick 45:35

residuals can be any more or less related than any other to residuals. And that's sometimes called the IID assumption. The residuals are independent and identically distributed. What else you got on your list? A couple of core ones are you have to have at least one more subject than number of predictors. This is another intolerably self righteous one where you're at a cocktail party and say, oh, did you know? And what that is, is I don't know what, ah, what that allows is you can invert the matrix, if you have five predictors, all you need are six people and you're fine. Well, okay, maybe not. But you get the point, there has to be variability in your predictor variables, that makes perfect sense, you can't have a constant, you have to have variability in your x 's, you can't have perfect collinearity. So you can't have two predictors that are correlated 1.0, they can get awfully close. And we can talk about that on another day. But they can't be perfect. You can't have a correlation between a residual and a predictor. And you have to have a correct model. And that's a weird one to get your head around. Does that mean a correct model? Okay, you realize that over 10 minutes, you haven't answered a single question, and you have black belt Judo meet on the buy your thing?

Greg 46:47

If you don't know, then that's fine. I don't Would you please tell me? Wait, what was the question? That's a matter. Yeah, the assumption about something being the right model kind of primitive pervades a lot of things that we have, there's either a lot we could unpack there or not much that we could unpack there. I'm gonna judo move you and not unpack it right there. But if all of these things are in place, then the estimator is said to be blue. There you go. Best linear, unbiased estimate, and we have statisticians are Hopelessly Devoted to blue.

Patrick 47:20

No, no, it's too early in the season.

Greg 47:23

Hopeless.

Patrick 47:25

No, you're not dude, I'm gonna tap out right here, if you.

Greg 47:30

Alright, what do you want to say about Blue

Patrick 47:32

Kelvin, double Gauss Markov is Kelvin bowl. And when you meet that there exists no other estimator in that class of linear estimators that results in a smaller sum of squared residuals, it is the best Linear Unbiased Estimator available, but it's just like saying, what is the best college? It depends, is a well less blue. It is under Gauss Markov.

Greg 47:57

And when those conditions don't hold, then what do we do have episode two or

Patrick 48:01

three or four? Where we say, what do we do? What if it is not homeless? gastic? What does fit is not independent? What if you don't have a continuous dependent variable, all of these things, say, Oh, we violated Calvin Ball, it is no longer our best Linear Unbiased Estimator. Then we get into this really interesting topographical space, which is well, is it close enough for government work? Or do we need to look ourselves in the mirror and say, Now I can't use this anymore, I got to move to something else. And so that's I think a topic for future conversations is when do we tap out and move to other things? But before we do that, I think we should all take a moment of deep and sincere respect for least squares estimation, because it is profoundly important and interesting, and cool. And rage Hulk and sassy Hulk will still give you that same be one that's the bottom of that parabola. Sexy Hulk doesn't much sexier than rage. Ciocca does. What I am rage Hulk, who can't benchpress his own bodyweight?

Greg 49:12

That's because you need to get physical, physical, you need to know everything that we've talked about so far, has fallen under the heading of ordinary least squares, which is in this linear family with certain assumptions that are being made. What is beyond the ordinary Patrick,

Patrick 49:34

who will What about something like to stage least squares, or partial least squares? All of these things are variants on the same theme, but allow us to do things that may be ordinary least squares does not. But it's still that same foundation. And I think we should address those on a later episode.

Greg 49:57

Let's do that. Thanks very much, everybody. Thank you for tolerating us and we're looking forward to season four.

Patrick 50:03

Take care, everybody. All right, bye bye,

Greg 50:05

see ya. 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 avoid class prep for the new school year. You can also follow us on Twitter where we are at quantity food pod and visit our website [quantity food pod.org](http://quantityfoodpod.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 cool back to school quantity merch, like shirts, mugs, stickers and spiral notebooks from [Red bubble.com](http://Redbubble.com) Where All proceeds go to donorschoose.org to help support low income schools. You've been listening to quantitative the podcast that not even fishers z prime transformation could make normal. Today's episode has been sponsored by US college ranking reports always handy to have around in case you need a random number generator. And by asymptotes your functions way of asking Are we there yet? Are we there yet? And finally by RI-CLPM. Believe it or not, this is not a Charles Dickens character. (giggles) Oh Mr. RickleIPim! This is most definitely not NPR. You have been listening to Quantitude: The top ranked quantitative methods podcast at East Chapel Hill High School. But I like psychologists Four beers!