Episode 109: Patrick Aldrich

# KL: Katie Linder

# MEDS: Mary Ellen Dello Stritto

# PA: Patrick Aldrich

# KL: You’re listening to “Research in Action”: episode one hundred and nine.

# [intro music]

# Segment 1:

# KL: Welcome to “Research in Action,” a weekly podcast where you can hear about topics and issues related to research in higher education from experts across a range of disciplines. Along with every episode, we post show notes with links to resources mentioned in the episode, full transcript, and an instructor guide for incorporating the episode into your courses. Visit our website at ecampus.oregonstate.edu/podcast to find all of these resources.

**MEDS:** I am your guest host, Dr. Mary Ellen Dello Stritto. I’m pleased to bring you the second episode in our periodic series focusing on quantitative methodology and statistics.

In this episode, I’m joined by Patrick Aldrich. Patrick received his bachelor’s degree in Wildlife biology and a minor in Entomology from the University of California, Davis. After graduation, he spent 5 years in various field biology positions, studying a wide array subjects from Bowerbird mating systems in Australia to integrated pest management of ground squirrels in Northern California. He subsequently decided to return to school to pursue a PhD at the University of Hawaii, Manoa, where he studied the spatio-temporal variation of pollination networks in Hawaiian tropical dry forests. Following his graduate work, he was the project director for a project that used spatial analyses to study the random correspondence of fingerprint patterns. Through his work, he has acquired extensive experience in biostatistics. He is currently the data manager and statistician for the Oregon Quality Rating and Improvement System for early childhood and other projects at The Research Institute at Western Oregon University. He continues to apply parametric, non-parametric and likelihood methodologies to analyze various datasets associated with early childhood and educational research.

**MEDS**: Thank you for joining me today, Patrick.

**PA**: Happy to be here.

**MEDS**: Great! So we are going to begin talking about some basics of statistical testing. Before we get into discussions of types of statistical tests, I would like to first talk about some basic statistical measurements. That most people have heard of. Um so let’s begin to talk about the mean and the median. So can you define these for us and then we will talk a little bit about their differences.

**PA**: Sure, so we you are looking at data about a population. There is multiple ways you want to measure that population. One is kind of the center of that distribution of that you are measuring that is in the population and then kind of the spread. So the center of the distribution can be measured in two different ways the median or the mean. Median is quite simply the mid-point of the distribution. So its half way between all the observations, it’s 50% below it and 50% above it. The mean on the other hand is the average as many people understand it. It’s a mathematical formula that you sum up all the observations and divide it by the total number of observations. And these are ways to kind of measure the center point of your data.

The important differences between this is if you have kind of a symmetric distribution these two measurements should be the same. If you have distributions that are not symmetric have long tails or highly skewed, that especially where you have data kind of bunched up at the bottom or at the top, and the tails are going out quite far. You can think of income or cost of houses are a great example of this where you have a lot kind of down at the bottom then these long kind of trailing ones that go up and up to millions of dollars versus the ones around the hundred or two hundred thousand range. If we are talking about houses. And what happens with the mean, because it is calculated it has those numbers at the higher end of the distribution can then start pulling they influence that number and it pulls the mean up. So it isn’t always quite accurately depicting what that distribution is. So the nice thing about the median is its robust and it doesn’t alter based on the actual number that the data point is. It’s just where it falls within the entire sample.

**MEDS**: Good. I think the example of housing prices really works there. Right, it’s one of situations where one of those statistics is better than the other. So with housing prices we are more likely to look at a median, is that right?

**PA**: Right, right. You usual here median home price as the main statistic as kind of where the center of that data is when it’s reported out in the media or other places.

**MEDS**: Right. And if we were to add in all of the millions of dollars of mansions in the world. That would really skew that mean or that average home price. So that’s why we use the median. [**PA:** Exactly, exactly]. Great and I think a lot of people are fairly familiar with the mean or the average or the median. But it is good to think about how they relate to one another. So I appreciate that. Any other things you want to talk about in terms of their differences?

**PA**: Um, that’s one of the main ones. Like I said if you have a nice symmetric distribution they should be approximately the same. Because it’s a great way they both measure that kind of midpoint of your distribution.

**MEDS**: Okay great. So now next I would like to talk about parametric and non-parametric statistical tests. So based on what we already talked about with some of these basic statistical measurements. What distinguishes these two types of tests? And then we will talk about what is there most appropriate uses.

**PA**: The easiest way to talk about it is that we have to define parametric tests and what a non-parametric test is. But the non-parametric test oddly is kind of a difficult thing to define. The definitions are all over the board in literature and things like that. So I think it is almost more easy to define it for what it’s not, and so what it’s not is a parametric statistic. A parametric statistic is a statistics where the samples come from a population where the sampling distribution is based on a fixed set of parameters. And what I mean by parameters are going back to defining the center and defining the spread. So like, for example, normal distribution is defined by its mean and standard deviation. And so a non-parametric test is often described as a distribution-free static. What that really means it doesn’t mean that the data is distributed in only way it just means we are not defining it by a specific distribution. Okay, so there is multiple ways to distinguish them.

**MEDS:** Okay.

**PA:** Like I said, parametric tests is based on, on parameters of a known distribution and usual a majority of tests that most people learn in there intro stats courses and things like that are parametric tests. So these are the T-tests your ANOVA, your regressions. And they are all based around analyzing the differences between means. And so a non-parametric tests, a lot of the ones that say are the analogs for say a T-test which is looking at two samples, is instead of looking at the mean is looking at the difference between the medians. And the reason why you look at that is because again the median its self is quite robust to any outliers or anything like that. So if you have data that is relatively small sample size and shows skew, these methods are much better at analyzing the differences between two samples than a parametric test would be.

**MEDS**: That really helps to kind of think about you know the traditional statistics that we learn in basic math classes as parametric. Then comparing parametric to them. That’s helpful; thank you. So what are the most appropriate uses then for these types of tests? You mentioned specific stats.

**PA**: Right. So uh there is a wide variety of non-parametric tests. And they have their parametric analogy’s so going back. We have multiple versions of a T test we have a one sample T, a paired T, and two sample T tests. So you have analytical analogs for that within the non-parametric world. You also have analogs for correlation and the parametric correlation is Pearson’s correlation. And you have a couple other non-parametric ones such as Spearman’s Rank correlation or Kendalls Tuo. Then you can even get up into there is regressions there is parametric, non-parametric, ANOVAs, that allow analysis of data that violates the underlying assumptions of that parametric test or model.

And so that’s really the key we are getting at there are a lot of assumptions built into parametric tests. And if your data fits those assumptions. You have a decent data size, you have data that is normally or approximately distributed then parametric tests are by far in a way the more powerful versions of these analyses. In a lot of times there is data that doesn’t fit those categories. And so non-parametric tests allow you to analyze data and get some conclusions about what might be going on without violating any of those assumptions and thus coming to erroneous conclusions.

**MEDS**: So can you give an example of a case in which you would typically go to a T-test comparing a couple of groups and seeing if there is a difference. What would be an example of a time when the T-test wouldn’t be quite as appropriate and you would use a non-parametric instead?

**PA**: Right. So for example if you have lets say a two sample T-test. If you have a small sample size and the data is highly skewed so you have heavy tails in the distribution, large outliers, things like that. That would be pulling your mean away from its center. The T-test, especially with a small sample size, it starts to lose power in its ability to analyze what is going on within that framework. And so you have things like sign based non-parametric tests and rank based non-parametric tests. Which then would allow you to actual analyze that data, see what might be going on. Even when that data is not following a normal distribution.

**MEDS**: So when you say the data is skewed, do you mean that there is a lot of scores that are similar to one another on one end or the other on the distribution?

**PA**: Yeah so, you might have a lot that are at the low end then you have very few that are at a higher end of the data. And so you have a lot let say your data is height and so you have a lot around the person is five foot four and 20 people are kind of in that range. Then you have someone like Shaquille O’Neal at what was he [soft laugh]… seven foot eight or something like that. So you have this one guy way over here who is pulling that mean up. And say you only have a sample of ten people that, that data point while in actual and true being is highly skewing you data. And so a better way to look at that distribution would be to use a non-parametric test, opposed to a parametric test.

**MEDS**: And especially because you only have one of those basketball players [**PA:** Right right] in a small sample. Okay okay [**PA:** Yes, exactly] that makes sense. That’s really helpful.

So, this kind of helps looking at examples like that really helps us to answer the next question which are what are the best practices for deciding to use these non-parametric tests? I think you have mentioned some of them. Are there other best practices?

**PA:** Um really the two...something I harp on quite a bit. Is always look at your data first. Kind of have an understanding of what your data looks like and kind of the role sample sizes paly into your data. Um for parametric tests you want a relatively robust sample size. Meaning something around 30 at least 30 individuals. But depending on how your data is distributed you many have less. If you have a nice kind of bell shape curve in the data set, something around 20 might be okay. But if you have a much larger data set, really symmetric a parametric test is great and works wonders. And you have a lot of nice underpinnings of the assumptions and everything built into the data.

It’s when you have smaller sample sizes. This can be maybe you are studying a rare population. I use to work on rare plants and rare insects. And so the numbers just didn’t exist in the wild. So I might have capture a, when I was studying bees in Hawaii. I might capture a bee species once in two years. That doesn’t mean that, that data point is bad it just means that it is extremely rare species and rare population. So I am just not going to be able to capture that single very often. So you need some ways to look at what is going on in the world without having these assumptions or need of big data or big sample sizes.

**MEDS**: Great. I think that is important because there is a lot of things that we study that you know maybe not have large samples. And that I think its important to know about non-parametric statistical tests. That they exist and because they are going to fill that need like you said they can maybe detect those differences in ways that those parametric tests cannot. Am I right? [**PA:** Right, right] Great.

And so you mentioned here a little bit about sample size. And there is the other statistical concept of power. And I thought maybe you could talk about what power is and how power relates in this context.

**PA**: So, generally speaking again going if you have a good sample size and the data fits kind of the parametric assumption. The parametric test is always going to be more powerful than a non-parametric test. Its when you start violating those assumptions of that parametric test that test starts to lose power, so either you are losing. Again you are losing sample size its getting smaller sample size. Your data is highly skewed that actually some of the more simple tests the assigned tests. Going back to the T-test. If the data is normally distributed and has a large sample size a T-test is more powerful. But if you have a small sample size with heavy tails again housing prices or something like that the sign based test becomes more powerful at that in these situations. And so they are much more unique in their needs. So in the needs to use them where the data fits some very specific kind of I was about to same parameters, but that would have been kind of confusing. But the data fits in very specific kind of details that then say. You should maybe start thinking about a non-parametric test, because in this situation this is actually a more power full test. It may show an alternative hypothesis than the parametric test would.

**MEDS**: And when we talk about power that maybe maybe this is a good moment to kind of define what we mean by statistical power. Its that ability to detect [PA: yes] a potential difference. Do you want to mention anything about that?

**PA**: Well yes. And you defined it perfectly that is exactly what we are trying to do is. How strong is our ability to say that this phenomenon we are looking at is true.

**MEDS**: Great and that’s the basis of our statistical testing we are trying to do. And hypothesis testing. Correct? [**PA:** Yes, yes] Okay great. And what about this issue of generalizability?

**PA:** And so, that is the interesting thing about non-parametric tests. Depending on which one they can be pretty generalizable. Generally speaking, especially if you are dealing with smaller sample sizes, I always want to caution people about generalizing beyond their data. Because while you can be finding a difference within your data set. Say you only have ten individuals and you use a non-parametric test. I would caution against generalizing beyond your data about to the general population about this phenomenon. But if, say, you do a study and ten other people independently do the study with ten individuals. So your sample is ten people in a parametric test. And each one keeps showing significance different within these samples, that starts to demonstrate a broader more robust demonstration that this phenomenon might actual be more generalizable to the population.

**MEDS**: Ah, that is really interesting. And a previous guest we talked about actual having the full population and that full population might be a really small group. So, this would apply in that situation as well. [**PA:** Yes, yes! Definitely.] Okay great, fantastic!

So we are going to take a brief break here and when we come back we will hear more from Patrick about using non-parametric tests.

[Music plays in background]

The research and action podcast is one of many innovative project that are a result of the hard work of a great team of people at Oregon State University Ecampus. A national ranked leader in online education. Oregon State Ecampus delivers nearly 50 degrees and programs online. Making an Oregon State University education accessible to learners worldwide. Learn more at ecampus.oregonstate.edu

# Segment 2:

**MEDS:** So we have been talking about sample sizes and generalizability. And these are all good issues to think about in terms of actually using these non-parametric tests. So I would like to have you talk about how you have used non-parametric tests in your own research.

**PA:** Okay so, I have used I use them quite a bit. And again going back to uh a lot of times its due to I am choosing them due to the kind of distribution of the data or the sample size that I have. And so I have used them in um looking at sometimes community. In the community ecology world. Looking at how communities change based on different regimes. And so, one of the tests that I used which is called perma NOVA, which is permutational analysis of variants. Or permutational multi analysis of variance.

And its used to look at basically communities, so lets say a bird community. I am going to kind of use a kind of example here. Say you have a forest like an old growth forest and you have three replicates of that old grove forest. You have a clear-cut, and three replicates of the clear cut and you have a 50-year-old regrowth forest. So you basically have three types of three forest types. And three replicates of it, and you go out and you do samples of the bird communities. And if the bird community changes within these types, these different forest types. Now while you are out collecting and doing bird observations you are doing point count or transaction or whatever. You are collecting a lot of data about the different bird species. But really when you think about it you are describing that community and kind of that diversity within each one of those groups. You only have now a sample size of nine. Like I said you have three of the old growth, three of the clear-cut and three of the 50-year-old forest. And then on top of that especially with in community ecology. When you look at a region pool of species that could possibly be there, each time you go out you are not going to get every single species that could possibly be in that forest type. You main only get say three or four and those three or four might be highly abundant and then the rest are maybe only one or two. Really you get lots of zeros in your data, because they just don’t exist. At that point in time when you were out looking. And so, this ANOVA this non-parametric nova was built by Marti Anderson in 2001. As a way to look at how these how to measure these communities like I said even though you have a lot of bird counts. The diversity indices are only measuring in each forest type. So your overall *n* is nine. And so what this test did, does is it looks at how these indices are dispersed within a non-euclidean space. And so, it’s a way to measure how similar or how distance these communities are in forest types.

So, I have used this actual in bird data, which is way I was kind of using that analysis. A friend of mine Dr. Balty right now is using it to look at the flora microbiome of invasive snails from his yard. And looking at how the introduction of an antibiotic will change those community types across a pre and post regime. And so you can see how a typical nova looking at means wouldn’t work for this type of data. But it is a way to analyze these multi variant data systems and see how differences change through space. Or through individuals

**MEDS**: So could can you example what you mean by non-euclidean space? That might be a term people don’t know.

**PA**: [Laughs] Yes, so this whole test is looking at how things map out in a spatial way. And euclidean space is what you really think of your x, y coordinates. So if you move one space over that is a measurable distance. So, you could say and if you put it out on a landscape it;s like if you walk three feet in one direction that is a measurable space. And I could say that you and I are three feet away from each other. Non-euclidean space is spatial difference that is not linear in that fashion. Right so it’s there it’s not that the physical difference say that you measure the two points on your screen, do not equal there actual distance. And so it’s a more conceptual mathematical way of looking at how things are related to each other even though the distance its self is not a straight measurement. Does that make sense?

**MEDS**: Interesting. So lets go back to the birds for a minute. So you were talking about the birds, so talk a little bit more about what that statistic is actually telling you.

**PA:** So it’s what it does it looks at say going back to let’s say those three forest types [**MEDS:** Right]. So what you would do is take all those bird species and you build a species diversity index and that index is kind of a flat number. You usually have taken hundreds of different numbers and you made it into one single number. And that tells you kind of what the species composition of that one forest is. And what you are kind of hoping to see in this type of analysis is the three indices from the old growth forest will cluster closer to each other. And the three say from the clear cut cluster with each other, but are away from the old growth. And hopefully maybe that 50- year-old forest kind of falls somewhere in between. So you can see that would be like the perfect way for the data to fall out. That there is some obvious relationship between no tree, full tree, and kind of the forest isn’t quite at its full growth yet, but the community is kind of in between those two things in that hypothetical.

**MEDS**: Okay that really helps kind of putting it in those terms. So you mentioned how this is um you know analogous in parametric tests to the ANOVA. Right? [**PA:** Right, yeah] And so with the ANOVA we are kind of looking at potentially you know the differences between three means [**PA:** Exactly]. And in this case a little more complex. Am I right?

**PA**: Right right. So, it’s a instead of comparing means of the groups you are comparing and looking at how these group kind of cluster with each other and how far apart they are in some space.

**MEDS**: Interesting okay, very interesting then you mentioned the other study about the snails. Can you go further?

**PA**: Yeah he just started doing it. So we were talking about this the other day. And I find it kind of interesting. He’s looking at gut flora of snails. So they took basically a gut sample and ran genetics sequences across all the things that were sitting in the gut. And you get spikes within the sequence. And those spikes kind of tell you what quote end quote species of bacteria are in there. So you have bacteria in snails prior to getting antibiotics. And then you can look at the spikes in those bacteria after antibiotics and how those communities might change because of that one treatment that happened to the snail.

**MEDS**: Interesting. And where is the non-parametric benefit here? So why is that the better test in this case?

**PA**: So again it’s going back it has the same issues with because it is community level data. You have this idea that there could be hundreds of species that could be there. But say in one individual you have 200 of species A and 20 of species B and then bunches of zeros. And so the data its self because you have all this quote end quote missing data of zeros. Makes it very difficult for to be used in parametric tests. So that’s we didn’t really go over some of the assumptions of the parametric tests. But that’s one especially with nova is missing data causes issues in calculations and so this is a way to get around that issue.

**MEDS**: Great that’s very interesting. And two really interesting examples in ecology, right? [**PA:** Yeah] So, do you have any examples of how you use non-parametric tests in program evaluation? Since you are a program evaluator as well.

**PA**: Yes, I have! I have used I was using. So, I work on the Oregon’s Quality Rating and Improvement System for early childhood. And early on we were looking at using an observation tool within the childcare programs. That looked at the environment known as the ECERS.

And we have data from ECERS and then we had the rating we gave programs. And we only had 14 observations of programs and there is a lot of long story why that happened. But so we had 14 observations and we had 14 programs that got ratings and so I wanted to see, because I couldn’t. I wanted to see if there was a correlation between the how programs landed on this independent scale known as the observation and our rating system. To see if you would except or at least you would hope that programs that we ranked as higher quality came out as higher quality within the ECERS.

So because it was a small sample of 14. I couldn’t really run a parametric correlation. So I used what is known as Kendall Tau as a way to test whether there was a relationship between those two detriments.

**MEDS**: Great, and that’s one that I think maybe people have heard of. Of non-parametric test folks have heard of so kendall pow is one. Are there other tests that you can mention that other we may have heard of.

**PA**: Yeah, so I the Mann-Whitney *U*, which is a test built for looking at it’s a non-parametric test used for looking at basically similar type data to T-tests. I have mentioned the sign test earlier. And there’s its equivalent and that’s usually for kind of a one sample data or paired data. If you have more than one sample so if you have two or three resamples its equivalent test would be known as mood’s median. So it’s a median test. And these are kind of like the lower ranked, I don’t want to say lower rank, but the less powerful tests the mood’s median and the sign test. Because there is almost no assumptions built into the models beyond your data is independent of each other.

And so they are great because they have no assumptions, but because of the lack of assumptions they are usually less powerful. Unless as I said earlier they have small samples and heavy tails.

**MEDS**: Okay, good. And perhaps maybe you could share with us. If people want to learn more about non-parametric tests. You could let us know what kind of resources are out there and we can put those into our show notes.

**PA**: Yes, I can defiantly do that!

**MEDS**: Great fantastic! So is there anything else related to non-parametric tests or statistical testing that you would like to add to our conversation?

**PA**: Eh, I think it just goes back to what I said earlier is I hope people who are doing research and analysis take the time and do that basic descriptive kind of work. Of actually thinking about looking at your data, thinking about the distribution that your data takes, and from there that helps you decide which test might be most appropriate for what you are trying to do.

**MEDS**: Absolutely and we here at the research unit have always said “Know your data”. You know really look at your data and really look at the distributions and really know what your data looks like before you make those kinds of decisions. So, I second that completely. [**PA:** Yeah totally.]

Anything else you want to mention?

**PA**: Ah no I think that is it.

**MEDS**: Alright, well thank you very much.

**PA**: Thank you.

**MEDS:** And thanks for talking to me today. And thanks also to our listeners for joining us for this week’s episode of Research in Action. I am Mary Ellen Dello Stritto and join us next week for another episode.

[Music plays]

# Show notes with links to resources mentioned in the episode, a full transcript, and an instructor’s guide for incorporating the episode into your courses, can be found at the show’s website at [ecampus.oregonstate.edu/podcast](http://www.ecampus.oregonstate.edu/podcast).

# There are several ways to connect with the “Research in Action” podcast. Visit the website to post a comment about a specific episode, suggest a future guest, or ask a question that could be featured in a future episode. Email us at riapodcast@oregonstate.edu. You can also offer feedback about “Research in Action” episodes or share research-related resources by contacting the Research in Action podcast via Twitter @RIA\_podcast. Finally, you can call the “Research in Action” voicemail line at 541-737-1111 to ask a question or leave a comment. If you listen to the podcast via iTunes, please consider leaving us a review.

# The “Research in Action” podcast is a resource funded by Oregon State University Ecampus – ranked one of the nation’s best providers of online education with more than 50 degree programs and over 1,000 classes online. Learn more about Ecampus by visiting ecampus.oregonstate.edu. This podcast is produced by the phenomenal Ecampus Multimedia team.

# “Research in Action” transcripts are sometimes created on a rush deadline and accuracy may vary. Please be aware that the authoritative record of the “Research in Action” podcast is the audio.