



Statistics By Jim

Making statistics intuitive

[Basics](#)[Hypothesis Testing](#)[Regression](#)[ANOVA](#)[Time Series](#)[Fun](#)[Glossary](#)[Blog](#)[My Store](#)

Confidence Intervals vs Prediction Intervals vs Tolerance Intervals

By [Jim Frost](#) — [32 Comments](#)

Intervals are estimation methods in statistics that use sample data to produce ranges of values that are likely to contain the population value of interest. In contrast, point estimates are single value estimates of a population value. Of the different types of statistical intervals, confidence intervals are the most well-known. However, certain kinds of analyses and situations call for other types of ranges that provide different information.

In this post, I'll compare confidence intervals, prediction intervals, and tolerance intervals, so you'll know when to use each type. I'll include an example of each type of range to make them easier to understand!

What are Confidence Intervals?

Confidence interval calculations take sample data and produce a range of values that likely contains the population parameter that you are interested in. For example, the confidence interval of the mean [9 11] suggests that the population mean is likely to be between 9 and 11.

Different random samples drawn from the same population are liable to produce slightly different confidence intervals. If you collect numerous random samples from the

same population and calculate a confidence interval for each sample, a certain proportion of the ranges contain the population parameter. That percentage is the confidence level.

For example, a 95% confidence level indicates that if you draw 20 random samples from the same population, you'd expect 19 of the confidence intervals to include the population value. The confidence interval procedure is useful because it produces ranges that usually contain the parameter.

19 out of 20 intervals (95%) contain the population parameter.

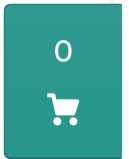
Use confidence intervals to produce ranges for all types of population parameters. A confidence interval for a population mean is probably the most common type, but you can also use these ranges for the standard deviation, proportions, rates of occurrence, regression coefficients, and the differences between populations.

Example of a Confidence Interval

Suppose that you randomly sample a product, measure the strength, and the 95% confidence interval is 100 – 120 units. You can be 95% confident that the mean strength of the entire population falls within this range. However, the 95% confidence level does not indicate that 95% of observations fall within this range. To draw that type of conclusion, we need to use a different kind of interval.

Here are some important considerations for confidence intervals.

- As you draw larger and larger random samples from the same population, the confidence intervals tend to become narrower.
- As you increase the confidence level for a given same sample, say from 95% to 99%, the range becomes wider. At first, this fact might seem counter-intuitive, but think about it. To have greater confidence that an interval contains the parameter, it makes sense that the range must become wider. Conversely, a narrower range is less likely to include the parameter, which lowers your confidence.



- A confidence interval for the mean says nothing about the dispersion of values around the mean.

For a graphical representation that makes these concepts more intuitive, please read blog post: [How Confidence Intervals and Confidence Levels Work](#).



What Are Prediction Intervals?

After you fit a regression model, you can obtain prediction intervals. These intervals predict the value of the dependent variable given specific settings of the independent variables. I'll cover two types of prediction intervals that provide different types of predictions.

Confidence interval of the prediction

A confidence interval of the prediction is a range that likely contains the mean value of the dependent variable given specific values of the independent variables. Like regular confidence intervals, these intervals provide a range for the population average. In this case, it's a particular population defined by the values of your independent variables. Similarly, these ranges don't tell you anything about the spread of the individual data points around the population mean.

Going back to our product strength example, let's assume it is a plastic product, and our independent variables are the plastic type (A or B) and the processing temperature. After we fit our model, the statistical software can produce the confidence interval of the prediction for specific settings.

We want to predict the mean strength for our product if we use plastic type A with a processing temperature of 125 degrees Celsius. The resulting confidence interval of the prediction is 140 – 150. These results indicate we can be 95% confident that the population defined by plastic type A and 125C has a mean that falls within this range. However, it provides no indication of the distribution of strength values for individual products.

Prediction interval

A prediction interval is a range that likely contains the value of the dependent variable for a single new observation given specific values of the independent variables. With this

type of interval, we're predicting ranges for individual observations rather than the mean value.

Let's use the same model and the same values that we used above. The statistical software produces a prediction interval of 130 – 160. We can be 95% confident that the strength of the next individual item produced using our settings will fall within this range.

There is greater uncertainty when you predict an individual value rather than the mean value. Consequently, a prediction interval is always wider than the confidence interval of the prediction.

We can predict the range for an individual observation, but we need a model. For more information, read my post about [using regression to make predictions](#).

What Are Tolerance Intervals?

Use tolerance intervals to answer the question, "what range of values covers X% of the population?" If you want to know the range where most values fall, use a tolerance interval.

A tolerance interval is a range that likely contains a specific proportion of a population. For example, you might want to know where 99% of the population falls for a particular characteristic. With tolerance intervals, we are specifically dealing with the spread of individual values around the mean.

To create a tolerance interval, you need to specify both the confidence level and the proportion. The confidence level is required because we're still working with samples and their inherent uncertainties.

For example, we want to create a tolerance interval where we'll be 95% confident that the interval contains 99% of the population.





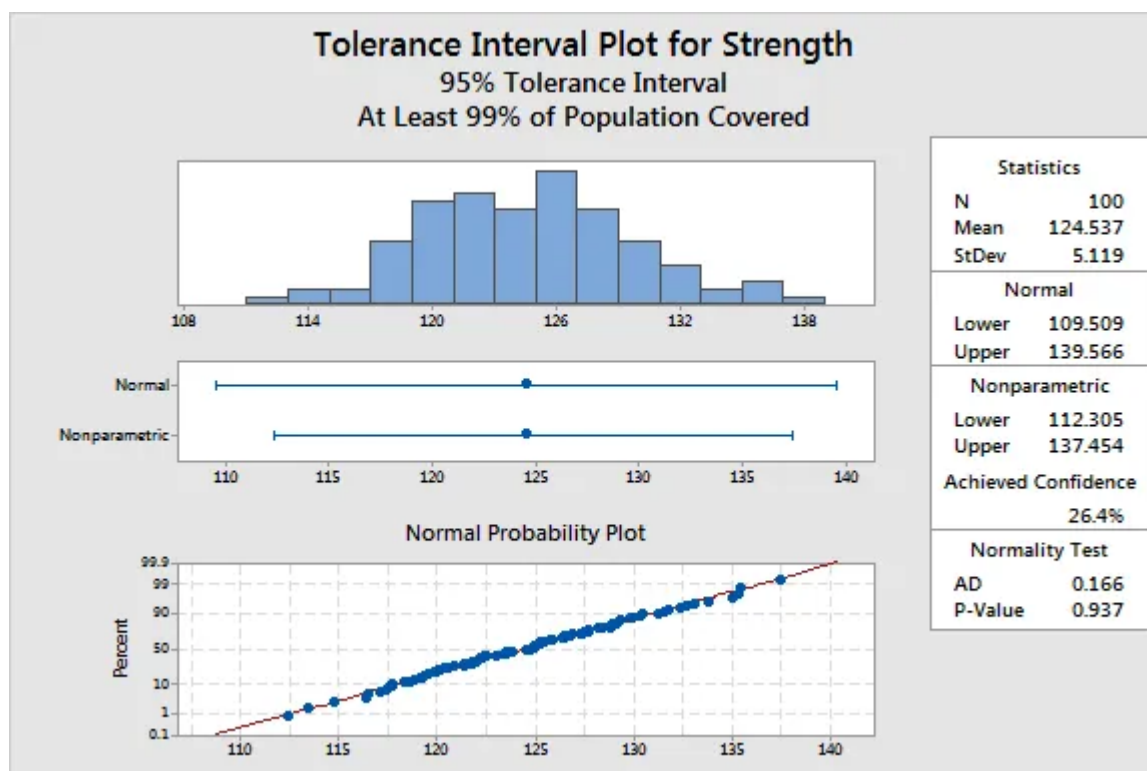
I think it's a lot easier to understand confidence intervals using an example!

Example of a tolerance interval

As the plastic manufacturer, we need to know the strength of our product. However, we need to know more than just the mean strength. It's important to understand the distribution of the individual values around the average.

For instance, the mean strength can be higher than our minimum requirement, which sounds great. However, if the spread around the average is too broad, too many products can fall below the minimum required strength.

To create a tolerance interval, we'll start by randomly sampling 100 plastic products and recording their strengths. Download the CSV data file: [Strength](#). Here is the statistical output for tolerance intervals.



Tolerance intervals are sensitive to the distribution of the data. In the output, the normality test indicates that our plastic strength data are normally distributed. Therefore, we'll use the Normal interval, which is 110—140 (rounded values). We can be 95% confident that at least 99% of all strength values for the product will be between 110 and 140.

How do we use these tolerance interval results? As the manufacturer, we need to compare the tolerance limits to our client's requirements. If our tolerance interval is broader than the requirements, our production process produces too many defects.

Tolerance Intervals vs Confidence Intervals

To help distinguish confidence intervals from tolerance intervals, here are some key differences.

A confidence interval estimates only the mean and the sampling error determines the width of a confidence interval. As the sample size approaches the whole population, the sample error decreases and the width of the CI approaches zero as it converges on the single value of the population mean.

A tolerance interval reflects the spread of values around the average. Both the sampling error and the dispersion of values in the entire population determine the widths of these ranges. As the sample size approaches the whole population, tolerance intervals don't converge on a zero width. Instead, they converge on the actual width of the population associated with the percentage you specify.



The width is based on percentiles. For example, to determine where 99% of the population lies, the software determines the data values that correspond to the 99.5th percentile and the 0.5th percentile ($99.5 - 0.5 = 99\%$ of the population). Tolerance interval calculations factor in the sampling error associated with the sample estimates of the percentiles.

For more information about percentiles, read my post: [Percentiles: Interpretations and Calculations](#).

Tolerance intervals can help you identify cases where excess variation can cause problems. Compare your requirements to the tolerance intervals to determine whether excessive variation is a problem for your study area.

Confidence intervals are the most well-known ranges in statistics. However, you might need to use a different type of range based on your specific needs.

Share this:

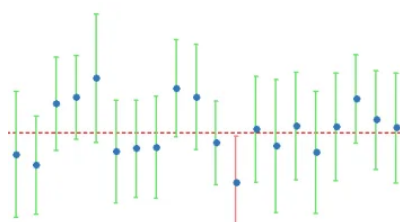
Share 251



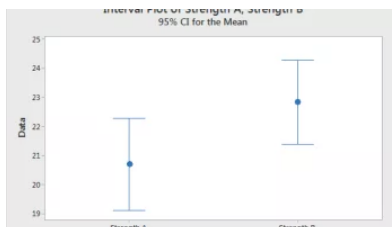
Tweet



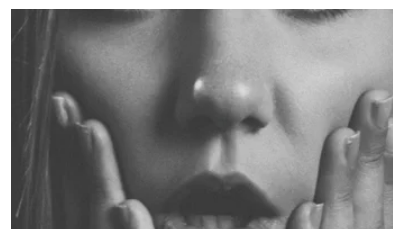
Related



How Hypothesis Tests Work:
Confidence Intervals and
Confidence Levels
In "Hypothesis Testing"



Using Confidence Intervals to
Compare Means
In "Hypothesis Testing"



Sample Statistics Are Always
Wrong (to Some Extent)!
In "Basics"

Filed Under: [Hypothesis Testing](#)Tagged With: [choosing analysis](#), [concept](#)

0



Comments

James Nishimuta says

July 29, 2020 at 7:07 pm

Thank you Jim for these articles, they've been very helpful.

I'm trying to figure out which method (tolerance interval, percentile analysis, regression analysis?) is appropriate for answering three questions. As an example, let's say the variable is measured brightness values for a thousand light bulbs with four different filament diameters. The questions I'm trying to answer are:

- 1- for a given filament diameter, what percentage of the population of light bulbs is expected to have at least a desired brightness? Eg for a wire diameter of 0.001", how many bulbs should have at least 600 lumens
- 2- for a desired percentile of bulbs, what is the expected measured brightness? Eg at least what brightness do 95% of bulbs with wire diameter of 0.001" have?
- 3- but the true objective here is to get a desired 95% of bulbs with at least 600 lumens, determine what the filament diameter should be?

For question 1, it seems that I should fit the brightness data for each filament diameter to a distribution function, plot it and select the desired x-value (lumens) to determine the percentile covered (which can be done in Minitab's Probability Distribution Plot graph function). So let's say that for 0.001" filament, 93.3% of the bulbs measured 600 lumens or more. (note, my data doesn't follow a standard distribution so I've been doing a Johnson transformation, do the percentile analysis, and inverse transform back).

For question 2, I used the tolerance interval function to determine what was the value that covers 95% of the population. When I do this however, the results are a bit different from simply looking at the distribution function and finding the value that yields 95%. But, I think this makes sense as there is uncertainty in that the population

fit, so it becomes more conservative (eg at least 95% of the bulbs measured at least 462 lumens)

For question 3, it seems like I need to do a regression analysis with prediction intervals. What I've been doing so far is just plotting the results for the tolerance interval as a function of filament diameter and, since my data seems to be simply linearly related, just interpolating the value that provides 95% tolerance interval at my desired output level. But seems like this a question that should be solved using regression, but not sure how to do this with a categorical independent variable instead of continuous.

★ Loading...

Reply

Jim Frost says

July 31, 2020 at 5:52 pm

Hi James,

If I'm understanding correctly, it seems to be you should use regression analysis/DOE with prediction intervals for all three questions.

I'm assuming that there is a relationship between filament diameters and lumens. So, you design an experiment that includes different filament diameters and include other factors as needed. Perform those experimental runs and measure the lumens. Fit the model. That provides you with ability to predict lumens for all the factors you include in the model, including filament diameter.

Using the prediction tools in Minitab that are based on your model, you can set the factors to specific values to obtain predictions. Perhaps filament thickness is either the only factor or the main one you're varying at this point. You can have Minitab construct a 95% lower bound for the prediction interval. You want to be sure that 95% of the new observations are above the bound. From there, try different values for the diameter thickness that produces a lower PI bound of 600 lumens. When you find a filament diameter that produces a lower bound of 600, that translates to 95% of the bulbs have lumens greater than 600. Additionally, the fitted value for that set of input values is the /expected/average brightness for that diameter (question 2).

I show a similar process to this in a post about [using prediction intervals to account for precision prediction](#). About halfway through that post I show an example of what I'm talking about. Although, instead of a lower bound, I need to use an upper bound. But it's the same idea. I don't want to go over a value while you don't want to go under a value.

I hope this helps! I think that approach should get you the answers you need.

★ Loading...

[Reply](#)

Julian says

June 25, 2020 at 2:32 pm

Hi Jim, what would be the process or what would change if I wanted to calculate prediction intervals for time series data? Would the independent variable become the time index or is this ignored? I would like to use this methodology for business forecasting essentially for predicting future sales demand.

★ Loading...

[Reply](#)

S K G says

June 13, 2020 at 2:11 am

0



What would be the confidence level of a point forecast? When would we be 100% confident?

★ Loading...

[Reply](#)

Jim Frost says

June 13, 2020 at 11:48 pm

Hi, if you mean a point estimate, you can't calculate a confidence level for an individual value—only for a range of values. To be 100% confident, you'd need a confidence interval that extends from $-\infty$ to $+\infty$!

★ Loading...

[Reply](#)

Sarah says

March 17, 2020 at 6:32 pm



Hi Jim,

How would you tell if a prediction interval is useful to your data? What would be benefits of using a prediction interval instead of a confidence interval or tolerance interval?

Looking forward to hearing from you! Your blog seems great!

THanks,
Sarah

★ Loading...

[Reply](#)

Jim Frost says

March 17, 2020 at 10:40 pm

Hi Sarah,

This post compares all three of those intervals! I'm not sure what more you need? It should be all covered in this post. Let me know if there's something specific you need to know or if something isn't clear. I describe why you'd use each type of interval.

★ Loading...

[Reply](#)

Priscilla Branch says

December 28, 2018 at 4:19 pm

Hey Professor Jim :o)!!! Priscilla again. The comment sections still appear in ascending order when viewed on the Chrome browser and Internet Explorer. So, the oldest comments appear first. I'm still an enthusiastic fan and follower if your intent is to show the most recent comments last. But, it feels like it would be more convenient for more recent respondents to see your feedback on the site earlier than scrolling all the way down.

★ Loading...

[Reply](#)

Jim Frost says

December 28, 2018 at 4:36 pm

Hi Priscilla! Ah, the comments! Yes, that's a great point. I will look into that. You can also subscribe to the comments for a specific post and then get an email when one is posted. I'll check to see if I can change the display order of the comment. What you say makes sense!

★ Loading...

[Reply](#)

Jim Frost says

December 28, 2018 at 4:38 pm

Success! Thanks for the suggestion, Priscilla!

★ Loading...

[Reply](#)



Sanjeev Gadre says

August 27, 2018 at 6:28 am

Thank you Jim. I have been struggling for a while to understand the difference between the confidence interval and prediction interval in the context of a linear regression and your explanation, especially your example, finally got me to understand the difference. Thank you once again.

★ Loading...

[Reply](#)

Jim Frost says

August 27, 2018 at 2:16 pm

You're very welcome. I'm glad it was helpful!

★ Loading...

[Reply](#)

John says

June 26, 2018 at 8:56 am

Hi Jim,

Quick question. Say one calculates the prediction interval for a sample population at hand. The next real individual value (ascertained through measurement/assaying etc.) can fall within or outside the previously calculated prediction interval. Does one superimpose the new point on top of the previous interval (thereby excluding it from the prediction calculation itself), or does one recalculate the prediction interval with the newly observed value. The reason I ask is because the new value has the ability to skew/widen the prediction that is intended to flag it as aberrant.

Thanks!

★ Loading...

Reply

Jim Frost says

June 27, 2018 at 4:25 pm

Hi John,

I'm not sure if there is a standard approach to this issue or not. I'm guessing that each area has its own standards.

In a general sense, what you say is correct. If a point falls outside the PI, it'll tend to widen it if you include it in the dataset. If it falls closer to the fitted value, it'll tend to tighten the PI. The degree of the change depends on the sample size and where the point falls exactly. The key thing to determine about each data point is that you want to include only valid data and exclude outliers that aren't representative of the process. That determination can be time consuming because it can involve investigation. Consequently, I'd be leery about automatically including new data points into the analysis to recalculate the PIs.

I think part of the answer depends on why you want to recalculate the PIs? If you have a good model with an adequate sample size, you're not necessarily going to improve the PIs by adding more data. However, if you're not sure that you have a good model or an adequate sample size, you might have reason to do something like that, but you also have reason to question the PIs in the first place! In that case, I would generally recommend performing a follow-up study rather than

adding data points in a continuous fashion like that and redoing the analysis over and over. However, again, I'm not sure of any conventions that are used in the field related to this issue.

Also, you wouldn't want to include only those that fall outside the PI (I'm not sure if that was what you were suggesting).

I hope this helps at least somewhat!

★ Loading...

[Reply](#)

Perry Sisk says

March 19, 2018 at 10:29 am

Although the Wallis paper is good, I would recommend using the methodology outlined in Chapter 3 in the book entitled Statistical Tolerance Regions by Krishnamoorthy and Mathew in order to determine a tolerance interval for a linear regression model.

★ Loading...

[Reply](#)

Jim Frost says

March 19, 2018 at 1:24 pm



Thanks for the tip!

★ Loading...

[Reply](#)

Nikhil Rai says

March 4, 2018 at 1:32 pm

Hey jim,

you made my life simple with your clear explanation. Keep the good work up

★ Loading...

[Reply](#)

Jim Frost says

March 4, 2018 at 4:16 pm

Hi Nikhil, thanks so much for the nice comment! I'm glad you found it helpful!

★ Loading...

[Reply](#)

João Luciano Skrock says

January 2, 2018 at 1:24 pm

Hi Jim.

I am thinking about use Percentile Regression (PR) instead of Linear Regression (LR) to do capacity analysis of IT infrastructure (for example: % CPU utilization for a user demand).

I use the models to 95 confidence interval of LR trying to estimate the worst cases.

Are there similarities between the model 95th PR and the LR model at 95% of confidence interval?

I am a system analyst not a statistician 😊

Your blog is fantastic.

★ Loading...

Reply

Jim Frost says

January 2, 2018 at 2:34 pm

Hi João, thanks so much for the kind words. I really appreciate them!

Conceptually, performing a 95th percentile regression and a linear regression with a Prediction Interval with a 95% Upper Bound sound very similar. I don't have a lot of experience with percentile regression so I'm not positive about how close the math works out. A key difference between the two is that percentile regression can be better when the relationship between each predictor and the response varies based on the percentile. If the relationships change based on the percentile, linear regression can over or underestimate the outcome.

Often, the goal of percentile regression is to show how the predictors' effects changes for different percentiles. For example, if you're looking at X and Y, X might have a larger effect in lower percentiles of Y than in higher percentiles. You can even graph out how the parameter estimates change by percentile to get some very useful information. Bear in mind that there is a confidence interval associated with the predicted percentile values just like there is a confidence interval for the predicted mean value in linear regression. So, you won't obtain only a single number but both a point estimate and a CI.

The typical use for prediction intervals is to model where individual responses will fall based on a linear regression model. Note—you'd want to use prediction intervals rather than confidence intervals. A confidence interval in this context is the range that the mean response is likely to fall within—which is what you're specifically NOT interested in. You'd want to use prediction interval with a 95% upper bound. This approach does give you a single number for the upper bound. 95% of new observations should fall below this value.

You can certainly try both techniques and see how they compare. I'm more familiar with using the linear model with prediction interval approach. However, if the relationship between each predictor and the response varies based on the percentile, then percentile regression might be a better approach because you can produce a model for the percentile that you are most interested in. You can then input values for the predictors and produce a predicted value for the 95th percentile.

If you try both, I'd be very interested in how the results compare!

★ Loading...

[Reply](#)

Steve Maggio says

November 9, 2017 at 1:16 pm

Thank you, Jim, I'm using a regression analysis to define the relationship between pay rate vs a job responsibility rating. I then use the residuals analysis, which is based on prediction intervals I believe, to identify any outliers. Is that correct?

★ Loading...

[Reply](#)

Jim Frost says

November 9, 2017 at 3:53 pm



I'm not sure about that. I've never used PIs to detect outliers. In fact, outliers can cause PIs to be wider, which would make the outliers less detectable. However, if there was a point that was far outside of the PI, it should make you wonder about it!

There are various other diagnostics for identifying outliers amongst the residuals that I'm more familiar with. You can look at the standardized value of the residuals. Standardized residuals greater than 2 and less than -2 are usually considered large. Although, you'd expect about 5% of the residuals to be unusual using this criterion, so it's really just identifying candidates for further investigation. This approach sounds similar to a 95% PI approach. There are other measures such as H_i (assesses leverage) and Cook's D (leverage and standardized value).

My preferred method is using the good old fashioned residuals by fitted values plot, along with the other residual plots. When it comes to assessing residuals in general, I place more weight on plots than the various numeric measures. It's very easy to see unusual values on graphs. If you haven't already, check out my post about [residual plots](#). I suppose when it comes to justifying removing an outlier, it's nice to have those numbers as support rather than just a perception from a graph. Although, even when you have the numbers, you need more of a justification than just the numbers. You need a reason for why the data point is truly invalid. At some point, I need to write a blog post about outliers!

★ Loading...

[Reply](#)

Steve Maggio says

[November 9, 2017 at 10:36 am](#)

Jim, How do you create a tolerance interval around a regression model. Most software adds a prediction interval around a regression model.

★ Loading...

Reply

0



Jim Frost says

November 9, 2017 at 11:39 am

Hi Steve, a prediction interval is kind of like a tolerance interval in that you're getting down to the distribution of individual observations with a given probability. Additionally, the standard error of the regression gives you similar information—the spread of the residuals around the fitted values. I've never heard of tolerance intervals for regression models though. I'm not sure if that's possible—but you can get very similar information using the other tools. I did a quick Google search and there seems to be research in that area but I can't point you to a specific statistical package that can do that for you. Here is what seems to be the major paper on [tolerance intervals for linear regression](#).

★ Loading...

Reply

akroy1946 says

September 18, 2017 at 2:41 pm

really knowledgeable writeup.

★ Loading...

Reply

Jim Frost says

September 18, 2017 at 6:01 pm

Thank you!

★ Loading...

[Reply](#)

0



Bruno says

August 24, 2017 at 9:21 am

Good job Jim! People usually hate statistics due the lack of simple and direct explanations like yours. I will be following your blog and will point it to some of my friends who are in need for it!

★ Loading...

[Reply](#)

Jim Frost says

August 24, 2017 at 3:15 pm

Thank you very much, Bruno! I always strive to provide clear explanations. I don't think statistics has to be hard!

★ Loading...

[Reply](#)

audiggerblog says

July 2, 2017 at 11:25 pm

I really like your discussions and thoughts. As a geostatistician in mining, we tend to over complicate things and get caught up in the theoretical nuances of things. Your

thoughts are giving me some clear and concise ways to think about some of these ideas in stats. Thank you.

★ Loading...

[Reply](#)



Jim Frost says

July 2, 2017 at 11:47 pm

Thank you very much for your kind comments. I really appreciate them!

★ Loading...

[Reply](#)

Limbu M. Limbu says

June 26, 2017 at 5:09 am

19 out of 25 intervals (95%) contain the population parameter (from the diagram above). I think should be 19 out of 20 intervals (95%)

★ Loading...

[Reply](#)

Jim Frost says

June 26, 2017 at 9:25 am

Yes, indeed! Thank you! I'm off to make the edit now.

★ Loading...

[Reply](#)

Leave a Reply

0



Enter your comment here...

Meet Jim

I'll help you intuitively understand statistics by focusing on concepts and using plain English so you can concentrate on understanding your results.

[Read More...](#)

Buy My Introduction to Statistics eBook!

0



Introduction to Statistics: An Intuitive Guide [ebook]

\$9.00 USD

[Buy it now](#)

New! Buy My Hypothesis Testing eBook!

Hypothesis Testing: An Intuitive Guide [ebook]

\$14.00 USD

Buy it now

VISA

AMEX

JCB

DISCOVER

PayPal

0

Buy My Regression eBook!

Regression
Analysis: An
Intuitive Guide
[ebook]

\$14.00 USD

Buy it now

VISA

AMEX

JCB

DISCOVER

PayPal

Search this website

Subscribe by Email

0



Enter your email address to receive notifications of new posts by email.

Subscribe

I won't send you spam. Unsubscribe at any time.

Follow Me



Facebook



RSS Feed



Twitter

Popular

Latest

[How To Interpret R-squared in Regression Analysis](#)

[How to Interpret P-values and Coefficients in Regression Analysis](#)

[Measures of Central Tendency: Mean, Median, and Mode](#)

[Normal Distribution in Statistics](#)

[Multicollinearity in Regression Analysis: Problems, Detection, and Solutions](#)

[Understanding Interaction Effects in Statistics](#)

[How to Interpret the F-test of Overall Significance in Regression Analysis](#)

0



Recent Comments

Kicab Castaneda-Mendez on [The Monty Hall Problem: A Statistical Illusion](#)

Asiah Collinson on [Answering the Birthday Problem in Statistics](#)

Emily Stern on [Introduction to Bootstrapping in Statistics with an Example](#)

Jim Frost on [The Monty Hall Problem: A Statistical Illusion](#)

Jim Frost on [Normal Distribution in Statistics](#)

0

