Choosing Between a Nonparametric Test and a Parametric Test

Jim Frost (http://blog.minitab.com/blog/adventures-in-statistics) · 19 February, 2015

It's safe to say that most people who use statistics are more familiar with parametric analyses than nonparametric analyses. Nonparametric tests are also called distribution-free tests because they don't assume that your data follow a specific distribution.

You may have heard that you should use nonparametric tests when your data don't meet the assumptions of the parametric test, especially the assumption about normally distributed data. That sounds like a nice and straightforward way to choose, but there are additional considerations.

In this post, I'll help you determine when you should use a:

- Parametric analysis to test group means.
- Nonparametric analysis to test group medians.

In particular, I'll focus on an important reason to use nonparametric tests that I don't think gets mentioned often enough!

Hypothesis Tests of the Mean and Median

Nonparametric tests are like a parallel universe to parametric tests. The table shows related pairs of hypothesis tests (http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/hypothesis-tests/basics/hypothesis-tests-in-minitab/) that Minitab statistical software (http://www.minitab.com/en-us/products/minitab/features/) offers.

Parametric tests (means)	Nonparametric tests (medians)
1-sample t test	1-sample Sign, 1-sample Wilcoxon
2-sample t test	Mann-Whitney test
One-Way ANOVA	Kruskal-Wallis, Mood's median test
Factorial DOE with one factor and one blocking variable	Friedman test

Reasons to Use Parametric Tests

Reason 1: Parametric tests can perform well with skewed and nonnormal

distributions

This may be a surprise but parametric tests can perform well with continuous data that are nonnormal if you satisfy these sample size guidelines.

Parametric analyses	Sample size guidelines for nonnormal data
1-sample t test	Greater than 20
2-sample t test	Each group should be greater than 15
One-Way ANOVA	• If you have 2-9 groups, each group should be greater than 15. • If you have 10-12 groups, each group should be greater than 20.

Reason 2: Parametric tests can perform well when the spread of each group is different

While nonparametric tests don't assume that your data follow a normal distribution, they do have other assumptions that can be hard to meet. For nonparametric tests that compare groups, a common assumption is that the data for all groups must have the same spread (dispersion). If your groups have a different spread, the nonparametric tests might not provide valid results.

On the other hand, if you use the 2-sample t test or One-Way ANOVA, you can simply go to the **Options** subdialog and uncheck *Assume equal variances*. Voilà, you're good to go even when the groups have different spreads!

Reason 3: Statistical power

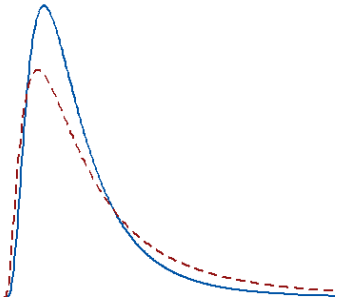
Parametric tests usually have more statistical power (http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/power-and-sample-size/what-is-power/) than nonparametric tests. Thus, you are more likely to detect a significant effect when one truly exists.

Reasons to Use Nonparametric Tests

Reason 1: Your area of study is better represented by the median

This is my favorite reason to use a nonparametric test and the one that isn't mentioned often enough!

The fact that you *can* perform a parametric test with nonnormal data doesn't imply that the mean is the best measure of the central tendency (http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/summary-statistics/measures-of-central-tendency/) for your data.



For example, the center of a skewed distribution, like income, can be better measured by the median where 50% are above the median and 50% are below. If you add a few billionaires to a sample, the mathematical mean increases greatly even though the income for the typical person doesn't change.

When your distribution is skewed enough, the mean is strongly affected by changes far out in the distribution's tail whereas the median continues to more closely reflect the center of the distribution. For these two distributions, a random sample of 100 from each distribution produces means that are significantly different, but medians that are not significantly different.

Two of my colleagues have written excellent blog posts that illustrate this point:

- Michelle Paret: Using the Mean in Data Analysis: It's Not Always a Slam-Dunk (http://blog.minitab.com/blog/michelle-paret/using-the-mean-its-not-always-a-slam-dunk)
- Redouane Kouiden: The Non-parametric Economy: What Does Average Actually Mean? (http://blog.minitab.com/blog/statistics-for-lean-six-sigma/the-non-parametric-economy-what-does-average-actually-mean)

Reason 2: You have a very small sample size

If you don't meet the sample size guidelines for the parametric tests and you are not confident that you have normally distributed data, you should use a nonparametric test. When you have a really small sample, you might not even be able to ascertain the distribution of your data because the distribution tests will lack sufficient power to provide meaningful results.

In this scenario, you're in a tough spot with no valid alternative. Nonparametric tests have less power to begin with and it's a double whammy when you add a small sample size on top of that!

Reason 3: You have ordinal data, ranked data, or outliers that you can't remove

Typical parametric tests can only assess continuous data and the results can be significantly affected by outliers. Conversely, some nonparametric tests can handle ordinal data, ranked data, and not be seriously affected by outliers. Be sure to check the assumptions for the nonparametric test because each one has its own data requirements.

Closing Thoughts

It's commonly thought that the need to choose between a parametric and nonparametric test occurs when your data fail to meet an assumption of the parametric test. This can be the case when you have both a small sample size and nonnormal data. However, other considerations often play a role because parametric tests can often handle nonnormal data. Conversely, nonparametric tests have strict assumptions that you can't disregard.

The decision often depends on whether the mean or median more accurately represents the center of your data's distribution.

- If the mean accurately represents the center of your distribution and your sample size is large enough, consider a parametric test because they are more powerful.
- If the median better represents the center of your distribution, consider the nonparametric test even when you have a large sample.

Finally, if you have a very small sample size, you might be stuck using a nonparametric test. Please, collect more data next time if it is at all possible! As you can see, the sample size guidelines aren't really that large. Your chance of detecting a significant effect when one exists can be very small when you have both a small sample size and you need to use a less efficient nonparametric test!

You Might Also Like:

A correspondence table for non parametric and parametric tests (http://blog.minitab.com/blog/applying-statistics-in-quality-projects/a-correspondence-table-for-non-parametric-and-parametric-tests)
Why You Should Use Non-parametric Tests when Analyzing Data with Outliers (http://blog.minitab.com/blog/applying-statistics-in-quality-projects/why-you-should-use-non-parametric-tests-when-analyzing-data-with-outliers)
Truth, Beauty, Nonparametrics & Symmetry Plots (http://blog.minitab.com/blog/statistics-and-quality-data-analysis/truth-beauty-nonparametrics-and-symmetry-plots)
The Non-parametric Economy: What Does Average Actually Mean? (http://blog.minitab.com/blog/statistics-for-lean-six-sigma/the-non-parametric-economy-what-does-average-actually-mean)

Comments

Zakariya • a month ago
Thank you for the info. Any reference for the minimum sample size required for Kruskal -Wallis test? Particularly, the second table indicates that the sample size per group to compare 2 to 9 groups should be greater than 15.

Thank you in advance!
• Reply • Share ›

Jim Frost At Minitab Mod Zakariya • a month ago
Hi Zakariya,

For the Kruskal-Wallis test, the sample size for each group should be at least 5. If a group has less than 5 observations, the p-value can be inaccurate. The number of groups for this test is not a consideration.

I hope this helps!
Jim
• Reply • Share ›

Zakariya Jim Frost At Minitab • a month ago
Hi Jim. Many thanks for your response. Any article or a book that I can refer to support the above info. Thank you once again!
• Reply • Share ›

Jim Frost At Minitab Mod Zakariya • a month ago
Hi again!

Here's the scoop! Under the null hypothesis, the distribution of the test statistic, H, can be approximated by a chi-square distribution with k - 1 degrees of freedom, where k is the number of groups. The approximation is reasonably accurate if no group has fewer than five observations.

Here are a couple of references:

E.L. Lehmann (1975). Nonparametrics: Statistical Methods Based on Ranks, Holden-Day.

M. Hollander and D.A. Wolfe (1973). Nonparametric Statistical Methods, John Wiley & Sons.

I hope this helps!
Jim

• Reply • Share ›

1_000_000 • a month ago

Thank you for the article! I 'd like to ask the following: I have 6 groups with 5-6 observations per group. I have run a Shapirow-Wilk test to examine distribution normality and was Ok (p>0.05 in all groups) However, after running a Levene's test, the result was p<0.01. How will I continue my analysis? With a parametric (ANOVA) or a non-parametric (Kruskal Wallis) test? Z(skewness) and Z(kurtosis) are all lower than 1.96.
Second question: Are both tests (ANOVA and Kruskal-Wallis) "reliable" enough in case I've got one group (out of 6) with 4 observations? The other groups have 5-6 observations.

• Reply • Share ›

Jim Frost At Minitab Mod ↗ 1_000_000 • a month ago

Hi,

One thing to keep in mind is that with such a small sample size per group, the power of the normality test can be very low. It's possible that your data are not normally distributed but the test lacks sufficient power to detect that departure from normality.

The Levene's test results indicates that the variances of the groups are not equal. That method doesn't assume normality. This test is a good choice for small samples like yours.

For your second question, I would say that given the characteristics of your data as you describe them, you don't have a large enough sample size per group to use ANOVA but you don't meet a requirement of the nonparametric analyses. Unfortunately, you're in a difficult spot.

Kruskal-Wallis or Mood's Median test would be the usual analysis that I'd recommend for such small sample sizes. However, both analyses assume that the shape of the distributions are the same between your groups. The fact that the Levene's test is significant suggests that this isn't the case.

If you had at least 15 observations per group, I'd recommend that you use One-Way ANOVA and

see more

• Reply • Share ›

1_000_000 ↗ Jim Frost At Minitab • a month ago

Thank you for your analytical answer !

• Reply • Share ›

Jim Frost At Minitab Mod ↗ 1_000_000 • a month ago

You're very welcome! Best of luck with your study!

Jim

• Reply • Share ›

Lanja Ibrahim • a month ago

thank you very much for this post, i have three groups each of 15 sample , i wanna test the difference in pre and post of 22 measurement and observation in all the withing groups and among groups to see which one had best response to treatment , but i test for normality by both Shapiro wilky and z value i see some are normally distributed and some are not , i am very confused which way to use , can you help? thank

are normally distributed and some are not , I am very confused which way to use , can you help? thank you

• Reply • Share ›

Jim Frost At Minitab Mod ↗ Lanja Ibrahim • a month ago

Hi Lanja,

It appears like you meet the sample size guidelines that I mention in this post so the normality of the groups may not be a concern, although I'm not certain how the repeated measures aspect of your design could change that aspect.

I suggest performing the analysis and assessing the normality of the residuals.

Best of luck with your analysis!
Jim

• Reply • Share ›

ng • 14 days ago

Hi Jim, thanks for this post. I would like to find out what is the rationale behind to support the sample size guidelines for nonnormal data? Thanks for your advice in advance

• Reply • Share ›

Jim Frost At Minitab Mod ↗ ng • 13 days ago

Hi, these sample size guidelines are based on simulation studies conducted here at Minitab. The goal of these particular simulation studies is to determine how well different hypothesis tests handle nonnormal data with various sample sizes.

Ideally, if you use a significance level of 0.05, your Type I error rate should be around 0.05. The simulation studies found that if you meet these sample size guidelines, the significance level matches the Type I error rate even when your data are not normally distributed. If you have smaller sample sizes, the actual error rate may not match the significance level if your data are not normally distributed.

If you're interested, you can read about these studies in our Technical Papers. If you want all of the real nitty gritty details, be sure to check out the appendices!

These guidelines were developed for use with Minitab's Assistant menu, which assesses the suitability of your data for a variety of conditions. However, the guidelines are applicable even if you're not using the Assistant.

Thanks for writing with the great question!

Who We Are

Minitab is the leading provider of software and services for quality improvement and statistics education. More than 90% of Fortune 100 companies use Minitab Statistical Software, our flagship product, and more students worldwide have used Minitab to learn statistics than any other package.

Minitab Inc. is a privately owned company headquartered in State College, Pennsylvania, with subsidiaries in the United Kingdom, France, and Australia. Our global network of representatives serves more than 40 countries around the world.

Visit Us at Minitab.com

Authors

Carly Barry (http://blog.minitab.com/blog/real-world-quality-improvement)
Patrick Runkel (http://blog.minitab.com/blog/statistics-and-quality-data-analysis)
Joel Smith (http://blog.minitab.com/blog/fun-with-statistics)
Kevin Rudy (http://blog.minitab.com/blog/the-statistics-game)

statistics-game)
Jim Frost (http://blog.minitab.com/blog/adventures-in-statistics)
Greg Fox (http://blog.minitab.com/blog/data-analysis-and-quality-improvement-and-stuff)
Eric Heckman (http://blog.minitab.com/blog/starting-out-with-statistical-software)
Dawn Keller (http://blog.minitab.com/blog/adventures-in-software-development)
()
Eston Martz (http://blog.minitab.com/blog/understanding-statistics)
()
Bruno Scibilia (http://blog.minitab.com/blog/applying-statistics-in-quality-projects)
Eduardo Santiago (http://blog.minitab.com/blog/understanding-statistics-and-its-application)
()
Cody Steele (http://blog.minitab.com/blog/statistics-and-quality-improvement)