Nonparametric Tolerance Intervals

Authored by:
Krista Wurscher,
Nicholas Jones,
Gina Sigler
12 Nov 2021

The goal of the STAT COE is to assist in developing rigorous, defensible test strategies to more effectively quantify and characterize system performance and provide information that reduces risk. This and other COE products are available at www.afit.edu/STAT.

STAT Center of Excellence
2950 Hobson Way – Wright-Patterson AFB, OH 45433

Table of Contents

Executive Summary

Many people are familiar with the concepts of confidence intervals or prediction intervals; however, tolerance intervals, while an extremely useful form of statistical intervals, are less widely known. Tolerance intervals quantify the proportion of a population that can be expected to meet some performance threshold, for a specified confidence level. The process of generating a tolerance interval depends on the assumed underlying distribution of the data, making them more difficult to calculate than either confidence or prediction intervals. When little information is known about the distribution of the data, a nonparametric tolerance interval may be the best approach. In this best practice, we address when to use nonparametric tolerance intervals, how to determine what sample size is needed to obtain a tolerance interval of a desired width, and how to find the bounds of the nonparametric tolerance interval once data has been collected. We also introduce an Excel-based tool created by the STAT COE for calculating nonparametric tolerance intervals and associated test sizes.

Keywords: tolerance interval, sample size, nonparametrics, STAT tools

Introduction to Tolerance Intervals

A tolerance interval is an interval that contains a specified proportion of the population with a certain degree of confidence. We use P to represent the proportion of the population we wish to contain within our interval and $1-\alpha$ to represent the confidence level. Thus, a $(P \cdot 100\%, 1-\alpha)$ tolerance interval refers to an interval we are $(1-\alpha)100\%$ confident encloses $P \cdot 100\%$ of the population.

Tolerance intervals are best used when answering questions regarding coverage (how much of the range of the possible space does a proportion of the population fall into?) (Meeker, Hahn, & Escobar, 2017). For example, you are concerned with the amount of sleep you get each night. You have taken measures to get more sleep and would like to know if you have reached your target of sleeping at least 7 hours a night 80% of the time. After collecting weeks-worth of data, you generate an 80% one-sided tolerance interval, expecting the values above the lower bound of your tolerance interval to include at least 80% of the distribution of hours of sleep. However, the bounds for this interval are based on a sample of nights, resulting in uncertainty in the estimate for the bound. This uncertainty is why tolerance intervals are also specified by a confidence level. You find an (80%, 95%) one-sided tolerance interval for the number of hours of sleep you get has a lower bound of 5.7 hours, concluding you are 95% confident that at least 80% of the time you sleep at least 5.7 hours. Recall that your goal was to be 95% confident that you sleep at least 7 hours 80% of the time. Comparing the tolerance interval with your success criteria, you can see that the lower bound falls below the threshold value. You cannot conclude that you have reached your target and should consider taking further steps to improve your sleep quantity.

For more information about the definition of a tolerance interval and the application of tolerance intervals, see Ortiz and Truett (2015). Likewise, more information on one-sided and two-sided tolerance intervals can be found in Tolerance Intervals Demystified by Splinter et al. (2020).

Nonparametric Approach

The process of calculating a tolerance interval depends on the underlying distribution of the data. Sometimes information is already known about the distribution, and in these cases it is appropriate to calculate a tolerance interval using methods specific to the distribution of your data. This knowledge may come from subject matter experts, or previous experiments on the same or similar system. For example, if you are working with data that you believe to be normally distributed, normal distribution tolerance interval calculation methods may be used (for information about tolerance interval methods using a normal distribution, see Splinter, Sigler, Harmon, and Kolsti, 2020). However, tolerance intervals are sensitive to deviations from the distributional assumption, so assumptions about the underlying distribution of the data should always be evaluated before proceeding with calculating the tolerance interval. If assumptions regarding the distribution of the data do not appear to be met, or if the underlying distribution is unknown, use a nonparametric tolerance interval (Ortiz and Truett, 2015).

Nonparametric approaches do not assume the data are from any particular distribution. So, in circumstances where the distributional assumption does not appear to be met or when the distribution of the data is very uncertain, it is best to use a nonparametric tolerance interval. The STAT COE has developed an Excel/VBA-based tool for calculating nonparametric tolerance intervals, which is available via the STAT COE website. While nonparametric tolerance intervals save us from needing to make assumptions about the distribution of the data, they do require more runs than parametric tolerance intervals to achieve the same tolerance and confidence levels (Kvam and Vidakovic, 2007). For this reason, it is important to understand the resource requirements before running an experiment with the intention of generating a nonparametric tolerance interval.

Determining Sample Size

If one of the desired goals of an experiment is to obtain a tolerance interval, it is important to establish what sample size is necessary to generate a tolerance interval with the desired coverage. The proportion of the population and the confidence level both affect the minimum sample size for tolerance intervals. If the conducted experiment has too small of a sample size, then the resulting tolerance interval at the desired confidence level will be very wide, making it much less useful. If the sample size is too large, then you will have used more resources on the experiment than necessary. The STAT COE nonparametric tolerance interval tool helps users determine the minimum number of runs necessary to obtain the desired $(P \cdot 100\%, 1 - \alpha\%)$ nonparametric tolerance interval within a specified margin of error.

Inputs

To obtain a minimum sample size for a nonparametric tolerance interval using the tool, there are four inputs that must be specified. Table 1 summarizes the tool inputs. The first two inputs are confidence and proportion ($1 - \alpha$ and P). The second two inputs, ε and α^*, are needed to provide acceptable margins of error. The parameter ε sets a margin of error profile for P, determining how accurate the tolerance interval can be. Tolerance intervals based on more precise values of ε will be more accurate, but will require more runs. If a system just barely fails to meet a tolerance interval for an imprecise ε, that system may actually be meeting the spec, but the tolerance interval is too imprecise to detect this. If a more precise test was done (which would require more runs), the system might be shown to actually be within tolerance. The value of ε must be between P and 1. α^* represents the risk that the true proportion included in the tolerance interval is greater than the desired proportion plus the allowable error, $P+\varepsilon$. Choosing smaller values of α^* means there is less risk of overshooting, approaching some minimum level set by the choice of ε. The proportion of values included in the tolerance interval should not exceed $P+\varepsilon$ with the confidence $1-\alpha^*$ (Meeker, Hahn, and Escobar, 2017, and Minitab, 2019).

Table 1: Inputs of the Nonparametric Tolerance Interval Tool

Name of Input	Description	Range of Common Values
$1-\alpha$	Confidence of nonparametric tolerance interval	0.70-0.99
P	Proportion of values to be within the bounds of the tolerance interval	0.70-0.99
ε	Margin of error on P, allows user to specify how close to the desired proportion is good enough	0.01-0.1
α^*	Risk of proportion of data in tolerance interval exceeding $P+\varepsilon$	0.05-0.10

Values of the confidence and proportion parameters for the tolerance interval should be chosen on the basis of system performance requirements. If 85% of engines produced must be capable of running continuously for 10 minutes at 4000 RPM with 90% confidence, then $1-\alpha=0.90$ and $P=0.85$. Choosing appropriate values of ε and α^* is more difficult and requires the consideration of SME judgement of what level of performance is required, what difference in performance is likely to be significant for the application, and the risk tolerance of both management and the end-user. ε may be thought of as a statistical noise floor for the proportion of population meeting the threshold. This noise floor should be chosen to be small enough to accurately reflect the desired proportion of the population, while being cognizant that smaller values of ε (i.e., very low noise tolerance) will require more runs to determine if the system is operating in tolerance. For a typical P value of 0.90, an appropriate ε might be 0.05. The value of α^* denotes the acceptable risk of being outside the allowable margin of error specified by ε, and should be chosen based on the program and users' risk tolerance. A typical value might range from 0.005-0.1. An example for how to enter the inputs to determine the sample size for a (85%, 90%) tolerance interval with $\varepsilon = .05$ and $\alpha^* = .10$ is given in Figure 1.

Some inputs or combinations of inputs are not possible. Table 2 lists all the constraints for the inputs. All the inputs for the nonparametric interval tool are proportions, so we require that all inputted values be between 0 and 1. The margin of error, ε, is additionally constrained because not only does ε need to be between 0 and 1, so does $P+\varepsilon$.

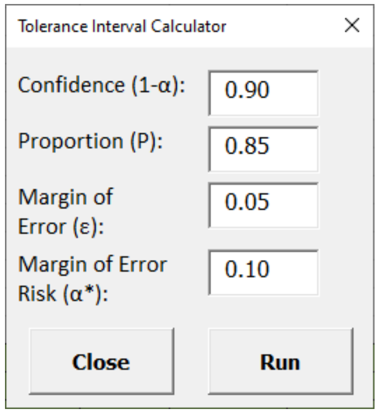

Figure 1: Tolerance Interval Tool Input Form

Table 2: Constraints of Inputs for the Nonparametric Tolerance Interval Tool

Constraint	Reason
$0 < 1 - \alpha < 1$	Probabilities must be between 0 and 1
$0 < P < 1$	Probabilities must be between 0 and 1
$0 < \varepsilon < 1 - P$	Both ε and $P + \varepsilon$ are probabilities and must be between 0 and 1
$0 < \alpha^* < 1$	Probabilities must be between 0 and 1

Outputs

The nonparametric tolerance interval tool uses a binary search function to identify a sample size for the tolerance subject to the convergence criteria specified below. The algorithm can calculate tolerance intervals of arbitrary precision, but due to the limitations of built-in Excel functions, there is a precision limit of approximately 6 significant figures for the tolerance interval proportion and confidence level. If a sample size is found that meets the specified criteria, then the nonparametric tolerance interval tool will provide the user with the minimum sample size, n, and the successful runs, k, needed to construct the tolerance interval. A sample output for the entry criteria in Figure 1 is shown in Figure 2.

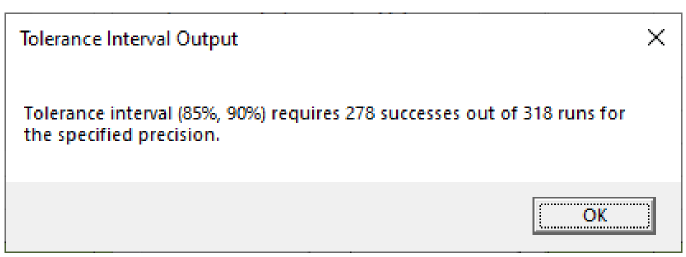

Figure 2: Example Tolerance Interval Tool Output Message

Within a sample of n data points, we would expect a tolerance interval to contain a subset, k, of those points, on average. The value of k represents the number of successful runs that we should expect to see out of n total runs required to say that the system meets the desired tolerance interval. In the engine example given earlier, if we conduct 318 engine runs for 10 minutes each, and at least 278 sustain 4000 RPMs in continuous operation, then we could say with 95% confidence that 85% of engines perform at the desired level (meeting an (85%, 90%) tolerance interval). Based on the definition of a tolerance interval, we would expect that $\frac{k}{n} \approx P$. If after n runs we observe fewer than k successes, then the system is likely (to the confidence level in the tolerance interval) not performing at the desired tolerance. The iterative process used by the nonparametric tolerance intervals tool selects n and k based on the following criteria:

1) The probability of having k or more points in the tolerance interval of proportion P is at least $1 - \alpha$
2) If the proportion of points the tolerance interval contains is $P + \varepsilon$, then the probability of having k or fewer points in the tolerance interval is less than α^*

The first criterion ensures, for a given n, that k is small enough that we are likely to have at least k values within a tolerance interval with a proportion P. The second criterion ensures, for a given n, that k is large enough that having less than k values within a tolerance interval with a proportion $P + \varepsilon$ is unlikely, that is, the probability of having less than k values within the interval is less than α^* (Meeker, Hahn, and Escobar, 2017, and Minitab, 2019).

The number of required successes is dependent on the number of runs executed. This means that the output of k applies only if the minimum number of runs provided by the nonparametric tolerance intervals tool is used. If data is collected from more than n runs, we should expect to see proportionally more successes to conclude that the tolerance interval is met. Figure 2 shows that an (85%, 90%) tolerance interval with $\varepsilon = .05$ and $\alpha^* = .10$ would require 318 runs with at least $k = 278$ (see below for details about calculating actual tolerance intervals for specific data).

For more information regarding the details of the algorithm used in this tool, see Minitab (2019). For a better understanding of the criteria for convergence described above, see Kvam and Vidakovic (2007). It should be noted that while the algorithm implemented in this tool will generate tolerance intervals for a wide range of proportion and confidence requirements, there might often be cases where the number of runs required to observe the tolerance interval is very large. This is particularly true for tolerance intervals with very precise requirements for proportion or margin of error. If the tolerance interval required for the user's specified parameters requires more than 1000 runs, the tool will advise the user to instead use a parametric tolerance interval to save resources. An example of this is shown in Figure 3.

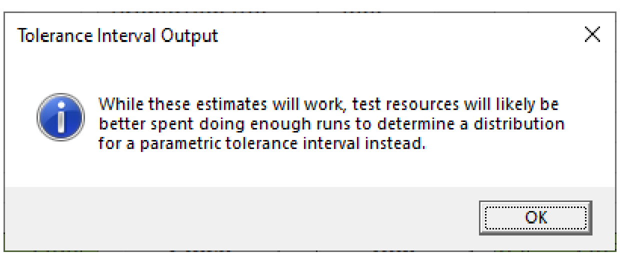

Figure 3. Large Sample Warning Example

Calculating Nonparametric Tolerance Intervals

A simple method for nonparametric calculations uses order statistics. If you expect to collect n observations of a random variable X, then an order statistic would be determined by ranking those n data points according to some order, and then choosing the ith order statistic (written $X_{(i)}$) as the i^{th} largest value. Nonparametric tolerance intervals use order statistics to bound the ends of the interval. In the case of one-sided tolerance intervals, an order statistic serves as a lower- or upper-bound on the interval. In the case of two-sided tolerance intervals, two order statistics are used to bound the interval. The minimum sample size returned by the nonparametric tolerance interval tool applies to both one-sided and two-sided tolerance intervals.

One-sided Intervals

One-sided tolerance intervals have only one bound. A lower tolerance interval has only an upper bound, below which is $P \cdot 100\%$ of the data. This tolerance interval has the form:

$$\left(-\infty, X_{(k)}\right)$$

where $X_{(k)}$ is the k^{th} order statistic (Kvam and Vidakovic, 2007). By $-\infty$, we mean the lowest possible value X could take. Depending on what measurements X represents, other values besides $-\infty$ may be more reasonable. To determine if there is a more appropriate value than $-\infty$, consider the domain of X. If X is a measurement that can only be positive, it would be appropriate to replace $-\infty$ with 0.

An upper tolerance interval has only a lower bound, above which is a proportion P of the data. The one-sided upper tolerance interval has the form:

$$(X_{(n-k+1)}, \infty)$$

where $X_{(n-k+1)}$ is the $(n-k+1)^{th}$ order statistic and again, ∞ should be thought of as the highest value in the domain of X (Kvam and Vidakovic, 2007).

Two-sided Intervals

A two-sided tolerance interval is defined by both a lower and upper limit, between which we expect to find about $P \cdot 100\%$ of the data. The two-sided interval is calculated using the form:

$$(X_{(r)}, X_{(s)})$$

where $k = s - r$ and $s = n - r + 1$. If k is an odd integer, then solving for r and s will result in non-integer solutions. If this occurs, round down to the nearest integer (Kvam and Vidakovic, 2007; and Minitab, 2019). Note that if r and s are rounded, then the resulting two-sided interval will not be symmetric (Meeker, Hahn, and Escobar, 2017).

Verifying Requirements without Calculation

Verifying program requirements using the nonparametric tolerance interval output by the tool is relatively straight-forward process. If at least as many runs as specified by the needed number of successes from the tool are within the requirement threshold, then the system meets or exceeds the requirement. It is important to note that all valid runs must be included in the data set, even those which fail to meet requirements. For example, in the case of the tolerance interval shown in Figure 2, it would not be valid to perform 600 total runs and then claim to meet the tolerance interval by choosing a subset of 318 of them which happens to contain at least 278 successes. This is important because the other runs provide evidence of the performance of the system, and since we have that evidence, we should expect that a corresponding proportion of the total 600 runs would be successful, not just 278. Running excess runs and cherry-picking a subset that meets the tolerance interval can lead to incorrect conclusions about system performance.

If the system is tested for the number of runs specified in the tool are performed and the number of observed successes does not meet the number specified, then if the performance on each run was measured and recorded, we can use that data to estimate both the proportion of the system that *does* meet the requirement, as well as what level of performance the needed proportion is actually meeting. We will give a brief description of how to calculate a tolerance interval once an experiment with the specified number of runs, n, has been conducted.

Applications

In practical situations, tolerance intervals are often most useful or necessary when dealing with systems where very little deviation from an expected performance range is allowed in operation. Some examples are message delivery times for critical messaging systems and machined part tolerances for hardware in mission critical applications. Such systems often require that a part consistently have very precise dimensions or that message delivery almost never fail. In these cases, tolerance intervals provide the capability to state with known confidence whether the system is expected to perform within acceptable bounds over all future operations. The use of tolerance intervals in these cases provide an advantage over confidence and prediction intervals, which only provide information regarding an estimate of the mean or a single future observation, respectively.

For new systems that are not yet well-understood (where the distribution of data is unknown), a nonparametric tolerance interval can provide a means of determining if the system meets requirements. However, because a nonparametric tolerance interval typically requires more runs than a parametric interval for the same probability and confidence, it can be difficult and expensive to perform the number of test runs required for mission-critical systems which need a very high probability of performing as desired (e.g., message delivery probability of 99.99%). In these cases, the nonparametric tolerance interval tool may still be useful to provide a roadmap for testing needed to show the system meets performance requirements. If there is an urgent need to field a system, a set of nonparametric tolerance intervals at increasing probabilities building up to the required value (e.g., tolerance intervals at 95% confidence for message delivery probabilities of 99%, 99.9%, and 99.99%) can allow an understanding of the system performance to be built and refined over time. This can provide decision makers with the framework to decide if it is reasonable to make a short-term fielding decision for an initial capability with the expectation that a better understanding of the system can be built towards over time using additional testing or operational data.

Conclusion

Tolerance intervals identify a set of bounds that contain a specified proportion of the population. There is always a risk that the sample is not representative of the population, so tolerance intervals include a confidence level to control this risk. This underutilized statistical interval is particularly useful for answering questions regarding ranges where data commonly falls. However, tolerance intervals are sensitive to deviations from assumptions regarding the underlying distribution of the data. If distributional assumptions do not hold, or if there is insufficient information about the distribution of the data, then it is best to use nonparametric strategies to build the tolerance interval. The nonparametric tolerance interval tool described in this report allows users to determine what minimum sample size is necessary to build a nonparametric tolerance interval with the characteristics they input. The tool also provides the number of data points that should be contained in the tolerance interval, allowing the user to create the desired nonparametric tolerance interval once data collection has been completed.

References

Kvam, Paul H. & Vidakovic, Brani. “Nonparametric Statistics with Applications to Science and Engineering”. Wiley Series in Probability and Statistics, 2007.

Meeker, William Q.; Hahn, Gerald J.; and Escobar, Luis A. “Statistical Intervals: A Guide for Practitioners and Researchers” (Second Ed.). Wiley Series in Probability and Statistics, 2017.

Minitab. “Methods and Formulas for Sample Size for Tolerance Intervals”. Minitab, LLC, 2019.

Ortiz, Francisco and Truett, Lenny. “Using Statistical Intervals to Assess System Performance”. Scientific Test and Analysis Techniques Center of Excellence (STAT COE), 2015.

Splinter, Kevin; Sigler, Gina; Harman, Michael; and Kolsti, Kyle. “Tolerance Intervals Demystified”. Scientific Test and Analysis Techniques Center of Excellence (STAT COE), 2020.

Appendix – Detailed Example

Imagine testing the notional New Secure Messaging System (NSMS). The designers of the system are sure that it can send messages with the desired level of security; however, the amount of time between sending and receiving the message is not well understood for this system. Subject matter experts say that being able to regularly send and receive a message within 20 seconds is required for the messaging system to be useable in the field.

Based on the information we have been given so far, we know that we wish to be able to place and upper bound on the amount of time NSMS takes to deliver a message. While we cannot say that ALL delivery times will be below a particular, we may be able to identify a bound below which a certain percentage of the delivery times, say 90%, fall. In other words, this situation is ideal for a one-sided tolerance interval.

No previous tests have been done on the delivery time of NSMS. Additionally, the updates that make this system more secure also make it different enough from previous systems that we are unable to use previous systems to inform our understanding of the distribution for this system. Therefore, we are unsure what distribution the delivery times will follow. In the absence of any information about the distribution of the delivery time of these messages, it will be most reasonable to use a nonparametric tolerance interval to determine if the majority of the messages are being delivered within the desired time limit.

To complicate matters, the test for NSMS delivery times has two factors that the testers must manually change between each run: network load and message size. Each factor has two levels (low and high), making four possible conditions. Leadership would like to understand how these factors impact message delivery time, so we wish to create a tolerance interval for each of the four conditions.

The lab tells us that because manually changing the factors between each run takes time, we only have enough time to do a maximum of 200 runs. Since we wish for our tolerance intervals to be comparable, we will begin by assuming that the number of runs are evenly divided between each of the conditions. Then we have a maximum of 50 runs to use to create a tolerance interval for each condition. We can use the nonparametric tolerance interval tool to see what precision can be achieved for 50 runs or less. We will start with an ideal nonparametric tolerance interval that is likely to require too many runs then try loosening some of our inputs.

Ideally, our tolerance interval would have a very large P because we would like the vast majority of the distribution to fall below our upper tolerance bound. One might imagine that a $P = .99$ might work well for this problem, however, we must keep in mind that nonparametric tolerance intervals generally require larger sample sizes. If $P = .99$, then $\varepsilon < .01$; this small margin of error will require a very large sample size. Instead, we will use the requirement that $P = .90$ so we have a more room to work with in terms of margin of error. The total run requirements given by the tool for various sets of nonparametric tolerance interval parameters are shown in Table 4. As we can see in the first row of Table 4, our first try required far too many runs to be feasible in this experiment. We then try loosening our requirements for $1 - \alpha$, ε, and α^*one-by-one as shown in intervals 1-4. We can only increase ε from .05 to .09 to conform to the constraint $0 < \varepsilon < 1 - P$. We can see that allowing a larger margin of error on P (i.e., increasing the value of ε) goes a long way towards reducing the required number of runs. After the 4th interval, we change the confidence level, ε, and α^* to try and find a combination that provides a useful tolerance interval while keeping n small. By modifying the confidence level and α^* we produce several intervals that have much smaller sample sizes. Only the 10th interval has a sample size less than 50, for a total of 156 runs across the four conditions. However, if it would be possible to convince the program office to increase the run cap by 8 (two more for each condition), then we could use the tolerance interval in row 9, which gives 0.5 α^* instead of a .10 α^*.

Table 4: Considered Tolerance Intervals

Interval #	Confidence $(1-\alpha)$	Proportion (P)	Margin of Error (ε)	Margin of Error Risk (α^*)	Runs Required (n)
1	.99	.90	.05	.01	612
2	.95	.90	.05	.01	425
3	.95	.90	.09	.01	89
4	.95	.90	.05	.05	310
5	.90	.90	.05	.01	368
6	.90	.90	.05	.05	245
7	.90	.90	.05	.10	210
8	.90	.90	.09	.05	52
9	.90	.90	.09	.01	65
10	.90	.90	.09	.10	38

After speaking to the program office and the lab, we are told it would be possible to do eight more tests to make the total number of tests run 208. Given this information, the tolerance interval in row 8 appears to be the best for our purposes.

The lab executes the test with 208 runs and returns the (notional) results to us. Below are the results of the high load, low file size condition. The results in Table 5 have been sorted from largest to smallest for our convenience.

Table 5: Delivery Times (in seconds) for High Load, Low File Size

5.126429	8.212724	9.106340	9.253784
10.423324	12.393966	12.609172	12.670580
12.972754	13.195848	13.204790	13.447015
13.603801	13.629823	13.735626	13.738818
13.827530	14.171431	14.184233	14.445580
14.912384	15.412042	15.435410	15.464379
15.777407	16.333406	16.359170	16.364353
16.420189	16.667285	16.967149	17.018127
17.118655	17.160076	17.164883	17.688404

17.768533	18.238697	18.257817	18.405642
18.580804	18.763560	19.348932	19.500355
19.745636	19.748467	19.881802	20.735652
20.819050	21.452909	21.660328	21.717389

Before making our nonparametric tolerance interval, we should confirm that a normal tolerance interval is not appropriate for this data. If the data follows a normal distribution, we should use a normal tolerance interval instead because a normal tolerance interval would allow us to have a higher confidence level or higher proportion, P, using the same sample size. Checking for normality in the data may be done more rigorously using a QQ-Plot, but for our purposes, it is sufficient to examine the histogram below. As we can see, the delivery times do not appear to follow a bell-shaped, symmetric distribution, so it would not be appropriate for us to use a normal tolerance interval. In the absence of more information about the distribution of delivery times, we will proceed with finding a nonparametric tolerance interval.

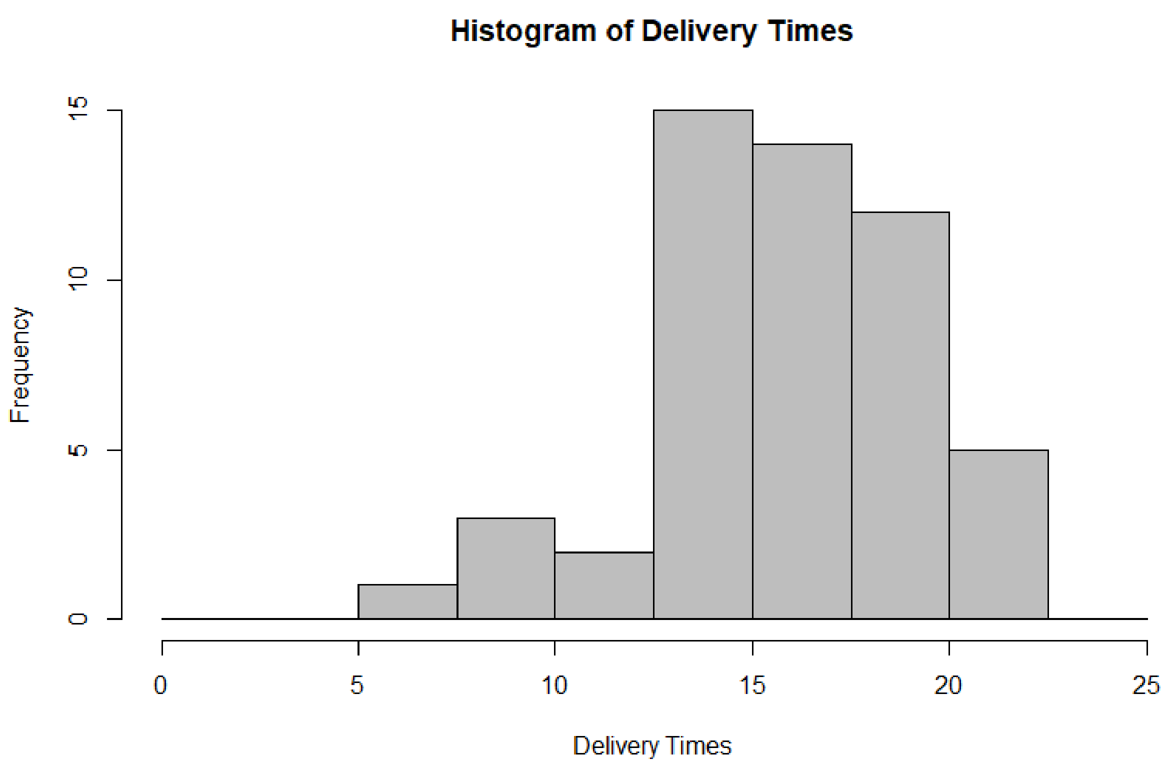

Figure 4: Histogram of Delivery Times

We will begin our analysis by finding the nonparametric tolerance interval for the high load, low file size condition (data given in Table 5). Our objective is to determine if the NSMS sends and delivers messages within 20 seconds or less most of the time, so we wish to find a one-sided tolerance interval which has the form $\left(-\infty, X_{(k)}\right)$. In this scenario, it is not possible for delivery times to be less than zero, so our one-sided tolerance interval will take the form $(0, X_{(k)})$. To find k, we return to the nonparametric tolerance interval tool. Table 6 shows the tool inputs and outputs for our tolerance interval.

Table 6: Nonparametric Tolerance Interval Tool Inputs and Outputs

Confidence $(1-\alpha)$	Proportion (P)	Margin of Error (ε)	Margin of Error Risk (α^*)	Runs Required (n)	Expected Successes (k)
.90	.90	.09	.05	52	49

Based on the output of k from the nonparametric tolerance interval tool, we have that our one-sided tolerance interval should take the form $(0, X_{(49)})$. All that remains for our calculation of this tolerance interval is to find the 49th largest collected value under this condition. Table 5 is already sorted from smallest to largest, making it easy for us to find 20.81905 to be the 49th largest value. Therefore, we find that for the high load, low file size condition, we are 90% confident that at least 90% of the delivery times for NSMS are less than 20.81905 seconds. Hence, at least for this condition, we do not appear to have met the objective of most delivery times being less than 20 seconds. This should lead to a discussion with leadership regarding how close this is to the requirement. A larger budget will be required to enable further testing, or a parametric tolerance interval should be used based on the distribution of the data already collected.

The values in Table 6 apply to data from all four conditions, so we my repeat the process of finding the 49th largest value to find the upper bound for the one-sided tolerance interval three more times (data from the remaining three conditions are not shown here). By finding the tolerance interval for each of the conditions, we might produce the results in Table 7.

Table 7: Nonparametric Tolerance Intervals for Each Condition

Condition	Tolerance Interval
Low load, low file size	(0 sec, 17.39657 sec)
Low load, high file size	(0 sec, 19.17505 sec)
High load, low file size	(0 sec, 20.81905 sec)
High load, high file size	(0 sec, 20.08922 sec)

The findings in Table 7 are all intervals we are 90% confident contain at least 90% of the delivery times under the specified network load and file size conditions. Based on these findings, we might have some concern over the high load, high file size condition, which has an upper bound slightly above our cut-off of 20 seconds in additional to the high load low file size condition. A discussion with subject matter experts and leadership may be necessary to decide if this tolerance interval indicates that NSMS is operating within acceptable limits, if further testing is needed, or if the new secure messaging system does not offer short enough messaging times to be reliable under some conditions.