5-2014

Lifetime Modeling of Deficient Bridges in New York

Levi Phippen

LIFETIME MODELING OF DEFICIENT BRIDGES IN NEW YORK

by

Levi Phippen

A research project submitted in partial fulfillment
Of the requirements for the degree

of

MASTER OF SCIENCE

in

Statistics

Approved:

Dr. Daniel C. Coster Major Professor	Dr. John R. Stevens Committee Member

Dr. Paul Barr
Committee Member

UTAH STATE UNIVERSITY
Logan, Utah

2014

ABSTRACT

Lifetime Modeling of Deficient
Bridges in New York

by

Levi Phippen, Master of Science
Utah State University, 2014

Major Professor: Dr. Daniel C. Coster
Department: Mathematics and Statistics

Given the importance of bridges to a state's economy and strength, and the costs involved in building and maintaining bridges, maximizing their service life is essential. In order to safely extend a bridge's utility as long as possible, an understanding of its lifetime processes is needed. This paper attempts to model the lifetime of a bridge in New York once it has become deficient. Lifetime is defined to be the length of time between deficiency classification and failure. A bridge is considered deficient when certain structural components receive a poor rating in the National Bridge Inventory, which is compiled annually by the Federal Highway Administration. A list of bridge failures is provided by the New York State Department of Transportation.

In 2012 the Federal Highway Administration database showed that 39.5 percent of New York's bridges were deficient. Using analysis of variance and considering a bridge failure to be a random Bernoulli trial, this paper shows that New York's deficient bridges are typically older than their non-deficient counterparts, and that they are also more susceptible to failure. From survival analysis techniques an estimate of the mean

time to failure of a deficient bridge is found to be 47.2 years. Finally, a statistical model is created to predict the lifetime of a deficient bridge, while accounting for influential factors such as average daily truck traffic, deck geometry and structural evaluation ratings.

ACKNOWLEGMENTS

Special thanks to Mr. Auyeung Winchell for providing the author with the New York Bridge Failure Database. And special thanks to Dr. Wesley Cook for providing expert opinion on different aspects of bridge structures as they related to this study. In addition, thank you to my major professor and committee members, Drs. Daniel Coster, John Stevens, and Paul Barr for their help and involvement in this study.

Levi Phippen

LIFETIME MODELING OF DEFICIENT
BRIDGES IN NEW YORK

by

Levi Phippen, Master of Science
Utah State University, 2014

Major Professor: Dr. Daniel C. Coster
Department: Mathematics and Statistics

Introduction:

America's infrastructure is aging, and resources to maintain the infrastructure are scarce. One particular infrastructure concern is the state of our nation's bridges. In the American Society of Civil Engineer's 2013 Report Card for America's Infrastructure, they note that roughly 24.9 percent of our bridges are considered deficient. In New York, the percentage is higher, 39.5 in 2012, according to the Federal Highway Administration (FHWA). Given the critical role bridges play in a nation's and state's economic strength, maintaining our bridges is critical. Part of maintaining a bridge properly is having an estimate of how long the bridge can be expected to last given certain conditions.

The primary condition of interest in this paper is a bridge's classification as either deficient or non-deficient. Deficiency status has received a lot of attention in the media lately, as two major bridge collapses have caught the public's attention. Both the I-5 Skagit River Bridge in Washington and the I-35 West Bridge in Minnesota were deficient when they collapsed in 2013 and 2007 respectively. In an effort to better understand characteristics of deficient bridges this paper will focus on answering four questions about deficient bridges in New York State. We restrict our analysis to bridges in New

York because only its department of transportation has kept a record of bridge failures in its state. No other state has kept a similar record. This paper focuses on the following objectives, all of which refer only to bridges in New York:

1. Are deficient bridges more vulnerable to collapse than non-deficient bridges?
2. Is there a significant difference in age between deficient and non-deficient bridges?
3. Once a bridge has become deficient, how long can it be expected to last before it fails, assuming no repairs are made to change its deficiency status?
4. Can we make a model that will predict how long a bridge will last once it becomes deficient, accounting for different bridge characteristics, such as design, construction material, traffic load, length, etc.?

Common statistical analysis tools will be used to answer all questions. In particular, to answer the third and fourth questions, survival analysis techniques will be employed. This is not the first time they have been applied to bridges. FHWA's bimonthly magazine *Public Roads* contained an article in its May/June edition of 2008 detailing a study using survival analysis to predict the length of time a deck remains in any given condition. In 2009, Habib Tabatabai, Mohammad Tabatabai and Chin-Wei Lee used similar methodology to build a deck reliability model and analyze failure rates of bridges in Wisconsin. Si Soon Beng and Takashi Matsumoto published a paper in 2012 detailing a bridge lifetime projection model for all bridges in general in the U.S. and Japan. Their model was derived using the techniques of survival analysis. In this paper we apply many of the same techniques to estimate how long a bridge lasts on average once it has become deficient, and to build a prediction model that refines this estimate by accounting for influential characteristics of the bridge, such as daily traffic load.

Data Sources and Definitions:

The data for this analysis come from two sources, the New York State Department of Transportation (NYSDOT), and the Federal Highway Administration's National Bridge Inventory, the NBI. In 1992 NYSDOT began keeping a record of all bridge failures that occurred within its state. They define failure in two ways, as either a partial collapse or a total collapse. A partial collapse is a bridge failure such that "all or some of the primary structural members of a span or multiple spans have undergone severe deformation such that the lives of those traveling on or under the structure would be in danger" (Auyenung). A total collapse is a bridge failure such that "all primary members of a span or several spans have undergone severe deformation such that no travel lanes are passable" (Auyenung). New York is the only state in the United States that has kept a record of bridge failures within its state. Therefore this study, of necessity, limits its analysis to only those bridge failures that occurred in New York.

The second source of data comes from the National Bridge Inventory, the NBI. Each bridge in the United States highway system is inspected yearly or biennially. Myriad parts of a bridge's structure and use are evaluated or measured, and then recorded in the Federal Highway Administration's NBI. Of particular interest in this study is the rating a bridge inventory inspector assigns to certain structural components. These ratings determine whether or not a bridge is deficient. There are two types of deficiency, structurally deficient and functionally obsolete. According to the Federal Highway Administration a bridge is structurally deficient if in the NBI it receives one or more of the following ratings:

1. A condition rating of four or less for Deck, Superstructure, Substructure, or

Culverts and Retaining Walls (if applicable)

2. An appraisal rating of two or less for Structural Condition or Waterway Adequacy (if applicable)

A bridge is functionally obsolete if in the NBI it receives one or more of the following ratings:

1. An appraisal rating of three or less for Deck Geometry, Underclearances (if applicable), or Approach Roadway Alignment
2. An appraisal rating of three for Structural Condition or Waterway Adequacy (if applicable)

These ratings are on a scale of zero to nine, with nine meaning excellent condition and zero meaning failed condition. If a bridge qualifies as both structurally deficient and functionally obsolete, it is counted as structurally deficient (U.S. DOT). Except for when investigating a significant difference in age between different deficiency statuses, this paper will not distinguish between structurally deficient and functionally obsolete bridges. If a bridge is either, it is simply considered deficient.

Before proceeding further, it is important to make clear that when NYSDOT lists a bridge collapse in its database, it does not necessarily mean that the bridge fell to the ground. It simply means that the bridge experienced a failure such that continued use without repair would be unsafe or impossible. This paper will therefore use the terms failure and collapse interchangeably.

In order to begin the analysis, as many bridges as possible from NYSDOT's failure database were found in the NBI. There were listed 112 failures between 1992 and 2012. Of these, five did not have NBI identification numbers and therefore could not be identified. Of the remaining 107, two were duplicates. Removing those left 105. Of these, one was a private bridge and five were for pedestrian use only. None of these six bridges had NBI data relevant to the study. Removing them left 99. From this group of 99 bridges, 64 were deficient at the time of collapse. These 64 bridges were used to model the lifetime of a bridge from deficiency classification to collapse, lifetime being measured in years.

Greater Risk of Failure:

Of particular interest is whether or not deficient bridges are at greater risk of failure than non-deficient bridges. If a bridge collapse is viewed as a random Bernoulli trial, and the collapse of a deficient bridge a success, then a bridge collapse follows a Bernoulli distribution. If we let C represent the event that a bridge collapses, and p represent the probability of success (a deficient bridge collapsing), then we can write $C \sim Bernoulli(p)$. If X is the random variable that counts the number of deficient bridge collapses in n collapses, then $X \sim Binomial(n, p)$. In order to determine which deficiency status is more susceptible to collapse, p needs to be estimated. Under the assumption that the two statuses have an equal probability of collapse, p can be estimated by the proportion of deficient bridges in New York between 1992 and 2012. For example, if the proportion during these years was 30 percent, and deficient and non-deficient bridges collapsed with equal probability, then about 30 percent of all collapses should be collapses of deficient bridges. Or in other words, the probability of a deficient bridge failure is 30 percent, 0.30.

The total number of bridges and total number of deficient bridges in New York for each of the years 1992 through 2012 can be found on the Federal Highway Administration's website. Using their figures it is easy to calculate the proportion of deficient bridges for each year in question. If we sum the number of deficient bridges across all years, and then divide by the sum of the total number of bridges across all years, we end up with the proportion of deficient bridges in New York from 1992 through 2012. The sum of deficient bridges across this time span is 181,969. The total sum of all bridges in New York during the same period is 364,431. Thus the proportion of deficient bridges from 1992 through 2012 is $\frac{181,969}{364,431} = 0.4994 \approx 0.5$. If we assume that the risk of collapse is equal between deficient and non-deficient bridges, then this proportion also represents the probability of a deficient bridge collapse during the same time period.

From 1992 through 2012 there were 99 bridge collapses, 64 of them were deficient. Using the binomial distribution we can now compute the probability of seeing 64 or more deficient bridges collapsing out of 99 total collapses. Earlier we said that X represented the total number of deficient bridge collapses and that $X \sim Binomial(n, p)$. Now that we have values for n and p, we can write $X \sim Binomial(n = 99, p = 0.5)$. The probability of seeing 64 or more deficient bridges collapsing out of 99 total collapses is then 0.0023. This probability means that if deficiency status does not affect probability of collapse, then the probability of seeing 64 or more deficient bridges collapsing out of 99 total collapses is 0.0023, or 0.23 percent. This is a very low probability, and a more reasonable explanation for seeing so many deficient bridges collapse is that their probability of collapse is higher than that of non-deficient bridges.

Significant Difference in Age:

Since there is strong evidence that deficient bridges are more susceptible to collapse than non-deficient bridges, an understanding of the various attributes of deficient bridges is important. One attribute studied in this paper is whether or not there is a significant difference in age between deficient and non-deficient bridges. If the data suggest that older bridges are more likely to be deficient than younger ones, then older bridges are also at greater risk of collapse. Knowing this gives government agencies greater knowledge to help them effectively manage bridge maintenance funds.

Analysis of variance tests were done to investigate differences in age between deficient and non-deficient bridges. The bridge population came from New York's 2012 bridge population. Bridge age was computed as the difference between 2012 and the year of its construction, or if the bridge had been reconstructed, it was taken to be the number of years between 2012 and its reconstruction. Two F-tests were done, the first one looking for a significant difference in the square root age between non-deficient, structurally deficient, and functionally obsolete bridges. The second test explored differences in the square root age among deficient and non-deficient bridges, without differentiating between types of deficiency. The square root of age was used in this analysis instead of an untransformed age, because without the transformation the distribution of ages was right skewed. Taking the square root of age normalized the distribution quite well.

The difference in square root age between groups of bridges was highly significant in both tests, each with a p-value of less than 2.2×10^{-16}. Below is given the analysis of variance table comparing square root age between non-deficient, structurally deficient and functionally obsolete bridges.

ANOVA for Difference in Square Root Age of Non-Def, SD, and FO Bridges

Source	Degrees of Freedom	Sums of Squares	Mean Squares	F-statistic	P-value
Treatments	2	7767	3883.7	1055	< 2.2e-16
Error	17488	64375	3.7		
Total	17490	72142			

Results of an analysis of variance test comparing square root bridge age between non-deficient and deficient bridges, with no distinction for type of deficiency is given below.

ANOVA for Difference in Square Root Age of Non-Def and Def Bridges

Source	Degrees of Freedom	Sums of Squares	Mean Squares	F-statistic	P-value
Treatments	1	5649	5649.4	1485.9	< 2.2e-16
Error	17489	66493	3.8		
Total	17490	72142			

The analysis of variance tests suggest that there is a strong association between age and deficiency status. It appears that in New York deficient bridges are generally older than non-deficient bridges. Figure 1 below shows a boxplot comparing square root age between non-deficient and deficient bridges.

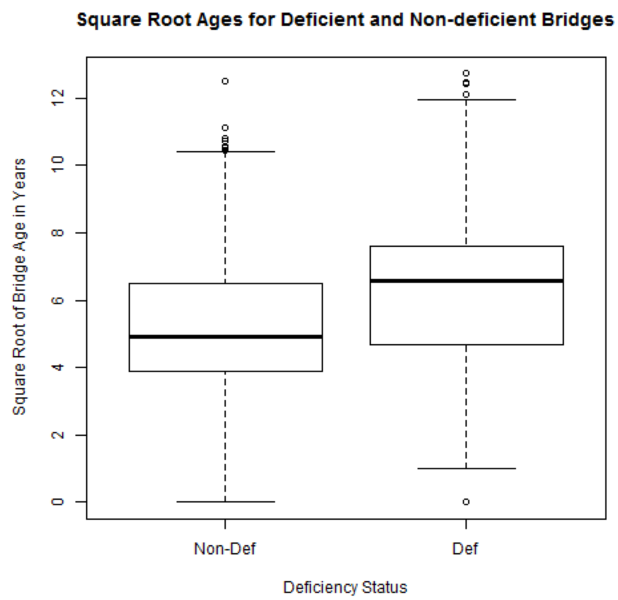

Figure 1: Boxplot showing the square root age distribution for non-deficient and deficient bridges.

On the untransformed scale, median age for non-deficient bridges was 24 years, and the mean age was 29.72 years. The median and mean ages for deficient bridges were 43 and 43.68 years respectively. And lastly, the median and mean ages for all bridges were 39 and 39.08 years.

Survival Analysis Censorship:

An influential aspect of this research results from the fact that the data is highly censored. Censoring is a phenomenon common in most survival analysis studies. It occurs when the lifetime of a subject is not completely known. Censoring in this study occurs because there is no available NBI data prior to 1992. Many bridges were presumably deficient before 1992, but we don't know for sure. For them exact lifetimes cannot be known. For example, suppose a bridge was built in 1975, collapsed in 1996, and was recorded in the NBI as being deficient in 1992, 1993, 1994, 1995, and 1996. Since there is no data available earlier than 1992, exactly when this bridge first became deficient cannot be determined. It is possible that the bridge had been deficient since 1981 or 1990. All that is known for sure, is that it was deficient from 1992 to 1996.

This bridge would be considered a censored observation, and we would say its lifetime is at least 5 years. As stated above, the data for this study are highly censored. Of the 64 bridges that were deficient when they collapsed, only 10 have lifetimes that can be determined exactly. With roughly 84 percent of the data censored, precise conclusions are difficult to make. This paper will make an attempt however, using common survival analysis techniques.

When conducting a survival analysis study, one of the first steps is to determine which type of censoring is present. Right censoring is most common. It occurs when the lifetime measurement is unknown on the right side of the time axis. For example, suppose 100 rats are administered a poison, and then monitored for 30 days. After 30 days 84 rats have died, and 16 are still alive. The most that can be said for the 16 surviving rats, is that they survived at least 30 days. Since we don't know their lifetimes exactly, these observations are considered censored. And since the incompleteness of their lifetime observations falls on the right of the time axis, they are considered right censored observations.

For the rats, lifetime was measured in days from poison administration. Their beginning time began when they were poisoned, and the event of interest was their death. For the bridges, their beginning time is classification as deficient, and the event of interest is their collapse. Unlike the rats however, every bridge under study has experienced the event of interest. What is unknown for most bridges is their time zero, or when they first became deficient. Therefore the incompleteness in the bridge lifetimes falls on the left side of the time axis and not on the right. Because of this, right censoring is perhaps not an appropriate way to account for the censoring present in this study.

Since the incompleteness of the observations occurs on the left, perhaps left censoring could account for the censoring present in this analysis. According to Melvin Moeschberger and John Klein, left censoring occurs "if the event of interest has already occurred" when the study begins (p. 62). This definition refers to a prospective study, where qualified subjects are brought in and watched until the event of interest happens. If a subject has already experienced the event of interest before coming under observation, that subject is left censored. For a left censored observation, the most you can say about its lifetime is that it is less than some time length.

The study in this paper is retrospective. We are studying bridges we know have collapsed and looking into their past for clues about their collapse. The most we can say about our censored observations is that their lifetimes are greater than or equal to, not less than, some number of years. Therefore left censoring does not make sense.

Another censoring scheme is called left truncation. David Hosmer and Stanley Lemeshow define left truncation as a situation where there exists a clear beginning time (or $t = 0$), but a subject doesn't come under observation until this time has passed. To further clarify their definition they give an example that is worth repeating here:

> In modeling age at menarche, suppose we define the zero value of time as 8 years. Suppose a subject enters the study at age 10, still not having experienced menarche. We know that this subject was "at risk" for experiencing menarche since age 8 but, due to the study design, was not enrolled in the study until age 10. This subject would not enter the analysis until time 10. This type of observation is called *left truncation* or *delayed entry* (p. 20).

This type of censoring is at first attractive as a means of handling the censoring present in this bridge study. It is tempting to say that all the bridges that were deficient back to 1992 did not come under observation until their time zero (t = 0) had passed. This would be appropriate except for the number of elapsing years between time zero and 1992 is unknown. A bridge deficient in 1992 may have been deficient for five years prior, or ten. Therefore, left truncation is not appropriate.

Another form of incomplete observation can happen when the entire study group has experienced the event of interest. This most often occurs when subjects are selected because it is known that they have experienced the event of interest. In such a situation, the data is said to be right truncated. In this case, researchers are concerned with time from the appearance of some risk factor to the occurrence of the event. According to Hosmer and Lemeshow, a defining characteristic of this type of censoring, is that “the time to event for each subject is known” (p. 21). Since the time to event of each bridge in this study is not known, a right truncation censoring scheme is not possible.

Interval censoring is a method to account for incomplete observations when the most that can be said about a lifetime is that it is greater than some time period, but less than some other period. Again, this does not fit our description because our lifetime is not known to be greater than some length of time and less than another. Each censored observation is only known to be greater than or equal to some number of years.

None of the above options for dealing with censorship is an exact fit to the data in this paper. The closest alignment is with right censoring. The defining characteristic of right censorship is that censored observations are only known to be at least as large as some value. Since this is all that can be said about the censored bridge lifetimes, we proceed with the analysis using a right censoring scheme.

The Survivorship Function:

One of the most commonly estimated functions from survival data is the Kaplan-Meier survivorship function (or the Kaplan-Meier product-limit estimator). It is quite simple to compute, and estimates the probability that a subject survives beyond some time t. The cumulative distribution function, $F(t)$ gives the probability that a subject's survival time is less than or equal to time t. The survivorship function is therefore the complement of the cumulative distribution function, $1 - F(t)$, and is denoted $S(t)$. A graph of the Kaplan-Meier estimator for the survival function $S(t)$ is given below in Figure 2 for New York's collapsed deficient bridges. The tick marks represent censored observations, and the dashed lines represent the upper and lower bounds of a 95 percent confidence interval for the function.

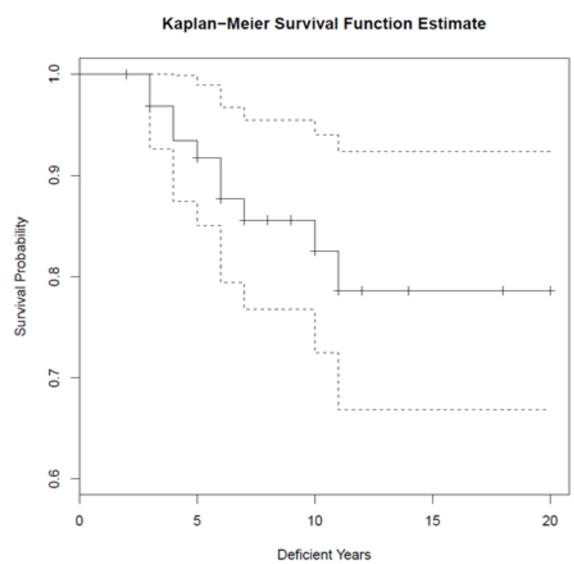

Figure 2: Kaplan-Meier survival function estimate with 95 percent confidence interval.

We can see that the function takes its lowest value at the line $y = 0.786$. To interpret this value one would say the probability that a bridge survives 11 years or more after it becomes deficient is 0.786. Note that in this context, a bridge surviving implies that it does not experience a failure. The survival probability estimate of 0.786 has an associated 95 percent confidence interval of [0.668, 0.924]. This interval tells us we can be 95 percent confident that the probability of a bridge surviving beyond 11 years after its classification as deficient is between 0.668 and 0.924. In other words, a bridge has about a 66.8 to 92.4 percent chance of surviving more than 11 years from the time it became deficient.

Besides the Kaplan-Meier survival curve estimated above, another common function estimated from survival data is the hazard function. If the random variable time is taken to be completely continuous, then the survivorship function can be expressed as:

$$S(t) = e^{-H(t)} \Rightarrow H(t) = -ln\big(S(t)\big).$$

$H(t)$ is known as the cumulative hazard function. It represents the probability that a bridge having survived up to time t will collapse in the next instant. There are a few different ways to estimate $H(t)$. Since the surivival function discussed above was derived using Kaplan and Meier's method, we will use A. V. Peterson's approach, proposed in 1977, for estimating the cumulative hazard function from the Kaplan-Meier curve. The derivation is quite simple and is given below:

$$\widehat{H}(t) = -ln\left(\hat{S}(t)\right) = -ln\left(\prod_{t_{(i)}\le t} \frac{n_i - d_i}{n_i}\right) = -\sum_{t_{(i)}\le t} ln\left(\frac{n_i - d_i}{n_i}\right) = -\sum_{t_{(i)}\le t} ln\left(1 - \frac{d_i}{n_i}\right),$$

where $t_{(i)}$ represents the i^{th}ordered survival time, d_i is the number of uncensored collapses at time $t_{(i)}$, and n_i represents the number of bridges at risk at $t_{(i)}$. In Figure 3 below is a graph of both the survival function estimator and the cumulative hazard function estimator.

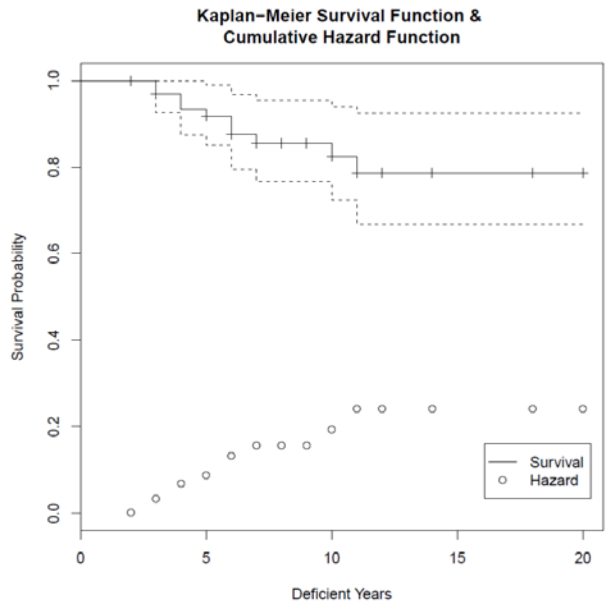

Figure 3: Kaplan-Meier survival curve with an estimate of the cumulative hazard function.

The estimated cumulative hazard function reaches its highest probability at 0.241, which corresponds to 11 deficient years. This number represents the probability that a bridge that has been deficient for 11 years collapses in the next instant. This is a rather large probability. It should be remembered that the word collapse in this context does not necessarily mean that the bridge falls to the ground, but rather something happens to the bridge that makes it unsafe for use. It should also be noted that after 11 years, the only bridge collapses are censored observations, and therefore there is no new information with which to update the cumulative hazard or survival functions. Thus the probability of a deficient bridge surviving to 20 years and then collapsing is also estimated to be 0.241, the same as for a bridge deficient for 11 years. In reality, these two probabilities are probably not equal, but this is the most that can be said from the available data.

Parametric Survival Regression:

A Weibull distribution is used to fit a parametric model to the Kaplan-Meier survival function. This distribution is chosen primarily, because unlike the Exponential, the Weibull distribution has a non-constant hazard function, as will be shown later. The Weibull model is fit using maximum likelihood parameter estimates calculated from the software package R, using the “survreg” function, available in the “survival” package. The estimated parameters are $\alpha = 1.18$ for the shape parameter, and $\lambda = 0.02$ for the scale parameter. The Wald test statistic for the log of the shape parameter is -0.618. Testing whether or not it is significantly different from zero gives a p-value of 0.536. This large p-value indicates that there is little evidence that the additional complexity of a shape parameter is justified by the data. This suggests that an exponential model would be sufficient. However, using the Exponential distribution to model the failure times of New York's deficient bridges results in a hazard function that is constant, meaning the probability of failure in the next instant is the same for all bridges, no matter for how long they have been deficient. Since any bridge deteriorates over time, and it is believed that deterioration increases risk of failure, a constant hazard rate seems unrealistic. Therefore we will use the more complex Weibull distribution to model deficient bridge lifetimes, even though the log of its scale parameter is not significantly different from zero. Doing so results in a hazard function that is monotonically increasing, implying that the probability of bridge failure in the next instant increases as deficient bridge lifetime increases. Below is the Kaplan-Meier survival curve with the Weibull survival function plotted in the same window. As can be seen in Figure 4, the Weibull model provides a reasonable fit to the data.

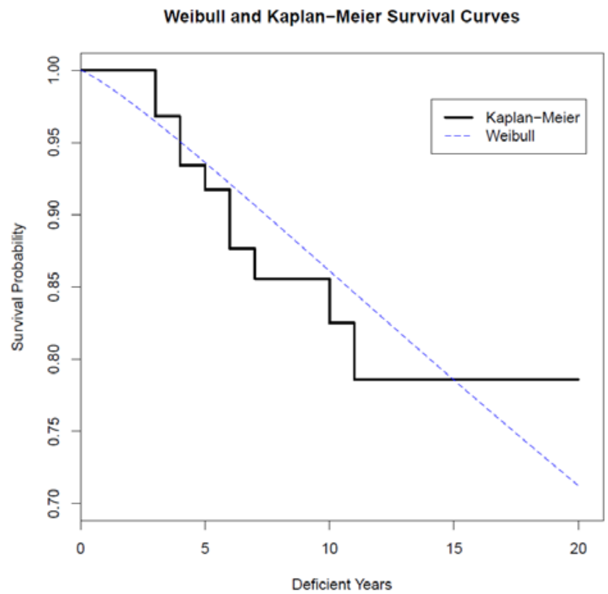

Figure 4: The Kaplan-Meier Survival Curve plotted with the estimated Weibull survival function.

Once the shape and scale parameters are estimated, we can derive the survival and probability density functions. Using these two functions we can calculate the hazard function from the formula $h(t) = \frac{f(t)}{S(t)}$. The plots for the survival, probability density and hazard functions are given in Figures 5, 6 and 7 below.

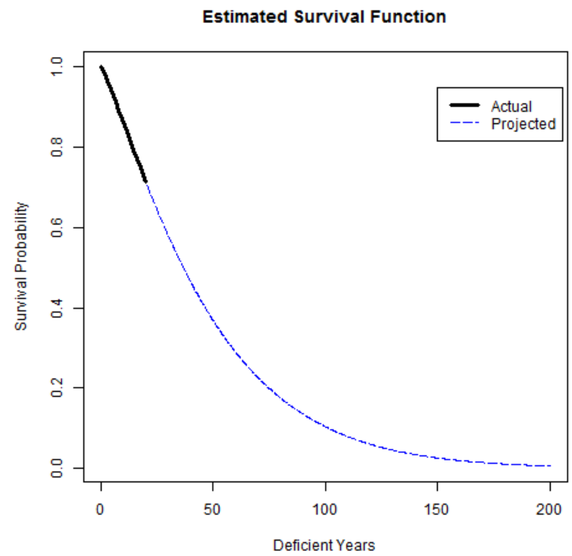

Figure 5: The survival function of a Weibull random variable with shape and scale parameters 1.18 and 0.02 respectively.

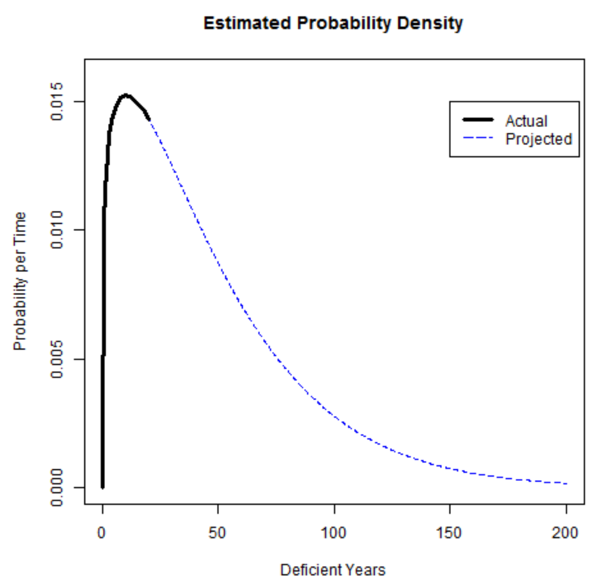

Figure 6: The probability density function of a Weibull random variable with shape and scale parameters 1.18 and 0.02 respectively.

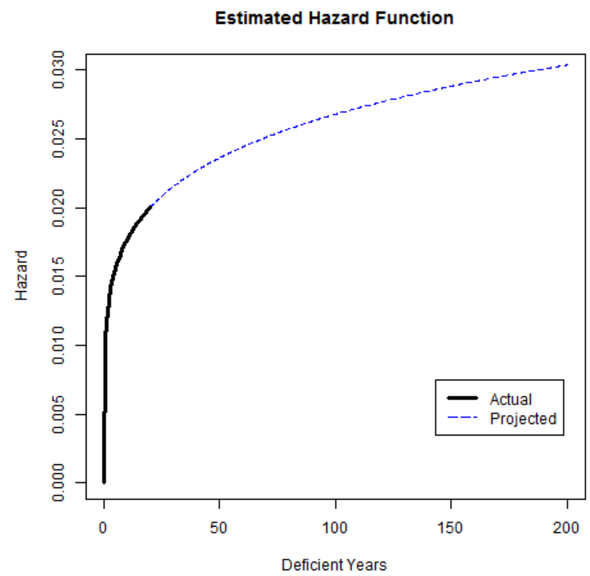

Figure 7: The hazard function of a Weibull random variable with shape and scale parameters 1.18 and 0.02 respectively.

From the hazard function in Figure 7 we can see that the risk of a bridge failure in the next instant increases as deficient years increases.

In each of these plots, the bold solid line represents years for which there is actual data. The rest of the curve, the blue dashed line, is an extrapolation from the data. The accuracy of any conclusions derived from these functions will therefore be suspect. However, given the current unavailability of pre-1992 NBI data, greater accuracy is not possible. Estimates from these functions, though perhaps not totally reliable, can at least give an idea of the lifetime processes of New York's deficient bridges.

Estimating Mean and Median Survival Times:

From the survival curve, we can estimate the median survival time of a deficient bridge. It is 36.6 years, meaning we would expect 50 percent of all deficient bridges to fail by the time they had been deficient for about 36.6 years. From the probability density function we can get an estimate of the mean survival time, or the expected survival time, of a deficient bridge by applying the formula for the expected value of a Weibull distribution: $E(T) = \frac{1}{\lambda}\Gamma\left(1+\frac{1}{\alpha}\right)$. We can also obtain the standard deviation for the random variable survival time. The sandard deviation is found by via the following formula: $SD(T) = \sqrt{\frac{1}{\lambda^2}\{\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma\left(1+\frac{1}{\alpha}\right)^2\}}$. The expected survival time, or mean lifetime, turns out to be 47.2 years. The standard deviation of survival time is 40.1 years. Beng and Matsumoto derived an estimate of the mean lifetime of all bridges in general in the US and Japan using the Weibull distribution. Their estimate was an expected lifetime of 93.5 years (p. 24). When compared to this figure, our estimate of the mean lifetime of a deficient bridge in New York seems reasonable, since time from deficiency classification to failure only represents a portion of the overall lifetime of a bridge.

Prediction Model Development:

The goal of this section of the research was to build a model which, given various levels of covariates, would predict how long a bridge would last once it was found to be deficient, assuming no repairs were made to change its deficiency status. The model building process was begun by selecting a subset of possible covariates believed to be correlated with the response, lifetime measured in years between deficiency classification and collapse. A subset of covariates was selected from among possible predictors given in the NBI. Dr. Wesley Cook, a civil engineer and former bridge inspector was consulted in the selection process.

In addition to selecting a subset of covariates, two data sets were created. A training set and a testing set. Forty two observations were selected at random from the 64 bridge failures. These observations were used as a training sample. The covariate values and lifetimes used to build the prediction model came from this data set. The remaining 22 observations were used to test the model, to see how accurately it predicted deficient lifetimes. By chance, it happened that each data set contained five uncensored observations.

It should be noted that the values of each covariate for each bridge are taken from the year in which the bridge collapsed, not the year in which it first became deficient, as would be preferable. Since the majority of the lifetimes in this study are censored, obtaining covariate values when a bridge first became deficient is not possible for most bridges. We could take covariate values for each censored lifetime from the year 1992, as close to time zero as we can get, but instead values are taken from the year of collapse. To take the values from 1992 would mean that we were taking values from bridges after they had been deficient for zero or ten or twenty or some other number of years. As opposed to taking covariate values at a different point in each bridge's lifetime, it was judged to be better to take the values from the one point common to all bridge lifetimes under study, their failure time.

The selected subset of covariates was regressed one at a time onto the response, using Weibull survival regression techniques. Those covariates found to be significant at the $\alpha = 0.25$ level were included in a preliminary model. Hosmer and Lemeshow argue that using a high significance level, such as $\alpha = 0.25$, at this point in model development provides a bulwark against missing a covariate that is either a confounder or is statistically significant (p. 160).

Table 1 in the Appendix shows the results of these bivariate analyses. The continuous and ordinal variables were analyzed in both their continuous or ordinal and discrete forms. If it was significant in either form, it was added to the preliminary model as a continuous or ordinal variable. Checking the scale of these covariates was carried out later. Observations with missing values of a covariate were not used in the analysis. The results show that the variables Average Daily Traffic, Average Daily Truck Traffic, Bridge Age, Maximum Span Length, Superstructure Condition, Deck Geometry Rating, Over Land or Water Indicator, Structural Evaluation, and Approach Roadway Alignment Evaluation were all significant at the $\alpha = 0.25$ level. These variables were added to a multiple survival regression model and checked for continued significance, this time at an $\alpha = 0.05$ level. The Fracture Critical Indicator was also significant at level $\alpha = 0.25$, but there were no observed survival times for a fracture critical bridge, only censored survival times. Thus the software was unable to estimate a standard error for the coefficient of this predictor. Because of this numerical problem it cannot be used further in the analysis.

The results of the multiple survival regression show that there are now numerical problems for the variables approach roadway alignment rating and over land or water indicator. Table 2 in the Appendix shows these results. Removing these two variables and refitting the model gives more stable estimates, with a log likelihood of -18.6. The results are shown in Table 3 below.

Table 3, Model 1

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	-0.000003	0.000044	-0.060900	0.951410
ADTTn	0.000018	0.000200	0.087400	0.930310
MaxSpan	-0.024700	0.021300	-1.157400	0.247090
Age	0.014700	0.011400	1.289300	0.197300
Sup	-0.214000	0.246000	-0.868500	0.385130
dgeom	-0.343000	0.171000	-2.002000	0.045280
strEval	0.590000	0.249000	2.372200	0.017680
shape	0.356000			
log(shape)	-1.030000	0.355000	-2.907600	0.003640
Log Likelihood	-18.6			

From Table 3, we can see that there are only two predictors that are significant at the $\alpha = 0.05$ level, deck geometry (dgeom) and structural evaluation (strEval). In order to not miss possible confounders, we will remove non-significant predictors one at a time, then refit the model, checking for significant changes in the coefficients and performing likelihood ratio tests to compare the new model to the previous one. A significant change in a coefficient will be a change of about 20 percent.

In the model above, it can be seen that average daily traffic (ADT) is the most non-significant coefficient, so we removed it first. After refitting the model there was a 67.4 percent change in the coefficient for average daily truck traffic. This is a very large change, however the coefficients for all other covariates changed by less than one percent, and the likelihood ratio test comparing the two models with and without average daily traffic gives a test statistic of 0.0036 with one degree of freedom and a p-value of 0.9524. This is a very large p-value, indicating that there is very little evidence that average daily traffic affects deficient bridge lifetime. Because of very little change in coefficients besides average daily truck traffic, the very high likelihood ratio test p-value, and because average daily traffic and average daily truck traffic may be collinear, we remove average daily traffic from the model. The results of fitting the new model are shown in Table 4 in the Appendix.

At this point, there are still only two variables that are significant, the same two as before, deck geometry and structural evaluation. The most non-significant covariate is now average daily truck traffic (ADTTn). This covariate is highly non-significant, with a Wald p-value of 0.91. We continue by removing ADTTn from the model and refitting. Comparing the new model to the previous one gives a log likelihood statistic 0.0129 with one degree of freedom and a p-value of 0.91, another highly non-significant result supporting our decision to remove ADTTn. Checking the percentage change for each of the coefficients, the coefficient with the greatest change was maximum span length (MaxSpan), with a reduction of about 7 percent. A table of these results is given in the Appendix in Table 5. Proceeding as before, we now remove superstructure rating (Sup), the now least significant coefficient. Results of the refit are in the Appendix under Table 6. Looking at the results we see that the maximum percentage change was for the structural evaluation coefficient (strEval). It dropped by roughly 16 percent. The likelihood ratio test comparing this model with the former gives a test statistic of 0.7019, with one degree of freedom and a p-value of 0.40, supporting the decision to remove it. There are still only two significant coefficients, structural evaluation (strEval) and deck geometry (dgeom). The current least significant coefficient is that of maximum span length (MaxSpan). Removing this from the model yields the results in Table 7 of the Appendix. With this removal there were no important changes in the coefficients. The likelihood ratio test comparing this model to the one in Table 6 gives a test statistic of 1.52 with one degree of freedom. The p-value is 0.22, and supports the removal of this covariate.

The only remaining covariates are Age, dgeom, and strEval. Age is the only non-significant coefficient (Wald p-value of 0.11), suggesting that deck geometry and the structural evaluation are driving the lifetime processes of deficient bridges in New York. Continuing in the usual manner, removing Age from the model gives the following result:

Table 8, Preliminary Model

Variable	Coeff.	Std. Err.	z	P>|z|
dgeom	-0.315000	0.116000	-2.710000	0.006680
strEval	0.510000	0.227000	2.250000	0.024700
shape	0.459000			
log(shape)	-0.778000	0.326000	-2.380000	0.017100
Log Likelihood	-21.300000			

No coefficient experienced an important change from the removal of age. The coefficient that changed the most was that of deck geometry, which changed from -0.29 to -0.32, a change of 7 percent. The likelihood ratio test comparing this model to the one including age has a test statistic of 3.2 with 1 degree of freedom. The p-value for the test is 0.0738, suggesting that age probably does not influence deficient bridge lifetime. This is an interesting result, given that it has already been shown that there is a significant difference in age between deficient and non-deficient bridges in New York.

At this point we have a final preliminary model, one containing only the covariates deck geometry rating and structural evaluation rating. The next step is to check the scale of these variables. Up to this point they have been treated as ordinal variables and not categorical, and we have assumed log survival time increases or decreases linearly as their ratings increase or decrease. To assess whether this assumption is true, each variable is broken into intervals, and then the model is refit with each interval used as a categorical variable. The estimated coefficients are shown below:

Estimated Coefficients for Categorical Variables

dgeom			strEval		
Interval	**Midpoint**	**Coefficient**	**Interval**	**Midpoint**	**Coefficient**
[0, 2)	1	0	[0, 3)	1.5	0
[2, 5)	3.5	1.382	[3, 4)	3.5	1.119
[5, 7)	6	0.032	[4, 5)	4.5	1.419
[7, 9]	8	-0.721	[5, 7]	6	1.567

The midpoint of each categorical variable is plotted against log time to assess its adherence to linearity. Note that the intervals and therefore midpoints are not evenly spaced. This is because when the model was fit with evenly spaced intervals, numerical instability caused standard errors of zero and infinity and very large coefficients. Again, this is due to there being very few uncensored data points with which to build the model. In order to avoid this instability it is necessary to allow the intervals to be unevenly spaced for the sake of having observed survival times in each interval. The graphs of log time versus the midpoints of these variables is given below in Figures 8 and 9.

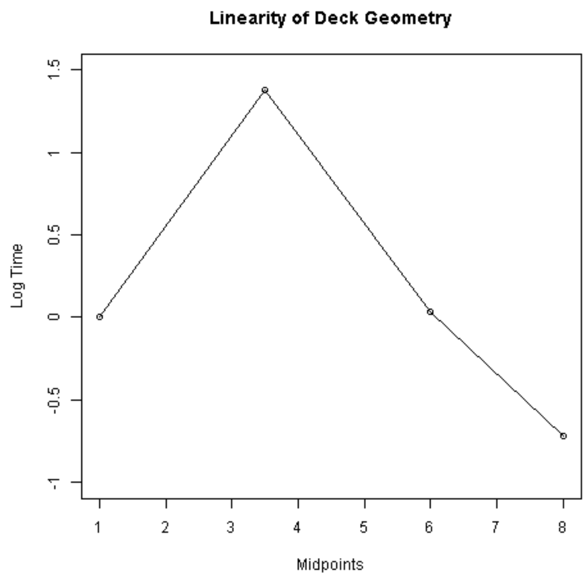

Figure 8: Log Time vs. Midpoints of Deck Geometry intervals.

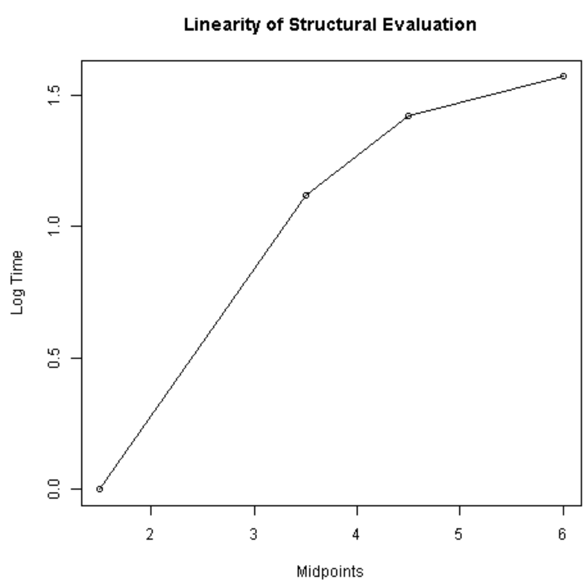

Figure 9: Log Time vs. Midpoints of Structural Evaluation intervals.

The graph for deck geometry increases at first, and then decreases linearly as the deck geometry rating increases. The likelihood ratio test comparing the model containing deck geometry as an ordinal variable to the one containing deck geometry as a categorical variable with the four different levels gives a test statistic of 0.954 with two degrees of freedom and a p-value of 0.621. This high p-value tells us that there is little evidence to suggest that use of the above design variables is better than the simpler model with deck geometry treated as an ordinal variable. The initial rise and then fall in the graph of deck geometry against log time is problematic. However, due to the greater simplicity in the model achieved by using a single linear version of deck geometry rating, and the fact that the graph shows a strong linear relationship after the initial rise, deck geometry rating will be treated as an ordinal variable.

The graph for structural evaluation rating appears to be linear in log time. A likelihood ratio test comparing the model with design variables to the one with only an ordinal variable gives a p-value of 0.05. This is a marginally significant result. Again, believing that a simpler model is better than a more complex one, and not having strong evidence that the more complicated model with categorical variables is better than a single ordinal variable, we will consider structural evaluation rating as a linear ordinal variable.

Now that the scale of these variables has been judged to be acceptable, we add back in the variables previously removed, one at a time, to ensure that they are neither significant nor confounders. After doing so, no variables were found to be significant. Average daily traffic and average daily truck traffic, however, both caused a 28 percent change in the value of the coefficient for deck geometry, signaling that they are confounders. Tables showing the results of adding back in each variable one at a time are given in the Appendix, tables 9 through 13. Adding average daily traffic and average daily truck traffic back into the model yields the one below:

Table 14, Final Preliminary Model Check 6

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	-0.000018	0.000044	-0.404000	0.686000
ADTTn	0.000033	0.000198	0.167000	0.867000
dgeom	-0.231000	0.125000	-1.840000	0.065800
strEval	0.453000	0.199000	2.278000	0.022700
shape	0.413000			
log(shape)	-0.884000	0.350000	-2.526000	0.011600
Log Likelihood	-20.600000			

Because adding back in ADT and ADTTn both caused the same percentage change in the coefficient for deck geometry, and because it is reasonable to assume that as average daily traffic increases on a road so does average daily truck traffic, we suspect again that these two variables may be collinear. To investigate this further, a scatter plot of average daily traffic versus average daily truck traffic is generated and shown below in Figure 10.

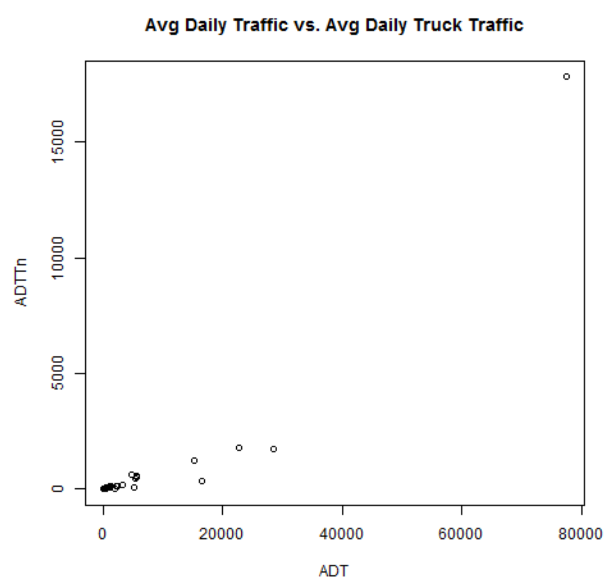

Figure 10: Scatter plot of average daily traffic vs. average daily truck traffic.

Looking at this scatter plot suggests a linear relationship between the two variables, however it is difficult to really see what's happening given the tight clustering of points near the origin. To try and spread things out and make the relationship more visible, the plot is generated again, this time using the logs of both variables. The result is shown in Figure 11.

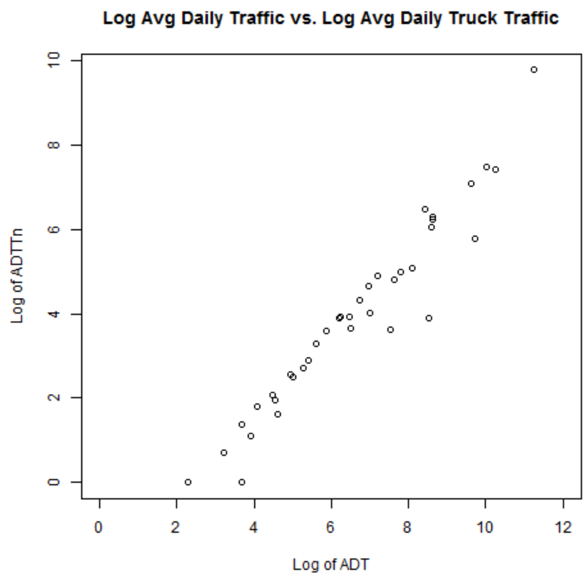

Figure 11: Plot of the log of average daily traffic vs. the log of average daily truck traffic. A very clear linear relationship is apparent in this graph.

In this graph the linear relationship between average daily traffic and average daily truck traffic is very clear. The coefficient of correlation between the logs of these two variables is 0.967, a very high value. Given these results, and since average daily truck traffic is believed by Dr. Cook to have a greater impact on bridge wear, average daily traffic will be dropped from the model, as it is collinear with average daily truck traffic. Table 15 below gives the results of refitting the model.

Table 15, Preliminary Model Check 7

Variable	Coeff.	Std. Err.	z	P>|z|
ADTTn	-0.000046	0.000040	-1.140000	0.254000
dgeom	-0.225000	0.124000	-1.820000	0.068400
strEval	0.460000	0.199000	2.310000	0.020900
shape	0.418000			
log(shape)	-0.873000	0.348000	-2.510000	0.012000
Log Likelihood	-20.700000			

The coefficient for average daily truck traffic experienced a significant change of 137 percent upon removal of average daily traffic. Neither coefficient for the other variables experienced an important change, however. The results from the likelihood ratio test comparing the models with and without average daily traffic gave a p-value of 0.72. Given this result and since we have already determined average daily traffic and average daily truck traffic to be collinear, we continue with the decision to not include average daily traffic in the model.

At this point in the model building process, we consider interactions. Each interaction term is added one at a time to the main effects model. Table 16 below shows every possible interaction with its respective likelihood ratio test p-value, comparing each interaction model to the main effects only model.

Table 16, Interactions

Interaction	Variables	df	LRT p-Value
ADTTn	dgeom	1	0.7403
	strEval	1	0.0003953
dgeom	strEval	1	0.3937

Only one interaction term is significant, ADTTn*strEval. This term will therefore be added to the model. Thus far, we have constructed a final preliminary model to predict the lifetime of a deficient bridge. Table 17 below gives the parameter estimates for each term in the model.

Table 17, Final Preliminary Model

Variable	Coeff.	Std. Err.	z	P>|z|
Intercept	1.770410	0.303600	5.830000	0.000000
ADTTn	0.014270	0.005930	2.410000	0.016200
dgeom	-0.337420	0.170390	-1.980000	0.047700
strEval	0.928820	0.383010	2.430000	0.015300
ADTTn*strEval	-0.003580	0.001480	-2.410000	0.015700
shape	0.274000			
log(shape)	-1.295910	0.342330	-3.790000	0.000153
Log Likelihood	-14.400000			

Goodness of Fit:

Next we perform a goodness of fit assessment to see how well our model fits the data. In a parametric survival regression setting, instead of testing for a significant interaction between time and each covariate, we compare the parametric cumulative hazard function to the non-parametric cumulative hazard function. This is done by plotting the values of the nonparametric cumulative hazard function at each time point, against the values of the parametric cumulative hazard function at the same time points. If the model is a good fit, then the points should line up roughly along the line $y = x$, indicating that the values of each version of the cumulative hazard are roughly the same, and thus the model fits the data well. Figure 12 shows a graph of the cumulative hazard functions plotted against each other. The dashed line is the line $y = x$, given as a reference. The points represent the ordered pairs $(\widehat{H}_{pm}, \widehat{H}_{KM})$for the different values of time, where $\widehat{H}_{pm}$ is the cumulative hazard for the parametric model, and $\widehat{H}_{KM}$ is the cumulative hazard from the Kaplan-Meier estimator. The blue line represents the Lowess smooth of these points.

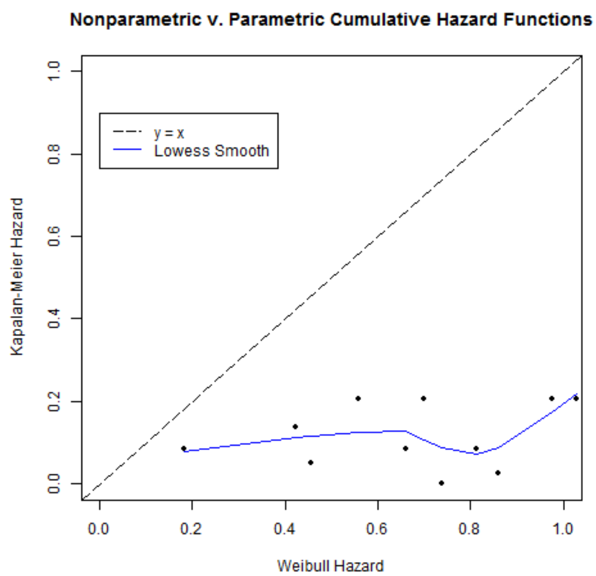

Figure 12: Plot of the estimated cumulative hazard function values from a Kaplan-Meier estimator vs. a Weibull estimator. Each dot represents an ordered pair of cumulative hazard function values. The blue line is their Lowess smooth. The dashed line is the reference line $y = x$.

It is clear that the parametric and Kaplan-Meier versions of the cumulative hazard function do not agree. If the model were a good fit to the data, all points would fall roughly on or near the dashed line. Such is not the case in this plot. The Weibull hazard is much greater than the Kaplan-Meier estimator at each time value. This indicates that the parametric model does not fit the data well.

In an effort to improve model fit, square root transformations were applied to each variable one at a time and the model refit. No improvement was made by so doing. The square root and log transformations were both applied to the response variable, but still the cumulative hazard estimates did not agree. It is believed that the poor fit results from attempting to build a model based on only five uncensored observation times. It is likely that a data set with a higher percentage of uncensored observations would result in a better fitting model.

Testing the Model:

Despite its shortcomings, the model was run on the test data set to see how closely it predicted the actual results. It should be noted that the values of each variable for each bridge fell within the range of values used in the training set to create the model. Table 18 gives the results, showing lifetimes in years. The bolded entries are the uncensored survival times.

Table 18, Test Results

Bridge	Actual Deficient Lifetime	Predicted Deficient Lifetime	Censor
1	3	1,313	0
2	3	57	0
3	3	25	0
4	5	782	0
5	5	28	0
6	5	13	0
7	8	55	0
8	9	0	0
9	10	87	0
10	10	738	0
11	11	4	0
12	12	35	0
13	14	3	0
14	**4**	**44**	**1**
15	**3**	**28**	**1**
16	18	293	0
17	**6**	**88**	**1**
18	20	86	0
19	**7**	**13**	**1**
20	20	168	0
21	**3**	**25**	**1**

From this table it is clear that our model is not giving accurate predictions. Some are obviously very high, some are high but plausible, and others are low. Since our model was built primarily from censored data, or lifetime values that were at least as large as some number of years, we would expect most predictions to err on the high side. This holds true for most cases, but there are three where it does not, bridges 8, 11, and 13. Most notably is the case for bridge number 8. Here, our model predicts a deficient lifetime of 0, but the true deficient lifetime is at least nine years.

Conclusions:

Through the application of common statistical techniques this paper has been able to show a significant difference in age between deficient and non-deficient bridges in New York for the 2012 bridge population, with deficient bridges tending to be older than non-deficient bridges. There was also found strong evidence that deficient bridges are more at risk of failure than non-deficient bridges. These two results were found using analysis of variance and the binomial distribution, respectively.

From survival analysis, a Weibull probability density function was derived modeling the distribution of deficient bridge lifetimes in New York. From this function, we were able to estimate the mean deficient bridge lifetime as 47.2 years. Using the probability density function, we were able to derive a survival function. This curve shows an estimate of the median survival time of New York's deficient bridges to be 36.6 years.

Lastly an attempt was made to construct a model to predict lifetimes of New York's deficient bridges based on a bridge's average daily truck traffic, its deck geometry and structural evaluation ratings. The model was found to poorly fit the data, and no transformations were identified to improve the fit. Nonetheless, a test data set was available, so the model was used to predict the lifetimes of 21 deficient bridges that had collapsed between 1992 and 2012. When comparing the predicted survival times to the actual censored and uncensored survival times, the predictions were not accurate.

Future Studies:

Redoing this study with pre-1992 NBI data would help immensely in the parametric regression and model building processes. It is believed that more complete data could lead to different estimated mean and median survival times and a different prediction model. It is believed that doing so would result in a model that better fit the data and gave better predictions.

A related study that would yield interesting information could be a survival analysis of bridges from construction to deficiency status. Still another similar study could look at the lifetime of deficient bridges from deficiency classification to either failure or repair. Yet another interesting study would be a logistic regression model describing the probability that a bridge is deficient based on various factors such as bridge age, traffic load, length, etc.

Appendix:

Table 1, Discrete Covariate Description

Variable	Category	Description	LRT p-Value
ADTcut	light	[0, 1000)	0.420
	medium	[1000, 10000)	
	heavy	[10000, 100000)	
ADTTncut	light	[0, 50)	0.180
	medium	[50, 500)	
	high	[500, 20000)	
AgeCut	young	[0, 39)	0.056
	old	[39, 250)	
MaxSpanCut	short	[0, 20.11m)	0.200
	long	[20.11m, 101m)	
Ddef	N	Non-Deficient	0.820
	Y	Deficient Deck	
SupDef	N	Non-Deficient	0.051
	Y	Deficient Superstructure	
SubDef	N	Non-Deficient	0.470
	Y	Deficient Substructure	
dgDef	N	Non-Deficient	0.006
	Y	Deficient Deck Geometry	
scourCut	N	Not Scour Critical	0.640
	Y	Scour Crititcal	
frac	N	Not Fracture Critical	0.100
	Y	Fracture Critical	
unwat	N	Underwater Inspection Not Required	0.350
	Y	Required	
over	land	Bridge is Over Land	0.073
	water	Bridge is Over Water	

Table 1, Discrete Covariate Description

Variable	Category	Description	LRT p-Value
mat2	steel	Bridge Made of Steel	0.730
	other	Bridge Made of Other Material	
dgn2	beam	Beam Design	0.800
	other	Other Design	
DT2	conc	Deck Structure is Concrete	0.870
	other	Deck Structure is Other Material	
SF2	conc	Wearing Surface is Concrete	0.510
	other	Wearing Surface is Other Material	
DP2	none	No Deck Protection System	0.330
	has	Has a Deck Protection System	
InpFrq	12	Bridge Inspected Every 12 Months	1.000
	24	Bridge Inspected Every 24 Months	
strEvalDef	N	Non-Deficient	0.120
	Y	Deficient Structural Evaluation	
apEvalDef	N	Non-Deficient	0.250
	Y	Deficient Approach Roadway Alignment	
wtEvalnDef	N	Non-Deficient	0.700
	Y	Deficient Waterway Adequacy	

Table 1, Continuous/Ordinal Covariate Description

Variable	Description	LRT
ADT	Average Daily Traffic	0.023
ADTTn	Average Daily Truck Traffic	0.023
MaxSpan	Maximum Span Length	0.700
Age	Age of Bridge	0.160
Deck	Deck Condition Rating	0.350
Sup	Superstructure Condition Rating	0.180
Sub	Substructure Condition Rating	0.870
dgeom	Deck Geometry Rating	0.450
scour	Scour Crital Briges	0.820
StrEval	Structural Evaluation	0.061
apprEval	Approach Roadway Alignment	0.016
waterEval	Water Way Adequacy	0.098

Table 2

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	0.000014	0.000046	0.298700	0.765000
ADTTn	0.000007	0.000222	0.030400	0.976000
MaxSpan	-0.014500	0.017733	-0.817500	0.414000
Age	0.021100	0.010248	2.058100	0.039600
Sup	-0.120000	0.121762	-0.987500	0.323000
dgeom	-0.507000	0.132693	-3.817800	0.000135
strEval	0.676000	0.162751	4.151900	0.000033
apprEval	-0.648000	0	-Inf	0
over (water)	-1.010000	0	-Inf	0
shape	0.248000			
log(shape)	-1.390000	0	-Inf	0
Log Likelihood	-14.8			

Table 3, Model 1

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	-0.000003	0.000044	-0.060900	0.951410
ADTTn	0.000018	0.000200	0.087400	0.930310
MaxSpan	-0.024700	0.021300	-1.157400	0.247090
Age	0.014700	0.011400	1.289300	0.197300
Sup	-0.214000	0.246000	-0.868500	0.385130
dgeom	-0.343000	0.171000	-2.002000	0.045280
strEval	0.590000	0.249000	2.372200	0.017680
shape	0.356000			
log(shape)	-1.030000	0.355000	-2.907600	0.003640
Log Likelihood	-18.6			

Table 4, Model 2

Variable	Coeff.	Std. Err.	z	P>|z|
ADTTn	0.000006	0.000051	0.113000	0.909890
MaxSpan	-0.024800	0.021200	-1.169000	0.242380
Age	0.014700	0.011300	1.298000	0.194350
Sup	-0.216000	0.243000	-0.890000	0.373370
dgeom	-0.342000	0.170000	-2.014000	0.044000
strEval	0.592000	0.247000	2.396000	0.016570
shape	0.357000			
log(shape)	-1.030000	0.354000	-2.911000	0.003610
Log Likelihood	-18.600000			

Table 5, Model 3

Variable	Coeff.	Std. Err.	z	P>|z|
MaxSpan	-0.023100	0.015000	-1.541000	0.123399
Age	0.014300	0.010600	1.346000	0.178420
Sup	-0.210200	0.239200	-0.879000	0.379494
dgeom	-0.326400	0.098100	-3.328000	0.000875
strEval	0.576600	0.205700	2.803000	0.005066
shape	0.356000			
log(shape)	-1.033100	0.354200	-2.917000	0.003536
Log Likelihood	-18.600000			

Table 6, Model 4

Variable	Coeff.	Std. Err.	z	P>|z|
MaxSpan	-0.019900	0.014400	-1.390000	0.166000
Age	0.015500	0.010300	1.490000	0.135000
dgeom	-0.285600	0.084200	-3.390000	0.000699
strEval	0.487100	0.172000	2.830000	0.004620
shape	0.373000			
log(shape)	-0.987400	0.343600	-2.870000	0.004050
Log Likelihood	-18.900000			

Table 7, Model 5

Variable	Coeff.	Std. Err.	z	P>|z|
Age	0.016400	0.010400	1.580000	0.114426
dgeom	-0.293600	0.103700	-2.830000	0.004634
strEval	0.500100	0.200800	2.490000	0.012773
shape	0.399000			
log(shape)	-0.917800	0.330300	-2.780000	0.005456
Log Likelihood	-19.700000			

Table 8, Preliminary Model

Variable	Coeff.	Std. Err.	z	P>|z|
dgeom	-0.315000	0.116000	-2.710000	0.006680
strEval	0.510000	0.227000	2.250000	0.024700
shape	0.459000			
log(shape)	-0.778000	0.326000	-2.380000	0.017100
Log Likelihood	-21.300000			

Table 9, Preliminary Model Check 1

Variable	Coeff.	Std. Err.	z	P>|z|
MaxSpan	-0.018400	0.012300	-1.500000	0.134000
dgeom	-0.309500	0.103200	-3.000000	0.002700
strEval	0.490900	0.209500	2.340000	0.019100
shape	0.427000			
log(shape)	-0.851300	0.334700	-2.540000	0.011000
Log Likelihood	-20.500000			

Table 10, Preliminary Model Check 2

Variable	Coeff.	Std. Err.	z	P>|z|
Sup	-0.170000	0.266000	-0.640000	0.522330
dgeom	-0.351000	0.134000	-2.620000	0.008760
strEval	0.571000	0.246000	2.320000	0.020500
shape	0.449000			
log(shape)	-0.800000	0.334000	-2.400000	0.016470
Log Likelihood	-21.100000			

Table 11, Preliminary Model Check 3

Variable	Coeff.	Std. Err.	z	P>|z|
ADTTn	-0.000046	0.000040	-1.140000	0.254000
dgeom	-0.225000	0.124000	-1.820000	0.068400
strEval	0.460000	0.199000	2.310000	0.020900
shape	0.418000			
log(shape)	-0.873000	0.348000	-2.510000	0.012000
Log Likelihood	-20.700000			

Table 12, Preliminary Model Check 4

Variable	Coeff.	Std. Err.	z	P>|z|
Age	0.016400	0.010400	1.580000	0.114426
dgeom	-0.293600	0.103700	-2.830000	0.004634
strEval	0.500100	0.200800	2.490000	0.012773
shape	0.399000			
log(shape)	-0.917800	0.330300	-2.780000	0.005456
Log Likelihood	-19.700000			

Table 13, Preliminary Model Check 5

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	-0.000011	0.000009	-1.190000	0.235000
dgeom	-0.226000	0.121000	-1.870000	0.061800
strEval	0.454000	0.199000	2.290000	0.022300
shape	0.414000			
log(shape)	-0.882000	0.349000	-2.520000	0.011600
Log Likelihood	-20.600000			

Table 14, Preliminary Model Check 6

Variable	Coeff.	Std. Err.	z	P>|z|
ADT	-0.000018	0.000044	-0.404000	0.686000
ADTTn	0.000033	0.000198	0.167000	0.867000
dgeom	-0.231000	0.125000	-1.840000	0.065800
strEval	0.453000	0.199000	2.278000	0.022700
shape	0.413000			
log(shape)	-0.884000	0.350000	-2.526000	0.011600
Log Likelihood	-20.600000			

Table 15, Preliminary Model Check 7

Variable	Coeff.	Std. Err.	z	P>|z|
ADTTn	-0.000046	0.000040	-1.140000	0.254000
dgeom	-0.225000	0.124000	-1.820000	0.068400
strEval	0.460000	0.199000	2.310000	0.020900
shape	0.418000			
log(shape)	-0.873000	0.348000	-2.510000	0.012000
Log Likelihood	-20.700000			

Table 16, Interactions

Interaction	Variables	df	LRT p-Value
ADTTn	dgeom	1	0.7403
	strEval	1	0.0003953
dgeom	strEval	1	0.3937

Table 17, Final Preliminary Model

Variable	Coeff.	Std. Err.	z	P>|z|
Intercept	1.770410	0.303600	5.830000	0.000000
ADTTn	0.014270	0.005930	2.410000	0.016200
dgeom	-0.337420	0.170390	-1.980000	0.047700
strEval	0.928820	0.383010	2.430000	0.015300
ADTTn*strEval	-0.003580	0.001480	-2.410000	0.015700
shape	0.274000			
log(shape)	-1.295910	0.342330	-3.790000	0.000153
Log Likelihood	-14.400000			

Table 18, Test Results

Bridge	Actual Deficient Lifetime	Predicted Deficient Lifetime	Censor
1	3	1,313	0
2	3	57	0
3	3	25	0
4	5	782	0
5	5	28	0
6	5	13	0
7	8	55	0
8	9	0	0
9	10	87	0
10	10	738	0
11	11	4	0
12	12	35	0
13	14	3	0
14	**4**	**44**	**1**
15	**3**	**28**	**1**
16	18	293	0
17	**6**	**88**	**1**
18	20	86	0
19	**7**	**13**	**1**
20	20	168	0
21	**3**	**25**	**1**

Bibliography:

1. 2013 Report Card for America's Infrastructure, Bridges. American Society of Civil Engineers, 2013. Web. 27 March 2014.
2. Beng, Si Soon and Matsumoto, Takashi. Survival analysis on bridges for modeling bridge replacement and evaluating bridge performance, Structure and Infrastructure Engineering 8.3 (2012): 251-268
3. Cook, Wesley PhD PE. Former Bridge Inspector for Center for Transportation and Research, New Mexico State University. Personal emails. December 17, 2013 – February 27, 2014.
4. Dekelbab, Waseem; Al-Wazeer, Adel; Harris, Bobby. "History Lessons from the National Bridge Inventory." *Public Roads.* Federal Highway Administration, May/June 2008, Web. 29 March 2014.
5. Hosmer, David W., Jr. and Stanley Lemeshow. *Applied Survival Analysis.* New York: John Wiley & Sons, Inc., 1999. Print.
6. Klein, John P., and Moeschberger, Melvin L. *Survival analysis: techniques for censored and truncated data.* New York: Springer, 1997. Web.
7. Peterson, A. V. "Expressing the Kaplan-Meier Estimation as a Function of Empirical Subsurvival Functions." *Journal of the American Statistical Association* 72 (1977): 854 – 858. Print.
8. Tabatabai, Habib; Tabatabai, Mohammad; Lee, Chin-Wei. "Reliability of Bridge Decks in Wisconsin." Journal of Bridge Engineering 16.1 (2011): 53 – 62. Print.
9. Therneau, T (2013). *A Package for Survival Analysis in R.* R package version 2.37-4, http://CRAN.R-project.org/package=survival.

10. U.S. Department of Transportation, Federal Highway Administration, Bridges and Structures, Deficient Bridges by State and Highway System. Web. 10 October 2013

11. Winchell, Auyeung. Bridge Safety Assurance Unit NYSDOT – Office of Structures. Personal email. September 12, 2013.