

**A Statistical Analysis of Witchcraft
Accusations in Colonial America: A Time
Series Count Data Analysis**

by

Tara M. McCart

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Mathematics

Program

YOUNGSTOWN STATE UNIVERSITY

May, 2014

A Statistical Analysis of Witchcraft Accusations in Colonial America:
A Time Series Count Data Analysis

Tara M. McCart

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Tara M. McCart, Student Date

Approvals:

Dr. G. J. Kerns, Thesis Advisor Date

Dr. Andy Chang, Committee Member Date

Dr. Diane Barnes, Committee Member Date

Dr. Sal Sanders, Associate Dean of Graduate Studies Date

©

Tara M. McCart

2014

ABSTRACT

The idea of magic and witches was not foreign to the society of Colonial America. Many historical studies have examined the social, political, and economic aspects of witch hunts in Colonial America, especially the infamous Salem witch hunt. During the Salem witch hunt of 1692, over 150 people were accused of practicing witchcraft, 19 people were executed, and countless others were victimized. The object of this paper is to analyze witchcraft accusations of seventeenth-century Colonial America to discover any underlying patterns. Time series count data of witchcraft accusations were used to conduct an intervention analysis to determine whether the Salem witch hunt had an effect on witchcraft accusations in Colonial America.

Acknowledgements

I would like to thank Dr. Kerns, Dr. Chang, and Dr. Barnes for their help with this thesis.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Methodology | 14 |
| 2.1 | Event Count Models | 14 |
| 2.2 | Poisson Regression Model | 15 |
| 2.2.1 | The Poisson Distribution | 15 |
| 2.2.2 | Specification of the Poisson Regression Model | 15 |
| 2.2.3 | Estimation of the Poisson Regression Model | 16 |
| 2.3 | Negative Binomial Model | 17 |
| 2.3.1 | Negative Binomial Distribution | 17 |
| 2.3.2 | Specification of the Negative Binomial Model | 17 |
| 2.3.3 | Estimation of the Negative Binomial Model | 19 |
| 2.4 | Poisson Autoregressive Model of Order p | 19 |
| 2.4.1 | Par(p) Model Identification and Specification | 20 |
| 2.4.2 | Par(p) Model Estimation | 23 |
| 2.4.3 | Interpretation of the PAR(p) Model | 26 |
| 2.5 | Intervention Analysis | 28 |
| 2.5.1 | Time Series Subject to Intervention | 29 |
| 2.6 | Diagnostic Tests | 32 |
| 2.6.1 | Normality | 32 |
| 2.6.2 | Independence | 33 |
| 3 | Analysis | 35 |
| 3.1 | Witchcraft Accusations September 1622 to December 1712 | 35 |

| | | |
|----------|--|-----------|
| 3.2 | Descriptive Statistics | 36 |
| 3.2.1 | Biographical Portrait | 36 |
| 3.2.2 | A Description of Accusations | 42 |
| 3.3 | Time Series of Event Counts | 47 |
| 3.3.1 | The Data | 47 |
| 3.3.2 | Model Specification | 51 |
| 3.4 | Intervention Analysis | 52 |
| 3.4.1 | Determining the Intervention Period | 52 |
| 3.4.2 | Estimating the Parameters of the Intervention Analysis | 53 |
| 3.4.3 | Model Diagnostics | 56 |
| 4 | Discussion | 63 |
| 5 | Conclusion | 68 |
| | References | 70 |
| A | Complete R Scripts | 75 |
| B | Complete List of Sources Used to Compile Data Set | 78 |

1 Introduction

In pursuit of a topic that merges history and statistics, the *Salem Witchcraft Site*, a web site dedicated to the study of the Salem witch hunt that was constructed and maintained by Dr. Richard Latner of Tulane University, inspired an exploration in to studying Colonial American witchcraft accusations with a statistical approach [24]. Dr. Latner, professor of history, specializes in Jacksonian America; Sectionalism and Civil War; Salem Witchcraft, and Digital Humanities. Dr. Latner used primary documents, many transcribed and edited by historians Paul Boyer and Steven Nissenbaum in their work *The Salem Witchcraft Papers: Verbatim Transcripts of the Legal Documents of the Salem Witchcraft Outbreak of 1692,3*, to construct data sets on the Salem witchcraft outbreak [3]. He used these data sets in his own work, *Salem Witchcraft, Factionalism, and Social Change Reconsidered: Were Salem's Witch-Hunters Modernization's Failures?* that was published in 2008 [23]. The Salem Witchcraft Site made these data sets available to the general public, and Dr. Latner encouraged examining them for further study. Having historical data available, in a ready to analyze form, is a rarity. Thus the Salem Witchcraft Site is a treasure for a descriptive statistics study on the events of the Salem witch hunt. These data sets inspired thoughts on how Colonial American witchcraft could be studied with advanced statistics, going beyond descriptive statistics.

During a historiographical study of witchcraft in Colonial America, it came to light that the most infamous episode of witchcraft accusations, the Salem witch hunt, has been extensively studied in comparison to witchcraft accusations on the whole throughout Colonial America. Historians agree that the origins of the Salem witchcraft outbreak are obscure but the original spark began in the home of Salem

Village's minister, Samuel Parris, when his daughter and niece seemed to be vexed by curious forces that could not be relieved by the usual means of fasting and prayer. These events led to more citizens exhibiting strange and unusual behavior and then led to a snowball effect of accusations and executions. In less than a years time, accusations spread over 20 communities in Massachusetts, over 150 people were jailed, a number of whom had confessed and many more were most likely accused, even though no legal record remains. Nineteen people were convicted and hanged, others found guilty and detained in prison, and Giles Corey, was pressed to death during interrogation. A few, including the baby of an accused witch, died from conditions in prison [24].

A want to study the social, economic, and political aspects of this hysteria motivated many historians to dedicate decades of hard work and resilience to the study of the 1692-93 episode in Salem, Massachusetts. Many historians posit that there were many underlying social and economic factors that added to the severity and extensiveness of the Salem witch hunt and they present well studied and thought out evidence to support these ideas. Mary Beth Norton's *In the Devil's Snare* [28] and Marion L. Starkey's *The Devil in Massachusetts* [33] provided solid historical analyses of multiple aspects of the Salem witchcraft outbreak including social, political, and religious aspects. These works were meticulously researched and provided arguments as to why the outbreak was so severe. Charles Upham, Nissenbaum, Boyer, and Latner's works, *Salem Witchcraft*, *Salem Possessed*, and *Salem Witchcraft, Factionalism and Socialism Reconsidered* focused on factionalism, or division within the community, as one reason for the wide spread influence of the Salem outbreak [35]. Historians like Latner, Nissenbaum, and Boyer among many others, discussed the Salem witchcraft outbreak as a unique episode in American history, that was unlike any other kind of

witchcraft accusations. Some historians question whether the Salem trials constituted anything more than an inconsequential episode in colonial history. The Salem witch hunt was studied extensively from different perspectives such as “divine retribution, fraud, class conflict, village factionalism, cultural provincialism, and hysteria [10].”

The ideas of magic and witchcraft crossed the Atlantic ocean with the first European colonists coming to America, over seventy years before the Salem witch hunt. It is important to note that ideas about magic and witchcraft were prevalent in the society of Colonial America. A large portion of Colonial America was populated by people of the Puritan faith. The Puritans were descendants of the Protestant Revolution and focused on more direct relationships with God, a deepened sense of human sinfulness, and simplified ritual. Citizens who followed Puritan ideas believed in living lives of virtue and that the Devil played a major role in “bad” deeds. The ideology that human history unrolled during a battle between good and evil, God versus the Devil, where the battle ended when God triumphed, was a main focus of the Puritan belief system [11].

This thesis examines witchcraft accusations throughout seventeenth-century Colonial America, their patterns, and how the Salem outbreak influenced the overall pattern of witchcraft accusations. John Demos presented an extensive study of witchcraft in the western world in his work *The Enemy Within: 2,000 Years of Witch-hunting in the Western World* [10]. Early cases of witchcraft accusations in Colonial America began in the early 1600s, occurred throughout the colonies in “specific, local, small-scale” incidents, and had roots in the local culture [10]. Unlike the Salem outbreak, witchcraft accusations occurred on a much smaller scale throughout the rest of the seventeenth-century in other parts of the colonies. In *Enemy Within*, Demos described the “typical witch,” the profile of typical accusers, and aspects of the communities

where accusations took place. His classifications of similarities within the groups of the accused, accusers, and communities ignites curiosity about possible mathematical patterns of the accusations. Demos also argued that the flow of witchcraft accusations had peaks, troughs, and oscillations caused by forms of community distress like extreme weather, factional struggle, and outside influences such as Indian Wars. This idea leads to the question of whether or not this theory can be modeled mathematically.

After an extensive literature review in order to study the historiography on the subject of witchcraft in Colonial America, the research question, how can witchcraft accusations be modeled mathematically, was posed. How did witchcraft accusations behave over time, as time series count data and can this data be modeled? How do witchcraft patterns in Colonial America differ when the witch hunt in Salem is taken out of the equation? Few works have gone beyond reporting frequencies as a statistical analysis dealing with witchcraft accusations in Colonial America. Some historians utilized descriptive statistics, rank and percentile analysis, and age analysis to aide in their historical analyses.

Stephen Nissenbaum and Paul Boyer edited and compiled a wealth of primary documents from the Salem outbreak in their work *The Salem Witchcraft Papers* [3]. Using these numerous primary documents, Nissenbaum and Boyer analyzed tax documents, court records, and wills to provide evidence of factionalism in Salem Village as one of the main reasons for the intensity of the Salem witch hunt. They argued that there was a social divide in Salem Village between those citizens who supported the traditional agricultural society and those citizens who supported the newer, more merchant capitalistic society. This divide was further instilled in Salem Village by the fact that many of the supporters of the accusers during the witch hunt were support-

ers of the traditional agricultural society and those who supported the accused tended to be supporters of the newer merchant society [2]. Their in-depth economic analysis, dealing with tax information from 1695 tax documents, employed rank and percentile analysis to determine that there was an economic divide between the two factions defined above. Similarly, Richard Latner also used some statistics in his analysis of Salem Village factionalism.

Latner's *Factionalism and Social Change Reconsidered* is an analysis on factionalism in Salem Village the decade before, during, and after the Salem witch hunt. This paper, published in *Mary and William Quarterly*, challenged Nissenbaum and Boyer's *Salem Possessed* [23]. Using tax documents from 1681, 1695, and 1701, and other primary sources, Latner used percentile and rank analysis to conclude that while factionalism did exist in Salem Village, those who supported the accusers of the 1692 outbreak were economically better off than those who supported the accused. This was the opposite of the findings of *Salem Possessed*. Latner's use of more cohesive data in his analysis grant him a stronger argument. Latner's Salem Witchcraft Site also employed descriptive statistics such as histograms and scatter plots when analyzing the events in the Salem outbreak [24].

When examining studies of witchcraft in all of Colonial America, Carol Karlsen, professor Emeritus of history and women studies at the University of Michigan, employed descriptive statistics to describe age, social status, and marital status among those accused of witchcraft in her book *The Devil in the Shape of a Woman: Witchcraft in Colonial New England* [19]. She examined the social and economic bases of witchcraft and how those factors influenced who was accused and executed. Her organized and visual presentation of frequencies of age, social status, and other descriptors of those accused provided a concise illustration of those victimized in witch

hunts. Another, among many, amazing historical accounts of witchcraft in Colonial America that incorporates descriptive statistics in order to illustrate the social demographics of witchcraft accusations is *Entertaining Satan: Witchcraft and the Culture of Early New England* by John Putnam Demos. In *Entertaining Satan*, Demos used descriptive statistics to convey age, gender, marital status, family size, type of crime committed, and social status of those accused of witchcraft, as what he called a “Collective Portrait” of those accused [11]. Demos’ “Collective Portrait” influenced the descriptive statistics of this thesis. The descriptive statistics, like those utilized by Karlsen and Demos, are great aides in understanding attributes of those accused of witchcraft.

Emily Oster, associate professor of economics at the University of Chicago, performed a more advanced statistical analysis of witchcraft, using time series analysis. *Witchcraft, Weather and Economic Growth in Renaissance Europe* focused on witchcraft trials in Europe over five centuries. Oster argued that there was a direct relationship between witchcraft trials in Europe and weather patterns [29]. She further extended her analysis to determine a direct relationship between witchcraft trials and periods of economic downturn. Due to the large amount of data available on European witchcraft trials and weather patterns, Oster was able to conduct a time series analysis that used data that spanned over multiple centuries. This paper was very influential for this thesis.

In order to study patterns in witchcraft accusations in Colonial America, a data set was needed for analysis. Inspired by the Salem Witchcraft Site, it was determined that certain information about those accused of witchcraft during the seventeenth-century would be needed. The main objective of data collection was to determine those who had been accused of witchcraft in Colonial America, the month and year they were

accused, their gender, their marital status, and attributes of said accusations like whether one was executed, imprisoned, or cleared. The goal is to use advanced time series analysis to examine patterns in witchcraft accusations in Colonial America, with and without data from the Salem witch hunt, in an effort to mathematically model how accusations occurred.

The *Accused Witches* data set from the Salem Witchcraft Site was used not only as inspiration, but also as a starting block for data collection on those accused of witchcraft in the Salem outbreak. Paul Boyer and Stephen Nissenbaum's *The Salem Witchcraft Papers: Verbatim Transcripts of the Legal Documents of the Salem Witchcraft Outbreak of 1692,3* and *Salem-Village Witchcraft: A Documentary Record of Local Conflict in Colonial New England* along with *Records of the Salem Witch-Hunt* edited by Bernard Rosenthal, aided Richard Latner in compiling the data sets for the Salem Witchcraft site [30]. Latner classified those included in the *Accused Witches* data set as those who were formally accused of witchcraft during the Salem episode. This means that there exists evidence of some form of direct legal involvement, or court record. For a more cohesive data set, the same classifications were used when compiling the data set for this thesis [24]. In order to gather data on seventeenth century witchcraft accusations for all of Colonial America, an extensive study of well researched, secondary sources was conducted.

Works by Richard Godbeer, Lyle Koehler, John Putnam Demos, and other leading historians on witchcraft in Colonial America were used to compile the data set. In Richard Godbeer's *The Devil's Dominion: Magic and Religion in Early New England* [13], Lyle Koehler's *A Search for Power: The "Weaker" Sex in Seventeenth-Century New England* [22], and John Putnam Demos' *Entertaining Satan: Witchcraft and the Culture of Early New England* [11] comprehensive lists of those people accused of

witchcraft in Colonial America from about 1622 to 1695 appeared in the appendices. Richard Godbeer, professor of history at University of Miami, specializes in Colonial and Revolutionary America, with an emphasis on religious culture, gender studies, and the history of sexuality. *The Devil's Dominion* is a well researched book about magic in the context of Colonial New England society. It contains two lists of those accused of witchcraft in seventeenth-century America, one dealing with those accused that were not in the Salem witch hunt and then one dealing with those accused during the Salem outbreak [13]. Historian Lyle Koehler extensively researched the role of women in Puritan society in Colonial America. In *A Search for Power*, Koehler included an appendix with an extensive list of those accused of witchcraft in Colonial America from 1638 to 1697 [22]. John Putnam Demos, American historian and retired Yale professor, extensively studied witchcraft in seventeenth century America. In *Entertaining Satan*, Demos compiled a list of known witchcraft cases in seventeenth-century New England from 1638 to the late 1690s but did not include those of the Salem witch hunt.

The data set for this thesis was compiled by cross referencing the list of accused witches in Godbeer's *The Devil's Dominion*, Koehler's *A Search for Power*, Demos' *Entertaining Satan*, and Richard Latner's *Accused Witches* data set. Once a single exhaustive list was finalized, a plethora of historical sources were used to determine the month and year of accusations, the accused's gender and marital status, and whether an accused person was imprisoned, executed, or cleared. A list of these sources are included in Appendix B. Like in Latner's *Accused Witches* data set, a person was included in the data set if there exists evidence of some form of formal accusation but we also include cases where there was record of filing for a suit of defamation of character due to being accused of witchcraft. Similarly, data on imprisonment, exe-

cution, and acquittal were recorded if evidence was found. If no record was available for a given variable, that entry was recorded as *not applicable*. Dates were recorded as they were found in record, meaning that original dates were used and not changed, as to avoid any kind of confusion dealing with Julian and Gregorian calendars. When dates were in question, the best historical estimate of month and year were used. This data set was meticulously compiled, but is by no means complete. There were names of those accused of witchcraft on the lists created by Godbeer, Koehler, and Demos that could not be found in records, thus those people were excluded from the data set.

The number of witchcraft accusations per month from September of 1622 to December 1712 was recorded. In this case, we have event count data, specifically time series of event counts. Event counts are dependent variables that take on non-negative integer values for each of the n observations. In the case of this thesis, an event is a witchcraft accusation and the time period being considered is a month. The statistical analysis of counts and time series of counts, has been studied extensively over the past fifty years.

When dealing with models of event count data, estimation and inference are concerned with unknown parameters given the probability distribution from which the counts arise. Often count data arises from a *Poisson distribution*. The Poisson distribution was derived from a special case of the binomial distribution by Simeon Denis Poisson in 1837 [7]. Much of the early literature on analysis, especially regression analysis, of count data assumes that events are *independently and identically distributed*, or an event is independent of the events that came before it. This assumption is violated when dealing with time series count data that are correlated [26]. Over the past half century, studies dealing with count data and time series of count data have

evolved and expanded to try and construct models appropriate for all types of event count data.

One of the earliest instances of count data analysis was done by the Polish statistician and economist Ladislaus von Bortkiewicz in his famous book *Das Gesetz der kleinen Zahlen (The Law of Small Numbers)*. Bortkiewicz analyzed the number of soldiers in each corps of the Prussian cavalry who were killed by horse kicks between 1875 and 1894. He showed that the data followed a Poisson distribution. About twenty years later, Greenwood and Yule were credited as having first derived the *Negative Binomial distribution*, a standard generalization of the Poisson [7]. During the 1930s and 1940s bio-statistic studies extensively used regression analysis of count data and explored issues that came from restrictive count data models. Over the next thirty years, new models for regression analysis of count data were developed and tested in many different fields.

Throughout the 1960s and 1970s different count data models such as the negative binomial model and *generalized linear models* were constructed and improved through studies in actuarial science, econometrics, demography, and political science. Generalized linear models, of which the Poisson regression is a special case, was first described by Nelder and Wedderburn in the early 1970s and marked an important point in the evolution of count data models [27]. McCullagh and Nelder further extended generalized linear models in the late 1980s. These models opened the door for many other scholarly articles on regression analysis of count data. The late 1970s and throughout the 1990s, was a prolific period in the study of count data analysis. Many models were proposed to remedy different issues that arose during the study of discrete count data analysis. Scholars like A. Colin Cameron, Pravin K. Trivedi, and Gary King researched and helped develop different methods for count data regression

analysis while exploring the limitations of different models and proposing solutions to those issues. A resurgence of the study of count data, in particular time series of count data, occurred in the the early 2000s when scholars delved into the arduous task of analyzing time series count data with *serial correlation*, or time dependence between observations.

One of the first resolute attempts to define a time series count model with auto-correlation structure like *ARMA models*, or autoregressive moving average models, was by Jacobs and Lewis in their work *Discrete Time Series Generated by Mixtures I: Correlations and Runs Properties* published in the *Journal of the Royal Statistical Society* in 1978 [18]. Jacobs and Lewis defined a group of discrete ARMA models (DARMA) where they noticed that the value y_t is a mixture of the past values and the current realization of a latent variable, ϵ_t . These models are restrictive because when there is high serial correlation, the data can be described by a series of runs of a single value, which is not always the case in time series data which often has more variability [7]. Time series count data appear often in social science fields. In the end of the twentieth-century many social science researchers were faced with the challenge of accounting for serial correlation when studying time series count data.

In the late 1980s through the beginning of the 2000s, many scholars dedicated relentless work to identify and construct models appropriate for modeling event count time series data with serial dependence. In the past, Ordinary Least Squares (OLS) regression, Poisson regression, and other techniques were used. In leading scholarly articles and texts published starting in the late 1980s, Gary King, Patrick Brandt, A. Colin Cameron and Pravin K. Trivedi, among many others, examined different models for event count time series data in social sciences, and proposed methods for modeling this kind of data with serial dependence. In these articles and texts, scholars

argued that the methods of OLS, Poisson regression, negative binomial regression, and other count models are inefficient in modeling time series event counts with serial dependence and produced estimates that could be flawed.

Gary King’s articles, *Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator* and *Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model*, both examine different models for event count data. In *Statistical Models for Political Science Event Counts*, King recognized the problems with ignoring serial dependence when dealing with count data and used Monte Carlo experiments to show that if OLS is used when event count data are Poisson-distributed the estimates can be insufficient and can lead to results that are not meaningful [21]. King also studied the effects of ignoring serial correlation when using a Poisson regression, concluding that ignoring autocorrelation can lead to incorrect findings. In *Variance Specification in Event Count Models*, King examined Poisson and negative binomial regression models for event count data and then proposed the *Generalized Event Count* (GEC) model. This model allows for the modeling of event counts with unknown dispersion, and the Poisson regression is a special case of the GEC model [20].

Brandt, et. al argued that time series event count models can be included in three different classes, integer-values ARMA or discrete autoregressive moving average models (DARMA), conditional autoregressive models or hidden Markov processes, and state-space or time varying parameter models [4]. DARMA models are based on probabilistic mixtures of different Poisson processes and are difficult to use in analyses. Conditional autoregressive models assume that the conditional distribution of the events depends on a static mean function and lagged-dependent variables or

serially correlated errors [17]. Inference based on these models can be viewed as imprecise. State-space models specify a measurement and transition equation that are based on the dynamics of the time series being studied. Brandt, et al. developed a state-space model for time series count data, and presented an argument as to why their model was more efficient than other contemporary state-space models.

Dynamic Models for Persistent Event-Count Time Series, an article by Patrick Brandt et. al elaborated on the findings of King's *Statistical Models for Political Science Event Counts* and argued that a Poisson Exponentially Weighted Moving Average (PEWMA) model, based on the work of Harvey and Fernandes (1989) [16], is a useful approach to event count time series data with persistent dynamics [4]. *A Linear Poisson Autoregressive Model: The Poisson AR(p) Model*, by Patrick Brandt and John T. Williams introduced a new method for modeling event count time series data with serial dependence [5]. They then used this model in an analysis of the effect of a community crime fighting group on the number of purse snatchings in Hyde Park from 1969 to 1974. In this analysis Brandt and Williams used their PEWMA and PAR(p) models and compared their findings against a similar analysis done with a Poisson and a negative binomial regression. Brandt and Williams also constructed R code that implemented their PEWMA and PAR(p) models. This code was utilized in this thesis.

Inspired by this PAR(p) model for event count time series, this thesis examines the effect of the Salem witch hunt on witchcraft accusations in seventeenth-century America by analyzing the event count time series data using a PAR(p) model, Poisson regression, and a negative binomial regression in an intervention analysis. A comparison of these three models, along with the findings on the effect of the Salem witch hunt, as an intervention period, are presented in the pages to follow.

2 Methodology

2.1 Event Count Models

Event counts are dependent variables that take on non-negative integer values for each of the n observations. These values represent the number of times an event occurs in a fixed time period. In the case of this thesis, an event is a witchcraft accusation and the time period being considered is a month. An intervention analysis was performed to examine the effect of the Salem witch hunt on witchcraft accusations in Colonial America. This analysis was performed using a Poisson regression, negative binomial regression, and a Poisson autoregressive model. The theory behind these models and the intervention analysis is presented in the following sections.

Over the past thirty years, scholars have dedicated years of research to the development and analysis of the best ways to model event counts and time series of event counts. This thesis examines two models that have an extended history in modeling event count data and one model that is still relatively young. The Poisson and negative binomial regressions have been used extensively to model event count data. In terms of time series event count data, scholars argue that these models fall short of modeling the true dynamics of the data because they assume independence of events. However, these two models are used frequently to model count data and can be used to model the time series event count data of this thesis if autocorrelation is ignored and the events are assumed to be independent. These two models are used in this thesis since the data are considered to be event count data, and are used as comparison to a newer model, the Poisson autoregressive model, when dealing with an intervention analysis.

In 2001, Patrick Brandt and John T. Williams introduced a linear Poisson au-

toregressive model, the PAR(p) model, to model time series event count data [5]. They argue that the PAR(p) better models the true dynamics of time series event count data when the autocorrelation function of the time series in question shows a dampening dynamic process. After examining the autocorrelation function for the time series for this thesis, prominent dampening dynamics were observed. This was motivation for using a PAR(p) model in the intervention analysis.

These three models were used to conduct an intervention analysis to examine the effect of the Salem Witch hunt on witchcraft accusations in Colonial America, outside of the Salem episode, and also to compare how the three different models performed with these data.

2.2 Poisson Regression Model

2.2.1 The Poisson Distribution

If a discrete random variable Y is Poisson distributed with rate μ , $\mu > 0$, and t is the length of time during which events are recorded, then Y has density

$$Pr[Y = y] = \frac{e^{-\mu t}(\mu t)^y}{y!}, \quad y = 0, 1, 2, \dots \quad (1)$$

The mean and variance of the Poisson distribution are equal [20],

$$E[Y] = Var[Y] = \mu t. \quad (2)$$

2.2.2 Specification of the Poisson Regression Model

In a Poisson regression, events are assumed to be independent and are generated from a Poisson density, thus the mean and variance are equivalent [4]. To specify a Poisson

regression model, let the event count data consist of n independent observations, the i^{th} of which is (y_i, \mathbf{x}_i) . The dependent variable, y_i , is the number of occurrences of the event being recorded, and \mathbf{x}_i is the vector of linearly independent regressors that are thought to determine y_i . The regression model follows by conditioning the distribution of y_i on a k - dimensional vector of covariates, $\mathbf{x}'_i = [x_{1i}, \dots, x_{ki}]$, and parameters β , through a continuous function $\mu(\mathbf{x}_i, \beta)$, such that $E[y_i|\mathbf{x}_i] = \mu(\mathbf{x}_i, \beta)$.

That is, y_i given \mathbf{x}_i is Poisson-distributed with density

$$f(y_i|\mathbf{x}_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \quad y_i = 0, 1, 2, \dots$$

with mean parameter $\mu_i = \exp(\mathbf{x}'_i\beta)$ [7].

This model implies that the conditional mean and conditional variance are given by

$$E[y_i|\mathbf{x}_i] = Var[y_i|\mathbf{x}_i] = \exp(\mathbf{x}'_i\beta).$$

2.2.3 Estimation of the Poisson Regression Model

To estimate β , the effect of the explanatory variables on the dependent variable, the method of maximum likelihood, or MLE, is used. If we assume independence of events, we can write the log-likelihood as

$$\ln \mathcal{L}(\beta) = \sum_{i=1}^n \{y_i \mathbf{x}'_i \beta - \exp(\mathbf{x}'_i \beta) - \ln y_i!\}$$

If we differentiate the log-likelihood function with respect to β we get the Poisson MLE, $\hat{\beta}$, as the solution to the first order conditions

$$\sum_{t=1}^n (y_t - \exp(\mathbf{x}_t' \beta)) \mathbf{x}_t = \mathbf{0}.$$

We then use iterative methods to estimate $\hat{\beta}$ [20].

2.3 Negative Binomial Model

The Poisson model assumes *equidispersion*, or the equality of mean and variance. Equidispersion is frequently violated in real-life data. *Overdispersion*, when the variance exceeds the mean, or *underdispersion*, when the variance is less than the mean, can occur in real-life data. To model data without the restriction of assuming equidispersion the *negative binomial* model, where the variance is assumed to be a quadratic function, is used [7].

2.3.1 Negative Binomial Distribution

The negative binomial density for the count random variable y with parameters α and P is

$$f(y) = \binom{\alpha + y - 1}{\alpha - 1} \left(\frac{P}{1 + P} \right)^y \left(\frac{1}{1 + P} \right)^\alpha, \quad y = 0, 1, \dots, \quad P, \alpha > 0. \quad (3)$$

2.3.2 Specification of the Negative Binomial Model

The Negative Binomial model is more dispersed than the Poisson model and can account for overdispersion. The most common enforcement of the negative binomial is the *NB2* model with mean $\mu_i = \exp(\mathbf{x}_i' \beta)$ and variance $\omega_i = \mu_i + \alpha \mu_i^2$. This model has density,

$$f(y|\mu, \alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\alpha^{-1} + \mu} \right)^y, \quad \alpha \geq 0, y = 0, 1, 2, \dots \quad (4)$$

If we take the limit as $\alpha \rightarrow 0$ then the above density reduces to the Poisson density [7].

The Gamma function: $\Gamma(c)$, is defined by

$$\Gamma(c) = \int_0^{\infty} e^{-t} t^{c-1} dt, \quad c > 0.$$

Properties of the Gamma function:

1. $\Gamma(c) = (c - 1)\Gamma(c - 1)$
2. $\Gamma(c) = (c - 1)!$ if $c \in \mathbb{Z}^+$
3. $\Gamma(0) = \infty$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
4. $\Gamma(nc) = (2\pi)^{\frac{(1-n)}{2}} (n)^{nc - \frac{1}{2}} \prod_{k=0}^{n-1} \Gamma(c + \frac{k}{n})$, where $n \in \mathbb{Z}^+$.
5. $\frac{\Gamma(y+c)}{\Gamma(c)} = \prod_{j=0}^{y-1} (j + c)$ if $y \in \mathbb{Z}$.

Using the properties of the Gamma function and natural logarithms, we can see that

$$\ln \left(\frac{\Gamma(y + \alpha^{-1})}{\Gamma(\alpha^{-1})} \right) = \sum_{j=0}^{y-1} \ln(j + \alpha^{-1}). \quad (5)$$

2.3.3 Estimation of the Negative Binomial Model

If we take Equation (5) and substitute it in to Equation (4) we can find the log-likelihood function for the exponential mean $\mu_i = \exp(\mathbf{x}'_i\beta)$ is

$$\ln L(\alpha, \beta) = \sum_{i=1}^n \left\{ \left(\sum_{j=0}^{y_i-1} \ln(j + \alpha^{-1}) \right) - \ln y_i! - (y_i + \alpha^{-1}) \ln(1 + \alpha \exp(\mathbf{x}'_i\beta)) + y_i \ln \alpha + y_i \mathbf{x}'_i\beta \right\}.$$

The NB2 MLE $(\hat{\beta}_{NB2}, \hat{\alpha}_{NB2})$ is the solution to the first-order conditions

$$\sum_{i=1}^n \frac{y_i - \mu_i}{1 + \alpha \mu_i} \mathbf{x}_i = \mathbf{0}, \quad (6)$$

$$\sum_{i=1}^n \left\{ \frac{1}{\alpha^2} \left(\ln(1 + \alpha \mu_i) - \sum_{j=0}^{y_i-1} \frac{1}{(j + \alpha^{-1})} \right) + \frac{y_i - \mu_i}{\alpha(1 + \alpha \mu_i)} \right\} = 0. \quad (7)$$

These estimates are determined numerically using iterative techniques.

2.4 Poisson Autoregressive Model of Order p

Typical approaches to modeling event count data are Poisson or negative binomial regression models. If the event counts are time series, a lagged dependent variable could be included in an analysis to deal with correlation in events. Another approach would be to ignore that the data follow a count distribution and use *ARIMA*, or integrated autoregressive moving average model, techniques to study the data. Both of these options tend to fail at adequately modeling different dynamics in the data [26]. For this thesis, an alternative approach is taken to model the time series event counts. The *Poisson Autoregressive model*, $\text{PAR}(p)$, developed by Brandt and Williams, is an

autoregressive state-space model for count data and is used to conduct an intervention analysis for this thesis [5].

2.4.1 Par(p) Model Identification and Specification

According to Brandt and Williams, authors of the PAR(p) model, the event count time series model specification steps are as follows:

1. **Identification:** Using the autocorrelation function, determine if there are any dynamics present in the time series.
2. **Specification:** If the autocorrelation function shows a dampening dynamic process, estimate a PAR(p) model. If the autocorrelation function does not display this type of dynamic, further analysis is needed to specify a more fitting model.
 - If the dynamics are strong and persistent, estimate a PEWMA model.
 - If the autocorrelation function shows no evidence of any dynamics, estimate a Poisson, negative binomial, or generalized event count model [5].

PAR(p) Model Assumptions: The Poisson autoregressive model assumes that the counts at time t are generated from a Poisson distribution with mean μ_t .

A Linear AR(p) Process:

Brandt and Williams define a linear AR(p) process as

$$E[y_t|Y_{t-1}] = \sum_{i=1}^p \rho_i Y_{t-i} + \left(1 - \sum_{i=1}^p \rho_i\right) \mu. \quad (8)$$

This linear AR(p) model and the assumption that the events are Poisson distributed are used to define the transition equation for the PAR(p) model. From there we can define the PAR(p) state-space model.

The PAR(p) model can be defined as follows:

1. **Measurement Equation:** The measurement equation describes how the observed number of events arise as a function of the mean number of events in the past [4].

Let y_t for $t = 1, 2, \dots$, the observed counts, be drawn from a Poisson distribution conditional on m_t :

$$Pr(y_t|m_t) = \frac{m_t^{y_t} e^{-m_t}}{y_t!} \quad (9)$$

Equation (9) is the measurement equation of the PAR(p) model.

2. **State Variable and Transition Equation:** The transition equation describes how important past events are for predicting the number of events in the current period. It describes the dynamic transition process from events in the past to those in the present.

Assume that m_t , the *state variable* for the model, is the conditional mean of the linear AR process $E[y_t|Y_{t-1}]$ as in Equation (8). Since the measurement equation is Poisson distributed, this state density is in the exponential family and can be described by its mean, m_t , and variance, σ_t .

The state variable of the measurement density evolves according to a stationary AR(p) process with autocorrelation parameters $\rho_i, i = 1, 2, \dots, p$. That is,

$$m_t = \sum_{i=1}^p \rho_i y_{t-i} + \left(1 - \sum_{i=1}^p \rho_i\right) \exp(X_t \delta). \quad (10)$$

Here X_t is a $T \times k$ matrix relating to y_{t-i} for all $i = 1, 2, \dots, p$ and δ is a $k \times 1$ vector of regression parameters.

3. **Conjugate Prior:** A *conjugate distribution* is a prior distribution that after being combined with a likelihood function yields a posterior distribution of the same form. A *conjugate prior* for a given distribution is that distribution that will produce a distribution in the same class of distributions when used to form a Bayesian update of the original distribution. For example, if y is sampled from a Poisson distribution with unknown mean λ , then the conjugate prior that describes λ is the gamma distribution [9]. Assume that the density of the state variable has a gamma-distributed conjugate prior, so

$$Pr(m_t | Y_{t-1}) = \Gamma(\sigma_{t-1} m_{t-1}, \sigma_{t-1}), \quad m_{t-1} > 0, \quad \sigma_{t-1} > 0 \quad (11)$$

with $m_{t-1} = E[y_t | Y_{t-1}]$ and $\sigma_{t-1} = Var[y_t | Y_{t-1}]$. Using the observed data, the prior is constructed by finding the conditional mean and variance at time t based on the previous $t - 1$ observations. This prior distribution is gamma distributed with mean m_{t-1} and variance $\frac{m_{t-1}}{\sigma_{t-1}}$.

Brandt and Williams further assumed that m_{t-1} is distributed gamma with location $(\sigma_{t-1} m_{t-1})$ and scale σ_{t-1} . They use this assumption to help construct the PAR(p) filter used in model estimation [4].

2.4.2 Par(p) Model Estimation

The computations for the measurement and transition equations require filtering the data using PAR(p) filter equations. To estimate the PAR(p) model, Brandt and Williams used an extended *Kalman Filter* to construct their PAR(p) filter. This filter was then used to construct the predictive distribution of the event counts and the log-likelihood function.

The Kalman Filter: The primary aims of an analysis dealing with state-space models is to produce estimators for the underlying unobserved indication \mathbf{x}_t , given the data $Y_t = \{y_1, \dots, y_t\}$. The Kalman filter helps accomplish this aim. The Kalman filter is a set of equations that provides an efficient computational way to estimate the state of a process, in a way that minimizes the mean of the squared error [31].

The PAR(p) Filter: Brandt and Williams constructed the PAR(p) filter using an extended Kalman filter. As presented in their article *A Linear Poisson Autoregressive Model*, Brandt and Williams assumed the following when constructing the PAR(p) filter [5]:

1. **Measurement Density:** The observed counts are generated by a marginal Poisson distribution, with state variable m_t , as listed above in the model specification.
2. **Transition Equation:** The state variable of the measurement density evolves according to a stationary AR(p) process with autocorrelation parameters $\rho_i, i = 1, 2, \dots, p$, as listed above.
3. **Conjugate Prior:** Assume that m_{t-1} is gamma distributed with location

($\sigma_{t-1}m_{t-1}$) and scale σ_{t-1} . Brandt and Williams posited that using the mean and variance of the state variable is equivalent to using the gamma distribution's location and scale, here described as a_{t-1} and b_{t-1} , because one can assume

$$m_{t-1}|Y_{t-1} \sim \Gamma(a_{t-1}, b_{t-1}). \quad (12)$$

Brandt and Williams determined that the most efficient and optimal estimator of the parameters a and b is equivalent to a second moment estimator. They used second moment conditions to solve and find the following:

$$E[m_t] = \frac{1}{n} \sum_{t=1}^{t-1} m_t = \frac{a_{t-1}}{b_{t-1}} = m_{t|t-1}, \quad (13)$$

$$Var[m_t] = \frac{1}{n} \sum_{t=1}^{t-1} (m_t - E[m_t])^2 = \frac{a_{t-1}}{b_{t-1}^2} = \frac{m_{t|t-1}}{\sigma_{t|t-1}}. \quad (14)$$

Thus $a_{t-1} = \sigma_{t-1}m_{t-1}$ and $b_{t-1} = \sigma_{t-1}$. These estimates were then used to compute the PAR(p) filter.

Brandt and Williams constructed the PAR(p) filter with the following steps:

1. First, they combined the prior with the transition equation. They assumed $m_{t-1}|Y_{t-1} \sim \Gamma(a_{t-1}, b_{t-1})$. Using properties of the gamma distribution and of a stationary AR(p) process, they found

$$E[m_{t-1}|Y_{t-1}] = \sum \rho_i m_{t-i} + \left(1 - \sum \rho_i\right) \exp(X_t \delta), \quad (15)$$

which is gamma distributed, $\Gamma(m_{t|t-1}\sigma_{t|t-1}, \sigma_{t|t-1})$.

2. Using Bayes rule to find $m_t|Y_t$, Brandt and Williams updated the value of m_t , and the conditional values of $m_{t|t-1}$ and $\sigma_{t|t-1}$. They found,

$$m_t|Y_t \sim \Gamma(m_{t|t-1}\sigma_{t|t-1} + y_t, \sigma_{t|t-1} + 1) \quad (16)$$

or

$$m_t|Y_t \sim \Gamma(m_t\sigma_t, \sigma_t). \quad (17)$$

3. Proceeding by induction, Brandt and Williams discovered that since the filter preserves the conjugate form, the model is recursive and at time period t they had the recursions that generate the prior for the observation at time period $t + 1$.

After defining the filter, a prior must be specified to put the filter in motion. These values would be m_0 and σ_0 . If one chose to use the sample mean and variance of the series in question, the filter would be the Kalman filter [4].

Log-Likelihood for the PAR(p) Model: The log-likelihood for the PAR(p) model was constructed using the forecast density for the one-step ahead predictive distribution:

$$\mathcal{L}(m_{t-1}, \sigma_{t-1} | y_t, \dots, y_T; Y_{t-1}) = \ln \prod_{t=1}^T Pr(y_t | Y_{t-1}).$$

This can be estimated using maximum likelihood techniques.

Forecast Function for PAR(p) Model: The forecast functions for the conditional mean and variance of the PAR(p) model were derived using properties from

the negative binomial distribution. For optimized values of ρ , δ , m_t , and σ_t the forecast functions for the conditional mean and variance of the PAR(p) model are,

$$E[y_{t+1}|Y_t] = m_{t+1|t} = \sum_{i=1}^p \rho_i m_{t+1|t-1} + \left(1 - \sum_{i=1}^p \rho_i\right) \mu \quad (18)$$

$$Var[y_{t+1}|Y_t] = \frac{1 + \sigma_{t+1|t}}{\sigma_{t+1|t}} m_{t+1|t}. \quad (19)$$

2.4.3 Interpretation of the PAR(p) Model

The interpretation of the PAR(p) model is different from that of the Poisson and negative binomial. Consider a PAR(p) model with a covariate matrix X_t and $\mu = \exp(X_t \delta)$. The effect of a change in X_t is given by an impact multiplier in the PAR(p) model. This impact multiplier measures the effect of a change in X_t on the mean number of counts at time t and is determined by calculating the first derivative of the mean function for this change.

For the PAR(p), this first derivative is

$$\frac{\partial m_t}{\partial X_t} = \frac{\partial(\sum_{i=1}^p \rho_i Y_{t-i} + (1 - \sum_{i=1}^p \rho_i) \exp(X_t \delta))}{\partial X_t} = \left(1 - \sum_{i=1}^p \rho_i\right) \exp(X_t \delta) \delta. \quad (20)$$

This measures the instantaneous effect of the change in X_t on the mean m_t . This instant effect of the change in the independent variable in the PAR(p) model, also known as the impact multiplier for the number of counts at time t , depends on the estimated value of the regression parameters and on the estimated values of ρ_i . The impact multiplier for the PAR(p) model can then be used to calculate the long-run

multiplier for the total effect of a shock to X_t . This long-run multiplier measures the effect of the change accounting for the dynamic effects of the change on the conditional mean number of events. It can be calculated with

$$\frac{\frac{\partial m_t}{\partial X_t}}{(1 - \sum_{i=1}^p \rho_i)} = \frac{(1 - \sum_{i=1}^p \rho_i) \exp(X_t \delta) \delta}{(1 - \sum_{i=1}^p \rho_i)} = \exp(X_t \delta) \delta. \quad (21)$$

The instantaneous and long-run estimated effect in the Poisson or negative binomial models are the same. This is in contrast to the PAR(p) multipliers. Thus, if the true data generation process is a PAR(p), the methods used to calculate and interpret instantaneous changes in event count models like the Poisson or negative binomial would not be valid.

To view the difference in the static versus dynamic event count models, that is the Poisson and negative binomial models versus the PAR(p) model, the total percentage change in the number of events after the intervention can be calculated. For the Poisson and the negative binomial models the total percentage change is calculated by taking the exponentiation of the intervention parameter, subtracting one, and multiplying by 100.

For the PAR(p) model, if $X_t = [x_t, z_t]$ and $\beta = [\beta_1, \beta_2]$ are the regression parameters for x_t and z_t , respectively, the estimated instantaneous percentage change in the counts for a change Δz_t is found using

$$100 \frac{m_{\Delta z} - m_z}{m_z} = 100 \left[\frac{(1 - \sum_{i=1}^p \rho_i) \exp(x_t \beta_1) \exp(z_t \beta_2) (\exp(z_t \beta_2) - 1)}{\sum_{i=1}^p \rho_i y_{t-i} + (1 - \sum_{i=1}^p \rho_i) \exp(x_t \beta_1 + z_t \beta_2)} \right] \quad (22)$$

From there, the long-run percentage change is calculated by dividing Equation

(22) by $\left(1 - \sum_{i=1}^p \rho_i\right)$.

2.5 Intervention Analysis

An intervention analysis, using a Poisson regression, a negative binomial regression, and a PAR(1) model, was performed to examine the Salem witch hunt as an intervention on witchcraft accusations of Colonial America. The Salem witch hunt was an episode in Colonial American history where a large number of people were accused and executed due to suspected associations with witchcraft. Between February and November of 1692, over 150 people in twenty or more towns of Essex County, Massachusetts, were accused of witchcraft and over 19 were executed or died in association with the Salem witch hunt [33]. Over more than a century, historians studied the events of the Salem witch hunt as an episode in Colonial American history that was isolated and influential to the culture of Colonial America. Using an intervention analysis the impact of the Salem witch hunt on Colonial American witchcraft accusations, outside of Essex County, was examined in this thesis.

Cryer and Chan explained that “*intervention analysis*, introduced by Box and Tiao (1975), provides a framework for assessing the effect of an intervention on a time series under study [8].” An intervention is assumed to affect the process being studied by changing the mean function or trend of the series. According to Cryer and Chan, interventions can be both man made and natural. The following methodology on intervention analysis, presented below, was originally presented in and studied from Jonathan A. Cryer and Kung-Sik Chan’s book *Time Series Analysis With Applications in R* [8].

2.5.1 Time Series Subject to Intervention

Assume there is a single intervention and consider a general time series model for $\{Y_t\}$,

$$Y_t = m_t + N_t \quad (23)$$

where m_t is the change in the mean function and $\{N_t\}$ is the time series with no intervention. Cryer and Chan define $\{N_t\}$ as the “*natural or unperturbed*” process [8]. $\{N_t\}$ could be stationary or non stationary and may or may not have seasonality.

Suppose the time series experiences an intervention at time T . Then m_t is assumed to be zero before time T . Thus the time series $\{Y_t, t < T\}$ is called the *preintervention data* and describes the series for the natural or unperturbed process $\{N_t\}$.

Specification in Intervention Analysis: A *step function* is often used in specification of intervention analysis. In this case, the function is 0 during the preintervention period and 1 during the postintervention period.

$$S_t^{(T)} = \begin{cases} 1 & , \text{if } t \geq T \\ 0 & , \text{otherwise} \end{cases}$$

A *pulse function*, $P_t^{(T)}$,

$$P_t^{(T)} = S_t^{(T)} - S_{t-1}^{(T)}$$

which equals 1 at $t = T$ and 0 otherwise, is the dummy variable that marks the time the intervention takes place.

An instantaneous and perpetual shift in m_t can be modeled by

$$m_t = \omega S_t^{(T)}$$

where ω is the unknown perpetual change in the mean caused by the intervention. If the pre- and postintervention data are assumed to be independent and identically distributed (iid), testing whether $\omega = 0$ or not is like testing whether population means of independent random samples are significantly different. Serial correlation that occurs in time series data makes the testing of whether $\omega = 0$ or not more arduous.

If the effects of the intervention do not occur instantaneously, say they begin after d time units, where d is known then

$$m_t = \omega S_{t-d}^{(T)}.$$

In many cases, the intervention affects the mean function over time, having more of an effect in the long run. This case can be modeled using an autoregressive, AR(1), type of model with the error term replaced by a multiple of the lag one of $S_t^{(T)}$:

$$m_t = \begin{cases} \omega \frac{1-\delta^{t-T}}{1-\delta} & , \text{ for } t > T \\ 0 & , \text{ otherwise} \end{cases}$$

with the initial condition $m_0 = 0$.

Often $1 > \delta > 0$, which results in m_t approaching $\frac{\omega}{(1-\delta)}$ for large t . Thus $\frac{\omega}{(1-\delta)}$ is the final change for the mean function, m_t .

If we limit $\delta = 1$ then $m_t = \omega(T - t)$ for $t \geq T$ and 0 otherwise. This specification produces an intervention effect that changes the mean function linearly in the

postintervention period.

For brief intervention effects, a pulse dummy variable, using the pulse function defined above, can aid in specifying the intervention consequences on the mean function. If the intervention alters the mean function only at time $T = t$, then

$$m_t = \omega P_t^{(T)}$$

where

$$P_t^{(T)} = \begin{cases} 1 & , \text{if } t = T \\ 0 & , \text{otherwise} \end{cases}$$

If intervention effects fade out steadily, those effects could be modeled using an AR(1) type of model, such as

$$m_t = \delta m_{t-1} + \omega P_t^{(T)}.$$

In this case, if $t \geq T$, $m_t = \omega \delta^{t-T}$. This means that the mean changes instantaneously by ω and the change in mean decreases geometrically by a factor of δ [8]. Delays, different intervention periods, and other specifications can be combined and analyzed to model more sophisticated intervention effects. To see an example of this, suppose the change in the mean takes place after a delay of one time unit and the effect fades out steadily. If we assume $m_0 = 0$, we could specify

$$m_t = \delta m_{t-1} + \omega P_{t-1}^{(T)}.$$

We could then use a *backshift operator* B , where $Bm_t = m_{t-1}$ and $BP_t^{(T)} = P_{t-1}^{(T)}$, to rewrite the above specification as,

$$m_t = \frac{\omega B}{1 - \delta B} P_t^{(T)}.$$

When modeling the intervention analysis for this thesis, a Poisson regression, negative binomial regression, and a PAR(1) model were used. In this case, an intervention variable called **Salem** was established. The preintervention period was from September of 1622 to April of 1692. The intervention occurred in May of 1692 and the postintervention period continued until December of 1712. Estimation of the parameters for this intervention analysis used maximum likelihood techniques for the Poisson and negative binomial models, and used a PAR(p) filter and maximum likelihood for the PAR(1) model. All of these models were fitted using **R** and the analysis is fully elaborated on in Section 3.

2.6 Diagnostic Tests

Diagnostic tests were used to assess the fit of the Poisson and negative binomial models. These tests included testing for normality of residuals using a Shapiro-Wilk test and testing independence of residuals using a Durbin Watson test. The normality assumption was visually assessed using quantile-quantile plots while the independence assumption was visually assessed using the autocorrelation function plot.

2.6.1 Normality

The normality assumption states that the population of error terms is normally distributed. To visually validate this assumption the order statistics of the sample are plotted against quantiles from a $N(0, 1)$ distribution, forming a quantile-quantile plot. If the normality assumption holds, the points would be randomly scattered around

the straight-line.

The Shapiro-Wilk normality test is used to statistically test this assumption. The test is based on the statistic:

$$W = \frac{\left(\sum_{i=1}^n c_i \epsilon_{(i)}\right)^2}{\sum_{j=1}^n \epsilon_j^2} \quad (24)$$

Here $\epsilon_{(i)}$ = ordered residuals and c_i = constants generated from means, variances, and covariances of the ordered statistics, sample size n .

We use the significance of the test statistic W to test the hypotheses:

H_0 : residuals are normally distributed

H_A : residuals are not normally distributed

2.6.2 Independence

Independence of the residuals is the assumption that every error term is statistically independent of every other error term. As with the normality assumption, we can visually assess the assumption of independence, using the autocorrelation function. Departures from independence are viewed as correlation amongst residuals, which are characterized by large spikes in the autocorrelation plot. The Durbin Watson test is used to statistically test the assumption of independence.

$$d = \frac{\sum_{i=1}^n (\epsilon_i - \epsilon_{i-1})^2}{\sum_{j=1}^n \epsilon_j^2}$$

Where $\epsilon_i = y_i - \hat{y}_i$ and y_i, \hat{y}_i are the observed and predicted values of the response variable for an individual i .

The significance of d is used to test the following hypotheses:

H_0 : residuals are not correlated

H_A : residuals are correlated

3 Analysis

3.1 Witchcraft Accusations September 1622 to December 1712

The data for this thesis were gathered from multiple historical sources including primary documents, such as court documents and secondary sources like the work of esteemed historians John Demos and Carol Karlsen. A complete list of the sources used for data collection are listed in Appendix B. These data included the name of the accused, their marital status, gender, and their colony of residence(colony). These data also included the month and year the victims were accused of witchcraft and whether or not they were executed. If a person was executed, the month and year of their execution was recorded. Data on imprisonment and acquittals were gathered for some of the people accused of witchcraft. A person was included in this data set if there exists evidence of some form of formal accusations or where there is some record of a suit of defamation of character due to being accused of witchcraft. If no record was available for a given variable, that entry was recorded as not applicable.

The data set included 272 people accused of witchcraft in Colonial America. The variables considered in this analysis were gender, marital status, month accused, season accused, execution status, and colony. All of these variables were categorical. The month and year of accusations were recorded as time series event count data, included 1084 observations, and was used in the intervention analysis, later in this section.

3.2 Descriptive Statistics

3.2.1 Biographical Portrait

With scattered, fragmented records, it is difficult to patch together a complete image of the victims accused of witchcraft in Colonial America. Many historians tirelessly researched and examined witchcraft in Colonial America to define the reach of witchcraft accusations, only to find it impossible to determine the exact number of those accused. In this thesis we consider a list of 272 people accused of witchcraft in seventeenth-century Colonial America, including victims of the Salem witch hunt. Due to lack of records, this list is not comprehensive and most likely under-represents those actually affected by witchcraft accusations in Colonial America. Often court records are sparse or hard to interpret whether witchcraft was part of a complaint. Unofficial materials, like letters or diary entries, are also limited by sparsity and lack of sufficient evidence [11]. Even though this data set does not include every person affected by witchcraft accusations in Colonial America, we take it as a sample that still merits investigation to help us understand accusations throughout seventeenth-century America.

Gender: In premodern times, the term “witch” applied to both men and women, thus witchcraft, or familiarity with the devil, did not, by definition, have specific distinctions between genders [19]. Many historians such as Karlsen, Demos, and Godbeer determined that the prevalence of women at the forefront of witchcraft accusations is historical commonplace. This is confirmed by the findings of this study.

In this data set, females outnumbered males by a ratio of 3 : 1. In his text *Entertaining Satan*, John Demos determined the ratio of females to males accused of

witchcraft is roughly 4 : 1 [11]. Demos’ ratio is supported by the findings of Carol Karlsen in her book *The Devil in the Shape of a Woman* [19]. While the ratio of females to males in this data set is not as pronounced as those of Demos and Karlsen, it still supports the historical idea that females were more likely the victims of witchcraft accusations than their male counterparts. Table 1 displays the number of males and females accused of witchcraft overall in seventeenth-century America, during the Salem witch hunt of 1692-93, and during the Connecticut witchcraft scare of 1662-63. There is slight deviation in the percentage of females and males accused of witchcraft during times of heightened hysteria. Demos explained this slight increase in male accusations as a result of the spreading of witchcraft accusations not only from one main suspect to another main suspect, but also to members of the suspect’s immediate family, often times including husbands and less often including male children [11]. Instances of husband and wife pairs being accused of witchcraft around the same time period, are evident in this data set. These include, but are not limited to, the cases of Nathaniel and Rebecca Greensmith who were accused and executed in Hartford, Connecticut in the 1660s and Hugh and Mary Parsons who were accused in Springfield, Connecticut in the 1650s [15]. Even in times of “intense witch fear” women were more likely suspects than men [19].

| | Males | Females |
|-----------------------|----------|-----------|
| Overall | 68 (25%) | 204 (75%) |
| Salem 1692-93 | 41 (26%) | 117 (74%) |
| Connecticut 1662-1663 | 5 (36%) | 9 (64%) |

During seventeenth-century America, witchcraft accusations occurred in small scale, local instances. The episode of the Salem witch hunt that resulted in large scale

accusations and executions went against the usual, small scale effect of witchcraft in Colonial America. The witch hunt in Connecticut 1662-63 was also uncharacteristic of “normal” witchcraft accusation rates, however it was not on the same scale as the Salem witch hunt. Table 2 displays witchcraft accusations by decades from before 1640 to after 1699, with the Salem witch hunt listed separately. One can see that in relation to other decades, 1660-1669 has high frequencies of accusations and the single year of the Salem witch hunt has significantly higher frequencies.

Table 2: Gender of Accused Witches, By Decades.

| | Males | Females | Total |
|------------------------|-------|---------|-------|
| Before 1640 | 0 | 2 | 2 |
| 1640-49 | 1 | 6 | 7 |
| 1650-59 | 7 | 23 | 30 |
| 1660-69 | 9 | 21 | 30 |
| 1670-79 | 4 | 9 | 13 |
| 1680-89 | 2 | 11 | 13 |
| 1690 – 99 ^a | 3 | 11 | 14 |
| After 1699 | 1 | 4 | 5 |
| Total | 27 | 87 | 114 |
| Salem 1692-93 | 41 | 117 | 158 |

^a. Not including those accused in Salem hysteria.

It is hard to ignore the fact that women were three to four times more likely to be accused of witchcraft than men in seventeenth-century America. This stark gender discrepancy led many historians to study how gender roles in Colonial American society influenced this large disproportion of accusations toward the female gender. Historians examined the idea of sexist oppression as an influence on witchcraft accusations. The stigma and danger of being accused of practicing witchcraft may have suppressed the behavior of women in Colonial America. In multiple documented cases, women were accused of being a witch when they had an “abrasive character,”

or spoke out stubbornly against societal standards [22]. At this point in history, men controlled political life, were religious leaders, and were seen as the head of families, thus there were distinct gender roles [12]. Lyle Koehler's *A Search for Power: The "Weaker Sex" in Seventeenth-Century New England* investigated the role of women in seventeenth-century New England, not only in relation to witchcraft accusations, but as an experience as a whole.

While distinct gender roles existed, John Demos argued that witchcraft accusations were not necessarily a means of sexist oppression. Demos pointed out that often women fell under accusation of witchcraft after being accused by another woman [11]. Carol Karlsen argued, in *The Devil in the Shape of a Woman*, that while witchcraft accusations were not used as sexist oppression, there was a difference in treatment of male and female suspected witches by citing the dissimilar treatment of Eunice Cole and John Godfrey of Essex County. She argued that women accused of witchcraft could expect harsher treatment than men accused of witchcraft [19]. This data set does not afford the information to draw conclusions on witchcraft accusations as sexist oppression, but it does allow us to look at different aspects of witchcraft dealing with gender. Table 3, illustrates the number of males and females accused, the number of those executed after being tried for witchcraft, and the percentage of each gender that was executed out of those accused. The percent of men executed out of those accused is slightly higher than the percent of women executed out of those accused even though almost three times more women were executed than men.

Table 3: Gender of Executed Witches

| Action | Males | Females | Total |
|----------|------------|------------|-----------------|
| Accused | 68 | 204 | 272 |
| Executed | 10 (14.7%) | 27 (13.2%) | 37 ^a |

^a. Including those in Salem 1692-93

Examining the geographic aspect of witchcraft accusations, as in Table 4, we find that a large portion of the witchcraft accusations in this data set were concentrated in the Massachusetts Bay colony and in the Connecticut colony. The category of *Other* includes New York, Pennsylvania, and other colonies. These findings are concurrent with the fact that New England colonies played a major role in the witchcraft demographics of early America. While Virginia was the site of earliest surviving records of witchcraft in America, dating back to 1626, throughout the rest of the seventeenth-century it experienced only about 10 cases of accusations [10]. New York experienced occasional witchcraft accusations throughout the seventeenth-century, including the trial of Ralph and Mary Hall in 1665 [15]. Pennsylvania, founded later than other colonies considered here, had a more truncated experience with witchcraft. Its Quaker majority did not seem to be as involved in witchcraft accusations as their Puritan contemporaries in the New England Colonies [13].

Table 4: Gender of Accused Witches, By Colonies.

| | Male | Female | Total |
|------------------------|------|--------|-------|
| Mass. Bay ^a | 51 | 147 | 198 |
| Conn. | 11 | 30 | 41 |
| New Hamp. | 1 | 6 | 7 |
| Virginia | 2 | 7 | 9 |
| Other | 3 | 14 | 17 |
| Total | 68 | 204 | 272 |

^a. Including the Salem Witch hunt 1692

Analyses of the Puritan religion and its relation to witchcraft are complex and extensive. Historians studied sermons, diaries, court records and other sources to build a wealth of literature about Puritanism and witchcraft in Colonial America. Their findings determined a direct connection between Puritan beliefs and witchcraft accusations. Richard Godbeer's *The Devil's Dominion: Magic and Religion in Early New England* is a meticulously detailed analysis of the relationship between religion and witchcraft in Colonial New England. Godbeer examined the role of witches in Puritan society and how they were viewed as "malevolent creatures who used occult means to harm their enemies, and are minions of the devil [13]." This relationship may have led to significantly higher frequencies of witchcraft accusations in the New England colonies.

Marital Status: Many historians studied the lives of those accused of witchcraft in order to understand what made someone more susceptible to witchcraft accusations, if there were actually any specific defining factors. Many historians agree on specific aspects of a "typical witch." A witch in seventeenth-century America was usually a middle-aged, married female of English and Puritan descent that had troubled family relationships, had fewer children than the societal norm, was of low social and economic position, and often practiced some kind of healing occupation [10]. An examination of these data supported some aspects of these "typical witch" descriptions. Tables 1 through 4 illustrate gender and geography of accused witches and Table 5 illustrates the marital status of accused witches. The findings of this study conclude that of those in this data set, most accused were married females from New England.

By no means is this a prosopography of those accused of witchcraft in seventeenth-century America, like those of Demos and Karlsen. However, it is an effort to draw

Table 5: Marital Status of Accused Witches

| | | |
|-----------|-----|--------|
| Married | 144 | (53 %) |
| Single | 71 | (26 %) |
| Widowed | 24 | (9 %) |
| Not Known | 33 | (12 %) |

Record taken at time of trial or suspicion.

similar conclusions to those of the well researched works like *Entertaining Satan*, *Enemy Within*, and *The Devil in the Shape of a Woman* in order to better understand the lives of the people who make up this data set.

3.2.2 A Description of Accusations

Before analyzing witchcraft accusations as time series count data it is important to see when these accusations occurred.

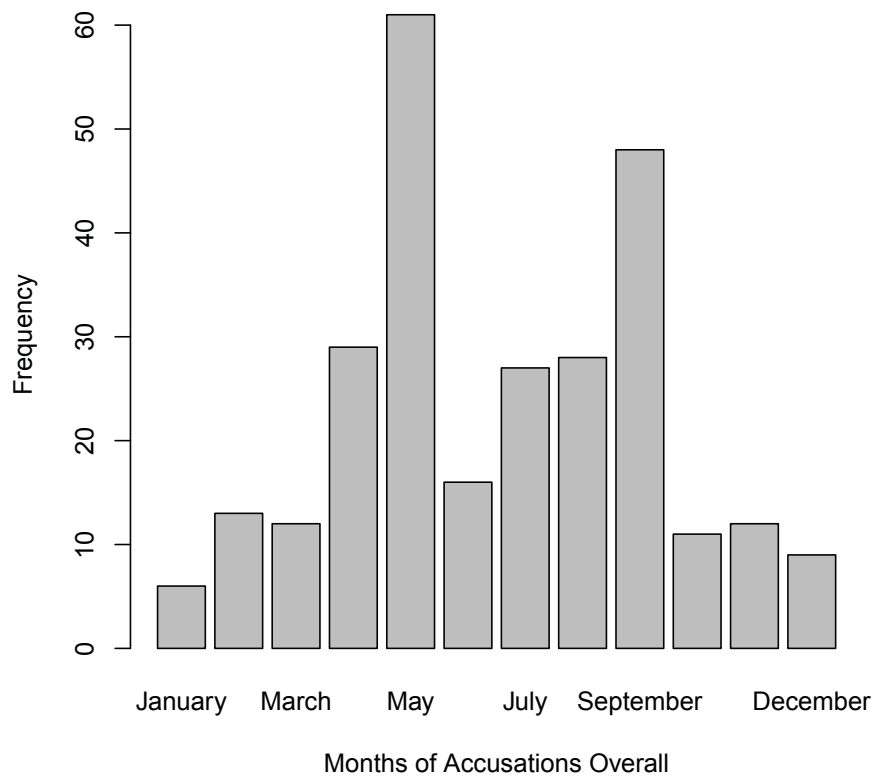


Figure 1: Months of Accusations Overall

We are interested to see if there are any patterns in accusations. From Figure 1 we can see that there are spikes in accusations in the months of May and September. Table 6 lists the frequencies of accusations for each month.

Table 6: Months When Accusations Occurred, Overall

| | | |
|-----------|----|--------|
| January | 6 | (2 %) |
| February | 13 | (5 %) |
| March | 12 | (4 %) |
| April | 29 | (11 %) |
| May | 61 | (23 %) |
| June | 16 | (6 %) |
| July | 27 | (10 %) |
| August | 28 | (10 %) |
| September | 48 | (18 %) |
| October | 11 | (4 %) |
| November | 12 | (4 %) |
| December | 9 | (3 %) |

Along with examining the months of accusations, we also examined the seasons in which accusations occurred. Table 7 displays the frequencies of accusations in different seasons. Spring saw the most accusations, with summer and fall with the next highest frequencies.

Table 7: Seasons When Accusations Occurred, Overall

| | | |
|--------|-----|--------|
| Winter | 28 | (10 %) |
| Spring | 102 | (38 %) |
| Summer | 71 | (26 %) |
| Fall | 71 | (26 %) |

The Salem witch hunt resulted in the most accusations and executions in an isolated witchcraft hysteria episode. In order to see how witchcraft accusations occurred in seventeenth-century America without the Salem witch hunt influence, we removed the Salem hysteria from the data. The data that corresponded to the hysteria included the accusations and executions that occurred in Salem Village plus the 20+ communities of Essex County that fell under the hysteria during 1692. These included 158 accusations and 22 executions. Figure 2 graphically represents the number of accusa-

tions that occurred in a given month and Table 8 reports the number of accusations for a given month. Comparing Figure 1 to Figure 2 we can see the spikes in accusations in May and September are less defined in Figure 2.

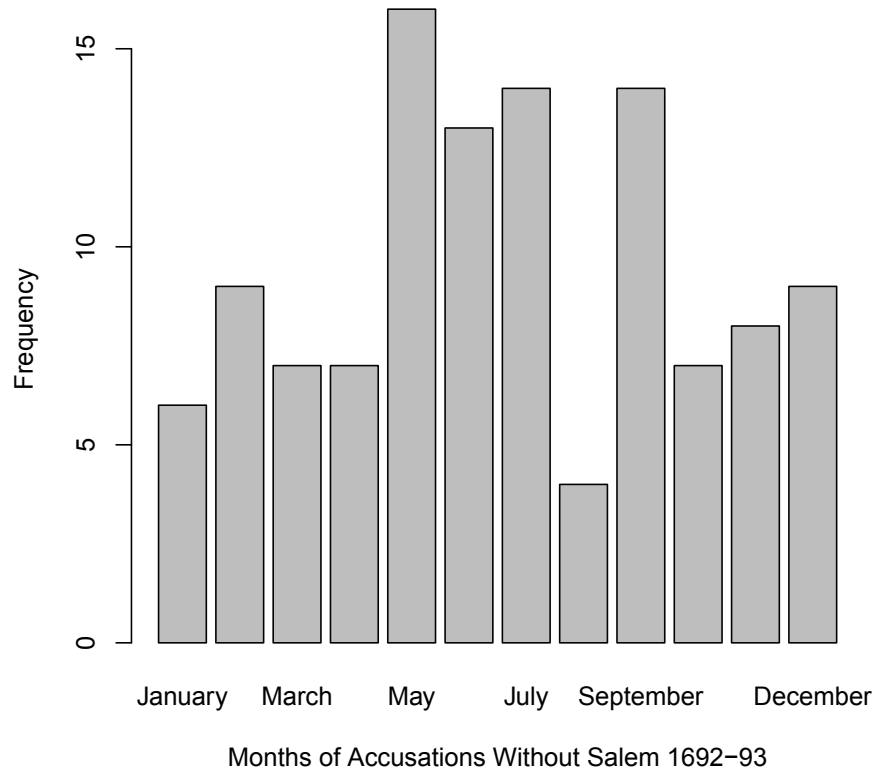


Figure 2: Months of Accusations Without the Salem Witch Hunt 1692-93

Table 8: Months When Accusations Occurred, Without Salem Witch Hunt 1692-93

| | | |
|-----------|----|--------|
| January | 6 | (5%) |
| February | 9 | (8 %) |
| March | 7 | (6 %) |
| April | 7 | (6 %) |
| May | 16 | (15 %) |
| June | 13 | (11 %) |
| July | 14 | (12 %) |
| August | 4 | (4 %) |
| September | 14 | (12 %) |
| October | 7 | (6 %) |
| November | 8 | (7 %) |
| December | 9 | (8%) |

Comparing Table 7 to Table 9 we can see that there is no longer a significant deviation in accusations in different seasons when the Salem witch hunt is removed from the data.

Table 9: Seasons When Accusations Occurred Without the Salem Witch Hunt 1692-93

| | | |
|--------|----|--------|
| Winter | 24 | (22 %) |
| Spring | 30 | (26 %) |
| Summer | 31 | (27 %) |
| Fall | 29 | (25 %) |

Many brilliant historians have studied the Salem witch hunt relentlessly and determined that its events were unlike other witchcraft accusations in Colonial America. The large number of accusations and executions from the Salem witch hunt left an impression on the rest of Colonial America, and still has people today questioning its origins and effects. Did the Salem witch hunt leave a lasting impression on the rest of witchcraft accusations in Colonial America? We explore this question through an intervention analysis using a Poisson regression, negative binomial regression, and a PAR(1) model.

3.3 Time Series of Event Counts

The number of witchcraft accusations per month from September of 1622 to December of 1712 made up time series event count data. These data were used in an intervention analysis to determine whether the Salem witch hunt had an effect on the number of witchcraft accusations in Colonial America.

3.3.1 The Data

The number of witchcraft accusations per month from September of 1622 to December of 1712 was recorded. This included witchcraft accusations over a 90 year period in nine different colonies. Counts ranged from 0 at the lowest to 48 at the highest. However, counts did not vary greatly each month. The data were characterized by a large number of 0 counts, a seven month period of counts higher than 15 but less than 48, and small counts ranging from 1 to 3. The seven month period of large counts corresponds to the Salem witch hunt.

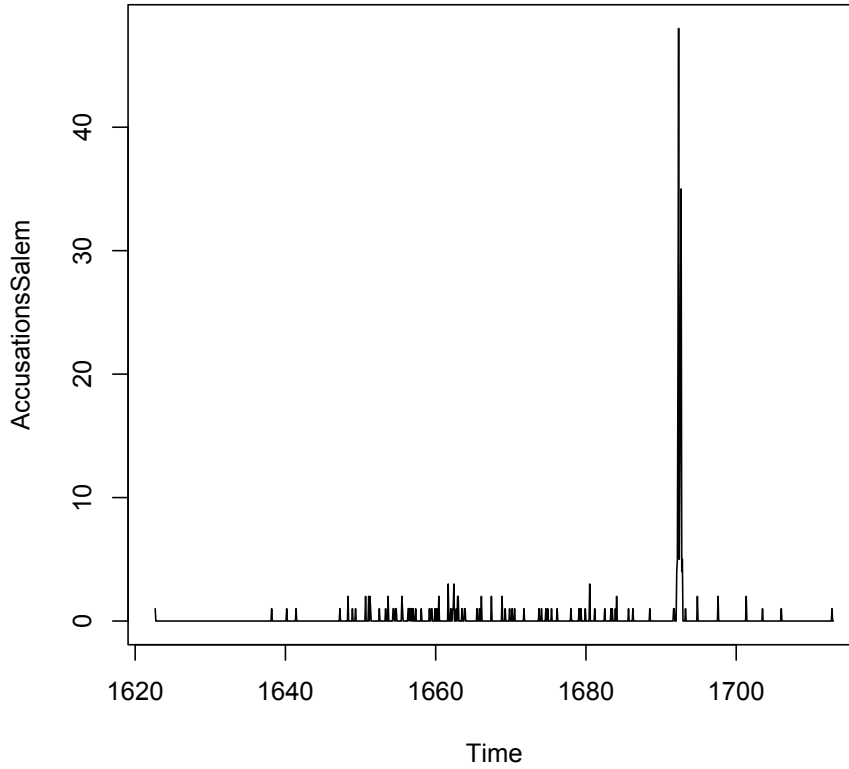


Figure 3: Witchcraft Accusations from September 1622 to December 1712

The plot of the data, Figure 3, displays low counts per month for the majority of the 90 year period with a large spike around 1692. This large spike corresponds to the Salem witch hunt. The seven month period, characterized by counts ranging between 15 and 48, was unlike any other time in the 90 year period. To examine witchcraft accusations in Colonial America outside of the Salem episode, the Salem counts were removed from the data set and the data were then plotted again.

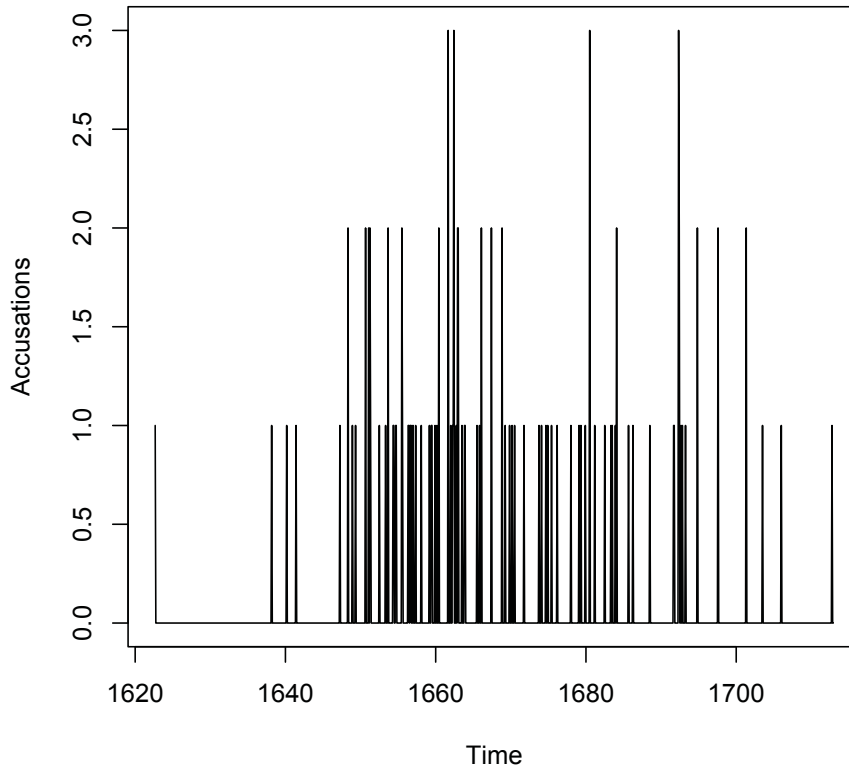


Figure 4: Witchcraft Accusations from September 1622 to December 1712 Without the Salem Witch Hunt

Figure 4 is the time series plot of the witchcraft accusation counts in Colonial America from September 1622 to December 1712, without the Salem witch hunt accusation counts. From Figure 4, one can see that all accusation counts range from zero to three per month over the 90 year period in question. The first step in trying to identify a model for this data was to examine whether the counts were correlated. This was explored using the autocorrelation function.

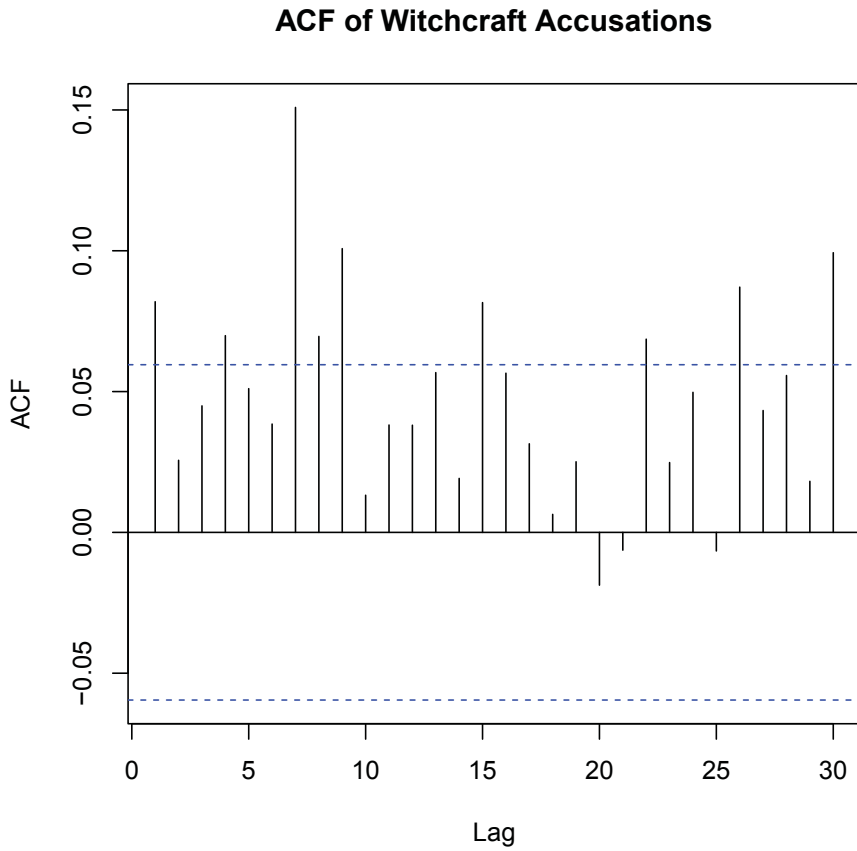


Figure 5: Autocorrelation Function for Witchcraft Accusations 1622 to 1712

The autocorrelation function, displayed in Figure 5, shows that the event counts are correlated. Thus the data has serial dependence, or the value at some time t in the series is statistically dependent on the value at another time s . Further investigation of the autocorrelation function showed that the data did not follow a seasonal pattern.

Prior to taking interest in exploring the effect of the Salem witch hunt on accusation counts in Colonial America, an in-depth exploration was performed to try and identify an appropriate model for these data. Since the data are event counts, event count models such as Poisson, negative binomial, zero-inflated Poisson, and

zero-inflated negative binomial were considered for these data. These models however do not take into consideration the serial dependence of the data. Efforts to find an appropriate model for the data then shifted to time series of event count data models.

The witchcraft accusation counts make up a discrete time series which is also count data. Taking into account the count properties of the data and the inherent time series attributes, finding an appropriate model for this time series proved to be an arduous task. After fitting multiple different models, studying models that are in their infancy and yet to be published, and trying to find models to study how the Salem witch hunt effected other witchcraft accusations in Colonial America, three different models were chosen for an intervention analysis. This intervention analysis was performed to discover whether or not the Salem witch hunt had an effect on witchcraft accusations in Colonial America.

3.3.2 Model Specification

Since the witchcraft accusation data are counts, Poisson and negative binomial regression models were used to model the data. The Poisson nor the negative binomial model account for serial correlation, so a time series model for count data was also specified.

Time series models for discrete counts do not exist in large numbers. After extensive research, two time series models for discrete counts were good candidates to model the witchcraft accusation data. These models were the Poisson Exponentially Weighted Moving Average model, PEWMA, and the Poisson Autoregressive model, $PAR(p)$. The PEWMA was originally constructed by Harvey and Fernandes in 1989 and then was improved by Brandt et. al in 2009 [16]. It models nonstationary time

series of counts with strong and persistent dynamics. The $\text{PAR}(p)$ model was constructed by Brandt and Williams and models stationary time series of counts with dampening dynamics [5]. After examining the autocorrelation function for the data, Figure 5, the presence of dampening dynamics led to the specification of a $\text{PAR}(p)$ model, specifically a $\text{PAR}(1)$ model.

A Poisson autoregressive model of order 1 was used to model the data to see if a model with autoregressive properties would better model the true dynamics of the series. This model along with a Poisson and a negative binomial were specified to model the witchcraft accusation counts in an intervention analysis.

3.4 Intervention Analysis

The main objective of this study is to learn more about witchcraft accusations in Colonial America, specifically witchcraft accusations that occurred outside of the infamous Salem witch hunt. The Salem witch hunt led to the accusation of over 150 people, the execution and deaths of more than nineteen people, the prosecution of many, and over a century of in-depth research to understand its social implications. After this infamous witch hunt occurred, did the rate at which people were accused of witchcraft change? An intervention analysis was performed on the witchcraft accusation data to try and answer this question.

3.4.1 Determining the Intervention Period

In order to examine the effect of the Salem witch hunt on witchcraft accusations in Colonial America, outside of the Salem episode, the witchcraft accusation counts without the Salem counts were used for the intervention analysis. The permanent

intervention period for this analysis began during the 839th period, which corresponds to May of 1692 and the establishment of the Court of Oyer and Terminer.

The establishment of the Court of Oyer and Terminer occurred in May of 1692 and was enacted by Sir William Phips [35]. Sir William Phips was the newly appointed first royal governor of Massachusetts in 1692 and as accusations of witchcraft escalated in Salem in May of 1692, he assigned the Court of Oyer and Terminer to decide the cases of those accused and jailed for witchcraft. The Court of Oyer and Terminer consisted of a few judges and a jury for each case. If someone was found guilty, they were sentenced to death for committing a capital offense [23]. This time period was chosen for the beginning of the permanent intervention because it is a valid historical marker. Also, the Court of Oyer and Terminer was influential in most of the executions during the Salem witch hunt. This time period was chosen over using the entire period of the Salem witch hunt as a non-permanent intervention because the effects of the Salem witch hunt would have resonated in the minds of the citizens of Colonial America longer than the time period in which the episode occurred. This same reasoning was the motivation to not end the intervention at the dissolution of the Court of Oyer and Terminer in October of 1692.

The permanent intervention for this analysis started at the establishment of the Court of Oyer and Terminer and continued until December of 1712. The intervention variable equals zero for the first 388 periods and one thereafter. Each model used in the analysis used the same intervention variable.

3.4.2 Estimating the Parameters of the Intervention Analysis

A Poisson regression, negative binomial regression, and a PAR(1) model was used to estimate the parameters of the intervention analysis. These computations were done

using R. The code for this analysis is listed in Appendix A.

The maximum likelihood estimate of the parameters was computed for each of the three models. These parameter estimates are presented in Table 10. All three models produced significant, negative intervention coefficients. Recall that the effects of a change in the regressors are estimated differently in the PAR(p) model than in a Poisson or negative binomial model, as discussed in Section 2. Thus to be able to compare the models, looking only at the coefficients is incorrect. The Long-run multiplier is of interest in the PAR(p) model. To compare the effects of the Salem witch hunt intervention across the estimators, the long-run multipliers need to be calculated. The long-run multiplier for the PAR(1) model is computed using Equation (21) and for the Poisson and negative binomial model a similar computation using $\exp(X_t\delta)\delta$ is used [5].

Table 10: Witchcraft Accusations Analysis.

| Model | PAR(1) | Poisson | Negative Binomial |
|-----------|----------------|----------------|-------------------|
| Salem | -1.762 (0.447) | -0.909 (0.305) | -0.910 (0.339) |
| Constant | 1.78(0.598) | -2.107 (0.099) | -2.110 (0.121) |
| ρ_1 | 0.9277 (0.048) | | |
| N | 1084 | 1084 | 1084 |
| Final LLF | -408.37 | -365.72 | -364.13 |
| AIC | 820.74 | 772.17 | 734.27 |

(Standard Error)

The total effect of the Salem witch hunt intervention is different in each of the models. Table 11 shows the estimated long-run effects of the Salem witch hunt intervention for the PAR(1), Poisson, and negative binomial models. The estimated long-run effect for the PAR(1) model is a total decline of 1.828 witchcraft accusations. For the Poisson and negative binomial models the long-run effects, which are the same as their instantaneous effects, are a decline of 0.971 and 0.972 witchcraft

accusations, respectively. However the interpretation of these effects is very different. For the PAR(1) model, this effect occurs over several periods. The impact multiplier, or the instantaneous effect, of the intervention for the PAR(1) model is -0.132. This instantaneous effect is less impactful than the instantaneous effects of the Poisson and negative binomial models. While the instantaneous effect of the PAR(1) model is less than that of the Poisson and negative binomial, the long-run effect, -1.828, illustrates a larger decrease in witchcraft accusations in the long-run. The Poisson and negative binomial models' interpretation is that the effect of the intervention is immediate. The omission of the dynamics from the Poisson and negative binomial models leads to an understatement of the magnitude of the intervention, since they fail to account for dynamics of a future drop in witchcraft accusations.

Table 11: Long Run Effects of Salem Witch Hunt.

| Model | Instantaneous Effect | Long-Run Effect |
|-------------------|----------------------|-----------------|
| PAR(1) | -0.132 | -1.828 |
| Poisson | -0.971 | -0.971 |
| Negative Binomial | -0.972 | -0.972 |

The difference in the static versus dynamic event counts models can be seen by computing the total percentage change in the number of witchcraft accusations after the Salem witch hunt intervention. For the Poisson and negative binomial models, the total percentage change is calculated by taking the exponentiation of the intervention parameter, subtracting one, and multiplying by 100. The total instantaneous percentage change and total long-run percentage change for the PAR(1) model were calculated using Equation (22), listed in Section 2. These percentage changes are shown in Table 12. For the Poisson and negative binomial models, the total change due to the intervention is a decline of 59.71 % and 58.01 % from the mean of the

series, respectively. The total instantaneous percentage change for the PAR(1) model is a decline of 5.99 %. This percentage change is starkly less than those of the Poisson and negative binomial models, just as in the instantaneous effects. However, the total long-run percentage change for the PAR(1) model is 82.82 %, a decline that is 23 and 24 points more than what we would have estimated from the Poisson and negative binomial models, respectively.

Table 12: Percent Change in Mean After the Salem Intervention.

| Model | Percentage Change in Mean | Long-Run Percent Change in Mean |
|-------------------|---------------------------|---------------------------------|
| PAR(1) | -5.989 | -82.824 |
| Poisson | -59.707 | -59.707 |
| Negative Binomial | -58.013 | -58.013 |

All three models produced negative intervention coefficients, and negative effects. This leads to the interpretation that the Salem witch hunt intervention led to a decrease in witchcraft accusations for the rest of the period in question. While all three models lead us to determine that the Salem witch hunt intervention led to a decrease in witchcraft accusations, the Poisson and negative binomial models illustrate a smaller impact of the intervention than that of the PAR(1) model.

3.4.3 Model Diagnostics

Finding an appropriate model to model the witchcraft accusation data proved to be a strenuous undertaking. In the end, a Poisson regression, negative binomial regression, and a Poisson autoregressive model of order one were used to model the data. Model diagnostics were performed to see how each of these models performed with these data.

Poisson Regression Residual Analysis: After examining the plot of the standardized residuals of the Poisson model, it was difficult to visualize the behavior of the residuals because they were compressed together. An examination of other residual plots, like the quantile-quantile plot and histogram also proved to be hard to interpret. To visually assess the residuals for this model, and for the negative binomial model, white noise was added to the residuals using a standard deviation of 0.4. With the added white noise, the behavior of the residuals is more clear. The residuals plus the white noise was then used for all of the residual analyses.

Examining the residual plots for the Poisson regression model, Figure 6, one can see that the residuals do not appear to be independent nor normal. The residual plot and the autocorrelation function for the residuals support a lack of independence in the residuals. The Durbin Watson test was performed to help determine whether the residuals of the Poisson model are independent.

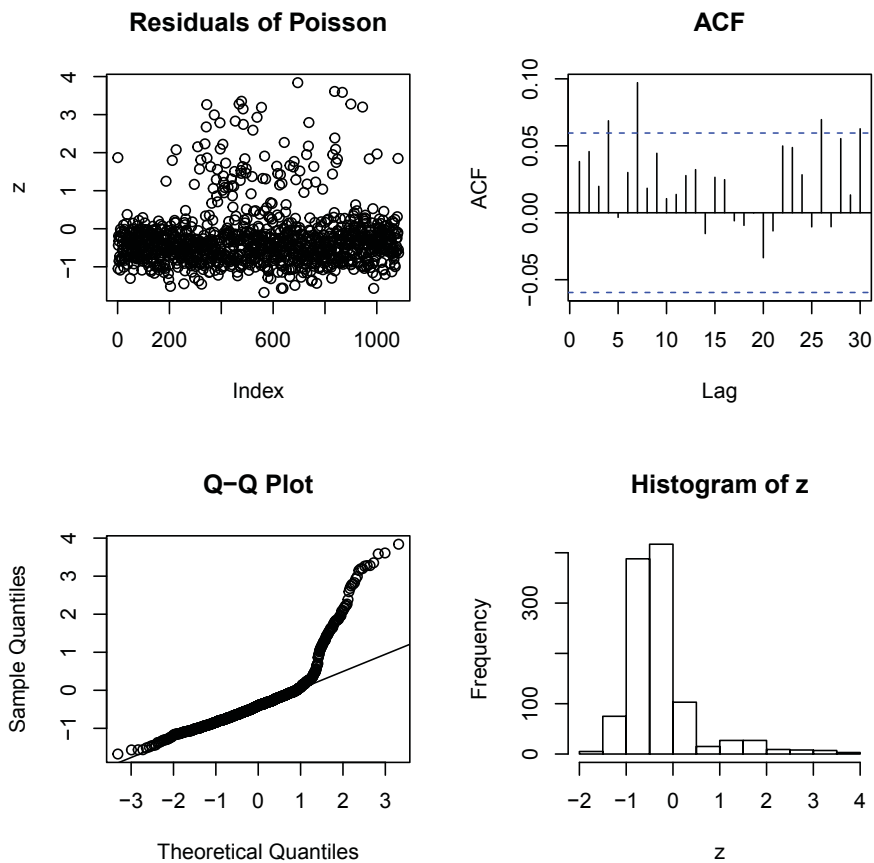


Figure 6: Residual Plots for the Poisson Model

Durbin-Watson test

data: witchesnosalem.Poisson

DW = 1.8427, p-value = 0.00872

alternative hypothesis: true autocorrelation is not 0

With a p-value of $0.00872 < 0.05$, we reject the null hypothesis that the residuals of the Poisson regression are independent. The quantile-quantile plot and histogram in Figure 6 leads one to believe that the residuals of the Poisson regression are not

normally distributed. A Shapiro-Wilk test was used to test whether the residuals of the Poisson model were normally distributed.

Shapiro-Wilk normality test

```
data: z
```

```
W = 0.7893, p-value < 2.2e-16
```

With a p-value of 2.2×10^{-16} , we reject the null hypothesis that the residuals of the Poisson regression are normally distributed. The residual analysis leads us to believe that the Poisson model does not fit the witchcraft accusation data well.

Negative Binomial Residual Analysis: Similar to the case of the Poisson model, the residual plots of the negative binomial model were compressed and hard to interpret. Again, white noise was added to the residuals to better assess their behavior. The residuals plus white noise with standard deviation 0.4, the same as in the Poisson residual analysis, was used in all the residual analyses for the negative binomial model.

Examining the residual plots for the negative binomial regression model, Figure 7, one can see that the residuals do not appear to be independent nor normal. The residual plot and the autocorrelation function for the residuals support a lack of independence in the residuals. The Durbin Watson test was performed to help determine whether the residuals of the negative binomial model are independent.

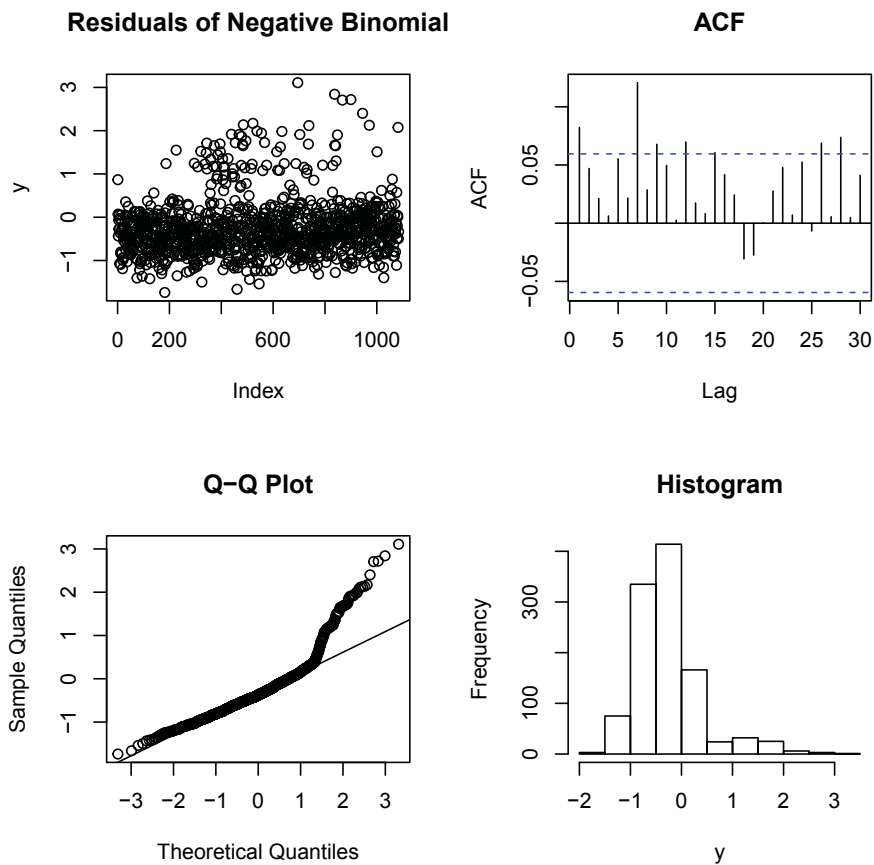


Figure 7: Residual Plots for Negative Binomial Model

Durbin-Watson test

data: witchesnosalem.nb

DW = 1.8427, p-value = 0.00872

alternative hypothesis: true autocorrelation is not 0

With a p-value of $0.00872 < 0.05$, we reject the null hypothesis that the residuals of the negative binomial regression are independent. The quantile-quantile plot and the histogram in Figure 7 leads one to believe that the residuals of the negative

binomial regression are not normally distributed. A Shapiro-Wilk test was used to test whether these residuals were normally distributed.

Shapiro-Wilk normality test

```
data: y
```

```
W = 0.8853, p-value < 2.2e-16
```

With a p-value of 2.2×10^{-16} , we reject the null hypothesis that the residuals of the negative binomial regression are normally distributed. The residual analysis leads us to believe that the negative binomial model does not fit the witchcraft accusation data well.

PAR(1) Model Diagnostics: Diagnostics for the PAR(1) are more difficult to evaluate. The authors of the PAR(p) model determined that the sample autocorrelation function provides a reliable diagnostic for count data with dynamics. Using the sample autocorrelation function of the data, Figure 5, one can see dynamics present in the time series. From this we can determine that a time series count model may better model these data than a Poisson or negative binomial model, since they cannot account for the correlation between event counts. Also, there is a dampening dynamic present in the autocorrelation function, which supports the PAR(p) model being a good model for the data. However the dampening dynamics does not revert to the mean quickly. The authors noted that these persistent type of dynamics can cause problems when modeling the PAR(p) model. As a way to remedy this, the Poisson exponentially weighted moving average model is often fitted. However, the PEWMA model is meant for non-stationary time series, so if the time series is stationary, fitting the PEWMA could lead to incorrect inferences. Determining whether a specific

event count time series model is better approximated either by $\text{PAR}(p)$ or PEWMA by significance test is not readily feasible to date.

When trying to determine which of these three models fit the data best, we cannot reach a definitive conclusion using common model comparison measures. In this case, a measure of comparison, such as comparing Akaike information criterion(AIC), can not be used to compare these models since they are estimated differently. From model diagnostics such as the residual analysis for the Poisson and negative binomial models and the autocorrelation function and model dynamics for the $\text{PAR}(1)$ model, we can conclude that while these models can fit and model the data, they may not be the best models for these data. However due to limitations of the study, which are discussed in Section 4, these models provided a means to conducting an intervention analysis that could take into account the count and serial dependence attributes of the data. So while these models do not fit the data well, they could be the first stepping stones for a more thorough analysis of these data. Once more appropriate models for this type of data are constructed and code for the implementation of those models is developed, there may be opportunity for better fitting models.

4 Discussion

Data on 272 people accused of witchcraft in Colonial America, were collected through resilient research of primary and secondary historical sources. With goals of learning about Colonial American witchcraft accusation patterns, those accused of witchcraft, and how the infamous Salem witch hunt effected Colonial America other than in a strictly social facet, analyses were performed to better understand the Colonial American experience.

A collective portrait of those accused of witchcraft, established through descriptive statistics, provided insight to the characteristics of the victims of witchcraft accusations. Analogous to historical record, those accused of witchcraft were most often women who were married or widowed of Puritan stock. Episodes of witchcraft hysteria, like the infamous Salem witch hunt, saw an increase in males being accused of witchcraft, most often due to guilt by association when a female family member was accused first. These results imply that witchcraft accusations were not necessarily gender specific but more so gender relative.

The number of witchcraft accusations per month from September 1622 to December 1712 were collected. These counts made up time series event counts and were used to do an intervention analysis. The goal of the intervention analysis was to examine whether the Salem witch hunt effected witchcraft accusations in Colonial America after the establishment of the Court of Oyer and Terminer.

The Court of Oyer and Terminer was established by the first royal governor of Massachusetts, Sir William Phips, in May 1692. This court was established to meet the huge demand of cases of those accused and jailed for being associated with witchcraft during the Salem witch hunt. The establishment of the Court of Oyer and Terminer

was designated as the beginning of the permanent intervention period for the intervention analysis because it is a valid historical marker and it played a major role in the trials and executions during the Salem witch hunt.

The social, political, and economic effects of the Salem witch hunt have been studied by many historians. This intervention analysis explored the effects of the Salem witch hunt on witchcraft accusations in the rest of Colonial America in the years after the Salem episode. A Poisson regression, negative binomial regression, and a PAR(1) model were used to estimate the effects of the Salem witch hunt intervention on witchcraft accusations in Colonial America.

The Poisson and negative binomial models were chosen to be used in the intervention analysis because the data consisted of event counts. While these models do not assess the serial correlation of the time series event counts they were still used to study the instantaneous effects of the intervention and to examine how inferences from models that do not model serial correlation can be different from a model designed to take serial correlation into account. The PAR(1) model was chosen to be used in the intervention analysis because it is a time series event count model that takes into account not only the count attributes of the data but also models dynamics like serial correlation. These three different models produced negative intervention coefficients and negative effects, implying that there was a decrease in the number of witchcraft accusations after the Salem witch hunt intervention.

After calculating the instantaneous and long-run effects of the intervention, using all three models, we found that the Poisson and negative binomial models estimated a larger negative instantaneous effect of the intervention than that estimated by the PAR(1) model. However, the long-run effect estimated by the PAR(1) model was larger than those estimated by the Poisson and negative binomial models. The

PAR(1) model took into account future event counts after the intervention to calculate the long-run effect, whereas the Poisson and negative binomial models only estimate the instantaneous effect of the intervention. Using the estimates from all three models, we can conclude that the Salem witch hunt intervention led to a decrease in the number of witchcraft accusations in Colonial America.

It is apparent through the volumes of historical literature that the events of the Salem witch hunt, especially those of the trials that condemned so many to death by hanging, had lasting effects on Colonial America. The events of accusations, trials, and executions of the victims of the Salem witch hunt resonated in the minds of the people of Colonial America. By reading the writings of some of the people directly involved in the Salem witch hunt, historians can understand how people viewed the events of the witch hunt from the perspective of a citizen who experienced the events firsthand. Looking at witchcraft accusations as time series event count data allows one to study Colonial America in a new way. It allows us to analyze the effect of the Salem witch hunt numerically.

The intervention analysis, using a Poisson, negative binomial, and PAR(1) model, determined that after the Salem witch hunt intervention, witchcraft accusations decreased in Colonial America. This decrease in witchcraft accusations could be linked to the social ramifications of the witch hunt. Possible further study could be researching changes in burden of proof standards in witchcraft cases and the lasting social impacts of the stigma of witchcraft. After the Salem episode, did burden of proof become more strict in witchcraft cases? After the Salem episode, were people more apprehensive about accusing others of witchcraft?

This project was an amazing learning experience, but was not without limitations. Data collection was a daunting effort that had limitations. The data sets used for

analysis in this project most likely under-represent the actual number of witchcraft accusations during Colonial America due to lack of available records of the time period in question. Once data was collected and compiled, the low frequencies and the discrete count nature of the data made model identification and specification difficult. After attempting to identify many different models for the time series event count data, with a lot of zero entries, and low frequencies, a well fitting, appropriate model for the data was not definitively found. The models used in the intervention analysis were chosen as the best models available for the analysis at this point in time. However, through research, some other models that are still in their infancy, were discovered that may be more appropriate for these data. Perhaps in the next few years these models will become more readily available with code for model specification and fitting.

Two papers recently published in 2009 and 2012 propose two models that may be appropriate for these data. Neither of these models have code available for analysis. Atanu Biswas and Peter X.K. Song published *Discrete-Valued ARMA Processes* in *Statistics and Probability Letters* in 2009 [1]. In this article, Biswas and Song introduced their model for discrete-valued stationary ARMA data using Pegram's mixing operator. These scholars then collaborated again in 2012 to publish *Statistical Analysis of Discrete-Valued Time Series Using Categorical ARMA Models* in *Computational Statistics and Data Analysis* [32]. In this paper, the scholars improved on their model from *Discrete-Valued ARMA Processes* and performed an analysis on an example problem dealing with data from computer hardware defects. This example included data that was characterized by a lot of zeros, low frequencies, and serial correlation. These are similar characteristics to the witchcraft accusation counts. If there was code available for this model, it could be tested to see if it is appropriate

for the witchcraft accusation counts.

A third paper, a working paper proposed for publishing in 2013, deals with zero-inflated time series count data. *Modeling and Coherent Forecasting of Zero-Inflated Time Series Count Data*, by Raju Maiti and Atanu Biswas et.al presents a Poisson Zero-Inflate Autoregressive model of order p to model time series count data characterized by a lot of zero counts and serial correlation [25]. This model could be tested to see if it is appropriate for the witchcraft accusation data, however code is not available for it at this time. Possible future work would come about if code was constructed and made available for any of these newer models.

While the Poisson, negative binomial, and PAR(1) models may not be the best models to model the witchcraft accusation data, they provided a means of studying witchcraft accusations in a way never before considered. In this case, they were the best models to use to explore and learn more about the time series count data for this thesis.

5 Conclusion

The infamous Salem witch hunt has perplexed and captivated many historians and scholars since its fateful events in 1692. Over 150 people were accused of witchcraft, nineteen people were executed by hanging, and three others died in prison during the Salem witch hunt. The social, political, and economic impacts of this witch hunt have been considered and studied for over a century, but how did this witch hunt effect witchcraft accusations throughout the rest of Colonial America? Who were the victims of witchcraft accusations and what were they like?

Data were collected on 272 victims of witchcraft accusations from September of 1622 to December 1712, including month and year of accusation, gender, marital status, execution status, and colony of residence. Descriptive statistics were used to present a collective portrait of those accused of witchcraft. Those most likely to be accused of witchcraft were married or widowed women from a Puritan background (New England Colonies). Most witchcraft accusations occurred in Spring and Summer months. The data on month and year of accusation made up a data set of time series event counts.

These data were used in an intervention analysis to determine whether or not the Salem witch hunt had an effect on witchcraft accusations in the rest of Colonial America. Three models were used to conduct this intervention analysis. Two of these models, a Poisson and negative binomial model, are used to model event counts and one model, a Poisson autoregressive model of order one is used to model time series event counts with strong dynamics. All three models produced negative intervention coefficients and negative effects. The Poisson and negative binomial models produced larger instantaneous effects than that estimated by the PAR(1) model. However the

PAR(1) model estimated a larger long-run effect than those of the Poisson and negative binomial models. The PAR(1) model estimated that witchcraft accusations decreased by 1.8 in the long run while the Poisson and negative binomial estimated a decline of 0.971 and 0.972, respectively. In the end, the intervention analysis determined that the Salem witch hunt intervention led to a decrease in witchcraft accusations in the rest of Colonial America. Thus the Salem witch hunt did indeed have an effect on witchcraft accusations that came after it in Colonial American history.

References

- [1] Biswas, Atanu and Peter X. K Song. (2009) *Discrete-Valued ARMA Process. Statistics and Probability Letters.* 79, 1884-1889.
- [2] Boyer, Paul and Stephen Nissenbaum. (1974) *Salem Possessed: The Social Origins of Witchcraft.* Cambridge: Harvard University Press.
- [3] Boyer, Paul and Stephen Nissenbaum. (1970) *The Salem Witchcraft Papers: Verbatim Transcripts of the Legal Documents of the Salem Witchcraft Outbreak of 1692,3* New York: Random House.
- [4] Brandt, Patrick, John T. Williams, Benjamin O. Fordham and Brian Pollins. (2000) *Dynamic Modeling for Persistent Event-Count Time Series.* American Journal of Political Science, Vol. 44, No.4, pp.823-843. <http://www.jstor.org/stable/2669284>.
- [5] Brandt, Patrick and John T. Williams. (2001) *A Linear Poisson Autoregressive Model: The Poisson Ar(p) Model.* Political Analysis , Vol. 9, No. 2 , pp. 164-184 <http://www.jstor.org/stable/25791638>.
- [6] Burr, George Lincoln, Ed. (1992) *Narratives of the Witchcraft Cases 1648-1706.* New York: The Notable Trials Library.
- [7] Cameron, A. Colin and Pravin K. Trivedi. (1998) *Regression Analysis of Count Data.* New York: Cambridge University Press.
- [8] Cryer, Jonathan A. and Kung-Sik Chan. (2008) *Time Series Analysis With Applications in R.* Second Edition. New York: Springer.

- [9] DeGroot, Morris H. (1970) *Optimal Statistical Decisions*. New York; McGraw Hill.
- [10] Demos, John. (2008) *The Enemy Within: 2,000 Years of Witch-hunting in the Western World*. New York: Penguin Group.
- [11] Demos, John Putnam. (1982) *Entertaining Satan: Witchcraft and the Culture of Early New England*. New York: Oxford University Press.
- [12] Demos, John. (1970) *A Little Common Wealth: Family Life In Plymouth Colony* New York: Oxford University Press.
- [13] Godbeer, Richard. (1992) *The Devil's Dominion: Magic and Religion in Early New England*. Cambridge: Cambridge University Press.
- [14] Godbeer, Richard. (2005) *Escaping Salem: The Other Witch Hunt of 1692*. New York: Oxford University Press.
- [15] Hall, David D. (1991) *Witch-Hunting in Seventeenth-Century New England: A Documentary History, 1638-1693*. Boston: Northeastern University Press.
- [16] Harvey, A.C. and C. Fernandes. (1989) *Time Series Models for Count or Qualitative Observations*. *Journal of Business and Economic Statistics*. Vol, 7. pp. 407-417.
- [17] Jackman, Simon. (2000) *In and Out of War and Peace: Transitional Models of International Conflict*. Stanford University. <http://jackman.stanford.edu/papers/inandout.pdf>.

- [18] Jacobs, P.A. and P. A. W. Lewis. *Discrete Time Series Generated by Mixtures. I: Correlational and Runs Properties. Journal of the Royal Statistical Society. Series B (Methodological)*. Vol. 40, No. 1 (1978) , pp. 94-105. <http://www.jstor.org/stable/2984870>.
- [19] Karlsen, Carol F. (1987). *The Devil in the Shape of a Woman: Witchcraft in Colonial New England*. New York: W.W. Norton and Co.
- [20] King, Gary. (1988) *Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model. American Journal of Political Science*, Vol. 32, No. 3, pp.838-863. Midwest Political Science Association. <http://www.jstor.org/stable/2111248>.
- [21] King, Gary. (1989) Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator. *American Journal of Political Science*, Vol.33. No.3 pp. 762-784. <http://www.jstor.org/stable/2111071>.
- [22] Koehler, Lyle. (1980) *A Search for Power. The "Weaker" Sex in Seventeenth-Century New England*. Chicago: University of Illinois Press.
- [23] Latner, Richard. (2008) *Salem Witchcraft, Factionalism, and Social Change Reconsidered: Were Salem's Witch-hunters Modernization's Failures?* *William and Mary Quarterly*, 3d Series, Volume LXv, Number 3, July.
- [24] Latner, Richard. *Salem Witchcraft Site*. Tulane University. <http://www.tulane.edu/salem/>

- [25] Maiti, Raju, Atanu Biswas et. al. (2013) *Modeling and Coherent Forecasting of Zero-Inflated Time Series Count Data*. Working Paper. Indian Institute of Management, Research and Publications.
- [26] McCleary, Richard and Richard A. Hay Jr. (1980) *Applied Time Series Analysis for the Social Sciences*. Beverly Hills: Sage Publications.
- [27] Nelder, J.A. and R.W.M. Wedderburn. (1972) *Generalized Linear Models*. *Journal of Royal Statistical Society: Series A (General)*. Vol. 135 No.3 pp.370-384. <http://www.jstor.org/stable/2344614>.
- [28] Norton, Mary Beth. (2002) *In The Devil's Snare: The Salem Witchcraft Crisis of 1692*. New York: Random House.
- [29] Oster, Emily. (2004) *Witchcraft, Weather and Economic Growth in Renaissance Europe*. *The Journal of Economic Perspectives*, Vol.18, No.1, pp.215-228.
- [30] Rosenthal, Bernard, ed. (2009) *Records of the Salem Witch-Hunt* New York: Cambridge University Press.
- [31] Shumway, Robert H. and David S. Stoffer. (2011) *Time Series Analysis and Its Applications: With R Examples*. New York: Springer.
- [32] Song, Peter X. K. and R. Keith Freeland, Atanu Biswas, Shulin Zhang.(2013) *Statistical Analysis of Discrete-Values Time Series Using Categorical ARMA Models*. *Computational Statistics and Data Analysis*, 57, 112-124.
- [33] Starkey, Marion L. (1949) *The Devil in Massachusetts: A Modern Enquiry into the Salem Witch Trials*. New York: Alfred A. Knopf Inc.

- [34] Taylor, John M. (1908) *The Witchcraft Delusion in Colonial Connecticut 1647-1697* Williamstown: Corner House.
- [35] Upham, Charles W. (1867) *Salem Witchcraft: With an Account of Salem Village and A History of Opinions on Witchcraft and Kindred Subjects*. New York: Wiggin and Lunt.
- [36] White, Halbert. (1994) *Estimation, Inference and Specification Analysis*. New York: Cambridge University Press.
- [37] Wood, Simon N. (2006) *Generalized Additive Models: An Introduction With R*. Boca Raton: Chapman and Hall.
- [38] *Records and Files of the Quarterly Courts of Essex County*. Vol IV, 129-133.
<https://archive.org/stream/recordsfilesqu04esse#page/130/mode/2up>
<http://salem.lib.virginia.edu/Essex/vol5/index/index.html>

A Complete R Scripts

The code for the $\text{PAR}(p)$ was created by Dr. Patrick Brandt, one of the co-authors of the $\text{PAR}(p)$ model. This code includes the $\text{PAR}(p)$ function, plus the code created for a different model, the Poisson Exponentially Weighted Moving Average model, PEWMA. This code is called “pests.r” and needs to be run in order for the $\text{PAR}(p)$ function to work. This code can be found at

<http://www.utdallas.edu/~pxb054000/code/pests.r>

Load the data:

```
tsnosalem <- read.table(file = "monthtimeseriesnosalem.csv", header=F,  
col.names=c("Accused.Witches1"))
```

```
attach(tsnosalem)
```

Define the intervention variable:

```
nosalem1 <- c(rep(0,839), rep(1,245))
```

Plot the time series:

```
plot(ts(tsnosalem, start=c(1622,1), freq=12), ylab="Accused Witches",  
main="Witchcraft Accusation Data")
```

Plot the ACF of the series:

```
acf(tsnosalem)
```

Perform a Poisson Regression:

```
witchesnosalem.Poisson <- glm(Accused.Witches1 nosalem1, family=poisson)  
print(summary(witchesnosalem.Poisson))
```

Perform a Negative Binomial Regression:

```
witchesnosalem.nb <- glm.nb(Accused.Witches1 nosalem1)  
print(summary(witchesnosalem.nb))
```

PAR(1) Model:

```
witchesnosalem.PAR1 <- Parp(Accused.Witches1 nosalem1, p=1, parp.init  
= rep(0.1,1), init.param=NULL)  
print(witchesnosalem.PAR1)
```

Calculate the PAR(1) Multipliers:

```
parp.multipliers(witchesnosalem.PAR1)
```

Poisson Residual Analysis:

Add noise to the residuals:

```
d=rstandard(witchesnosalem.Poisson)
```

```
z= d + rnorm(1084, sd=0.4)
```

Look at the residual plots:

```
plot(z)
```

```
acf(z)
```

Durbin Watson Test:

```
dwtest(witchesnosalem.Poisson, alternative="two.sided")
```

```
qqnorm(z)
```

```
qqline(z)
```

Shapiro- Wilk Test:

```
shapiro.test(z)
```

Negative Binomial Residual Analysis:

Add noise to the residuals:

```
x=rstandard(witchesnosalem.nb)
```

```
y= x + rnorm(1084, sd=0.4)
```

Look at the residual plots:

```
plot(y)
```

`acf(y)`

Durbin Watson Test:

`dwtest(witchesnosalem.nb, alternative="two.sided")`

`qqnorm(y)`

`qqline(y)`

Shapiro- Wilk Test:

`shapiro.test(y)`

Diagnostics for the PAR(1) Model:

`acf(tsnosalem)`

B Data Set References

1. Burr, George Lincoln, Ed. *Narratives of the Witchcraft Cases 1648-1706*.
2. Demos, John. *The Enemy Within: 2,000 Years of Witch-hunting in the Western World*.
3. Godbeer, Richard. *The Devil's Dominion: Magic and Religion in Early New England*.
4. Godbeer, Richard. *Escaping Salem: The Other Witch Hunt of 1692*
5. Hall, David D. *Witch-Hunting in Seventeenth-Century New England: A Documentary History, 1638-1693*.
6. Karlsen, Carol F. *The Devil in the Shape of a Woman: Witchcraft in Colonial New England*.
7. Koehler, Lyle. *A Search for Power. The "Weaker" Sex in Seventeenth-Century New England*.
8. Latner, Richard. *Salem Witchcraft Site*. Tulane University.
<http://www.tulane.edu/salem/>
9. Norton, Mary Beth. *In The Devil's Snare: The Salem Witchcraft Crisis of 1692*.
10. Starkey, Marion L. (1949). *The Devil in Massachusetts: A Modern Enquiry into the Salem Witch Trials*.
11. Taylor, John M. (1908) *The Witchcraft Delusion in Colonial Connecticut 1647-1697*.

12. *Records and Files of the Quarterly Courts of Essex County, Volumes I-IX.*