Parametric Estimation of Storm IBNR (Incurred But Not Reported Claims)

An Honors Thesis by

Bradford K. Hoagland

Thesis Advisor
John A. Beekman, Ph.D.

John A. Beekman, Ph.D.
(Advisor's Signature)

Ball State University
Muncie, IN
April, 1994

Expected date of graduation -- May 1994

Acknowledgements

The research and development of this thesis was quite involved. Without the assistance of several helpful individuals, progress would have been much slower than it was.

I would like to thank Kevin Murphy of American States Insurance for devoting considerable amounts of time explaining to me the details of IBNR and helping me to stay on track in my development of the equations involved in this thesis. Mr. Murphy taught me almost everything I know about IBNR.

I would also like to thank Larry Haefner of American States Insurance for suggesting the topic of parametric estimation of storm IBNR and allowing me to use some of company time for my research.

For devoting a portion of his day to aid me in speeding up the process of solving for the parameters of the discrete approximation of the Pareto distribution, I would like to show my appreciation to Dr. Earl McKinney of Ball State University.

Dr. John Beekman of Ball State University provided the necessary guidance and support I needed to complete this thesis. For this he deserves much thanks.

Table of Contents

1. Introduction

Incurred but not reported claims (IBNR claims) are claims where the accident has occurred, but the claims have not yet been reported. It is necessary for actuaries to be able to accurately estimate these IBNR claims in order to properly adjust reserve amounts. Actuaries currently use prior experience to project the number of claims to be reported in the future. An example of one method of using this prior experience to estimate ultimate values is shown in *Table 1.1*.

Estimated Ultimates by Accident Year
Based upon Incurred Losses

Acc. Year	Beginning Incurred	Days Early	Adjusted Incurred	12 Mo.	24 Mo.	36 Mo.	48 Mo.	Latest	Bondy Factor	Ultimates
12/85	29,406	0	29,406	34,031	33,828	33,536	33,459	33,409		33,409
				1.157	0.994	0.991	0.998	0.999	1.000	
12/86	28,278	0	28,278	32,755	32,437	32,271	32,195	32,135		32,135
				1.158	0.990	0.995	0.998	0.998	1.000	
12/87	43,083	0	43,083	48,828	48,422	48,194	48,178	48,123		48,123
				1.133	0.992	0.995	1.000	0.999	1.000	
12/88	50,311	1	50,435	55,960	55,274	55,146	55,045	55,017		55,017
				1.110	0.988	0.998	0.998	0.999	1.000	
12/89	57,906	2	58,162	57,588	57,109	56,928	56,894			56,837
				0.990	0.992	0.997	0.999	0.999	1.000	
12/90	58,417	3	58,811	58,546	58,157	57,902				57,786
				0.995	0.993	0.996	0.999	0.999	1.000	
12/91	53,219	0	53,219	53,082	52,699					52,436
				0.997	0.993	0.997	0.999	0.999	1.000	
12/92	48,834	0	48,834	48,885						48,301
				1.001	0.993	0.997	0.999	0.999	1.000	
12/93	48,394	0	48,394							47,720
				0.998	0.993	0.997	0.999	0.999	1.000	

Table 1.1

The numbers in green are calculated by dividing the incurred losses immediately above them by the incurred losses to the left (12 months prior). The numbers in red are estimated factors based on the observed factors listed above them. The estimated ultimates are calculated by multiplying the most recent observed amount by each of the 12-month factors to its right.

I recently heard an accountant ask an actuary how IBNR claims are estimated for storms. The actuary jokingly responded, "That's when we put away our little dart board and get out our big dart board." His response does show that storm IBNR estimation is much more difficult than estimation for other forms of accidents. Actuaries currently use sophisticated models combined with past experience to estimate storm IBNR. The purpose of this paper is to offer another method of storm IBNR estimation.

2. The Basic Method

In Edward W. Weissner's article in *Proceedings*, Casualty Actuarial Society, [4], titled "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood," he discusses how using the basic technique of maximum likelihood estimation on a specified truncated distribution, one could estimate the distribution of claims for a specific accident period and, ultimately, determine the number of claims yet to be reported (IBNR claims). Using the exponential distribution as an example, Weissner let the random variable X denote the "report lag," or the time elapsed between the accident and the moment the claim was reported, and let the variable c denote the maximum possible report lag.

In order to employ the maximum likelihood method he developed the distribution in the following way: The probability density function of a claim being reported at exact moment x is

$$f(x) = \theta e^{-\theta x},\ 0 < x < \infty. \tag{2.1.1}$$

Because we have only received claims reported before time c, we need to truncate the p.d.f. to only include the claims reported between time 0 and time c.

$$\begin{aligned} Pr[X \leq c] &= \int_0^c f(x)\,dx \\ &= 1 - e^{-\theta c} \end{aligned} \tag{2.1.2}$$

$$\therefore\ f(x \mid x \leq c) = \frac{\theta e^{-\theta x}}{1 - e^{-\theta c}}. \tag{2.1.3}$$

After performing the method of maximum likelihood and utilizing the Newton-Raphson method, Weissner determined that

$$\hat{\theta}_{m+1} = \hat{\theta}_m - \frac{\dfrac{n}{\hat{\theta}_m} - \sum_{i=1}^{n} x_i - \dfrac{nce^{-c\hat{\theta}_m}}{1 - e^{-c\hat{\theta}_m}}}{\dfrac{-n}{\hat{\theta}_m^2} + \dfrac{nc^2 e^{-c\theta_m}}{\left(1 - e^{-c\hat{\theta}_m}\right)^2}}, \tag{2.1.4}$$

where x_i is the report lag for claim i, n is the total number of claims, and m denotes the m^{th} iteration. Once θ has been determined, the IBNR claims could be calculated using the following formulas:

$$\textit{Total Claims} = \frac{n}{1 - e^{-\theta c}}. \tag{2.1.5}$$

$$\textit{IBNR Claims} = \textit{Total Claims} - n. \tag{2.1.6}$$

Because the case explored by Weissner deals with claims for several different accidents that occur in a given period, it is unavoidable that all accidents and report lags must be assumed to occur in the middle of the period. For example, if the report lag is in terms of months, then all accidents in the specified accident month are assumed to occur in the middle of the month, and all claims are assumed to be reported in the middle of their respective months.

When we are dealing with claims resulting from a storm, however, the assumption that the storm occurred in the middle of the specified period need not be made. We know the exact date of the storm, and can thus calculate the associated claim report lags exactly. A problem that occurs when exact lags are used, however, is that data may not exist for each claim, but only for aggregate claims for a given period such as a week. Another problem might be that if a storm occurs on a weekend, the Monday immediately following may have an unusually high number of claims. For example, if a storm occurs on a Friday at 6:00 p.m., all of the claims that normally would have been reported Friday evening, Saturday, and Sunday, may be reported on Monday, along with the regularly expected Monday claims. If the assumption is made that all claims are reported at one point in the specified period, the IBNR claims estimate may be thrown off. If claims are assumed to fit the exponential distribution, the parameter θ would be under-estimated, and the IBNR claims would be over-estimated.

As an example of this problem let us create a hypothetical storm where claims are distributed exponentially with parameter $\theta = 0.3$, and the total number of claims to be reported is 100,000. Let us use five weeks of claim data. Claims are reported according to *Table 2.1*. These "reported claims" were created using the formula $claims = 100{,}000 \cdot (e^{-0.3 \cdot (Week - 1)} \cdot e^{-0.3 \cdot Week})$, then rounded. Using equation (2.1.4), with the assumption that claims are reported in the middle of each time period, we estimate θ to be approximately 0.286703. This causes our estimated claims to appear as they do in *Table 2.1*.

Reported and Estimated Claims

Week	Reported Claims	Estimated Claims
1	25,918	25,428
2	19,201	19,090
3	14,224	14,332
4	10,538	10,759
5	7,806	8,077

Table 2.1

As shown in the graph in *Figure 2.1*, the estimated distribution decreases less rapidly than the observed distribution. It is obvious, then, that our estimated IBNR claims will be greater than the true IBNR claims. We know that there will be a total of 100,000 claims (because we set our example up that way), so we know that there are 22,313 IBNR claims. Using our estimated parameter, however, we estimate the number of IBNR claims to be 24,327. This estimated number is 2,014 or 9% greater than the true value, which is far from sufficient (we note, also, that this estimated distribution is rejected using the chi-square goodness of fit test to be discussed later).

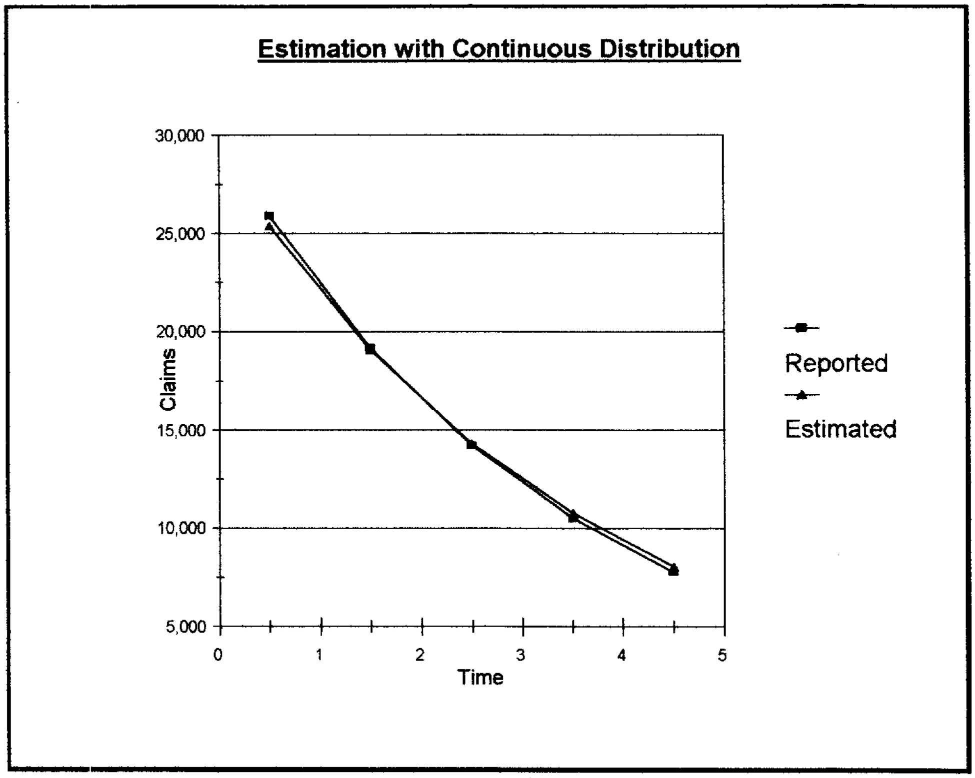

Figure 2.1

A way around this problem would be to replace a continuous distribution with a discrete approximation to it. We do this by letting the random variable X denote the period, such as a week, in which the claim was reported. What this does is use a value obtained over time (number of claims reported in a period) as a measure over time rather than being used as an instantaneous measure. We will now explore this method in greater detail.

3. Discrete Approximation of the Exponential Distribution

3.1. Introduction

As many know, the exponential distribution serves very limited practical use in claim estimation. We will use the exponential distribution, however, as a simpler means of presenting the basic methods of storm IBNR claims estimation. The equations developed for the exponential distribution are much less involved than much more practical distributions, and, hence, reduce confusion while learning the ideas presented. But as we soon will see, these equations are far from simple.

In creating a discrete approximation of the exponential distribution for the purpose of storm IBNR estimation, we let

$$\begin{aligned} p(x) &= Pr[x-1 \leq X \leq x] \\ &= \int_{x-1}^{x} \theta e^{-\theta t} dt \qquad (3.1.1) \\ &= e^{-(x-1)\theta} - e^{-x\theta},\ x \in 1,2,3,\ldots,\quad \theta > 0. \end{aligned}$$

What this does is define a probability function where $f(x)$ is the probability of a claim being reported during period x. The truncated distribution then becomes

$$p(x \mid x \leq c) = \frac{e^{-(x-1)\theta} - e^{-x\theta}}{1 - e^{-c\theta}},\ x \in 1,2,3,\ldots,c. \qquad (3.1.2)$$

A graph of such a truncated distribution with $\theta = 0.3$ and $c = 5$ is illustrated in *Figure 3.1.*

Weissner used only maximum likelihood estimation in his article. We will explore two other methods of estimation, the method of moments and least squares estimation, in addition to the maximum likelihood estimation. Thus, it is necessary for us to calculate the mean of this truncated distribution. We will also calculate the variance for those who are curious.

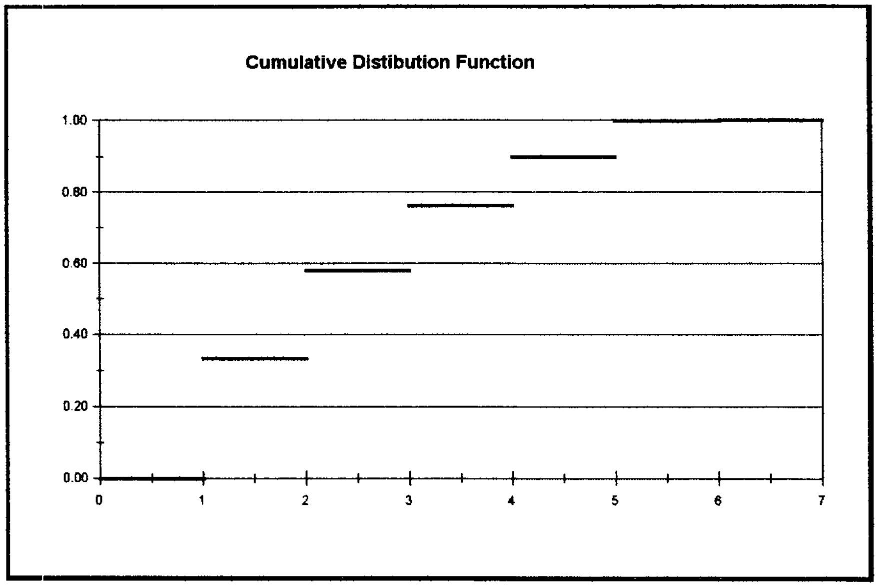

Figure 3.1

$$
\begin{aligned}
E[X|x \le c] &= \sum_{x=1}^{c} \frac{xe^{-(x-1)\theta} - xe^{-x\theta}}{1-e^{-c\theta}} \\
&= \frac{1-e^{-\theta}+2e^{-\theta}-2e^{-2\theta}+\dots+ce^{-(c-1)\theta}-ce^{-c\theta}}{1-e^{-c\theta}} \qquad \textbf{(3.1.3)} \\
&= \frac{1+e^{-\theta}+e^{-2\theta}+e^{-3\theta}+\dots+e^{-(c-1)\theta}}{1-e^{-c\theta}} - \frac{ce^{-c\theta}}{1-e^{-c\theta}}
\end{aligned}
$$

Using the formula

$$
1+r+r^2+r^3+\dots+r^{n-1} = \frac{1-r^n}{1-r}, \qquad \textbf{(3.1.4)}
$$

where $-1 < r < 1$, we find that

$$E[X|x \le c] = \frac{1-e^{-c\theta}}{(1-e^{-\theta})(1-e^{-c\theta})} - \frac{ce^{-c\theta}}{1-e^{-c\theta}}$$

$$= \frac{1}{1-e^{-\theta}} - \frac{ce^{-c\theta}}{1-e^{-c\theta}}. \qquad \textbf{(3.1.5)}$$

To determine the variance we subtract the square of the mean from $E[X^2|x \le c]$.

$$E[X^2|x \le c] = \sum_{x=1}^{c} \frac{x^2e^{-(x-1)\theta} - x^2e^{-x\theta}}{1-e^{-c\theta}}$$

$$= \frac{1-e^{-\theta}+4e^{-\theta}-4e^{-2\theta}+\dots+c^2e^{-(c-1)\theta}-c^2e^{-c\theta}}{1-e^{-c\theta}}$$

$$= \frac{1+3e^{-\theta}+5e^{-2\theta}+7e^{-3\theta}+\dots+(2c-1)e^{-(c-1)\theta}-c^2e^{-c\theta}}{1-e^{-c\theta}} \qquad \textbf{(3.1.6)}$$

We can now simplify the numerator in the following way:

$$1+3e^{-\theta}+5e^{-2\theta}+7e^{-3\theta}+\dots+(2c-1)e^{-(c-1)\theta}$$

$$\begin{aligned} = 1 + e^{-\theta} + e^{-2\theta} + e^{-3\theta} + \dots + e^{-(c-1)\theta} \\ +2e^{-\theta} + 4e^{-2\theta} + 6e^{-3\theta} + \dots + (2c-2)e^{-(c-2)\theta} \end{aligned}$$

$$\begin{aligned} = \frac{1-e^{-c\theta}}{1-e^{-\theta}} \\ +2e^{-\theta}(1+2e^{-\theta}+3e^{-2\theta}+4e^{-3\theta}+\dots+(c-1)e^{-(c-2)\theta}) \end{aligned} \qquad \textbf{(3.1.7)}$$

Continuing to apply this method to each term, the portion of equation (3.1.7) to the right side of the equal sign simplifies to

$$\left(\frac{1-e^{-c\theta}}{1-e^{-\theta}}\right)+2e^{-\theta}\left(\frac{1-e^{-(c-1)\theta}}{1-e^{-\theta}}\right)+2e^{-2\theta}\left(\frac{1-e^{-(c-2)\theta}}{1-e^{-\theta}}\right)+2e^{-3\theta}\left(\frac{1-e^{-(c-3)\theta}}{1-e^{-\theta}}\right)+\dots+2e^{-(c-1)\theta}\left(\frac{1-e^{-\theta}}{1-e^{-\theta}}\right)$$

$$= \frac{(1-e^{-c\theta})+2e^{-\theta}\left(\frac{1-e^{-(c-1)\theta}}{1-e^{-\theta}}\right)-2(c-1)e^{-c\theta}}{1-e^{-\theta}}$$

$$= \frac{1+e^{-\theta}-(2c+1)e^{-c\theta}+(2c-1)e^{-(c+1)\theta}}{(1-e^{-\theta})^2}. \tag{3.1.8}$$

$$\therefore E[X^2|x\leq c] = \frac{1+e^{-\theta}-(2c+1)e^{-c\theta}+(2c-1)e^{-(c+1)\theta}}{(1-e^{-\theta})^2(1-e^{-c\theta})} - \frac{c^2e^{-c\theta}}{1-e^{-c\theta}}. \tag{3.1.9}$$

Thus, we determine the variance to be

$$\frac{1+e^{-\theta}-(2c+1)e^{-c\theta}+(2c-1)e^{-(c+1)\theta}}{(1-e^{-\theta})^2(1-e^{-c\theta})} - \frac{c^2e^{-c\theta}}{1-e^{-c\theta}} - \left(\frac{1}{1-e^{-\theta}} - \frac{ce^{-c\theta}}{1-e^{-c\theta}}\right)^2. \tag{3.1.10}$$

3.2. Method of Moments

In applying the method of moments, we set the sample mean equal to the mean of the distribution and solve for our parameter θ.

$$\begin{aligned}\bar{x} &= E[X|x\leq c]\\ &= \frac{1}{1-e^{-\theta}} - \frac{ce^{-c\theta}}{1-e^{-c\theta}}\end{aligned} \tag{3.2.1}$$

$$\therefore \frac{1}{1-e^{-\theta}} - \frac{ce^{-c\theta}}{1-e^{-c\theta}} - \bar{x} = 0. \tag{3.2.2}$$

Solving this equation for θ can prove to be quite difficult. Resorting to numerical methods, such as the Newton-Raphson method [1], might be extremely helpful. Recall that the Newton-Raphson method states that

$$\hat{\theta}_{m+1} = \hat{\theta}_m - \frac{g(\hat{\theta}_m)}{g'(\hat{\theta}_m)}. \qquad \textbf{(3.2.3)}$$

Setting $g(\theta)$ equal to equation (3.2.2) we find that

$$g'(\theta) = \frac{c^2 e^{-c\theta}}{(1-e^{-c\theta})^2} - \frac{e^{-\theta}}{(1-e^{-\theta})^2}. \qquad \textbf{(3.2.4)}$$

Thus,

$$\theta_{m+1} = \theta_m - \frac{\dfrac{1}{1-e^{-\theta_m}} - \dfrac{ce^{-c\theta_m}}{1-e^{-c\theta_m}} - \bar{x}}{\dfrac{c^2 e^{-c\theta_m}}{(1-e^{-c\theta_m})^2} - \dfrac{e^{-\theta_m}}{(1-e^{-\theta_m})^2}}. \qquad \textbf{(3.2.5)}$$

To demonstrate the use of equation (3.2.5), we create the following example: Suppose a storm occurs and within five weeks an insurance company has received 719 claims according to the schedule in *Table 3.1.* We find the sample mean to be approximately 1.88873. Therefore, we start with θ_1 = reciprocal of the sample mean $\doteq$ 0.529455. Applying equation (3.2.5) until we achieve a tolerance of 5×10^{-8}, we calculate our final estimate of θ to be 0.64147761 (Our estimates at each iteration are shown in *Table 3.2.*). Using equations (2.1.5) and (2.1.6) we estimate a total of 30 IBNR claims. After several more weeks we find that we have an ultimate total of 743 claims for this storm; thus, our actual IBNR claims at the end of the fifth week was 24 claims -- very close to our estimate of 30 claims. (The data used for this example was created using a random number generator for the exponential distribution with parameter θ = 0.673 and a total number of claims of 743.)

Claim Schedule	
Week	Claims
1	364
2	181
3	97
4	44
5	33

Table 3.1

Iterations
$\theta_1 = 0.52945508$
$\theta_2 = 0.64293257$
$\theta_3 = 0.64143044$
$\theta_4 = 0.64147761$
$\theta_5 = 0.64147761$

Table 3.2

3.3. Maximum Likelihood Estimation

In using the maximum likelihood method, we first need to develop the likelihood function.

$$L(\theta) = \prod_{i=1}^{n} f(x_i | x_i \leq c) = \frac{\prod_{i=1}^{n} \left(e^{-(x_i-1)\theta} - e^{-x_i\theta}\right)}{(1-e^{-c\theta})^n}, \tag{3.3.1}$$

where x_i is the period in which claim i was reported and n is the number of claims reported within c periods. Finding the maximum of this equation is made somewhat easier if we first take the natural log of both sides of the equation.

$$\ln L(\theta) = \sum_{i=1}^{n} \ln\left(e^{-(x_i-1)\theta} - e^{-x_i\theta}\right) - n\ln(1-e^{-c\theta}) \tag{3.3.2}$$

$$\frac{d}{d\theta}\ln L(\theta) = \sum_{i=1}^{n} \frac{x_i e^{-x_i\theta} - (x_i - 1)e^{-(x_i-1)\theta}}{e^{-(x_i-1)\theta} - e^{-x_i\theta}} - \frac{nce^{-c\theta}}{1 - e^{-c\theta}}$$

$$= \sum_{i=1}^{n} \frac{-x_i\left(e^{-(x_i-1)\theta} - e^{-x_i\theta}\right) + e^{-(x_i-1)\theta}}{e^{-(x_i-1)\theta} - e^{-x_i\theta}} - \frac{nce^{-c\theta}}{1 - e^{-c\theta}}$$

$$= \sum_{i=1}^{n} \left(\frac{e^{-(x_i-1)\theta}}{e^{-(x_i-1)\theta} - e^{-x_i\theta}} - x_i \right) - \frac{nce^{-c\theta}}{1 - e^{-c\theta}}$$

$$= \frac{n}{1 - e^{-\theta}} - \frac{nce^{-c\theta}}{1 - e^{-c\theta}} - \sum_{i=1}^{n} x_i = 0 \qquad \textbf{(3.3.3)}$$

Dividing both sides by n we achieve the same equation as equation (3.2.2) in the method of moments. Therefore, the Newton-Raphson method determines that

$$\theta_{m+1} = \theta_m - \frac{\dfrac{1}{1 - e^{-\theta_m}} - \dfrac{ce^{-c\theta_m}}{1 - e^{-c\theta_m}} - \bar{x}}{\dfrac{c^2 e^{-c\theta_m}}{\left(1 - e^{-c\theta_m}\right)^2} - \dfrac{e^{-\theta_m}}{\left(1 - e^{-\theta_m}\right)^2}}, \qquad \textbf{(3.3.4)}$$

exactly as in the method of moments. It is important to note that the method of moments and maximum likelihood estimation do not always produce the equivalent results for all distributions.

3.4. Least Squares Estimation

In applying the least squares method, we first sum the squares of the differences between the actual probability of a claim being reported in each time period and the observed probability of a claim being reported in each time period.

$$SS = \sum_{x=1}^{c} \left(\frac{e^{-(x-1)\theta} - e^{-x\theta}}{1 - e^{-c\theta}} - f^{\circ}(x) \right)^2 \qquad \textbf{(3.4.1)}$$

The process of maximizing this formula is greatly simplified by instead maximizing

$$(1 - e^{-c\theta})^2 SS = \sum_{x=1}^{c} \left(e^{-(x-1)\theta} - e^{-x\theta} - (1 - e^{-c\theta}) f^{\circ}(x)\right)^2. \qquad \textbf{(3.4.2)}$$

Taking the derivative of this equation with respect to θ we get

$$2\sum_{x=1}^{c} \left[-(x-1)e^{-(x-1)\theta} + xe^{-x\theta} - ce^{-c\theta} f^{\circ}(x)\right] \cdot \left[e^{-(x-1)\theta} - e^{-x\theta} - (1 - e^{-c\theta}) f^{\circ}(x)\right] = 0. \qquad \textbf{(3.4.3)}$$

We will again apply the Newton-Raphson method. Setting $g(\theta)$ equal to equation (3.4.3) we find that

$$g'(\theta) = 2\sum_{x=1}^{c} \left(\begin{array}{l} \left[(x-1)^2 e^{-(x-1)\theta} - x^2 e^{-x\theta} + c^2 e^{-c\theta} f^{\circ}(x)\right] \cdot \left[e^{-(x-1)\theta} - e^{-x\theta} - (1 - e^{-c\theta}) f^{\circ}(x)\right] \\ + \left[-(x-1)e^{-(x-1)\theta} + xe^{-x\theta} - ce^{-c\theta} f^{\circ}(x)\right]^2 \end{array} \right). \qquad \textbf{(3.4.4)}$$

$$\therefore \hat{\theta}_{m+1} = \hat{\theta}_m - \frac{\sum_{x=1}^{c} \left[-(x-1)e^{-(x-1)\hat{\theta}_m} + xe^{-x\hat{\theta}_m} - ce^{-c\hat{\theta}_m} f^{\circ}(w_i)\right] \cdot \left[e^{-(x-1)\hat{\theta}_m} - e^{-x\hat{\theta}_m} - (1 - e^{-c\hat{\theta}_m}) f^{\circ}(w_i)\right]}{\sum_{i=1}^{c} \left(\begin{array}{l} \left[(x-1)^2 e^{-(x-1)\hat{\theta}_m} - x^2 e^{-x\hat{\theta}_m} + c^2 e^{-c\hat{\theta}_m} f^{\circ}(x)\right] \cdot \left[e^{-(x-1)\hat{\theta}_m} - e^{-x\hat{\theta}_m} - (1 - e^{-c\hat{\theta}_m}) f^{\circ}(x)\right] \\ + \left[-(x-1)e^{-(x-1)\hat{\theta}_m} + xe^{-x\hat{\theta}_m} - ce^{-c\hat{\theta}_m} f^{\circ}(x)\right]^2 \end{array} \right)} \qquad \textbf{(3.4.5)}$$

Using the same data and tolerance as in the example for the method of moments, we estimate θ to be 0.66868855. We then estimate that there are 26 IBNR claims -- two claims more than what the actual IBNR claims are. Even though our estimate of IBNR claims is a little better using the least squares method, the least squares method is not always the best approximation.

4. Discrete Approximation of the Pareto Distribution

4.1. Introduction

As stated earlier, the exponential distribution serves very little practical purpose in storm IBNR estimation. One usually needs a more complex distribution such as the Pareto distribution or the log normal distribution. If one decides that the Pareto distribution is most likely to fit the given data, he or she needs to develop a discrete approximation in the same manner as with the exponential distribution. The continuous Pareto distribution has a p.d.f. of the following form:

$$f(x) = \alpha \lambda^{\alpha} (\lambda + x)^{-\alpha-1}, \ \alpha > 0, \ \lambda > 0. \tag{4.1.1}$$

To develop the necessary discrete approximation of this distribution we integrate this p.d.f. from x - 1 to x.

$$\begin{aligned} p(x) &= \int_{x-1}^{x} \alpha \lambda^{\alpha} (\lambda + t)^{-\alpha-1} dt \\ &= \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}, \ x \in 1,2,3,\ldots \end{aligned} \tag{4.1.2}$$

Therefore,

$$p(x \mid x \leq c) = \frac{\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}, \quad x \in 1,2,3,\ldots,c. \tag{4.1.3}$$

When we apply the method of moments, it will be necessary for us to use both the mean and the variance of this truncated distribution.

$$E[X \mid x \leq c] = \sum_{x=1}^{c} \frac{x\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - x\left(\frac{\lambda}{\lambda + x}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} \tag{4.1.4}$$

$$= \frac{1 - \left(\frac{\lambda}{\lambda+1}\right)^{\alpha} + 2\left(\frac{\lambda}{\lambda+1}\right)^{\alpha} - 2\left(\frac{\lambda}{\lambda+2}\right)^{\alpha} + 3\left(\frac{\lambda}{\lambda+2}\right)^{\alpha} - 3\left(\frac{\lambda}{\lambda+3}\right)^{\alpha} + \cdots - (c-1)\left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} + c\left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}$$

$$= \frac{1 + \left(\frac{\lambda}{\lambda+1}\right)^{\alpha} + \left(\frac{\lambda}{\lambda+2}\right)^{\alpha} + \cdots + \left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}$$

$$= \frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} \tag{4.1.5}$$

$$E[X^2 \mid x \leq c] = \sum_{x=1}^{c} \frac{x^2\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - x^2\left(\frac{\lambda}{\lambda + x}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} \tag{4.1.6}$$

$$= \frac{1 - \left(\frac{\lambda}{\lambda+1}\right)^{\alpha} + 4\left(\frac{\lambda}{\lambda+1}\right)^{\alpha} - 4\left(\frac{\lambda}{\lambda+2}\right)^{\alpha} + 9\left(\frac{\lambda}{\lambda+2}\right)^{\alpha} - 9\left(\frac{\lambda}{\lambda+3}\right)^{\alpha} + \cdots - (c-1)^2\left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} + c^2\left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}$$

$$= \frac{1 + 3\left(\frac{\lambda}{\lambda+1}\right)^{\alpha} + 5\left(\frac{\lambda}{\lambda+2}\right)^{\alpha} + \cdots + (2c-1)\left(\frac{\lambda}{\lambda+c-1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}$$

$$= \frac{\sum_{x=1}^{c} (2x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}} \qquad \textbf{(4.1.7)}$$

Therefore,

$$\mathrm{VAR}[X|x \le c] = \frac{\sum_{x=1}^{c} (2x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}} - \left(\frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}}\right)^2 . \qquad \textbf{(4.1.8)}$$

4.2. Method of Moments

The method of moments for two parameter distributions is a little more complicated than for one parameter distributions. In the two parameter situation, we need to set the sample mean equal to the mean of the truncated distribution, and the sample variance equal to the variance of the truncated distribution.

$$\bar{x} = E[X \mid x \le c]$$
$$s^2 = VAR[X^2 \mid x \le c]$$

$$\bar{x} = \frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} \qquad \textbf{(4.2.1)}$$

$$s^2 = \frac{\sum_{x=1}^{c} (2x-1)\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} - \left(\frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}\right)^2 \qquad \textbf{(4.2.2)}$$

$$\frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} - \bar{x} = 0 \qquad \textbf{(4.2.3)}$$

$$\frac{\sum_{x=1}^{c} (2x-1)\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} - \left(\frac{\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}\right)^2 - s^2 = 0 \qquad \textbf{(4.2.4)}$$

Solving for α and λ is made a little easier by multiplying equation (4.2.3) by the probability of a claim being reported before time c, and multiplying equation (4.2.4) by the square of the probability of a claim being reported by time c.

$$\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha} - \bar{x}\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right) = 0 \qquad \textbf{(4.2.5)}$$

$$\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right)\cdot\left(\sum_{x=1}^{c} (2x-1)\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c^2\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right) - \left(\sum_{x=1}^{c} \left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - c\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right)^2 - s^2\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right)^2 = 0 \qquad \textbf{(4.2.6)}$$

Because this system of equations is so complex, we must turn to a numerical method to solve for our parameters. Simply applying the Newton-Raphson method will not work in this case; we have two parameters for which to solve this system of equations. One numerical method which we may use is the multi-dimensional Newton method [1]. The iterative formula for this method is as follows:

$$\begin{bmatrix} \hat{\alpha}_{m+1} \\ \hat{\lambda}_{m+1} \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_{m} \\ \hat{\lambda}_{m} \end{bmatrix} - J\left(\hat{\alpha}_m, \hat{\lambda}_m\right)^{-1} \mathbf{G}\left(\hat{\alpha}_m, \hat{\lambda}_m\right), \tag{4.2.7}$$

where

$$J(\alpha, \lambda) = \begin{bmatrix} \dfrac{\partial g_1(\alpha,\lambda)}{\partial \alpha} & \dfrac{\partial g_1(\alpha,\lambda)}{\partial \lambda} \\ \dfrac{\partial g_2(\alpha,\lambda)}{\partial \alpha} & \dfrac{\partial g_2(\alpha,\lambda)}{\partial \lambda} \end{bmatrix}, \tag{4.2.8}$$

the Jacobian, and

$$\mathbf{G}(\alpha, \lambda) = \begin{bmatrix} g_1(\alpha,\lambda) \\ g_2(\alpha,\lambda) \end{bmatrix}. \tag{4.2.9}$$

We determine

$$J(\alpha,\lambda)^{-1} = \left(\frac{1}{\dfrac{\partial g_1(\alpha,\lambda)}{\partial \alpha}\dfrac{\partial g_2(\alpha,\lambda)}{\partial \lambda} - \dfrac{\partial g_1(\alpha,\lambda)}{\partial \lambda}\dfrac{\partial g_2(\alpha,\lambda)}{\partial \alpha}} \right) \cdot \begin{bmatrix} \dfrac{\partial g_2(\alpha,\lambda)}{\partial \lambda} & \dfrac{-\partial g_1(\alpha,\lambda)}{\partial \lambda} \\ \dfrac{-\partial g_2(\alpha,\lambda)}{\partial \alpha} & \dfrac{\partial g_1(\alpha,\lambda)}{\partial \alpha} \end{bmatrix}. \tag{4.2.10}$$

Thus,

$$\begin{bmatrix}\hat{\alpha}_{m+1}\\ \hat{\lambda}_{m+1}\end{bmatrix} = \begin{bmatrix}\hat{\alpha}_{m}\\ \hat{\lambda}_{m}\end{bmatrix} - J(\hat{\alpha}_m, \hat{\lambda}_m)^{-1}\mathbf{G}(\hat{\alpha}_m, \hat{\lambda}_m)$$

$$= \begin{bmatrix}\hat{\alpha}_{m}\\ \hat{\lambda}_{m}\end{bmatrix} - \left(\frac{1}{\frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda} - \frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}}\right)\cdot\begin{bmatrix}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda} & \frac{-\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda}\\ \frac{-\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha} & \frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}\end{bmatrix}\cdot\begin{bmatrix}g_1(\hat{\alpha}_m, \hat{\lambda}_m)\\ g_2(\hat{\alpha}_m, \hat{\lambda}_m)\end{bmatrix} \quad \textbf{(4.2.11)}$$

$$\hat{\alpha}_{m+1} = \hat{\alpha}_m - \frac{g_1(\hat{\alpha}_m, \hat{\lambda}_m)\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda} - g_2(\hat{\alpha}_m, \hat{\lambda}_m)\frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda}}{\frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda} - \frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}} \quad \textbf{(4.2.12)}$$

$$\hat{\lambda}_{m+1} = \hat{\lambda}_m - \frac{g_2(\hat{\alpha}_m, \hat{\lambda}_m)\frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha} - g_1(\hat{\alpha}_m, \hat{\lambda}_m)\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}}{\frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda} - \frac{\partial g_1(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\lambda}\frac{\partial g_2(\hat{\alpha}_m, \hat{\lambda}_m)}{\partial\alpha}}. \quad \textbf{(4.2.13)}$$

We will now apply this to the method of moments. Setting $g_1(\alpha, \lambda)$ equal to equation (4.2.5) and $g_2(\alpha, \lambda)$ equal to equation (4.2.6), we determine our partial derivatives.

$$\frac{\partial}{\partial\alpha}g_1(\alpha,\lambda) = \sum_{x=1}^{c}\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right) - (c-\bar{x})\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right) \quad \textbf{(4.2.14)}$$

$$\frac{\partial}{\partial\lambda}g_1(\alpha,\lambda) = \alpha\,\lambda^{-2}\left(\sum_{x=1}^{c}(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1} - c(c-\bar{x})\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right) \quad \textbf{(4.2.15)}$$

$$\frac{\partial}{\partial\alpha}g_2(\alpha,\lambda) = +\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)\left(\sum_{x=1}^{c}(2x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right)-(c^2-2s^2)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\right)$$

$$-2\left(\sum_{x=1}^{c}\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}-c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)\left(\sum_{x=1}^{c}\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right)-c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\right) \quad \textbf{(4.2.16)}$$

$$-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\left(\sum_{x=1}^{c}(2x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}-c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)$$

$$\frac{\partial}{\partial\lambda}g_2(\alpha,\lambda) = \alpha\lambda^{-2}\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)\left(\sum_{x=1}^{c}(2x-1)(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1}-(c^3-2s^2c)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right)$$

$$+\alpha\lambda^{-2}c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\left(\sum_{x=1}^{c}(2x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}-c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right) \quad \textbf{(4.2.17)}$$

$$-2\alpha\lambda^{-2}\left(\sum_{x=1}^{c}(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1}-c^2\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right)\left(\sum_{x=1}^{c}\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}-c\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)$$

Before we apply these formulas, we need to determine a sufficiently close initial approximation for each of our parameters to increase our chance of convergence. To do this we employ the steepest descent method [1].

The steepest descent method is based on the idea of finding the parameters of the function $h(\alpha,\lambda) = [g_1(\alpha,\lambda)]^2 + [g_2(\alpha,\lambda)]^2$ where $h(\alpha,\lambda)$ equals zero, the minimum of the function. We start with an initial approximation for both α and λ. From here we determine the path of steepest descent down $h(\alpha,\lambda)$, which after a few iterations will lead to a much better initial approximation to use in our two-dimensional Newton method. The algorithm for the steepest descent method is explained in *Table 4.1*.

Steepest Descent Algorithm

1) Input initial approximations to α and λ.
2) While a possible solution has not been found do:
   a) Set:
      $h_1 = h(\alpha,\lambda)$
      $z = \nabla h(\alpha,\lambda)$ [The gradient of $h(\alpha,\lambda)$]
      $z_0 = \|z\|_2$
   b) If $z_0 = 0$, then STOP. [Minimum may have been determined.]
   c) Set:
      $z = z / z_0$
      $p_1 = 0$
      $p_3 = 1$
      $h_3 = h(\alpha - p_3 z_1, \ \lambda - p_3 z_2)$
   d) While $h_3 \geq h_1$ do:
      i) Set:
         $p_3 = p_3 / 2$
         $h_3 = h(\alpha - p_3 z_1, \ \lambda - p_3 z_2)$
      ii) If p_3 < Tolerance / 2, then STOP. [Minimum may have been determined -- Improvement in estimation not expected]
   e) Set:
      $p_2 = p_3 / 2$
      $h_2 = h(\alpha - p_2 z_1, \ \lambda - p_2 z_2)$
      $j_1 = (h_2 - h_1) / p_2$
      $j_2 = (h_3 - h_2) / (p_3 - p_2)$
      $j_3 = (j_2 - j_1) / p_3$
      $p_0 = (p_2 - j_1) / (2 j_3)$
      $h_0 = h(\alpha - p_0 z_1, \ \lambda - p_0 z_2)$
   f) If $h_0 < h_3$, then set:
      $p' = p_0$
      $h' = h_0$
      Else set:
      $p' = p_3$
      $h' = h_3$
   g) Set:
      $\alpha = \alpha - p' z_1$
      $\lambda = \lambda - p' z_2$
   h) If $| h' - h_1 |$ < Tolerance, then Stop. [Minimum has been found.]

Table 4.1

If the steepest descent algorithm and multi-dimensional Newton's method are applied, one will quickly notice that these have an extremely slow rate of convergence. The multi-dimensional Newton's method may take several million iterations to converge! This is due to the great complexity of our equations. One way of solving this problem involves summing the squares of our two functions, $g_1(\alpha, \lambda)$ and $g_2(\alpha, \lambda)$, and minimizing this new function, just like in the steepest descent method. Let us call this new function $h(\alpha, \lambda)$. Because both $g_1(\alpha, \lambda)$ and $g_2(\alpha, \lambda)$ intersect when they are both equal to zero, the square of each of these functions and, most importantly, our function $h(\alpha, \lambda)$ have a minimum of zero. Therefore,

$$\frac{\partial}{\partial \alpha} h(\alpha, \lambda) = \frac{\partial}{\partial \lambda} h(\alpha, \lambda) = 0 \qquad \textbf{(4.2.18)}$$

at this minimum. If we hold λ constant and perform the bisection method on

$$\frac{\partial}{\partial \lambda} h(\alpha, \lambda) = 2\left(g_1(\alpha, \lambda) \frac{\partial}{\partial \lambda} g_1(\alpha, \lambda) + g_2(\alpha, \lambda) \frac{\partial}{\partial \lambda} g_2(\alpha, \lambda) \right), \qquad \textbf{(4.2.19)}$$

we will find the value for λ where

$$\frac{\partial}{\partial \lambda} h(\alpha, \lambda) = 0.$$

Most likely, however,

$$\frac{\partial}{\partial \alpha} h(\alpha, \lambda) \neq 0$$

at our current estimates of α and λ, so we hold λ constant and perform the bisection method on

$$\frac{\partial}{\partial \alpha} h(\alpha, \lambda) = 2\left(g_1(\alpha, \lambda) \frac{\partial}{\partial \alpha} g_1(\alpha, \lambda) + g_2(\alpha, \lambda) \frac{\partial}{\partial \alpha} g_2(\alpha, \lambda) \right). \tag{4.2.20}$$

Now,

$$\frac{\partial}{\partial \alpha} h(\alpha, \lambda) = 0 \text{ and } \frac{\partial}{\partial \lambda} h(\alpha, \lambda) \neq 0,$$

but $h(\alpha, \lambda)$ has a value closer to zero. If we continue alternating which parameter is held constant and for which parameter the respective partial derivative is solved, estimates of α and λ of the desired degree of accuracy can quickly be found.

What we are basically doing is using an initial $\hat{\alpha}$ to minimize $h(\alpha, \lambda)\,|_{\alpha = \hat{\alpha}}$ and find a new $\hat{\lambda}$, using this $\hat{\lambda}$ to minimize $h(\alpha, \lambda)\,|_{\lambda = \hat{\lambda}}$ and find a new $\hat{\alpha}$, etc., until we have found the estimates of α and λ where $h(\alpha, \lambda)\,|_{\alpha = \hat{\alpha},\, \lambda = \hat{\lambda}}$ is sufficiently close to zero. Note that the Newton-Raphson method may be used in place of or in conjunction with the bisection method, but this will slow the process greatly, due to the complexity of our equations.

As a demonstration of the use of the method of moments, we will create the following example: A storm occurs, and within four weeks an insurance company has received 972 claims according to the schedule in *Table 4.2.* The sample mean is 1.557613, and the sample variance is 0.69595641. Using the method just described and the equations developed for the method of moments we determine $\hat{\alpha} = 9.680182$ and $\hat{\lambda} = 10.075520$. Thus, our IBNR claims estimate is 40 claims. After the fourth week following the storm, 28 more claims are reported. This shows our IBNR estimate is twelve claims too high. This does not mean that this method is a failure;

If we look at the estimated claims for each week in *Table 4.2* we find that our estimates are very close to the observed data. Data for this example were created using a random number generator for the Pareto distribution with parameters $\alpha = \lambda = 8$ and a total number of claims of 1000. Using these parameters we would expect 39 claims to be reported after the fourth week. If we ignore the set total number of claims of 1000 and use the 972 claims reported before the end of the fourth week we would expect 40 claims to be reported after the fourth week. In either case our estimated IBNR is extremely close to what we would expect with our set parameters, the former case off by one claim and the latter case exact. We can conclude that the large difference between our estimated and true IBNR claims is due to chance. We expect such events to occur occasionally.

Observed and Estimated Claims

Week	Observed	Estimated
1	606	607
2	232	229
3	92	94
4	42	41

(NOTE: one estimated claim was lost due to rounding error. Total estimated claims should equal observed claims.)

Table 4.2

4.3. Maximum Likelihood Estimation

For two parameter maximum likelihood estimation we create a likelihood function just as we did for the discrete approximation of the exponential distribution.

$$L(\alpha,\lambda) = \prod_{i=1}^{n} f(x_i \mid x_i \le c)$$

$$= \frac{\prod_{i=1}^{n}\left(\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha}\right)}{\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right)^{n}} \qquad \textbf{(4.3.1)}$$

To simplify our calculations we take the natural log of both sides of the equation.

$$\ln L(\alpha,\lambda) = \sum_{i=1}^{n} \ln\left(\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha}\right) - n\ln\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right) \qquad \textbf{(4.3.2)}$$

The derivatives are as follows:

$$\frac{\partial}{\partial\alpha}\ln L(\alpha,\lambda) = \sum_{i=1}^{n} \frac{\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda + x_i - 1}\right) - \left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda + x_i}\right)}{\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha}} + \frac{n\left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda + c}\right)}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} = 0 \qquad \textbf{(4.3.3)}$$

$$\frac{\partial}{\partial\lambda}\ln L(\alpha,\lambda) = \alpha\lambda^{-2}\left(\sum_{i=1}^{n} \frac{(x_i - 1)\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha+1} - x_i\left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha+1}}{\left(\frac{\lambda}{\lambda + x_i - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x_i}\right)^{\alpha}} + \frac{nc\left(\frac{\lambda}{\lambda + c}\right)^{\alpha+1}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}}\right) = 0 \qquad \textbf{(4.3.4)}$$

In the event that Newton's method is chosen as the numerical method used to solve for our parameters, we set $g_1(\alpha, \lambda)$ equal to equation (4.3.3) and $g_2(\alpha, \lambda)$ equal to equation (4.3.4) and take the following necessary partial derivatives:

$$\frac{\partial}{\partial\alpha}g_1(\alpha,\lambda) = -\sum_{i=1}^{n}\frac{\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}\left(\ln\left(\frac{\lambda}{\lambda+x_i-1}\right)-\ln\left(\frac{\lambda}{\lambda+x_i}\right)\right)^{2}}{\left(\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}\right)^{2}} + \frac{n\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\left(\ln\left(\frac{\lambda}{\lambda+c}\right)\right)^{2}}{\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)^{2}} \tag{4.3.5}$$

$$\begin{aligned}\frac{\partial}{\partial\alpha}g_2(\alpha,\lambda) = \frac{\partial}{\partial\lambda}g_1(\alpha,\lambda) = {} & \sum_{i=1}^{n}\frac{\lambda^{-2}\left((x_i-1)\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha+1}\left(1+\alpha\ln\left(\frac{\lambda}{\lambda+x_i-1}\right)\right)-x_i\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha+1}\left(1+\alpha\ln\left(\frac{\lambda}{\lambda+x_i}\right)\right)\right)}{\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}} \\ & -\sum_{i=1}^{n}\frac{\alpha\lambda^{-2}\left((x_i-1)\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha+1}-x_i\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha+1}\right)\cdot\left(\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x_i-1}\right)-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x_i}\right)\right)}{\left(\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}\right)^{2}} \\ & +\frac{nc\lambda^{-2}\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}+\alpha\ln\left(\frac{\lambda}{\lambda+c}\right)\right)}{\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)^{2}}\end{aligned} \tag{4.3.6}$$

$$\begin{aligned}\frac{\partial}{\partial\lambda}g_2(\alpha,\lambda) = {} & \alpha\sum_{i=1}^{n}\frac{\lambda^{-4}(\alpha-1)\left((x_i-1)^2\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha+2}-x_i^2\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha+2}\right)-2\lambda^{-3}\left((x_i-1)\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha+2}-x_i\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha+2}\right)}{\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}} \\ & -\alpha^2\lambda\sum_{i=1}^{n}\left(\frac{(x_i-1)\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha+1}-x_i\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha+1}}{\left(\frac{\lambda}{\lambda+x_i-1}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+x_i}\right)^{\alpha}}\right)^{2} \\ & +\frac{nc\alpha\lambda^{-4}\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+2}\left((c+2\lambda)\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)-c\alpha\right)}{\left(1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right)^{2}}\end{aligned} \tag{4.3.7}$$

If the method discussed earlier involving the bisection method is chosen as the numerical method used to solve for our parameters, we define the function $h(\alpha, \lambda)$ to be $ln\ L(\alpha, \lambda)$. In this case we are searching for the maximum of $h(\alpha, \lambda)$, rather than the minimum, but the process is not changed. (NOTE:

$$\frac{\partial}{\partial\alpha}h(\alpha,\lambda) \text{ is no longer } 2\left(g_1(\alpha,\lambda)\frac{\partial}{\partial\alpha}g_1(\alpha,\lambda) + g_2(\alpha,\lambda)\frac{\partial}{\partial\alpha}g_2(\alpha,\lambda)\right), \text{ and}$$

$$\frac{\partial}{\partial\lambda}h(\alpha,\lambda) \text{ is no longer } 2\left(g_1(\alpha,\lambda)\frac{\partial}{\partial\lambda}g_1(\alpha,\lambda) + g_2(\alpha,\lambda)\frac{\partial}{\partial\lambda}g_2(\alpha,\lambda)\right), \text{ but}$$

$$\frac{\partial}{\partial\alpha}h(\alpha,\lambda) = \frac{\partial}{\partial\alpha}\ln L(\alpha,\lambda) \text{ and } \frac{\partial}{\partial\lambda}h(\alpha,\lambda) = \frac{\partial}{\partial\lambda}\ln L(\alpha,\lambda)$$

in this case.)

Using the same data as in our example for the method of moments and the equations we developed from the maximum likelihood method, we find $\hat{\alpha} = 10.337055$ and $\hat{\lambda} = 10.784898$. The estimated claims using these parameters look identical to the estimated claims calculated using the parameters determined by the method of moments, with the exception that there are 230 claims in week 2 and 39 IBNR claims as opposed to 229 claims in week 2 and 40 IBNR claims.

4.4. Least Squares Method

As with the least squares method for the discrete approximation to the exponential distribution, we need to sum the squares of the differences between the actual probability of a claim being reported in each time period and the observed probability.

$$SS = \sum_{x=1}^{c} \left(\frac{\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}}{1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}} - f^{o}(x) \right)^{2} \tag{4.4.1}$$

where $f^{o}(x)$ is the observed probablity. Again we simplify the process by maximizing

$$\left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right)^{2} SS = \sum_{x=1}^{c} \left(\left(\frac{\lambda}{\lambda + x - 1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha} - \left(1 - \left(\frac{\lambda}{\lambda + c}\right)^{\alpha}\right) f^{o}(x) \right)^{2}. \tag{4.4.2}$$

Taking our partial derivatives with respect to α and λ, respectively,

$$g_1(\alpha,\lambda) = \frac{\partial}{\partial \alpha}\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]^{2} SS/2 = \sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha} - f^{o}(x)\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]\right] \cdot \left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right) - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x}\right) + f^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\right] \tag{4.4.3}$$

$$g_2(\alpha,\lambda) = \frac{\partial}{\partial \lambda}\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]^{2} SS/2 = \alpha\lambda^{-2}\sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha} - f^{o}(x)\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]\right] \cdot \left[(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1} - x\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+1} + cf^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right] \tag{4.4.4}$$

Again, to employ the two-dimension Newton method, we take the partial derivatives with respect to α and λ for each of these functions.

$$\begin{aligned} \frac{\partial}{\partial \alpha} g_1(\alpha,\lambda) = & \sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right) - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x}\right) + f^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\right]^{2} \\ & + \sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha} - f^{o}(x)\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]\right] \cdot \left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\left(\ln\left(\frac{\lambda}{\lambda+x-1}\right)\right)^{2} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}\left(\ln\left(\frac{\lambda}{\lambda+x}\right)\right)^{2} + f^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\left(\ln\left(\frac{\lambda}{\lambda+c}\right)\right)^{2}\right] \end{aligned} \tag{4.4.5}$$

$$\begin{aligned} \frac{\partial}{\partial \lambda} g_1(\alpha,\lambda) = \frac{\partial}{\partial \alpha} g(\alpha,\lambda) = & \ \alpha\lambda^{-2}\sum_{x=1}^{c}\left[\begin{array}{l}\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x-1}\right) - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+x}\right) \\ + f^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\ln\left(\frac{\lambda}{\lambda+c}\right)\end{array}\right] \cdot \left[(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1} - x\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+1} + cf^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right] \\ & + \sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha} - f^{o}(x)\left[1 - \left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]\right] \cdot \left[\begin{array}{l}\alpha(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1}\ln\left(\frac{\lambda}{\lambda+x-1}\right) + (x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+2} \\ - \alpha x\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+1}\ln\left(\frac{\lambda}{\lambda+x}\right) - x\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+2} \\ + \alpha cf^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\ln\left(\frac{\lambda}{\lambda+c}\right) + cf^{o}(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\end{array}\right] \end{aligned} \tag{4.4.6}$$

$$\frac{\partial}{\partial\lambda} g_2(\alpha,\lambda) = \alpha^2\lambda^{-4}\sum_{x=1}^{c}\left[(x-1)\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+1} - x\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+1} + cP(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+1}\right]^2$$
$$+ \alpha(\alpha+1)\lambda^{-4}\sum_{x=1}^{c}\left[\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha} - P(x)\left[1-\left(\frac{\lambda}{\lambda+c}\right)^{\alpha}\right]\right]\cdot\left[(x-1)^2\left(\frac{\lambda}{\lambda+x-1}\right)^{\alpha+2} - x^2\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+2} + c^2P(x)\left(\frac{\lambda}{\lambda+c}\right)^{\alpha+2}\right] \quad (4.4.7)$$

Again, using the data from our prior examples, we estimate α to be approximately 12.760960 and λ to be approximately 13.375895. Thus our IBNR estimate is 36 claims -- a little closer to our true IBNR. One will note while using the least squares approach that convergence is extremely slow. The IBNR claims estimate will not, most likely, be significantly better using this method than if we use the method of moments or the maximum likelihood method. Therefore, the estimate may not be worth the time involved in calculating it.

5. Testing Goodness of Fit

5.1. Chi-Square Goodness of Fit Test

In order to determine whether or not the estimated distribution is accurate enough, we use a common goodness of fit test called the chi-square test [3]. This is a very simple statistical test that involves summing the quotients of the squared differences between the reported number of claims and the estimated number of claims in each time period divided by the estimated number of claims in that time period, and comparing it to a value obtained from a chi-square table.

$$X^2 = \sum_{i=1}^{c}\frac{(E_i - n_i)^2}{E_i} \sim \chi^2(c-1-r), \quad (5.1.1)$$

where E_i is the expected number of claims reported in period i, n_i is the actual number of claims reported in period i, r is the number of estimated parameters, and

c - 1 - r is the number of degrees of freedom of the chi-square distribution. If $X^2 \leq \chi^2(c - 1 - r)$ of the desired significance level, then we do not reject the null hypothesis that our estimated claim distribution is equal to the actual claim distribution. Otherwise, we reject this null hypothesis.

An example of this process is shown in *Table 5.1*, where our hypothesized distribution is rejected. Looking at a graph of the reported and estimated claims (*Figure 5.1*) it is not completely apparent that we reject the estimated distribution shown in *Table 5.1*. However, taking into consideration the number of claims with which we are dealing, the differences between our estimated and reported claims may be great enough to cause our estimated distribution to be rejected, even though the percent differences are extremely small, if we are looking for accuracy to the nearest claim. If we are not interested in individual claims, but thousands of claims, an estimation based on the exponential distribution for our current example is not rejected at a 5% significance level as shown in *Table 5.2*.

Chi-Square Test

Period	Reported Claims	Estimated* Claims	Contribution to X^2 Statistic
1	45,171	44,664	5.75517
2	24,492	24,963	8.88679
3	14,017	13,952	0.30282
4	7,622	7,798	3.97230
5	3,865	4,358	55.77077
6	3,004	2,436	132.44007
			207.12792

*Estimation of claims calculated with exponential distribution then rounded.

$X^2 = 207.12792$
$\chi^2_{.05}(4) = 9.488$
$\therefore X^2 > \chi^2_{.05}(4)$, and we reject the null hypothesis that our estimated distribution of claims equals the true distribution of claims.

Table 5.1

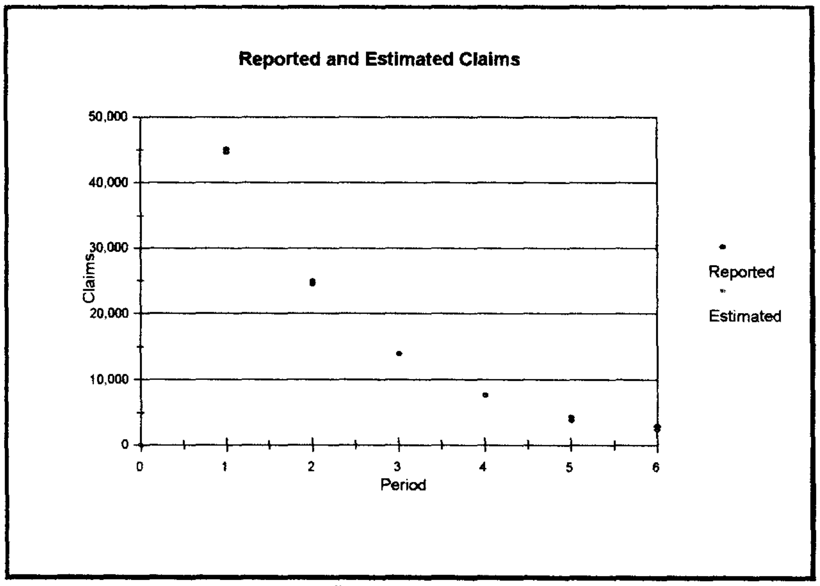

Figure 5.1

Chi-Square Test (claims in terms of thousands)

Period	Reported Claims	Estimated* Claims	Contribution to X^2 Statistic
1	44	44	0.02273
2	24	25	0.40000
3	14	14	0.00000
4	8	8	0.00000
5	4	4	0.00000
6	3	3	0.00000
			0.06273

*Estimation of claims calculated with exponential distribution then rounded.

$X^2 = 0.06273$
$\chi^2_{.05}(4) = 9.488$
$\therefore X^2 < \chi^2_{.05}(4)$, so we do not reject the null hypothesis that our estimated distribution of claims equals the true distribution of claims.

Table 5.2

5.2. Kolmogorov-Smirnov Test

Another useful goodness of fit test is called the Kolmogorov-Smirnov test [3]. This test involves calculating the absolute difference between the empirical distribution and the estimated cumulative distribution at each point and comparing the greatest of these values with the proper Kolmogorov-Smirnov acceptance limit (taken from a table of Kolmogorov-Smirnov acceptance limits).

$$D_c = sup_x [\,|F_c(x) - F_0(x)|\,], \tag{5.2.1}$$

where c is the maximum possible lag, x is the period, $F_c(x)$ is the empirical distribution, and $F_0(x)$ is the estimated cumulative distribution. We do not reject the estimated distribution if $D_c \leq d$, the value of the proper acceptance limit.

Kolmogorov-Smirnov Test

Period	Empirical Distribution	Estimated Distribution	Absolute Difference
1	0.460126	0.454964	0.005161
2	0.709609	0.709245	0.000364
3	0.852390	0.851364	0.001027
4	0.930030	0.930794	0.000764
5	0.969400	0.975188	**0.005788**
6	1.000000	1.000000	0.000000

Estimation of claims distribution calculated with exponential distribution.

$D_6 = 0.005788$
$d = 0.52$
$\therefore D_6 < d$, and we do not reject the null hypothesis that our estimated distribution of claims equals the true distribution of claims at the 5% significance level.

Table 5.3

Using our data from our original example for the chi-square goodness of fit test, we find that we do not reject our estimated distribution (*Table 5.3*). We are well within the bounds of the Kolmogorov-Smirnov acceptance limit, whereas we were well above the bounds of the chi-square test. This shows that the Kolmogorov-Smirnov test may not be the best test of fit when we are looking for a high degree of accuracy. If we are not looking for a high degree of accuracy, as when we tested accuracy to the thousands of claims in the chi-square test, then the Kolmogorov-Smirnov test is quite suitable.

6. Sample of Actual Data

The following is an example of parametric estimation of storm IBNR for a storm that occurred in Chicago. Claims were reported according to *Table 6.1*. We see that our sample mean is 1.574746 and our sample variance is 1.29709387.

Chicago Storm

Week	Claims
1	491
2	103
3	39
4	23
5	19
6	14

Table 6.1

Claims estimated by the exponential distribution are shown in *Table 6.2*. The method of moments and the maximum likelihood method both estimate θ to be approximately 0.99083335. The least squares method estimates θ to be approximately 1.306591. Thus, our IBNR estimates are 1.809162 and 0.271468, respectively. We see that both of these estimates are rejected.

When we attempt to estimate our IBNR claims using the Pareto distribution we find that our formulas will not converge for any of the three estimation methods. Because of this lack of convergence and the fact that the exponential distribution does not fit the observed data, one may assume that these equations are useless. This may be true for this particular case, but these formulas may be quite useful on other storm cases. Also, there are many other distributions that the data may fit. Using

Chicago Storm
Estimated Claims

	Meth. Mom./Max. Lik.		Least Squares	
Week	Claims	Cont. to X^2	Claims	Cont. to X^2
1	434.3344	7.392898	502.6571	0.270341
2	161.2541	21.044677	136.0901	8.045807
3	59.6684	7.274097	36.8452	0.126015
4	22.2272	0.026872	9.9755	17.005304
5	8.2522	13.998056	2.7008	98.365467
6	3.0638	39.037130	0.7312	240.778158
		88.773703		364.591094

88.773703 > 9.488 and 364.591094 > 9.488, so we **reject both** of our estimated distributions.

Table 6.2

the methods described in this thesis one can determine the formulas for parametric estimation of storm IBNR using these distributions. It only takes time, patience, and a handy computer.

7. Bibliography

[1]. Burden, Richard L., and Faires, J. Douglas, *Numerical Analysis*, Fifth Edition, PWS-KENT Publishing Company, Boston, 1993, pp. 56-65, 553-557, 568-573.

[2]. Hogg, Robert V., and Klugman, Stuart A., *Loss Distributions*, Wiley, New York, 1984, pp.222-223.

[3]. Hogg, Robert V., and Tanis, Elliot A., *Probability and Statistical Inference*, Third Edition, Macmillan Publishing Company, New York, 1988, pp. 224, 330, 334-339, 514-520, 591-596.

[4]. Weissner, Edward W., "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood," *Proceedings*, Volume LXV, Part 1, No. 123, Casualty Actuarial Society, 1978, pp. 1-9.