# POISSON FORECASTING MODELS: AN INTRODUCTION

Jambu Nathan, University of Iowa


## Abstract

This paper presents the use of the Poisson probability distribution for forecasting discrete time series. The properties of the weights using the Poisson distribution are discussed. These weights provide alternatives, not attainable by exponential smoothing models. As an introduction to Poisson smoothing process, a constant and linear trend correction models are presented. For some of the time series tests, the Poisson forecasting models show slightly improved forecast accuracies compared to exponential forecasting models.


## Introduction

The literature on business forecasting models are numerous and varied in nature. The most widely used approaches in the area of forecasting can be classified into exponential, Box-Jenkins, and adaptive forecasting methods [1, 2]. Both exponential [3, 4] and Box-Jenkins approaches are well established as sound techniques for forecasting discrete time series. A detailed description of the model identification, model estimation, diagnostic checking, and forecasting can be found in the text written by Box and Jenkins. In this paper a comparison between exponential and Poisson models are presented, while the comparison between Poisson and Box-Jenkins models is in order for future research.

## The Constant Exponential and Poisson Forecasting Models

The basic single parameter exponential smoothing process is defined as follows:

$$\hat{Y}(t) = \alpha Y(t) + (1-\alpha)\hat{Y}(t-1) \tag{1}$$

where, $Y(t)$ = actual time series value for period t

$\hat{Y}(t)$ = smoothed time series value for period t (used as the forecast for period t + 1)

$\alpha$ = smoothing parameter $0 \leq \alpha \leq 1$.

The above model provides the following weighting scheme for t periods of historical time series data:

$$Y(t) = w_t Y(t) + w_{t-1}Y(t-1) + w_{t-2}Y(t-2) + \dots w_2 Y(2) + w_1 Y(1) \tag{2}$$

where

$$w_i = \alpha(1-\alpha)^{t-i} \qquad i = 1,2,\dots,t \tag{3}$$

and

$$\sum_{i=1}^{t} w_i = 1 - (1-\alpha)^t \tag{4}$$

where

$(1-\alpha)^t$ approaches zero as t gets large.

This paper introduces a Poisson smoothing model which, while maintaining most of the advantages of the Exponential Smoothing model, provides a new range of weighting scheme alternatives for the t periods of time series data. The Poisson Smoothing process is offered as an alternative to, rather than a replacement for, exponential smoothing.

The basic single parameter Poisson Smoothing process is defined as follows:

$$\hat{Y}(t) = e^{-\lambda}\left[Y(t) + \lambda Y(t-1) + \frac{\lambda^2}{2!}Y(t-2) + \dots + \frac{\lambda^{(t-1)}}{(t-1)!}Y(1)\right] \tag{5}$$

where $Y(t)$ and $\hat{Y}(t)$ are defined above.

Given a discrete time series of t periods, the Poisson Smoothing model defines the weights for each period as follows:

$$w_i = \frac{e^{-\lambda}\lambda^{(t-i)}}{(t-i)!} \qquad i = 1,2,\dots,t \tag{6}$$

where $\lambda$ = Smoothing parameter, $\lambda > 0$.

Since these weights are recognizable as the probabilities of the Poisson distribution [5], the forecasting model is termed a Poisson Smoothing process.

Note that in the above model, $e^{-\lambda}$ is the weight applied to the most recent time series value. The corresponding exponential smoothing model weight for the same time series value is $\alpha$, which in most practical forecasting problems is .05 or greater. If we restrict our Poisson Smoothing parameter $\lambda$ to cases where $e^{-\lambda}$ is also .05 or greater, it is found that the sum of the Poisson weights for the most recent time series observations rapidly approaches one. This has practical significance in that the older time series data carries essentially negligible weights and can be removed from the model without significantly affecting the smoothed time series value $\hat{Y}(t)$. Thus, while the Poisson weights of equation [6] can be defined for any t = 1,2, 3,..., a t of 10 accounts for at least .9997 of the total weighting of the past data. In fact, as t becomes large, it is only necessary to consider the weights for the most recent k periods ($k < t$).

## The Exponential Forecasting Model With Linear Trend Correction

Let a time series have a constant linear trend b. Applying the constant exponential smoothing model (equation [1]) to the time series provides the following smoothed value expression:

$$\hat{Y}(t) = \alpha V + \alpha(1-\alpha)(V-b) + \alpha(1-\alpha)^2(V-2b) + \dots + \alpha(1-\alpha)^n(V-nb) + \dots \tag{7}$$

where

$Y(t) = V$

$Y(t-1) = Y(t) - b = V - b$

$$Y(t-2) = Y(t-1) - b = V - 2b$$

etc.

Therefore,

$$\hat{Y}(t) = \alpha V[1 + (1-\alpha) + (1-\alpha)^2 + \ldots]$$

$$- \alpha(1-\alpha)b[1 + 2(1-\alpha) + 3(1-\alpha)^2 + \ldots$$

$$+ n(1-\alpha)^{n-1} + \ldots] \quad (8)$$

$$= \alpha V \frac{1}{\alpha} - \alpha(1-\alpha)b[\{1 + (1-\alpha) + (1-\alpha)^2 + \ldots\}$$

$$+ \{(1-\alpha) + (1-\alpha)^2 + \ldots\} + \ldots]$$

$$= \alpha V \frac{1}{\alpha} - \alpha(1-\alpha)b\left[\frac{1}{\alpha} + \frac{(1-\alpha)}{\alpha} + \frac{(1-\alpha)^2}{\alpha} + \ldots\right]$$

$$= \alpha V \frac{1}{\alpha} - \alpha(1-\alpha)b\left[\frac{1 + (1-\alpha) + (1-\alpha)^2 + \ldots}{\alpha}\right]$$

$$= V - \frac{(1-\alpha)}{\alpha} b \quad (9)$$

Hence the Constant Exponential model lags the linear trend or ramp series by $\frac{(1-\alpha)}{\alpha} b$. The exponential smoothed estimate of the linear trend can be obtained as follows:

$$\hat{b}(t) = \alpha b(t) + \alpha(1-\alpha)b(t-1) + \ldots$$

$$+ \alpha(1-\alpha)^n b(t-n) + \ldots \quad (10)$$

where

$$b(t) = \hat{Y}(t) - \hat{Y}(t-1) \quad (11)$$

Note that in equation [10] the trend estimate in period t is defined as the difference between the two most recent smoothed estimates of the series.

Therefore, the smoothed value of the time series including the linear trend correction is best given by F(t) where

$$F(t) = \hat{Y}(t) + \frac{(1-\alpha)}{\alpha} \hat{b}(t) \quad (12)$$

The forecast for $\ell$ periods in the future (assuming no seasonality) is given by:

$$F(t+\ell) = F(t) + \ell\hat{b}(t). \quad (13)$$

## The Poisson Forecasting Model With Linear Trend Correction

Suppose we have a time series with a constant linear trend b. Applying the constant complete Poisson Smoothing model (equation (5)) to the time series provides the following smoothed value expression:

$$Y(t) = e^{-\lambda}[V + \lambda(V-b) + \frac{\lambda^2}{2!}(V-2b) + \ldots$$

$$+ \frac{\lambda^n}{n!}(V-nb) + \ldots] \quad (14)$$

where

$$Y(t) = V$$

$$Y(t-1) = Y(t) - b = V - b$$

$$Y(t-2) = Y(t-1) - b = V - 2b$$

etc.

Therefore,

$$\hat{Y}(t) = e^{-\lambda}V[1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^n}{n!} + \ldots]$$

$$- e^{-\lambda}b[\lambda + \frac{2\lambda^2}{2!} + \frac{n\lambda^n}{n!} + \ldots]$$

$$= e^{-\lambda}\cdot V\cdot e^{\lambda} - e^{-\lambda}b\lambda[1 + \lambda + \ldots$$

$$+ \frac{\lambda^{n-1}}{(n-1)!} + \ldots] \quad (15)$$

Since $e^{\lambda} = [1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^{n-1}}{(n-1)!} + \frac{\lambda^n}{n!} + \ldots]$

Therefore,

$$\hat{Y}(t) = V - \lambda b \quad (16)$$

Hence the constant Poisson smoothing model lags the linear trend or ramp time series by approximately $\lambda b$. The Poisson smoothed estimate of the linear trend can be obtained as follows:

$$\hat{b}(t) = e^{-\lambda}[b(t) + \lambda b(t-1) + \frac{\lambda^2}{2!}b(t-2) + \ldots$$

$$+ \frac{\lambda^k}{k!}b(t-k)] \quad (17)$$

where

$$b(t) = \hat{Y}(t) - \hat{Y}(t-1). \quad (18)$$

Note that in equation (17) we are applying Poisson smoothing to our most recent k + 1 trend estimates where the trend estimate in period t is defined as the difference between the two most recent smoothed estimates of the series.

Therefore, the smoothed value of the time series including the linear trend correction is best given by F(t) where

$$F(t) = \hat{Y}(t) + \lambda\hat{b}(t). \quad (19)$$

The forecast for $\ell$ periods in the future (assuming no seasonality) is given by:

$$F(t+\ell) = F(t) + \ell\hat{b}(t) \quad (20)$$

## Poisson And Exponential Forecasting Models: A Comparison

The Poisson and exponential smoothing models are similar in that they are both single parameter models that provide weighting schemes for the previous time series data. A trial-and-error procedure for finding the best smoothing parameter $\lambda$ for the Poisson process would be very similar to the procedure used to find the best smoothing parameter $\alpha$ for the exponential smoothing model.

With regards to the data handling requirements, the exponential smoothing model is superior. The Poisson model may require as many as seven more data values than the exponential smoothing model. The data handling and storage disadvantages of the Poisson model may be compensated for by the fact that this model enables a new variety of weighting scheme alternatives for the past data. In particular, the exponential model is limited to cases where the more recent data always receives more weight. While the Poisson model can also exhibit this behavior, it offers some additional flexibility in that certain values of $\lambda$ will allow the older data to receive higher weights than the most recent data. For example, a $\lambda$ of 1.2 with $e^{-\lambda}$ equal to .3, provides a weight of .36 for data two periods old and a weight of .30 for the most recent observation. It seems reasonable and conceivable that the weighting scheme flexibility of the Poisson model might improve the overall forecasting accuracy for some specific time series. Examples of the three possible Poisson weighting schemes are shown in Figure 1.

The overall value of the Poisson model will however have to be measured in terms of forecasting accuracy. In other words, is there empirical evidence that in some cases the Poisson weighting scheme provides smaller forecasting errors than the exponential smoothing model? To obtain an empirical comparison of the two smoothing processes, we have selected in the first data set three fairly divergent time series: industrial production, stock prices, and unemployment rates. These series, which consist of monthly values over a ten-year period, are shown in Figure 2.

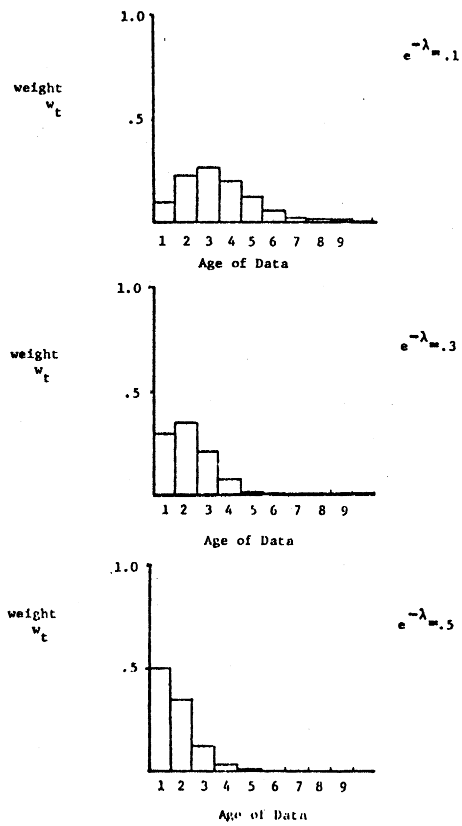


Figure 1. Poisson Weighting Schemes with $e^{-\lambda}$ Values of .1, .3, and .5.

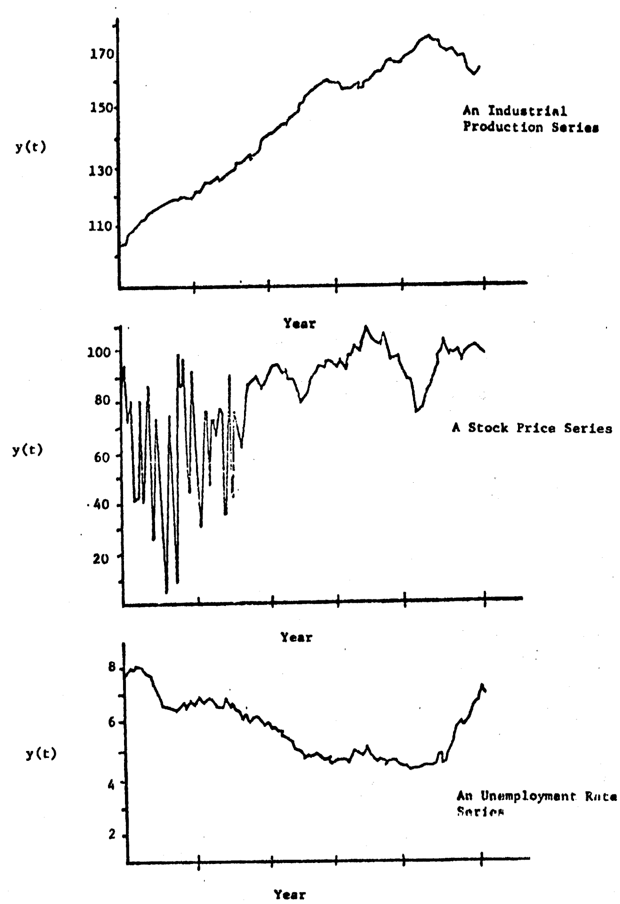


Figure 2. Time Series Used to Compare the Exponential and Poisson Smoothing Models

The forecasting accuracy, which is evaluated for each of three future periods, is measured in terms of the mean square error (MSE). The mean square error is defined as follows:

$$MSE_{\ell} = \frac{1}{T}\sum_{t=1}^{T} (y(t+\ell) - F(t+\ell))^2 \quad \text{for } \ell = 1,2, \text{ and } 3 \qquad (21)$$

where

T = total number of time series values.

For both the Poisson and exponential smoothing constant models, $F(t+\ell) = \hat{y}(t)$ for all $\ell$.

All time series were forecasted using the exponential constant and linear trend models and the Poisson constant and linear trend models. Each forecasting model was evaluated with nineteen different smoothing parameters. For the exponential smoothing models we used $\alpha$ = .05, .10, .15,..., .90, and .95. For the Poisson smoothing models we used values of $\lambda$ such that $e^{-\lambda}$ = .05, .10, .15, ..., .90, and .95. In this way, the weight applied to the most current observation y(t) is equivalent between the two models and therefore serves as a basis for comparing the two models.

The best exponential and Poisson forecasting models (i.e., the best $\alpha$ and $\lambda$) were identified for each time series. The α and λ values which yield the smallest MSE are defined to be the "best" α and λ values. The forecasting accuracies for constant exponential and Poisson models are summarized in Table 1.

TABLE 1
MEAN SQUARE ERRORS OF THE BEST FORECASTING MODELS FOR THREE TIME SERIES

| Time Series | Best Model | Mean Square Error |
| --- | --- | --- |
| Industrial Production | Exponential Constant ($\alpha = .60$) | 1.08 |
| | Poisson Constant ($e^{-\lambda} = .65$) | 1.09 |
| Stock Prices | Exponential Constant ($\alpha = .35$) | 181.69 |
| | Poisson Constant ($e^{-\lambda} = .10$) | 179.97 |
| Unemployment Rate | Exponential Constant ($\alpha = .35$) | .05 |
| | Poisson Constant ($e^{-\lambda} = .35$) | .05 |

The industrial production series favor the exponential constant model; the stock prices series shows the Poisson constant model providing the better forecasts; the unemployment rate series shows identical results for the forecasting models. Overall, based on this rather limited empirical comparison, the exponential and Poisson smoothing models provide quite similar forecasting results.

Table 2 summarizes the results for the exponential and Poisson linear trend models for the data set. The industrial production series favor the Poisson trend correction models for all the three time periods. Except for the first period, the stock prices series favor the Poisson trend correction model. In the case of unemployment rate series, both models indicate the same forecasting accuracies.

TABLE 2

MEAN SQUARE ERRORS FO THE BEST FORECASTING MODELS FOR THE THREE TIME SERIES

| Time Series | Best Model | Mean Square Error | | |
| --- | --- | --- | --- | --- |
| | | Period 1 | Period 2 | Period 3 |
| Industrial Production | Exponential Trend ($\alpha_1 = .55$, $\alpha_2 = .25$) | 1.06 | 3.15 | 5.77 |
| | Poisson Trend ($\lambda_1 = .36$, $\lambda_2 = .92$) | 1.03 | 2.91 | 5.25 |
| Stock Prices | Exponential Trend ($\alpha_1 = .25$, $\alpha_2 = 10$) | 196.48 | 221.35 | 208.29 |
| | Poisson Trend ($\lambda_1 = 1.39$, $\lambda_2 = 2.30$) | 202.04 | 208.29 | 198.24 |
| Unemployment Rate | Exponential Trend ($\alpha_1 = .55$, $\alpha_2 = .10$) | .03 | .04 | .05 |
| | Poisson Trend ($\lambda_1 = .60$, $\lambda_2 = 2.30$) | .03 | .04 | .05 |

## Conclusions

This article introduces the smoothing process based upon the Poisson probability distribution. The results of the Poisson forecasting models are compared with those of the exponential forecasting models. Even though the single parameter Poisson smoothing model requires additional data storage, it allows past data weighting scheme alternatives that may in some cases yield better forecasting accuracies than those observed by similar exponential smoothing models. However, the forecasts using this model cannot be updated recursively. This disadvantage is minimized due to the improved data handling and reduced storage costs of the available computers. Perhaps more important, the Poisson model provides an opportunity to use a weighting logic which does not monotonically decrease the weight applied to older observations, and this may have some real appeal to the practicing manager who attempts to forecast time series values.

## References


Box, George E. P. and Gwilym M. Jenkins, Time Series Analysis: Forecasting and Control, Revised Edition, San Francisco: Holden-Day, 1976.

Trigg, D. W. and A. G. Leach, "Exponential Smoothing with an Adaptive Response Rate," Operational Research Quarterly, 1967, 18:53-59.

Brown, R. G., Smoothing, Forecasting and Prediction of Discrete Time Series, Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

Brown, R. G., Decision Rules for Inventory Management, New York: Holt, Rinehart and Winston, 1967.

Haight, F. A., Handbook of the Poisson Distribution, New York: John Wiley and Sons., 1967.