TRANSFORMING THE POISSON DISTRIBUTION FOR SPC APPLICATIONS

Rudolf G. Kittlitz Jr.
Consultant, Rudy Kittlitz & Associates
2006 Ceredo Drive
Alpine Texas 79830-8501, U.S.A.
432 837 9937
rgke300@hughes.net

Abstract

A power transformation with an exponent of 2/3 for Poisson-distributed data achieves symmetry for improved statistical process control (SPC) applications whether it is for: an individuals chart, an exponentially weighted moving average chart, or a cumulative sum chart. The usual square-root transformation, used to stabilize the variance for analysis of variance problems, produces a negatively skewed distribution and tends to give false SPC signals. Two simple equations are proposed for calculating the lower control limit (LCL) and the upper control limit (UCL) for Poisson type data. Agreement between the exact LCL and UCL, as determined by the lower and upper tail area, is excellent. A set of published data will be used to illustrate the procedure for an individuals chart of the data.

Introduction

The fundamentals of the Poisson distribution are used as a basis for the c-chart primarily because of its simplicity in that the standard deviation can be calculated quickly from the mean. The usual lower and upper control limits, LCL and UCL, are:

\[
\text{LCL} = \bar{c} - 3\sqrt{\bar{c}}
\]

and

\[
\text{UCL} = \bar{c} + 3\sqrt{\bar{c}}
\]

A Lower Control Limit (LCL) is truncated to zero when the calculated LCL is negative and thus there is no LCL.

The Poisson distribution is a positively skewed distribution and these two simple equations ignore that. Therefore there will be times when limits calculated from these equations, for a c-chart, will either give false signals or will not signal. If a simple transformation could be obtained and achieve symmetry, then the possibility of the previously discussed problem would be eliminated or greatly diminished.

A variance stabilizing transformation, \(\sqrt{c + 0.375}\), does an excellent job of stabilizing the variance to 0.25 for a \(\bar{c}\) of 4 and greater. However it produces a negative skewness that defeats our objective. This negative skewness is approximately one-half of the original Poisson skewness.

Transforming The Poisson To Symmetry

It was shown by Haldane (1938)\(^7\) that the \(2/3\)-th-power would produce a symmetric transformed Poisson distribution. However there appears to be a scarcity of applications on this point. In Read and Cressie (1988, p. 96)\(^10\) a remark attributed to Anscombe\(^2\) implies the use of a small constant to be added to the Poisson number before performing the \(2/3\)-th-power transformation.

In completely separate private work, the author (Kittlitz, 2003)\(^8\) derived the following equations.

For symmetry,

\[
y_i = (c + 1/4)^{2/3} \tag{1}
\]

Expected mean on the transformed scale,
$$m_i = (\tau + 1/12)^{\phi_i/3}$$  \hspace{1cm} (2)

Expected standard deviation on the transformed scale,
$$s_i = (2/3)(\tau)^{\phi_i/6}$$  \hspace{1cm} (3)

**Expected Transformed Parameters**

The Poisson equation can be written as:
$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$  \hspace{1cm} (4)

Now with the proposed transformation (1) we can obtain the following:
$$x = y^{\phi/2} - 1/4$$  \hspace{1cm} (5)

As stated (Brownlee, 1965, p. 45) “We assume that $y = f(x)$ is a strictly monotonic function of $x$, and therefore $x = g(y)$ is a strictly monotonic function of $y$. In the discrete case things are simple. Let $p_x\{y\}$ be the probability function of $Y$. Whenever $X$ takes a particular value $x$, which it does with probability $p_x\{x\}$, so $p_x\{x\} = p_x\{g(y)\} = p_y\{y\}$”. Applying this to equations (4) and (5)
$$g(y) = \Pr(Y = y) = \Pr(X = y^{\phi/2} - 1/4) = \frac{\mu^{\phi/2} e^{-\mu}}{(y^{\phi/2} - 1/4)^{\phi/2}}$$  \hspace{1cm} (6)

To obtain the $E(Y)$, we have to sum the series as:
$$E(Y) = \sum_{y} y e^{-\mu} \left(\frac{\mu^{\phi/2} e^{-\mu}}{(y^{\phi/2} - 1/4)^{\phi/2}}\right)$$  \hspace{1cm} (7)

An example will illustrate. Let $\mu = 5$. What is $E(Y)$? Recall, when $x = 0, y = (0 + 1/4)^{\phi/2} = 0.39685…$; when $x = 1, y = (1 + 1/4)^{\phi/2} = 1.160397…$; etc. Thus, factoring out the $e^{-\mu}$ term and substituting for $y$ in equation (1) we have:
$$E(Y) = e^{-\mu} \sum_{y} \left(0 + 1/4\right)^{\phi/2} \frac{\mu^0}{0!} + \left(1 + 1/4\right)^{\phi/2} \frac{\mu^1}{1!} + \left(2 + 1/4\right)^{\phi/2} \frac{\mu^2}{2!} + K$$  \hspace{1cm} (8)

A Texas Instruments TI-83®, with internal accuracy to 14 digits and a two-digit exponent, was programmed to perform this summation for various numbers of terms. $E(Y) = 2.955286267$ for 25 terms. Using the Poisson distribution in Minitab®, for $\mu = 5$ and $n = 1,000,000$, the $E(Y)$ calculated to be 2.9552826 or in agreement.

To obtain the $E(Y^2)$ needs to be obtained. A similar series equation was developed and solved.
$$E(Y^2) = e^{-\mu} \sum_{y} \left(0 + 1/4\right)^{\phi/2} \frac{\mu^1}{1!} + \left(1 + 1/4\right)^{\phi/2} \frac{\mu^1}{1!} + \left(2 + 1/4\right)^{\phi/2} \frac{\mu^2}{2!} + K$$  \hspace{1cm} (9)

Using 25 terms again, $E(Y^2) = 9.495939968$. $\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = 9.495939968 - 2.955286267^2 = 0.7627230473$. This equates to a standard deviation of 0.8730538627, which compares well with 0.873046 obtained in the Minitab® calculations.

**Comparison of Various Transformations**

Previously I’ve mentioned the square-root transformation of Poisson data to stabilize the variance and there have been variations on it. Likewise I’ve just mentioned looking at the skewness as criteria, but the kurtosis is important also. In my experience I’ve found it very beneficial to examine also the standardized 5th and 6th moments in assessing various transformations.

The transformations to be considered are:
- No transformation
- $\sqrt{c}$
- $\sqrt{c + 0.375}$
For the Poisson, the theoretical skewness = $1/\sqrt{\mu}$ and the theoretical kurtosis = $3 + 1/\mu$. Likewise the theoretical standardized 5th and 6th moments are: $10/\sqrt{\mu} + 1/\mu^{1.5}$ and $15 + 25/\mu + 1/\mu^2$. For $\mu = 1$: skewness = 1, kurtosis = 4, standardized 5th = 11, and standardized 6th = 41. When $\mu = 25$, the respective values are: 0.2, 3.04, 2.008, and 16.0016. For a Gaussian distribution, the theoretical values are: 0, 3, 0, and 15, which will be the comparison values. Two graphs will be shown illustrating the skewness and kurtosis properties with the previously mentioned transformations.

Note in Figure 1 for skewness that the recommended transformation [left-pointed triangles] approaches 0 very quickly and is essentially 0 at a Poisson mean of ~3. The Haldane 2/3rd transformation [right-pointed triangles] is essentially 0 at a Poisson mean of ~6. The variance stabilizing transformation [angled square] becomes negative at a Poisson mean of ~1 and remains negative.

The calculated kurtosis values are shown in Figure 2 and once again the recommended transformation [left-pointed triangles] approaches 3 very quickly and is essentially 3 at a Poisson mean of ~4. The Haldane 2/3rd transformation [right-pointed triangles] is essentially 0 at a Poisson mean of ~8. The simple square-root transformation [squares] and the Tukey transformation [up-pointed triangles] have larger kurtosis values than the “No Transformation” [circles] for Poisson means up to ~15.

Not much is said these days about the standardized 5th and 6th moments, but my experience indicates that they should be examined also when trying to obtain a bell-shaped (i.e., Gaussian) distribution. For the Gaussian distribution, the standardized 5th is 0 and the standardized 6th is 15. Due to size limitations for this paper, these figures had to be left out. However they approached 0 and 15 at a Poisson mean of ~4.
The graphs shown, I believe, illustrate very well that the recommended transformation accomplishes the purpose of obtaining an essential bell-shaped distribution for SPC applications above a Poisson mean of ~4. When the Poisson mean is less than 4, I would recommend using the “time between Poisson events” distribution or the Exponential to analyze count data. A simple transformation to make Exponential data essentially symmetrical was described in my Journal of Quality Technology article in July 1999 and utilizes the fourth-root of the data.

\[ m_t = \left( \bar{x} + \frac{1}{12} \right)^{\frac{3}{5}} \]  
(2)

Why the constant 1/12\textsuperscript{th} ? In The Advanced Theory of Statistics, Volume 3, 4th Edition by Kendall, Stuart, and Ord is a discussion on “Removal of transformation bias”, paragraph 37.17. I’ll not go into the details, but once a transformation has been made, to obtain the original mean from the transformed data is not just a simply “inverse transformation”. Although there are some equations listed, I choose to take the simple approach and regress \( m_t^{1.5} \) versus \( \bar{x} \). The slope term was 0.9997 and I chose 1 and the constant was 0.8661 or 1/0.8661 = 11.55, which I rounded to 12.

The expected standard deviation on the transformed scale was determined to be,

\[ s_t = \left( \frac{2}{3} \right)^{\frac{1}{6}} \]  
(3)

How was this obtained? I initially regressed \( \ln[m_t] \) versus \( \ln[\bar{x}] \) and with the exception of the first two data points, the fit is very good. Remember in the previous discussions on the parameters of skewness, kurtosis, etc. that the recommended transformation did an excellent performance above a Poisson mean of ~3 or ~4. Therefore I set aside these two data points and regressed again. The slope term was 0.1663 or 1/0.1663 = 6.013 and I chose 6 to give an exponent of 1/6\textsuperscript{th}. The constant was −0.4043 and \( e^{-0.4043} = 0.6674 \). I chose to call this 2/3\textsuperscript{rd}.
An Application To SPC

My paper “Calculating The (Almost) Exact Control Limits For A C-Chart” was recently published in *Quality Engineering, Volume 18, Number 3, July-September 2006* and thus I’ll not go into a lot of details, but extract a few items from it.

Based on the equations previously discussed, the following two equations were developed to give (almost) exact 0.00135 (LCL) and the 0.99865 (UCL) limits.

\[ \text{LCL} = \left( m + \frac{1}{2}c \right)^{\frac{1}{2}} - 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} \] \hspace{1cm} (10)

\[ \text{UCL} = \left( m + \frac{1}{2}c \right)^{\frac{1}{2}} + 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} \] \hspace{1cm} (11)

Of course if the calculated LCL is negative, then there is no LCL.

An example of calculations for the LCL. If \( \bar{c} = 6.607675 \), a lower tail area of 0.00135 corresponds to a LCL of 1. Equation (10) calculates 1.0 and the traditional calculation, mentioned in the Introduction, calculates –1.1. If \( \bar{c} = 8.900233 \), a lower tail area of 0.00135 corresponds to a LCL of 2. Equation (10) calculates 2.0 and the traditional calculation calculates 0.0. The traditional calculation gives a LCL that is ~2 units low. Thus some low values will not plot “Out-of-Control” due to an incorrect LCL.

A few calculations for the UCL. If \( \bar{c} = 0.052883 \), an upper tail area of 0.99865 corresponds to an LCL of 1. Equation (11) calculates 1.1 and the traditional calculation calculates 0.7. If \( \bar{c} = 0.211682 \), an upper tail area of 0.99865 corresponds to an UCL of 2. Equation (11) calculates 2.1 and the traditional calculation calculates 1.6. As the Poisson mean increases, the traditional calculation limit deviates more. Thus at \( \bar{c} = 25.959924 \), an upper tail area of 0.99865 corresponds to an UCL of 42. Equation (11) calculates 42.0 and the traditional calculation calculates 41.2. This is almost a deviation of ~1 which results in some data points plotting “Out-of-Control”, but they are not due to an incorrect UCL.

An Example

Ryan (2000, p.172)11 presents a set of 25 nonconformity data points that will be used to illustrate these calculations. The nonconformity data have a mean of 7.56 and a range of 1 to 17. The assumption made is that the data follows a Poisson distribution. The calculations use the proposed equations 10 and 11:

\[ \text{LCL} = \left( m + \frac{1}{2}c \right)^{\frac{1}{2}} - 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} = \left[ 7.56 + \frac{1}{2}c \right]^{\frac{1}{2}} - 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} = 1.37 \]

\[ \text{UCL} = \left( m + \frac{1}{2}c \right)^{\frac{1}{2}} + 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} = \left[ 7.56 + \frac{1}{2}c \right]^{\frac{1}{2}} + 3\left( \frac{1}{2}c \right)^{\frac{1}{2}} = 16.52. \]

Figure 3 is a plot of the data, which shows two “Out-of-Control” signals. However the “trained eye” would see that there was a major reduction in the nonconformities and accordingly surmise that there appears to be two populations with means of ~11 and ~5, rather than one Poisson distribution. The plot, however, has served its initial purpose in graphically portraying this information. An exponentially weighted moving average (EWMA) and/or a cumulative sum (CUSUM) analysis could be applied to refine the conclusions.

Conclusions

This paper has illustrated a relatively simple power transformation of Poisson type data so that it can be made symmetric for cases of application in Statistical Process Control work. Two simple equations are proposed for calculating the LCL and UCL for Poisson type data to be used in place of the traditional \( c \)-chart calculations. The agreement between the exact LCL and UCL, as determined by the lower tail area and upper tail area, is excellent.
When performing a CUSUM analysis or an EWMA analysis of Poisson type data, it is recommended to do that on the transformed data since it is symmetric.

Endnotes

Bibliography