Tina Memo No. 2001-010
Internal Report

The Effects of a Square Root Transform on a Poisson Distributed Quantity.

N.A. Thacker and P.A. Bromiley

Last updated
25 / 01 / 2009

This document forms part of the **Statistics and Segmentation** Series (2008-001) available from www.tina-vision.net.

2007-008	Tutorial: Defining Probability for Science.
2001-007	Performance Characterisation in Computer Vision: The Role of Statistics in Testing and Design.
2002-007	The Effects of an Arcsin Square Root Transform on a Binomial Distributed Quantity.
2001-010	The Effects of a Square Root Transform on a Poisson Distributed Quantity.
2004-004	Shannon Entropy, Renyi Entropy, and Information.
2002-002	Validating MRI Field Homogeneity Correction Using Image Information Measures.
2004-001	Empirical Validation of Covariance Estimates for Mutual Information Coregistration.
2004-005	The Equal Variance Domain: Issues Surrounding the Use of Probability Densities in Algorithm Design.
2009-008	Avoiding Zero and Infinity in Sample Based Algorithms.
2001-008	Derivation of the Renormalisation Formula for the Product of Uniform Probability Distributions and Extension to Non-Integer Dimensionality.
2001-005	Model Selection and Convergence of the EM Algorithm.
2003-007	Noise Filtering and Testing for MR Using a Multi-Dimensional Partial Volume Model.
2002-004	A Novel Method for Non-Parametric Image Subtraction: Identification of Enhancing Lesions in Multiple Sclerosis from MR Images.
2001-014	Bayesian and Non-Bayesian Probabilistic Models for Image Analysis.
1997-001	The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data.
1999-001	The Bhattacharyya Measure requires no Bias Correction.
1999-004	B-Fitting: An Estimation Technique With Automatic Parameter Selection.
2005-008	Tutorial: Beyond Likelihood.

WWW.TINA-VISION.NET

Imaging Science and Biomedical Engineering,
School of Cancer and Imaging Sciences,
University of Manchester, Stopford Building,

Oxford Road, Manchester, M13 9PT.

The Effects of a Square Root Transform on a Poisson Distributed Quantity

N.A. Thacker and P.A. Bromiley
Imaging Science and Biomedical Engineering,
School of Cancer and Imaging Sciences,
University of Manchester, Stopford Building,
Oxford Road, Manchester, M13 9PT.
paul.bromiley@man.ac.uk

Abstract

The shape of the Poisson distribution is controlled by a single parameter λ, which determines both the mean and the variance. It is well known that, at large values of λ, the Poisson distribution is well approximated by a Gaussian distribution, due to the central limit theorem. However, it has been asserted that, after transformation into square-root space, this is more accurate at lower values of λ, and furthermore the variance of the distribution is made constant across the space, making the distribution a more valid basis for a similarity function. This document provides more proof for this assertion based upon an estimate of the errors associated with both approaches.

1 Approximation of a Poisson to a Gaussian

We start by calculating the errors associated with approximating a Poisson distribution with a Gaussian distribution with variance equal to the mean. Let $P(x|\lambda)$ be a Poisson distribution with mean λ,

$$P(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Applying Stirling's approximation,

$$\alpha! = \alpha^\alpha e^{-\alpha}\sqrt{2\pi\alpha}$$

which is accurate to better than a few percent for $\lambda > 5$ and becomes exact as $\lambda \to \infty$[1], gives

$$P(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x^x e^{-x}\sqrt{2\pi x}} \tag{1}$$

This is generally approximated by a Gaussian distribution $G(x|\lambda)$ with mean λ and standard deviation $\sqrt{\lambda}$

$$P(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x^x e^{-x}\sqrt{2\pi x}} = G(x|\lambda)e^{h(x,\lambda)} = \frac{e^{-(x-\lambda)^2)/2\lambda}}{\sqrt{2\pi\lambda}}e^{h(x,\lambda)} \tag{2}$$

where $e^{h(x,\lambda)}$ is the approximation error.

Taking logarithms of both sides gives

$$x - \lambda + x\ln(\lambda) - x\ln(x) - \ln(\sqrt{x}) = -\frac{(x-\lambda)^2}{2\lambda} - \ln\sqrt{\lambda} + h(x,\lambda) \tag{3}$$

which can be re-written as

$$(1 + \frac{1}{2x})\ln(\frac{\lambda}{x}) = -\frac{(x-\lambda)}{x} - \frac{(x-\lambda)^2}{2\lambda x} + \frac{h(x,\lambda)}{x}$$

Making the substitution $\frac{\lambda}{x} = z + 1$ gives

$$(1 + \frac{1}{2x})\ln(z+1) = \frac{1}{2}(z+1) - \frac{1}{2(z+1)} + \frac{h(x,\lambda)}{x}$$

[1] This approximation introduces a multiplicative error that is the same in the following analyses.

Using the Mercator series expansion (which is valid for $-1 < z < 1$ or $\lambda/2 < x < \infty$) for $ln(1+z)$

$$ln(1+z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4... \tag{4}$$

and the binomial expansion

$$(1+\alpha)^\beta = 1 + \beta\alpha + \frac{\beta(\beta-1)\alpha^2}{2!} + \frac{\beta(\beta-1)(\beta-2)\alpha^3}{3!}... \tag{5}$$

for the reciprocal z term gives

$$(1+\frac{1}{2x})(z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + ...) = z - \frac{z^2}{2} + \frac{z^3}{2} - \frac{z^4}{2} + ... + \frac{h(x,\lambda)}{x}$$

Thus the error term is given by

$$h(x,\lambda) = \frac{z}{2} - \frac{z^2}{4} - \frac{(x-1)z^3}{6} + \frac{(2x-1)z^4}{8}...$$

$$= \frac{(\lambda-x)}{2x} - \frac{(\lambda-x)^2}{4x^2} - \frac{(x-1)(\lambda-x)^3}{6x^3} + \frac{(2x-1)(\lambda-x)^4}{8x^4}... \tag{6}$$

which has large linear factors in z for small x but is dominated by the cubic terms for large x. It is also worth pointing out here that in practical circumstances we measure x as an estimate of λ and must approximate the probability of λ given x. Under these circumstances Eq. **??** becomes

$$G'(\lambda|x) = \frac{e^{-(x-\lambda)^2)/2x}}{\sqrt{2\pi x}} e^{h'(x,\lambda)}$$

There is thus a different approximation error for this case which follows from a similar analysis and is

$$h'(x,\lambda) = \frac{(\lambda-x)^3}{3x^2} - \frac{(\lambda-x)^4}{4x^3} + ...$$

2 Approximation of a Square-Root Poisson to a Gaussian

We now wish to estimate the error associated with approximating the distribution of the square root of a Poisson variable with a Gaussian. Starting from Eq. **??** the transformation from x to $\sqrt{x}$ satisfies the conditions for the use of the well-known change of variables formula for a differentiable function,

$$g(y) = f[r^{-1}(y)]\frac{dr^{-1}y}{dy} \quad if \quad y = r(x) \tag{7}$$

giving

$$P(y|\lambda) = \frac{2ye^{-\lambda}\lambda^{y^2}}{((y^2)^{y^2})e^{-y^2}\sqrt{2\pi y^2}} = \frac{2}{\sqrt{2\pi}}e^{y^2(1-\ln\frac{y^2}{\lambda})-\lambda}$$

Now substitute

$$\lambda = \gamma^2 \tag{8}$$

giving

$$P(y|\gamma) = \frac{2}{\sqrt{2\pi}}e^{y^2(\ln\frac{\gamma^2}{y^2}+1)-\gamma^2} \tag{9}$$

Approximating this with a Gaussian distribution $G(y|\gamma, \sigma = 0.5)$ with mean γ and standard deviation 0.5

$$G(y|\gamma, \sigma = 0.5) = \frac{2}{\sqrt{2\pi}}e^{-2(y-\gamma)^2} \tag{10}$$

gives

$$P(y|\gamma) = \frac{2}{\sqrt{2\pi}}e^{y^2(\ln\frac{\gamma^2}{y^2}+1)-\gamma^2} = G(y|\gamma, \sigma = 0.5)e^{g(\gamma,y)} = \frac{2}{\sqrt{2\pi}}e^{-2(y-\gamma)^2}e^{g(\gamma,y)}$$

where $e^{g(\gamma,y)}$ is the approximation error. Taking logarithms of both sides gives

$$y^2 \ln \frac{\gamma^2}{y^2} + y^2 - \gamma^2 = -2(y-\gamma)^2 + g(\gamma, y) \quad or \quad ln \frac{\gamma^2}{y^2} = \frac{4\gamma}{y} - 3 - \frac{\gamma^2}{y^2} + \frac{g(\gamma, y)}{y^2}$$

As in the previous derivation, substitute

$$z = \frac{\lambda}{x} - 1 = \frac{\gamma^2}{y^2} - 1 \tag{11}$$

giving

$$\ln(z+1) = 4\sqrt{z+1} - 3 - (z+1) + \frac{g(\gamma, y)}{y^2}$$

Substituting the Mercator series expansion for the $\ln(z+1)$ term and the binomial expansion of $\sqrt{z+1}$

$$\sqrt{z+1} = 1 + \frac{z}{2} - \frac{z^2}{8} + \frac{z^3}{16} - \frac{5z^4}{128} ...$$

gives

$$z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} ... = 4[1 + \frac{z}{2} - \frac{z^2}{8} + \frac{z^3}{16} - \frac{5z^4}{128} ...] - 3 - z - 1 + \frac{g(\gamma, y)}{y^2}$$

and therefore

$$g(\gamma, y) = y^2 (\frac{z^3}{12} - \frac{3z^4}{32} + ...)$$

the lack of terms below z^3 is consistent with our assumption of equal variances on the Gaussian approximation. Substituting back our original variables gives

$$g(x, \lambda) = \frac{(\lambda - x)^3}{12x^2} - \frac{3(\lambda - x)^4}{32x^3} + ... \tag{12}$$

In this case, the symmetry of Eq. **??** means that $G = G'$. This demonstrates that errors on both the Poisson approximation and the distribution of λ have equivalent errors $g' = g$.

3 Conclusions

Comparison of $g(x, \lambda)$ with $h(x, \lambda)$ demonstrates the superiority of the new approximation for small x and that the leading order (cubic) terms differ by a factor of two for large x. Comparison with h' demonstrates a consistent improvement of a factor of 4.

There are several practical implications of this result. The first is in the construction of χ^2 similarity measures, which rely on the approximation of a Poisson by a Gaussian, such as

$$\chi^2 = \sum_i \frac{(h_i - t_i)^2}{2t_i} \quad and \quad \chi^2 = \sum_i \frac{(h_i - k_i)^2}{2(h_i + k_i)}$$

where h_i and k_i are frequency measures such as histogram values and t_i is the theoretical estimate. These functions are better estimated (particularly for small values of h_i and k_i) as

$$\chi^2 = 2 \sum_i (\sqrt{h_i} - \sqrt{t_i})^2 \quad and \quad \chi^2 = 2 \sum_i (\sqrt{h_i} - \sqrt{k_i})^2$$

These results for statistical similarity measures also extend to measures for probability similarity, with the Bhattacharyya (and equivalent Matusita) measure being most suitable for the comparison of highly dissimilar probability densities (where the uniform variance of the new measure makes Euclidean difference measures appropriate).

$$B = \int \sqrt{P_1(x) P_2(x)} dx$$

In addition, the Poisson distribution can be approximated better by a function of the form

$$P' = \frac{1}{\sqrt{2\pi x}} e^{-2(\sqrt{\lambda} - \sqrt{x})^2}$$

than the standard Gaussian with width equal to the mean.

Finally, a Gaussian random variable x with variance equal to 1/4 and mean γ can be used to approximate a Poisson random variable of mean γ^2 by squaring (x^2), or conversely the variance of a Poisson random variate can be stabilised at 1/4 by applying a square-root transform.