An estimator for ratios of Poisson distributed quantities

Pijpers, F.P.

Publication date
2021

Document Version
Final published version

License
Unspecified

Link to publication

Discussion Paper

An estimator for ratios of Poisson distributed quantities

Frank P. Pijpers

November 2021

Contents

A common task for national statistical institutes is to estimate the size of a subpopulation, with certain specific characteristics, as a proportion of a larger population or subpopulation, in particular in a time series setting where this proportion is reported on a regular basis. This paper discusses the probability density function for the stochastic variable that describes this proportion, where the numerator is treated as realisations of a Poisson process, and the denominator is the sum of a fixed value and realisations of a Poisson process that is independent from the numerator.

1 Introduction

In official statistics it is common to publish tables of the number of individuals or firms that belong to certain subgroups of a population, categorised by properties. For people such categorization could be for example gender, age, or level of education. For businesses it would be more appropriate to consider properties such as turnover, business sector or number of people employed. Either might also be appropriate to regionalise, i.e. separate such groups out by region, and of course it is possible and usual to have multiple subdivisions, leading to multidimensional tables. Often it is appropriate to express the entries in these tables not in the form of absolute numbers but in terms of a proportion or percentage, which means that the size of the subgroup is compared to the size of some reference group or larger subgroup of that population. It is well-established that when such counts are in fact estimates, based on surveys, that there is a margin of uncertainty which depends on the survey size, and possibly also on biases. Conceptually, when calculating proportions of a particular subgroup compared to the whole population, the denominator is set to the size of the statistical reference population, and possibly also the inclusion probability of the group surveyed, both of which are assumed to be fixed and deterministic.

However, even when the numbers are based on administrative registers there are also margins of uncertainty to take into account. Administrative registers are not free of errors but even if they were error free, the counts are always performed over a finite time interval and within a certain spatial region. Due to unobserved behaviours of people or businesses these counts should therefore always be treated as realisations of a stochastic counting process even if the data source is administrative and even if that administrative source were perfectly accurate.

If it can be assumed that the behaviours of the people or businesses involved do not induce significant correlated statistics then the appropriate stochastic process is a Poisson process with the associated distribution function. The counts are then treated as estimators for the expectation value, the λ parameter of the Poisson process $Pois(\lambda)$. If the tabulated data are to be published in the form of ratios, it would appear that the above argument implies that numerator and denominator are each to be treated as realisations of Poisson processes. However, there is an important issue with this.

This issue is that for any Poisson process, regardless of the value of the parameter λ there is a finite probability for that process to produce a realisation, ie. a count, of 0. Since this can occur for the denominator in the ratio of two Poisson distributed variables there is a finite likelihood for the ratio to become infinite, and even a finite likelihood for the ratio to be undefined because it is $\frac{0}{0}$. That means that even an expectation value for this ratio is impossible to establish let alone any confidence margins. Treating the denominator as a fixed, non-stochastic, quantity is also not appropriate and would lead to an underestimation of the margins of uncertainty for the ratio. In the literature where counting statistics are discussed in completely different contexts (eg. Park et al. (2006)), in some cases a solution is chosen where the numerator is assumed to satisfy truncated Poisson statistics. Essentially the probability for a count of 0 is set to 0 and the remaining probabilities for any finite count are re-scaled in order to ensure proper normalisation of the probability distribution function. In Coath et al. (2013) a slightly different problem is adressed. In that paper both numerator and denominator contain a Poisson distributed variable the ratio of which is the variable of interest, but both require subtraction of separate noise terms. There is a parallel between Coath et al. (2013) and the present paper although the approaches to deal with 0 denominators are different.

While the approach of Coath et al. (2013) can also be applied here, this ignores that in the present application, the role of the denominator in the ratio is that of a reference: a more stable population than the numerator. The implication is that for this reference population in the denominator some subset might have the more volatile behaviour where a Poisson process is appropriate as a model, but part of that population can be regarded as fixed. For instance, out of the total population inhabiting a particular region during a particular time frame, such as a month for instance, the majority will have a 0 likelihood of altering at all from one month to the next. This would mean that for the denominator it is more appropriate to model this as the sum of some non-stochastic reference value N and a stochastic, Poisson behaved, part X. With the numerator modelled as an independent Poisson process Y the ratio is then a stochastic variable Z:

$$Z \equiv \frac{Y}{N+X} \tag{1}$$

2 the problem

For many counting processes the ratio Z is required, together with confidence intervals, of two separate Poisson distributed variables $X \equiv \text{Pois}(\lambda_x)$ and $Y \equiv \text{Pois}(\lambda_y)$, but with a denominator that also has a non-stochastic fixed term:

$$Z \equiv \frac{Y}{N_0+X} \tag{2}$$

In other words the distribution function is needed for $P(Z=z)$.

In what follows it is assumed that $z \equiv \frac{n}{m}$ with $n, m \in \mathbb{N}^+$ and where n and m are relative prime. Also $N_0 \in \mathbb{N}^+$ to ensure that Z never has its denominator $= 0$. For validity of approximations to be made, it is further assumed that $\lambda_x, \lambda_y, > 10$

$$\begin{aligned} P(Z=z) &= P\left(\frac{Y}{N_0+X}=z\right) \\ &= P\left(\frac{Y}{N_0+X}=\frac{n}{m}\right) \\ &= P\left(mY-nX=nN_0\right) \\ &= \sum_{l\geq N_0/m} P(Y=ln)P(X=lm-N_0) \\ &= \sum_{l\geq N_0/m} \frac{\lambda_y^{ln}e^{-\lambda_y}}{(ln)!}\frac{\lambda_x^{lm-N_0}e^{-\lambda_x}}{(lm-N_0)!} \end{aligned} \tag{3}$$

While the summation can be written out exactly, as shown here, a closed form solution is intractable for $m \neq n$. Therefore the Poisson distributions will be treated as approximately normal distributions with the appropriate mean and standard deviation. Given the assumptions for the λ, this is an allowable approximation. In addition, while l is a discrete variable it is treated as a continuous variable here, so that summations over l become integrations. The probability

$P(z)$ is therefore also replaced with a probability density $f(z)$, with appropriate normalisation:

$$
\begin{aligned}
f(z)\mathrm{d}z &\approx (N_0+\lambda_x)\int_{N_0/m}^{\infty}\frac{1}{\sqrt{2\pi\lambda_y}}e^{-\frac{1}{2}\frac{(ln-\lambda_y)^2}{\lambda_y}}\frac{1}{\sqrt{2\pi\lambda_x}}e^{-\frac{1}{2}\frac{(lm-N_0-\lambda_x)^2}{\lambda_x}}m\mathrm{d}l \\
&= \frac{m(N_0+\lambda_x)}{2\pi\sqrt{\lambda_x\lambda_y}}\int_{N_0/m}^{\infty} e^{-\frac{1}{2}\left[\frac{l^2n^2-2ln\lambda_y+\lambda_y^2}{\lambda_y}+\frac{l^2m^2-2lm(N_0+\lambda_x)+(N_0+\lambda_x)^2}{\lambda_x}\right]}\mathrm{d}l \\
&= \frac{m(N_0+\lambda_x)}{2\pi\sqrt{\lambda_x\lambda_y}}\int_{N_0/m}^{\infty} e^{-\frac{1}{2}\left[l^2\left(\frac{n^2}{\lambda_y}+\frac{m^2}{\lambda_x}\right)-2l\left(n+m(1+\frac{N_0}{\lambda_x})\right)+\lambda_y+\frac{(N_0+\lambda_x)^2}{\lambda_x}\right]}\mathrm{d}l \\
&= \frac{m(N_0+\lambda_x)}{2\pi\sqrt{\lambda_x\lambda_y}}\int_{N_0/m}^{\infty} e^{-\frac{1}{2}\left(\frac{n^2}{\lambda_y}+\frac{m^2}{\lambda_x}\right)\left[\left(l-\lambda_x\lambda_y\frac{n+m(1+\frac{N_0}{\lambda_x})}{n^2\lambda_x+m^2\lambda_y}\right)^2+\right.} \\
&\qquad\qquad e^{\left.+\lambda_y\frac{\lambda_y\lambda_x+(N_0+\lambda_x)^2}{n^2\lambda_x+m^2\lambda_y}-\left(\lambda_x\lambda_y\frac{n+m(1+\frac{N_0}{\lambda_x})}{n^2\lambda_x+m^2\lambda_y}\right)^2\right]}\mathrm{d}l \\
&= \frac{m(N_0+\lambda_x)}{2\pi\sqrt{\lambda_x\lambda_y}}e^{-\frac{1}{2}\left[\lambda_y+\frac{(N_0+\lambda_x)^2}{\lambda_x}-\left(n+m(1+\frac{N_0}{\lambda_x})\right)^2\frac{\lambda_x\lambda_y}{n^2\lambda_x+m^2\lambda_y}\right]}\sqrt{\frac{\lambda_x\lambda_y}{n^2\lambda_x+m^2\lambda_y}}\int_{x_0}^{\infty}e^{-\frac{1}{2}x^2}\mathrm{d}x \\
&= \frac{(N_0+\lambda_x)}{2\pi\sqrt{z^2\lambda_x+\lambda_y}}e^{-\frac{1}{2}\left[\lambda_y+\frac{(N_0+\lambda_x)^2}{\lambda_x}-\left(z+(1+\frac{N_0}{\lambda_x})\right)^2\frac{\lambda_x\lambda_y}{z^2\lambda_x+\lambda_y}\right]}\int_{x_0}^{\infty}e^{-\frac{1}{2}x^2}\mathrm{d}x \\
&= \frac{(N_0+\lambda_x)}{2\pi\sqrt{z^2\lambda_x+\lambda_y}}e^{-\frac{1}{2}\left[\frac{\left(z(N_0+\lambda_x)-\lambda_y\right)^2}{z^2\lambda_x+\lambda_y}\right]}\int_{x_0}^{\infty}e^{-\frac{1}{2}x^2}\mathrm{d}x
\end{aligned}
\tag{4}
$$

in which the integration variable x is defined as:

$$
x \equiv \left[l-\lambda_x\lambda_y\frac{n+m(1+\frac{N_0}{\lambda_x})}{n^2\lambda_x+m^2\lambda_y}\right]\sqrt{\frac{n^2\lambda_x+m^2\lambda_y}{\lambda_x\lambda_y}} \tag{5}
$$

and the integration limit x_0 satisfies:

$$
\begin{aligned}
x_0 &= \left[\frac{N_0}{m}-\lambda_x\lambda_y\frac{n+m(1+\frac{N_0}{\lambda_x})}{n^2\lambda_x+m^2\lambda_y}\right]\sqrt{\frac{n^2\lambda_x+m^2\lambda_y}{\lambda_x\lambda_y}} \\
&= \left[N_0-\lambda_x\lambda_y\frac{z+(1+\frac{N_0}{\lambda_x})}{z^2\lambda_x+\lambda_y}\right]\sqrt{\frac{z^2\lambda_x+\lambda_y}{\lambda_x\lambda_y}} \\
&= \frac{\lambda_x(z^2N_0-z\lambda_y-\lambda_y)}{\sqrt{\lambda_x\lambda_y(z^2\lambda_x+\lambda_y)}}
\end{aligned}
\tag{6}
$$

It can be convenient to re-scale the Poisson parameters with N_0:

$$
\begin{aligned}
\widetilde{\lambda_x} &\equiv \lambda_x/N_0 \\
\widetilde{\lambda_y} &\equiv \lambda_y/N_0
\end{aligned}
\tag{7}
$$

so that:

$$
f(z)\mathrm{d}z = \frac{\sqrt{N_0}(1+\widetilde{\lambda_x})}{2\pi\sqrt{z^2\widetilde{\lambda_x}+\widetilde{\lambda_y}}}e^{-\frac{1}{2}N_0\left[\frac{\left(z(1+\widetilde{\lambda_x})-\widetilde{\lambda_y}\right)^2}{z^2\widetilde{\lambda_x}+\widetilde{\lambda_y}}\right]}\int_{x_0}^{\infty}e^{-\frac{1}{2}x^2}\mathrm{d}x \tag{8}
$$

with

$$x_0 = \sqrt{N_0} \frac{\widetilde{\lambda_x}(z^2 - z\widetilde{\lambda_y} - \widetilde{\lambda_y})}{\sqrt{\widetilde{\lambda_x}\widetilde{\lambda_y}(z^2\widetilde{\lambda_x} + \widetilde{\lambda_y})}} \tag{9}$$

A practical problem is that while the numerator Y is measured, for instance as an item retrieved from administrative data, the denominator is only available in the summed form $\text{Num}_i = N_0 + X_i$ so that N_0 and X_i are not known separately. This means that a separate step is still required to estimate N_0, if there is no a-priori value available. One route can be if the denominator is determined repeatedly. For administrative data this repeated measurement is not unreasonable to assume: national statistical institutes that make use of administrative registers tend to obtain extracts from such registers on a regular basis, which can be annually, but monthly and even daily updates are not at all unusual. This means it is feasible to obtain repeated measurements Num, where by assumption N_0 is constant and the variations in the values arise from X. If this is the case then the average of these repeated measurements corresponds to the expectation value:

$$\overline{\text{Num}} = N_0 + E(X) = N_0 + \lambda_X \tag{10}$$

The variance of these measurements, given the assumption that X follows a Poisson distribution, satisfies:

$$Var(\text{Num}) = Var(X) = \lambda_X \tag{11}$$

Combining Eqs. (10) and (11) leads to an estimator for N_0 from such data:

$$N_0 = \overline{\text{Num}} - Var(\text{Num}) \tag{12}$$

Asymptotically, i.e. for many samples, this is unbiased. For finite samples there might be some issues in that the precision with which the average and variance are estimators of λ_X can be low if only very few samples are available. If in addition N_0 is not much larger than λ_X, that, in combination with the finite precision, might result in an estimated value of $N_0 \leq 0$ which would violate the assumptions of this paper. Operationally it would therefore be advisable to use instead:

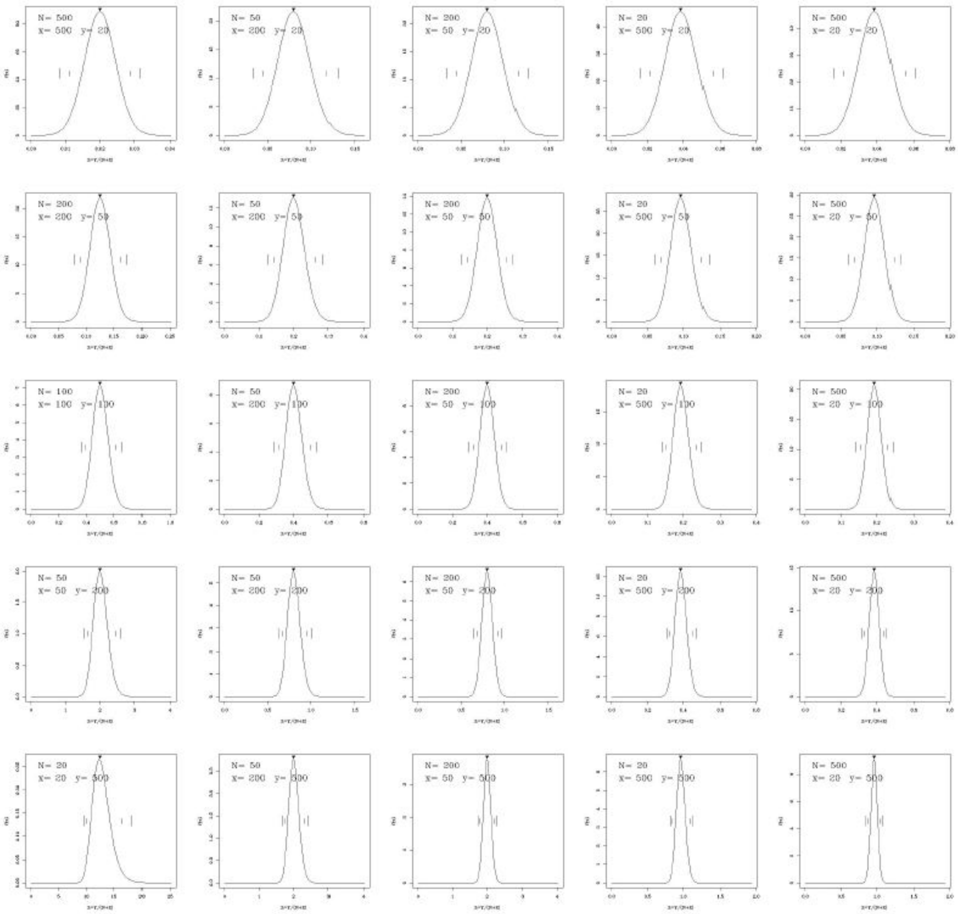
$$N_0 = \max\left(1, \overline{\text{Num}} - Var(\text{Num})\right) \tag{13}$$

so that even for small numbers of samples for the denominator a solution can always be obtained.

It is important to mention at this point that it may not be appropriate to regard the repeated extracts from registration data as fully independent. For instance there is an important demographic register - the base register of persons or 'basisregistratie personen' (BRP) - in which for any given person no changes are made if there is no life-course event to register (such as a birth or death, a marriage or legal partnership, a change of residence, etc.). Other longitudinal registers tend to suffer from the same problem, where registration errors, once made, may remain undetected for multiple months. While this should be captured in the formulation (10) it can mean that nevertheless with this approach the variance λ_X is slightly underestimated. In some cases it may be possible to correct for this if multiple sources are available for the same data but such procedures have their own pitfalls and shortcomings. From experience with the major registers in the Netherlands the effect is modest in most cases.

3 Some sample distributions

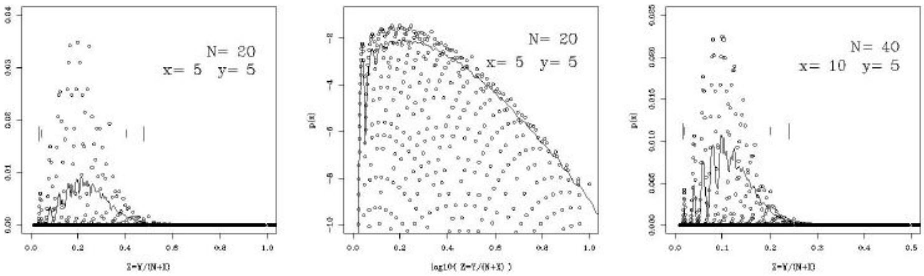
Figure 3.1 For a number of values of the three parameters N, λ_X and λ_Y, the resulting probability densities are shown. The parameter values are shown in each panel. Also indicated are the expectation value with a downward pointing triangle near the top of the panel and the 95% confidence interval with the shorter vertical lines and the 99% confidence interval with the longer vertical lines.



In figure 1 the resulting distribution function is shown for various values of the parameters N, λ_X, and λ_Y. Within each row the value of λ_Y is constant and going down the panels row by row, it increases in value. In the left-most column N and X are identical in value, both decreasing going down the panels. In the second and third column N and λ_X are unequal and the values are switched around between the two. In the fourth and fifth column the difference between N and X is larger, with again switching the values around between column 4 and column 5. The cases that are more likely to be encountered in official statistics are relatively large values of N, and smaller values for λ_X and λ_Y. In these cases the probability density is relatively symmetric such as the panel in the top right corner of Fig 1. The more asymmetric probability density, such as in the bottom left panel occurs for small values of N and λ_X and large values of λ_Y which are less likely to occur in typical settings in official statistics.

For smaller values of λ_X and λ_Y the explicit calculation (3) can be carried out. Some examples are shown in fig. 3.2. The open circles are the actual values of $P(z)$. The vertical lines indicating the 95 and 99 percentiles in fig. 3.2 are determined from these exact probabilities. Of course both the numerator and denominator in Z are integers, and therefore $P(z)$ is only defined for values

Figure 3.2 For two combinations of values of the three parameters N, λ_X and λ_Y, the resulting probabilities are shown by explicit summation. The parameter values are shown in each panel, where the middle panel has the same parameter values as the left panel, but the ordinate is a logarithmic scale. Also indicated are the 95% confidence interval with the shorter vertical lines and the 99% confidence interval with the longer vertical lines.



of $z \in \mathbb{Q}$. This means that it is not straightforward to see the connection between the continuous distributions of figure 3.1 and fig. 3.2. To facilitate this, a smoothed version is also shown as a solid line. This is constructed by first calculating the cumulative probability function from the exact $P(z)$. The cumulative function is monotonic which implies that a more stable interpolation can be obtained using a spline function. The spline is constructed using $\ln(z)$ as independent variable and $\ln(P(Z \leq z))$ as dependent variable. Using the spline, the function is resampled on a regularly spaced grid in $\ln(z)$. This resampled function is smoothed using a Gaussian kernel. after which the transformation back from the logarithmic domain and taking the derivative is carried out. The result is shown as a solid line in all panels of fig. 3.2. Evidently the peak of the probability density function is much lower than the highest value of the exact $P(z)$ which is entirely due to the smoothing. It is also worth noting that even the smoothed distribution function has a considerable amount of structure. Evidently this structure becomes less pronounced as λ_X and λ_Y increase.

4 more general case

As outlined in section 1, the formulation of the problem (2) implies a setting where the subpopulation of particular interest is represented by Y which is quite volatile, i.e. completely following a Poisson process, whereas the reference population in the denominator is a more stable sum of a fixed population and a smaller part behaving as a Poisson process. A natural extension of the formalism would be a setting in which also the numerator is partly fixed and partly Poissonian, so that the ratio Z would become:

$$Z \equiv \frac{M_0 + Y}{N_0 + X} \tag{14}$$

This leads to:

$$\begin{aligned} f(z)\mathrm{d}z &\approx (N_0 + \lambda_x) \int_{N_0/m}^{\infty} \frac{1}{\sqrt{2\pi\lambda_y}} e^{-\frac{1}{2}\frac{(ln - M_0 - \lambda_y)^2}{\lambda_y}} \frac{1}{\sqrt{2\pi\lambda_x}} e^{-\frac{1}{2}\frac{(lm - N_0 - \lambda_x)^2}{\lambda_x}} m dl \\ &= \frac{(N_0 + \lambda_x)}{2\pi\sqrt{z^2\lambda_x + \lambda_y}} e^{-\frac{1}{2}\left[\frac{\left(z(N_0+\lambda_x) - M_0 - \lambda_y\right)^2}{z^2\lambda_x + \lambda_y}\right]} \int_{x_0}^{\infty} e^{-\frac{1}{2}x^2} \mathrm{d}x \end{aligned} \tag{15}$$

in which the integration variable x is now defined as:

$$x \equiv \left[l - \lambda_x \lambda_y \frac{n(1+\frac{M_0}{\lambda_y}) + m(1+\frac{N_0}{\lambda_x})}{n^2\lambda_x + m^2\lambda_y} \right] \sqrt{\frac{n^2\lambda_x + m^2\lambda_y}{\lambda_x \lambda_y}} \tag{16}$$

and the integration limit from eq. (6) is adjusted accordingly. Note that (14) can trivially be reformulated as:

$$Z \equiv \frac{M_0}{N_0 + X} + \frac{Y}{N_0 + X} \tag{17}$$

which is the sum of two stochastic variables $Z_1 \equiv M_0/(N_0 + X)$ and $Z_2 \equiv Y/(N_0 + X)$. If Z_1 and Z_2 could be treated as independent variables the result would be simpler than eq. (15), but of course the presence of the same X in both denominators prevents this. However an analysis similar to that shown in the appendix demonstrates that the variance of Z_1 is much smaller than the variance of Z_2 for realistic situations, so that one would expect the resulting probability distribution function to be rather similar to the one with $M_0 = 0$, but shifted by an amount $M_0/(N_0 + \lambda_x)$.

5 Conclusions

It is common for national statistical institutes to report proportions of populations or subpopulations that staisfy certain properties of interest. Where such proportions are inferred from survey data, there is a natural way to also determine the margins of uncertainty of that estimate. However, if such proportions are determined from counts in numerator and denominator that are extracted from registers of administrative data, this does not mean that they are free of uncertainty. The processes controlling the registration have stochastic components. The distribution function most appropriate to describe the stochastic components is the Poisson distribution. For the numerator this is straightforward. For the denominator this can be problematic since there would be a finite probability for the denominator to be 0. Also conceptually it is more sensible to treat the denominator as the sum of a constant value > 0 and a stochastic component that follows a Poisson distribution, distinct from the numerator. This paper discusses the resulting distribution function appropriate for the ratio, enabling the reporting of margins of uncertainty as well as point estimates of ratios and proportions of populations are reported from source material that is not surveys but registers or administrative data. The expressions in the appendix were kindly provided by Sander Scholtus in the course of reviewing an earlier version of this paper, which I am happy to acknowledge.

References

Coath, C., R. Steele, and W. Lunnon (2013). Statistical bias in isotope ratios. *J. Anal. At. Spectrom. 28*, 52–58.

Park, T., V. Kashyap, A. Siemiginowska, D. van Dyk, Z. A., H. C., and B. Wargelin (2006, November). Bayesian Estimation of Hardness Ratios: Modeling and Computations. *Astroph.J. 652*(1), 610–628.

appendix

For the purposes of obtaining an approximate expression just for the expectation value and variance of Z one can also follow another route, making use of a Taylor series expansion of Z around the expectation value. The reasoning is as follows. Starting point is expression (2) which is written as $Z = f(X, Y)$, and the point around which the Taylor expansion is to be constructed is:

$$z_0 = f\left(E(X), E(Y)\right) = f(\lambda_X, \lambda_Y) = \frac{\lambda_Y}{N_0 + \lambda_X} \tag{18}$$

Then:

$$\begin{aligned} Z &\approx z_0 + \frac{\partial f}{\partial X}|_{(\lambda_X,\lambda_Y)}\,(X-\lambda_X) + \frac{\partial f}{\partial Y}|_{(\lambda_X,\lambda_Y)}\,(Y-\lambda_Y) \\ &= z_0 - \left[\frac{\lambda_Y}{(N_0+\lambda_X)^2}\right](X-\lambda_X) + \left[\frac{1}{(N_0+\lambda_X)}\right](Y-\lambda_Y) \\ &= z_0 + \frac{\lambda_Y}{N_0+\lambda_X}\left[\frac{Y-\lambda_Y}{\lambda_Y} - \frac{X-\lambda_X}{N_0+\lambda_X}\right] \end{aligned} \tag{19}$$

With this expression in hand, and using the assumption that X and Y are independent and therefore uncorrelated, the expression for the variance becomes:

$$\begin{aligned} \mathrm{var}\,(Z) &\approx \left(\frac{\lambda_Y}{N_0+\lambda_X}\right)^2\left[\mathrm{var}\left(\frac{Y}{\lambda_Y}\right) + \mathrm{var}\left(\frac{X}{N_0+\lambda_X}\right)\right] \\ &= \left(\frac{\lambda_Y}{N_0+\lambda_X}\right)^2\left[\left(\frac{\lambda_Y}{\lambda_Y^2}\right) + \frac{\lambda_X}{(N_0+\lambda_X)^2}\right] \\ &= \frac{\lambda_Y}{(N_0+\lambda_X)^2}\left[1 + \frac{\lambda_X\lambda_Y}{(N_0+\lambda_X)^2}\right] \end{aligned} \tag{20}$$

where the approximate equality arises because only the first term in the Taylor series is taken into account.

Figure 5.1 For the parameter values corresponding to the panels of figure 3.2 the expectation value of Z and the approximate variance and standard deviation are shown in the table.

λ_y	λ_x	N_0	E(Z)	var(Z)	σ(Z)
20	500	500	0,020	0,0000	0,004
20	200	50	0,080	0,0003	0,018
20	50	200	0,080	0,0003	0,018
20	500	20	0,038	0,0001	0,009
20	20	500	0,038	0,0001	0,009
50	200	200	0,125	0,0003	0,018
50	200	50	0,200	0,0009	0,030
50	50	200	0,200	0,0008	0,029
50	500	20	0,096	0,0002	0,014
50	20	500	0,096	0,0002	0,014
100	100	100	0,500	0,0031	0,056
100	200	50	0,400	0,0021	0,046
100	50	200	0,400	0,0017	0,042
100	500	20	0,192	0,0004	0,021
100	20	500	0,192	0,0004	0,019
200	50	50	2,000	0,0400	0,200
200	200	50	0,800	0,0052	0,072
200	50	200	0,800	0,0037	0,061
200	500	20	0,385	0,0010	0,032
200	20	500	0,385	0,0008	0,027
500	20	20	12,500	2,2656	1,505
500	200	50	2,000	0,0208	0,144
500	50	200	2,000	0,0112	0,106
500	500	20	0,962	0,0036	0,060
500	20	500	0,962	0,0019	0,044

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information