One in nine nurses will go to jail

Richard D. Gill* and Piet Groeneboom †

August 20, 2007

Abstract

It has often been noticed that, in observing the number of incidents that nurses experience during their shifts, there is a large variation between nurses. We propose a simple statistical model to explain this phenomenon and apply this to the Lucia de Berk case.

Introduction

We model the incidents a nurse experiences as a homogeneous Poisson process on the positive halfline, with a nurse-dependent intensity λ. As is well-known, a Poisson process is used to model incoming phone calls during non-busy hours, fires in a big city, etc. Since we believe incidents to be rare, a Poisson process is also a natural choice for modeling the incidents a nurse experiences.

Our model is parametric, and we take as the distribution of the intensity λ over nurses the Gamma$(\rho, \rho/\mu)$ distribution. Using this model, our sample (of one person!) consists of a realization of the random variable

$$(L, T, N),$$

where L has a Gamma$(\rho, \rho/\mu)$ distribution, and N, conditionally on $L = \lambda$ and $T = t$, has a Poisson distribution with parameter λt. The random variable T represents the time interval in which incidents take place (for a particular nurse).

If we would have n nurses, we (conceivably) would have a sample

$$(L_1, T_1, N_1), \ldots, (L_n, T_n, N_n),$$

of independent random variables, all having the same distribution as (L, T, N). The random variable N_i represents the number of incidents nurse i experiences in the time interval T_i. The trouble is, however, that nurses have overlapping services, hence nurses in the same hospital would not generate a sample of this type.

We use Derksen and de Noo's revised data set, taking account of incidents among the other nurses (which formerly were taken to be, by definition, *not suspicious*), and removing incidents and deaths for which Lucia was deemed innocent (not charged with murder or attempted murder, presumably because these events were medically speaking "expected to happen, when they actually did").

*Mathematical Institute, Leiden University; http://www.math.leidenuniv.nl/∼gill
†DIAM, Delft University; http://ssor.twi.tudelft.nl/∼pietg

To simplify matters, we take (for the time being) $\rho = 1$. This means, among other things, that it can easily happen that one nurse has twice the incident rate of another nurse. The statistical problem boils down to the estimation of the parameter μ.

The numbers

Combining the Juliana Children's Hospital and the two wards of the Red Cross Hospital, Lucia had 203 shifts, 7 incidents. It is not clear whether this combination works out pro or contra Lucia (this depends on whether she did proportionately more or less shifts at the different wards, and whether the overall mean incident rate is larger or smaller at each ward).

We do need to do a combined analysis; the alternative to quick and dirty "just add everything up" is a much more complicated analysis and certainly requires making more, unsupportable, assumptions. We'll take the overall probability of an incident per shift to be the ratio of total incidents to total shifts, $\mu = 23/1681$. If we take a shift to be our unit time interval, then this would be a moment estimate of the mean intensity of incidents EL. This means, that, conditionally on $T = 203$, the number of incidents for Lucia follows a mixture of Poisson random variables with parameter $203L$, where the intensity L has a $\Gamma(1, 1/\mu)$ distribution, which is in fact the exponential distribution on $[0, \infty)$ with first moment μ. Thus on average, an innocent Lucia would experience $203 \cdot \mu = 203 \cdot 23/1681 \approx 2.777513$ incidents.

The probability of having 7 or more incidents is given by:

$$\frac{1}{\mu}\int_0^\infty \mathbb{P}\left\{N \geq 7 | L = \ell, T = 203\right\} e^{-\ell/\mu}\, d\ell$$

It is well-known that, for a random variable P, which is distributed according to a Poisson(λ) distribution, we have:

$$\mathbb{P}\{P \geq n\} = \frac{1}{(n-1)!}\int_0^\lambda e^{-x}x^{n-1}\, dx,$$

see, e.g., Feller (1968), Problem 46, Chapter VI.10, p. 173. So we find:

$$\begin{aligned}
&\frac{1}{\mu}\int_0^\infty \mathbb{P}\left\{N \geq 7 | L = \ell, T = 103\right\} e^{-\ell/\mu}\, d\ell \\
&= \frac{1}{6!\mu}\int_0^\infty \left\{\int_0^{203\ell} e^{-y}y^6\, dy\right\} e^{-\ell/\mu}\, d\ell \\
&= \frac{1}{6!}\int_0^\infty e^{-\{1+1/(203\mu)\}y}y^6\, dy = \{1 + 1/(203\mu)\}^{-7} \approx 0.116185,
\end{aligned}$$

where we use:

$$\int_0^\infty e^{-ay}y^6\, dy = a^{-7}\int_0^\infty e^{-x}x^6\, dx = a^{-7}\Gamma(7) = a^{-7}6!,$$

by the definition of the gamma function. A picture of the probalilities $\mathbb{P}\{N \geq k | T = 203\}$, $k = 1, 2, \ldots$ is shown in Figure 1.

Smaller than one in eight, larger than one in nine. So we'll go for *one in nine*.

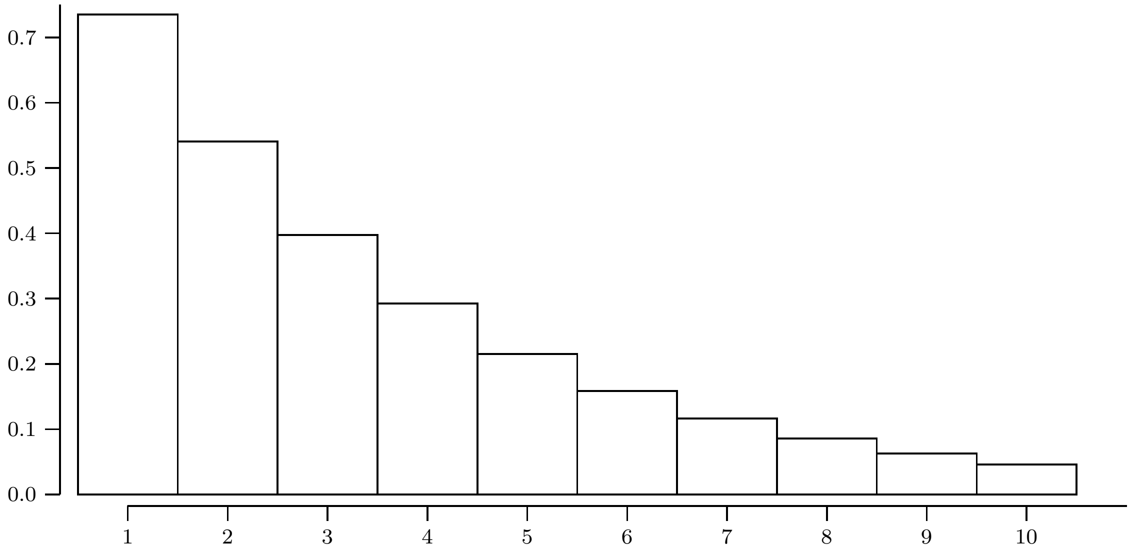

Figure 1: Probabilities (in the model) that the number of incidents in 203 shifts for one nurse is at least 1,2,3,..., if $\mu = 23/1681$. The probabilities are given by the heights of the columns above $1, 2, 3, \ldots$, respectively.

Conclusion: One in Nine will go to jail

A modest amount of variation makes the chance that an innnocent nurse experiences at least as many incidents as the number Lucia actually did experience, the somewhat unremarkable *one in nine*.

The fact that this amount of heterogeneity easily explains the coincidence, is strong evidence that it does truly exist: the intuition of medical specialists needs to be challenged and the experience of nurses and nursing specialists is supported.

References

Feller, W. (1968). An introduction to Probability Theorey and Its Applications, 3rd edition. Wiley. New York.