Updating a Classic: "The Poisson Distribution and the Supreme Court" Revisited

KEYWORDS:
Teaching;
Poisson distribution;
Goodness of fit test;
Supreme Court.

Julio H. Cole
Universidad Francisco Marroquín, Guatemala.
e-mail: jhcole@ufm.edu

Summary
W.A. Wallis studied vacancies in the US Supreme Court over a 96-year period (1837–1932) and found that the distribution of the number of vacancies per year could be characterized by a Poisson model. This note updates this classic study.

◆ INTRODUCTION ◆

One way to motivate students' interest in probability theory is to illustrate its applicability in novel and unexpected contexts. In this regard, the Poisson distribution is a guaranteed winner, since there are so many strange and wonderful real-world examples that can be used for classroom instruction. Indeed, ever since Bortkiewicz (1898) (cf. Keynes 1971) published his classic analysis of deaths-by-horse-kick data, statisticians have been coming up with ever more intriguing applications over a wide range of natural and social phenomena. One of the most interesting of these is Wallis's famous application of the Poisson to describe the occurrence of vacancies in the US Supreme Court (Wallis 1936).

These vacancies have occurred because of death or retirement of Supreme Court justices. Historically, the proportions of vacancies due to death and vacancies due to retirement have been roughly equal in number. (This is nicely illustrated by the two most recent vacancies: Chief Justice William Rehnquist died in 2005, while Justice Sandra Day O'Connor retired in 2006 due to personal, family reasons.) A third possible cause of vacancy, impeachment of a sitting justice, has never occurred.

Studying vacancies from 1837 to 1932, Wallis found that a Poisson process with parameter $\mu = 0.5$ vacancies/year gave a remarkably good fit to the distribution of the number of vacancies per year over the sample period. Wallis, by the way, studied vacancies from 1837 onwards because the size of the Court varied since its inception in 1789, but has remained stable at nine members since 1837 (with a minor exception in 1863–1867, which he chose to ignore). I have used this numerical example many times in teaching the Poisson distribution to students of introductory statistics, and it has never failed to intrigue them. In fact, for both students and teacher, it is often one of the most enjoyable classes of the entire semester.

There are two problems, however, that mar somewhat the pedagogic utility of this example. The first and most obvious problem, of course, is that Wallis's paper is now over 70 years old, so as time goes by students cannot fail to wonder if the good fit was a particularity of this specific sample period. (The teacher wonders too.) In other words, what happened after 1932?

The second problem is that, to compute the expected frequencies with a Poisson model, we need a value for the parameter μ, which in most applications is estimated from the sample itself, since usually no other information is available. Wallis obtained his value for μ by comparing the total number of vacancies over his sample period (48) with the total number of years in the sample (96), yielding an average of 0.5 vacancies/year. The problem here is that the parameter used to compute the expected frequencies under the null hypothesis (i.e. that the distribution is Poisson) is estimated from the same data that are used to test the null hypothesis. For some students, this might raise a suspicion that the data are being 'over-fitted' and

Number of vacancies (x)	Probability	Number of years in which x vacancies occurred			
		1837–1932		1933–2007	
		Observed	Expected	Observed	Expected
0	0.6065	59	58.227	47	45.490
1	0.3033	27	29.113	21	22.745
2	0.0758	9	7.278	7	5.686
3	0.0126	1	1.213	0	0.948
>3	0.0018	0	0.168	0	0.131
Totals	1.0000	96	96.000	75	75.000
Chi-square goodness of fit tests:					
1837–1932	$\chi_1^2 = 0.371$	5% critical value = 3.841			
1933–2007	$\chi_2^2 = 0.192$	5% critical value = 5.991			

Table 1. US Supreme Court vacancies, 1837–1932 and 1933–2007

that the test is somehow biased in favour of acceptance.

In this note I will try to deal with these problems by updating Wallis's data and validating his results by checking them against an independent sample. As it turns out, once this is done, this classic case study becomes an even more powerful classroom example.

◆ LITERATURE REVIEW ◆

This is not the first time this case has been updated. Callen and Leidecker (1971) extended Wallis's sample to 1970 and found that the Poisson continued to provide a good fit for the enlarged, 134-year sample. Kinney (1973) extended the sample both forwards (to 1972) and backwards (to 1790), and found that for the full period (1790–1972) and for the restricted period (1837–1972) the distributions were essentially the same, which suggests that Wallis's misgivings about the pre-1837 years were perhaps unwarranted. Morrison (1977) retained 1837 as the starting year but extended the sample to 1975, while Ulmer (1982) started with 1790 and extended the sample to 1980. In all of these studies, the authors fitted a Poisson model directly to the enlarged samples. The results are all qualitatively similar to Wallis's original finding (the distribution is Poisson), and the numerical results are similar as well.

In this study I complement Wallis's data set by counting the number of vacancies per year since 1933 through the last available full year, 2007. I base my counts on the table entitled "Members of the Supreme Court of the United States" (United States Supreme Court 2008). This new, post-1932 sample is then used to validate Wallis's results by comparing the observed frequencies with expected frequencies under a Poisson distribution with Wallis's original parameter estimate.

◆ DATA AND RESULTS ◆

Table 1 reports Wallis's data for 1837–1932 and the corresponding data for 1933–2007, as well as the expected frequencies under the hypothesis that the probability of each number of vacancies per year follows a Poisson distribution with parameter $\mu = 0.5$:

$$p(x) = \frac{(0.5)^x e^{-0.5}}{x!}, \; x = 0, 1, 2, \ldots$$

The table also reports the results for the chi-square goodness-of-fit tests, which were computed by combining the 2, 3 and 'over 3' categories following the well-known rule that all expected frequencies under the null hypothesis should equal 5 or more. In a conventional chi-square test, the degrees of freedom equals the number of categories minus one, but in the 1837–1932 sample an additional degree of freedom is lost because μ was estimated from the same sample. The test for the 1933–2007 sample has two degrees of freedom since the parameter was estimated independently. In both cases the hypothesis that the underlying distribution is Poisson with parameter $\mu = 0.5$ is not rejected.

◆ CONCLUSIONS ◆

The remarkably close fit in Wallis's original data is of course what gave rise to this whole exercise. What is even more remarkable, however, is the close fit for the second set of data, since the parameter for the Poisson probabilities was not estimated from the sample data – 'the observations whose conformity is to be tested', as Wallis puts it (Wallis 1936, p. 379). Instead, the expected frequencies for the second sample period were computed using the parameter that was estimated from the first period. This is a much stronger test.

Wallis's classic analysis of Supreme Court vacancies provides an excellent classroom application of a process modelled well by a Poisson. My hope, in updating this example and showing that it still applies, is that it will continue to amaze and instruct future generations of students for many years to come.

References

Bortkiewicz, L. von (1898). *Das Gesetz der Kleinen Zahlen*. Leipzig: B. G. Teubner.

Callen, E. and Leidecker, H. (1971). A mean life on the Supreme Court. *American Bar Association Journal*, **57**(12), 1188–1192.

Keynes, J.M. (1971). *The Collected Writings of John Maynard Keynes*, vol. 8, *A Treatise on Probability*. London: Macmillan, 439–440 (original work published 1921).

Kinney, J. (1973). Poisson updated. *The American Statistician*, **27**(4), 195.

Morrison, R.J. (1977). Franklin D. Roosevelt and the Supreme Court: An example of the use of probability theory in political history. *History and Theory*, **16**(2), 137–146.

Ulmer, S.S. (1982). Supreme Court appointments as a Poisson distribution. *American Journal of Political Science*, **26**(1), 113–116.

United States Supreme Court (2008). Members of the Supreme Court of the United States. http://www.supremecourtus.gov/about/members.pdf (accessed 1 November 2008).

Wallis, W.A. (1936). The Poisson distribution and the Supreme Court. *Journal of the American Statistical Association*, **31**(194), 376–380.

LOOK AHEAD

Buffon's coin problem and beyond
Business statistics at the top 50 US business programs
Cheating partners, conditional probability and contingency tables
Creating, implementing and integrating a first-year statistics requirement
Data and dialogue in primary school
Distance learning for teacher professional development in statistics education
Employing introductory statistics students at "Stats Dairy"
Formulating and using probability density functions with familiar geometric shapes
Hit me with your best plot
Looking at debit and credit card fraud
One is the likeliest number
Two applications of Venn diagrams
A visual model for the variance and standard deviation
Using pictures to enhance students' understanding of Bayes' Theorem
Why is S a biased estimate of σ?