4-2017

Poisson Windowing

David L. Farnsworth
Rochester Institute of Technology

Math. Scientist **42,** 43–50 (2017)
Printed in England

POISSON WINDOWING

DAVID L. FARNSWORTH,* *Rochester Institute of Technology*

Abstract

The statistical quality-control technique of windowing takes advantage of the additive property of the Poisson probability distribution and the geometry of the data. There are abundant applications of windowing to observations that naturally occur in arrays or that are produced in them, such as computer chips, which are manufactured many at a time on wafers. Standard statistical procedures are used.

Keywords: Poisson distribution; probability model; quality control; industrial statistics; windowing; additive property

2010 Mathematics Subject Classification: Primary 62-07
Secondary 97K80; 62P30

1. Introduction

A practical way to determine whether there are clusters of events is presented. The statistics used are relatively simple. More complicated analyses are possible, but the present methodology is useful, transparent, and usually sufficient.

Detecting clusters can be very important. They might be clusters of cancer sufferers in a neighborhood or of successful basketball shots by a player with a hot hand. The ideas presented are illustrated with a worked out example from quality control in which a quality engineer wants to determine whether there are clusters of imperfections in a large piece of carpet.

The testing process involves using a chi-square goodness-of-fit test to determine if the defects follow a Poisson distribution, which is equivalent to no overall pattern. Computing probabilities using a Poisson distribution or a normal-distribution approximation to a Poisson distribution is used to identify unusual observations, which might be identified with a cluster. The same process is repeated on a larger scale by combining neighboring regions into new, larger regions.

The prototypical example is from quality control. One goal of a quality engineer is to distinguish between chance events and events that have assignable causes. Events that concern the quality engineer are errors, defects, or items that are too large or too small. Chance events are scattered about in a nonsystematic fashion and arise from the general random variability of the process or product. Chance events are a feature of the overall process, and the cause or causes are intrinsic to the particular manufacturing process. Usually, those causes are correctable only by altering the whole process. On the other hand, events that have assignable causes recur periodically and show that a portion of the process needs to be fixed, thus preventing future bad events from that cause from occurring.

The main example is the production of rolls of carpet. The bad events are pulled tufts in the product. The carpet is an inexpensive variety, and some pulled tufts are expected at random in the production and do little harm. However, if there are clusters of pulled tufts, the carpet may not be useable. Regions containing a cluster will produce waste. In the particular example, it is

Received 4 November 2016.
* Postal address: School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY 14623, USA.
Email address: dlfsma@rit.edu

seen that a simple statistical test says that the pulled tufts are randomly distributed on a piece of carpet. However, two particular square feet of carpet appear to have a large number of pulled tufts. Then, a scanning technique called *windowing* is used on the same piece of carpet and helps to re-examine whether the pulled tufts are randomly distributed on new larger regions or windows. In the example, the windowing process reveals that there are two anomalous regions, or windows, where there appears to be a cluster of a large number of pulled tufts, which could render that part of the carpet useless. Knowledge of those locations can be used for targeted repairs to the machinery for future production or noted for future testing. Additional windowing tests can be performed. The size and shape of each window can be selected from past experience with the location and contours that had been encountered.

In Section 2, Poisson variables and processes and the goodness-of-fit methodology are reviewed. Section 3 contains a description of windowing. Section 4 contains the carpet example. Closing remarks are given in Section 5.

2. Poisson variables and goodness-of-fit testing

The Poisson probability distribution is covered in many introductory statistics textbooks [6, pp. 195–199], [10, pp. 341–346], [17, pp. 738–742], [20, pp. 228–235]. The case can be made that it is the central discrete distribution in the same way that the normal distribution is the central continuous distribution [19, pp. 26–27]. The distribution's random variable is a count expressed as a rate.

An important property is that the probability of the occurrence of an event is proportional to the size of the interval, area, or other increment in the denominator of the rate. This property is usually called *additivity*, but *scalability* might be a better word, because the variable scales both up and down. It is the property that is used here in the search for clusters. Sometimes, the property is expressed by saying that the events are uniformly distributed over the interval, area, or other measure.

The probability mass function of the Poisson random variable is

$$\mathbb{P}(X = x) = p(x) = p(x; \mu) = \frac{\mathrm{e}^{-\mu}\mu^x}{x!} \quad \text{for } x = 0, 1, 2, \ldots. \tag{1}$$

The positive parameter μ is the mean number of events or expected value of the random variable.

The test statistic for the chi-square goodness-of-fit test compares the expected number E of regions, which are intervals, areas, or other designated regions for the rate in the particular application, having a given count with the number O of regions observed with the same count in the data. Sometimes, the regions are called *cells*. The null hypothesis is that the data is from a particular hypothesized distribution, which gives the expected numbers. The hypothesized distribution is the Poisson distribution. The alternative hypothesis is that the distribution is not a Poisson distribution. The test statistic for the chi-square goodness-of-fit statistic is

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}, \tag{2}$$

where i indexes the (E, O) pairs and k is the number of pairs after those with very small expected values are combined, if necessary, to ensure all expected counts that are used in (2) are at least 1. The requirement that all expected frequencies be at least 1 is adopted, because of the firm theoretical foundation in [2], which is also specific to testing for the Poisson distribution. This statistic is presented in most textbooks for first courses in statistics [6, pp. 350–355], [10, pp. 631–643], [17, pp. 589–594], [20, pp. 564–613].

Statistic (2) has a chi-square distribution with degrees of freedom equal to $k - m - 1$, where m is the number of intermediate statistics that is estimated from the data in the calculation of the E_i.

One test for a single region having an unusually large count is to use the criteria that the region's count is more than three standard deviations from the mean count. Any value greater than

$$\bar{x} + 3\sqrt{\bar{x}} \tag{3}$$

is considered to be large. This formula is the upper control limit for a c control chart for data modeled with a Poisson distribution [4, pp. 697–698], [13, pp. 418–419]. Formula (3) is based on a normal-distribution approximation to a Poisson distribution [12, p. 223], so it is used only when $\mu \geq 5$.

Unusual or discrepant cells with a large number of events are those that have x-values such that $\mathbb{P}(X \geq x)$ is very small. A criterion is that the count x is unusually large if, using (1),

$$\mathbb{P}(X \geq x) \leq 0.003. \tag{4}$$

The value 0.003 was elected somewhat arbitrarily, but it gives results close to criterion (3).

The Poisson distribution and goodness-of-fit testing are illustrated with the carpet example in Section 4.

3. Windowing

Scalability of the Poisson distribution is used jointly with the spatial aspect of the observations. Events that tend to occur in neighboring regions disrupt the uniform nature of events, which is required for a Poisson variable or process. If these neighboring positions have concentrations of events in the subsequent samples, then a systematic problem with the process may have been discovered.

Windowing the array of events into new, larger regions or windows produces a new data set from the aggregated counts. The data in those new regions may or may not appear to be from a Poisson process, even if the original data had tested as being from a Poisson process. A chi-square goodness-of-fit test can be used for the new regions to see if the windowed data appear to follow a Poisson distribution. If so, the individual new regions can be examined for unusually large number of events.

Windowing is an important tool in the computer chip manufacturing process, since chips are produced in two-dimensional arrays on wafers [5], [11], [15], [16].

The distribution of pulled tufts in a piece of carpet is examined in the next section. Also, the windowing procedure is demonstrated.

4. The carpet example

Consider carpet that is made in huge rolls that are 18 feet wide. The carpet is sold relatively cheaply, and some imperfections, namely pulled tufts, are expected and are not of great concern. The distribution of these errors might be expected to have a Poisson distribution. One reason is that in the past, they have had that distribution. Another reason is that the errors are a rate per square foot or square yard.

A few pulled tufts are not harmful if they are scattered. If the mean or average number of pulled tufts per unit area is small and the distribution of pulled tufts is a Poisson distribution, then the quality engineer might believe that the pulled tufts are not a problem unless some

particular regions are flagged by criteria (3) or (4). The goal is to use statistical procedures to detect systematic problems in an environment that is not free of error or imperfections.

An 18-foot by 18-foot piece of a roll of carpet is examined. The number of pulled tufts in each 1-foot by 1-foot region appears in Table 1. This is simulated data. Each 18 foot long piece is made with the same set of parts, that is, there is a period of 18 feet in the production of the rolls. Any systematic imperfections are likely to appear periodically with period 18 feet. The particular piece of carpet has sample mean $\bar{x} = 2$ pulled tufts per square foot, which is used to estimate μ. The analysis below shows that the number of pulled tufts per square foot appears to follow a Poisson distribution with mean 2. However, two 1-foot square regions have an unusually large number of pulled tufts.

Further testing using windowing follows this initial testing. The scalability property of the Poisson distribution is invoked. It takes advantage of the geometrical configuration. Instead of $18 \cdot 18 = 324$ regions or cells, for illustrative purposes, the carpet is considered to be an array of 54 regions or cells, which are 2 feet by 3 feet. Selecting the window size can be an art and is based on past experience. Perhaps, in the past, clusters have occurred in such regions. Subsequent windowing might use other configurations of windows, as discussed in Section 4.3.

Since the number of pulled tufts in these new 6 square foot regions is hypothesized to be a Poisson process, the implied mean rate is (6 square feet)(2 pulled tufts per square foot) $= 12$ pulled tufts per 6 square feet. The chi-square goodness-of-fit test is used to check this hypothesis using the aggregated data.

In Section 4.1, the chi-square goodness-of-fit test is performed on the 324 1-foot by 1-foot regions of carpet. Each of the 324 regions is examined for unusually large numbers of pulled tufts using criterion (4). In Section 4.2, 2-foot by 3-foot windows are created and goodness-of-fit testing is performed to check whether the distribution of the number of pulled tufts might follow a Poisson distribution. Regions with large numbers of pulled tufts are then identified. Section 4.3 contains a discussion of subsequent windowing.

TABLE 1: The number of pulled tufts in each square foot of carpet.

0	4	1	4	0	4	0	1	2	6	5	6	0	1	1	5	0	1
3	1	2	2	1	1	2	3	2	3	3	4	1	1	1	0	2	0
1	2	0	4	2	4	2	0	3	1	2	1	4	0	4	1	3	1
2	3	2	2	0	2	1	2	0	2	1	4	0	0	4	2	1	2
0	1	3	3	4	2	3	2	2	3	2	1	2	6	2	1	2	1
1	3	3	3	4	1	1	2	1	1	1	1	2	1	2	2	1	4
1	2	2	2	1	1	2	4	2	4	1	2	1	3	2	1	2	2
2	1	2	1	2	1	0	2	2	4	2	4	1	1	2	1	0	2
2	2	0	4	0	2	3	4	3	2	4	2	1	3	1	2	2	2
2	0	2	4	2	0	5	0	3	1	2	2	3	1	2	1	2	0
3	1	3	0	1	2	1	1	0	5	0	4	4	2	6	4	4	3
1	3	1	2	2	4	1	0	1	4	1	2	3	2	3	3	4	1
2	1	1	3	3	1	3	2	1	0	4	2	3	3	0	1	1	1
1	2	0	3	3	4	2	1	1	2	0	4	0	2	3	1	1	1
5	1	3	2	1	3	2	1	2	2	2	0	3	2	2	1	3	2
2	1	2	1	1	2	1	2	3	2	0	2	1	3	2	1	2	1
2	0	4	2	0	0	1	2	1	5	0	1	4	3	6	1	3	2
9	1	2	3	8	0	3	1	2	0	2	4	5	4	3	3	1	1

TABLE 2: Statistics for Table 1.

Number of pulled tufts, x	Observed number of cells, O	$x \times O$	$\mathbb{P}(X = x)$ for $\mu = 2$	Expected number of cells, E	$(O - E)^2/E$
0	40	0	0.135 33	43.849	0.3379
1	91	91	0.270 67	87.697	0.1244
2	97	194	0.270 67	87.697	0.9869
3	47	141	0.180 45	58.465	2.2483
4	35	140	0.090 22	29.232	1.1381
5	7	35	0.036 09	11.693	1.8835
6	5	30	0.012 03	3.898	0.3115
7	0	0	0.004 54 (7 or more)	1.471 (7 or more)	0.1902 (7 or more)
8	1	8			
9	1	9			
Sum	324	648	1.000 00	324.002	7.2208

4.1. Initial testing

Table 2 contains the number of regions or cells for each count of the number of pulled tufts. These are the observed values. The sum of the 324 entries in Table 1 is 648, giving the sample mean $\bar{x} = \frac{648}{324} = 2$. The sample mean is used for μ to compute the probabilities for each x-value using (1). The expected values are 324 times these probabilities. The terms in the chi-square goodness-of-fit statistic (2) appear in the right-most column of Table 2, whose sum is 7.2208. Since the expected number of cells with count 7 is less than 1, the expected counts and the observed counts for 7 or more are summed, as shown in Table 2. The degrees of freedom are equal to $k - m - 1 = 8 - 1 - 1 = 6$, where $m = 1$ because the sample mean is used from the data. Since $\mathbb{P}(\chi^2 \geq 7.2208) = 0.301$, there is no evidence that this distribution is not Poisson.

Using criterion (3) is not appropriate since $\mu = 2 < 5$. Using criterion (4) with $\mu = 2$ gives $\mathbb{P}(X \geq 9) = 0.00024 \leq 0.003$ and $\mathbb{P}(X \geq 8) = 0.00109 \leq 0.003$, but $\mathbb{P}(X \geq 7) \approx 0.0045$. Referring to the first and second columns of Table 2, there is one cell each with 8 and 9 pulled tufts. Therefore, those two regions have an unusually large number of pulled tufts.

The conclusion is that there is not sufficient evidence to say that the process is not a Poisson process or out of control, but that the two cells with counts 8 and 9 should be examined. The quality engineer might expect that edges of the roll of carpet would be more susceptible to pulled tufts. The outside edges of the roll of carpet are at the top and bottom of Table 1, and the cells with 8 and 9 pulled tufts are along the bottom edge in row 18.

4.2. First windowing

Windows are created by combining cells in Table 1 into new cells that are 2 feet by 3 feet. The counts for the new cells are given in Table 3. With the combining of counts for small and large values of x, in Table 4 we give the statistics for Table 3. The chi-square statistic (2) is 18.4095 with degrees of freedom equal to $15 - 1 - 1 = 13$. Since $\mathbb{P}(\chi^2 \geq 18.4095) \approx 0.143$, there is little evidence against the null hypothesis that the windowed data in Table 3 was the outcome of a Poisson process.

Using criterion (3), $\bar{x} + 3\sqrt{\bar{x}} = 12 + 3\sqrt{12} \approx 22.39$ says that any count of 23 or more is unusually large. There is one cell with 25 pulled tufts and one cell with 27 pulled tufts, and the next largest value is 20. Using criterion (4) with $\mu = 12$ yields similar results, since

TABLE 3: Counts in the 2-foot by 3-foot regions in Table 1.

11	12	10	27	5	8
10	14	8	11	12	10
11	17	11	9	15	11
10	8	12	17	10	8
8	12	18	13	11	9
12	11	4	16	20	19
7	17	10	12	11	6
14	10	11	8	13	10
18	13	10	12	25	11

TABLE 4: Statistics for Table 3.

Number of pulled tufts, x	Observed number of cells, O	$x \times O$	$\mathbb{P}(X = x)$ for $\mu = 12$	Expected number of cells, E	$(O - E)^2/E$
$0 \le x \le 5^\star$	2	9	0.020 342	1.098 47	0.7399
6	1	6	0.025 481	1.375 97	0.1027
7	1	7	0.043 682	2.358 83	0.7828
8	6	48	0.065 523	3.538 24	1.7128
9	2	18	0.087 364	4.717 66	1.5655
10	9	90	0.104 837	5.661 20	1.9691
11	10	110	0.114 368	6.175 87	2.3679
12	7	84	0.114 368	6.175 87	0.1100
13	3	39	0.105 570	5.700 78	1.2795
14	2	28	0.090 489	4.886 41	1.7050
15	1	15	0.072 391	3.909 11	2.1649
16	1	16	0.054 293	2.931 82	1.2729
17	3	51	0.038 325	2.069 55	0.4183
18	2	36	0.025 550	1.379 70	0.2789
$19 \le x^\dagger$	4	91	0.037 417	2.020 52	1.9393
Sum	54	648	1.000 000	54.000 00	18.4095

* The two counts 4 and 5 are included.
† The four counts 19, 20, 25, and 27 are included.

$\mathbb{P}(X \ge 27) = 0.000\,134 \le 0.003$, $\mathbb{P}(X \ge 25) = 0.000\,686 \le 0.003$, and $\mathbb{P}(X \ge 20) = 0.021\,280 > 0.003$. From Table 3, the cells with counts of 25 and 27 are on the edges of the carpet. They do not contain the one-square-foot cells identified in the initial analysis.

With this windowing, the quality engineer would again decide that there is not sufficient evidence to say that the process is not a Poisson process or out of control, but the two cells with counts 25 and 27 should be monitored.

4.3. Additional windowing and the experiment-wise error rate

The original sample piece of carpet can be repeatedly windowed in many different ways. For example, the next set of windows could be 3 feet by 3 feet or 3 feet by 2 feet. Examining a piece of carpet with a number of different window sizes is like taking an object and examining it from different angles. However, there is a potential problem, which is encountered in hypothesis testing. Sensitivity to the overall type-I error rate is important. A type-I error is making the mistake of rejecting a true null hypothesis. The probability of a type-I error is designated α,

which is called the *level of significance* and is sometimes called the *false-alarm rate*, which is an appropriate name here. Choosing a too-large value of α produces many false alarms, which might cause unnecessary adjustments and shut downs of the machinery. A too small α can yield false negatives and missed effects. The value $\alpha = 0.05$ is usually the default value, but values between 0.01 and 0.10 are common.

The type-I error rate for all the tests combined is called the *experiment-wise error rate* or *family-wise error rate* [10, pp. 686–688], [20, pp. 608–609]. Performing many windowings would almost certainly lead to rejection of this being a Poisson process, even if actually it was a Poisson process, since the type-I error rates accumulate.

Suppose that r chi-square tests are to be performed, counting the original test, on the same piece of carpet, and the goal is an experiment-wise type-I error rate equal to α for the goodness-of-fit tests. If each test uses the same level of significance γ, then a formula for α is

$$\alpha = 1 - (1 - \gamma)^r. \tag{5}$$

Solving (5) for γ gives

$$\gamma = 1 - (1 - \alpha)^{1/r}.$$

An approximation of γ is α/r [10, p. 686], [14, p. 850]. For the example in Section 4, if $\alpha = 0.05$ then $\gamma \approx 0.05/2 = 0.025$ for two goodness-of-fit tests.

5. Discussion

The topic of identifying clusters is very deep and broad, even when limited to Poisson processes.

One-dimensional ordered data, such as data taken over time, are often modeled with a Poisson process. Examples are calls received per minute by a call center and clicks per minute of a radiation detector. The windowing is by subintervals.

Besides windowing, many methods for the identification of clusters are available [7]. The nonparametric one-sample runs test [3, pp. 63–66], [20, pp. 675–682] can be used to detect grouping.

There are alternative goodness-of-fit statistics to identify clusters. Some of these are relatively simple [1, pp. 337–343], [3, pp. 343–346], [18]. An interesting technique is to create a quantile–quantile plot for a Poisson distribution, which is analogous to a normal probability plot. A Poisson probability plot is described in [8] and [9].

The analysis in Section 4 is at the intersection of quality control, pattern recognition, exploratory data analysis, and hypothesis testing. The analyst is exploring the observations and looking for patterns in the data, which is not exactly the same as doing mathematical or statistical model building. In the carpet example, the quality engineer is using statistical methodology in a practical and dynamic manner for reaching a quick decision in an industrial setting.

References

[1] Cochran, W. G. (1952). The χ^2 test of goodness of fit. *Ann. Math. Statist.* **23,** 315–345.
[2] Cochran, W. G. (1954). Some methods for strengthening the χ^2 common tests. *Biometrics* **10,** 417–451.
[3] Daniel, W. W. (2000). *Applied Nonparametric Statistics*, 2nd edn. Cengage, Boston, MA.
[4] Devore, J. L. (2015). *Probability and Statistics for Engineering and the Sciences*, 9th edn. Cengage, Boston, MA.
[5] Flack, V. F. (1985). Introducing dependency into IC yield models. *Solid-State Electron.* **28,** 555–559.
[6] Freund, J. E. (2004). *Modern Elementary Statistics*, 11th edn. Pearson, Upper Saddle River, NJ.

[7] GLAZ, J. AND NAUS, J. (1983). Multiple clusters on the line. *Commun. Statist. Theory Meth.* **12,** 1961–1986.
[8] HOAGLIN, D. C. (1980). A Poissonness plot. *Amer. Statistician* **34,** 146–149.
[9] HOAGLIN, D. C. AND TUKEY, J. W. (2006). Checking the shape of discrete distributions. In *Exploring Data Tables, Trends and Shapes*, revised edn, eds D. C. Hoaglin, F. Mosteller and J. W. Tukey, John Wiley, Hoboken, NJ, pp. 345–416.
[10] LAROSE, D. T. (2016). *Discovering Statistics*, 3rd edn. Freeman, New York.
[11] LONG, M. E. AND FARNSWORTH, D. L. (2010). Modeling the random component of manufacturing yield of integrated circuits. *Internat. J. Eng. Technology* **2,** 402–405.
[12] MHATRE, S., SCHEAFFER, R. L. AND LEAVENWORTH, R. S. (1981). Acceptance control charts based on normal approximations to the Poisson distribution. *J. Quality Technology* **13,** 221–227.
[13] MONTGOMERY, D. C., RUNGER, G. C. AND HUBELE, N. F. (2004). *Engineering Statistics*, 3rd edn. John Wiley, New York.
[14] ROHATGI, V. K. (2003). *Statistical Inference*. Dover, Mineola, NY.
[15] STAPPER, C. H. (1989). Large-area fault clusters and fault tolerance in VLSI circuits: a review. *IBM J. Res. Develop.* **33,** 162–173.
[16] STAPPER, C. H., ARMSTRONG, F. M. AND SAJI, K. (1983). Integrated circuit yield statistics. *Proc. IEEE* **71,** 453–470.
[17] SULLIVAN, M., III (2013). *Statistics: Informed Decisions Using Data*, 4th edn. Pearson, Boston, MA.
[18] TALLIS, G. M. (1983). Goodness of fit. In *Encyclopedia of Statistical Sciences*, Vol. 3, eds S. Kotz, N. L. Johnson, and C. B. Read, John Wiley, New York, pp. 451–461.
[19] TAYLOR, H. M. AND KARLIN, S. (1998). *An Introduction to Stochastic Modeling*, 3rd edn. Academic Press, San Diego, CA.
[20] TRIOLA, M. F. (2014). *Elementary Statistics*, 12th edn. Pearson, Boston, MA.