Journal of Quality Technology

A Quarterly Journal of Methods, Applications and Related Topics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

Monitoring proportions with two components of common cause variation

Rob Goedhart & William H. Woodall

To cite this article: Rob Goedhart & William H. Woodall (2022) Monitoring proportions with two components of common cause variation, Journal of Quality Technology, 54:3, 324-337, DOI: 10.1080/00224065.2021.1903823

To link to this article: https://doi.org/10.1080/00224065.2021.1903823

Published online: 14 Apr 2021.

JOURNAL OF QUALITY TECHNOLOGY
2022, VOL. 54, NO. 3, 324–337
https://doi.org/10.1080/00224065.2021.1903823

Monitoring proportions with two components of common cause variation

Rob Goedhart[a] and William H. Woodall[b]

[a]Department of Operations Management, University of Amsterdam, Amsterdam, Netherlands; [b]Department of Statistics, Virginia Tech, Blacksburg, Virginia

ABSTRACT

We propose a method for monitoring proportions when the in-control proportion and the sample sizes vary over time. Our approach is able to overcome some of the performance issues of other commonly used methods, as we demonstrate in this paper using analytical and numerical methods. The derivations and results are shown mainly for monitoring proportions, but we show how the method can be extended to the monitoring of count data.

KEYWORDS

Attribute control chart; control charts; Laney method; *p*-chart; statistical process monitoring

1. Introduction

Control charts are a common tool to monitor processes. When it comes to attribute data, the *p*- and *np*-charts are widely used for monitoring proportions such as the fraction of non-conforming products, while other charts such as the *c*- and *u*-charts are used to monitor counts such as the number of non-conformities. See, for example, Woodall (1997) for a bibliography and review on control charts for attributes.

The basic available control charts for attributes are based on either the binomial or the Poisson distribution with the assumption of a constant in-control parameter for the mean. The corresponding classical control limits are then determined by the expected sampling variation only. If common cause variability is present between subgroups, these control limits could be very misleading, as addressed by Alwan and Roberts (1995), Heimann (1996), Laney (2002) and Woodall (1997), among others. This issue is more relevant when sample sizes are large, because then the sampling variation diminishes and the control limits move toward the center line, resulting in misleading out-of-control signals. For small sample sizes, the sampling variation can be relatively large compared to the variation between subgroups, making this less of an issue.

In practice, sample sizes can be quite large. The application in Heimann (1996) involved monitoring failure rates of circuits in a business telephone service maintenance process. The average sample sizes were in the thousands and hundreds of thousands. Hagan and Li (2018) considered a pharmacy application where the weekly proportions of narcotic prescriptions involving hydrocodone were of interest and the sample sizes varied from 550 to 751. Mohammed and Laney (2006) considered a healthcare application in which hospital readmission rates were of interest and sample sizes ranged up to 100,000. Laney (2002) mentioned that in some service industries sample sizes can even run into the millions.

Laney (2002) proposed adjusted versions of the *p*- and *u*-charts that aim to handle the case of varying subgroup sizes when there is some common cause variation between subgroups. The adjusted charts were named the p'- and u'-chart. Laney's methods have been widely accepted. His charts have been included in several textbooks, including Crossley (2007), Provost and Murray (2011), and Montgomery (2020). As pointed out by Vidmar and Blagus (2014), his approach has also been incorporated into the following software packages: CHARTrunner 3.6, October 2009; WinChart Professional 4, February 2010; QI Macros, October 2010; Minitab 16.2, October 2011; and SigmaXL 6.1, December 2011. In addition, a control chart selection flowchart in Suman and Prajapati (2018) directed practitioners to the Laney attribute methods whenever the sample sizes were very large.

In our paper we show that the performance of the method of Laney (2002) deteriorates when subgroup sizes vary widely. In addition, we present an

CONTACT Rob Goedhart R.Goedhart2@uva.nl Department of Operations Management, University of Amsterdam, Amsterdam, Netherlands.

alternative estimation method that does not lead to this problem, and demonstrate its good performance in a variety of settings. We focus primarily on the monitoring of proportions, but our method can easily be applied to monitoring counts as well.

This paper is organized as follows. In the next section we provide a motivational example to illustrate the performance issues of the p-chart. In Section 3 we elaborate on a more general model for variation and the p'-chart from Laney (2002). In Section 4 we explain our proposed method, and illustrate its performance through a simulation study. In Section 5 we provide a discussion as well as an illustrative comparison between our proposed method and the one of Laney (2002). In Section 6 we briefly discuss the extension to the Poisson case. Finally, in Section 7 we provide some concluding remarks.

2. Motivation and background

The p-chart is used to detect special causes by monitoring the fraction of nonconforming products. An important assumption for the construction of this chart is that the data should follow a binomial distribution with a constant in-control parameter. While violation of this assumption does not necessarily lead to issues directly, performance issues of the chart become more noticeable if sample sizes are very large, as recognized by Alwan and Roberts (1995), Heimann (1996), Laney (2002) and Woodall (1997), among others. As an illustrative example we consider nonconformance data from 20 consecutive weeks of hospital emergency department data records, as available in Table 7.7 of Montgomery (2020) and displayed in our Table 1. In Figure 1 the fraction nonconforming is plotted, along with the classic p-chart control limits (see e.g. Montgomery (2020)).

As can be observed, the control limits are very narrow compared to the variation in the data. Such narrow control limits, especially for larger sample sizes, are a likely consequence of overdispersion, i.e. more variability in the data than is predicted by the binomial distribution (see e.g. Montgomery (2020)). This additional variability is in turn a consequence of many (small) influences varying between samples, i.e. between-subgroup (also known as inter-subgroup) common cause variation.

One possible approach to deal with this effect is to treat the proportions as individual variables, as suggested for example by the Western Electric Company (1956) and Heimann (1996). However, one must be careful not to over-aggregate the data over long

Table 1. Nonconformance data from 20 consecutive weeks of hospital emergency data from Montgomery (2020).

Week	Sample size	Errors	Fraction nonconforming
1	2500	187	0.0748
2	3000	345	0.1150
3	2245	210	0.0935
4	2900	185	0.0638
5	3650	376	0.1030
6	3119	412	0.1321
7	2415	241	0.0998
8	1985	156	0.0786
9	2430	200	0.0823
10	3620	412	0.1138
11	2765	254	0.0919
12	3800	275	0.0724
13	2600	185	0.0712
14	1875	210	0.1120
15	3125	298	0.0954
16	3900	450	0.1154
17	3850	325	0.0844
18	2350	256	0.1089
19	2145	198	0.0923
20	3450	300	0.0870

aggregation intervals, e.g., using weekly data as opposed to daily or hourly data. As discussed by Zwetsloot and Woodall (2021), over-aggregation can lead to the masking of the effects of assignable causes and loss of information about the process. This issue was recognized in the Western Electric Company (1956) referring to the use of sample sizes of hundreds or thousands for monitoring proportions:

> *"When large amounts of data are combined on an overall p-chart, these data may include many different variations in raw material, part numbers, code numbers, processing batches or inspection lots. These are assignable causes that cannot be studied properly except with individual p-charts at various operations. Super-imposed on these ordinary types of causes, however, may be large general shifts of trends that affect the whole shop. The overall p-chart can be used to study these broad shifts and trends by calculating its control limits… "*

Their recommended method for determining the control limits was to use the control chart for individuals with the estimate of variability based on the average moving range. This often recommended approach, however, does not take into account varying sample sizes and cannot be used to separate the sampling variation from the variation in the in-control parameter value.

To that end, Laney (2002) proposed an alternative method for modifying the control limits of the p-chart to account for overdispersion, while allowing varying subgroup sizes. His approach introduces a multiplicative correction of the estimated standard deviation, based on the relative amount of process variation not explained by the binomial assumption alone (a detailed description follows in Section 3). The chart resulting from this approach is often referred to as the

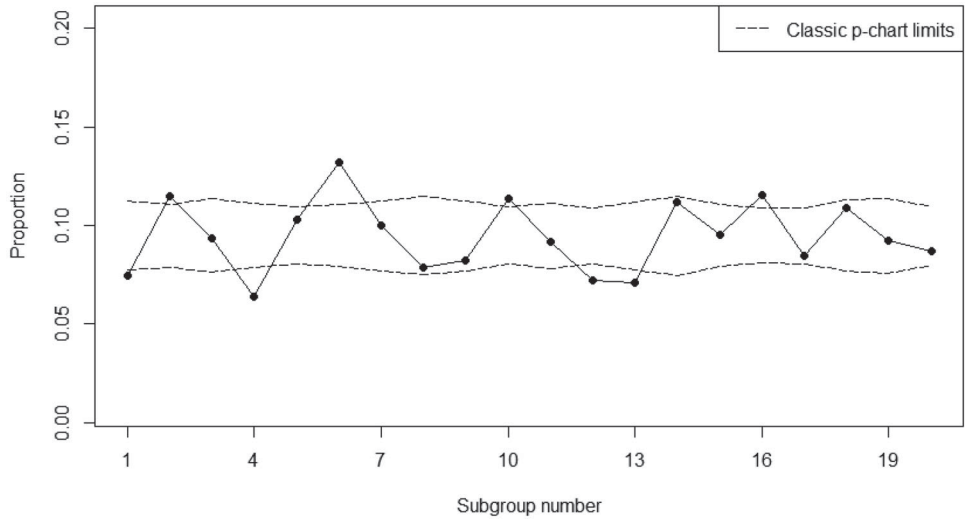


Figure 1. Fraction of nonconforming in hospital emergency department data records using classic p-chart limits.

p'-chart. In Figure 2 we illustrate the control limits obtained using the method of Laney (2002) for the dataset in Table 1. At first glance, the estimated control limits appear to reflect a much more suitable estimation of common cause variation compared to the p-chart limits.

In Montgomery (2020) it is shown that when applying the method of Laney (2002) for this dataset, the estimated standard deviation for each subgroup in the p-chart has to be multiplied by a factor 3.87 to account for the variability not explained by the binomial distribution alone. In terms of the variance, this means a multiplication by 14.98 ($= 3.87^2$). In other words, with the method of Laney (2002) it is estimated that the total variance is 14.98 times larger than the intra-subgroup (or within-subgroup) variance. That means that the intra-subgroup variance is estimated to be just 6.7% of the total variance. Thus, it would be expected that the differences between control limits for different subgroups would be very small as well. However, when taking a closer look at the estimated control limits in Figure 2, we observe substantial variation in the limits still. In particular, consider for each subgroup the distance between the upper control limit and the center value of the chart. It turns out that the largest distance (subgroup 14) is 44% larger than the smallest distance (subgroup 16). This is rather surprising given the small impact (6.7%) of the intra-subgroup variation estimated earlier. In the next section we elaborate in more detail on the causes for this phenomenon with the method of Laney (2002), and the problems it can cause for the control chart performance when sample sizes vary.

3. General model and Laney method

In this section we explain the general model used in this paper, which includes the possibility of between-subgroup common cause variation in addition to sampling variation. Next, we describe the the p-chart and Laney's p'-chart in detail. Finally, we illustrate the performance of the latter method for various settings. Note that in this paper we use a slightly adjusted notation compared to Laney (2002) to make the distinction between estimators and known parameters more clear.

3.1. General model

We assume that in Phase I there are m independent subgroups available, where for each subgroup i the number of occurrences of the attribute of interest (X_i) is drawn from a binomial distribution with parameter p_i. The corresponding subgroup size is denoted as n_i. Note that we do not assume p_i to be constant over i in the in-control state. Instead, we consider p_i to be a random variable itself with mean p_0 and variance σ_p^2. This allows for the presence of some common cause inter-subgroup variation. A common estimator of p_i is $\hat{p}_i = X_i/n_i$. Note that the variation in $\hat{p}_i$ values is then caused by two factors: the variation in p_i values (inter) and sampling variation (intra). This approach is similar to the approach of Woodall and Thomas (1995), who described an estimation method for Shewhart $\bar{X}$ control charts based on several components of common cause variability. Some other similar methods can be found in Chang and Gan (2004) and Yashchin (1994).

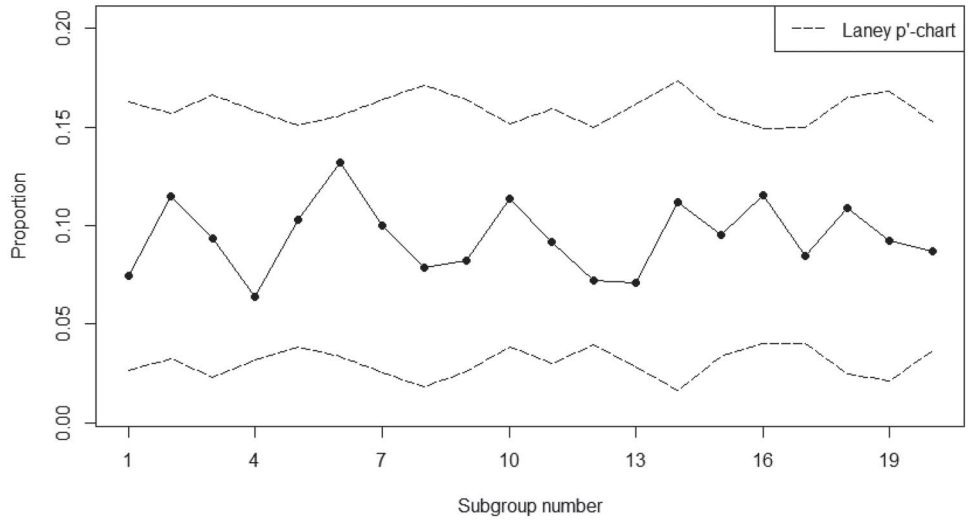


Figure 2. Fraction of nonconforming in hospital emergency department data records using the p'-chart limits of the method of Laney (2002).

For $X_i \sim Bin(n_i, p_i)$ we know that $E[X_i|p_i] = n_i p_i$ and $Var(X_i|p_i) = n_i p_i(1-p_i)$, so that $E[\hat{p}_i|p_i] = p_i$ and $Var(\hat{p}_i|p_i) = \frac{p_i(1-p_i)}{n_i}$. Then, using the law of total expectation and the law of total variance we find that $E[\hat{p}_i] = p_0$ and

$$\begin{aligned} Var(\hat{p}_i) &= \sigma^2_{\hat{p}_i} \\ &= E\left[\frac{p_i(1-p_i)}{n_i}\right] + Var(p_i) \\ &= \frac{1}{n_i}\left(E[p_i] - E\left[p_i^2\right]\right) + \sigma_p^2 \\ &= \frac{1}{n_i}\left(p_0 - \sigma_p^2 - p_0^2\right) + \sigma_p^2 \\ &= \frac{p_0(1-p_0)}{n_i} + \sigma_p^2\left(1 - \frac{1}{n_i}\right) \\ &= \sigma^2_{W,i} + \sigma^2_{B,i} \end{aligned} \tag{1}$$

where $\sigma^2_{W,i} = \frac{p_0(1-p_0)}{n_i}$ and $\sigma^2_{B,i} = \sigma_p^2\left(1 - \frac{1}{n_i}\right)$. Here $\sigma^2_{W,i}$ represents the within (intra) subgroup variation if no between (inter) subgroup variation were present (i.e., $p_i = p_0$ for all i), while $\sigma^2_{B,i}$ represents the additional variance due to the inter-subgroup variation in p_i values. These two components determine the total variance of $\hat{p}_i$ in our model. Note that $\sigma^2_{\hat{p}_i}$ varies over i due to varying sample sizes.

In the classic p-chart it is assumed that in the in-control state all subgroups have the same parameter $p_i = p_0$, such that $\sigma_p^2 = 0$. Common estimators of p_0 and $\sigma^2_{\hat{p}_i}$ in that case are $\hat{p}_0 = \frac{\sum_{i=1}^m X_i}{\sum_{i=1}^m n_i}$ and $\hat{\sigma}^2_{\hat{p}_i} = \hat{\sigma}^2_{W,i} = \frac{\hat{p}_0(1-\hat{p}_0)}{n_i}$ respectively, which are also used by Laney (2002). The upper control limit (UCL) and lower control limit (LCL) of the p-chart are typically set at

$$\hat{p}_0 + k\hat{\sigma}_{\hat{p}_i} \text{ and } \hat{p}_0 - k\hat{\sigma}_{\hat{p}_i}, \tag{2}$$

respectively, with k some positive constant most often set to $k = 3$.

3.2. Laney's p'-chart

Because the p-chart is based on the assumption of a constant in-control parameter p_0, it follows that all the variation is assumed to be caused by intra-subgroup variation. As argued by Laney (2002), this assumption is seldom true in applications. Violation of this assumption becomes a bigger problem when subgroup sizes are very large, as this lowers the sampling variation toward zero. Because the used estimated control limits are based on sampling variation only, they will move close to the center line. In that situation, even minor inter-subgroup variation will lead to an abundance of misleading out-of-control signals due to these narrow control limits. This was the motivation for Laney (2002) to propose adjusted control limits, which incorporate inter-subgroup variation as well as intra-subgroup variation. Advantages of his method are also discussed in Mohammed and Laney (2006) and Mohammed et al. (2013).

3.2.1. Laney's adjustments

Consider $\hat{p}_0 = \frac{\sum_{i=1}^m X_i}{\sum_{i=1}^m n_i}$ and $\hat{\sigma}^2_{W,i} = \frac{\hat{p}_0(1-\hat{p}_0)}{n_i}$ as estimates of p_0 and $\sigma^2_{W,i}$, respectively, and consider $\hat{p}_i = X_i/n_i$. In order to make the $\hat{p}_i$ values more comparable (due to different subgroup sample sizes n_i, some samples will have more sampling uncertainty than others), Laney (2002) stated that one should "adjust each point

for its unique intra-subgroup variation". A z-score is calculated as

$$z_i = \frac{\hat{p}_i - \hat{p}_0}{\hat{\sigma}_{W,i}}. \quad (3)$$

Because the assumption of zero mean and unit variance for the z_i values relies on the assumption of no inter-subgroup variation (which may not be a reasonable assumption), Laney (2002) proposed to estimate the actual variation in z_i values. This was done by calculating the average moving range (AMR) of the z_i's to obtain the estimator

$$\hat{\sigma}_z = AMR(z_i)/1.128. \quad (4)$$

As final step, Laney (2002) argued that the z_i transformation should be unraveled back into the p-plane. This was done by expressing $\hat{p}_i$ in terms of the components $\hat{p}_0, \hat{\sigma}_{W,i}$ and z_i, which leads to $\hat{\sigma}_{\hat{p}_i} = \hat{\sigma}_z \hat{\sigma}_{W,i}$ as an estimator of the standard deviation of $\hat{p}_i$. Consequently, the control limits for the p'-chart at time i are constructed according to

$$\begin{aligned} UCL_i &= \hat{p}_0 + 3\hat{\sigma}_z \hat{\sigma}_{W,i} \\ LCL_i &= \hat{p}_0 - 3\hat{\sigma}_z \hat{\sigma}_{W,i}. \end{aligned} \quad (5)$$

In our view, this final step actually seems to rely on the first step, as it incorporates the rationale behind the choice of z_i. The factor $\hat{\sigma}_z$ exists to compensate for the violation of the assumption of a constant in-control value for p_i in the p-chart. Note that for the example p'-chart illustrated in Figure 2 $\hat{\sigma}_z = 3.87$.

The method of Laney (2002) is based on the assumption of a constant variance of the z_i values, namely $\sigma_{z_i} = \sigma_z$ for all i. In order to evaluate the appropriateness of this assumption, first consider replacing $\hat{p}_0$ and $\hat{\sigma}_{W,i}$ by their assumed known parameter values p_0 and $\sigma_{W,i}$, respectively, in Eq. [3] to obtain

$$z_i^* = \frac{\hat{p}_i - p_0}{\sigma_{W,i}}. \quad (6)$$

From Eq. [1] it follows that z_i^* has expectation $E[z_i^*] = 0$ and variance

$$\begin{aligned} \sigma^2_{z_i^*} &= 1 + \frac{\sigma^2_{B,i}}{\sigma^2_{W,i}} = 1 + \frac{\sigma_p^2}{\sigma^2_{W,i}}\left(1 - \frac{1}{n_i}\right) \\ &= 1 + \frac{\sigma_p^2}{p_0(1-p_0)}(n_i - 1). \end{aligned} \quad (7)$$

This variation of z_i^* is not constant over i, but increases as n_i increases. Note that when there is no inter-subgroup variation (i.e., $\sigma_p^2 = 0$), this is not an issue because the z_i^* values then have (constant) unit variance. However, when $\sigma_p^2 \neq 0$, it is clear that the z_i^* values do not have constant variance when n_i varies, which is implicitly assumed in the approach of Laney (2002).

3.2.2. Laney's p′-chart performance

To illustrate the non-constant variance of z_i values explicitly, we performed a simulation study. In particular, we consider a Phase I sample consisting of $m = 100$ subgroups of size n_i each. In our simulations, we consider a simplified example where there are (at most) two different values for n_i in order to illustrate the effect of the subgroup size. We consider the p_i values to be drawn from a uniform distribution $U(p_0 - r_p, p_0 + r_p)$, where r_p represents the "radius" of p_i. This implies that $\sigma_p = \frac{2r_p}{\sqrt{12}}$. Note that other distributions of p_i may also be used, but this is not of importance in the illustration considered here. The first $m/2$ subgroups ($i = 1, \ldots, 50$) have sample size N_1, while the second $m/2$ subgroups ($i = 51, \ldots, 100$) have sample size N_2. Then, for $p_0 = 0.1$ and various combinations of N_1, N_2, and r_p, we applied the following simulation procedure:

1. Draw m independent values of p_i from a uniform distribution $U(p_0 - r_p, p_0 + r_p)$
2. Use the obtained p_i values to draw m independent values of X_i from a binomial distribution with parameters n_i and p_i and calculate the corresponding $\hat{p}_i = X_i/n_i$. Note that $n_i = N_1$ for $i = 1, \ldots, m/2$ and $n_i = N_2$ for $i = m/2 + 1, \ldots, m$.
3. Calculate all m values of z_i according to Eq. [3].
4. Calculate $\hat{\sigma}_z$ according to Eq. [4] for …
   4(a). all m values of z_i
   4(b). the first $m/2$ values of z_i (i.e., where $n_i = N_1$).
   4(c).the last $m/2$ values of z_i (i.e., where $n_i = N_2$).
5. Repeat steps 1 to 4 for 10,000 times, and calculate the averages of the values from 4(a), 4(b) and 4(c).

The results of the simulation procedure are shown in Table 2 for the cases $r_p = 0, 0.025$ and 0.05. Because exchanging N_1 and N_2 leads to equivalent results, Table 2 shows only unique combinations of these two (e.g., $N_1 = 500$ and $N_2 = 100$ yields equivalent results to $N_1 = 100$ and $N_2 = 500$). In the case that $r_p = 0$, it can be seen that the z_i values have constant unit variance regardless of subgroup sizes. When $r_p \neq 0$, this is no longer the case. As long as subgroup sizes are constant (i.e., $N_1 = N_2$), z_i values will have constant variance, which means that the method of Laney (2002) works well in estimating the variance of the $\hat{p}_i$ values. However, when the difference in n_i values becomes larger, the assumption of constant variance of the z_i values is no longer appropriate. This deviation from a constant variance also becomes

Table 2. Average of $\hat{\sigma}_z$ values of the method of Laney (2002) for $p_0 = 0.1$ and various combinations of r_p, N_1, and N_2.

Average of $\hat{\sigma}_z$ values										
		$r_p = 0$			$r_p = 0.025$			$r_p = 0.05$		
N_1	N_2	all n_i	$n_i = N_1$	$n_i = N_2$	all n_i	$n_i = N_1$	$n_i = N_2$	all n_i	$n_i = N_1$	$n_i = N_2$
100	100	1.00	1.00	1.00	1.10	1.11	1.10	1.39	1.39	1.39
100	500	1.00	1.00	1.00	1.29	1.10	1.48	1.90	1.39	2.41
100	1,000	1.00	0.99	1.00	1.48	1.11	1.85	2.33	1.39	3.27
100	5,000	1.00	1.00	1.00	2.36	1.10	3.62	4.22	1.39	7.03
100	10,000	1.00	1.00	1.00	3.07	1.11	5.03	5.66	1.39	9.90
500	500	1.00	1.00	1.00	1.48	1.48	1.48	2.41	2.40	2.42
500	1,000	1.00	1.00	1.00	1.66	1.48	1.84	2.84	2.41	3.27
500	5,000	1.00	1.00	1.00	2.56	1.48	3.63	4.74	2.42	7.04
500	10,000	1.00	1.00	1.00	3.25	1.48	5.02	6.17	2.42	9.90
1,000	1,000	1.00	1.00	1.00	1.84	1.84	1.84	3.27	3.27	3.28
1,000	5,000	1.00	1.00	1.00	2.74	1.85	3.63	5.16	3.28	7.03
1,000	10,000	1.00	1.00	1.00	3.44	1.84	5.03	6.60	3.27	9.90
5,000	5,000	1.00	1.00	1.00	3.62	3.62	3.62	7.05	7.06	7.04
5,000	10,000	1.00	1.00	1.00	4.32	3.62	5.02	8.47	7.04	9.90
10,000	10,000	1.00	1.00	1.00	5.02	5.03	5.02	9.90	9.90	9.91

larger when there is more inter-subgroup variation present (i.e., as r_p becomes larger).

The consequence is that estimated control limits for the p'-chart will be too wide for smaller samples, and too narrow for larger samples. For example, consider the case that $N_1 = 100$ and $N_2 = 10,000$ for $r_p = 0.05$. In that situation, the average estimate of σ_z from the method of Laney (2002) equals 5.66. This implies that one should multiply $\hat{\sigma}_{W,i}$ by 5.66 for each subgroup i to calculate the control limits according to Eq. [5]. However, from Table 2 we also observe that the actual variation of z_i values is very different for the subgroups with $n_i = N_1$ compared to the ones with $n_i = N_2$. In this situation, the subgroups with $n_i = N_1$ would only require a multiplication of $\hat{\sigma}_{W,i}$ with around 1.39 instead of 5.66, such that using the latter results in control limits that are far too wide. At the same time, the subgroups with $n_i = N_2$ require a multiplication of $\hat{\sigma}_{W,i}$ with around 9.90, such that 5.66 is not sufficient, and control limits will be too narrow.

The issues with the Laney (2002) approach are a result of modeling a correction for inter-subgroup variation in a multiplicative manner relative to the intra-subgroup variation. For small subgroup sample sizes, one would expect intra-subgroup variation to be more dominant compared to when subgroup sample sizes become very large. In the method of Laney (2002) however, the adjustment based on multiplication of $\hat{\sigma}_{W,i}$ with $\hat{\sigma}_z$ leads to standard deviation estimates with an equal proportion of estimated inter-subgroup variation, regardless of sample size.

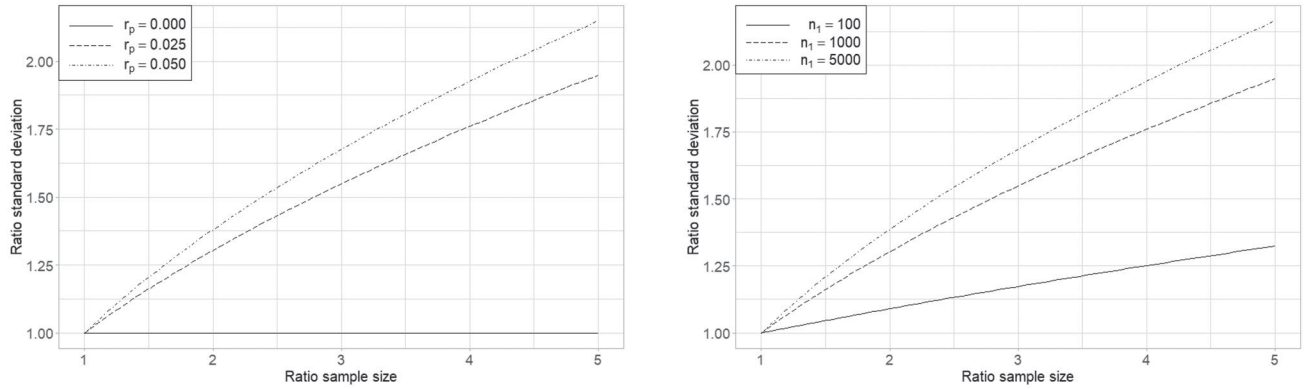
As an additional illustration, consider $\sigma_{z_i^*}$, the standard deviation of the z_i^* values, as can be obtained from Eq. [7]. For two fixed values n_1 and n_2 it is possible to determine the ratio of standard deviations $\sigma_{z_2^*}/\sigma_{z_1^*}$. In Figure 3 this ratio is plotted on the vertical axes, against the ratio of sample sizes n_2/n_1 on the horizontal axes, for various combinations of r_p and n_1, and $p_0 = 0.1$. As can be observed, the ratio $\sigma_{z_2^*}/\sigma_{z_1^*}$ is far from one when sample sizes vary and inter-subgroup variation is present. This deviation is larger the further away the ratio of sample sizes is from one. Also, the deviation is larger when more inter-subgroup variation is present (larger values of r_p), or when sample sizes are larger (i.e. for the same sample size ratio, but a larger absolute difference between n_2 and n_1). The method of Laney (2002) only leads to appropriate limits when there is no inter-subgroup variation present ($r_p = 0$). This indicates the need for an alternative way to adjust control limits for control charts for proportions when a practitioner believes that there is inter-subgroup common cause variation.

4. New proposal

We propose adjusting for inter-subgroup variation in an additive manner, rather than a multiplicative manner. In this section we provide the required calculations for our proposed method, and illustrate its performance by means of a simulation study.

4.1. Estimation of the control limits

Recall from Eq. [1] that $\hat{p}_i$ has expectation zero and variance $\sigma^2_{\hat{p}_i} = \sigma^2_{W,i} + \sigma^2_{B,i} = \frac{p_0(1-p_0)}{n_i} + \sigma^2_p\left(1 - \frac{1}{n_i}\right)$, where it is important to note that $\sigma^2_{\hat{p}_i}$ varies over i because of varying sample sizes. Consider $d_i = \hat{p}_i - \hat{p}_{i-1}$, which has expectation $E[d_i] = 0$ and variance $\sigma^2_{d_i} = \sigma^2_{W,i} + \sigma^2_{W,i-1} + \sigma^2_{B,i} + \sigma^2_{B,i-1} = p_0(1 - p_0)\left(\frac{1}{n_i} + \frac{1}{n_{i-1}}\right) + \sigma^2_p\left(2 - \frac{1}{n_1} - \frac{1}{n_{i-1}}\right)$, because $\hat{p}_i$ and $\hat{p}_{i-1}$



(a) $n_1 = 1{,}000$, varying r_p (b) $r_p = 0.025$, varying n_1

Figure 3. Ratio of standard deviations $\sigma_{z_2^*}/\sigma_{z_1^*}$ (vertical axis) versus ratio of sample sizes n_2/n_1 (horizontal axis) for $p_0 = 0.1$ and various values of n_1 and r_p.

are assumed to be independent. Next, consider the case that m is even. We define an adjusted version of the mean square successive differences (cf. Roes, Does, and Schurink 1993) as

$$MSD^* = \frac{1}{m}\sum_{i=1}^{m/2} d_{2i}^2 = \frac{1}{m}\sum_{i=1}^{m/2}\left(\hat{p}_{2i} - \hat{p}_{2i-1}\right)^2, \tag{8}$$

which is based on differences of non-overlapping sets of two successive observations. This ensures that the d_i^2 terms in the sum are independent. Next, since $E[d_i^2] = \sigma_{d_i}^2 + (E[d_i])^2 = \sigma_{d_i}^2$, it can then be determined that

$$\begin{aligned} E[MSD^*] &= \frac{1}{m}\sum_{i=1}^{m/2} E\left[d_{2i}^2\right] = \frac{1}{m}\sum_{i=1}^{m}\sigma_{B,i}^2 + \frac{1}{m}\sum_{i=1}^{m}\sigma_{W,i}^2 \\ &= \sigma_p^2\left(1 - \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}\right) + \frac{1}{m}\sum_{i=1}^{m}\frac{p_0(1-p_0)}{n_i}. \end{aligned} \tag{9}$$

We can then estimate σ_p^2 by

$$\hat{\sigma}_p^2 = \frac{\frac{1}{m}\sum_{i=1}^{m/2}\left(\hat{p}_{2i} - \hat{p}_{2i-1}\right)^2 - \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{p}_0(1-\hat{p}_0)}{n_i}}{1 - \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}}. \tag{10}$$

Note that there is a possibility for this estimator to become negative, in which case we recommend to use $\hat{\sigma}_p^2 = 0$ instead. The obtained estimate of the inter-subgroup variation can be used to provide an estimate $\hat{\sigma}_{\hat{p}_i}^2$ of the variance for any subgroup i using

$$\hat{\sigma}_{\hat{p}_i}^2 = \hat{\sigma}_{W,i}^2 + \hat{\sigma}_{B,i}^2 = \frac{\hat{p}_0(1-\hat{p}_0)}{n_i} + \hat{\sigma}_p^2\left(1 - \frac{1}{n_i}\right). \tag{11}$$

The estimate from Eq. [11] can be used to calculate the required control limits using Eq. [2].

4.2. Performance evaluation

In order to demonstrate the performance of the proposed estimation method, we apply a similar simulation method as in Section 3.2.2 with the same combinations of N_1, N_2 and r_p. In particular, the applied simulation procedure is as follows:

1. Draw m independent values of p_i from a uniform distribution $U(p_0 - r_p, p_0 + r_p)$
2. Use the obtained p_i values to draw m independent values of X_i from a binomial distribution with parameters n_i and p_i and calculate the corresponding $\hat{p}_i = X_i/n_i$. Note that $n_i = N_1$ for $i = 1, \ldots, m/2$ and $n_i = N_2$ for $i = m/2 + 1, \ldots, m$.
3. Calculate the value of $\hat{\sigma}_p^2$ and the resulting values of $\hat{\sigma}_{\hat{p}_i}^2$ for each subgroup i, as described in Section 4.1.
4. Calculate the ratio $Y_i = \hat{\sigma}_{\hat{p}_i}/\sigma_{\hat{p}_i}$ for all m subgroups, with $\sigma_{\hat{p}_i}$ as in Eq. [1].
5. Calculate the average of …
   5(a). the first $m/2$ values of Y_i (i.e., where $n_i = N_1$).
   5(b). the last $m/2$ values of Y_i (i.e., where $n_i = N_2$).
6. Repeat steps 1 to 4 for 10,000 times, and calculate the averages of the resulting values from 5(a) and 5(b).

The results of the simulation procedure are shown in Table 3 for the cases $r_p = 0, 0.025$, and 0.05. It can be observed that when there is some inter-subgroup variation present ($r_p = 0.025$ and $r_p = 0.05$), there is little to no bias in the estimated standard deviations, as the average ratios Y_i are close to 1 for any combination of N_1 and N_2. Our method provides a fairly unbiased estimate of $\sigma_{\hat{p}_i}$ in all cases.

In the special case that there is absolutely no inter-subgroup variation ($r_p = 0$), there can be some bias in

Table 3. Average of Y_i values of the proposed method for $p_0 = 0.1$ and various combinations of r_p, N_1, and N_2.

Average of Y_i values							
		$r_p = 0$		$r_p = 0.025$		$r_p = 0.05$	
N_1	N_2	$n_i = N_1$	$n_i = N_2$	$n_i = N_1$	$n_i = N_2$	$n_i = N_1$	$n_i = N_2$
100	100	1.04	1.04	1.01	1.01	1.00	1.00
100	500	1.03	1.12	1.00	0.99	1.00	0.99
100	1,000	1.03	1.21	1.00	0.98	1.00	0.99
100	5,000	1.03	1.70	1.00	0.95	1.00	0.99
100	10,000	1.03	2.11	1.00	0.94	1.00	0.99
500	500	1.04	1.04	1.00	1.00	0.99	0.99
500	1,000	1.03	1.06	1.00	0.99	1.00	1.00
500	5,000	1.03	1.21	1.00	0.99	1.00	1.00
500	10,000	1.03	1.36	1.00	0.99	1.00	1.00
1,000	1,000	1.04	1.04	0.99	0.99	1.00	1.00
1,000	5,000	1.03	1.12	1.00	1.00	1.00	1.00
1,000	10,000	1.03	1.21	1.00	0.99	1.00	1.00
5,000	5,000	1.04	1.04	1.00	1.00	1.00	1.00
5,000	10,000	1.03	1.06	1.00	1.00	1.00	1.00
10,000	10,000	1.04	1.04	1.00	1.00	1.00	1.00

the proposed estimators. This bias is mostly present for widely varying sample sizes, such as the situation where $N_1 = 100$ and $N_2 = 10,000$. Note that this is a potential consequence of requiring $\hat{\sigma}_p$ to be non-negative. The larger intra-subgroup variation of the smaller subgroups (e.g., $N_1 = 100$) can sometimes lead to positive estimates from $\hat{\sigma}_p$. Due to the small intra-subgroup variation of the larger subgroups (e.g. $N_2 = 10,000$), this can lead to a positive bias in the estimation when $\sigma_p = 0$. The standard p-chart would be used, however, if there were no inter-subgroup variation.

5. Comparison and discussion

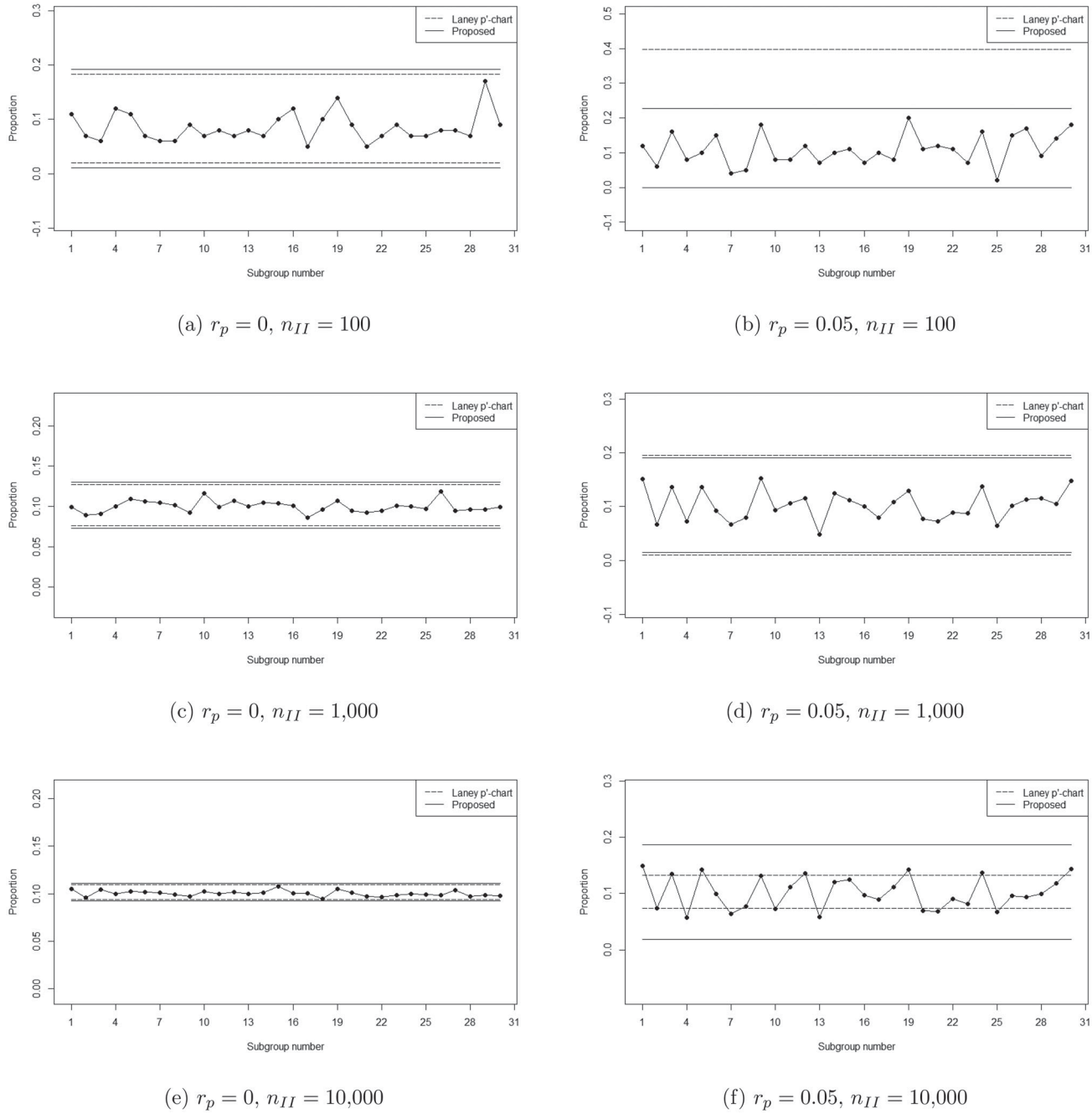
As observed in Section 3.2.2, the p'-charts as proposed by Laney (2002) can have performance issues when subgroup sizes vary substantially. To overcome this issue, we have proposed an alternative control chart design that accounts for inter-subgroup variation in Section 4. Our proposed method is able to handle varying subgroup sizes well. To illustrate the advantage of the newly proposed method, we simulated data in various scenarios and compared the control limits of both methods. We also evaluated and compared the false alarm rate (FAR) by means of simulation study.

5.1. Comparison of control limits

To illustrate the difference in control limits, we simulated Phase I samples consisting of $m = 100$ subgroups of size $n_i = n_I = 1,000$ each ($i = 1, ..., 100$), with $p_0 = 0.1$ and either no ($r_p = 0$) or substantial ($r_p = 0.05$) inter-subgroup variation present. These data were used for estimating the control limits for both methods. We considered the same uniform distribution as in the simulation procedures from the previous sections, and used $k = 3$ in Eq. [2] for both methods. In Phase II, we simulated $m = 30$ subgroups of constant size $n_i = n_{II}$ each ($i = 101, ..., 130$), with several values of n_{II} considered in the simulations. In particular, we consider the values 100, 1,000 and 10,000 for n_{II}, such that the situations $n_I > n_{II}, n_I = n_{II}$, and $n_I < n_{II}$ are all illustrated. The results are shown in Figure 4. Note that the dashed control limits represent the results of the method of Laney (2002), and the solid control limits represent the results of our method.

In the situation that there is zero inter-subgroup variation (Figures 4a, c, and e), it can be observed that there is not much difference in control limits between the two methods. The difference depends slightly on the Phase I sample, but other simulated Phase I samples have led to similar results. However, when there is substantial inter-subgroup variation, the effect of varying subgroup sizes becomes quite clear. In Figures 4b, d, and f it can easily be observed that the method of Laney (2002) provides limits that are too wide when n_i is smaller in Phase II compared to Phase I ($n_I > n_{II}$), and limits that are too narrow when this is reversed ($n_I < n_{II}$). When the sample sizes are identical in Phase I and Phase II ($n_I = n_{II}$), the difference between the two methods is not very large, and is again subject to some sampling variation.

Although the considered scenarios are obviously simplified, it is clear that the method of Laney (2002) suffers from performance issues when sample sizes are different in Phase II than in Phase I. This is due to the multiplicative nature of the variance adjustment, such that the adjustment is too little for subgroups of larger size (small intra-subgroup variation), and too



(a) $r_p = 0$, $n_{II} = 100$

(b) $r_p = 0.05$, $n_{II} = 100$

(c) $r_p = 0$, $n_{II} = 1{,}000$

(d) $r_p = 0.05$, $n_{II} = 1{,}000$

(e) $r_p = 0$, $n_{II} = 10{,}000$

(f) $r_p = 0.05$, $n_{II} = 10{,}000$

Figure 4. Binomial-based Phase II control limits based on $m = 100$ simulated subgroups of size $n_i = n_I = 1{,}000$ in Phase I ($i = 1, \ldots, 100$), with $p_0 = 0.1$ and various settings for r_p and subgroup size $n_i = n_{II}$ in Phase II ($i = 101, \ldots, 130$).

much for subgroups of smaller size (large intra-subgroup variation). In applications where the subgroup sizes remain constant over time, both our method and the p'-chart of Laney (2002) work well. However, our method provides a more suitable alternative to deal with inter-subgroup variation when subgroup sizes vary. Note also that, similarly to the approach of Laney (2002), our method can easily be extended to the monitoring of count data, as discussed in Section 6.

To relate it back to the motivational example in section 2, consider again the hospital emergency data from Table 1. Recall from section 2 that with the method of Laney (2002) it was estimated that only around 6.7% of the total variance is intra-subgroup variance. In Figure 5 we plot the data and the control limits of Laney (2002) again, but this time with the control limits from our proposed method included in the graph. As observed, the proposed limits are much more stable in value, which should be the case when

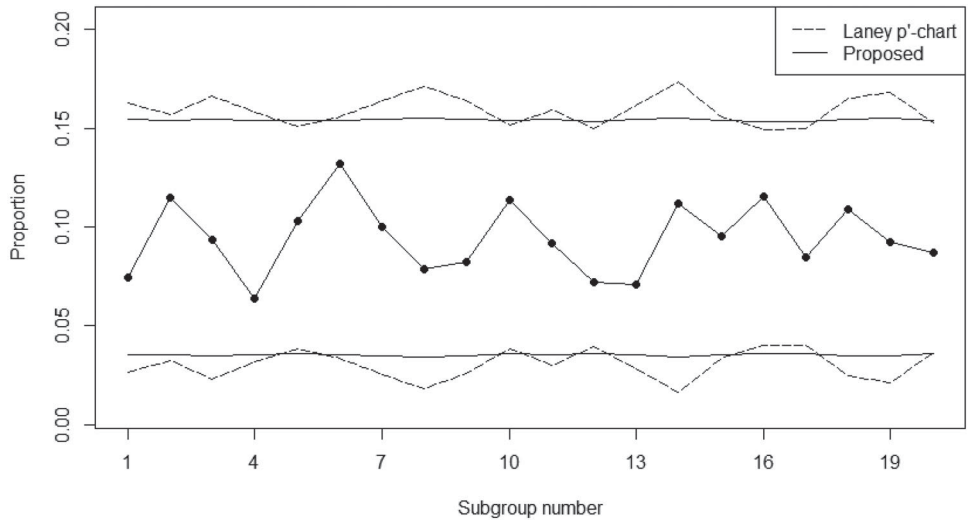


Figure 5. Fraction of nonconforming in hospital emergency department data records using the p'-chart limits of the method of Laney (2002) and our proposed method.

Table 4. Simulated Phase I false alarm rates for control limits based on the method of Laney (2002) and our proposed method, using Phase I samples with $m = 100$ and sample sizes n_i varying from 100 to 1,000.

		Laney			Proposed		
i	n_i	$\sigma = 0$	$\sigma = 0.025$	$\sigma = 0.05$	$\sigma = 0$	$\sigma = 0.025$	$\sigma = 0.05$
1 to 10	100	0.0034	0.0000	0.0000	0.0022	0.0032	0.0037
11 to 20	200	0.0031	0.0002	0.0001	0.0021	0.0031	0.0032
21 to 30	300	0.0030	0.0005	0.0003	0.0019	0.0032	0.0030
31 to 40	400	0.0029	0.0012	0.0010	0.0018	0.0033	0.0027
41 to 50	500	0.0028	0.0025	0.0022	0.0017	0.0034	0.0027
51 to 60	600	0.0028	0.0045	0.0043	0.0017	0.0034	0.0027
61 to 70	700	0.0027	0.0075	0.0069	0.0016	0.0037	0.0026
71 to 80	800	0.0028	0.0109	0.0099	0.0016	0.0036	0.0024
81 to 90	900	0.0027	0.0153	0.0141	0.0015	0.0036	0.0024
91 to 100	1,000	0.0027	0.0205	0.0192	0.0015	0.0038	0.0024

most of the variance is explained outside of the intra-subgroup variation. This further illustrates that our method is better able to handle data with varying subgroup sizes when inter-subgroup variation is present.

5.2. Comparison of false alarm rates

As an additional comparison we evaluated and compared the false alarm rates (FARs) of both methods. When considering this, it is important to realize that the actual FAR of any control chart depends on the distribution of the charting statistic. In our case the charting statistic is $\hat{p}_i$, for which the distribution in turn depends on the distribution of p_i. While the uniform distribution is very suitable to demonstrate and compare the accuracy of the estimation of the control limits under various scenarios, it may not be representative when considering the FAR. Instead, in order to use a bell-shaped distribution for p_i which still adheres to the restriction $0 \leq p_i \leq 1$, we consider the truncated normal distribution for the p_i values here. Additionally, while we used substantial difference in sample sizes in the previous examples to emphasize the estimation issues, smaller differences may be more likely in practice. Therefore, we have also considered smaller stepwise differences in sample sizes in this comparison.

In particular, we considered a Phase I scenario with $m = 100$ subgroups with varying sample sizes. The first ten subgroups $(i = 1, \ldots, 10)$ have sample size $n_i = n_{base}$, the second ten subgroups $(i = 11, \ldots, 20)$ have sample size $n_i = 2n_{base}$, and this continues up until the last ten subgroups $(i = 91, \ldots, 100)$ which have sample size $n_i = 10n_{base}$. We considered $n_{base} = 100$ and $n_{base} = 1,000$. For the parameters of the truncated normal distribution we considered $\mu = 0.1$, the values 0, 0.025, and 0.05 for σ, and truncation at the values 0 (left) and 1 (right). For each combination of n_{base} and σ we simulated 100,000 Phase I samples. For each simulated sample the control limits were calculated for the Laney method as well as the method proposed in this paper. The FAR was then evaluated by determining the proportion of false alarms at each

Table 5. Simulated Phase I false alarm rates for control limits based on the method of Laney (2002) and our proposed method, using Phase I samples with $m = 100$ and sample sizes n_i varying from 1,000 to 10,000.

		Laney			Proposed		
i	n_i	$\sigma = 0$	$\sigma = 0.025$	$\sigma = 0.05$	$\sigma = 0$	$\sigma = 0.025$	$\sigma = 0.05$
1 to 10	1,000	0.0028	0.0000	0.0000	0.0018	0.0028	0.0024
11 to 20	2,000	0.0028	0.0000	0.0000	0.0018	0.0028	0.0023
21 to 30	3,000	0.0028	0.0001	0.0002	0.0017	0.0030	0.0022
31 to 40	4,000	0.0027	0.0009	0.0007	0.0017	0.0030	0.0023
41 to 50	5,000	0.0026	0.0026	0.0019	0.0016	0.0030	0.0022
51 to 60	6,000	0.0028	0.0054	0.0039	0.0016	0.0029	0.0021
61 to 70	7,000	0.0027	0.0097	0.0066	0.0016	0.0028	0.0021
71 to 80	8,000	0.0027	0.0148	0.0104	0.0015	0.0028	0.0020
81 to 90	9,000	0.0025	0.0215	0.0153	0.0014	0.0028	0.0020
91 to 100	10,000	0.0026	0.0286	0.0217	0.0015	0.0026	0.0020

sampling time in the simulation. The results are summarized separately for each unique n_i to illustrate how the varying sample sizes influence the number of false alarms in both methods. The results of our simulation are shown in Tables 4 and 5 for $n_{base} = 100$ and $n_{base} = 1,000$, respectively.

It can be observed from Tables 4 and 5 that if there is no inter-subgroup variation ($\sigma = 0$), both methods have a rather stable FAR for the different subgroup sizes. The FAR values are slightly lower for our proposed method, which is in line with the results of Table 3 that show that the estimated limits can be slightly overestimated in this situation. When inter-subgroup variation is introduced ($\sigma = 0.025$ and $\sigma = 0.05$), substantial differences between both methods become visible. When considering the method of Laney (2002), Tables 4 and 5 clearly show the variation in FAR values increase steadily as the sample sizes increase. For example, when considering $\sigma = 0.025$ in Table 4, the FAR varies from 0.0000 for $n_i = 100$ to 0.0205 for $n_i = 1,000$. In line with the results of the previous sections, this shows that applying the method of Laney (2002) in a situation with varying sample sizes and inter-subgroup variation present leads to control limits that are too wide for smaller samples (resulting in smaller FAR), and too narrow for larger samples (resulting in larger FAR). In this same scenario, the FAR of our proposed method varies much less (between 0.0031 and 0.0038). When comparing the FAR values across all the considered scenarios we also observe that the FAR varies substantially less for our proposed method compared to the method of Laney (2002), which means a more stable and predictable control chart performance. Similar results were found for other values for n_i and σ.

6. Extension to Poisson data

The extension of the binomial case to the Poisson case requires a few small adjustments in the procedure. In order to illustrate the performance of the method in this setting, we reproduced Tables 2 and 3 and Figure 4 by making the following adjustments to the (simulation) procedures of Sections 3 and 4:

1. Consider u_i to be a random variable with mean u_0 and variance σ_u^2, similar to p_i. In our simulations we consider u_i to be drawn from the uniform distribution with radius r_u, e.g. $u_i \sim U(u_0 - r_u, u_0 + r_u)$, such that $\sigma_u = \frac{2r_u}{\sqrt{12}}$. (Note: contrary to p_i, the Poisson rates u_i could be larger than 1, but the essence of the approach remains unchanged).
2. Change X_i to $C_i \sim \text{Poisson}(n_i u_i)$, the total number of nonconformities in subgroup i. Note that in that case $E[C_i|u_i] = Var(C_i|u_i) = n_i u_i$, and consequently $E[C_i] = n_i u_0$ and $Var(C_i) = n_i u_0 + n_i^2 \sigma_u^2$.
3. Replace $\hat{p}_i$ by $\hat{u}_i = C_i/n_i$. Then, $\hat{u}_i$ has mean $E[\hat{u}_i] = u_0$ and variance $Var(\hat{u}_i) = \frac{u_0}{n_i} + \sigma_u^2 = \sigma_{W,i}^2 + \sigma_{B,i}^2$, where $\sigma_{W,i}^2 = u_0/n_i$ and $\sigma_{B,i}^2 = \sigma_u^2$. Note that in this case $\sigma_{B,i}^2$ is constant for all i, regardless of n_i. Note also that this step changes d_i into $\hat{u}_i - \hat{u}_{i-1}$, such that Eq. [8] becomes

$$MSD^* = \frac{1}{m}\sum_{i=1}^{m/2} d_{2i}^2 = \frac{1}{m}\sum_{i=1}^{m/2} (\hat{u}_i - \hat{u}_{i-1})^2 \qquad (12)$$

and Eq. [9] becomes

$$E[MSD^*] = \frac{1}{m}\sum_{i=1}^{m/2} E\left[d_{2i}^2\right] = \frac{1}{m}\sum_{i=1}^{m} \sigma_{B,i}^2 + \frac{1}{m}\sum_{i=1}^{m} \sigma_{W,i}^2 = \sigma_u^2 + \frac{1}{m}\sum_{i=1}^{m} \frac{u_0}{n_i}. \qquad (13)$$

4. Replace $\hat{p}_0$ by $\hat{u}_0 = \frac{\sum_{i=1}^{m} C_i}{\sum_{i=1}^{m} n_i}$.

Table 6. Average of $\hat{\sigma}_z$ values of the method of Laney (2002) for the Poisson case with $u_0 = 0.1$ and various combinations of r_u, N_1, and N_2.

Average of $\hat{\sigma}_z$ values										
		$r_u = 0$			$r_u = 0.025$			$r_u = 0.05$		
N_1	N_2	all n_i	$n_i = N_1$	$n_i = N_2$	all n_i	$n_i = N_1$	$n_i = N_2$	all n_i	$n_i = N_1$	$n_i = N_2$
100	100	0.99	0.99	0.99	1.09	1.09	1.09	1.35	1.35	1.35
100	500	1.00	0.99	1.00	1.27	1.09	1.44	1.84	1.35	2.32
100	1,000	1.00	1.00	1.00	1.44	1.09	1.78	2.24	1.35	3.13
100	5,000	1.00	1.00	1.00	2.27	1.09	3.45	4.03	1.35	6.68
100	10,000	1.00	1.00	1.00	2.94	1.09	4.78	5.39	1.35	9.39
500	500	1.00	1.00	1.00	1.44	1.44	1.44	2.31	2.31	2.31
500	1,000	1.00	1.00	1.00	1.61	1.44	1.78	2.72	2.31	3.12
500	5,000	1.00	1.00	1.00	2.45	1.44	3.45	4.52	2.32	6.70
500	10,000	1.00	1.00	1.00	3.11	1.44	4.78	5.87	2.31	9.40
1,000	1,000	1.00	1.00	1.00	1.78	1.78	1.78	3.12	3.12	3.12
1,000	5,000	1.00	1.00	1.00	2.62	1.78	3.45	4.90	3.11	6.68
1,000	10,000	1.00	1.00	1.00	3.28	1.78	4.78	6.28	3.12	9.42
5,000	5,000	1.00	1.00	1.00	3.45	3.45	3.45	6.68	6.69	6.68
5,000	10,000	1.00	1.00	1.00	4.12	3.45	4.78	8.04	6.69	9.39
10,000	10,000	1.00	1.00	1.00	4.78	4.78	4.78	9.40	9.40	9.40

Table 7. Average of Y_i values of the proposed method for the Poisson case with $u_0 = 0.1$ and various combinations of r_u, N_1, and N_2.

Average of Y_i values							
		$r_u = 0$		$r_u = 0.025$		$r_u = 0.05$	
N_1	N_2	$n_i = N_1$	$n_i = N_2$	$n_i = N_1$	$n_i = N_2$	$n_i = N_1$	$n_i = N_2$
100	100	1.04	1.04	1.01	1.01	0.99	0.99
100	500	1.03	1.12	1.00	0.99	1.00	0.99
100	1,000	1.03	1.22	1.00	0.97	1.00	0.99
100	5,000	1.03	1.71	1.00	0.94	1.00	0.99
100	10,000	1.03	2.11	1.00	0.94	1.00	0.99
500	500	1.04	1.04	1.00	1.00	1.00	1.00
500	1,000	1.03	1.06	1.00	0.99	1.00	1.00
500	5,000	1.03	1.21	1.00	0.99	1.00	1.00
500	10,000	1.03	1.36	1.00	0.99	1.00	1.00
1,000	1,000	1.04	1.04	1.00	1.00	1.00	1.00
1,000	5,000	1.03	1.12	1.00	0.99	1.00	1.00
1,000	10,000	1.03	1.21	1.00	0.99	1.00	1.00
5,000	5,000	1.04	1.04	1.00	1.00	1.00	1.00
5,000	10,000	1.03	1.06	1.00	1.00	1.00	1.00
10,000	10,000	1.04	1.04	1.00	1.00	1.00	1.00

5. Estimate σ_u^2 by

$$\hat{\sigma}_u^2 = \frac{1}{m}\sum_{i=1}^{m/2}(\hat{u}_i - \hat{u}_{i-1})^2 - \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{u}_0}{n_i}. \qquad (14)$$

Similarly to the binomial case, we advise to use $\hat{\sigma}_u^2 = 0$ in the event that this estimator turns out to become negative.

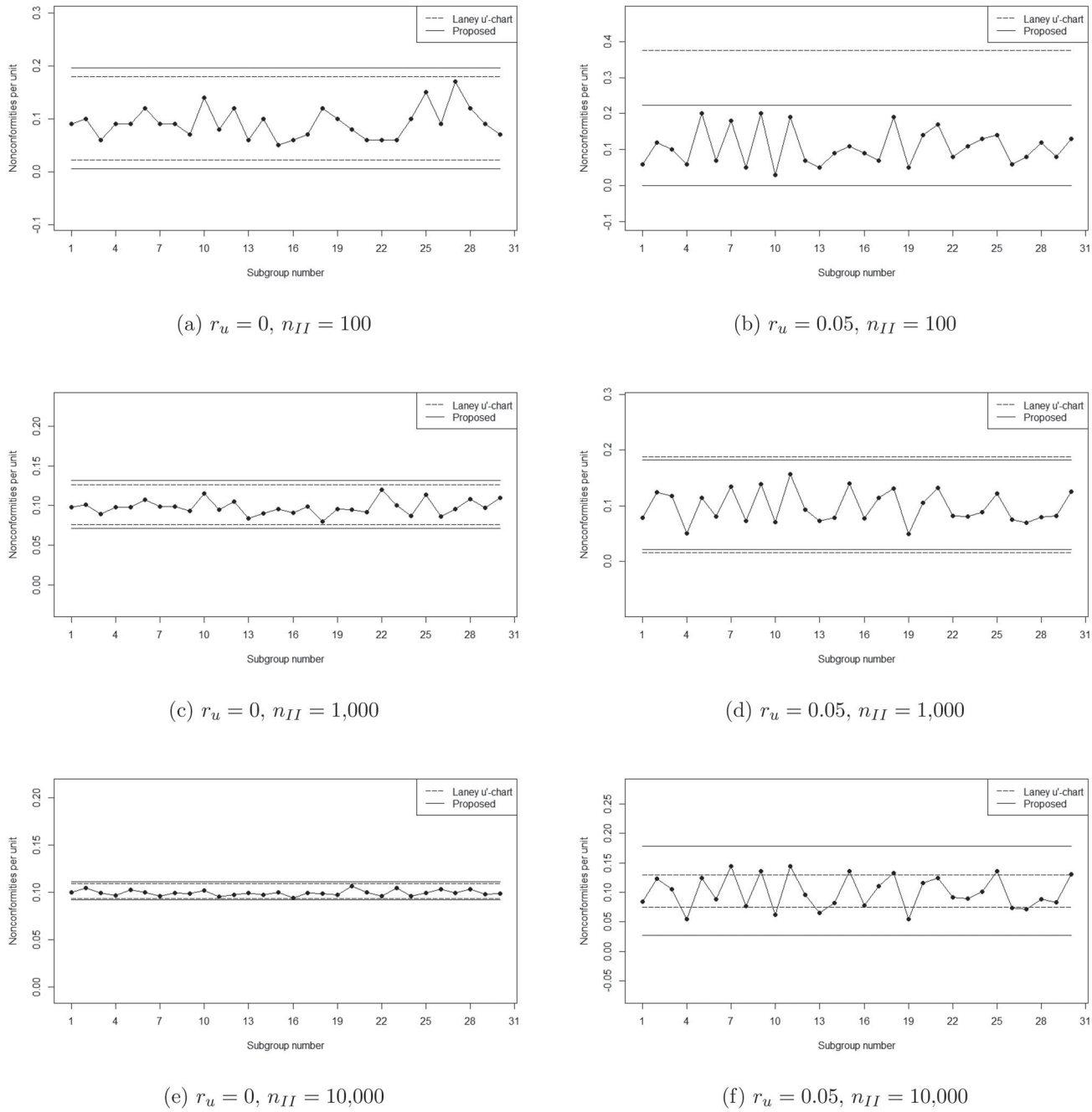
6. Estimate $\sigma_{W,i}^2$ and $\sigma_{B,i}^2$ by $\hat{\sigma}_{W,i}^2 = \frac{\hat{u}_0}{n_i}$ and $\hat{\sigma}_{B,i}^2 = \hat{\sigma}_u^2$, respectively. For the method of Laney (2002), consider $z_i = \frac{\hat{u}_i - \hat{u}_0}{\hat{\sigma}_{W,i}}$ and $\hat{\sigma}_z = AMR(z_i)/1.128$ as before.

7. Use the obtained estimates in the previous step to calculate $\hat{\sigma}_{\hat{u}_i}^2 = \hat{\sigma}_z \hat{\sigma}_{W,i}$ for the method of Laney (2002), and $\hat{\sigma}_{\hat{u}_i}^2 = \hat{\sigma}_{B,i}^2 + \hat{\sigma}_{W,i}^2$ for our proposed method. The UCL and LCL of each method can then be calculated according to

$$\hat{u}_0 + k\hat{\sigma}_{\hat{u}_i} \text{ and } \hat{u}_0 - k\hat{\sigma}_{\hat{u}_i}, \qquad (15)$$

respectively, with k some positive constant most often set to $k = 3$. We also consider $k = 3$ in our simulations.

8. Run all further simulation procedures as before. The results are given in Tables 6, 7, and Figure 6, and are analogous to the results of the binomial case in Tables 2, 3, and Figure 4 respectively. Our method provides accurate estimates of the variance of the counts whereas in some cases the method of Laney (2002) does not. In particular, our method provides a substantial improvement when sample sizes vary and there is inter-subgroup variation present.



(a) $r_u = 0$, $n_{II} = 100$

(b) $r_u = 0.05$, $n_{II} = 100$

(c) $r_u = 0$, $n_{II} = 1{,}000$

(d) $r_u = 0.05$, $n_{II} = 1{,}000$

(e) $r_u = 0$, $n_{II} = 10{,}000$

(f) $r_u = 0.05$, $n_{II} = 10{,}000$

Figure 6. Poisson-based Phase II control limits based on $m = 100$ simulated subgroups of size $n_i = n_I = 1{,}000$ in Phase I ($i = 1, \ldots, 100$), with $u_0 = 0.1$ and various settings for r_u and subgroup size $n_i = n_{II}$ in Phase II ($i = 101, \ldots, 130$).

7. Concluding remarks

We proposed a new method to calculate the control limits when monitoring proportions. When inter-subgroup common cause variation is present, the p-chart is not able to provide suitable control limits as subgroup sizes become larger. Laney (2002) provided an alternative method to compensate for this variation. However, as shown in our paper, the proposed method of Laney (2002) has performance issues when subgroup sizes vary. We provide an alternative procedure to estimate the control limits of p-charts to overcome this issue. The results in this paper show that the method is well able to handle situations of varying subgroup sizes and inter-subgroup variation. The proposed method can easily be extended to other control charts for attribute data, such as we demonstrated for Poisson count data.

About the authors

Dr Rob Goedhart is an assistant professor in the Department of Operations Management and senior consultant at the Institute for Business and Industrial Statistics of the University of Amsterdam, the Netherlands. He received his MS (cum laude) in econometrics in 2014 and his PhD in statistics in 2018, both from the University of Amsterdam.

Dr William H. Woodall is an emeritus professor in the Department of Statistics at Virginia Tech. He is a former editor of the *Journal of Quality Technology* (2001–2003). He is the recipient of the Box Medal (2012), Shewhart Medal (2002), Hunter Award (2019), Youden Prize (1995, 2003), Brumbaugh Award (2000, 2006), Bisgaard Award (2012), Nelson Award (2014), Ott Foundation Award (1987), and best paper award for IIE Transactions on Quality and Reliability Engineering (1997). He is a Fellow of the American Statistical Association, a Fellow of the American Society for Quality, and an elected member of the International Statistical Institute.

References

Alwan, L. C., and H. V. Roberts. 1995. The problem of misplaced control limits. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 44 (3):269–78.

Chang, T. C., and F. F. Gan. 2004. Shewhart charts for monitoring the variance components. *Journal of Quality Technology* 36 (3):293–308. doi: 10.1080/00224065.2004.11980275.

Crossley, M. L. 2007. *The desk reference of statistical quality methods* 2nd ed. Milwuakee, WI: ASQ Quality Press.

Hagan, J., and B. Li. 2018. Phase II performance of p-charts and p'-charts. *Journal of Medical Statistics and Informatics* 6 (1):3. doi: 10.7243/2053-7662-6-3.

Heimann, P. A. 1996. Attributes control charts with large sample sizes. *Journal of Quality Technology* 28 (4):451–9. doi: 10.1080/00224065.1996.11979703.

Laney, D. B. 2002. Improved control charts for attributes. *Quality Engineering* 14 (4):531–7. doi: 10.1081/QEN-120003555.

Mohammed, M. A., and D. B. Laney. 2006. Overdispersion in health care performance data: Laney's approach. *Quality & Safety in Health Care* 15 (5):383–4. doi: 10.1136/qshc.2006.017830.

Mohammed, M. A., J. S. Panesar, D. B. Laney, and R. Wilson. 2013. Statistical process control charts for attribute data involving very large sample sizes: A review of problems and solutions. *BMJ Quality & Safety* 22 (4):362–8. doi: 10.1136/bmjqs-2012-001373.

Montgomery, D. C. 2020. *Introduction to statistical quality control*. 8th ed. Hoboken, NJ: John Wiley & Sons.

Provost, L. P., and S. Murray. 2011. *The health care data guide: Learning from data for improvement*. San Francisco, CA: Jossey-Bass.

Roes, K. C. B., R. J. M. M. Does, and Y. Schurink. 1993. Shewhart–type control charts for individual observations. *Journal of Quality Technology* 25 (3):188–98. doi: 10.1080/00224065.1993.11979453.

Suman, G., and D. Prajapati. 2018. Control chart applications in healthcare: A literature review. *International Journal of Metrology and Quality Engineering* 9 (5):5. doi: 10.1051/ijmqe/2018003.

Vidmar, G., and R. Blagus. 2014. Outlier detection for healthcare quality monitoring–A comparison of four approaches to over-dispersed proportions. *Quality and Reliability Engineering International* 30 (3):347–62. doi: 10.1002/qre.1581.

Western Electric Company. 1956. *Statistical quality control handbook*. Indianapolis, IN: AT&T.

Woodall, W. H. 1997. Control charts based on attribute data: Bibliography and review. *Journal of Quality Technology* 29 (2):172–83. doi: 10.1080/00224065.1997.11979748.

Woodall, W. H., and E. V. Thomas. 1995. Statistical process control with several components of common cause variability. *IIE Transactions* 27 (6):757–64. doi: 10.1080/07408179508936792.

Yashchin, E. 1994. Monitoring variance components. *Technometrics* 36 (4):379–93. doi: 10.1080/00401706.1994.10485844.

Zwetsloot, I. M., and W. H. Woodall. 2021. A review of some sampling and aggregation strategies for basic statistical process monitoring. *Journal of Quality Technology* 53 (1):1–16. doi: 10.1080/00224065.2019.1611354.