Binomial Capability and Poisson Capability

Overview

Capability analysis is used to evaluate whether a process is capable of producing output that meets customer requirements. When it is not possible to represent the quality of a product or service with continuous data, attribute data is often collected to assess its quality. The Minitab Assistant includes two analyses to examine the capability of a process with attribute data:

- Binomial Capability: This analysis is used when a product or service is characterized as defective or not defective. Binomial capability evaluates the chance (p) that a selected item from a process is defective. The data collected are the number of defective items in individual subgroups, which is assumed to follow a binomial distribution with parameter p.
- Poisson Capability: This analysis is used when a product or service can have multiple defects and the number of defects on each item is counted. Poisson capability evaluates the number of defects per unit. The data collected are the total number of defects in k units contained in individual subgroups, which is assumed to follow a Poisson distribution with an unknown mean number of defects per unit (u).

To adequately estimate the capability of the current process and to reliably predict the capability of the process in the future, the data for these analyses should come from a stable process (Bothe, 1991; Kotz and Johnson, 2002). In addition, there should be enough subgroups collected over time to ensure that the capability estimates represent the process capability over a long period of time. Even if a process is in control, it may experience input and environmental changes over time. Therefore, using an adequate number of subgroups can better enable you to capture the different sources of variation over time (Bothe, 1997; AIAG, 1995). Finally, there should be enough data to ensure that the capability statistics have good precision, as indicated by the width of the confidence interval for the key capability measure reported by both analyses.

Based on these requirements, the Assistant Report Card automatically performs the following checks on your data:

- Stability of process
    - Tests for special causes
    - Subgroup size
- Number of subgroups
- Expected variation
- Amount of data

In this paper, we investigate how these requirements relate to capability analysis in practice and we describe how we established the guidelines to check for these requirements in the Assistant.

We also explain the Laney P' and U' charts that are recommended when the observed variation in the data doesn't match the expected variation and Minitab detects overdispersion or underdispersion.

Note Binomial and Poisson capability analyses include the P and U attribute control charts, respectively, to check process stability. These two charts depend on additional assumptions that either cannot be checked or are difficult to check. See Appendix A for details.

Data checks

Stability (Part I) – Test for special causes

To estimate process capability accurately, your data should come from a stable process. You should verify the stability of your process before you evaluate its capability. If the process is not stable, you should identify and eliminate the causes of the instability.

The P chart and the U chart are the most widely used attribute control charts to evaluate the stability of a process. The P chart plots the proportion of defective items per subgroup and is used with data that follow a binomial distribution. The U chart plots the number of defects per unit and is used with data that follow a Poisson distribution. Four tests can be performed on these charts to evaluate the stability of the process. Using these tests simultaneously increases the sensitivity of the control chart. However, it is important to determine the purpose and added value of each test because the false alarm rate increases as more tests are added to the control chart.

Objective

We wanted to determine which of the four tests for stability to include with the attribute control charts in the Assistant. Our first goal was to identify the tests that significantly increased sensitivity to out-of-control conditions without significantly raising the false alarm rate. Our second goal was to ensure the simplicity and practicality of the charts.

Method

The four tests for stability for attribute charts correspond with tests 1-4 for special causes for variables control charts. With an adequate subgroup size, the proportion of defective items (P chart) or the number of defects per unit (U chart) follow a normal distribution. As a result, simulations for the variables control charts that are also based on the normal distribution will yield identical results for the sensitivity and false alarm rate of the tests. Therefore, we used the results of a simulation and a review of the literature performed for variables control charts to evaluate how the four tests for stability affect the sensitivity and the false alarm rate of the attribute charts. In addition, we evaluated the prevalence of special causes associated with the test. For details on the method(s) used for each test, see the Results section below and Appendix B.

Results

Of the four tests used to evaluate stability in attribute charts, we found that tests 1 and 2 are the most useful:

TEST 1: IDENTIFIES POINTS OUTSIDE OF THE CONTROL LIMITS

Test 1 identifies points > 3 standard deviations from the center line. Test 1 is universally recognized as necessary for detecting out-of-control situations. It has a false alarm rate of only 0.27%.

TEST 2: IDENTIFIES SHIFTS IN THE PROPORTION OF DEFECTIVE ITEMS (P CHART) OR THE MEAN NUMBER OF DEFECTS PER UNIT (U CHART)

Test 2 signals when 9 points in a row fall on the same side of the center line. We performed a simulation to determine the number of subgroups needed to detect a signal for a shift in the proportion of defective items (P chart) or a shift in the mean number of defects per unit (U chart). We found that adding test 2 significantly increases the sensitivity of the chart to detect small shifts in the proportion of defective items or the mean number of defects per unit. When test 1 and test 2 are used together, significantly fewer subgroups are needed to detect a small shift compared to when test 1 is used alone. Therefore, adding test 2 helps to detect common out-of-control situations and increases sensitivity enough to warrant a slight increase in the false alarm rate.

Tests not included in the Assistant

TEST 3: K POINTS IN A ROW, ALL INCREASING OR ALL DECREASING

Test 3 is designed to detect drifts in the proportion of defective items or in the mean number of defects per unit (Davis and Woodall, 1988). However, when test 3 is used in addition to test 1 and test 2, it does not significantly increase the sensitivity of the chart. Because we already decided to use tests 1 and 2 based on our simulation results, including test 3 would not add any significant value to the chart.

TEST 4: K POINTS IN A ROW, ALTERNATING UP AND DOWN

Although this pattern can occur in practice, we recommend that you look for any unusual trends or patterns rather than test for one specific pattern.

Stability (Part II) - Subgroup size

Although the P chart and the U chart monitor the stability of the process with attribute data, the normal distribution is used to approximate the distribution of the proportion of defective items ($\hat{p}$) in the P chart and the distribution of the number of defects per unit ($\hat{u}$) in the U chart. As the subgroup size increases, the accuracy of this approximation improves. Because the criteria for the tests used in each control chart are based on the normal distribution, increasing the subgroup size to obtain a better normal approximation improves the chart’s ability to accurately identify out-of-control situations and reduces the false alarm rate. When the proportion of defective items or the number of defects per unit is low, you need larger subgroups to ensure accurate results.

Objective

We investigated the subgroup size that is needed to ensure that the normal approximation is adequate enough to obtain accurate results for the P chart and the U chart.

Method

We performed simulations to evaluate the false alarm rates for various subgroup sizes and for various proportions (p) for the P chart and for various mean numbers of defects per subgroup (c) for the U chart. To determine whether the subgroup size was large enough to obtain an adequate normal approximation and thus, a low enough false alarm rate, we compared the results with expected false alarm rate under the normal assumption (0.27% for Test 1 and 0.39% for test 2). See Appendix C for more details.

Results

P CHART

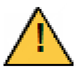
Our research showed that the required subgroup size for the P chart depends on the proportion of defective items (p). The smaller the value of p, the larger the subgroup size (n) that is required. When the product np is greater than or equal to 0.5, the combined false alarm rate for both test 1 and test 2 is below approximately 2.5%. However, when the product np is less than 0.5, the combined false alarm rate for tests 1 and 2 can be much higher, reaching levels well above 10%. Therefore, based on this criterion, the performance of the P chart is adequate when the value of $np \geq 0.5$.

U CHART

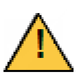
Our research showed that the required subgroup size for the U chart depends on the number of defects per subgroup (c), which equals the subgroup size (n) times the number of defects per unit (u). The percentage of false alarms is highest when the number of defects c is small. When $c = nu$ is greater than or equal to 0.5, the combined false alarm rate for both test 1 and test 2 is below approximately 2.5%. However, for values of c less than 0.5, the combined false alarm rate for tests 1 and 2 can be much higher, reaching levels well above 10%. Therefore, based on this criterion, the performance of the U chart is adequate when the value of $c = nu \geq 0.5$.

Based on the above results for the tests for special causes (Part I) and for the subgroup size (Part II), the Assistant Report Card displays the following status indicators when checking stability in the attribute control charts that are used in binomial and Poisson capability:

P chart – Binomial capability

Status	Condition
	No test 1 or test 2 failures on the chart and a$n_i\ \bar{p}\ \geq 0.5$ for all i where $n_i =$ subgroup size for the ith subgroup $\bar{p}$ = mean proportion of defective items
	Test 1 or test 2 reveals one or more out-of-control points that may be due to special causes.
	The subgroup size may be too small. a$n_i\ \bar{p} < 0.5\ for\ at\ least\ one\ i$

U chart - Poisson capability

Status	Condition
	No test 1 or test 2 failures on the chart and a$n_i\ \bar{u}\ \geq 0.5$ for all i where $n_i =$ subgroup size for the ith subgroup $\bar{u}$ = mean number of defects per unit
	Test 1 or test 2 reveals one or more out-of-control points that may be due to special causes.
	The subgroup size may be too small. a$n_i\ \bar{\bar{u}}\ < 0.5\ for\ at\ least\ one\ i$

Number of subgroups

To ensure that the capability estimates accurately reflect your entire process, you should try to capture all the likely sources of variation in your process over time. If you increase the number of subgroups you collect, you are likely to increase the chance that you are capturing the different sources of variation. Collecting an adequate number of subgroups also helps to improve the precision of the limits of the control charts that are used to evaluate the stability of your process. However, collecting more subgroups requires more time and resources; therefore, it is important to know how the number of subgroups affects the reliability of the capability estimates.

Objective

We investigated how many subgroups are needed to adequately represent the process and provide a reliable estimate of process capability.

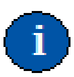
Method

We reviewed the literature to find out the number of subgroups that is generally considered adequate for estimating process capability.

Results

According to the Statistical Process Control (SPC) manual, the number of subgroups you collect should be based on how long it takes to collect data that is likely to reflect the different sources of variation in your process (AIAG, 1995). That is, you should collect as many subgroups as is necessary to adequately represent your entire process. In general, to provide accurate tests of stability and a reliable estimate of process performance, AIAG (1995) recommends that you collect at least 25 subgroups.

Based on these recommendations, the Assistant Report Card displays the following status indicator when checking the number of subgroups for binomial or Poisson capability analysis:

Status	Condition
	**Number of subgroups ≥ 25** The number of subgroups should be enough to capture different sources of process variation when collected over an adequate period of time. **Number of subgroups < 25** Generally, you should collect at least 25 subgroups over an adequate period of time to capture different sources of process variation.

Expected Variation

The traditional P charts and U charts that are used to assess the stability of the process prior to evaluating its capability assume the variation in the data follows the binomial distribution for defectives or a Poisson distribution for defects. The charts also assume that your rate of defectives or defects remains constant over time. When the variation in the data is either greater than or less than expected, your data may have overdispersion or underdispersion and the charts may not perform as expected.

Overdispersion

Overdispersion exists when the variation in your data is more than expected. Typically, some variation exists in the rate of defectives or defects over time, caused by external noise factors that are not special causes. In most applications of these charts, the sampling variation of the subgroup statistics is large enough that the variation in the underlying rate of defectives or defects is not noticeable. However, as the subgroup sizes increase, the sampling variation becomes smaller and smaller and at some point the variation in the underlying defect rate can become larger than the sampling variation. The result is a chart with extremely narrow control limits and a very high false alarm rate.

Underdispersion

Underdispersion exists when the variation in your data is less than expected. Underdispersion can occur when adjacent subgroups are correlated with each other, also known as autocorrelation. For example, as a tool wears out, the number of defects may increase. The increase in defect counts across subgroups can make the subgroups more similar than they would be by chance. When data exhibit underdispersion, the control limits on a traditional P chart or U chart may be too wide. If the control limits are too wide the chart will rarely signal, meaning that you can overlook special cause variation and mistake it for common cause variation.

If overdispersion or underdispersion is severe enough, Minitab recommends using a Laney P′ or U′ chart. For more information, see Laney P′ and U′ charts below.

Objective

We wanted to determine a method to detect overdispersion and underdispersion in the data.

Method

We performed a literature search and found several methods for detecting overdispersion and underdispersion. We selected a diagnostic method found in Jones and Govindaraju (2001). This method uses a probability plot to determine the amount of variation expected if the data were from a binomial distribution for defectives data or a Poisson distribution for defects data. Then, a comparison is made between the amount of expected variation and the amount of observed variation. See Appendix D for details on the diagnostic method.

As part of the check for overdispersion, Minitab also determines how many points are outside of the control limits on the traditional P and U charts. Because the problem with overdispersion is a high false alarm rate, if only a small percentage of points are out of control, overdispersion is unlikely to be an issue.

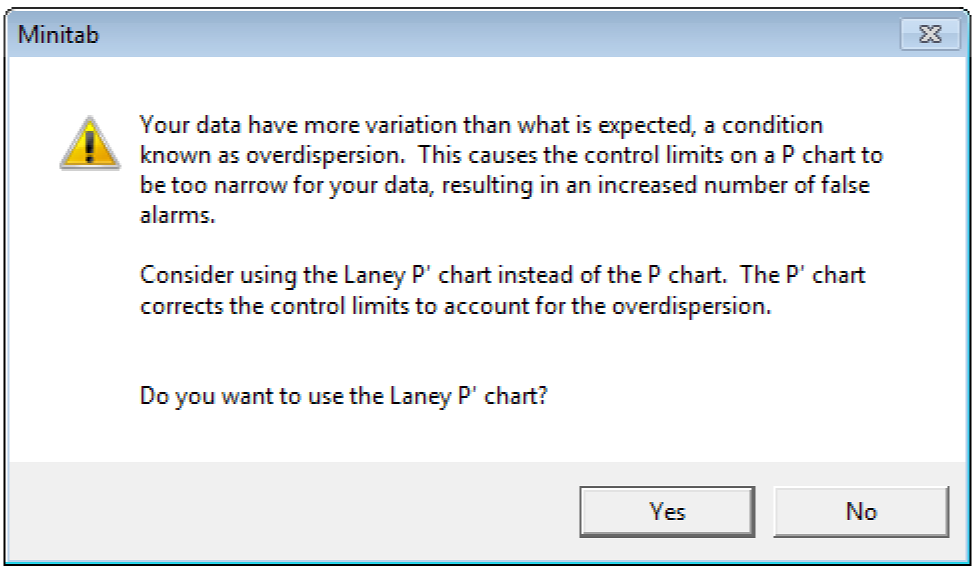
Results

Minitab performs the diagnostic check for overdispersion and underdispersion after the user selects OK in the dialog box for the P or U chart before the chart is displayed.

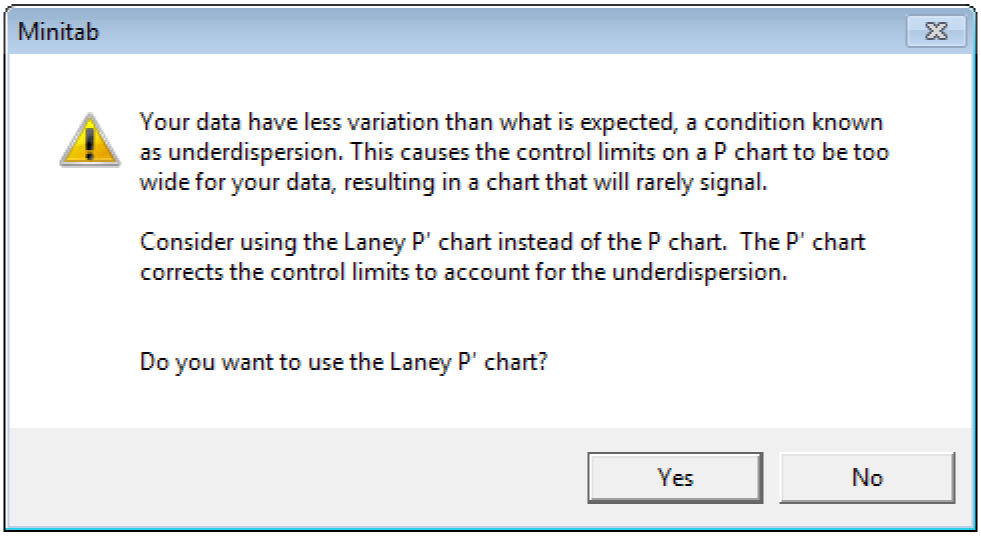
Overdispersion exists when these following conditions are met:

- The ratio of observed variation to expected variation is greater than 130%.
- More than 2% of points are outside the control limits.
- The number of points outside the control limits is greater than 1.

If overdispersion is detected, Minitab displays a message that asks if the user wants to display a Laney P′ or U′ chart. Shown below is the message for the P′ chart:

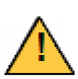


Underdispersion exists when the ratio of observed variation to expected variation is less than 75%. If underdispersion is detected, Minitab displays a message that asks if the user wants to display a Laney P′ or U′ chart. Shown below is the message for the P′ chart:



If the user chooses to use the Laney chart, Minitab displays the Laney charts in the Diagnostic report. If the user chooses not to use the Laney chart, Minitab displays both the traditional chart and the Laney chart in the Diagnostic report. Showing both charts allows the user to see the effect of overdispersion or underdispersion on the traditional P or U chart and determine whether the Laney chart is more appropriate for their data.

Additionally, when checking for overdispersion or underdispersion, the Assistant Report Card displays the following status indicators:

Status	Condition
	Dispersion ratio > 130%, less than 2% of points outside control limits or number of points outside control limits = 1 Dispersion ratio > 75% and <= 130% Dispersion ratio > 130%, more than 2% of points outside control limits and number of points outside control limits > 1 and user chose to use Laney P’ or U’ Dispersion ratio < 75% and user chose to use Laney P’ or U’ Where Dispersion ratio = 100*(observed variation)/(expected variation)
	Dispersion ratio > 130%, more than 2% of points outside control limits and number of points outside control limits > 1 and user did not choose to use Laney P’ or U’ Dispersion ratio < 75% and user did not choose to use Laney P’ or U’

Amount of data

The Assistant reports for binomial and Poisson capability analyses also include a 95% confidence interval for the percentage of defective items or the number of defects per unit, respectively. This interval is calculated using standard statistical methodology and did not require any special research or simulations.

The Assistant Report Card displays the following status indicator when checking the amount of data:

Status	Condition
i	**Binomial capability** The 95% confidence interval for % defective is (a, b). If this interval is too wide for your application, you can gather more data to increase the precision. **Poisson capability** The 95% confidence interval for the number of defects per unit is (a, b). If this interval is too wide for your application, you can gather more data to increase the precision.

Laney P’ and U’ Charts

Traditional P charts and U charts assume the variation in the data follows the binomial distribution for defectives data or a Poisson distribution for defect data. The charts also assume that your rate of defectives or defects remains constant over time. Minitab performs a check to determine whether the variation in the data is either greater than or less than expected, an indication the data may have overdispersion or underdispersion. See the Expected Variation data check above.

If overdispersion or underdispersion are present in the data, the traditional P and U charts may not perform as expected. Overdispersion can cause the control limits to be too narrow, resulting in a high false alarm rate. Underdispersion can cause the control limits to be too wide, which can cause you to overlook special cause variation and mistake it for common cause variation.

Objective

Our objective was to identify an alternative to the traditional P and U charts when overdispersion or underdispersion is detected in the data.

Method

We reviewed the literature and determined that the best approach for handling overdispersion and underdispersion are the Laney P’ and U’ charts (Laney, 2002). The Laney method uses a revised definition of common cause variation, which corrects the control limits that are either too narrow (overdispersion) or too wide (underdispersion).

In the Laney charts, common cause variation includes the usual short-term within subgroup variation but also includes the average short-term variation between consecutive subgroups. The common cause variation for Laney charts is calculated by normalizing the data and using the average moving range of adjacent subgroups (referred to as Sigma Z on the Laney charts) to adjust the standard P or U control limits. Including the variation between consecutive subgroups helps correct the effect when the variation in the data across subgroups is greater than or less than expected due to fluctuations in the underlying defect rate or a lack of randomness in the data.

After Sigma Z is calculated, the data are transformed back to the original units. Using the original data units is beneficial because if the subgroup sizes are not the same, the control limits are allowed to vary just as they are in the traditional P and U charts. For more details on Laney P' and U' charts, see Appendix E.

Results

Minitab performs a check for overdispersion or underdispersion and if either condition is detected, Minitab recommends a Laney P' or U' chart.

References

AIAG (1995). *Statistical process control (SPC) reference manual.* Automotive Industry Action Group.

Bischak, D.P., & Trietsch, D. (2007). The rate of false signals in $\overline{\mathrm{X}}$ control charts with estimated limits. *Journal of Quality Technology*, *39*, 55–65.

Bothe, D.R. (1997). *Measuring process capability: Techniques and calculations for quality and manufacturing engineers*. New York: McGraw-Hill.

Bowerman, B.L., & O' Connell, R.T. (1979). *Forecasting and time series: An applied approach*. Belmont, CA: Duxbury Press.

Chan, L. K., Hapuarachchi K. P., & Macpherson, B.D. (1988). Robustness of $\bar{X}$ and *R* charts. *IEEE Transactions on Reliability, 37*, 117–123.

Davis, R.B., & Woodall, W.H. (1988). Performance of the control chart trend rule under linear shift. *Journal of Quality Technology*, *20*, 260–262.

Laney, D. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531-537.

Montgomery, D.C. (2001). *Introduction to statistical quality control*, 4^{th} edition. New York: John Wiley & Sons, Inc.

Schilling, E.G., & Nelson, P.R. (1976). The effect of non-normality on the control limits of $\bar{X}$ charts. *Journal of Quality Technology*, *8*, 183–188.

Trietsch, D. (1999). *Statistical quality control: A loss minimization approach*. Singapore: World Scientific Publishing Co.

Wheeler, D.J. (2004). *Advanced topics in statistical process control. The power of Shewhart's charts*, 2^{nd} edition. Knoxville, TN: SPC Press.

Yourstone, S.A., & Zimmer, W.J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences*, *23*, 1099–1113.

Appendix A: Additional assumptions for attribute control charts

The P chart and the U chart require additional assumptions that are not evaluated by data checks:

P Chart	U Chart
• The data consists of n distinct items, with each item classified as either defective or not defective. • The probability of an item being defective is the same for each item within a subgroup. • The likelihood of an item being defective is not affected by whether the preceding item is defective or not.	• The counts are counts of discrete events. • The discrete events occur within some well-defined finite region of space, time, or product. • The events occur independently of each other and the likelihood of an event is proportional to the size of area of opportunity.

For each chart, the first two assumptions are an inherent part of the data collection process; the data itself cannot be used to check whether these assumptions are satisfied. The third assumption can be verified only with a detailed and advanced analysis of data, which is not performed by the Assistant.

Appendix B: Stability - Tests for special causes

Simulation B1: How adding test 2 to test 1 affects sensitivity

Test 1 detects out-of-control points by signaling when a point is greater than 3 standard deviations from the center line. Test 2 detects shifts in the proportion of defective items or the number of defects per unit by signaling when 9 points in a row fall on the same side of the center line.

To evaluate whether using Test 2 with Test 1 improves the sensitivity of the attribute charts, we established control limits based on a normal (p, $\sqrt{\frac{p(1-p)}{n}}$) (p is the proportion of defective items and n is the subgroup size) distribution for the P chart and on a normal (u $\sqrt{u}$) (u is the mean number of defects per unit) distribution for the U chart. We shifted the location (p or u) of each distribution by a multiple of the standard deviation (SD) and then recorded the number of subgroups needed to detect a signal for each of 10,000 iterations. The results are shown in Table 1.

Table 1 Average number of subgroups until a test 1 failure (Test 1), test 2 failure (Test 2) or test 1 or test 2 failure (Test 1 or 2). The shift equals a multiple of the standard deviation (SD).

Shift	Test 1	Test 2	Test 1 or 2
0.5 SD	154	84	57
1 SD	44	24	17
1.5 SD	15	13	9
2 SD	6	10	5

As shown in the table, when both tests are used (*Test 1 or 2* column) an average of 57 subgroups are needed to detect a 0.5 standard deviation shift in the location, compared to an average of 154 subgroups needed to detect a 0.5 standard deviation shift when test 1 is used alone. Therefore, using both tests significantly increases sensitivity to detect small shifts in the proportion of defective items or the mean number of defects per unit. However, as the size of the shift increases, adding test 2 does not increase the sensitivity as significantly.

Appendix C: Stability - Subgroup size

The central limit theorem states that the normal distribution can approximate the distribution of the average of an independent, identically distributed random variable. For the P chart, $\hat{p}$ (subgroup proportion) is the average of an independent, identically distributed Bernoulli random variable. For the U chart, $\hat{u}$ (subgroup rate) is the average of an independent, identically distributed Poisson random variable. Therefore, the normal distribution can be used as an approximation in both cases.

The accuracy of the approximation improves as the subgroup size increases. The approximation also improves with a higher proportion of defective items (P chart) or a higher number of defects per unit (U chart). When either the subgroup size is small or the values of p (P chart) or u (U chart) are small, the distributions for $\hat{p}$ and $\hat{u}$ are right skewed, which increases the false alarm rate. Therefore, we can evaluate the accuracy of the normal approximation by looking at the false alarm rate and we can also determine the minimum subgroup size necessary to obtain an adequate normal approximation.

To do this, we performed simulations to evaluate the false alarm rates for various subgroup sizes for the P chart and the U chart and compared the results with the expected false alarm rate under the normal assumption (0.27% for Test 1 and 0.39% for test 2).

Simulation C1: Relationship between subgroup size, proportion, and false alarm rate of the P chart

Using an initial set of 10,000 subgroups, we established the control limits for various subgroup sizes (n) and proportions (p). We also recorded the percentage of false alarms for an additional 2,500 subgroups. We then performed 10,000 iterations and calculated the average percentage of false alarms from test 1 and test 2, as shown in Table 2.

Table 2 % of false alarms due to test 1, test 2 (np) for various subgroup sizes (n) and proportions (p)

	p				
Subgroup Size (n)	0.001	0.005	0.01	0.05	0.1
10	0.99, 87.37 (0.01)	4.89, 62.97 (0.05)	0.43, 40.14 (0.1)	1.15, 1.01 (0.5)	1.28, 0.42 (1)
50	4.88, 63.00 (0.05)	2.61, 10.41 (0.25)	1.38, 1.10 (0.5)	0.32, 0.49 (2.5)	0.32, 0.36 (5)
100	0.47, 40.33 (0.10)	1.41, 1.12 (0.5)	1.84, 0.49 (1)	0.43, 0.36 (5)	0.20, 0.36 (10)

	p				
Subgroup Size (n)	0.001	0.005	0.01	0.05	0.1
150	1.01, 25.72 (0.15)	0.71, 0.43 (0.75)	0.42, 0.58 (1.5)	0.36, 0.42 (7.5)	0.20, 0.36 (15)
200	1.74, 16.43 (0.2)	1.86, 0.50 (1.00)	0.43, 0.41 (2)	0.27, 0.36 (10)	0.34, 0.36 (20)
500	1.43, 1.12 (0.5)	0.42, 0.50 (2.5)	0.52, 0.37 (5)	0.32, 0.37 (25)	0.23, 0.36 (50)

The results in Table 2 show that the percentage of false alarms is generally highest when the proportion (p) is small, such as 0.001 or 0.005, or when the sample size is small (n = 10). Therefore, the percentage of false alarms is highest when the value of the product np is small, and lowest when np is large. When np is greater or equal to 0.5, the combined false alarm rate for both test 1 and test 2 is below approximately 2.5%. However, for values of np less than 0.5, the combined false alarm rate for tests 1 and 2 is much higher, reaching levels well above 10%. Therefore, based on this criterion, the performance of the P chart is adequate when the value of np ≥ 0.5. Thus, the subgroup size should be at least $\frac{0.5}{\bar{p}}$.

Simulation C2: Relationship between subgroup size, number of defects per unit, and false alarm rate of the U chart

Using an initial set of 10,000 subgroups, we established the control limits for various subgroup sizes (n) and number of defects per subgroup (c). We also recorded the percentage of false alarms for an additional 2,500 subgroups. We then performed 10,000 iterations and calculated the average percent of false alarms from test 1 and test 2, as shown in Table 3.

Table 3 % of false alarms due to test 1, test 2 for various number of defects per subgroup (c = nu)

c	0.1	0.3	0.5	0.7	1.0	3.0	5.0	10.0	30.0	50
% False alarms	0.47, 40.40	3.70, 6.67	1.44, 1.13	0.57, 0.39	0.36, 0.51	0.38, 0.40	0.54, 0.38	0.35, 0.37	0.29, 0.37	0.25, 0.37

The results in Table 3 show that the percentage of false alarms is highest when the product of the subgroup size (n) times the number of defects per unit (u), which equals the number of defects per subgroup (c), is small. When c is greater or equal to 0.5, the combined false alarm rate for both test 1 and test 2 is below approximately 2.5%. However, for values of c less than 0.5, the combined false alarm rate for tests 1 and 2 is much higher, reaching levels well above

10%. Therefore, based on this criterion, the performance of the U chart is adequate when the value of c = nu ≥ 0.5. Thus, the subgroup size should be at least $\frac{0.5}{\bar{u}}$.

Appendix D: Overdispersion/Underdispersion

Let d_i be the defective count from subgroup i, and n_i be the subgroup size.

First, normalize the defective counts. To account for possibly different subgroup sizes, use adjusted defective counts ($adjd_i$):

$adjd_i$ = adjusted defective count for subgroup i = $\frac{d_i}{n_i}(\bar{n})$, where

$\bar{n}$ = average subgroup size

$X_i = \sin^{-1} \sqrt{\frac{adjd_i + {}^3/_8}{\bar{n}+0.75}}$

The normalized counts (X_i) will have a stdev equal to $\frac{1}{\sqrt{4 * \bar{n}}}$. This means that 2 standard deviations is equal to $\frac{1}{\sqrt{\bar{n}}}$.

Then, generate a standard normal probability plot using the normalized counts as data. A regression line is fit using only the middle 50% of the plot points. Find the 25th and 75th percentiles of the transformed count data and use all X-Y pairs ≥ 25th percentile and ≤ 75th percentile. This line is used to obtain the predicted transformed count values corresponding to Z values of -1 and +1. The “Y” data in this regression are the normal scores of the transformed counts and the “X” data are the transformed counts.

Calculate the observed variation as follows:

Let Y(-1) be the predicted transformed count for Z = -1

Let Y(+1) be the predicted transformed count for Z = +1

Observed estimate of 2 standard deviations = Y(+1) – Y(-1).

Calculate the expected variation as follows:

Expected estimate of 2 standard deviations = $\frac{1}{\sqrt{\bar{n}}}$

Calculate the ratio of observed variation to expected variation and convert to a percentage. If the percentage is > 130%, more than 2% of the points are outside the control limits, and the number of points outside the control limits > 1, there is evidence of overdispersion. If the percentage is < 75%, there is evidence of underdispersion.

Appendix E: **Laney P' and U' Charts**

The concept behind the Laney P' and U' charts is to account for cases where the observed variation between subgroups does not match the expected variation if the subgroup data were from a random process with a constant rate of defects or defectives. Small changes in the underlying rate of defects or defectives occur normally in every process. When subgroup sizes are relatively small, the sampling variation in the subgroups is large enough so that these small changes are not noticeable. As subgroup sizes increase, the sampling variation decreases, and the small changes in the underlying rate of defects or defectives become large enough to adversely affect the standard P and U charts by increasing the false alarm rate. Some examples have shown false alarm rates to be as high as 70%. This condition is known as overdispersion.

An alternative method was developed to remedy this issue, which normalizes the subgroup p or u values and plots the normalized data in an I Chart. The I Chart uses a moving range of the normalized values to determine its control limits. Thus, the I Chart method changes the definition of common cause variation by adding in the variation in the defectives or defect rate from one subgroup to the next.

The Laney method transforms the data back to the original units. The advantage of this is that if the subgroups are not all the same size, the control limits will not be fixed, as they are with the I Chart method.

The P' and U' charts combine the new definition of common cause variation with the variable control limits one would expect from having different subgroup sizes. Thus, the key assumption for these charts is that the definition of common cause variation is changed – it includes the usual short-term variation that is present within the subgroups plus the average short-term variation one would expect to see between consecutive subgroups.

Laney P' chart

Let

X_i = number of defectives in subgroup i

n_i = subgroup size for subgroup i

p_i = proportion defective for subgroup i

$$\bar{p} = \frac{\sum X_i}{\sum n_i}$$

$$\sigma p_i = \sqrt{\frac{\bar{p} * (1 - \bar{p})}{n_i}}$$

First, convert the p_i to z-scores:

$$Z_i = \frac{p_i - \bar{p}}{\sigma p_i}$$

Next, a moving range of length 2 is used to evaluate the variation in the z-scores and calculate Sigma Z (σz).

$$\sigma z = \frac{\overline{MR}}{1.128}$$

where 1.128 is an unbiasing constant.

Transform the data back to original scale:

$$p_i = \bar{p} + \sigma p_i * \sigma z$$

Thus, the standard deviation of p_i is:

$$sd(p_i) = \sigma p_i * \sigma z$$

The control limits and center line are calculated as:

Center line = $\bar{p}$

UCL= $\bar{p} + 3 * sd(p_i)$

LCL = $\bar{p} - 3 * sd(p_i)$

Laney U' chart

Let

X_i = number of defectives in subgroup i

n_i = subgroup size for subgroup i

u_i = proportion defective for subgroup i

$\bar{u} = \frac{\sum X_i}{\sum n_i}$

$$\sigma u_i = \sqrt{\frac{\bar{u} * (1 - \bar{u})}{n_i}}$$

First, convert the p_i to z-scores:

$$Z_i = \frac{u_i - \bar{u}}{\sigma u_i}$$

Next, a moving range of length 2 is used to evaluate the variation in the z-scores and calculate Sigma Z (σz).

$$\sigma z = \frac{\overline{MR}}{1.128}$$

where 1.128 is an unbiasing constant.

Transform the data back to original scale:

$$u_i = \bar{u} + \sigma u * \sigma z$$

Thus, the standard deviation of p_i is:

$$sd(u_i) = \sigma u_i * \sigma z$$

The control limits and center line are calculated as:

Center line = $\bar{u}$

UCL= $\bar{u} + 3 * sd(u_i)$

LCL= $\bar{u} - 3 * sd(u_i)$