14 Visualizing quality control

This chapter covers

- Quality control measures
- Control charts for attributes (or qualitative data)
- Control charts for variables (or quantitative data)

Quality is a fundamental concept that permeates every aspect of our lives, from the products we use daily to the services we rely on. Defining quality, however, is not always straightforward. In a broad sense, quality can be understood as the degree to which a product or service meets or exceeds customer expectations. This encompasses attributes such as reliability, durability, and performance. But quality is not merely an attribute of a finished product; it is an integral aspect of the entire production and service delivery process. When organizations focus on quality, they aim to ensure that every component of their operations contributes to the creation of a superior final product or service.

To maintain and improve quality, organizations implement controls. Controls are systematic measures put in place to regulate processes, ensuring they function within defined parameters and produce consistent results. Controls can be preventive, aimed at stopping problems before they occur, or detective, identifying issues as they arise. Effective control mechanisms are essential for maintaining standards, reducing variability, and minimizing defects. Without proper controls, processes can deviate, leading to variations that compromise the quality of the final product.

This is where statistical quality control comes into play. Statistical quality control involves the use of statistical methods to monitor and control a process. It is a branch of quality management that employs statistical tools, including control charts for attributes and variables, to identify and rectify sources of variability, ensuring processes remain stable and predictable. The foundation of statistical quality control lies in the collection and analysis of data. By examining data patterns and trends, organizations can make informed decisions that enhance process performance and product quality. Techniques include control charts, process capability analysis, and design of experiments, all of which are geared toward achieving and maintaining high quality.

Visualization plays a crucial role in statistical quality control. The human brain processes visual information far more efficiently than numerical data. This is why charts and graphs are invaluable tools in quality management. They provide a clear and immediate understanding of complex data, making it easier to identify trends, outliers, and patterns. Visual tools like control charts transform abstract numbers into tangible insights, enabling quick and effective decision-making.

Control charts are a quintessential example of how visualization and quality control are inherently linked. These charts, whether for attributes or variables, graphically display process data over time, helping to monitor process stability and control. For instance, a p-chart (proportion chart) tracks the proportion of defective items in a process, whereas an np-chart monitors the number of defective items in a sample. By plotting data points against control limits, control charts highlight any deviations from the norm, signaling when a process is going out of control. This visual representation allows quality managers to quickly pinpoint issues, investigate causes, and implement corrective actions.

The marriage of visualization and quality control is not just a matter of convenience; it is a powerful synergy that enhances the effectiveness of quality management. Visual tools like control charts simplify the interpretation of data, making it accessible to a broader audience, from shop floor operators to senior management. They facilitate a common understanding of process performance and foster a culture of continuous improvement. In the technology industry, similar visualizations are widely used to monitor key performance indicators—such as latency, error rates, and availability—providing guardrails that help teams detect quality issues and maintain service reliability.

In the context of this last chapter, we will explore various quality control charts, including p-charts and np-charts, demonstrating how they can be constructed and interpreted. These tools will serve as practical examples of how visualizing quality control can drive better decision-making and improve overall process quality. By integrating statistical methods with visual tools, organizations can not only maintain but also elevate their quality standards, ensuring that they consistently meet or exceed customer expectations.

To effectively utilize quality control charts, it is crucial to first identify and understand the key measures that are foundational to these tools. These measures provide the necessary insights into process performance and help in making informed decisions to maintain or improve quality.

14.1 Quality control measures

Defining key measures used in common quality control reports is essential for accurately interpreting various types of control charts. Although not all of these measures will appear in every chart demonstrated, they play a crucial role in providing a comprehensive understanding of quality control practices. Gaining familiarity with these metrics ensures a stronger foundation for analyzing and maintaining process quality effectively.

14.1.1 Upper control limit and lower control limit

In quality control charts, the upper control limit (UCL) and lower control limit (LCL) are the most critical measures. These limits are essential for distinguishing between normal process variation and signals that a process may be out of control.

The UCL and LCL are typically defined mathematically as being three standard deviations above and below the mean, respectively. This conventional setting ensures that about 99.73% of all data points should fall within these limits if the process is in control. Imagine a chart where the mean is represented as a central horizontal line, with the UCL and LCL as parallel lines above and below it; data points that fall within these limits indicate a stable process, whereas those outside suggest potential issues that require further investigation. However, it's worth noting that not all industries follow this exact three-sigma convention. In the technology sector, for instance, teams often define custom guardrails for key metrics—such as latency thresholds or error rates—based on operational goals rather than statistical distributions. This ties back to chapter 3, where we noted that approximately 99.73% of a normally distributed data set lies within three standard deviations of the mean—a useful baseline that some industries adapt or reinterpret based on context.

However, the control limits can and should be adjusted based on specific use cases and requirements. For example,

- In highly critical processes, such as pharmaceutical manufacturing or aerospace component production, the control limits might be set at two standard deviations from the mean. This tighter range helps to detect smaller shifts or trends in the process more quickly, ensuring that any potential issues are identified and corrected promptly.
- In processes where some variability is inherent and acceptable, such as agricultural products or construction materials, the control limits might be set at four

standard deviations. This broader range accommodates natural variations without signaling unnecessary alarms.

In quality control charts, the UCL and LCL are typically depicted as horizontal lines above and below the center line (which represents the process mean). Data points are plotted over time to monitor the process performance.

Understanding how to interpret and use control limits is critical for effective quality control. The following points outline the key considerations for evaluating data points in relation to the UCL and LCL:

- When all data points fall within the UCL and LCL, the process is considered to be in control. This indicates that any variations are due to common causes inherent to the process.
- If a data point falls outside these limits, it signals that the process may be out of control. This could be due to special cause variation, which requires investigation and corrective action.
- Even if data points are within control limits, specific patterns, such as runs of seven or more points on one side of the mean or consistent upward or downward trends, can indicate potential issues.

By setting appropriate control limits, organizations can effectively monitor process performance, quickly identify any deviations, and implement corrective actions to maintain high-quality standards. The choice of control limits should be based on the specific context, goals, and requirements of the process, ensuring a balanced approach to quality control.

14.1.2 Mean and center line

The mean is the arithmetic average of a set of values and serves as a central value for the data being analyzed. It is calculated by summing all the individual values and dividing by the number of values. The mean provides a measure of central tendency, representing the expected value in a set of data.

The center line (CL) in a control chart is the horizontal line that represents the mean of the process being monitored. It serves as a reference point for evaluating process performance. Data points are plotted in relation to this line to determine whether the process is stable and in control. The CL helps identify variations from the mean, facilitating the detection of trends and patterns that may indicate issues in the process.

14.1.3 Standard deviation

In the context of quality control charts, the standard deviation is used to measure the variability of a process. For attribute data, such as in p-charts, the formula for standard deviation is different from the general formula used for continuous data. This is because attribute data represents counts or proportions, which have different statistical properties compared to continuous measurements.

The formula for the standard deviation (σ) of a proportion in a p-chart is

$$\sigma = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where

- $\bar{p}$ represents the average proportion of defective items over all samples. It is calculated as the total number of defective items divided by the total number of items inspected.
- $1 - \bar{p}$ represents the proportion of nondefective items in the sample. Together with $\bar{p}$ it accounts for the total variability in the sample.
- n is the sample size, indicating the number of items inspected in each sample. The variability of the proportion decreases with larger sample sizes.

Otherwise, the formula uses the square root to transform the variance (the average of the squared differences from the mean) into the standard deviation, thereby providing a measure of spread in the same units as the data.

For continuous data, which involves measurements on a continuous scale, the standard deviation is calculated using a different formula:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

where

- x_i represents the individual data points or measurements.
- $\bar{x}$ is the mean of all the data points.
- $\sum(x_i - \bar{x})^2$ represents the sum of the squared differences between each data point and the mean. It measures the total variability in the data.
- n is the number of data points in the data set.

The square root of the variance converts it into the standard deviation, which is in the same units as the data points.

Using the correct formula for standard deviation in quality control charts ensures that control limits are appropriately set, allowing for effective monitoring of process performance. By understanding the nature of the data and applying the right statistical measures, quality control charts can accurately detect variations, identify potential issues, and maintain high standards of quality.

14.1.4 Range

Range in the context of quality control charts is a measure of the dispersion or variability within a set of data points. It is calculated as the difference between the highest and lowest values in a sample. The range provides a simple yet effective way to assess the spread of data and is often used in r-charts (range charts) to monitor process variability over time.

The formula for calculating the range (R) is

$$R = X_{max} - X_{min}$$

where

- X_{max} is the maximum value in the sample.
- X_{min} is the minimum value in the sample.

In quality control, the range is particularly useful for detecting changes in the consistency of the process. It helps identify variations that may indicate a loss of control, allowing for timely interventions to maintain process stability and quality.

14.1.5 Sample size

Sample size in the context of quality control charts refers to the number of observations or data points collected from a process during a specific sampling period. It is denoted by n and plays a crucial role in the accuracy and reliability of the control charts.

The sample size affects several aspects of quality control charts, including

- *Control limits*—UCL and LCL calculations are sometimes influenced by the sample size. Larger sample sizes tend to produce narrower control limits, making the control chart more sensitive to small shifts in the process.
- *Statistical power*—A larger sample size increases the ability to detect true variations in the process, enhancing a chart's effectiveness in identifying out-of-control conditions.
- *Variation*—With larger sample sizes, the estimates of process parameters (such as the mean and standard deviation) become more accurate, providing a better representation of the actual process behavior.

In summary, the sample size is a fundamental parameter in quality control charts that affects the precision, sensitivity, and reliability of the monitoring process. It determines the extent to which the control chart can accurately reflect the process's stability and detect deviations from the desired quality standards.

14.1.6 Proportion defective

Proportion defective is a fundamental metric in quality control charts that quantifies the percentage of units within a sample that fail to meet established quality standards. This measure, denoted by p, is particularly critical in attribute control charts like p-charts, where the focus is on tracking the quality of a process over time. By calculating the proportion defective, quality control practitioners gain immediate insight into the overall performance of a process. The proportion defective is determined using a straightforward formula:

$$p = \frac{D}{n}$$

where

- D represents the number of defective units.
- n is the total number of units in the sample.

In the context of p-charts, the proportion defective serves as the primary measure for monitoring the quality of a process over time. By plotting the proportion of defective items in each sample, p-charts help identify variations and trends that may indicate underlying issues in the process. Although the proportion defective itself does not determine the control limits, it is used to calculate the standard deviation, which in turn helps establish the upper and lower control limits. These control limits are essential for distinguishing between common cause variation, which is inherent to the process, and special cause variation, which signals a potential problem that warrants investigation.

Monitoring the proportion defective is crucial for maintaining consistent quality standards and ensuring that the process remains in control. By identifying trends, shifts, or cycles in the proportion of defective items, quality control practitioners can implement timely corrective actions and continuous improvements. This proactive approach helps organizations meet customer expectations, comply with regulatory requirements, and maintain a competitive edge in the market.

14.1.7 Number of defective items

The number of defective items is a crucial measure in quality control, representing the count of units in a sample that fail to meet specified quality standards. This measure, denoted by D, is fundamental in attribute control charts such as np-charts and c-charts. It provides a straightforward way to assess the quality of a process by counting the actual number of defective units.

In the context of np-charts, the number of defective items is the primary measure used to monitor the quality of a process over time. Unlike p-charts, which track the proportion of defective items, np-charts focus on the actual count of defects in a sample. This approach is particularly useful when the sample size remains constant, allowing for a direct and intuitive interpretation of the number of defects. For instance, if a sample of 100 items contains 5 defective units, then $D = 5$, providing a clear picture of the defect level.

For c-charts, the number of defective items is used to count defects per unit in situations where each unit can have multiple defects. This type of chart is helpful for processes where defects can occur in various forms on a single unit, such as scratches, dents, or other imperfections. By tracking the count of defects per unit, c-charts help in identifying patterns and trends in the process, highlighting areas that may require attention.

Calculating the number of defective items involves simply counting the units that do not meet quality criteria in each sample. This measure helps directly assess the level of defects, making it easy for operators and managers to understand and interpret. By plotting the number of defective items on control charts, quality control practitioners can quickly identify when a process is out of control, necessitating investigation and corrective actions.

Monitoring the number of defective items is essential for maintaining high quality standards. It allows organizations to identify variations and trends in the production process, implement timely corrective actions, and ensure customer satisfaction. By using the number of defective items in quality control charts, organizations can effectively track defects, identify areas for improvement, and maintain consistent quality in their processes, thereby meeting regulatory requirements and achieving operational excellence.

14.1.8 Number of defects

The number of defects is a critical measure in quality control that quantifies the total instances of nonconformance or imperfections within a sample. This measure, often denoted by C, plays a pivotal role in attribute control charts such as c-charts and u-charts. Unlike the number of defective items, which counts the units that fail to meet quality standards, the number of defects focuses on the total count of defects within those units, providing a more detailed picture of the quality issues present in the process.

In the context of c-charts, the number of defects is used to monitor the count of defects per unit or per sample over time. This type of chart is particularly useful in processes where each unit can have multiple defects, such as manufacturing, where a single product might have several types of flaws, like scratches, dents, or misalignments. By tracking the total number of defects, c-charts help identify trends and variations that may indicate problems in the production process, enabling timely interventions to maintain quality standards.

Similarly, u-charts utilize the number of defects but normalize this count based on the number of units inspected, providing a defect rate per unit. This normalization is crucial when the sample size varies, allowing for a fair comparison of defect levels across different samples. For example, if 100 units are inspected in one sample and 200 units in another, the u-chart adjusts for this difference, ensuring that the analysis remains consistent and meaningful.

Calculating the number of defects involves counting all instances of nonconformance within a sample. For instance, if a sample of 50 products has a total of 20 scratches and 10 dents, the number of defects C is 30. This measure offers a granular view of quality issues, making it easier to pinpoint specific problems and areas that require improvement.

14.1.9 Defects per unit

Tracking defects per unit is a vital aspect of quality control, particularly in processes where multiple defects can occur in a single unit. Control charts such as c-charts and u-charts are used to monitor the number of defects per unit over time. By capturing and analyzing this data, organizations can quickly identify when the process is deviating from acceptable quality standards. The number of defects per unit can range from zero (indicating no defects) to higher numbers, depending on the complexity of the product and the production process. Monitoring these variations helps maintain process stability, implement corrective actions when needed, and ensure that the final products meet quality expectations and regulatory requirements.

14.1.10 Moving range

Moving range is an important concept in quality control, particularly in the analysis of process variability through control charts. It refers to the difference between consecutive data points in a sequence, providing a measure of how much the process varies from one observation to the next. The moving range (MR) is calculated by taking the absolute difference between successive observations:

$$\mathrm{MR} = |X_i - X_{i-1}|$$

where

- X_i is the current data point.
- X_{i-1} is the previous data point.

In the context of quality control charts, the moving range is often used in conjunction with individuals-moving range (I-MR) charts. Although the individuals (I) chart monitors the central tendency of a process (i.e., the mean), the MR chart focuses on the variability between consecutive measurements. By tracking these two aspects, I-MR charts provide a comprehensive view of the process, allowing quality control practitioners to monitor both the process average and the consistency of the data.

A stable process typically exhibits a consistent MR, with most values falling within established control limits. When the MR remains stable, it suggests that the process variability is under control and that any observed differences between consecutive data points are within expected limits. However, significant deviations in the MR, such as large spikes, can indicate unusual variability or special cause variation. These deviations may signal potential issues in the process, requiring further investigation to identify and address the root cause.

14.1.11 z-score

The z-score is a statistical measure that quantifies the distance of a specific data point from the mean of a data set, expressed in terms of standard deviations. This measure is crucial in both attribute and variable control charts, as it helps determine whether a process is operating within its expected range or if a particular data point indicates a potential issue. The z-score (z) is calculated using the formula

$$z = \frac{X - \mu}{\sigma}$$

where

- X represents the individual data point.
- μ is the mean of the data set.
- σ is the standard deviation.

Despite the different methods for calculating the standard deviation depending on the type of data, the z-score formula itself remains consistent. This consistency allows quality control practitioners to assess how far any given data point deviates from the mean and whether the data pertains to measurable quantities or defect rates. A z-score of 0 indicates that the data point is exactly at the mean, whereas positive or negative z-scores show how many standard deviations the data point is above or below the mean. This makes the z-score an essential tool in monitoring process stability and identifying when corrective action may be necessary to maintain quality.

14.1.12 Process capability indices

Process capability indices are statistical measures used in quality control to assess the ability of a process to produce output that meets specified quality standards. These indices compare the spread and centering of the process data relative to the specified limits, providing a quantitative measure of how well a process can produce products within the desired specifications.

The process capability index (C_p) measures the potential capability of a process by comparing the width of the process variation (measured by six standard deviations, or the process spread) to the width of the specification limits. The formula for C_p is

$$C_p = \frac{\text{USL} - \text{LSL}}{6\sigma}$$

where

- USL is the upper specification limit.
- LSL is the lower specification limit.
- σ is the standard deviation of the process.

A higher *Cp* value means the process has a greater potential to produce output within the specification limits. However, *Cp* assumes that the process is perfectly centered between the specification limits, which is often not the case in real-world scenarios.

The process capability index adjusted for centering (*Cpk*) accounts for the actual centering of the process relative to the specification limits. It adjusts the *Cp* index to reflect the position of the process mean relative to the specification limits. The formula for *Cpk* is

$$Cpk = min\left(\frac{\text{USL} - \mu}{3\sigma}, \frac{\mu - \text{LSL}}{3\sigma}\right)$$

Therefore, it provides a more realistic assessment of process capability by considering both the process variation and how well the process is centered within the specification limits. A higher *Cpk* value indicates that the process is both capable of meeting the specifications and is properly centered within the specification limits.

With a solid understanding of the key quality control measures, we're now ready to see how these elements come together in practice. Next, we'll explore control charts for attributes, where we'll apply many of these measures to monitor and control processes involving categorical data.

14.2 Control charts for attributes

Control charts for attributes are essential tools in quality management, used to monitor and control processes where the data is categorical in nature, such as the number of defective items or the proportion of defective units in a sample. Unlike control charts for variables, which deal with continuous data, attribute control charts focus on tracking the quality of a process based on the presence or absence of specific characteristics. These charts provide valuable insights into the stability and capability of a process, helping to identify trends, variations, and potential issues that could affect overall quality. In this section, we will explore some of the key attribute control charts, demonstrating how they are constructed and interpreted to ensure that processes remain within acceptable limits and consistently meet quality standards.

The choice of which control chart to use, or whether to use multiple charts, depends on the specific characteristics of the process and the nature of the data being monitored. For example, a p-chart, which tracks the proportion of defective items in a sample, is particularly useful when sample sizes vary from one period to the next. However, if sample sizes are consistent, an np-chart, which tracks the number of defective items rather than the proportion, might be more straightforward and effective. Similarly, although multiple control charts might be necessary to monitor different aspects of a complex process, a single chart could suffice for simpler processes with fewer variables to track. The key is to select the most relevant chart or combination of charts that best suit the specific quality control needs of the process.

14.2.1 p-charts

Proportion charts, typically referred to as p-charts, are a fundamental tool in quality control for monitoring the proportion of defective items in a process over time. They are particularly useful when dealing with attribute data, where each item is classified as either defective or nondefective. The p-chart helps to determine whether the process is stable and operating within acceptable limits by tracking the proportion of defective items in successive samples.

In practice, p-charts are essential for processes where the focus is on the rate of defects rather than the number of defects. For example, in a manufacturing process, a p-chart might be used to monitor the percentage of faulty products in daily production batches. By plotting the proportion of defects and comparing it against established control limits, quality control practitioners can quickly identify trends, shifts, or any unusual patterns that could indicate potential issues with the process.

We have a pandas data frame called `oj` containing synthetic data representing the number of defective units across 64 batches of orange juice cans. Potential defects could include physical damage, improper sealing, incorrect or missing labels, rust or corrosion, or the presence of foreign objects or contaminants inside the can or under the lid. Before accessing and analyzing the data, the `pd.read_csv()` method must be called to import the data set. In this example, we've used a discretionary parameter to load only three specific columns from the .csv file, thereby focusing our analysis on just the relevant variables.

The `info()` method displays a concise summary of a pandas data frame, including important details such as the column names, the number of non-null entries in each column, and the data types of each column. This information is valuable for understanding the structure and contents of a data frame before performing further analysis:

```
>>> import pandas as pd
>>> print(oj.info())
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 64 entries, 0 to 63
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   sample  64 non-null     int64
 1   D       64 non-null     int64
 2   size    64 non-null     int64
dtypes: int64(3)
memory usage: 1.6 KB
None
```

Each batch, or sample, is represented as a single row in the `oj` data frame, which includes the following attributes:

- `sample`—A numeric variable that serves as a unique identifier for each batch of orange juice cans, numbered sequentially from 31 through 94. This column contains no null values. In our p-chart, the `sample` variable will be used as the *x* axis to plot the proportion of defective units across the batches.
- `D`—A numeric variable indicating the number of defective orange juice cans in each sample. This column also contains no null values.
- `size`—A numeric variable representing the total number of cans inspected in each sample; consistently set at 50 across all entries. This column contains no null values.

A subsequent call to the `head()` method returns, by default, the first five rows of the `oj` data frame. This allows us to visually inspect the initial data, providing context and confirming the integrity of the data frame beyond the summary provided by the `info()` method:

```
>>> print(oj.head())
   sample   D  size
0      31   9    50
1      32   6    50
2      33  12    50
3      34   5    50
4      35   6    50
```

The `oj` data frame provides the defect count for each sample, but to better analyze the quality control process, we want to examine the *fraction* of defective units per sample. To do this, we will create a derived variable called `proportion`, which represents the ratio of defective cans (`D`) to the total cans inspected (`size`). This can be achieved with the following line of code:

```
>>> oj['proportion'] = oj['D'] / oj['size']
```

Now we have the *y*-axis variable for our p-chart.

A p-chart must include a center line, which represents the average proportion of defective units across all samples. Additionally, it requires an UCL and LCL, which are calculated to identify the thresholds beyond which the process is considered out of control.

The next line of code creates a variable called `p_bar`, which stores the computed mean of the `proportion` values across all rows in the `oj` data frame. This average serves as the CL in the p-chart, representing the overall defect rate across all samples:

```
>>> p_bar = oj['proportion'].mean()
```

The following snippet of code calculates the upper and lower control limits for the p-chart, which are critical in determining whether the process is within control. The UCL and LCL are derived from the standard deviation of the proportion of defective units, calculated using the `numpy` library and the following formula:

```
>>> import numpy as np
>>> oj['sigma'] = np.sqrt((p_bar * (1 - p_bar)) / oj['size'])
```

This standard deviation accounts for the variability in the proportion of defects. The UCL and LCL are then determined by adding and subtracting three times the standard deviation from the mean proportion (`p_bar`), respectively:

```
>>> oj['UCL'] = p_bar + 3 * oj['sigma']
>>> oj['LCL'] = p_bar - 3 * oj['sigma']
```

With all the necessary attributes and measures calculated, we are now ready to show the p-chart using the `matplotlib` library. This chart will plot the proportion of defective units for each sample, along with the CL (`p_bar`), UCL, and LCL, allowing us to assess the stability and control of the process over time:

```
>>> import matplotlib.pyplot as plt                   <- Imports the matplotlib library
>>> plt.figure()                                      <- Initializes a matplotlib figure, creating a blank canvas for the upcoming plot
>>> plt.plot(oj['sample'], oj['proportion'],
>>>          marker = 'o',
>>>          linestyle = '-',
>>>          label = 'Proportion Defective')          <- Plots proportion against sample using a solid line and circular markers; adds a label
>>> plt.axhline(y = p_bar,
>>>             color = 'r', linestyle = '--',
>>>             label = 'p-bar (Mean)')               <- Draws a horizontal dashed red line at the height of p_bar on the plot; adds a label
```

```
>>> plt.plot(oj['sample'], oj['UCL'],
>>>          color = 'g', linestyle = '--',
>>>          label = 'UCL')
>>> plt.plot(oj['sample'], oj['LCL'],
>>>          color = 'g', linestyle = '--',
>>>          label = 'LCL')
>>> plt.fill_between(oj['sample'], oj['LCL'], oj['UCL'],
>>>                  color = 'g', alpha = 0.1)
>>> plt.title('p-Chart')
>>> plt.xlabel('Sample')
>>> plt.ylabel('Proportion Defective')
>>> plt.legend(loc = 'upper right',
>>>            bbox_to_anchor = (1, .95))
>>> plt.grid(True)
>>> plt.show()
```

Draws a green dashed line representing the UCL across all samples; adds a label

Draws a green dashed line representing the LCL across all samples; adds a label

Adds a label to the x axis

Fills the area between the control limits across all samples with a light green color

Adds a label to the y axis

Adds a title atop the plot

Adds horizontal and vertical lines to the plot

Adds a legend in the upper right corner of the plot, nudged downward

Displays the plot

Finally, we can now display the p-chart to show the proportion of defective units across the samples: see figure 14.1. It reveals that the process of producing orange juice cans is generally in control, as all the data points fall within the upper and lower control

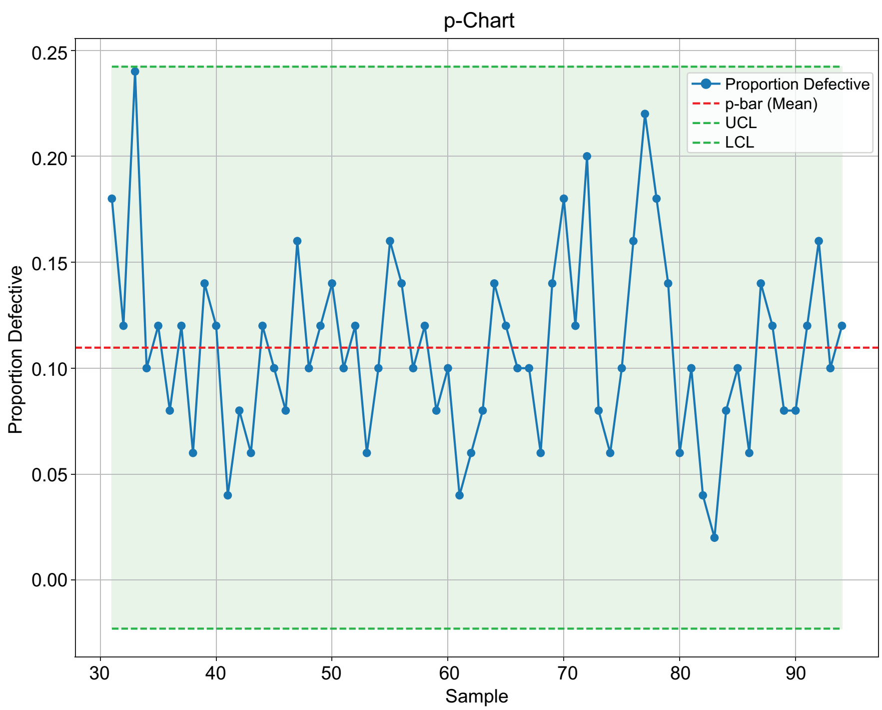

Figure 14.1 This p-chart illustrates the proportion of defective units in each sample of orange juice cans, with the dashed line across the middle representing the mean proportion. The dashed lines at the top and bottom indicate the UCL and LCL, providing boundaries for expected variation in the process. Data points within these limits suggest that the process is in control, and any points outside these limits would signal potential issues requiring investigation.

limits. This indicates that the variations in the proportion of defective units across the samples are due to common causes inherent to the process rather than any special cause variations. However, the proportion defective fluctuates noticeably around the mean, with some samples showing higher defect rates. These fluctuations, although within the control limits, suggest that there could be areas for improvement to reduce variability and enhance overall process stability. Continuous monitoring is essential to ensure that the process remains stable and that any potential issues are identified and addressed promptly before they lead to defects that fall outside the control limits.

14.2.2 np-charts

In the field of quality control, np-charts serve as a valuable tool for monitoring the number of defective items in a sample, particularly when the sample size remains consistent across different observations. Unlike p-charts, which focus on the proportion of defects relative to the sample size, np-charts directly track the count of defective items, making them simpler to interpret when dealing with uniform sample sizes. Given that the `oj` data frame maintains a constant sample size of 50 for each batch of orange juice cans, it is well-suited for demonstrating how to construct and interpret an np-chart. This chart will allow us to easily observe and analyze variations in the number of defects across the various samples, providing insights into the process's stability and quality over time.

To refine our analysis and make the np-chart more sensitive to variations in the process, we've adjusted the control limits to be within two standard deviations from the mean, rather than the typical three standard deviations. This change allows for earlier detection of potential issues, providing a more conservative approach to quality control. Additionally, we've incorporated logic into our Matplotlib code to highlight any data points that fall outside these new control limits, making it easier to identify when the process is out of control. The following code demonstrates these enhancements and their implementation:

```
>>> np_bar = oj['D'].mean()                          # Computes the mean of D, which represents the number of defective units
>>> UCL_np = np_bar + 2 * np.sqrt(np_bar)            # Computes the UCL to be two standard deviations above the mean
>>> LCL_np = np_bar - 2 * np.sqrt(np_bar)            # Computes the LCL to be two standard deviations below the mean
>>> plt.figure()                                     # Initializes a matplotlib figure, creating a blank canvas for the upcoming plot
>>> plt.plot(oj['sample'], oj['D'],
>>>          marker = 'o',
>>>          linestyle = '-',
>>>          label = 'Number of Defective Items')    # Plots D against sample using a solid line and circular markers; adds a label
>>> out_of_control = \
>>>         (oj['D'] > UCL_np) | (oj['D'] < LCL_np)  # Establishes out-of-control conditions
>>> plt.plot(oj['sample'][out_of_control],
>>>          oj['D'][out_of_control],
>>>          marker = 'o', color = 'r', linestyle = 'None',
>>>          label = 'Out of Control')               # Plots out-of-control data points in red; adds a label
>>> plt.axhline(y = np_bar,
```

```
>>>             color = 'r', linestyle = '--',
>>>             label = 'np-bar (Mean)')
>>> plt.axhline(y = UCL_np,
>>>             color = 'g', linestyle = '--',
>>>             label = 'UCL')
>>> plt.axhline(y = LCL_np,
>>>             color = 'g', linestyle = '--',
>>>             label = 'LCL')
>>> plt.fill_between(oj['sample'],
>>>                  LCL_np, UCL_np,
>>>                  color = 'g', alpha = 0.1)
>>> plt.title('np-Chart')
>>> plt.xlabel('Sample')
>>> plt.ylabel('Number of Defective Items')
>>> plt.legend(loc = 'upper left',
>>>            bbox_to_anchor = (0.1, 0.9))
>>> plt.grid(True)
>>> plt.show()
```

Draws a horizontal dashed red line at the height of np_bar on the plot; adds a label

Draws a horizontal dashed green line at the height of UCL_np on the plot; adds a label

Draws a horizontal dashed green line at the height of LCL_np on the plot; adds a label

Fills the area between the control limits across all samples with a light green color

Adds a title atop the plot

Adds a label to the x axis

Adds a label to the y axis

Adds a legend in the upper left corner of the plot, nudged downward and to the right

Adds horizontal and vertical lines to the plot

Displays the plot

The np-chart in figure 14.2 visually represents the number of defective items per sample, with control limits set at two standard deviations from the mean. It demonstrates

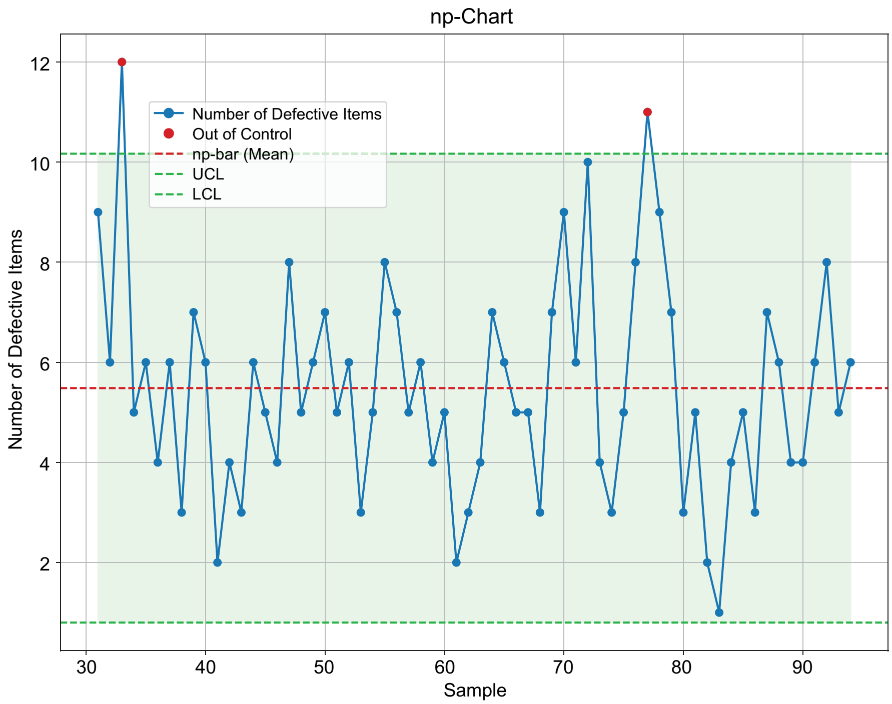

Figure 14.2 This np-chart displays the number of defective units in each sample of orange juice cans, including a pair of data points above the UCL. This type of chart is a better option than a p-chart when the sample sizes are consistent.

that even with the tighter control limits set at two standard deviations from the mean, the process is largely within control. Only a pair of data points fall above the UCL, indicating that the process is generally stable with occasional variations that may require attention. An np-chart, as shown here, is particularly effective when sample sizes are consistent across observations, offering a clear view of the number of defective items per sample. The out-of-control points suggest potential issues that could be due to special cause variation, warranting further investigation to maintain process quality. Overall, this chart is a useful tool for monitoring process stability and identifying areas that may need corrective actions to keep the process within desired quality standards.

14.2.3 c-charts

A c-chart is a type of control chart specifically designed to monitor the number of defects per unit in a process. Unlike p-charts and np-charts, which focus on the proportion or count of defective items within a sample, a c-chart tracks the total number of defects found in individual units over time. This makes c-charts particularly valuable when defects are counted rather than measured as a proportion of the whole. For instance, if you're inspecting units of a product for multiple types of defects—such as scratches, dents, or other flaws—a c-chart helps you determine whether the variation in the number of defects per unit is consistent with an in-control process or if it signals an underlying issue that needs correction. Because c-charts plot the actual count of defects rather than a proportion, they provide a straightforward way to assess the quality of a process where the number of defects is expected to be variable but within a predictable range.

We will generate a synthetic data set to demonstrate how to create and plot a c-chart. To simulate the number of defects per unit, we use a Poisson distribution, which is commonly applied in quality control to model the occurrence of rare events (like defects) over a fixed interval. This approach allows us to create a realistic data set that reflects the expected variability in a process where defects are infrequent but possible. We'll then use this data set to plot a c-chart and analyze the results:

```
>>> np.random.seed(0)
>>> units = 20
>>> defects = np.random.poisson(lam = 4,
>>>                             size = units)
>>> c_bar = np.mean(defects)
>>> UCL = c_bar + 3 * np.sqrt(c_bar)
>>> LCL = max(0, c_bar - 3 * np.sqrt(c_bar))
>>> plt.figure()
>>> plt.plot(range(1, units + 1), defects,
>>>          marker = 'o', linestyle = '-', color = 'b',
>>>          label = 'Number of Defects')
```

Sets the seed for the numpy random number generator, ensuring reproducible results

Defines the total number of units (or samples) for which defects will be simulated

Generates an array of defect counts per unit using a Poisson distribution

Computes the mean number of defects per unit

Computes the UCL to be three standard deviations above the

Computes the LCL to be three standard deviations below the mean. Note that the LCL cannot be negative and will default to zero.

Initializes a matplotlib figure, creating a blank canvas for the upcoming plot

Plots the defect counts using a solid line and circular markers; adds a label

```
>>> plt.axhline(c_bar,
>>>             color = 'r', linestyle = '--',
>>>             label = 'Center Line')
>>> plt.axhline(UCL,
>>>             color = 'g', linestyle = '--',
>>>             label = 'UCL')
>>> plt.axhline(LCL,
>>>             color = 'g', linestyle = '--',
>>>             label = 'LCL')
>>> plt.title('c-Chart')
>>> plt.xlabel('Unit Number')
>>> plt.ylabel('Number of Defects')
>>> plt.xticks(range(1, units + 1))
>>> plt.legend(loc = 'upper right',
>>>            bbox_to_anchor = (1, .95))
>>> plt.grid(True)
>>> plt.show()
```

Draws a horizontal dashed red line at the height of c_bar on the plot; adds a label

Draws a horizontal dashed green line at the height of UCL on the plot; adds a label

Draws a horizontal dashed green line at the height of LCL on the plot; adds a label

Sets the title, labels the axes, adjusts the x-axis ticks, places the legend in the upper right corner, enables the grid, and displays the plot

The c-chart in figure 14.3 shows the number of defects per unit, helping to identify any patterns or trends in the defect counts across the 20 units. It shows the number of defects identified across 20 units, with the CL (center dashed line) representing the average number of defects per unit. The chart also includes a UCL and a LCL (defaulted to zero because it can't be negative), marked by the upper and lower

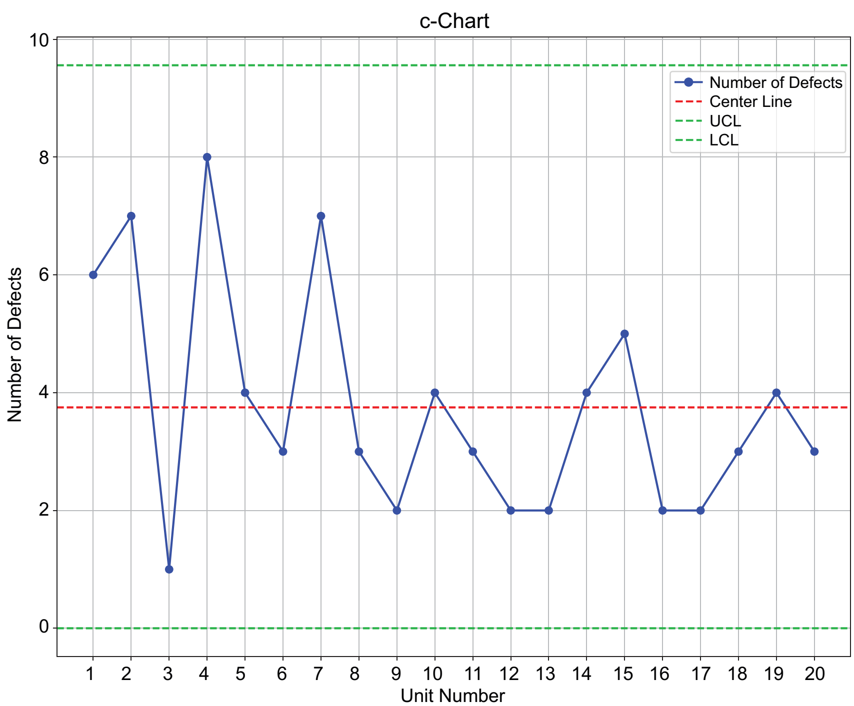

Figure 14.3 This c-chart displays the number of defects per unit for 20 units. All data points fall within the control limits, indicating a stable process.

dashed lines, which are set to monitor process variation. All data points fall within these control limits, indicating that the process is stable and under control. Although there is some variability in the number of defects, with several data points above the CL, no points exceed the control limits, suggesting that the variations observed are within the expected range of process behavior. This consistency reinforces that the process is operating as expected, without any signals of special cause variation.

14.2.4 g-charts

A g-chart is a type of quality control chart specifically designed to monitor the number of units or opportunities between occurrences of a specific event, such as defects or failures. Unlike p-charts, np-charts, and c-charts, which focus on the proportion or count of defects in a sample, the g-chart tracks the gaps or intervals between such events. This approach is particularly valuable when the events of interest are relatively rare and it is more informative to measure the time or number of units between occurrences. G-charts are crucial in identifying patterns that suggest whether a process is improving or deteriorating over time, offering a different perspective on process stability and performance. As the last type of control chart for attributes that we will explore, the g-chart provides a unique and powerful tool for quality control, especially in processes where defects are infrequent but critical to monitor.

To demonstrate how to create and plot a g-chart, we'll generate a synthetic data set that simulates the number of units between defects in a process. This type of data is ideal for a g-chart, which is used to monitor the intervals between occurrences of rare events, such as defects or failures. By applying this data to the g-chart, we can analyze the process's performance and detect any unusual patterns or trends that might indicate process instability. The following code will create the necessary data and plot the g-chart to show these intervals:

Creates a pandas data frame, df, with a defect sequence and the corresponding number of units produced between defects

Computes the average number of units between defects from the units_between_defects column in df and stores it in g_bar

Computes the UCL to be three standard deviations above the mean

Computes the LCL to be three standard deviations below the mean. Note that the LCL cannot be negative and will default to zero.

Initializes a matplotlib figure, creating a blank canvas for the upcoming plot

Plots units_between_defects against defect_sequence, using a solid line and circular markers; adds a label

```
>>> data = {
>>>     'defect_sequence': np.arange(1, 21),
>>>     'units_between_defects': [50, 30, 45, 60, 55, 35, 40,
>>>                               70, 65, 80, 75, 50, 85, 90,
>>>                               95, 60, 55, 100, 105, 110]
>>> }
>>> df = pd.DataFrame(data)
>>> g_bar = df['units_between_defects'].mean()
>>> UCL = g_bar + 3 * np.sqrt( \
>>>     g_bar * (g_bar + 1))
>>> LCL = max(g_bar - 3 * np.sqrt( \
>>>     g_bar * (g_bar + 1)), 0)
>>> plt.figure()
>>> plt.plot(df['defect_sequence'], df['units_between_defects'],
>>>          marker = 'o', linestyle = '-',
>>>          label = 'Units Between Defects')
```

```
>>> plt.axhline(y = g_bar,
>>>             color = 'r', linestyle = '--',
>>>             label = 'g-bar (Mean)')
>>> plt.axhline(y = UCL,
>>>             color = 'g', linestyle = '--',
>>>             label = 'UCL')
>>> plt.axhline(y = LCL,
>>>             color = 'g', linestyle = '--',
>>>             label = 'LCL')
>>> plt.fill_between(df['defect_sequence'], LCL, UCL,
>>>                  color = 'g', alpha = 0.1)
>>> plt.title('g-Chart')
>>> plt.xlabel('Defect Sequence')
>>> plt.ylabel('Units Between Defects')
>>> plt.legend()
>>> plt.grid(True)
>>> plt.show()
```

Draws a horizontal dashed red line at the height of g_bar on the plot; adds a label

Draws a horizontal dashed green line at the height of UCL on the plot; adds a label

Draws a horizontal dashed green line at the height of LCL on the plot; adds a label

Fills the area between the control limits across all samples with a light green color

Sets the title, labels the axes, adds a legend, enables the grid, and displays the plot

The g-chart in figure 14.4 shows the number of units between defects over the defect sequence (see figure 14.4), with the defect sequence plotted on the *x* axis and the units between defects on the *y* axis. The center dashed line represents the average (g-bar) of units between defects, and the upper and lower dashed lines denote the UCL and LCL. This chart is significant in monitoring the frequency of defects over time, helping to

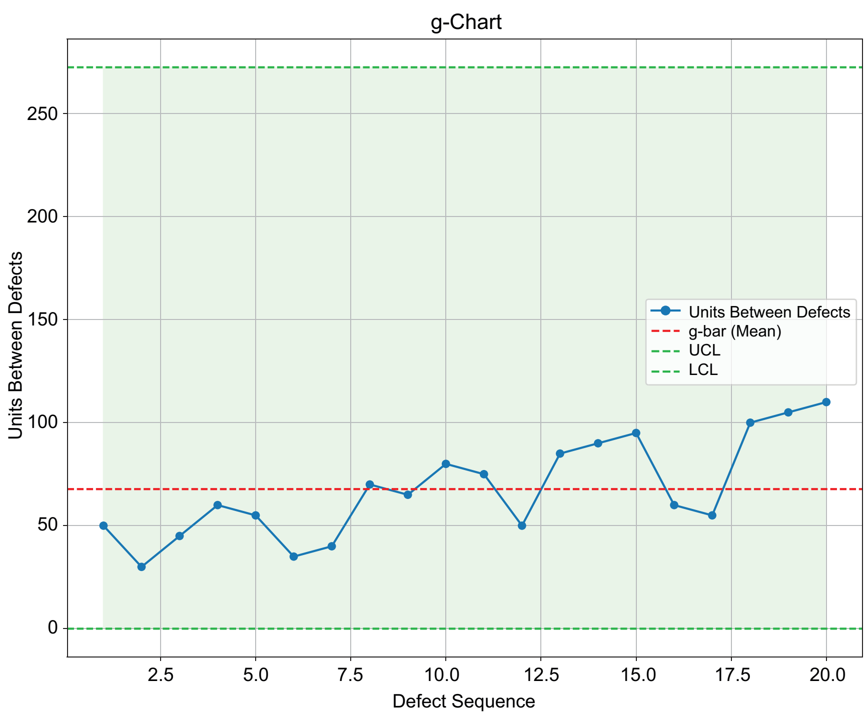

Figure 14.4 This g-chart displays the number of units between defects; it highlights process stability, with all points falling within the control limits. Additionally, it shows an improving process, as the number of manufactured units between defects is trending upward.

identify whether a process is consistently producing a certain number of units between defects or if there are any unusual patterns that indicate the process is out of control. In this case, all the data points fall within the control limits, suggesting that the process is stable and within expected variability. The g-chart's value add lies in its ability to highlight shifts in process reliability, providing a visual tool to detect and address issues before they lead to significant quality problems.

So far, we've explored the most common control charts for attributes, including p-charts, np-charts, c-charts, and g-charts. Using the Matplotlib library, we demonstrated how to create these charts step by step, providing the necessary tools and knowledge to plot and interpret them effectively. By understanding these charts, you're now equipped to monitor and control processes where the quality of products or services is measured in terms of discrete attributes, such as the number of defects or the proportion of defective items.

As we move forward to control charts for variables, it's important to note that although these charts differ in what they measure, the underlying concepts remain consistent. Due to the repetitive nature of the plotting process and because the essential techniques have already been thoroughly covered, we will not be sharing the Matplotlib code for each subsequent chart. Instead, the focus will be on understanding the specific applications and interpretations of these charts, using the skills you have already developed. This approach allows us to avoid redundancy while ensuring that you grasp the critical aspects of quality control for both attributes and variables.

14.3 Control charts for variables

Control charts for variables are crucial tools in quality management, designed to monitor and control processes where the data is continuous and can be measured on a numerical scale. Unlike control charts for attributes, which focus on counting the occurrence of defects, variable control charts track the actual values of a process characteristic to assess its stability and consistency over time. These charts provide a more detailed understanding of process performance, allowing for the detection of small shifts or trends that might indicate underlying issues. In this section, we will delve into some of the most commonly used variable control charts, demonstrating their construction, interpretation, and application to ensure processes remain under control and within specified tolerance limits.

Control charts for variables are typically used in combination, depending on the specific needs of the process being monitored. For example, an x-bar chart might be used alongside an r-chart to monitor both the process mean and variability. Although multiple charts can be employed together to provide a comprehensive view of process performance, in some cases, a single chart may be sufficient to monitor the most critical aspect of the process. The choice of charts depends on the complexity of the process and the specific characteristics that need to be controlled

We'll be using a pandas data frame named `pistons` as our data source throughout, which contains the diameter measurements of 200 piston rings produced in a manufacturing process. These measurements are equally divided across 40 batches, or samples. Before working with the data, the `pd.read_csv()` method must be called to import the data set into Python. A subsequent call to the `info()` method provides a summary of the data frame's essential details:

```
>>> print(pistons.info())
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 2 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   diameter  200 non-null    float64
 1   sample    200 non-null    int64
dtypes: float64(1), int64(1)
memory usage: 3.3 KB
None
```

Each measurement of piston ring diameter is represented as a single row in the `pistons` data frame, with five rows corresponding to each batch or sample. Here's a brief overview of its attributes:

- `diameter`—Represents the measured diameter of each piston ring, recorded as a floating-point number (or real number with fractional parts). The measurements are taken in millimeters. There are no null values.
- `sample`—A numeric variable that uniquely identifies each batch of measurements, sequentially numbered from 1 through 40. Each sample number corresponds to five individual diameter measurements.

The `describe()` function in pandas provides a quick statistical summary of numerical data in a data frame. When applied to the `diameter` variable in the pistons data frame, it will return key metrics such as the mean, standard deviation, and range, giving us a comprehensive overview of the distribution of piston ring diameters:

```
>>> stats = pistons['diameter'].describe()
>>> print(stats)
count    200.000000
mean      74.003605
std        0.011417
min       73.967000
25%       73.995000
50%       74.003000
75%       74.010000
max       74.036000
Name: diameter, dtype: float64
```

The summary statistics for the `diameter` variable reveal minimal differences between the minimum (73.967 millimeters) and maximum (74.036 millimeters) values, with a very small standard deviation of approximately 0.011 millimeters. This suggests that the piston ring manufacturing process might be highly consistent, producing rings with very little variation in diameter. As we move forward to create control charts for variables, these small variations indicate that the process is (presumably) stable and tightly controlled, which is a positive sign for maintaining quality standards.

Finally, a call to the `head()` method gives us a glimpse at the top 10 observations in the `pistons` data frame:

```
>>> print(pistons.head(10))
   diameter  sample
0    74.030       1
1    74.002       1
2    74.019       1
3    73.992       1
4    74.008       1
5    73.995       2
6    73.992       2
7    74.001       2
8    74.011       2
9    74.004       2
```

These initial rows confirm the consistent measurements within each sample, reinforcing the earlier observation of minimal variation in the manufacturing process. With the foundational data now explored and understood, we are ready to transition to our first control chart: the x-bar chart, which will help us visualize the stability and consistency of the piston ring manufacturing process.

14.3.1 x-bar charts

An x-bar chart is used to monitor the central tendency of a process over time. Specifically, it tracks the mean of a sample, or subgroup, across a series of measurements, allowing quality practitioners to assess whether the process remains stable and predictable. In the context of our piston ring manufacturing process, the x-bar chart will help us determine whether the average diameter of piston rings remains consistent across different batches, ensuring that the process is producing parts within the desired specifications.

The x-bar chart is constructed by first calculating the mean diameter of piston rings within each sample or batch. This is done using the `groupby()` method in Python, which groups the data by sample and then calculates the mean diameter for each group, resulting in a series of sample means. The overall mean, or the central line on the chart, is then determined by taking the average of these sample means. This overall mean represents the expected average diameter if the process is in control.

To assess the variability of the sample means, we calculate the standard deviation of the sample means. This measure of dispersion is then used to establish the control limits, which are typically set at three standard deviations above and below the overall mean. If a sample mean falls outside these limits, it signals that the process may be experiencing an assignable cause of variation, which requires further investigation. The x-bar chart thus provides a clear visual representation of the process's stability, highlighting any shifts or trends that may indicate a need for corrective action (see figure 14.5). By closely monitoring the x-bar chart, manufacturers can ensure that their processes remain consistent, producing high-quality products that meet the desired specifications.

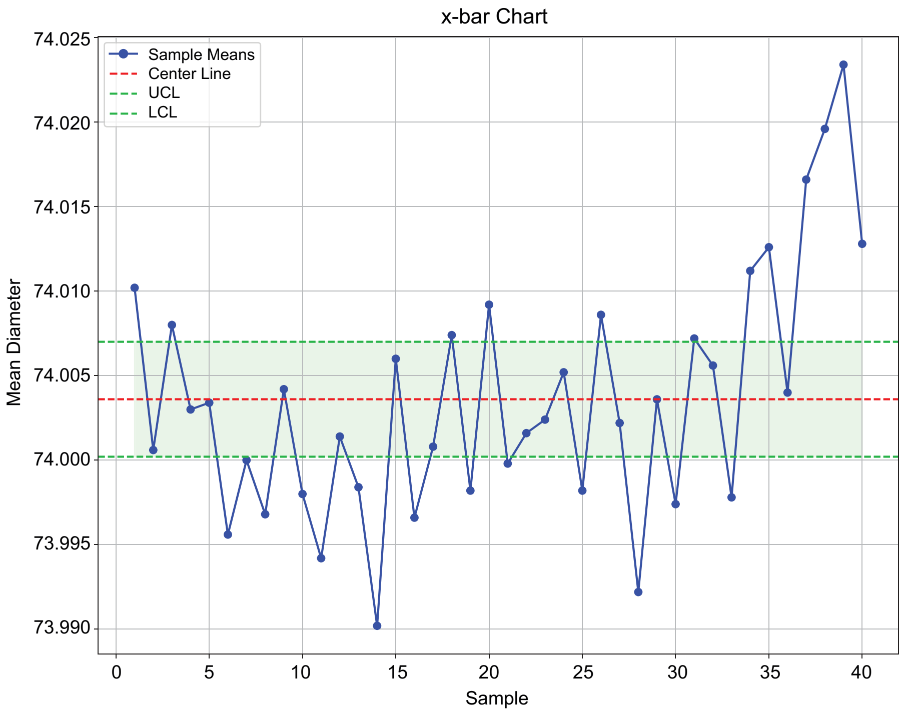

Figure 14.5 This x-bar chart displays the mean diameter of piston rings across 40 samples, with control limits set at three standard deviations from the overall mean. The chart helps monitor the consistency of the manufacturing process, ensuring that the average diameter remains within acceptable limits, signaling that the process is stable and in control.

This x-bar chart shows the mean diameters of piston rings across 40 samples, with the control limits set at three standard deviations above and below the overall mean. Several data points fall above or below the upper and lower control limits, respectively, indicating that the process may be out of control. Specifically, the data points toward the end of the sequence exhibit a rising trend, with several points exceeding the UCL. This suggests that there might be a systematic issue in the manufacturing process that requires investigation and corrective action to bring the process back under control and ensure consistent product quality. The x-bar chart is crucial in identifying these variations and maintaining the stability of the process.

14.3.2 r-charts

An r-chart, also known as a range chart, is used to monitor the variability within a process. Unlike x-bar charts, which track the mean values of a process, r-charts focus on the spread of data within each sample, making them particularly useful for detecting changes in the dispersion of the process over time. The range of a sample is calculated as the difference between the maximum and minimum values within that sample. By plotting these ranges on an r-chart, we can visualize how the variability within the process shifts, helping us to identify any points where the process might be going out of control.

The key measures in an r-chart include the average range, which is calculated as the mean of all the individual sample ranges, and the control limits, which are set at three standard deviations above and below the average range. The UCL and LCL help determine whether the variability of the process is within acceptable bounds. If the range for any sample exceeds the UCL or falls below the LCL, it signals a potential issue with the process that warrants further investigation.

Displaying these measures on an r-chart allows for a straightforward interpretation of the data (see figure 14.6). When all the points fall within the control limits, it suggests that the process variability is stable and consistent. However, if points fall outside these limits or show a pattern, such as increasing or decreasing trends, it indicates that the process variability may be changing, which could affect the overall quality of the product. By using r-charts in conjunction with x-bar charts, quality control practitioners can gain a comprehensive understanding of both the central tendency and variability of the process, enabling them to take timely corrective actions to maintain process control.

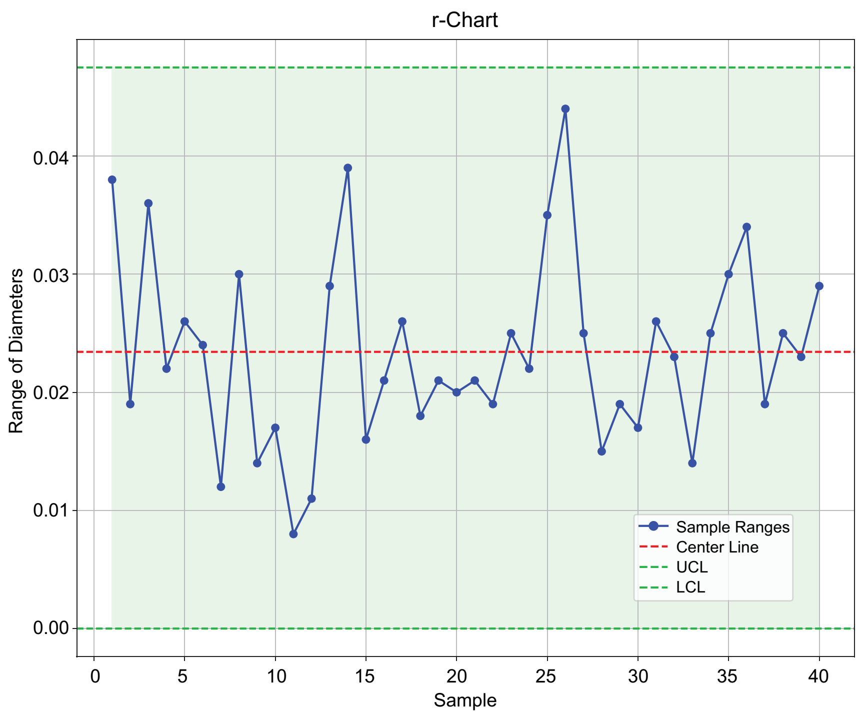

Figure 14.6 This r-chart displays the range of piston ring diameters across 40 samples, with control limits set at three standard deviations from the mean range. The chart helps monitor the variability in the process, indicating that all of the sample ranges fall within the control limits, suggesting the variability in the process is stable.

The r-chart shows that the variability in the diameter measurements of the piston rings appears to be stable, with all data points falling within the control limits. This indicates that the process's range of variability is under control. However, it is important to cross-reference these results with the x-bar chart, where several points were out of control. The discrepancy between the x-bar chart and the r-chart suggests that although the overall process variability is stable, there might be shifts in the process mean that need to be addressed. This could indicate issues such as tool wear or calibration drift that affect the central tendency without significantly affecting the range of measurements. Further investigation is necessary to ensure that the process remains within both range and mean control.

14.3.3 s-charts

An s-chart, or standard deviation chart, is a critical tool in quality control that monitors the variability within a process by tracking the standard deviations of sample measurements over time. Unlike the r-chart, which focuses on the range of data within each sample, the s-chart provides a more accurate representation of variability, especially when sample sizes are large. This chart is particularly valuable when consistent precision is required, as it helps identify shifts or drifts in the process that could lead to defects or non-conformance.

To construct an s-chart, we first need to calculate the standard deviation of the measured characteristic (in this case, the diameter of piston rings) for each sample. This is done by grouping the data by sample and then calculating the standard deviation of the measurements within each group. The average of these standard deviations across all samples serves as the central line of the chart, or the expected average variability if the process is in control.

Control limits are established to help identify when the process variability is outside of acceptable bounds. The control limits are derived from the following formulas:

$$\mathrm{UCL} = \bar{s} + 3\left(\frac{\bar{s}}{c4}\right)$$

$$\mathrm{LCL} = \bar{s} - 3\left(\frac{\bar{s}}{c4}\right)$$

where

- $\bar{s}$ is the average standard deviation.
- $c4$ is a correction factor that accounts for bias in the standard deviation estimates. It is derived by taking the square root of the square root of the ratio of 2 to the sample size minus 1.

If the standard deviations of the samples fall outside these control limits, it indicates that the process variability is out of control and requires investigation.

An s-chart therefore offers a precise method for monitoring the consistency of a process, ensuring that variations remain within predictable limits (see figure 14.7). It is particularly useful in manufacturing processes where maintaining a consistent level of quality is crucial. By understanding the distribution of standard deviations and their control limits, quality managers can detect early signs of process instability and take corrective actions before significant quality issues arise.

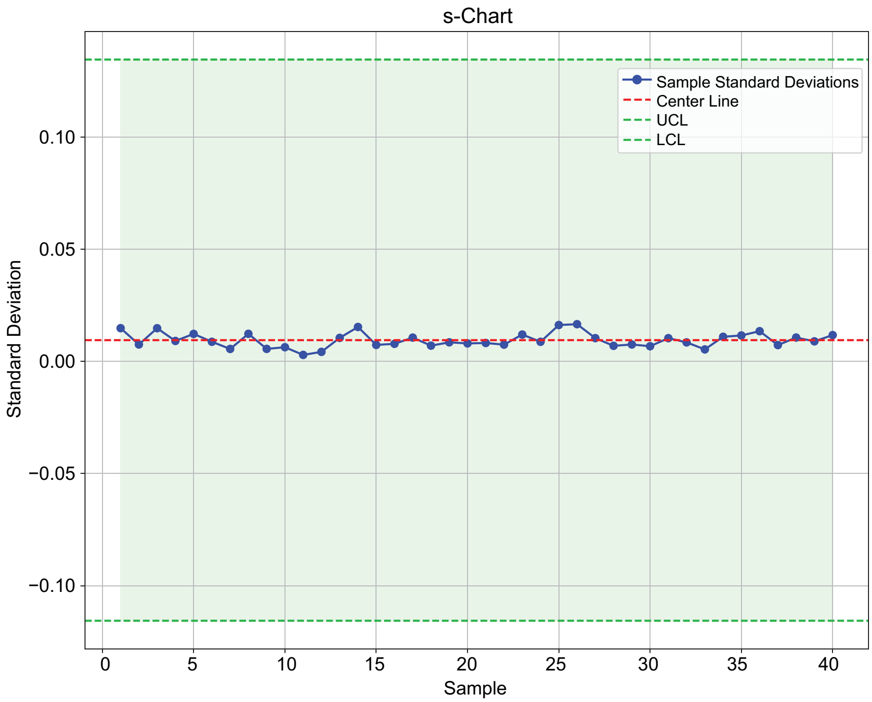

Figure 14.7 This s-chart displays the sample standard deviations for each batch of piston rings, with the CL representing the average standard deviation across all samples. The control limits (UCL and LCL) are set to identify any variations in the process, ensuring the process remains stable within the expected range.

The s-chart shows the standard deviation of piston ring diameters across 40 samples, with the center dashed line representing the average standard deviation (CL). All data points are well within the control limits (UCL and LCL), indicating that the variation in the process is stable and consistent across the samples. This stability suggests that the process is in control, with no signs of significant variability that would indicate a problem. The close clustering of the points near the CL further reinforces that the process is producing consistent results, with minimal variation from sample to sample.

Reconciling the results of the s-chart with the x-bar chart reveals an interesting dynamic in the manufacturing process. Although the s-chart provides another indication that the variability in the piston ring diameters is stable and within control limits, the x-bar chart shows several data points outside the control limits, suggesting that the process mean is not consistently centered. This discrepancy implies that although the process variation is well-controlled, the process may still be drifting from the target mean, leading to potential issues with meeting specific quality standards. Therefore, even though the variability is stable, the process might require adjustments to bring the mean back within control limits, ensuring that the product consistently meets the desired specifications.

14.3.4 I-MR charts

An I-MR chart is a powerful tool in statistical process control, particularly useful when measurements are taken individually rather than in subgroups. This type of control chart is ideal for monitoring processes where data points are not grouped into batches but instead are collected in a time-ordered sequence. The I-MR-chart is composed of two distinct charts: the individual (I) chart and the moving range (MR) chart. The I-chart tracks individual data points over time, allowing for the detection of shifts or trends in the process mean, whereas the MR-chart monitors the variability between consecutive data points, providing insights into the stability of the process variation.

To construct an I-MR chart, we first calculate the mean of the individual measurements, which forms the CL of the I-chart. The control limits for the I-chart are then determined by adding and subtracting three standard deviations of the individual data points from the process mean. This gives us the UCL and LCL, which define the range within which the data points should fall if the process is in control.

Next, we calculate the MR, which is the absolute difference between consecutive data points. The average of these moving ranges forms the CL of the MR-chart. The control limits for the MR-chart are derived using constants $D4$ and $D3$, which are specifically chosen based on the number of data points used to calculate the moving range. The control limits for the MR-chart are calculated from the following formulas:

$$\mathrm{UCL} = (D4)\left(\bar{R}\right)$$

$$\mathrm{LCL} = (D3)\left(\bar{R}\right)$$

where

- $D4$ is a constant derived from statistical tables used specifically to calculate the UCL in quality control charts, based on the sample size; it equals 3.267.
- $D3$ is another constant derived similarly to $D4$; equals 0.
- $\bar{R}$ represents the average of the MRs.

These limits provide a visual reference for the variability in the process; if the MR exceeds these limits, it signals potential instability in the process.

The I-MR chart is particularly valuable because it enables the detection of both small shifts in the process mean and changes in process variability that might not be evident with other types of control charts (see figure 14.8). By analyzing the I-chart and the MR-chart together, quality control practitioners can gain a deeper understanding of the process's overall stability. If either chart shows points outside the control limits, it indicates that the process may be out of control, necessitating further investigation and potential corrective actions. This makes the I-MR chart an essential tool for maintaining high quality standards in processes where individual data points are critical to monitoring and controlling production.

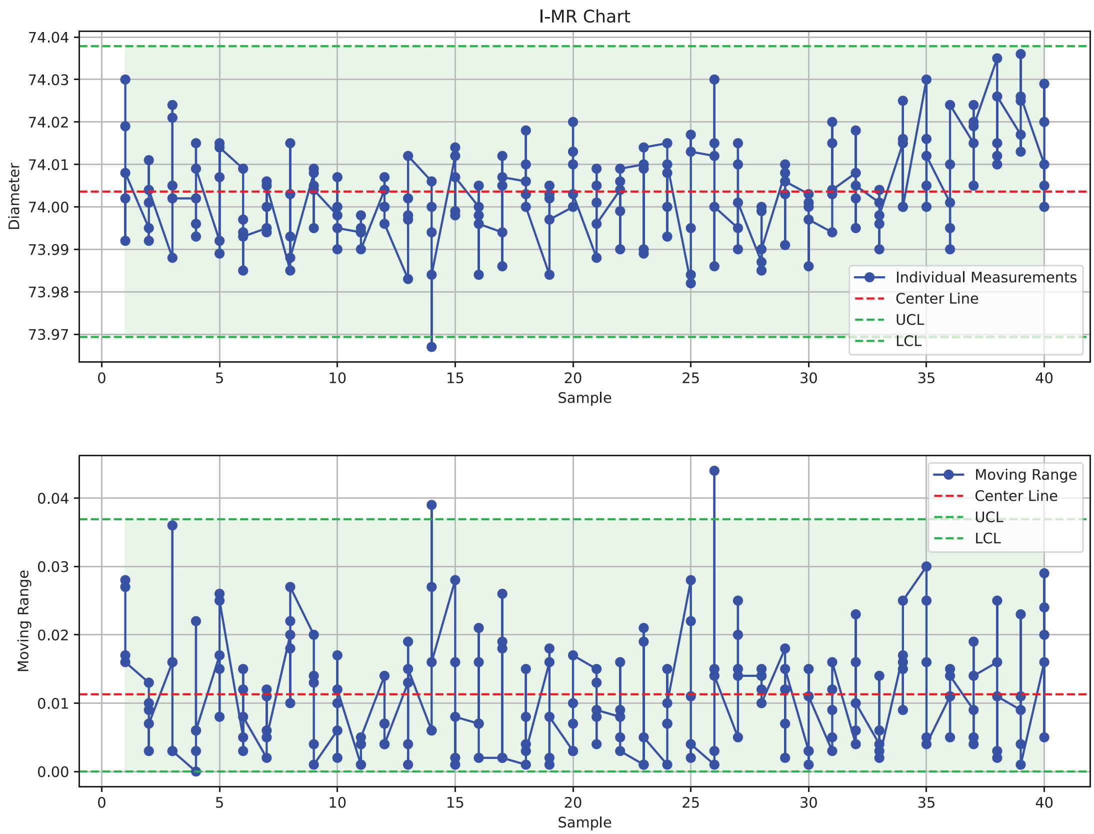

Figure 14.8 An I-MR chart, with the I-chart on top and the MR-chart below. The I-chart plots individual piston ring diameters for each sample, highlighting variation around the process mean. UCL and LCL lines indicate acceptable limits. The MR-chart tracks the moving range, or the difference between successive measurements, reflecting process variability. Lines connect data points from one sample to the next, illustrating changes in measurements and variability across samples.

The I-MR-chart provides a detailed view of both individual measurements and the variability within a process. The top part of the chart, the I-chart, displays individual piston ring diameters across 40 samples. Each point represents a single measurement, and the lines connect these points from one sample to the next. The chart's UCL and

LCL define the control limits, within which most data points should ideally fall. In this case, the measurements are generally within control, but there are some points close to the limits, indicating potential areas to monitor for process drift.

The bottom part of the chart, the MR-chart, shows the moving range, which is the absolute difference between consecutive measurements. This chart helps to detect sudden shifts in process variability that might not be visible in the individual measurements alone. The MR-chart reveals that the variability is mostly stable, with the moving ranges staying within the control limits. Together, the I-chart and MR-chart offer a comprehensive picture of the process, making it easier to identify any unusual patterns that could indicate a problem, such as a shift in the process mean or an increase in variability. This dual-chart approach is valuable in monitoring processes where individual data points and their variations need to be closely controlled.

14.3.5 EWMA charts

An exponentially weighted moving average (EWMA) control chart is a powerful tool used in quality control to monitor the performance of a process over time. Unlike traditional control charts that treat each data point equally, the EWMA chart places more emphasis on recent observations while still considering past data. This approach makes it particularly sensitive to small shifts in the process mean, which can be crucial in detecting trends or gradual changes that might go unnoticed with other types of control charts.

The key measures displayed in an EWMA chart, such as the CL (EWMA mean) and control limits (UCL and LCL), are derived using a smoothing constant, often denoted as λ. The smoothing constant controls how much weight is given to the most recent data point versus the historical data. In the calculation process, the first EWMA value is set to the first measurement in the data set. Each subsequent EWMA value is calculated as a weighted average of the current data point and the previous EWMA value, with λ determining the weighting. This process ensures that the EWMA line is less sensitive to noise and more reflective of genuine shifts in the process.

The control limits on the EWMA chart are calculated similarly to those in traditional control charts but take into account the smoothing constant. The formula for the control limits includes the standard deviation of the original data and adjusts it using λ, making the limits narrower and more responsive to recent changes. The UCL and LCL help determine whether the process is staying within acceptable boundaries or if there are signals that the process is shifting. If the EWMA crosses these control limits, it indicates that the process may be moving out of control, necessitating further investigation and possible corrective actions (see figure 14.9).

The EWMA chart offers a powerful way to monitor subtle shifts in a process over time, particularly when dealing with processes that require early detection of trends or small shifts. In this chart, each point represents the EWMA of the piston ring diameters across 40 samples, giving more weight to recent observations. The CL, or the overall EWMA mean, serves as a benchmark against which individual data points are compared. The UCL and LCL define the range within which the process is expected to operate under normal conditions. In this EWMA chart, most data points fall within these control limits, indicating that the process is generally in control. However, as we approach the latter samples, there's a noticeable upward trend, similar to what our x-bar chart first revealed, suggesting a potential shift in the process. This could be an early warning of a systematic change, warranting further investigation to determine whether corrective actions are necessary to maintain process stability.

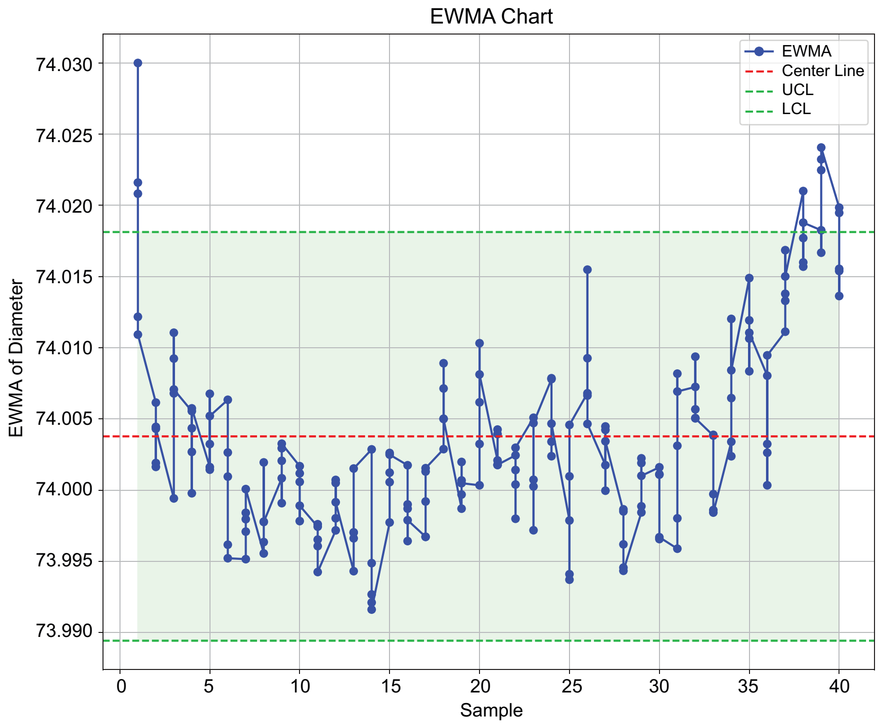

Figure 14.9 This EWMA chart displays the exponentially weighted moving average of piston ring diameters across 40 samples. The chart shows how the process mean evolves over time, with recent data points having more influence on the chart's trajectory than older ones. The CL represents the overall EWMA mean, whereas the UCL and LCL indicate the expected range of variation. Data points within these limits suggest the process is under control, whereas any point outside may signal a shift or trend that requires further investigation.

As we close this book, it's our hope that the statistical and other quantitative techniques explored have equipped you with a solid foundation for tackling real-world challenges. These methods are not just tools but a lens through which you can make informed decisions and drive meaningful outcomes. Thank you for embarking on this journey and continuing to develop your expertise.

Summary

- Understanding key quality control measures, such as control limits, mean, and standard deviation, is crucial for interpreting control charts effectively. These measures provide the foundation for analyzing process stability and identifying areas for improvement.
- Control charts for attributes, including p-charts, np-charts, c-charts, and g-charts, are essential tools for monitoring categorical data. These charts help in tracking defects, proportions, and other attribute-based metrics to ensure processes remain within acceptable limits.
- Control charts for variables, which focus on continuous data such as measurements and process variations, were also covered. These charts, including x-bar charts, r-charts, s-charts, I-MR charts, and EWMA charts, are valuable for assessing the precision and consistency of manufacturing processes.
- It's crucial to use multiple control charts to comprehensively assess whether a process is in or out of control. Relying on just one chart can lead to incomplete or misleading conclusions, as different charts highlight different aspects of process performance. By using a combination of charts, such as x-bar and r-charts, you gain a more accurate and complete understanding of process stability and potential areas for improvement.