“Over 560 billion texts are sent every month. Texting could lead to De Quervain’s disease. Social media is very popular. This tweet is out of

The first statement above is from https://www.textrequest.com/blog/texting-statistics-answer-questions/. The second statement may be true, but pregnancy could also lead to the same disease. I have no source for the third statement, but I believe it to be true. The fourth statement ended due to a 140 character limit, but may have continued “. . . this world” or “. . . characters.”

So much valuable, trivial, fascinating, misleading information can be conveyed in a tweet. Remove the character restriction and the magnitude of output clearly increases. World-altering “real news” speeds by us alongside “fake news.” Associated with both comes secondary data about the source and the recipient. Textrequest.com does not know I visited their website (at least I do not think they do), but they know my IP address and other things associated with my browsing. My wearable fitness device died, but my inexpensive cell phone knows how far and fast I bicycle, as well as where I went. The hotel I am staying at knows when I entered and left the fitness center. You may or may not believe Big Brother is here, but there is no denying the era of Big Data is.

What does this have to do with the ASQ Statistics Division? Everything. We are quality professionals, dedicated to making processes more efficient and effective. We are data driven, using basic statistical and visual tools to analyze data and convert them into information. We use sophisticated algorithms to perform predictive analytics. Whether the analysis is basic or complex, it is not very useful if the data are inaccurate or misleading. Who is better prepared, or should be, to help ensure that the massive amounts of data being generated are representing what they should, and are being used to their fullest potential? We may not be the only folks that have an important role here, but we certainly are key players in Big Data, data science, and analytics. More on this in a moment.

As I write this I am in Charlotte, NC for the 2017 World Conference on Quality Improvement. WCQI is always an exciting time for those of us in the quality profession. Not only do we get to hear from great speakers, but we get to spend time with colleagues we do not see on a regular basis. We also make new connections, expanding our quality network. Much of this networking occurs at the Statistics Division booth. Having “worked the booth” for multiple years, I am used to visitors walking up, noticing what booth they are near, and making comments that generally fall into one of three categories.
Submission Guidelines

Mini-Paper
Interesting topics pertaining to the field of statistics; should be understandable by non-statisticians with some statistical knowledge. Length: 1,500-4,000 words.

Feature
Focus should be on a statistical concept; can either be of a practical nature or a topic that would be of interest to practitioners who apply statistics. Length: 1,000-3,000 words.

General Information
Authors should have a conceptual understanding of the topic and should be willing to answer questions relating to the article through the newsletter. Authors do not have to be members of the Statistics Division. Submissions may be made at any time to newsletter@asqstatdiv.org.

All articles will be reviewed. The editor reserves discretionary right in determination of which articles are published. Submissions should not be overly controversial. Confirmation of receipt will be provided within one week of receipt of the email. Authors will receive feedback within two months. Acceptance of articles does not imply any agreement that a given article will be published.

Disclaimer
The technical content of material published in the ASQ Statistics Division Newsletter may not have been refereed to the same extent as the rigorous refereeing that is undergone for publication in Technometrics or JQT. The objective of this newsletter is to be a forum for new ideas and to be open to differing points of view. The editor will strive to review all articles and to ask other statistics professionals to provide reviews of all content of this newsletter. We encourage readers with differing points of view to write to the editor and an opportunity to present their views via a letter to the editor. The views expressed in material published in this newsletter represents the views of the author of the material, and may or may not represent the official views of the Statistics Division of ASQ.

Vision, Mission, and Strategies of the ASQ Statistics Division

Vision
The ASQ Statistics Division promotes innovation and excellence in the application and evolution of statistics to improve quality and performance.

Mission
The ASQ Statistics Division supports members in fulfilling their professional needs and aspirations in the application of statistics and development of techniques to improve quality and performance.

Strategies
1. Address core educational needs of members
   - Assess member needs
   - Develop a “base-level knowledge of statistics” curriculum
   - Promote statistical engineering
   - Publish featured articles, special publications, and webinars

2. Build community and increase awareness by using diverse and effective communications
   - Webinars
   - Newsletters
   - Body of Knowledge
   - Web site
   - Blog
   - Social Media (LinkedIn)
   - Conference presentations (Fall Technical Conference, WCQI, etc.)
   - Short courses
   - Mailings

3. Foster leadership opportunities throughout our membership and recognize leaders
   - Advertise leadership opportunities/positions
   - Invitations to participate in upcoming activities
   - Student grants and scholarships
   - Awards (e.g., Youden, Nelson, Hunter, and Bisgaard)
   - Recruit, retain and advance members (e.g., Senior and Fellow status)

4. Establish and Leverage Alliances
   - ASQ Sections and other Divisions
   - Non-ASQ (e.g. ASA)
   - CQE Certification
   - Standards
   - Outreach (professional and social)

Updated October 19, 2013

The Statistics Division was formed in 1979 and today it consists of both statisticians and others who practice statistics as part of their profession. The division has a rich history, with many thought leaders in the field contributing their time to develop materials, serve as members of the leadership council, or both. Would you like to be a part of the Statistics Divisions’ continuing history? Feel free to contact chair@asqstatdiv.org for information or to see what volunteer opportunities are available. No statistical knowledge is required, but a passion for statistics is expected.
The first category will be familiar to most of you. “Oh. Statistics. Yeah . . . I had a stats class in school. I’m not a numbers person, but I survived.” We booth workers, having received this ringing endorsement, engage in some lively banter with our guests about how interesting and useful the field of statistics is. Our visitors thank us for our time, grab some Statistics Division sticky notes (always useful) and head down the aisle. A second category includes the many members that stop by to say hello and hear what is happening with the Division. “Where is the Fall Technical Conference this year? Do you have another great lineup of speakers?” The third category falls in between the other two. In my opinion, the conversation with these visitors has morphed from “So what do you people do?” to “You must be heavily involved in Big Data, the Internet of Things, and data analytics, right?”

While we clearly have members involved in this area, and we will sponsor a session on it at this year’s FTC (www.falltechnicalconference.org), I cannot say we are heavily involved at this point. But I believe we do need to be heavily involved in this space. Accordingly, we have begun laying the groundwork for our new technical community in Big Data/data science/analytics. We need to solidify a mission and strategic goals, but I envision members of this community ranging from those that want to keep abreast of developments in the field without being hands on, to technical experts developing and using cutting edge quality improvement tools on massive data sets. All will be welcome.

Our time table is to have the group up and running in the third quarter of 2017. I will be in touch on our progress via this digest in the Fall as well as through the LinkedIn group. If you have suggestions about this or other matters, please let me know!

Welcome to the second issue of Statistics Digest for 2017. This issue’s Mini-Paper and Feature both pertain to statistical process control. We have Forrest Breyfogle III providing us with a method for using control charts for management reporting and Ryan Monson presenting a case study in which things went wrong in the implementation of statistical process control and the lessons that were learned.

This issue’s DoE column is “Sudoku and Design of Experiments” by Bradley Jones. Jack B. ReVelle’s Stats 101 column is a discussion of “Control Charts,” and Mark Johnson tells us what “Membership in a Technical Advisory Group” is like in his Statistics InSide-Out column. I’m pleased to announce we will be running a new column, Hypothesis Testing, written by Jim Frost. Jim has extensive experience in academic research and consulting projects. He’s been performing statistical analysis on-the-job for over 20 years. For 10 of those years, he was with a statistical software company helping others make the most out of their data. He loves sharing the joy of statistics and has his own website to help with this process: statisticsbyjim.com. His first Statistics Digest column is “Introduction to Hypothesis Testing Terminology.” Donald J. Wheeler discusses “Finding the Signals” in his SPC column.

Nominations for FTC Student Grant and FTC Early Career Grant

Nominations for FTC Student Grant and FTC Early Career Grant are due Setmember 1st 2017.

For details, please go to http://asq.org/statistics/about/awards-statistics.html
MINI-PAPER
Performance Metric Reporting at the 30,000-Foot-Level: Resolving Issues with $x$ and $R$ Control Chart and Process Capability Indices Reporting
by Forrest Breyfogle III
CEO, Smarter Solutions Inc.

Wikipedia states:

Control charts are a statistical process control tool used to determine if a manufacturing or business process is in a state of control. If analysis of the control chart indicates that the process is currently under control (i.e., is stable with variation only coming from sources common to the process), then no corrections or changes to process control parameters are needed or desired. In addition, data from the process can be used to predict the future performance of the process. If the chart indicates that the monitored process is not in control, analysis of the chart can help determine the sources of variation, as this will result in degraded process performance. A process that is stable but operating outside of desired (specification) limits (e.g., scrap rates may be in statistical control but above desired limits) needs to be improved through a deliberate effort to understand the causes of current performance and fundamentally improve the process.

Many organizations have benefited from control charts since the tool was invented by Walter A. Shewhart while working for Bell Labs in the 1920s; however, over the years it seems to me that there has been a decline in the tool’s usage. The question that one might ask is what could drive a usage-frequency decrease?

One potential underlying reason for this lack-of-usage question is that control-charting usages can be violating underlying mathematical presumptions for the chart’s application. The misalignment of control charting math with what is needed for a process stability assessment can result in firefighting common-cause variation as though it were special cause. If this were the case, an organization might conclude after some period of time that control charting techniques were not beneficial.

In addition to assessing process stability, one should also note that Wikipedia also states “data from the process (i.e., when stable) can be used to predict the future performance of the process.” A follow-up question to this statement might then be: how should this prediction statement be determined and reported to others?

This paper addresses the above-described potential control charting assumptions and an approach to provide a prediction statement. Focus in this article will be given to data that would typically be charted using an $x$ and $R$ control chart; however, many of the basic concepts also apply to individuals charts, p-charts, c-charts, and u-charts.

This document describes a 30,000-foot-level alternative reporting approach\(^1\) that addresses how to:

- Create control charts so that the chart-creation mathematics is consistent with a team’s belief as to what fundamentally should be considered the source of common-cause variability and special-cause occurrence(s).
- Determine and report-out a process prediction statement whenever a process is stable when viewed from the 30,000-foot-level. This prediction statement is to be written so that everyone can easily interpret its meaning and be provided even when there is no specification.
- Provide in one visual report-out the process stability assessment along with a prediction statement, when appropriate.

**Separation of Special-cause from Common-cause Variability**

For a given process, one would think that everyone (when creating a control chart) would make a similar conclusion relative process stability and its capability/performance, where the only difference is from sampling probability (i.e., samples will differ by “luck of the draw”). However, this is not necessarily true. With a traditional approach, process statements are not only a function of process characteristics and sampling-chance differences but can also be dependent upon sampling approach.
This sampling-difference issue can have dramatic implications:

- One person could describe a process as not being in control, which would lead to activities that are to immediately address process perturbations as abnormalities, while another person could describe the process as being stable. For this second interpretation, the described perturbations would be perceived as fluctuations typically expected within the process. Any long-lasting improvement effort for this second type of process response involves looking at the whole process to determine what might be done differently to enhance the process' performance.

- One person could describe a process as not predictable, while another person could describe output response as predictable.

To illustrate how different interpretations can occur, let’s consider the process time series data in Table 1. In this table, each subgrouping interval contained five samples, where only ten subgroups are shown to simply the illustration.

This type of data traditionally leads to an $\bar{x}$ and $R$ control chart, as shown in Figure 1 when determining if the process is in control or not. If the process were determined to be in-control, process capability indices are then might be used to provide

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Sample One</th>
<th>Sample Two</th>
<th>Sample Three</th>
<th>Sample Four</th>
<th>Sample Five</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>102.7</td>
<td>102.2</td>
<td>102.7</td>
<td>103.3</td>
<td>103.6</td>
</tr>
<tr>
<td>2</td>
<td>108.2</td>
<td>108.8</td>
<td>106.7</td>
<td>106.6</td>
<td>109.1</td>
</tr>
<tr>
<td>3</td>
<td>101.9</td>
<td>103.0</td>
<td>100.6</td>
<td>101.4</td>
<td>101.3</td>
</tr>
<tr>
<td>4</td>
<td>103.9</td>
<td>105.5</td>
<td>104.3</td>
<td>104.5</td>
<td>104.5</td>
</tr>
<tr>
<td>5</td>
<td>97.2</td>
<td>99.0</td>
<td>96.5</td>
<td>94.9</td>
<td>96.5</td>
</tr>
<tr>
<td>6</td>
<td>94.4</td>
<td>93.0</td>
<td>93.0</td>
<td>95.2</td>
<td>93.6</td>
</tr>
<tr>
<td>7</td>
<td>104.7</td>
<td>103.6</td>
<td>103.7</td>
<td>104.7</td>
<td>104.5</td>
</tr>
<tr>
<td>8</td>
<td>102.5</td>
<td>102.7</td>
<td>101.2</td>
<td>100.6</td>
<td>103.1</td>
</tr>
<tr>
<td>9</td>
<td>101.9</td>
<td>103.1</td>
<td>101.0</td>
<td>101.2</td>
<td>101.4</td>
</tr>
<tr>
<td>10</td>
<td>95.0</td>
<td>95.3</td>
<td>95.3</td>
<td>94.4</td>
<td>94.2</td>
</tr>
</tbody>
</table>

Table 1: Process Time Series Data

![Figure 1: $\bar{x}$ and $R$ Control Chart](image)
MINI-PAPER

how the process is performing relative to customer specifications; e.g., 95 and 105 lower and upper specification limits for this example.

Whenever a measurement on a control chart is beyond the upper control limit (UCL) or lower control limit (LCL), the process is said to be out of control. Out of control conditions are considered special cause events, where these out-of-control conditions can trigger a causal problem investigation. Since so many out-of-control conditions are apparent in Figure 1, many causal investigations could have been initiated. When a process is out of control, no process capability statement should be made about how the process is performing relative to its specification limits.

However, when creating a sampling plan for the above described situation, one could have select only one sample, instead of several samples, for each sampling time period. For this sampling approach illustration, consider that only the first measurement was observed for each of the 10 sampling intervals. An $XmR$ control chart, as shown in Figure 2, would be an appropriate method to assess process stability for this situation.

Relative to providing a process control statement, the $XmR$ control chart in Figure 2 provides a very different view when assessing process stability assessment than that provided by the $\bar{x}$ and $R$ chart shown in Figure 1. Since the plotted values in Figure 2 are within the control limits, one would conclude from this plot that only common cause variability exists and the process should be considered to be stable or predictable.

The dramatic difference between the limits of these two control charts is caused by the differing approaches to determining sampling standard deviation, which is a term involved in the calculation of statistical control limits. To illustrate this difference, let’s examine how these two control chart limit calculations are determined.

![Figure 2: XmR Chart of Sample One](image-url)
MINI-PAPER

For $\bar{x}$ charts (from an $\bar{x}$ and $R$ chart pair), the UCL (upper control limit) and LCL (lower control limit) are calculated from the relationships

\[
\begin{align*}
\text{UCL} &= \bar{x} + A_2 \bar{R} \\
\text{LCL} &= \bar{x} - A_2 \bar{R}
\end{align*}
\]

where $\bar{x}$ is the overall average of the subgroups, $A_2$ is a constant depending upon subgroup size and $\bar{R}$ is the average range within subgroups.

For an $X$ chart (from an $XmR$ chart pair), the UCL and LCL values are calculated from the relationships

\[
\begin{align*}
\text{UCL} &= \bar{x} + 2.66(\overline{MR}) \\
\text{LCL} &= \bar{x} - 2.66(\overline{MR})
\end{align*}
\]

where $\overline{MR}$ is the average moving range between adjacent readings.

The control limits for the $\bar{x}$ chart are derived from within-subgroup variability ($\bar{R}$), while sampling standard deviations for an $X$ chart is calculated from between-subgroup variability ($\overline{MR}$).

A process capability analysis of the data collectively would yield the output shown in Figure 3.

From Figure 3, one observes two standard deviations. The distribution that has a small standard deviation originates from the within subgroup variability of the process. Within subgroup variability is the variation used to determine the control limits for an $\bar{x}$ and $R$ control chart. This is the basic reason that the control limits in Figure 1 are narrow relative to the spread of the plotting for the data’s mean and range statistics in the control-chart plots.

**Figure 3: Process Capability Analysis of All Data**
The 30,000-Foot-Level View

Which of the two control charting techniques is most appropriate? It depends on how one categorizes the source of variability relative to common and special causes. To explain this point, a manufacturing situation will be used, although the same thought process applies to transactional environments.

Let’s consider that a new supplier’s shipments are received daily and there is a difference between the shipments, which unknowingly affects the output of the process. To answer our original common versus special cause question, we need to decide if the impact of day-to-day differences on our process should be considered a component of common or special cause variability. If these day-to-day differences should be considered a process noise variable, a control charting procedure that considers the day-to-day variability as a source of common cause variability would be most appropriate.

For this to occur in the creation of a control chart, one needs a sampling plan where the impact from this type of noise variable occurs between subgroupings. I refer to the sampling plan that accomplishes this need infrequent sampling/sub-grouping and this high-level process-output perspective the 30,000-foot-level view. When creating time-series control charts at the 30,000-foot-level, one needs to include between-subgroup variability within the control chart limit calculations, as was achieved in the earlier XmR charting procedure.

One needs to highlight that a 30,000-foot-level objective is different from the original intent of control charts. In my opinion, a primary intent for traditional control charts is that the tool was used to timely identify unusual events that violated the mathematical properties inherent within the chart calculations. When this situation occurred, the observation was termed an assignable cause and action was to be taken to address the condition, if appropriate, in a timely fashion. A 30,000-foot-level chart to provide a high-level view of a process that assess both process stability and, when only common-causes variability currently exists (e.g., since the staging of a control chart because a process change occurred), provides a prediction statement.

30,000-foot-level Charts with Subgrouping

One would like to have flexibility to use a methodology that has infrequent sampling or subgrouping with multiple samples for each reported time interval but still addresses between-subgroup variability when calculating individuals-chart limits. For this situation, one can track the subgroup mean and log standard deviation using two individuals control charts to assess whether the process is stable/predictable. The reason that a log transformation (i.e., Box-Cox transformation lambda equals zero) can be needed for the standard deviation plot is that standard deviation is bounded by zero and a log transformation of standard deviation tends to follow a normal distribution.

For in-control or predictable processes, the original data is then used in a 30,000-foot-level report-out to determine the overall process capability and performance; e.g., percent non-conformance.

One point that should be highlighted is that a moving range chart (in an XmR pair of charts) only tracks the differences in adjacent-responses over time, which can also be observed in the corresponding individuals chart. Because of this and to make the chart-output response simpler to read, the moving range chart is not included in a 30,000-foot-level report out’s individuals chart.

For the data in Table 1, Figure 4 shows a 30,000-foot-level report-out. Rather than report process capability indices such as $C_p$, $C_{pk}$, $P_p$, and $P_{pk}$, a 30,000-foot-level report-out provides a percentage non-conformance estimate whenever a specification is provided (See Figure 4). This value is determined from a probability plot and should be automatically reported at the bottom of the chart. A 30,000-foot-level process performance statement is easier to understand and is similar to a process expected overall performance report-out from a process capability analysis; i.e., 268,525.98 from Figure 3 is similar to a 26.852% nonconformance rate in Figure 4. In addition, if no specification exists, a 30,000-foot-level report-out would provide a best estimate for the process median and 80% frequency of occurrence.

One should also note from Figure 4 how a 30,000-foot-level report-out provides an estimate of how this stable process is not only performing now but in the future unless something were to change. If this prediction is undesirable, the 30,000-foot-level metric...
enhancement need pulls for a process improvement effort to enhance the process-output performance. One infers that a process was improved if the 30,000-foot-level individuals chart transitions to an enhanced level of performance.

Cost of Doing Nothing Differently
To determine if a process should be improved, one could calculate the cost of poor quality (COPQ) or the cost of doing nothing different (CODND). If the COPQ/CODND amounts are unsatisfactory, this metric improvement need should pull (used as a lean term) for creating a Lean Six Sigma project or other process improvement effort.

With this overall approach for metric reporting, an entire enterprise set of metrics could be collectively evaluated to determine where it is most beneficial to give focused improvement efforts. This differs from the use of Figure 1, which could lead one to create activities that are unproductive if one is attempting to understand the cause of isolated events, which are really common cause; i.e., firefighting common-cause (i.e., variability that occur between subgroups) issues as though they were special cause.

Applications
Metrics drive behavior; however, one needs to insure that not only the applicable metrics are tracked but that the report-out format for the measurements lead to the most appropriate behaviors. The 30,000-foot-level individuals chart(s) and corresponding capability/performance report-out (in one chart) provide a high-level view of process output, which is different from the application of traditional control charting. From an elevated organizational-performance-tracking viewpoint, common cause process variability can be separated from special cause conditions. From this separation viewpoint, the most appropriate action can occur for a given set of conditions.

The 30,000-foot-level process-output tracking system can be used for:
- Baselining the output response from a process, which is to be improved
- Identification that an enhancement has been made to a process from a targeted improvement effort; i.e., a process shift would be staged in the 30,000-foot-level individuals chart and the prediction statement be made using the data that was present after the staging point
MINI-PAPER

• Visual process control mechanism for maintaining the gain from an improved process
• An alternative to traditional organizational performance-metric reporting that addresses the issues described in a one-minute performance-scorecard reporting video
• An alternative to red-yellow-green stoplight reporting and a table of numbers report-outs as described in a 15-minute video. Traditional scorecards don’t typically report-out performance metrics from a process point of view, which can lead to unhealthy, if not destructive behaviors; a 30,000-foot-level report-out overcomes these issues.
• An alternative to various scorecard report-outs (eight example illustrations are shown in the reference)
• The structured reporting of performance metrics in an Integrated Enterprise Excellence (IEE) operational excellence system with an alignment to the processes that created them through an organizational value chain

The 30,000-foot-level reporting format has application to a variety of process-output tracking situations:

• Continuous data: Subgroups (illustrated in this paper)
• Continuous data: No subgroup
• Attribute non-conformance rate data
• Attribute count data
• Non-normal data
• Seasonal data

Creation procedures for 30,000-foot-level report-outs is described for a variety of situations in chapters 12 and 13 of Integrated Enterprise Excellence Volume III.

How to Easily Create a 30,000-foot-level Chart
A Minitab add-in for the easy creation of 30,000-foot-level report-outs can be downloaded at no charge.

Organizations benefit when they link 30,000-foot-level performance metrics with the processes that created them. This objective can be achieved through an Integrated Enterprise Excellence (IEE) value chain. Additional benefits can be achieved when software is used so that 30,000-foot-level metrics are automatically updated and become clickable to those in management and others who have an Internet connection and access authorization.

References
5. Fifteen minute video describing issues with traditional red-yellow-green scorecard reporting and 30,000-foot-level alternative, also application of this reporting format to eight example scorecard report-outs http://www.ieesystem.com/scorecard-examples/.

About the Author
Forrest Breyfogle III is an ASQ Fellow who has a MSME degree from the University of Texas in Austin. He is CEO of Smarter Solutions Inc. (www.SmarterSolutions.com). He authored or co-authored over a dozen books. ASQ Body of Knowledge references are made to several of his books. His five-book Integrated Enterprise Excellence series provides the details for implementing an Operational Excellence system that structurally integrates predictive scorecards with the processes that created them. This IEE system provides the guidance for the selection and execution of improvement project efforts that benefit the big picture.
Sudoku and Design of Experiments

Sudoku is a very popular puzzle and pastime. What may come as a surprise is that a completed Sudoku can be viewed as an interesting kind of designed experiment. Figure 1 is a table showing a completed Sudoku.

A Sudoku has a very interesting structure. It is a 9 by 9 grid of numbers from 1 to 9 where every row and column has all 9 numbers. A Latin Square design has these same elements.

I have separated 9 smaller 3 by 3 cells, each of which also contains all 9 numbers. So, a Sudoku is a very special case of a Latin Square design. Most Latin Square designs do not have this extra feature.

Why were Latin Square Designs Invented?
Ronald Fisher invented these designs in the 1930s for agricultural experiments. Imagine dividing a large square field into a 9 by 9 grid of smaller plots. Suppose that you wanted to test 9 different treatments on some crop to see which one resulted in the best yield. A Latin Square design allows the experimenter to independently estimate the row effect, the column effect, and the treatment effect. For this scenario the Latin Square Design is the most efficient design possible in the sense of minimizing the variance of all these estimates.

What is the Design Value of Further Grouping the 9 by 9 grid into Nine 3 by 3 Cells?
You can imagine that there might be some fertility gradient across the field. The row and column effects can adjust the treatment effect estimate for any fertility gradient effect. However, suppose the northwest corner of the large field (upper left corner) had some unique damage so that the yield in that whole 3 by 3 grid was zero. Losing that 3 by 3 square would still leave eight squares,

<table>
<thead>
<tr>
<th>Rows</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>c4</th>
<th>c5</th>
<th>c6</th>
<th>c7</th>
<th>c8</th>
<th>c9</th>
</tr>
</thead>
<tbody>
<tr>
<td>r1</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>r2</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>r3</td>
<td>2</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>r4</td>
<td>9</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>r5</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>r6</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>r7</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>r8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>r9</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 1: A Completed Sudoku
Design of Experiments

Design of Experiments

each employing all nine treatments. In a standard Latin Square design, the result of losing this square generally would be some lack of balance in the number of replications of each treatment.

More importantly, an assumption behind using a Latin Square design is that the row and column factors only have additive effects (i.e. they do not have a two-factor interaction). The Sudoku design can provide a kind of test of this assumption. You could do this by comparing each cell’s predicted value with the average of the predicted values of the row and column entries making up the cell.

**Can We Estimate the Nine Cell Effects Along with all the Other Effects?**
Sad, no. The cell effects are confounded with the row and column effects if we treat all these effects as fixed effects. However, if we treat the row and column effects as random effects, we can estimate the variance of these effects as well as the fixed cell and treatment effects. We call a design with two crossed random effects like this a strip-plot design.

**Is the Sudoku Design the Best Strip-plot Design?**
It depends . . .

If the goal is to minimize the variances of the treatment effects, while maintaining the meaning of a cell in terms of its location, then the answer is, yes.

However, if you wanted to estimate both the cell effects and the treatment effects with maximum efficiency, there is a better design. I constructed a D-optimal strip-plot design using a Covering Array design. A D-optimal design minimizes an omnibus scalar measure of the standard errors of the parameter estimates for a specific model. Many textbook designs are also D-optimal for an appropriate model. Examples are full factorial, fractional factorial, Latin Square, and balanced incomplete block designs. The D-optimal design here is actually a Graeco-Latin Square design.

Figure 2 shows the relative standard errors of the estimates of the Sudoku to the D-optimal design for estimating the cell effects (Cell 1 through Cell 8) and the treatment effects (designated N 1 through N 8).

![Figure 2: Relative Standard Errors](image-url)
Design of Experiments

You can see that the standard errors of the cell effect estimates for the Sudoku design are at least three times larger than for the D-optimal design.

Final Thoughts

The D-optimal design accomplishes this by removing the interpretation of a cell in terms of its location as a group of rows and columns. In agricultural experiments this would be undesirable.

In an industrial setting with two crossed random blocking factors, the D-optimal design could be very useful.

COLUMN
Statistical Process Control
Donald J. Wheeler

This article is based on material found in Advanced Topics in Statistical Process Control, Second Edition © 2004 SPC Press, and is used with permission.

Finding the Signals

A fundamental principle for understanding data is that while some data contain signals, all data contain noise, therefore, before we can detect the signals we have to filter out the noise. This act of filtration is the essence of all our data analysis techniques. Here we will look at the mechanism used by all modern techniques to filter out the noise. An understanding of this mechanism is important because most software programs offer options involving both correct and incorrect computations.

Given a collection of data it is common to begin with the computation of some summary statistics for location and dispersion. Averages and medians are used to characterize location, while either the range statistic or the standard deviation statistic is used to characterize dispersion. This much is taught in every introductory class. However, what is usually not taught is that the structures within our data will often create alternate ways of computing these measures of dispersion. Understanding the roles of these different methods of computation is essential for anyone who wishes to analyze data. In this paper we shall look at three different methods of computation. To illustrate how these methods work we shall use them with two different data sets.

Data Set One

Perhaps the most common type of structure for a data set is to have \( k \) subgroups of size \( n \) where the \( n \) values within each subgroup were collected under the same set of conditions. This structure is found in virtually all types of experimental data, and in most types of data coming from a production process. To illustrate the alternate ways of computing measures of dispersion we shall use a simple data set consisting of \( k = 3 \) subgroups of size \( n = 8 \) as shown in Table 1.

| Subgroup One | 4 5 5 4 8 4 3 7 |
| Subgroup Two | 2 4 3 7 5 4 2 5 |
| Subgroup Three | 3 6 6 4 5 4 6 6 |

Table 1: Data Set One
Data Set One with Method One

The first method of computing a measure of dispersion is the method taught in introductory classes in statistics and shown schematically in Figure 1. All of the data from the \( k \) subgroups of size \( n \) are collected into one group of size \( nk \) and a single dispersion statistic is found using all \( nk \) values. This dispersion statistic is then used to estimate a dispersion parameter such as the standard deviation for the distribution of \( X \) which we shall denote by the symbol \( SD(X) \).

As shown in Figure 2 the range of all 24 values is 6. The bias correction factor for ranges of 24 values is 3.895. Dividing 6 by 3.895 yields an unbiased estimate of the standard deviation parameter of the distribution of \( X \) of 1.540.

The global standard deviation statistic in Figure 2 is 1.551. The bias correction factor for this statistic based on 24 values is 0.9892. Dividing 1.551 by 0.9892 yields an unbiased estimate of the standard deviation parameter of the distribution of \( X \) of 1.568.

Since the original data are given to the nearest whole number, there is no practical difference between the two estimates of \( SD(X) \) shown in Figure 2. Whether we use the range or the standard deviation statistic will not substantially affect our subsequent computations.
Statistical Process Control

Data Set One with Method Two

While Method One ignores the subgroups, Method Two respects the subgroup structure within the data. Here we calculate a dispersion statistic for each subgroup as indicated in Figure 3. These separate dispersion statistics are then averaged, and the average dispersion statistic is used to form an unbiased estimate for the standard deviation parameter of the distribution of $X$.

Using Data Set One, we compute a dispersion statistic for each of the three subgroups. Since the subgroups are all the same size we can average the statistics prior to dividing by the common bias correction factor.

As shown in Figure 4, the subgroup ranges are respectively 5, 5, and 3. The average range is 4.333 and the bias correction factor for ranges of eight data is 2.847. Dividing 4.333 by 2.847 we estimate the standard deviation parameter for the distribution of $X$ to be 1.522.

![Figure 3: Method Two for Estimating Dispersion](image)

**Method Two**

<table>
<thead>
<tr>
<th>Subgroup One:</th>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1 = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_2 = 2.847$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$s_1 = 1.690$
$c_4 = 0.9650$

<table>
<thead>
<tr>
<th>Subgroup Two:</th>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_2 = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_2 = 2.847$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$s_2 = 1.690$
$c_4 = 0.9650$

| Subgroup Three: |   | X | X |   |
|-----------------|---|---|---|
| $R_3 = 3$       |   |   |   |
| $d_2 = 2.847$   |   |   |   |
| 1 2 3 4 5 6 7 8 9 |   |

$s_3 = 1.195$
$c_4 = 0.9650$

**Average Range** = 4.333
**Est. SD$(X)$** = 1.522

**Average Std. Dev.** = 1.525
**Est. SD$(X)$** = 1.580

![Figure 4: Method Two with Data Set One](image)
Statistical Process Control

The subgroup standard deviation statistics are respectively 1.690, 1.690, and 1.195. The average standard deviation statistic is 1.525 and the bias correction factor is 0.9650. Dividing 1.525 by 0.9650 we estimate the standard deviation parameter for the distribution of $X$ to be 1.580.

As before, there is no practical difference between the two estimates shown in Figure 4. Neither is there any practical difference between the estimates in Figure 2 and those in Figure 4. The four estimates of $SD(X)$ obtained using the two different measures of dispersion and the two different methods are all very similar for Data Set One.

Data Set One with Method Three

The third method will probably seem rather strange. It is certainly indirect. Instead of working with the individual values as the first two methods do, the third method works with the subgroup averages as shown in Figure 5. These subgroup averages are used to obtain a dispersion statistic, and this dispersion statistic is then used to estimate the standard deviation parameter of the distribution of $X$.

For Data Set One the subgroup averages are respectively 5.0, 4.0, and 5.0. The range of these three averages is 1.00. The bias correction factor for the range of three values is 1.693. Since this range was based on the subgroup averages we obtain an estimate of the standard deviation of the averages when we divide the range by 1.693. To convert this into an estimate for the standard deviation parameter for the distribution of $X$ we will have to multiply by the square root of 8 since each average represents eight original data. When we do this with the values above we obtain an estimate of $SD(X)$ of 1.671.

The standard deviation statistic for the three subgroup averages is 0.5774. Dividing by the bias correction factor of 0.8862 and multiplying by the square root of 8 we obtain an unbiased estimate of the standard deviation of the distribution of $X$ of 1.843.
Statistical Process Control

Once again, there is no practical difference between using the range and using the standard deviation statistic. While these two estimates are slightly larger than the first four, they are not appreciably different. It takes about 1.5 to 1.8 measurement units to equal one standard unit of dispersion.

Before we attempt to draw any lesson from this example we need to know that Data Set One has a very special property. When we place Data Set One on an average and range chart we end up with Figure 7. There we see no evidence of any differences between the three subgroups. Data Set One contains no signals. It is pure noise.

Therefore, at this point we can reasonably conclude that when the data are homogeneous and contain no signals the three methods will yield similar values for unbiased estimates of $SD(X)$, and this is true regardless of whether we use the range or the standard deviation statistic.

Data Set Two

But what happens in the presence of signals? After all, the objective is to filter out the noise so we can detect any signals that may be present. To see how signals affect our estimates of $SD(X)$ we shall modify Data Set One by inserting two signals. Specifically we shall shift subgroup two down by two units while we shift subgroup three up by four units. This will result in Data Set Two which is shown in Table 3. As may be seen in the average and range chart in Figure 8, these changes have introduced two distinct signals.

<table>
<thead>
<tr>
<th>Method One</th>
<th>Range Based</th>
<th>Std. Dev. Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est. $SD(X)$</td>
<td>1.540</td>
<td>1.563</td>
</tr>
<tr>
<td>Method Two</td>
<td>1.522</td>
<td>1.580</td>
</tr>
<tr>
<td>Method Three</td>
<td>1.671</td>
<td>1.845</td>
</tr>
</tbody>
</table>

Table 2: Summary of Three Methods for Data Set One

<table>
<thead>
<tr>
<th>Averages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Figure 7: Average and Range Chart for Data Set One

<table>
<thead>
<tr>
<th>Subgroup One</th>
<th>4 5 5 4 8 4 3 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subgroup Two</td>
<td>0 2 1 5 3 2 0 3</td>
</tr>
<tr>
<td>Subgroup Three</td>
<td>7 10 10 8 9 8 10 10</td>
</tr>
</tbody>
</table>

Table 3: Data Set Two
Method One with Data Set Two

Method One uses all 24 values in Data Set Two to compute a global measure of dispersion. As shown in Figure 9, the global range is 10.0 which results in an unbiased estimate of the standard deviation parameter of 2.567. The global standard deviation statistic is 3.279 which gives an unbiased estimate of the standard deviation parameter of 3.315.

These estimates of $SD(X)$ are roughly twice the size of those found in Figure 2. Thus, the signals introduced by shifting the subgroup averages have inflated both of the Method One estimates by an appreciable amount.

Method Two with Data Set Two

Using Method Two, we compute a dispersion statistic for each of the three subgroups. Since the subgroups are all the same size we can average the statistics prior to dividing by the common bias correction factor. As shown in Figure 10, the average range is 4.333 and the bias correction factor for ranges of eight data is 2.847. Dividing 4.333 by 2.847 we estimate the standard deviation for the distribution of $X$ to be 1.522.

The average standard deviation statistic is 1.525 and the bias correction factor is 0.9650. Dividing 1.525 by 0.9650 we estimate the standard deviation for the distribution of $X$ to be 1.580.

The Method Two estimates of $SD(X)$ for Data Set Two are exactly the same as those obtained for Data Set One in Figure 4. Thus, the Method Two estimates are not affected by the signals introduced by shifting the subgroup averages.

Method Three with Data Set Two

For Data Set Two the subgroup averages are respectively 5.0, 2.0, and 9.0. The range of these three averages is 7.00. The bias correction factor for the range of three values is 1.693. Since this range was based on subgroup averages we obtain an estimate of the standard deviation of the averages when we divide the range by 1.693. To convert this into an estimate for the standard deviation parameter for the distribution of $X$ we will have to multiply by the square root of 8 since each average represents eight original data. When we do this with the values above we obtain an estimate of $SD(X)$ of 11.693.
Statistical Process Control

Method Two

Subgroup One:
\[ R_1 = 5 \]
\[ d_2 = 2.847 \]
\[
\begin{array}{cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
X & X & X & s_1 = 1.690 & & & & & \\
\end{array}
\]
\[ c_4 = 0.9650 \]

Subgroup Two:
\[ R_2 = 5 \]
\[ d_2 = 2.847 \]
\[
\begin{array}{cccccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
X & X & X & & s = 1.690 & & & & & \\
\end{array}
\]
\[ c_4 = 0.9650 \]

Subgroup Three:
\[ R_3 = 3 \]
\[ d_2 = 2.847 \]
\[
\begin{array}{ccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
X & X & & X & & s = 1.195 & & & c_4 = 0.9650 \\
\end{array}
\]

Average Range = 4.333
Est. SD(X) = 1.522

Average Std. Dev. = 1.525
Est. SD(X) = 1.580

Figure 10: Method Two with Data Set Two

The standard deviation statistic for the three subgroup averages is 3.512. Dividing by the bias correction factor of 0.8862 and multiplying by the square root of 8 we obtain an unbiased estimate of the standard deviation of the distribution of \( X \) of 11.209.

These Method Three estimates of \( SD(X) \) are seven times larger than values found in Figure 6. Thus, the signals introduced by shifting the subgroup averages have severely inflated both of the Method Three estimates.

When we summarize the results of the three methods with Data Set Two we get Table 4. We have obtained six unbiased estimates of \( SD(X) \) using three different methods and two different statistics, yet these six values differ by almost an order of magnitude!

Method Three

\[ R_x = 7.000 \]
\[ d_2 = 1.693 \]
\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 \\
X & X & X & s_x = 3.512 & & & \\
\end{array}
\]
\[ c_4 = 0.8862 \]

Est. SD(X) = 11.693

Est. SD(X) = 11.209

Figure 11: Method Three with Data Set Two

<table>
<thead>
<tr>
<th>Method</th>
<th>Est. SD(X)</th>
<th>Est. SD(X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range Based</td>
<td>Std. Dev. Based</td>
<td></td>
</tr>
<tr>
<td>Method One</td>
<td>2.567</td>
<td>3.315</td>
</tr>
<tr>
<td>Method Two</td>
<td>1.522</td>
<td>1.580</td>
</tr>
<tr>
<td>Method Three</td>
<td>11.693</td>
<td>11.209</td>
</tr>
</tbody>
</table>

Table 4: Summary of Three Methods for Data Set Two
Statistical Process Control

The differences left to right in Table 4 show the effects of using the different dispersion statistics. The differences top to bottom reveal the differences due to using the different methods. Clearly, the differences left to right pale in comparison with those top to bottom. The key to filtering out the noise so we can detect the signals does not depend upon whether we use the standard deviation statistic or the range, but rather upon which method we employ to compute that dispersion statistic.

Method One estimates of dispersion are commonly known as the Total Variation or the Overall Variation. Method One is used for description. It implicitly assumes that the data are globally homogeneous. When the data are not globally homogeneous this method will be inflated by the signals contained within the data and the value obtained will no longer estimate \( SD(X) \).

Method Two estimates of dispersion are commonly known as the Within-Subgroup Variation. Method Two is used for analysis. Whenever we seek to filter out the noise in order to detect signals we use Method Two to establish the filter. Method Two implicitly assumes that the data are homogeneous within the subgroups, but it places no requirement of homogeneity upon the different subgroups. Thus, even when the subgroups differ, Method Two will provide a useful estimate of \( SD(X) \).
**Statistical Process Control**

**Method Three** estimates of dispersion are commonly known as the Between-Subgroup Variation. Method Three is used for comparison purposes. It assumes that the subgroup averages are globally homogeneous. When Method Three is computed it is generally compared with Method Two; the idea being that any signals present in the data will affect Method Three more than they affect Method Two. When the subgroups have different averages, Method Three will *not* provide an estimate of $SD(X)$.

**Separating the Signals from the Noise**

The essence of every statistical analysis is the separation of the signals from the noise. We want to find the signals so that we can use this knowledge constructively. We want to ignore the noise where there is nothing to be learned. To this end we begin by filtering out the noise. And for the past 100 years the standard technique for filtering out the noise has been Method Two! To illustrate this point Figure 15 shows the average chart for Data Set Two with limits computed using each of the three methods. Only Method Two correctly identifies the two signals we deliberately buried in Data Set Two.

So when it comes to filtering out the noise you have a choice between Method Two, Method Two, or Method Two. Any method for filtering out the noise is right as long as it is Method Two!

Method One is inappropriate for filtering out the noise because it gets inflated by the signals. Method One has always been wrong for analysis, and it will always be wrong. Trying to use Method One for analysis is so wrong that it has a name. It is known as Quetelet’s Fallacy and it is the reason there was so little progress in statistical analysis in the Nineteenth Century.

![Figure 14: Between-Subgroup Variation (Used to Compare with Method Two)](image1)

![Figure 15: Average Charts for Data Set Two](image2)
Method Three is completely inappropriate for filtering out the noise because it will be severely inflated in the presence of signals. If you use Method Three to filter out the noise you will have to wait a very long time before you detect a signal. So while there are analysis techniques that make use of the Method Three (Between Subgroup) estimate of dispersion, they do so only in order to compare it with a Method Two (Within Subgroup) estimate of dispersion.

Thus, the foundation of all modern data analysis techniques is the use of Method Two to filter out the noise. This is the foundation for the Analysis of Variance. This is the foundation for the Analysis of Means. And this is the foundation for Shewhart’s process behavior charts. Ignore this foundation and you will undermine your whole analysis.

Many analysis techniques from the Nineteenth Century, such as Franklin Pierce’s test for outliers, are built on the use of Method One to filter out the noise. As may be seen in Figure 15, this approach will let you occasionally detect a signal, but it will cause you to miss other signals.

In fact, many techniques developed in the Twentieth Century also suffer from Quetelet’s Fallacy. Among these are Grubb’s test for outliers, the Levey-Jennings control chart, and the Tukey control chart. Moreover, virtually every piece of statistical software available today allows the user to choose Method One for creating control charts and performing various other statistical tests. Nevertheless, this error on the part of naive programmers does not make it right or even acceptable to use Method One for analysis.

So while there are proper uses of Method One and Method Three, they are never appropriate for filtering out the noise. The only correct method for filtering out the noise is Method Two. Understanding this point is the beginning of competence for every data analyst.

You now know the difference between modern data analysis techniques and naive analysis techniques. Naive techniques use Method One or Method Three to filter out the noise. Today all sorts of new naive techniques are being created by those who know no better. Let the reader beware.

To help with this problem of identifying naive techniques Table 5 contains a listing of 27 of the more commonly encountered within-subgroup estimators of both the standard deviation parameter and the variance parameter. There we see the hallmark of the within-subgroup approach: Each estimator is based on either the average or the median of a collection of \( k \) within-subgroup

<table>
<thead>
<tr>
<th>Name of Estimator</th>
<th>Estimators for SD(X)</th>
<th>Estimators for V(X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Range</td>
<td>Biased</td>
<td>Unbiased</td>
</tr>
<tr>
<td>Median Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Moving Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Moving Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Root Mean Square Dev.</td>
<td>( s_n^- )</td>
<td></td>
</tr>
<tr>
<td>Median Root Mean Square Dev.</td>
<td>( \bar{s}^- )</td>
<td></td>
</tr>
<tr>
<td>Average Standard Deviation</td>
<td>( \bar{s}^- )</td>
<td></td>
</tr>
<tr>
<td>Median Standard Deviation</td>
<td>( s^- )</td>
<td></td>
</tr>
<tr>
<td>Pooled Variance</td>
<td>( s^2^- )</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Some Within-Subgroup Estimators
Identifying Bottleneck Operations in a Manufacturing Systems

Lin, Qing et al (2009) define a bottleneck in a production process as a machine which reduces the overall systems performance in the strongest manner. As a result, improvements in bottleneck machines result in significantly higher overall throughput than performance improvement on non-bottleneck machines. If machine $k$ is the bottleneck, Chang et al 2007 identify it by,

$$\frac{\Delta TP_{sys,k}}{\Delta TP_k} = \max \left( \frac{\Delta TP_{sys,1}}{\Delta TP_1}, \frac{\Delta TP_{sys,2}}{\Delta TP_2}, \ldots, \frac{\Delta TP_{sys,n}}{\Delta TP_n} \right),$$

where $\Delta TP_{sys,i}$ is the change in system throughput due to reducing the downtime of machine $i$ and $\Delta TP_i$ is the isolated throughput improvement of machine $i$.

Chang et al (2007) states that manufacturing systems with unreliable machines and finite buffers present a need for a control policy that can adjust to short term needs. Variations in processing times and machine failures cause system variability. A question we will ask is whether an analysis that reacts to short-term bottleneck changes over a limited time interval such as a day will be more effective in increasing system throughput.

This paper describes the use of a data driven bottleneck identification method called the turning-point method (Li, Qing et al. 2009). Analytical methods use simplified models that are not generally useful. Analyses using discrete-event simulation models are capable of analyzing complex systems and identifying bottlenecks. However, the lengthy simulation development time prevents...
Statistics for Quality Improvement

its use in many potential applications. Chiang et al (2000) describe a method for using online data to identify bottlenecks by comparing two adjacent machines. If the blockage time of the upstream machine is higher than the starvation time of the downstream machine, then the bottleneck must be downstream. Otherwise, the bottleneck must be upstream. The approach described by Chiang et al is called an arrow-based method. The turning-point method is similar to the arrow-based method but the analysis is different.

**Turning Point Method**

A production line uses buffers between machines to store parts from upstream machines and provide parts to downstream machines. Buffers reduce the effects of machine failures. If a buffer is never empty or full during a specified time period, it is effective because the downstream machine is never starved and the upstream machine is never blocked. An effective buffer decouples a production line so it might have multiple turning points.

The “turning point” definition uses machine blockage and starvation information. A bottleneck machine will tend to cause upstream machines to be blocked and downstream machines to be starved. Also, the bottleneck machine will frequently have less total blockage plus starvation time than its adjacent machines. The following figure illustrates a seven machine line with M4 being the turning point and the bottleneck machine.

For an observed time period, let $TB_i$ be the machine i blockage time and $TS_i$ be the corresponding starvation time for machine i. Machine j is the turning point when:

$TB_i - TS_i > 0: 0 < i < j, j ≠ 1, j ≠ n$

$TB_i - TS_i < 0: j < i < n+1, j ≠ 1, j ≠ n$

$TB_i + TS_i < TB_{i+1} + TS_{i+1}, j ≠ 1, j ≠ n$

$TB_i + TS_i < TB_{i+1} + TS_{i-1}, j ≠ 1, j ≠ n$

If $j = 1$: $TB_1 - TS_1 > 0$ and $TB_2 - TS_2 < 0$ and $TB_1 + TS_1 < TB_2 + TS_2$

If $j = n$: $TB_{n-1} - TS_{n-1} > 0$ and $TB_n - TS_n < 0$ and $TB_n + TS_n < TB_{n-1} + TS_{n-1}$

---

**Figure 1: Blockage and Starvation Times**

- % Blocked
- % Starved

---

Blockage and Starvation Times
When multiple turning points exist they are ranked by a heuristic index. That is, \( I_i = (TB_{i+1} - TS_{i+1})/(TB_i - TS_i) \). The denominator of the index \( I_i \) is the sum if machine \( i \)'s blockage and starvation time. A lower value of the denominator implies a higher value of machine \( i \)'s operating time plus downtime. That indicates a more significant bottleneck. Li, Qing et al. 2009 report that an analysis of over 1000 cases shows that the index is effective and provides good guidelines for the primary bottleneck in over 85% of the cases.

Simulation Verification
Li, Qing et al. 2009 report results supporting the hypothesis that system bottlenecks are turning points. One set of results consisted of simulations of serial production lines with finite buffers. Example lines included two-machine-one-buffer (2M1B), three-machine-two-buffers (3M2B), up to six-machine-five-buffers (6M5B), and several-machine-no-buffers. In more than 95% of over 1000 cases a turning point is the correct bottleneck.

Case Study Based on Industrial Data
The case study involves an automobile production line with 17 stations and 4 buffers described by Li, Qing et al. 2009. They operated a simulation to estimate a long-term static bottleneck. Its estimate was based on 200 replications, and it identified station 15 as the long term bottleneck. In this case the long term is one week of operation. In addition to testing the turning-point in identifying the long term bottleneck, they used the turning point method to update the bottleneck each day.

This system had one maintenance engineer and three dispatching policies for the maintenance were compared by simulation. The baseline policy was for the maintenance engineer to work on the machine that failed first without considering bottleneck location. With bottleneck identification, the maintenance engineer worked on the bottleneck machine first. If the bottleneck machine was the long-term bottleneck, the maintenance engineer always preferred station 15. With daily bottleneck identification, that could change.

<table>
<thead>
<tr>
<th>Dispatching Policy</th>
<th>Average Throughput</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>6320</td>
<td>6273, 6367</td>
</tr>
<tr>
<td>Static long-term bottleneck identification</td>
<td>6521</td>
<td>6469, 6573</td>
</tr>
<tr>
<td>Daily turning-point identification</td>
<td>6917</td>
<td>6855, 6979</td>
</tr>
</tbody>
</table>

This case study clearly indicates the potential improvement of using a data-driven bottleneck identification method such as the turning-point method.

References
Control Charts

When processes require close surveillance of specific attributes as well as critical dimensions or rates during production, this status assessment is best accomplished with the application of control charts. The primary purpose of control charts is to graphically show trends in the frequency of occurrence of the attributes or in the magnitude of the dimensions or rates with respect to maximum and minimum limit lines (referred to as upper and lower control limits).

A process that has reached a state of statistical control has a predictable spread of quality and a predictable output. Control charts provide a means of anticipating and correcting whatever special causes, as opposed to common causes, may be responsible for the generation of defects or unacceptable variation.

Control charts are used to direct the efforts of engineers (design, process, quality, manufacturing, industrial, etc.), operators, supervisors, and other Process Improvement Team (PIT) members toward special causes when, and only when, a control chart detects the presence and influence of a special cause. The ultimate power of control charts lies in their ability to separate out special (assignable) causes of defects and variation.

The objective of control chart analysis is to obtain evidence that the inherent process variability (known as the range or R) and the process average (known as the average of the averages or X-double bar) are no longer operating at stable (predictable) levels. When this occurs, one or both are out-of-control (not stable) and appropriate and timely corrective action should be introduced.

The purpose of establishing control charts is to distinguish between the inherent random variability of a process (known as common causes) and the variability attributed to an assignable (or special) cause. Common causes are those that cannot be readily changed without significant process reengineering.

Control charts are applied to study specific ongoing processes to keep them operating with minimum defects and variation. This operating state is known as Six Sigma and can be verified using performance measures such as Cp, the process capability index, and Cpk, the mean-sensitive process capability index. This is in contrast to downstream inspection and test which aim to detect defects and variation after they have been generated by the process. In other words, control charts are focused on prevention rather than detection and rejection.

It has been confirmed in practice, that economy and efficiency are better served by the prevention concept than by the detection concept. It costs as much to make a bad part as a good part, and the cost of applying control charts is repaid many times over as a direct result of improving production quality.

Types of Control Charts

Just as there are two types of data, continuous and discrete, there are also two types of control charts, variables charts for use with continuous data and attributes charts for use with discrete data.

Variables control charts should be used whenever measurements from a product or process are available. The following are examples of continuous data: diameters, torque values, durations, temperatures and electrical performance values such as voltages and amperages.

Whenever possible, variables control charts are preferred to attribute control charts because they provide greater insights regarding product or process status.
Stats 101

When continuous data are not available, but discrete data are, then it is appropriate to use attributes control charts. There are only two levels for attribute data: i.e., conform/nonconform, pass/fail, go/no-go, present/absent; however, they can be counted, recorded and analyzed. Examples of discrete data include the following: presence or location/position of a label, installation of all required fasteners, presence of excess or insufficient solder, and measurements when recorded as accept/reject.

The first of the following figures summarizes the various types of variables and attribute control charts. The second figure indicates when they should be applied.

Figure 1: Control Chart Flow Chart

<table>
<thead>
<tr>
<th>Characteristic Value</th>
<th>Name</th>
</tr>
</thead>
</table>
| Variable (Continuous) | \( \bar{x} \)-R Chart (average value and range)  
|                      | \( x \)-MR Chart (measured value and moving range) |
| Attribute (Discrete)  | np Chart (number of defective units)  
|                      | p Chart (fraction defective)  
|                      | c Chart (number of defects)  
|                      | u Chart (number of defects per unit) |

Table 1: Types of Control Charts
**COLUMN**

**Hypothesis Testing**

*by Jim Frost*

Originally published at [http://statisticsbyjim.com](http://statisticsbyjim.com)

---

**Introduction to Hypothesis Testing Terminology**

In this column, I explain why you need to use hypothesis testing and cover some basic hypothesis testing terminology. Hypothesis testing is a crucial procedure to perform when you want to make inferences about a population using a sample. These inferences include estimating population properties such as the mean, differences between means, proportions, and the relationships between variables.

**Why You Should Use Hypothesis Tests**

There are tremendous benefits that you gain by working with a random sample drawn from a population. In most cases, it is simply not possible to observe the entire population to understand its properties. The only alternative is to assess a sample.

While samples are much more practical and less expensive to work with, there are tradeoffs. When you estimate the properties of a population from a sample, the sample statistics are unlikely to equal the actual population value exactly. For instance, your sample mean is unlikely to equal the population mean. The difference between the sample statistic and the population value is the sample error.

An effect observed in any sample might due to sample error rather than representing a true effect at the population level. If the difference caused by sample error, the next time someone performs the same experiment the results might be different. Hypothesis tests incorporate estimates of the sampling error to help you make the correct decision.

For example, if you are studying the proportion of defects produced by two manufacturing methods, any difference you observe between the two sample proportions might be sample error rather than a true difference. If the difference does not exist at the population level, you won’t obtain the benefits that you expect based on the sample statistics. That can be a costly mistake!

Let’s cover some basic terms that you need to know.

**Hypothesis Tests**

A hypothesis test uses sample data to assess two mutually exclusive theories about the properties of a population. Statisticians call these theories the null hypothesis and the alternative hypothesis. The hypothesis test assesses your sample statistic and factors in an estimate of the sample error to determine which hypothesis the data support.

When you can reject the null hypothesis, the results are statistically significant, and your data support the hypothesis that an effect exists at the population level.

**Effect**

The effect is the difference between the population value and the null hypothesis value. The effect is also known as population effect or the difference. For example, the mean difference between the health outcome for a treatment group and a control group is the effect.

Typically, you do not know the size of the actual effect. However, you can use a hypothesis test to help you determine whether an effect exists and to estimate its size.
Hypothesis Testing

Null Hypothesis
The null hypothesis is one of two mutually exclusive theories about the properties of the population. Typically, the null hypothesis states that there is no effect (i.e., the effect size equals zero). The null hypothesis is often signified by $H_0$.

In all hypothesis tests, the researchers are testing an effect of some sort. The effect can be the effectiveness of a new vaccination, the durability of a new product, the proportion of defect in a manufacturing process, and so on. There is some benefit or difference that the researchers hope to identify.

However, it’s possible that there is no effect or no difference between the experimental groups. In statistics, we call this lack of an effect the null hypothesis. Therefore, if you can reject the null hypothesis, you can favor the alternative hypothesis, which states that the effect exists (doesn’t equal zero) at the population level.

You can think of the null hypothesis as the default theory that requires strong enough evidence before you can reject it.

For example, in a 2-sample t-test, the null hypothesis often states that the difference between the two means equals zero.

Alternative Hypothesis
The alternative hypothesis is the other theory about the properties of the population. Typically, the alternative hypothesis states that a population parameter does not equal the null hypothesis value. In other words, there is a non-zero effect. If your sample contains sufficient evidence, you can reject the null hypothesis and favor the alternative hypothesis. The alternative hypothesis is often identified with $H_1$ or $H_A$.

For example, in a 2-sample t-test, the alternative hypothesis often states that the difference between the two means does not equal zero.

You can specify either a one- or two-tailed alternative hypothesis:

If you perform a two-tailed hypothesis test, the alternative hypothesis states that the population parameter does not equal the null hypothesis value. For example, when the alternative hypothesis is $H_A: \mu \neq 0$, the test can detect differences both greater than and less than the null value.

A one-tailed alternative hypothesis has more power to detect an effect but it can test for a difference in only one direction. For example, $H_A: \mu > 0$ can only test for differences that are greater than zero.

P-value
P-values are the probability that you would obtain the effect observed in your sample, or larger, if the null hypothesis is correct. In simpler terms, p-values tell you how strongly your sample data contradict the null hypothesis. Lower p-values represent stronger evidence against the null.

Significance Level
The significance level, also known as alpha or $\alpha$, is an evidentiary standard that a researcher sets before the study. It specifies how strongly the sample evidence must contradict the null hypothesis before you can reject the null hypothesis for the entire population. This standard is defined by the probability of rejecting a null hypothesis that is true. In other words, it is the probability that you say there is an effect when there is no effect. Lower significance levels indicate that you require stronger evidence before you will reject the null hypothesis.

For instance, a significance level of 0.05 signifies a 5% risk of deciding that an effect exists when it does not exist.
Hypothesis Testing

Use p-values and significance levels together to help you determine which hypothesis the data support. If the p-value is less than your significance level, you can reject the null hypothesis and conclude that the effect is statistically significant. In other words, the evidence in your sample is strong enough to be able to reject the null hypothesis at the population level.

Types of Error

Hypothesis tests are not 100% accurate because they use a random sample to draw conclusions about entire populations. When you perform a hypothesis test, there are two types of errors related to drawing an incorrect conclusion.

- False positives: You reject a null hypothesis that is true. Statisticians call this a Type I error. The Type I error rate equals your significance level.
- False negatives: You fail to reject a null hypothesis that is false. Statisticians call this a Type II error. Generally, you do not know the Type II error rate. However, it is a great risk when you have a small sample size, noisy data, or a small effect size.

Which Type of Hypothesis Test is Right for You?

There are many different types of hypothesis tests you can use. The correct choice depends on your research goals and the data you collect. Do you need to understand the mean or the differences between means? Or, perhaps you need to assess proportions. You can even use hypothesis tests to determine whether the relationships between variables are statistically significant.

To choose the proper hypothesis test, you’ll need to assess your study objectives and collect the correct type of data. This background research is necessary before you begin a study.

Hypothesis tests are crucial when you want to use sample data to make conclusions about a population because these tests account for sample error. Using significance levels and p-values to determine when to reject the null hypothesis improves the probability that you will draw the correct conclusion.

Nominations Sought for 2017 William G. Hunter Award

The ASQ Statistics Division is pleased to announce that nominations are now open for its William G. Hunter Award for 2017.

Procedure: The nominator must have the permission of the person being nominated and letters from at least two other people supporting the nomination. Claims of accomplishments must be supported with objective evidence. Examples include publication lists and letters from peers. Nominators are encouraged to read the accompanying article “William G. Hunter: An Innovator and Catalyst for Quality Improvement” written by George Box in 1993 (http://williamghunter.net/george-box-articles/william-hunter-an-innovator-and-catalyst-for-quality-improvement) to get a better idea of the characteristics this award seeks to recognize.

Nominations: Nominations for the current year will be accepted until June 30. Those received following June 30 will be held until the following year. A committee of past leaders of the Statistics Division selects the winner. The award is presented at the Fall Technical Conference in October.

The award criteria and nomination form can be downloaded from the “Awards” page of the ASQ Statistics Division website (http://asq.org/statistics/about/awards-statistics.html) or may be obtained by contacting Necip Doganaksoy (ndoganaksoy@siena.edu).
Membership in a Technical Advisory Group

Congratulations! Your application for membership on the US TAG (Technical Advisory Group) of ISO TC69 on Application of Statistical Methods has been approved following a vote of its current members. Better go tell your supervisor since your Agency/Company agreed to support your participation. Now that you are “in” what exactly are you expected to do???

The first thing that is likely to occur is a friendly email such as the following:

“Hello TAG 69 members,

Please click here to review and vote on this ballot by April 2.

The question is; “Do you approve the technical content of the draft international standard (DIS) 11843-7 Capability of detection—Part 7: Methodology based on stochastic properties of instrumental noise?”

This email also contains a recommendation from the US Expert. So, all that remains is to download and to review the DIS, note any comments, arrive at a decision and vote electronically. That isn’t so bad!

The previous example referred to a specific document but the vote could have been, for example, on a New Work Item, membership for a potential new TAG member, or creation of a new working group. The actual vote choices on progressing a document is from one of the following possibilities:

- Approve (recommended)
- Approve with comments
- Disapprove (provide justification)
- Abstain

For a “Disapprove” vote, justification must be given (preferably with specific difficulties beyond “I don’t like this document.”). Approval with comments to improve the document are also welcomed. A vote of abstain is a participation vote which lets Standards know that you have received the request. One is not expected to provide an in-depth opinion on every vote!

Perhaps the most useful comments include identification of errors (calculations or formulae), recognition of incorrect usage of a statistical method or interpretation, or noting that a so-called “standard” approach as given in the document under review is not a recognized standard but one of several possibilities. For the last possibility the document is likely more appropriate as a technical report rather than a statistical standard.

After some experience in reviewing documents, down the road you may wish to consider serving as a project leader for herding a document through the phases to publication as an ISO standard. In addition to commenting on your own document, you would be assimilating the comments from the voting countries for disposition at a future meeting (either during an interim meeting...
Standards InSide-Out

or at the annual international meeting). Speaking of the annual international meeting, all this pro bono work can pay off if you have the resources to attend this meeting which rotates among the member bodies. Recent venues have included London, Dalian (China), Vienna, Tokyo, Kuala Lumpur and Paris. At these meetings you get the chance to meet in person the folks behind the member body comments (usually).

The process of reviewing the collected comments is interesting. Each comment is identified by member body (i.e., country), line number, clause/sub-clause, paragraph/figure/table, Comments, Proposed change, Observations of the secretariat. Since multiple countries may comment on the same item in the document, the Comment Template facilitates collecting comments on like items. An absolutely essential aspect of the comment row is the “Proposed change”. If no change is proposed then the comment can be immediately rejected! Although this might seem a bit cruel, it is necessary to avoid endless iterations of the working group posing proposed changes that do not capture the intent of the commenter. I have found this process incredibly helpful in expediting document revisions by committee. It is easy to criticize a sentence, but providing a viable alternative may be challenging.

Eventually the big day arrives where your document receives the requisite vote totals from the ISO TC69 P-member countries and the document is then off to publication by ISO.

Looking further down the road, a role within the US TAG as the lead delegate for a Subcommittee is possible (there is turnover in the US delegation). The subcommittees are:

- SC1 Terminology and symbols
- SC4 Applications of statistical methods in process management
- SC5 Acceptance sampling
- SC6 Measurement methods and results
- SC7 Applications of statistical and related techniques for the implementation of Six Sigma
- SC8 Application of statistical and related methodology for new technology

Within each official Subcommittee there are multiple working groups some with multiple documents under their aegis. Hence, there are ample opportunities to take the lead on a specific document either overall or as a US delegate. We also have a Working Group 3 under TC69 dealing with legacy standards and more recently an Ad hoc Group on Big Data.

The above discussion was intended to highlight the basic tasks involved in being a member of the US TAG, written under the presumption that you have joined it and want further details on the specific work program. I am grateful to Matthew Barsalou for suggesting the topic of this column.

The next column will report on the accomplishments at our next ISO TC69 International meeting June 19–23, 2017 in Cape Town, South Africa!
Years ago I was involved in the application of an Xbar Control Chart in a Manufacturing company. The chart led to a significant positive impact on the organization, even though our Control Chart limit calculation was “wrong”. In this article I’ll present my experience and the lessons I learned.

Background

I was working for a company that manufactured high purity silicon rods for the photovoltaic and silicon wafer industries. The rods were made using the Siemens process, which consists of resistance heating small caliber high purity silicon rods (Slim Rods) enclosed in a bell jar reactor. Silane gas (SiH₄) is introduced to the reactor where the silane molecule dissociates into silicon depositing on the Slim Rods and hydrogen (H₂) gas venting from the reactor.

After many days of continuous silane flow and silicon deposition the Slim Rods had grown to the desired diameter. The process was stopped, bell jar removed and the rods harvested for further processing.

After harvesting two critical quality checks on all rods in the reactor run were performed: morphology and powder fall. The morphology measurement addressed the surface texture of the rod. A rod surface similar to the surface of a head of cauliflower realized a lower selling price than a surface that looked like sandpaper.

Powder built up on the walls of the reactor from undesirable side reactions. Powder falling on the rods also lowered the selling price.
Feature

Silane flow rate into the reactor and electrical power into the rods were the process inputs. Silane flow was continuously monitored with a mass flow controller (MFC), while for process control purposes power input was indirectly monitored by temperature: a sight glass allowed measurement of one reactor rod temperature by an infrared thermometer.

All the mounted thermometers were adjusted against a hand held calibrated master. The master measured temperature through the sight glass, and the IR thermometer electronically adjusted to match the master temperature.

Initiating the Control Chart

A colleague and I were hired into the department and within months we questioned this control method. We specifically asked “Why are we so concerned about the absolute value of a rod temperature? Isn’t our ultimate concern the quality and volume of product coming out of each reactor?” The morphology and powder fall checks were performed on the rods after harvesting, but they were only used to indicate product quality. They were not used as feedback for the next reactor run.

We had an old presentation about Control Charts lying around the office. The presentation was specific to Zone Control Charts. I read over the presentation and thought this could be used on our reactors. The operators would have a simple chart to plot reactor run output and adjust inputs as needed for the next run. I was reading about control charts but using them to adjust the process toward a target value. Reactor run morphology of 3 and no powder fall was the desirable target. A morphology Xbar control chart could be used to decide when to make process adjustments. Temperature or silane flowrate adjustments would be made only if the Zone rules indicated it or if powder fall occurred. The centerline would be established at morphology of 3.

Having no previous knowledge of control charts and no one to guide me, I calculated the standard deviation for morphology on each reactor, lumping together the within batch and between batch data for the calculation. The necessary Zone Chart bands for the reactor average were established based on this total standard deviation.

We established a reaction plan based on the reactor output. If the zone chart indicated a morphology shift, temperature was incrementally decreased or increased, depending on the direction of the shift (the morphology scale was 1–6). Powder fall led to an incremental decrease in silane flow rate.

After much internal debate and resistance to change we were able to implement the charts. No one questioned the standard deviation calculation. No one questioned the lack of a Range or Standard deviation chart. No one commented on how we were mixing SPC and EPC by forcing the reactor output to move toward a target. No one recognized that one of the inputs (Slim Rods) was being completely ignored. There was so much focus on the IR thermometers and MFC’s that no one considered the reactor output variation that could be caused by Slim Rod variation.
Feature

What Happened

In spite of the standard deviation utilized for the Xbar limits and the lack of a Range or Standard Deviation chart, within a few runs on each reactor we were able to reduce reactor run variability around the target morphology. Powder fall was reduced significantly. With the tighter control of morphology our yield to “sandpaper” surface product increased by 10%.

Engineering believed that with reduced variation of morphology we could run the reactors faster by increasing silane flow rate and rod temperature. They were right. Not only were we able to significantly increase our yield to sandpaper product but were able to reduce the time to attain the target diameter by 15%. Our yield and throughput improved significantly. The accountants determined that these improvements were worth $1MM in profits. This change, in addition to another change implemented by my co-worker, pushed the company from breaking even to being profitable.

I was transferred to the Quality Department where I read a little more about control charts and was convinced I calculated the limits incorrectly. I recalculated them, using within reactor batch variation from previous runs to determine $\bar{s}$, which I just multiplied it by 3 to get the limits. I had new upper and lower limits for the average chart (still no range or std deviation chart). I started plotting the data on my own using the new limits and contacted Production when Special Causes occurred.

With the new limits on all the reactors Special Causes occurred regularly. Every call to Production on a Special Cause was followed by a “we don’t see anything wrong” response. Within two months I realized I was labeled as “the boy who cried wolf” and stopped plotting the data on my own.

Lessons Learned

As I reflect back on this experience, I see a number of valuable lessons:

- **Have an out-of-control reaction plan.** We knew what to do after every reactor run. Control Charts need a reaction plan. If you’re control charting and don’t have a reaction plan, start building the knowledge needed to have a reaction plan. This was something we did right and was an essential contributor to our success.

- **Make sure your out-of-control reaction plan is doable and justifiable.** Adjusting silane flow rate and rod temperature was easily done and had solid engineering justification. This justification was important in bringing about the change.

- **Make sure your out-of-control reaction plan is palatable.** Back then if I’d known about surface science and suggested part of our reaction plan would be adjustments/corrections to the slim rod process, I would have lost credibility and may not have been allowed to implement the control chart. Messing with the slim rod pullers was just not acceptable to Management. The best I could have done is identify which slim rod puller(s) produced the slim rods that were in a special cause run. If certain pullers always produced high morphology I think I would have been allowed to do some investigation to determine why.

- **When investigations lead nowhere, you may be looking in the wrong place.** After gaining some knowledge about the physics of surfaces, I’m convinced there was variation in our process due to the Slim Rods that could have been reduced. The crystal growth leading to high morphology was not only affected by rod temperature and gas flow rate but also by the Slim Rod surface, and I think there were occasional slim rod surfaces that varied from the norm. If our organization could have identified those events and taken effective corrective action, we could have further increased our throughput and yield to desirable product.

- **Don’t lose sight of the end goal.** Control charts build confidence that the process is predictable and stable. They identify undesirable disruptions so you can correct them. As valuable as these features are, to me the end goal of the chart is to make money. Private enterprises are not paid for correct control chart limits; they are paid based on the quality and quantity of their output. In spite of the flaws, our control charts led to us make more money.

- **Remember Shewhart chose 3σ because it made economic sense.** Shewhart wanted to separate common and special cause variation. Whatever line of separation is chosen some of the variation will be misidentified. Effort will be wasted investigating common cause variation. Sometimes a special cause will not be investigated. He chose 3σ for economic, not statistical reasons (see his book “Economic Control of Manufactured Product”). Unfortunately, in our modern day we emphasize the statistics and have forgotten the economics. Statisticians have provided technical explanations.
of why $3 \sigma$ works so well, and those limits should be the standard. If, however, the special causes are coming fast and furious and you’re upsetting Management with your investigations maybe the control chart limit should be greater than $3 \sigma$ (at least initially). I’m suggesting this idea only from a practicality, dealing-with-people perspective. Investigating and correcting only some special causes is better than not correcting any. In order to do that you may need to limit your special cause investigation to the “big ones” that show up when using expanded control limits. Both “big” special causes and “little” ones come from the same process, so they aren’t really any different. But the human psychology of finding small changes more palatable than big ones has to be considered in order to implement process improvement.

• **Think about what source of variation in your process is appropriate for calculating control limits.** I’ve already discussed the lack of a range or std deviation chart accompanying the average chart for the reactor runs. The “industry standard” would have had me develop the correct chart limits based on the variation *within* a reactor run. The rational subgrouping should be within a reactor run. There should only be common cause variation within a run. But the reactors do not run continuously. The rational subgroups are not coming from a steady stream of production. This is a batch operation with a new set up and it’s associated variation for every reactor run. A range or std deviation for within run variation may have been beneficial, but a mR chart for run to run variation would also have been helpful. And, given the batch nature of the process, the mR chart would have been the more appropriate source of variation for calculating the limits on the Xbar chart. It may be that my 1st set of incorrectly calculated limits were close to limits that would be calculated with reactor run to reactor run mR. It seems I got lucky and stumbled into limits that were reasonably appropriate and the later “corrected” limits were not.

• **Control Chart SPC can help you drive toward the target.** Even with the control chart flaws, we were able to move operation of the reactors to the desired target morphology with minimal powder fall. We made adjustments to hit the target morphology, using the control chart to decide when to adjust and when to leave the process alone.

This was one of the great experiences of my career. To make that significant of a difference for a company was extremely gratifying. Sadly, the company over-expanded, incurred a mountain of debt and got into serious financial trouble. I got out about a year before they were bought out. Fortunately most of my co-workers kept their jobs and the new company appears to be doing well.

Over the years I’ve decided that conditions within an organization have to be right in order to make substantial change of this kind. I believe that the combination of:

1. The company losing money,
2. 1 or 2 experienced engineers having solid rationale for believing there was potential gains in production quantity and quality without buying new equipment, and
3. New managers not afraid of change* were critical to trying a brand new process control approach. It just took a new, inexperienced guy to have an idea.

**About the Author**

Ryan Monson has 25 years’ experience improving processes in the Semiconductor, Textile and Medical Device industries in Elkton, MD. He holds a BS and MS in Chemical Engineering from Brigham Young University. He can be reached at ryandmonson@gmail.com. Or visit his LinkedIn profile at [https://www.linkedin.com/in/ryan-monson-02391212/](https://www.linkedin.com/in/ryan-monson-02391212/).

---

*i* Over the years I’ve learned that the social factors of change implementation require much greater effort and time than the technical factors.

*ii* The literature I read did not identify the need for bias correction.

*iii* Given how pleased Management was with the yield and throughput improvements using the current limits, I was not going to suggest a change until I demonstrated the new limits could make further improvements.

*iv* Japanese upper management, imported from the parent company, ensured that we tried this new process control approach. All the resistance came from the American management and a senior engineer. Their resistance ended once the improvements showed up on the balance.
Upcoming Conference Calendar

**Joint Statistical Meeting**  
29 July–3 August, 2017  
Baltimore, Minn.  

JSM (the Joint Statistical Meetings) is the largest gathering of statisticians held in North America. The JSM program consists not only of invited, topic-contributed, and contributed technical sessions, but also poster presentations, roundtable discussions, professional development courses and workshops, award ceremonies, and countless other meetings and activities.

**RSS 2017 Annual Conference**  
4–7 September 2017  
Glasgow, Scotland  

A three-day professional and academic conference with a mix of keynote presentations and talks/workshops. The latter run in parallel and are organized by topics, including Data science, Medical statistics, Official statistics, Methods and theory and many more.

**ENBIS-17**  
9–14 September 2017  
Naples, Italy  

The annual conference takes place every September and features presentations from a wide variety of sectors, ranging from manufacturing to service, from private to public sector. The conference booklet containing the program and abstracts will be published before the conference and inserted in delegates’ packs. A special number of *Quality and Reliability Engineering International (QREI)* featuring selected papers presented in ENBIS-17 will be published in 2018 following a special post-conference Call for Papers to the conference participants and the wider ENBIS community.

**Fall Technical Conference**  
5–6 October, 2017  
Philadelphia, PA  
[http://asq.org/conferences/fall-technical/](http://asq.org/conferences/fall-technical/)

The 2017 Fall technical Conference them is Statistics: Powering a Revolution in Quality Improvement. The goal of this conference is to engage researchers and practitioners in a dialogue that leads to more effective use of statistics to improve quality and foster innovation. Attendees should have a desire to participate, network, and discover new techniques to solve real-world problems!

**Conference on Statistical Practice**  
15–17 February 2018  
Portland, OR  
[https://ww2.amstat.org/meetings/csp/2018/](https://ww2.amstat.org/meetings/csp/2018/)

The Conference on Statistical Practice aims to bring together hundreds of statistical practitioners—including data analysts, researchers, and scientists—who engage in the application of statistics to solve real-world problems on a daily basis. The goal of the conference is to provide participants with opportunities to learn new statistical methodologies and best practices in statistical analysis, design, consulting, and statistical programming. The conference also will provide opportunities for attendees to further their career development and strengthen relationships in the statistical community.
Statistics Division Committee Roster 2017

OFFICERS

CHAIR
Herb McGrath
chair@asqstatdiv.org

CHAIR-ELECT
Steve Schuelka
chair@asqstatdiv.org
219-689-3804

TREASURER
Mindy Hatchkiss
mindy.hatchkiss@rocket.com
561-882-5331

SECRETARY
Jennifer Williams
JenniferStatistics@outlook.com
209-877-7784

PAST CHAIR
Theresa Utlaut
pastchair@asqstatdiv.org
503-613-7763

APPOINTED

Auditing

AUDIT CHAIR
Steve Schuelka
outreach@asqstatdiv.org
219-689-3805

By-Laws

BY-LAWS CHAIR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

Nominating

NOMINATING CHAIR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

Planning

PLANNING CHAIR
Theresa Utlaut
Theresa.L.Utlaut@intel.com
503-613-7763

Operations

OPERATIONS CHAIR
Joel Smith
joelmcquarriesmith@gmail.com
814-753-3224

MEMBERSHIP CHAIR
Gary Gehring
ggehring.q@gmail.com
306-787-8418

VOICE OF THE CUSTOMER CHAIR
Joel Smith
joelmcquarriesmith@gmail.com
814-753-3224

CERTIFICATION CHAIR
Brian Sersion
bsersion@gmail.com
513-363-0177

STANDARDS CHAIR
Mark Johnson
mejohnson@mail.ucf.edu
407-888-1746

Content

CONTENT CHAIR
Amy Ste. Croix
croix.amy@asqstatdiv.org
850-324-0904

NEWSLETTER EDITOR
Matthew Barsalou
mathew.barsalou@gmail.com
+49-152-05421794

WEBINAR COORDINATOR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

SOCIAL MEDIA MANAGER
Brian Sersion & Joshua Briggs
social@asqstatdiv.org
513-363-0177

WEBSITE AND INTERNET LIAISON
Landon Jensen
lu.jensen@colubus.rr.com
801-767-3328

Awards

AWARDS CHAIR
Peter Parker
peter.o.carper@nasa.gov
757-864-4709

OTT SCHOLARSHIP CHAIR
Lynne Hare
lynnehare@verizon.net
774-413-5268

FTC STUDENT/EARLY CAREER GRANTS
Jennifer Williams
JenniferStatistics@outlook.com
209-877-7784

HUNTER AWARD CHAIR
Open

NELSON AWARD CHAIR
Open

BISGAARD AWARD CHAIR
Open

Conferences

WCQI
Gary Gehring
membershipchair@asqstatdiv.org
306-787-8418

FTC STEERING COMMITTEE
Bill Myers
myers.wr@pg.com

FTC PROGRAM REPRESENTATIVE
Mindy Hatchkiss
mindy.hatchkiss@rocket.com
561-882-5331

FTC SHORT COURSE CHAIR
Yangtai Cao
yciao@up.edu
724-357-4767

Member Development

MEMBER DEVELOPMENT CHAIR
Gary Gehring
membershipschair@asqstatdiv.org
306-787-8418

OUTREACH/SPEAKER LIST CHAIR
Steve Schuelka
sschuelka@yahoo.com
219-689-3804

EXAMINING CHAIR
Daksha Chokshi
daksha.chokshi@rocket.com
561-796-8373

APPOINTED

Auditing

AUDIT CHAIR
Steve Schuelka
outreach@asqstatdiv.org
219-689-3805

By-Laws

BY-LAWS CHAIR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

Nominating

NOMINATING CHAIR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

Planning

PLANNING CHAIR
Theresa Utlaut
Theresa.L.Utlaut@intel.com
503-613-7763

Operations

OPERATIONS CHAIR
Joel Smith
joelmcquarriesmith@gmail.com
814-753-3224

MEMBERSHIP CHAIR
Gary Gehring
ggehring.q@gmail.com
306-787-8418

VOICE OF THE CUSTOMER CHAIR
Joel Smith
joelmcquarriesmith@gmail.com
814-753-3224

CERTIFICATION CHAIR
Brian Sersion
bsersion@gmail.com
513-363-0177

STANDARDS CHAIR
Mark Johnson
mejohnson@mail.ucf.edu
407-888-1746

Content

CONTENT CHAIR
Amy Ste. Croix
croix.amy@asqstatdiv.org
850-324-0904

NEWSLETTER EDITOR
Matthew Barsalou
mathew.barsalou@gmail.com
+49-152-05421794

WEBINAR COORDINATOR
Adam Pintar
adam.pintar@nist.gov
301-975-4554

SOCIAL MEDIA MANAGER
Brian Sersion & Joshua Briggs
social@asqstatdiv.org
513-363-0177

WEBSITE AND INTERNET LIAISON
Landon Jensen
lu.jensen@colubus.rr.com
801-767-3328

Awards

AWARDS CHAIR
Peter Parker
peter.o.carper@nasa.gov
757-864-4709

OTT SCHOLARSHIP CHAIR
Lynne Hare
lynnehare@verizon.net
774-413-5268

FTC STUDENT/EARLY CAREER GRANTS
Jennifer Williams
JenniferStatistics@outlook.com
209-877-7784

HUNTER AWARD CHAIR
Open

NELSON AWARD CHAIR
Open

BISGAARD AWARD CHAIR
Open

Conferences

WCQI
Gary Gehring
membershipchair@asqstatdiv.org
306-787-8418

FTC STEERING COMMITTEE
Bill Myers
myers.wr@pg.com

FTC PROGRAM REPRESENTATIVE
Mindy Hatchkiss
mindy.hatchkiss@rocket.com
561-882-5331

FTC SHORT COURSE CHAIR
Yangtai Cao
yciao@up.edu
724-357-4767

Member Development

MEMBER DEVELOPMENT CHAIR
Gary Gehring
membershipschair@asqstatdiv.org
306-787-8418

OUTREACH/SPEAKER LIST CHAIR
Steve Schuelka
sschuelka@yahoo.com
219-689-3804

EXAMINING CHAIR
Daksha Chokshi
daksha.chokshi@rocket.com
561-796-8373
The ASQ Statistics Division Newsletter is published three times a year by the Statistics Division of the American Society for Quality.

All communications regarding this publication, EXCLUDING CHANGE OF ADDRESS, should be addressed to:
Matthew Barsalou
Editor
email: newsletter@asqstatdiv.org

Other communications relating to the ASQ Statistics Division should be addressed to:
Richard Herb McGrath
Division Chair
email: chair@asqstatdiv.org

Communications regarding change of address should be sent to ASQ at:
ASQ
P.O. Box 3005
Milwaukee, WI 53201-3005

This will change the address for all publications you receive from ASQ. You can also handle this by phone (414) 272-8575 or (800) 248-1946.

Upcoming Newsletter Deadlines for Submissions

<table>
<thead>
<tr>
<th>Issue</th>
<th>Vol.</th>
<th>No.</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>36</td>
<td>3</td>
<td>August 16</td>
</tr>
</tbody>
</table>

To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.

Sir Ronald Fisher

VISIT THE STATISTICS DIVISION WEBSITE:

www.asq.org/statistics

ASQ Periodicals with Applied Statistics content

Journal of Quality Technology
http://www.asq.org/pub/jqt/

Quality Engineering
http://www.asq.org/pub/qe/

Six Sigma Forum
http://www.asq.org/pub/sixsigma/

STATISTICS DIVISION RESOURCES:

LinkedIn Statistics Division Group
https://www.linkedin.com/groups/ASQ-Statistics-Division-2115190

Scan this to visit our LinkedIn group!

Connect now by scanning this QR code with a smartphone (requires free QR app)

Check out our YouTube channel at youtube.com/asqstatsdivision