

What is the Probability That Your Process is Predictable?

How to compute a p -value for your process behavior chart.

Donald J. Wheeler

Software packages use p -values to report the results of many statistical procedures. As a result some people have come to expect a p -value as the outcome of any statistical analysis. This column will tell you how to compute and use a p -value for a process behavior chart.

BACKGROUND

Process behavior charts allow you to listen to the voice of your process. They allow you to characterize the process as being operated predictably or unpredictably. They help you to find assignable causes of exceptional variation and thereby to reduce the process variation and increase productivity. However, in this day of instant pudding, rather than taking the time to listen to the voice of the process, people want everything boiled down to a number they can put in the monthly report. At the risk of contributing to this practice, this column tells you how to compute and use a p -value for your process behavior chart.

Process behavior charts were designed for the sequential analysis of a continuing stream of observational data. Here the data generally represent one condition and the purpose of the chart is to identify unplanned changes in the underlying process. After the limits have been computed for some baseline period, they are extended forward and additional points are added to the chart as they become available. Each time we add another point to the chart we are performing an act of analysis, and each of these sequential analyses asks if the current value is consistent with the baseline period.

Because of this sequential nature, a process behavior chart has no fixed risk of a false alarm and no fixed risk of a missed signal. So how can we talk about a p -value for a process behavior chart? We can do this in the same way that we compute all of the other values associated with a process behavior chart—we use the baseline period. We compute the average, the average range, the limits, the capability indexes, and the performance indexes using that fixed amount of data we define as the baseline period. We shall do the same for the p -value.

WHAT IS THE PROBABILITY THAT YOUR PROCESS IS PREDICTABLE?

A p -value is a test statistic that is expressed as a probability. Under the condition that there is no difference between two quantities, a p -value is the probability of getting a result that is more unlikely than the observed result. So small p -values are associated with unlikely events and large p -values are associated with likely events (under the condition that no difference exists).

Here we shall use a p -value to ask the question: “What is the probability that this process

was operated predictably during the baseline period?" Since predictable operation provides a rational basis for using that product which we have measured to characterize the product that was not measured, this question of predictability is very important in practice. A small p -value will be an indication that the process is unlikely to have been operated predictably and our extrapolation to the unmeasured product becomes questionable.

Say we have a baseline that consists of k subgroups of size n , so that the total number of data in the baseline is $N = nk$. (In the case of an XmR chart we define $n = 1$.) We define the capability and performance indexes as follows:

$$\text{Capability Ratio} = C_p = \frac{USL - LSL}{6 \text{ Sigma}(X)} = \frac{\text{space available}}{\text{space required}}$$

$$\text{Performance Ratio} = P_p = \frac{USL - LSL}{6 s} = \frac{\text{space available}}{\text{space used in past}}$$

$$\text{Centered Capability Ratio} = C_{pk} = \frac{2 \text{ DNS}}{6 \text{ Sigma}(X)} = \frac{\text{effective space available}}{\text{space required}}$$

$$\text{Centered Performance Ratio} = P_{pk} = \frac{2 \text{ DNS}}{6 s} = \frac{\text{effective space available}}{\text{space used in past}}$$

The quantities in these formulas are defined as follows. The difference between the specification limits, $USL - LSL$, is the specified tolerance. The distance to the nearer specification, DNS , is the distance from the average to the nearer specification limit. $\text{Sigma}(X)$ denotes any one of several within-subgroup measures of dispersion, such as the average of the subgroup ranges divided by the appropriate bias correction factor, d_2 . And s is the global standard deviation statistic computed using all N data in the baseline period.

THE PREDICTABILITY RATIO

Using a baseline consisting of N data, define the *predictability ratio* as the capability ratio, C_p , divided by the performance ratio, P_p .

$$\text{Predictability Ratio} = PR = \frac{C_p}{P_p}$$

Next we compare our observed value for the predictability ratio with the maximum one-percent critical value for N data (found in Figure 2). If your computed predictability ratio exceeds the maximum critical value for N data, you have less than a one percent chance that your process has been operated predictably. This is evidence that is strong enough to convince a skeptic that the process was operated unpredictably.

If your predictability ratio is noticeably smaller than the maximum critical value you can say that its p -value will be larger than one percent. However, in this case the process may, or may not, have been operated predictably. The only way to judge that a process displays a reasonable degree of predictability is to use the process behavior chart.

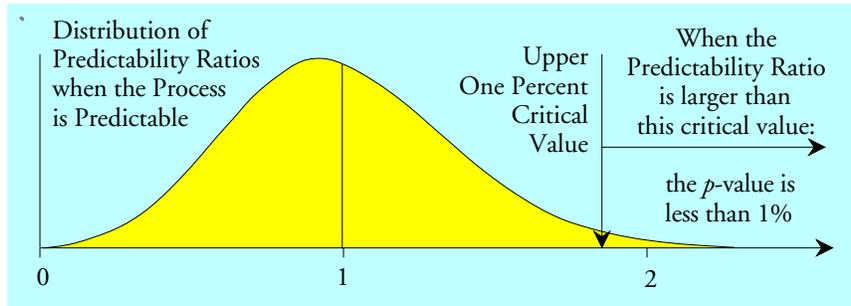


Figure 1: The Relationship between the Predictability Ratio and the *p*-Value

| <i>N</i> | 1% <i>CIP</i> |
|----------|---------------|----------|---------------|----------|---------------|----------|---------------|
| 10 | 3.40 | 50 | 1.59 | 90 | 1.40 | 300 | 1.19 |
| 12 | 2.95 | 52 | 1.57 | 92 | 1.39 | 340 | 1.18 |
| 14 | 2.66 | 54 | 1.56 | 94 | 1.39 | 380 | 1.17 |
| 16 | 2.48 | 56 | 1.55 | 96 | 1.38 | 420 | 1.16 |
| 18 | 2.32 | 58 | 1.53 | 98 | 1.38 | 460 | 1.15 |
| 20 | 2.20 | 60 | 1.52 | 100 | 1.37 | 500 | 1.15 |
| 22 | 2.10 | 62 | 1.51 | 110 | 1.35 | 600 | 1.13 |
| 24 | 2.03 | 64 | 1.50 | 120 | 1.33 | 700 | 1.12 |
| 26 | 1.96 | 66 | 1.49 | 130 | 1.32 | 800 | 1.11 |
| 28 | 1.91 | 68 | 1.48 | 140 | 1.30 | 1000 | 1.10 |
| 30 | 1.86 | 70 | 1.47 | 150 | 1.29 | 1200 | 1.09 |
| 32 | 1.82 | 72 | 1.46 | 160 | 1.28 | 1400 | 1.08 |
| 34 | 1.78 | 74 | 1.45 | 170 | 1.27 | 1800 | 1.07 |
| 36 | 1.75 | 76 | 1.44 | 180 | 1.26 | 2600 | 1.06 |
| 38 | 1.72 | 78 | 1.44 | 190 | 1.25 | 3400 | 1.05 |
| 40 | 1.69 | 80 | 1.43 | 200 | 1.25 | 5000 | 1.04 |
| 42 | 1.67 | 82 | 1.42 | 220 | 1.23 | 10000 | 1.03 |
| 44 | 1.65 | 84 | 1.42 | 240 | 1.22 | 15000 | 1.02 |
| 46 | 1.63 | 86 | 1.41 | 260 | 1.21 | 45000 | 1.01 |
| 48 | 1.61 | 88 | 1.40 | 280 | 1.20 | | |

Figure 2: Maximum 1% Critical Values for the Predictability Ratio

In the interest of simplicity, Figure 2 assumes that either an average range or average standard deviation has been used to compute the limits. (For charts that use a median range, you will need to compute an exact *p*-value using the stability ratio defined below.)

EXAMPLE ONE

The first example will use the ball-joint socket thickness data. Ninety-six values were collected over the course of one week and organized into 24 subgroups of size 4. The capability and performance indexes were $C_p = 1.48$, $P_p = 1.43$, $C_{pk} = 0.95$, and $P_{pk} = 0.92$. Thus the predictability ratio is:

$$Predictability\ Ratio = PR = \frac{1.48}{1.43} = 1.035$$

For $N = 96$ values the maximum one-percent critical value is 1.38. Since the observed

predictability ratio of 1.035 is smaller than the one-percent critical value of 1.38, these data have a *p*-value larger than one percent, and this process might be predictable. The average and range chart in Figure 3 confirms that this process was indeed operated predictably during the baseline period.

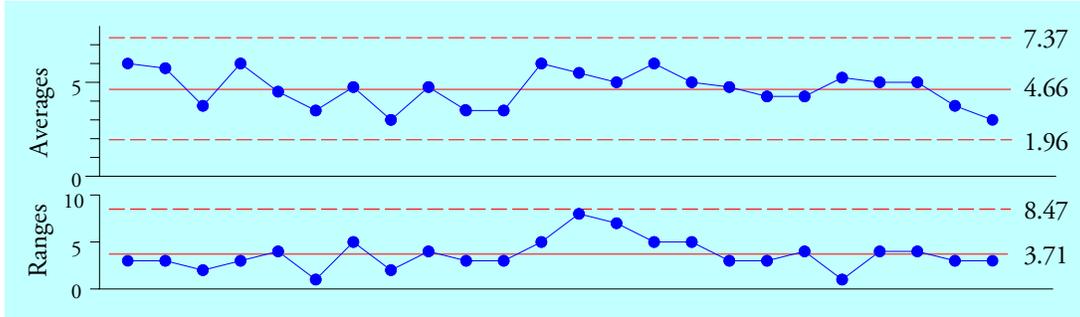


Figure 3: Average and Range Chart for Ball Joint Socket Thicknesses

EXAMPLE TWO

The creel yield data for one week consist of thirty-three values placed on an *XmR* chart. The capability and performance indexes are $C_p = 5.38$, $P_p = 2.40$, $C_{pk} = 2.00$, and $P_{pk} = 0.90$. Thus the predictability ratio is:

$$Predictability\ Ratio = PR = \frac{5.38}{2.40} = 2.242$$

With $N = 33$ values we use the maximum critical value for $N = 32$ values which is 1.82. Since our predictability ratio of 2.24 exceeds this critical value of 1.82, we know that these data have a *p*-value that is smaller than one percent, and this process is very unlikely to have been operated predictably. The *XmR* chart in Figure 4 confirms this interpretation.

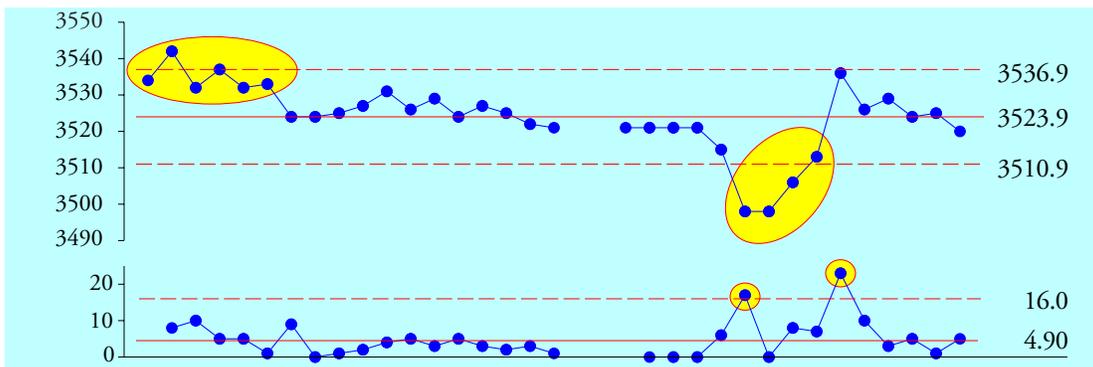


Figure 4: *XmR* Chart for Creel Yields

Inspection of the formulas given earlier for the capability and performance indexes will quickly reveal that the predictability ratio may be computed using any one of three ratios:

$$\text{Predictability Ratio} = PR = \frac{C_p}{P_p} = \frac{C_{pk}}{P_{pk}} = \frac{s}{\text{Sigma}(X)}$$

When computing a ratio of ratios it is possible for round-off to produce small differences in the results from the different formulas above. When no specifications are given you can still compute a predictability ratio by dividing the global standard deviation statistic, s , by your within-subgroup measure of dispersion, $\text{Sigma}(X)$. Common within-subgroup formulas for $\text{Sigma}(X)$ are:

$$\text{Sigma}(X) = \frac{\text{average range}}{d_2} \text{ or } \frac{\text{median range}}{d_4} \text{ or } \frac{\text{avg. std. dev.}}{c_4}$$

where d_2 , d_4 , and c_4 are the usual bias correction factors found in most SPC books.

EXAMPLE THREE

The camshaft bearing diameter data consist of 50 values placed on an XmR chart. The average moving range is 1.510. Dividing by the bias correction factor of $d_2 = 1.128$ we get a $\text{Sigma}(X)$ of 1.3388. The global standard deviation statistic is $s = 1.6807$. Thus, the predictability ratio is:

$$\text{Predictability Ratio} = \frac{s}{\text{Sigma}(X)} = \frac{1.6807}{1.3388} = 1.255$$

From Figure 2, the maximum one-percent critical value for $N = 50$ is 1.59. Since the predictability ratio of 1.255 is less than this critical value we conclude that the p -value for these data is greater than 0.01. However, this large p -value does not guarantee that this process was operated predictably. It just means that *this numerical summary* does not provide strong evidence of unpredictability (notice the double negative).

The X chart for the camshaft bearing diameters in Figure 5 is much more informative than the predictability ratio. While this process is not terribly unpredictable, it does show evidence of occasional excursions. Since each point on the chart represents 50 parts produced, these excursions represent potential problems.

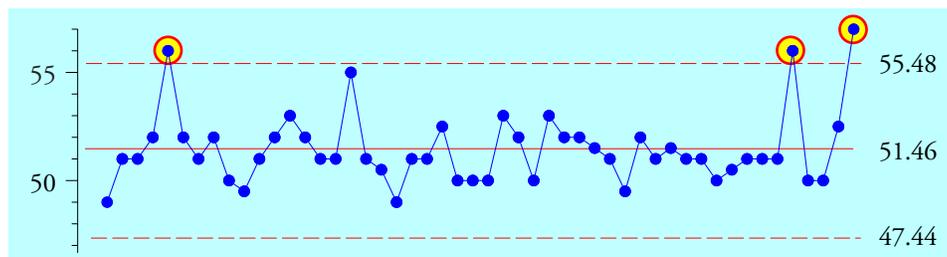


Figure 5: X Chart for Camshaft Bearing Diameter Data

The predictability ratio uses values that are commonly available to provide a quick check on the predictability of your process. Knowledge that the p -value is less than one percent is sufficient to provide reasonable certainty that your process is not being operated up to its full potential. However, if an exact p -value is desired you will need to use the *stability ratio*.

THE STABILITY RATIO

In 2006 Brenda Ramirez and George Runger suggested using the square of the predictability ratio as a measure of process stability over time. They noted that the stability ratio, *SR*, defined as:

$$SR = [PR]^2 = \frac{s^2}{[\text{Sigma}(X)]^2}$$

will behave as a pseudo-F statistic. The numerator degrees of freedom for this F-distribution will be $[N-1]$. The denominator degrees of freedom will be the degrees of freedom for the within-subgroup statistic used to compute *Sigma(X)*. So an exact *p*-value will depend upon three values: the value of the stability ratio, *SR*; the numerator degrees of freedom, $[N-1]$; and the denominator degrees of freedom based on how we computed *Sigma(X)*. The next three sections will provide ways to find the denominator degrees of freedom.

CASE ONE: SIGMA(X) IS BASED ON THE AVERAGE RANGE

When we use the average range (or the average moving range) to compute *Sigma(X)* for a baseline period consisting of *k* subgroups of size *n*, we can look up the degrees of freedom from Figure 6, or approximate them using the formulas in the last row.

| Effective Degrees of Freedom for an Average Range Statistic | | | | | | | | | | |
|---|--------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | <i>n</i> = 1 | <i>n</i> = 2 | <i>n</i> = 3 | <i>n</i> = 4 | <i>n</i> = 5 | <i>n</i> = 6 | <i>n</i> = 7 | <i>n</i> = 8 | <i>n</i> = 9 | <i>n</i> = 10 |
| <i>k</i> = 2 | 1.0 | 1.9 | 3.8 | 5.7 | 7.5 | 9.2 | 10.8 | 12.3 | 13.8 | 15.2 |
| <i>k</i> = 3 | 1.6 | 2.8 | 5.7 | 8.4 | 11.1 | 13.6 | 16.0 | 18.3 | 20.5 | 22.6 |
| <i>k</i> = 4 | 2.2 | 3.7 | 7.5 | 11.2 | 14.7 | 18.1 | 21.3 | 24.4 | 27.3 | 30.1 |
| <i>k</i> = 5 | 2.8 | 4.6 | 9.3 | 13.9 | 18.4 | 22.6 | 26.6 | 30.4 | 34.0 | 37.5 |
| <i>k</i> = 6 | 3.4 | 5.5 | 11.1 | 16.7 | 22.0 | 27.0 | 31.8 | 36.4 | 40.8 | 45.0 |
| <i>k</i> = 7 | 4.0 | 6.4 | 12.9 | 19.4 | 25.6 | 31.5 | 37.1 | 42.5 | 47.6 | 52.4 |
| <i>k</i> = 8 | 4.6 | 7.2 | 14.8 | 22.1 | 29.3 | 36.0 | 42.4 | 48.5 | 54.3 | 59.9 |
| <i>k</i> = 9 | 5.3 | 8.1 | 16.6 | 24.9 | 32.8 | 40.4 | 47.7 | 54.5 | 61.1 | 67.3 |
| <i>k</i> = 10 | 5.9 | 9.0 | 18.4 | 27.6 | 36.5 | 44.9 | 52.9 | 60.6 | 67.8 | 74.8 |
| <i>k</i> = 11 | 6.5 | 9.9 | 20.2 | 30.3 | 40.1 | 49.4 | 58.2 | 66.6 | 74.6 | 82.2 |
| <i>k</i> = 12 | 7.1 | 10.7 | 22.0 | 33.0 | 43.7 | 53.8 | 63.5 | 72.6 | 81.4 | 89.7 |
| <i>k</i> = 13 | 7.7 | 11.6 | 23.8 | 35.8 | 47.3 | 58.3 | 68.7 | 78.6 | 88.1 | 97.2 |
| <i>k</i> = 14 | 8.3 | 12.5 | 25.7 | 38.6 | 51.0 | 62.8 | 74.0 | 84.7 | 94.9 | 104.6 |
| <i>k</i> = 15 | 8.9 | 13.4 | 27.5 | 41.3 | 54.6 | 67.2 | 79.3 | 90.7 | 101.6 | 112.1 |
| <i>k</i> = 16 | 9.5 | 14.2 | 29.3 | 44.0 | 58.2 | 71.7 | 84.5 | 96.7 | 108.4 | 119.5 |
| <i>k</i> = 17 | 10.1 | 15.1 | 31.1 | 46.8 | 61.8 | 76.2 | 89.8 | 102.8 | 115.1 | 127.0 |
| <i>k</i> = 18 | 10.7 | 16.0 | 32.9 | 49.5 | 65.5 | 80.6 | 95.1 | 108.8 | 121.9 | 134.4 |
| <i>k</i> = 19 | 11.3 | 16.9 | 34.7 | 52.3 | 69.1 | 85.1 | 100.3 | 114.8 | 128.7 | 141.9 |
| <i>k</i> = 20 | 11.9 | 17.8 | 36.5 | 55.0 | 72.7 | 89.6 | 105.6 | 120.9 | 135.4 | 149.3 |
| <i>k</i> = 21 | 12.5 | 18.6 | 38.4 | 57.7 | 76.3 | 94.0 | 110.9 | 126.9 | 142.2 | 156.8 |
| <i>k</i> = 22 | 13.1 | 19.5 | 40.2 | 60.5 | 80.0 | 98.5 | 116.1 | 132.9 | 148.9 | 164.2 |
| <i>k</i> = 23 | 13.7 | 20.4 | 42.0 | 63.2 | 83.6 | 103.0 | 121.4 | 139.0 | 155.7 | 171.7 |
| <i>k</i> = 24 | 14.3 | 21.3 | 43.8 | 66.0 | 87.2 | 107.4 | 126.7 | 145.0 | 162.4 | 179.2 |
| <i>k</i> = 25 | 15.0 | 22.1 | 45.6 | 68.7 | 90.8 | 111.9 | 131.9 | 151.0 | 169.2 | 186.6 |
| any <i>k</i> | below | 0.88 <i>k</i> | 1.82 <i>k</i> | 2.74 <i>k</i> | 3.62 <i>k</i> | 4.47 <i>k</i> | 5.27 <i>k</i> | 6.03 <i>k</i> | 6.76 <i>k</i> | 7.45 <i>k</i> |

Figure 6: Denominator Degrees of Freedom based on the Average Range Statistic

The formulas in the last row allow you to extend the table in Figure 6 to larger numbers of

subgroups. For XmR charts the degrees of freedom for the average moving range can be approximated by the formula:

$$\text{degrees of freedom} \approx 1 + 0.605 [k - 2]$$

CASE TWO: $SIGMA(X)$ IS BASED ON THE AVERAGE STANDARD DEVIATION

When we use the average standard deviation statistic to compute $Sigma(X)$ for a baseline period consisting of k subgroups of size n , we can look up the degrees of freedom from Figure 7, or approximate them using the formulas in the last row.

Effective Degrees of Freedom for an Average Standard Deviation Statistic

| | $n = 2$ | $n = 3$ | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ | $n = 10$ |
|--------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| $k = 2$ | 1.9 | 3.9 | 5.8 | 7.8 | 9.8 | 11.8 | 13.8 | 15.8 | 17.8 |
| $k = 3$ | 2.8 | 5.7 | 8.7 | 11.6 | 14.6 | 17.6 | 20.6 | 23.6 | 26.6 |
| $k = 4$ | 3.7 | 7.6 | 11.5 | 15.4 | 19.4 | 23.4 | 27.4 | 31.3 | 35.3 |
| $k = 5$ | 4.6 | 9.4 | 14.3 | 19.2 | 24.2 | 29.2 | 34.1 | 39.1 | 44.1 |
| $k = 6$ | 5.5 | 11.2 | 17.1 | 23.0 | 29.0 | 34.9 | 40.9 | 46.9 | 52.9 |
| $k = 7$ | 6.4 | 13.0 | 19.9 | 26.8 | 33.8 | 40.7 | 47.7 | 54.7 | 61.6 |
| $k = 8$ | 7.2 | 14.9 | 22.7 | 30.6 | 38.5 | 46.5 | 54.5 | 62.5 | 70.4 |
| $k = 9$ | 8.1 | 16.7 | 25.5 | 34.4 | 43.3 | 52.3 | 61.3 | 70.2 | 79.2 |
| $k = 10$ | 9.0 | 18.5 | 28.3 | 38.2 | 48.1 | 58.0 | 68.0 | 78.0 | 88.0 |
| $k = 11$ | 9.9 | 20.4 | 31.1 | 42.0 | 52.9 | 63.8 | 74.8 | 85.8 | 96.7 |
| $k = 12$ | 10.7 | 22.2 | 34.0 | 45.8 | 57.7 | 69.6 | 81.6 | 93.6 | 105.5 |
| $k = 13$ | 11.6 | 24.0 | 36.8 | 49.6 | 62.5 | 75.4 | 88.4 | 101.3 | 114.3 |
| $k = 14$ | 12.5 | 25.9 | 39.6 | 53.4 | 67.3 | 81.2 | 95.2 | 109.1 | 123.0 |
| $k = 15$ | 13.4 | 27.7 | 42.4 | 57.2 | 72.1 | 87.0 | 101.9 | 116.9 | 131.8 |
| $k = 16$ | 14.2 | 29.5 | 45.2 | 61.0 | 76.9 | 92.7 | 108.7 | 124.7 | 140.6 |
| $k = 17$ | 15.1 | 31.3 | 48.0 | 64.8 | 81.7 | 98.5 | 115.5 | 132.5 | 149.3 |
| $k = 18$ | 16.0 | 33.2 | 50.8 | 68.6 | 86.4 | 104.3 | 122.3 | 140.3 | 158.1 |
| $k = 19$ | 16.9 | 35.0 | 53.6 | 72.4 | 91.2 | 110.1 | 129.0 | 148.0 | 166.9 |
| $k = 20$ | 17.8 | 36.8 | 56.4 | 76.2 | 96.0 | 115.9 | 135.8 | 155.8 | 175.7 |
| $k = 21$ | 18.6 | 38.7 | 59.2 | 80.0 | 100.8 | 121.6 | 142.6 | 163.6 | 184.4 |
| $k = 22$ | 19.5 | 40.5 | 62.1 | 83.8 | 105.6 | 127.4 | 149.4 | 171.4 | 193.2 |
| $k = 23$ | 20.4 | 42.3 | 64.9 | 87.6 | 110.4 | 133.2 | 156.2 | 179.2 | 202.0 |
| $k = 24$ | 21.3 | 44.2 | 67.7 | 91.4 | 115.2 | 139.0 | 163.0 | 186.9 | 210.7 |
| $k = 25$ | 22.1 | 46.0 | 70.4 | 95.1 | 119.9 | 144.8 | 169.7 | 194.6 | 219.5 |
| <i>any k</i> | $0.88 k$ | $1.83 k$ | $2.81 k$ | $3.80 k$ | $4.80 k$ | $5.79 k$ | $6.79 k$ | $7.78 k$ | $8.78 k$ |

Figure 7: Denominator Degrees of Freedom based on the Average Standard Deviation Statistic

The formulas in the last row allow you to extend the table in Figure 7 to larger numbers of subgroups. For subgroup sizes greater than 10 the degrees of freedom for the average standard deviation statistic may be approximated using the formula:

$$\text{degrees of freedom} \approx 0.98 k [n - 1]$$

CASE THREE: $SIGMA(X)$ IS BASED ON THE MEDIAN RANGE

When we use a median range (or a median moving range) to compute $Sigma(X)$ for a baseline period consisting of k subgroups of size n , we can look up the degrees of freedom from Figure 8, or approximate them using the formulas in Figure 9.

| Effective Degrees of Freedom for a Median Range Statistic | | | | | | | | | | |
|---|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|
| | $n = 1$ | $n = 2$ | $n = 3$ | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ | $n = 10$ |
| $k = 2$ | 1.0 | 1.9 | 3.8 | 5.7 | 7.5 | 9.2 | 10.8 | 12.3 | 13.8 | 15.2 |
| $k = 3$ | 1.6 | 1.9 | 3.9 | 5.9 | 8.0 | 9.8 | 11.6 | 13.2 | 14.8 | 16.5 |
| $k = 4$ | 1.5 | 2.7 | 5.7 | 8.8 | 11.9 | 14.6 | 17.2 | 19.8 | 22.2 | 24.6 |
| $k = 5$ | 2.2 | 2.6 | 5.7 | 9.0 | 12.2 | 14.9 | 17.8 | 20.5 | 23.0 | 25.5 |
| $k = 6$ | 2.2 | 3.4 | 7.6 | 12.0 | 16.0 | 19.8 | 23.4 | 26.5 | 29.2 | 32.0 |
| $k = 7$ | 2.7 | 3.3 | 7.6 | 12.0 | 16.2 | 20.0 | 23.5 | 26.7 | 29.2 | 32.0 |
| $k = 8$ | 2.7 | 4.1 | 9.5 | 15.0 | 20.2 | 24.6 | 29.0 | 32.6 | 35.3 | 38.6 |
| $k = 9$ | 3.3 | 4.1 | 9.5 | 15.1 | 20.3 | 24.6 | 28.7 | 32.6 | 35.1 | 38.3 |
| $k = 10$ | 3.4 | 4.8 | 11.3 | 18.1 | 24.2 | 29.0 | 33.7 | 38.1 | 40.8 | 44.6 |
| $k = 11$ | 3.9 | 4.8 | 11.4 | 18.1 | 24.1 | 28.8 | 33.3 | 38.1 | 40.8 | 44.4 |
| $k = 12$ | 4.0 | 5.6 | 13.2 | 21.1 | 27.7 | 32.9 | 37.8 | 43.2 | 46.3 | 50.0 |
| $k = 13$ | 4.5 | 5.5 | 13.2 | 21.1 | 27.5 | 32.7 | 37.5 | 43.1 | 46.4 | 49.8 |
| $k = 14$ | 4.6 | 6.2 | 15.1 | 24.0 | 31.0 | 36.7 | 42.0 | 48.0 | 51.5 | 55.0 |
| $k = 15$ | 5.1 | 6.2 | 15.0 | 23.9 | 30.7 | 36.5 | 41.7 | 47.7 | 51.5 | 54.9 |
| $k = 16$ | 5.2 | 7.0 | 16.9 | 26.7 | 34.3 | 40.5 | 46.2 | 52.6 | 56.3 | 60.1 |
| $k = 17$ | 5.7 | 7.0 | 16.8 | 26.7 | 34.0 | 40.3 | 45.9 | 52.3 | 56.3 | 60.0 |
| $k = 18$ | 5.8 | 7.6 | 18.7 | 29.4 | 37.6 | 44.3 | 50.4 | 57.1 | 61.1 | 65.1 |
| $k = 19$ | 6.3 | 7.6 | 18.6 | 29.4 | 37.3 | 44.1 | 50.1 | 56.8 | 61.1 | 65.0 |
| $k = 20$ | 6.4 | 8.4 | 20.5 | 32.2 | 40.9 | 48.1 | 54.6 | 61.7 | 65.9 | 70.2 |
| $k = 21$ | 6.9 | 8.4 | 20.4 | 32.1 | 40.6 | 47.9 | 54.3 | 61.3 | 65.9 | 70.1 |
| $k = 22$ | 7.0 | 9.1 | 22.3 | 34.9 | 44.1 | 51.9 | 58.8 | 66.2 | 70.7 | 75.3 |
| $k = 23$ | 7.6 | 9.1 | 22.2 | 34.8 | 43.9 | 51.7 | 58.5 | 65.9 | 70.7 | 75.1 |
| $k = 24$ | 7.6 | 9.8 | 24.1 | 37.6 | 47.4 | 55.7 | 63.0 | 70.7 | 75.5 | 80.3 |
| $k = 25$ | 8.2 | 9.8 | 24.0 | 37.5 | 47.1 | 55.5 | 62.7 | 70.4 | 75.5 | 80.2 |

Figure 8: Denominator Degrees of Freedom based on the Median Range Statistic

The stair-step nature of the values in each column of Figure 8 complicates the problem of approximating the degrees of freedom. For XmR charts (where $n = 1$) the formulas will be given in terms of odd values of k , and the degrees of freedom for an even value of k will be approximately the same as for $k-1$.

For average and range charts (where $n \geq 2$) the formulas will be given in terms of even values for k , and the degrees of freedom for odd values of k will be approximately the same as for $k-1$.

The formulas for approximating and extending Figure 8 may be found in Figure 9.

FINDING AN EXACT p -VALUE

Thus, the p -value for your process behavior chart will depend upon three quantities: the observed value for the stability ratio, SR ; the numerator degrees of freedom, $[N-1]$; and the appropriate denominator degrees of freedom from Figure 6, 7, 8, or 9.

In Excel you can use the FDIST function to obtain the p -value for the computed stability ratio by entering the following formula in a cell:

$$= \text{FDIST} (SR , N-1 , \text{denominator } d.f.)$$

and Excel will return the p -value.

Formulas for Effective Degrees of Freedom for a Median Range Statistic

| Chart | n | values for k | Approx. Degrees of Freedom |
|-----------------|----------|-----------------------|--------------------------------------|
| $\bar{X}mR$ | $n = 1$ | odd $k \leq 49$ | $d.f. \approx 1.6 + 0.30 [k - 3]$ |
| | $n = 1$ | odd $k \geq 51$ to 99 | $d.f. \approx 16.0 + 0.27 [k - 51]$ |
| | $n = 1$ | odd $k \geq 101$ | $d.f. \approx 29.5 + 0.25 [k - 101]$ |
| Average & Range | $n = 2$ | even k | $d.f. \approx 2.0 + 0.36 [k - 2]$ |
| Average & Range | $n = 3$ | even k | $d.f. \approx 4.2 + 0.90 [k - 2]$ |
| Average & Range | $n = 4$ | even $k \leq 12$ | $d.f. \approx 5.7 + 1.53 [k - 2]$ |
| | $n = 4$ | even $k \geq 14$ | $d.f. \approx 24.0 + 1.36 [k - 14]$ |
| Average & Range | $n = 5$ | even $k \leq 12$ | $d.f. \approx 7.5 + 2.02 [k - 2]$ |
| | $n = 5$ | even $k \geq 14$ | $d.f. \approx 31 + 1.64 [k - 14]$ |
| Average & Range | $n = 6$ | even $k \leq 10$ | $d.f. \approx 9.6 + 2.50 [k - 2]$ |
| | $n = 6$ | even $k \geq 12$ | $d.f. \approx 32.9 + 1.90 [k - 12]$ |
| Average & Range | $n = 7$ | even $k \leq 8$ | $d.f. \approx 10.8 + 3.03 [k - 2]$ |
| | $n = 7$ | even $k \geq 10$ | $d.f. \approx 33.5 + 2.10 [k - 10]$ |
| Average & Range | $n = 8$ | even $k \leq 8$ | $d.f. \approx 12.8 + 3.40 [k - 2]$ |
| | $n = 8$ | even $k \geq 10$ | $d.f. \approx 38.6 + 2.27 [k - 10]$ |
| Average & Range | $n = 9$ | $k = 2$ | $d.f. = 13.8$ |
| | $n = 9$ | $k = 3$ | $d.f. = 14.8$ |
| | $n = 9$ | even $k \leq 10$ | $d.f. \approx 22.6 + 3.10 [k - 4]$ |
| | $n = 9$ | even $k \geq 12$ | $d.f. \approx 46.5 + 2.40 [k - 12]$ |
| Average & Range | $n = 10$ | $k = 2$ | $d.f. = 15.2$ |
| | $n = 10$ | $k = 3$ | $d.f. = 16.5$ |
| | $n = 10$ | even $k \leq 10$ | $d.f. \approx 25 + 3.30 [k - 4]$ |
| | $n = 10$ | even $k \geq 12$ | $d.f. \approx 50 + 2.53 [k - 12]$ |

Figure 9: Degrees of Freedom Formulas for Median Range Statistics

EXACT p -VALUE FOR EXAMPLE ONE

Recall that the ball-joint socket data were organized in an average and range chart with $n = 4$, $k = 24$, and $N = 96$. The stability ratio is:

$$\text{Stability Ratio} = SR = PR^2 = (1.035)^2 = 1.071$$

The numerator d.f. is $N-1 = 95$, and from Figure 6 the denominator d.f. is 66. From these three quantities we get a p -value of 0.387. This value may be interpreted as the likelihood that these baseline data came from a predictable process. This was confirmed by what we found in Figure 3.

EXACT p -VALUE FOR EXAMPLE TWO

Recall that the creel yield data were placed on an XmR chart with $n = 1$, $k = 33$, and so $N = 33$. The stability ratio is:

$$\text{Stability Ratio} = SR = PR^2 = (2.242)^2 = 5.025$$

The numerator d.f. is $N-1 = 32$, and from the formula following Figure 6, the denominator d.f. is $[1 + 0.605(k-2)] = 19.8$. From these three quantities we get a p -value of 0.00025 for this chart. This tiny p -value represents the astronomically remote likelihood that these baseline data came from a predictable process. Thus, we conclude that it is more likely that these data came from an unpredictable process, which is what we found in Figure 4.

EXACT p -VALUE FOR EXAMPLE THREE

The camshaft bearing diameter data were placed on an XmR chart with $n = 1$, $k = 50$, and so $N = 50$. The stability ratio is:

$$\text{Stability Ratio} = SR = PR^2 = (1.255)^2 = 1.576$$

The numerator d.f. is $N-1 = 49$, and from the formula following Figure 6 the denominator d.f. is $[1 + 0.605(k-2)] = 30.0$. From these three quantities we get a p -value of 0.093 for this chart. A predictable process with a baseline of 50 values could have a stability ratio this size or greater about nine percent of the time. So while the stability ratio does not provide strong evidence of unpredictability, the X chart in Figure 5 shows 3 out of 50 values outside the limits, and this process is properly judged to be unpredictable. This is why only very small p -values provide an unequivocal interpretation, and larger p -values are ambiguous.

THE ORIGIN OF FIGURE TWO

Instead of computing specific p -values, Figure 2 provides cut-offs that allow you to classify a p -value as larger or smaller than 0.01. To get the values in Figure 2 the one-percent critical values for the stability ratio were computed for different combinations of n and k . Next these critical values were converted into critical values for the predictability ratio and plotted versus the value for N . For each value of N these critical values turned out to all be very similar. This similarity allowed the simplification of tabling the maximum one-percent critical value for each value of N to produce the table in Figure 2.

When the critical values for limits based on a median range were added to the mix the strong similarity between critical values observed earlier for each value of N was no longer present. (This was due to the substantial differences in degrees of freedom when using median ranges.) So, when using a median range you cannot use Figure 2 to characterize the predictability ratio, but will instead need to find an exact p -value using the stability ratio.

SUMMARY

The p -value for either the predictability ratio or the stability ratio may be used as a one-number summary to characterize the predictability of a process during a baseline period. These ratios may be computed from capability and performance ratios, or they may be computed directly using the global standard deviation statistic and a within-subgroup measure of dispersion based on the average range, the median range, or the average standard deviation.

The predictability ratio, PR , may be used with the table in Figure 2 to characterize the p -value as being larger or smaller than 0.01. When the p -value can be shown to be smaller than 0.01 the unpredictability of the process is beyond reasonable doubt.

The square of the predictability ratio is known as the stability ratio, SR . It can be used to obtain an exact p -value using an F-distribution. This requires finding the denominator degrees of freedom, but the values and formulas in Figures 6, 7, 8, and 9 simplify this computation. Finding an exact p -value for the stability ratio provides a numerical summary that quantifies in a general way the likelihood that a particular process is being operated predictably.

While a very small p -value is a sure sign of an unpredictable process, a larger p -value is no guarantee of predictability. This is because no formula or algorithm can detect all types of unpredictable behavior. Every formula will have its blind spots, and aggregate summaries like the stability ratio are no exception.

When your stability ratio or predictability ratio has a small p -value you should know that the summary and descriptive statistics you have computed using your historical baseline data *will not characterize the future operation of the process*. When the p -value is small it means that the process average and the process standard deviation are changing over time, and thus the capability and performance indexes will also be changing. However, when your current p -value is small, you can expect that future p -values for your process are likely to remain small until you take action to operate the process predictably.

When process behavior charts are used as a sequential procedure to listen to the voice of the process in real time, reasonable baselines will generally contain somewhere between 20 and 150 data. However, many software packages dump all of the historical data into the baseline. This practice treats a process behavior chart as a one-time analysis procedure. When this happens you may have baselines consisting of thousands of data. This is why Figure 2 contains such large values for N .

There are two drawbacks to using a process behavior chart as a one-time analysis. First, when your baseline contains more than a few hundred data you are virtually guaranteed to find a small p -value. Second, by the time you have found the signals you will usually have forgotten what happened to create the shifts and spikes seen on your chart, making it impossible for you to use the chart to learn how to improve your process.

However, now you have a p -value for the monthly report that quantifies the probability that your process was operated predictably. The next step is to figure out what to do about your unpredictable processes.

