The Estimated Standard Deviation and Control Charts

November 2012

One of the purposes of control charts is to estimate the average and standard deviation of a process. The average is easy to calculate and understand – it is just the average of all the results. The standard deviation is a little more difficult to understand – and to complicate things, there are multiple ways that it can be determined – each giving a different answer.

σ

Sometimes people ask why some software packages give different values for the control limits. Which program is correct? The answer is probably both. The difference is simply how the standard deviation is estimated. The objective of this newsletter is to show three different, but common, ways that the standard deviation may be estimated. We will look at data that are formed into subgroups and the control limits on the $\overline{X}$ chart.

In this issue:

- **Data Subgroups**
- **Three Ways to Estimate the Standard Deviation**
- **Average of Subgroup Ranges**
- **Average of Subgroup Standard Deviations**
- **Pooled Standard Deviation**
- **Summary**

- **Quick Links**

As always, you can leave comments at the end of the newsletter.

DATA SUBGROUPS

The data we will use are shown in the table. We have 10 subgroups, each containing 3 observations or results.

Table 1: Subgroup Data

Subgroup	X_1	X_2	X_3	Subgroup Average	Subgroup R	Subgroup s
1	80.3	86.9	108.0	91.73	27.7	14.47
2	99.4	89.5	96.4	95.10	9.9	5.08
3	95.1	95.9	85.3	92.10	10.6	5.90
4	99.0	123.9	100.6	107.83	24.9	13.94
5	97.1	98.6	107.7	101.13	10.6	5.74
6	97.4	105.5	104.5	102.47	8.1	4.42
7	97.9	106.0	95.6	99.83	10.4	5.46

8	81.6	99.9	101.1	94.20	19.5	10.93
9	90.8	90.1	95.1	92.00	5.0	2.71
10	107.3	102.7	92.5	100.83	14.8	7.57
			Sum	977.23	141.50	76.21
			Average	97.72	14.15	7.62

So, our subgroup size is constant for each of the 10 subgroups. The subgroup average, range and standard deviation have also been calculated for use below. The overall sum and average are given for subgroup averages, subgroup ranges and subgroup standard deviations – again for use below. There may be some minor differences due to rounding.

THREE WAYS TO ESTIMATE THE STANDARD DEVIATION

We will look at three different ways to estimate the standard deviation. These impact how control limits are calculated. Control limits for the $\overline{X}$ chart are given by:

$$UCL=\overline{\overline{X}} + 3\frac{\sigma}{\sqrt{n}}$$

$$LCL=\overline{\overline{X}} - 3\frac{\sigma}{\sqrt{n}}$$

where UCL and LCL are the upper and lower control limits, n is the subgroup size, and σ is the estimated standard deviation of the individual values. Remember: the standard deviation of the subgroup averages is equal to the standard deviation of the individual values divided by square root of the subgroup size. These control limit equations may be different from the ones you normally use. They are not, as will be shown below.

The value of σ depends on the method you use to estimate it. We will look at three methods for estimating σ for subgroup data:

1. Average of the subgroup ranges
2. Average of the subgroup standard deviations
3. Pooled standard deviation

AVERAGE OF SUBGROUP RANGES

The average of the subgroup ranges is the classical way to estimate the standard deviation. The average range is simply the average of the subgroup averages when the subgroup size is constant:

$$\overline{R}=\frac{\sum R_i}{k}$$

where R_i is the range of the i^{th} subgroup and k is the number of subgroups. The standard deviation is then estimated from the following equation:

$$\sigma=\frac{\overline{R}}{d_2}$$

where d_2 is a constant that depends on subgroup size.   Table 2 shows the values of d_2 based on subgroup sizes up to 20.   From the table, you can see that d_2 for a subgroup size of 3 is 1.693.

Table 2: Constants for Control Charts

n	d_2	c_4
2	1.128	0.7979

3	1.693	0.8862
4	2.059	0.9213
5	2.326	0.9400
6	2.534	0.9515
7	2.704	0.9594
8	2.847	0.9650
9	2.970	0.9693
10	3.078	0.9727
11	3.173	0.9754
12	3.258	0.9776
13	3.336	0.9794
14	3.407	0.9810
15	3.472	0.9823

16	3.532	0.9835
17	3.588	0.9845
18	3.640	0.9854
19	3.689	0.9862
20	3.735	0.9869

For the data in Table 1, the average range and σ are given by:

$$\overline{R} = \frac{\sum R_i}{k} = \frac{141.50}{10} = 14.15$$

$$\sigma = \frac{\overline{R}}{d_2} = \frac{14.15}{1.693} = 8.36$$

Using the estimate of the standard deviation from the average range, we can now calculate the control limits:

$$UCL = \overline{\overline{X}} + 3\frac{\sigma}{\sqrt{n}} = 97.72 + 3\frac{8.36}{\sqrt{3}} = 112.20$$

$$LCL = \overline{\overline{X}} - 3\frac{\sigma}{\sqrt{n}} = 97.72 - 3\frac{8.36}{\sqrt{3}} = 83.24$$

You may not be used to calculating control limits this way for the $\overline{X}$ chart. You probably use the following equations:

$$\mathrm{UCL}=\overline{\overline{X}} + A_2\overline{R}$$

$$\mathrm{LCL}=\overline{\overline{X}} - A_2\overline{R}$$

where A_2 is a constant that depends on subgroup size. Consider just the UCL. There are two different equations for the UCL above, which must give the same result. So,

$$\mathrm{UCL}=\overline{\overline{X}} + 3\frac{\sigma}{\sqrt{n}} = \overline{\overline{X}} + A_2\overline{R}$$

This can be rearranged to the following:

$$3\frac{\sigma}{\sqrt{n}} = A_2\overline{R}$$

Substituting for $\overline{R}$ and solving for A_2 gives:

$$3\frac{\sigma}{\sqrt{n}} = A_2(d_2)\sigma$$

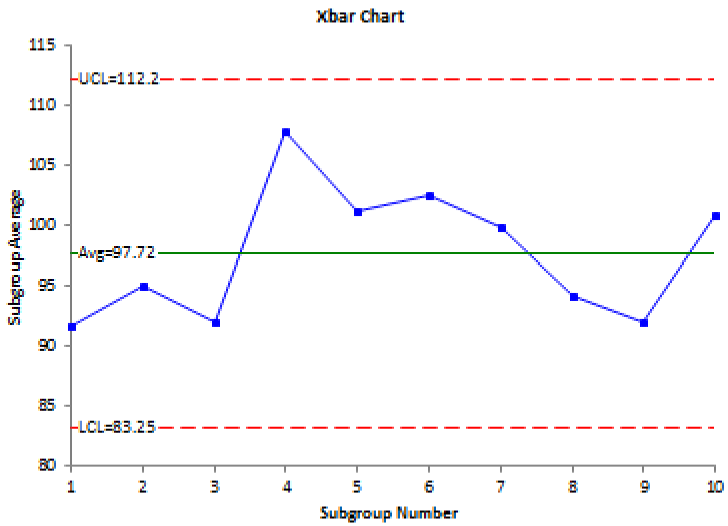
$$A_2 = \frac{3}{d_2\sqrt{n}}$$

Substituting in d_2 and n for our example gives:

$$A_2 = \frac{3}{d_2\sqrt{n}} = \frac{3}{(1.693)\sqrt{3}} = 1.023$$

This is the value of A_2 for a subgroup size of 3 that you find in the tabulated control chart constants for A_2. For a table of these values, please see our newsletter our **$\overline{X}$-R control charts (https://www.spcforexcel.com/xbar-r-charts-part-1)**. So, both methods for calculating the control limits are equivalent. The $\overline{X}$ control chart for these data is shown in Figure 1.

Figure 1: $\overline{X}$ Based on Sigma from Average Range



AVERAGE OF SUBGROUP STANDARD DEVIATIONS

The average of the subgroup standard deviations could also be used to estimate the standard deviation. When the subgroup size is constant, the average of the subgroup standard deviations is given by:

$$\bar{s}=\frac{\sum s_i}{k}$$

where s_i is the standard deviation of the i^{th} subgroup and k is the number of subgroups. The standard deviation is then estimated from the following equation:

$$\sigma=\frac{\bar{s}}{c_4}$$

where c_4 is constant that depends on subgroup size. The values of c_4 are shown in Table 2 above. For n = 3, the value of c_4 is 0.8862. For the data in Table 1, the average standard deviation and σ are given by:

$$\bar{s} = \frac{\sum s_i}{k} = \frac{76.21}{10} = 7.62$$

$$\sigma = \frac{\bar{s}}{c_4} = \frac{7.62}{0.8862} = 8.60$$

This value of σ is different than that estimated by the average range, which was 8.36. Thus, the control limits will be different also. The control limits based on the standard deviation estimated from the subgroup standard deviations are:

$$UCL = \bar{\bar{X}} + 3\frac{\sigma}{\sqrt{n}} = 97.72 + 3\frac{8.60}{\sqrt{3}} = 112.61$$

$$LCL = \bar{\bar{X}} - 3\frac{\sigma}{\sqrt{n}} = 97.72 - 3\frac{8.60}{\sqrt{3}} = 82.82$$

The differences are not large, but there are differences. This is why people wonder why the control limits can be slightly different. It is usually the way the standard deviation is estimated. The control chart with these limits will look about the same as in Figure 1 - just with the control limits a little wider in this example.

POOLED STANDARD DEVIATION

The pooled standard deviation, s_p, can also be used to estimate the standard deviation. The standard deviation, σ, is equal to the pooled standard deviation divided by c_4:

$$\sigma = \frac{s_p}{c_4}$$

where:

$$s_p = \sqrt{\frac{\sum(X_{ij} - \overline{X}_i)^2}{\sum(n_i - 1)}}$$

where X_{ij} is the j^{th} observation in the i^{th} subgroup, $\overline{X}_i$ is the average of the observations in the i^{th} subgroup, n_i is the number of observations in the i^{th} subgroup, c_4 is the constant defined above but this time it depends on the degrees of freedom (df); which is given by the sum of the n_i -1 values (the denominator under the square root sign. Thus,

$$s_p = \sqrt{\frac{\sum(X_{ij} - \overline{X}_i)^2}{\sum(n_i - 1)}} = \sqrt{\frac{1461.15}{20}} = \sqrt{73.0575} = 8.547$$

For df = 20, the value of c_4 is 0.9869. The estimated standard deviation is then given by:

$$\sigma = \frac{s_p}{c_4} = \frac{8.547}{0.9869} = 8.66$$

This third method of estimating the standard deviation gives another value for σ. The control limits will be different as well:

$$UCL = \overline{\overline{X}} + 3\frac{\sigma}{\sqrt{n}} = 97.72 + 3\frac{8.66}{\sqrt{3}} = 112.72$$

$$LCL = \overline{\overline{X}} - 3\frac{\sigma}{\sqrt{n}} = 97.72 - 3\frac{8.66}{\sqrt{3}} = 82.72$$

The control chart will look the same as Figure 1 - again with slightly different control limits.

SUMMARY

This newsletter has looked at the three different methods of estimating the standard deviation from data that are in subgroups. Each method gives a different value for the estimate standard deviation:

- σ from the average range = 8.36
- σ from the average standard deviation = 8.60
- σ from the pooled standard deviation = 8.66

This leads to different values for the control limits. Which method is correct? All three are correct. Dr. Donald Wheeler has suggested that the average range method is more robust than the pooled standard devaition - so there may be something **there (http://www.qualitydigest.com/inside/quality-insider-column/right-and-wrong-ways-computing-limits.html)**. But in the end, the important thing is the story that the control chart is telling you about your process. Minor changes in the estimate of the standard deviation will not change this in most cases.

QUICK LINKS

SPC for Excel Software (/spc-for-excel-software)

Visit our home page (/home)

SPC Training (/seminars-on-site-training)

SPC Consulting (/implementation)

Ordering Information (/ordering-information)

Thanks so much for reading our publication. We hope you find it informative and useful. Happy charting and may the data always support your position.

Sincerely,

Dr. Bill McNeese
BPI Consulting, LLC

(https://www.linkedin.com/in/billmcneese)

Connect with Us

(https://www.facebook.com/spcforexcel) **(https://www.linkedin.com/company/3569889)** **(https://www.youtube.com/channel/UCnogTilST4u1z_TQFops_LQ)**

Add new comment (/knowledge/control-chart-basics/estimated-standard-deviation-and-control-charts#comment-form)

Comments (20)

reply (/comment/reply/340/667)

Anonymous

Good morning, Bill -The newsletters are great but it would be nice to be able to print them and HTMLs do not print well. Any chance that you could either publish in PDF format or add a "printer friendly" option?Thanks.dave

Dec 10, 2012

reply (/comment/reply/340/1121)

Bill McNeese

Hi Dave,
I will take a look at doing the PDF format. You can always copy the newsletter and paste into Word in the meantime. Thanks for the comment on the newsletters.

Bill

Jan 12, 2013

reply (/comment/reply/340/6207)

Sachin Bakshi

Very informative article , had observed the difference in control its but known the reason yet..thanks for sharing the underlying reason..

Mar 11, 2016

CIES

reply (/comment/reply/340/6344)

How do you calculate the 1461.15?? From the initial data, how should I interpret the Xij???

Feb 23, 2017

Bill McNeese

reply (/comment/reply/340/6345)

There should really be a second summation sign in the numerator. Below are the calculations. Xij is the ith observation in the jth subgroup.

Subgroup	X1	X2	X3	Xbar	$(X1-Xbar)^2$	$(X2-Xbar)^2$	$(X3-Xbar)^2$	Sum
1	80.3	86.9	108.0	91.73	130.72	23.36	264.6	418.68
2	99.4	89.5	96.4	95.10	18.49	31.36	1.69	51.54
3	95.1	95.9	85.3	92.10	9	14.44	46.24	69.68
4	99.0	123.9	100.6	107.83	78.03	258.14	52.32	388.49
5	97.1	98.6	107.7	101.13	16.27	6.42	43.12	65.81
6	97.4	105.5	104.5	102.47	25.67	9.2	4.13	39
7	97.9	106.0	95.6	99.83	3.74	38.03	17.92	59.69
8	81.6	99.9	101.1	94.20	158.76	32.49	47.61	238.86
9	90.8	90.1	95.1	92.00	1.44	3.61	9.61	14.66
10	107.3	102.7	92.5	100.83	41.82	3.48	69.44	114.74
							Sum	1461.15

Feb 23, 2017

Carlos E

reply (/comment/reply/340/6347)

How to calculate C4?? Is there a ful table available??

Feb 28, 2017

Bill McNeese **reply (/comment/reply/340/6348)**

There is a table in the article. What are you looking for?

Feb 28, 2017

Viet Nguyen **reply (/comment/reply/340/6384)**

Dear Dr. Bill McNeeseI have a question about Control Limit UCL & LCL, Sigma = Rbar/d2 ,and Theory is Control Limit is 3Sigma (+/-) so, Can we just use UCL=Xbar+3*Rbar/d2 and LCL=Xbar-3*Rbar/d2 (3*Rbar/d2 = 3Sigma), Hope you can help me to have more clear about it ,I known that correct Formula shows ,Formula ,UCL = Xbar + 3 Sigma/SQRT(n) ,LCL = Xbar-3Sigma/SQRT(n) ,But it will not be 3Sigma because still devide to SQRT(n), Why is it? ,Thank you so much

Jun 18, 2017

Bill McNeese **reply (/comment/reply/340/6387)**

Please see this link for information on where control limits come from:
https://www.spcforexcel.com/knowledge/control-chart-basics/control-limits (https://www.spcforexcel.com/knowledge/control-chart-basics/control-limits)
Control limits are +/- three standard deviations of what is being plotted. So, with subgroup averages, it is +/- three standard devations of the subgroup averages. The subgroup averages standard devaition comes from sigma/sqrt(n) where n is the subgroup size and sigma is the standard devaition of the individual values (estimated by Rbar/d2).

Jun 18, 2017

Viet Nguyen **reply (/comment/reply/340/6388)**

Understood now, thanks so much Dr Bill

Jun 19, 2017

M. Senthil Kumar **reply (/comment/reply/340/6416)**

Dear Sir, Greetings, I have a doubt, is it calculate the control limits for population method.I have 125 nos samples. i will not seperate this for sub groups. How to calculate the UCL & LCL. Is it possible.

Aug 21, 2017

Bill McNeese **reply (/comment/reply/340/6417)**

Use an individuals control chart if you are not going to subgroup.
https://www.spcforexcel.com/knowledge/variable-control-charts/individual... (https://www.spcforexcel.com/knowledge/variable-control-charts/individuals-control-charts)

Aug 21, 2017

Marcus **reply (/comment/reply/340/6448)**

Hi there, i was wondering if u could help me in the following problemI have been given 50 numbers in an excel sheet.. each number represents an average of 5 subsample observations. I have not been given the individual 5 numbers, just their average. and therefore am unable to calculate the range.Ive also been told the 50 average of subsamples follow a normal distribution following N(2,3), sample size n = 5 , the standard deviation = 3, and to follow a 2-sigma deviation.The average of the 50 averages i calculated =1.94How would i go about calculating the LCL and UCL, given the information i have.? many thanks in advance mate

Oct 14, 2017

Bill McNeese **reply (/comment/reply/340/6449)**

All the information you need in the article. You have the average. Are you saying the distribution of subgroup averages has a standard deviaiton of 3? Or is the individual values? What do you mean by a 2 sigma deviation?

Oct 15, 2017

Ratul **reply (/comment/reply/340/6474)**

In X-R chart the value of A2R is 1.8 calculate the value of sigma i.e standard deviation?

Nov 29, 2017

Bill McNeese

reply (/comment/reply/340/6475)

Not clear to me what you are asking.

Nov 29, 2017

Eagles

reply (/comment/reply/340/6480)

Why don't we estimate standrad deviation by using standard deviation of all samples?

Dec 12, 2017

Bill McNeese

reply (/comment/reply/340/6481)

If you use the calculated standard deviation of all the range, it will inflated when the data are not in control. The purpose of a control chart is to determine if the data are homogeneous. Calculating the standard deviation assumes that the data are homogenous.

Dec 12, 2017

siddharth

reply (/comment/reply/340/6650)

For PpK calculation, overall standard deviation is used. Can you please explain how to calculate for above datas.

May 30, 2018

Bill McNeese

reply (/comment/reply/340/6651)

The calculated standard deviation is the same as the STDEV function in excel. It is the square root of the sum of the (Xi-Xavg)^2 divided by n - 1.

May 30, 2018

Leave a comment

Your name

Comment *

Paragraph

Styles

Path: p

Disable rich-text

Text format Filtered HTML ▼

More information about text formats (/filter/tips)

- Web page addresses and e-mail addresses turn into links automatically.
- Allowed HTML tags: <a> <em> <strong> <cite> <code>   <dl> <dt> <dd> <h1> <h2> <h3> <h4> <h5> <h6> <img> <hr> <div> <span> <strike> <b> <i> <u> <table> <tbody> <tr> <td> <th>
- Lines and paragraphs break automatically.

CAPTCHA

This question is for testing whether you are a human visitor and to prevent automated spam submissions.

I'm not a robot

reCAPTCHA
Privacy - Terms

Preview

Control Charts - Basics

- **<< Return to Categories (/spc-for-excel-publications-category)**
- **Interpreting Control Charts (/knowledge/control-charts-basics/interpreting-control-charts)**
- **Control Strategies (/knowledge/control-chart-basics/control-strategies)**
- **Use of Control Charts (/knowledge/control-chart-basics/use-control-charts)**
- **The Impact of Statistical Control (/knowledge/control-chart-basics/impact-statistical-control)**
- **Selecting the Right Control Chart (/knowledge/control-chart-basics/selecting-right-control-chart)**
- **Control Limits - Where Do They Come From? (/knowledge/control-chart-basics/control-limits)**
- **The Purpose of Control Charts (/knowledge/control-chart-basics/purpose-control-charts)**
- **When to Calculate, Lock, and Recalculate Control Limits (/knowledge/control-chart-basics/when-calculate-lock-and-recalculate-control-limits)**
- **My Process is Out of Control! Now What Do I Do? (/knowledge/control-chart-basics/my-process-out-control-now-what-do-i-do)**
- **The Estimated Standard Deviation and Control Charts (/knowledge/control-chart-basics/estimated-standard-deviation-and-control-charts)**
- **Applying the Out of Control Tests (/knowledge/control-chart-basics/applying-out-of-control-tests)**
- **Control Chart Rules and Interpretation (/knowledge/control-chart-basics/control-chart-rules-interpretation)**
- **Control Charts and the Central Limit Theorem (/knowledge/control-chart-basics/control-charts-and-central-limit-theorem)**
- **Control Charts, ANOVA, and Variation (/knowledge/control-chart-basics/control-charts-ANOVA-variation)**

- **How Much Data Do I Need to Calculate Control Limits? (/knowledge/control-chart-basics/how-much-data-do-i-need-calculate-control-limits)**
- **How to Mess Up Using Control Charts (/knowledge/how-to-mess-up-control-charts)**
- **The Difficulty of Setting Baseline Data for Control Charts (/knowledge/control-chart-basics/difficulty-baseline-control-charts)**
- **Three Sigma Limits and Control Charts (/knowledge/control-chart-basics/three-sigma-limits-control-charts)**

Catalog (/catalog)

- **Software (3) (/catalog/software)**
- **Complete Teaching Guides to SPC (11) (/catalog/complete-teaching-guides-spc)**

Connect with Us

(https://www.facebook.com/spcforexcel) **(https://www.linkedin.com/company/3569889)** **(https://www.youtube.com/channel/UCnogTilST4u1z_TQFops_LQ)**

SPC Knowledge Base Sign-up

Click here (/newsletter-signup) to sign up for our FREE monthly publication, featuring SPC and other statistical topics, case studies and more!

SPC Around the World

SPC for Excel is used in over 60 countries internationally. **Click here for a list of those countries. (https://www.spcforexcel.com/spc-for-excel-used-internationally)**