POLITECNICO
MILANO 1863

DEPARTMENT OF MECHANICAL
ENGINEERING

# EXERCISE CLASS 2
## SPC for iid data

Name Surname

DIPARTIMENTO DI ECCELLENZA
MIUR 2018-2022

# Exercise 1

**The gears in wind turbine gearboxes are essential for converting rotational energy from the turbine blades into electrical energy. Data in "gears_phase1.csv" represent measurements of the gear diameters. These diameters are sequentially sampled in groups of n = 5 from the manufacturing process and must meet a tolerance of 24 ± 1 mm to ensure reliable performance and durability.**

**1.** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

**2.** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length (ARL0) of 1000 (assuming that the normal approximation applies for both of them).

**3. Determine the operating characteristic curve (OC) for the X-bar chart** (by using K=3 and expressing the shift of the mean in standard deviation units)

**4.** Determine the corresponding ARL curve.

**5.** Estimate the standard deviation through the statistic R (consider original Phase I data).

**6. Design the confidence interval on the process mean that corresponds to the control limits computed in point 1** (consider original Phase I data).

# Exercise 1

**The gears in wind turbine gearboxes are essential for converting rotational energy from the turbine blades into electrical energy. Data in "gears_phase1.csv" represent measurements of the gear diameters. These diameters are sequentially sampled in groups of n = 5 from the manufacturing process and must meet a tolerance of 24 ± 1 mm to ensure reliable performance and durability.**

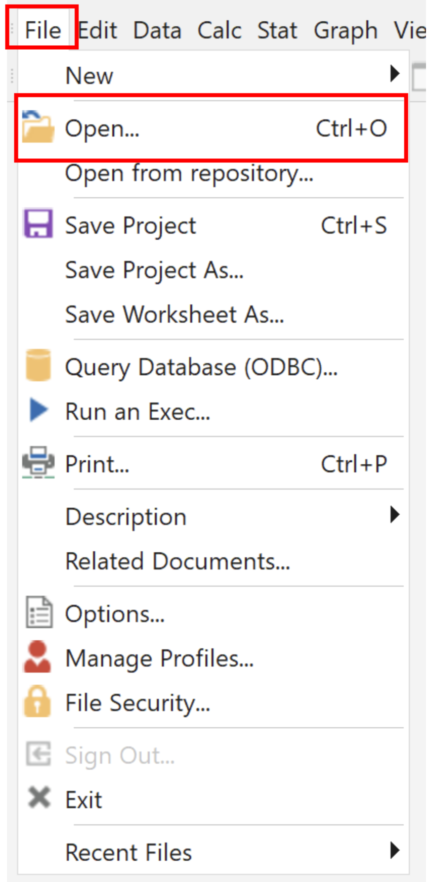
**7.** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

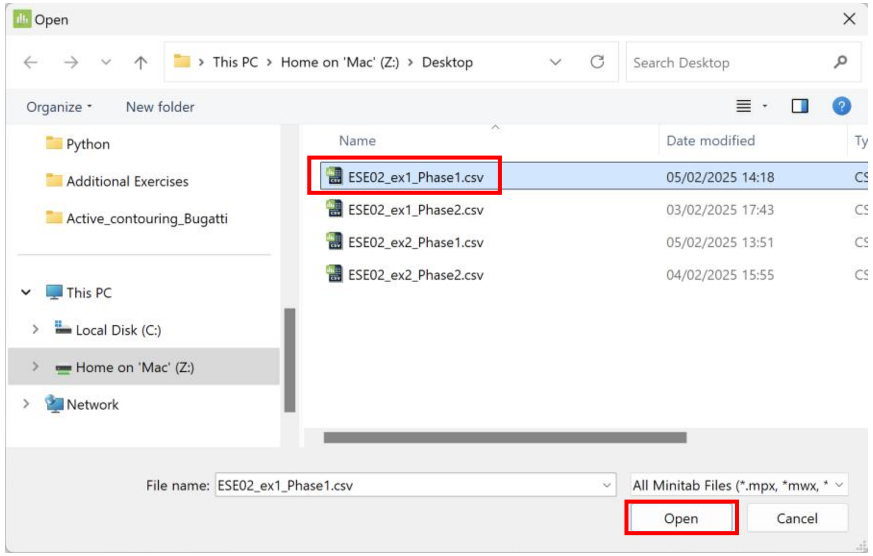
**8.** Knowing that the gear diameter is distributed as a normal distribution with mean 24 mm and standard deviation 0.26 mm, design an Xbar and S chart. For any out-of-control points detected, assume that an assignable cause is found. Check if Phase II data is in control.

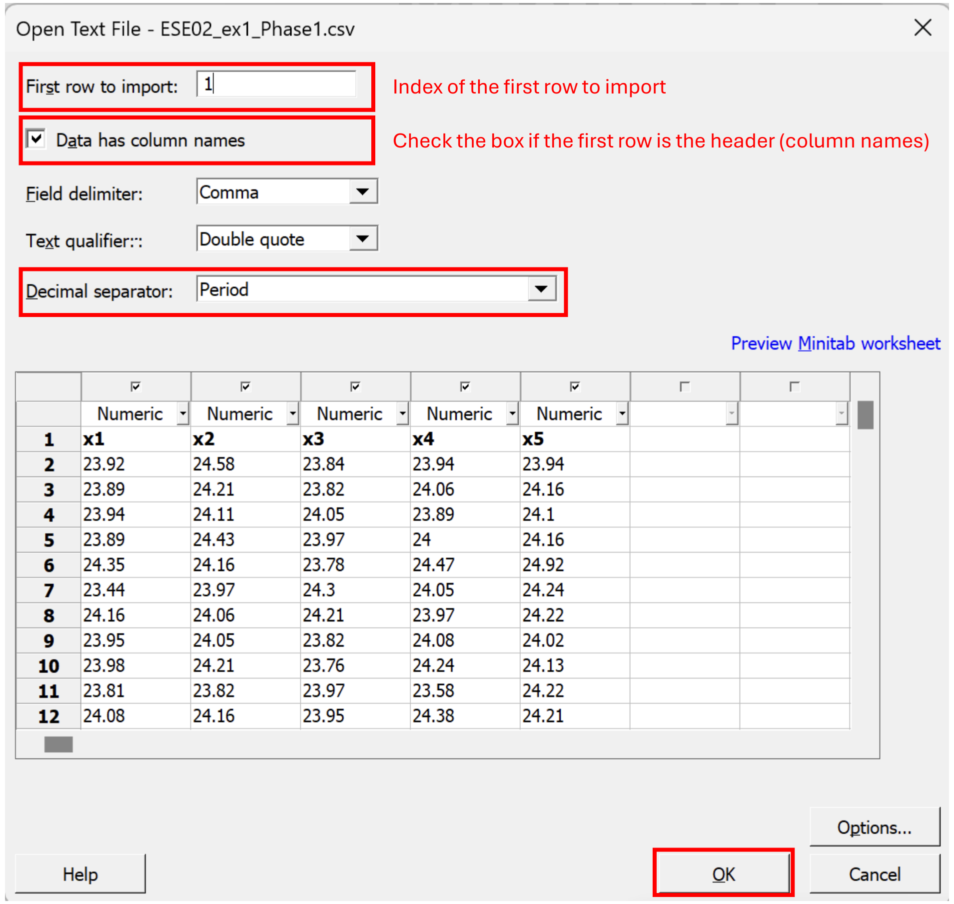
**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 1) Import the data in Minitab:

*File → Open → Locate “ESE02_ex1_PhaseI.csv” and Open*



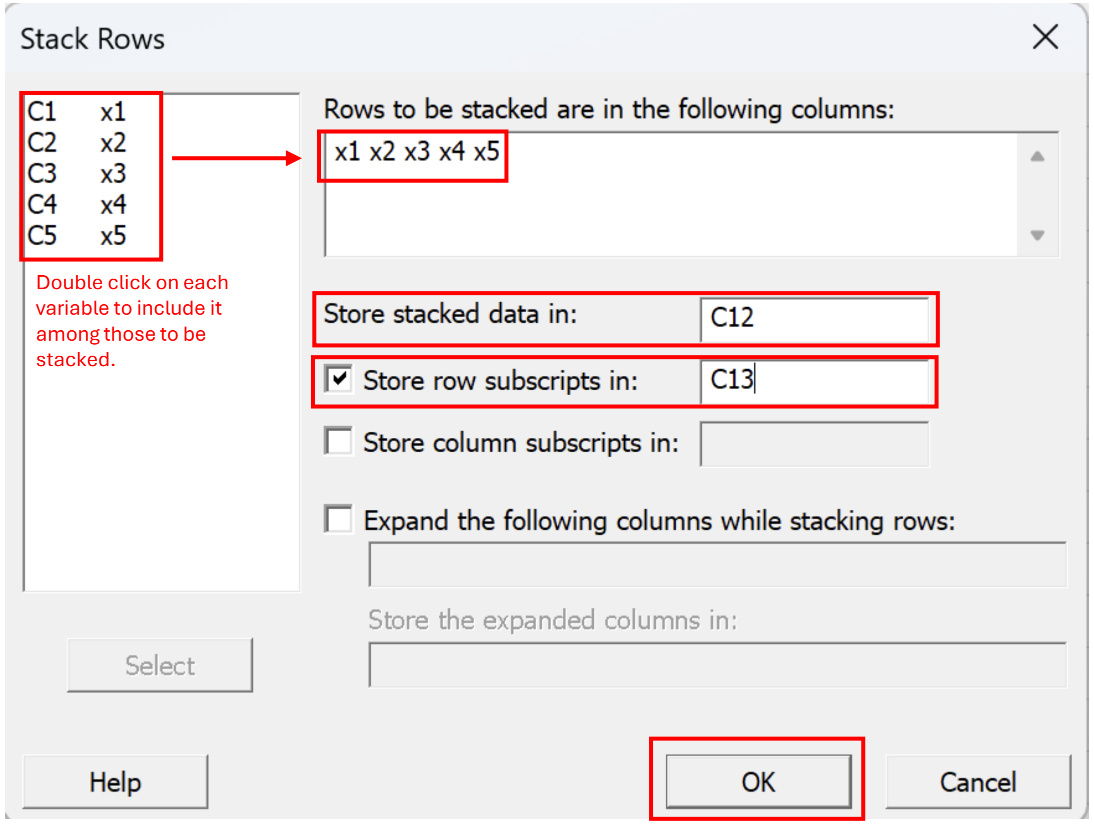




Alternative way: Open *“ESE02_ex1_PhaseI.csv”* in Excel, then copy/paste cells into a Minitab worksheet.

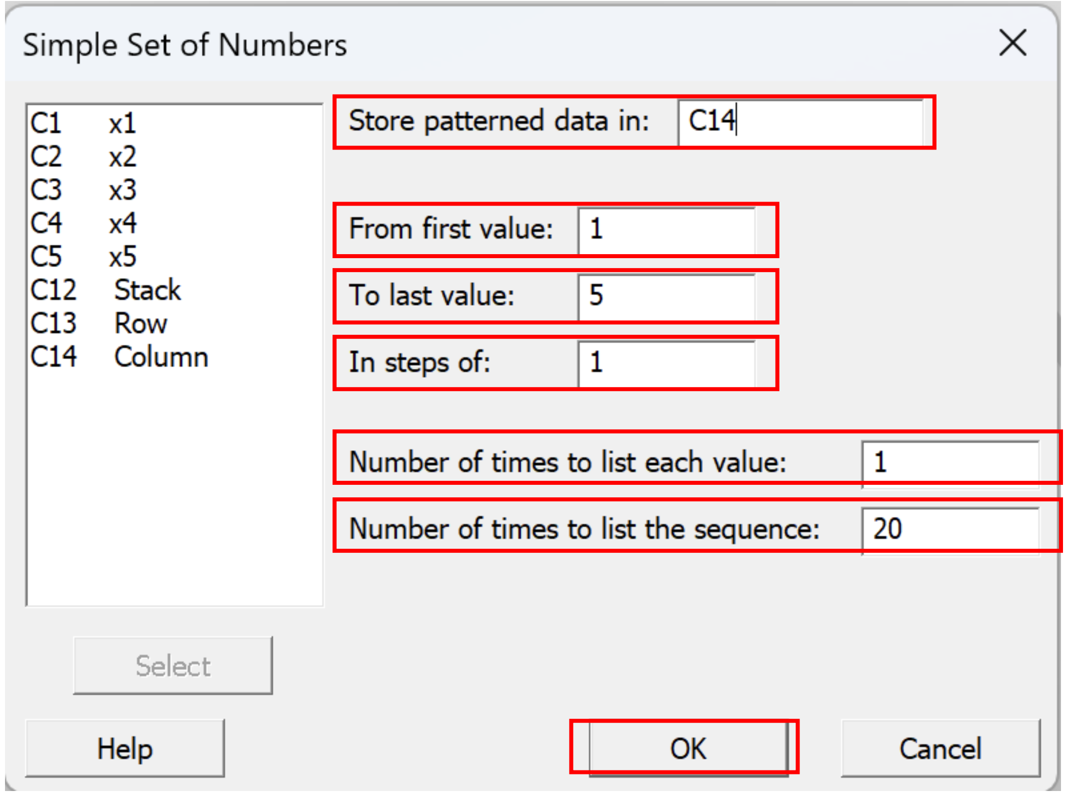
**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 2) Inspect the data by plotting the individual datapoints:

First, let’s stack the data into a single column for further analysis:
*Data → Stack → Rows*



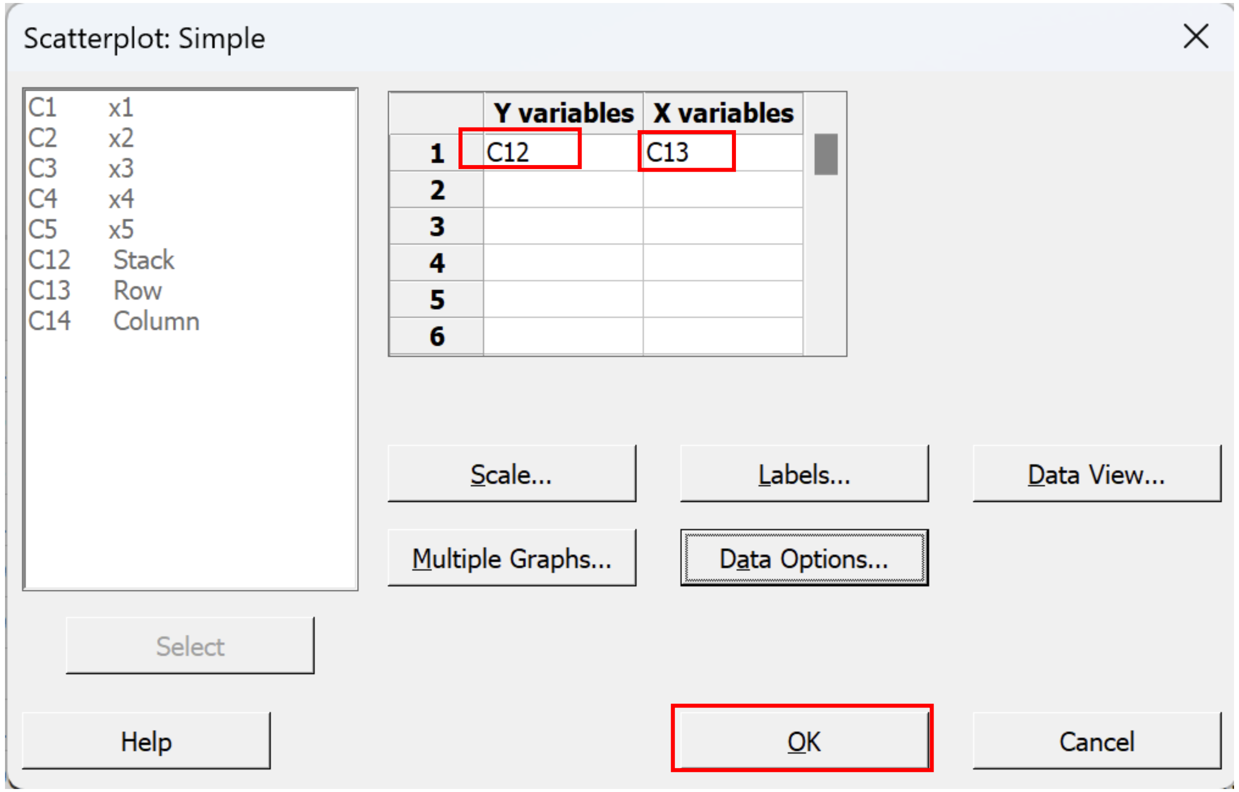

Let’s also create a column for column subscripts:
*Calc → Make Patterned Data → Simple Set of Numbers*



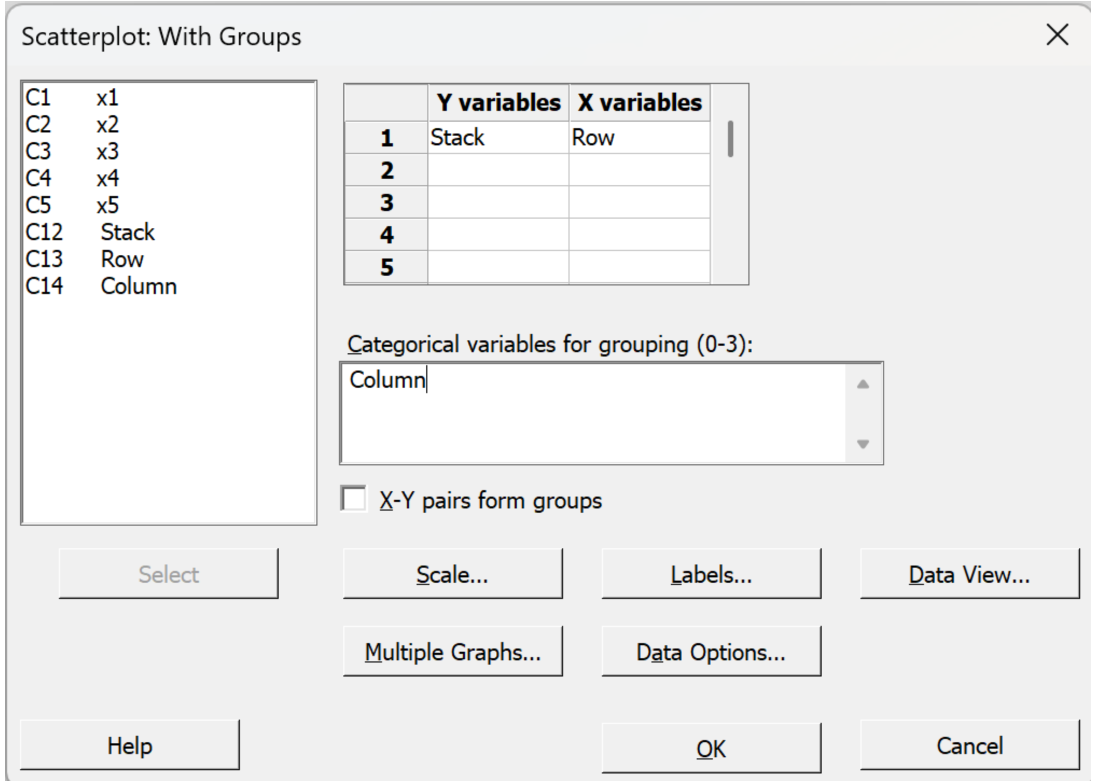
**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 2) Inspect the data by plotting the individual datapoints:

To create a scatterplot:
*Graph → Scatterplot → With groups*

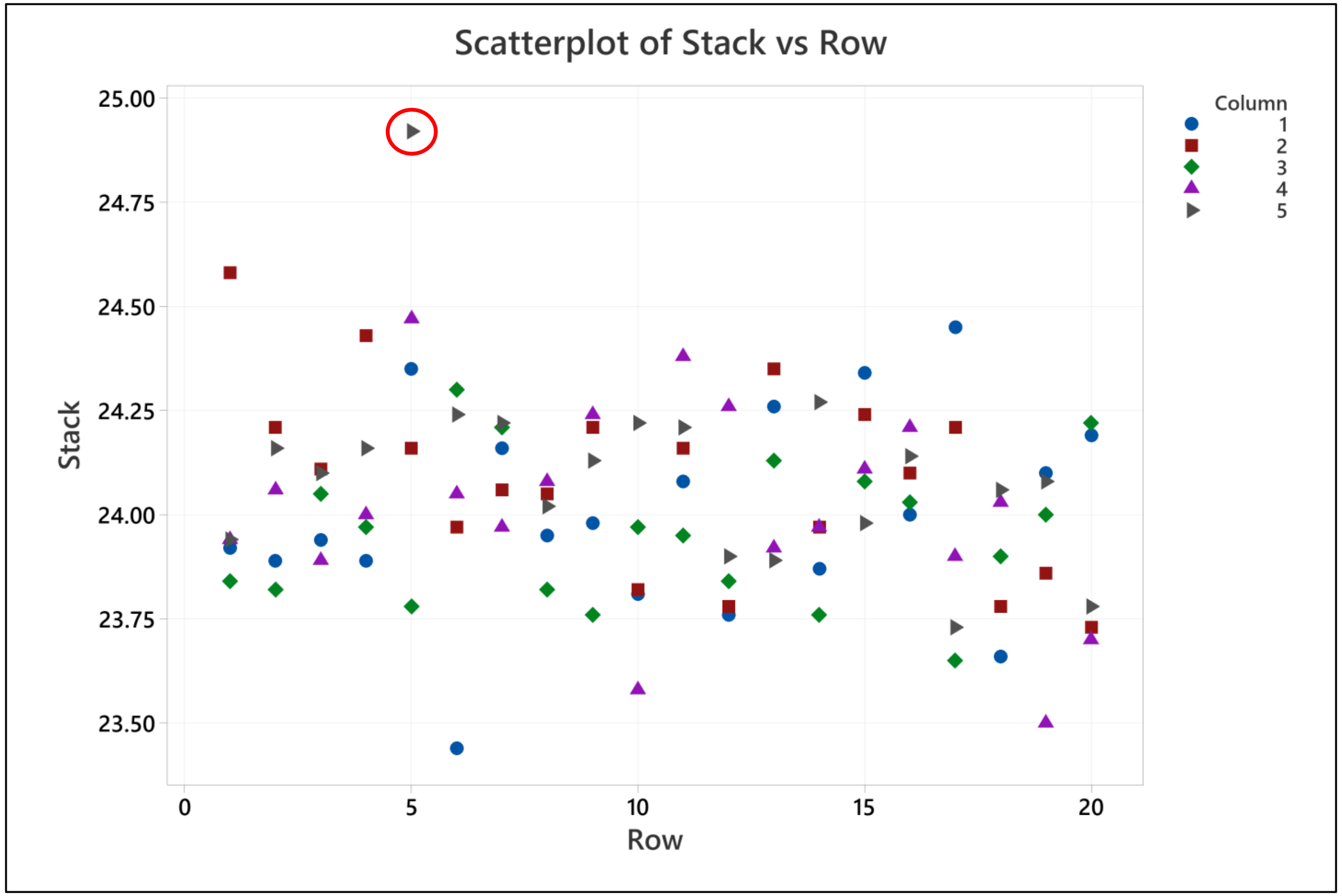



Let's edit the scatterplot to add group attributes:
*Y = C12 (Stack), X = C13 (Row), grouping: C14 (Colum*



**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 2) Inspect the data by plotting the individual datapoints:

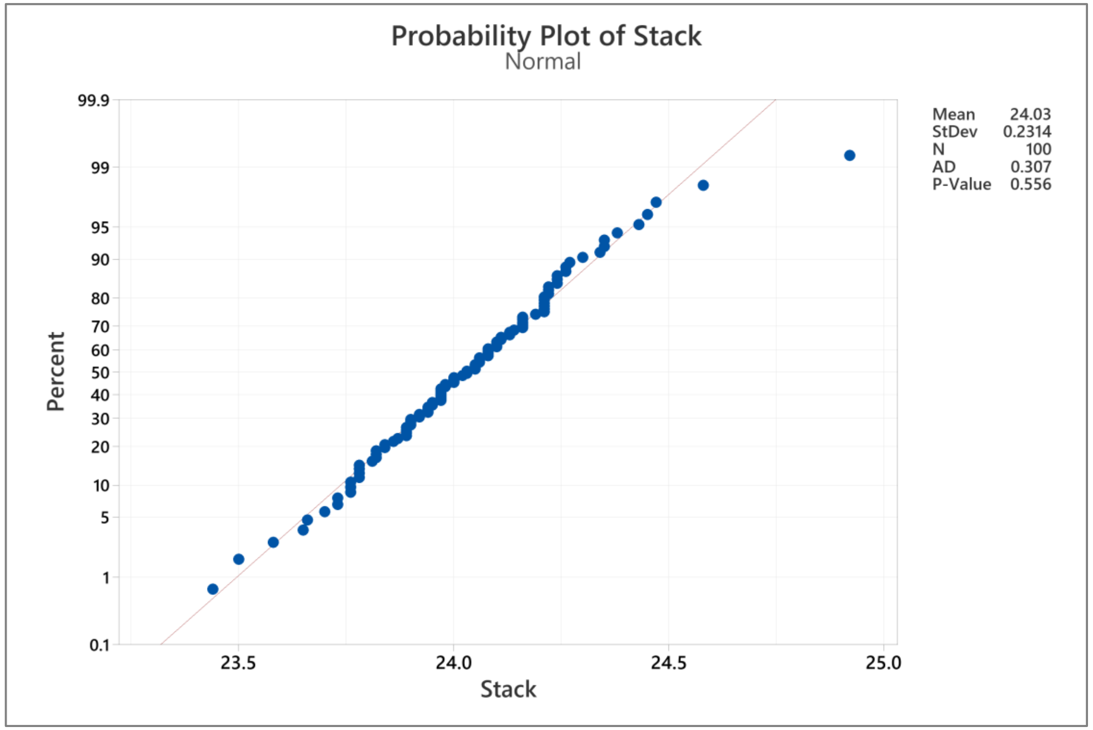



Doesn't look like strange patterns or outliers are present except for one extreme value.

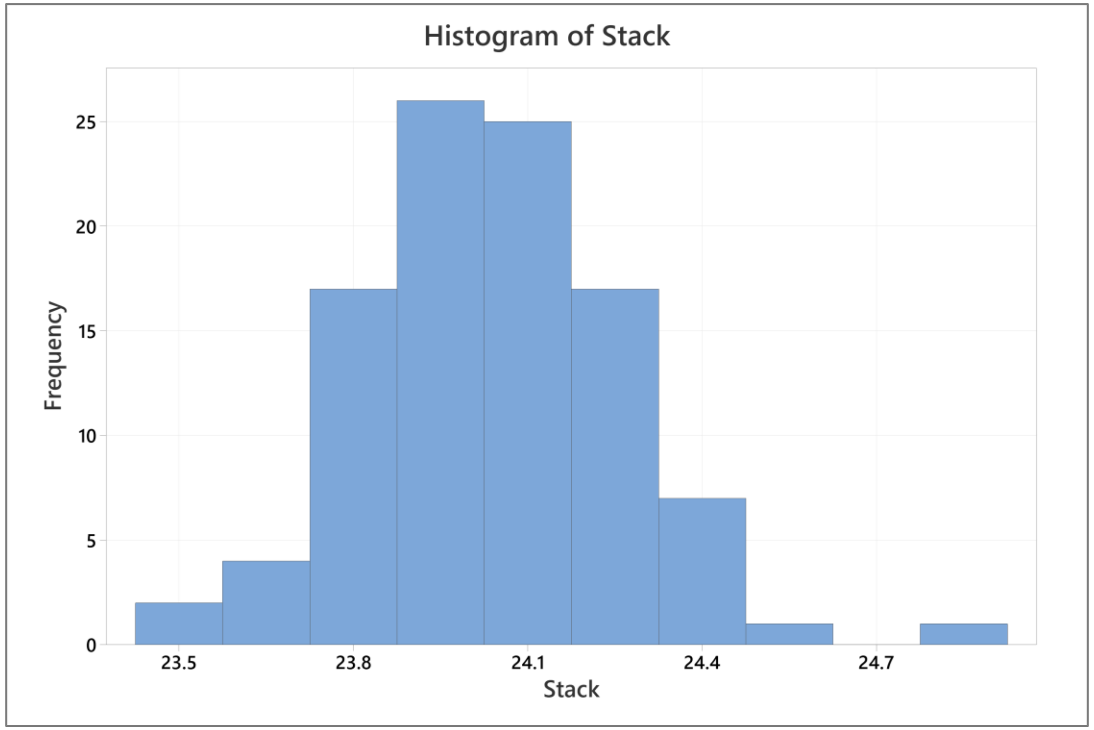
**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 3) Check the normality assumption:

*Stat -> Basic Statistics -> Normality Test*
Variable: *C12*, Test: *Anderson-Darling;*




*Graph -> Histogram -> Simple*
Graph variables: *C12;*
*Double click on graph -> Double Click on bars*
*In "Binning" section, set the Number of Intervals to 10*




With a significance level of 5% (0.05) we fail to reject the null hypothesis of the Anderson-Darling test. Note that one extreme value is responsible for borderline normality.
We might also check randomness, but we must know within-sample order!

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's design the Xbar-R chart (Phase I):

Since there is no constraint on the choice of Type I error $\alpha$, we can set K = 3 ($\alpha$ = 0.0027). Therefore:

**Xbar Chart** (unknown parameters, K=3):

$$UCL = \overline{\overline{X}} + A_2(n)\overline{R}$$

$$CL = \overline{\overline{X}}$$

$$LCL = \overline{\overline{X}} - A_2(n)\overline{R}$$

**R Chart** (unknown parameters, K=3):

$$UCL = D_4(n)\overline{R}$$

$$CL = \overline{R}$$

$$LCL = D_3(n)\overline{R}$$

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's design the Xbar-R chart (Phase I):

You can refer to statistical tables to retrieve control chart constants :

**Factors for Constructing Variables Control Charts**

| Observations in Sample, $n$ | Chart for Averages: Factors for Control Limits | | | Chart for Standard Deviations: Factors for Center Line | | Factors for Control Limits | | | | Chart for Ranges: Factors for Center Line | | Factors for Control Limits | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $A$ | $A_2$ | $A_3$ | $c_4$ | $1/c_4$ | $B_3$ | $B_4$ | $B_5$ | $B_6$ | $d_2$ | $1/d_2$ | $d_3$ | $D_1$ | $D_2$ | $D_3$ | $D_4$ |
| 2 | 2.121 | 1.880 | 2.659 | 0.7979 | 1.2533 | 0 | 3.267 | 0 | 2.606 | 1.128 | 0.8865 | 0.853 | 0 | 3.686 | 0 | 3.267 |
| 3 | 1.732 | 1.023 | 1.954 | 0.8862 | 1.1284 | 0 | 2.568 | 0 | 2.276 | 1.693 | 0.5907 | 0.888 | 0 | 4.358 | 0 | 2.574 |
| 4 | 1.500 | 0.729 | 1.628 | 0.9213 | 1.0854 | 0 | 2.266 | 0 | 2.088 | 2.059 | 0.4857 | 0.880 | 0 | 4.698 | 0 | 2.282 |
| 5 | 1.342 | 0.577 | 1.427 | 0.9400 | 1.0638 | 0 | 2.089 | 0 | 1.964 | 2.326 | 0.4299 | 0.864 | 0 | 4.918 | 0 | 2.114 |
| 6 | 1.225 | 0.483 | 1.287 | 0.9515 | 1.0510 | 0.030 | 1.970 | 0.029 | 1.874 | 2.534 | 0.3946 | 0.848 | 0 | 5.078 | 0 | 2.004 |
| 7 | 1.134 | 0.419 | 1.182 | 0.9594 | 1.0423 | 0.118 | 1.882 | 0.113 | 1.806 | 2.704 | 0.3698 | 0.833 | 0.204 | 5.204 | 0.076 | 1.924 |
| 8 | 1.061 | 0.373 | 1.099 | 0.9650 | 1.0363 | 0.185 | 1.815 | 0.179 | 1.751 | 2.847 | 0.3512 | 0.820 | 0.388 | 5.306 | 0.136 | 1.864 |
| 9 | 1.000 | 0.337 | 1.032 | 0.9693 | 1.0317 | 0.239 | 1.761 | 0.232 | 1.707 | 2.970 | 0.3367 | 0.808 | 0.547 | 5.393 | 0.184 | 1.816 |
| 10 | 0.949 | 0.308 | 0.975 | 0.9727 | 1.0281 | 0.284 | 1.716 | 0.276 | 1.669 | 3.078 | 0.3249 | 0.797 | 0.687 | 5.469 | 0.223 | 1.777 |
| 11 | 0.905 | 0.285 | 0.927 | 0.9754 | 1.0252 | 0.321 | 1.679 | 0.313 | 1.637 | 3.173 | 0.3152 | 0.787 | 0.811 | 5.535 | 0.256 | 1.744 |
| 12 | 0.866 | 0.266 | 0.886 | 0.9776 | 1.0229 | 0.354 | 1.646 | 0.346 | 1.610 | 3.258 | 0.3069 | 0.778 | 0.922 | 5.594 | 0.283 | 1.717 |
| 13 | 0.832 | 0.249 | 0.850 | 0.9794 | 1.0210 | 0.382 | 1.618 | 0.374 | 1.585 | 3.336 | 0.2998 | 0.770 | 1.025 | 5.647 | 0.307 | 1.693 |
| 14 | 0.802 | 0.235 | 0.817 | 0.9810 | 1.0194 | 0.406 | 1.594 | 0.399 | 1.563 | 3.407 | 0.2935 | 0.763 | 1.118 | 5.696 | 0.328 | 1.672 |
| 15 | 0.775 | 0.223 | 0.789 | 0.9823 | 1.0180 | 0.428 | 1.572 | 0.421 | 1.544 | 3.472 | 0.2880 | 0.756 | 1.203 | 5.741 | 0.347 | 1.653 |
| 16 | 0.750 | 0.212 | 0.763 | 0.9835 | 1.0168 | 0.448 | 1.552 | 0.440 | 1.526 | 3.532 | 0.2831 | 0.750 | 1.282 | 5.782 | 0.363 | 1.637 |
| 17 | 0.728 | 0.203 | 0.739 | 0.9845 | 1.0157 | 0.466 | 1.534 | 0.458 | 1.511 | 3.588 | 0.2787 | 0.744 | 1.356 | 5.820 | 0.378 | 1.622 |
| 18 | 0.707 | 0.194 | 0.718 | 0.9854 | 1.0148 | 0.482 | 1.518 | 0.475 | 1.496 | 3.640 | 0.2747 | 0.739 | 1.424 | 5.856 | 0.391 | 1.608 |
| 19 | 0.688 | 0.187 | 0.698 | 0.9862 | 1.0140 | 0.497 | 1.503 | 0.490 | 1.483 | 3.689 | 0.2711 | 0.734 | 1.487 | 5.891 | 0.403 | 1.597 |
| 20 | 0.671 | 0.180 | 0.680 | 0.9869 | 1.0133 | 0.510 | 1.490 | 0.504 | 1.470 | 3.735 | 0.2677 | 0.729 | 1.549 | 5.921 | 0.415 | 1.585 |
| 21 | 0.655 | 0.173 | 0.663 | 0.9876 | 1.0126 | 0.523 | 1.477 | 0.516 | 1.459 | 3.778 | 0.2647 | 0.724 | 1.605 | 5.951 | 0.425 | 1.575 |
| 22 | 0.640 | 0.167 | 0.647 | 0.9882 | 1.0119 | 0.534 | 1.466 | 0.528 | 1.448 | 3.819 | 0.2618 | 0.720 | 1.659 | 5.979 | 0.434 | 1.566 |
| 23 | 0.626 | 0.162 | 0.633 | 0.9887 | 1.0114 | 0.545 | 1.455 | 0.539 | 1.438 | 3.858 | 0.2592 | 0.716 | 1.710 | 6.006 | 0.443 | 1.557 |
| 24 | 0.612 | 0.157 | 0.619 | 0.9892 | 1.0109 | 0.555 | 1.445 | 0.549 | 1.429 | 3.895 | 0.2567 | 0.712 | 1.759 | 6.031 | 0.451 | 1.548 |
| 25 | 0.600 | 0.153 | 0.606 | 0.9896 | 1.0105 | 0.565 | 1.435 | 0.559 | 1.420 | 3.931 | 0.2544 | 0.708 | 1.806 | 6.056 | 0.459 | 1.541 |

For $n > 25$.

$$A = \frac{3}{\sqrt{n}} \quad A_3 = \frac{3}{c_4\sqrt{n}} \quad c_4 \cong \frac{4(n-1)}{4n-3}$$

$$B_3 = 1 - \frac{3}{c_4\sqrt{2(n-1)}} \quad B_4 = 1 + \frac{3}{c_4\sqrt{2(n-1)}}$$

$$B_5 = c_4 - \frac{3}{\sqrt{2(n-1)}} \quad B_6 = c_4 + \frac{3}{\sqrt{2(n-1)}}$$

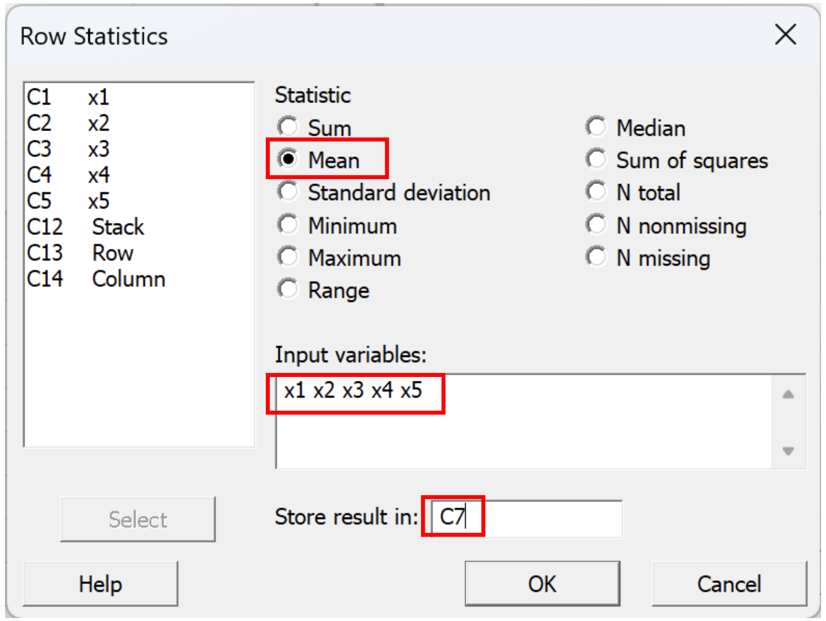
$$n = 5$$
$$A_2(n) = 0.577$$
$$D_3(n) = 0$$
$$D_4(n) = 2.114$$

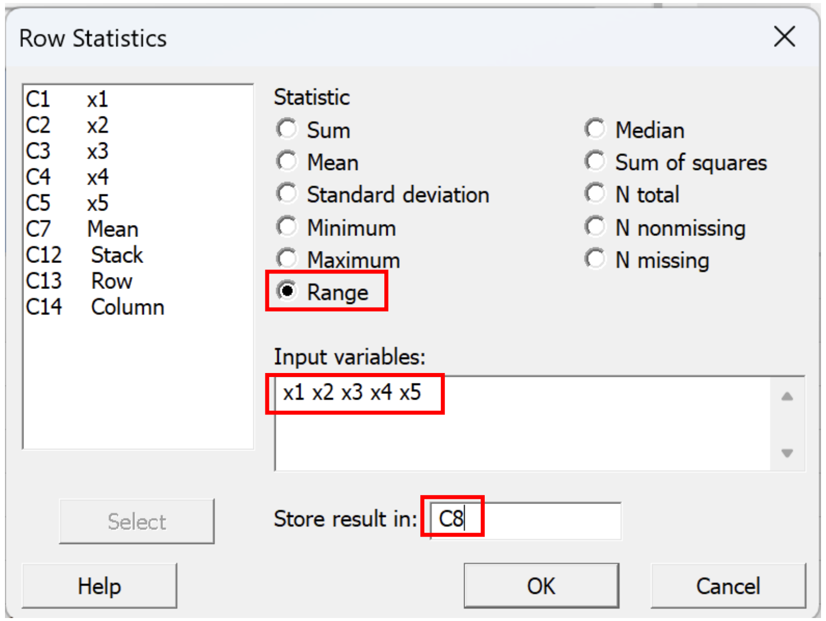
**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's design the Xbar-R chart (Phase I):

Let's compute the mean and range statistic for each sample:

*Calc → Row Statistics*



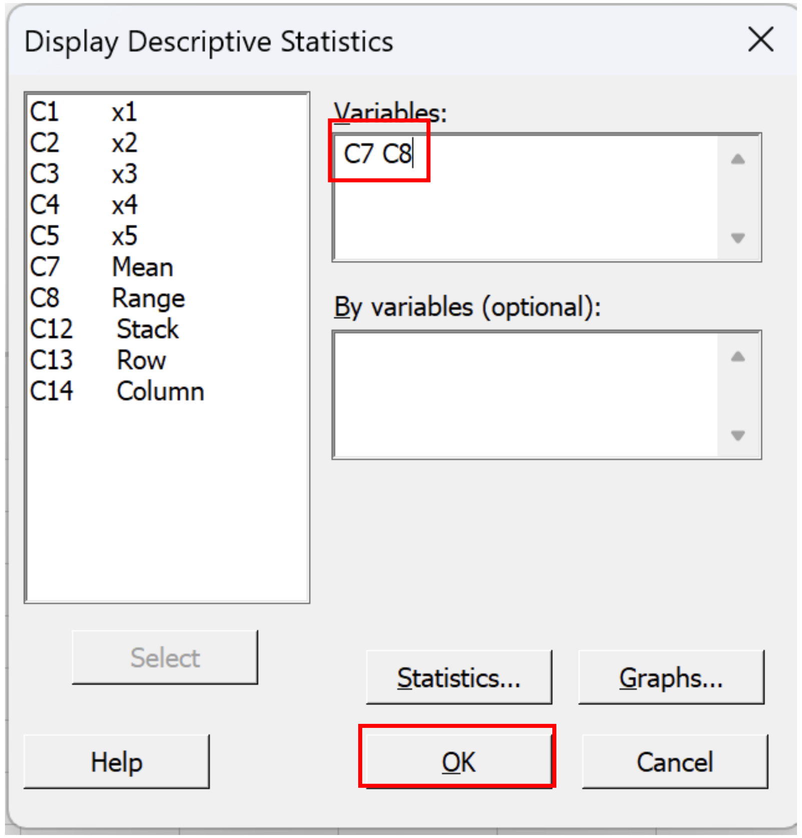

| C7 | C8 |
|---|---|
| Mean | Range |
| 24.044 | 0.74 |
| 24.028 | 0.39 |
| 24.018 | 0.22 |
| 24.090 | 0.54 |
| 24.336 | 1.14 |
| 24.000 | 0.86 |
| 24.124 | 0.25 |
| 23.984 | 0.26 |
| 24.064 | 0.48 |
| 23.880 | 0.64 |
| 24.156 | 0.43 |
| 23.908 | 0.50 |
| 24.110 | 0.46 |
| 23.968 | 0.51 |
| 24.150 | 0.36 |
| 24.096 | 0.21 |
| 23.988 | 0.80 |
| 23.886 | 0.40 |
| 23.908 | 0.60 |
| 23.924 | 0.52 |

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's design the Xbar-R chart (Phase I):

Now we can compute the grand mean (mean of sample means) and the mean of ranges:

*Stat → Basic Statistics → Display Descriptive Statistics*




**Statistics**

| Variable | N | N* | Mean | SE Mean | StDev | Minimum | Q1 | Median | Q3 | Maximum |
|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 20 | 0 | 24.0331 | 0.0251979 | 0.112689 | 23.88 | 23.935 | 24.023 | 24.1065 | 24.336 |
| Range | 20 | 0 | 0.5155 | 0.0524378 | 0.234509 | 0.21 | 0.3675 | 0.49 | 0.63 | 1.14 |

$$\overline{\overline{X}} = 24.033$$

$$\overline{R} = 0.516$$

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, **assume that an assignable cause is found**. Check if data contained in "gears_phase2.csv" is in control.

**4) Let's design the Xbar-R chart (Phase I):**

Since there is no constraint on the choice of Type I error $\alpha$, we can set K = 3 ($\alpha$ = 0.0027). Therefore:

**Xbar Chart** (unknown parameters, K=3):

$$UCL = \bar{\bar{X}} + A_2(n)\bar{R} = 24.033 + 0.577 \times 0.516 = 24.331$$

$$CL = \bar{\bar{X}} = 24.033$$

$$LCL = \bar{\bar{X}} - A_2(n)\bar{R} = 24.033 - 0.577 \times 0.516 = 23.736$$

**R Chart** (unknown parameters, K=3):

$$UCL = D_4(n)\bar{R} = 2.114 \times 0.516 = 1.091$$

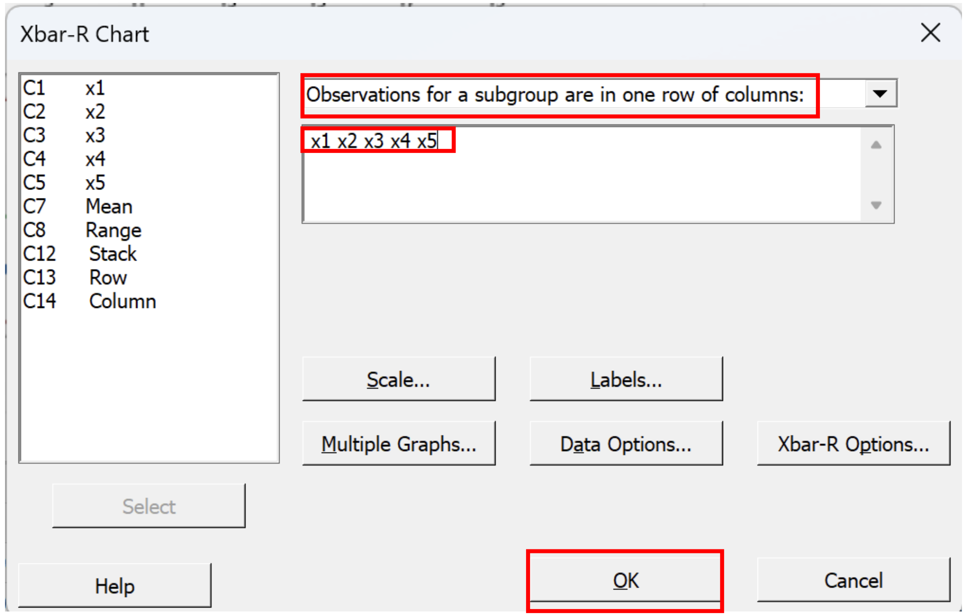
$$CL = \bar{R} = 0.516$$

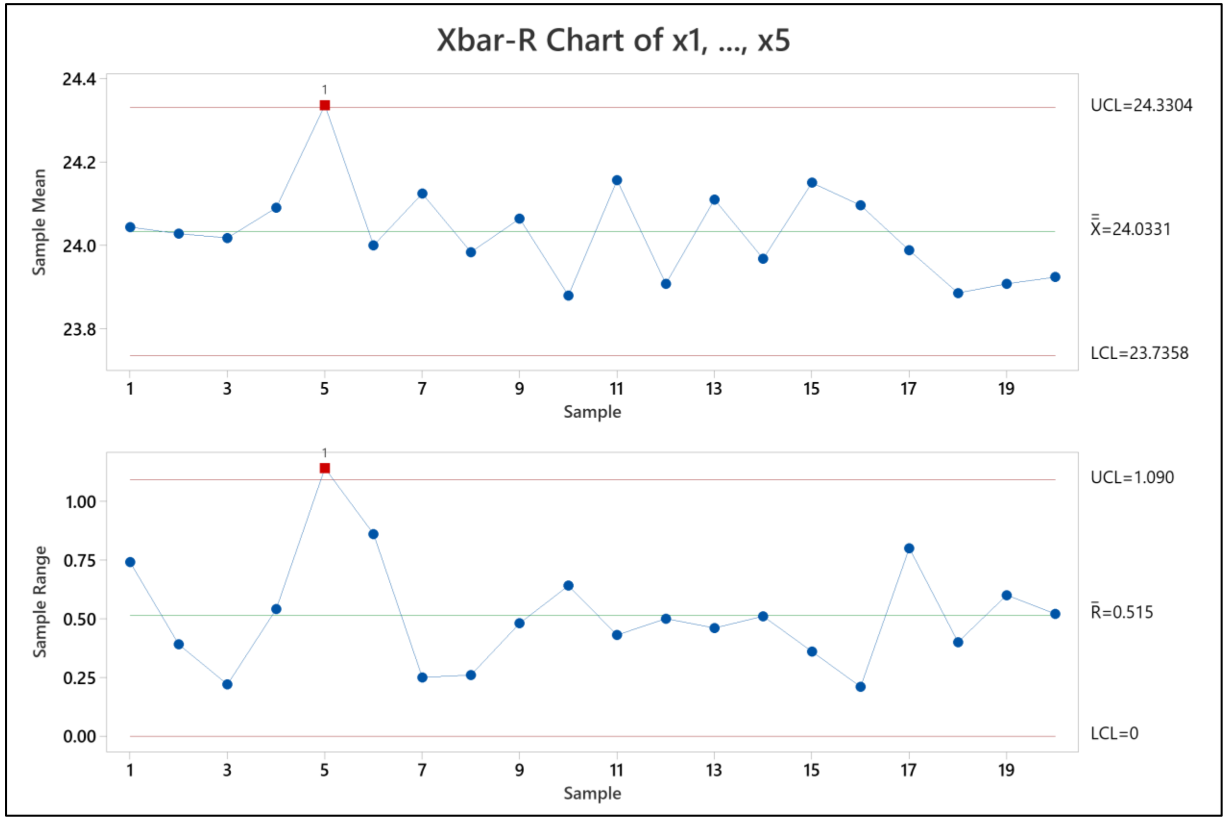
$$LCL = D_3(n)\bar{R} = 0 \times 0.516 = 0$$

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's design the Xbar-R chart (Phase I):

Let's build the Xbar-R chart in Minitab:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R*




One observation is signalled as out-of-control in both the X-bar and R charts. According to the text, when an out-of-control observation is detected, we must assume that an assignable cause has been identified. Consequently, we cannot attribute this alarm to a random false positive (i.e. false alarm). We must remove the observation and recalculate the control limits.

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

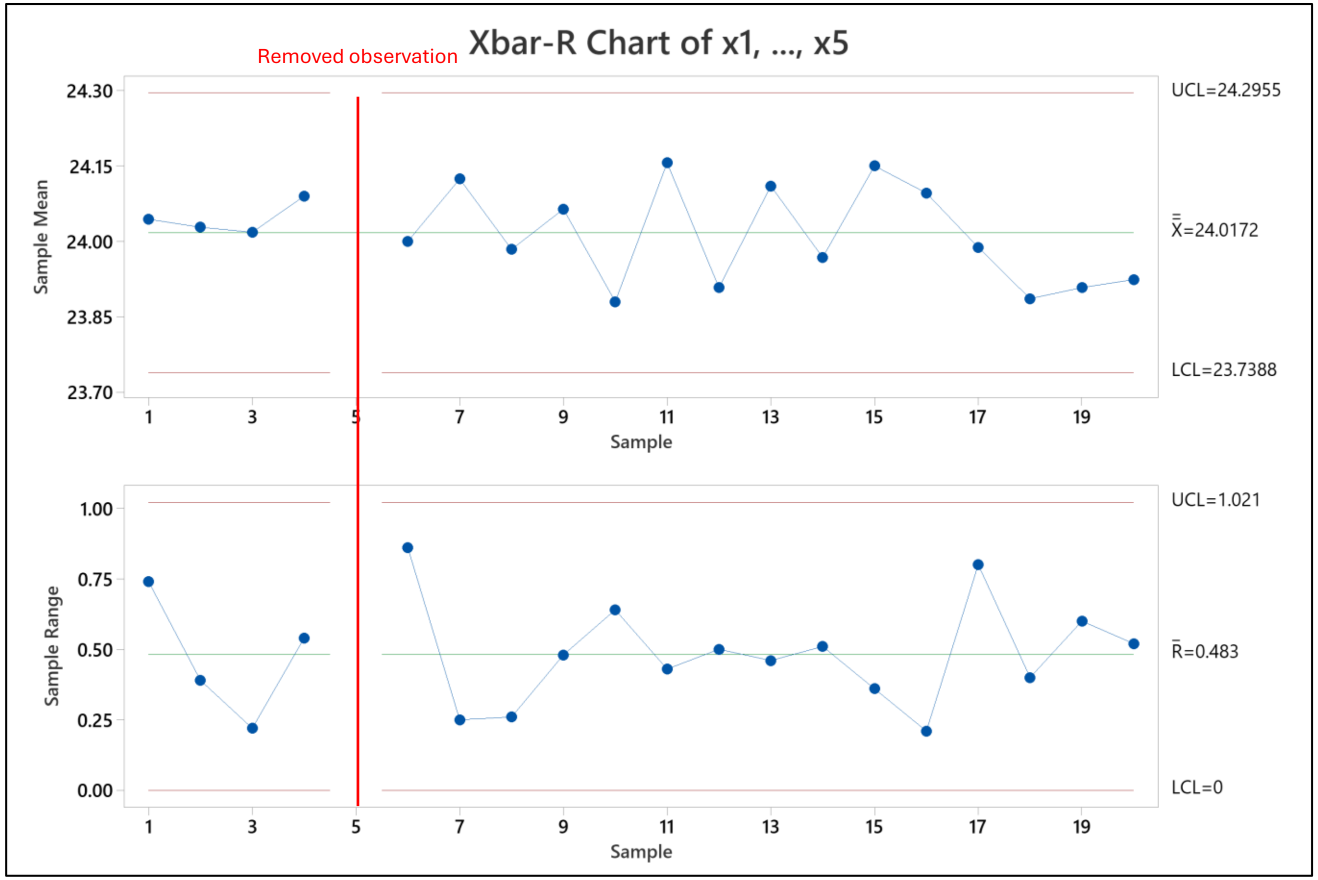
## 4) Let's design the Xbar-R chart (Phase I):

We must remove the OOC observation, which corresponds to the fifth sample:
*Select the cells of the fifth observation → Right click → Clear Cells*

| | x1 | x2 | x3 | x4 | x5 |
|---|---|---|---|---|---|
| 1 | 23.92 | 24.58 | 23.84 | 23.94 | 23.94 |
| 2 | 23.89 | 24.21 | 23.82 | 24.06 | 24.16 |
| 3 | 23.94 | 24.11 | 24.05 | 23.89 | 24.10 |
| 4 | 23.89 | 24.43 | 23.97 | 24.00 | 24.16 |
| 5 | * | * | * | * | * |
| 6 | 23.44 | 23.97 | 24.30 | 24.05 | 24.24 |
| 7 | 24.16 | 24.06 | 24.21 | 23.97 | 24.22 |
| 8 | 23.95 | 24.05 | 23.82 | 24.08 | 24.02 |

Cleared observation

Now we can recompute control limits:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R*




Removing as observation affects the grand mean and range statistics and, therefore, control limits.
The process is in-control.

**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 4) Let's proceed with Phase II and check if the new data is in-control:

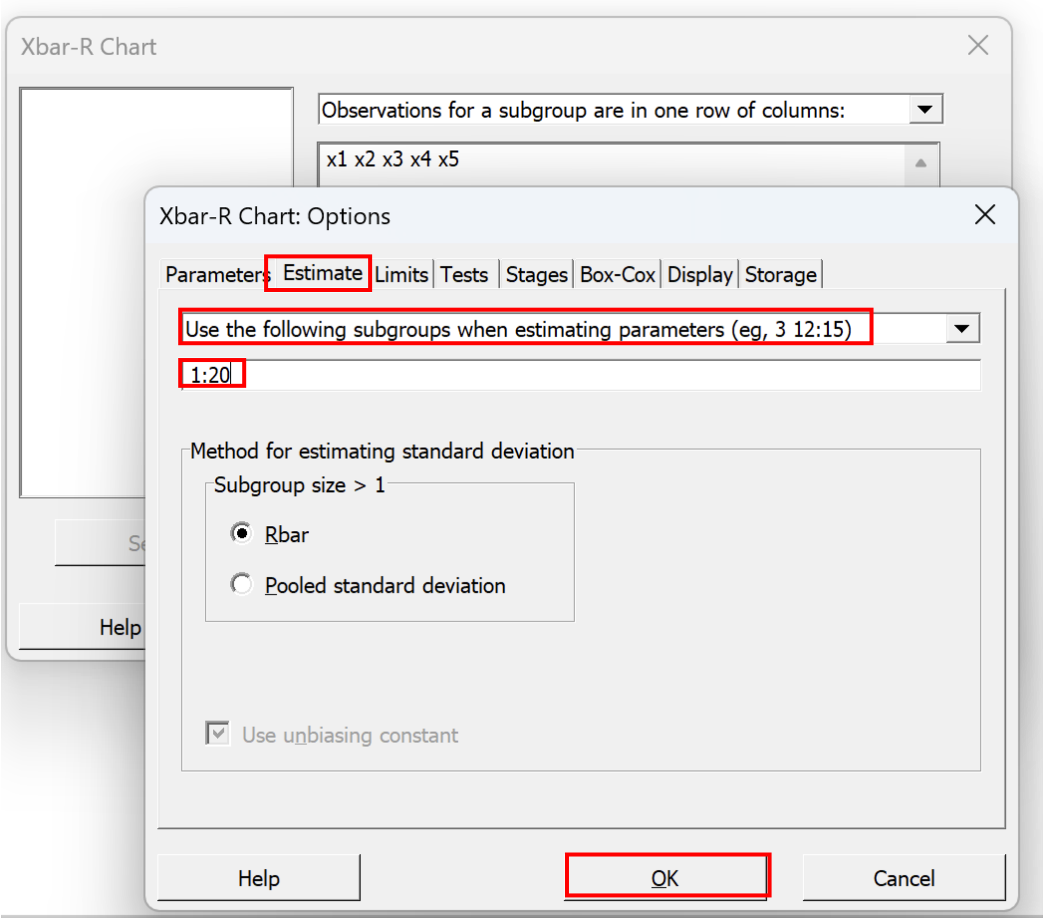
Import the data contained in "ESE02_ex1_PhaseII.csv" in a new worksheet:
*File → Open → Locate "ESE02_ex1_PhaseII.csv" → Open → Set import preferences → Ok*

Copy the cells of Phase II data and paste them in the first worksheet as the tail of Phase I data.

| x1 | x2 | x3 | x4 | x5 |
|---|---|---|---|---|
| 23.92 | 24.58 | 23.84 | 23.94 | 23.94 |
| 23.89 | 24.21 | 23.82 | 24.06 | 24.16 |
| 23.94 | 24.11 | 24.05 | 23.89 | 24.10 |
| 23.89 | 24.43 | 23.97 | 24.00 | 24.16 |
| * | * | * | * | * |
| 23.44 | 23.97 | 24.30 | 24.05 | 24.24 |
| 24.16 | 24.06 | 24.21 | 23.97 | 24.22 |
| 23.95 | 24.05 | 23.82 | 24.08 | 24.02 |
| 23.98 | 24.21 | 23.76 | 24.24 | 24.13 |
| 23.81 | 23.82 | 23.97 | 23.58 | 24.22 |
| 24.08 | 24.16 | 23.95 | 24.38 | 24.21 |
| 23.76 | 23.78 | 23.84 | 24.26 | 23.90 |
| 24.26 | 24.35 | 24.13 | 23.92 | 23.89 |
| 23.87 | 23.97 | 23.76 | 23.97 | 24.27 |
| 24.34 | 24.24 | 24.08 | 24.11 | 23.98 |
| 24.00 | 24.10 | 24.03 | 24.21 | 24.14 |
| 24.45 | 24.21 | 23.65 | 23.90 | 23.73 |
| 23.66 | 23.78 | 23.90 | 24.03 | 24.06 |
| 24.10 | 23.86 | 24.00 | 23.50 | 24.08 |
| 24.19 | 23.73 | 24.22 | 23.70 | 23.78 |
| 24.08 | 24.16 | 23.95 | 24.38 | 24.21 |
| 23.44 | 23.97 | 24.30 | 24.05 | 24.24 |
| 23.76 | 23.78 | 23.84 | 24.26 | 23.90 |
| 23.81 | 23.82 | 23.97 | 23.58 | 24.22 |
| 23.95 | 24.05 | 23.82 | 24.08 | 24.02 |
| 23.92 | 24.58 | 23.84 | 23.94 | 23.94 |
| 23.87 | 23.97 | 23.76 | 23.97 | 24.27 |
| 24.35 | 24.16 | 24.14 | 23.27 | 23.97 |
| 23.89 | 24.21 | 23.82 | 24.06 | 24.16 |
| 24.10 | 23.86 | 24.00 | 23.50 | 24.08 |

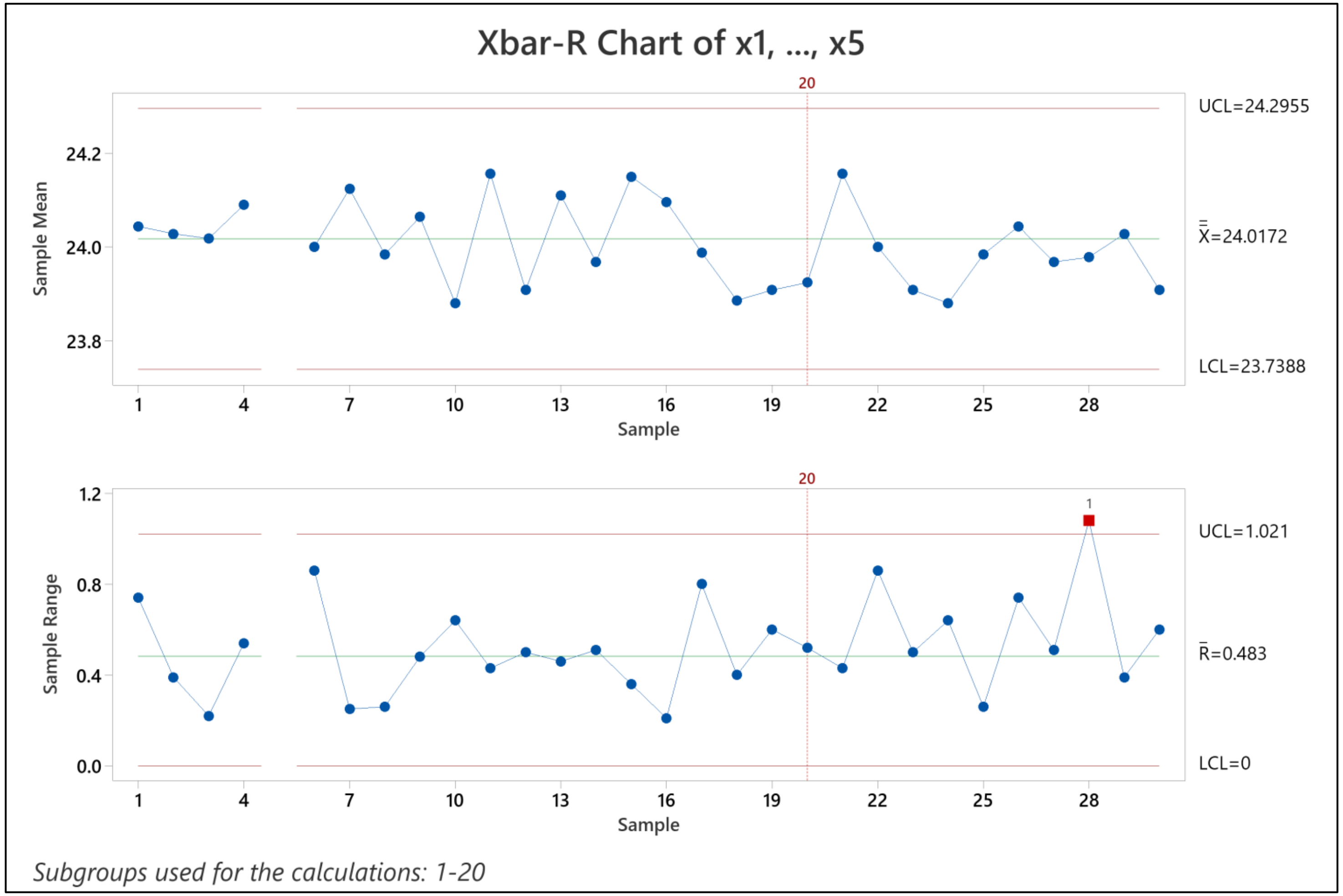
Let's plot the Xbar-R chart including data from Phase II while maintaining control limits from Phase I:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R → Options*

We specify to use only the sample groups from row 1 to 20 (Phase I data) for computing control limits.



**Exercise 1.1** Assuming that the distribution of the gear diameters is unknown, design an Xbar-R chart to verify if the process is in control. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

**4) Let's proceed with Phase II and check if the new data is in-control:**




How to add the reference line to distinguish Phase I and Phase II data? For both Xbar and Range charts:
*Double Click on graph → Click on a data point → Add Item → Show reference lines at time scale positions: 20 → OK*

The R chart signals an out-of-control point, indicating that one sample is showing unusual variability. Since we are in Phase II, we should investigate the process to determine if it is a false alarm or if there is an assignable cause that requires intervention.

**Exercise 1.2** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length ($ARL_0$) of 1000 (assuming that the normal approximation applies for both of them).

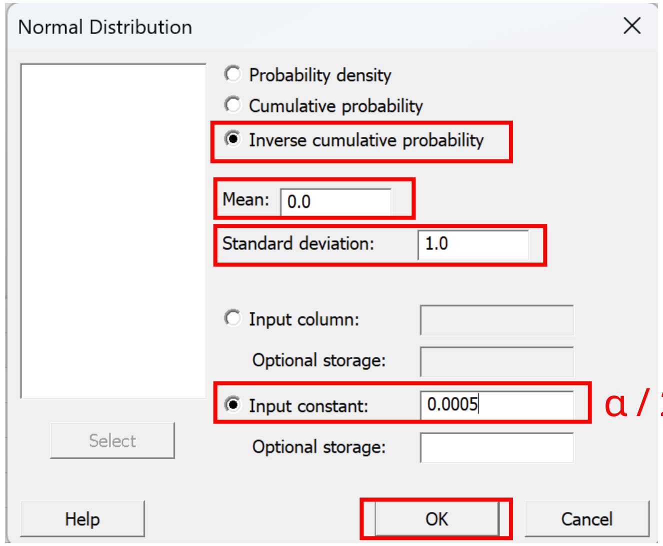
1) **Compute the Type I error ($\alpha$):**

$$\alpha = \frac{1}{ARL_0} = \frac{1}{1000} = 0.001$$

2) **Compute the Type I error ($\alpha$):**

$$K = |z_{\alpha/2}| = \boxed{3.291}$$

To compute K from the Inverse cumulative normal:
*Calc → Probability Distributions → Normal*




α / 2

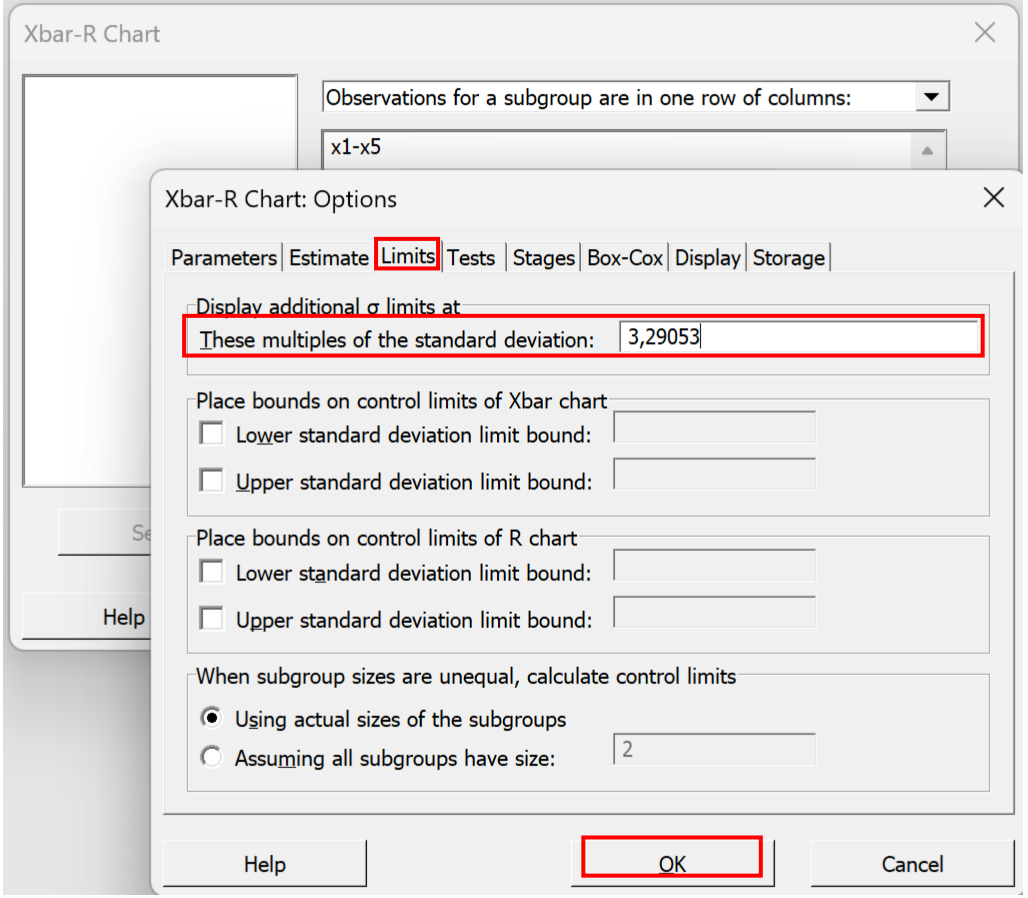
Normal with mean = 0 and standard deviation = 1

| P( X ≤ x ) | x |
|---|---|
| 0.0005 | $\boxed{-3.29053}$ |

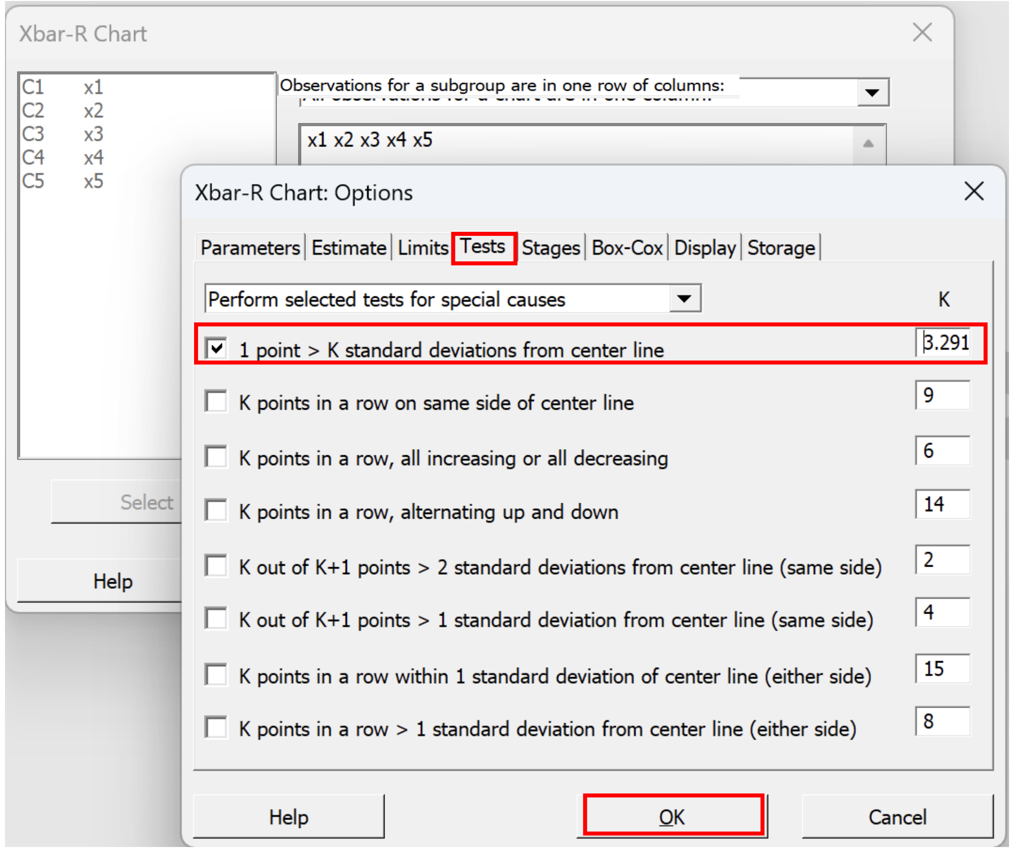
**Exercise 1.2** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length ($ARL_0$) of 1000 (assuming that the normal approximation applies for both of them).

## 3) Build the Xbar-R chart with K=3.291 on Phase I original data:

To build the Xbar-R chart with K different from 3:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R → Xbar-R Options*

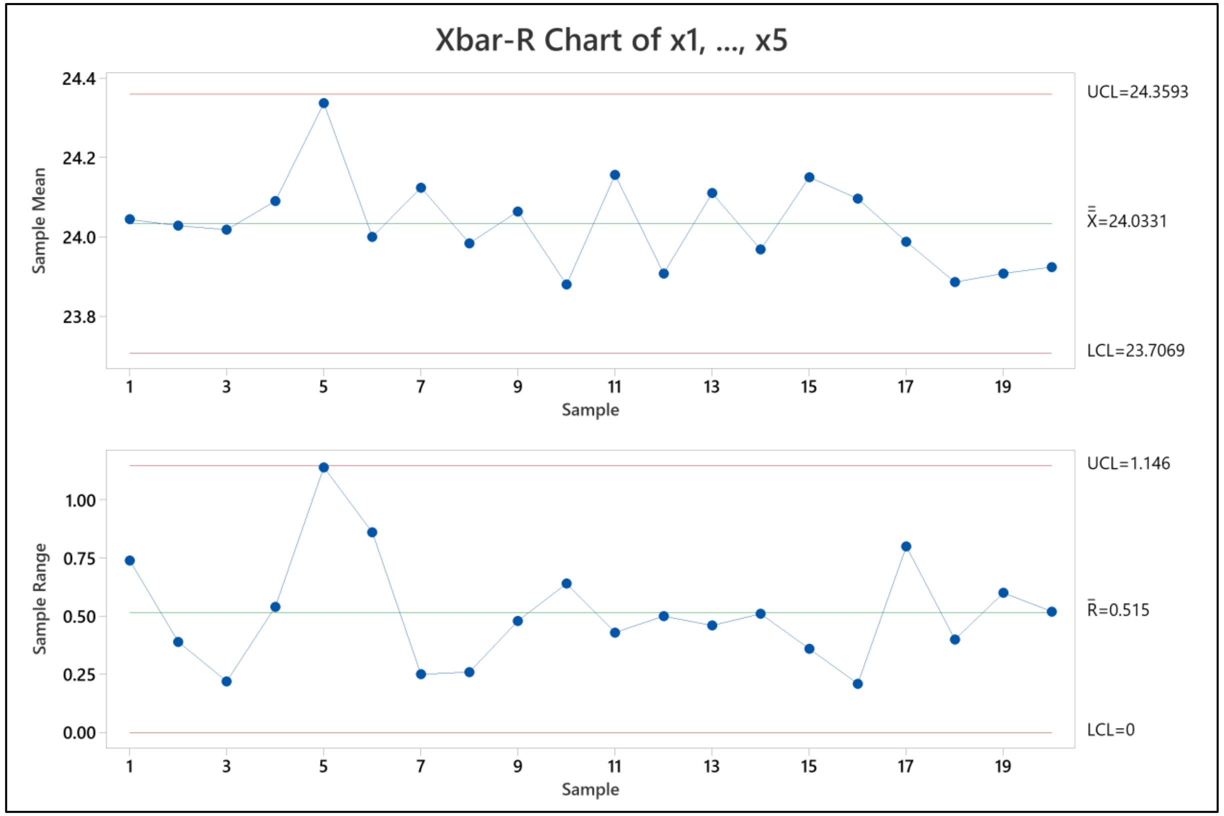



To build the Xbar-R chart with K different from 3:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R → Xbar-R Options*



**Exercise 1.2** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length ($ARL_0$) of 1000 (assuming that the normal approximation applies for both of them).

3) **Build the Xbar-R chart with K=3.291 on Phase I original data:**




Note that using alpha = 0.001 has resulted in both X-bar and R charts without any out-of-control samples. These charts will signal fewer false alarms, but they may miss process non-random variations.

**Let's proceed with Phase II.**

**Exercise 1.2** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length ($ARL_0$) of 1000 (assuming that the normal approximation applies for both of them).

## 3) Build the Xbar-R chart with K=3.291 on Phase I original data:

$$
\begin{aligned}
z_{\alpha/2} &= 3.291 \\
n &= 5 \\
d_2(n) &= 2.326 \\
d_3(n) &= 0.864 \\
\bar{\bar{x}} &= 24.033 \\
\bar{R} &= 0.516
\end{aligned}
$$

**Xbar Chart** (unknown parameters, K≠3):

$$\text{UCL} = \bar{\bar{x}} + z_{\alpha/2}\frac{1}{d_2\sqrt{n}}\bar{R}$$

$$= 24.033 + 3.291\left(\frac{0.516}{2.326\sqrt{5}}\right) = 24.360$$

$$\text{CL} = \bar{\bar{x}} = 24.033$$

$$\text{LCL} = \bar{\bar{x}} - z_{\alpha/2}\frac{1}{d_2\sqrt{n}}\bar{R}$$

$$= 24.033 - 3.291\left(\frac{0.516}{2.326\sqrt{5}}\right) = 23.706$$

**R Chart** (unknown parameters, K≠3):

$$\text{UCL} = \bar{R} + z_{\alpha/2}\frac{d_3(n)}{d_2(n)}\bar{R}$$

$$= 0.516 + 3.291\left(\frac{0.864}{2.326} \times 0.516\right) = 1.144$$

$$\text{CL} = \bar{R} = 0.516$$

$$\text{LCL} = \max\left\{0, \bar{R} - z_{\alpha/2}\frac{d_3(n)}{d_2(n)}\bar{R}\right\}$$

$$= \max\left\{0, 0.516 - 3.291\left(\frac{0.864}{2.326} \times 0.516\right)\right\} = 0$$

**Exercise 1.2** Redesign the X-bar and R chart in order to achieve in both the charts an Average Run Length ($ARL_0$) of 1000 (assuming that the normal approximation applies for both of them).

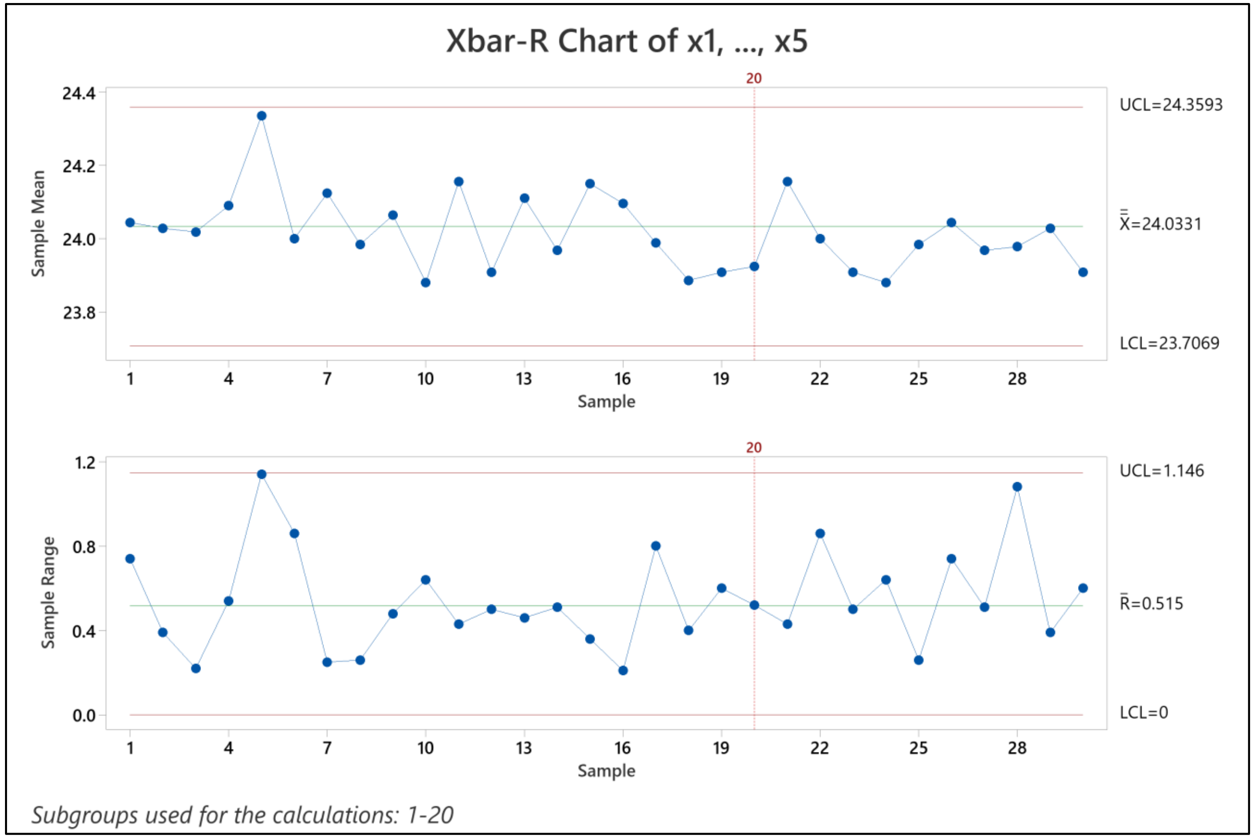
## 4) Plot Phase I and Phase II data with control limits computed in Phase I:

Append Phase II data to Phase I data.

Build the Xbar-R chart with K different from 3 including variables x1, x2, x3, x4, x5:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-R*

*→ Xbar-R Options → Estimate → Use the subgroups from 1 to 20 to compute control limits*

*→ Xbar-R Options → Estimate → Tests → K = 3.291*




The observation previously signaled as out-of-control in Phase II (using alpha = 0.0027) is now in-control (using alpha = 0.001). The choice of the tolerated Type I error depends on the balance between detecting true out-of-control conditions and minimizing false alarms. A lower alpha (e.g., 0.001) reduces false alarms but may miss subtle process variations. The decision should be based on the specific requirements and tolerance for risk in the process being monitored.

**Exercise 1.3** Determine the operating characteristic curve (OC) for the X-bar chart (by using K=3 and expressing the shift of the mean in standard deviation units).

To determine the OC curve, we need to compute the probability of $\beta$ for each value of the shift $\mu$.

We are testing the null hypothesis $H_0$ that the sample mean $\bar{X}$ is normally distributed with mean $\mu_0$ and variance $\frac{\sigma^2}{n}$.

$$H_0 : \bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$$

The alternative hypothesis is that the sample mean is normally distributed with mean $\mu_1$ and variance $\frac{\sigma^2}{n}$.

$$H_1 : \bar{X} \sim N(\mu_1, \frac{\sigma^2}{n})$$

So $\beta$ is the probability of not rejecting $H_0$ when $H_1$ is true.

$$\beta = P(LCL \leq \bar{X} \leq UCL \mid H_1)$$

$$\beta = P\left(Z \leq \frac{UCL - \mu_1}{\sigma/\sqrt{n}}\right) - P\left(Z \leq \frac{LCL - \mu_1}{\sigma/\sqrt{n}}\right)$$

If we define $\delta = \frac{\mu_1 - \mu_0}{\sigma}$, we can write:

$$\beta = P\left(Z \leq 3 - \delta\sqrt{n}\right) - P\left(Z \leq -3 - \delta\sqrt{n}\right)$$

**Exercise 1.3** Determine the operating characteristic curve (OC) for the X-bar chart (by using K=3 and expressing the shift of the mean in standard deviation units).

Create a "delta" column:

*Calc → Make patterned data → Simple Set of Numbers*

From first value: *0;* To last value: *4;* In steps of: *0.001;*

Create two columns of values $3-\delta\sqrt{n}$ and $-3-\delta\sqrt{n}$

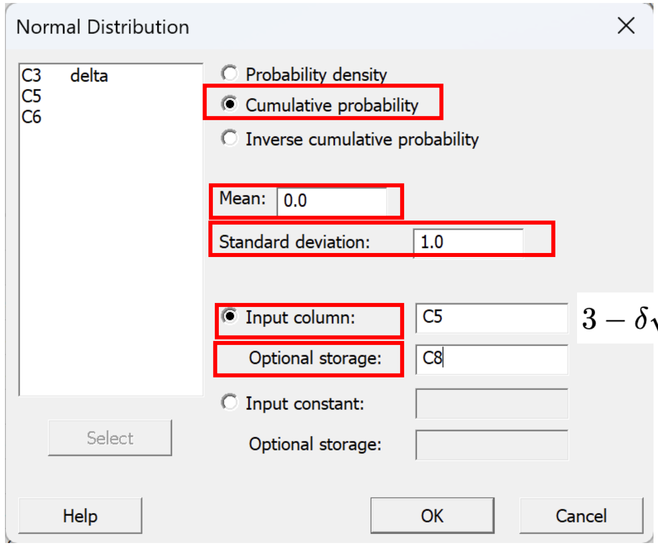
*Calc → Calculator*

Expression:
3 - 'delta' * SQRT(5)

Expression:
-3 - 'delta' * SQRT(5)

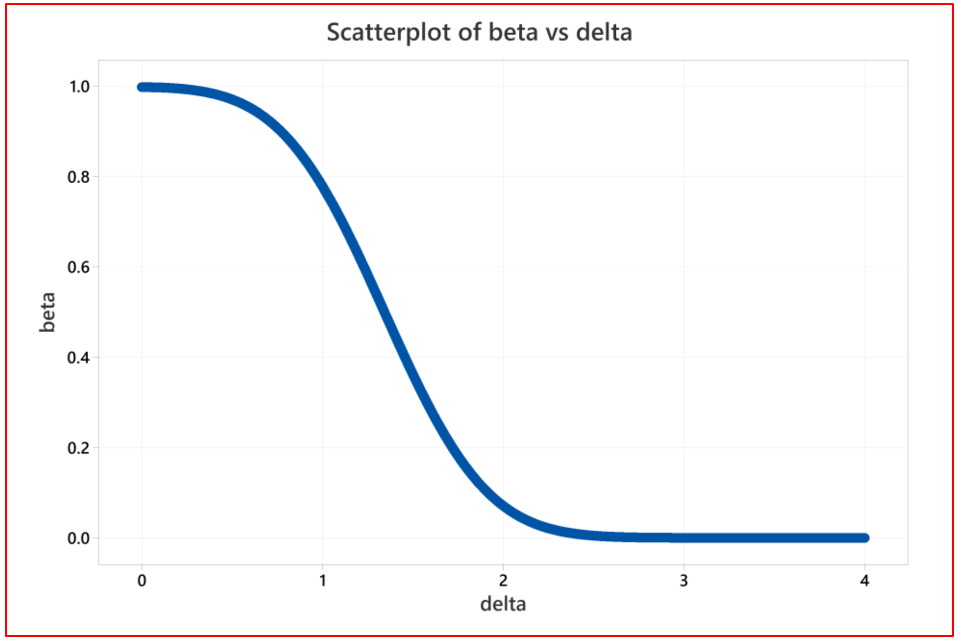
Create two columns of values $P\left(Z \leq 3-\delta\sqrt{n}\right)$ and $P\left(Z \leq -3-\delta\sqrt{n}\right)$

*Calc → Probability Distributions → Normal*




$3-\delta\sqrt{n}$ / $-3-\delta\sqrt{n}$

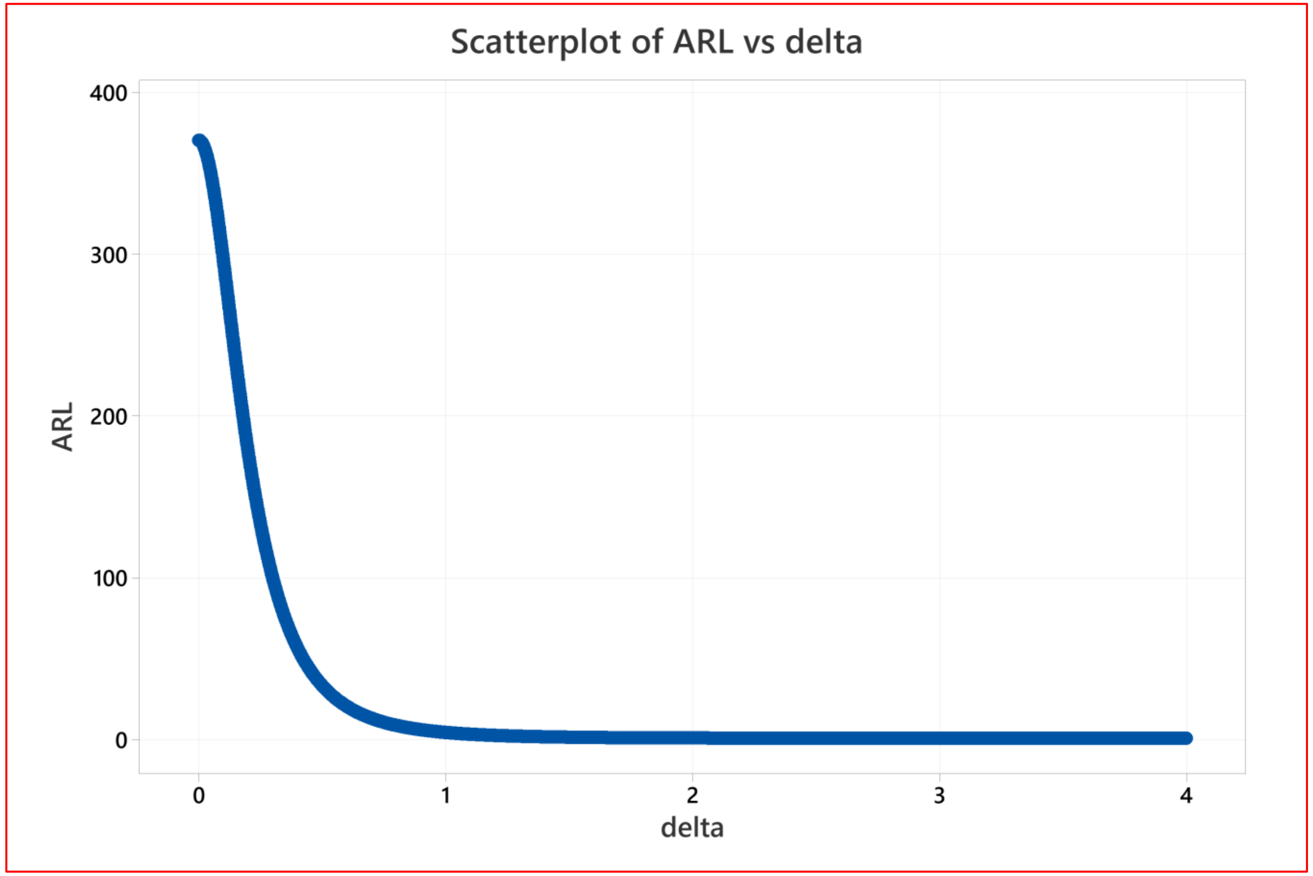
Compute $\beta$ and plot $\beta$ against $\delta$



**Exercise 1.4** Determine the corresponding ARL curve.

$$ARL = \frac{1}{1 - \beta}$$

Compute ARL and plot ARL against $\delta$



**Exercise 1.5** Estimate the standard deviation through the statistic R (consider original Phase I data).

$$\bar{R} = 0.516$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2(n)} = \frac{0.516}{2.326} = 0.222$$

**Exercise 1.6** Design the confidence interval on the process mean that corresponds to the control limits computed in point 1 (consider original Phase I data).

The confidence interval corresponding to the control limits computed in point 1 uses:

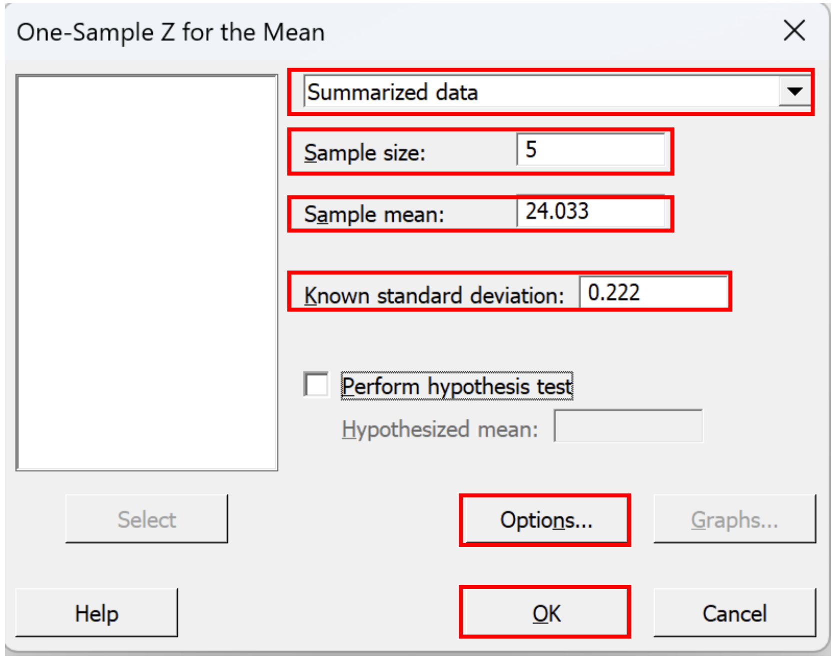
- $n = 5$
- $\alpha = 0.0027$
- $\hat{\sigma} = 0.222$ (computed from the data)
- $\overline{X} = 24.033$ (computed from the data)

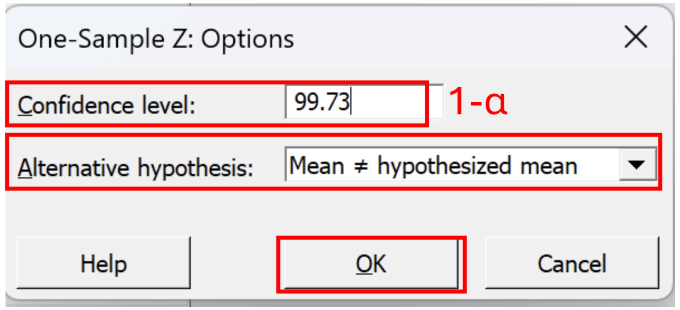
Remember the formula of the confidence interval (assume that $\hat{\sigma}$ is the real population variance):

$$\bar{X} - z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$$

**Exercise 1.6** Design the confidence interval on the process mean that corresponds to the control limits computed in point 1 (consider original Phase I data).

Let's exploit a One-Sample Z test (assumed known standard deviation) to compute the confidence interval:
*Stat → Basic Statistics → 1-Sample Z*




## Descriptive Statistics

| N | Mean | SE Mean | 99.73% CI for μ |
|---|---|---|---|
| 5 | 24.0330 | 0.0993 | (23.7352, 24.3308) |

*μ: population mean of Sample*
*Known standard deviation = 0.222*

See the UCL and LCL of slide 13.

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

*Xbar chart* (in *Xbar*-S) K=3

$$\text{UCL} = \hat{\mu} + \text{K}\frac{\hat{\sigma}}{\sqrt{n}} = \overline{\overline{x}} + 3\frac{1}{c_4\sqrt{n}}\overline{s} = \overline{\overline{x}} + A_3(n)\overline{s}$$

$$\text{CL} = \hat{\mu} = \overline{\overline{x}}$$

$$\text{LCL} = \hat{\mu} - \text{K}\frac{\hat{\sigma}}{\sqrt{n}} = \overline{\overline{x}} - 3\frac{1}{c_4\sqrt{n}}\overline{s} = \overline{\overline{x}} - A_3(n)\overline{s}$$

Analogously: *S chart*

*parameters*

| *known* | *unknown* |
|---|---|
| $\text{UCL} = B_6(n)\sigma$ | $\text{UCL} = B_4(n)\overline{s}$ |
| $\text{CL} = c_4(n)\sigma$ | $\text{CL} = \overline{s}$ |
| $\text{LCL} = B_5(n)\sigma$ | $\text{LCL} = B_3(n)\overline{s}$ |

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

*S chart*
*Known parameters*

$$UCL = \mu_S + K\sigma_S = c_4\sigma + 3\sqrt{1-c_4^2}\,\sigma = B_6\sigma \Rightarrow B_6 = c_4 + 3\sqrt{1-c_4^2}$$

$$CL = \mu_S = c_4\sigma$$

$$LCL = \mu_S - K\sigma_S = c_4\sigma - 3\sqrt{1-c_4^2}\,\sigma = B_5\sigma \Rightarrow B_5 = c_4 - 3\sqrt{1-c_4^2}$$

*Unknown parameters*

$$UCL = c_4\hat{\sigma} + 3\sqrt{1-c_4^2}\,\hat{\sigma} = \bar{s} + 3\frac{\sqrt{1-c_4^2}}{c_4}\bar{s} = B_4\bar{s} \Rightarrow B_4 = 1 + 3\frac{\sqrt{1-c_4^2}}{c_4}$$

$$CL = c_4\hat{\sigma} = \bar{s}$$

$$LCL = c_4\hat{\sigma} - 3\sqrt{1-c_4^2}\,\hat{\sigma} = \bar{s} - 3\frac{\sqrt{1-c_4^2}}{c_4}\bar{s} = B_3\bar{s} \Rightarrow B_3 = 1 - 3\frac{\sqrt{1-c_4^2}}{c_4}$$

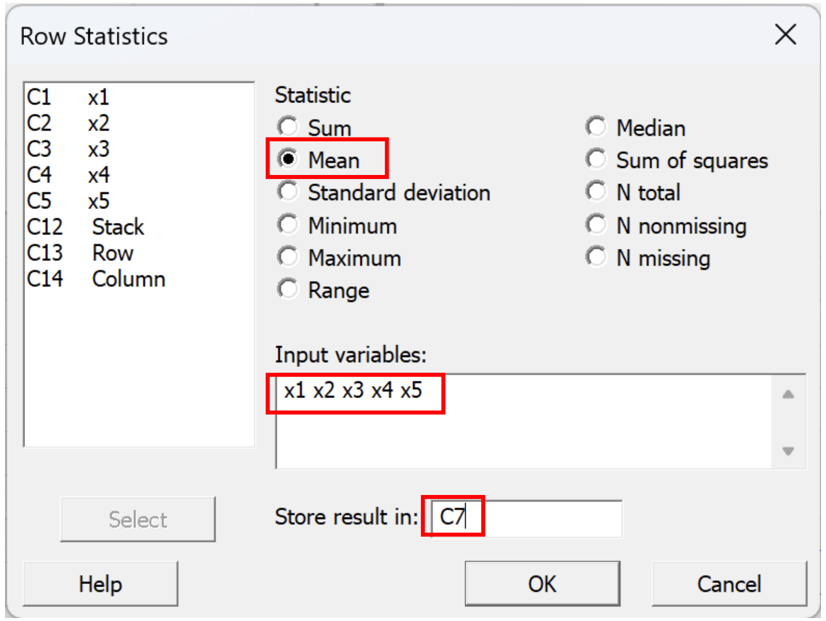
**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

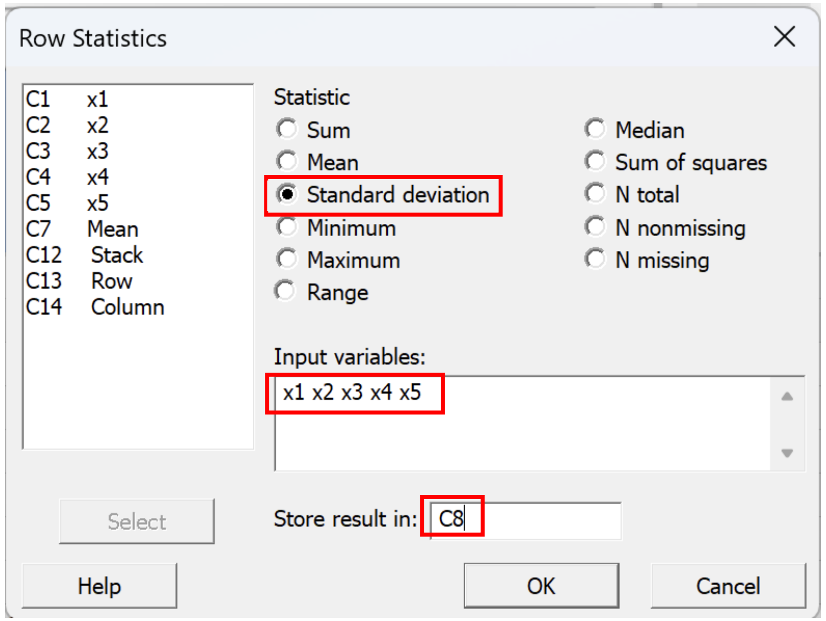
We have already checked normality of data during point 1.1, so we can directly design the X-bar and S charts.

## 1) Design the Xbar-S chart (Phase I):

To compute the mean and standard deviation statistic for each sample:

*Calc → Row Statistics*



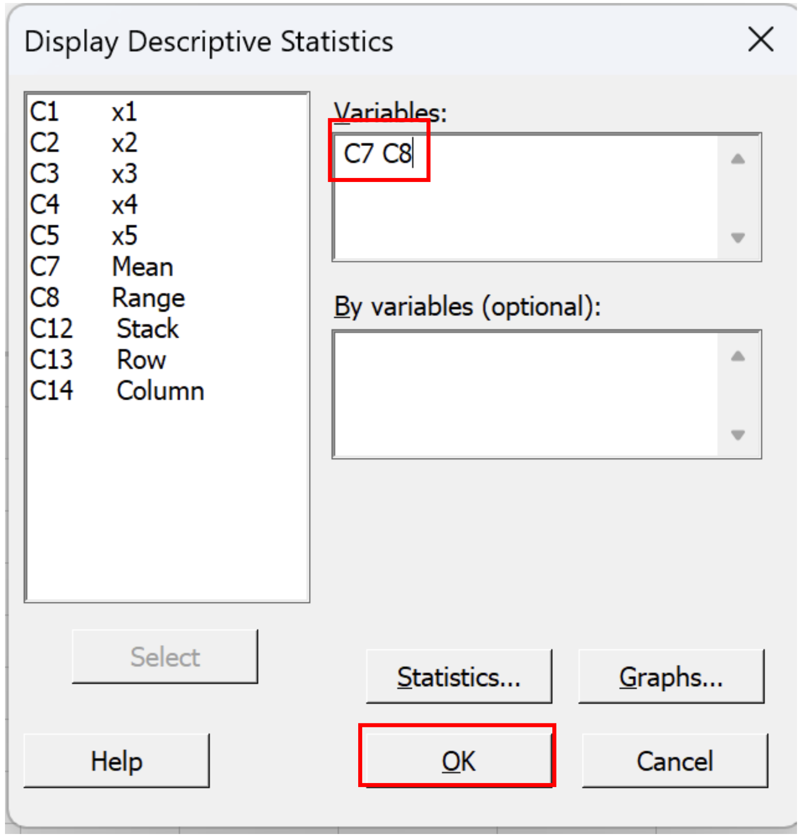

| C7 | C8 |
|---|---|
| Mean | St. dev. |
| 24.044 | 0.302457 |
| 24.028 | 0.168731 |
| 24.018 | 0.098336 |
| 24.090 | 0.213892 |
| 24.336 | 0.418127 |
| 24.000 | 0.340808 |
| 24.124 | 0.106911 |
| 23.984 | 0.103586 |
| 24.064 | 0.197560 |
| 23.880 | 0.235690 |
| 24.156 | 0.159154 |
| 23.908 | 0.204255 |
| 24.110 | 0.203101 |
| 23.968 | 0.189789 |
| 24.150 | 0.141067 |
| 24.096 | 0.084439 |
| 23.988 | 0.335887 |
| 23.886 | 0.168464 |
| 23.908 | 0.246820 |
| 23.924 | 0.258322 |

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 1) Let’s design the Xbar-S chart (Phase I):

Now we can compute the grand mean (mean of sample means) and the mean of sample standard deviations:

*Stat → Basic Statistics → Display Descriptive Statistics*




**Statistics**

| Variable | N | N* | Mean | SE Mean | StDev | Minimum | Q1 | Median | Q3 |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 20 | 0 | $\boxed{24.0331}$ | 0.0251979 | 0.112689 | 23.88 | 23.935 | 24.023 | 24.1065 |
| St. dev. | 20 | 0 | $\boxed{0.208870}$ | 0.0199181 | 0.0890766 | 0.0844393 | 0.145589 | 0.200331 | 0.255446 |

$$\bar{\bar{X}} = 24.033$$

$$\bar{S} = 0.209$$

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in "gears_phase2.csv" is in control.

## 1) Let's design the Xbar-S chart (Phase I):

You can refer to statistical tables to retrieve control chart constants :

**Factors for Constructing Variables Control Charts**

| | Chart for Averages | | | | | Chart for Standard Deviations | | | | | | Chart for Ranges | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Observations in Sample, $n$ | Factors for Control Limits | | | Factors for Center Line | | Factors for Control Limits | | | | Factors for Center Line | | Factors for Control Limits | | | | |
| | $A$ | $A_2$ | $A_3$ | $c_4$ | $1/c_4$ | $B_3$ | $B_4$ | $B_5$ | $B_6$ | $d_2$ | $1/d_2$ | $d_3$ | $D_1$ | $D_2$ | $D_3$ | $D_4$ |
| 2 | 2.121 | 1.880 | 2.659 | 0.7979 | 1.2533 | 0 | 3.267 | 0 | 2.606 | 1.128 | 0.8865 | 0.853 | 0 | 3.686 | 0 | 3.267 |
| 3 | 1.732 | 1.023 | 1.954 | 0.8862 | 1.1284 | 0 | 2.568 | 0 | 2.276 | 1.693 | 0.5907 | 0.888 | 0 | 4.358 | 0 | 2.574 |
| 4 | 1.500 | 0.729 | 1.628 | 0.9213 | 1.0854 | 0 | 2.266 | 0 | 2.088 | 2.059 | 0.4857 | 0.880 | 0 | 4.698 | 0 | 2.282 |
| 5 | 1.342 | 0.577 | 1.427 | 0.9400 | 1.0638 | 0 | 2.089 | 0 | 1.964 | 2.326 | 0.4299 | 0.864 | 0 | 4.918 | 0 | 2.114 |
| 6 | 1.225 | 0.483 | 1.287 | 0.9515 | 1.0510 | 0.030 | 1.970 | 0.029 | 1.874 | 2.534 | 0.3946 | 0.848 | 0 | 5.078 | 0 | 2.004 |
| 7 | 1.134 | 0.419 | 1.182 | 0.9594 | 1.0423 | 0.118 | 1.882 | 0.113 | 1.806 | 2.704 | 0.3698 | 0.833 | 0.204 | 5.204 | 0.076 | 1.924 |
| 8 | 1.061 | 0.373 | 1.099 | 0.9650 | 1.0363 | 0.185 | 1.815 | 0.179 | 1.751 | 2.847 | 0.3512 | 0.820 | 0.388 | 5.306 | 0.136 | 1.864 |
| 9 | 1.000 | 0.337 | 1.032 | 0.9693 | 1.0317 | 0.239 | 1.761 | 0.232 | 1.707 | 2.970 | 0.3367 | 0.808 | 0.547 | 5.393 | 0.184 | 1.816 |
| 10 | 0.949 | 0.308 | 0.975 | 0.9727 | 1.0281 | 0.284 | 1.716 | 0.276 | 1.669 | 3.078 | 0.3249 | 0.797 | 0.687 | 5.469 | 0.223 | 1.777 |
| 11 | 0.905 | 0.285 | 0.927 | 0.9754 | 1.0252 | 0.321 | 1.679 | 0.313 | 1.637 | 3.173 | 0.3152 | 0.787 | 0.811 | 5.535 | 0.256 | 1.744 |
| 12 | 0.866 | 0.266 | 0.886 | 0.9776 | 1.0229 | 0.354 | 1.646 | 0.346 | 1.610 | 3.258 | 0.3069 | 0.778 | 0.922 | 5.594 | 0.283 | 1.717 |
| 13 | 0.832 | 0.249 | 0.850 | 0.9794 | 1.0210 | 0.382 | 1.618 | 0.374 | 1.585 | 3.336 | 0.2998 | 0.770 | 1.025 | 5.647 | 0.307 | 1.693 |
| 14 | 0.802 | 0.235 | 0.817 | 0.9810 | 1.0194 | 0.406 | 1.594 | 0.399 | 1.563 | 3.407 | 0.2935 | 0.763 | 1.118 | 5.696 | 0.328 | 1.672 |
| 15 | 0.775 | 0.223 | 0.789 | 0.9823 | 1.0180 | 0.428 | 1.572 | 0.421 | 1.544 | 3.472 | 0.2880 | 0.756 | 1.203 | 5.741 | 0.347 | 1.653 |
| 16 | 0.750 | 0.212 | 0.763 | 0.9835 | 1.0168 | 0.448 | 1.552 | 0.440 | 1.526 | 3.532 | 0.2831 | 0.750 | 1.282 | 5.782 | 0.363 | 1.637 |
| 17 | 0.728 | 0.203 | 0.739 | 0.9845 | 1.0157 | 0.466 | 1.534 | 0.458 | 1.511 | 3.588 | 0.2787 | 0.744 | 1.356 | 5.820 | 0.378 | 1.622 |
| 18 | 0.707 | 0.194 | 0.718 | 0.9854 | 1.0148 | 0.482 | 1.518 | 0.475 | 1.496 | 3.640 | 0.2747 | 0.739 | 1.424 | 5.856 | 0.391 | 1.608 |
| 19 | 0.688 | 0.187 | 0.698 | 0.9862 | 1.0140 | 0.497 | 1.503 | 0.490 | 1.483 | 3.689 | 0.2711 | 0.734 | 1.487 | 5.891 | 0.403 | 1.597 |
| 20 | 0.671 | 0.180 | 0.680 | 0.9869 | 1.0133 | 0.510 | 1.490 | 0.504 | 1.470 | 3.735 | 0.2677 | 0.729 | 1.549 | 5.921 | 0.415 | 1.585 |
| 21 | 0.655 | 0.173 | 0.663 | 0.9876 | 1.0126 | 0.523 | 1.477 | 0.516 | 1.459 | 3.778 | 0.2647 | 0.724 | 1.605 | 5.951 | 0.425 | 1.575 |
| 22 | 0.640 | 0.167 | 0.647 | 0.9882 | 1.0119 | 0.534 | 1.466 | 0.528 | 1.448 | 3.819 | 0.2618 | 0.720 | 1.659 | 5.979 | 0.434 | 1.566 |
| 23 | 0.626 | 0.162 | 0.633 | 0.9887 | 1.0114 | 0.545 | 1.455 | 0.539 | 1.438 | 3.858 | 0.2592 | 0.716 | 1.710 | 6.006 | 0.443 | 1.557 |
| 24 | 0.612 | 0.157 | 0.619 | 0.9892 | 1.0109 | 0.555 | 1.445 | 0.549 | 1.429 | 3.895 | 0.2567 | 0.712 | 1.759 | 6.031 | 0.451 | 1.548 |
| 25 | 0.600 | 0.153 | 0.606 | 0.9896 | 1.0105 | 0.565 | 1.435 | 0.559 | 1.420 | 3.931 | 0.2544 | 0.708 | 1.806 | 6.056 | 0.459 | 1.541 |

For $n > 25$.

$$A = \frac{3}{\sqrt{n}} \quad A_3 = \frac{3}{c_4\sqrt{n}} \quad c_4 \cong \frac{4(n-1)}{4n-3}$$

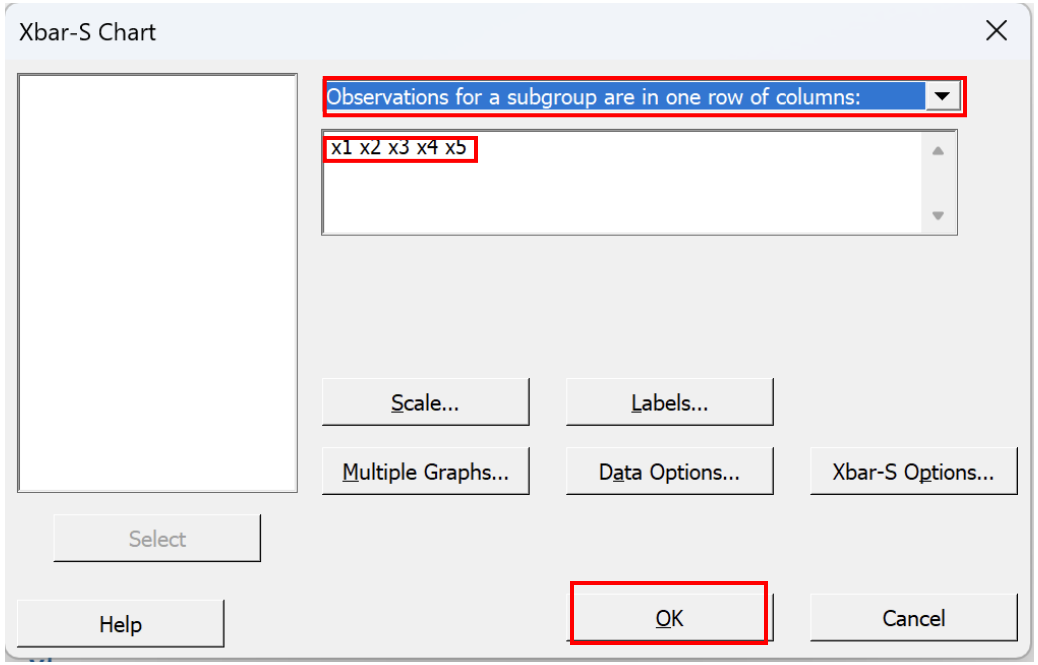
$$B_3 = 1 - \frac{3}{c_4\sqrt{2(n-1)}} \quad B_4 = 1 + \frac{3}{c_4\sqrt{2(n-1)}}$$

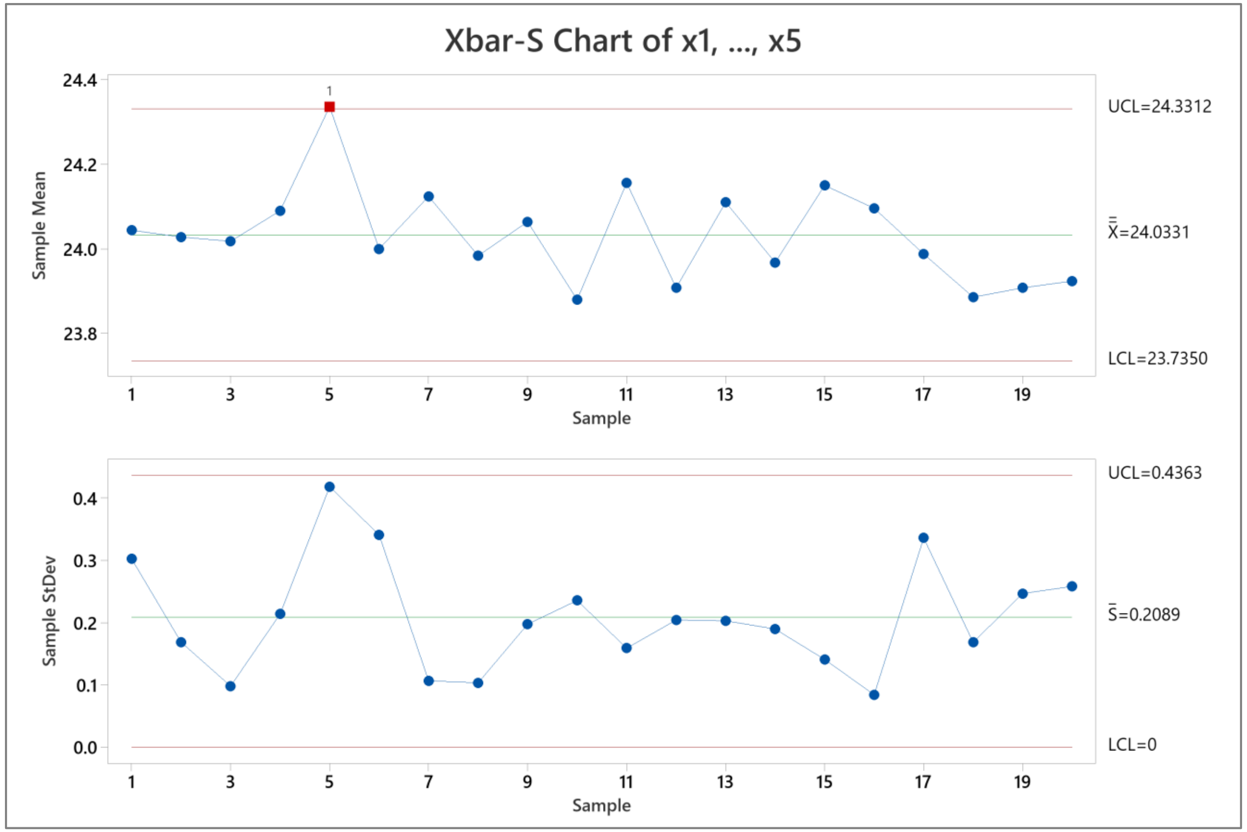
$$B_5 = c_4 - \frac{3}{\sqrt{2(n-1)}} \quad B_6 = c_4 + \frac{3}{\sqrt{2(n-1)}}$$

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 1) Let’s design the Xbar-S chart (Phase I):

Let’s build the Xbar-R chart in Minitab:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-S*



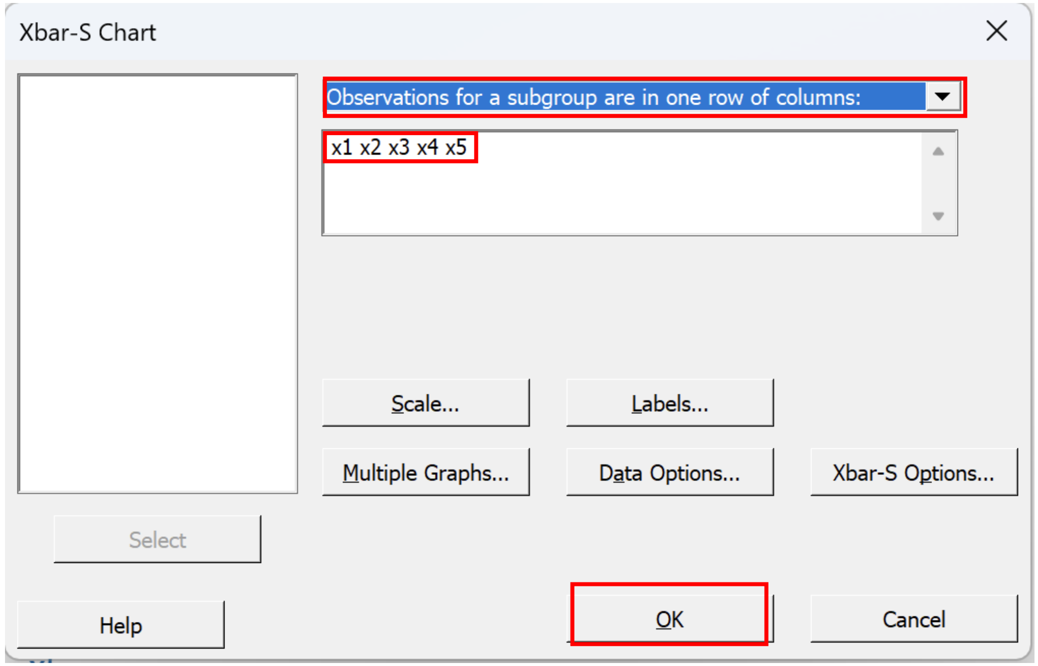



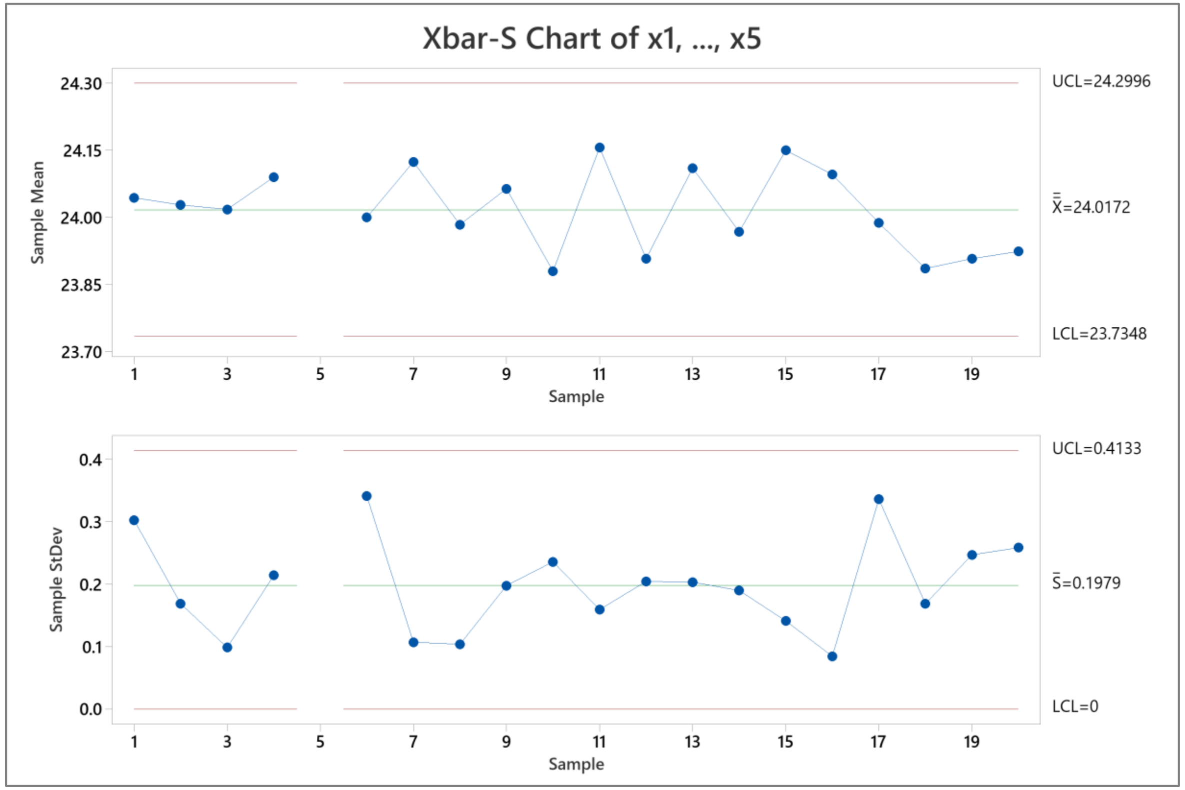

The X-bar chart is signalling an out-of-control point (observation 5), the same detected by the Xbar chart at point 1.1. However, the S chart is not signalling an alarm for observation 5, unlike the R chart. This is because the S chart considers the overall spread of data points within each sample and is less influenced by extreme values. In contrast, the range statistic is highly sensitive to extreme values within the sample since it considers only the lowest and highest values.

**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 1) Let’s design the Xbar-S chart (Phase I):

Remove the OOC observation and recompute control limits:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-S*



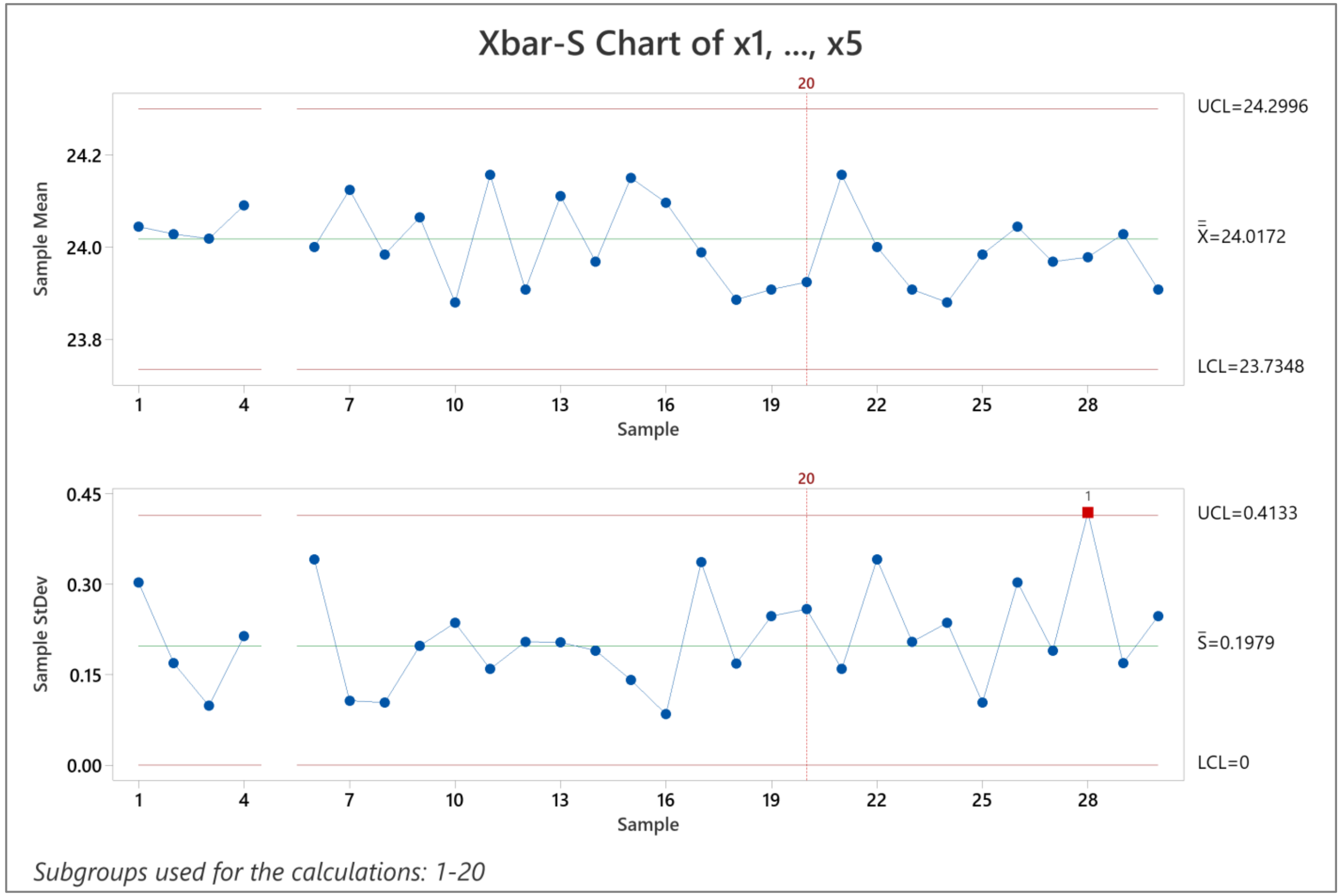



**Exercise 1.7** Verify if the process is in control by using an X-bar and S chart. For any out-of-control points detected in Phase I, assume that an assignable cause is found. Check if data contained in “gears_phase2.csv” is in control.

## 2) Let’s design the Xbar-S chart (Phase II):

To check if Phase II data is in control, append Phase II data to Phase I data. Then, plot the whole dataset using control limits computed on Phase I data: *Stats → Control Charts → Variable Charts for Subgroups → Xbar-S*

*→ Xbar-R Options → Estimate → Use the subgroups from 1 to 20 to compute control limits*




The S chart signals an out-of-control point, the same detected by the R chart, indicating that one sample is showing unusual variability. As already discussed, we should investigate the process to determine if it is a false alarm or if there is an assignable cause that requires intervention.

**Exercise 1.8** Knowing that the gear diameter is distributed as a normal distribution with mean 24 mm and standard deviation 0.26 mm, design an Xbar and S chart. For any out-of-control points detected, assume that an assignable cause is found. Check if Phase II data is in control.

Since there is no constraint on the choice of Type I error $\alpha$, we can set $K = 3$ ($\alpha = 0.0027$). Remember the formulas for the control limits for known parameters (k=3):

$\bar{\mathbf{X}}$ **chart**:

$$UCL = \mu + K\frac{\sigma}{\sqrt{n}} = \mu + A(n)\sigma$$

$$CL = \mu$$

$$LCL = \mu - K\frac{\sigma}{\sqrt{n}} = \mu - A(n)\sigma$$

$S$ **chart**:

$$UCL = B_6(n)c_4\sigma$$

$$CL = c_4\sigma$$

$$LCL = B_5(n)c_4\sigma$$

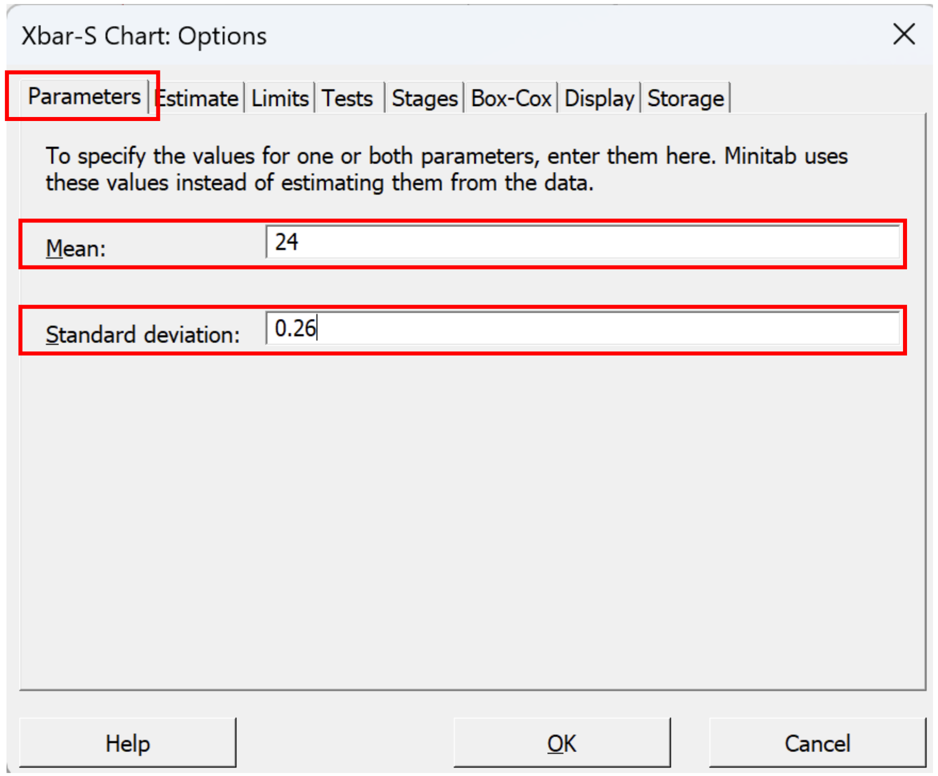
**Exercise 1.8** Knowing that the gear diameter is distributed as a normal distribution with mean 24 mm and standard deviation 0.26 mm, design an Xbar and S chart. For any out-of-control points detected, assume that an assignable cause is found. Check if Phase II data is in control.

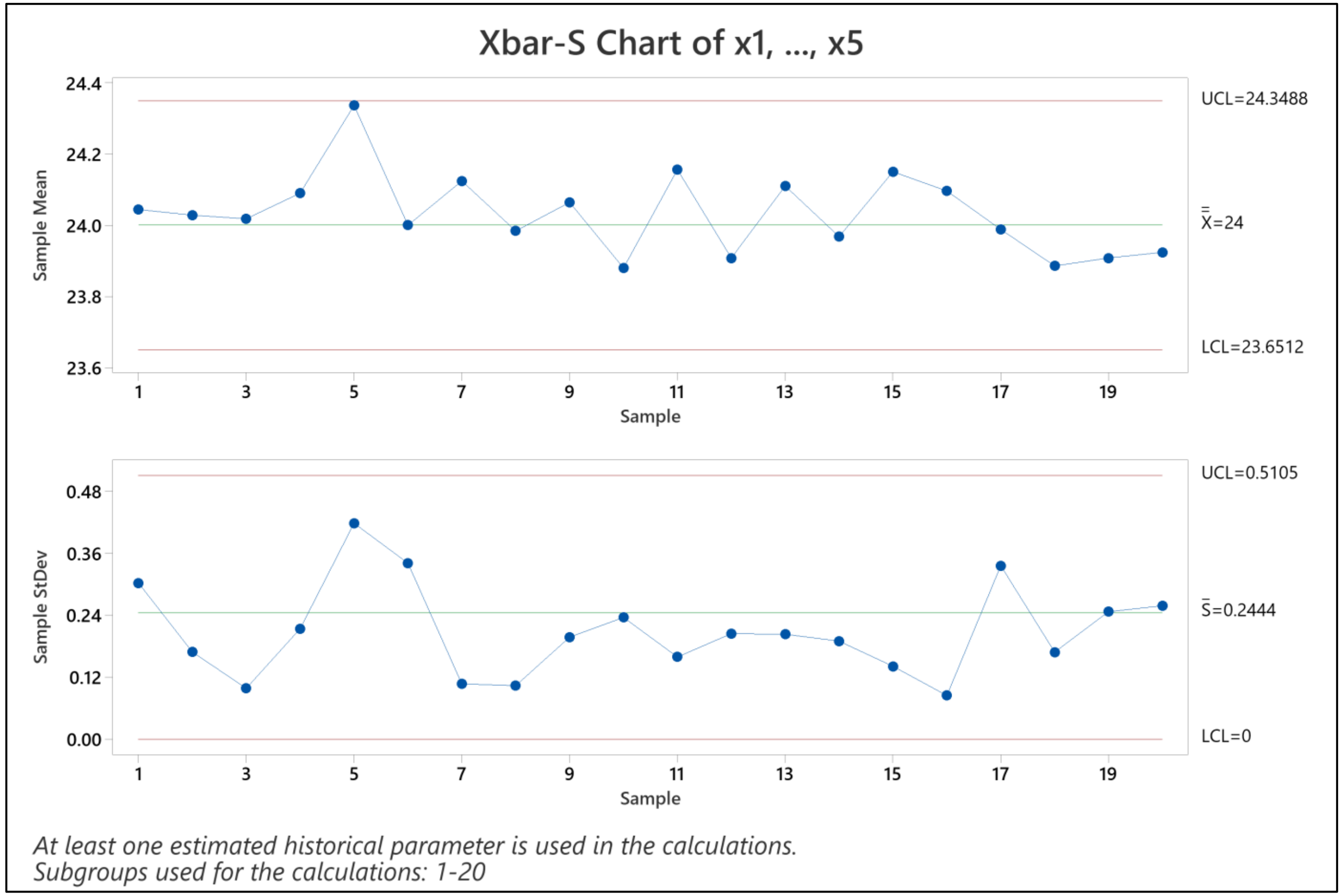
## 1) Let's design the Xbar-S chart (Phase I):

Let's build the Xbar-R chart in Minitab:

*Stats → Control Charts → Variable Charts for Subgroups → Xbar-S*

*→ Xbar-R Options → Parameters*




Even if one point is borderline in the Xbar chart, no observation is OOC when designing the charts with the known parameters.

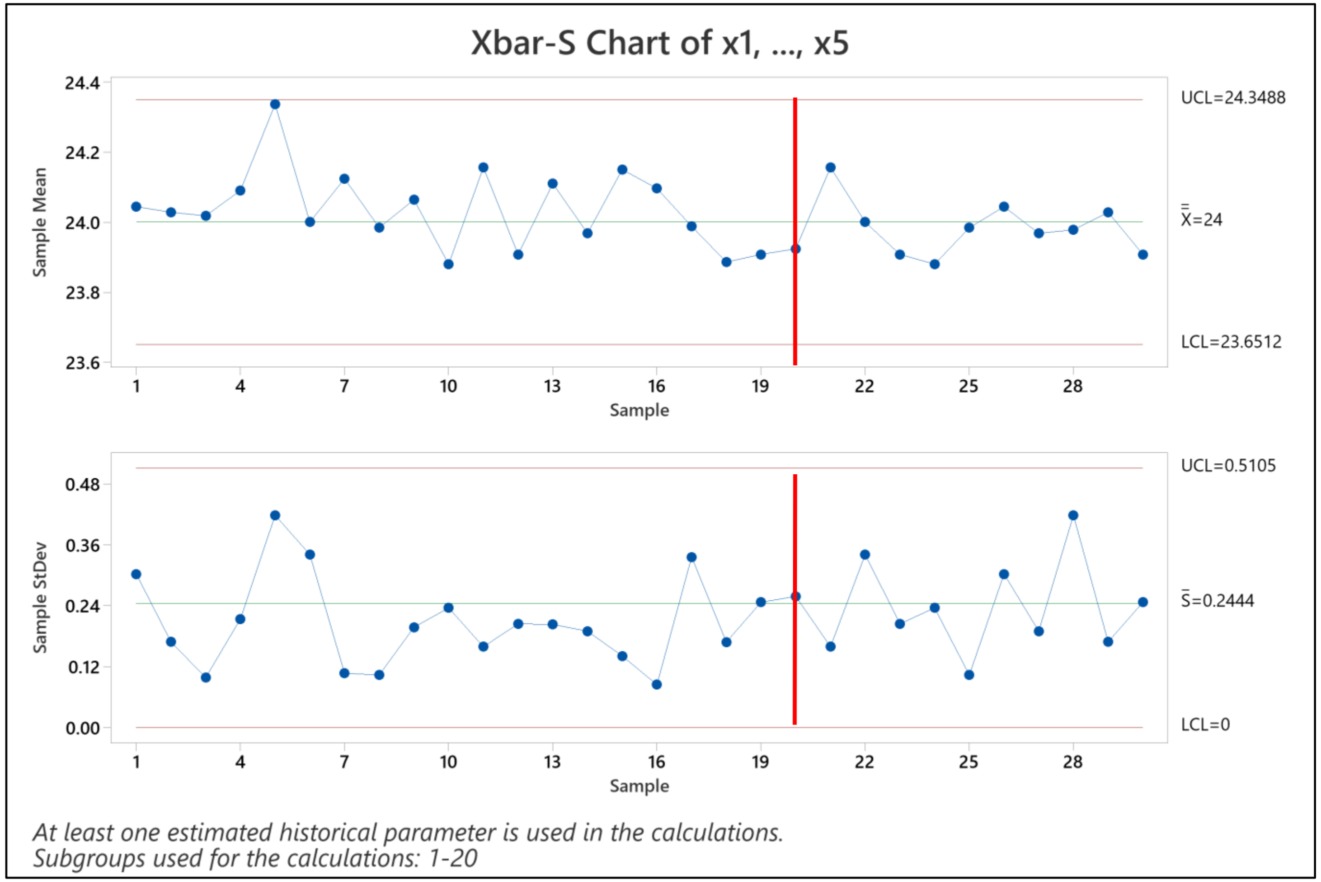
**Exercise 1.8** Knowing that the gear diameter is distributed as a normal distribution with mean 24 mm and standard deviation 0.26 mm, design an Xbar and S chart. For any out-of-control points detected, assume that an assignable cause is found. Check if Phase II data is in control.

## 2) Let's design the Xbar-S chart (Phase II):

To check if Phase II data is in control, append Phase II data to Phase I data. Then, plot the whole dataset using control limits computed on Phase I data:
*Stats → Control Charts → Variable Charts for Subgroups → Xbar-S*

→ *Xbar-R Options → Estimate → Use the subgroups from 1 to 20 to compute control limits*

→ *Xbar-R Options → Parameters → Set the known mean and standard deviation*




The process is in-control.

# Exercise 2

**Data reported in “bpm_phase1.csv” represent daily resting heart measurements (BPM) of a professional athlete, recorded at the same hour each day.**

**1.** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

**2.** Determine if the values reported in “bpm_phase2.csv” are in-control.

**3.** Design an I-MR control chart with probability limits (i.e., use the true distribution of both statistics) alpha=0.01. With regard to the MR chart, use the half-normal distribution.

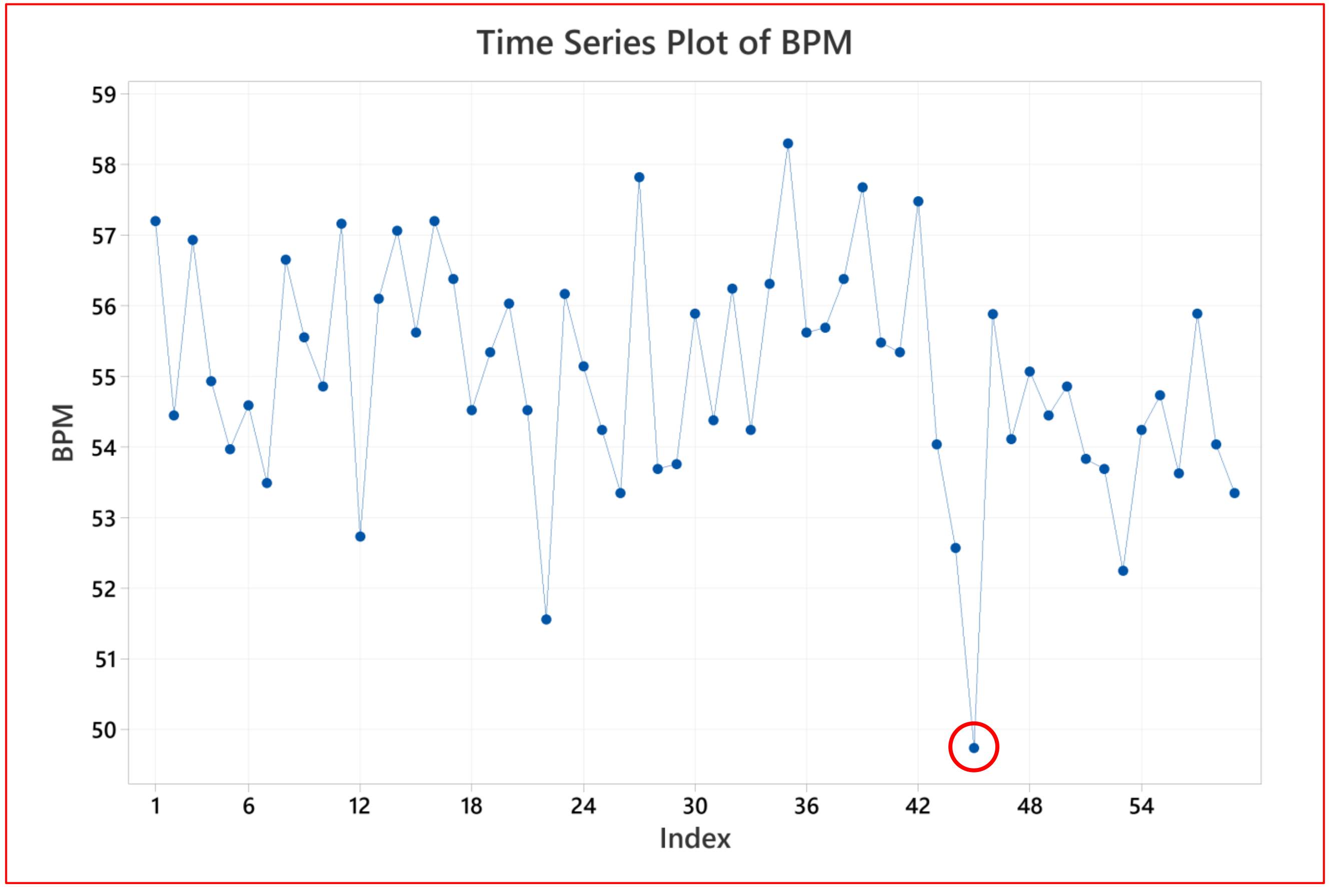
**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

1) **Import the data in Minitab:**

*File → Open → Locate "ESE02_ex2_PhaseI.csv" and Open*

2) **Inspect the data:**

Since we are dealing with a time series, let's plot a time series plot:
*Graph → Time Series Plot → Simple*




Looks like there's one point with a value much lower than the others.
But let's test all assumptions first.

**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

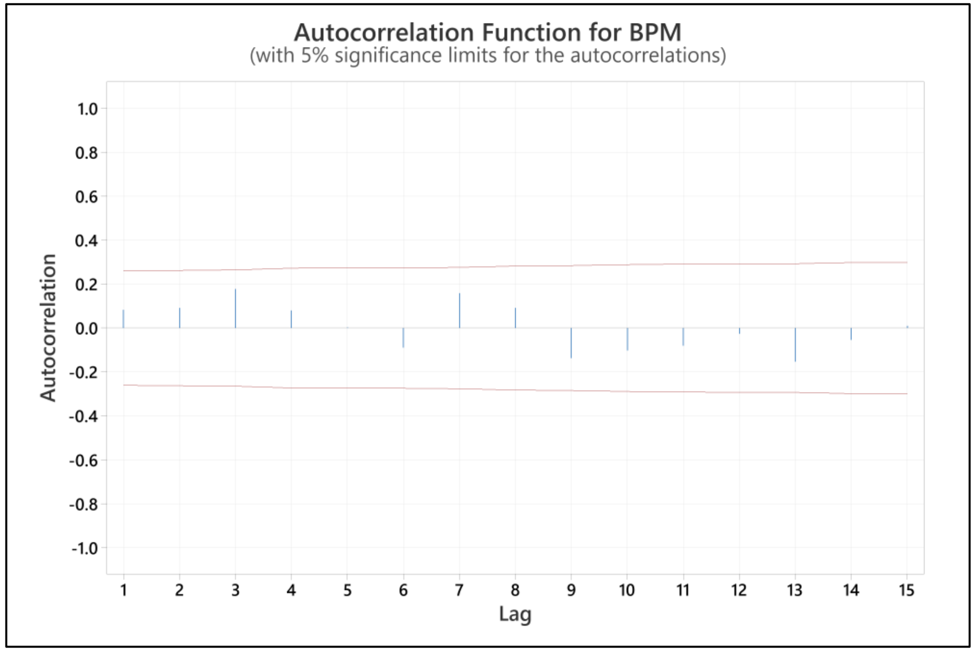
## 3) Test of independence and normality assumptions:

We can check the independence assumption through the runs test, autocorrelation function and partial autocorrelation function:

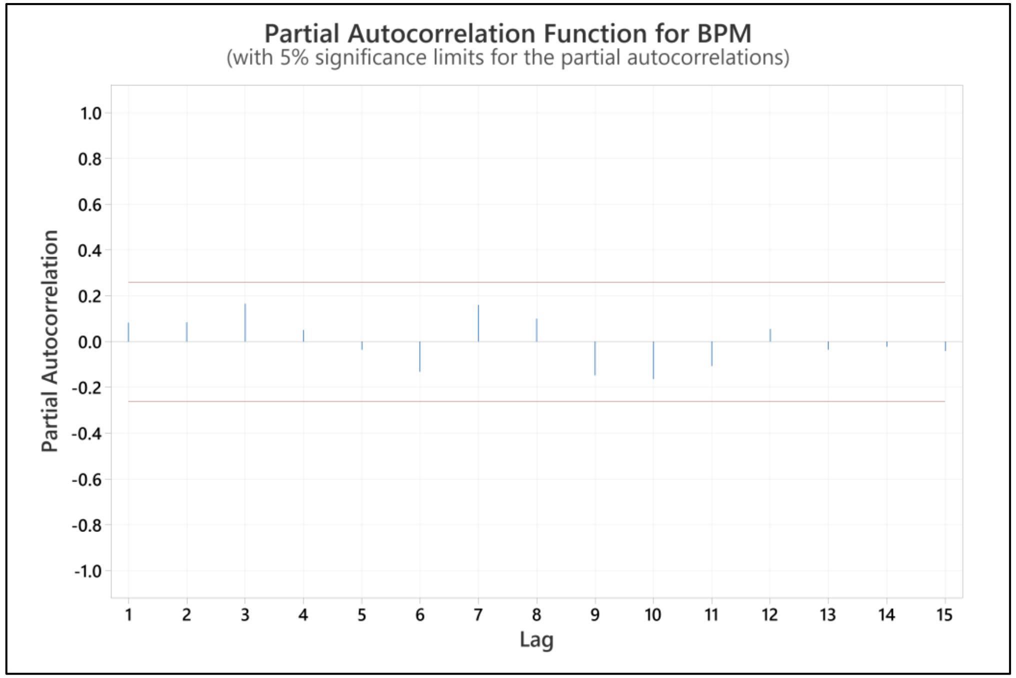
*Stat → Nonparametrics → Runs Test*

| Number of Runs | | |
| --- | --- | --- |
| Observed | Expected | P-Value |
| 28 | 30.49 | $\boxed{0.513}$ |

*Stat → Time Series → Autocorrelation*




*Stat → Time Series → Partial Autocorrelation*



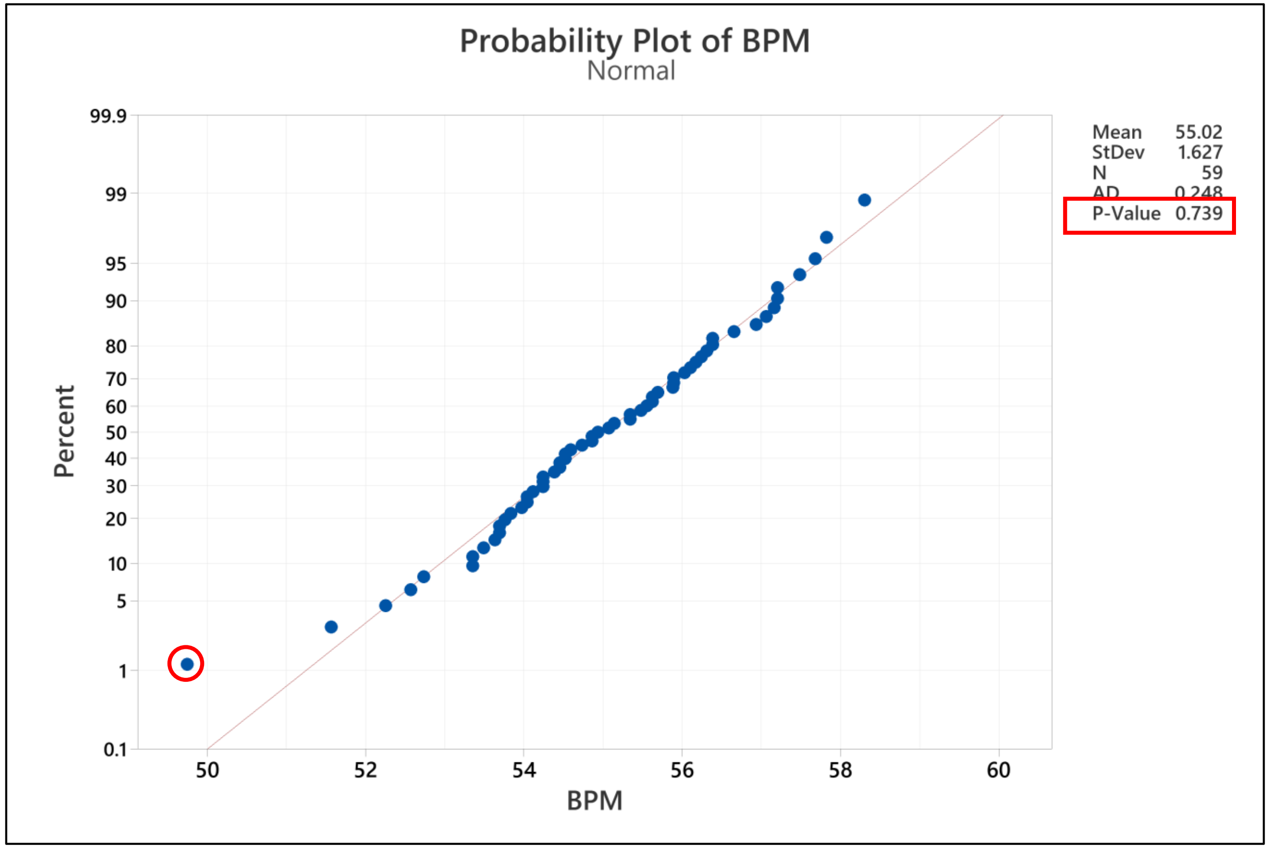

From the results of the autocorrelation and the runs tests, there is no statistical evidence to assume non randomness of the process

**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

## 3) Test of independence and normality assumptions:

We will assess normality through the Anderson-Darling Test and Normal Probability Plot:

*Stat → Basic Statistics → Normality Test*




We cannot reject the null hypothesis that the data are normally distributed with confidence 95%. However, one point deserves attention.

**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

**4) Since we are dealing with individual observations, let's design an I-MR chart:**

**I Chart** (unknown parameters, K=3):

$$UCL = \bar{x} + 3\left(\frac{\bar{MR}}{d_2}\right)$$

$$CL = \bar{x}$$

$$LCL = \bar{x} - 3\left(\frac{\bar{MR}}{d_2}\right)$$

**MR Chart** (unknown parameters, K=3):

$$UCL = D_4\bar{MR}$$

$$CL = \bar{MR}$$

$$LCL = 0$$

$$D_i = X_{i+1} - X_i$$

$$MR_i = |D_i|$$

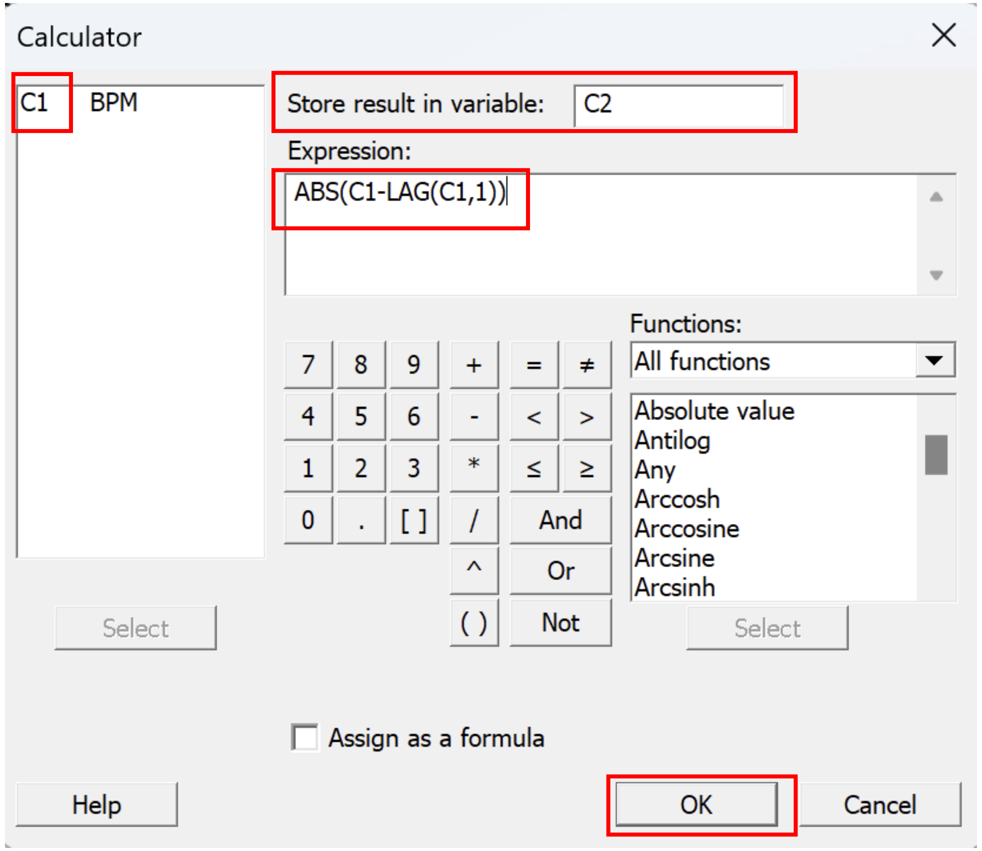
$$\bar{MR} = \frac{\sum MR_i}{n-1}$$

**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

## 4) Since we are dealing with individual observations, let's design an I-MR chart:

To compute the Moving Range in Minitab:

*Calc → Calculator*




ABS(): returns the absolute value of a variable.

LAG(C1, n): return the lag n of variables stored in C1

| BPM | MR |
|---|---|
| 57.20 | * |
| 54.45 | 2.75 |
| 56.93 | 2.48 |
| 54.93 | 2.00 |
| 53.97 | 0.96 |
| 54.59 | 0.62 |
| 53.49 | 1.10 |
| 56.65 | 3.16 |
| 55.55 | 1.10 |
| 54.86 | 0.69 |
| 57.16 | 2.30 |
| 52.73 | 4.43 |
| 56.10 | 3.37 |
| 57.06 | 0.96 |
| 55.62 | 1.44 |
| 57.20 | 1.58 |
| 56.38 | 0.82 |
| 54.52 | 1.86 |
| 55.34 | 0.82 |
| 56.03 | 0.69 |
| 54.52 | 1.51 |
| 51.56 | 2.96 |
| 56.17 | 4.61 |
| 55.14 | 1.03 |
| 54.24 | 0.90 |
| 53.35 | 0.89 |

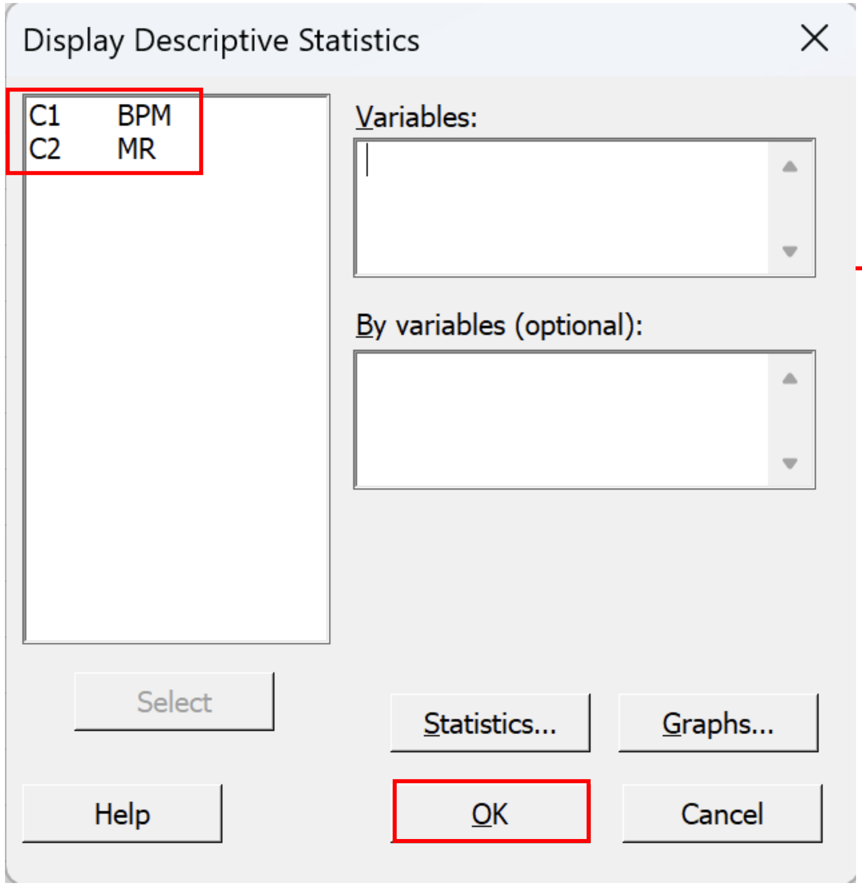
...

**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

## 4) Let's design the I-MR chart (Phase I):

Now we can compute mean (mean of individual observations) and the mean of moving ranges:

*Stat → Basic Statistics → Display Descriptive Statistics*




**Statistics**

| Variable | N | N* | Mean | SE Mean | StDev | Minimum | Q1 | Median | Q3 | Maximum |
|---|---|---|---|---|---|---|---|---|---|---|
| BPM | 59 | 0 | 55.0239 | 0.211794 | 1.62682 | 49.74 | 54.04 | 54.93 | 56.17 | 58.3 |
| MR | 58 | 1 | 1.77776 | 0.165757 | 1.26237 | 0.07 | 0.8725 | 1.51 | 2.27 | 6.14 |

$$\overline{X} = 55.024$$

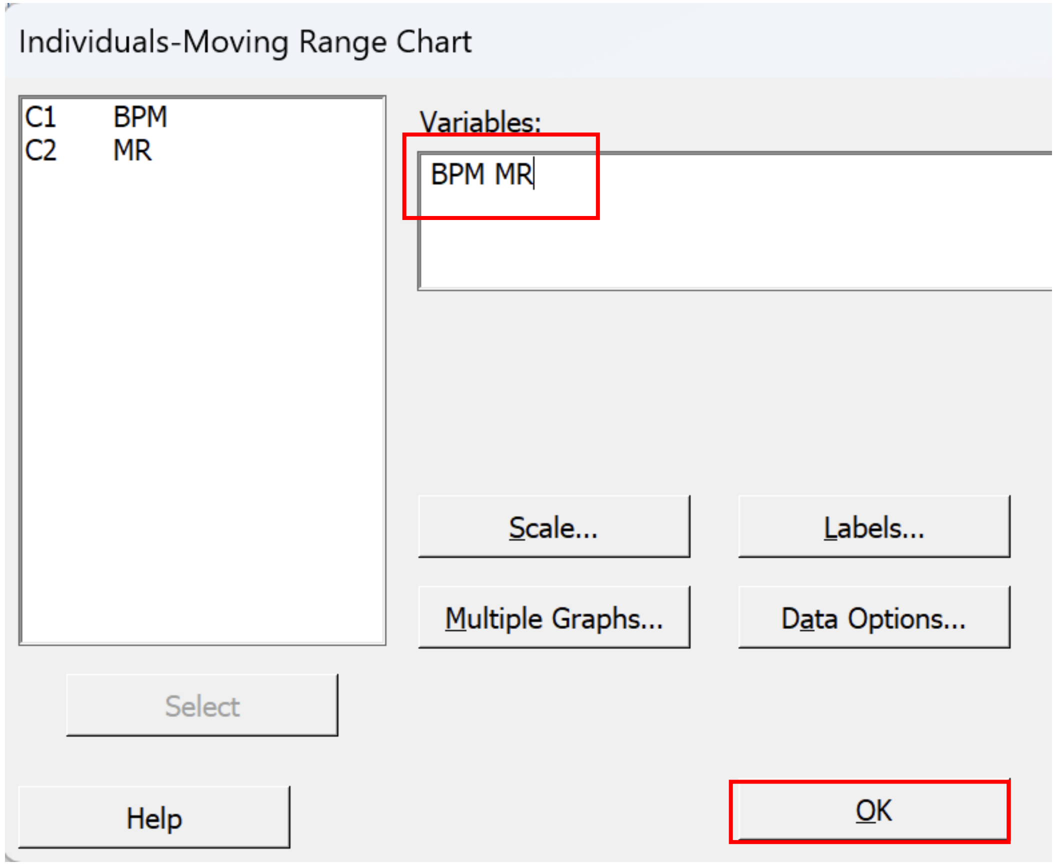
$$\overline{MR} = 1.778$$

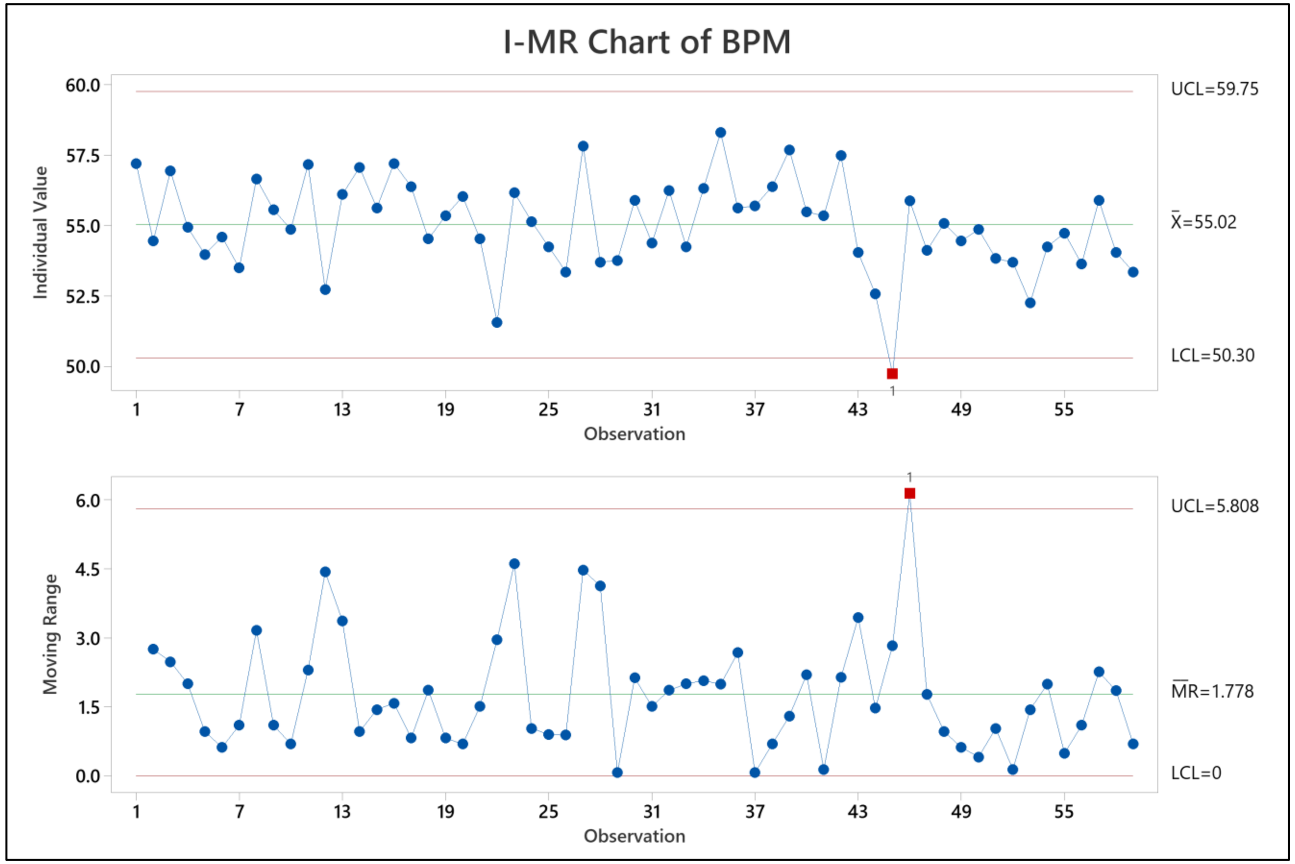
**Exercise 2.1** Design an appropriate control chart to monitor the athlete's physical fitness. No information is available regarding potential causes of any out-of-control observation.

## 4) Let's design the I-MR chart (Phase I):

Let's compute control limits in Minitab:

*Stat → Control Charts → Variables Charts for Individuals*



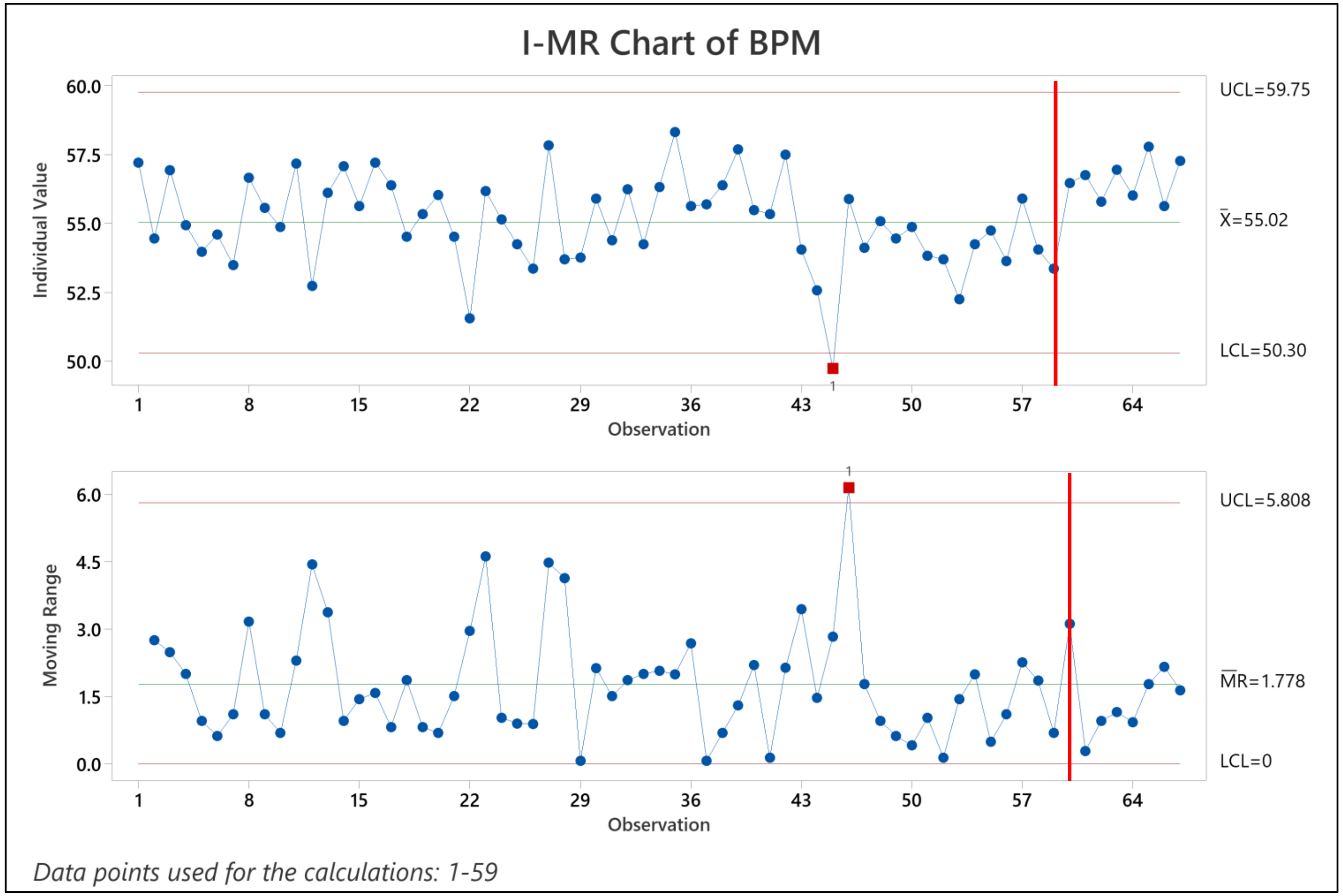

There is one point outside the control limits. Since no information is available regarding the nature of the out-of-control condition, we assume that it is due to the natural variability of the process. Therefore, we de not recompute the control limits.

**Exercise 2.2** Determine if the values reported in “bpm_phase2.csv” are in-control.

## 5) Let’s design the I-MR chart (Phase II):

To check if Phase II data is in control, append Phase II data to Phase I data. Compute again MR values including Phase II. Then, plot the whole dataset using control limits computed on Phase I data:
*Stats → Control Charts → Variable Charts for Individuals → I-MR*

→ *Xbar-R Options → Estimate → Use the values from 1 to 59 to compute control limits*




Phase II data is in control.

**Exercise 2.3** Design an I-MR control chart with probability limits (i.e., use the true distribution of both statistics) alpha=0.01. With regard to the MR chart, use the half-normal distribution.

## 1) Compute the Type I error (α):

$$K = |z_{\alpha/2}|$$

To compute K from the Inverse cumulative normal:
*Calc → Probability Distributions → Normal*

**Normal with mean = 0 and standard deviation = 1**

| P( X ≤ x ) | x |
|---|---|
| 0.005 | -2.57583 |

$$\begin{matrix} UCL \\ LCL \end{matrix} = \bar{x} \pm z_{\alpha/2}\left(\frac{\overline{MR}}{d_2}\right) = 55.0239 \pm 2.57583\frac{1.778}{1.128}$$

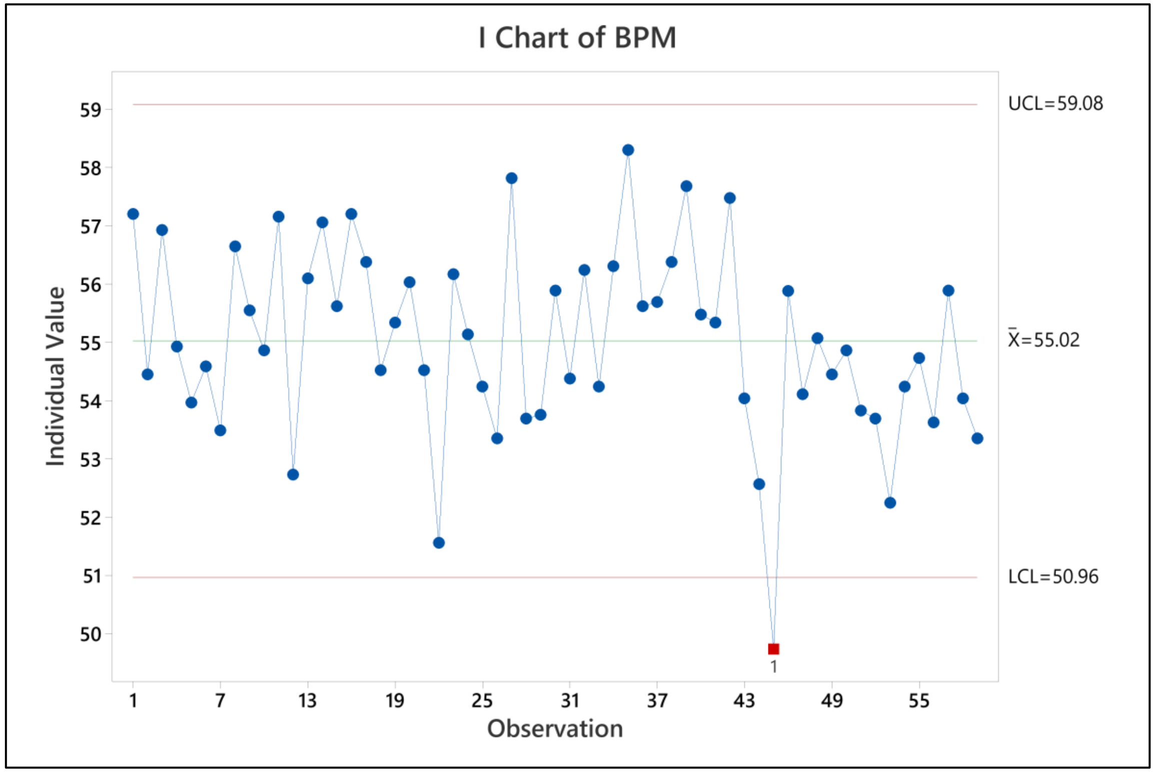
$$UCL = 59.08$$

$$LCL = 50.96$$

## 2) Compute the new I chart (Phase I):

*Stats → Control Charts → Variable Charts for Individuals → Individuals*

*→ I-Chart Options → Limits → Display additional limits at... K = 2.576*



**Exercise 2.3** Design an I-MR control chart with probability limits (i.e., use the true distribution of both statistics) alpha=0.01. With regard to the MR chart, use the half-normal distribution.

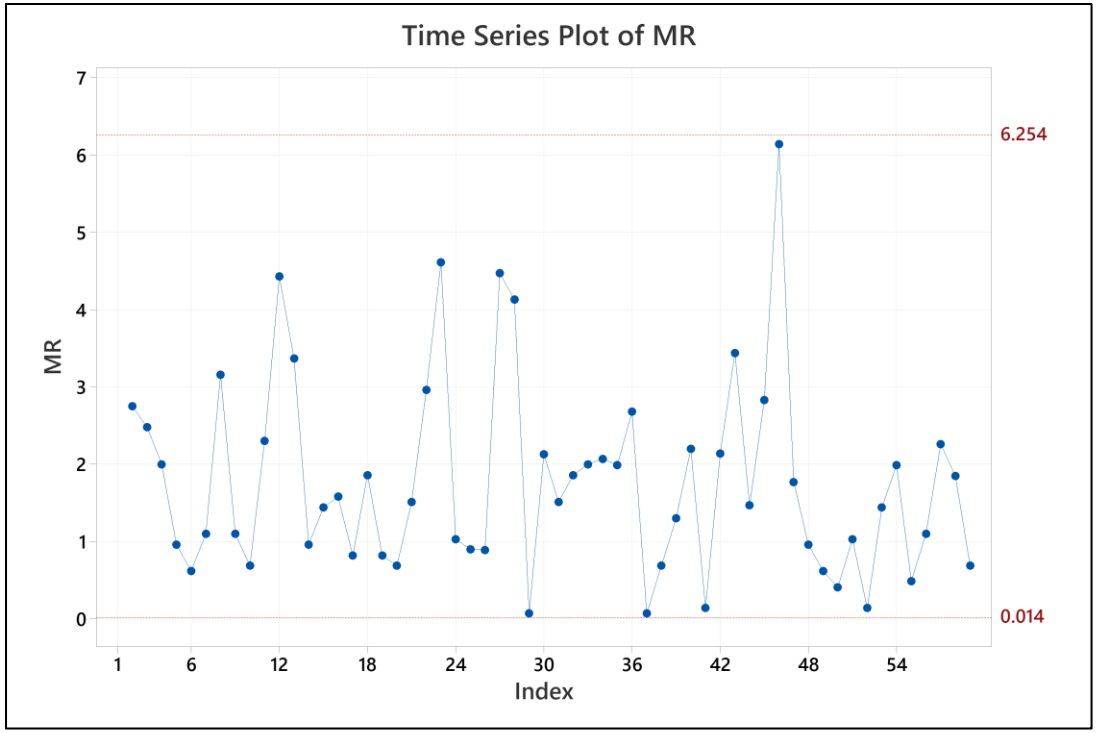
**3) Redesign the MR chart by using the half-normal approximation:**

$$\text{UCL} = D_{1-\alpha/2}\frac{\overline{MR}}{d_2} \qquad \text{LCL} = D_{\alpha/2}\frac{\overline{MR}}{d_2}$$

= 6.254 = 0.014

For n = 2

$$D_{1-\alpha/2} = \sqrt{2}z_{\alpha/4} \qquad D_{\alpha/2} = \sqrt{2}z_{1/2-\alpha/4}$$




$$UCL = \sqrt{2}\frac{2.80703\cdot 1.778}{1.128} = 6.254 \qquad \left|z_{0.0025}\right| = 2.80703$$

$$LCL = \sqrt{2}\frac{0.00627\cdot 1.778}{1.128} = 0.014 \qquad \left|z_{0.4975}\right| = 0.00627$$

Plot the time series of the MR and add two reference lines on Y, one for each control limits.