EXERCISE CLASS 2 – SPC for iid data

Additional Exercices

EXERCISE 1

Data reported below represent the diameters of spheres in a recirculating ball mechanism widely used in machine tool industry. The component is very critical and the tolerance is $10 \pm 0.5mm$:

spheres.csv

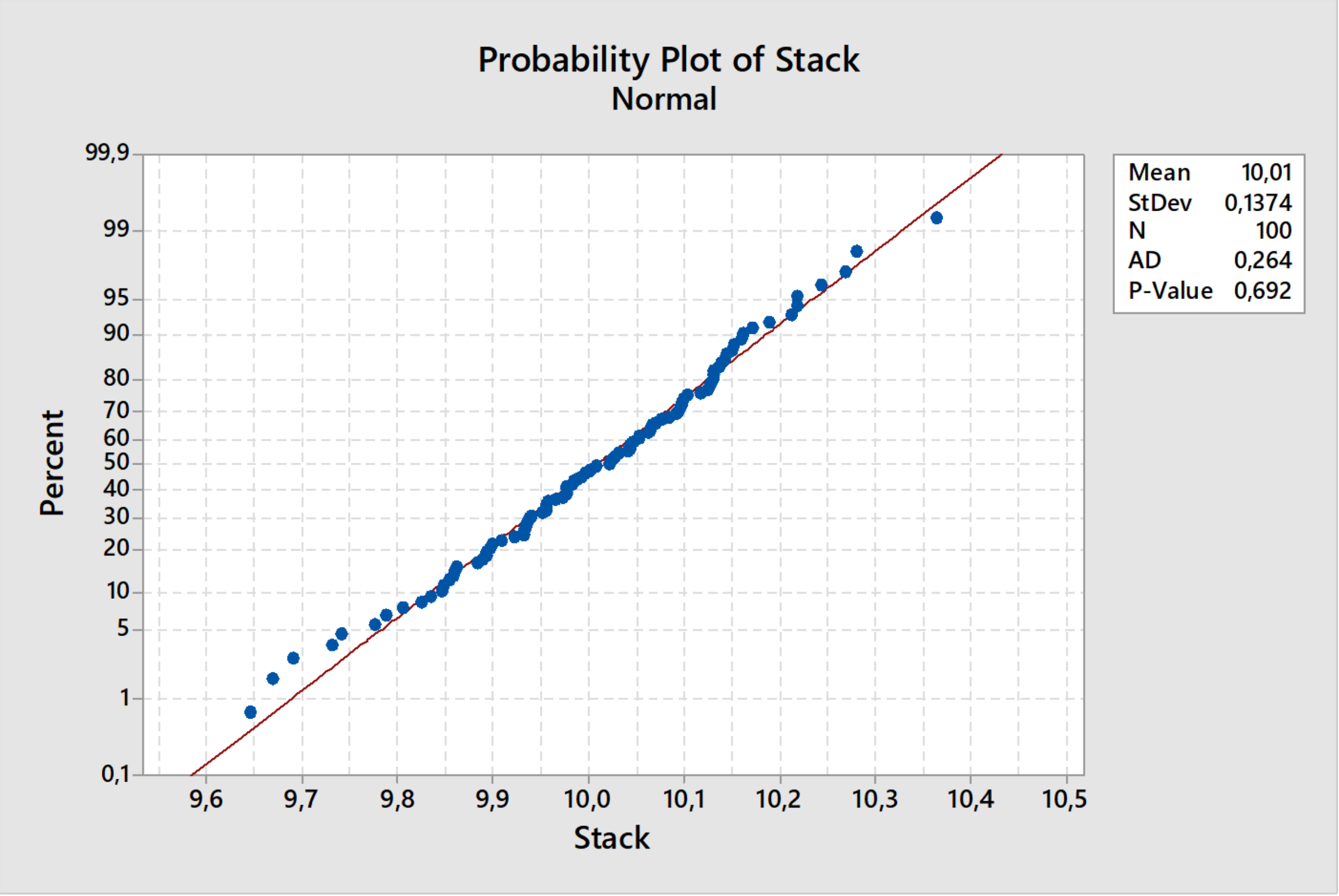
x1	x2	x3	x4	x5
10.10	10.04	10.13	9.98	10.14
10.28	10.13	9.78	9.94	9.83
9.95	10.36	9.90	9.96	9.96
9.88	9.89	9.98	9.74	10.14
10.12	9.83	10.14	9.81	9.86
10.16	10.22	10.08	9.95	9.93
9.65	9.98	10.19	10.03	10.15
9.85	9.86	9.90	10.16	9.94
9.93	10.13	9.89	10.04	10.10
9.79	9.86	9.94	10.02	10.04
9.92	9.98	9.85	9.98	10.17
9.97	10.03	9.89	10.05	10.01
10.06	9.91	10.00	9.69	10.05
9.93	10.27	9.98	10.00	10.10
10.21	10.15	10.05	10.07	9.99
10.00	10.06	10.02	10.13	10.09
10.05	10.10	9.97	10.24	10.13
9.96	10.07	10.03	9.93	10.06
10.22	10.10	10.09	9.67	9.73
9.99	10.13	9.85	10.15	10.08

Verify if the process is in-control by using an $\bar{X} - R$ chart

EXERCISE 1 (SOLUTION)

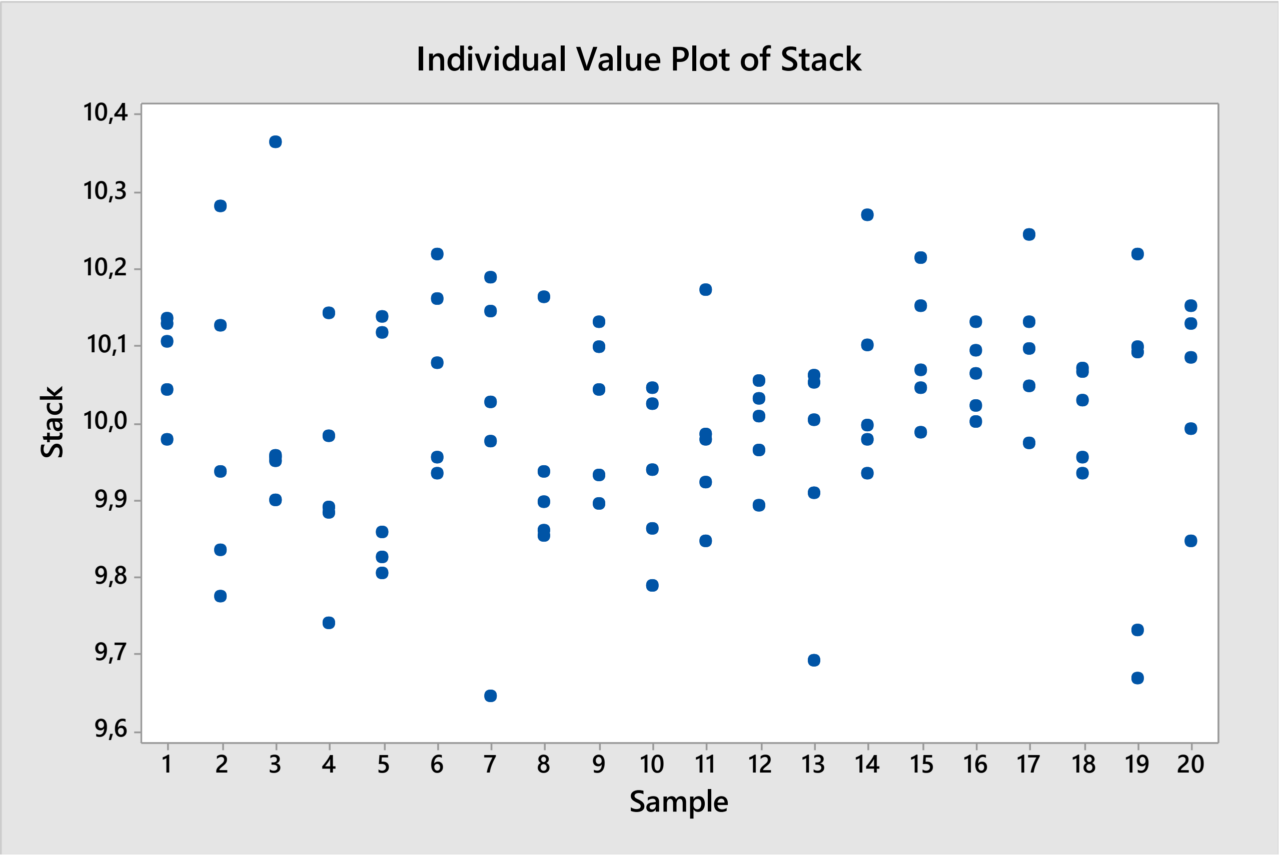
Checking assumptions: NORMALITY

Data-> Stack -> Rows
(all data in a single column; label=sample ID)



EXERCISE 1 (SOLUTION)

a) Data snooping (strange data? Patterns?...)



No strange data

We might also check randomness but we must know within-sample order!

EXERCISE 1 (SOLUTION)

Calc->Row Statistics (sample mean and range)

$$\bar{X} = \frac{1}{n} \sum_{j=1,\ldots,n} X_j \qquad R = \max_j X_j - \min_j X_j$$

Use 'row statistics' command

mean_j	range_j
10.078	0.16
9.992	0.5
10.026	0.46
9.926	0.4
9.952	0.33
10.068	0.29
10	0.54
9.942	0.31
10.018	0.24
9.93	0.25
9.98	0.32
9.99	0.16
9.942	0.37
10.056	0.34
10.094	0.22
10.06	0.13
10.098	0.27
10.01	0.14
9.962	0.55
10.04	0.3

Descriptive Statistics: Xbar; R

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Xbar	10,008	0,0124	0,0555	9,928	9,952	10,003	10,060	10,098
R	0,3149	0,0287	0,1284	0,1301	0,2275	0,3068	0,3939	0,5507

$$\boxed{\bar{\bar{X}} = 10.008 \qquad \bar{R} = 0.3149}$$

EXERCISE 1 (SOLUTION)

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

$\bar{X}$ Control Chart

$$\begin{cases} UCL = \bar{\bar{X}} + A_2(n)\bar{R} \\ CL = \bar{\bar{X}} \\ LCL = \bar{\bar{X}} - A_2(n)\bar{R} \end{cases}$$

R Control Chart

$$\begin{cases} UCL = D_4(n)\bar{R} \\ CL = \bar{R} \\ LCL = D_3(n)\bar{R} \end{cases}$$

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

EXERCISE 1 (SOLUTION)

Factors for contructing variable control charts

	Chart for Averages			Chart for Standard Deviations						Chart for Ranges						
Observations in	Factors for Control Limits			Factors for Center Line		Factors for Control Limits				Factors for Center Line		Factors for Control Limits				
Sample, n	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 $n > 25$.

EXERCISE 1 (SOLUTION)

$\bar{X}$ Control Chart

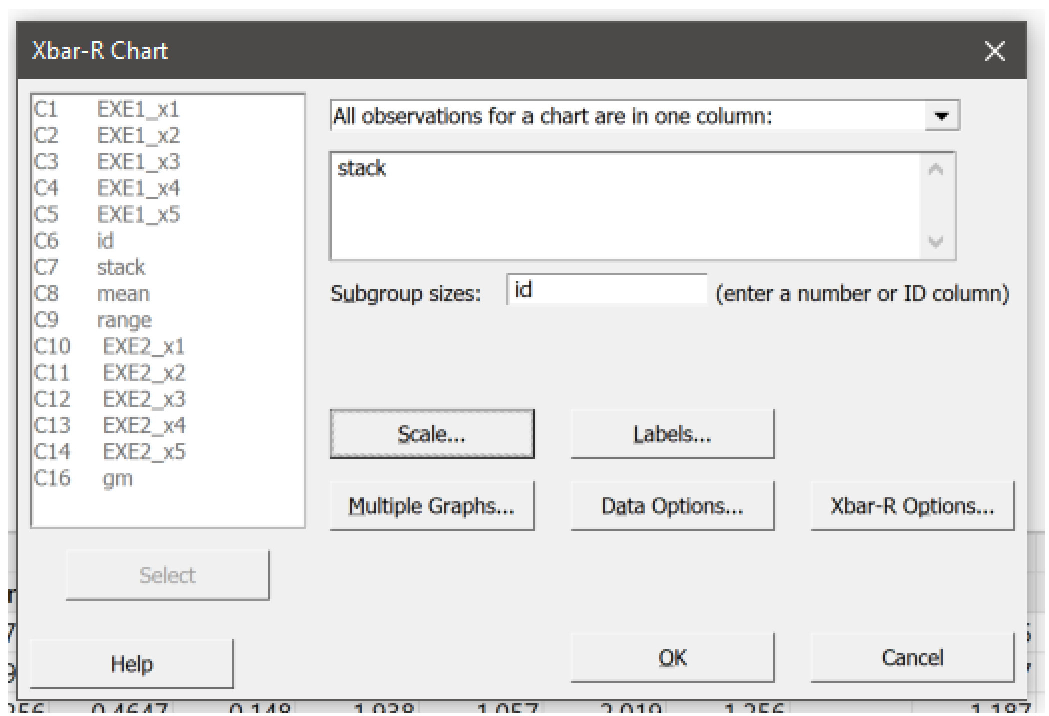
$$\begin{cases} UCL = \bar{\bar{X}} + A_2(n)\bar{R} \\ CL = \bar{\bar{X}} \\ LCL = \bar{\bar{X}} - A_2(n)\bar{R} \end{cases}$$

R Control Chart

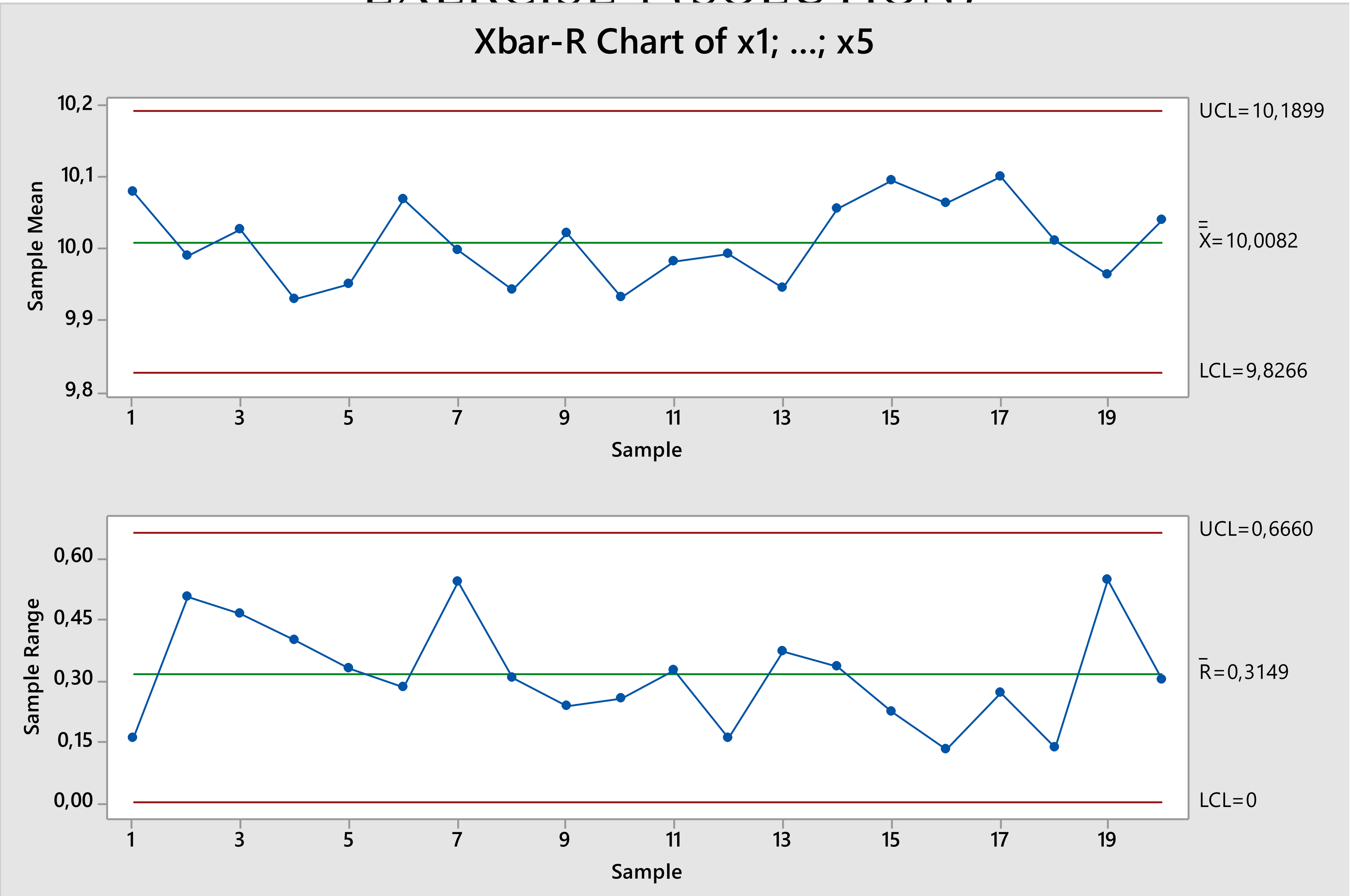
$$\begin{cases} UCL = D_4(n)\bar{R} \\ CL = \bar{R} \\ LCL = D_3(n)\bar{R} \end{cases}$$

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

Stat → Control charts → Variable control charts for subgroups



EXERCISE 1 (SOLUTION)



EXERCISE 1 (Follow on)

Given the previous dataset:

a. Redesign the $\bar{X} - R$ control chart in order to achieve in both the charts a Type I error equal to 0.002 (assuming that the normal approximation applies for both of them)

b. Determine the operating characteristic curve (OC) for the $\bar{X}$ chart (by using K=3 and expressing the shift of the mean in standard deviation units)

c. Determine the corresponding ARL curve

d. Estimate the standard deviation through the statistic R

e. Design (with Minitab) the confidence interval on the process mean that corresponds to the control limits computed in point a).

EXERCISE 1 (Follow on) (SOLUTION)

a)

$\alpha = 0.002 \Rightarrow z_{0.002/2} = 3.09$

$$\mathrm{UCL} = \bar{\bar{x}} + z_{\alpha/2}\frac{1}{d_2\sqrt{n}}\bar{R} = 10.008 + 3.09\left(\frac{0.3140}{2.326\sqrt{5}}\right) = 10.195$$

$$\mathrm{CL} = \bar{\bar{x}} = 10.008$$

$$\mathrm{LCL} = \bar{\bar{x}} - z_{\alpha/2}\frac{1}{d_2\sqrt{n}}\bar{R} = 10.008 - 3.09\left(\frac{0.3140}{2.326\sqrt{5}}\right) = 9.821$$

$$\boxed{\begin{array}{l} n = 5 \\ d_2(n) = 2.326 \\ d_3(n) = 0.864 \end{array}}$$

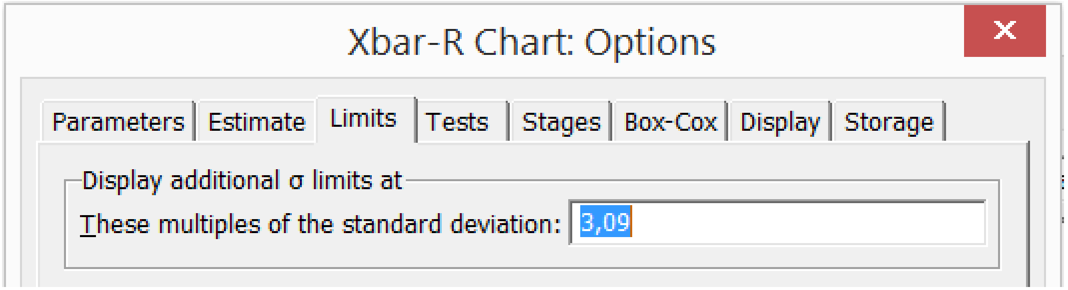
$$\mathrm{UCL} = d_2(n)\hat{\sigma} + z_{\alpha/2}d_3(n)\hat{\sigma} = \bar{R} + z_{\alpha/2}\frac{d_3(n)}{d_2(n)}\bar{R} = 0.674$$

$$\mathrm{CL} = d_2(n)\hat{\sigma} = \bar{R} = 0.3140$$

$$\mathrm{LCL} = MAX\{0; d_2(n)\hat{\sigma} - z_{\alpha/2}d_3(n)\hat{\sigma}\} = MAX\left\{0; \bar{R} - z_{\alpha/2}\frac{d_3(n)}{d_2(n)}\bar{R}\right\} = 0$$

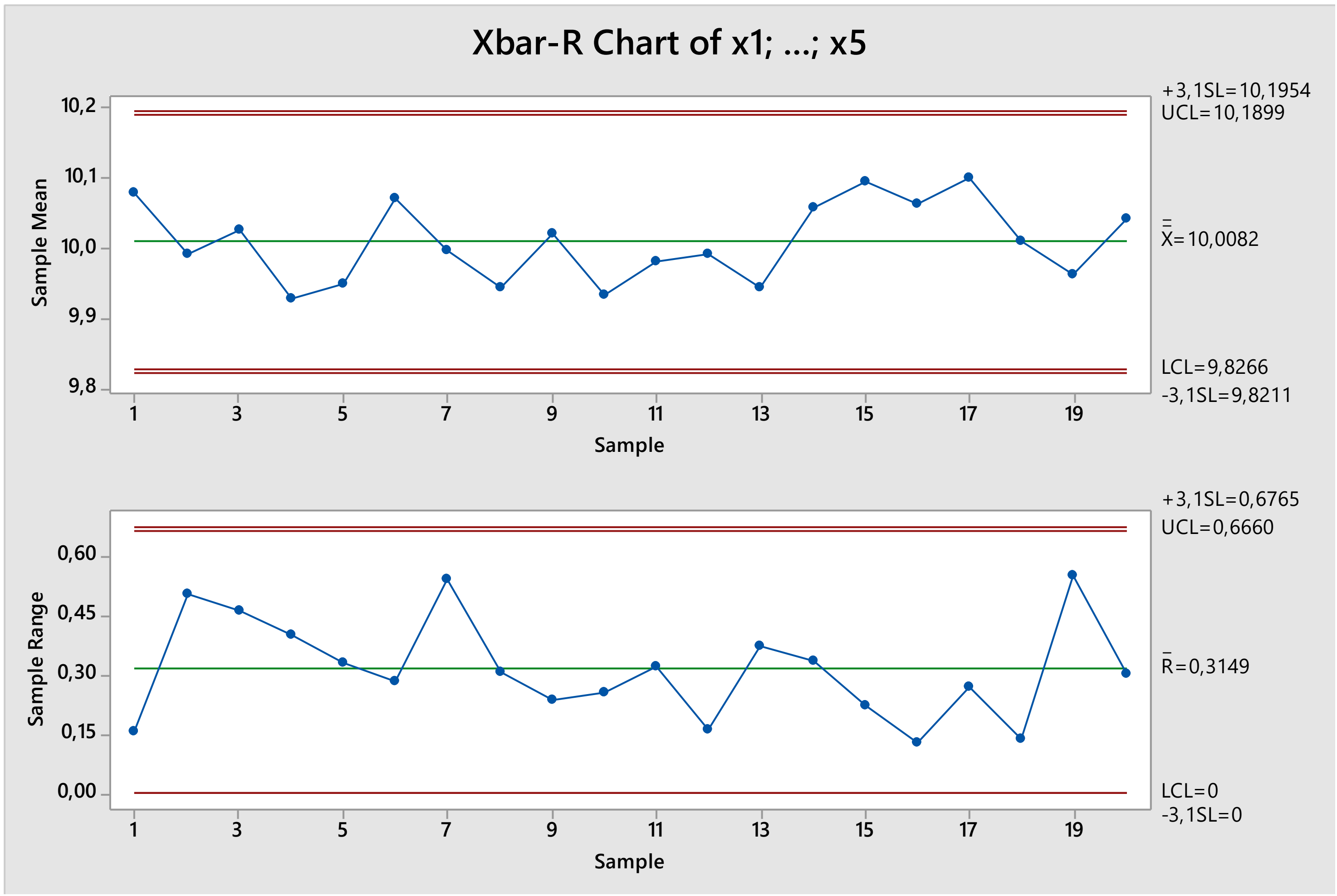
EXERCISE 1 (Follow on) (SOLUTION)

a)



When using K ≠ 3, Minitab shows both limits at desired K and the ones with K = 3

Just ignore the ones at K = 3



EXERCISE 1 (Follow on) (SOLUTION)

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

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

$$\beta = \Pr(LCL \le \bar{X} \le UCL \mid H_1)$$

$$= \Pr\left(Z \le \frac{UCL - \mu_1}{\sigma / \sqrt{n}}\right) - \Pr\left(Z \le \frac{LCL - \mu_1}{\sigma / \sqrt{n}}\right) =$$

$$\boxed{= \Pr\left(Z \le 3 - \delta\sqrt{n}\right) - \Pr\left(Z \le -3 - \delta\sqrt{n}\right)}$$

where $\delta = \dfrac{\mu_1 - \mu_0}{\sigma}$

Create a 'delta' column.

Create two columns of values:

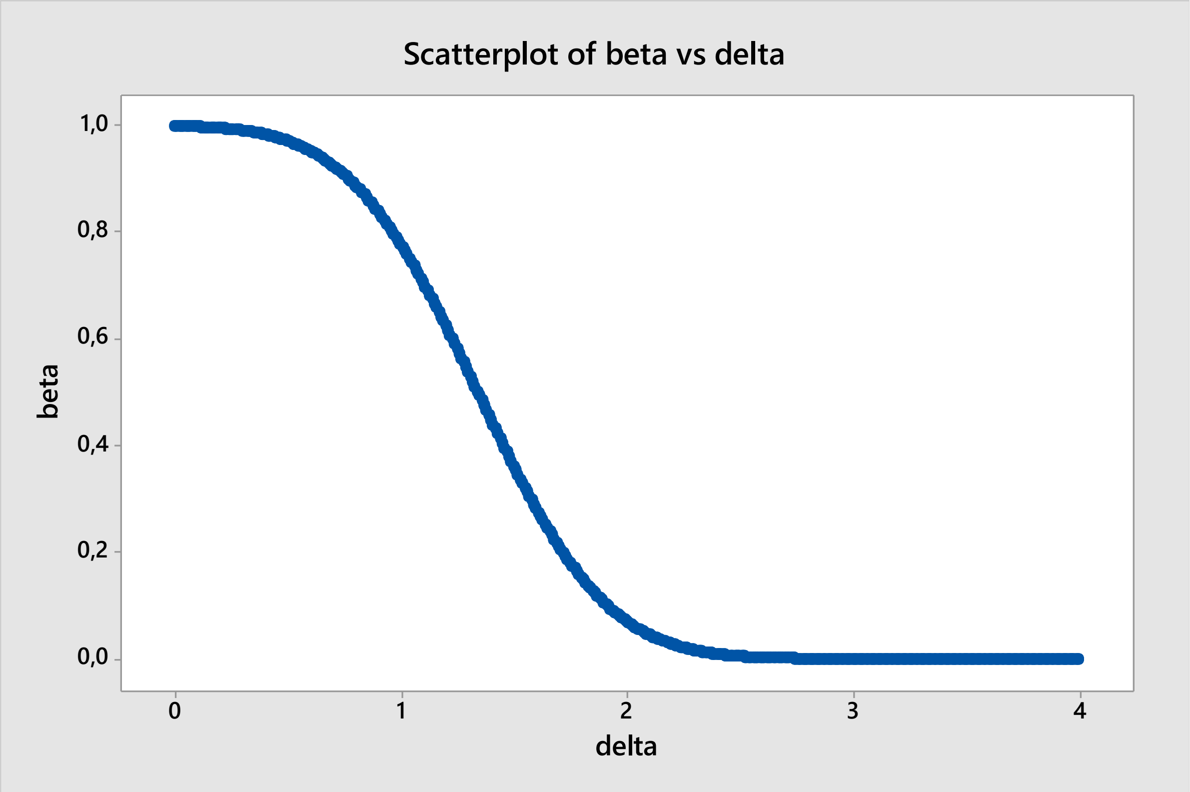
- $3 - \delta\sqrt{n}$
- $-3 - \delta\sqrt{n}$

Create two columns of values:

- $P(Z \le 3 - \delta\sqrt{n})$
- $P(Z \le -3 - \delta\sqrt{n})$

Create a column for β

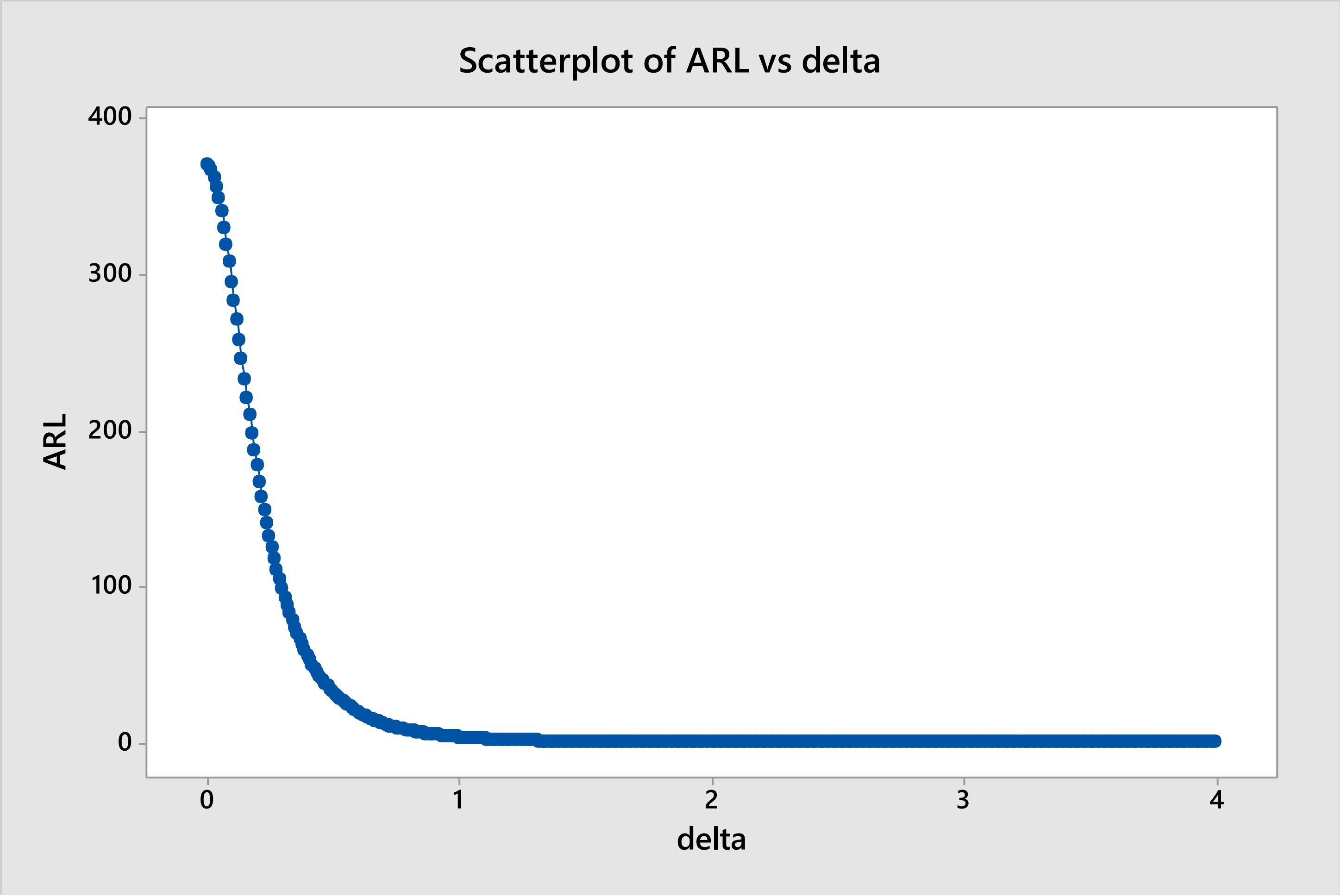
Plot β vs δ



EXERCISE 1 (Follow on) (SOLUTION)

c)

$$ARL = 1/(1-\beta)$$



EXERCISE 1 (Follow on) (SOLUTION)

d)

Estimate the standard deviation through the statistic R

$$\bar{R} = 0.3140$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2(n)} = \frac{0.314}{2.326} = 0.13499$$

EXERCISE 1 (Follow on) (SOLUTION)

e) Confidence interval at 99,8% (REMIND: Type I error = 0.002)

With:

$n = 5$

$\hat{\mu} = \bar{\bar{X}} = 10.008$

$$\hat{\sigma} = \frac{\bar{R}}{d_2(n)} = \frac{0.3140}{2.326} = 0.13499$$

CI for the means of each variable (i.e., column means)

One-Sample Z

(assumed known variance)

```
The assumed standard deviation = 0,13499

N      Mean  SE Mean        99,8% CI
5  10,0080   0,0604  (9,8214; 10,1946)
```

Control limits: $LCL = 9.821, UCL = 10.195$

EXERCISE 2

The following table (read from the left to the right and then top-down) shows the daily changes of the General Motors Co. closing prices since September 4, 1998 to November 27, 1998

general_motors.csv

-0.875	2.437	-1.187	-2.063	0.938	3.000	-0.563	-0.062	-1.375	-0.375
-0.125	0.062	1.750	-1.500	1.250	1.000	-1.500	-3.125	-0.687	0.812
-1.250	-2.500	0.125	-1.125	0.625	1.063	-0.875	1.125	5.875	2.875
0.812	1.875	-0.687	0.562	-1.187	0.437	1.188	0.500	-1.250	2.000
2.000	-0.438	2.250	-0.250	0.313	-1.063	1.250	0.563	-0.813	3.250
2.563	0.312	-0.437	-0.125	-0.688	1.500	-0.500	-0.937	-0.500	

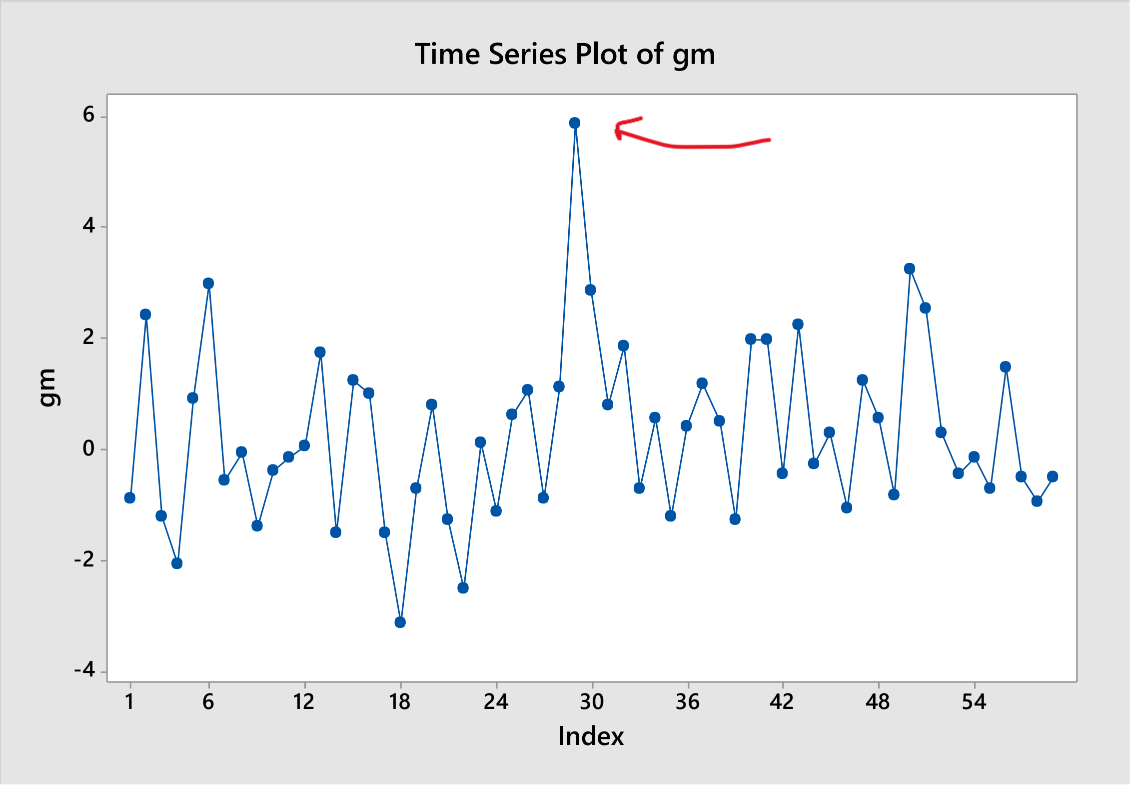
a) Design a suitable quality control tool by assuming the existence of an assignable cause for the OOC observations (if any)

b) Determine if the following values are IC (use the previously designed control chart – point a))

1.327 1.594 0.716 1.767 -0.915 -2.524 -0.563 2.053

EXERCISE 2 (SOLUTION)

a) DATA SNOOPING and Hypothesis Tests



Descriptive Statistics: gm

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
gm	0,292	0,210	1,615	-3,125	-0,875	0,062	1,188	5,875

EXERCISE 2 (SOLUTION)

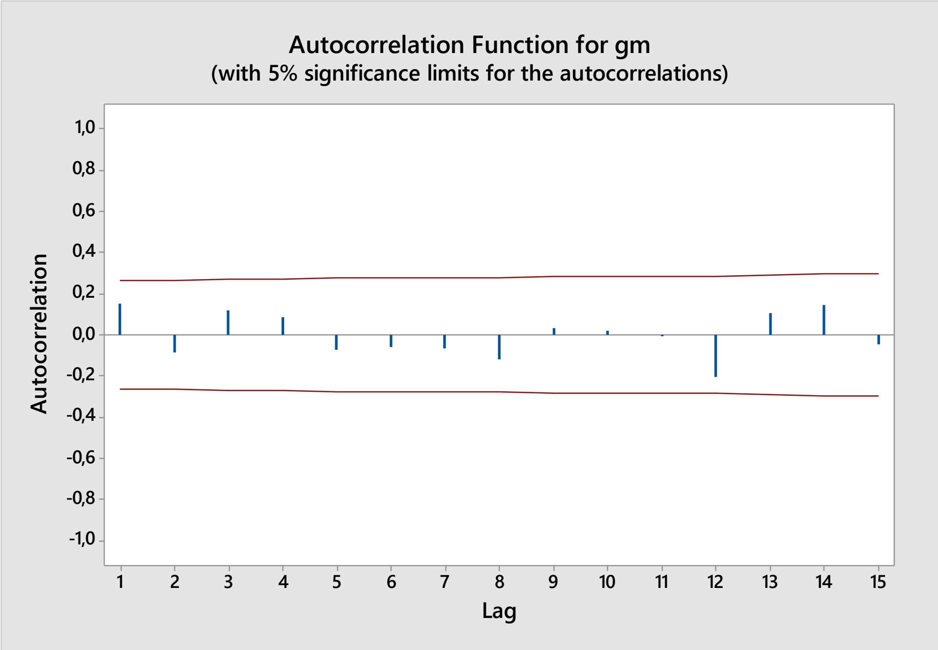
a) DATA SNOOPING and Hypothesis Tests

Test

Null hypothesis	H_0: The order of the data is random
Alternative hypothesis	H_1: The order of the data is not random

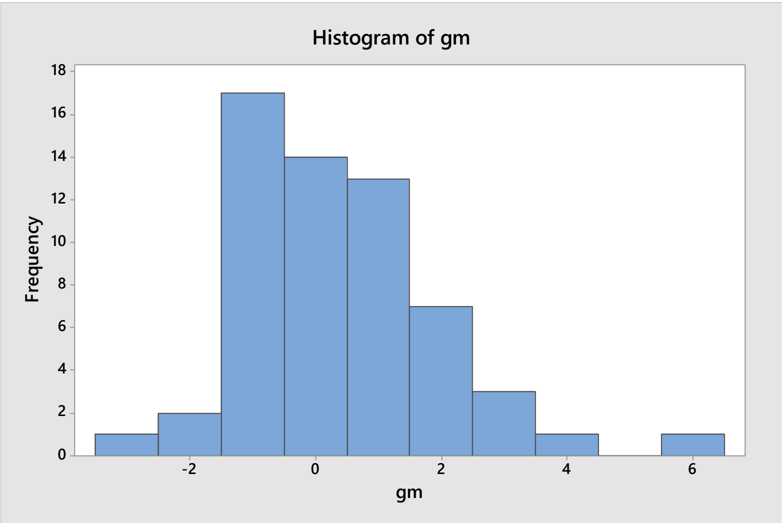
Number of Runs		
Observed	Expected	P-Value
31	30,42	0,879

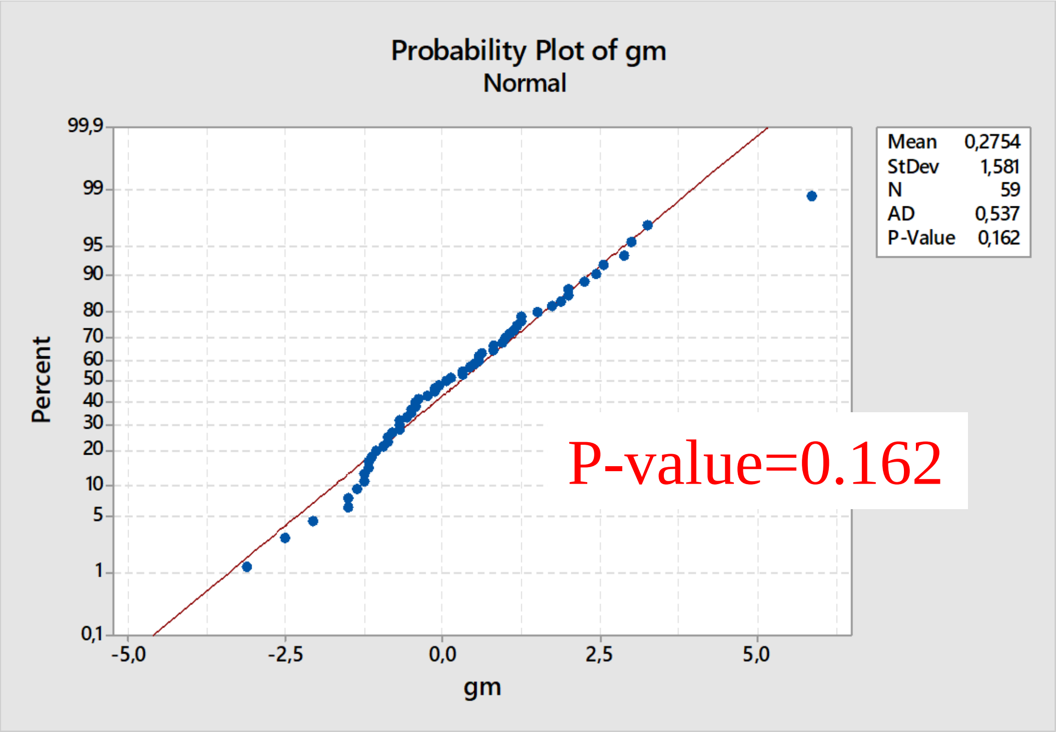
There is no statistical evidence to assume non-randomness of the process



EXERCISE 2 (SOLUTION)

a) DATA SNOOPING and Hypothesis Tests





There is no statistical evidence to assume non-normality of process data

EXERCISE 2 (SOLUTION)

a)

Quality control tool: $I - MR$ control chart

Let's compute the moving ranges:
Stat-> Time series-> Differences
Calc->Abs

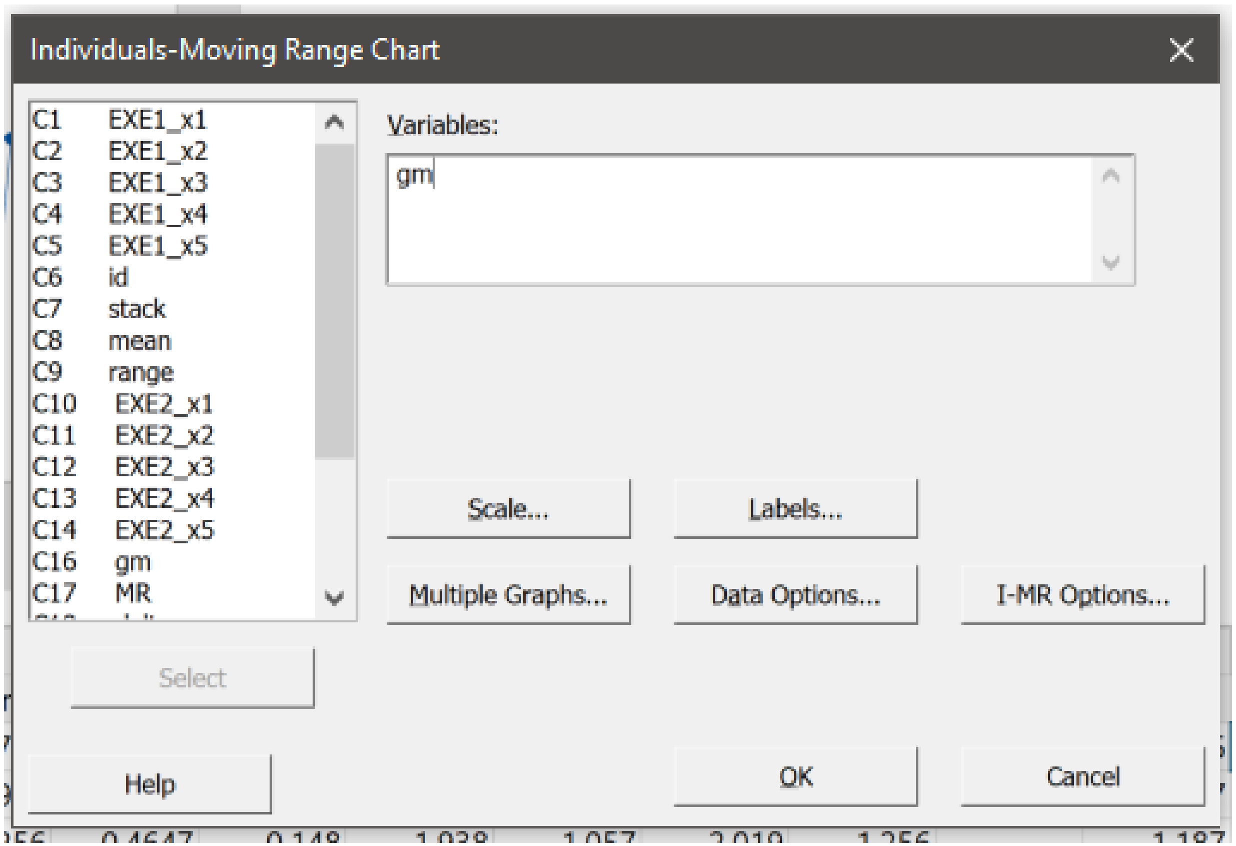
Descriptive Statistics: MR

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
MR	1,744	0,143	1,092	0,000	0,734	1,719	2,500	4,750

$$\overline{MR} = 1.744$$

EXERCISE 2 (SOLUTION)

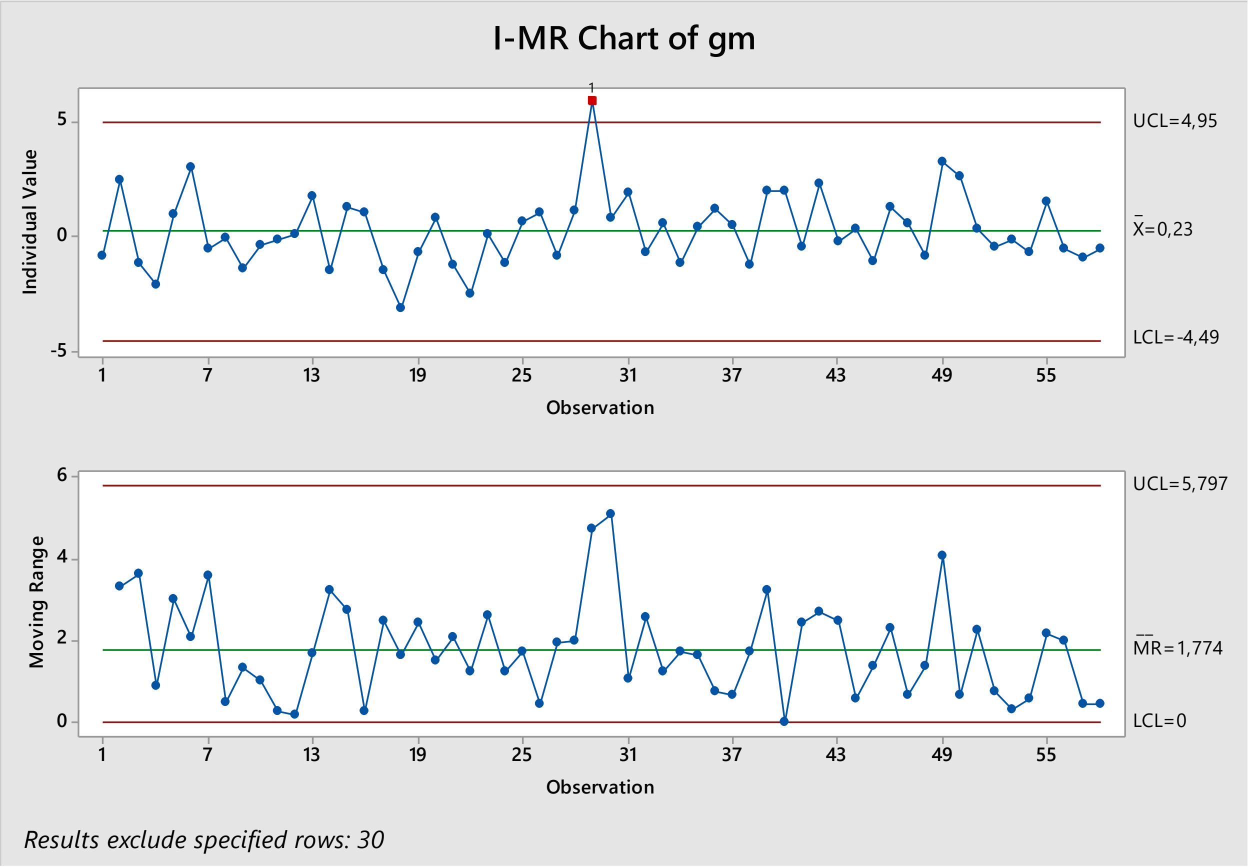
a) Stat → Control charts → Variable control charts for individuals



EXERCISE 2 (SOLUTION)

a)

One OOC observation



Assume that we found an **assignable cause** for observation 29

We have to remove the datum from the dataset

EXERCISE 2 (SOLUTION)

a)

Before removing data #29

Descriptive Statistics: gm

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
gm	0,275	0,206	1,581	-3,125	-0,875	0,062	1,188	5,875

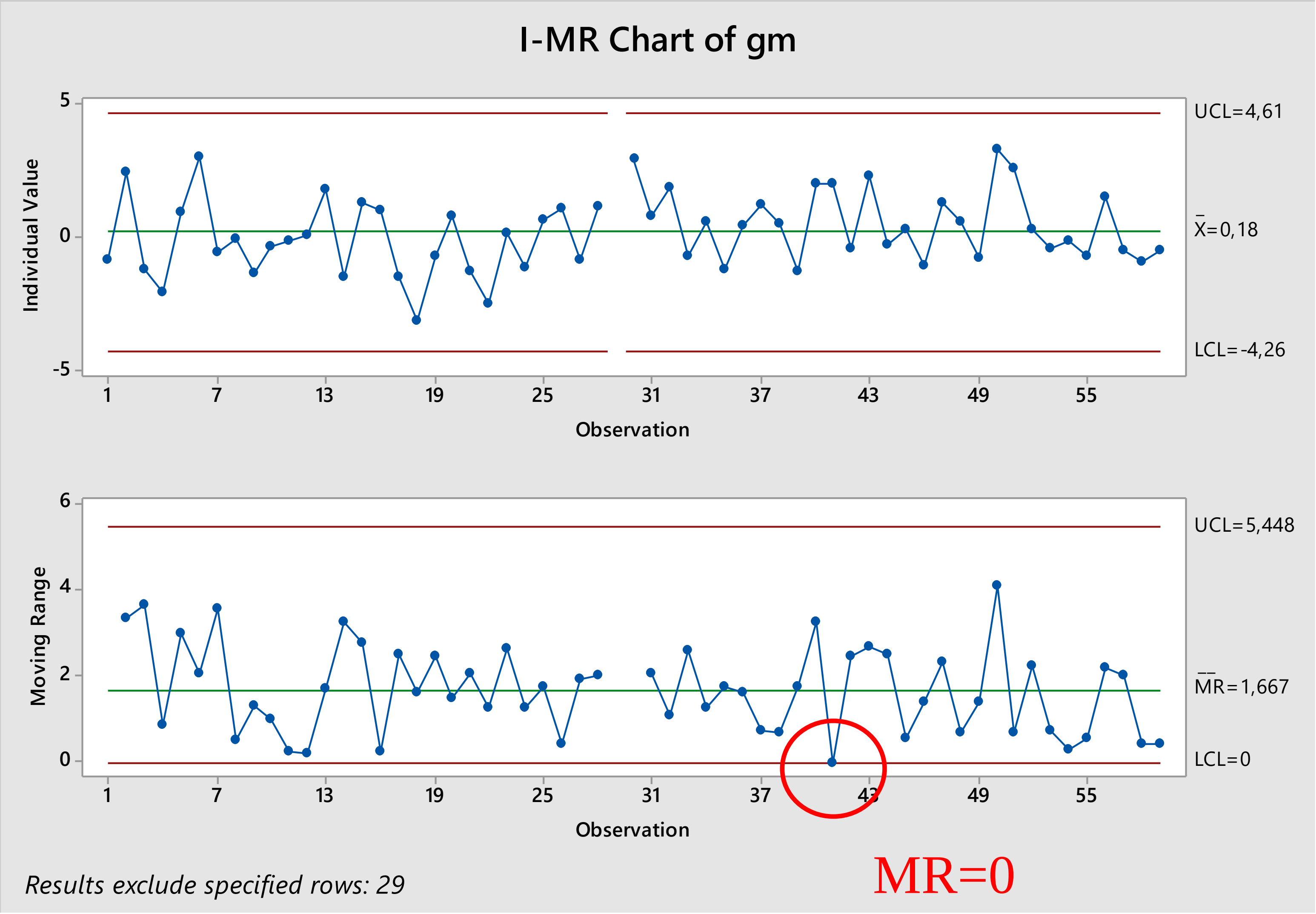
After removing data #29

Descriptive Statistics: gm_1

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
gm_1	0,179	0,185	1,409	-3,125	-0,875	0,000	1,141	3,250

EXERCISE 2 (SOLUTION)

a)



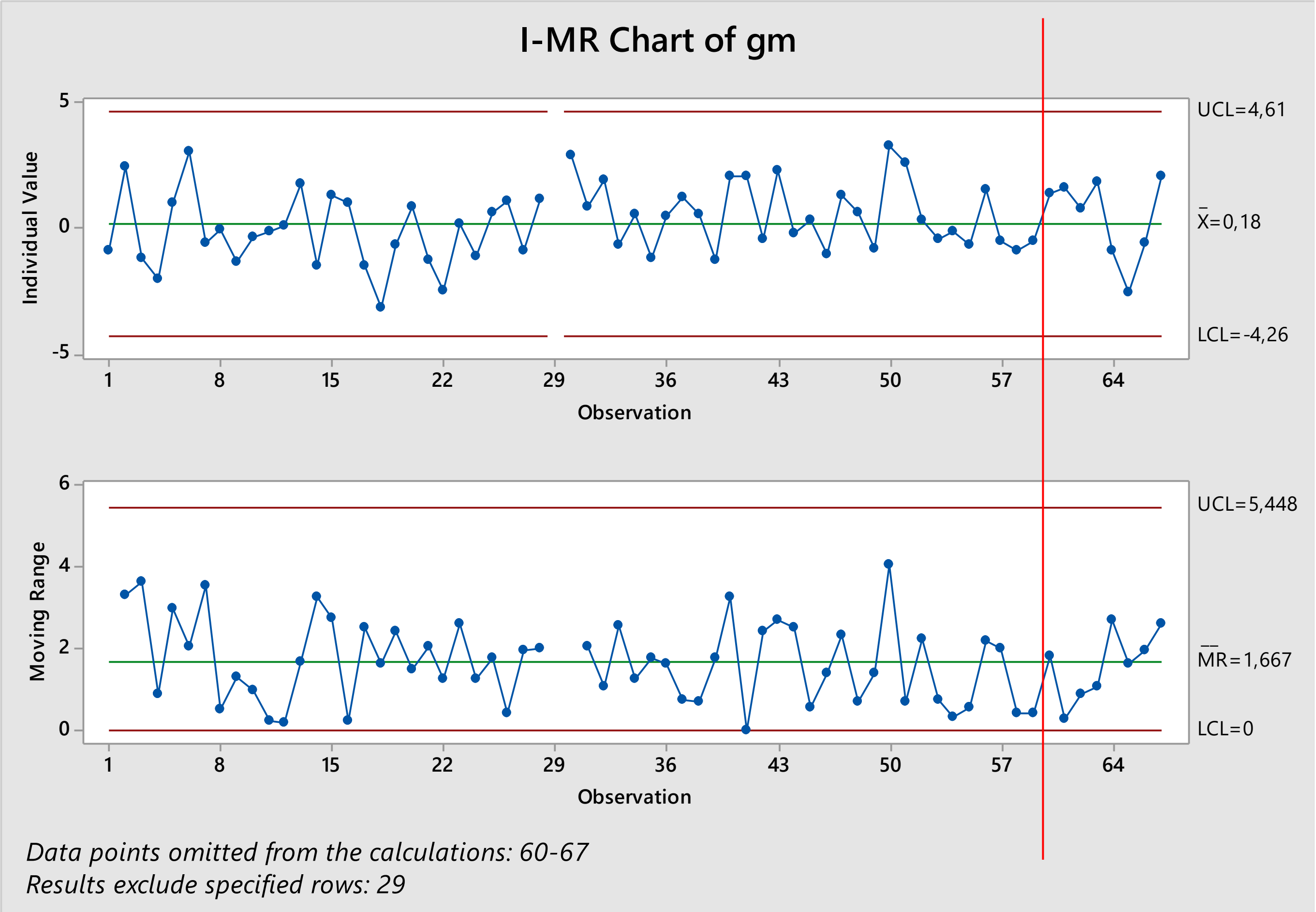
MR=0

EXERCISE 2 (SOLUTION)

b) Phase II

Create a single column that includes both Phase I and Phase II data, and specify the subset of data to be used for control chart design (SEE NEXT SLIDE)

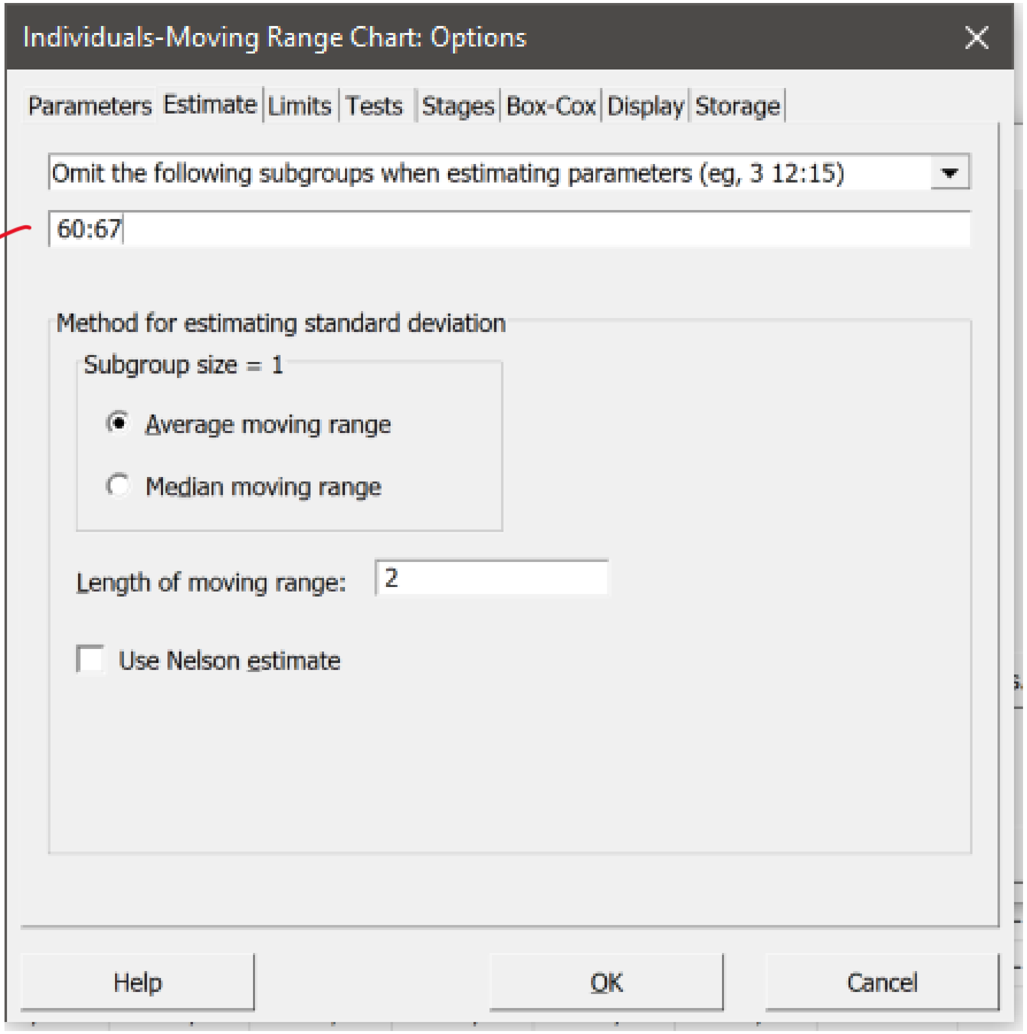
New samples are IC



EXERCISE 2 (SOLUTION)

To exclude Phase II data from the design of the control chart!

Rows corresponding to Phase 2 data



EXERCISE 3

Using data in general_motors.csv, design an I-MR control chart with probability limits (i.e., use the true distribution of both statistics) with $\alpha = 0.01$.

With regard to the MR chart, use the half-normal distribution.

EXERCISE 3 (SOLUTION)

I Control chart

Inverse Cumulative Distribution Function

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

P( X ≤ x )          x
     0,005  -2,57583
```

$$z_{\alpha/2} = 2.57583$$

$$\begin{matrix} UCL \\ LCL \end{matrix} = \overline{x} \pm 2.57583 \left(\frac{\overline{MR}}{d_2} \right)$$

$$= 0.292 \pm 2.57583 \left(\frac{1.744}{1.128} \right)$$

$$\boxed{\begin{matrix} UCL = 4.2745 \\ LCL = -3.6905 \end{matrix}}$$

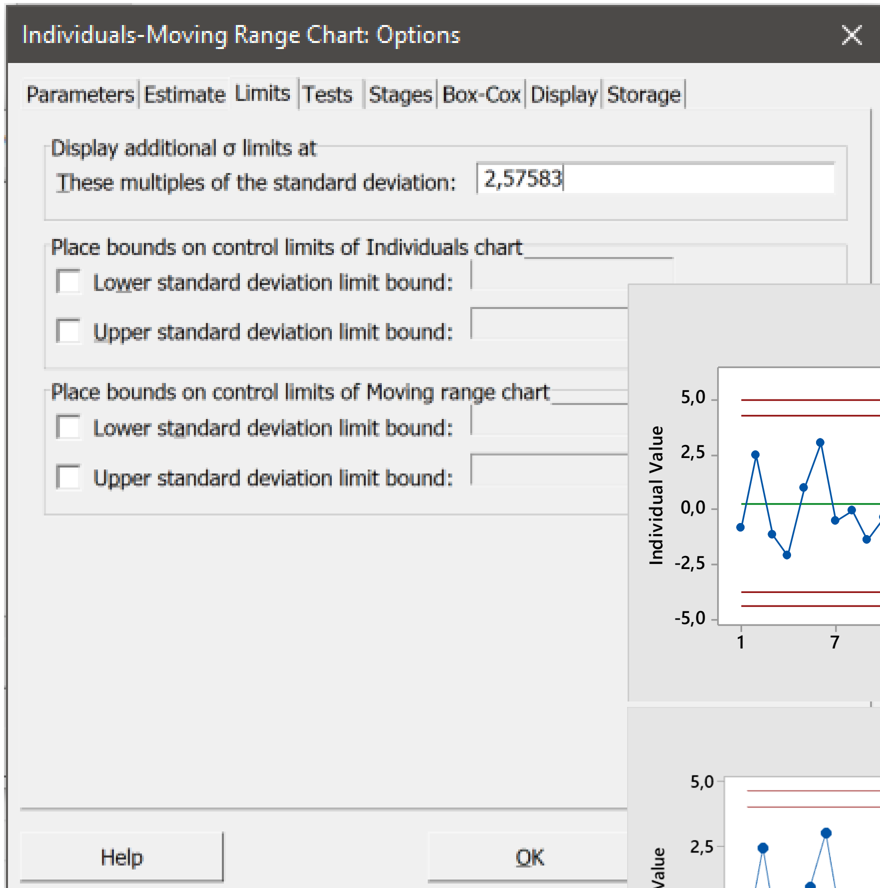
REMIND: moving range of length n=2

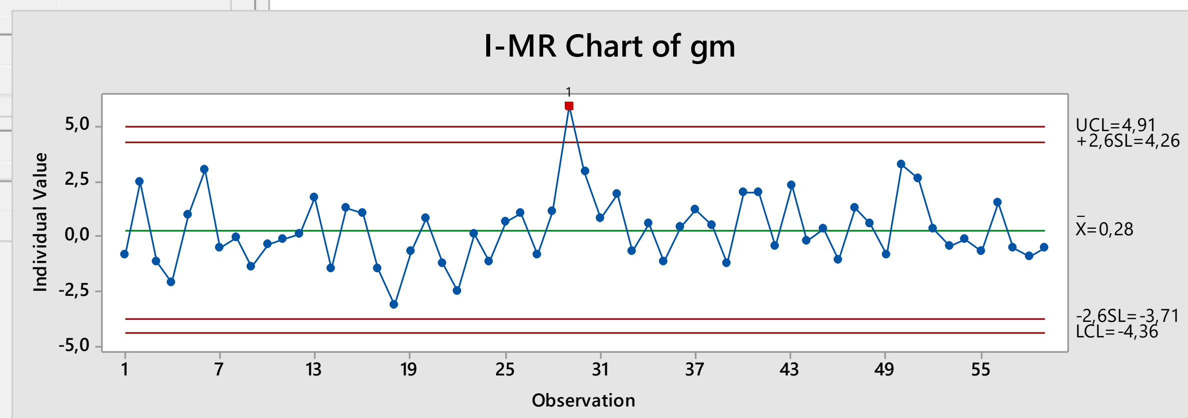
EXERCISE 3 (SOLUTION)

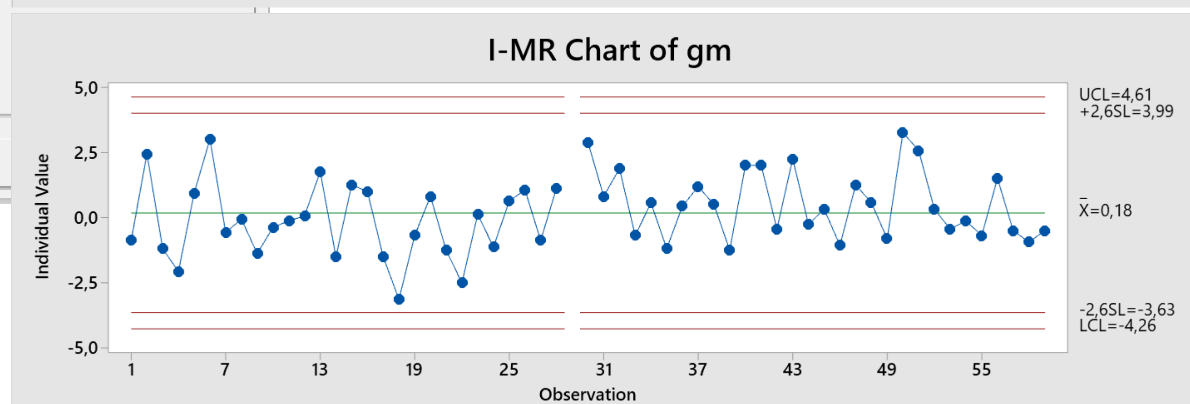
When using K ≠ 3, Minitab shows both limits at desired K and the ones with K = 3
Just ignore the ones at K = 3

I Control chart

In Minitab…







EXERCISE 3 (SOLUTION)

MR Control chart – by using the half-normal approximation

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

For n = 2

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

Alwan – Appendix A

LCL =0,0131

UCL=5,8683

Excluding row 29

Out-of-control MR=0