

A Look at Statistical Process Control

Dr. Lee Fawcett



Introduction

Introduction

Operations managers are responsible for developing and maintaining the production processes that deliver quality products and services.

Introduction

Operations managers are responsible for developing and maintaining the production processes that deliver quality products and services.

Once the production process is operating, it is often necessary to constantly monitor the process to ensure that it functions the way it was designed.

Introduction

The statistical methods we introduce in this lecture represent the most commonly-used applications of Statistics in many industries – at any point in time, there are *thousands* of organisations applying these methods.

Introduction

The statistical methods we introduce in this lecture represent the most commonly-used applications of Statistics in many industries – at any point in time, there are *thousands* of organisations applying these methods.

In this lecture, we deal with the subject of **Statistical Process Control**.

Process variation

All production processes result in variation; that is, no product is exactly the same as another.

Process variation

All production processes result in variation; that is, no product is exactly the same as another.

For example, if you weigh two boxes of 500 gram breakfast cereal,

Process variation

All production processes result in variation; that is, no product is exactly the same as another.

For example, if you weigh two boxes of 500 gram breakfast cereal,

- it is unlikely that either of them will weight exactly 500 grams

Process variation

All production processes result in variation; that is, no product is exactly the same as another.

For example, if you weigh two boxes of 500 gram breakfast cereal,

- it is unlikely that either of them will weight exactly 500 grams
- it is unlikely that they will have the same weight

Process variation

All products exhibit some degree of variation.

Process variation

All products exhibit some degree of variation. This could be

Process variation

All products exhibit some degree of variation. This could be

- 1 **Chance variation** – caused by a number of randomly occurring events that are part of the production process and that cannot be eliminated without changing the process
- 2 **Assignable variation** – caused by specific events or factors that are frequently temporary and that can usually be identified and eliminated

Example: Tablet production

A pharmaceutical company is required to make tablets of a nominal 500mg weight.

The company acknowledge that there will be some variation in the weight of the tablets produced, *even when the production process is working as it was designed to work*. This is **chance variation**.

Occasionally, there will be a malfunction somewhere in the production process, causing the variation in the weight of tablets to increase.

Example: Tablet production

A pharmaceutical company is required to make tablets of a nominal 500mg weight.

The company acknowledge that there will be some variation in the weight of the tablets produced, *even when the production process is working as it was designed to work*. This is **chance variation**.

Occasionally, there will be a malfunction somewhere in the production process, causing the variation in the weight of tablets to increase. This increase is the **assignable variation**.

Example: Tablet production

The best way to understand what is happening is to consider the weight of each tablet as a **random variable**.

Example: Tablet production

The best way to understand what is happening is to consider the weight of each tablet as a **random variable**.

If the only sources of variation are caused by chance, then each tablet's weight is drawn from identical distributions –

Example: Tablet production

The best way to understand what is happening is to consider the weight of each tablet as a **random variable**.

If the only sources of variation are caused by chance, then each tablet's weight is drawn from identical distributions – that is, each distribution has the same *shape*, *mean* and *standard deviation*.

Example: Tablet production

The best way to understand what is happening is to consider the weight of each tablet as a **random variable**.

If the only sources of variation are caused by chance, then each tablet's weight is drawn from identical distributions – that is, each distribution has the same *shape*, *mean* and *standard deviation*.

Under such circumstances, the production process is said to be **under control**.

Process variation

In recognition of the fact that variation in output will occur even when the process *is* under control and operating properly, most processes are designed so that their products will fall within designated **specification limits** or “**specs**”.

Process variation

In recognition of the fact that variation in output will occur even when the process *is* under control and operating properly, most processes are designed so that their products will fall within designated **specification limits** or “**specs**”.

For example, the process that makes the tablets may be designed so that the tablets can weight between 498 and 502 mg.

Process variation

Inevitably, some event or combination of factors in a production process will cause the process distribution to change.

Process variation

Inevitably, some event or combination of factors in a production process will cause the process distribution to change.

When it does, the process is said to be **out of control**.

There are several ways for the process to go out of control:

There are several ways for the process to go out of control:

- 1 Level shift

There are several ways for the process to go out of control:

- ① Level shift
- ② Instability

There are several ways for the process to go out of control:

- ① Level shift
- ② Instability
- ③ Trend

Process variation

There are several ways for the process to go out of control:

- ① Level shift
- ② Instability
- ③ Trend
- ④ Cycle

1. Level shift

- ▶ This is a change in the *mean* of the process distribution
- ▶ Assignable causes: Machine breakdown, new machine and/or operator, or a change in the environment

Process variation

1. Level shift

- ▶ This is a change in the *mean* of the process distribution
- ▶ Assignable causes: Machine breakdown, new machine and/or operator, or a change in the environment

2. Instability

- ▶ This is when the *standard deviation* increases
- ▶ Assignable causes: A machine in need of repair, defective materials, poorly trained operator

3. Trend

- ▶ When there is a slow, steady shift (either up or down) in the process distribution mean.
- ▶ Assignable causes: less-than-regular maintenance, operator fatigue, residue or dirt build up

3. Trend

- ▶ When there is a slow, steady shift (either up or down) in the process distribution mean.
- ▶ Assignable causes: less-than-regular maintenance, operator fatigue, residue or dirt build up

4. Cycle

- ▶ This is a repeated series of small observations followed by large observations
- ▶ Assignable causes: environmental changes, worn parts, or operator fatigue

Process variation

The key to quality is to detect when the process goes out of control so that we can correct the malfunction and restore control of the process.

Process variation

The key to quality is to detect when the process goes out of control so that we can correct the malfunction and restore control of the process.

The **control chart** is the statistical method that we use to detect such problems.

Control charts

A **control chart** is a plot of statistics over time.

Control charts

A **control chart** is a plot of statistics over time.

For example, an **$\bar{x}$ -chart** plots a series of sample means taken over a period of time.

Control charts

We will now break for a problem session covering this section.

Please pause the video here and go to **Problem Session 1** of **Lesson 10**.

See you in a bit!

Control charts

Control charts

Each control chart contains a **centreline** and **control limits**.

Control charts

Each control chart contains a **centreline** and **control limits**.

The control limit above the centreline is called the **upper control limit** (UCL) and that below the centreline is called the **lower control limit** (LCL).

Control charts

Each control chart contains a **centreline** and **control limits**.

The control limit above the centreline is called the **upper control limit** (UCL) and that below the centreline is called the **lower control limit** (LCL).

If, when the sample statistics are plotted, all points are randomly distributed between the control limits, we conclude that the process is **under control**.

Control charts

Each control chart contains a **centreline** and **control limits**.

The control limit above the centreline is called the **upper control limit** (UCL) and that below the centreline is called the **lower control limit** (LCL).

If, when the sample statistics are plotted, all points are randomly distributed between the control limits, we conclude that the process is **under control**.

If the points are not randomly distributed between the control limits, we conclude that the process is **out of control**.

Example: Tablet production

Suppose that in the tablet production example described previously, we want to determine whether the central location (mean) of the distribution has changed from one period to another.

Example: Tablet production

Suppose that in the tablet production example described previously, we want to determine whether the central location (mean) of the distribution has changed from one period to another.

We will draw our conclusion from an $\bar{x}$ -chart.

Example: Tablet production

Suppose that in the tablet production example described previously, we want to determine whether the central location (mean) of the distribution has changed from one period to another.

We will draw our conclusion from an $\bar{x}$ -chart.

For the moment, let us assume that we know the mean μ and the standard deviation σ of the process when it is under control; $\mu = 500$ mg and $\sigma = 15$ mg.

Example: Tablet production

Samples of 20 tablets are taken every hour for a period of 24 hours; the resulting sample means (in mg) are shown in the table below.

Example: Tablet production

Samples of 20 tablets are taken every hour for a period of 24 hours; the resulting sample means (in mg) are shown in the table below.

Time $\bar{x}$	13:03 501	14:03 494	15:03 496	16:03 502	17:03 505	18:03 488	19:03 487	20:03 503
Time $\bar{x}$	21:03 500	22:03 502	23:03 498	00:03 491	01:03 509	02:03 495	03:03 495	04:03 508
Time $\bar{x}$	05:03 490	06:03 500	07:03 503	08:03 507	09:03 486	10:03 480	11:03 485	12:03 482

Example: Tablet production

Recall the **Central Limit Theorem** from Lecture 5.

Example: Tablet production

Recall the **Central Limit Theorem** from Lecture 5.

This states that, no matter what the distribution of our random variable, the mean $\bar{x}$ of a random sample $x_1, x_2, \dots, x_n$ from this distribution (provided n is large) is approximately Normal with mean μ and variance σ^2/n , i.e.

Example: Tablet production

Recall the **Central Limit Theorem** from Lecture 5.

This states that, no matter what the distribution of our random variable, the mean $\bar{x}$ of a random sample $x_1, x_2, \dots, x_n$ from this distribution (provided n is large) is approximately Normal with mean μ and variance σ^2/n , i.e.

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

where μ and σ^2 are the mean and variance (respectively) of the distribution of our random variable.

Example: Tablet production

In the tablet production, we are told that $\mu = 500$ mg and $\sigma = 15$ mg. Means are taken every hour from samples of size 20, and so we have

$$\bar{x} \sim N\left(500, \frac{15^2}{20}\right) \quad \text{i.e.}$$

$$\bar{x} \sim N(500, 11.25)$$

Example: Tablet production

In the tablet production, we are told that $\mu = 500$ mg and $\sigma = 15$ mg. Means are taken every hour from samples of size 20, and so we have

$$\bar{x} \sim N\left(500, \frac{15^2}{20}\right) \quad \text{i.e.}$$

$$\bar{x} \sim N(500, 11.25)$$

Also, we have the following properties of the Normal distribution, in terms of areas under its probability density function:

Example: Tablet production

In the tablet production, we are told that $\mu = 500$ mg and $\sigma = 15$ mg. Means are taken every hour from samples of size 20, and so we have

$$\bar{x} \sim N\left(500, \frac{15^2}{20}\right) \quad \text{i.e.}$$

$$\bar{x} \sim N(500, 11.25)$$

Also, we have the following properties of the Normal distribution, in terms of areas under its probability density function:

- $\bar{x} \pm 1$ standard deviation gives about 68% coverage,

Example: Tablet production

In the tablet production, we are told that $\mu = 500$ mg and $\sigma = 15$ mg. Means are taken every hour from samples of size 20, and so we have

$$\bar{x} \sim N\left(500, \frac{15^2}{20}\right) \quad \text{i.e.}$$

$$\bar{x} \sim N(500, 11.25)$$

Also, we have the following properties of the Normal distribution, in terms of areas under its probability density function:

- $\bar{x} \pm 1$ standard deviation gives about 68% coverage,
- $\bar{x} \pm 2$ standard deviations gives about 95% coverage, and

Example: Tablet production

In the tablet production, we are told that $\mu = 500$ mg and $\sigma = 15$ mg. Means are taken every hour from samples of size 20, and so we have

$$\bar{x} \sim N\left(500, \frac{15^2}{20}\right) \quad \text{i.e.}$$

$$\bar{x} \sim N(500, 11.25)$$

Also, we have the following properties of the Normal distribution, in terms of areas under its probability density function:

- $\bar{x} \pm 1$ standard deviation gives about 68% coverage,
- $\bar{x} \pm 2$ standard deviations gives about 95% coverage, and
- $\bar{x} \pm 3$ standard deviations gives about 99% coverage

Control charts

Thus, the lower and upper control limits for an $\bar{x}$ -chart are often defined as:

Control charts

Thus, the lower and upper control limits for an $\bar{x}$ -chart are often defined as:

$$\text{Lower control limit} = \mu - 3 \text{ standard errors} = \mu - 3 \sqrt{\frac{\sigma^2}{n}} = \mu - 3 \frac{\sigma}{\sqrt{n}}$$

Control charts

Thus, the lower and upper control limits for an $\bar{x}$ -chart are often defined as:

$$\text{Lower control limit} = \mu - 3 \text{ standard errors} = \mu - 3\sqrt{\frac{\sigma^2}{n}} = \mu - 3\frac{\sigma}{\sqrt{n}}$$

$$\text{Upper control limit} = \mu + 3 \text{ standard errors} = \mu + 3\sqrt{\frac{\sigma^2}{n}} = \mu + 3\frac{\sigma}{\sqrt{n}}$$

Control charts

For the tablet production example, this would give:

For the tablet production example, this would give:

$$\text{Lower control limit} = 500 - 3 \times \frac{15}{\sqrt{20}} = 489.94$$

Control charts

For the tablet production example, this would give:

$$\text{Lower control limit} = 500 - 3 \times \frac{15}{\sqrt{20}} = 489.94$$

$$\text{Upper control limit} = 500 + 3 \times \frac{15}{\sqrt{20}} = 510.06$$

Control charts

For the tablet production example, this would give:

$$\text{Lower control limit} = 500 - 3 \times \frac{15}{\sqrt{20}} = 489.94$$

$$\text{Upper control limit} = 500 + 3 \times \frac{15}{\sqrt{20}} = 510.06$$

Also, we have

$$\text{Centreline} = 500$$

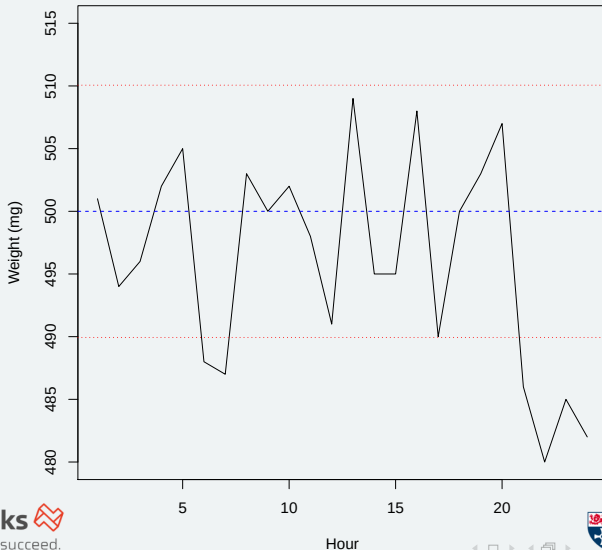
Control charts

We can produce the $\bar{x}$ -chart in R very easily, given the data provided:

```
1 hour = 1:24
2 xbar = c(501, 494, 496, 502, 505, 488, 487, 503, 500, 502, 498, 491, 509, 495, 495,
3         508, 490, 500, 503, 507, 486, 480, 485, 482)
4 LCL = 500 - 3*15/sqrt(20)
5 UCL = 500 + 3*15/sqrt(20)
6 CL = 500
7 plot(hour, xbar, type="l", main="x-bar chart for tablet weights", xlab="Hour", ylab="
8      Weight (mg)", ylim=c(480, 515))
9 abline(h = LCL, col="red", lty=3)
10 abline(h = UCL, col="red", lty=3)
11 abline(h = CL, col="blue", lty=2)
```

Control charts

x-bar chart for tablet weights



2-sigma and 3-sigma control limits

The control limits used in the previous example are called “3-sigma” control limits – we used a distance of 3 standard errors either side of the mean as a tool to detect the process being out of control.

2-sigma and 3-sigma control limits

The control limits used in the previous example are called “3-sigma” control limits – we used a distance of 3 standard errors either side of the mean as a tool to detect the process being out of control.

Why not use “2-sigma” control limits?

2-sigma and 3-sigma control limits

The control limits used in the previous example are called “3-sigma” control limits – we used a distance of 3 standard errors either side of the mean as a tool to detect the process being out of control.

Why not use “2-sigma” control limits?

Sometimes these *are* used, although the following reasoning should help explain why 3-sigma control limits are often preferred.

2-sigma and 3-sigma control limits

Suppose we have

2-sigma and 3-sigma control limits

Suppose we have

H_0 : The process is under control

2-sigma and 3-sigma control limits

Suppose we have

H_0 : The process is under control

H_1 : The process is out of control

2-sigma and 3-sigma control limits

Suppose we have

H_0 : The process is under control

H_1 : The process is out of control

In statistical process control, a **Type I error** occurs if we conclude that the process is out of control when in fact it is not (*i.e. reject H_0 when in fact it is true*).

2-sigma and 3-sigma control limits

Suppose we have

H_0 : The process is under control

H_1 : The process is out of control

In statistical process control, a **Type I error** occurs if we conclude that the process is out of control when in fact it is not (*i.e. reject H_0 when in fact it is true*).

Expensive error: The production process is usually stopped and the causes of the variation are found and repaired –

2-sigma and 3-sigma control limits

Suppose we have

H_0 : The process is under control

H_1 : The process is out of control

In statistical process control, a **Type I error** occurs if we conclude that the process is out of control when in fact it is not (*i.e. reject H_0 when in fact it is true*).

Expensive error: The production process is usually stopped and the causes of the variation are found and repaired – stopping the production process is costly, especially if there was no need to do so!

2-sigma and 3-sigma control limits

So, we would like the probability of a Type I error to be as small as possible. For the tablet production example, we can work out the probability of a Type I error as follows.

2-sigma and 3-sigma control limits

Using 3-sigma control limits:

2-sigma and 3-sigma control limits

Using 3-sigma control limits:

$$P(\bar{x} < \text{LCL}) = P(\bar{x} < 489.94)$$

$$P(\bar{x} > \text{UCL}) = P(\bar{x} > 510.06)$$

Therefore,

$$P(\text{Type I Error}) = P(\bar{x} < 489.94) + P(\bar{x} > 510.06),$$

which we can find very easily using techniques from lecture 4:

```
1 pnorm(489.94, 500, sqrt(11.25)) + (1-pnorm(510.06, 500, sqrt(11.25)))
2 [1] 0.002705896
3
4 #Or, by the symmetry of the Normal distribution:
5 2*pnorm(489.94, 500, sqrt(11.25))
6 [1] 0.002705896
```

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

$$\text{Lower control limit} = 500 - 2 \times \frac{15}{\sqrt{20}} = 493.29$$

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

$$\text{Lower control limit} = 500 - 2 \times \frac{15}{\sqrt{20}} = 493.29$$

$$\text{Upper control limit} = 500 + 2 \times \frac{15}{\sqrt{20}} = 506.71$$

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

$$\text{Lower control limit} = 500 - 2 \times \frac{15}{\sqrt{20}} = 493.29$$

$$\text{Upper control limit} = 500 + 2 \times \frac{15}{\sqrt{20}} = 506.71$$

This would give:

```
1 2*pnorm(493.29, 500, sqrt(11.25))  
2 [1] 0.04544247
```

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

$$\text{Lower control limit} = 500 - 2 \times \frac{15}{\sqrt{20}} = 493.29$$

$$\text{Upper control limit} = 500 + 2 \times \frac{15}{\sqrt{20}} = 506.71$$

This would give:

```
1 2*pnorm(493.29, 500, sqrt(11.25))  
2 [1] 0.04544247
```

- Using 3-sigma control limits will give a **false alarm** in about 0.26% of samples

2-sigma and 3-sigma control limits

Similarly, if we use 2-sigma control limits, we would have:

$$\text{Lower control limit} = 500 - 2 \times \frac{15}{\sqrt{20}} = 493.29$$

$$\text{Upper control limit} = 500 + 2 \times \frac{15}{\sqrt{20}} = 506.71$$

This would give:

```
1 2*pnorm(493.29, 500, sqrt(11.25))  
2 [1] 0.04544247
```

- Using 3-sigma control limits will give a **false alarm** in about 0.26% of samples
- Using 2-sigma control limits will give a **false alarm** in about 4.6% of samples

2-sigma and 3-sigma control limits

A **Type II error** occurs when we do not detect when the process is out of control (*i.e. we retain H_0 when in fact it is false*).

Average run length

The **Average Run Length** (ARL) is the expected number of samples that must be taken before the control chart indicates that the process has gone out of control.

Average run length

The **Average Run Length** (ARL) is the expected number of samples that must be taken before the control chart indicates that the process has gone out of control.

The ARL is determined by

Average run length

The **Average Run Length** (ARL) is the expected number of samples that must be taken before the control chart indicates that the process has gone out of control.

The ARL is determined by

$$ARL = \frac{1}{P}$$

Average run length

The **Average Run Length** (ARL) is the expected number of samples that must be taken before the control chart indicates that the process has gone out of control.

The ARL is determined by

$$ARL = \frac{1}{P}$$

where P is the probability that a sample mean falls outside the control limits.

Average run length

Assuming that 3-sigma control limits are used, this would give

Average run length

Assuming that 3-sigma control limits are used, this would give

$$ARL = \frac{1}{0.0026} = 385.$$

Average run length

Assuming that 3-sigma control limits are used, this would give

$$ARL = \frac{1}{0.0026} = 385.$$

This means that when the process is under control, the $\bar{x}$ -chart will erroneously conclude that it is *out* of control once every 385 samples, on average.

Average run length

Assuming that 3-sigma control limits are used, this would give

$$ARL = \frac{1}{0.0026} = 385.$$

This means that when the process is under control, the $\bar{x}$ -chart will erroneously conclude that it is *out* of control once every 385 samples, on average.

If the sampling plan calls for samples to be taken once every hour (as in the tablet production example), on average there will be a **false alarm** once every 385 hours.

Control charts

We will now break for a problem session covering this section.

Please pause the video here and go to **Problem Session 2** of **Lesson 10**.

See you in a bit!

Control charts: Unknown population parameters

Control charts

In the examples presented so far, we have assumed that the process parameters (μ and σ) were known.

Control charts

In the examples presented so far, we have assumed that the process parameters (μ and σ) were known.

When the parameters are *unknown*, we estimate their values from the sample data.

Control charts

In the examples presented so far, we have assumed that the process parameters (μ and σ) were known.

When the parameters are *unknown*, we estimate their values from the sample data.

In the next section, we discuss how to construct and use control charts in more realistic situations.

Control charts

In the examples presented so far, we have assumed that the process parameters (μ and σ) were known.

When the parameters are *unknown*, we estimate their values from the sample data.

In the next section, we discuss how to construct and use control charts in more realistic situations.

We will re-visit $\bar{x}$ -charts, as well as S -charts and R -charts, for the case when the process parameters (μ and σ) are unknown.

Control charts

There are several ways to determine whether a change in the process distribution has occurred.

Control charts

There are several ways to determine whether a change in the process distribution has occurred.

To determine whether the distribution means have changed, we employ the $\bar{x}$ -chart.

Control charts

There are several ways to determine whether a change in the process distribution has occurred.

To determine whether the distribution means have changed, we employ the $\bar{x}$ -chart.

To determine whether the process distribution standard deviation has changed, we can use the S -chart or the R -chart (S for standard deviation, R for range).

In the previous section we determined the centreline and the control limits of an $\bar{x}$ -chart using the mean and standard deviation of the process distribution.

In the previous section we determined the centreline and the control limits of an $\bar{x}$ -chart using the mean and standard deviation of the process distribution.

However, it is unrealistic to believe that the mean (μ) and standard deviation (σ) of the process distribution will be known.

In the previous section we determined the centreline and the control limits of an $\bar{x}$ -chart using the mean and standard deviation of the process distribution.

However, it is unrealistic to believe that the mean (μ) and standard deviation (σ) of the process distribution will be known.

Thus, to construct the $\bar{x}$ -chart, we need to **estimate** the relevant parameters from the data.

We begin by drawing samples when we have determined that the process is under control.

$\bar{x}$ -chart

We begin by drawing samples when we have determined that the process is under control.

For each sample, we compute the mean and standard deviation.

$\bar{\bar{x}}$ -chart

We begin by drawing samples when we have determined that the process is under control.

For each sample, we compute the mean and standard deviation.

The estimator of the mean of the distribution is the mean of the sample means (denoted $\bar{\bar{x}}$ and pronounced “x bar bar” or “x double bar”), given by

We begin by drawing samples when we have determined that the process is under control.

For each sample, we compute the mean and standard deviation.

The estimator of the mean of the distribution is the mean of the sample means (denoted $\bar{\bar{x}}$ and pronounced “x bar bar” or “x double bar”), given by

$$\bar{\bar{x}} = \frac{\sum_{j=1}^k \bar{x}_j}{k},$$

We begin by drawing samples when we have determined that the process is under control.

For each sample, we compute the mean and standard deviation.

The estimator of the mean of the distribution is the mean of the sample means (denoted $\bar{\bar{x}}$ and pronounced “x bar bar” or “x double bar”), given by

$$\bar{\bar{x}} = \frac{\sum_{j=1}^k \bar{x}_j}{k},$$

where $\bar{x}_j$ is the mean of the j th sample and there are k samples.

To estimate the standard deviation of the process distribution, we calculate the “pooled standard deviation”:

To estimate the standard deviation of the process distribution, we calculate the “pooled standard deviation”:

$$S = \sqrt{\frac{\sum_{j=1}^k s_j^2}{k}}.$$

To estimate the standard deviation of the process distribution, we calculate the “pooled standard deviation”:

$$S = \sqrt{\frac{\sum_{j=1}^k s_j^2}{k}}.$$

Previously, we used:

$$\begin{aligned}\text{Centreline} &= \mu \\ \text{Lower Control Limit} &= \mu - 3 \frac{\sigma}{\sqrt{n}} \\ \text{Upper Control Limit} &= \mu + 3 \frac{\sigma}{\sqrt{n}}\end{aligned}$$

Because the values of μ and σ are now unknown, we *now* use:

Because the values of μ and σ are now unknown, we *now* use:

- $\bar{\bar{x}}$ in place of μ
- S in place of σ

Because the values of μ and σ are now unknown, we *now* use:

- $\bar{\bar{x}}$ in place of μ
- S in place of σ

Therefore, the centreline and control limits are now:

$$\text{Centreline} = \bar{\bar{x}}$$

$$\text{Lower Control Limit} = \bar{\bar{x}} - 3 \frac{S}{\sqrt{n}}$$

$$\text{Upper Control Limit} = \bar{\bar{x}} + 3 \frac{S}{\sqrt{n}}$$

Tablet production

Another pharmaceutical company produces tablets of nominal weight 500 mg.

Every hour, over the last 25 hours, the company have randomly sampled four tablets off their production line.

The raw data are shown in the table overleaf.

Unlike the previous company, this company do not know what the standard deviation of their production process is.

Example: Tablet production

Sample	Tablet 1	Tablet 2	Tablet 3	Tablet 4	Mean ($\bar{x}_j$)	St. dev. (s_j)
1	501.02	501.65	504.34	501.10	502.027	1.56689
2	499.80	498.89	499.47	497.90	499.015	0.83309
3	497.12	498.35	500.34	499.33	498.785	1.37556
4	500.68	501.39	499.74	500.41	500.555	0.68267
5	495.87	500.92	498.00	499.44	498.558	2.15203
6	497.89	499.22	502.10	500.03	499.810	1.76324
7	497.24	501.04	498.74	503.51	500.132	2.74084
8	501.22	504.53	499.06	505.37	502.545	2.93381
9	499.15	501.11	497.96	502.39	500.152	1.97782
10	498.90	505.99	500.05	499.33	501.068	3.31578
11	497.38	497.80	497.57	500.72	498.368	1.57771
12	499.70	500.99	501.35	496.48	499.630	2.21626
13	501.44	500.46	502.07	500.50	501.118	0.77993
14	498.26	495.54	495.21	501.27	497.570	2.81996
15	497.57	497.00	500.32	501.22	499.027	2.05853
16	500.95	502.07	500.60	500.44	501.015	0.73487
17	499.70	500.56	501.18	502.36	500.950	1.11887
18	501.57	502.09	501.18	504.98	502.455	1.72411
19	504.20	500.92	500.02	501.71	501.712	1.79632
20	498.61	499.63	498.68	501.84	499.690	1.50694
21	499.05	501.82	500.67	497.36	499.725	1.94344
22	497.85	494.08	501.79	501.95	498.918	3.74114
23	501.08	503.12	503.06	503.56	502.705	1.10603
24	500.75	501.18	501.09	502.88	501.475	0.95479
25	502.03	501.44	498.76	499.39	500.405	1.57566

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\bar{\bar{x}} = \frac{502.027 + 499.015 + \dots}{25}$$

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\begin{aligned}\bar{\bar{x}} &= \frac{502.027 + 499.015 + \dots}{25} \\ &= 500.296.\end{aligned}$$

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\begin{aligned}\bar{\bar{x}} &= \frac{502.027 + 499.015 + \dots}{25} \\ &= 500.296.\end{aligned}$$

Also, the pooled standard deviation S is found as

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\begin{aligned}\bar{\bar{x}} &= \frac{502.027 + 499.015 + \dots}{25} \\ &= 500.296.\end{aligned}$$

Also, the pooled standard deviation S is found as

$$S = \sqrt{\frac{1.56689^2 + 0.83309^2 + \dots}{25}}$$

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\begin{aligned}\bar{\bar{x}} &= \frac{502.027 + 499.015 + \dots}{25} \\ &= 500.296.\end{aligned}$$

Also, the pooled standard deviation S is found as

$$\begin{aligned}S &= \sqrt{\frac{1.56689^2 + 0.83309^2 + \dots}{25}} \\ &= \sqrt{\frac{97.125}{25}}\end{aligned}$$

Example: Tablet production

The mean of the means – $\bar{\bar{x}}$ – is given by

$$\begin{aligned}\bar{\bar{x}} &= \frac{502.027 + 499.015 + \dots}{25} \\ &= 500.296.\end{aligned}$$

Also, the pooled standard deviation S is found as

$$\begin{aligned}S &= \sqrt{\frac{1.56689^2 + 0.83309^2 + \dots}{25}} \\ &= \sqrt{\frac{97.125}{25}} \\ &= 1.971\end{aligned}$$

Example: Tablet production

Thus, the centreline and control limits are

Example: Tablet production

Thus, the centreline and control limits are

$$\text{Centreline} = \bar{\bar{x}} = 500.296$$

Example: Tablet production

Thus, the centreline and control limits are

$$\text{Centreline} = \bar{\bar{x}} = 500.296$$

$$\text{Lower Control Limit} = \bar{\bar{x}} - 3 \frac{S}{\sqrt{n}} = 500.296 - 3 \times \frac{1.971}{\sqrt{4}} = 497.34$$

Example: Tablet production

Thus, the centreline and control limits are

$$\text{Centreline} = \bar{\bar{x}} = 500.296$$

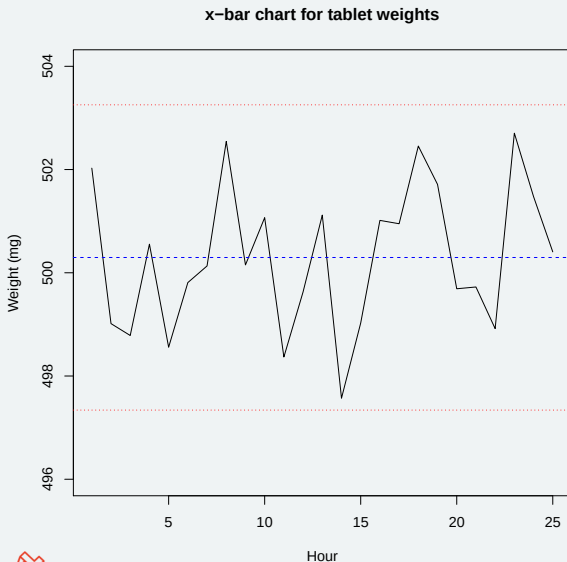
$$\text{Lower Control Limit} = \bar{\bar{x}} - 3 \frac{S}{\sqrt{n}} = 500.296 - 3 \times \frac{1.971}{\sqrt{4}} = 497.34$$

$$\text{Upper Control Limit} = \bar{\bar{x}} + 3 \frac{S}{\sqrt{n}} = 500.296 + 3 \times \frac{1.971}{\sqrt{4}} = 503.253$$

Example: Tablet production

```
1 hour = 1:25
2 xbar = c(502.027, 499.015, ..., 500.405)
3 LCL = 497.34
4 UCL = 503.253
5 CL = 500.296
6
7 plot(hour, xbar, type="l", main="x-bar chart for tablet weights", xlab="Hour", ylab="
  Weight (mg)", ylim=c(496, 504))
8 abline(h = LCL, col="red", lty=3)
9 abline(h = UCL, col="red", lty=3)
10 abline(h = CL, col="blue", lty=2)
```

Control charts

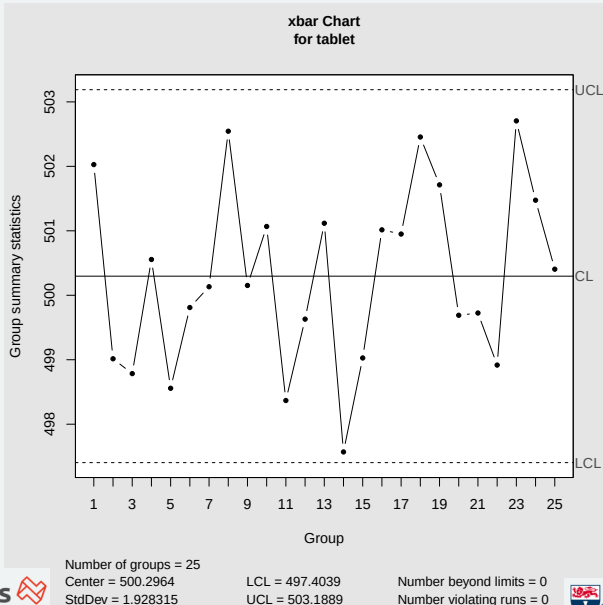


Control charts

There is, of course, a much easier way to do all of this in R – we can use the **qcc** package!

```
1 install.packages("qcc")
2 library(qcc)
3
4 / _ | / _/ _| Quality Control Charts and
5 | (-| | (-| (- Statistical Process Control
6 \-- | \---\---|
7   | _| version 2.7
8 #If the data are stored in a dataframe, in which the rows represent samples:
9 qcc(tablet, type="xbar")
10 List of 11
11 $ call      : language qcc(data = tablet, type = "xbar")
12 $ type      : chr "xbar"
13 $ data.name : chr "tablet"
14 $ data      : num [1:25, 1:4] 501 500 497 501 496 ...
15 ..- attr(*, "dimnames")=List of 2
16 $ statistics: Named num [1:25] 502 499 499 501 499 ...
17 ..- attr(*, "names")= chr [1:25] "1" "2" "3" "4" ...
18 $ sizes     : int [1:25] 4 4 4 4 4 4 4 4 4 ...
19 $ center    : num 500
20 $ std.dev   : num 1.93
21 $ nsigmas   : num 3
22 $ limits    : num [1, 1:2] 497 503
23 ..- attr(*, "dimnames")=List of 2
24 $ violations:List of 2
25 - attr(*, "class")= chr "qcc"
```

Control charts



S -charts and R -charts

The S -chart graphs sample standard deviations to determine whether the process distribution standard deviation has changed. The R -chart graphs the range.

S -charts and R -charts

The S -chart graphs sample standard deviations to determine whether the process distribution standard deviation has changed. The R -chart graphs the range.

The format is similar to that of the $\bar{x}$ chart – the S and R charts display a centreline and control limits.

S -charts and R -charts

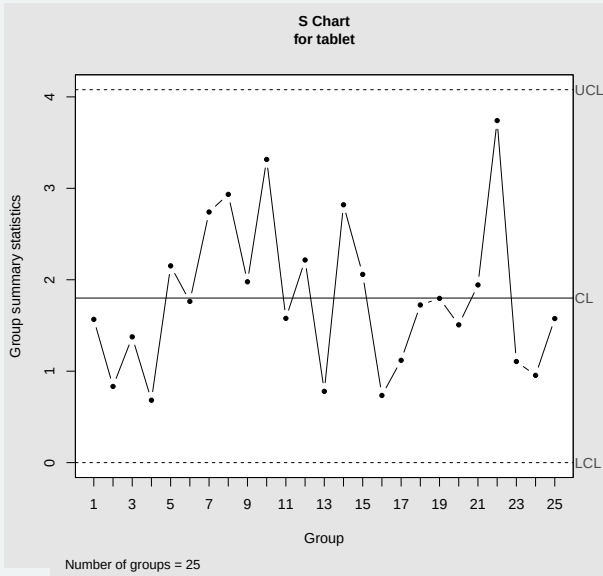
The S -chart graphs sample standard deviations to determine whether the process distribution standard deviation has changed. The R -chart graphs the range.

The format is similar to that of the $\bar{x}$ chart – the S and R charts display a centreline and control limits.

However, the formulae for the centreline and control limits are more complicated than those for the $\bar{x}$ -chart – consequently, we will let the **qcc** package in R do the work here!

```
1 qcc(tablet, type="S")  
2 qcc(tablet, type="R")
```

S -charts and R -charts



Number of groups = 25

Center = 1.799851

StdDev = 1.953562

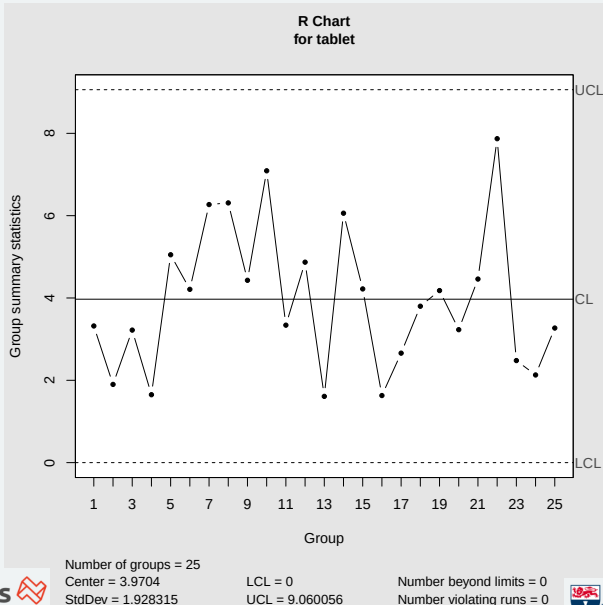
LCL = 0

UCL = 4.078548

Number beyond limits = 0

Number violating runs = 0

S -charts and R -charts



Monitoring the production process

So far, we have thought about $\bar{x}$ -charts and S/R -charts as separate procedures.

Monitoring the production process

So far, we have thought about $\bar{x}$ -charts and S/R -charts as separate procedures.

In actual practice, the charts are often drawn and assessed together – usually, the $\bar{x}$ -chart is used alongside the S -chart.

Monitoring the production process

So far, we have thought about $\bar{x}$ -charts and S/R -charts as separate procedures.

In actual practice, the charts are often drawn and assessed together – usually, the $\bar{x}$ -chart is used alongside the S -chart.

The reason for this is that the $\bar{x}$ -chart uses S to calculate the control limits; consequently, if the S -chart indicates that the process is out of control, the value of S will not lead to an accurate estimate of the standard deviation of the process distribution.

Monitoring the production process

The usual procedure is:

Monitoring the production process

The usual procedure is:

- 1 Draw the S -chart first;

Monitoring the production process

The usual procedure is:

- 1 Draw the S -chart first;
- 2 If this indicates that the process is under control, then draw the $\bar{x}$ -chart;

Monitoring the production process

The usual procedure is:

- 1 Draw the S -chart first;
- 2 If this indicates that the process is under control, then draw the $\bar{x}$ -chart;
- 3 If the $\bar{x}$ -chart also indicates that the process is under control, we can then use both charts to maintain control;

Monitoring the production process

The usual procedure is:

- 1 Draw the S -chart first;
- 2 If this indicates that the process is under control, then draw the $\bar{x}$ -chart;
- 3 If the $\bar{x}$ -chart also indicates that the process is under control, we can then use both charts to maintain control;
- 4 If either chart shows that the process was out of control at some time during the creation of the charts, we can detect and fix the problem and then re-draw the charts with new data.

Back to tablet production example

We have determined that the process is under control – none of the points in the S -chart nor the $\bar{x}$ -chart fell outside the control limits.

Back to tablet production example

We have determined that the process is under control – none of the points in the S -chart nor the $\bar{x}$ -chart fell outside the control limits.

We now use the statistics generated in the creation of these charts to monitor the production process.

Back to tablet production example

Recall that the sampling plan calls for samples of size 4 every hour.

Back to tablet production example

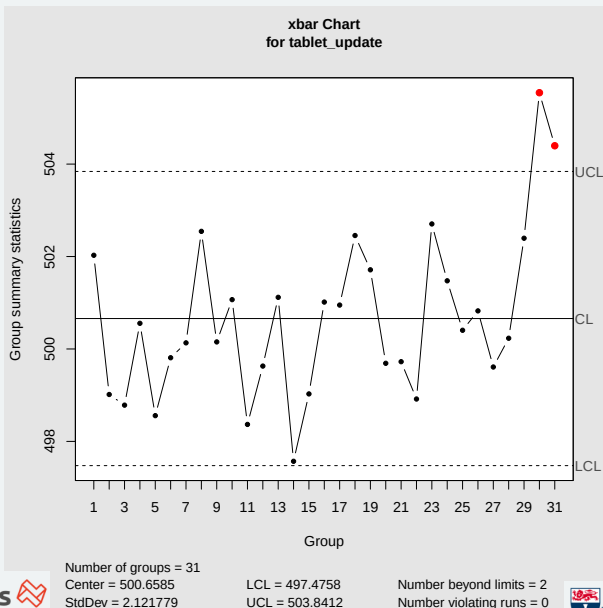
Recall that the sampling plan calls for samples of size 4 every hour.

The following table lists the weight of tablets from samples taken during the *next* six hours.

Sample	Tablet1	Tablet 2	Tablet 3	Tablet 4	$\bar{x}_j$	s_j
26	502.653	498.354	502.209	500.080	500.824	1.99315
27	501.212	494.454	500.918	501.855	499.610	3.45937
28	500.086	500.826	496.426	503.591	500.232	2.95199
29	502.994	500.481	502.996	503.113	502.396	1.27788
30	506.549	508.708	502.480	504.442	505.545	2.68474
31	500.441	508.887	505.005	503.254	504.397	3.53488

```
1 qcc(tablet_update, type="xbar")
```

Back to tablet production example



Control charts

We will now break for a problem session covering this section.

Please pause the video here and go to **Problem Session 3** of **Lesson 10**.

See you in a bit!

Conclusions

- ① Process variation – **chance** and **assignable**
- ② Process going “out of control” – **level shift**, **instability**, **trend** and **cycle**

Conclusions

- ① Process variation – **chance** and **assignable**
- ② Process going “out of control” – **level shift**, **instability**, **trend** and **cycle**
- ③ Control charts
 - ▶ $\bar{x}$ -charts (known and unknown process variation)
 - ▶ S -charts
 - ▶ R -charts

Conclusions

- ① Process variation – **chance** and **assignable**
- ② Process going “out of control” – **level shift**, **instability**, **trend** and **cycle**
- ③ Control charts
 - ▶ $\bar{x}$ -charts (known and unknown process variation)
 - ▶ S -charts
 - ▶ R -charts
- ④ **Average run lengths** and probability of **false alarms**