air pollution training institute

7.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

introduction to environmental statistics

self-instructional course

SI 473

MODULE VII

QUALITY CONTROL CHARTS

UNIT 1 - INTRODUCTION TO STATISTICAL QUALITY CONTROL CHARTS

EPA is primarily concerned with environmental quality. To assess the quality of local environments, EPA relies upon local measurements. As a consequence, EPA is also concerned with the quality of measurements. Because of the importance of decisions based on these data, a system of quality control for measurements is essential. To ensure that the data obtained are both accurate and precise, most of the same statistical quality control techniques that are used in industrial production can be applied. One of the statistical techniques is that of quality control charts. This unit introduces you to the types of control charts that are most applicable to measurement systems. The following topics are covered:

1. The concept of quality control.
2. The difference between quality control and quality assurance.
3. A discussion of accuracy and precision.
4. A brief discussion of bias, both analytical and sampling.
5. The use of percentages in speaking of confidence limits.

16

Let's tap your memory once again. In the introduction to hypothesis testing, you were exposed to "critical values." If your obtained statistic was beyond these values (assume $\alpha = .05$), you would be able to say that you were 95% confident that your obtained value was due to something other than chance. This 95% then termed your degree of confidence. What is the degree of confidence when $\alpha = .01$?

____

2

Every measurement process is subject to variation. Some of the variation is a result of determinable or "assignable" causes. Other variation results from indeterminate, "unassignable" or random causes. The variability resulting from these random causes is considered the inherent or uncontrollable variation of the measurement process. Statistical quality control charts based upon the inherent variability are used to detect when excessive variation has occurred in the process. The resulting investigation of the process to identify and correct any non-random or controllable causes should improve the quality of the measurement process.

16a

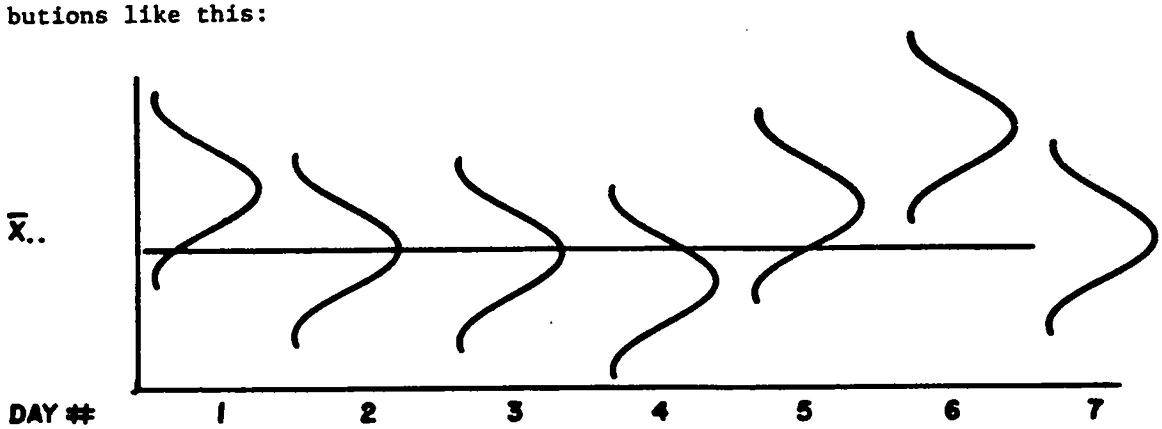
Answer: 99%

If you did not get the correct answer, refer back to Module II.

17

When we get to the construction of control charts, we will again encounter the concept of confidence in determining our limits, which when exceeded will result in an "out of control" conclusion.

Let's take a look at this variation, using a graphical representation. Assume we take many readings on successive days and plot the distribution of data for each day. In the following graphical example, the variability within each day is considered as the inherent variability. We could show the distributions like this:



For review purposes, what is $\bar{\bar{x}}$ .. and how is it calculated? † __________

____

____

____

____

18

Let's review:

A) Can measurements be precise, but not accurate? (yes/no)

B) Can measurements be neither accurate nor precise? (yes/no)

C) Can measurements be both accurate and precise? (yes/no)

3a

Answer: $\bar{x}$.. is the grand mean. It is calculated by combining all the samples into one large sample and calculating the mean. (If the samples are all of equal size, we could find the mean of the sample means.) If this still doesn't ring a bell, the grand mean was introduced in Module IV and you would be wise to re-read Module IV, page IV-12.

18a

Answer:

A) yes

B) yes

C) yes (In fact, the general objective of quality control of a measurement process is to assure that measurements are both accurate and precise.)

19

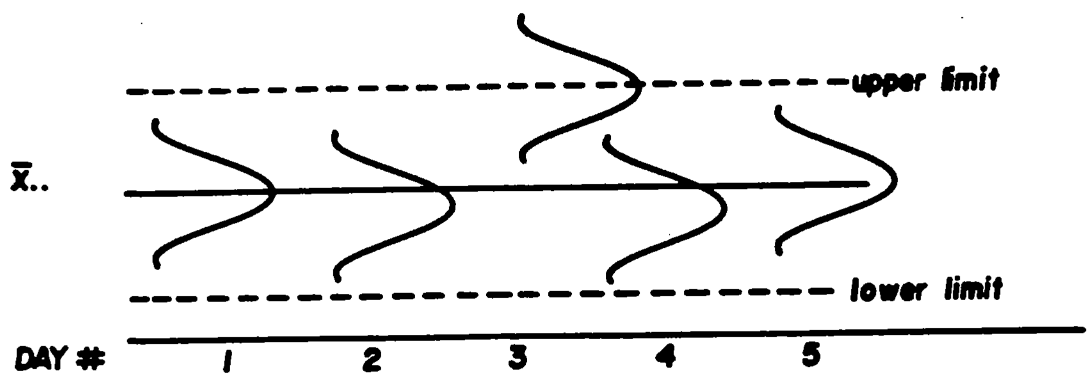
In quality control terms define:

A) α

B) β

To determine whether or not the variation is being controlled, limits are calculated. Whenever the data of the within-day sample distributions cross these limits, the process is termed "out of control". Converseely, if the sample distributions remain within the limits, the measurement process is "in control."

Which of the following were "in control" and which were "out of control?"



19a

Answer: α - the probability of concluding that a process is out of control when it is actually in control.

β - the probability of concluding that a process is in control when, in fact, it is out of control.

20

When a reading is beyond the control limits already set for that data, and when α = .01, we are #_____% confident that the reading is due to something other than chance.

4a

Answer: 1 - in control

2 - in control

3 - out of control

4 - in control

5 - in control

5

As you have probably already guessed, the use of charts similar to the ones you've just seen increases the ease with which quality control decisions can be made. The actual construction of quality control charts will be discussed in Unit 2 of this module.

20a

Answer: 99%

21

Because of concern over the quality of our environment, EPA must consider the quality of environmental measurements taken. Enforcement personnel and personnel involved in pollution control strategy base their actions on these measurements. Furthermore, the measurements must stand up to any challenges that may follow enforcement. For these reasons, we must be certain (at least with a specified degree of confidence) that the measurements are both accurate and precise.

At this point, let us distinguish between quality control and quality assurance. According to one source (APTD 1132 - Quality Control Practices in Processing Air Pollution Samples), quality control is defined as "the program applied to routinized systems (i.e., systems composed of methods, equipment, materials, and people) in order to evaluate and document the ability of a function, activity, or person to produce results which are valid within predetermined acceptance limits." Quality assurance is viewed as a higher-level non-routine auditing-type function to assure that the routine quality control system is working satisfactorily. In other words, quality assurance is quality control of quality __________. For a further discussion see Section 1.3 of Reference 1.

Pollution measurement systems can be subdivided into two separate but related processes; namely, sampling and analysis. Biases in final results may arise from biasing effects in either process. The resulting biases may be called:

1. Sampling bias
2. Analytical bias

Analytical bias may result from any of the elements of the analytical process, including analysts, procedures, materials, equipment, reagents, laboratory conditions, computational procedures, etc. For example, an analyst may employ a procedure variation, or a measuring device may consistently provide inaccurate data. Since this type of bias may be revealed by trends or shifts in level appearing on control charts for accuracy presented in the next unit, we will consider analytical bias in detail at that time.

6a

Answer: control

7

Another distinction is the difference between precision and accuracy. Precision refers to a measure of uniformity in a measurement process. Precision is determined from the agreement among repeated observations made under similar conditions. In other words, if four analysts of the same laboratory make a measurement of something with the same device and obtain readings in close agreement, these measurements (are/are not) precise.

23

The second type of bias, sampling bias, results when the sample that has been taken is not representative of the population about which inferences are to be made. An error of this sort was responsible for the prediction that Alf Landon would win the 1936 presidential election. In this famous case of improper sampling, the sample was selected from the mailing list of The Literary Digest, and from telephone directories. Unfortunately for the Digest, this sample consisted mostly of Republicans, who would have elected Landon if the remainder (and majority) of the country's voters had not been allowed to vote. However, they did vote, and Franklin Roosevelt was elected. Similar sampling bias results can occur in pollution measurements, such as when a sample taken from near the riverbank is used to represent the average for the entire river, or when an air sample taken from the top of a building is used to represent the air breathed by pedestrians. Sampling bias may also result when stack source samples are taken at times not representative of specified plant operation.

Answer: are

8

On the other hand, accuracy refers to the degree of difference between (1) the observed value and (2) the known, actual or "true" value. If an analyst measures the pH of a sample of water whose pH is known to be 7.2, and the analyst obtains a pH close to 7.2, then his measurement is _______________. Bias is the magnitude of the difference between the observed and the true value, expressed as a positive or negative value when the true value is subtracted from the observed value.

24

To help solve the problem of sampling bias, many sampling plans have been developed to increase the likelihood that a sample actually represents the population from which it was drawn. You have already been introduced to the random sample (one in which each member of the target population has an equal chance of being selected). Other types of sampling plans exist.

Some of the more common sampling plans include systematic sampling, stratified sampling, and cluster sampling. More information on statistical sampling plans is available in many statistics texts.

In addition to the statistical aspects of sampling, there are many practical considerations which must be made relating to the physical arrangement and equipment used for sampling, the schedule for sampling, and the detailed procedure employed.

8a

Answer: accurate

9

If four researchers measure SO_2 concentrations in the same air sample (known SO_2 concentration 35 ppm) and all four determine that the sample concentration is close to 35 ppm, the measurements are ________________ and ________________.

25

To help you understand why a consideration of bias is necessary, explain why each of the following samples does not qualify as an <u>un</u>biased sample from the target population:

(1) To estimate annual income of Princeton graduates 10 years after graduation, a questionnaire is sent in 1974 to graduates of 1964, and the estimate is based on those returned.

†____

____

(2) To predict local election results, a telephone poll is taken of a random sample of people listed in the phone directory.

†____

____

9a

Answer: accurate and precise

10

On the other hand, if the four analysts measure the SO_2 concentration in the same sample (known SO_2 concentration 35 ppm) and all four determine that the sample concentration is near 45 ppm, the measurements are ___________, but not ____________________. The bias, in this example is, ____________________________.

25a

Answer:

(1) Those people likely to return the questionnaire will have different incomes than those who did not. The very rich (who would not spend the time to fill out the questionnaire) and the very poor (who might be too embarassed to fill it out) are likely to be under-represented in the sample. Also, the selection of 1964 graduates may not represent other years' graduates.

(2) This one is a little tricky. This sample might represent the population well, but a few questions need to be answered first. Are the persons in the phone book truly representative of the persons who will vote in the election? In other words, are phone subscribers similar to voters on vote-related dimensions such as party registration, age, sex, religion, ethnic background, occupation, and income? If you can determine that the phone subscribers are truly representative of the voters, a random sampling from the phone book would be an efficient, unbiased procedure.

Answer: precise

accurate

+ 10 ppm

11

Now let's discuss some statistical concepts that you've seen before with a different viewpoint: We'll talk in terms of quality control applications. In hypothesis testing, when we made the decision to accept or reject the null hypotheses, we had the following possible outcomes:

	H_o True	H_o False
Accept H_o	No Error	Type II Error
Reject H_o	Type I Error	No Error

If H_o is true and we accept it, we have made * ______________.

If H_o is false and we accept it, we have made* ______________.

If H_o is true and we reject it, we have made * ______________.

If H_o is false and we reject it, we have made* ______________.

26

With further thought, you can think of similar sample bias situations in pollution sampling. As you can see from this exercise, it is easy to obtain biased data by failing to give the sample selection problem sufficient thought.

11a

Answer: no error

a Type II error

a Type I error

no error

12

You may also recall that we used Greek letters to designate the probabilities of making the two types of errors. The probability of making a Type I error is _____________(Greek Letter).

The probability of making a Type II error is ___________(Greek letter).

27

Another topic of concern to the evaluation of air-pollution data is the shape of the data distributions that are most frequently observed. Nearly all introductory statistics courses limit their discussions to those techniques that are applicable to normally distributed data, or to techniques that are "robust" (a term used by statisticians to describe the insensitivity to assumptions about the data collected) enough that lack of normality is not critical. Much of the pollution monitoring data are not normally distributed. These data would yield a curve that (is/is not) bell-shaped, and symmetrical.

12a

Answer: α

β

13

In a like manner, when the null hypothesis is that the process is in control, deciding whether or not a process is in control has the same four possibilities:

		Actually in control	Actually out of control
Decision	in control	no error	(________)
	out of control	(________)	no error

In the above matrix, fill in the Greek letter designating the probabilities of making that error.

27a

Answer: is not

28

As you may recall, in Module I we introduced the geometric mean as a measure of central tendency that is often used on data that are log-normally distributed (e.g., with a large number of low values, and a few high values).

The use of the geometric mean tends to (reduce/increase) the impact of these high values.

13a

Answer:

		Actually in control	Actually out of control
Decision	in control	no error	β
	out of control	α	no error

28a

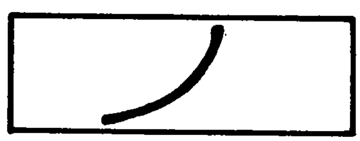
Answer: reduce

29

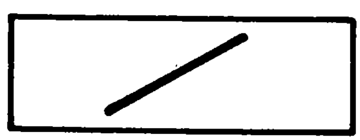
The reason for the reduction of impact of the high values lies in the transformation to logarithms in the calculation of the geometric mean

$$\text{Log } \bar{x}_g = \frac{\Sigma \text{ Log } x_i}{n}$$

The effect of the Log transformation can be demonstrated graphically. Remember in Module I, in the discussion of graphical analysis of data, we encountered a curve instead of a straight line?



And, when we plotted the data on log normal paper, it straightened the curve?



This was due to the fact that when you plot data on log normal paper, you are actually performing a log transformation (the spatial construction of the graph paper is responsible for the transformation). (If you don't recall this, briefly review Module I, Unit 5.)

Therefore, in quality control terms, if we want to talk about the probability of making an "out of control" decision when, in fact, the process is in control, we use ________. (Greek letter)

Conversely, if we want to describe the probability of making an "in control" decision when the process is actually "out of control," we use ________. (Greek letter)

These two probabilities are often referred to as risks by quality control statisticians. In most quality control chart applications, the emphasis or interest is placed upon the α risk. That is, the statistical control limits are established based upon the α risk, with little or no direct consideration for the β risk.

The analysis of pollution monitoring data would not be too complicated if the log normal distribution were the only one (other than normal) that you might encounter. Unfortunately, there are many others, including the hypergeometric and Poisson. Discussion of these and others, however, will be left to the advanced course. It is important to remember that monitoring data may not be normally distributed, and that certain transformation may be necessary. It is emphasized that the above discussion of non-normal distributions is concerned primarily with the aggregate distributions of monitoring data. When, if ever, quality control charting is done with monitoring data, the above considerations must be taken into account.

Answer: α

β

Fortunately, in most quality control chart work, actual monitoring data are not directly used, since there is no information in the monitoring data themselves to reveal or indicate the quality of the data, from either precision or accuracy standpoints. For example, to measure and control accuracy, periodic measurements of a standard are made. To measure and control precision, the information in the differences between duplicate samples or duplicate analyses, for example, are used. And fortunately, most quality control data are normal or near-normal either in the original units of measurement, or as percentages based upon the level of the measurements. These points will be reinforced by the particular examples used in preparing control charts.

Because you select α when you select a significance level for a statistical test, you can control the probability of making a Type I error. It is important to remember, however, that and time you reduce α with the same amount of data, (e.g., from .05 to .01), you are increasing β. The exact amount of this increase in β is calculable, but this calculation will be left for an advanced course. The point is, that by decreasing α substantially, you may be increasing β enough to make committing a Type II error likely.

GO TO:

Frame 16

Page VII-1

32

To review:

This unit introduced the application of statistical concepts to the quality control of pollution measurement systems. Most of the unit dealt with definitions of terms used for statistical quality control charts. These definitions also appear in the glossary section of the Guide.

The most important point to remember is that the quality of pollution data are always open to challenge. To ensure that the data are not erroneous, certain precautions must be taken to promote their accuracy and precision.

Now, go to Unit 2.

UNIT 2 - CHARTING TECHNIQUES

1

Unit 1 of this module introduced the basic concepts of statistical quality control charts. The two most important of these basic concepts were precision and accuracy. Both can be measured relatively easily through the use of special measurement procedures and the charting techniques described in this unit, namely,

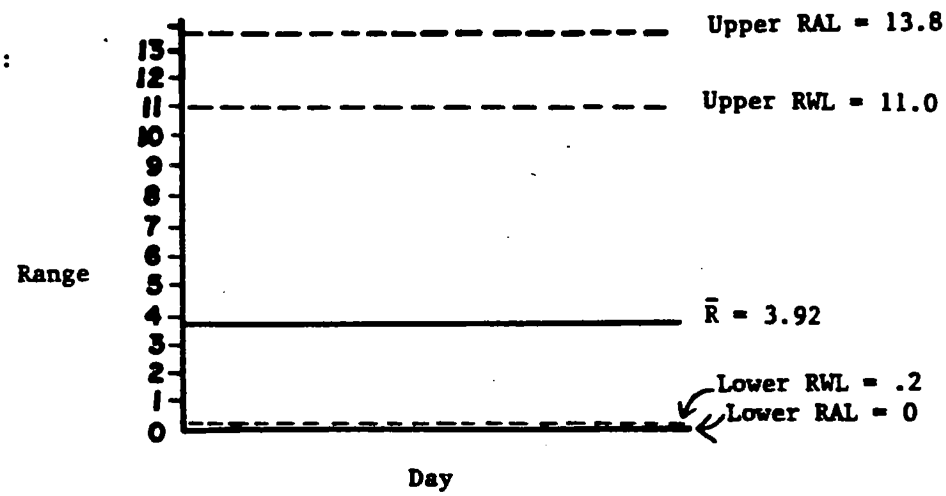
(a) Control charts for the range, relating to precision

(b) Control charts for the mean, or for individual values, relating to accuracy or bias

Continue now by reading pages 235-243 in the text, and then go to frame 2 of this unit. In your reading, pay particular attention to the setting of "warning" and "action" limits, and to the reasons behind "in control" and "out of control" decisions.

41a

Answer:



Note that although the original observed measurements were made to check the accuracy of the analytical process, the range chart itself is still only a check on the within-day precision of the method. Note that all of the individual range values are well within the warning limits.

The section of the text you just read discussed the use of control charts in production processes. Since quality control of measurement systems differs only slightly from production quality control, the techniques we will discuss later in this unit will also be only slightly different. There are two basic approaches in establishing and maintaining control through the use of control charts.

1. A simple approach based simply on data from some past period of satisfactory performance.
2. The classical approach based on the rational subgroup concept and range values obtained from these subgroups.

Let's consider the classical approach first.

G. To provide a check on the day-to-day accuracy of the analytical process we must plot the mean chart. Calculate the following limits:

Upper $\bar{D}$WL = $A_W \times \bar{R}$ = # ____________ Upper $\bar{D}$AL = $A_A \times \bar{R}$ = # ____________

Lower $\bar{D}$WL = $-A_W \times \bar{R}$ = # ____________ Lower $\bar{D}$AL = $-A_A \times R$ = # ____________

3

You're probably wondering what is meant by the "rational subgroup" concept.

Under the rational subgroup concept, the inherent variability of a process is reflected by the uniformity existing under some specified local conditions. For example, if the repeatability of a method within a day for a given sample is considered the inherent and random variation of the method, then it should be possible to control the day-to-day variability within limits corresponding to the within-day variability. In this case, data from analyses made within a day are considered a "rational subgroup" and are used as a basis for control chart limits applying to averages (or individual values) from day to day. This approach is the classical or traditional approach which has been used very successfully in industry and which applies equally well to many pollution measurement situations. The objective of this approach is to detect any group-to-group (such as day-to-day) variations which are not expected when considering only the inherent within-group variability.

Now, let's go back to the simple approach, based upon data from some past period of satisfactory performance.

42a

Answer:

Upper $\bar{D}$WL = 1.23 x 3.92 = 4.82

Lower $\bar{D}$WL = -1.23 x 3.92 = -4.82

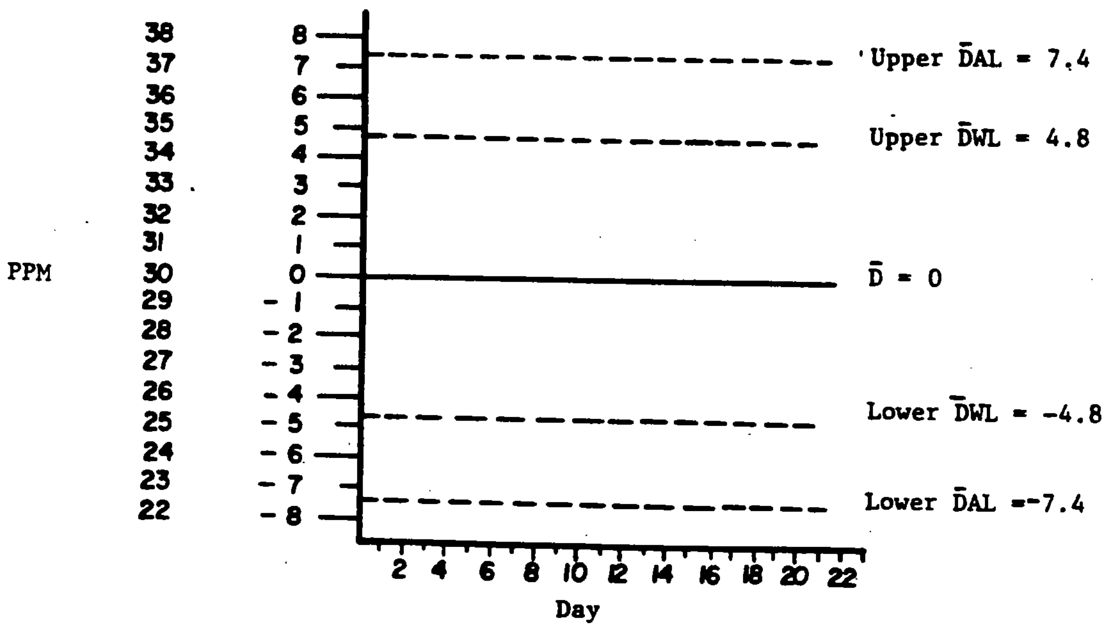
Upper $\bar{D}$AL = 1.88 x 3.92 = 7.37

Lower $\bar{D}$AL = -1.88 x 3.92 = -7.37

The simple approach will permit some between-group (such as between-day) variation beyond that expected from within-group (within-day) considerations. Although some significant between group variations occur these also are considered to be inherent and uncorrectable in the method and the extent of these variations are considered acceptable and permissable from a practical standpoint.

With this approach, data from some past period of acceptable performance, with any outliers removed, are used as a basis for establishing control limits to detect any possible future excursions exceeding the variations of the past acceptable performance.

H. Now we graph the limits in the following manner: 43



The zero line is termed the Standard Nominal Value. Note that the grand average of all the original measurements is 29.93 ppm, very close to the "Known" value of 30 ppm. (The 29.93 ppm corresponds to the grand average of the $\bar{D}$ of -.07 ppm) It is consequently concluded that <u>on the average</u> for the 17 days the analytical process was very accurate, and therefore justifies the use of zero (corresponding to 30 ppm) as the center line for the control chart. Using this chart we can now plot D values. Accuracy is acceptable, provided the points remain within the DAL's. Plot the data (in frame 38a) on the chart. Circle any points out of control.

The statistical computations involved in establishing control limits from some past period of acceptable performance (the simple approach) are as follows:

1. Calculate the average, $\bar{x}$.

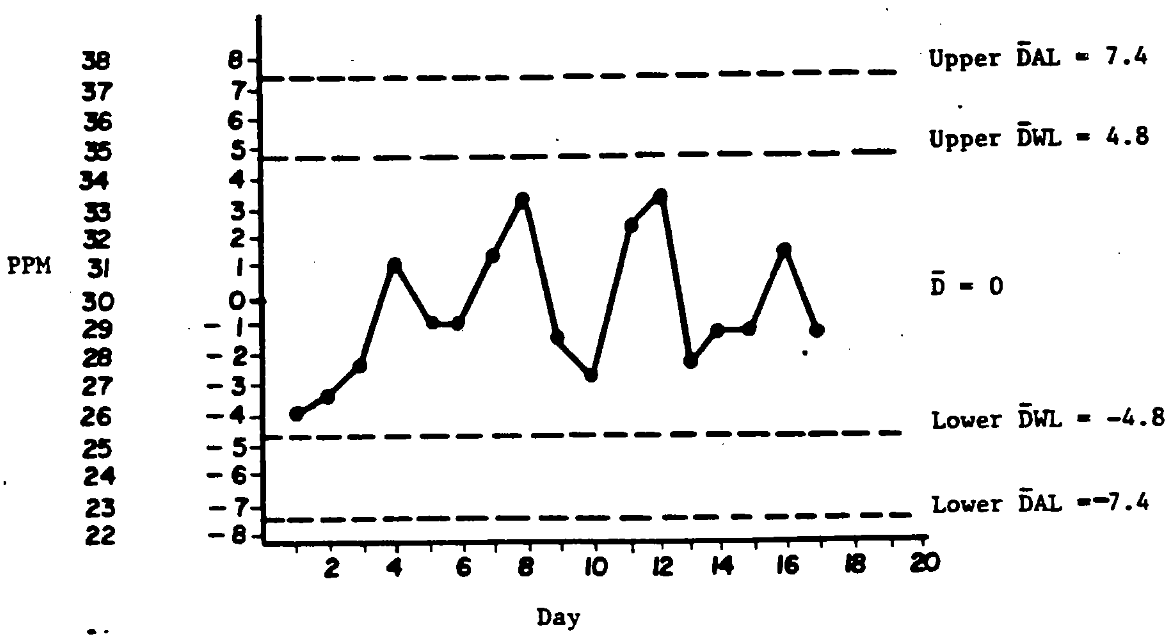
2. Calculate the standard deviation, s, of all the individual data values.

3. Compute $\bar{x} \pm 2s$ to be used as warning limits.

4. Compute $\bar{x} \pm 3s$ to be used as action limits.

The + 2s limits would constitute approximately 95% limits (actually 95.4%) and the + 3s limits would constitute 99.7% limits. The individual data values would be plotted on these control charts. Such charts could be used whether or not any rational subgroups exist. However, if there is a basis for rational subgroups (such as within-day subgroups) range charts for the within-subgroup ranges could also be used to control the within-subgroups (e.g., within-day) precision, and mean charts could also be used to control subgroup-to-subgroup (day-to-day) variation. Both the simple and classical approaches will be described in this module.

Answer: No points are out of control



First, let us look at the data on frame 1000. Fold out frame 1000 and review the situation and the data. The first thing which ought to be done with the data is to look at its distribution by preparing a histogram tally.

Tally the histogram on the scale below:

105
104
103
102
101
100
99
98
97
96
95

Note that all of the points a⁻e within the action limits. Note also that we went to a lot of trouble getting the "Observed-Known" differences when we could have ended up with the same result had we worked with the averages for each day. However, we needed the differences in order to compute the control limits. It would be necessary to use the "Observed-Known" difference, however, if the level of the knowns changed as in the following problem.

6a

Answers:

105	
104	////
103	///
102	///// ////
101	///// ///// //
100	///// /////
99	////
98	/////
97	////
96	////
95	//

7

Note that the results appear to be fairly normal, although there appears to be a slight aversion to reporting 99! Also, there does not appear to be any outliers in the data. The next step would be to compute the average and the standard deviation. Record your answers below:

$\bar{X}$ =

s =

45

Fold frame 4000 back in, and fold out frame 5000. The data shown are readings taken daily during the month of August. The simple approach will be taken to establish control chart limits for this situation. First, there are no rational subgroups with which to compute precision estimates since only one result is obtained each day and the laboratory operation is continual, i.e., seven days each week. Further, the variability exhibited in the results are considered acceptable.

Note: In this case, there are no rational subgroups or ranges on which to base control limits. Consequently no range chart can be prepared for precision. Actually, in this case, the control chart which is prepared is a control of accuracy and precision combined.

7a

Answer: $\bar{X}$ = 100.00

S = 2.34

8

Next, compute the warning limits, $\bar{X} \pm 2s$, and the action limits, $\bar{X} \pm 3s$.

$\bar{X}$ + 2s = __________ ± 2 (   ) = ________ and __________

$\bar{X}$ + 3s = __________ ± 3 (   ) = ________ and __________

46

As described earlier, in such situations, it is necessary only to calculate the average, $\bar{x}$, and the standard deviation, s, of the past history and compute $\bar{x} \pm 2s$ for warning limits, and $\bar{x} \pm 3s$ for action limits. (If necessary, refer back to Module I pages I-24 through I-28 for the method of computing the standard deviation). To perform the necessary calculations, fill in the necessary information to complete frame 5000.

(Check your answers in Frame 46a)

GO TO

Frame: 46a
Page: VII-29

8a

Answers: $\bar{X} + 2s = 100.00 \pm 2(2.34)$

$= 104.68$ Upper Warning Limit

$= 95.32$ Lower Warning Limit

$\bar{X} + 3s = 100.00 \pm 3(2.34)$

$= 107.02$ Upper Action Limit

$= 92.98$ Lower Action Limit

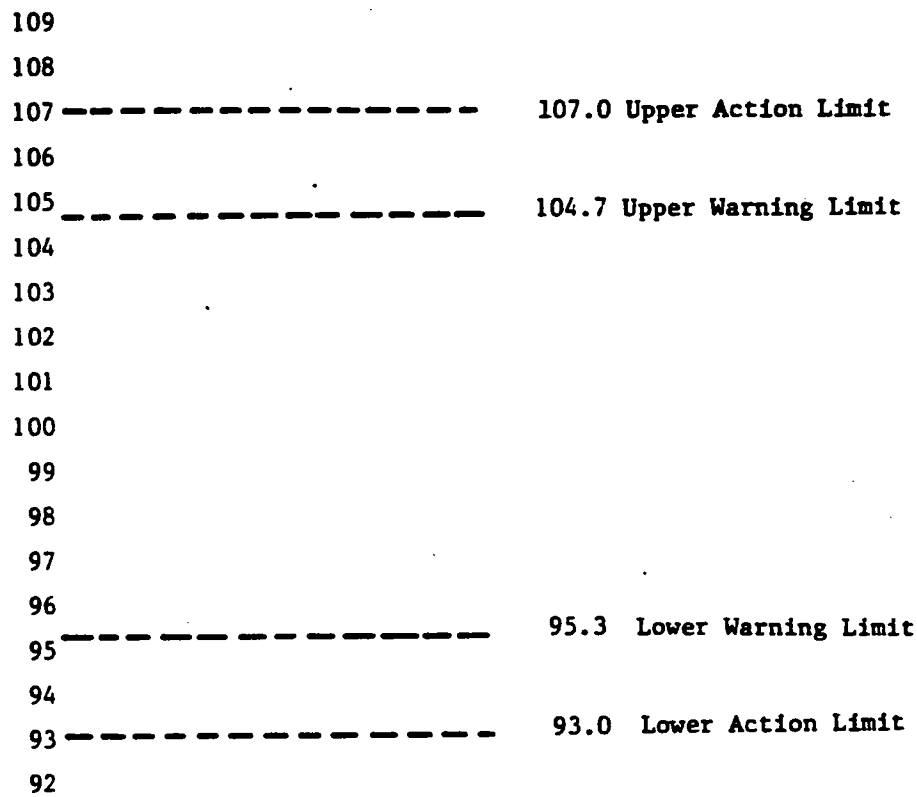
9

Thus, all of the values are within the warning and action limits. Except for indications of trends (to be discussed later), any future value outside the action limits would indicate an out-of-control situation which requires corrective action. Construct the control chart with the above limits.

108
107
106
105
104
103
102
101
100
99
98
97
96
95
94
93
92

9a

Answer:



10

Next, let's consider the construction of the typical mean and range control charts using the classical approach based on rational subgroups. In control chart terminology the combination of mean, $\bar{X}$, and range, R, charts are commonly referred to as "X-bar, R" charts.

Control charts based upon rational subgroups have been in use since the 1920's. As originally developed, the control limits were calculated from standard deviations. Subsequent statistical developments made possible the use of the range, rather than the standard deviation, for the calculation of control limits. Since the range is much easier to calculate (with little loss of efficiency) use of the range accelerated the application of control charts on a wide scale.

The specific development that made control charts more easily constructed from ranges was the derivation of a series of constants, based on sample size, that could be used in the calculation of warning and action limits. Some of these constants to be used later are listed in Table K in the Guide.

Answer:

D = Observed - Known

-5
2
0
-3
1
-1
0
3
-2
5
-4
3
-2
0
-1

Note that all of the constants in Table K do not agree with those listed in Tables A-12 and A-13 of the text, Neville and Kennedy. The reason the constants are different is because Neville and Kennedy uses 95% limits as warning limits and 99.8% limits as action limits, whereas this self study course uses 95% and 99%, respectively. In contrast, most American industrial quality control texts and practice simply use 2-sigma (2 standard deviations) and 3-sigma limits, respectively.

	Industrial (American Practice)		Neville & Kennedy (British Practice)		Table K, SI-473	
	Probability	Std. Dev.	Probability	Std. Dev.	Probability	Std. Dev.
Warning Limits	.954	2σ	.95	1.96σ	.95	1.96σ
Action Limits	.997	3σ	.998	3.09σ	.99	2.58σ

For practical purposes, any of the above sets of limits are equally satisfactory, as long as the appropriate probabilities are understood.

D. Calculate the average difference, $\bar{D}$, and the standard deviation, s.

Note: The signed D values should be used for these calculations.

$$D = \frac{\Sigma D}{n}$$ n = no. of individual data values (D values)

$\bar{D}$ = ______________

$$s = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n - 1}}$$

= ______________

13

Let's continue by calculating the limits for the data contained in frame 2000 located at the end of this unit. Please unfold this frame now.

In the example in frame 2000, the five (5) measured values each day are considered a rational subgroup since greater agreement might be expected from the values within each day than across all days involved. With limits established on the within-day variability, any significant between-day change in level will be most effectively detected.

A. The first step in the calculation of the control limits is to determine the mean and range of each subgroup. We simply use the same formula that we've used before.

(1) $\bar{x} = \frac{\Sigma x_i}{n}$

(2) R = maximum value minus minimum value

Fill in these values on frame 2000. (To check your answers see frame 13a).

47a

Answer: $\bar{D} = \frac{\Sigma D}{n} = \frac{-5 + 2 + \ldots + (-1)}{15} = \frac{-4}{15} = -.27$

$$S = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n-1}} = \sqrt{\frac{108 - \frac{(-4)^2}{15}}{14}} = 2.76$$

Answer:

Day	$\bar{x}_5$	R_5
1	99.6	3
2	100.6	9
3	100.4	5
4	99.0	6
5	99.8	5
6	100.0	8
7	99.4	5
8	100.4	7
9	100.2	4
10	99.6	9
11	100.0	8
12	101.0	2

Note that we use the subscript, 5, to indicate the sample size of the subgroup.

E. Calculate the control limits according to the formulas:

Upper action limit = $\bar{D} + 3s$

Upper warning limit = $\bar{D} + 2s$

Lower warning limit = $\bar{D} - 2s$

Lower action limit = $\bar{D} - 3s$

B. Next, find the grand mean ($\bar{\bar{x}}$ or $\bar{x}_{..}$). Because the subgroups are the same size, we can simply add the subgroup means and divide by the number of subgroups (k).

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \ldots\ldots\ldots\ldots\ldots + \bar{x}_k}{k}$$

$\bar{\bar{x}}$ = # ______________________

Answer: Upper action limit = -.27 + 3(2.76) = 8.0
Upper warning limit = -.27 + 2(2.76) = 5.3
Lower warning limit = -.27 - 2(2.76) = -5.8
Lower action limit = -.27 - 3(2.76) = -8.6

Note: In the absence of a bias between the supervisor (or auditor) and the technician, the average would be expected to be zero. In this example, the actual bias, -.27, is not significantly different from zero. (See Module III, Unit 2 for the appropriate t test.) Consequently, more appropriate limits of

zero $\pm$ 2s for warning limits

and

zero $\pm$ 3s for action limits

or

$0 \pm 2(s) = \pm 5.5$

and

$0 \pm 3(s) = \pm 8.3$

should be used.

14a

Answer: $\bar{\bar{x}} = 100.00$

15

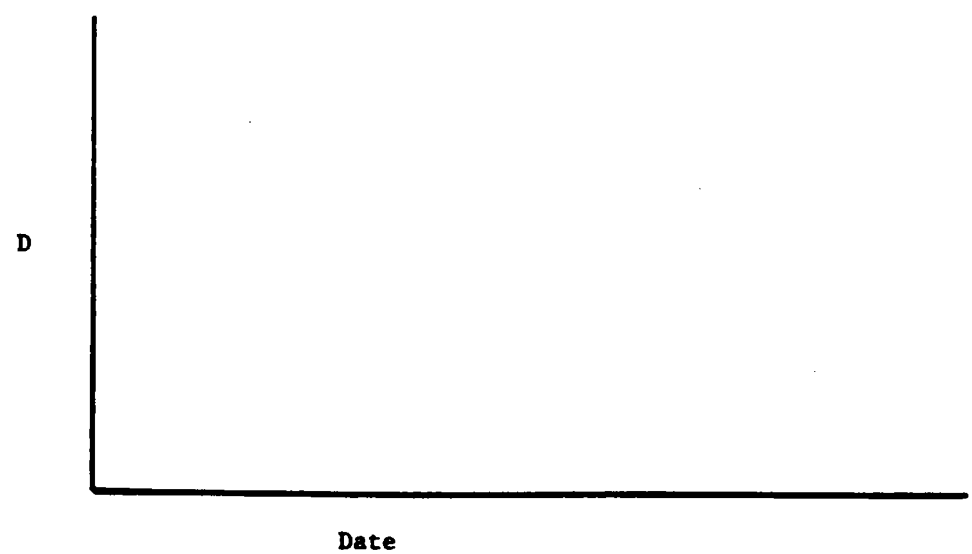
C. Now we want to find the average range ($\bar{R}$)

$$R_5 = \frac{\Sigma R_5}{k}$$

$\bar{R}_5 = \#$ ___________

49

F. Construct the chart using the limits just calculated.



Answer: $\bar{R}_5 = 5.92$

D. Now we can determine our warning and action limits by the following formulas:

Upper Mean Warning Limit = $\bar{\bar{x}} + A_W\bar{R}$

Lower Mean Warning Limit = $\bar{\bar{x}} - A_W\bar{R}$

Upper Mean Action Limit = $\bar{\bar{x}} + A_A\bar{R}$

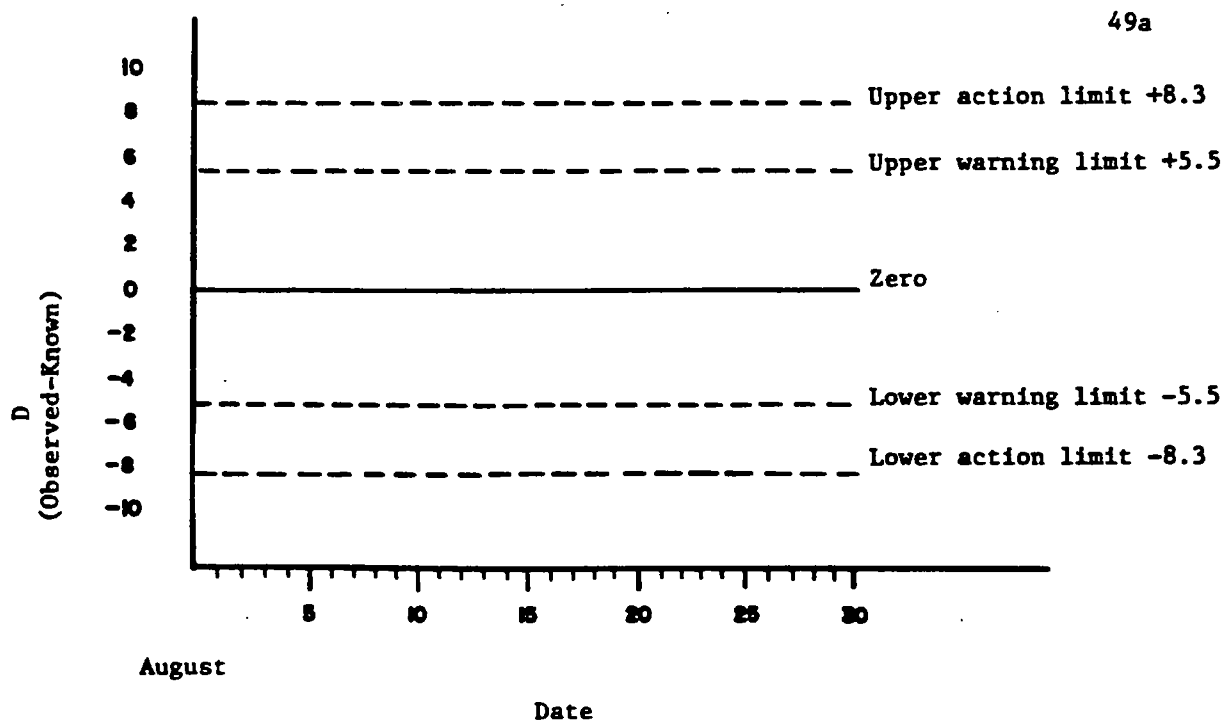
Lower Mean Action Limit = $\bar{\bar{x}} - A_A\bar{R}$

The values for the constants A_W and A_A are found in Table K. (Since we have five values in each subgroup, n = 5.)

Upper MWL = ____________________ Upper MAL = ____________________

Lower MWL = ____________________ Lower MAL = ____________________

Answer:



Answer: $A_W = 0.38$

$A_A = 0.58$

Upper MWL = 100 + (0.38 x 5.92) = 102.2

Lower MWL = 100 - (0.38 x 5.92) = 97.8

Upper MAL = 100 + (0.58 x 5.92) = 103.4

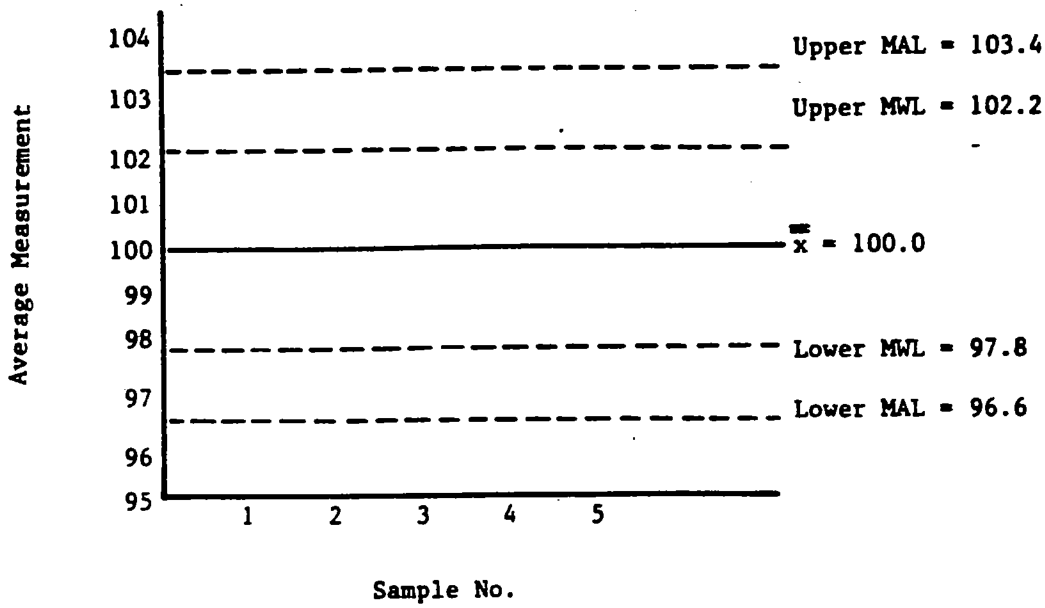
Lower MAL = 100 - (0.58 x 5.92) = 96.6

Note that these limits are considerably tighter than the limits previously calculated because these limits apply to the daily averages of 5 values each, whereas the previously calculated limits applied to individual values.

50

Now plot the fifteen data points on the chart. Circle any points out-of-control.

E. We now construct a graph as follows:



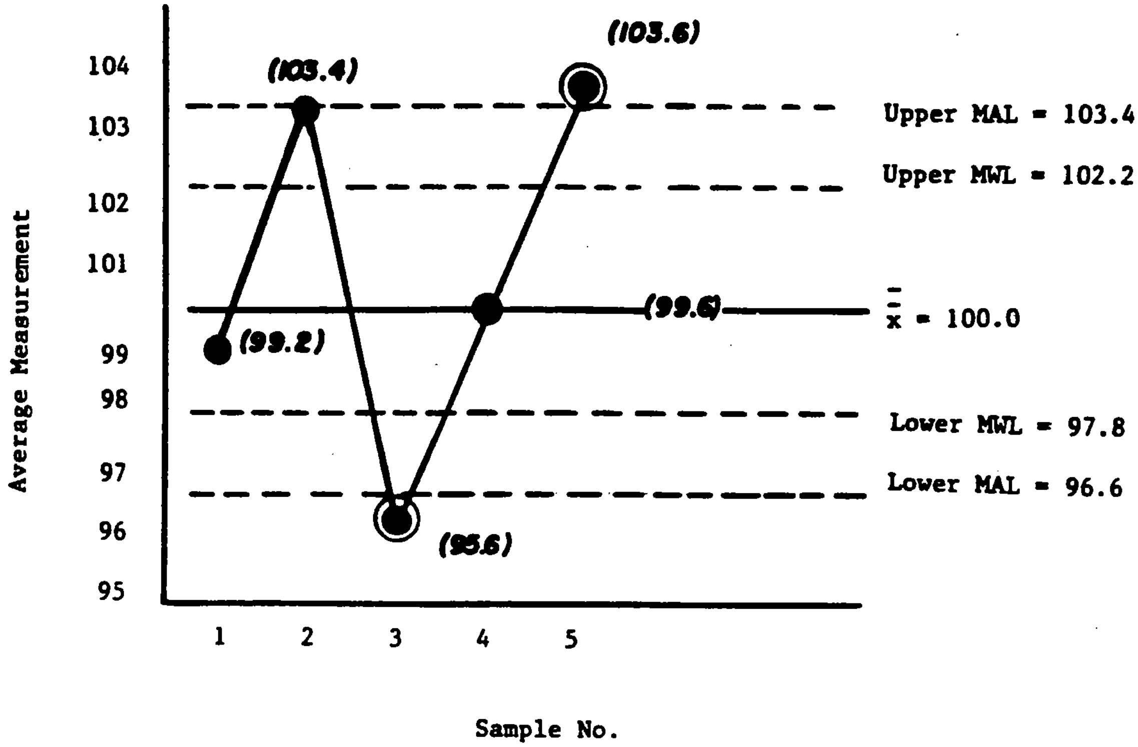
On this graph we can now plot any new data we collect from the same measurement process to test for control. Plot the means of the following subgroups and circle any that represent an out-of-control condition:

Sample	Measurements					$\bar{x}$
1	97	101	100	102	96	#________
2	102	104	103	105	103	#________
3	95	96	100	93	94	#________
4	101	103	98	97	99	#________
5	106	102	103	104	103	#________

GO TO

Frame: 50a
Page: VII-41

Answer:



The sample means are plotted above and listed below:

$\bar{x}_1 = 99.20$

$\bar{x}_2 = 103.40$

$\bar{x}_3 = 95.60$ out of control (below lower control limit)

$\bar{x}_4 = 99.60$

$\bar{x}_5 = 103.60$ out of control (above upper control limit)

NOTE: A better picture of the variability is obtained when the plotted points are connected by lines as shown above.

Since we discovered two (however, one is sufficient) "out-of-control" subgroups (i.e., they exceeded the action limits which are based on the 99% level of confidence), we should stop the analytical process and check it to isolate and correct the cause of the problem. Also, some appropriate action needs to be taken for the routine samples analyzed on the out-of-control days, such as rerunning the samples if still available, or invalidating the data if the samples are no longer available.

It is emphasized that quality control data should be obtained and plotted in a timely manner, i.e., as soon as the data are available, so that appropriate corrective action can be taken as soon as possible to prevent the loss of future data. In addition, timely action may permit the correction of the routine data in question such as reanalyzing the samples before the samples have been discarded.

F. Let's now use the data in frame 2000 to construct a range chart.

The formulas for the limits on the range chart are:

Upper RWL = $D_{WU} \times \bar{\bar{R}}$

Lower RWL = $D_{WL} \times \bar{\bar{R}}$

Upper RAL = $D_{AU} \times \bar{\bar{R}}$

Lower RAL = $D_{AL} \times \bar{\bar{R}}$

We already know $\bar{\bar{R}}$ = 5.92 and the values for D_{WU}, D_{WL}, D_{AU}, and D_{AL} are located in Table K. (Again, n = 5.)

Upper RWL = # ________________

Lower RWL = # ________________

Upper RAL = # ________________

Lower RAL = # ________________

19a

Answer: Upper RWL = 1.81 x 5.92 = 10.72

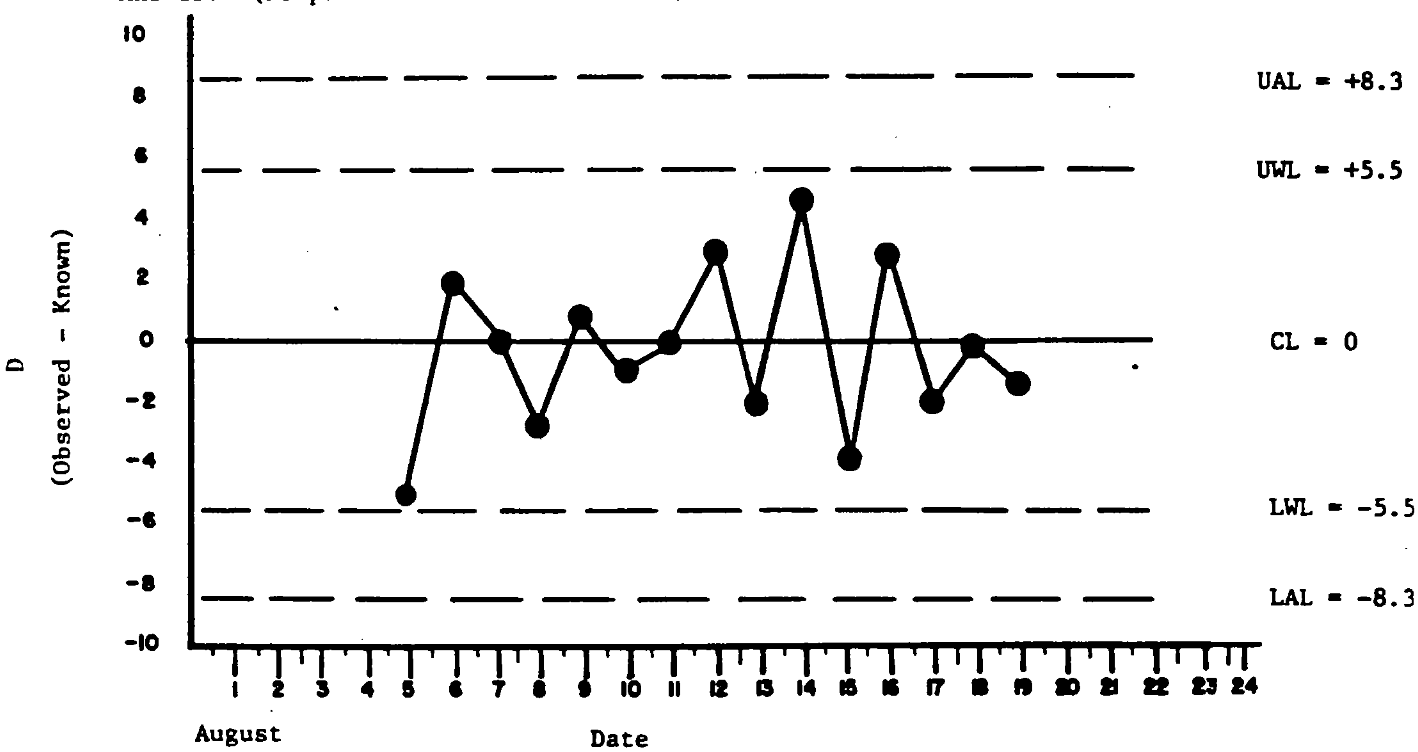
Lower RWL = .37 x 5.92 = 2.19

Upper RAL = 2.10 x 5.92 = 12.43

Lower RAL = 0.24 x 5.92 = 1.42

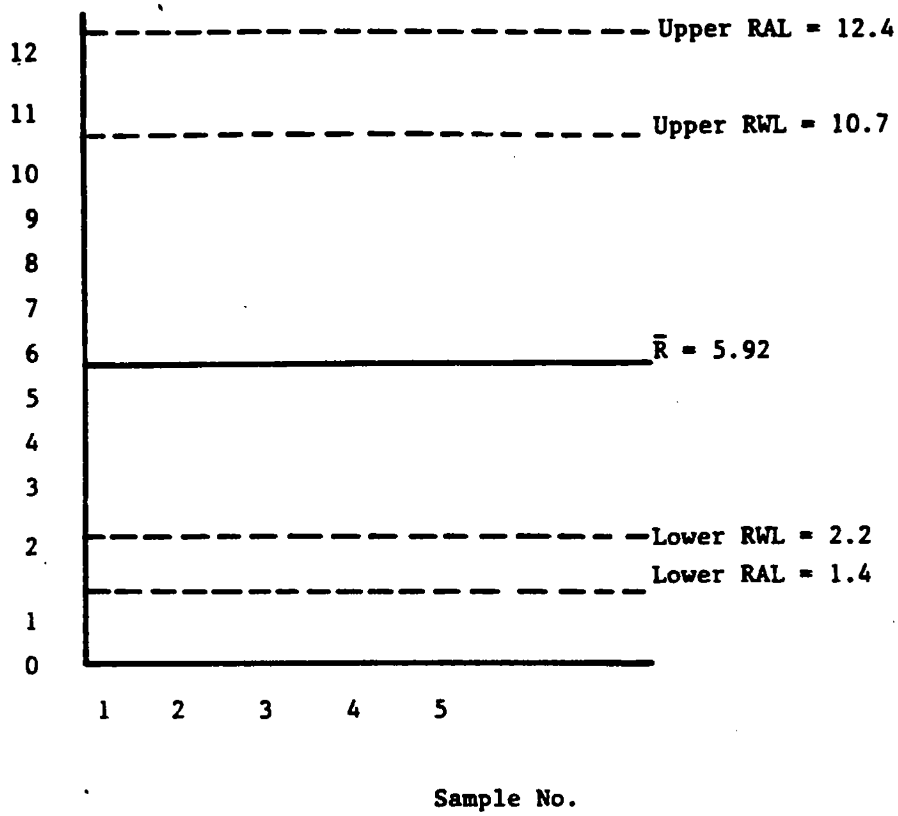
50a

Answer: (No points are out-of-control)



20

As before, we construct a range chart:



51

Data collected for the next five days is shown below. Calculate the data points for the control chart and plot them on the chart in frame 50a. Circle any points out-of-control.

	Observed	Known	D, Observed - Known
August 20	85	87	
21	64	67	
22	43	42	
23	74	75	
24	77	71	

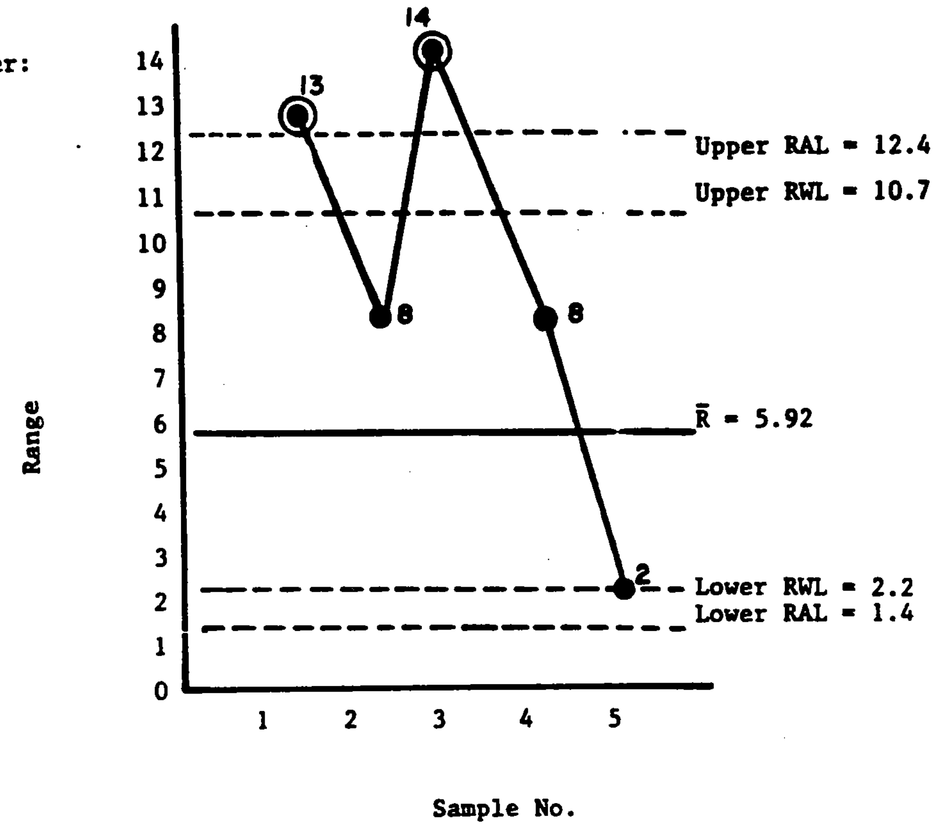
GO TO

Frame: 51a
Page: VII-46

Using the chart in Frame 20, we can now plot any new range data from this analytical process to test for control. For each of the following samples, calculate R, plot them on the chart in frame 20, and circle any "out-of-control" samples.

Sample	Measurements					R
1	106	103	93	97	101	# ______
2	104	97	103	96	100	# ______
3	107	100	94	93	106	# ______
4	102	98	104	96	100	# ______
5	100	101	99	100	100	# ______

Answer:



R_1 = 13 out-of-control

R_2 = 8

R_3 = 14 out-of-control

R_4 = 8

R_5 = 2

<u>You may have noticed</u>: the mean of each sample in the previous example is exactly equal to 100. Therefore, if we were to plot these means on the mean chart, we would not obtain an "out-of-control" condition, even though the variation with each sample is very high for samples 1 and 3. This points out the need for BOTH mean and range charts to determine control.

<u>Also Note:</u> The data used in the previous example to establish the control limits, by design, resulted in no points being out-of-control for that set of data. If the original set of data would have resulted in out-of-control points the data from those particular days should have been eliminated and the limits recomputed.

In other words, all outliers should be eliminated from the body of data used to compute control limits. Obviously, it doesn't make sense to compute limits representing in-control conditions from out-of-control data.

We're through with frame 2000; please fold it back in.

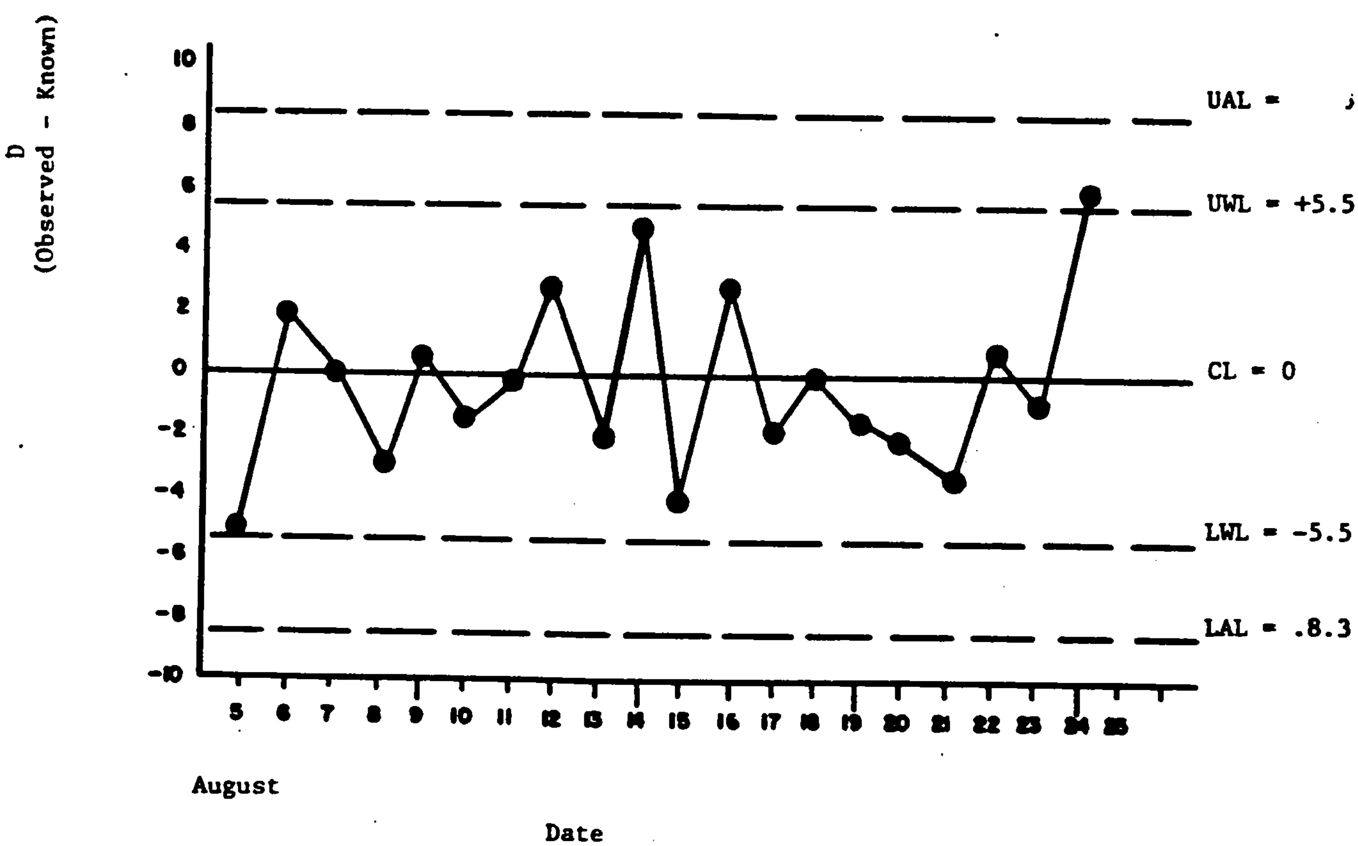
Wait a second--quick review. Remember, we are interested in controlling <u>precision</u> and <u>accuracy</u>:

(<u>Precision/Accuracy</u>) refers to the repeatability among repeated observations of a phenomenon.

<u>(Precision/Accuracy)</u> refers to a degree of difference between observed and known, or actual, values.

			D
Answer:	August	20	-2
		21	-3
		22	1
		23	-1
		24	6

No points are out of control.



Fold Frame 5000 Back In

23a

Answer: Precision

Accuracy

24

In the previous example, the mean or $\bar{x}$ chart would be used to control accuracy since the measured values were of a standard. The range or R chart would be used to control precision, since the agreement within a subgroup is a measure of precision. In some cases, the $\bar{x}$ chart may relate only to relative accuracy or bias. Tests for precision require the running of samples in replicate or the measurement of a standard of some type, the value of which is considered "Known". Tests for relative bias are involved when for example, one of a pair of duplicates is analyzed under one condition (e.g., by one analyst) and the second of the pair is analyzed under another condition (e.g., by a second analyst.)

52

Fold out frame 6000, another example of typical quality control data. In this particular example, we do not necessarily have rational subgroups since we might expect or suspect by eyeball analysis that a bias exists between the results of Sampler A and Sampler B (or between the results of Analyst A and Analyst B). Also since the concentration levels vary considerably, we feel that we should compute percent differences rather than actual differences. And now, we introduce a new consideration. Since we suspect a bias, we are interested in computing the <u>signed</u> percent differences. Compute these <u>signed</u> percent differences using the following formula.

$$\% D = \frac{(X_A - X_B) \times 100}{(X_A + X_B)/2}$$

and write these values on frame 6000.

Next, let's look further at the construction of precision control charts. In data quality control, range charts are used to test precision. Typically, precision charts are developed by collecting data for many samples, usually a minimum of 15 or 20, each run in replicate, under assumed controlled conditions. In laboratory work, duplicates are usually adequate and less costly than running triplicates, etc. Once these data have been collected, we follow steps very similar to those we used before to construct range charts. Let's use an example again. Unfold frame 3000.

A. List the range (R_2) for each sample in frame 3000. (Check your answers in frame 25a.)

52a

Answers: %D

-9.1
-5.7
0
-1.4
-6.7
7.1
-22.2
-9.8
50.7
-17.4
1.8
-11.8
6.7

53

Do you notice anything unusual about the data as a group?

1. ____________________

Do you notice anything unusual about a particular value in the group?

2. ____________________

GO TO

Frame: 53a
Page: VII-50

Answer:

Sample	A	B	R_2
1	20	22	2
2	15	17	2
3	19	16	3
4	18	13	5
5	21	22	1
6	9	11	2
7	13	13	0
8	17	15	2
9	16	19	3
10	21	19	2
11	7	8	1
12	22	19	3
13	14	16	2
14	12	15	3
15	19	20	1

26

You may now be asking, "Isn't the R we just calculated really a difference?" The answer is a qualified yes, in a sense. It is the absolute difference between paired observations. The reasons for using R are (1) it eliminates negative numbers, since it is the maximum minus the minimum; and (2) given samples of size two, the two members of each sample automatically become the maximum and minimum values of that sample, so we are really figuring the range of each sample.

Later, we will use the signed difference in constructing another type of control chart.

Let's move on.

B. Calculate the average range ($\bar{R}$)

$$\bar{R}_2 = \frac{\Sigma R_2}{k}$$

k = number of subgroups

$\bar{R}_2$ = #______________

53a

Answers:

1. <u>Most of the values are negative</u>, indicating a bias between the results of Sampler A and Sampler B (or between Analyst A and Analyst B).
2. The Value of plus 50.7% seems to be an outlier compared to the rest of the group. The fact that it is an outlier can be confirmed by testing, using an appropriate test of Module I pages 121-131.

54

A graphical distribution of the data can be made on a scale plot. Place an X on the scale below where each value falls.

27a

Answer: $\bar{R}_2 = \frac{32}{15}$

$= 2.1$

28

C. Calculate the limits according to the formulas:

Upper RWL = $D_{WU} \times \bar{R}$

Lower RWL = $D_{WL} \times \bar{R}$

Upper RAL = $D_{AU} \times \bar{R}$

Lower RAL = $D_{AL} \times \bar{R}$

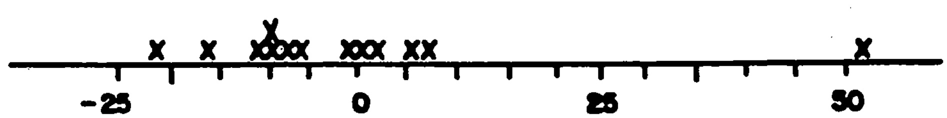
The values for D_{WU}, D_{WL}, D_{AU}, and D_{AL} are in Table K. This time we have two observations in each sample, so n = 2.

Upper RWL = #____________ Upper RAL = #____________

Lower RWL = #____________ Lower RAL = #____________

54a

Answer:



28a

Answer: Upper RWL = 2.81 x 2.1 = 5.90

Lower RWL = 0.04 x 2.1 = .08

Upper RAL = 3.52 x 2.1 = 7.39

Lower RAL = 0.01 x 2.1 = .02

29

Note that with one exception, all of the range values from the data fall within the limits. The highest value, 5, is less than the upper warning limit. The zero value is less than lower limits, however, probably due to rounding off in obtaining the individual results. Besides, there is not as much concern about range values below the lower limits, indicating better-than-normal precision, as there is with range values above the upper limits, indicating poorer-than-normal precision.

55

The outlier is even more evident when looking at the plot in Frame 54a. It should be noted in this situation that the results from one of the two samplers is very questionable, although it is not known which one, without any additional facts or information. Therefore, both values 9.4 and 5.6 should be invalidated, or some other action, such as reanalysis, should be taken in an attempt to resolve the large difference. If the large difference can not be appropriately corrected, the values and the difference should be eliminated from consideration in establishing control limits.

Control limits are determined by computing the average and the standard deviation of the signed percentage differences. Compute these values.

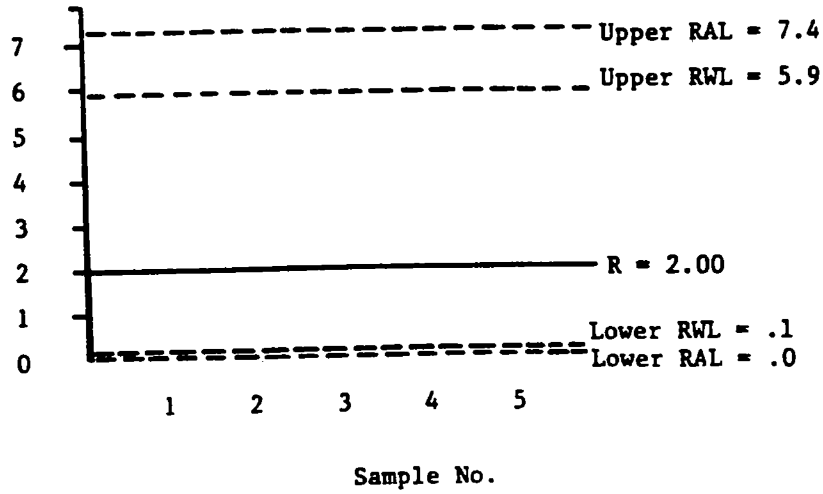
Before omitting outlier $\bar{X}$ = __________

$\bar{s}$ = __________

After omitting outlier $\bar{X}$ = __________

$\bar{s}$ = __________

D. Let's construct the chart



Assuming we now have a new set of samples (of the same procedure) run in duplicate, determine on the above chart whether the measurements are precise (i.e., fall within the RAL's).

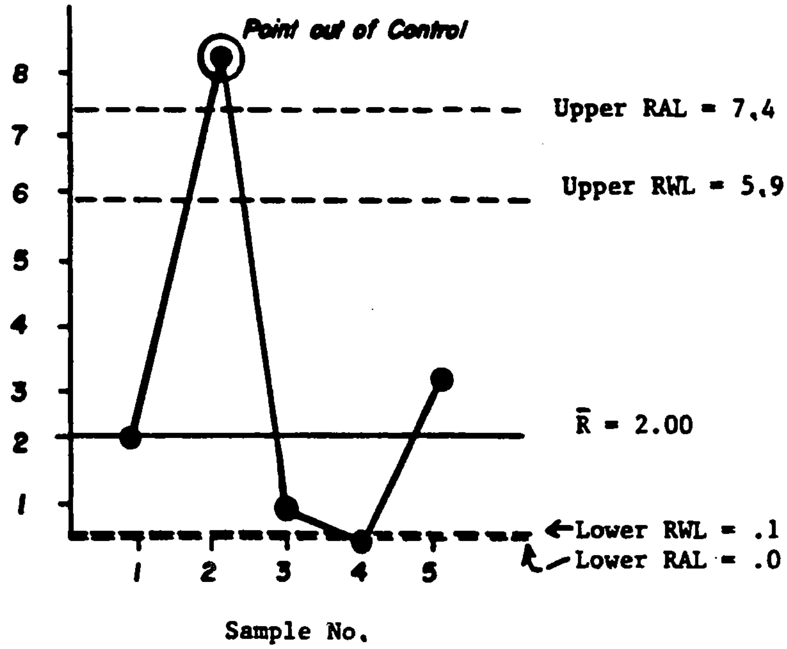
Sample	A	B	R
1	17	19	______
2	26	18	______
3	15	16	______
4	19	19	______
5	18	21	______

55a

Answers:	Before	After
	$\bar{X}$ = -1.37	$\bar{X}$ = -5.71
	s = 17.88	s = 9.05

ANSWER:

SAMPLE	R
1	2
2	8
3	1
4	0
5	3



The sample No. 2 is not precise. The system should be shut down and the cause determined. After determining the cause of imprecision, for sample 2, consideration should be given to re-running old routine samples analyzed in that same period of time, i.e., between the time of analyzing samples 1 and 3.

Please fold frame 3000 back in.

56

The appropriate t-test (Module III) can now be computed to confirm the suspected bias that Sampler (or Analyst) A produces results which are generally lower than those from Sampler (or Analyst) B. Corrective action should be initiated to determine and eliminate the cause of the bias since the results of the paired duplicate data should agree within the limits of precision of the operations involved.

The next question: What should the control limits be?

either 1. $\bar{x} \pm 2s$ Warning limits
$\bar{x} \pm 3s$ Action limits

or 2. zero $\pm$ 2s Warning limits
zero $\pm$ 3s Action limits

Your answer ______________________.

GO TO

FRAME: 56a
PAGE: VII-56

Let's try another one. In this example, we will introduce an additional complexity, namely the need to work with percentage or relative range rather than the actual range values. As in much pollution data, the level of the samples may vary widely. When analyzing duplicates we expect to get better agreement at low values. However the percentage range is expected to be stable.

A. List the range and % range for each set of samples:

Sample	X_1	X_2	R	%R*
1	16	19		
2	74	81		
3	23	21		
4	34	38		
5	98	84		
6	9	11		
7	18	17		
8	26	29		
9	67	74		
10	12	9		
11	50	41		
12	107	91		
13	22	21		
14	19	15		
15	68	61		

* The %R is calculated by the following formula

$$\%R = \frac{|x_1 - x_2|}{(x_1 + x_2)/2} \times 100 = \frac{R}{\bar{x}} \ (100)$$

31a

Sample	X_1	X_2	R	%R
1	16	19	3	17.1
2	74	81	7	9.0
3	23	21	2	9.1
4	34	38	4	11.1
5	98	84	14	15.4
6	9	11	2	20.0
7	18	17	1	5.7
8	26	29	3	10.9
9	67	74	7	9.9
10	12	9	3	28.6
11	50	41	9	19.8
12	107	91	16	16.2
13	22	21	1	4.7
14	19	15	4	23.5
15	68	61	7	10.9

Note that the higher range values are associated with the higher level samples.

32

B. Calculate $\overline{\%R}$

$$\overline{\%R} = \frac{\Sigma \%R}{K} \quad K = \text{number of samples}$$

$\overline{\%R}$ = # ____________

56a

Answer: 2. If the cause of the bias can or should be eliminated, the answer is 2, since the significant bias should not exist in the measurement system.

Therefore, zero $\pm$ 2s = ____________ $\pm$ 2(   ) = $\pm$ ____________

zero $\pm$ 3s = ____________ $\pm$ 3(   ) = $\pm$ ____________

57

Compute the above limits.

Answer: $\overline{\%R} = \frac{211.9}{15}$

$= 14.13$

The limits for a percent or relative range chart are computed in the same way as for range charts except that the average percent range is used instead of the average range.

C. Calculate the limits for the chart.

Upper RWL = $D_{WU} \times \overline{\%R}$

Lower RWL = $D_{WL} \times \overline{\%R}$

Upper RAL = $D_{AU} \times \overline{\%R}$

Lower RAL = $D_{AL} \times \overline{\%R}$

Values for D_{WU}, D_{WL}, D_{AU}, and D_{AL} are in Table K.

Upper RWL = # ____________

Lower RWL = # ____________

Upper RAL = # ____________

Lower RAL = # ____________

Answer: zero $\pm$ 2s = 0 $\pm$ 2(9.05) = $\pm$ 18.10

zero $\pm$ 3s = 0 $\pm$ 3(9.05) = $\pm$ 27.15

Note: It is very important to work with signed differences when working with paired differences, or, as in this case, signed percentage differences. Where one member of each pair is consistently associated with one variable and the second member of each pair is associated with a second variable, use the signed difference for plotting on the control chart will facilitate detection of a bias whenever one develops.

Fold Frame 6000 back in.

GO TO

Frame: 58
Page: VII-59

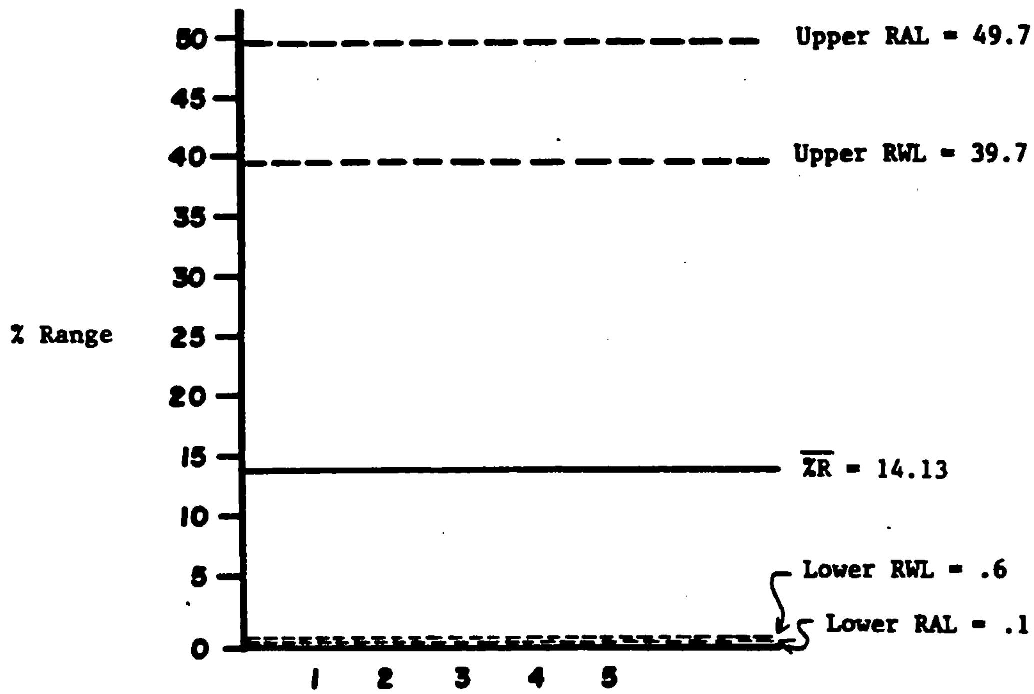
Answer: Upper RWL = 2.81 x 14.31 = 39.7

Lower RWL = 0.04 x 14.13 = 0.6

Upper RAL = 3.52 x 14.13 = 49.7

Lower RAL = 0.01 x 14.13 = 0.1

D. Construct the chart.

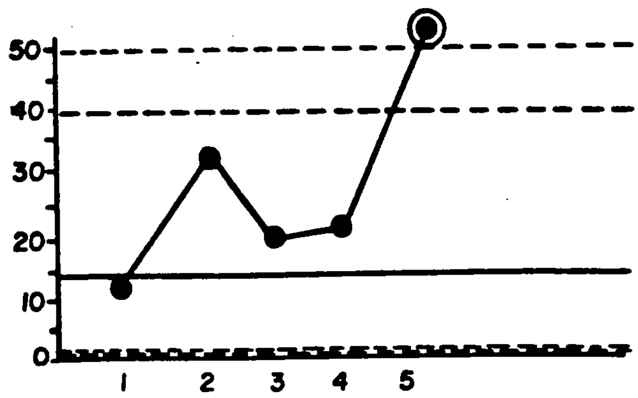


E. Using this chart, test the precision of the following data.

Circle any points that are not precise.

Sample	A	B	R	%R
1	75	68	7	9.8
2	19	14	5	30.3
3	65	54	11	18.5
4	85	70	15	19.4
5	48	28	20	52.6

Answer:



Sample 5 shows imprecision since it is outside the upper RAL.

58

The previous part of this module covers the statistical aspects of control charts. In addition to the statistical aspects, there are a number of practical hints on the preparation and use of statistical control charts:

1. Where possible and practical, control charts should be plotted by the analyst, and preferably posted on a wall in the operating area. Notes of corrective action taken when indicated by the charts can be made directly on the charts.

2. Many different charts can be prepared. Careful selection should be made to plot only the more critical and troublesome items.

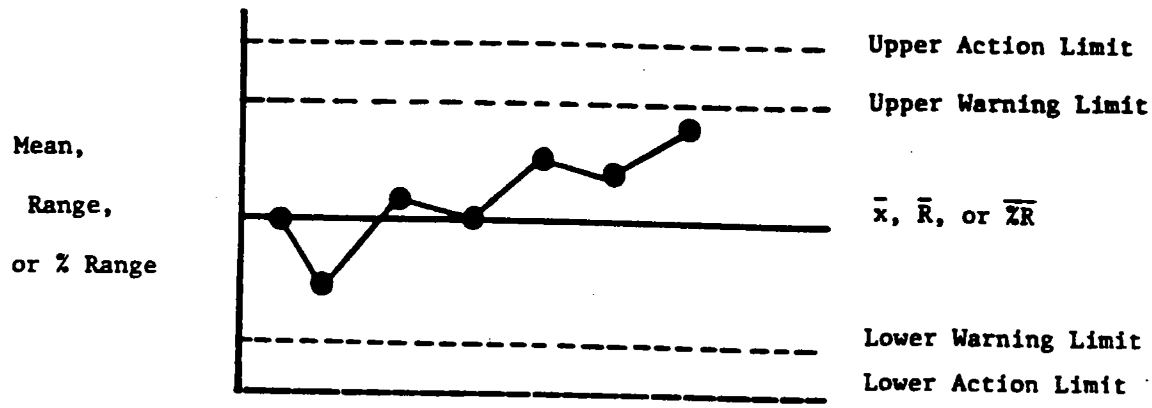
3. The following rules should be followed for determining when to investigate for out-of-control causes and take corrective action:
   a. When a single point falls outside the action limits.
   b. When two or more consecutive points fall outside the warning limits.
   c. When seven or more points fall on the same side of the central line, indicating a trend.

4. For simplicity, on charts for the less critical items, work only with the action limits.

GO TO

Frame: 59
Page: VII-62

Another valuable piece of information that you can observe from mean and range charts is evidence of a trend. Below is an example in which the system is "in control" but a trend toward high out-of-control points is indicated. We should try to find a way to halt this trend and thereby prevent the production of questionable data.



Let's summarize the steps (The steps below are shown for R; however, they apply in exactly the same way for %R):

A. List R for each sample (i.e., each subgroup)

B. Calculate $\bar{R}$. $\bar{R} = \frac{\Sigma R}{k}$ k = no. of samples

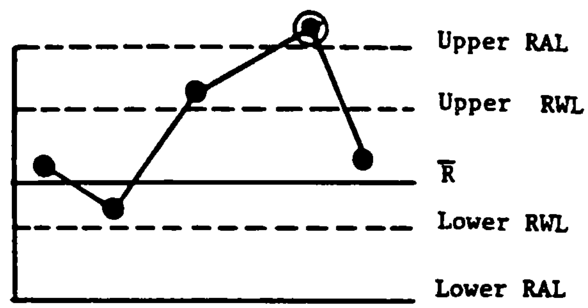
C. Calculate the limits.

Upper RWL = $D_{WU} \times \bar{R}$ Upper RAL = $D_{AU} \times R$

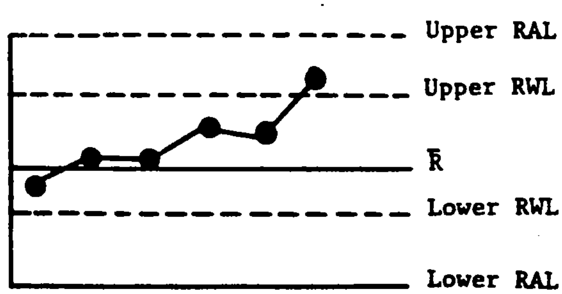
Lower RWL = $D_{WL} \times \bar{R}$ Lower RAL = $D_{AL} \times \bar{R}$

D. Prepare a chart with $\bar{R}$, Upper RWL, Upper RAL, Lower RWL, and Lower RAL.

E. Plot the new data to be tested for precision to reveal imprecision.



or trends toward imprecision:



Let's turn now to control charts for accuracy data. (Actually, the mean control chart for frame 1000 was a control chart for accuracy). Another type of chart is developed by collecting data for many samples, usually a minimum of 15 to 20, but this time we are concerned with the difference between the observed measurement and the known value. So, in this type of accuracy control chart we are comparing the observed measurement with a standard, one whose value is known. Let's use the data in frame 4000 to construct control charts for accuracy of data.

Please unfold frame 4000.

In review: Because of the importance of pollution data, the accuracy and precision of this data must be ensured. To maintain a high degree of accuracy and precision, a quality-control scheme can be applied. While similar to production quality control in many ways, the control of data quality necessitates some alterations to the standard quality-control techniques.

The major types of control charts developed for maintaining the quality of data were discussed. Further discussion of these can be found in:

(1) Quality Assurance Handbook for Air Pollution Measurement Systems, Volume I, Principles, Appendix H, Control Charts; Environmental Protection Agency. Research Triangle Park, N.C., EPA-600/9-76-005, March 1976

(2) Quality Control Practices in Processing Air Pollution Samples. Research Triangle Park, N.C.: Environmental Protection Agency, March 1973. APTD-1132.

Additional references for quality control statistics in general are:

(1) Duncan, A.J. Quality Control and Industrial Statistics, 3rd Ed. Homewood, Ill.: R.D. Irwin, Inc., 1965, Chap. 18.

(2) Grant, E.G., Leavenworth, R. S. Statistical Quality Control, 4th Ed. New York: McGraw-Hill, 1972.

(3) ASTM Manual on Quality Control of Materials, Special Technical Publication 15-D, American Society for Testing and Materials, Philadelphia, 1977.

A. First we need to list the "observed - known" differences by subtracting the known value, 30 ppm, from each of the observed measurements. (Check your answers in frame 38a). Note that the differences are much more easily analyzed visually, than the original measurements.

B. Next, compute the daily averages, $\bar{D}$, of the 2 "observed - known" differences, (Check your answers on frame 38a). Note that these values indicate the daily bias of the measurements from the known value, and if significant for a given day would indicate the bias which is assumed to apply to all routine measurements made that same day.

C. Third, compute the range for each day. (Check your answers in frame 38a). Note that the same answers are obtained whether you get the difference between the observed measurements, e.g., 29.2 - 22.7 = 6.5, or whether you get the difference of the "Observed - known", e.g., -.8 - (-7.3) = 6.5.

This completes the Introduction to Environmental Statistics. Congratulations on finishing a course covering what is commonly thought to be one of the most difficult subject matters. Of course, you should realize that this course dealt with the most basic topics of statistical analysis and that there are many topics that were left untouched. We hope that you found taking the course pleasant and informative; perhaps you will want to continue your study of statistics in the near future.

Go now to the Introduction of the <u>Guide to Statistical Problem Solving</u>.

Answer:

Day	Observed Measurements, ppm		D = Observed-Known		D_2	R_2
	a.m.	p.m.	a.m.	p.m.		
1	29.2	22.7	-.8	-7.3	-4.1	6.5
2	28.4	25.2	-1.6	-4.8	-3.2	3.2
3	29.2	26.4	-.8	-3.6	-2.2	2.8
4	32.9	30.1	2.9	0.1	1.5	2.8
5	27.9	30.2	-2.1	0.2	-1.0	2.3
6	26.4	31.8	-3.6	1.8	-0.9	5.4
7	31.8	31.5	1.8	1.5	1.7	0.3
8	39.4	29.1	9.4	-.9	4.3	10.3
9	28.6	29.2	-1.4	-.8	-1.1	0.6
10	28.0	26.2	-2.0	-3.8	-2.9	1.8
11	31.2	35.2	1.2	5.2	3.2	4.0
12	37.6	31.8	7.6	1.8	4.7	5.8
13	26.9	29.0	-3.1	-1.0	-2.1	2.1
14	30.7	28.0	0.7	-2.0	-0.7	2.7
15	31.9	26.8	1.9	-3.2	-0.7	5.1
16	28.9	36.2	-1.1	6.2	2.6	7.3
17	27.8	31.4	-2.2	1.4	-0.4	3.6

39

D. Next, calculate the average range, $\bar{R}_2$

$R_2 = \frac{\Sigma R}{N}$ N = number of subgroups (days)

$$\bar{R}_2 = \frac{6.5 + 3.2 + \ldots + 3.6}{17}$$

= # ____________

Answer: $\bar{R}_2 = \frac{66.6}{17} = 3.92$

E. All the preparatory calculations are complete, and we can now calculate the limits for our range and mean charts. Let's construct the range chart first.

Upper RWL = $D_{WU} \times \bar{R}$ = # ____________

Lower RWL = $D_{WL} \times \bar{R}$ = # ____________

Upper RAL = $D_{AU} \times \bar{R}$ = # ____________

Lower RAL = $D_{AL} \times \bar{R}$ = # ____________

Values for D_{WU}, D_{WL}, D_{AU}, and D_{AL}, are in Table K. The number of observations is 2 per subgroup.

Answer:

Upper RWL = 2.81 x 3.92 = 11.02

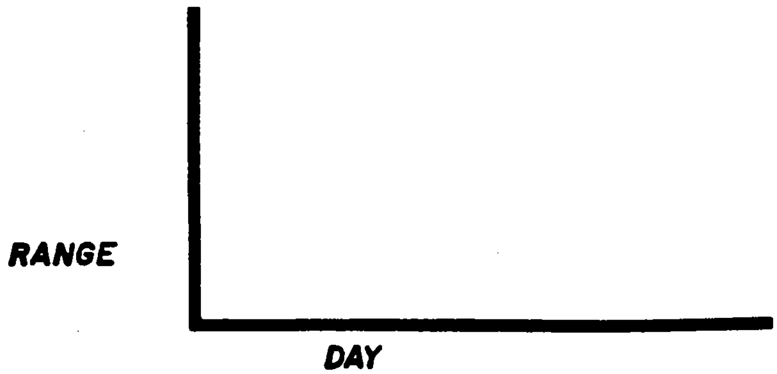
Lower RWL = .04 x 3.92 = .16

Upper RAL = 3.52 x 3.92 = 13.80

Lower RAL = 0.01 x 3.92 = .04

41

F. Now we construct the chart using these values: $\bar{R}$, Upper RWL, Lower RWL, Upper RAL, Lower RAL. Construct the R chart.



GO TO

Frame: 41a
Page: VII-19

Suppose the following measurements were made of a standard having a nominal or assessed value of 100. Suppose also that these values were considered acceptable and that it is desired to control the measurement process in the future to assure that it does not get worse.

Measurements of a Standard

100	98	101	101	98
95	103	100	101	104
103	101	102	98	98
100	102	97	100	96
98	101	97	102	101
104	100	96	101	99
97	98	101	99	102
102	96	103	100	101
101	100	100	98	102
104	95	99	101	99
96	104	102	101	97
102	100	100	101	102

Average ______________________

Standard Deviation ______________

Note that the following data are exactly the same as that for frame 1000. However, we have determined that the standard was analyzed five times each day for a period of 12 days. The same data would appear as follows:

Day	MEASUREMENTS OF A STANDARD					$\bar{x}$	R
1	100	98	101	101	98	____	____
2	95	103	100	101	104	____	____
3	103	101	102	98	98	____	____
4	100	102	97	100	96	____	____
5	98	101	97	102	101	____	____
6	104	100	96	101	99	____	____
7	97	98	101	99	102	____	____
8	102	96	103	100	101	____	____
9	101	100	100	98	102	____	____
10	104	95	99	101	99	____	____
11	96	104	102	101	97	____	____
12	102	100	100	101	102	____	____

The following data represent duplicate measurements (A and B) made by the same analyst of a series of 15 separate routine samples.

Sample	A	B	R_2
1	20	22	____
2	15	17	____
3	19	16	____
4	18	13	____
5	21	22	____
6	9	11	____
7	13	13	____
8	17	15	____
9	16	19	____
10	21	19	____
11	7	8	____
12	22	19	____
13	14	16	____
14	12	15	____
15	19	20	____

A standard solution known to correspond to 30 ppm, is checked once each morning and once each afternoon for purposes of accuracy control. Since shifts in level might be expected to occur between days, and close agreement is expected within a given day, the rational subgroup is two, the number of analyses each day. The data obtained during a period of acceptable performance were as follows:

Day	Observed Measurements, ppm	
	a.m.	p.m.
1	29.2	22.7
2	28.4	25.2
3	29.2	26.4
4	32.9	30.1
5	27.9	30.2
6	26.4	31.8
7	31.8	31.5
8	39.4	29.1
9	28.6	29.2
10	28.0	26.2
11	31.2	35.2
12	37.6	31.8
13	26.9	29.0
14	30.7	28.0
15	31.9	26.8
16	28.9	36.2
17	27.8	31.4

A supervisor or auditor prepares only <u>one</u> known sample at a different concentration level each day for analysis by a technician. Further, the laboratory works each day so there would be no logical or rational subgrouping unless it were to be done arbitrarily. The following data were obtained over one consecutive 20-day period during which time the results of the check were considered acceptable or satisfactory. It is desired to establish a quality control chart based upon this past data so as to provide a means of assuring that future checks will be as good as in the past.

Date		Observed	Known	D, Observed-Known
August	5	87	92	
	6	64	62	
	7	88	88	
	8	52	55	
	9	61	60	
	10	54	55	
	11	71	71	
	12	86	83	
	13	56	58	
	14	61	56	
	15	55	59	
	16	87	84	
	17	50	52	
	18	70	70	
	19	68	69	

The following results were obtained from colocated samplers for a particular measurement method. The results could also represent analyses by two different analysts of the same routine monitoring samples.

Sampler A (or Analyst A) ppm	Sampler B (or Analyst B) ppm	Signed %D
6.3	6.9	
82.2	87.0	
1.8	1.8	
27.6	28.0	
7.2	7.7	
34.9	32.5	
1.6	2.0	
11.6	12.8	
9.4	5.6	
2.1	2.5	
5.7	5.6	
.8	.9	
1.6	1.4	

$\bar{x}$ = $\overline{\%D}$ =

s = s =

TABLE K

FACTORS FOR CONSTRUCTING CONTROL CHARTS

Observations in a sample or subgroup	A^*_A	$A^\dagger_W$	D^*_{AU}	$D^\dagger_{WU}$	$D^\dagger_{WL}$	D^*_{AL}
	For Control Chart for Averages		For Control Chart For Ranges			
2	1.88	1.23	3.52	2.81	0.04	.01
3	1.02	0.67	2.61	2.17	0.18	.08
4	0.73	0.48	2.28	1.93	0.29	.17
5	0.58	0.38	2.10	1.81	0.37	.24
6	0.48	0.32	1.99	1.72	0.42	.30
7	0.42	0.27	1.90	1.66	0.46	.34
8	0.37	0.24	1.85	1.62	0.50	.38
9	0.34	0.22	1.80	1.58	0.52	.41
10	0.31	0.20	1.76	1.56	0.54	.43
11	0.29	0.19	1.73	1.53	0.56	.46
12	0.27	0.17	1.70	1.51	0.58	.48

* These factors are for the 99% confidence level.

† These factors are for the 95% confidence level.