Statistical Process Control: Applications and Examples for Forest Products Industries

Henry J. Quesada-Pineda

Associate Professor

Department of Sustainable Biomaterials

Virginia Tech

quesada@vt.edu

Edgar Arias

Post-doctoral Researcher

Department of Sustainable Biomaterials

Virginia Tech

earias@vt.edu

Table of Contents

Virginia Cooperative Extension
Virginia Tech · Virginia State University

www.ext.vt.edu

Introduction

Creating value requires managers to master quantitative and qualitative techniques to document and analyze information used in the decision making process. Statistical Process Control (SPC) is a tool that allows administrators both in service and manufacturing industries to monitor process capability to ensure customer requirements are met efficiently and effectively.

This material has been prepared to assist entrepreneurs in forest products industries to understand the relevance of statistical process control (SPC) in the decision making process. Supplementary examples of how to implement SPC tools for both continuous and categorical variables are presented. In addition, examples of tools used to conduct root cause analysis of potential quality problems are included too.

The use of an information technology package such as MS Excel is recommended for managers who are in the process of implementing SPC. An information technology too is critical to reduce errors and also to increase the capabilities of SPC. It is also an excellent tool to document process data for further analysis.

1   Importance of Quality Control

Quality Control falls within the philosophy of continuous improvement promoted by WE Deming. This philosophy is oriented to the ongoing struggle to improve quality and eliminate waste. Quality Control aims to establish the philosophy of continuous improvement at the institutional level so that all members of the organization work together to improve processes and the work environment. Continuous improvement specifically targets:

- Improvement of production management
- Increasing of flexibility and adaptability of the organization
- Improvement of the work environment to inspire incremental and radical innovation
- Developing of positive attitudes in all employees
- Increasing of the efficiency of the production processes

2   Sampling and related concepts

2.1   Sampling

Sampling consists in observing part of the process in order to make inferences about the performance of a process or a population. Usually, parameters are used to describe a population, such as the arithmetic mean, mode, and the median to measure the central tendency or variance to understand the dispersion of data.

2.2   Arithmetic Mean

It is computed as follows:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

Table 1 shows width measurements of a sample of sawn cypress boards. To calculate the arithmetic mean or average we use Equation 1.

Table 1. Width measures of sawn cypress boards

Sample	1	2	3	4	5	6	7	8	9
Board Width (cm)	25.5	24.8	25.3	25.0	25.1	24.8	25.5	25.1	25.2

$$\bar{X} = \frac{25.5 + 24.8 + 25.3 + 25.0 + 25.1 + 24.8 + 25.5 + 25.1 + 25.2}{9} = 25.14$$

The average width of a sample of 9 wooden boards in a sawn cypress is 25.14 cm.

2.3   Variance

The variance is used to measure the dispersion of the data. It is computed as following:

$$s_x^2 = \sum_{i=1}^{n} \frac{(X-\bar{X})^2}{n-1} \qquad (2)$$

In the calculation of variance $(n-1)$ is used in the denominator because there are n deviations $(X-\bar{X})$ but only $(n-1)$ are independent. The square root of the variance is known as standard deviation and represented with s_x.

For the above case in Table 1, the variance of the width for the cypress board is calculated using Equation 2:

$$s_x^2 = \sum_{i=1}^{n} \frac{(25,5-25,14)^2 + (24,8-25,14)^2 + (25,5-25,14)^2 + \cdots + (25,5-25,14)^2}{8} = 0.060$$

And the standard deviation is the square root of the variance:

$$s_x = 0.245$$

2.4   Range of sample

The range can be used to obtain an estimate of the standard deviation of the population of individual X points (σ_x).

It is computed as following:

$$R = X_{largest} - X_{smallest} \qquad (3)$$

Using Equation 3, we can now calculate the R range of the sample of cypress boards' widths in Table 1 as:

$$R = 25.5 - 24.8 = 0.7$$

2.5   Variation Coefficient

The Variation Coefficient or VC can be used as a measure of the quality or productivity of a process over time. The smaller the VC, the less dispersion on the data.

The VC is computed as following:

$$VC = \frac{s_x}{\bar{X}} \qquad (4)$$

2.6   Frequency plot

Frequency plots are useful to investigate the spreading of continuous data and also to see if the data cluster around a particular shape. Figure 1 shows a frequency plot of the density measurement of various samples of teak (*Tectona grandis*), a tropical hardwood species commonly used in plantations.

As can be seen from the data plotted in Figure 1, most of the values are grouped between 690 a 720 kg/m³.
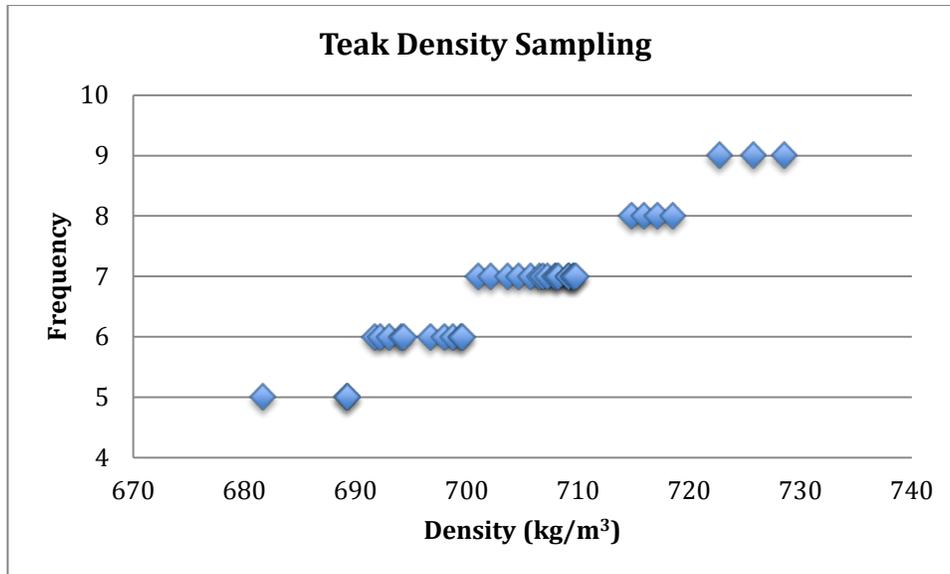
Figure 1. Sampling of Teak Density

3 Probability

A randomized experiment has a total of n possible outcomes A_i, where $i = 1, 2, ..., n$. If m experiments result in outcomes $A_1, ..., A_n$ which occur $m_1, ..., m_n$ times, then the probability of event A_i to occur can be calculated as:

$$p_i = \frac{m_i}{m}, i = 1, 2, ..., n \qquad\qquad (5)$$

Because $m_i \leq m$, then $0 \leq p_i \leq 1$ where $i = 1, 2, ..., n$. And since $m_1 + \cdots + m_n = m$, then we have that:

$$\sum_{i=1}^{n} p_i = 1 \qquad\qquad (6)$$

- If you roll a dice, the probability that the result is 1 is 1/6 = 0.16. In this case the probability of getting a 2, 3, 4, 5 or 6 is equal in all cases.
- A bag contains 10 marbles, 5 white, 3 red and 2 blue. The probability of drawing a blue marble is 2/10 = 0.2
- The probability of drawing an ace from a deck is 4/52 = 0.0769
- The probability of getting heads in a flip of a coin is a ½ = 0.5
- If two coins are thrown together, the probability of getting tails is ¾ = 0.75

3.1 Random Variable

If a real number is assigned to each possible outcome of an experiment, the set of all assigned values is known as a random variable. For example, the possible outcomes of flipping two coins are (H, H), (H, T), (T, H) and (T, T) where H: heads and T Tails. If $1,000 are paid for each head (H), then you can define the following random variable as shown in Table 2.

In this particular experiment (Table 2), a player will be paid $0 if both coins land on tails (T, T), $1,000 if at least one coin lands on heads (H,T) or (T, H) or $2,000 if both coins land on heads (H,H).

Table 2. Possible outcomes of flipping two coins

X ($)	0	1,000	2,000
Results	(T,T)	(H,T) or (T,H)	(H,H)
Probability f(x)	0,25	0,5	0,25

3.2 Cumulative Probability Function

This function describes the probability that a random variable X with a given probability distribution will be found to a have a value less than or equal to x. The cumulative distribution function of a random variable is defined as:

$$F(x) = P(X \leq x) = P(X \ taccumulates \ values \ below \ or \ equal \ x)$$

$$F(x) = \sum_{t \leq x} f(t)$$

The the probability -f(x)- and cumulative probability -F(x)- functions of the event to earn $1,000 every time you obtain heads is shown in Table 3.

Table 3. Probability distribution functions of a coin tossing experiment

X ($)	0	1,000	2,000
Results	(T,T)	(H,T) or (T,H)	(H,H)
Probability f(x)	0,25	0,5	0,25
Cumulative Probability F(x)	0,25	0,75	1

The single probabilities for each event are calculated as follow:

- Event when the two coins land on tails (T, T). The probability of landing on heads (H) is 0.5. Therefore the probability of landing on tails (T) is 1-0.5=0.5. So we have,

$$(1-0.5)*(1-0.5)=0.25$$

- Event when at least one of the two coins lands on heads (H,T) or (T, H). So we have,

$$(0.5)(1-0.5) + (1-0.5)(0.5) = 0.5$$

- Event when the two coins land on heads (H,H). So we have,

$$(0.5)*(0.5)=0.25$$

The cumulative probability function F(x) is calculated as follow:

- No heads at all

$$(T,T) = 0.25$$

- At least one coin landing on heads plus the event where (T,T)

$$(T,T) + (H,T) \text{ o } (T,H) = 0.25 + 0.25 *2 = 0.75$$

- Two coins landing on heads (H,H) plus the previous to events

$$(T,T) + (H,T) \text{ o } (T, H) + (H,H) = 0.25 + 0.5 + 0.25 = 1$$

3.3 Expected Value

The expected value provides a measure of the center of the distribution variable. By taking the expected value of a distribution other features such as spread, skewness, kurtosis and correlation can also be measured. The expected value of a function $g(x)$ of a discrete random variable *X* is defined as:

$$E\big(g(x)\big) = \sum g(k)f(x), \; X: discrete \; variable \qquad (7)$$

The expected value of the game to earn a thousand dollars if heads are obtained by flipping two coins can be calculated using Equation 7 as:

$$E\big(g(x)\big) = \$0 * 0{,}25 + \$1{,}000 * 0{,}5 + \$2{,}000 * 0{,}25 = \$1{,}000$$

The expected value of rolling a die is:

$$E\big(g(x)\big) = (1 + 2 + 3 + 4 + 5 + 6)\frac{1}{6} = 3{,}5$$

Notice that to calculate the expected value of any random variable you need to know the probabilities of each outcome.

3.4 Binomial Distribution

This distribution is very important since in *p* control charts, the defective fraction is based on this distribution to calculate the average proportion of defects and control limits.

Suppose you have 10 marbles in a bag, and that 7 are clear glass (C) and 3 white (W). If we draw a marble from the bag, the probability of drawing a white marble is:

$$P(W) = \frac{3}{10} = 0.33$$

Now, instead of drawing one marble, we will draw three marbles, one after the other, without replacement. What is the probability of getting one white marble? The order of the marbles could be depicted like:

$$WCC, CWC, or \; CCW$$

The probability of drawing a white marble (W) is *0.33* and the probability of drawing a non-white marble is (*1-0.33*). In any of the three possible scenarios, the probability of each combination is:

$$(0.33)(1 - 0.33)(1 - 0.33) = 0.15$$

You should consider that when drawing a white marble in a sample of three, is necessary to count in how many ways a white marble can be drawn. As it is a small sample, we can see that there are 3 different ways. But in case of larger samples, the number of combinations is calculated as:

$$\binom{n}{d} = \frac{n!}{d!(n-d)!} \qquad\qquad (8)$$

The notation $n!$ is pronounced "n factorial". If $n * 3$ then $3! = 1 * 2 * 3 = 6$. The number of successes is represented with the variable *d*. In the case of drawing three marbles with one white (the success), the number of combinations is calculated using Equation 8 as:

$$\binom{3}{1} = \frac{3!}{1!\,(3 - 1)!} = \frac{6}{2} = 3$$

Keep in mind that the probability of a combination (WCC, CWC or CCW) is:

$$0.33(1 - 0.33)(1 - 0.33) = 0.15$$

The result *0.15* is multiplied by the number of possible combinations, which gives:

$$P(W = 1) = \binom{3}{1} 0.33^1 (1 - 0.33)^{3-1}$$

$$P(W = 1) = 3(0.15) = 0.45$$

The general formula $P(d) = \binom{n}{d} p^d (1 - p)^{n-d}$, is known as the general formula of the binomial distribution. If you draw three marbles from a bag where d is the probability of drawing a white marble (W), the probability distribution is represented in Table 4:

Table 4. Probability distribution of drawing three marbles from a bag with 10.

Result	d	Combinations	Probability
CCC	0	1	0.343
WCC	1	3	0.441
WWC	2	3	0.189
WWW	3	1	0.027
Total			1.000

Note that we want to know all possible outcomes of drawing three marbles of the bag, where d is the presence of a white marble (W) and C is a clear one.

If we want to calculate the probability of getting at least 2 white marbles, then add the results BBT + BBB in Table 4, which is the same as:

$$P(W \le 2) = WWC + WWW = 0.189 + 0.027 = 0.21$$

The same result can be directly obtained by using the formula for the binomial distribution, we have:

$$P(d \leq 2) = \binom{3}{2} 0.33^2 (1 - 0.33)^{3-2} + \binom{3}{3} 0.33^3 (1 - 0.33)^{3-3} = 0.21$$

3.5 Poisson Distribution

This distribution is used to calculate probabilities of the number of occurrences of certain events in a given time or a given area. Examples of occurrences that can be modeled with a Poisson distribution are:

- Number of cars entering the car wash per hour
- Number of defects per square meter of plywood
- Number of saws that break in a sawmill per week
- Number of defects found in a wood board of a given area

In quality control, this distribution is important because is used for *c charts*, which are utilized to study number of defects over a period of time.

The Poisson distribution is defined as:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, where\ x = 0,1,2,\ldots (9)$$

It can be seen from Equation 9 that the Poisson distribution has a parameter λ (Greek letter lambda), which represents the occurrence of events per unit time or area.

Suppose that in a sawmill, the log loader can service a maximum of 15 trucks per day and that the average number of trucks entering the logyard is $\lambda = 11$ per day. What is the probability that the loader process no more than 2 trucks in a given day?

Using the formula for the Poisson distribution we have:

$$P(X \leq 2) = \frac{11^0 e^{-11}}{0!} + \frac{11^1 e^{-11}}{1!} + \frac{11^2 e^{-11}}{2!} = 0.001$$

What is the probability of unloading exactly 15 trucks?

$$P(X = 15) = \frac{11^{15} e^{-11}}{15!} = 0.05$$

What is the probability that the charger exceeds its capacity? That is, arriving more than 15 trucks.

$$P(X \geq 15) = 1 - \sum_{i=0}^{15} \frac{11^i e^{-11}}{i!} = 0.09$$

3.6 Normal Distribution

The bell-shaped distribution makes an excellent approximation of the relative frequencies of many natural events and other phenomena where man intervenes (weights, dimensions, forces, etc.). Under this distribution, the probability of a given event can be determined using as inputs the population average μ and variance σ^2. Figure 2 shows a

normal probability distribution with parameters $\mu = \bar{X} = 15.5$ and $\sigma^2 = 0.0144$. The area under the curve represents the probability.
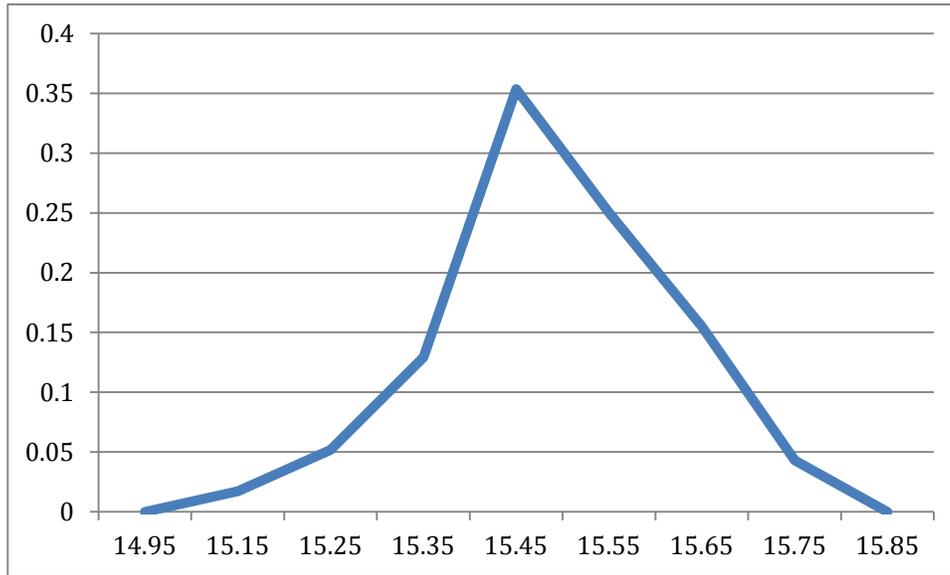
Figure 2. Example of a normal distribution

3.7 Standard Normal Distribution

A convenient way to get the values or probabilities of a normal distribution is using a standard normal distribution. In this case, the value of the parameters is $\mu_x = 0$ and $\sigma_x = 1$. This makes the job of finding the area under the curve for a given case easier. The formula of the standard normal distribution is:

$$Z = \frac{X - \mu_x}{\sigma_x} \qquad\qquad (10)$$

Where Z indicates the distance or number standard deviations to the right or left of μ_x. Considering the probability distribution of Figure 2 with parameters $\mu_x = 15.5$ and $\sigma_x = 0.12$, we want to know the probability of X being less than 15.61. We have:

$$Z = \frac{15{,}61 - 15{,}5}{0{,}12}$$

$$Z = 0{,}917$$

$$P(X \le 15{,}61) = P(Z \le 0{,}92)$$

The standard normal table shows the values of the cumulative distribution function of the normal distribution. A copy of this table can be found in the Appendix section. For this particular example, the standard normal table indicates that the probability of X being less that 15.61 is 0.8212 or 82.12%. If you are interesting in knowing the probability that X is between 15.4 and 15.61, then proceed as follows:

$$Z_1 = \frac{15{,}4 - 15{,}5}{0{,}12} = -0{,}83$$

$$Z_2 = \frac{15{,}61 - 15{,}5}{0{,}12} = 0{,}92$$

$$P(15{,}4 \leq X \leq 15{,}61) = P(-0{,}83 \leq Z \leq 0{,}92)$$

In this case, the value of the standard normal table for $Z_1 = 0.917$ is 0.2033, and 0.8212 for $Z_2 = 0.917$. Then we have:

$$P(15{,}4 \leq X \leq 15{,}61) = P(Z \leq 0{,}917) - P(Z \leq -0{,}833)$$

$$P(15{,}4 \leq X \leq 15{,}61) = 0{,}8212 - 0{,}2033 = 0{,}6179$$

This indicates that the probability of X is between 15.61 and 15.4 is a 61.79%.

3.8   Central Limit Theorem

The distribution of the means of a sample is always normal even if the sample comes from a non-normal population. As the size or number of samples n increases, the normality of distribution increases.
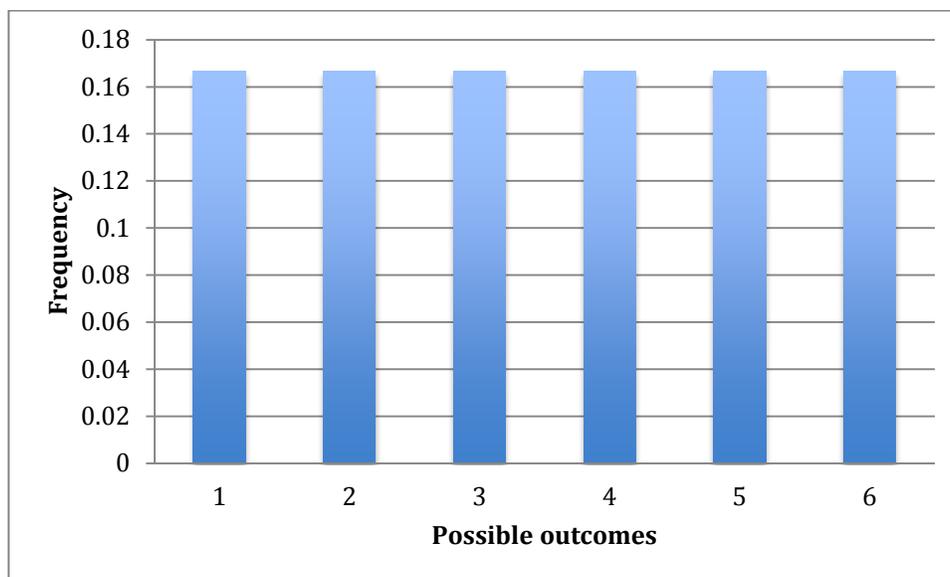
Figure 3. Probability distribution of throwing a dice

Figure 3 shows the probability distribution of rolling a dice. Note that the probability of the six possible outcomes $(1, 2, 3, 4, 5, 6)$ is 1/6 = 0.1667. In this case, this distribution is considered uniform, meaning that each result has exactly the same chance to occur. Figure 4 shows the probability distribution of the averages of 91 samples of size 10 for the game of throwing a dice. When calculating the frequency distribution for the averages of samples of size 10, the distribution of mean gets closer to a normal distribution. The larger the amount of sampled means, the better approximation to a normal distribution.
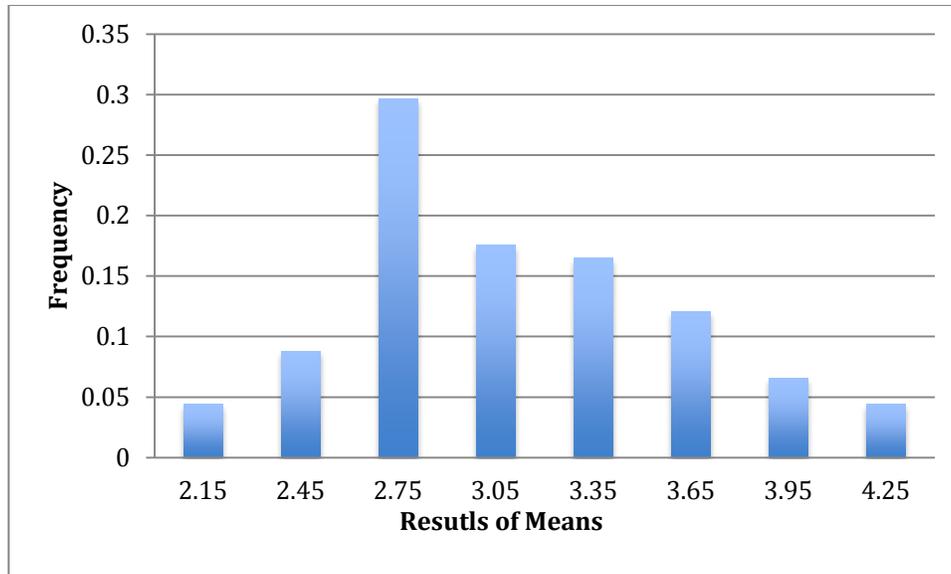
Figure 4. Probability distribution of means

4   Process Capability and Control

According to Taguchi, variability in products or processes comes from three possible sources (Devor, Chang, and Sutherland 1992):

- External noise: they are external forces or factors operating in the environment where the product is functioning and whose variation is transmitted through the design to the product performance. Included here are: temperature, humidity, pollutants, fluctuations in voltage, vibration of machinery, human performance variations.
- Internal noise: internal changes in product characteristics such as displacement from the nominal value over time caused by deterioration.
- Noise due to variations: variation in the product parameters from one unit to another as a result of the manufacturing process. It is also known as manufacturing imperfections.

4.1   Process Control

A process is said to be under control if through past experience or historical data a prediction can be made on how the process or phenomenon will behave or vary in the future. This prediction can be done by the probability that the process will be within the process control limits.

4.2   Process Capability

It indicates that the quality characteristics of the final product or process are within the specification limits, implying that the product is meeting customer requirements.

5   Control Charts

5.1   Control Chart $\bar{X}$ (x bar)

In a process, there are two features of interest that are measured: the process average and the process variation. Of particular interest it is the measurement in their changes of

magnitude. Table 5 shows data collected when measuring the width of certain batch of bald cypress boards (in inches). The measurements were made using a caliper with an accuracy of 1/100 inches. As shown in Table 5, a total of 30 samples were taken every half hour and each sample has a total of 5 measurements. The average $(\bar{X}_i)$ and (R_i) range for each sample is also included in Table 5.

Table 5. Board Widths Measurements

Sample (i)	1	2	3	4	5	$\bar{X}_i$	R_i
1	24.52	21.96	23.96	22.74	26.34	23.90	4.38
2	26.09	24.86	26.16	23.49	24.80	25.08	2.66
3	23.15	24.92	24.32	25.67	25.32	24.68	2.51
4	25.73	26.89	24.13	27.87	23.66	25.66	4.21
5	24.71	26.76	26.09	24.45	24.12	25.23	2.64
6	25.21	25.65	29.98	23.02	26.57	26.08	6.96
7	26.59	26.99	22.91	26.61	25.71	25.76	4.08
8	25.01	25.46	26.30	25.55	28.50	26.16	3.49
9	23.51	26.25	24.94	23.28	23.83	24.36	2.97
10	28.77	26.81	34.00	25.45	22.12	26.07	6.65
11	24.12	26.05	25.72	22.37	22.91	24.24	3.68
12	23.99	24.74	24.78	26.66	24.41	24.92	2.67
13	21.22	24.17	25.67	26.16	24.86	24.41	4.94
14	23.76	25.37	26.74	26.60	25.08	25.51	2.98
15	26.10	23.85	23.26	24.07	22.87	24.03	3.23
16	25.54	26.46	24.11	22.47	24.84	24.69	3.99
17	28.84	23.50	22.50	25.63	28.20	25.73	6.35
18	25.68	23.96	24.08	23.55	25.20	24.50	2.13
19	25.96	24.42	24.65	26.68	28.28	26.00	3.86
20	25.81	24.78	26.84	27.33	26.70	26.29	2.55
21	26.91	22.78	27.55	27.55	25.10	25.98	4.77
22	25.79	26.17	27.89	27.21	24.72	26.36	3.18
23	25.71	25.37	26.49	25.26	24.48	25.46	2.01
24	23.15	23.98	26.49	20.68	23.94	23.65	5.81
25	26.35	26.54	26.43	23.39	23.60	25.26	3.15
26	26.33	26.47	25.13	24.80	27.71	26.09	2.91
27	26.08	25.07	26.66	22.35	24.90	25.01	4.30
28	25.85	24.41	24.42	25.69	26.59	25.39	2.18
29	25.40	23.73	23.94	25.60	23.09	24.35	2.51
30	27.53	26.88	24.42	26.24	24.89	25.99	3.11

The average for each sample i is calculated using the formula:

$$\bar{X}_i = \frac{\sum_{j=1}^{n} X_{ij}}{n}$$

And to calculate the ranges R_i, the following formula is used:

$$R_i = X_{\max i} - X_{\min i}$$

The first step to plot the graph $\bar{X}$ is to calculate the average of the samples, along with the average range of the same. We proceed as follows:

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{k} \bar{X}_i}{k} = 25{,}25$$

To calculate the average range of samples, the formula used is:

$$\bar{R} = \frac{\sum_{i=1}^{k} R_i}{k} = 3{,}70$$

The statistic $\bar{\bar{X}}$ is used to estimate the population average, just as $\bar{R}$ is computed to estimate the population variance. In this case, it should be noted that we have 30 samples, which is considered sufficient to have a good approximation of the average and variance of the population. If less than 30 samples were taken, the risk of sampling errors would be too large to make a good estimation of population parameters.

The next step is to calculate the process' test limits for the $\bar{X}$ chart. The limits to calculate are of the form:

$$Average\ of\ samples\ \pm\ 3\ standard\ deviations$$

Is important to mention that the centerline of $\bar{X}$ graph is the average of the samples $(\bar{\bar{X}})$. To estimate the population standard deviation, the following formula is used:

$$\widehat{\sigma_{\bar{x}}} = \frac{\bar{R}}{d_2 \sqrt{n}}$$

The variable d_2 is a function of the sample size, assuming that the measurements follow a normal distribution. It is obtained from a special table in the Appendix section. Therefore, the upper control limit (*UCL*) and the lower control limit (*LCL*) can be estimated as:

$$UCL = \bar{\bar{X}} + \frac{3\bar{R}}{d_2 \sqrt{n}}$$

$$LCL = \bar{\bar{X}} - \frac{3\bar{R}}{d_2 \sqrt{n}}$$

The term $\frac{3\bar{R}}{d_2 \sqrt{n}}$ depends solely on the sample size (n), allowing the rest of terms to be grouped in a single constant A_2. The values A_2 can also be found in the Appendix section. To calculate the control limits for the X chart, to the following simplified formulas can be used:

$$UCL = \bar{\bar{X}} + A_2 \bar{R}$$

$$LCL = \bar{\bar{X}} - A_2 \bar{R}$$

At this point we know the necessary formulas and reasoning for calculating control limits of a graphic $\bar{X}$. Hence, the control limits for the board widths of the Cypress boards measurements can be computed as:

$$UCL = \bar{\bar{X}} + A_2\bar{R} = 25.27 + 0.577(3.87)$$

$$UCL = 27.46$$

$$LCL = \bar{\bar{X}} - A_2\bar{R} = 25.27 - 0.577(3.87)$$

$$LCL = 22.99$$

The initial $\bar{X}$ chart is shown in Figure 5.
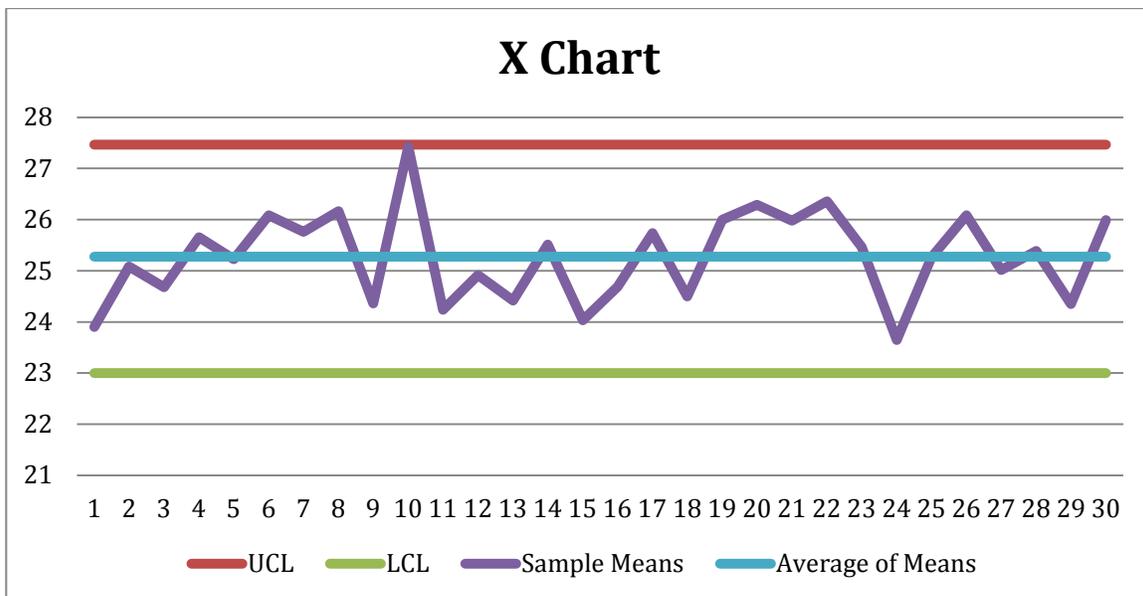
Figure 5. X bar chart for the widths of cypress boards

5.2   Control Chart for Ranges (*R Chart*)

For *R charts*, it should be noted that the distribution of these is not easy to determine, nor necessarily symmetric around the value of $\bar{R}$. To simplify the computation of the control limits, the following formulation has been stablished: $\bar{R} \pm 3\sigma_R$. For those cases where the value obtained for the lower control limit is negative, the value of zero should be adopted as the actual limit. The formula to estimate σ_R is:

$$\sigma_R = d_3\sigma_X$$

Where d_3 is a function of the sample size (*n*). Values for d_3 can be found in the Appendix section. Given the relationship between *R* and σ_X, we have that:

$$\hat{\sigma}_R = \frac{d_3\bar{R}}{d_2}$$

Therefore, the control limits for the *R chart* can be computed using the following formula:

$$\bar{R} \pm 3 \left(\frac{d_3}{d_2}\right) \bar{R}$$
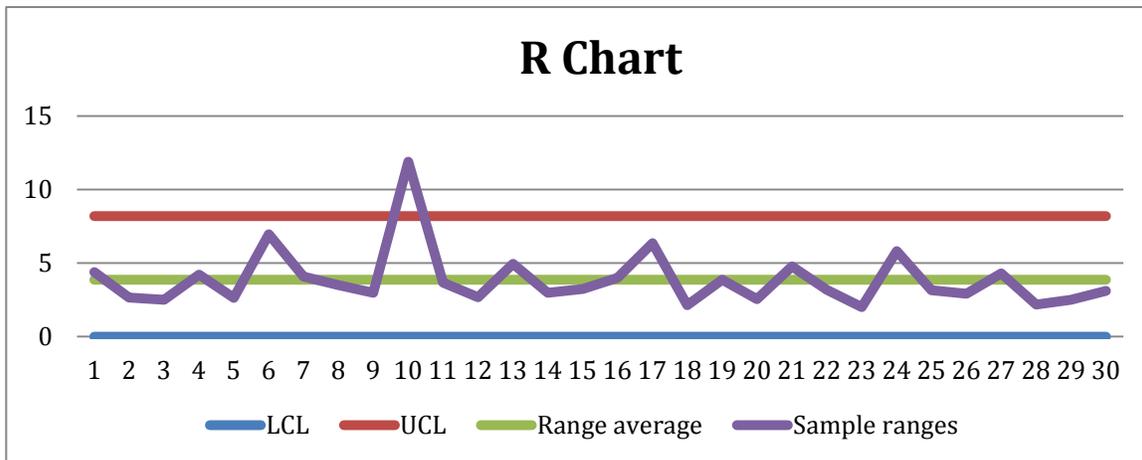
Figure 6. R chart for the widths of cypress boards

Here the proportion d_3/d_2 can be substituted with the values D_3 or D_4 as following:

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

$$LCL = D_3 \bar{R}$$

The values of D_3 and D_4 depend on the sample size n and these values are tabulated in the Appendix section. Therefore, the calculation of the limits would be:

$$UCL = D_4 \bar{R} = 2.115(3.87) = 8.19$$

$$LCL = D_3 \bar{R} = 0(3.87) = 0$$

Figure 6 shows the resulting R Chart.

5.2.1   Interpretation of $\bar{X}$ and R charts

Looking at the $\bar{X}$ Chart and the R Chart it can be noted that there is a sample outside the control limits of the R Chart and the same data point is right on the edge of the $\bar{X}$ Chart. These anomalies could be an indication that the process might be out of control and we need to understand what may have occurred. After reviewing all samples, it is noticed that the point out of control occurred in sample 10. Fortunately for us, the operator keeps a log or shift report. At this time, it was reported that a change of saws was performed, so it is possible that this adjustment was not performed correctly, causing this point out of control. Checking further, the operator who made the change of saws does not have much experience in this setup procedure. To prevent this problem from re-appearing in the future, we developed a standard procedure for changing saws, to avoid the same error and make sure the operator has the proper training.

The nature of this type of chart is to capture a warning of a particular situation. As we were able to identify the source of the problem and proceeded to correct it, we now need to recalculate the $\bar{X}$ and $\bar{R}$ Charts. We do this by removing the sample that caused the problem (sample 10). The new values of $\bar{\bar{X}}$ and $\bar{R}$ are:

$$\bar{\bar{X}} = 25.20$$

$$\bar{R} = 3.59$$

The new control limits for the $\bar{\bar{X}}$ chart are:

$$UCL = \bar{\bar{X}} + A_2\bar{R} = 25.20 + 0.577(3.59) = 27.30$$

$$LCL = \bar{\bar{X}} - A_2\bar{R} = 25.20 - 0.577(3.59) = 23.16$$
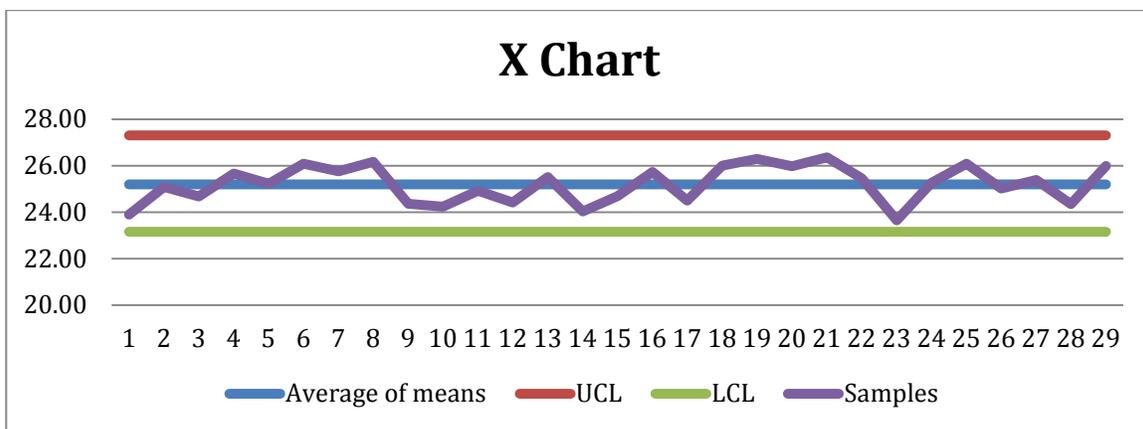
Figure 7. Revised X Chart with sample 10 removed

Figure 7 shows the recalculated X Chart. In this particular case, no abnormalities are evident indicating that the process is under control.
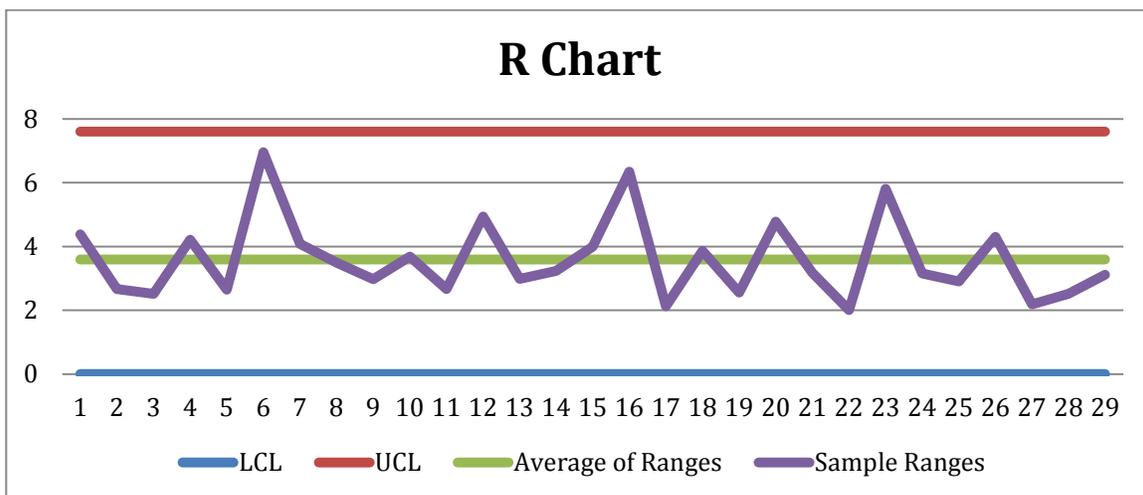
Figure 8. Revised R chart with sample 10 removed

The control limits for the revised R chart are:

$$UCL = D_4 \bar{R} = 2.115(3.59) = 7.60$$

$$LCL = D_3 \bar{R} = 0(3.59) = 0$$

The new R chart is shown in Figure 8. Similarly to X Chart, no abnormalities are observed.

5.3 Control Charts for Attributes (Attribute Charts)

In quality control, we are interested in determining the presence of defects. These non-conformities can be determined either visually or through special instruments. In this context, quality is evaluated via attributes. These are some key definitions to be considered:

- Defect: failure or nonconformity that causes an item not to meet a specification or requirement. Every instance of such nonconformity found in an item is considered a separate defect.
- Defective: an article having one or more defects is considered a defective item
- Number of defective items: in a sample of size n, d is the number of defective items
- Number of defects: in a sample of n items, c is the number of defects in the sample. An item can have many different types of defects. The sum of all defects in the sample is c.
- Defective Fraction: the ratio of the number of defective items in the sample d, to the number of items in the sample n. The defective fraction of the sample is p:

$$p = \frac{d}{n} \qquad (11)$$

5.4 Control chart for the proportion of nonconforming units in a sample: *p-chart*

Table 6 shows data from a sample of defects in lumber boards. The defects were: stain, bark, knot, twist, or split. If any board in each sample had at least one occurrence of any of these defects, it was counted as defective board. Each sample consisted of 10 tables.

To find the control limits of this type of chart, you must estimate the defective fraction within of the population first. For this, the following procedure is used:

$$\mu_p = E\left(\frac{d}{n}\right) = \frac{1}{n}E(d) = \frac{1}{n}np' = p'$$

Since there is no way of knowing p', it is then estimated from the obtained sample. Therefore $p' = \bar{p}$, where:

$$\bar{p} = \frac{Number\ of\ defective\ boards}{Number\ of\ inspected\ boards} = \frac{32}{200} = 0.16$$

To find the variance, proceed as follows:

$$\sigma_p^2 = Var\left(\frac{d}{n}\right) = \frac{1}{n^2}Var(d)$$

$$\sigma_p^2 = \frac{1}{n^2}np'(1-p')$$

Virginia Cooperative Extension

Virginia Tech • Virginia State University

www.ext.vt.edu

$$\sigma_p^2 = \frac{p'(1 - p')}{n}$$

$$\sigma_p = \sqrt{\frac{p'(1 - p')}{n}}$$

Table 6. Data of defects in sample of wood boards

Sample (n=10)	Defects	Defective Fraction (p)
1	1	0.1
2	2	0.2
3	0	0
4	0	0
5	1	0.1
6	3	0.3
7	1	0.1
8	1	0.1
9	0	0
10	2	0.2
11	1	0.1
12	3	0.3
13	1	0.1
14	1	0.1
15	4	0.4
16	1	0.1
17	2	0.2
18	6	0.6
19	1	0.1
20	1	0.1
Total	32	

The center line of the graph is $\mu_p = p' = \bar{p}$, and because the control limits are $\bar{p} \pm 3\sigma_p$, we have that:

$$\sigma_p = \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} = \sqrt{\frac{0.16(1 - 0.16)}{10}} = 0.13$$

$$UCL_p = \bar{p} + 3\sigma_p$$

$$0.16 + 3(0.13) = 0.55$$

$$LCL_p = \bar{p} - 3\sigma_p$$

$$0.16 - 3(0.13) = -0.23 = 0$$
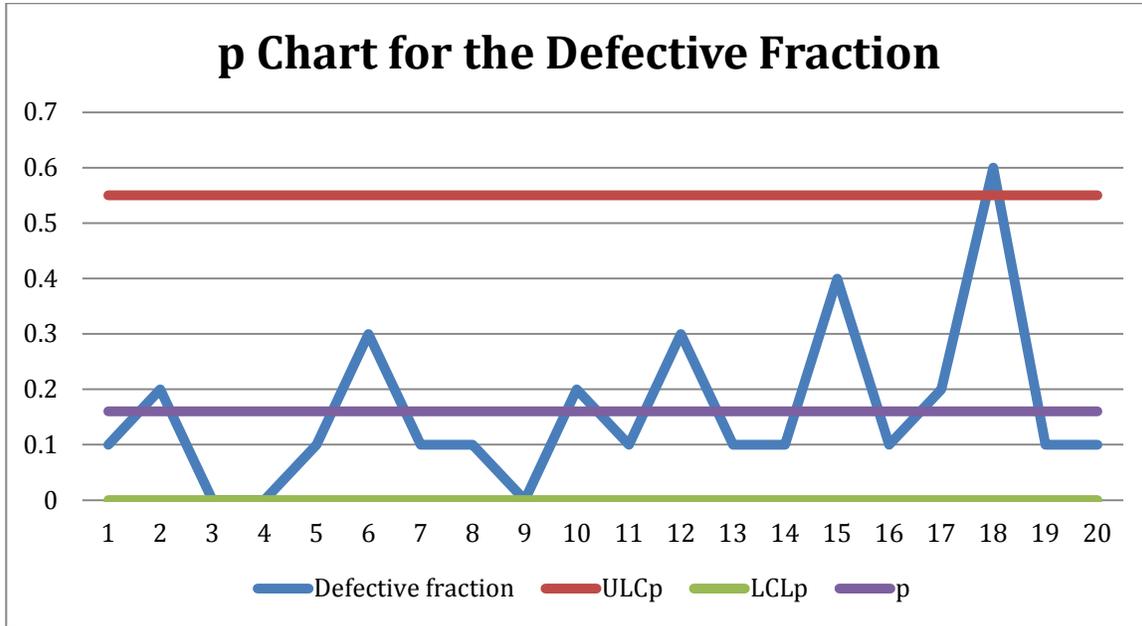
Figure 9. *P chart* for the Fraction Defective Data in Table 6

The *p Chart* for the defective fraction of sampled the boards in Table 6 is shown in Figure 9. The Interpretation of a p Chart, is very similar to that of the *X bar Chart*.

- No points should get out of the control limits
- The points should be randomly distributed within the control limits
- The points should not show evidence of trends or cycles
- The points should be random in time (no trends above or below the center line)

In this particular case, there is a point out of control (sample 18). Furthermore, the fraction defective is quite high (16%). This is an indication that the production system has problems of variability and that there might be systemic symptoms.

5.5   Control chart for Number of Defects:

The *p Chart* means that the product or item is defective because at least one defect is present at the time of the inspection. Sometimes in the lumber industry is of interest to examine the occurrence of each type of defect in an article (number of stains, splits, knots, etc. in a board for instance), and the number of defects per sample as well.

In this case the number of defects *c*, follows a Poisson distribution. An important property of the Poisson distribution is that the mean and variance are equal. To estimate c', the average per sample $\bar{c}$ defects is used:

$$c' = \pm 3, \quad where$$

$$c' = \bar{c} = \sum_{i=1}^{k} \frac{c_i}{k}$$

Here, c_i is the number of defects in the sample and k is the sample size.

Table 7. Defect data for a Sample of 20 wooden boards

Sample #	Stain	Wane	Knot	Twist	Split	Number of defects (c)
1	0	0	0	0	0	0
2	1	0	0	0	0	1
3	0	0	4	0	0	4
4	2	0	1	0	0	3
5	1	0	0	0	2	3
6	1	0	1	0	0	2
7	1	0	3	1	0	5
8	0	2	1	0	0	3
9	3	0	1	2	3	9
10	1	0	1	0	0	2
11	0	0	1	0	0	1
12	1	0	2	0	4	7
13	0	2	1	1	0	4
14	1	0	1	0	0	2
15	1	0	0	0	2	3
16	0	0	2	0	0	2
17	1	0	1	1	0	3
18	1	0	2	0	1	4
19	0	0	1	0	0	1
20	1	0	1	0	0	2
Total	16	4	24	5	12	61

The control limits for the control of the number of defects are calculated as:

$$UCL_c = c' + 3\sqrt{c'}$$

$$LCL_c = c' - 3\sqrt{c'}$$

Table 7 shows data on the number of defects found in 20 samples, each with five boards. Then we have:

$$c' = \bar{c} = \sum_{i=1}^{k} \frac{c_i}{k} = \frac{61}{20} = 3.05$$

$$UCL_c = 3.05 + 3\sqrt{3.05} = 4.8$$

$$LCL_c = 3.05 - 3\sqrt{3.05} = 1.3$$
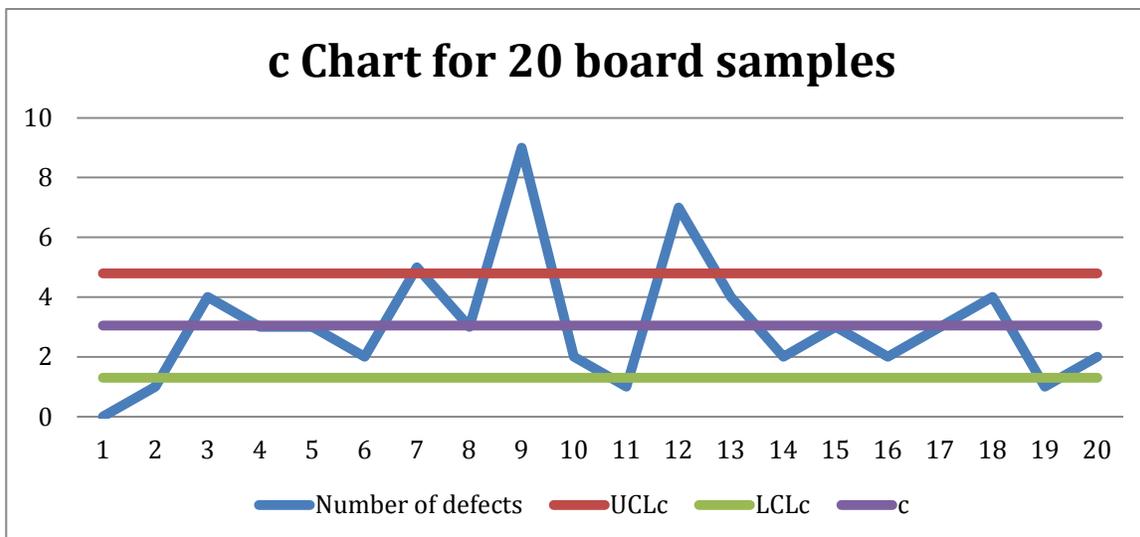
Figure 10. Control Chart *c* for the Number Defects

The c Chart for the number of defects on the wood boards is shown in Figure 10. Note that the graph shows a situation out of control (there are several points that exceed the limits). Table 7 shows that knots represent the majority of defects, followed by stains. This analysis may indicate a better selection of the raw material should be made (with less knots) and also that the conveyor chains might be dirty.

6   Diagnosing problems

There are several tools that are very useful for troubleshooting and continuous process improvement when working with Control Charts in SPC. Below is a description of the most important tools used in detecting aspects and events relevant to the process under study.

6.1   Scatter Plots

The scatter diagram is used to discover correlations that may exist between a measurement of quality or performance.

Figure 11 depicts the scatter plot of the averages of samples from wood boards measurements by operator. Note that for the case of operator 2 (Marcos), the data has a higher dispersion and in the case of operator 1 (Javier), the data is less disperse or variable.

Virginia Cooperative Extension

Virginia Tech • Virginia State University

www.ext.vt.edu
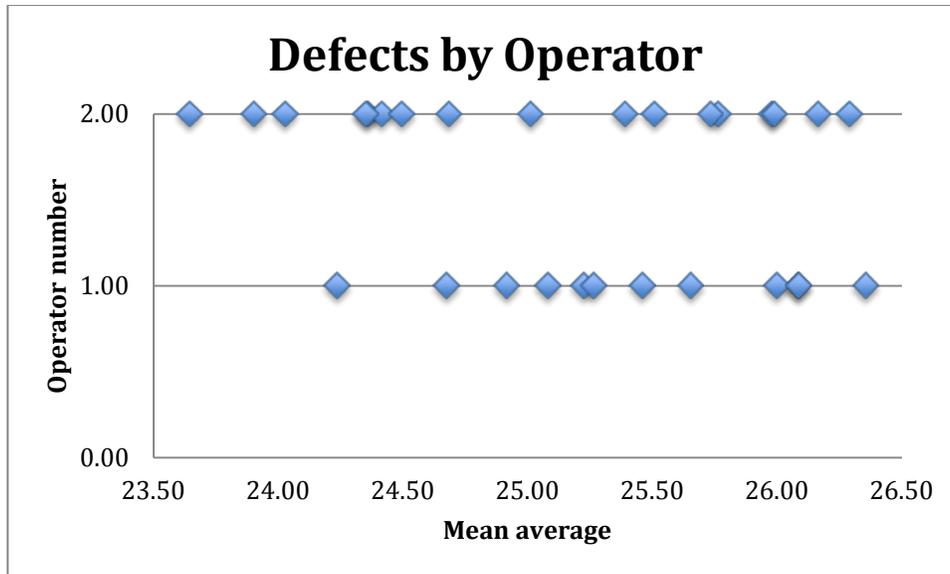
Figure 11. Scatterplot of Sample Averages for Operator

6.2   Pareto chart

The Pareto chart follows the principle that the sum total of a certain type of occurrence can come from different sources and it is very likely that most of these occurrences come from a relatively small set of sources.

Table 8 shows the results of sampling to quantify the presence of certain defects in Cypress boards. Each sampling was conducted by taking five consecutive boards at the end of a kiln drying cycle. The inspector reviewed both sides of each board, and if one of the defects (stain, wane, knot, twist, or split) is present in any of the boards being inspected, then a value of 1 is annotated, or 0 otherwise. We are not interested in the number of defects for each type table or the total, but the presence of at least one of these defects in any of the five boards that form a sampling unit.

After taking 20 samples, the inspector should quantify the total defects by type and grand total. The relative frequency of each defect is then obtained as shown at the end of Table 8. These frequencies are used to build the Pareto chart shown in Figure 12.

In the Pareto chart you can see that two defects (knot and stain) make up 82% of the defects found in all samples. In this particular case, the process manager must focus on reducing or eliminating these two defects that represent over 80% of defects found in the samples.

Table 8. Defects in Sampled Wood Boards

Sample #	Stain	Wane	Knot	Twist	Split
1	0	0	0	0	0
2	1	0	0	0	0
3	0	0	1	0	0
4	1	0	1	0	0
5	1	0	0	0	1
6	1	0	1	0	0
7	1	0	1	1	0
8	0	0	1	0	0
9	1	0	1	1	0
10	1	0	1	0	0
11	0	0	1	0	0
12	1	0	0	0	0
13	0	1	1	1	0
14	1	0	1	0	0
15	1	0	0	0	0
16	0	0	1	0	0
17	1	0	1	1	0
18	1	0	1	0	0
19	0	0	1	0	0
20	1	0	1	0	0
Total	13	1	15	4	1
Frequency	38.2%	2.9%	44.1%	11.8%	2.9%
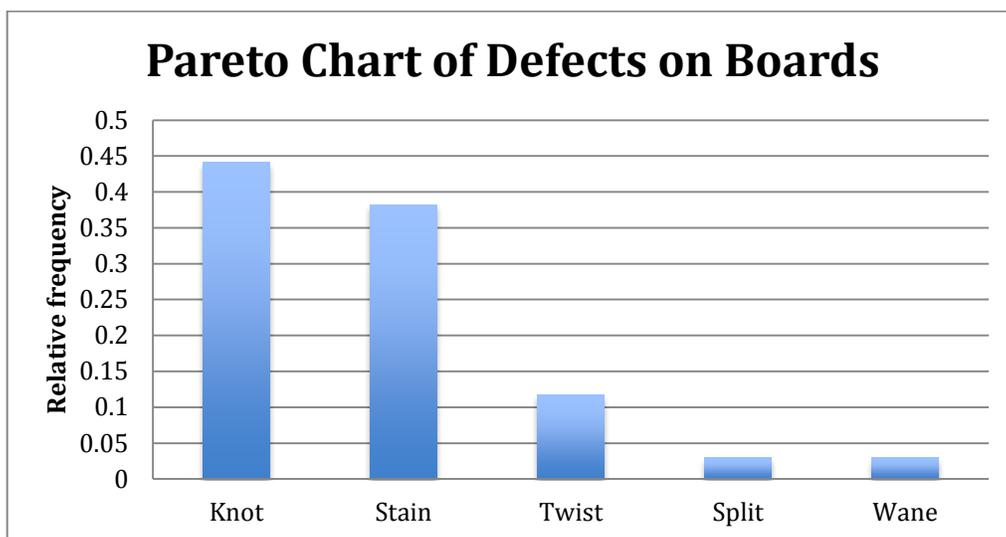
Figure 12. Pareto Diagram of Defects in Cypress Boards

6.3 Cause and Effect (Fishbone, Ishikawa) Diagram

The cause-effect diagram is a graphic tool to find the root cause or causes of problems. This diagram allows a graphical representation of the possible systematic path that ultimately will help to identify the root cause of the quality problem. The idea is to associate symptoms with possible factors or causes that produce these symptoms. To do this, you start with the most general categories and hierarchically continue to the most detailed levels.

Figure 13 shows a diagram to find the root cause of the stains on wooden boards. In this case, 4 possible general categories were established: materials, machinery, human resources and working methods. Then we arranged a meeting with a group of people, including operators and supervisors to jointly identify possible causes of the defect. Those possible causes can be seen in the diagram. Kaizen meetings are an effective way to identify problems. A cause-effect diagram is excellent visual aid to help Kaizen team members in the process of identifying problems and finding potential solutions.
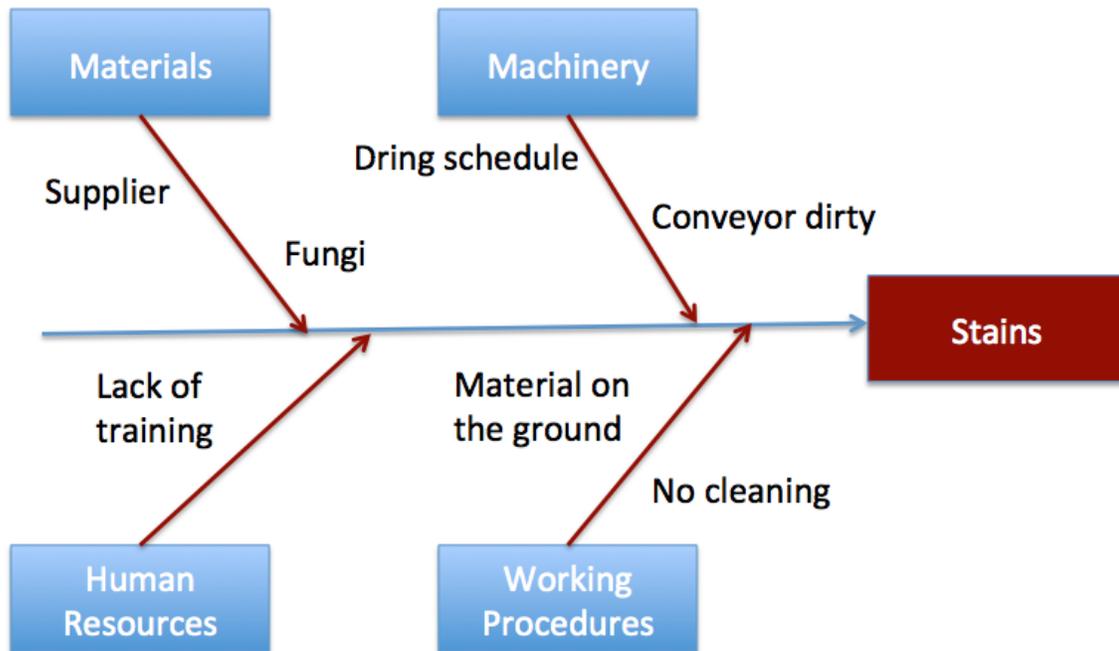
Figure 13. Cause-effect Diagram for Stains in Wooden Boards

7 Appendix

7.1 Area under the normal curve, cumulative.

Z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
-4.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.90	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.80	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.70	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.60	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
-3.50	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.40	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.30	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.20	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.10	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.00	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.90	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.80	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.70	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.60	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.50	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.40	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.30	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.20	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.10	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.00	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.90	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.80	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.70	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.60	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.50	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668

Continuation

Z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
-1.40	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.30	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.20	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.10	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.00	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.90	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.80	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.70	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.60	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.50	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.40	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.30	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.20	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.10	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
0.00	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

Continuation

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.60	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.70	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.80	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.20	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.30	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.40	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.50	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

Continuation

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.60	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.80	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.90	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.00	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.10	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.20	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.30	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.40	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.50	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.60	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.70	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.80	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.90	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

7.2 Constants for determining control limits charts in graphs $\bar{X}$ and R

n	d2	A2	D3	D4
2	1.128	1.881	0	3.267
3	1.693	1.023	0	2.575
4	2.059	0.729	0	2.282
5	2.326	0.577	0	2.115
6	2.534	0.483	0	2.004
7	2.704	0.419	0.076	1.924
8	2.847	0.373	0.136	1.864
9	2.97	0.337	0.184	1.816
10	3.078	0.308	0.223	1.777
15	3.472	0.223	0.348	1.652
20	3.735	0.180	0.414	1.586

References

- Devor, R., Chang T., and Sutherland, J. 1992. Statistical Quality Design and Control: Contemporary Concepts and Methods. Macmillan Publishing Company. Ontario, Canada