

Estatística e Qualidade



Fabrício Bueno

Sumário

1	Introdução.....	1
2	Folhas de verificação.....	2
2.1	Proporção de itens não-conformes.....	2
2.2	Inspeção de atributos.....	3
2.3	Localização de defeitos.....	4
2.4	Causa dos defeitos.....	4
2.5	Distribuição de uma variável.....	5
2.6	Monitorar processo de fabricação.....	6
3	Diagrama de Pareto.....	8
3.1	Construindo um diagrama de Pareto.....	8
3.2	Análise do diagrama de Pareto.....	11
4	Histograma.....	12
4.1	Construindo um Histograma.....	12
5	Gráfico de Controle.....	17
5.1	Indicadores de processos fora de controle.....	18
5.1.1	Periodicidade.....	18
5.1.2	Tendência.....	19
5.1.3	Deslocamento.....	19
5.1.4	Pontos fora dos limites.....	20
5.2	Tipos de gráficos de controle.....	21
5.2.1	Gráfico de Controle np.....	21
5.2.2	Gráfico de Controle p.....	24
5.2.3	Gráfico de Controle c.....	27
5.2.4	Gráfico de Controle $\bar{x} - R$	30
5.2.5	Gráfico de Controle $\bar{x} - S$	34
5.2.6	Gráfico de Controle \bar{u}	37
5.3	Considerações finais.....	41
6	Diagrama de Dispersão.....	42
6.1	Construindo um diagrama de dispersão.....	42
6.2	Calculando a correlação entre variáveis.....	46
6.2.1	Cuidados na interpretação da correlação.....	47
6.3	Análise de Regressão	48
	Anexo.....	51
	Bibliografia.....	52

Índice definido pelo usuário

<i>Tabela 2.1 - Levantamento de itens conformes e não-conformes.....</i>	<i>3</i>
<i>Tabela 2.1 - Levantamento de itens conformes e não-conformes.....</i>	<i>3</i>
<i>Tabela 2.2 - Levantamento tipos de defeitos.....</i>	<i>3</i>
<i>Tabela 2.3 – Levantamento de causas e efeitos.....</i>	<i>5</i>
<i>Tabela 2.4 – Distribuição de uma variável.....</i>	<i>6</i>
<i>Tabela 2.5 – Monitoramento de um processo de fabricação.....</i>	<i>7</i>
<i>Tabela 3.1 – Distribuição dos tipos de defeitos gerados por máquinas.....</i>	<i>9</i>
<i>Tabela 3.2 – Distribuição dos tipos de defeitos investigados.....</i>	<i>9</i>
<i>Tabela 3.3 – Custo de correção de defeitos.....</i>	<i>11</i>
<i>Tabela 4.1 – Dados organizados em classe.....</i>	<i>13</i>
<i>Tabela 5.1 – Medidas de amostras.....</i>	<i>23</i>
<i>Tabela 5.3 – Tabela com exemplos de monitoramento de fabricação.....</i>	<i>31</i>
<i>Tabela 5.4 - Uma comparação dos resultados de ambos os gráficos.....</i>	<i>33</i>
<i>Tabela 5.5 – Tabela com exemplos de monitoramento de fabricação.....</i>	<i>35</i>
<i>Tabela 5.6 – Número de defeitos por rolo de tecido.....</i>	<i>38</i>
<i>Tabela 5.7 – Número médio de defeitos por unidade.....</i>	<i>38</i>
<i>Tabela 5.8 – Cálculo dos limites de controle.....</i>	<i>39</i>

Índice definido pelo usuário

<i>Figura 3.1 – Gráfico de Barras.....</i>	<i>10</i>
<i>Figura 3.2 - Diagrama de Pareto.....</i>	<i>10</i>
<i>Figura 4.1 – Exemplo de Histograma.....</i>	<i>14</i>
<i>Figura 5.1 – Exemplo de um gráfico de controle.....</i>	<i>17</i>
<i>Figura 5.2 – Gráfico de controle indicando periodicidade.....</i>	<i>18</i>
<i>Figura 5.3 – Gráfico de controle indicando tendência.....</i>	<i>19</i>
<i>Figura 5.4 – Gráfico de controle indicando deslocamento.....</i>	<i>20</i>
<i>Figura 5.5 – Gráfico de controle indicando pontos fora dos limites.....</i>	<i>21</i>
<i>Figura 5.6 – Gráfico de controle do exemplo.....</i>	<i>24</i>
<i>Figura 5.7 – Gráfico de controle do exemplo.....</i>	<i>26</i>
<i>Tabela 5.2 – Defeitos em dez unidades produzidas.....</i>	<i>28</i>
<i>Figura 5.8 – Gráfico de controle c.....</i>	<i>29</i>
<i>Figura 5.9 – Gráfico de controle para médias das médias das amostras.....</i>	<i>32</i>
<i>Figura 5.10 – Gráfico de controle para médias das amplitudes.....</i>	<i>32</i>
<i>Figura 5.11 – Gráfico de Controle para média das médias das amostras.....</i>	<i>36</i>
<i>Figura 5.12 – Gráfico de Controle para média dos desvios padrões das amostras.....</i>	<i>36</i>
<i>Figura 5.13 – Gráfico de controle u</i>	<i>39</i>
<i>Figura 5.14 – Gráfico de controle u com limites fixos.....</i>	<i>40</i>
<i>Figura 6.1 – Gráfico de dispersão de defeitos por distância de entrega de produtos.....</i>	<i>43</i>
<i>Figura 6.2 – Exemplo de gráfico de dispersão indicando correlação positiva fraca.....</i>	<i>44</i>
<i>Figura 6.3 – Exemplo de gráfico de dispersão indicando correlação negativa forte.....</i>	<i>44</i>
<i>Figura 6.4 – Exemplo de gráfico de dispersão indicando correlação positiva perfeita.....</i>	<i>45</i>

1 Introdução

O Controle Estatístico do Processo (CEP) é uma poderosa ferramenta de qualidade. Ao se implementar CEP, é necessário a coleta de dados, para posterior análise estatística. A coleta pode ser destinada à inspeção (aceite ou rejeite de produtos), monitoração (acompanhamento de desempenho do processo) e controle (visando a diminuição de perdas).

A coleta de dados deve ser estratificada, ou seja, deve ser feita entre processos de produção homogêneos. Isto significa que não se pode analisar estatisticamente a qualidade de processos que usam máquinas, insumos, matérias primas, métodos e outras características diferentes. Portanto, antes de iniciar uma coleta, certifique-se que a amostra está estratificada.

Além da estratificação é necessário que a coleta de dados seja planejada, de forma que o uso de dados seja fácil e não necessite de etapas de correção e filtragem. Uma ferramenta bastante útil para planejar a coleta é a elaboração de folhas de verificação.

2 Folhas de verificação

Folhas de verificação são planilhas para registro de dados, que tornam a coleta rápida e automática. Por exemplo, imagine um supervisor de linha de produção que anota todas as informações sobre o processo de produção em folhas de papel, sem padronização ou planejamento. No fim de uma semana ele já terá enorme dificuldade em organizar e interpretar os vários conjuntos de dados obtidos: horários, máquinas, peças, defeitos, operadores, e outros. O uso de folhas de verificação poderia evitar este inconveniente.

Uma folha de verificação deve conter informações como local e data de coleta, nome do responsável pela coleta, nome do operador, máquina, identificação do produto/peça, defeitos, medidas, e outros.

A organização e as informações de uma folha de verificação dependem da sua finalidade. Vejamos alguns exemplos.

2.1 *Proporção de itens não-conformes*

Quando a análise da produção consiste na simples classificação de conforme ou não-conforme, a folha de verificação deverá conter o número de itens inspecionados (n), o número de itens não-conformes (d) e a proporção de itens não conformes (p), que pode ser calculada de forma automatizada em planilhas eletrônicas. Um exemplo desta folha de verificação pode ser visto na Tabela 2.1.

Tabela 2.1 - Levantamento de itens conformes e não-conformes

Peça/Produto:					Processo:				
Operador:					Máquina:				
Data:					Local:				
	Amostra								
	1	2	3	4	5	6	7	8	9
n									
d									
p									

2.2 Inspeção de atributos

Analisar se um produto ou peça está ou não conforme pode ser insuficiente para detecção de falhas na linha de produção, e até mesmo para planejamento de relacionamentos pós-venda. Portanto, pode-se usar folhas de verificação para registros de tipos de defeitos, conforme a Tabela 2.2.

Tabela 2.2 - Levantamento tipos de defeitos

Peça/Produto:										Processo:									
Operador:										Máquina:									
Data:										Local:									
Tipo de defeito		Contagem																Total	
Risco																			
Pintura																			
Amassado																			
Funcionamento																			
Outro																			

Esta folha de verificação permite que se conte as ocorrências dos defeitos em um lote. A partir destes dados, pode-se obter a porcentagem, média, mediana, moda, desvio padrão e outras medidas sobre os defeitos.

2.3 Localização de defeitos

Os defeitos externos, quando devidamente documentados, também podem fornecer importantes pistas para detecção de problemas em linhas de produção. Neste caso as folhas de verificação possuem um croqui do produto, bastando marcar o tipo de defeito e o local onde ocorreu. Cada folha normalmente é aplicável a um único produto ou peça.

2.4 Causa dos defeitos

Há várias causas possíveis para um defeito: desajuste de máquinas, desgaste de ferramentas, método, operador, matéria prima, dia da semana, horário, e outros. A folha de verificação para causas dos defeitos deve ser feita a partir de um levantamento de causas prováveis. Um exemplo pode ser visto na Tabela 2.3.

Tabela 2.3 – Levantamento de causas e efeitos

Peça/Produto:		Processo:				
Operador:		Máquina:				
Data:		Local:				
		Dia				
Máquina	Matéria-prima	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a
A	1					
	2					
	3					
B	1					
	2					
	3					
C	1					
	2					
	3					

Esta folha de verificação, por exemplo, pode apontar defeitos mais frequentes nas segundas-feiras para a máquina A usando o insumo 1. A partir desta constatação, pode-se ampliar o monitoramento do processo neste dia da semana, verificando se há alguma característica na máquina, na operação ou na matéria-prima que diferencie o processo dos demais dias.

2.5 Distribuição de uma variável

Pode-se querer estudar como uma variável é distribuída em uma faixa de valores para verificar o quão corretamente as especificações estão sendo atendidas. Por exemplo, pode-se estudar as diversas variações de espessuras de um fio produzido para verificar o quanto e com que frequência se obtém fios fora da especificação. Para tanto poderia ser elaborada uma folha de dados conforme a apresentada na Tabela 2.4.

Tabela 2.4 – Distribuição de uma variável

Peça/Produto:		Processo:						
Operador:		Máquina:						
Data:		Local:						
Espessura	Contagem							Total
	1	2	3	4	5	6	7	
menos de 0,005								
0,0055								
0,006								
0,0065								
0,007								
0,0075								
0,008 ou mais								

Na coluna de espessuras são exibidos, com exceção da primeira linha, os limites inferiores dos intervalos. A cada medida que se enquadra numa linha, marca-se o contador (que evidentemente pode ter mais unidades que neste exemplo). Ao fim da inspeção, pode-se verificar a quantidade de fios dentro da especificação (entre 0,0065 e 0,0069) e quantos e como se distribuíram os fios fora de especificação, e a partir daí tomar as devidas providências, se necessário.

2.6 Monitorar processo de fabricação

Para monitorar a variação de alguma característica de um produto ou peça, pode-se coletar diariamente uma amostra de tamanho n . Os dados coletados devem ser organizados em uma folha de verificação conforme na Tabela 2.5.

Tabela 2.5 – Monitoramento de um processo de fabricação

Peça/Produto:		Processo:						
Operador:		Máquina:						
Data:		Local:						
	Amostra							
	1	2	3	4	5	6	7	8
X ₁								
X ₂								
X ₃								
X ₄								
Média								
Amplitude								

Nesta tabela, cada amostra possui quatro unidades que terão alguma característica medida e anotada, e serão calculadas a média e amplitude para posterior análise, conforme será visto na seção de gráficos de controle.

3 Diagrama de Pareto

A qualidade nas linhas de produção tem como finalidade a redução de perdas. As ferramentas de qualidade visam a identificação das causas destas perdas e a ordem em que devem ser sanadas. De acordo com estes pressupostos, analisaremos a ferramenta de qualidade diagrama de Pareto.

3.1 Construindo um diagrama de Pareto

A construção deste diagrama inicia-se com a coleta e organização de dados, que deve ocorrer após serem definidos:

- Tipo de perda que se quer investigar. Por exemplo: peças fora da especificação, com defeitos, acidentes;
- Quais informações sobre as perdas se quer obter. Por exemplo: em produtos com defeitos pode se querer saber o local do defeito, o tipo do defeito ou máquinas/operadores que produzem o defeitos;

Estabelecidos os tipos de perdas e que informações sobre elas se deseja obter, deve-se organizar uma folha de verificação com as categorias das informações a serem investigadas (conforme exemplificado na [seção 2.2](#)). Após o preenchimento da folha de verificação, deve-se fazer as contagens e organizar as categorias por ordem decrescente de frequência. As categorias com frequências baixas podem ser agrupadas sobre o nome de *outros*, sempre na última linha. Um exemplo de contagem e organização de tipo de defeitos pode ser visto na tabela 3.1.

A seguir, deve-se calcular as frequências relativas, acumuladas e relativas acumuladas de cada categoria. Conforme exibido na tabela 3.2.

Tabela 3.1 – Distribuição dos tipos de defeitos gerados por máquinas

Tipo de Defeito	Frequência
Risco	25
Pintura	20
Amassado	17
Funcionamento	11
Mancha	7
Outros	15
Total	95

Tabela 3.2 – Distribuição dos tipos de defeitos investigados

Tipo de Defeito	Frequência	Frequência relativa	Frequência acumulada	Frequência relativa acumulada
Risco	25	26,32%	25	26,32%
Pintura	20	21,05%	45	47,37%
Amassado	17	17,89%	62	65,26%
Funcionamento	11	11,58%	73	76,84%
Mancha	7	7,37%	80	84,21%
Outros	15	15,79%	95	100,00%
Total	95	100,00%	-	-

A frequência relativa é calculada pela divisão de cada frequência pelo total de ocorrências do defeito. Ou seja, a frequência relativa de Amassados é dada por: $[17/95]*100$ (uma vez que se trata de porcentagem). Já a frequência acumulada é calculada pela soma de cada frequência com as anteriores. Por exemplo, a frequência acumulada de Funcionamento é dada por: $25+20+17+11$. A frequência relativa acumulada é calculada de forma análoga à a frequência relativa.

A partir destes dados é possível desenhar o gráfico de barras baseado na frequência das categorias. Este gráfico é exibido na figura 3.1. A este gráfico pode-se adicionar a curva de Pareto, que corresponde à curva das frequências acumuladas (ou frequências acumuladas relativas). Na

Figura 3.2 é exibido o diagrama de Pareto.

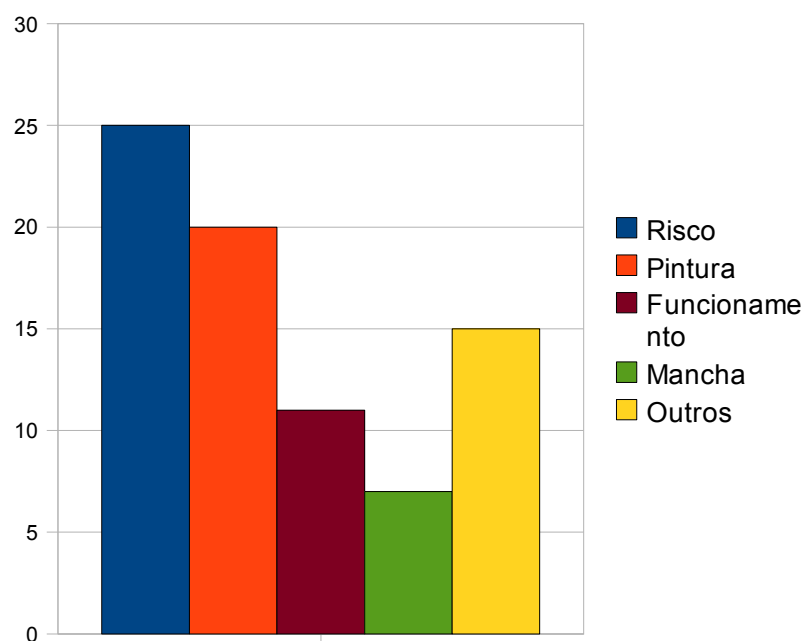


Figura 3.1 – Gráfico de Barras

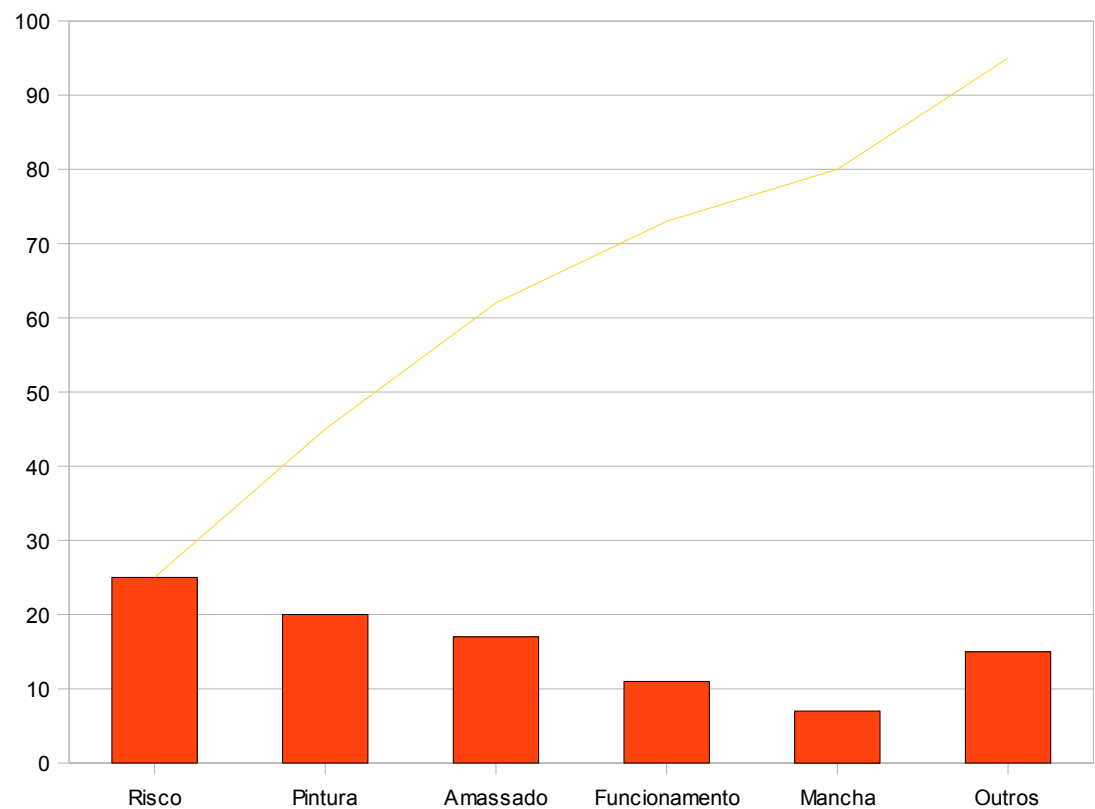


Figura 3.2 - Diagrama de Pareto

3.2 Análise do diagrama de Pareto

À análise dos gráficos da seção anterior, fazem-se necessárias as seguintes considerações:

1. O diagrama de Pareto mostra, visualmente, os defeitos mais frequentes que, a princípio devem ser tratados com prioridade: risco e pintura;
2. Há casos em que este diagrama pode não ser suficiente. Por exemplo, caso os defeitos apresentassem frequências similares, não seria possível uma conclusão precisa. Nestes casos é recomendável usar diagramas de Pareto para todas as possíveis fontes de falhas: operadores, máquinas, insumos, dias da semana. A partir da análise conjunta destes dados, pode-se obter melhores conclusões;
3. A frequência de defeitos podem não ser a única fonte na tomada de decisão. Os custos desencadeados pelos defeitos (seja pelo desperdício, repetição de tarefas, ou pós-vendas), combinados às frequências, podem fornecer informações mais precisas. Por exemplo, a tabela 3.3 pode indicar que o defeito mais frequente nem sempre é o mais prioritário em termos de custos. É recomendável ainda, que se gere um diagrama de Pareto para custos, visando uma melhor representação gráfica dos dados;
4. A curva de Pareto mostra, visualmente, a proporção dos defeitos mais frequentes em relação ao total. No exemplo analisado, risco, amassado e pintura correspondem a cerca de 65% dos defeitos. Logo, seria uma medida interessante buscar causas em comum entre estes três tipos de defeitos;
5. Caso a categoria outros possua uma porcentagem muito alta, deve-se rever a classificação das categorias, ou seja, algumas categorias não deveriam estar sobre a discriminação Outros.

Tabela 3.3 – Custo de correção de defeitos

Tipo de Defeito	Frequência	Custo Unitário(R\$)	Custo Total
Risco	25	10	250
Pintura	20	12	240
Amassado	17	5	85
Funcionamento	11	40	440
Mancha	7	10	70
Outros	15	5	75
Total	95	82	1160

4 Histograma

Histograma é uma representação gráfica da distribuição de frequência de uma massa de dados divididos em categorias¹, normalmente um gráfico de barras verticais. É uma importante ferramenta de qualidade que permite uma fácil visualização de grandes amostras.

O histograma é um gráfico composto por retângulos justapostos em que a base de cada um deles corresponde ao intervalo de classe e a sua altura à respectiva frequência. A construção de histogramas tem caráter preliminar em qualquer estudo e é um importante indicador da distribuição de dados, podendo indicar se uma distribuição aproxima-se de uma função normal e se há mistura de populações quando as amostras se apresentam bimodais.

4.1 Construindo um Histograma

Após a coleta de dados, estes devem ser agrupados em classes e organizados em uma tabela de distribuição de frequências. Não há um número “certo” de classes, mas existem fórmulas que fornecem uma estimativa de uma quantidade “ideal” de classes, embora nem sempre sejam necessárias. Uma delas é:

$$k = \sqrt{n}$$

onde k é o número de classes e n é o número de dados da amostra.

Por exemplo, para uma amostra com 150 dados pode-se utilizar aproximadamente, 12 classes. Porém, nada impede que, dependendo das particularidades das amostras, se adotem valores diferentes dos obtidos pelo uso de fórmulas.

Uma vez definido o número de classes, deve-se definir o tamanho do intervalo de cada

¹Quando o número de dados aumenta indefinidamente e o intervalo de classe tende a zero, a distribuição de frequência passa para uma distribuição de densidade de probabilidades.

classe, que é calculado pela divisão da amplitude da amostra (ou seja, o valor máximo *MAX* subtraído do valor mínimo *MIN*) pelo número de classes *k*.

$$I = \frac{MAX - MIN}{k}$$

Tomemos como exemplo uma amostra com tempos que operadores levam para montar um produto. Sabe-se que o tamanho da amostra é de 120, o maior tempo é 59 e o menor 10 (medidas de tempo não costumam ser tão discrepantes, usamos estes apenas com fins didáticos). Logo, pode-se calcular um número de classes equivalente a aproximadamente 11. Porém, este pode não ser um valor interessante, pois a amplitude da amostra é 49 (59-10), e o intervalo calculado não seria um valor exato. Portanto, poderia-se adotar um número de 7 classes, com intervalos exatos de valor 7, coincidentemente. Mas, ainda assim não teríamos classes de fácil visualização, pois em intervalos de 7 não teríamos um padrão visual. Pode-se, mais uma vez, sugerir uma alteração nestes valores. Para haver um padrão visual, intervalos de 5 ou 10 anos seria mais recomendável. Concluindo, há casos em que as fórmulas não são perfeitamente aplicáveis, sendo mais indicado o bom senso. A Tabela 4.1 traz os dados organizados desta amostra.

Tabela 4.1 – Dados organizados em classe

Classe	Ponto médio	Frequência	Frequência relativa
10 - 15	12,5	4	3,33%
15 - 20	17,5	5	4,17%
20 - 25	22,5	23	19,17%
25 - 30	27,5	27	22,50%
30 - 35	32,5	21	17,50%
35 - 40	35,5	13	10,83%
40 - 45	42,5	10	8,33%
45 - 50	47,5	7	5,83%
50 - 55	52,5	6	5,00%
55 - 60	57,5	4	3,33%
Total:		120	

A partir dos dados da Tabela 4.1, têm-se o gráfico exibido na Figura 4.1. Onde as barras são as frequências com os respectivos pontos médios da categoria.

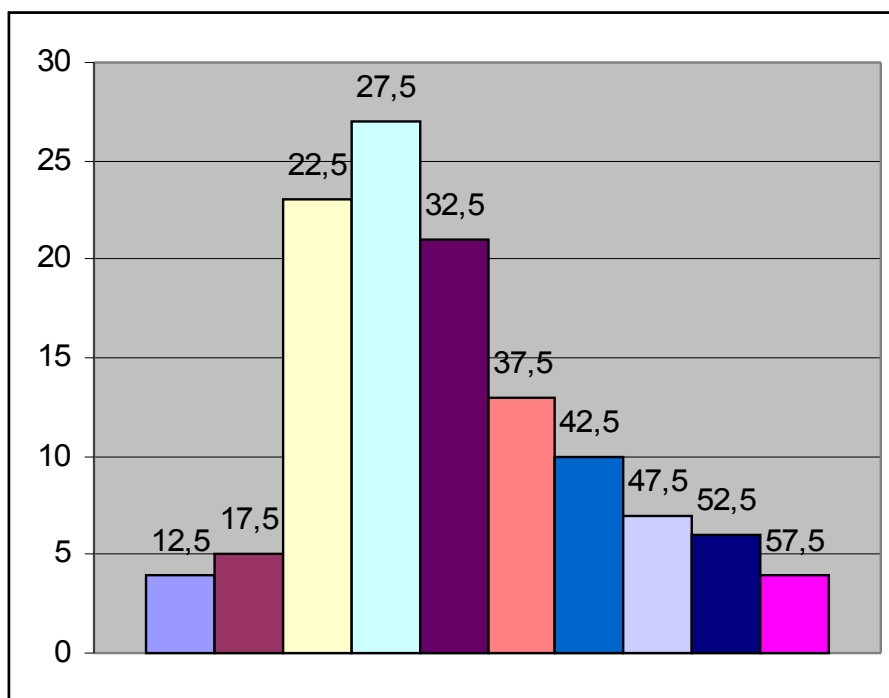


Figura 4.1 – Exemplo de Histograma

A interpretação deste Histograma depende dos controles de limites de tempo pela fábrica. Por exemplo, se é estabelecido um limite máximo de 50 minutos, e um limite mínimo de 15 minutos, podemos verificar que há três categorias fora de especificação. A média aritmética ponderada destes dados é aproximadamente 32 que bate perfeitamente com a média dos valores extremos permitidos. Isto quer dizer que não são necessários esforços para corrigir a média dos tempos dos operadores. Porém há necessidade de diminuir a variação dos tempos medidos, sendo necessárias ações corretivas na linha de produção.

Outro fato positivo apontado pelo histograma é a assimetria positiva, indicando que a maioria das medidas estão abaixo da média, o que é interessante em um sistema que visa minimização de tempo como uma linha de produção.

Generalizando a interpretação de histogramas:

1. Nos processos industriais a forma desejável do Histograma é que ele se aproxime ao máximo da distribuição normal, sendo, portanto, simétrico. Em outras palavras, é uma situação desejável que a média dos dados esteja no centro do gráfico e que as frequências mais altas também estejam no centro;
2. Histogramas com assimetria positiva, ou seja, a média dos dados está à esquerda do centro da figura e a cauda à direita é alongada, é um forma muito comum quando o limite inferior é controlado e não podem ocorrer valores abaixo de determinado limite;
3. Histogramas com assimetria negativa, ou seja, a média dos dados está localizada à direita do centro da figura e a cauda à esquerda é alongada. Esta forma é muito comum quando o limite superior é controlado ou quando não podem ocorrer valores acima de certo limite;
4. Histogramas em *plateau*, isto é, com várias classes com frequências quase iguais, com exceção das primeiras e últimas classes, indica que várias distribuições (ou amostras) com diferentes médias foram misturadas;
5. Histogramas com dois picos, ou duas modas, e com frequências baixas no centro, indica que duas distribuições (amostras) com médias bem diferentes foram misturadas. Isto pode ocorrer, por exemplo, quando são misturados produtos de dois turnos de trabalho;
6. Histogramas também mostram o grau de dispersão da variável. Em Histogramas cujas categorias mais afastadas do centro possuem frequências altas e aproximadas entre si, indicam alto grau de dispersão.
7. Para melhor interpretação de um Histograma, deve-se traçar os limites de especificação superior e inferior. O que permitirá a identificação de folga (quando as categorias extremas estão relativamente distantes dos limites), o que representa uma situação desejada, pois, neste caso, a probabilidade de ocorrência de valores fora de especificação é baixa. Caso

contrário, se não existir folga (quando as categorias extremas estão sobre ou além dos limites de especificação), não há margem de segurança, havendo maior probabilidade de valores fora da especificação. Neste último caso, uma ação corretiva se faz necessária para diminuir a variabilidade dos valores.

Resumindo, um Histograma ideal deve ter forma simétrica, dispersão pequena e média no centro. Um exercício recomendável seria encontrar situações que reproduzam as configurações de Histograma citadas acima.

5 Gráfico de Controle

Gráfico de controle é uma ferramenta que permite a observação da variação de características de qualidade ao longo do tempo. É constituído por uma linha central cercada por um par de limites de controle, e por pontos que representam valores de estados do processo, ou seja, amostras tomadas em momentos diferentes. Os pontos normalmente são unidos por segmentos de reta, o que permite melhor visualização da evolução da característica analisada ao longo do tempo.

Se todos os pontos estiverem marcados dentro dos limites de controle, sem apresentarem tendências (ou seja, dispostos aleatoriamente), o processo é considerado sob controle. Porém, se houver pontos fora dos limites de controle ou uma disposição que indique tendência a possíveis falhas, o processo é visto como fora de controle e deve ser verificado. A figura 5.1 exibe um exemplo de um gráfico de controle.

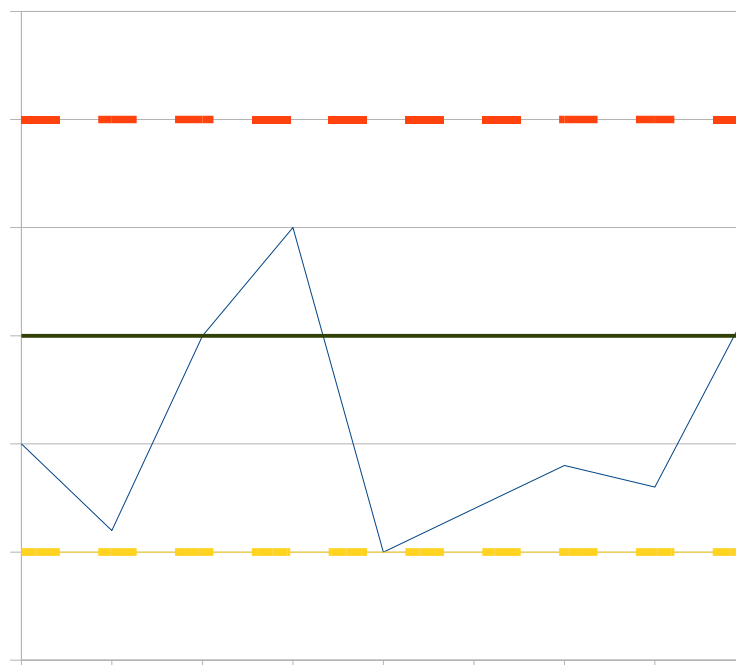


Figura 5.1 – Exemplo de um gráfico de controle

5.1 – Indicadores de processos fora de controle

Os gráficos de controle podem indicar “anomalias” no processo. São indicadores:

5.1.1 Periodicidade

“Subidas” e “descidas” em intervalos regulares de tempo, conforme exibido na Figura 5.2, podem indicar mudanças periódicas no processo. Estas mudanças podem se dever, por exemplo, a trocas de máquinas, operadores, variações de condições de operação (temperatura, pressão, voltagem) ou outros fatores que possam promover instabilidade ao processo.

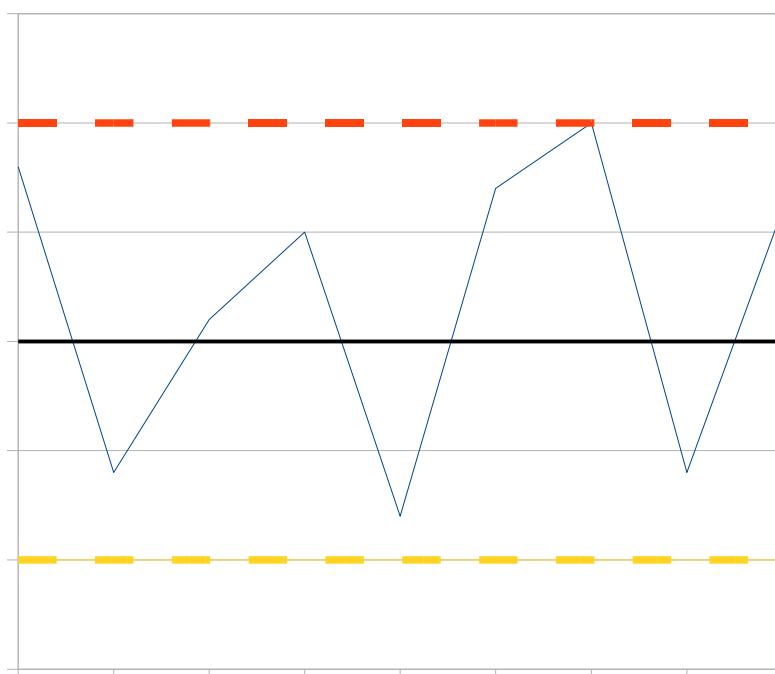


Figura 5.2 – Gráfico de controle indicando periodicidade

Se, como exemplo, considerar-se que a Figura 5.2 trata-se de uma medição de um produto ao longo do dia de um processo de produção, poderia-se verificar variações entre picos extremos em relação à média em determinadas horas do dia. A partir desta informação, poderia-se, nas horas indicadas, investigar quais fatores poderiam estar causando estas variações e buscar corrigi-los, antes que desencadeiem possíveis falhas na produção.

5.1.2 Tendência

Alguns pontos em seqüência podem indicar alguma tendência, conforme exibido na Figura 5.3. Pode-se notar uma tendência para baixo, ou seja, para o limite inferior de valores.

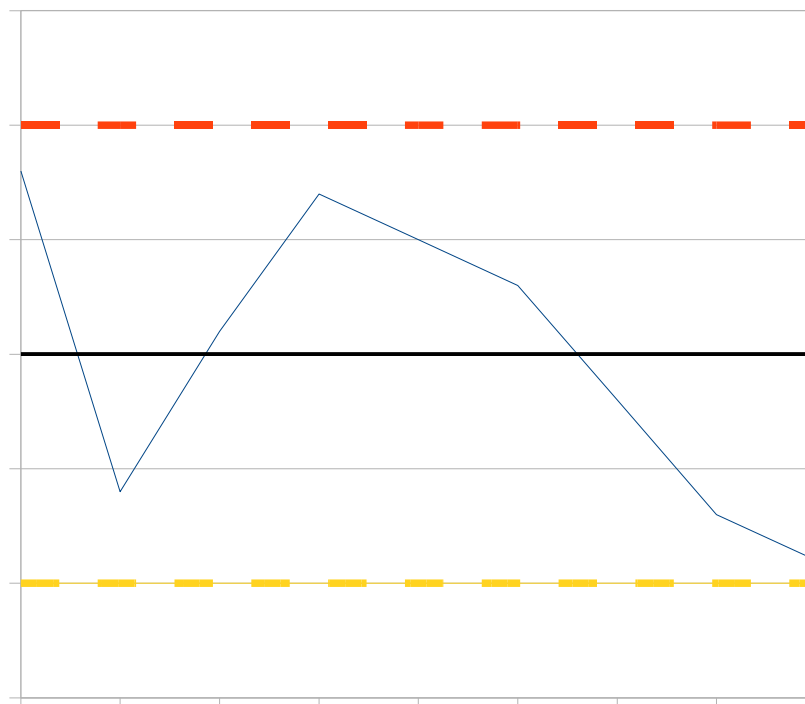


Figura 5.3 – Gráfico de controle indicando tendência

Em geral, uma tendência pode significar uma deterioração gradual de um fator crítico do processo. Pode-se dever ao desgaste de uma máquina, cansaço do operador, deposição de resíduos, e outros fatores a serem corrigidos.

5.1.3 Deslocamento

Deslocamentos são mudanças no nível de desempenho de um processo, conforme exibido na Figura 5.4. Tais mudanças de desempenho podem ser causadas por adoção de novas metodologias,

troca de operadores ou máquinas, e outros fatores que devem ser investigados.

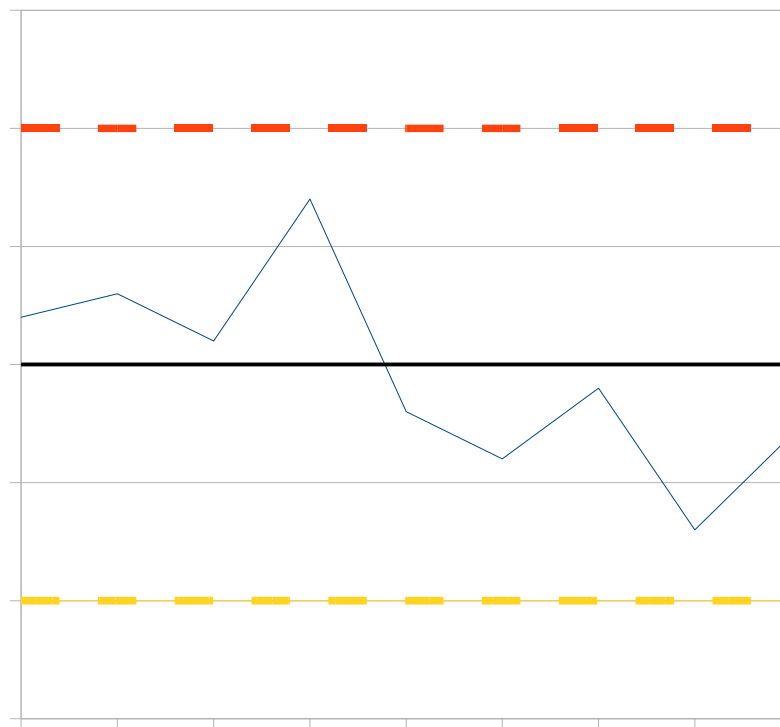


Figura 5.4 – Gráfico de controle indicando deslocamento

5.1.4 Pontos fora dos limites

Esta é a situação mais evidente de um processo fora de controle. Caso uma porção significativa de pontos ultrapassem os limites, deve-se investigar e sanar as causas deste descontrole.

A Figura 5.5 exibe uma situação em que algumas amostras, recolhidas ao longo de um dia de produção, extrapolam os limites de controle.

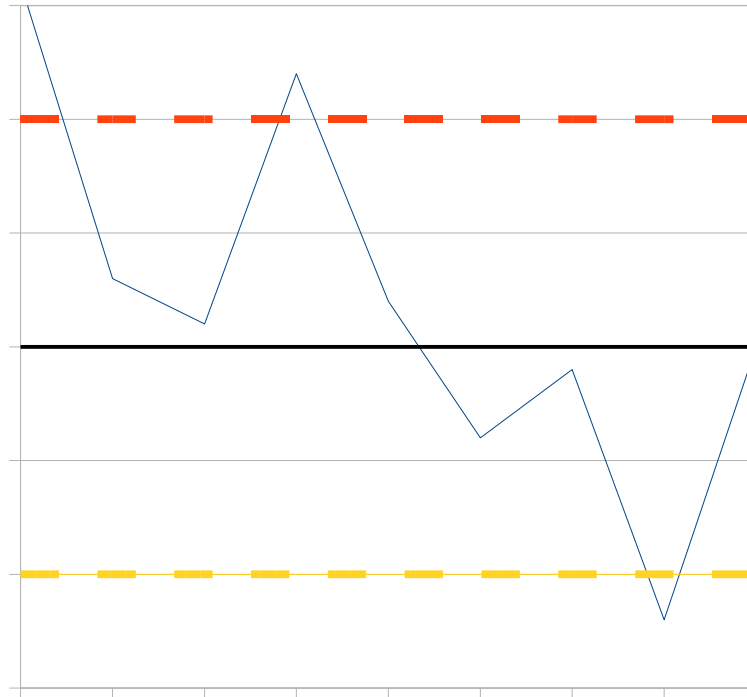


Figura 5.5 – Gráfico de controle indicando pontos fora dos limites

Nesta situação, deve-se investigar quais causas podem ter influenciado as falhas, nos horários em que elas ocorreram.

5.2 Tipos de gráficos de controle

Dentre os tipos de gráficos de controle, analisaremos dois: gráfico np e gráfico $\bar{x} - R$.

5.2.1 Gráfico de Controle np

Este gráfico monitora a variação de um número de itens não-conformes em amostras de tamanho constante. Para montar este gráfico, será necessário:

1. Usar uma folha de verificação da [Tabela 2.1](#). Com base nesta folha, têm-se o número de

itens não-conformes (d) e sua proporção (p) em relação ao tamanho das amostras(n).

2. Em seguida deve-se calcular a média das proporções de itens não-conformes em cada amostra.

$$\bar{p} = \frac{1}{m} \sum p_i = \frac{1}{m} \sum \frac{d_i}{n} = \frac{1}{mn} \sum d_i$$

onde m é o número de amostras. Resumindo, a média é obtida pela soma de todos os itens não-conformes das amostras, dividida pelo tamanho das amostras vezes o número de amostras.

3. Calcular o número médio de itens não-conformes:

$$\bar{d} = \frac{1}{m} \sum d_i$$

4. Calcular o limite superior de controle (LSC) e o limite inferior de controle (LIC). Se o limite inferior for negativo, iguale-o a zero:

$$LCS = \bar{d} + 3\sqrt{\bar{d}(1-\bar{p})}$$

$$LIC = \bar{d} - 3\sqrt{\bar{d}(1-\bar{p})}$$

5. Trace o sistema de eixos cartesianos (no eixo das abscissas deve estar o número das amostras (ou o tamanho da amostra, caso se queira analisar uma única amostra isoladamente) e no eixo das ordenadas, o número de itens não-conforme);
6. Trace no gráfico uma linha central, com base no número médio de itens não-conformes;
7. Trace o LSC e o LIC;
8. Faça um ponto para cada par de valores (itens não-conforme e número da amostra);

9. Una os pontos por segmentos de reta;

Com base nestes passos, façamos um exemplo com base nos dados da Tabela 5.1.

Tabela 5.1 – Medidas de amostras

	Amostras				
	1	2	3	4	5
n	10	10	10	10	10
d	2	4	1	3	5
p	0,2	0,4	0,1	0,3	0,5

Para estes valores temos os cálculos:

$$\bar{p} = \frac{2+4+1+3+5}{5*10} = \frac{15}{50} = 0,3$$

$$\bar{d} = \frac{15}{5} = 3$$

$$LCS = 3 + 3\sqrt{3(1-0,3)} \approx 7,35$$

$$LIS = 3 - 3\sqrt{3(1-0,3)} \approx -1,35$$

Como o limite inferior é negativo, será tomado como zero. O gráfico de controle é exibido na Figura 5.6.

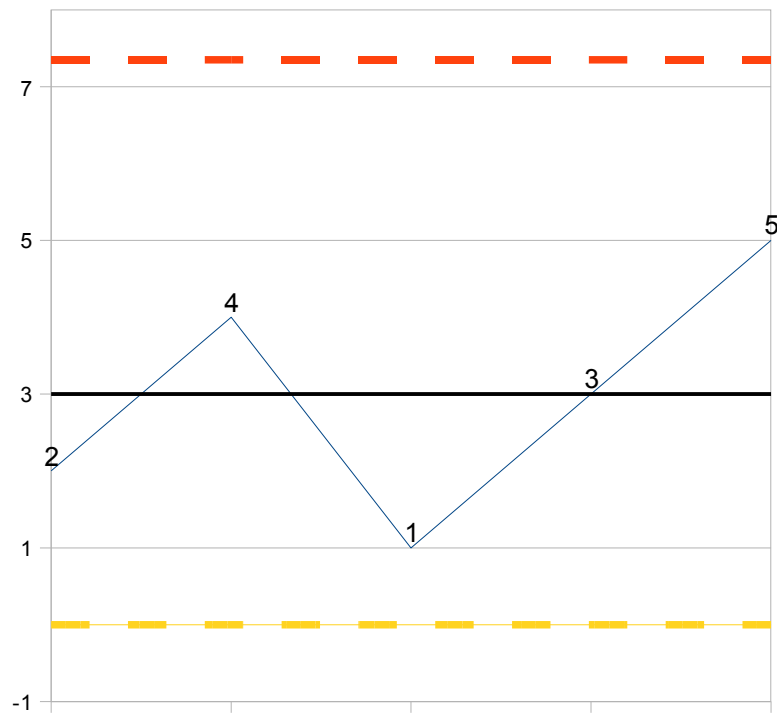


Figura 5.6 – Gráfico de controle do exemplo

O número e tamanho das amostras são pequenos pois este exemplo visa não reproduzir uma situação real, mas ser didático. Portanto, temos um gráfico de controle muito pequeno, que não possibilita interpretações precisas. Porém, estes poucos dados sugerem periodicidade e tendência.

5.2.2 Gráfico de Controle p

Este gráfico monitora a variação da proporção de números de itens não-conformes e o total de itens inspecionados em amostras de tamanho constante ou variável. Para montar este gráfico, será necessário:

1. Usar uma folha de verificação da [Tabela 2.1](#). Com base nesta folha, têm-se o número de itens não-conformes (d) e sua proporção (p) em relação ao tamanho das amostras(n);

2. Em seguida deve-se calcular a média das proporções de itens não-conformes em cada amostra.

$$\bar{p} = \frac{1}{m} \sum p_i = \frac{1}{m} \sum \frac{d_i}{n} = \frac{1}{mn} \sum d_i$$

onde m é o número de amostras. Resumindo, a média é obtida pela soma de todos os itens não-conformes das amostras, dividida pelo tamanho das amostras vezes o número de amostras;

3. Calcular o limite superior de controle (LSC) e o limite inferior de controle (LIC). Se o limite inferior for negativo, iguale-o a zero:

$$LCS = \bar{p} + 3 \sqrt{\bar{p} \frac{(1-\bar{p})}{n}}$$
$$LIC = \bar{p} - 3 \sqrt{\bar{p} \frac{(1-\bar{p})}{n}}$$

5. Trace o sistema de eixos cartesianos (no eixo das abscissas devem estar o número das amostras (ou o tamanho da amostra, caso se queira analisar uma única amostra isoladamente) e no eixo das ordenadas, a proporção de itens não-conforme);
6. Trace no gráfico uma linha central, com base na proporção média de itens não-conformes;
7. Trace o LSC e o LIC;
8. Faça um ponto para cada par de valores (proporção de não conformes e número da amostra);
9. Una os pontos por segmentos de reta.

Com base nestes passos, façamos um exemplo com base nos dados da Tabela 5.1.

Para estes valores temos os cálculos:

$$\bar{p} = \frac{2+4+1+3+5}{5*10} = \frac{15}{50} = 0,3$$

$$LCS = 0,3 + 3\sqrt{0,3 \frac{(1-0,3)}{10}} \approx 0,735$$

$$LIS = 0,3 - 3\sqrt{0,3 \frac{(1-0,3)}{10}} \approx -0,135$$

Como o limite inferior é negativo, será tomado como zero. O gráfico de controle é exibido na Figura 5.7.

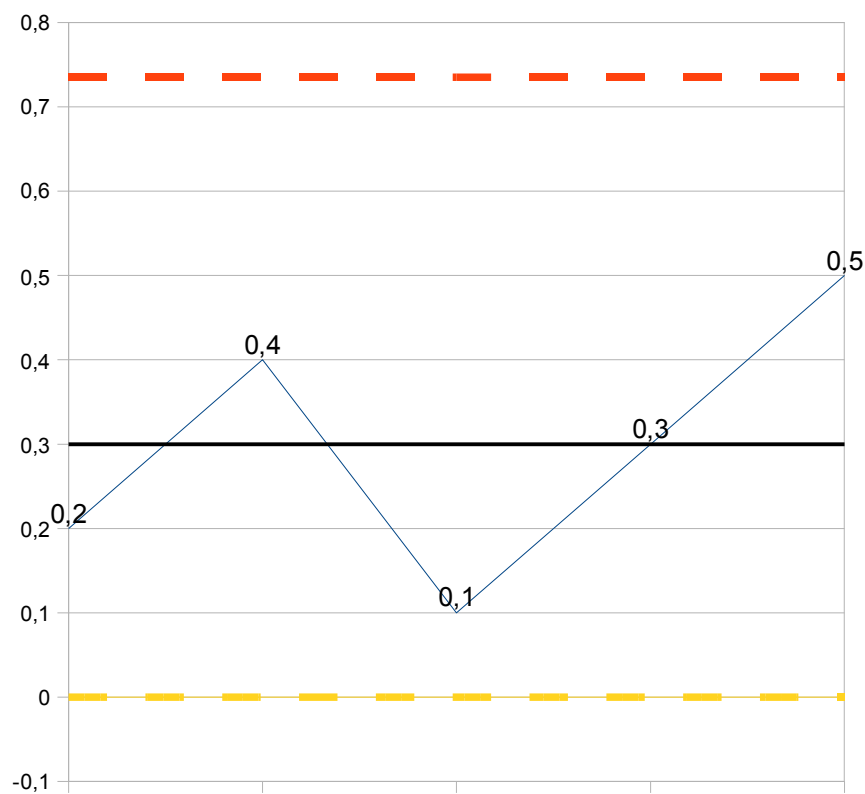


Figura 5.7 – Gráfico de controle do exemplo

Para novos dados, este método, quando aplicado em um processo sob controle (conforme visto na seção 5.1), não requer o recálculo das médias das proporções de itens não-conformes, nem dos limites. Caso o processo não esteja sob controle, a média e os limites devem ser recalculados a

cada amostra.

O gráfico de controle p é equivalente ao gráfico np quando se trata de amostras de tamanho constante. Caso as amostras possuam tamanho variável, é aconselhável o uso do gráfico de controle p , uma vez que ele trabalha com proporções das amostras.

5.2.3 Gráfico de Controle c

Este gráfico monitora o número médio de defeitos por unidade, uma vez que é necessário analisar a quantidade de especificações não atendida em unidades produzidas. Para montar este gráfico, será necessário:

1. Usar uma folha de verificação similar a apresentada na [Tabela 2.2](#). Com base nesta folha, têm-se o número de defeitos por unidade;
2. Em seguida deve-se calcular o número médio de defeitos nas m unidades, usando a fórmula:

$$\bar{c} = \frac{\sum c_i}{m}$$

3. Calcular o limite superior de controle (LSC) e o limite inferior de controle (LIC). Se o limite inferior for negativo, iguale-o a zero:

$$LCS = \bar{c} + 3\sqrt{\bar{c}}$$

$$LCI = \bar{c} - 3\sqrt{\bar{c}}$$

4. Trace o sistema de eixos cartesianos (no eixo das abscissas devem estar o número das unidades e no eixo das ordenadas, o número de não-conformidades);

5. Trace no gráfico uma linha central, com base na proporção média de itens não-conformes;
6. Trace o LSC e o LIC;
7. Faça um ponto para cada par de valores (defeitos e número da unidade);
8. Una os pontos por segmentos de reta.

Com base nestes passos, façamos um exemplo com base nos dados da Tabela 5.2.

Tabela 5.2 – Defeitos em dez unidades produzidas

Amostras										
	1	2	3	4	5	6	7	8	9	10
c	10	3	8	13	11	7	1	2	6	10

Para estes valores temos os cálculos:

$$\bar{c} = \frac{10+3+8+13+11+7+1+2+6+10}{10} = 7,1$$

$$LCS = 7,1 + 3\sqrt{7,1} \approx 15,1$$

$$LCI = 7,1 - 3\sqrt{7,1} \approx -0,1$$

Como o limite inferior é negativo, será tomado como zero. O gráfico de controle é exibido na Figura 5.8.

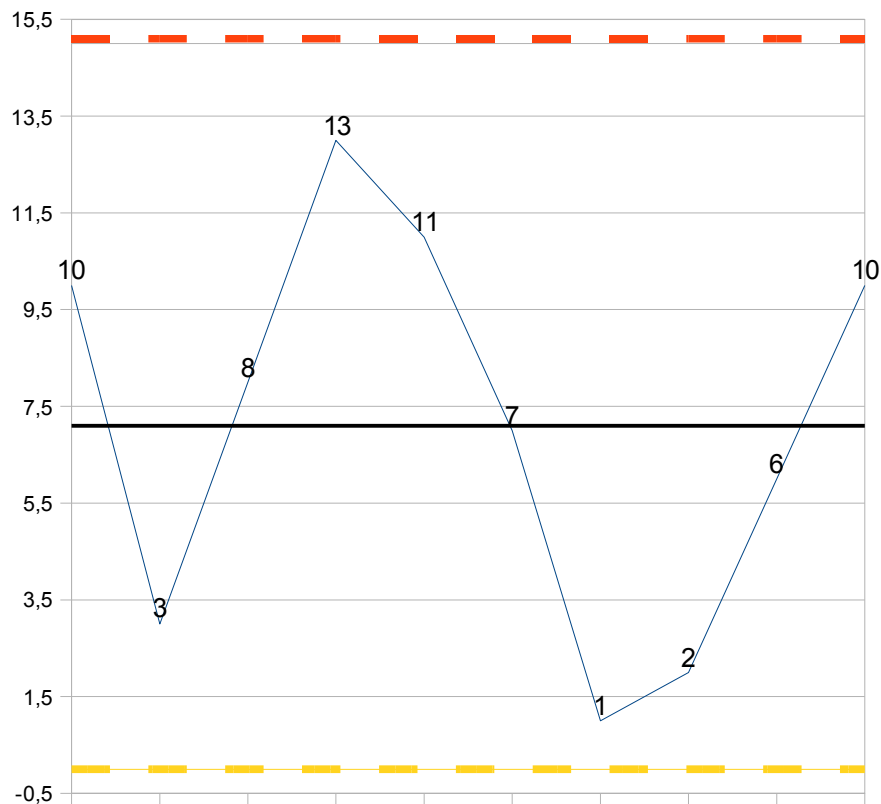


Figura 5.8 – Gráfico de controle c

Além das verificações apresentadas na seção 5.1, neste gráfico deve-se atentar para todos os pontos extremos. Os extremos acima da média indicam um número alto de defeitos e é de senso comum que estes casos devam ser analisados em busca de causas. Já os extremos abaixo da média podem indicar inspeções falhas, em que defeitos possam passar despercebidos.

Em alguns produtos, há defeitos mais relevantes que outros. Por exemplo: em um eletrônico, um defeito em uma placa de circuito impresso que impeça o funcionamento do produto, é um defeito mais relevante que um risco na sua pintura. Nestes casos, é indicado o uso de pesos para defeitos. Esta ponderação é chamada sistema de deméritos.

5.2.4 Gráfico de Controle $\bar{x} - R$

Este gráfico de controle mostra a variação de variáveis como peso, comprimento, e outras. É um gráfico recomendável para amostras pequenas e de mesmo tamanho. Para desenhar este gráfico é necessário usar a folha de verificação apresentada na [Tabela 2.5](#). Uma vez coletados os dados, preenchida a folha, e calculadas as médias e amplitudes das amostras, têm-se os passos:

1. Calcular a média das médias das amostras:

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots + \bar{x}_n}{m}$$

2. Calcular a média das amplitudes:

$$\bar{R} = \frac{R_1 + R_2 + R_3 + \dots + R_n}{m}$$

3. Calcular os limites superior e inferior de controle para a média das médias das amostras:

$$LSC = \bar{\bar{x}} + A_2 \bar{R}$$

$$LIC = \bar{\bar{x}} - A_2 \bar{R}$$

onde A_2 depende do tamanho da amostra, e pode ser obtido na Tabela 1 do Anexo.

4. Calcular os limites superior e inferior de controle para a média das amplitudes:

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

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

onde D_3 e D_4 dependem do tamanho da amostra, e podem ser obtidos na Tabela 1 do Anexo.

5. Os passos para desenhar os gráficos de controle para a média das médias e para a média das amplitudes das amostras, são semelhantes aos passos do gráfico np , a partir do item 5.

Para exemplificar a construção destes gráficos, tomemos como base os dados da Tabela 5.2, que contém medidas de peso de um produto.

Tabela 5.3 – Tabela com exemplos de monitoramento de fabricação

	Amostra							
	1	2	3	4	5	6	7	8
x_1	15	18	13	15	12	15	16	12
x_2	18	14	15	15	14	16	15	19
x_3	12	16	16	17	15	17	15	20
x_4	17	15	20	13	14	15	14	13
Média	15,5	15,75	16	15	13,75	15,75	15	16
Amplitude	6	4	7	4	3	2	2	8

Média das Médias:	15,34	Média das amplitudes	4,50
-------------------	-------	----------------------	------

$$LSC_{\bar{x}} = \bar{\bar{x}} + A_2 \bar{R} = 15,34 + 0,729 * 4,5 = 18,62$$

$$LIC_{\bar{x}} = \bar{\bar{x}} - A_2 \bar{R} = 15,34 - 0,729 * 4,5 = 12,06$$

$$LSC_R = D_4 \bar{R} = 2,282 * 4,5 = 10,27$$

$$LIC_R = D_3 \bar{R} = 0 * 4,5 = 0$$

A partir destes dados temos os gráficos das Figuras 5.9 e 5.10.

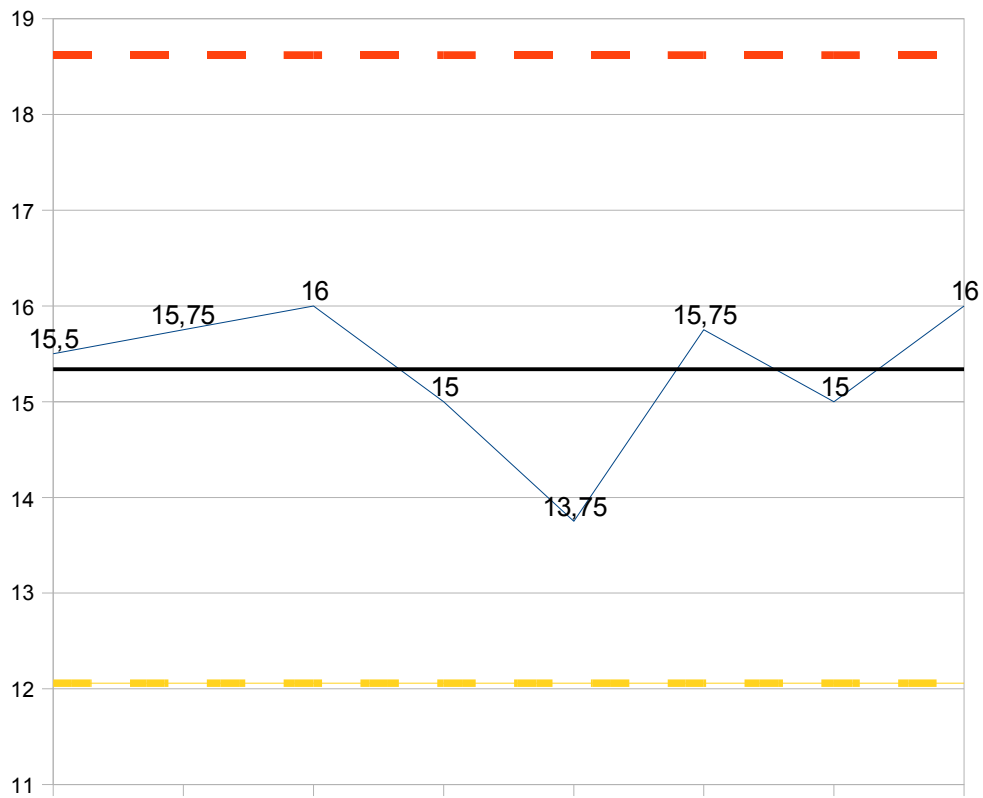


Figura 5.9 – Gráfico de controle para médias das médias das amostras

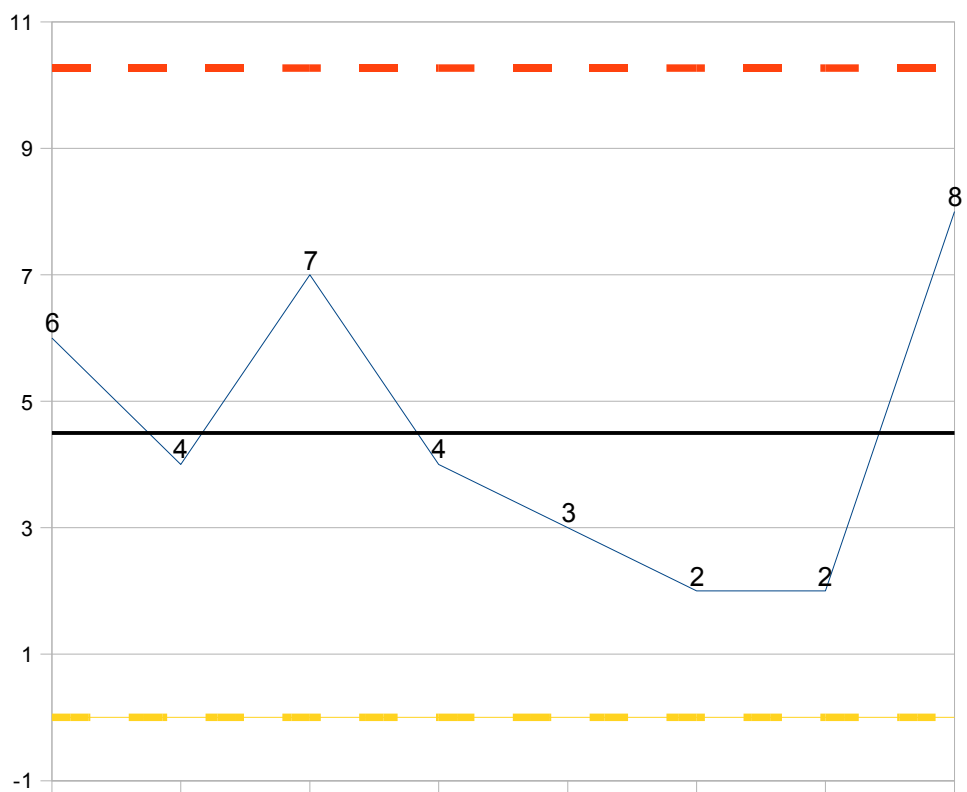


Figura 5.10 – Gráfico de controle para médias das amplitudes

No gráfico de controle para as médias, podemos notar uma certa periodicidade nas amostras. Entretanto a variabilidade é pequena (havendo apenas um ponto mais afastado da média 13,75) e há uma grande folga em relação aos limites. Porém, é preciso verificar se a média das médias das amostras se aproxima das especificações de qualidade do produto. Caso este quesito não seja atendido, a interpretação do gráfico fica comprometida.

O gráfico de controle para médias exhibe apenas o comportamento das amostras como um todo, e não a variação dentro das amostras. Já o gráfico de controle para amplitudes, exhibe a variação de valores dentro de todas as amostras. É desejável que a média e os pontos estejam o mais próximo possível de zero. Neste caso, uma tendência para zero é bem vinda. Porém, o gráfico do exemplo esboça uma periodicidade (embora haja poucas amostras para confirmar esta suspeita), e há picos de amplitude, ou seja, dentro de algumas amostras há valores muito discrepantes, cujas causas devem ser investigadas e corrigidas. Um ponto positivo é que a maioria dos valores está abaixo da média, que deve baixar bastante se forem corrigidos os problemas que causam picos de amplitude.

Tabela 5.4 - Uma comparação dos resultados de ambos os gráficos.

		Gráfico das médias	
		Sob Controle	Fora de Controle
Gráfico das Amplitudes	Sob Controle	Bom desempenho do processo	<ul style="list-style-type: none"> - Desempenho do processo afetado ao ponto de uma ou mais amostras apresentarem problemas, embora a variabilidade seja aceitável dentro de todas as amostras. - Possíveis soluções : substituição ou aprimoramento de máquinas, operadores ou material;
	Fora de Controle	<ul style="list-style-type: none"> - Embora as amostras não estejam fora dos limites, há problemas dentro delas, devendo ser investigadas as causas de variabilidade interna; - Possíveis soluções: ajuste de máquinas, e metodologia, e treinamento de operadores; 	<ul style="list-style-type: none"> - Todo o processo deve ser reestruturado, começando das causas especiais de variação dentro das amostras; - A diminuição da variação pode corrigir as médias das amostras;

5.2.5 Gráfico de Controle $\bar{x} - S$

Este gráfico de controle deve ser usado quando o tamanho das amostras é relativamente grande e/ou quando as amostras têm tamanho variável.

Para desenhar este gráfico é necessário usar uma folha de verificação similar à apresentada na [Tabela 2.5](#). Uma vez coletados os dados, preenchida a folha, e calculadas as médias e os desvios padrões de cada uma das amostras, têm-se os passos:

1. Calcular a média das médias das amostras:

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots + \bar{x}_n}{m}$$

2. Calcular a média dos desvios padrões das amostras:

$$\bar{s} = \frac{s_1 + s_2 + s_3 + \dots + s_n}{m}$$

3. Calcular os limites superior e inferior de controle para a média das médias das amostras:

$$LSC = \bar{\bar{x}} + A_3 \bar{s}$$

$$LIC = \bar{\bar{x}} - A_3 \bar{s}$$

onde A_3 depende do tamanho da amostra, e pode ser obtido na Tabela 1 do Anexo.

4. Calcular os limites superior e inferior de controle para a média dos desvios padrões:

$$LSC = B_4 \bar{s}$$

$$LIC = B_3 \bar{s}$$

onde B_3 e B_4 dependem do tamanho da amostra, e podem ser obtidos na Tabela 1 do Anexo.

5. Os passos para desenhar os gráficos de controle para a média das médias e para a média das amplitudes das amostras, são semelhantes aos passos do gráfico np , a partir do item 5.

Para exemplificar a construção destes gráficos, tomemos como base os dados da Tabela 5.5 (contendo os mesmos dados da Tabela 5.4, porém com o cálculo do desvio padrão).

Tabela 5.5 – Tabela com exemplos de monitoramento de fabricação

	Amostra							
	1	2	3	4	5	6	7	8
x_1	15	18	13	15	12	15	16	12
x_2	18	14	15	15	14	16	15	19
x_3	12		16	17	15	17	15	20
x_4	17		20	13		15	14	13
Média	15,5	16	16	15	13,67	15,75	15	16
Desvio Padrão	2,65	2,82	2,94	1,63	1,53	0,96	0,82	4,08

Média das Médias:	15,36	Média dos desvios padrões	2,18
-------------------	-------	---------------------------	------

$$LSC_{\bar{x}} = \bar{\bar{x}} + A_3 \bar{s} = 15,36 + 1,628 * 2,18 = 18,90$$

$$LIC_{\bar{x}} = \bar{\bar{x}} - A_3 \bar{s} = 15,36 - 1,628 * 2,18 = 11,81$$

$$LSC_{\bar{s}} = B_4 \bar{s} = 2,266 * 2,18 = 4,94$$

$$LIC_{\bar{s}} = B_3 \bar{s} = 0 * 2,18 = 0$$

A partir destes dados temos os gráficos das Figuras 5.11 (que é idêntico ao gráfico 5.9, apresentado anteriormente) e 5.12.

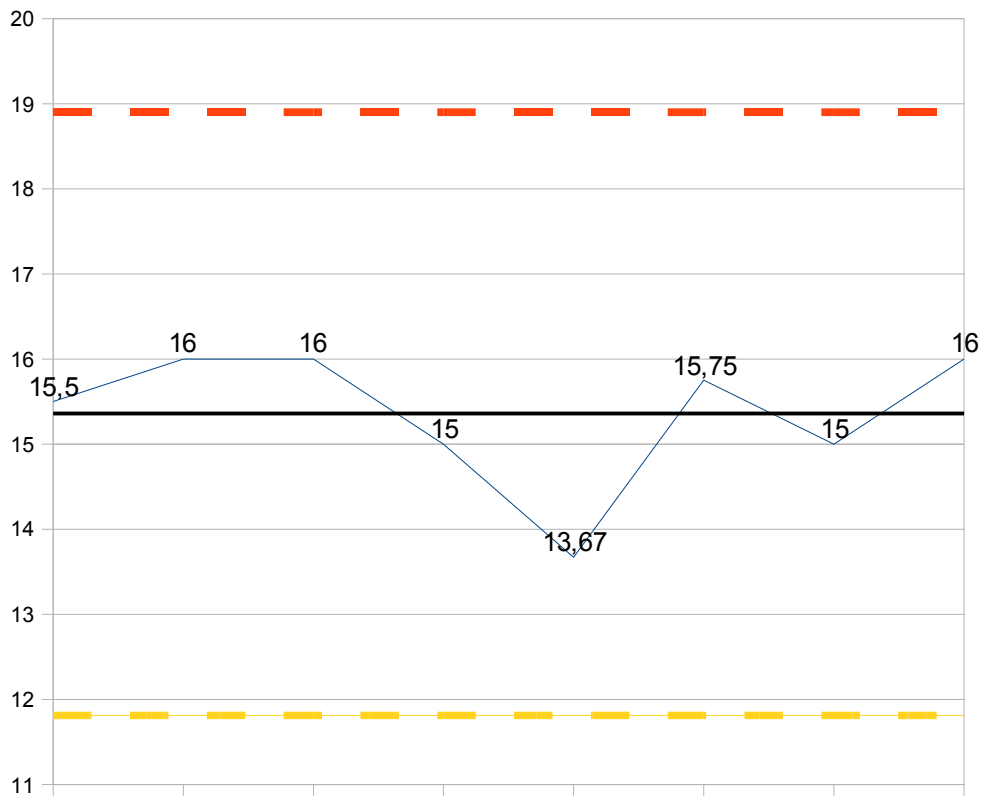


Figura 5.11 – Gráfico de Controle para média das médias das amostras

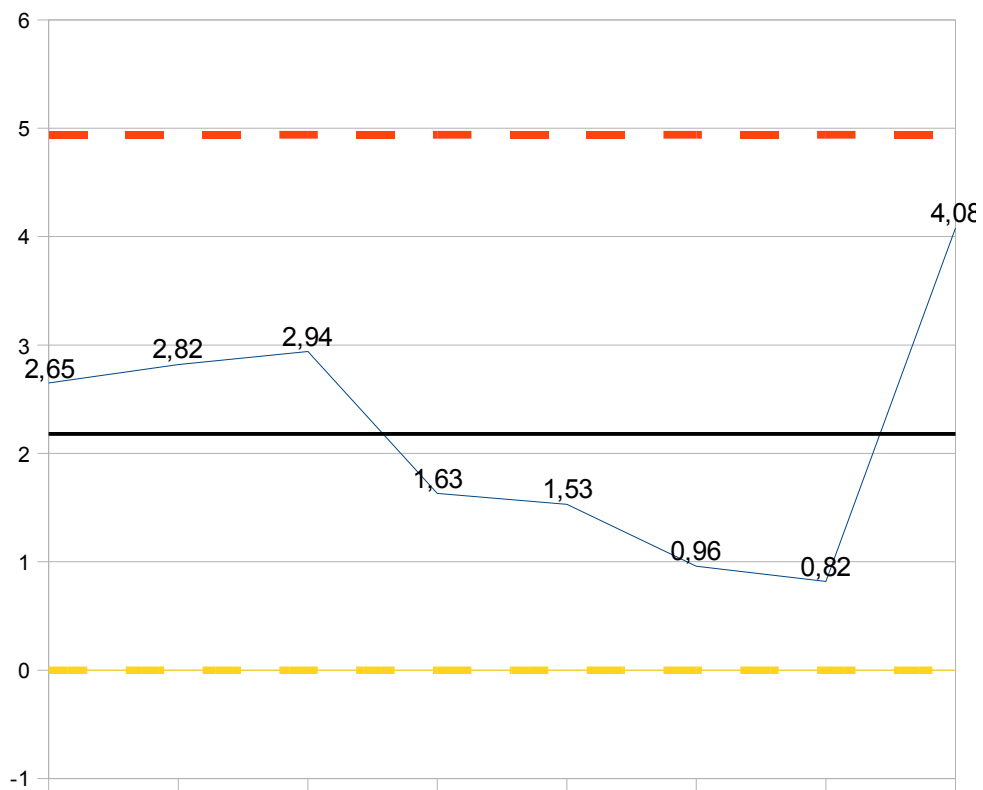


Figura 5.12 – Gráfico de Controle para média dos desvios padrões das amostras

O primeiro gráfico de controle mostra certa periodicidade das médias das amostras, sugerindo alguma instabilidade periódica do processo produtivo. Já o gráfico de controle para média dos desvios padrões das amostras mostra o princípio de um deslocamento e uma tendência ao limite superior (suspeitas que só podem ser comprovadas com amostras maiores. A análise conjunta destes é análoga à feita na Tabela 5.4.

5.2.6 Gráfico de Controle u

O gráfico de controle u monitora o número médio de defeitos por unidade. Por exemplo, inspeção em rolos de tecido de tamanhos variáveis em uma indústria têxtil.

Para montar este gráfico, será necessário:

1. Usar uma folha de verificação para registrar o tamanho de cada amostra e o número de defeitos por amostra;
2. Estabelecer a unidade e calcular o número de unidades (n_i) em cada amostra;
3. Calcular o número médio de defeitos (u_i) por unidade em cada amostra;
4. Dividir o número total de defeitos pelo número total de unidades, obtendo o número médio de defeitos por unidade (\bar{u});
5. Calcular os limites superiores e inferiores de controle para cada amostra:

$$LCS = \bar{u} + 3\sqrt{\frac{\bar{u}}{n_i}}$$

$$LIC = \bar{u} - 3\sqrt{\frac{\bar{u}}{n_i}}$$

6. Trace o sistema de eixos cartesianos (no eixo das abscissas devem estar o número das amostras e no eixo das ordenadas, o número de defeitos por unidade);

5. Trace no gráfico uma linha central, com base na proporção média de itens não-conformes;
6. Trace o LSC e o LIC;
7. Faça um ponto para cada par de valores (amostra e número de defeitos por unidade);
8. Una os pontos por segmentos de reta.

Com base nestes passos, façamos um exemplo a partir dos dados da Tabela 5.6.

Tabela 5.6 – Número de defeitos por rolo de tecido

	Amostras									
	1	2	3	4	5	6	7	8	9	10
Defeitos por rolo	4	15	7	18	27	1	13	2	1	10
Tamanho do rolo (m ²)	200	300	200	450	450	100	500	500	400	400

Pode-se estabelecer a unidade em 100m² de tecido. Desta forma, o número de unidades por rolo (n_i) é calculada pela divisão do tamanho de cada rolo por 100m². A tabela 5.7 mostra estes valores e o número médio de defeitos por rolo (obtido pela divisão do número de defeitos por rolo pelo número de unidades do rolo).

Tabela 5.7 – Número médio de defeitos por unidade

	Amostras									
	1	2	3	4	5	6	7	8	9	10
Unidades do rolo (n_i)	2	3	2	4,5	4,5	1	5	5	4	4
Número médio de defeitos em cada unidade (u_i)	2	5	3,5	4	6	1	2,6	0,4	0,25	2,5

O número médio de defeitos por unidade (levando em conta os defeitos e unidades de todas amostras) é dado pela equação:

$$\bar{u} = \frac{4+15+7+18+27+1+13+2+1+10}{2+3+2+4,5+4,5+1+5+5+4+4} = \frac{98}{35} = 2,8$$

A tabela 5.8 exibe o número de unidades em cada amostra e os limites de controle calculados para cada rolo.

Tabela 5.8 – Cálculo dos limites de controle

Amostras	n_i	LSC	LIC
1	2	6,35	-0,75
2	3	5,7	-0,1
3	2	6,35	-0,75
4	4,5	5,17	0,43
5	4,5	5,17	0,59
6	1	7,82	-2,22
7	5	5,04	0,56
8	5	5,04	0,56
9	4	5,31	0,29
10	4	5,31	0,29

O gráfico de controle é exibido na figura 5.13.

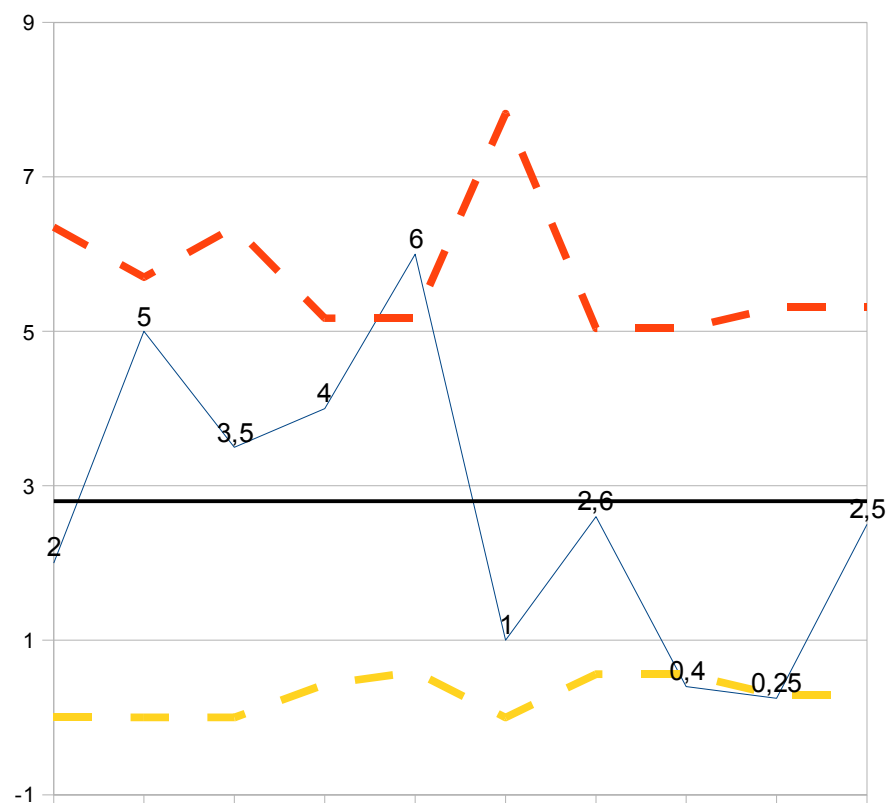


Figura 5.13 – Gráfico de controle u

Os limites deste gráfico variam de acordo com o tamanho de cada amostra. Porém este gráfico pode ser simplificado com um só limite superior e um só limite inferior de controle. Para tanto, deve-se calcular o tamanho médio das amostras:

$$\bar{n} = \sum \frac{n_i}{m} = \frac{2+3+2+4,5+4,5+1+5+5+4+4}{10} = 3,5$$

Logo, os limites recalculados são:

$$LCS = 2,8 + 3 \sqrt{\frac{2,8}{3,5}} \approx 5,48$$

$$LCI = 2,8 - 3 \sqrt{\frac{2,8}{3,5}} \approx 0,11$$

Uma vez calculados os limites a partir do tamanho médio das amostras, o gráfico de controle é exibido na figura 5.14.

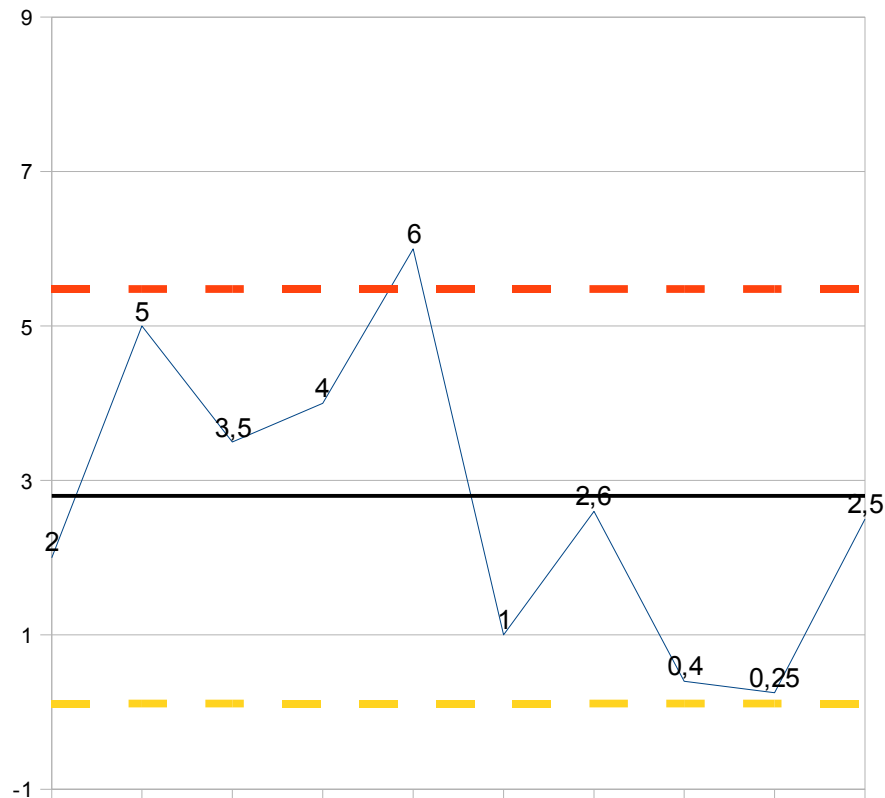


Figura 5.14 – Gráfico de controle u com limites fixos

5.3 Considerações finais

Uma prática usual nas indústrias é utilizar amostras pequenas e freqüentes para a elaboração de gráficos de controle. Esta estratégia se justifica pelo menor custo da amostra e, principalmente, a rápida detecção de produtos não-conformes. Caso o tempo decorrido entre duas amostras seja grande, é possível que muitos itens não-conforme sejam produzidos até a detecção do problema.

Não há um tamanho ideal que sirva para qualquer taxa de produção e para qualquer tipo de controle que se queira fazer. Porém, é comum a coleta de 20 a 25 amostras de tamanho 4 ou 5, para se obter gráficos de controle mais precisos e a custo moderado.

6 Diagrama de Dispersão

Muitas variáveis possuem relações entre si. Por exemplo: as falhas em uma linha de montagem podem estar relacionadas a falha em outra linha específica (seja pela atuação de um mesmo operador, tipo de manutenção ou de máquinas utilizadas); defeitos em peças podem estar relacionados ao tempo de montagem das mesmas; o número de defeitos pode estar relacionado à distância de entrega.

A detecção de relações entre duas variáveis pode ser auxiliada pela ferramenta diagrama de dispersão.

6.1 Construindo um diagrama de dispersão

Para construir um diagrama de dispersão é necessário a coleta de dados de duas variáveis. A tabela a seguir traz dados coletados das variáveis distância de entrega e número de defeitos de 10 amostras.

Tabela 6.1 – Distância de entrega e número de defeitos em amostras 10 amostras de produtos

Distância (km)	Número de defeitos
750	7
300	5
100	1
150	3
400	6
20	0
1000	10
500	6
350	4
1300	11

Estes dados devem ser representados nos eixos cartesianos de um gráfico. Para comprovar a influência da distância de entrega no número de defeitos das amostras, deve-se tomar a primeira como uma variável independente X (representada no eixo das abcissas) e a segunda como uma variável dependente Y (representada no eixo das ordenadas). Deve-se fazer um ponto para cada par de valores x e y .

A figura 6.1 apresenta o gráfico de dispersão dos referidos dados.

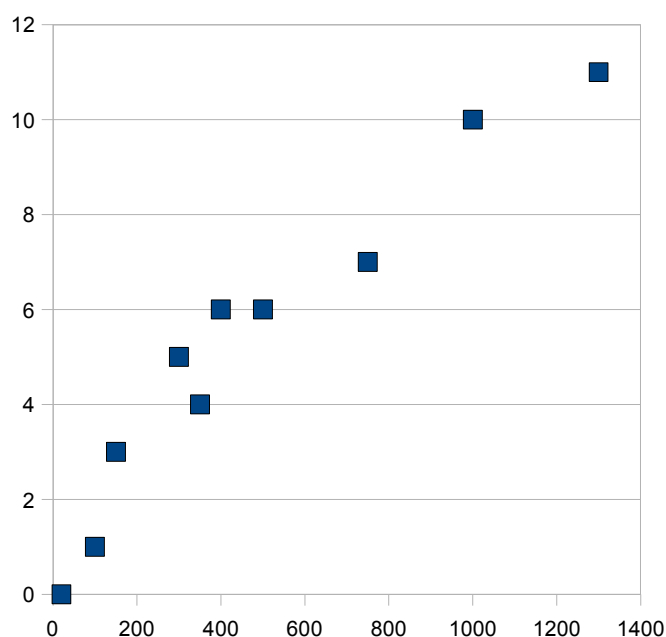


Figura 6.1 – Gráfico de dispersão de defeitos por distância de entrega de produtos

Uma vez desenhado o diagrama de dispersão deve-se atentar para a direção e dispersão dos pontos. Se X e Y crescem no mesmo sentido, conforme pode ser visto na Figura 6.1, têm-se uma correlação positiva. Esta correlação aumenta à medida que a dispersão dos pontos diminui. No exemplo acima, a correlação é positiva e forte, pois os pontos estão pouco dispersos.

Já na Figura 6.2, temos uma correlação positiva fraca, pois os dados, embora cresçam no mesmo sentido, estão bastante dispersos. A Figura 6.3, por sua vez, exibe uma correlação negativa forte, uma vez que os dados variam em sentidos contrários e há pouca dispersão. Uma correlação

positiva perfeita é exibida na Figura 6.4, onde os dados crescem linearmente. Esta seria a correlação ideal, pois permite previsões precisas, porém não é muito comum encontrá-la em dados reais.

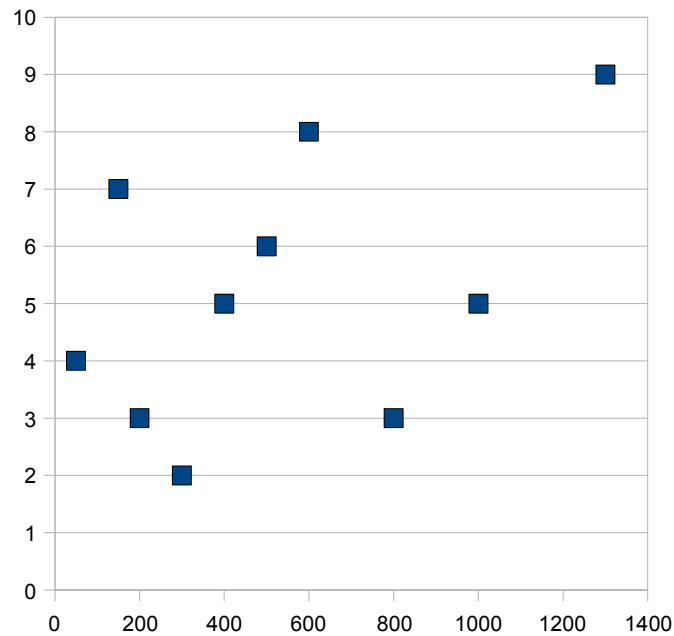


Figura 6.2 – Exemplo de gráfico de dispersão indicando correlação positiva fraca

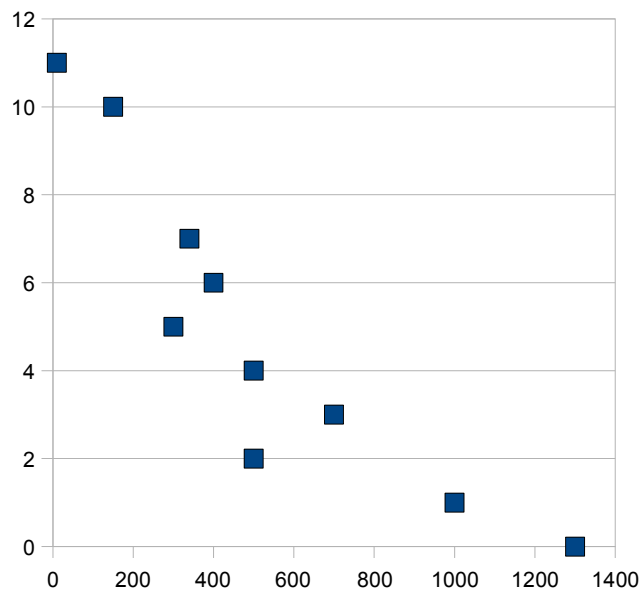


Figura 6.3 – Exemplo de gráfico de dispersão indicando correlação negativa forte.

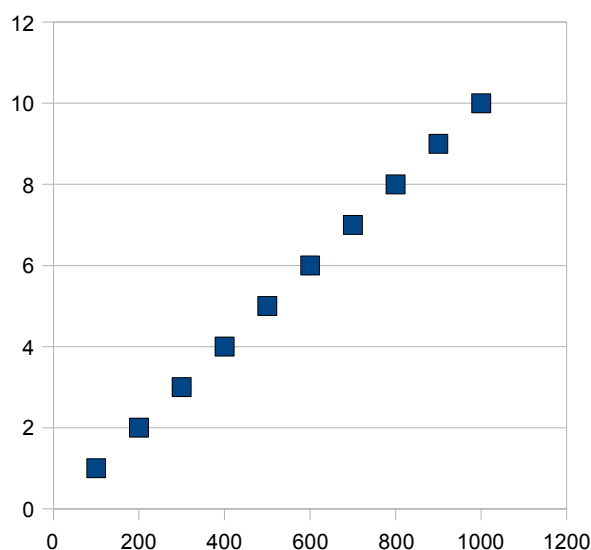


Figura 6.4 – Exemplo de gráfico de dispersão indicando correlação positiva perfeita.

A Tabela 6.2 resume várias possibilidades apresentadas em um diagrama de dispersão.

Tabela 6.2 – Análises de um diagrama de dispersão

	Correlação Positiva	Correlação Negativa
Fraca	Dados crescem no mesmo sentido, porém dispersos	Dados variam em sentidos diferentes, dispersos
Forte	Dados crescem no mesmo sentido, pouco dispersos	Dados variam em sentidos diferentes, pouco dispersos
Perfeita	Dados crescem no mesmo sentido linearmente	Dados variam em sentidos diferentes, porém linearmente

As interpretações das combinações apresentadas na Tabela 6.2 são resumidas na tabela 6.3.

Tabela 6.3 – Interpretações de um diagrama de dispersão

	Correlação Positiva	Correlação Negativa
Fraca	A variável X tem relação pouco relevante com Y; são diretamente proporcionais	A variável X tem pouca relação com Y; são inversamente proporcionais
Forte	A variável X tem relação relevante com Y; são diretamente proporcionais	A variável X tem relação relevante com Y; são inversamente proporcionais
Perfeita	A variável X tem relação perfeita com Y; são diretamente proporcionais	A variável X tem relação perfeita com Y; são inversamente proporcionais

Com base no diagrama de dispersão apresentado na Figura 6.1 e nas interpretações da Tabela 6.3, pode-se afirmar que os dados de distância de entrega e defeito dos produtos, apresentados na Tabela 6.1, têm forte relação e são diretamente proporcionais. Ou seja, há uma tendência de que haja mais produtos defeituosos à medida que a distância percorrida para entrega aumenta. Este fato pode ser corrigido com uma reestruturação da cadeia de distribuição dos produtos.

6.2 Calculando a correlação entre variáveis

O gráfico de dispersão permite identificar, visualmente, o quanto e como duas variáveis estão relacionadas entre si. Há também um coeficiente matemático que permite a avaliação do grau de correlação entre duas variáveis. É o coeficiente de correlação, que se representa por r , que pode ser obtido pela fórmula:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}} \sqrt{\sum y^2 - \frac{(\sum y)^2}{n}}}$$

A partir dos dados da tabela 6.1, pode-se calcular o coeficiente de correlação.

$$r = \frac{38400 - \frac{4870 * 53}{10}}{\sqrt{3907900 - \frac{23716900}{10}} \sqrt{393 - \frac{2809}{10}}} = 0,96$$

O valor de r varia entre -1 e +1. Quanto mais próximo dos extremos, maior a correlação entre as variáveis (e quanto mais próximo de zero, menor a correlação). O coeficiente positivo indica que as variáveis crescem no mesmo sentido, e o coeficiente negativo indica que variam em

sentidos diferentes. No exemplo acima, há uma forte correlação positiva entre distância de entrega e número de defeitos, conforme o gráfico de dispersão indica.

6.2.1 Cuidados na interpretação da correlação

Apesar de ser evidente a correlação entre as duas variáveis da Tabela 6.1, nem sempre esta relação pode ocorrer de forma direta. Podem haver outras variáveis não estudadas que interfiram em Y. Um exemplo bastante difundido desta situação é a correlação entre o número de cegonhas e a taxa de natalidade em algumas cidades escandinavas. Seria uma interpretação ingênua relacionar o número de ninhos de cegonhas com o número de nascimentos, embora o coeficiente de correlação indicasse esta situação. Na verdade há uma terceira variável neste caso: o aumento da população tem uma relação de causa e efeito com o número de casas, e na Escandinávia as casas têm chaminés, perto de onde as cegonhas fazem ninhos. Portanto, a relação entre a taxa de natalidade e o número de cegonhas é indireta.

Outro risco na interpretação da correlação é a extrapolação. Não se pode estimar o que acontecerá em um período com base no que foi observado em outro período. Por exemplo, se a correlação entre índice de desenvolvimento humano e a taxa de natalidade de um país de terceiro mundo, obtida com base em dados da década de 80, for usada para prever a taxa de natalidade na quinta década do século atual, será obtido um número absurdamente irreal.

Correlações não confiáveis também podem ser obtidas a partir de um número relativamente pequeno de observações ou do cálculo do coeficiente baseado em observações estratificadas analisadas em conjunto.

6.3 Análise de Regressão

O cálculo do coeficiente de correlação dos dados da Tabela 6.1, referente à distância de entrega e número de defeitos em um produto, evidencia uma forte relação entre as variáveis. Além da correlação, a análise destes dados pode ser útil em previsões de períodos próximos aos observados, tomando-se cuidado em evitar possíveis extrapolações.

A análise de regressão permite a identificação de uma relação linear entre as variáveis. Esta relação linear fica evidente ao se observar a Figura 6.1. Os pontos descrevem, aproximadamente, uma reta, a reta de regressão.

A equação de uma reta é:

$$Y = \alpha + \beta X$$

onde:

- Y é a variável dependente;
- X é a variável independente, também chamada de variável explanatória ou de controle;
- α é o coeficiente linear, ou seja, a altura em que a reta toca o eixo das ordenadas;
- β é o coeficiente angular da reta, isto é, o ângulo de inclinação da reta em relação ao eixo das abcissas.

O coeficiente angular β pode ser obtido pela equação:

$$\beta = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

O coeficiente linear α pode ser obtido pela fórmula:

$$\alpha = \bar{y} - \beta \bar{x}$$

onde \bar{y} e \bar{x} , são, respectivamente, a média de Y e X.

Com base nos dados da Tabela 6.1, pode-se calcular os coeficientes.

$$\beta = \frac{38400 - \frac{4870 \cdot 53}{10}}{3907900 - \frac{23716900}{10}} \approx 0,0082$$

$$\alpha = 5,3 - 0,0082 \cdot 487 = 1,31$$

Logo a equação da reta de regressão é:

$$Y = 1,31 + 0,0082 X$$

Esta equação seria útil, por exemplo, para estimar o número aproximado de defeitos para uma dada distância. A Tabela 6.4 exibe uma série de estimativas aproximadas.

Tabela 6.4 – Estimativas de números de defeitos com base na distância (usando análise de regressão)

Distância (km)	Número de defeitos
0	1,31
250	3,36
550	5,82
800	7,87
950	9,1
1050	9,92
1250	11,56
1400	12,79
1600	14,43
2000	17,71

Estas estimativas devem ser encaradas apenas como projeções aproximadas, uma vez que têm-se uma quantidade pequena de dados para estimar os coeficientes com maior precisão, e a correlação não é perfeita neste caso. Portanto, têm-se números de defeitos aproximados mesmo para as distâncias apresentadas na Tabela 6.1, conforme exibido na Tabela 6.5.

Tabela 6.5 – Estimativas do números de defeitos com base nas distância da Tabela 6.1 (usando análise de regressão)

Distância (km)	Número de defeitos
750	7,46
300	3,77
100	2,13
150	2,54
400	4,59
20	1,47
1000	9,51
500	5,41
350	4,18
1300	11,97

Apesar de não fornecer um número exato, a regressão pode permitir que seja feita uma estimativa bastante confiável, desde que seja baseada em observações adequadas, ou seja, das variáveis corretas, em número adequado, sem extrapolações, estratificações e valores discrepantes.

Anexo

Tabela para construção de gráficos de controle

n	A ₂	A ₃	D ₃	D ₄	d ₂	B ₃	B ₄
2	1,880	2,659	0	3,267	1,128	0	3,267
3	1,023	1,957	0	2,575	1,693	0	3,568
4	0,729	1,628	0	2,282	2,059	0	2,266
5	0,577	1,427	0	2,115	2,326	0	2,089
6	0,483	1,287	0	2,004	2,534	0,030	1,970
7	0,419	1,182	0,076	1,924	2,704	0,118	1,882
8	0,373	1,099	0,136	1,816	2,847	0,185	1,815
9	0,337	1,032	0,184	1,816	2,970	0,239	1,761
10	0,308	0,975	0,223	1,777	3,078	0,284	1,716
11	0,285	0,927	0,256	1,744	3,173	0,321	1,679
12	0,266	0,886	0,283	1,717	3,258	0,354	1,646
13	0,249	0,850	0,307	1,693	3,336	0,382	1,646
14	0,235	0,817	0,328	1,672	3,407	0,406	1,594
15	0,223	0,789	0,347	1,653	0,347	1,653	3,472
16	0,212	0,763	0,363	1,637	3,532	0,448	1,552
17	0,203	0,739	0,378	1,637	3,532	0,448	1,552
18	0,194	0,718	0,391	1,608	3,640	0,482	1,518
19	0,187	0,698	0,403	1,597	3,689	0,497	1,503
20	0,180	0,680	0,415	1,585	3,735	0,510	1,490
21	0,173	0,663	0,425	1,575	3,778	0,523	1,477
22	0,167	0,647	0,434	1,566	3,819	0,534	1,466
23	0,162	0,633	0,443	1,557	3,858	0,545	1,455
24	0,157	0,619	0,451	1,548	3,895	0,555	1,445
25	0,153	0,606	0,459	1,541	3,391	0,565	1,435

Bibliografia

KUME, Hitoshi. **Métodos Estatísticos para Melhoria da Qualidade**. Rio de Janeiro: Editora Gente, 1993, 11ª Edição

VIEIRA, Sônia. **Estatística para a Qualidade**. Rio de Janeiro: Editora Campus, 1999, 1ª Edição.