Controle estatístico de processo

Prof. Ramon Gomes da Silva

Controle estatístico de processo

1. Conceitos básicos;
2. Medidas descritivas e gráficos básicos;
3. Gráficos - caixa de medianas e histograma.

1. Conceitos básicos

➔ A ideia principal do Controle Estatístico de Processo (CEP) é que melhores processos de produção com menos variabilidade propiciam níveis melhores de qualidade nos resultados da produção;

➔ Quando se fala em melhores processos isso significa não somente qualidade melhor, mas também custos menores. Os custos diminuem principalmente em função de duas razões: a amostragem e a redução de rejeito;

➔ Uma segunda razão pela qual a aplicação de CEP impulsiona os custos para baixo é que o número e a porcentagem de peças defeituosas produzidas na fábrica vão diminuir com as melhorias na linha de produção.

1. Conceitos básicos

- ➔ Portanto, com menos refugo e menos retrabalho, o custo por peça produzida diminui;
- ➔ Enfatiza-se que existe somente uma razão para utilizar CEP na fábrica: aumentar o resultado financeiro da empresa, se possível no curto prazo, e também, talvez mais importante, no longo prazo;
- ➔ No entanto, CEP não é nenhum milagre e consequentemente ele deve ser abordado na empresa como qualquer projeto de investimento em que os custos são contabilizados e os benefícios previstos e medidos.

2. Medidas descritivas e gráficos básicos

1. Média;
2. Mediana;
3. Medidas de variabilidade - desvio-padrão;

2.1. Média

- Em qualquer área de investigação onde números aparecem com frequência, os profissionais da área estudam maneiras e metodologias gráficas e estatísticas para expressar esses números mais clara e resumidamente;
- Existem várias maneiras de medir a tendência central dos dados, e nenhuma maneira é necessariamente a melhor, tudo depende da situação;
- O cálculo de uma tendência central é importante porque ela consegue condensar uma série de dados em um único número;
- Certamente, a mais popular é a **média**, a soma de uma série de dados dividida pelo número de dados na soma.

2.1. Média

$$\text{Média} = \frac{\sum_{i=1}^{n} \text{valor}_i}{n}$$

- ➔ Uma tabela de números não é nada interessante para o gestor;
- ➔ Por outro lado, com a média das medidas o gestor pode saber se o produto está sendo fabricado dentro do desejado;
- ➔ Um problema que pode ocorrer às vezes é que a média perde a sua representatividade, quando, entre os números, existem valores muito diferentes dos outros;
- ➔ Esses valores levam a média para um valor muito longe da tendência central dos dados, e não muito perto dos outros números;
- ➔ Uma maneira de resolver o problema dessa distorção seria simplesmente eliminar esses números; no entanto, o gerente não recomenda esse caminho por causa de certo grau de arbitrariedade.

2.2. Mediana

➔ Para resolver a distorção de números discrepantes e assimétricos, utiliza-se a **mediana**, o número no meio dos números (ou a média dos dois números no meio);

➔ Numa relação de números ordenados do maior para o menor, existe um número que separa todos os números em dois grupos, os números maiores que a mediana e os números menores;

➔ Tendo um conjunto com oito elementos (1, 2, 3, **4**, **5**, 6, 8, 9), a mediana é a média:

$$\frac{4+5}{2} = 4,5$$

➔ Na lista dos oito elementos, há quatro números menor que 4,5 e quatro maiores.

2.2. Mediana

- Quando o número de dados é ímpar, a mediana é exatamente o número no meio dos números ordenados, sem haver necessidade de calcular a média dos dois números no meio;

- Na amostra de sete elementos (1, 3, 3, **6**, 7, 8, 9), a mediana é 6;

- A fórmula usada para encontrar a posição de um valor do meio em uma amostra de **n** elementos organizados em ordem crescente é:

$$\frac{n+1}{2}$$

- Para uma população de 5 elementos, a posição do valor médio é: $\frac{n+1}{2} = \frac{5+1}{2} = 3$

- Para uma população de 6 elementos, a posição do valor médio é 3,5. Logo a mediana usará os elemento das posições 3 e 4.

2.2. Mediana

- ➔ Os analistas argumentam que a mediana é melhor do que a média para representar a tendência central dos números na presença de dados muito diferentes dos outros. Isso ocorre porque a mediana é insensível aos valores muito grandes ou muito pequenos;
- ➔ Partindo da mediana, os quartis são calculados;
- ➔ Com a mediana, os dados foram divididos em dois subgrupos, acima e abaixo da mediana. Para cada subgrupo há sua própria mediana e essa mediana se chama de quartil. Obviamente tem um quartil inferior, o primeiro quartil, e um quartil superior, o terceiro quartil;
- ➔ Para completar o raciocínio, pode-se chamar a mediana de segundo quartil. Os quartis dividem os dados em quatro grupos distintos, cada grupo tem exatamente um quarto dos dados;

2.2. Mediana

➔ Os quartis podem ser utilizados também para definir a variabilidade dos dados.

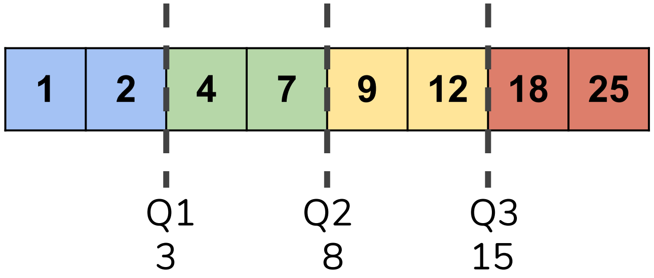
➔ **Exemplo:**

1	**2**	**4**	**7**	**9**	**12**	**18**	**25**

➔ Calculando a mediana ou segundo quartil: (7+9)/2 = 8

➔ O primeiro quartil: (2+4)/2 = 3

➔ O terceiro quartil: (12+18)/2 = 15



2.3. Medidas de variabilidade - desvio-padrão

➔ Tão importantes como as medidas de tendência central são as medidas de dispersão, que mostram como os dados se espalham ao redor da média;

➔ Quando os números são sempre próximos à média, isso significa que a tendência central representa bem os dados. No entanto, se alguns números ficam longe da média, então a média não representa muito bem todos os dados;

➔ Peças fabricadas que exibem mensurações muito espalhadas não têm qualidade, pois muitas peças vão acabar rejeitadas e retrabalhadas, significando custos altos de fabricação e uma posição fraca em termos da competição empresarial do mercado;

➔ O desvio ao redor da média é definido como a diferença entre um número individual e a média de todos os dados; $\text{desvio} = (X_i - \overline{X}_i)$

2.3. Medidas de variabilidade - desvio-padrão

- ➔ Assim, quando a média é menor que o dado individual, o desvio é positivo, e vice-versa. É muito interessante calcular a média dos desvios que representaria a variabilidade dos dados;
- ➔ Como é demonstrada na Tabela, a soma dos desvios é sempre igual a zero, é uma fatalidade matemática, e, portanto, a média dos desvios também é sempre igual a zero;
- ➔ Para resolver o problema do sinal do desvio, é preferível utilizar o quadrado do desvio, também sem sinal, todos somados como antes, e a média deles calculada;

Observação	Produção	Desvio	Desvio quadrado
1	124	-35,90	1.288,81
2	219	59,10	3.492,81
3	183	23,10	533,61
4	188	28,10	789,61
5	100	-59,90	3.588,01
6	139	-20,90	436,81
7	140	-19,90	396,01
8	153	-6,90	47,61
9	221	61,10	3.733,21
10	132	-27,90	778,41
Média =	**159,90**	0,00	1.508,49
		Raiz da média quadrada =	38,84
		Desvio-padrão =	**40,94**

$$\frac{\sum_{i=1}^{n}(X_i - \overline{X}_i)^2}{n-1}$$

2.3. Medidas de variabilidade - desvio-padrão

➔ A média dos quadrados dos desvios leva o nome técnico de variância.

➔ Para chegar a uma medida do desvio médio, é necessário aplicar a raiz quadrada à variância. Esse desvio médio recebe outro nome em estatística, o **desvio-padrão**.

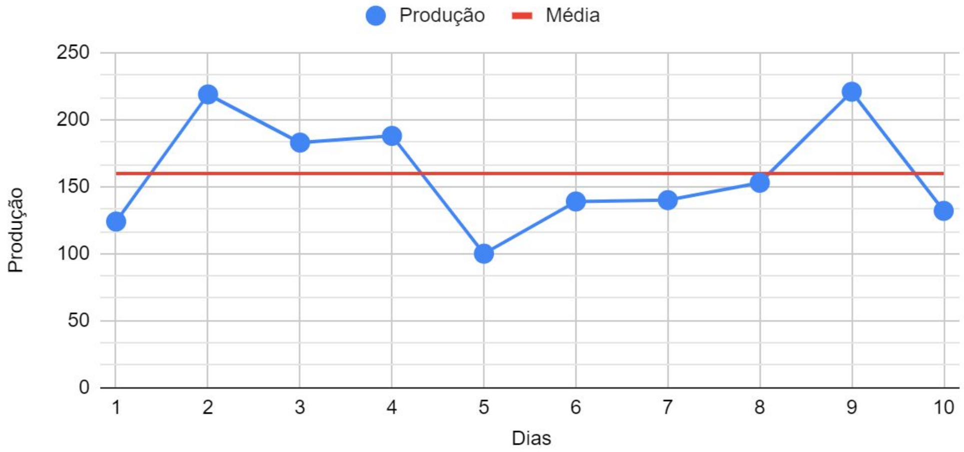
➔ O desvio-padrão muitas vezes é simbolizado com sigma (σ):

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X}_i)^2}{n-1}}$$

➔ O desvio-padrão para os dados apresentados é 40,94.

3. Gráficos - caixa de medianas e histograma

Sem a menor dúvida, a melhor maneira de analisar uma série de dados é graficamente. A tentativa de ver padrões e tendências em uma relação de dados escritos em uma tabela certamente resultará em fracasso, especialmente quando o número de dados é grande.



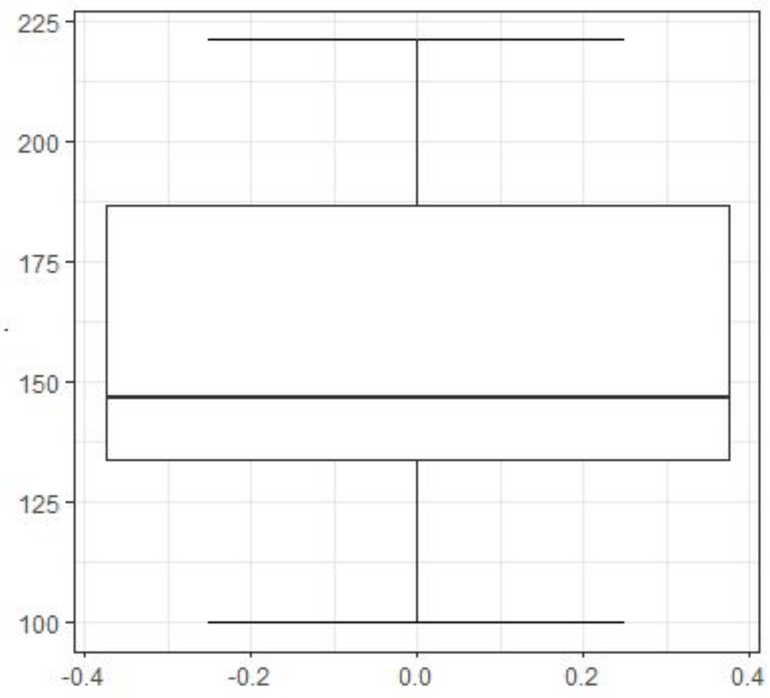
3. Gráficos - caixa de medianas e histograma

➔ Um gráfico que reúne as informações da mediana e dos quartis de uma maneira fácil de entender é a caixa das medianas;

➔ As duas linhas horizontais representam os valores mínimos e máximos de toda a série, ou em outras palavras, a distância entre elas é a amplitude geral dos dados.

➔ A caixa no meio da figura representa o quartil inferior e o superior, onde fica agrupada a metade central dos dados, e a distância entre esses valores é o desvio quartílico.

➔ Finalmente, a linha dentro da caixa é a mediana.

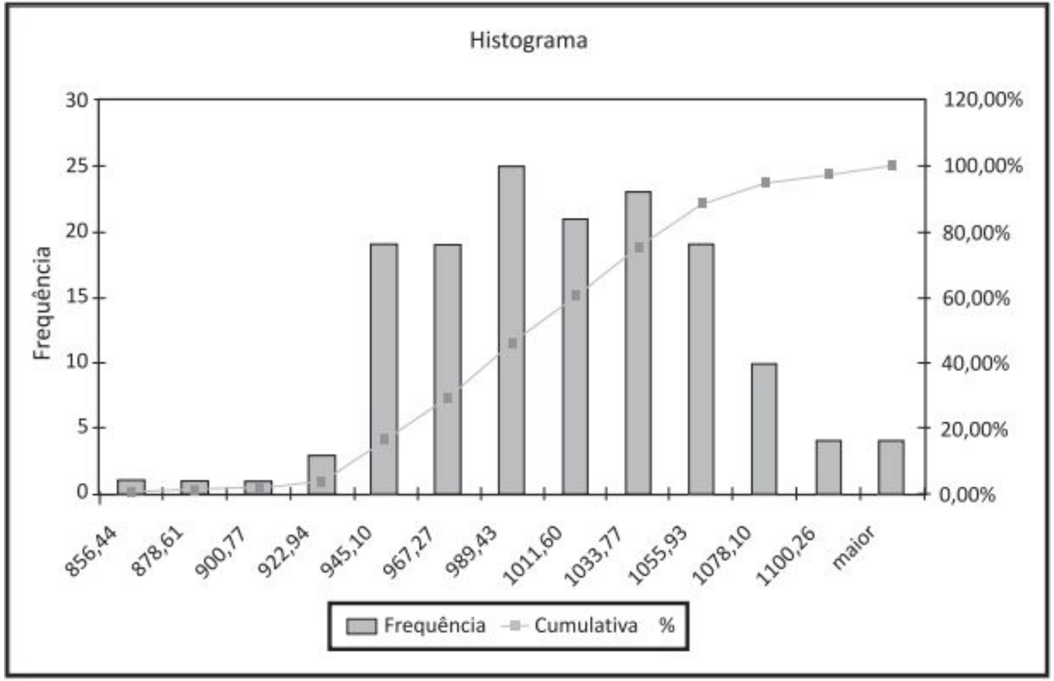


3. Gráficos - caixa de medianas e histograma

Finalmente, apresenta-se o histograma, um gráfico que tem todas as boas características da caixa de medianas, mas exibe muito mais informação sobre a distribuição dos dados. Foram amostrados em um laticínio 150 sacos de leite contendo por lei 1 litro do alimento. O histograma é um retrato dos dados na Tabela logo em seguida.

Classes até	Frequência	Cumulativa %
856,44	1	0,67%
878,61	1	1,33%
900,77	1	2,00%
922,94	3	4,00%
945,10	19	16,67%
967,27	19	29,33%
989,43	25	46,00%
1011,60	21	60,00%
1033,77	23	75,33%
1055,93	19	88,00%
1078,10	10	94,67%
1100,26	4	97,33%
maior	4	100,00%

3. Gráficos - caixa de medianas e histograma



3. Gráficos - caixa de medianas e histograma

- Na primeira linha da Tabela 9.4, dos 150 sacos investigados, um saco entra na classe de pesos de zero a 864,44ml. Na próxima linha, tem a classe de sacos entre 864,44ml a 878,61ml, e nessa classe tem de novo um saco. A frequência mais comum onde caíram 25 sacos de leite é a de 967,27 a 989,43. A última coluna da Tabela mostra a porcentagem cumulativa de frequências até o tamanho daquela classe. Por exemplo, de todos os sacos amostrados, 16,67% têm volume até 945,10ml. É claro que isso significa que aproximadamente 83% dos sacos têm tamanho maior.

- Toda essa informação também consta na Figura, mas de uma maneira mais clara e mais fácil, graficamente. Por sinal, a forma do histograma, com frequências altas no meio dos números e frequências mais baixas em números distantes da tendência central, é muito comum. Essa constatação é a base da famosa distribuição normal. O histograma apresenta de forma quase completa a tendência central dos dados e a variabilidade, melhor que a caixa das medianas.

Espaço para dúvidas

Prof. Ramon Gomes da Silva, MSc.

ramongs1406@gmail.com
https://ramongss.github.io