



Fundamentos do Controle Estatístico de Qualidade



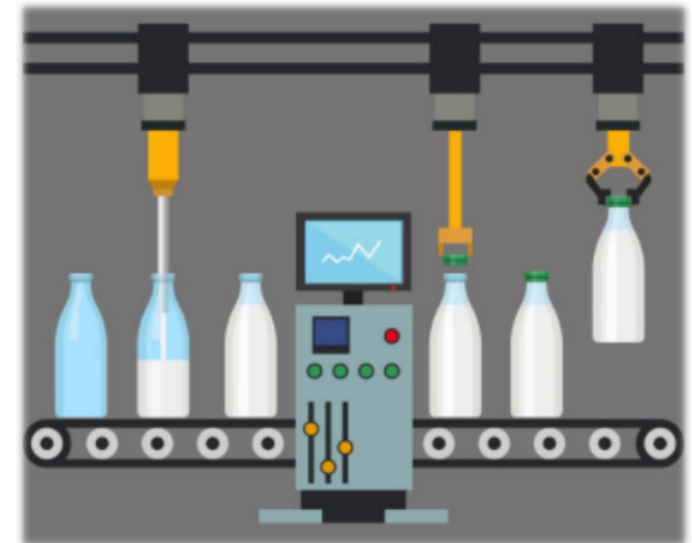
Não há na literatura uma definição uma **definição única** e universal **para qualidade**. Segundo **Falconi (1992)**, qualidade pode ser definida como: “Um **produto** ou um **serviço** de qualidade é aquele que atende perfeitamente, de forma **confiável**, de forma **acessível**, de forma **segura** e no **tempo certo** às necessidades do cliente.

Qualidade é um conceito que agrega valor. O desafio é descobrir os investimentos que levam a uma qualidade mais elevada e com custos mais baixos por unidade produzida.

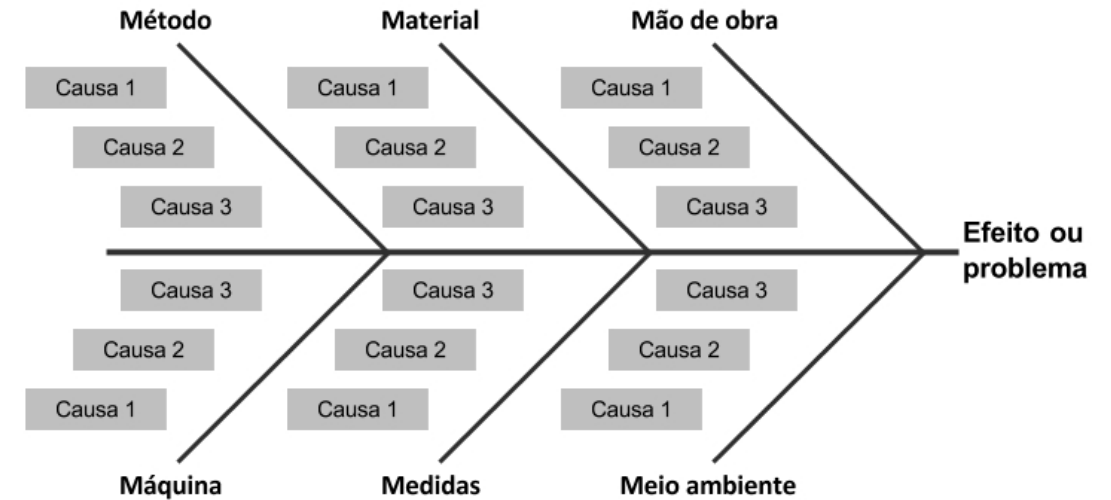
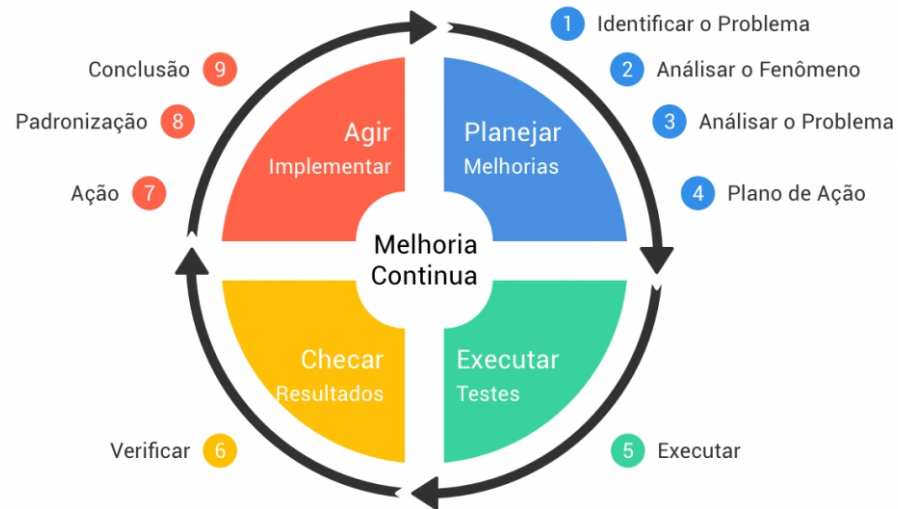
A **variabilidade do processo** tem a ver com as **diferenças entre as unidades produzidas** - Todo e qualquer processo, por mais bem projetado e controlado, possui uma **variabilidade natural**, fruto de pequenas **perturbações ou causas aleatórias**.

Causas especiais (perturbações maiores) tem o efeito de **deslocar a distribuição** da variável de interesse e em geral podem ser corrigidas.

Para construir os **gráficos de controle**, precisamos estimar o **desvio padrão** do processo e sua respectiva **média** (ou usar um valor teórico pré-estabelecido), garantir que as **observações** sejam **independentes** e estimar os **limites de controle**.



Antes de tudo é necessário conhecer o processo, buscando-se conhecer os fatores que afetam a característica de interesse, e eliminando as causas especiais conhecidas.



O **PDCA** é uma metodologia para melhoria contínua de processos e produtos, e que visa garantir o alcance das metas necessárias a “sobrevivência” da organização/conceito/atividade.

- **Subgrupos racionais:** Preconiza a retirada de pequenas **amostras** em **intervalos** de tempo **regulares**, minimizando a chance de amostras com elementos de diferentes populações.
- **Amostragem Estratificada:** Os dados precisam ser provenientes da mesma fonte. Caso haja **mais de uma máquina** no processo, é necessário **monitorar** cada máquina **individualmente**, onde cada máquina constitui um estrato.



Exemplo: Suponha uma linha de empacotamento de leite. Se medirmos o volume de cada saquinho, vamos descobrir que nenhum deles contém exatamente a mesma quantidade de leite. Para fins de comercialização, pouco importa se um saquinho contém um pouco mais ou um pouco menos de leite em relação ao valor especificado, no entanto, existem restrições que devem ser respeitadas: se o “pouco a mais” não for assim tão pouco, o invólucro pode estourar; por outro lado, o “pouco a menos” poderá resultar em multa ao produtor. Assim, a especificação determina que cada saquinho contenha 1000 ml de leite (valor alvo). Considere uma situação em que foram colhidas 25 amostras (m) de tamanho 5 (n) para o volume dos saquinhos de leite. Deseja-se estimar os limites de controle para esse processo.

Limites de controle para a amplitude (variabilidade do processo)

$$LSC_R = (d_2 + 3d_3)\hat{\sigma}_0 \quad LIC_R = (d_2 - 3d_3)\hat{\sigma}_0 \quad \bar{R} = \sum_{i=1}^m \frac{R_i}{m} \quad \hat{\sigma}_0 = S_D = \frac{\bar{R}}{d_2}$$

R_i e \bar{X}_i correspondem a amplitude e média da i – ésima amostra, respectivamente
 d_2 e d_3 são valores tabelados

Limites de controle para a Média

$$LSC_{\bar{X}} = \hat{\mu}_0 + 3 \frac{\hat{\sigma}_0}{\sqrt{n}} \quad LIC_{\bar{X}} = \hat{\mu}_0 - 3 \frac{\hat{\sigma}_0}{\sqrt{n}} \quad \bar{\bar{X}} = \frac{\sum_{i=1}^m \bar{X}_i}{m}$$

Gráficos de Controle

Gráfico de Controle para Amplitude

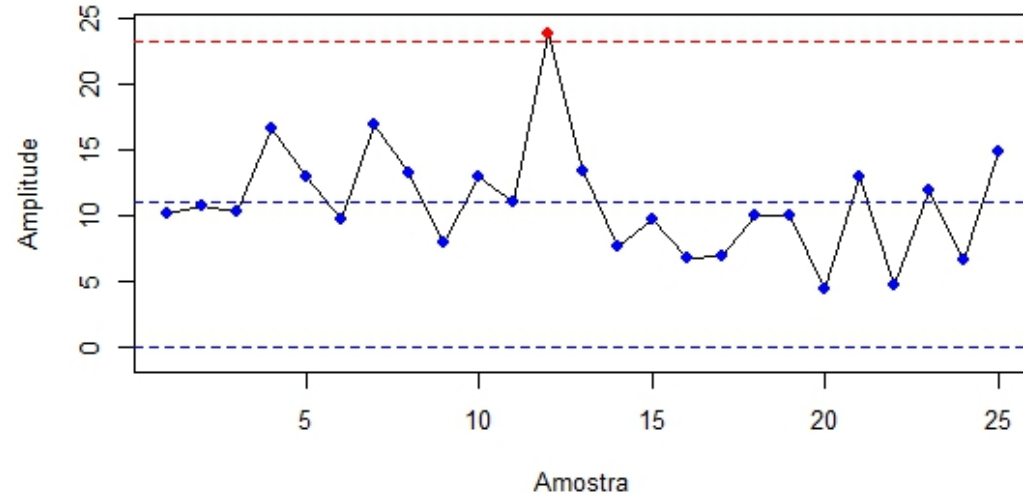


Gráfico de Controle para Amplitude

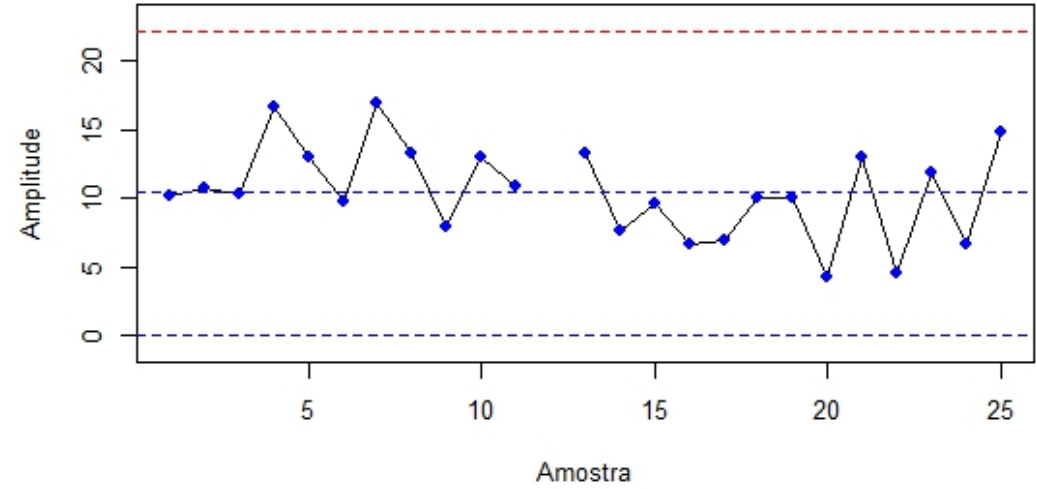


Gráfico de Controle para Média

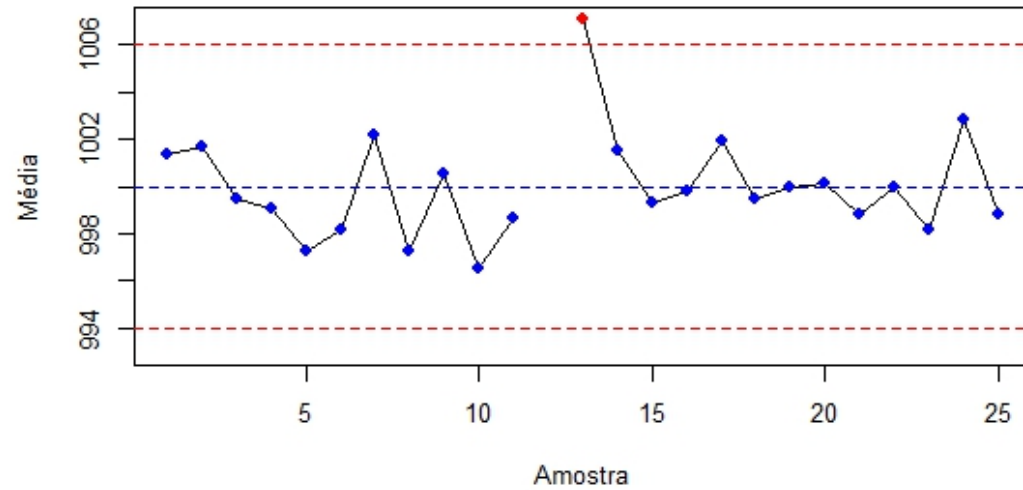
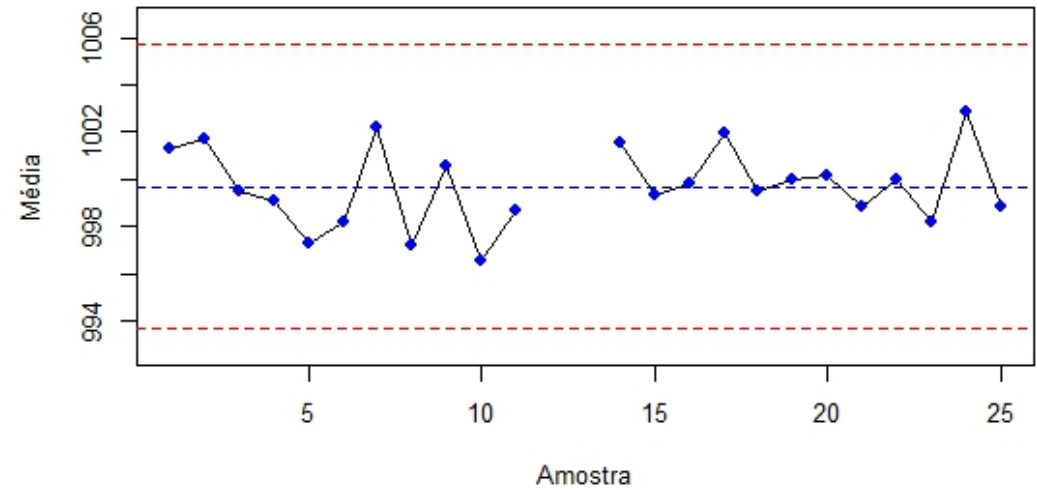


Gráfico de Controle para Média



Verdade	Decisão	
	Aceitar H_0	Rejeitar H_0
H_0 verdadeiro	— $(1 - \alpha)$	Erro Tipo I (α)
H_0 falso	Erro Tipo II β	— $(1 - \beta)$

A consequência de ordem prática associada ao **erro tipo I (alarme falso)** é **intervir** no processo na **hora errada**, onde a interrupção do processo acarreta em custos e perdas. Por outro lado, a **consequência** associada ou **erro tipo II (não detecção)** é **não intervir** no processo quando o mesmo está sob influências de causas especiais.

Dado que o processo é considerado em controle (H_0 verdadeira) quando \bar{X} cai dentro dos limites do gráfico e fora de controle caso contrário: tem-se que:

$$\alpha = Pr[\bar{X} < LIC_{\bar{X}} \text{ ou } \bar{X} > LSC_{\bar{X}} \mid \mu = \mu_0] \text{ e } \beta = Pr[LIC_{\bar{X}} \leq \bar{X} \leq LSC_{\bar{X}} \mid \mu \neq \mu_0]$$

O poder do gráfico de controle, Pd , é definido como a probabilidade de detecção ($Pd=1-\beta$). A distribuição de \bar{X} é assumida normal pelo Teorema Central do Limite, definindo a variável aleatória Z como: $Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$.

Tradicionalmente, adota-se um valor de $z=3$ na construção de limites de controle. Pela tabela da distribuição normal é possível concluir que o risco $\alpha=0,0027$, ou seja, enquanto o processo permanecer estável e ajustado, essa é a chance que uma amostra tem de gerar um alarme falso.

A **distribuição** do número de **amostras** que antecedem um alarme falso é **geométrica** (parâmetro $p = \alpha$). A média da distribuição geométrica é igual a $1/p$, ou seja, teremos em média um alarme falso a cada $1/0,0027=370,4$ pontos plotados.

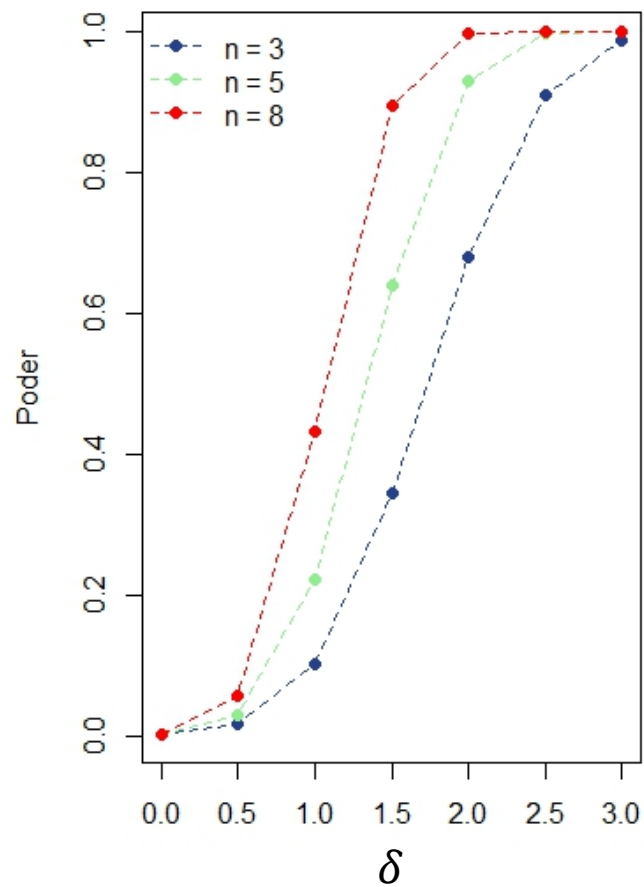
Caso o usuário considere essa **frequência de alarmes inaceitável**, pode-se **estreitar ou alargar os limites de controle** de acordo com o z de α escolhido.

Note que o **tamanho da amostra n não afeta a probabilidade de um alarme falso**, por outro lado tem uma grande **influência** no risco de ocorrência do **erro tipo II** e no **poder** do gráfico de controle.

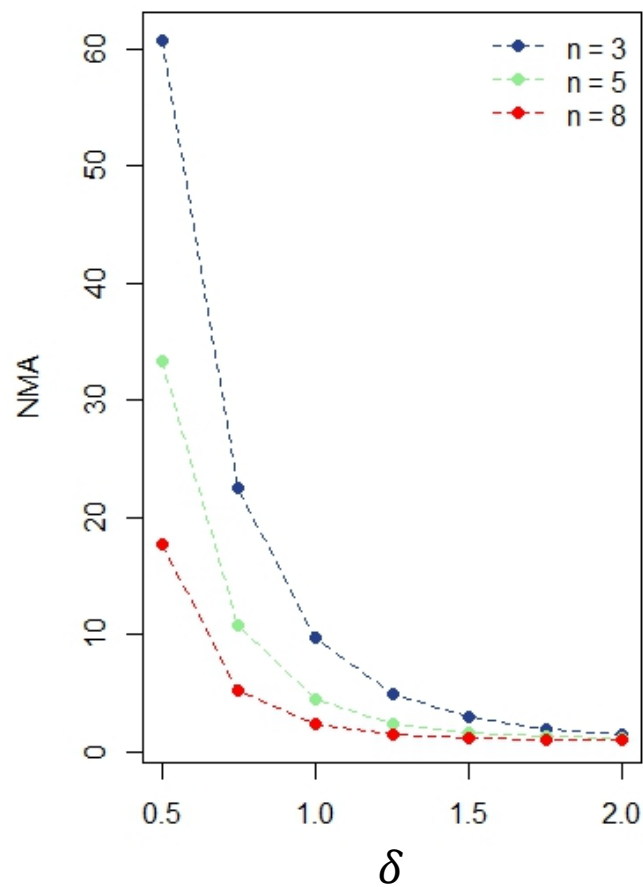
Quando o **processo** está sob a **influência de causas especiais**, o **ideal** seria que o **primeiro ponto** já caísse **fora dos limites de controle**. Nesse contexto, é usual expressar o deslocamento da média do processo segundo o desvio padrão da variável X , portanto $\mu_1 = \mu_0 + \delta\sigma_0$ e $\delta = \frac{\mu_1 - \mu_0}{\sigma_0}$.

Após simplificações, tem-se que: $Pd = Pr[Z < -k + \delta\sqrt{n}] + Pr[Z < -k - \delta\sqrt{n}]$. Quando $k = z = 3$, $\delta = 1$ e $n = 5$, teremos $Z_{LSC} = -0,7639$ e $Z_{LIC} = -5,2361$, logo, $Pd = 0,2236 + 0,0 = 0,2236$. Assim como antes, o número medio de amostras até um alarme verdadeiro segue uma distribuição geométrica e pode ser obtido por $1/Pd = 1/0,2236 = 4,4723$. Ou seja, são necessárias 4,47 amostras para se detectar um deslocamento da média em 1 desvio padrão.

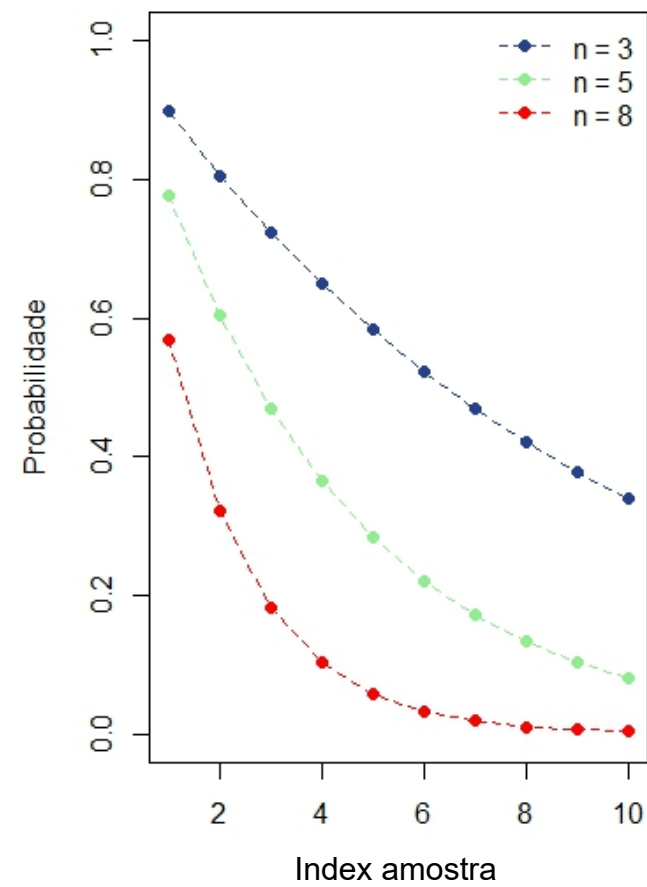
Curvas de Poder



Rapidez de Detecção



Não Detecção



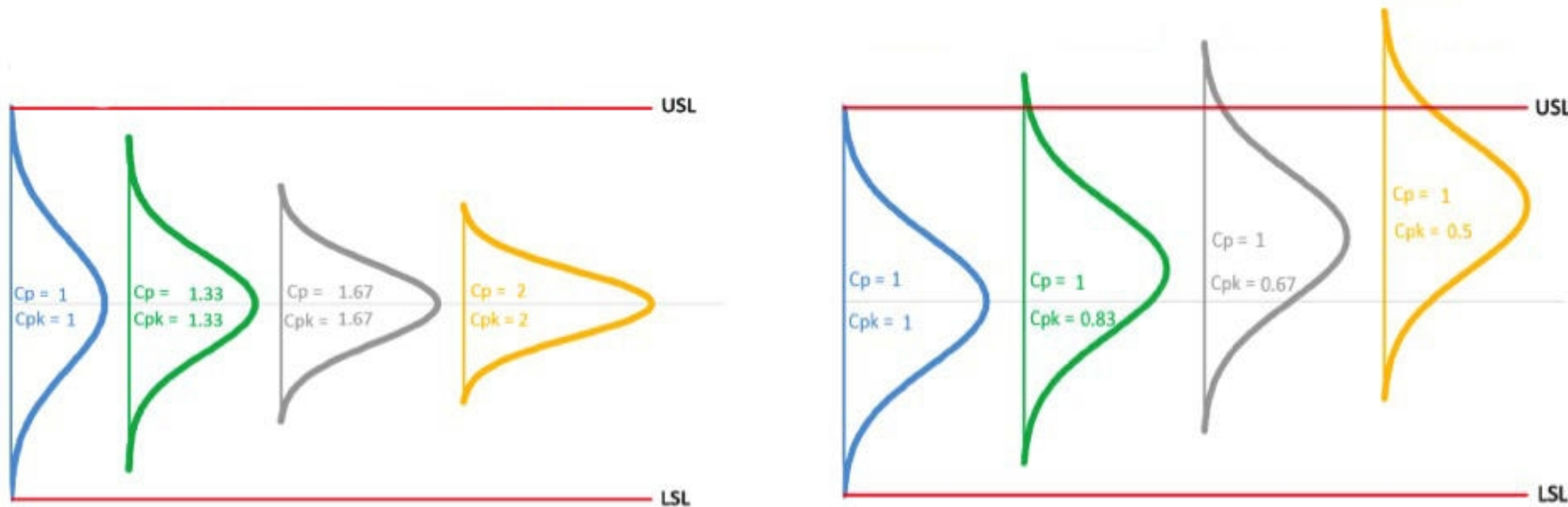
- **Aumentando o tamanho das amostras, aumenta-se o poder** e, conseqüentemente, **reduz-se o número de amostras e o tempo** necessários para que se **detecte** a presença de **causas especiais** processo.
- A **curva de probabilidade de não detecção** visa calcular a **chance** que os **novos valores** caiam **dentro dos limites do processo** mesmo o processo estando fora de controle. As velocidades de detecção estão associadas ao valor δ . Fixando $\delta = 1$, e subgrupo racional de tamanho $n = 8$, tal desvio tem altíssima chance de ter sido tectado até a amostra de index 8, enquanto que para $n=3$, a chance do desvio passar despercebido ainda será de aproximadamente 40%.
- Análises semelhantes podem e devem ser feitas com o gráfico de amplitude/variabilidade, assim como conjuntamente em ambos. Para mais detalhes, consultar as referências desse material.



Existem uma grande possibilidade de utilização de outros tipos de gráficos de controle. Não é intuito desse material passar por todos eles, mas ficam como sugestões de pesquisas à quem possa interessar.

- Gráficos de controle por atributos - proporção e frequência esperada;
- Gráficos de controle de CUSUM e EWMA;
- Gráficos de controle para medidas individuais e autocorrelacionadas;
- Gráficos de controle multivariados - múltiplos atributos;

A **capacidade de um processo** pode ser definido como a **capacidade de produzir itens** conformes, ou seja, **de acordo com suas especificações**. Alterações na variabilidade do processo e na sua média, assim como a definição dos limites de especificação podem interferir nas conformidades do mesmo.



Índices de Capacidade

$$C_p = \frac{LSE - LIE}{6\sigma}$$

$$C_{pk} = \text{Min} \left\{ \frac{LSE - \mu}{3\sigma}, \frac{\mu - LIE}{3\sigma} \right\}$$

$$C_{pm} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + (d - \mu)^2}}$$

LSE e LIE são os limites de especificação do processo, definidos arbitrariamente, e d é o ponto médio do intervalo de especificação

Valores de Referência para Capacidade

- *Capaz:* $C_{pk} \geq 1,33$
- *Razoavelmente capaz:* $1 \leq C_{pk} < 1,33$
- *Incapaz:* $C_{pk} < 1$
- *No caso do C_{pm} ,* quanto maior, mais capaz é o processo;

Especificação bilateral e processo centrado ($C_p = C_{pk}$)

- *Capaz:* 70 (ppm)
- *Razoavelmente capaz:* De 70 a 2700
- *Incapaz:* Acima de 2700

Especificação unilateral e/ou processo não centrado

- *Capaz:* 35 (ppm)
- *Razoavelmente capaz:* De 35 a 1350
- *Incapaz:* Acima de 1350

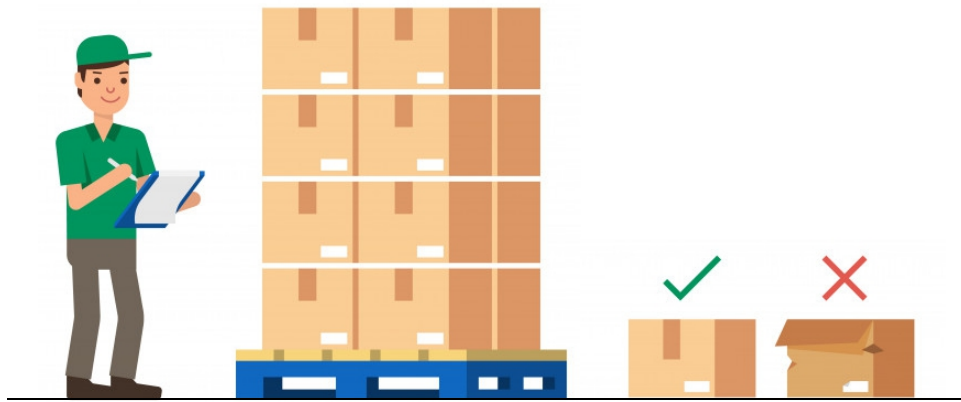
Considerações importantes sobre capacidade:

- Um processo pode estar em controle e ser pouco capaz, assim como o contrário;
- C_p é insensível a mudanças na média;
- C_{pk} pode assumir valores negativos se a média do processo não pertence ao intervalo das especificações.
- C_{pm} tem o problema de que processos que produzem itens não conformes em proporções muito diferentes podem ter valores de C_{pm} próximos, assim como o contrário.

μ	σ	C_p	C_{pk}	C_{pm}	ppm
1000	2	2,50	2,50	2,50	0,00
1002	2	2,50	2,17	1,77	0,00
1000	4	1,25	1,25	1,25	176,83
1002	4	1,25	1,08	1,12	587,71



No caso de um processo para envazamento de garrafas com limites de especificações de LIE e LSE de 985 ml e 1015 ml, respectivamente, os valores estimados de capacidade podem ser consultados na tabela acima conforme alguns cenários.



A **inspeção de qualidade** consiste na seleção aleatória de itens de um lote de produtos, afim de compor uma **amostra de avaliação da qualidade** dos mesmos, podendo **acarretar** ou não, na **rejeição** de um **lote** inteiro de mercadorias.

Se m for o tamanho do lote e D seu número de defeituosos, a probabilidade de uma amostra aleatória de tamanho n retirada do lote conter exatamente d itens defeituosos segue uma distribuição hipergeométrica:

$$Pd[d = d_0] = \frac{\binom{D}{d_0} \binom{m - D}{n - d_0}}{\binom{m}{n}}$$

Considere um lote de 200 itens, onde 10 são defeituosos. a probabilidade de uma amostra de 20 unidades conter exatamente 2 itens defeituosos será?

$$Pd[d = 2] = \frac{\binom{10}{2} \binom{200 - 10}{20 - 2}}{\binom{200}{20}} = 0,1975$$

A variável D, número de itens defeituosos no lote, tem distribuição binomial:

$$Pd[D = d_0] = \binom{m}{D_0} p^{D_0} (1 - p)^{m - D_0}$$

Supondo $p=0,1$, $Pd[D = 10] = \binom{200}{10} 0,1^{10} (1 - 0,1)^{200 - 10} = 0,0045$. A função de probabilidade do número de defeituosos na amostra é então dada pela regra da probabilidade do total:

$$Pd[d = 2] = \sum_{D_0}^{200} Pr[d = 2 | D = D_0] Pr[D = D_0] = 0,285$$

Nesse exemplo, o número esperado de itens não conformes aceitável na amostra de inspeção seria $0,1 \cdot 20 = 2$. Logo, a probabilidade de se aceitar o lote seria dada por $Pd[d = 0] + Pd[d = 1] + Pd[d = 2] = 0,6769$, logo, há probabilidade de se rejeitar o lote de mercadorias é de 0,3231.

Inspeção Retificadora e Planos de amostragem são outras duas abordagens que tangibilizam o conceito de inspeção de qualidade, mas que não serão abordadas nesse material.

Técnicas de controle de qualidade poderiam ser utilizadas em Machine Learning?



- Costa, A. F. B. , Epprecht, E. K., Carpinetti, L. C. R. Controle Estatístico da Qualidade. 1ª edição, São Paulo: Atlas, 2004.
- Notas de aula da professora Sueli Mingoti (DEST-UFMG).