ENGENHARIA DA QUALIDADE A
ENG 09008

AULA 8
PROJETOS DE EXPERIMENTOS - ANOVA

PROFESSORES:

CARLA SCHWENGBER TEN CATEN

Tópicos desta aula

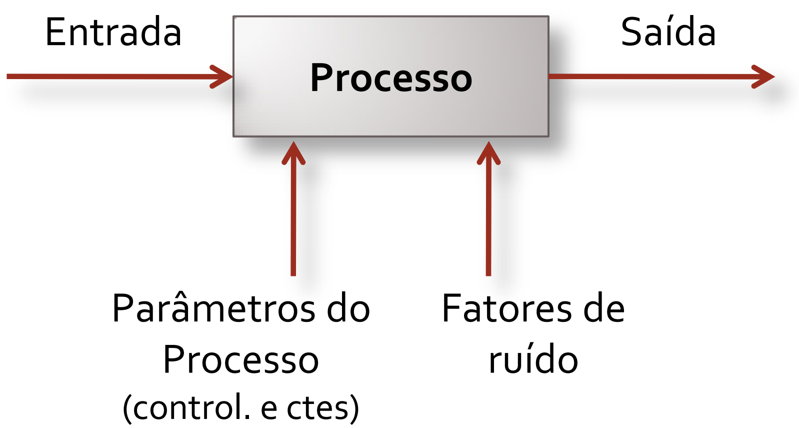
- Projetos de Experimentos
  - Terminologia
  - Passos de implementação
  - Princípios fundamentais
  - ANOVA
  - Otimização

Projeto de Experimentos

- Projeto de Experimentos é uma atividade que em geral envolve o estudo sistêmico de diversos fatores que podem afetar uma ou mais características de qualidade.
- Apesar do título Projeto de Experimentos, essa é uma atividade que envolve:
  - O planejamento do experimento.
  - A sua execução.
  - A análise dos resultados.
  - A otimização do produto/processo em estudo.

Projeto de Experimentos

- O uso de Projeto de Experimentos conduz a uma seqüência estruturada de ensaios, que assegura o máximo de informação com um gasto mínimo de tempo/dinheiro.



Testes onde são feitos mudanças propositais nos fatores controláveis

Terminologia

Parâmetros do processo

- Todas as variáveis da linha de produção que podem ser alteradas e que talvez tenham um efeito sobre as variáveis de resposta.

Fatores controláveis

- São um subconjunto dos parâmetros do processo; são aqueles parâmetros do processo que foram eleitos para serem estudados a vários níveis no experimento.

Terminologia

Fatores constantes

- São os parâmetros do processo que não entrarão no experimento e que são mantidos constantes durante o experimento.

Fatores não controláveis (Ruído)

- São as variáveis que não podem ser controladas pela equipe técnica. Também são responsáveis pelo erro experimental ou variabilidade residual.

Passos no projeto de experimentos

1. Estabelecimento do problema
   - Definição dos objetivos

2. Formulação das hipóteses
   - Brainstorming

Passos no projeto de experimentos

3. Planejamento do Experimento
   - Escolha das variáveis de resposta.
   - Escolha dos fatores controláveis
   - Escolha do nº de níveis para cada fator controlável
   - Identificação das restrições experimentais
   - Escolha das técnicas experimentais
   - Escolha do modelo estatístico

4. Coletas de dados
   - Realização dos experimentos

Passos no projeto de experimentos

5. Análise Estatísticas
   - Tabelas e Testes de significância
6. Análise técnica e tomada de decisão
   - Gráficos
7. Verificação dos resultados
   - Confirmação em campo dos resultados
8. Conclusões finais e recomendações
   - Documentação

Princípios fundamentais

Um bom experimento é aquele que de forma eficiente traz a tona informações precisas que permitem verificar as hipóteses formuladas.

Há quatro princípios fundamentais a serem observados:

1. **Aleatorização**
   - A aleatorização da seqüência de ensaios assegura que as respostas estarão livres de qualquer viés. Quando a aleatorização não é empregada, corre-se o risco de coletar dados viciados, que podem falsear a análise de algum fator.

Princípios fundamentais

Por exemplo

- Imagine que todos os ensaios com o fator A no nível baixo sejam feitos no início da manhã (quando a temperatura do dia está baixa), e todos os ensaios com o fator A no nível alto sejam feitos no início da tarde (quando a temperatura do dia está alta).
- Nesse caso, o efeito do fator A estará confundido com o efeito da temperatura (se este existir) e a análise deste fator estará prejudicada.

Princípios fundamentais

2. Repetições

- As repetições fornecem uma base para o cálculo do erro experimental, o qual, por sua vez, fornece uma base para o julgamento da significância do efeito dos diversos fatores.
- Se um fator tem um efeito significativo, este efeito deve ser muito maior que a magnitude do erro experimental, e os resultados do experimento irão revelar isso.

Princípios fundamentais

3. Controle local

- O controle local se refere a execução de experimentos divididos em blocos equilibrados de ensaio. O objetivo dos blocos é tornar o experimento mais eficiente e mais prático.

4. Trabalho em equipe

- Conhecimentos mercadológicos.
- Conhecimentos técnicos.
- Conhecimentos estatísticos.

Análise de Variância - (Analysis of variance - ANOVA)

- A análise de variância é a metodologia estatística que avalia a significância dos diversos fatores e interações.
- Há suposições básicas para validar a análise de variância:
  - Distribuição normal dos dados.
  - Homogeneidade das variâncias (em cada tratamento ou grupo) - aleatoriedade dos erros.
  - Aditividade dos efeitos.
  - Independência estatística dos valores observados (sem correlação).

Análise de Variância

- Se as suposições de normalidade e homogeneidade não forem satisfeitas, o resultado da análise de variância deixa de ser exato, e passa a ser aproximado.
- Em raras situações a suposição de aditividade dos efeitos não é satisfeita. Nesse caso, uma transformação dos dados (log, √, etc.) pode recuperar a aditividade e permitir uma análise mais precisa.
- A independência estatística dos valores observados é obtida com o uso da aleatorização.

One-Way ANOVA

Experimentos que envolvem:

- 1 Variável de Resposta
- 1 Fator Controlável a vários níveis

  Os ensaios feitos em cada nível do fator controlável configuram um grupo.

Objetivo:

- Identificar se os valores da variável de resposta medidos nos diversos níveis diferem entre si.

One-Way ANOVA

2 tipos de experimentos:

- **Fatores Controláveis a níveis fixos**

  (Por ex., 5 valores de temperatura)

É possível repetir o ensaio tempos depois, basta utilizar os níveis dos FC escolhidos

- **Fatores Controláveis a níveis aleatórios**

  (Por ex., 3 lotes de MP escolhidas ao acaso)

Nunca mais será possível ter os mesmos fatores controláveis

One-Way ANOVA

Disposição dos dados:

Fator A		A1	A2	...	Ak	
		y11	y12	...	y1k	
		y21	y22	...	y2k	
		:	:	yij	:	
		:	:	...	:	
		yn1	yn2	...	ynk	
Totais	T.j	T.1	T.2	...	T.k	T.. =
No.Obs.	nj	n1	n2	...	nk	N =
Médias	$\bar{y}_{.j}$	$\bar{y}_{.1}$	$\bar{y}_{.2}$	...	$\bar{y}_{.k}$	$\bar{\bar{y}}_{..}$ =

Formulação matemática da One-Way ANOVA

Modelo Estatístico: $y_{ij} = \mu + \tau_j + \varepsilon_{ij}$

onde: μ é a média geral;

τ_j é o efeito do grupo j;

ε_{ij} é um erro aleatório.

Hipóteses:

$$H_0: \quad \tau_j = 0$$

$$H_1: \quad \tau_j \neq 0 \quad \forall\ j$$

Objetivo:

- Rejeitar H_0 a um nível de significância α
  - Ou seja, se valor-**p** observado for menor que α, o erro incorrido de rejeitar H_0 (sendo este verdadeiro) em favor de H_1, é muito pequeno.

Formulação matemática da One-Way ANOVA

Exemplo:

Um profissional deseja estudar se a temperatura ambiente influencia na produtividade dos funcionários. Para isso realizou três medidas de produtividade (peças/hora) em três temperaturas diferentes.

Temperatura		
15°C	25°C	35°C
12	20	17
13	19	16
11	18	18

Fator controlável

Níveis do fator controlável

Repetições

Medição da variável de resposta

One-Way ANOVA

Exemplo:

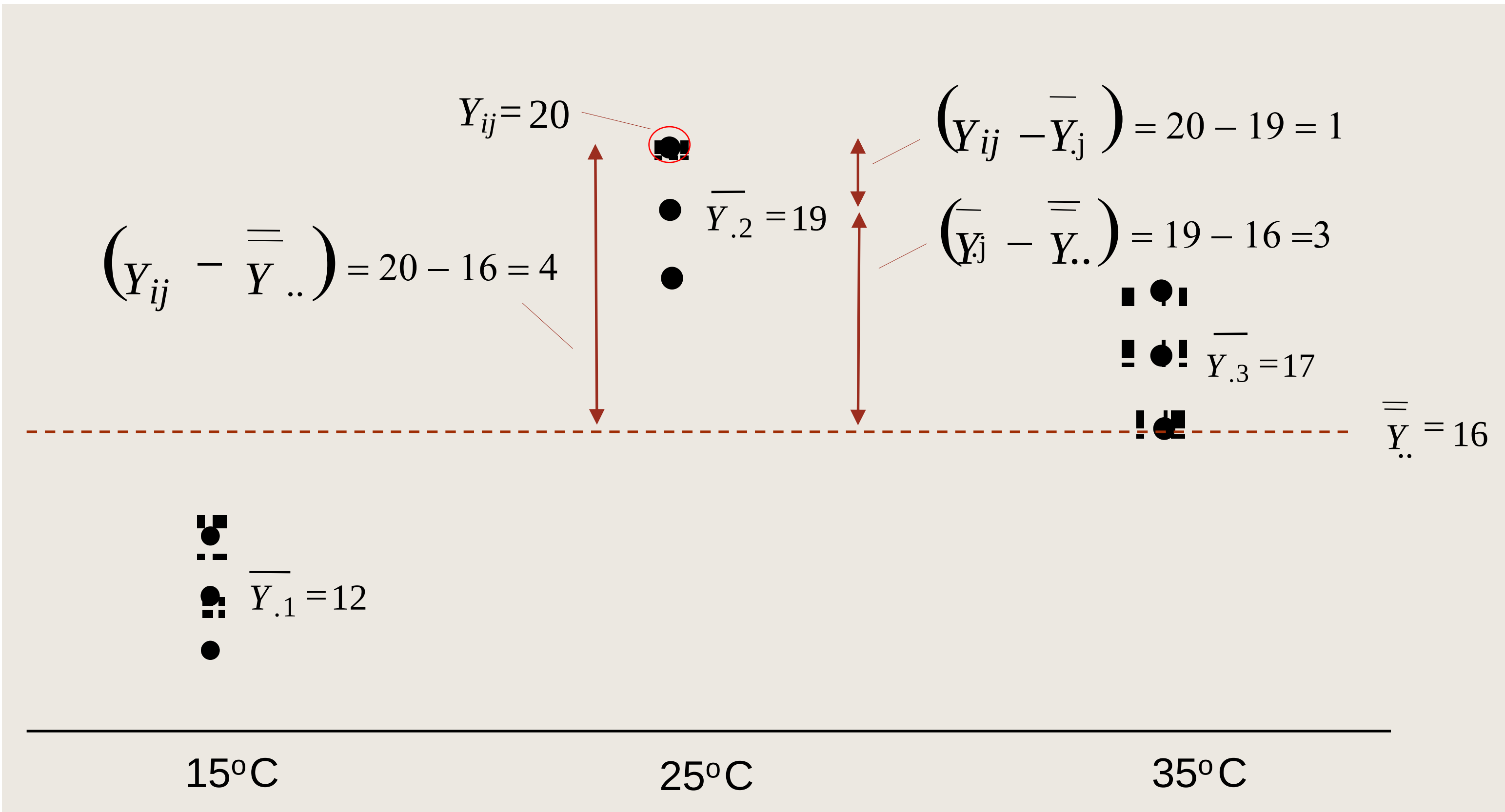
	Temperatura			
	15°C	**25°C**	**35°C**	
	12	20	17	
	13	19	16	
	11	18	18	
$T_{.j} =$	36	57	51	$T_{..} = 144$
$n_j =$	3	3	3	$N = 9$
$\overline{Y}_{.j} =$	12	19	17	$\overline{\overline{Y}}_{..} = 16$

SQR dentro do grupo
Variabilidade residual (erro)

SQG= Variabilidade entre grupos (fator controlável)

SQT = SQG grupo + SQR resíduo

Decomposição da variabilidade



$$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

20 = 16 + 3 + 1

SQT = SQG + SQR

$$\left(Y_{ij} - \bar{\bar{Y}}_{..}\right) = \left(\bar{Y}_{.j} - \bar{\bar{Y}}_{..}\right) + \left(Y_{ij} - \bar{Y}_{.j}\right)$$

4 = 3 + 1

Formulação matemática da One-Way ANOVA

Decomposição dos resíduos:

$$\left(Y_{ij} - \overline{\overline{Y}}_{..}\right) = \left(\overline{Y}_{.j} - \overline{\overline{Y}}_{..}\right) + \left(Y_{ij} - \overline{Y}_{.j}\right)$$

Total *Grupos* *Resíduos*

Elevando ao quadrado e somando:

$$\sum_{i=1}^{n}\sum_{j=1}^{k}(Y_{ij} - \overline{\overline{Y}}_{..})^2 = n\sum_{j=1}^{k}(\overline{Y}_{.j} - \overline{\overline{Y}}_{..})^2 + \sum_{i=1}^{n}\sum_{j=1}^{k}(Y_{ij} - \overline{Y}_{.j})^2$$

SQT = *SQG* + *SQR*

Graus de liberdade:

(N - 1) = *(K - 1)* + *(N - K)*

Formulação matemática da One-Way ANOVA

Médias quadradas:

As MQs são as variâncias

MQG (fator controlável) = SQG / (K - 1)

MQR (residuo) = SQR / (N - K)

- Médias quadradas são variâncias que podem ser comparadas entre si, pois já foram padronizadas devido a divisão pelos graus de liberdade.

SQT = SQG + SQR

(N - 1) = (K - 1) + (N - K)

MQT ≠ MQG + MQR

Teste F

$$F_{calc} = \frac{Variância \;\; entre\ grupos}{Variância \;\; dentro\ do\ grupo} = \frac{MQG(\,fator\,)}{MQR(residuo)}$$

$F_{tabelado} = F_{\alpha,K\text{-}1,N\text{-}k}$

α é um limite de probabilidade aceitável de se cometer o erro do tipo I, ou seja, rejeitar H_0 sendo que a hipótese é verdadeira

Logo α é a probabilidade de errar na conclusão e o intervalo de confiança na decisão é (1- α)

- Comparar Fcalculado com Ftabelado
  - Se o valor calculado for maior que o valor tabelado, descarta-se H_0, ou seja, existe diferenças significativas entre os grupos devido ao fator controlável.
- Se o Teste Fcalc = 1 não há diferenças significativas entre os grupos MQG = MQR e não há evidências suficientes para rejeitar Ho
- Se o Teste Fcalc > Ftab ou valor-p<0,05 rejeita-se Ho e conclui-se existem diferenças significativas entre os grupos devido ao fator controlável MQG > MQR

Fórmulas para os cálculos (simplificadas)

TC = termo de correção

$$TC = T..^2 / N$$

$$SQT = \sum(Y_{ij}^2) - TC$$

$$SQG = \sum(T._j^2 / n_j) - TC$$

$$SQR = SQT - SQG$$

Tabela ANOVA:

Fonte	SQ	GDL	MQ	Teste F
Entre Grupos	SQG	K - 1	MQG	F = MQG / MQR
Dentro Grupos	SQR	N - K	MQR	
Total	SQT	N - 1		

Exemplo

- Os dados a seguir representam o alongamento (maior é melhor) medido sobre um composto de borracha, em função da quantidade de agente de processo adicionado durante a mistura (considere que o LIE = 46).

Exemplo

Agente	0	5	10	15	20	
	43	47	55	50	52	
	47	53	50	54	49	
	46	52	54	54	54	
	45	50	55	55	55	
	45	49	52	56	55	
	46	51	53	52	56	
	47	55	55	57	56	
	44	48	56	57	53	
	42	49	59	55	57	
	48	50	56	60	60	
	49	47	57	56	57	
	44	49	54	58	55	
Totais	546	600	656	664	659	**T..= 3125**
No.Obs.	12	12	12	12	12	**N = 60**
Médias	45,5	50,0	54,7	55,3	54,9	**Y.. = 52,08**

Exemplo

Cálculos iniciais:

TC = $T_{..}^2 / N$ = $(3125)^2 / 60 = 162.760,42$

SQT = $\Sigma (Y_{ij})^2 - TC$ = $163.971,00 - 162.760,42 = 1210,58$

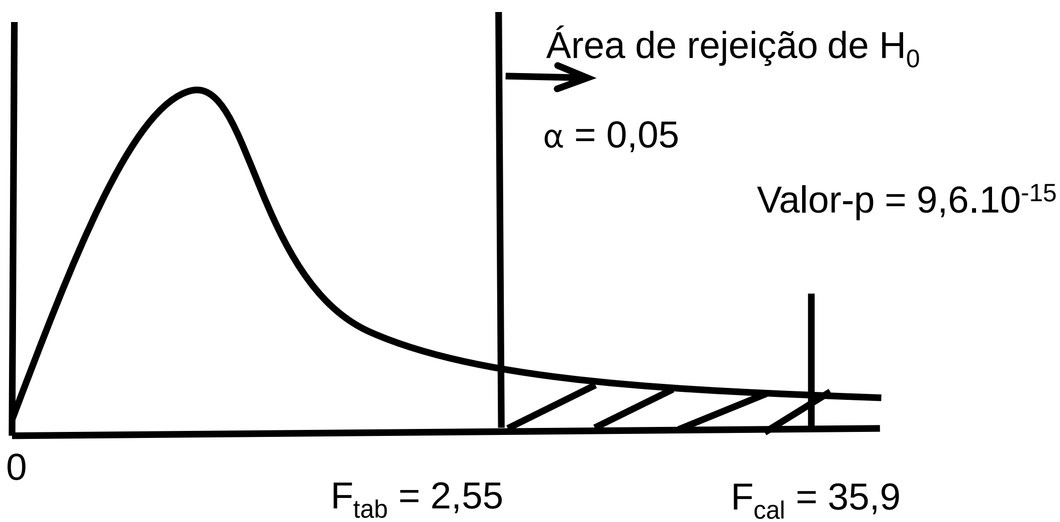
SQG = $\Sigma (T_{.j}^2 / nj) - TC$

$= [(546)^2 / 12] + \ldots + [(659)^2 / 12] - 162.760,42 = 875,33$

SQR = SQT - SQG = $1210,58 - 875,33 = 335,25$

Tabela ANOVA

Fonte	SQ	GDL	MQ	Teste F
Entre Grupos (Agente de processo)	875,33	4	218,83	35,9
Dentro Grupos (Residual)	335,25	55	6,09	
Total	1210,58	59		



Há diferenças significativas entre os grupos quando

$F_{cal} > F_{tab}$

Valor-p < α (0,05 ou 5%)

Existe efeito significativo do fator controlável Agente (valor-p=$9,6\times10^{-15}$) sobre a VR alongamento da borracha

Comparação múltipla de médias (CMM)

- Existem diversos métodos de CMM: Duncan, Tukey, Scheffé, Bartlett, etc
- Deseja-se comparar se existe diferença entre as médias e
- considera-se que elas diferem ser elas vierem de populações diferentes, ou seja as curvas não estão sobrepostas
- Como compara-se médias é necessário calcular o tamanho da curva usando 3 x desvio-padrão das médias
- Considera-se que as médias diferem se uma média estiver fora da curva da outra média, ou seja, a distância entre elas deve ser maior do que 3 x desvio-padrões da médias

Comparação múltipla de médias (CMM)

- O método apresentado a seguir é o de Duncan

1. Calcular o desvio padrão das médias pelo Teorema do Limite Central

$$S_{\overline{Y}_{.j}} = \frac{S}{\sqrt{n}} = \frac{\sqrt{MQR}}{\sqrt{n}} = \frac{\sqrt{6{,}09}}{\sqrt{12}} = 0{,}71$$

onde $n_c = (n_1 + n_2 + ... + n_k) / k$

2. Calcular o limite de decisão

$L_d = 3 \times S_{\overline{Y}_{.j}} = 3 \times 0{,}71 = 2{,}13$

Comparação múltipla de médias (CMM)

3. Escrever as médias em ordem crescente ou decrescente e comparà-las duas a duas (todas combinações). A diferença será significativa se for maior que o Ld (=2,13)

45,5	50,0	54,7	54,9	55,3
Y.1	Y.2	Y.3	Y.5	Y.4

4. Usar barras contínuas sobre as médias que não diferem entre si

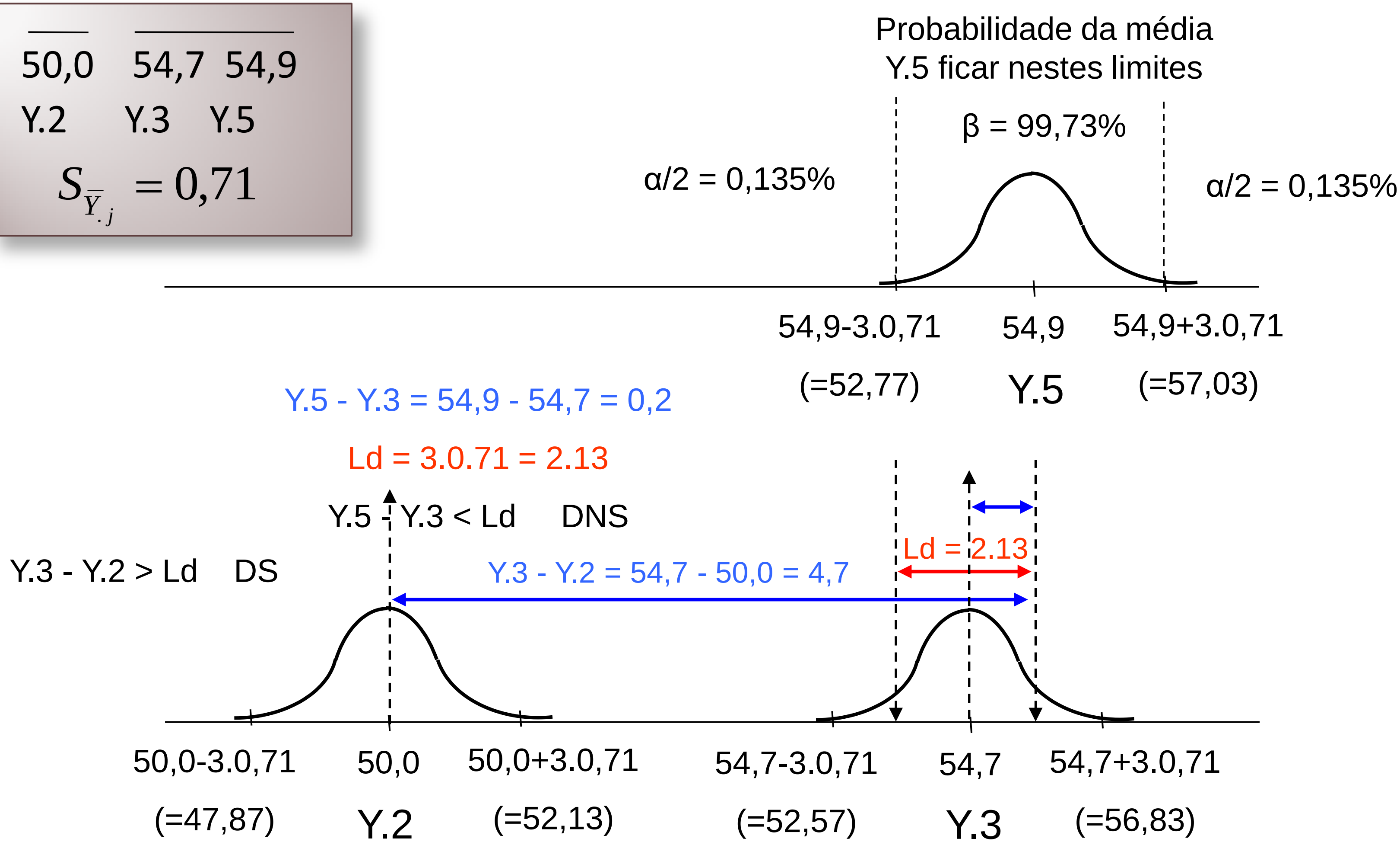
$\overline{\text{Y.1}}$ $\overline{\text{Y.2}}$ $\overline{\text{Y.3 Y.5 Y.4}}$

50,0 – 45,5 = 4,5 > Ld=2,13	DS
54,7 – 50,0 = 4,7 > Ld=2,13	DS
54,9 – 54,7 = 0,2 < Ld=2,13	DNS
55,3 – 54,9 = 0,4 < Ld=2,13	DNS

Comparação múltipla de médias (CMM)

$\overline{50{,}0}$	$\overline{54{,}7 \quad 54{,}9}$	
Y.2	Y.3	Y.5

$$S_{\overline{Y}_{.j}} = 0{,}71$$



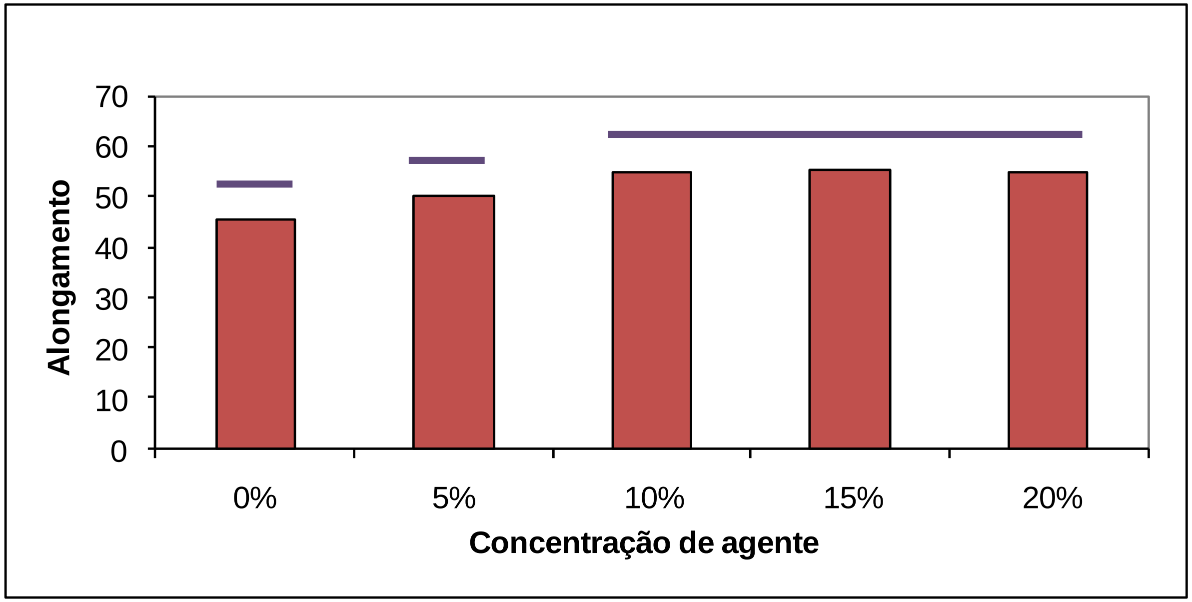
Otimização

- Em geral, deve ser considerado o binômio:
  - Qualidade
  - Custo
- Os resultados estatísticos, em conjunto com a análise gráfica dão suporte à tomada de decisão a respeito do processo.
- Via de regra, o experimento revela opções para a redução de custos e melhoria da qualidade, simultaneamente.

Otimização

- Para realizar a otimização do produto ou processo é feita uma análise técnica.
- A análise técnica deve acompanhar e completar a análise estatística.
- Para isso é recomendável representar graficamente os dados.
- Para os dados do experimento anterior, poderia se usar, por exemplo, um boxplot ou um gráfico de barras.

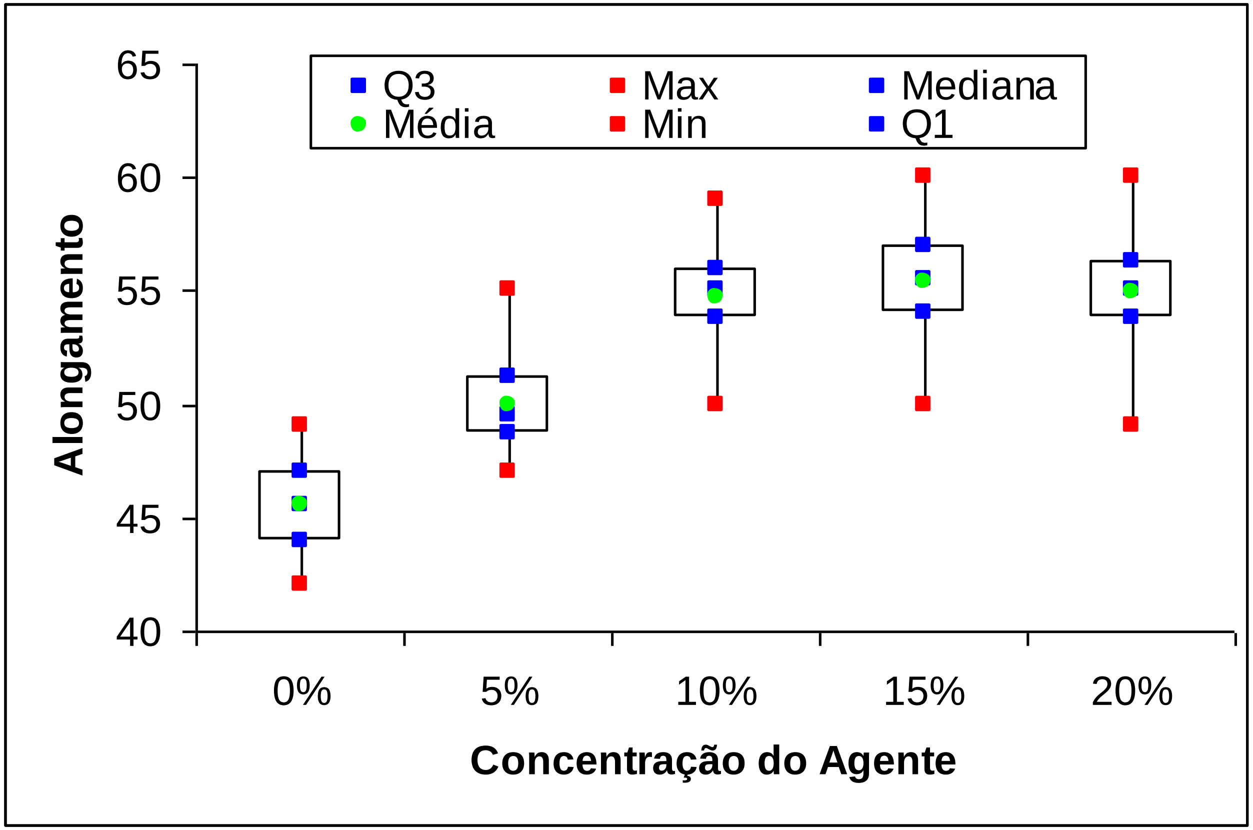
Otimização observando gráfico de barras



Como a CQ alongamento é do tipo maior-é-melhor, então a escolha da concentração de agente pode ser 10, 15 e 20%, pois produzem a mesma resposta média.

Contudo, quanto maior a concentração do agente, mas caro é a borracha, então a escolha recomendada é a concentração de 10%.

Otimização observando o boxplot



Como o LIE do cliente é 46, então a escolha da concentração de agente pode ser 5, 10, 15 e 20%, pois atendem a especificação.

Tópicos próxima aula

- Projetos de Experimentos com dois fatores
- Two-way ANOVA
- Experimentos com ou sem repetição