

**Universidade de São Paulo**  
**Instituto de Ciências Matemáticas e de Computação**  
**Departamento de Sistemas de Computação**

# SSC546 – Avaliação de Sistemas Computacionais

## Parte 1 - Aula 4

Sarita Mazzini Bruschi

Material baseado nos slides de:  
Marcos José Santana  
Regina Helena Carlucci Santana

# Conteúdo

1. Planejamento de Experimentos
2. Técnicas para Avaliação de Desempenho
3. **Análise de resultados**

# Etapas de um Experimento

- Caracterização do problema
- Escolha dos fatores de influência
- Escolha dos níveis, para cada fator
- Seleção das variáveis de resposta
- Determinar um projeto experimental
- Determinar um modelo de planejamento de experimento
- Condução do experimento
- **Planejamento de como será a análise dos dados do experimento**
- Conclusões e recomendações

# Análise de Resultados

- Considera-se que alguma técnica para avaliação de desempenho e obtenção dos resultados tenha sido utilizada

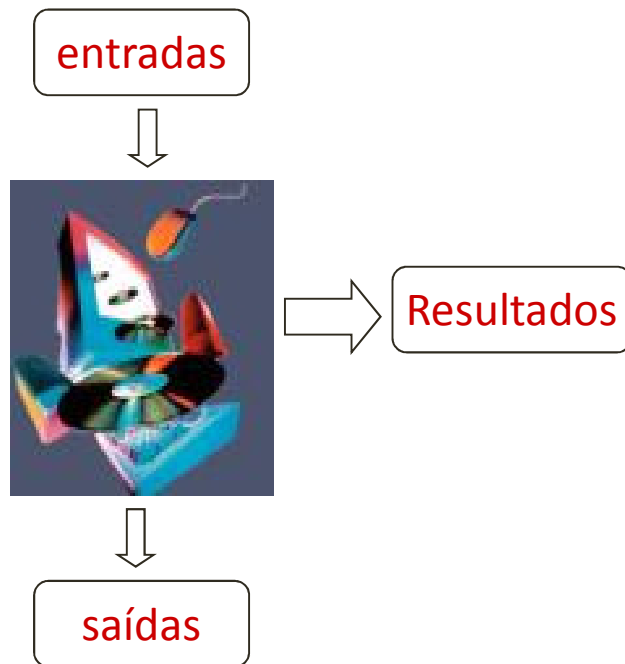
# Análise de Resultados

- Procedimento **errado**, normalmente utilizado para uma avaliação:
  1. Desenvolvimento de um procedimento para avaliação
  2. Validação e verificação do sistema de avaliação
  3. Obtenção dos resultados através de **uma** execução da forma de avaliação escolhida
  4. Conclusões sobre o sistema em estudo

# Análise de Resultados

- Por que esse procedimento está **errado**?

## Aferição



**Sistema a ser Avaliado**

Tem-se controle de todo o sistema?

Como são controladas as entradas do sistema?

Quais as condições iniciais do sistema?

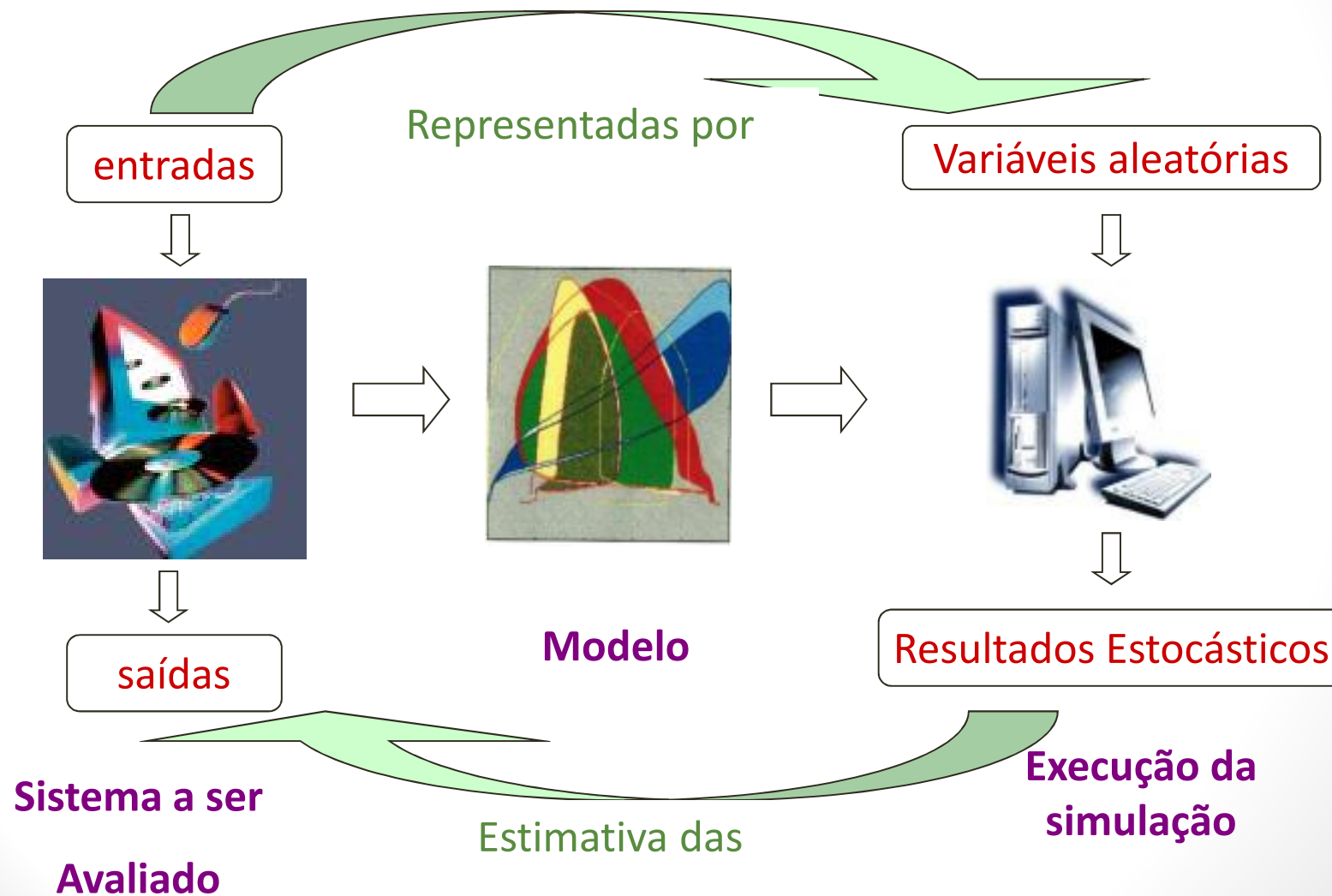
O que mais o sistema está processando no momento da avaliação?

Como controlar as interrupções?

Diferentes características a serem consideradas, p.ex. onde estão localizadas as informações no disco?

# Análise dos resultados

## Modelagem



# Análise de Resultados

- Portanto, em uma simulação estocástica, tem-se uma variabilidade inerente ao processo estocástico da simulação



# Análise de Resultados

- Em qualquer experimentação, três problemas devem ser considerados:
  1. Condições iniciais da experimentação
  2. Quando parar uma experimentação
  3. Resultado de uma execução oferece um resultado dentre muitos outros possível

# Análise de Resultados

- **Primeiro problema:** quais condições iniciais que deve-se ter para iniciar uma avaliação?
- Três condições possíveis:
  1. Início no estado vazio
  2. Início no estado de maior probabilidade de ocorrência
  3. Início na média do estado de equilíbrio

# Análise de Resultados

## 1. Início no estado vazio:

- Simples
- Todos os servidores estão desocupados, portanto filas vazias
- Importante para a análise do período transitório mas não do comportamento estacionário
- Ex.: Banco
  - Comportamento normal x inicial
- Solução:
  - Valores iniciais
  - Truncamento
  - Experimentação muito grande

# Análise de Resultados

## 2. Início no estado de maior probabilidade de ocorrência

- Vantagem:

- Começar em um estado representativo do sistema

- Desvantagem:

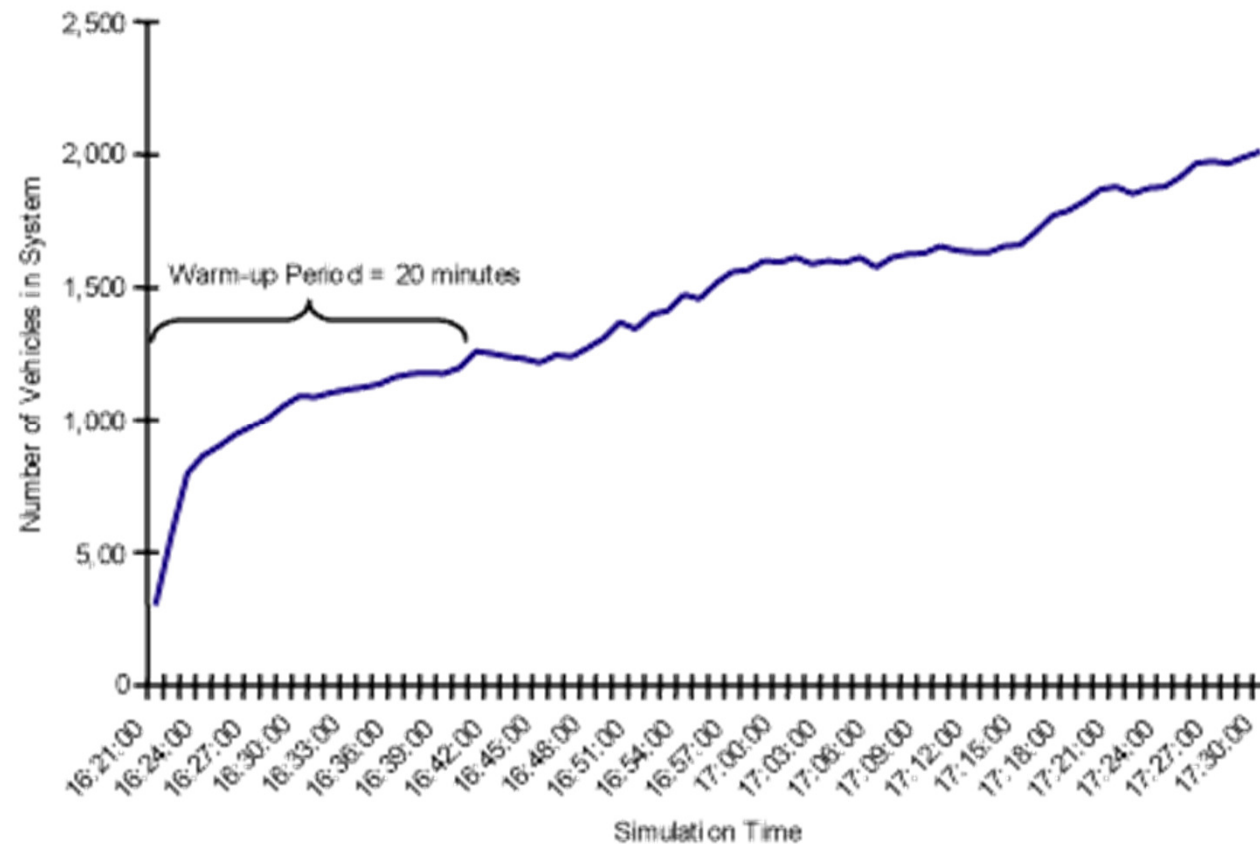
- Como determinar o estado mais provável?
  - Como levar o sistema até esse estado?

# Análise de Resultados

## 3. Início na média do estado de equilíbrio

- Começar a coleta para estatísticas quando os resultados se estabilizam
- Truncamento de dados
  - Período de aquecimento - *warm-up*
  - Retardar a coleta de estatísticas por um período de aquecimento
  - Problema: quando trucar? 10% do valor total (??)

# Análise de Resultados



# Análise de Resultados

- **Segundo problema:** quando parar a avaliação?
- Algumas possibilidades:
  1. Limitar o tempo de experimentação
  2. Limitar o número de elementos que entram no sistema avaliado
  3. Limitar o número de entidades processadas por um servidor
  4. Parada automática

# Análise de Resultados

1. Limitar o tempo de experimentação
  - Desvantagem: número de amostras coletadas será diferente em cada caso
2. Limitar o número de elementos que entram no sistema
  - Termina em estado vazio e ocioso, voltando ao problema inicial
3. Limitar o número de entidades processadas por um servidor
  - Problema: sistemas com prioridades, como por exemplo, o experimento terminar com tarefas longas na fila e isso pode camuflar os resultados



# Análise de Resultados

## 4. Parada automática

- Trabalham com os resultados da experimentação em intervalos selecionados
- Calculam a média e a variância
- Experimentação para quando a estimativa da variância da média está dentro de uma certa tolerância

# Análise de Resultados

- Primeiro e segundo problemas:
  - Relacionados com a condução do experimento
  - Dependem muito da técnica de avaliação que está sendo utilizada
  - Depende bastante do sistema que deve ser avaliado
  - Devem ser melhor discutidos em cada técnica de avaliação e/ou aplicação considerada

# Análise de Resultados

- **Terceiro problema:** durante a obtenção de dados sobre sistemas computacionais, tem-se que os resultados de uma medição oferecem um resultado dentre muitos outros possíveis
  - Deve ser utilizado em qualquer experimento que gere um conjunto de resultados possíveis
  - Após a obtenção dos resultados, estes devem ser analisados independente da técnica utilizada
  - Qual resultado deve ser considerado?
  - Como comparar dois conjuntos de resultados?

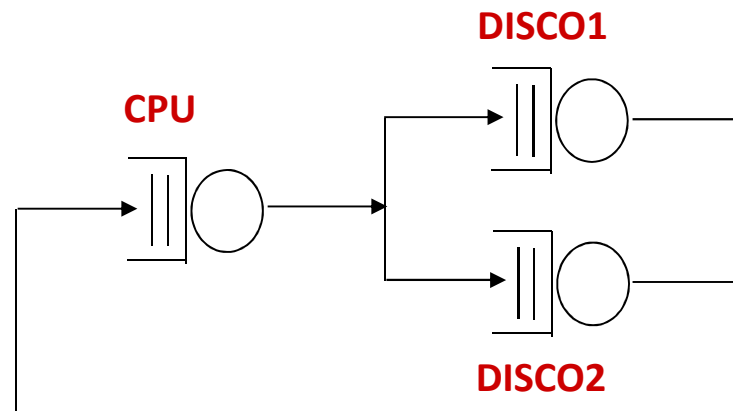
# Análise de Resultados

- Como analisar os diferentes resultados de uma avaliação?
  - Primeiro objetivo da análise estatística dos resultados é estimar o erro ou intervalo de confiança
  - Deve-se considerar diversos resultados provenientes de diferentes execuções:
    - Para simulação: utilizar conjuntos de números aleatórios sem correção (diferentes sementes)
    - Para aferição: considerar diversas medidas

# Análise de Resultados

## Exemplo

- Utilização da CPU
  - Execução 1: 0,36123
  - Execução 2: 0,32571
  - Execução 3: 0,32510
  - Execução 4: 0,29999
  - Execução 5: 0,35813
- Como analisar estes resultados?
- Perigo utilizar resultados de uma única execução

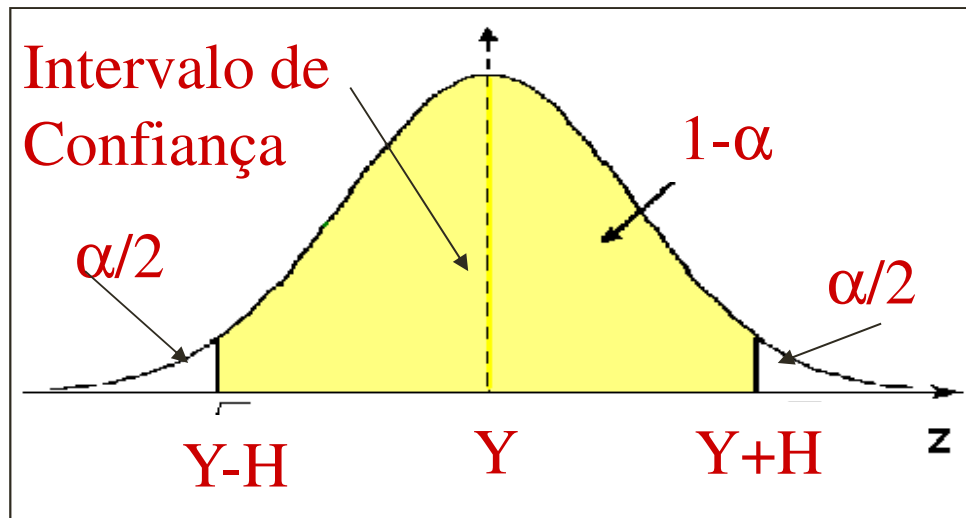


# Análise de Resultados

- Utilização de Intervalos de Confiança
  - A partir de um conjunto de resultados possíveis, queremos estimar o comportamento de um sistema
  - Podemos utilizar:
    - Valores fixos – Estimativa pontual
    - Intervalos – Estimativa intervalar
  - Valores fixos, como por exemplo, a média dos valores, não permitem estimar o erro cometido
  - Intervalos de confiança permitem avaliar a confiança no resultado

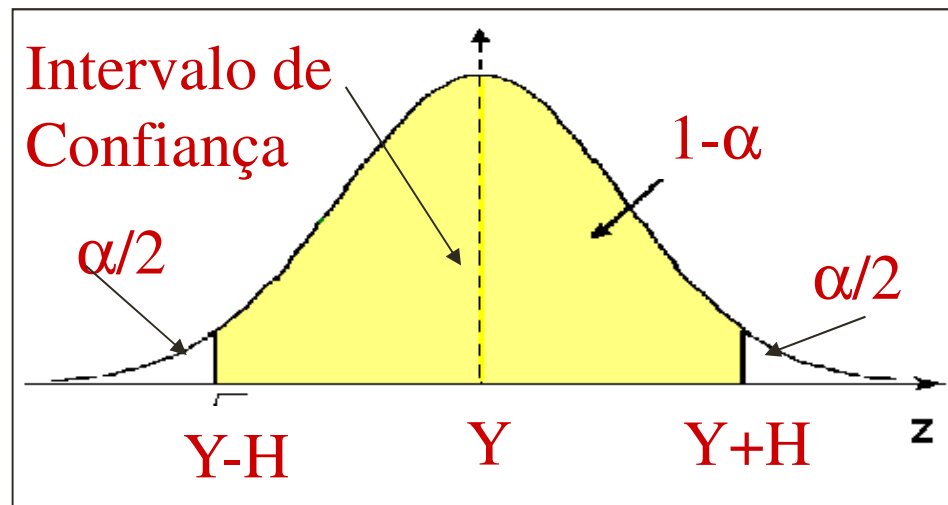
# Análise de Resultados

- O que é o Intervalo de Confiança?
  - Intervalo que com uma determinada probabilidade (confiança) contém o valor do parâmetro estudado
  - Nível de confiança:  $100 * (1-\alpha)\%$
  - $\alpha$  = probabilidade de erro
  - $Y$  = média da amostra
  - $H$  = largura do Intervalo de Confiança



# Análise de Resultados

- O que significa Intervalo de Confiança?
- Se Confiança = 95%, tem-se 95% de chances de que o parâmetro estará dentro do intervalo
- Nada garante que o resultado de uma única execução ( $Y_i$ ) cairá no intervalo
- O resultado de uma única execução poderá estar na área definida por  $\alpha/2$



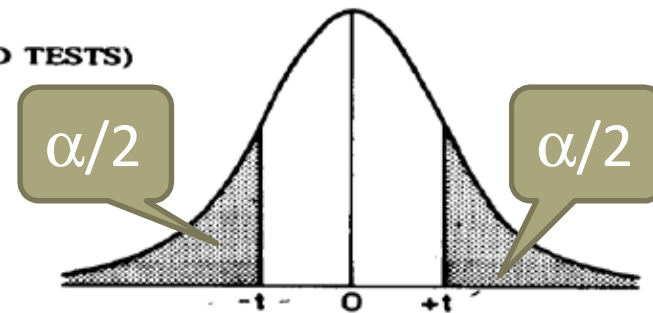


# Análise de Resultados

- Como determinar o Intervalo de Confiança, com menos do que 30 amostras?
  1. Ordenar os valores obtidos
  2. Eliminar os  $\alpha/2$  maiores valores
  3. Eliminar os  $\alpha/2$  menores valores
  4. Obtém-se o intervalo procurado

Ou então... Utiliza-se o Teorema do Limite Central e a Tabela t-Student

**TABLE A 4**  
**DISTRIBUTION OF  $t$  (TWO-TAILED TESTS)**



Degrees of Freedom	Probability of a Larger Value, Sign Ignored								
	0.500	0.400	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	1.000	1.376	3.078	6.314	12.706	25.452	63.657		
2	0.816	1.061	1.886	2.920	4.303	6.205	9.925	14.089	31.598
3	.765	0.978	1.638	2.353	3.182	4.176	5.841	7.453	12.941
4	.741	.941	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	.727	.920	1.476	2.015	2.571	3.163	4.032	4.773	6.859
6	.718	.906	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	.711	.896	1.415	1.895	2.366	2.841	3.499	4.029	5.405
8	.706	.889	1.397	1.860	2.306	2.752	3.355	3.832	5.041
9	.703	.883	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	.700	.879	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	.697	.876	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	.695	.873	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	.694	.870	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	.692	.868	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	.691	.866	1.341	1.753	2.131	2.490	2.947	3.286	4.073
16	.690	.865	1.337	1.746	2.120	2.473	2.921	3.252	4.015
17	.689	.863	1.333	1.740	2.110	2.458	2.898	3.222	3.965
18	.688	.862	1.330	1.734	2.101	2.445	2.878	3.197	3.922
19	.688	.861	1.328	1.729	2.093	2.433	2.861	3.174	3.883
20	.687	.860	1.325	1.725	2.086	2.423	2.845	3.153	3.850

$\alpha$

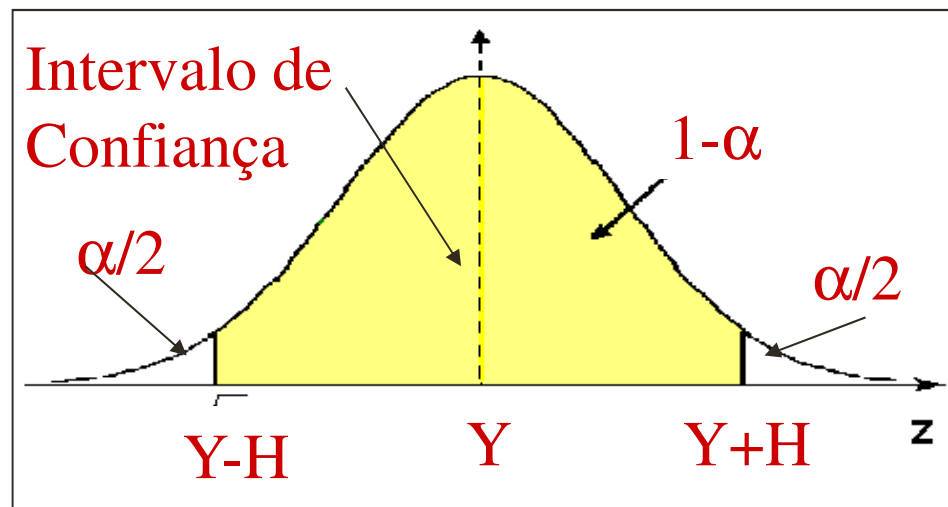
# Análise de Resultados

- Como determinar o Intervalo de Confiança?
  - Média Amostral  $\rightarrow Y_i$  = Média das observações de uma execução
  - Estimativa Global  $\rightarrow Y$  = Média das médias amostrais
  - Variância Amostral  $\rightarrow s^2 = \sum_{i=1}^N (y_i - y)^2 / (n - 1)$
  - Desvio Padrão (DP)  $\rightarrow \sqrt{s^2}$
- $t_{1-\alpha/2, N-1}$   $\rightarrow$  distribuição t-Student com N-1 graus de liberdade e nível de confiança igual a  $1 - \alpha$

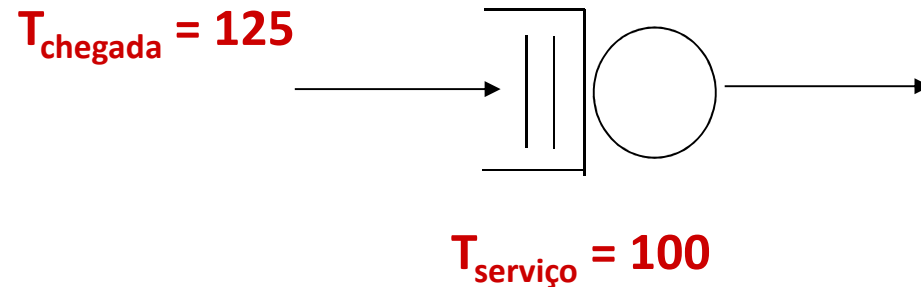
# Análise de Resultados

- Como determinar o Intervalo de Confiança?
  - Largura do Intervalo de Confiança de  $100 * (1-\alpha)\%$
  - $H = t * \text{desvio} / \sqrt{N}$
  - $H = t_{\alpha/2, N-1} * \sqrt{\frac{s^2}{N}}$
  - Intervalo de Confiança:  $Y \pm H$


$$\sqrt{s^2}$$



# Análise de Resultados



- $X_i$ : tempo na fila para o cliente  $i$
- $X$ : tempo médio na fila para 5.000 clientes
- $\mu$  : média real para o tempo na fila ... não conhecido

# Análise de Resultados

- Simulação executada 10 vezes para diferentes conjunto de números aleatórios, obtendo-se  $Y = 406,554$
- Esse valor está suficientemente próximo de  $\mu$ ?

---

331,993	447,532
366,052	420,858
403,524	355,959
464,856	492,144
393,393	389,200

---

# Análise de Resultados

- Simulação executada 10 vezes

$$1-\alpha = 0,95 \rightarrow \alpha = 0,05$$

$$\text{Média} = Y = 406,551$$

$$\text{Variância} = S^2 = \sum_{i=1}^{10} \frac{(331,993 - 406,551)^2 + \dots + \dots}{9} = 2539,082$$

$$\text{DP} = \sqrt{s^2} = \sqrt{2539,082} = 50,39$$

$$t_{0,05/2;9} = 2,26$$

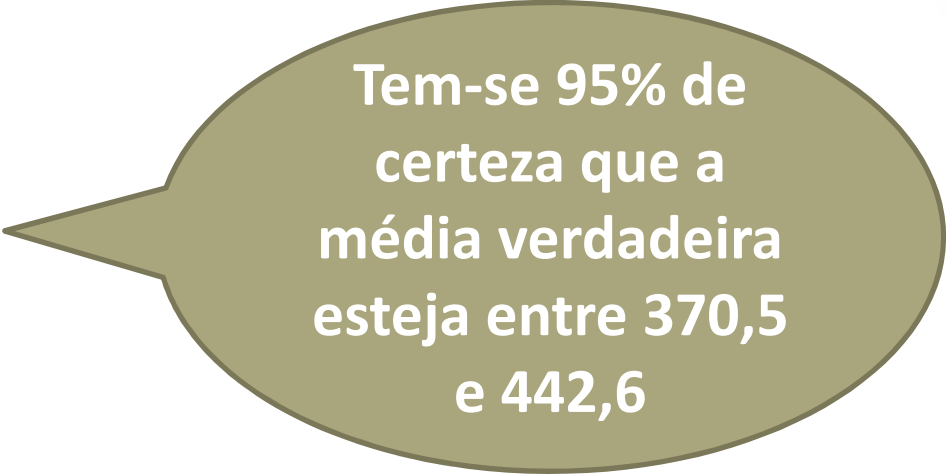
$$H = t_{0,05/2;9} * \text{DP} / \sqrt{N} = 36,04$$

$$\text{IC: } 370,5 - 442,6$$

331,993	447,532
366,052	420,858
403,524	355,959
464,856	492,144
393,393	389,200

# Análise de Resultados

- $H = 72,08$
- IC: 370,5 – 442,6



Tem-se 95% de certeza que a média verdadeira esteja entre 370,5 e 442,6

- Se a simulação for repetida várias vezes e em cada vez for determinado o intervalo de confiança, 95% destes intervalos irão conter a média verdadeira
- Valores individuais podem estar fora do intervalo de confiança



# Análise de Resultados

- Média:  $Y = 406,554$
- $H = 36,04$
- IC:  $370,5 - 442,6$
- Amplitude do intervalo de confiança =  $72,08$ 
  - $17,7\%$  do valor médio
- Não é um valor muito grande?
- Como diminuir?
  - Aumentando o número de replicações

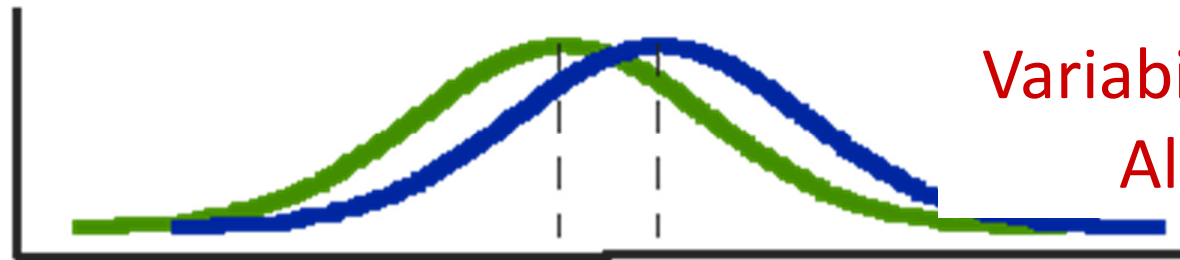
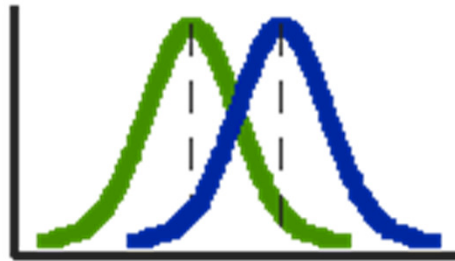
# Análise de Resultados

- Comparação entre dois experimentos
  - Testes estatísticos podem ser utilizados para definir se os resultados provenientes de dois experimentos são conclusivos
  - Definem se os resultados são estatisticamente diferentes

# Análise de Resultados

- Comparação entre dois experimentos: **teste visual**

Variabilidade  
Média



Variabilidade  
Alta

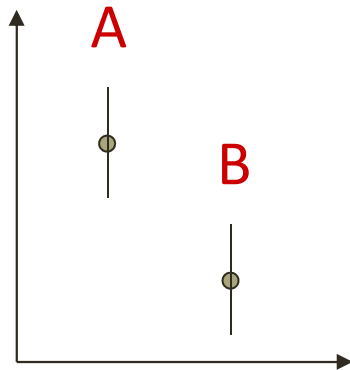
Variabilidade  
Baixa



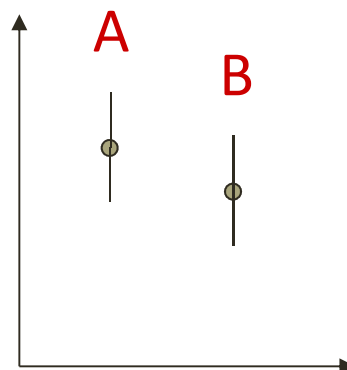
# Análise de Resultados

- Comparação entre dois experimentos: **teste visual**

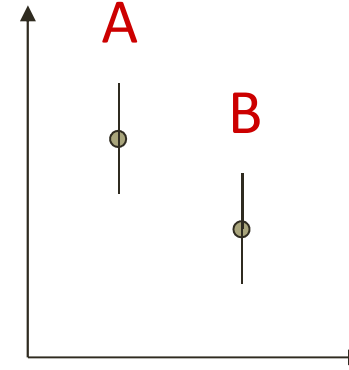
**Caso 1**



**Caso 2**



**Caso 3**



- **Caso 1** – ICs não sobrepostos  $\Rightarrow A > B$
- **Caso 2** – Média de um está inserida no IC do outro  $\Rightarrow A = B$
- **Caso 3** – ICs sobrepostos mas média está fora  $\Rightarrow$  necessário outro teste

# Análise de Resultados

- Comparação entre dois experimentos
  - Área de Estatística oferece um grande número de testes para comparação entre experimentos:
    - Teste t-student – para comparar a média de duas amostras
    - Teste para amostras pareadas
    - Teste para amostras não pareadas
    - Análise de Variância - para comparar média de três ou mais amostras
    - Chi-Quadrado e Poisson - para valores não contínuos