

**Universidade de São Paulo**

**Instituto de Ciências Matemáticas e de Computação**

**Departamento de Sistemas de Computação**

**SSC643 – Avaliação de**

**Desempenho de Sistemas**

**Computacionais**

**Aula 8**

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**
  - Introdução
  - Medidas de Desempenho Frequentemente Utilizadas
  - Análise Estatística dos Resultados
  - Comparação de Resultados
  - Procedimento para análise de resultados

# Conteúdo

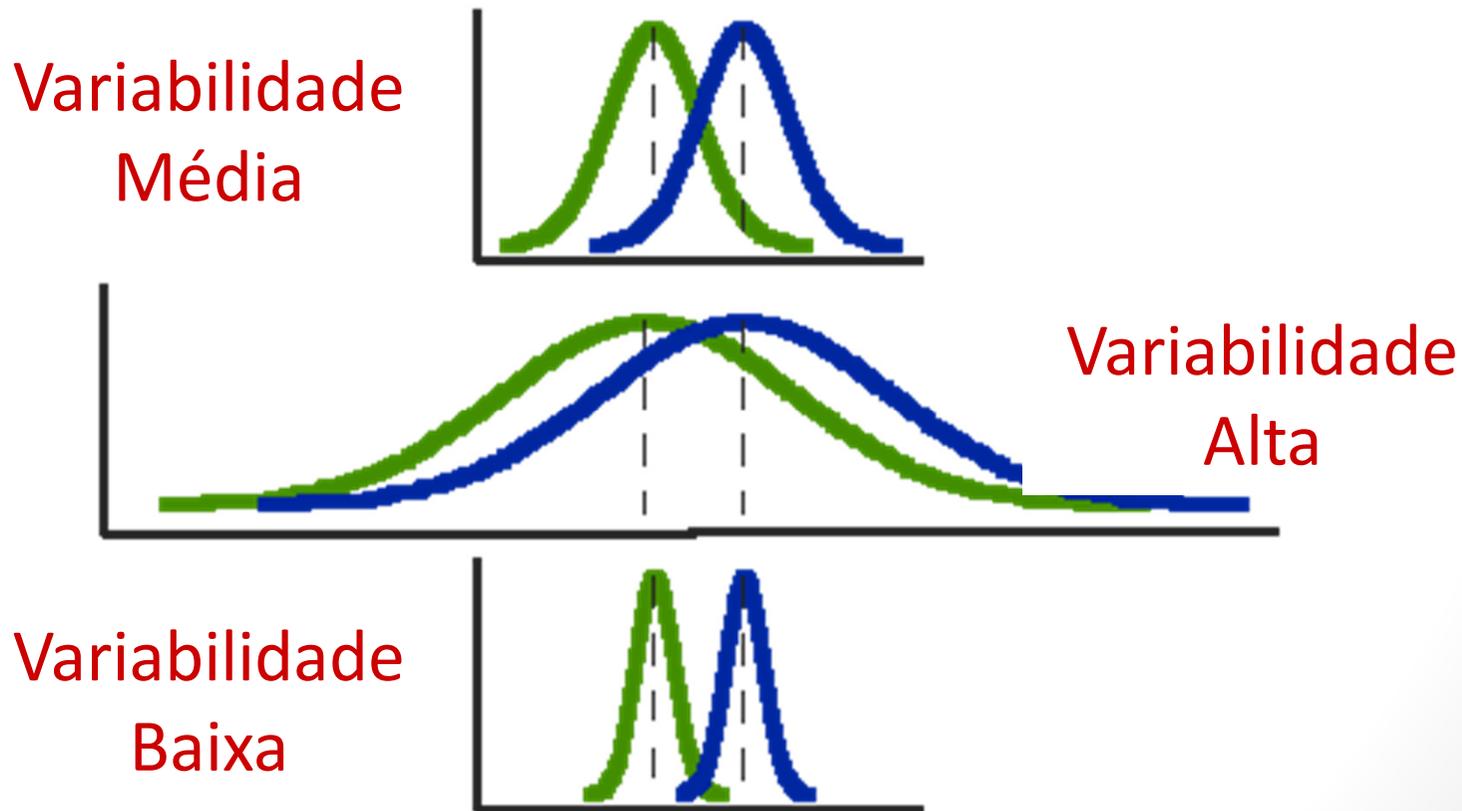
1. Planejamento de Experimentos
2. Técnicas para Avaliação de Desempenho
3. **Análise de Resultados**
  - Introdução
  - Medidas de Desempenho Frequentemente Utilizadas
  - Análise Estatística dos Resultados
  - **Comparação de Resultados**
  - Procedimento para análise de resultados

# 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

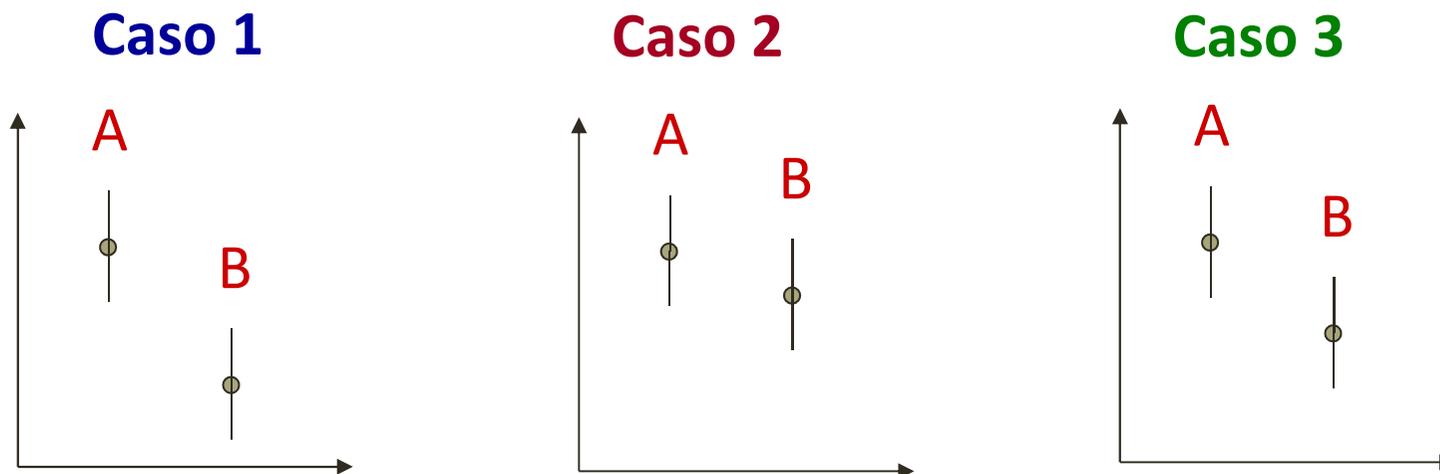
# Análise de Resultados

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



# Análise de Resultados

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



- **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

# Comparação entre dois experimentos

- Área de Estatística oferece grande número de testes para comparação entre experimentos:
  - Teste de hipótese:
    - Teste-t: para comparar a média de duas amostras
    - Análise de variância (ANOVA): para comparar média de três ou mais amostras
  - Teste para amostras pareadas
  - Teste para amostras não pareadas
  - Chi-Quadrado e Poisson - para valores não contínuos

# Teste de hipótese

- Assume-se que uma proposição inicial é verdadeira e então testa essa proposição com os dados amostrais
  - Hipótese nula ( $H_0$ ): proposição inicial, normalmente definida por pesquisas prévias ou pelo senso comum
  - Hipótese alternativa ( $H_1$ ): a que se acredita que pode ser verdadeira
  - Valor de probabilidade (valor-p):
    - Se o valor-p é menor ou igual a um determinado nível de significância (nível  $\alpha$ ), a hipótese nula é rejeitada e a hipótese alternativa é apoiada.
    - Se o valor-p é maior que o nível  $\alpha$ , a hipótese nula não pode ser rejeitada e a hipótese alternativa não tem apoio.

# Teste de hipótese

## Teste-t

- Cálculo do valor de t para a amostra:  $t = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{var_T}{n_T} + \frac{var_C}{n_C}}}$
- Cálculo do número de graus de liberdade ->  
 $n = n_T + n_C - 2$
- Entrar na tabela t-student com  $n$  e confiança desejada:  $t_{tab}$ 
  - Se  $t > t_{tab}$  : médias são diferentes
  - Se  $t < t_{tab}$  : não existe diferença significativa entre as médias

# Análise de Resultados

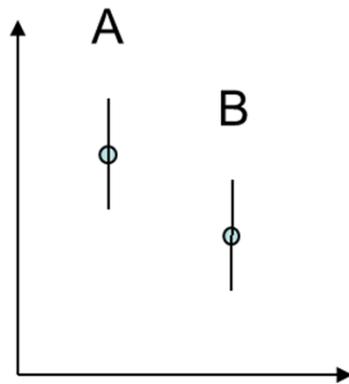
- Após 10 replicações de um programa de simulação, avaliando-se o tempo médio na fila de um recurso, obtiveram-se as médias das amostras e intervalo de confiança para 95% e para 90% representados na tabela a seguir:

	<b>A</b>	<b>B</b>
<b>Média</b>	<b>12</b>	<b>11</b>
<b>H<sub>1</sub> (0,05)</b>	<b>0,8</b>	<b>0,9</b>
<b>H<sub>2</sub> (0,1)</b>	<b>0,4</b>	<b>0,5</b>

- O que se pode concluir?

# Análise de Resultados

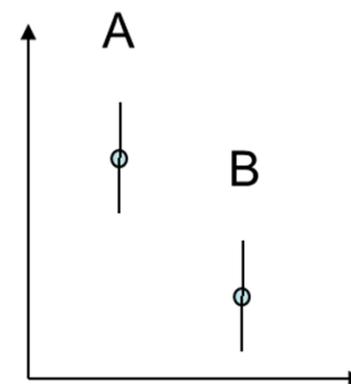
	A	B
<b>Média</b>	<b>12</b>	<b>11</b>
<b>H<sub>1</sub> (0,05)</b>	<b>0,8</b>	<b>0,9</b>
<b>H<sub>2</sub> (0,1)</b>	<b>0,4</b>	<b>0,5</b>



H<sub>1</sub> (0,05)

A - 11,2 - 12,8

B - 10,1 - 11,9



H<sub>2</sub> (0,1)

A - 11,6 - 12,4

B - 10,5 - 11,5

# Análise de Resultados

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

	A	B
Média	12	11
H <sub>1</sub> (0,05)	0,8	0,9
H <sub>2</sub> (0,1)	0,4	0,5

$$\text{Estatística } t_{obs} = \frac{X_A - X_B}{\sqrt{\frac{var_A}{n_A} + \frac{var_B}{n_B}}}$$

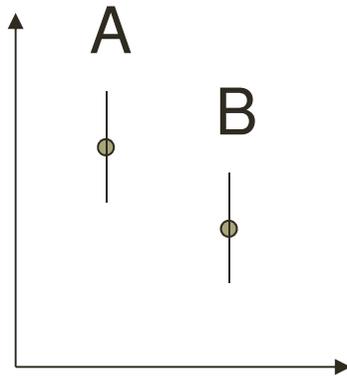
$$var_A = 1,13$$

$$var_B = 1,27$$

$$n = n_A + n_B - 2 = 18$$

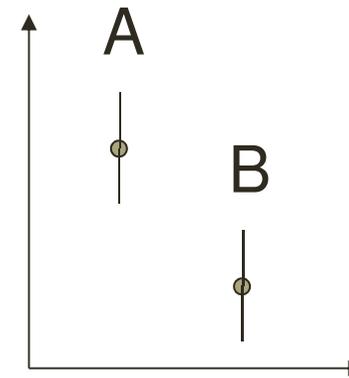
$$t_{obs} = \frac{12 - 11}{\sqrt{\frac{1,13}{10} + \frac{1,27}{10}}} = 2,040$$

# Análise de Resultados



$$H_1 (0,05) \rightarrow t_{\text{obs}} = 2,040 < 2,101 = t_2$$

Não existe diferença significativa

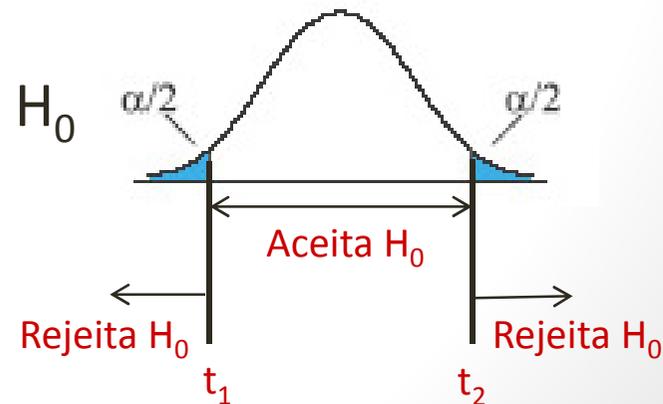


$$H_1 (0,1) \rightarrow t_{\text{obs}} = 2,040 > 1,734$$

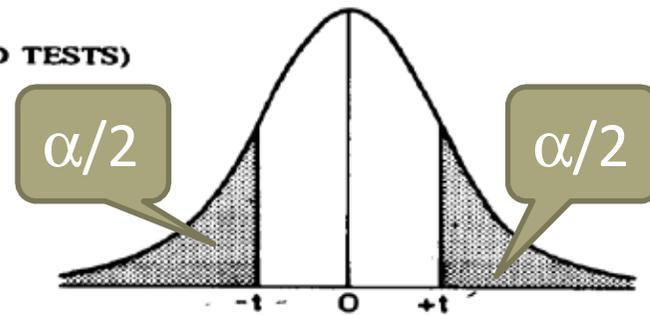
Médias diferentes

Região Crítica (RC): se o  $t_{\text{obs}}$  não pertencer a reunião crítica, aceita-se  $H_0$

$$\{t \in \mathbb{R}; t < t_1 \text{ ou } t > t_2\}$$



**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	2.365	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

$\alpha$

# Teste de hipótese X Intervalo de Confiança

- Teste de hipótese
  - Resposta: aceita ou rejeita a hipótese
  - Conclusivo: não deixa dúvida
  - Não oferece maiores informações
  - Difícil de interpretar
  - O que significa  $t_{obs} = 2,101 > 2,040$ ?

# Teste de hipótese X Intervalo de Confiança

- Intervalo de confiança
  - Informações adicionais
  - Intervalo pequenos -> parâmetro bem estimado
  - Valores com o mesmo significado que as medidas originais -> mais fácil de entender e analisar
  - Significado de Média=12 e  $H(95\%)=0,6$

# Conteúdo

1. Planejamento de Experimentos
2. Técnicas para Avaliação de Desempenho
3. **Análise de Resultados**
  - Introdução
  - Medidas de Desempenho Frequentemente Utilizadas
  - Análise Estatística dos Resultados
  - Comparação de Resultados
  - **Procedimento para análise de resultados**

# Procedimento para análise de resultados

1. Escolher as variáveis de resposta adequadas e suficientes para atingir o objetivo
2. Escolher os fatores e níveis adequadamente
3. Realizar o experimento quantas vezes forem necessárias
4. Pensar na melhor forma de apresentar os dados
5. Fazer o tratamento estatístico adequado para os resultados
6. Observar os resultados e correlaciona-los com o que se conhece do sistema sendo avaliado

# Procedimento para análise de resultados

1. Escolher as variáveis de resposta adequadas e suficientes para atingir o objetivo
  - Não considerar variáveis essenciais pode levar a erros na análise
  - Considerar que variáveis desnecessárias contribuem para aumentar a complexidade da análise
  - Sempre tentar analisar conjuntos não muito grandes de variáveis e, se necessário, realizar a análise em diversas fases

# Procedimento para análise de resultados

2. Escolher os fatores e níveis adequadamente
  - Escolher poucos fatores e, se possível, apenas dois níveis por fator
  - Para os fatores com grande influência nas variáveis de resposta, detalhar processo separadamente.
  - Considerar um grande número de fatores e de níveis em um primeiro momento da avaliação, torna a análise suscetível a erros.

# Procedimento para análise de resultados

3. Realizar o experimento quantas vezes forem necessárias
  - Utilizar um dos métodos apresentados para determinar o ponto de parada de coleta de dados
  - Não tirar conclusões baseando-se em um único resultado

# Procedimento para análise de resultados

4. Pensar na melhor forma de apresentar os dados
  - Tabelas são ótimas para observar detalhes e valores precisos
  - Gráficos são adequados para melhor visualizar os resultados
  - Nos gráficos, cuidado com escalas e origem dos eixos
  - Muitos valores em uma tabela ou em um gráfico tornam a análise mais complexa, e possíveis resultados mais difíceis de serem identificados

# Procedimento para análise de resultados

5. Fazer o tratamento estatístico adequado para os resultados
  - Não tirar conclusões considerando-se apenas médias
  - Valores médios só fazem sentido quando acompanhados de desvio padrão, variância, intervalo de confiança, etc.
  - Valores médios com máximo e mínimo podem ajudar na análise mas não levam a resultados conclusivos, sem a presença de uma métrica que indique a dispersão dos dados

# Procedimento para análise de resultados

6. Observar os resultados e correlaciona-los com o que se conhece do sistema sendo avaliado
  - Desconfie de resultados não esperados
  - Tente relacionar os diferentes resultados obtidos
  - Tente explicar os resultados obtidos