

# Sistemas de Informação



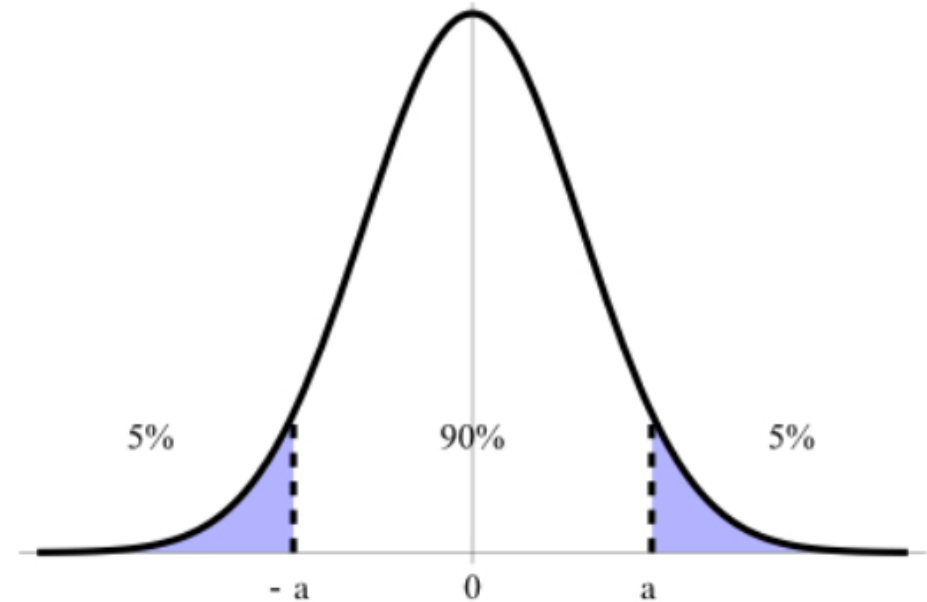
## Planejamento de Capacidade e Avaliação de Sistemas Computacionais

### Aula 4: Intervalo de Confiança

Prof. Fábio Leandro Rodrigues Cordeiro, Me.

# Sumário

- Geração de Resultados Confiáveis
- Comparação de Alternativas
- Intervalo de Confiança
- Determinando o Tamanho da Amostra
- Referências



# Comparando alternativas

## Conceito

- Uma premissa é que sempre estamos comparando alternativas diferentes, sejam elas algoritmos, modelos, implementações etc.
- Geralmente um computador (sistema complexo) está sujeito a todo tipo de variação. Portanto, se utilizamos experimentação ou qualquer modelo matemático ou de simulação que tente ser preciso, cada alternativa deve ser testada diversas vezes, para minimizar as variações.

# Exemplo da Cache

## Exemplo anterior

- No nosso exemplo de cache, não existe variação, pois a simulação testa a cache isoladamente de forma bastante simplificada.
- Mas se desejássemos simular a cache considerando todo o processador, a hierarquia de memória e o sistema operacional com os seus vários processos?

# Exemplo da Cache

## Exemplo 1

- Variações e interferências possíveis:
  - A cache seria esvaziada a cada novo processo que fosse escalonado.
  - Existiriam duas caches separadas, uma para dados e outra para instruções.
  - A latência e largura de banda dos barramentos seriam fundamentais para o tempo de acesso a cache.
  - Os barramentos poderiam ser compartilhados, poderia ou não existir um controlador de DMA etc.
  - Tudo isso afetaria o tempo de execução de um programa na memória cache.

- Como podemos garantir que nenhuma interferência externa vai atrapalhar os nossos resultados?
  - Simples: testamos a mesma alternativa várias vezes para uma mesma configuração.
  - Argumento (Exemplo)
    - Se em todas as vezes que testarmos a alternativa, ocorrer uma falta de página, então ela deixou de ser uma interferência e passou a ser relevante e deve ser considerada.
    - Senão, mesmo que ela tenha ocorrido uma ou outra vez por acaso, a sua influência seria minimizada considerando todos os repetidos testes.

# Como evitar interferências não desejáveis



- Em um simulador, onde temos total controle, permitimos apenas as variações desejadas, que fazem parte do experimento.
- Em um sistema real, onde temos baixo controle, podemos minimizar o problema desativando todos os hardwares e softwares desnecessários.
  - Ninguém deseja que no meio da execução de um programa importante, a impressora mostre uma mensagem de falta de papel.

# Intervalo de Confiança

## IC

- Ao testarmos uma mesma alternativa várias vezes para uma mesma configuração, podemos calcular o seu intervalo de confiança (IC).
- Porque utilizamos um IC?
  - Sumarizar o erro na média da amostra
  - Prover elementos para saber se a amostra é significativa
  - Permitir comparações à luz dos erros



# Intervalo de Confiança

## IC

- O intervalo de confiança é dado pelas seguintes fórmulas:
  - Limite superior:  $M + d(nc)*s$
  - Limite inferior:  $M - d(nc)*s$
  - Onde:
    - $M$  = média da amostra
    - $d$  = distribuição
      - T student ( $< 30$  elementos) e Z normal ( $\geq 30$  elementos)
    - $Nc$  = nível de confiança
    - $S$  = erro padrão = desvio padrão/raiz(n. elementos)

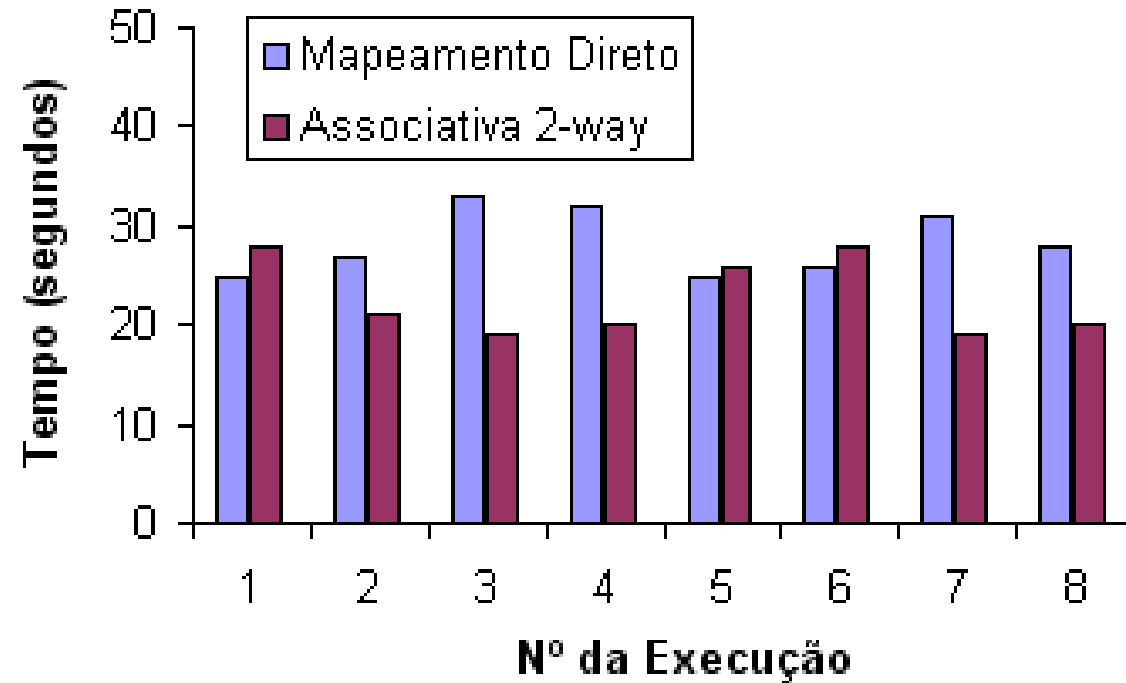
# Intervalo de Confiança

## Exemplo 1

- Suponha que desejamos comparar a eficiência de duas organizações de cache para um mesmo programa, considerando todos os aspectos do computador.
  - Métrica: Tempo em segundos
  - Organizações: Mapeamento Direto e Associativa 2-way
  - Número de Execuções (repetições): 8

# Resultados obtidos

## Exemplo 1



	1	2	3	4	5	6	7	8
MD	25	27	33	32	25	26	31	28
2-way	28	21	19	20	26	28	19	20

# Calculando Média e Desvio Padrão

## Exemplo 1

- Como dizer que uma é melhor que a outra, já que cada hora deu um resultado diferente?
  - Precisamos calcular a média, o desvio padrão e finalmente o intervalo de confiança

	Média	Desvio
MD	28,375	2,997395
2-way	22,625	3,739569

- **Mas e o Intervalo de confiança? Como calculá-lo?**

# Calculando os Intervalos de Confiança

## Exemplo 1

- A distribuição é lida de uma tabela padrão.
- Então o cálculo fica assim, considerando um nível de confiança de 90%:
  - Para MD:  $28,375 \pm t[0.9;7] \cdot 2,99/\sqrt{8}$   
 $IC = (26,87 < x < 29,87)$
  - Para 2-way:  $22,625 \pm t[0.9;7] \cdot 3,73/\sqrt{8}$   
 $IC = (20,75 < x < 24,49)$

**Obs:** i) distribuição t usa nível de confiança, e  $n - 1$  elementos;  
 $t[0.9;7] = 1.415$

# Calculando os Intervalos de Confiança

## Exemplo 1

- O que esses intervalos significam?
  - Que com 90% de certeza, o tempo médio do programa para a organização mapeamento direto está entre 26,87 e 29,87 segundos.
  - Mas quem é melhor? MD ou 2-way? Temos a impressão que é o 2-way, mas ainda precisamos calcular o intervalo de confiança da diferença dos dois para ter certeza.

# Calculando as Diferenças

## Exemplo 1

- Basta subtrair uma linha pela outra.

	1	2	3	4	5	6	7	8
MD	25	27	33	32	25	26	31	28
2-way	28	21	19	20	26	28	19	20
Diferença	-3	6	14	12	-1	-2	12	8

- Média = 5,75 e Desvio = 6,45
- Intervalo de Confiança =  $(1,27 < x < 10,22)$

## Exemplo 1

- O que significa o IC das diferenças?
  - Se o intervalo contém zero, então as soluções não são diferentes, ou seja, precisa de mais testes para diferenciá-las.
  - Se o intervalo é positivo, neste exemplo, a segunda solução é a melhor.
  - Se o intervalo é negativo, neste exemplo, a primeira solução é a melhor.



## Exemplo 1

- Neste exemplo, com 90% de confiança podemos afirmar que a solução com cache 2-way é melhor que mapeamento direto para o programa em questão.
- A maioria dos trabalhos na área estão sujeitos a variação e necessitam do cálculo dos intervalos de confiança.

# Dicas Gerais

- Quantos testes fazer?
  - Comece com pouco  $< 10$  e depois vai incrementando caso você não consiga diferenciar as soluções.
- E se tiver que comparar mais de dois?
  - As comparações são pareadas, logo você fará todas as combinações dois a dois.
- Qual o nível de confiança normalmente utilizado?
  - Na área, o mínimo é 90%, mas os mais comuns são 95% e 99%.
- Como diminuir o intervalo de confiança?
  - Realizar mais testes

# Como escolher o tamanho da amostra?

- Para uma exatidão de  $\pm r\%$ , use a seguinte fórmula:

$$n = \left( \frac{100zs}{r\bar{x}} \right)^2$$

- Onde  $z$  é a distribuição normal ou  $t$ ;  $s$  é o desvio padrão;  $\bar{x}$  é a média
- $r$  é a exatidão, por exemplo, a média de uma amostra deu 10 e o desvio padrão 2. Se você quer uma variação de apenas 1 segundo, a exatidão seria de  $1/10 = 10\%$ . Ou seja,  $r$  controla o tamanho do intervalo.

# Como escolher o tamanho da amostra?

## Exemplo

- Suponha cinco execuções de um programa que gastaram: 22.5, 19.8, 21.1, 26.7 e 20.2 segundos
- Quantas execuções devem ser executadas para obter  $\pm 5\%$  de exatidão e num nível de confiança de 90%?  $\bar{x} = 22.1$ ,  $s = 2.8$ ,  $t_{0.95;4} = 2.132$

$$n = \left( \frac{(100)(2.132)(2.8)}{(5)(22.1)} \right)^2 = 5.4^2 = 29.2$$

# Como escolher o tamanho da amostra?

## Exemplo

- Para calcular o  $r$ , geralmente analisamos o desvio padrão da amostra
- Caso ele esteja muito alto, calculamos o  $r$  em função do desvio desejado (não quer dizer que ele será alcançado)
- Se  $s = 10.0$  e  $x = 20.0$  (para uma amostra de tamanho 10) e desejamos que o desvio seja  $s = 5.0$ , então  $r = 5/20 = 25\%$ . Pois atualmente ele é de 50%.

## Exercício

- Baseado em um teste preliminar, a média do tempo de resposta de um programa é 20s e o erro padrão é 5. Quantas repetições são necessárias para conseguir um exatidão de 5% com confiança de 95%?
- Onde  $z_{95\%} = 1.960$

## Exercício

- Seis cargas de trabalho similares foram executadas em dois sistemas. As observações foram:  $\{(5.4, 19.1), (16.6, 3.5), (0.6, 3.4), (1.4, 2.5), (0.6, 3.6), (7.3, 1.7)\}$ .
- Qual sistema é o melhor?

Table of Probabilities for Student's t-Distribution								
df	0.600	0.700	0.800	0.900	0.950	0.975	0.990	0.995
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617



# Referências

- Jain, R. K., “The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling”, John Wiley & Sons, 1991.
- Material da disciplina Métodos Quantitativos do professor Virgílio Almeida