IND 1115
Inferência Estatística
Aula 12

Novembro 2004
Revista em Novembro de 2005

Mônica Barros

Conteúdo

- Intervalos de Confiança – Motivação
- Intervalos de Confiança para Médias

Intervalos de Confiança

- Até agora estivemos interessados em encontrar uma estimativa pontual para um parâmetro desconhecido θ.
- Também enumeramos algumas propriedades desejáveis de estimadores pontuais.
- Agora tentaremos obter não apenas uma estimativa pontual, mas um intervalo que contenha o parâmetro de interesse com uma probabilidade especificada. Este intervalo será chamado de "Intervalo de Confiança".

Intervalos de Confiança

O intervalo de confiança 100(1-α)% para θ é dado por:

$$L(\underset{\sim}{X}) \le \theta \le U(\underset{\sim}{X})$$

Onde $L(\underset{\sim}{X})$ (limite inferior) e $U(\underset{\sim}{X})$ (limite superior) são tais que:

$$\mathrm{Prob}[L(\underset{\sim}{X}) \le \theta \le U(\underset{\sim}{X})] = 1-\alpha$$

Onde α é um número especificado pelo usuário.

Intervalos de Confiança

- Note que o intervalo $\left[L\left(\underset{\sim}{X}\right), U\left(\underset{\sim}{X}\right)\right]$ é aleatório, e a cada amostra obtida iremos encontrar valores diferentes para os limites L e U.

- A notação $\underset{\sim}{X}$ indica todos os elementos da amostra aleatória, isto é:

$$\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$$

Intervalos de Confiança – Média da Normal

- Consideraremos agora o caso mais comum na prática onde os dados são supostos NORMAIS e θ é média da distribuição.

- Serão abordados dois casos: variância do modelo conhecida e variância do modelo desconhecida.

Intervalos de Confiança – Média da Normal

- Argumento intuitivo....
- Suponha que você tem uma amostra aleatória da Normal, em que a média é desconhecida.

- Se você precisasse achar um estimador pontual de θ (a média), usaria a média amostral $\overline{X}$.

Intervalos de Confiança – Média da Normal

- E se agora você precisar encontrar um intervalo que contenha θ com uma probabilidade especificada?

- Parece natural que este intervalo tenha a forma: $\left(\overline{X} - c, \overline{X} + c\right)$ onde c é uma constante a ser especificada. Veremos que os intervalos encontrados para a média da Normal têm exatamente esta dorma!

Intervalo de Confiança – Média da Normal

Caso I

X~ NORMAL(θ,σ^2); σ^2 conhecido

- Seja $\underset{\sim}{X} = (X_1, \ldots, X_n)$ uma a.a. de tamanho n da distribuição Normal acima.
- Já vimos que $\bar{X} = \sum \frac{X_i}{n}$ é o estimador de máxima verossimilhança de θ. Além disto, é fácil provar que:

$$\bar{X} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

Intervalo de Confiança – Média da Normal

- Logo, podemos padronizar a média amostral, transformando-a numa v.a. com densidade N(0,1) da seguinte maneira:

$$Z = \frac{\bar{X} - \theta}{\sigma / \sqrt{n}} = \frac{\sqrt{n}\left(\bar{X} - \theta\right)}{\sigma} \sim N(0,1)$$

- Usando uma tabela da Normal podemos encontrar, por exemplo, a probabilidade desta nova variável estar entre -2 e +2.

Intervalo de Confiança – Média da Normal

Prob (-2 < Z < 2) = Φ(2) - Φ(-2)= 0.954

- Substituindo Z na expressão anterior leva a:

$$-2 < \frac{\bar{X} - \theta}{\sigma / \sqrt{n}} < +2 \Leftrightarrow \bar{X} - \frac{2\sigma}{\sqrt{n}} < \theta < \bar{X} + \frac{2\sigma}{\sqrt{n}}$$

- Daí:

$$\Pr ob\{-2 < Z < +2\} = \Pr ob\left\{\bar{X} - \frac{2\sigma}{\sqrt{n}} < \theta < \bar{X} + \frac{2\sigma}{\sqrt{n}}\right\} = 0.954$$

- O intervalo que acabamos de encontrar é um intervalo de confiança 95.4% para θ.

Intervalo de Confiança – Média da Normal

- Ou seja, na notação mostrada antes:

$$L(\underset{\sim}{X}) = \bar{X} - \frac{2\sigma}{\sqrt{n}}$$

$$U(\underset{\sim}{X}) = \bar{X} + \frac{2\sigma}{\sqrt{n}}$$

$$1 - \alpha = 0.954$$

- A seguir exibimos uma "receita de bolo" para obter o IC da média de uma Normal com variância conhecida.

Intervalo de Confiança – Média da Normal

- **Receita de Bolo**
- Seja $\underset{\sim}{X} = (X_1, \ldots\ldots, X_n)$ uma a.a. de tamanho n da distribuição Normal com média desconhecida θ e variância conhecida σ^2.
- Um intervalo de confiança 100(1- α)% para θ é dado por: $[\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$
- Onde $z_{1-\alpha/2}$ é obtido da função de distribuição N(0,1) e é tal que $\Pr(Z < z_{1-\alpha/2}) = 1-\alpha/2$.

Intervalo de Confiança – Média da Normal

- Note que, pela simetria em torno de zero da distribuição N(0,1):
- $z_{1-\alpha/2}$ é o ponto tal que, a probabilidade de estar ACIMA dele é $\alpha/2$ usando uma distribuição N(0,1).
- Também é fácil perceber que, se Z é N(0,1):

$$\Pr\left\{ -z_{1-\frac{\alpha}{2}} < Z < +z_{1-\frac{\alpha}{2}} \right\} = 1-\alpha$$

- E esta última expressão foi empregada para obter o IC para a média.

IC para a média da Normal com σ conhecido

- Exemplo
- Considere a população de alunos da PUC. Para uma amostra de 50 alunos obtivemos uma altura média de 1,68m.
- Sabe-se que o desvio-padrão da altura da população de alunos da PUC é o mesmo que o da população de jovens cariocas com menos de 25 anos: 0,11m.
- Suponha que as alturas dos alunos são Normalmente distribuídas.
- Determine, com um nível de confiança de 95%, o intervalo onde a real altura média da população de alunos da PUC deve estar localizada.

IC para a média da Normal com σ conhecido

- Solução
  - Note que a amostra é Normal com variância conhecida, e assim a distribuição de $\overline{X}$ também é Normal.
  - Da tabela da Normal, ou usando a função INV.NORMP do Excel, procuramos um valor z_0 tal que $\Pr(Z < z_0) = 1- \alpha/2 = 97.5\%$, isto é, $\Phi(z_0) = 97.5\%$. A função INV.NORMP fornece $z_0 = 1.96$.

IC para a média da Normal com σ conhecido

- Solução
- O IC 95% (para as alturas em cm) é então:

$$\left(\overline{X} - z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}\right) = \left(168 - 1.96\frac{11}{\sqrt{50}}, 168 + 1.96\frac{11}{\sqrt{50}}\right)$$

$$= (164.95 \text{ cm},\ 171.05 \text{ cm})$$

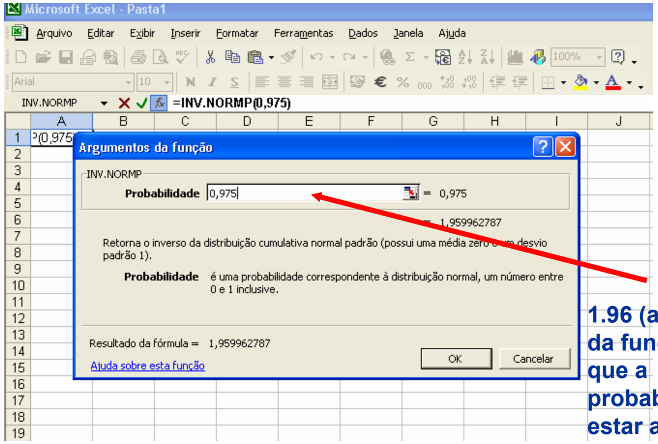
IC para a média da Normal com σ conhecido

- Receita de bolo – qual valor de $z_{\alpha/2}$ usar?

Coeficiente de Confiança	valor tabelado de z
80.0%	1.282
90.0%	1.645
95.0%	1.960
97.0%	2.170
97.5%	2.241
99.0%	2.576

Estes pontos são encontrados através da função INV.NORMP do Excel – Note que, se o coeficiente de confiança é 1- α, devemos buscar um ponto na tabela da Normal tal que a probabilidade de estar ACIMA dele é α/2, ou seja, a probabilidade de estar ABAIXO dele é 1 - α/2 (o argumento da função INV.NORMP é 1 - α/2).

IC para a média da Normal com σ conhecido



1.96 (a "resposta da função" é tal que a probabilidade de estar abaixo deste valor é 0,975

IC para a média da Normal com σ conhecido

- Exemplo
- Numa amostra de 36 postos de gasolina no Rio de Janeiro, o preço médio do litro da gasolina aditivada foi de R$ 1.78. Sabe-se, por experiências anteriores, que o desvio padrão é R$ 0.20.
- Encontre intervalos de confiança 90%, 95% e 99% para o preço médio da gasolina aditivada no Rio de Janeiro supondo que a amostra é Normal.
- Solução
- Aqui estamos supondo que o desvio padrão é conhecido, e assim podemos usar um intervalo baseado na densidade Normal.

IC para a média da Normal com σ conhecido

- Os IC têm a forma geral: $\left(\overline{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$
- O IC 90% é: $\left(1.78 - 1.645\frac{(0.20)}{6}, 1.78 + 1.645\frac{(0.20)}{6}\right) = (\text{R\$ }1.725, \text{R\$ }1.835)$
- O IC 95% é: $\left(1.78 - 1.96\frac{(0.20)}{6}, 1.78 + 1.96\frac{(0.20)}{6}\right) = (\text{R\$}1.715, \text{R\$}1.845)$
- O IC 99% é: $\left(1.78 - 2.576\frac{(0.20)}{6}, 1.78 + 2.576\frac{(0.20)}{6}\right) = (\text{R\$}1.694, \text{R\$}1.866)$

Note que, à medida que o coeficiente de confiança aumenta, a largura do intervalo também aumenta!

IC para a média da Normal com σ conhecido

- Exemplo (para casa)
- O preço médio de um automóvel Palio ELX 1.0 4 portas ano 2001 é R$ 17727 (segundo o Jornal Valor Econômico de 07/07/2003).
- Suponha que o desvio padrão REAL dos preços seja R$ 1500 e o tamanho da amostra é n = 25 carros.
- Encontre intervalos de confiança 95% e 99% para os preços de Palios ELX 1.0 quatro portas ano 2001 supondo que os preços são Normalmente distribuídos.

IC para a média da Normal com σ conhecido

- Exemplo (para casa)
- Toma-se uma amostra de 25 usuário de um cartão de crédito e observa-se que o gasto médio mensal é R$ 600.
- O desvio padrão é conhecido e igual a R$ 250.
- Encontre intervalos de confiança 95 e 99% para o gasto médio com cartão na população de usuários.

PIVOT

- Seja $\underset{\sim}{X} = (X_1, \ldots\ldots, X_n)$ uma a.a. de tamanho n de uma densidade (ou função de probabilidade) f(x,θ).
- Seja Q = q(X_1,....,X_n, θ) uma função dos elementos da amostra e do parâmetro desconhecido θ.
- **Q é chamado de PIVOT se sua distribuição não depende de θ.**
- Um PIVOT é usado para encontrar intervalos de confiança para parâmetros desconhecidos.

PIVOT

- No exemplo do IC da média da Normal com variância conhecida, a quantidade:

$$\boxed{Z = \frac{\sqrt{n}\left(\overline{X} - \theta\right)}{\sigma}}$$

- é um PIVOT, pois depende de $\underset{\sim}{X} = (X_1, \ldots\ldots, X_n)$ e θ, sua distribuição não depende de θ (pois é N(0,1)) e assim pode ser usada na construção de um IC para θ.

IC para a média da Normal com σ desconhecido

Caso II

X ~ NORMAL(θ, σ²); σ² DESCONHECIDO

- Seja $\underset{\sim}{X} = (X_1, \ldots\ldots, X_n)$ uma a.a. de tamanho n da distribuição Normal acima.
- Os estimadores **não tendenciosos** de θ e σ² são: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ e $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

$$\text{onde } \overline{X} \sim N\left(\theta, \frac{\sigma^2}{n}\right) \text{ e } \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

IC para a média da Normal com σ desconhecido

- Também, $\overline{X}$ e S^2 são independentes.
- Pela definição de uma v.a. t de Student:

$$T = \frac{\dfrac{\sqrt{n}(\overline{X} - \theta)}{\sigma}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}} = \sqrt{n}.\frac{\overline{X} - \theta}{S} \sim t_{n-1}$$

- Onde: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$
- Assim da tabela da distribuição t de Student com n-1 graus de liberdade podemos obter dois números <u>a</u> e <u>b</u> tais que: Pr(a < T < b) = 1- α

IC para a média da Normal com σ desconhecido

- Para encontrar um intervalo simétrico fazemos a = -b e assim:

$$\text{Prob}[a < T < b] = \text{Prob}\{-b < \text{T} < +b\} = \text{Prob}\left(-b < \sqrt{n}\left(\frac{\overline{X} - \theta}{S}\right) < b\right) = 1 - \alpha$$

$$\Leftrightarrow \text{Prob}\left(-b\frac{S}{\sqrt{n}} < \overline{X} - \theta < +b\frac{S}{\sqrt{n}}\right) =$$

$$= \text{Prob}\left(-\overline{X} - b\frac{S}{\sqrt{n}} < -\theta < -\overline{X} + b\frac{S}{\sqrt{n}}\right) =$$

$$= \text{Prob}\left(\overline{X} - b\frac{S}{\sqrt{n}} < \theta < \overline{X} + b\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

IC para a média da Normal com σ desconhecido

- **Portanto:**
- **O intervalo** $\left(\bar{X} - b\frac{S}{\sqrt{n}}, \bar{X} + b\frac{S}{\sqrt{n}}\right)$
- **é um intervalo aleatório com probabilidade 1- α de incluir o parâmetro desconhecido θ.**
- **O ponto <u>b</u> que aparece na definição do IC é obtido da distribuição t com n-1 graus de liberdade, e é tal que Pr(T > b) = α/2.**

IC para a média da Normal com σ desconhecido

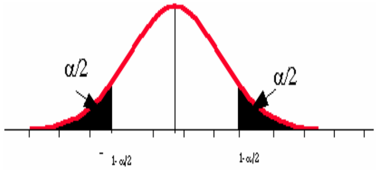
- **Receita de Bolo**
- **Seja $X_1, X_2, ..., X_n$ uma a.a. de tamanho n da distribuição Normal com média desconhecida θ e variância *desconhecida* σ^2.**
- **Um intervalo de confiança 100(1- α)% para θ é dado por:** $\left(\bar{X} - b\frac{S}{\sqrt{n}}, \bar{X} + b\frac{S}{\sqrt{n}}\right)$
- **Onde <u>b</u> é obtido da função de distribuição t de Student com n-1 graus de liberdade e é tal que Pr(T > b) = α/2.**

IC para a média da Normal com σ desconhecido

- **O IC 100(1-α)% para θ é:**

$$\left(\bar{X} - t_{n-1;1-\alpha/2}\frac{S}{\sqrt{n}}, \bar{X} + t_{n-1;1-\alpha/2}\frac{S}{\sqrt{n}}\right)$$

- **Onde S é o desvio padrão amostral e $t_{n-1,1-\alpha/2}$ é um ponto da distribuição t de Student com n-1 graus de liberdade tal que Pr(T > $t_{n-1,1-\alpha/2}$) = α/2, como no gráfico a seguir:**



IC para a média da Normal com σ desconhecido

- **O valor $t_{n-1,1-\alpha/2}$ é obtido de uma tabela da distribuição t com n-1 graus de liberdade. Pode-se, alternativamente, usar a função INVT do Excel.**

IC para a média da Normal com σ desconhecido

- **Exemplo**
- **Numa amostra de 16 postos de gasolina no Rio de Janeiro, o preço médio do litro da gasolina aditivada foi de R$ 1.78.**
- **O desvio padrão dos preços estimado na amostra é R$ 0.20. Encontre intervalos de confiança 90%, 95% e 99% para o preço médio da gasolina aditivada no Rio de Janeiro e compare-os com os encontrados no exemplo da página 18.**

IC para a média da Normal com σ desconhecido

- **Solução**
- **Aqui deve-se usar a distribuição t para encontrar o IC, pois o desvio padrão é desconhecido. A forma do intervalo é:**

$$IC = \overline{X} \pm t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} = \left(\overline{X} - t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \right)$$

- **Pela função INVT do Excel com 15 graus de liberdade obtemos os pontos percentuais para os IC 90, 95 e 99%, que são, respectivamente: 1.753, 2.131 e 2.947.**

IC para a média da Normal com σ desconhecido

- **O IC 90% é:** $\left(1.78 - 1.753\frac{(0.20)}{\sqrt{16}}, 1.78 + 1.753\frac{(0.20)}{\sqrt{16}}\right) = (\text{R\$ } 1.692,\ \text{R\$ } 1.868)$
- **O IC 95% é:** $\left(1.78 - 2.131\frac{(0.20)}{\sqrt{16}}, 1.78 + 2.131\frac{(0.20)}{\sqrt{16}}\right) = (\text{R\$ } 1.673,\ \text{R\$ } 1.887)$
- **O IC 99% é:** $\left(1.78 - 2.947\frac{(0.20)}{\sqrt{16}}, 1.78 + 2.947\frac{(0.20)}{\sqrt{16}}\right) = (\text{R\$ } 1.633,\ \text{R\$ } 1.927)$

Note que os intervalos de confiança são mais largos que os correspondentes para a Normal

Nota IMPORTANTE – uso de INVT no Excel

- **Suponha que você quer encontrar um intervalo de confiança 100*(1 – α)%.**
  - **Então para obter o ponto t $_{1-\alpha/2}$ que entra no cálculo do IC, use a função INVT com os argumentos:**
    - **α e**
    - **n – 1 graus de liberdade**
    - **Pois a função INVT do Excel fornece a o ponto tal que a probabilidade de estar ACIMA dele é especificada.**
  - **Isso se deve ao fato do primeiro argumento da função no Excel ser, na verdade, o valor para o intervalo bilateral.**

Utilizando o Excel

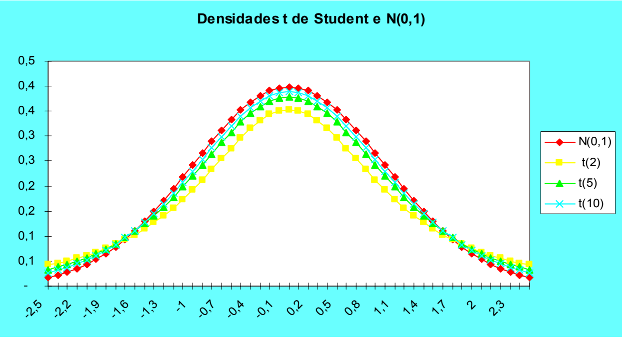
- Funções do Excel para a distribuição t

Função	Descrição
invt(p; gl)	Para a distribuição t de Student, calcula o valor t para **p = 2.α**, com gl graus de liberdade

- Por exemplo, INVT(0.05, 20) = 2.086 calcula o valor na tabela t com 20 graus de liberdade e é tal que Pr(T > 2.086) = 0.05/2 = 0.025

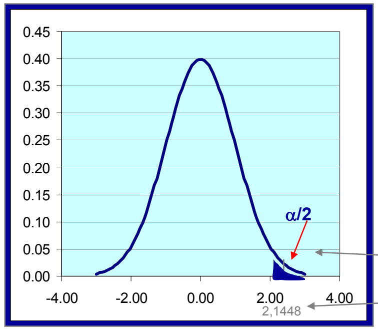
Distribuição t de Student

- Quando <u>n</u> (número de graus de liberdade) cresce, a densidade t de Student se torna cada vez mais parecida com uma N(0,1)



A distribuição t de Student

- Exemplo: para uma amostra com 15 elementos (14 graus de liberdade) e para um nível de confiança de 5% ($\alpha/2$ = 0,025), t é igual a 2,1448



G.L	0.100	0.075	0.050	0.025	0.020
1	3.0777	4.1653	6.3137	12.7062	15.8945
2	1.8856	2.2819	2.9200	4.3027	4.8487
3	1.6377	1.9243	2.3534	3.1824	3.4819
4	1.5332	1.7782	2.1318	2.7765	2.9985
5	1.4759	1.6994	2.0150	2.5706	2.7565
6	1.4398	1.6502	1.9432	2.4469	2.6122
7	1.4149	1.6166	1.8946	2.3646	2.5168
8	1.3968	1.5922	1.8595	2.3060	2.4490
9	1.3830	1.5737	1.8331	2.2622	2.3984
10	1.3722	1.5592	1.8125	2.2281	2.3593
11	1.3634	1.5476	1.7959	2.2010	2.3281
12	1.3562	1.5380	1.7823	2.1788	2.3027
13	1.3502	1.5299	1.7709	2.1604	2.2816
14	1.3450	1.5231	1.7613	2.1448	2.2638
15	1.3406	1.5172	1.7531	2.1315	2.2485
16	1.3368	1.5121	1.7459	2.1199	2.2354

Comparação: IC Normais x IC t de Student

- A distribuição t nos fornece intervalos de comprimento maior que os intervalos Normais com a mesma probabilidade.

- À medida que o número de graus de liberdade da densidade t cresce, a densidade se torna mais e mais parecida com uma N(0,1), e conseqüentemente, os intervalos se tornam mais próximos dos encontrados através da distribuição N(0,1).

Comparação: IC Normais x IC t de Student

- Também, o comprimento dos intervalos diminui à medida que aumentamos o número de observações.

- Isto é intuitivamente razoável, pois à medida que o tamanho da amostra cresce, $\overline{X}$ "converge" para μ e temos cada vez mais "certeza" de que a média amostral está num intervalo de pequeno comprimento em torno de μ com alta probabilidade (este resultado é conhecido como "lei dos grandes números").

Utilizando o Excel

- O Excel também pode ser utilizado para o cálculo do intervalo de confiança para σ desconhecido (para qualquer tamanho de amostra)
  - Selecione no menu Ferramentas a opção Análise de Dados;
  - Escolha a opção Estatística Descritiva;
  - Na caixa Intervalo de Entrada, selecione os dados da amostra;
  - Selecione a opção Intervalo de Confiança para a Média e coloque o intervalo de confiança desejado;
  - Na caixa Intervalo de Saída, selecione o local da planilha onde os resultados serão colocados;
  - Clique em Ok.

Utilizando o Excel

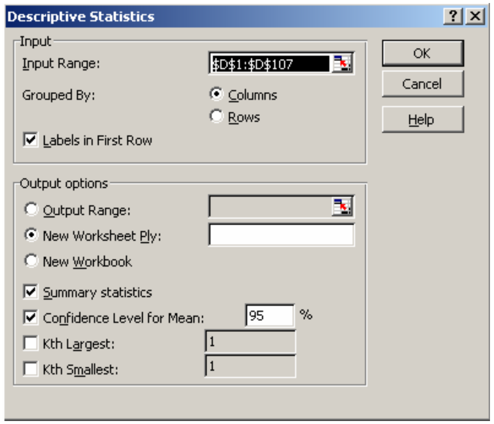
- A saída Erro padrão fornece o valor de $\sigma/\sqrt{n}$ para n grande.

- Para obter o intervalo de confiança baseado na Normal, calcule $z_{1-\alpha/2}$ utilizando a função apropriada, multiplique pelo Erro padrão, e faça: média amostral + e - o resultado encontrado.

- A saída Intervalo de Confiança já fornece o valor de $(t_{1-\alpha/2,n-1})\sigma/\sqrt{n}$ (ou seja, já fornece o que deve ser somado e subtraído da média amostral), bastando apenas subtrair e somar à média.

Utilizando o Excel

- A seguir aplicamos esta análise para o preço da gasolina em 106 postos do Rio de Janeiro em Agosto de 2002.



Utilizando o Excel

Gas. Comum	
Média	**1.725**
Erro Padrão	**0.007**
Mediana	1.725
Moda	1.749
Desvio Padrão	**0.075**
Variância Amostral	0.006
Curtose	1.082
Assimetria	0.386
Amplitude (Máx - Mín)	0.410
Mínimo	1.520
Máximo	1.930
Soma	182.847
n	**106**
IC 95%	0.014

O erro padrão é apenas o desvio padrão dividido por $\sqrt{n} = \sqrt{106}$

$(t_{0.025})\sigma/\sqrt{n}$ – basta subtrair e somar este valor à média para encontrar o IC 95%

Utilizando o Excel

- **Nota:**
- Como o tamanho da amostra é grande, poderíamos ter usado um IC baseado na distribuição Normal.
- Na verdade, a diferença praticamente inexiste, pois o número de graus de liberdade da distribuição t neste caso (105) a torna, para todos os efeitos, indistigüível da Normal.

Forma Alternativa para um IC baseado na distribuição t

- Se definirmos a variância amostral como:

$$S^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

$$\text{e então } \frac{(n)S^{*2}}{\sigma^2} \sim \chi^2_{n-1}$$

- Daí a variável T torna-se:

$$T = \frac{\frac{\sqrt{n}(\bar{X}-\theta)}{\sigma}}{\sqrt{\frac{(n)S^{*2}}{\sigma^2(n-1)}}} = \sqrt{n-1}.\frac{\bar{X}-\theta}{S^*} \sim t_{n-1}$$

Forma Alternativa para um IC baseado na distribuição t

- E aí o intervalo de confiança torna-se:

$$IC = \bar{X} \pm t_{n-1,1-\frac{\alpha}{2}}.\frac{S^*}{\sqrt{n-1}} = \left(\bar{X} - t_{n-1,1-\frac{\alpha}{2}}.\frac{S^*}{\sqrt{n-1}}, \bar{X} + t_{n-1,1-\frac{\alpha}{2}}.\frac{S^*}{\sqrt{n-1}}\right)$$

- **Qual intervalo é "melhor"? Nenhum** – são equivalentes, o importante é saber se você está calculando a variância amostral com denominador n ou (n-1), para ser coerente na sua escolha.

IC para a média de uma distribuição qualquer – GRANDES AMOSTRAS

- **Intervalo de confiança aproximado para as médias de distribuição não-normais (baseado no Teorema Central do Limite).**
- **Considere a v.a. X com densidade ou função de probabilidade f(x), não necessariamente Normal.**
- **Tome uma a.a. de tamanho n desta densidade.**

IC para a média de uma distribuição qualquer – GRANDES AMOSTRAS

- **Se *n (o tamanho da amostra)* é grande o Teorema Central do Limite estabelece que:**

$$S^2 \xrightarrow{P} \sigma^2 \qquad \sqrt{n}\frac{\left(\overline{X}-\theta\right)}{\sigma} \xrightarrow{d} N(0,1)$$

$$\frac{\sqrt{n}(\overline{X}-\theta)/\sigma}{\sqrt{(n-1)S^2/(n-1)\sigma^2}} = \sqrt{n}\frac{(\overline{X}-\theta)}{S} \xrightarrow{d} N(0,1)$$

IC para a média de uma distribuição qualquer – GRANDES AMOSTRAS

- **Daí, um intervalo de confiança aproximado para θ quando a variância é desconhecida e X_i é não- Normal é:**

$$\boxed{\left(\overline{X} - z_{1-\alpha/2}.\frac{S}{\sqrt{n}} ; \overline{X} + z_{1-\alpha/2}.\frac{S}{\sqrt{n}}\right)}$$

onde $z_{1-\alpha/2}$ é obtido de uma N(0,1) tal que:

Prob [- $z_{1-\alpha/2}$ < Z < $z_{1-\alpha/2}$] = 1- α sendo Z ~ N (0,1)