

INTERVALO DE CONFIANÇA 4

Dirce Maria Trevisan Zanetta

- 4.1** Introdução
 - 4.2** Intervalo de confiança (IC)
 - 4.3** Erro-padrão da média e da proporção
 - 4.4** Precisão da estimativa dos parâmetros
 - 4.5** Conclusão
- Referências

LICENCIATURA EM CIÊNCIAS · USP/UNIVESP

Métodos estatísticos: coleta, tratamento e análise de dados; aplicação nas pesquisas de saúde pública, controle de doenças e epidemias.

4.1 Introdução

Na aula passada, você estudou que o **teste de hipóteses** é muito utilizado para fazer a inferência estatística. Vimos que no início do teste são estabelecidas as **hipóteses nula** (H_0) e **alternativa** (H_A), sendo a nula aquela que supõe que não existe diferença e é considerada a verdadeira inicialmente. Também se estabelece o nível *alfa*, que é o nível que se aceita errar se a hipótese nula é rejeitada e ela é a verdadeira.

Uma forma de decidir por uma das hipóteses é verificar na curva de distribuição de probabilidades dos valores do teste estatístico se o resultado do teste está na região de aceitação ou na de rejeição da hipótese nula, que são delimitadas pelo nível *alfa*. Outra forma é calcular o valor p , que é a chance de observar o resultado obtido ou mais extremo que ele com uma amostra, supondo que a hipótese nula seja verdadeira. A hipótese nula é rejeitada se **o valor p for menor que o nível alfa** e não é rejeitada se o valor p for igual ou maior que *alfa*. Nesta aula, nós vamos estudar outro tipo de inferência estatística – **a estimação**.

Existem **dois tipos básicos de inferência estatística**: o **teste de hipóteses** e a **estimação**.

- No **teste de hipóteses**, calcula-se qual a chance de uma diferença entre dois parâmetros ser observada pelo acaso. Quando essa chance é pequena (menor que o nível alfa), rejeita-se H_0 (hipótese de não diferença) e aceita-se H_A (concluindo que os dados são compatíveis com uma diferença entre os parâmetros comparados).
- A inferência por meio da **estimação** é feita quando se deseja conhecer, isto é, saber o valor de um **parâmetro** ou conjunto deles.

Você já estudou a estimação feita por ponto. Por exemplo, uma média de uma amostra, \bar{X} , calculada para estimar o parâmetro populacional m é uma estimação por ponto. Quando a estimação é feita por intervalo, é calculado o intervalo de confiança. O **intervalo de confiança (IC)** é um intervalo estimado de um parâmetro de interesse de uma população. Em vez de estimar o parâmetro por um único valor, é dado um intervalo de estimativas prováveis.

4.2 Intervalo de confiança (IC)

O intervalo de confiança 95% (IC95%) indica a região onde se tem uma confiança de 95% de que cubra o valor verdadeiro do parâmetro populacional, por exemplo, a média verdadeira m ou a proporção π .

Como vimos, os testes estatísticos de hipóteses em geral são feitos para nível alfa de 5% (o nível aceitável de erro, quando se rejeita H_0 e ela é verdadeira) na área da saúde. Embora esse nível seja muito comum, esse valor é arbitrário e pode ser qualquer outro valor. A decisão depende principalmente da **consequência do erro do tipo I** ou **erro alfa**.

○○○○

Por exemplo, se o teste é para avaliar um possível benefício de uma dieta rica em fibras e houver erro na conclusão, esse erro terá **consequências menores** do que se concluísse erradamente que determinada técnica cirúrgica seja melhor que outra, o que pode colocar em risco a vida do paciente que se submete à cirurgia.

○○○○○

Por isso, algumas vezes o nível alfa pode ser estabelecido em outros valores, como 10% ou mesmo 1%, quando o pesquisador precisa diminuir a possibilidade de erro se rejeitar a hipótese nula.

Da mesma forma que o nível *alfa* em geral é de 5%, o intervalo de confiança em geral é calculado para uma confiança de 95%, isso porque, se é aceitável errar em 5%, a confiança deve incluir os outros 95%, que é coeficiente de confiança ($1-\alpha$). Devemos lembrar, porém, que esse coeficiente de confiança é estabelecido pelo pesquisador e, como o nível *alfa*, também pode variar, podendo ser de 99%, por exemplo, ou qualquer outro valor.

Existe um estimador para calcular o IC para estimativas amostrais de cada tipo de parâmetro populacional, mas todos têm uma **fórmula geral**:

$$\text{IC} = \text{estatística amostral} \pm (\text{valor crítico da distribuição} * \text{erro-padrão})$$

A **estatística amostral** é o estimador de um parâmetro amostral desconhecido, como média ou proporção, por exemplo. O valor crítico corresponde ao valor da curva de probabilidades

do teste bicaudal (pode ser testes z ou t , por exemplo) que delimita a região de aceitação da hipótese nula no mesmo nível de confiança em que se quer calcular o intervalo de confiança. O erro-padrão é outra estimativa de dispersão, que vai ser definido a seguir. Para cada parâmetro ou comparação de grupos existe uma fórmula adequada, que pode ser encontrada nos livros de estatística ou de epidemiologia.

4.3 Erro-padrão da média e da proporção

Vamos ver aqui duas situações comuns como exemplo: estimativa do IC95% de **uma média** e de **uma proporção**.

O IC 95% para a média verdadeira da população é calculado com os dados de uma amostra, sendo:

$$\bar{X} \pm 1,96 \text{ EPM}$$

onde \bar{X} é a média observada da amostra e EPM é o erro-padrão da média. 1,96 é o nível crítico da distribuição normal de probabilidades, que delimita 2,5% dos valores em cada extremo da curva e que, portanto, deixa em sua porção central 95% dos valores. O erro-padrão da média é o desvio-padrão (S) da amostra dividido pela raiz quadrada do tamanho da amostra (n):

$$\text{EPM} = \frac{S}{\sqrt{n}}$$

O erro-padrão dá uma informação sobre a distribuição de todas as médias amostrais possíveis de mesmo tamanho de uma população. Vamos ver um exemplo para entender isso melhor.

O **intervalo de confiança espelha o tamanho da amostra**, isto é, quanto maior o n , mais estreito é o IC 95% e, se o intervalo for mais largo, sabemos que a amostra foi pequena em praticamente uma relação de **proporcionalidade inversa** e isso porque o EP é menor quando ocorre aumento do n e vice-versa, uma vez que no seu cálculo é feita a divisão pela raiz quadrada do n .

Para o cálculo do intervalo de confiança de uma proporção, a fórmula será:

$$p = \pm 1,96 \text{ EPP}$$

onde p é a proporção observada na amostra e o EPP é o **erro-padrão da proporção**:

$$\text{EPP} = \sqrt{\frac{p \times (1 - p)}{n}}$$

O intervalo de confiança depende dessas duas quantidades (a **estimativa pontual** e o **erro-padrão**, por exemplo, para médias de \bar{X} e EPM e proporções, de p e EPP) que são calculadas a partir da própria amostra, e devemos esperar resultados diferentes com diferentes amostras.

O intervalo de confiança nos diz que temos 95% de “chance” de estar contido nele o verdadeiro valor do parâmetro populacional (média, proporção etc.) que queremos estimar com a amostra. Em outras palavras, se produzirmos diversos intervalos de confiança, provenientes de diferentes amostras independentes de mesmo tamanho feitas em uma população, podemos esperar que aproximadamente 95% desses intervalos devem conter o verdadeiro valor do parâmetro populacional.

Deve-se notar, contudo, que isso não é garantia de que um determinado intervalo de uma amostra contém o valor verdadeiro, mas que esperamos que o parâmetro esteja contido nele, uma vez que isso ocorre em 95% dos intervalos de confiança.



Intervalos de confiança são usados para **indicar a confiabilidade de uma estimativa**. Então, um IC pode ser usado para descrever o quanto **os resultados** de uma pesquisa são confiáveis. Podemos interpretar o intervalo de confiança como um intervalo que contém os valores “possíveis” que o parâmetro pode assumir. Assim, **quanto maior o intervalo, maior a incerteza** que temos a respeito do parâmetro.

Uma pesquisa que resulte num IC pequeno indica que a estimativa é feita com precisão e tem seu resultado mais confiável do que uma que resulte num IC maior, menos preciso.

Vamos supor que se deseja estimar a média μ de uma população normal com desvio-padrão σ conhecido. Por exemplo, uma população em que a altura média dos homens é de 170,0 cm e com desvio-padrão de 3,2 cm. Esses valores, entretanto, são desconhecidos e, para medi-los com segurança, haveria a necessidade de um censo, o que não é possível quando a população é grande ou de difícil acesso.

Para estimar a altura média dos homens dessa população, é feita uma amostra aleatória, por exemplo, de 100 homens, obtendo-se os valores de \bar{X} e do desvio-padrão S nessa amostra.

Vamos repetir esse processo muitas vezes, isto é, retiramos muitas amostras dessa população, todas com o mesmo tamanho n , que em nosso exemplo é de 100, e cada uma das amostras é selecionada da população inteira, isto é, os indivíduos selecionados em uma amostra anterior estão presentes e têm chance de serem selecionados novamente.

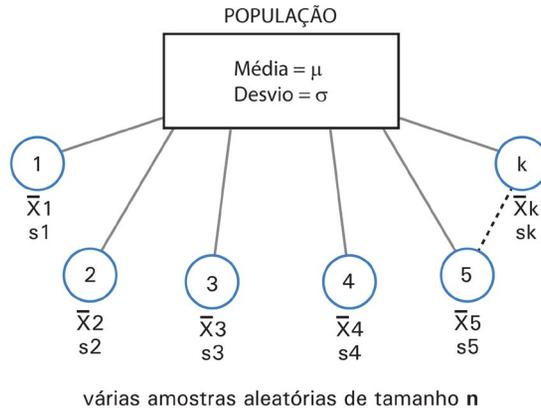


Figura 4.1: Distribuição das médias amostrais para n diferentes.

Para cada amostra, calculamos a **média** e o **desvio-padrão** da altura dos indivíduos sorteados. O erro-padrão, como vimos, nos dá informação sobre a distribuição das médias calculadas em todas as amostras possíveis com o mesmo tamanho em uma determinada população. Ele é o desvio-padrão dessas médias.

É esperado que a variabilidade das médias seja menor que a variabilidade das medidas individuais e vemos que isso se confirma, uma vez que o desvio-padrão das médias (o erro-padrão) é o desvio-padrão da amostra dividido pela raiz quadrada de n . Os intervalos de confiança das estimativas das médias de 20 dessas amostras são mostrados no **Gráfico 4.1**. Observe a flutuação aleatória dos limites do intervalo e que, na grande maioria das vezes, ele passa pelo verdadeiro valor da média da população. O intervalo n° 3 por pouco deixa de cobrir o valor verdadeiro da média.

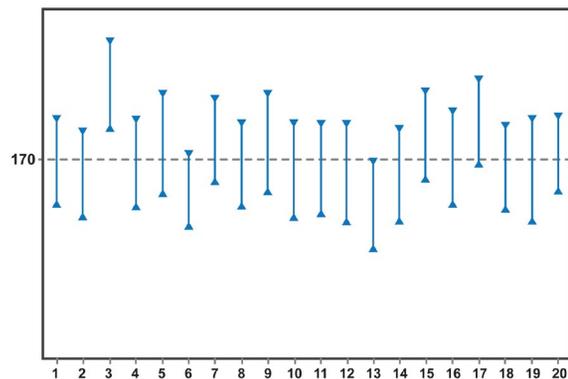


Gráfico 4.1: Intervalos de confiança das estimativas em 20 amostras.

Se não há erro sistemático nas medidas, então, os intervalos de confiança derivados de uma série infinita de amostras deverão conter o **verdadeiro parâmetro estimado** com uma frequência não menor do que o seu nível de confiança (como vimos, na maioria das vezes é de 95%).

No exemplo citado, conhecemos os valores populacionais e, por meio de simulação, obtivemos as várias amostras. Na prática, isso não ocorre. O parâmetro estimado tem um valor desconhecido, que estamos querendo estimar por meio de uma amostra e não temos a oportunidade de verificar se ele contém a média m porque não a conhecemos. A única certeza que nos resta na prática é a de que a probabilidade de o verdadeiro valor do parâmetro estar no intervalo é $1-\alpha$.

A questão, então, é: nunca sabemos se o intervalo de nossa amostra contém ou não o verdadeiro valor que está sendo estimado, mas esperamos, com base na teoria, que sim e assumimos isso na prática, correndo os riscos de uma decisão errada (com possibilidade máxima de erro do α estabelecido para seu cálculo).



O IC é importante para estimar a precisão estatística da estimativa feita. No período eleitoral, são comuns as pesquisas sobre intenções de voto. Sempre que um resultado é divulgado, é ressaltada a sua margem “de segurança” (para mais ou para menos 3%, por exemplo). Esse valor corresponde ao intervalo de confiança, isto é, as pesquisas são feitas para que o verdadeiro valor que está sendo estimado esteja nesse intervalo com a confiança estabelecida (em geral, 95%).

4.4 Precisão da estimativa dos parâmetros

Quando se estabelece o intervalo de confiança aceitável no planejamento da pesquisa, ele é considerado para o cálculo de tamanho da amostra. Como vimos, quanto maior a precisão que se deseja da estimativa, isto é, quanto menor o intervalo de confiança aceitável, maior deve ser o tamanho da amostra.

Além de estimar a precisão da estimativa dos parâmetros, é possível também fazer **testes de significância estatística** com os intervalos de confiança, onde se rejeita a H_0 se, no nível α de significância, o parâmetro de hipótese não estiver contido no intervalo $100(1-\alpha)$ por cento de confiança. Se estiver contido, a H_0 não pode ser rejeitada.

Para a comparação de dois grupos, pode-se estimar o intervalo de confiança da diferença dos valores estimados em cada grupo. A H_0 nesse caso é: a diferença entre eles é igual a zero (se os parâmetros nos dois grupos são iguais, a diferença vai ser igual a zero) e a H_A é: a diferença é diferente de zero, isto é, eles são diferentes. Para fazer essa comparação pelo **teste de hipóteses** que estudamos na aula passada, seria calculada qual a probabilidade de encontrar o resultado obtido pela diferença apenas pelo acaso e, se essa probabilidade fosse grande, nós aceitaríamos H_0 ; e se fosse pequena, menor que *alfa*, nós rejeitaríamos H_0 e aceitaríamos H_A .

Por meio de intervalo de confiança, a **inferência** pode ser feita calculando o intervalo de confiança da diferença, em que H_0 é rejeitada se o intervalo de confiança não contiver o valor zero. Isso é interpretado como se o zero não fosse o verdadeiro valor da diferença, e que, portanto, os dois grupos são diferentes, uma vez que temos 95% de confiança de que o intervalo contém o verdadeiro valor.

○○○○○

Um exemplo é a comparação de **intenção de votos** estimada em uma pesquisa de dois candidatos à eleição. Suponha que o candidato X tem 36% de intenções de voto e o candidato Y tem 32%. Nesse caso, é calculado o intervalo de confiança para a diferença entre os votos dos dois candidatos (que é 4%). Se o intervalo contiver o valor zero, a conclusão é por aceitar H_0 , isto é, concluir que a intenção de votos dos dois candidatos é igual, pois um dos valores possíveis para a diferença entre eles é zero. Se não contiver o valor zero, rejeita-se H_0 e aceita-se H_A , de que há diferença entre a intenção de votos dos dois candidatos.

○○○○○

Outra forma de fazer a **comparação entre grupos**, mais comum, é verificar se os intervalos de confiança de cada estimativa se sobrepõem ou não. Quando não há sobreposição, rejeita-se H_0 e se aceita H_A , uma vez que temos 95% de confiança de que os verdadeiros valores de cada grupo não são coincidentes. No caso da intenção de votos dos candidatos X e Y, se a pesquisa teve margem de “segurança” de 3%, diz-se que os candidatos estão empatados. Isso significa que o intervalo de confiança da intenção de votos do candidato X vai de 33% a 39% e do candidato Y vai de 29% a 35%. Vemos que há sobreposição dos intervalos dos dois candidatos: entre 33% e 35% estão os verdadeiros valores possíveis para ambos. Portanto, não é possível afastar que eles tenham a mesma intenção de votos e a conclusão é a de que eles estão empatados.

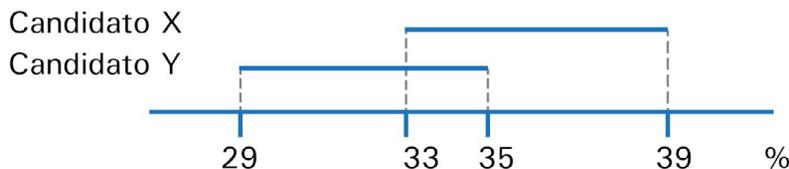


Gráfico 4.2: Intervalo de confiança da frequência de intenção de votos dos candidatos X e Y.

Vamos ver outro exemplo: foram feitos dois inquéritos telefônicos avaliando uma amostra da população adulta que mora em algumas capitais do Brasil, um em 2006 e outro em 2009. Nesses inquéritos, investigou-se a ocorrência de fatores considerados de risco para doenças crônicas. Veja alguns dos resultados na **Tabela 4.1**.

Tabela 4.1: Distribuição de adultos residentes das capitais brasileiras segundo a ocorrência de fatores de risco para doenças crônicas selecionadas, 2006 e 2009.

	2006	2009
Fumante atual	16,2% (15,4% – 17,0%)	15,5% (14,5% – 16,5%)
Presença de obesidade	11,4% (10,8% – 12,0%)	13,9% (13,1% – 14,7%)

Fonte: Adaptado de The Lancet, 2011.

Nessa tabela, são mostradas as frequências de adultos da amostra avaliada que eram fumantes e obesos no momento da entrevista em 2006 e 2009. Entre parênteses após as frequências estão os IC 95% dessas frequências. Para avaliar se existe mudança ao longo do tempo na frequência desses fatores de risco são testadas H_0 , de que não existe diferença, e a H_A , de que existe a diferença. Observando os valores de fumantes atuais, vemos que há **sobreposição dos intervalos de confiança** estimados nos dois anos. A conclusão, então, é a de que não houve mudança na frequência de fumantes nos dois anos avaliados. Já a conclusão para a avaliação da frequência de obesidade é a de que houve aumento em 2009, em comparação com 2006, uma vez que o limite superior do intervalo de confiança dos dados de 2006 (12,0%) é menor do que o limite inferior do IC 95% dos dados de 2009 (13,9%), isto é, não há sobreposição dos IC e, portanto, a conclusão mostra que ocorreu um aumento na frequência de obesidade em adultos nas capitais avaliadas em 2009 em relação a 2006.

4.5 Conclusão

Nesta aula, nós vimos outra forma de inferência estatística, que é por meio da estimação do intervalo de confiança. O intervalo de confiança (IC) é um intervalo de estimativas prováveis de um parâmetro de interesse de uma população. O intervalo de confiança $(1-\alpha)$ indica a região onde se tem uma confiança de $1-\alpha$ de que cubra o valor verdadeiro do parâmetro populacional. Em geral, é calculado para o coeficiente de confiança de 95%. Existe um estimador para calcular o intervalo de confiança para estimativas amostrais de cada tipo de parâmetro populacional, mas todos têm uma fórmula geral, que utiliza o valor crítico da distribuição utilizada que corresponde ao coeficiente de confiança estabelecido e o erro-padrão.

O erro-padrão é o desvio-padrão das médias amostrais de todas as possíveis amostras de mesmo tamanho selecionadas de uma população. Como exemplos, vimos como calcular o IC95% de uma média e de uma proporção. Os intervalos de confiança indicam a confiabilidade de uma estimativa, pois estimam a precisão estatística dessa estimativa. Quanto maior a precisão que se deseja da estimativa, isto é, quanto menor o intervalo de confiança aceitável, maior deve ser o tamanho da amostra. Os intervalos de confiança também são utilizados para testes de significância estatística, onde se rejeita a H_0 se, ao nível α de significância, o parâmetro de hipótese não estiver contido no intervalo $100(1-\alpha)$ por cento de confiança. Se estiver contido, a H_0 não pode ser rejeitada.

É importante entender que a ideia ou os princípios gerais para a comparação de grupos, por exemplo, da proporção encontrada nos dois grupos, ou das médias de cada grupo etc. são os mesmos e válidos em todos os tipos de pesquisa que envolvem **testes de significância**.

Tanto a estatística descritiva como a de inferência são instrumentos importantes para o planejamento de ações que promovam a saúde, para identificar fatores de risco para ocorrência de doenças e poder intervir neles, para a escolha do melhor tratamento etc.

A quantificação da ocorrência da doença também é muito importante. Na próxima aula, nós vamos aprender quais as medidas utilizadas para isso.



Agora é a sua vez...

Agora é sua vez, entre no ambiente virtual e realize as atividades desta aula. Em caso de dúvida, entre em contato com o tutor pelo Fórum. Bons estudos!

Referências

- BONITA, R.; BEAGLEHOLE, R.; KJELLSTRÖM, T. **Epidemiologia Básica**. 2. ed. São Paulo: Santos, 2010.
- DAWSON-SANDERS, B.; TRAPP, R. G. **Bioestatística Básica e Clínica**. 3. ed. Rio de Janeiro: Lange-Appleton & Lange/ McGraw-Hill, 2001.
- JEKEL, J. F.; KATZ, D. L.; ELMORE, J. G. **Epidemiologia, Bioestatística e Medicina Preventiva**. 2. ed. Porto Alegre: Artmed, 2005.
- LOPES, A. P. **Probabilidades e Estatística**. Rio de Janeiro: Reichmann & Affonso, 2000.
- MAGALHÃES, M. N. **Noções de Probabilidade e Estatística**. 6. ed. São Paulo: Edusp, 2008.
- MASSAD, E. et al. **Métodos Quantitativos em Medicina**. São Paulo: Manole, 2004.
- PAGANO, M. et al. **Princípios de Bioestatística**. Tradução da 2. ed. norte-americana. São Paulo: Thompson Learning, 2006.
- SCHMIDT, M. I. et al. Doenças crônicas não transmissíveis no Brasil: carga e desafios atuais. **The Lancet - Saúde no Brasil**. Rio de Janeiro: Elsevier, fasc.1, mai. 2011. Disponível em: <http://189.28.128.100/dab/docs/geral/the_lancet_05_2011.pdf>. Acesso em: 28 fev. 2013.

Glossário

- Coefficiente de confiança (1- α):** é interpretado da forma seguinte: se fossem realizadas amostragens do mesmo modo um número muito grande de vezes é esperado que 1- α % dos “intervalos de confiança 1- α %” construídos a partir delas contenham o verdadeiro valor do parâmetro. Em geral alfa é 5% e, portanto, o IC é de 95%.
- Desvio-padrão:** é a raiz quadrada da variância, a medida de dispersão dos valores da amostra que estima o desvio médio dos valores dos indivíduos em relação à média do grupo.
- Erro-padrão da média:** é desvio-padrão (S) da amostra dividido pela raiz quadrada do tamanho da amostra (n):

$$EPM = \frac{S}{\sqrt{n}}$$

- Erro-padrão da proporção:** é $EPP = \sqrt{\frac{p \times (1-p)}{n}}$, onde p é a proporção estimada na amostra e n é o seu tamanho.

- Erro-padrão:** é outra estimativa de dispersão, que dá uma informação sobre distribuição das médias amostrais de todas as possíveis amostras de mesmo tamanho selecionadas de uma população.

Intervalo de confiança (IC): é um intervalo de estimativas prováveis de um parâmetro de interesse de uma população. O intervalo de confiança nos diz que temos 1-alfa, em geral 95%, de “chance” do verdadeiro valor do parâmetro populacional (média, proporção etc.), que queremos estimar com a amostra, estar contido nele. O intervalo de confiança é utilizado para estimar a precisão estatística da estimativa dos parâmetros e para fazer testes de significância estatística. Para o teste de hipóteses, H_0 é rejeitada, aceitando-se então H_A , se o parâmetro não estiver contido no intervalo 100(1-alfa)% de confiança. Se estiver contido, a H_0 não pode ser rejeitada. Ou para comparação de dois grupos, verifica-se se os intervalos de confiança de cada estimativa se sobrepõem ou não. Quanto não há sobreposição, rejeita-se H_0 e se aceita H_A .