ME951 - Estatística e Probabilidade I

Parte 14

Notas de aula de ME414 produzidas pelos professores **Samara Kiihl**, **Tatiana Benaglia** e **Benilton Carvalho** modificadas e alteradas pela Profa. **Larissa Avila Matos**

Fundamentos de Inferência

Introdução

Um dos principais objetivos da Estatística é tirar conclusões a partir dos dados.

Introdução

Um dos principais objetivos da Estatística é tirar conclusões a partir dos dados.

Dados em geral consistem de uma amostra de elementos de uma população de interesse.

Introdução

Um dos principais objetivos da Estatística é tirar conclusões a partir dos dados.

Dados em geral consistem de uma amostra de elementos de uma população de interesse.

O objetivo é usar a amostra e tirar conclusões sobre a população.

Introdução

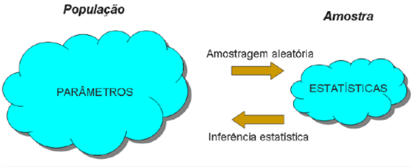
Um dos principais objetivos da Estatística é tirar conclusões a partir dos dados.

Dados em geral consistem de uma amostra de elementos de uma população de interesse.

O objetivo é usar a amostra e tirar conclusões sobre a população.

Quão confiável será utilizar a informação obtida apenas de uma amostra para concluir algo sobre a população?

Inferência Estatística

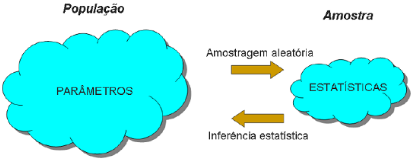


Variável Aleatória: Característica numérica do resultado de um experimento.

População: todos os elementos ou resultados de um problema que está sendo estudado.

Amostra: qualquer subconjunto da população que contém os elementos que podem ser observados e é onde as quantidades de interesse podem ser medidas.

Inferência Estatística



Parâmetros: Característica numérica (desconhecida) da distribuição dos elementos da população.

Estimador/Estatística: Função da amostra, construída com a finalidade de representar, ou estimar um parâmetro de interesse na população.

Estimativa: Valor numérico que um estimador assume para uma dada amostra.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + ... + X_n)$: a média amostral é uma estatística.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

- $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + ... + X_n)$: a média amostral é uma estatística.

- $X_{(1)} = min\{X_1, ..., X_n\}$.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + ... + X_n)$: a média amostral é uma estatística.
- $X_{(1)} = min\{X_1, ..., X_n\}$.
- $X_{(n)} = max\{X_1, ..., X_n\}$.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + ... + X_n)$: a média amostral é uma estatística.
- $X_{(1)} = min\{X_1, ..., X_n\}$.
- $X_{(n)} = max\{X_1, ..., X_n\}$.
- $X_{(i)}$ é o i-ésimo valor da amostra ordenada.

Estatística

Seja $X_1, ..., X_n$ uma amostra, $T = f(X_1, ..., X_n)$ é uma estatística.

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + ... + X_n)$: a média amostral é uma estatística.
- $X_{(1)} = min\{X_1, ..., X_n\}$.
- $X_{(n)} = max\{X_1, ..., X_n\}$.
- $X_{(i)}$ é o i-ésimo valor da amostra ordenada.

Note que uma estatística é uma função que em uma determinada amostra assume um valor específico (estimativa).

Estatística

Para que serve uma estatística? Para "estimar" os valores de uma distribuição, ou características de uma população.

Estatística

Para que serve uma estatística? Para “estimar” os valores de uma distribuição, ou características de uma população.

- **População:**
    - média$_P$.
    - variância$_P$.

Estatística

Para que serve uma estatística? Para "estimar" os valores de uma distribuição, ou características de uma população.

- **População:**
  - média_P.
  - variância_P.

- **Amostra:**
  - $\text{média}_A = \sum_{i=1}^{n} \frac{X_i}{n}$ "estima" a média_P.
  - $\text{variância}_A = \sum_{i=1}^{n} \frac{(X_i - \text{média}_A)^2}{n}$ "estima" a variância_P

Exemplo

Temos interesse em conhecer a média e variância das alturas dos brasileiros adultos. Sabemos que a distribuição das alturas pode ser representada por um modelo normal.

Exemplo

Temos interesse em conhecer a média e variância das alturas dos brasileiros adultos. Sabemos que a distribuição das alturas pode ser representada por um modelo normal.

- Solução 1: Medir a altura de todos os brasileiros adultos.

Exemplo

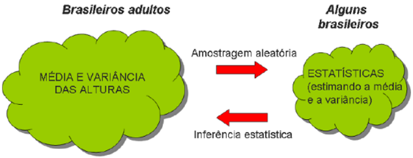
Temos interesse em conhecer a média e variância das alturas dos brasileiros adultos. Sabemos que a distribuição das alturas pode ser representada por um modelo normal.

- Solução 1: Medir a altura de todos os brasileiros adultos.
- Solução 2: Selecionar de forma aleatória algumas pessoas (amostra), analisá-las e inferir propriedades para toda a população.

Exemplo

Temos interesse em conhecer a média e variância das alturas dos brasileiros adultos. Sabemos que a distribuição das alturas pode ser representada por um modelo normal.

- Solução 1: Medir a altura de todos os brasileiros adultos.
- Solução 2: Selecionar de forma aleatória algumas pessoas (amostra), analisá-las e inferir propriedades para toda a população.



Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Inviável perguntar para todos os estudantes: coleta-se uma amostra.

Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Inviável perguntar para todos os estudantes: coleta-se uma amostra.

Planejamento amostral: obter uma amostra aleatória simples de tamanho $n = 100$ alunos, sem reposição.

Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Inviável perguntar para todos os estudantes: coleta-se uma amostra.

Planejamento amostral: obter uma amostra aleatória simples de tamanho $n = 100$ alunos, sem reposição.

Cada X_i, $i = 1, ..., 100$, vai assumir o valor 1 se o aluno i concorda com presença da PM, e 0 se não.

Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Inviável perguntar para todos os estudantes: coleta-se uma amostra.

Planejamento amostral: obter uma amostra aleatória simples de tamanho $n = 100$ alunos, sem reposição.

Cada X_i, $i = 1, ..., 100$, vai assumir o valor 1 se o aluno i concorda com presença da PM, e 0 se não.

Estatística: $T = \frac{X_1 + ... + X_{100}}{100}$.

Exemplo

Seja θ a proporção de alunos na Unicamp que concorda com a presença da PM no campus.

Inviável perguntar para todos os estudantes: coleta-se uma amostra.

Planejamento amostral: obter uma amostra aleatória simples de tamanho $n = 100$ alunos, sem reposição.

Cada X_i, $i = 1, ..., 100$, vai assumir o valor 1 se o aluno i concorda com presença da PM, e 0 se não.

Estatística: $T = \frac{X_1 + ... + X_{100}}{100}$.

Uma vez que a coleta foi implementada, T assume um valor, por exemplo, 0.63, que será usado para estimar θ, ou seja, $\hat{\theta} = 0.63$.

Parâmetro

Cada quantidade de interesse (como θ no exemplo anterior) é chamada de parâmetro da população.

Parâmetro

Cada quantidade de interesse (como θ no exemplo anterior) é chamada de parâmetro da população.

Para apresentar uma estimativa de um parâmetro ($\hat{\theta}$), devemos escolher uma estatística (T).

Parâmetro

Cada quantidade de interesse (como θ no exemplo anterior) é chamada de parâmetro da população.

Para apresentar uma estimativa de um parâmetro ($\hat{\theta}$), devemos escolher uma estatística (T).

Note que da maneira que o plano amostral foi executado (amostra aleatória simples), a estatística T é uma variável aleatória, visto que cada vez que executarmos o plano amostral poderemos obter resultados diversos.

Parâmetro

Cada quantidade de interesse (como θ no exemplo anterior) é chamada de parâmetro da população.

Para apresentar uma estimativa de um parâmetro ($\hat{\theta}$), devemos escolher uma estatística (T).

Note que da maneira que o plano amostral foi executado (amostra aleatória simples), a estatística T é uma variável aleatória, visto que cada vez que executarmos o plano amostral poderemos obter resultados diversos.

Portanto, a estatística T possui uma distribuição de probabilidade, chamada de **distribuição amostral** de T.

Distribuição Amostral

Exemplo

Imagine um fenômeno de interesse que possa ser representado por uma v.a. X que assume os valores 1 ou 2 com igual probabilidade.

Exemplo

Imagine um fenômeno de interesse que possa ser representado por uma v.a. X que assume os valores 1 ou 2 com igual probabilidade.

$$\begin{aligned}\mu = E(X) &= 1 \times P(X = 1) + 2 \times P(X = 2) \\ &= 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}\end{aligned}$$

Exemplo

Imagine um fenômeno de interesse que possa ser representado por uma v.a. X que assume os valores 1 ou 2 com igual probabilidade.

$$\begin{aligned}\mu = E(X) &= 1 \times P(X=1) + 2 \times P(X=2) \\ &= 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}\end{aligned}$$

$$\begin{aligned}\sigma^2 &= Var(X) = E[(X-\mu)^2] \\ &= (1-1.5)^2 \times P(X=1) + (2-1.5)^2 \times P(X=2) \\ &= \frac{1}{4}\end{aligned}$$

Exemplo

Imagine que uma população de interesse tenha distribuição como a de X definida anteriormente.

Exemplo

Imagine que uma população de interesse tenha distribuição como a de X definida anteriormente.

Imagine também que, embora saibamos que os valores possíveis sejam 1 e 2, não tenhamos conhecimento sobre suas respectivas probabilidades.

Exemplo

Imagine que uma população de interesse tenha distribuição como a de X definida anteriormente.

Imagine também que, embora saibamos que os valores possíveis sejam 1 e 2, não tenhamos conhecimento sobre suas respectivas probabilidades.

Isto é, se temos N elementos nessa população, podemos pensar que a característica de interesse de cada elemento i segue uma v.a. X_i em que $P(X_i = 1) = P(X_i = 2) = 1/2$, mas nós não sabemos disso.

Exemplo

Imagine que uma população de interesse tenha distribuição como a de X definida anteriormente.

Imagine também que, embora saibamos que os valores possíveis sejam 1 e 2, não tenhamos conhecimento sobre suas respectivas probabilidades.

Isto é, se temos N elementos nessa população, podemos pensar que a característica de interesse de cada elemento i segue uma v.a. X_i em que $P(X_i = 1) = P(X_i = 2) = 1/2$, mas nós não sabemos disso.

Imagine que o interesse seja μ.

Exemplo

Vamos coletar uma amostra aleatória simples com reposição (AAS_c) de tamanho $n = 2$ e calcular a média amostral.

Exemplo

Vamos coletar uma amostra aleatória simples com reposição (AAS_c) de tamanho $n = 2$ e calcular a média amostral.

Usaremos esta média amostral para estimar μ.

Exemplo

Vamos coletar uma amostra aleatória simples com reposição (AAS_c) de tamanho $n = 2$ e calcular a média amostral.

Usaremos esta média amostral para estimar μ.

Quão útil é esta estimativa que se baseia em apenas 2 elementos da população?

Exemplo

Vamos coletar uma amostra aleatória simples com reposição (AAS_c) de tamanho $n = 2$ e calcular a média amostral.

Usaremos esta média amostral para estimar μ.

Quão útil é esta estimativa que se baseia em apenas 2 elementos da população?

Quão precisa?

Exemplo

Imagine que o aluno A coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

Exemplo

Imagine que o aluno A coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

O aluno B coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

Exemplo

Imagine que o aluno A coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

O aluno B coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

As duas médias amostrais serão necessariamente iguais?

Exemplo

Imagine que o aluno A coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

O aluno B coleta uma AAS_c com $n = 2$ a partir da população, obtém os dados e calcula $\bar{x}$.

As duas médias amostrais serão necessariamente iguais?

A média amostral é uma v.a. e, portanto, tem uma distribuição de probabilidade.

Exemplo

Todas as combinações possíveis de valores para o primeiro e para o segundo elemento amostrados segundo o plano AAS_c com $n = 2$ são:

Possibilidades	$(X_1 = 1, X_2 = 1)$	$(X_1 = 1, X_2 = 2)$
$\bar{x}$	1	1.5
$P(X_1 = i, X_2 = j)$	0.25	0.25

Possibilidades	$(X_1 = 2, X_2 = 1)$	$(X_1 = 2, X_2 = 2)$
$\bar{x}$	1.5	2
$P(X_1 = i, X_2 = j)$	0.25	0.25

$$E(\bar{X}) = 1 \times \frac{1}{4} + 1.5 \times \frac{1}{2} + 2 \times \frac{1}{4} = \frac{3}{2}$$

$$E(\bar{X}) = 1 \times \frac{1}{4} + 1.5 \times \frac{1}{2} + 2 \times \frac{1}{4} = \frac{3}{2}$$

$$\begin{aligned} Var(\bar{X}) &= E\left[(\bar{X} - E(\bar{X}))^2\right] \\ &= (1 - 1.5)^2 \times \frac{1}{4} + (1.5 - 1.5)^2 \frac{1}{2} + (2 - 1.5)^2 \frac{1}{4} = \frac{1}{8} \end{aligned}$$

$$E(\bar{X}) = 1 \times \frac{1}{4} + 1.5 \times \frac{1}{2} + 2 \times \frac{1}{4} = \frac{3}{2}$$

$$\begin{aligned} Var(\bar{X}) &= E\left[(\bar{X} - E(\bar{X}))^2\right] \\ &= (1-1.5)^2 \times \frac{1}{4} + (1.5-1.5)^2\frac{1}{2} + (2-1.5)^2\frac{1}{4} = \frac{1}{8} \end{aligned}$$

Repare que:

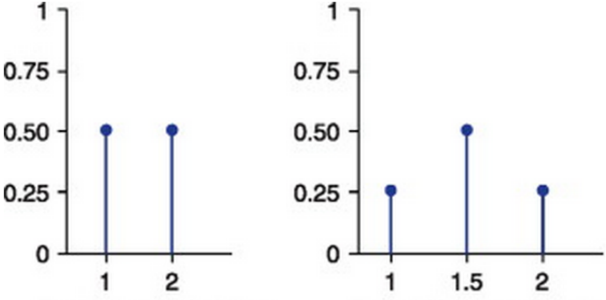
$$E(\bar{X}) = \mu = E(X)$$

e

$$Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{Var(X)}{n}.$$

Exemplo

Distribuição de probabilidade de X (esquerda) e de $\bar{X}$ (direita):



Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Seja $X_1, ..., X_n$ uma amostra aleatória simples de X.

Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Seja $X_1, ..., X_n$ uma amostra aleatória simples de X.

$\bar{X}_n = \frac{X_1 + ... + X_n}{n}$.

Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Seja $X_1, ..., X_n$ uma amostra aleatória simples de X.

$\bar{X}_n = \frac{X_1 + ... + X_n}{n}$.

$E\left(\bar{X}_n\right) = \mu$.

Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Seja $X_1, ..., X_n$ uma amostra aleatória simples de X.

$\bar{X}_n = \frac{X_1 + ... + X_n}{n}$.

$E\left(\bar{X}_n\right) = \mu$.

$Var\left(\bar{X}_n\right) = \frac{\sigma^2}{n}$.

Distribuição Amostral

Resultado:

Seja X uma v.a. com média μ e variância σ^2.

Seja $X_1, ..., X_n$ uma amostra aleatória simples de X.

$\bar{X}_n = \frac{X_1 + ... + X_n}{n}$.

$E\left(\bar{X}_n\right) = \mu$.

$Var\left(\bar{X}_n\right) = \frac{\sigma^2}{n}$.

Ou seja, embora μ seja desconhecido, sabemos que o valor esperado da média amostral é μ. Além disso, conforme o tamanho amostral aumenta, a imprecisão da média amostral para estimar μ fica cada vez menor, pois $Var(\bar{X}) = \sigma^2/n$.

Exemplo

Exemplo: X_1, X_2, X_3 ensaios de Bernoulli(p) independentes.

Exemplo

Exemplo: X_1, X_2, X_3 ensaios de Bernoulli(p) independentes.

$$\mu = E\left(X_i\right) = 0.3 \quad \Rightarrow \quad E\left(\bar{X}_3\right) = 0.3.$$

Exemplo

Exemplo: X_1, X_2, X_3 ensaios de Bernoulli(p) independentes.

$\mu = E\left(X_i\right) = 0.3 \quad \Rightarrow \quad E\left(\bar{X}_3\right) = 0.3.$

$\sigma^2 = Var\left(X_i\right) = p(1-p) = 0.3(0.7) = 0.21 \quad \Rightarrow \quad Var\left(\bar{X}_3\right) = \frac{0.21}{3} = 0.07$

Teorema do Limite Central

Teorema do Limite Central

Usando o resultado enunciado anteriormente, temos a esperança e a variância da média amostral $\bar{X}$: $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{\sigma^2}{n}$.

Teorema do Limite Central

Usando o resultado enunciado anteriormente, temos a esperança e a variância da média amostral $\bar{X}$: $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{\sigma^2}{n}$.

No entanto, para conhecermos a distribuição de probabilidade de $\bar{X}$, como foi feito no exemplo anterior, é preciso conhecer todos os valores possíveis de X e suas respectivas probabilidades.

Teorema do Limite Central

Usando o resultado enunciado anteriormente, temos a esperança e a variância da média amostral $\bar{X}$: $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{\sigma^2}{n}$.

No entanto, para conhecermos a distribuição de probabilidade de $\bar{X}$, como foi feito no exemplo anterior, é preciso conhecer todos os valores possíveis de X e suas respectivas probabilidades.

Mas, se conhecermos tudo isso, não precisamos fazer amostragem nem inferência: saberemos tudo o que desejarmos daquela população!

Teorema do Limite Central

Usando o resultado enunciado anteriormente, temos a esperança e a variância da média amostral $\bar{X}$: $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{\sigma^2}{n}$.

No entanto, para conhecermos a distribuição de probabilidade de $\bar{X}$, como foi feito no exemplo anterior, é preciso conhecer todos os valores possíveis de X e suas respectivas probabilidades.

Mas, se conhecermos tudo isso, não precisamos fazer amostragem nem inferência: saberemos tudo o que desejarmos daquela população!

O exemplo anterior foi um caso hipotético apenas para demonstrar como a média amostral $\bar{X}$ se comporta quando realizamos a amostragem.

Teorema do Limite Central

Usando o resultado enunciado anteriormente, temos a esperança e a variância da média amostral $\bar{X}$: $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \frac{\sigma^2}{n}$.

No entanto, para conhecermos a distribuição de probabilidade de $\bar{X}$, como foi feito no exemplo anterior, é preciso conhecer todos os valores possíveis de X e suas respectivas probabilidades.

Mas, se conhecermos tudo isso, não precisamos fazer amostragem nem inferência: saberemos tudo o que desejarmos daquela população!

O exemplo anterior foi um caso hipotético apenas para demonstrar como a média amostral $\bar{X}$ se comporta quando realizamos a amostragem.

Na prática, não teremos informações suficientes para de fato descrevermos a distribuição exata de $\bar{X}$.

Teorema Central do Limite (TLC)

Resultado

Para uma amostra aleatória simples $X_1, ..., X_n$ coletada de uma população com média μ e variância σ^2, a distribuição amostral de $\bar{X}_n$ aproxima-se de uma **distribuição Normal** de média μ e variância $\frac{\sigma^2}{n}$, quando n for suficientemente grande.

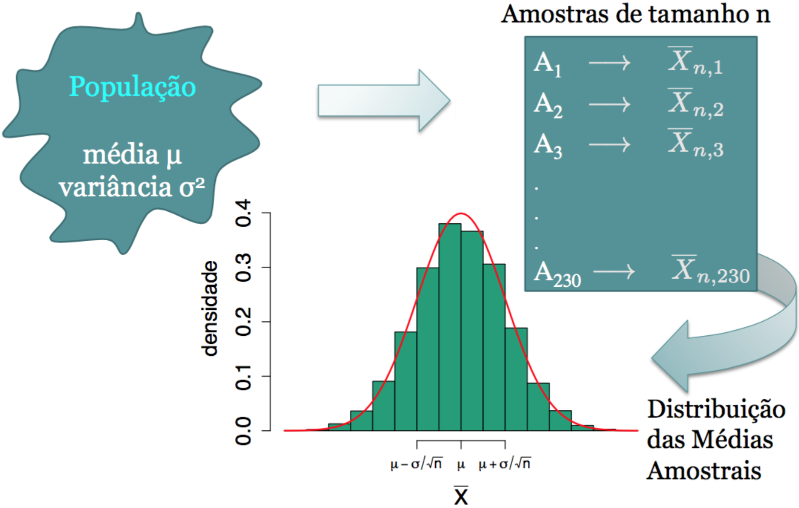
Resultado

Para uma amostra aleatória simples $X_1, ..., X_n$ coletada de uma população com média μ e variância σ^2, a distribuição amostral de $\bar{X}_n$ aproxima-se de uma **distribuição Normal** de média μ e variância $\frac{\sigma^2}{n}$, quando n for suficientemente grande.

Definimos também:

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

Teorema do Limite Central



Exemplo

Seja $X_1, ..., X_n$ uma amostra aleatória de tamanho n tal que $X \sim Exp(2)$:

$$f_{X_i}(x) = 2e^{-2x}, \qquad \text{para } x \geq 0$$

Então $E\left(X_i\right) = \frac{1}{2}$ e $Var\left(X_i\right) = \frac{1}{4}$.

Suponha que X_i modela o tempo de vida de um transistor em horas. Os tempos de vida de 100 transistores são coletados. Desejamos estudar a variável aleatória $\bar{X}_{100}$ (média amostral de uma amostra de tamanho 100). Sabemos:

$$E\left(\bar{X}_{100}\right) = \frac{1}{2} \qquad \text{e} \qquad Var\left(\bar{X}_{100}\right) = \frac{1/4}{100} = \frac{1}{400}.$$

Pelo TLC, temos que: $\bar{X}_n \sim N\left(\frac{1}{2}, \frac{1}{400}\right)$

Exemplo

$$
\begin{aligned}
F_{\bar{X}_{100}}(x) = P\left(\bar{X}_{100} \leq x\right) &= P\left(\frac{\bar{X}_{100} - (1/2)}{(1/2)/\sqrt{100}} \leq \frac{x - (1/2)}{(1/2)/\sqrt{100}}\right) \\
&= P\left(Z \leq 10(2x-1)\right)
\end{aligned}
$$

e

$$
\begin{aligned}
P\left(\bar{X}_{100} \geq x\right) &= 1 - P\left(\bar{X}_{100} < x\right) \\
&= 1 - P\left(\frac{\bar{X}_{100} - (1/2)}{(1/2)/\sqrt{100}} \leq \frac{x - (1/2)}{(1/2)/\sqrt{100}}\right) \\
&= 1 - P\left(Z \leq 10(2x-1)\right)
\end{aligned}
$$

Exemplo

X = resultado obtido no lançamento de um dado honesto.

x	1	2	3	4	5	6
$p(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Exemplo

X = resultado obtido no lançamento de um dado honesto.

x	1	2	3	4	5	6
$p(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(X) = \frac{1}{6} \times (1+2+3+4+5+6) = \frac{21}{6} = 3.5$

$Var(X) = \frac{1}{6}[(1+4+9+16+25+36) - \frac{1}{6} \times (21)^2] = \frac{35}{2} = 17.5$

Exemplo

$X =$ resultado obtido no lançamento de um dado honesto.

x	1	2	3	4	5	6
$p(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(X) = \frac{1}{6} \times (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$

$Var(X) = \frac{1}{6}[(1 + 4 + 9 + 16 + 25 + 36) - \frac{1}{6} \times (21)^2] = \frac{35}{2} = 17.5$

X_i: resultado do i-ésimo lançamento de um dado honesto.

Exemplo

X = resultado obtido no lançamento de um dado honesto.

x	1	2	3	4	5	6
$p(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(X) = \frac{1}{6} \times (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$

$Var(X) = \frac{1}{6}[(1 + 4 + 9 + 16 + 25 + 36) - \frac{1}{6} \times (21)^2] = \frac{35}{2} = 17.5$

X_i: resultado do i-ésimo lançamento de um dado honesto.

X_i tem distribuição uniforme discreta $\forall i$.

Exemplo

X = resultado obtido no lançamento de um dado honesto.

x	1	2	3	4	5	6
$p(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(X) = \frac{1}{6} \times (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$

$Var(X) = \frac{1}{6}[(1 + 4 + 9 + 16 + 25 + 36) - \frac{1}{6} \times (21)^2] = \frac{35}{2} = 17.5$

X_i: resultado do i-ésimo lançamento de um dado honesto.

X_i tem distribuição uniforme discreta $\forall i$.

$\mu = E(X_i) = 3.5$ e $\sigma^2 = Var(X_i) = 17.5, \forall i.$

Exemplo

Se temos uma amostra aleatória simples de tamanho n: $X_1, X_2, \ldots, X_n$, pelo TLC sabemos que a distribuição amostral de $\bar{X}_n$ é aproximadamente Normal$\left(3.5, \frac{17.5}{n}\right)$.

Exemplo

Se temos uma amostra aleatória simples de tamanho n:
$X_1, X_2, \ldots, X_n$, pelo TLC sabemos que a distribuição amostral de $\bar{X}_n$ é aproximadamente Normal$\left(3.5, \frac{17.5}{n}\right)$.

O primeiro histograma a seguir mostra o resultado de 10000 repetições do seguinte experimento: observar o resultado do lançamento de 1 dado. Repare que é muito próximo de uma distribuição uniforme discreta (chance 1/6 para cada resultado).

Exemplo

Se temos uma amostra aleatória simples de tamanho n: $X_1, X_2, \ldots, X_n$, pelo TLC sabemos que a distribuição amostral de $\bar{X}_n$ é aproximadamente Normal$\left(3.5, \frac{17.5}{n}\right)$.

O primeiro histograma a seguir mostra o resultado de 10000 repetições do seguinte experimento: observar o resultado do lançamento de 1 dado. Repare que é muito próximo de uma distribuição uniforme discreta (chance 1/6 para cada resultado).

O segundo histograma mostra o resultado de 10000 repetições do seguinte experimento: observar a média do lançamento de 2 dados (equivalente a observar a média de 2 lançamentos de um dado).

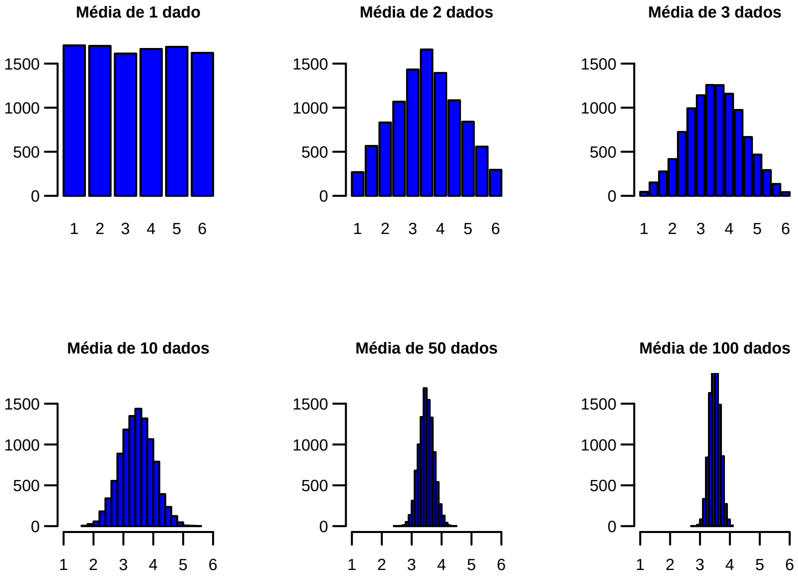
Exemplo

Se temos uma amostra aleatória simples de tamanho n: $X_1, X_2, \ldots, X_n$, pelo TLC sabemos que a distribuição amostral de $\bar{X}_n$ é aproximadamente Normal$\left(3.5, \frac{17.5}{n}\right)$.

O primeiro histograma a seguir mostra o resultado de 10000 repetições do seguinte experimento: observar o resultado do lançamento de 1 dado. Repare que é muito próximo de uma distribuição uniforme discreta (chance 1/6 para cada resultado).

O segundo histograma mostra o resultado de 10000 repetições do seguinte experimento: observar a média do lançamento de 2 dados (equivalente a observar a média de 2 lançamentos de um dado).

Exemplo


Média de 1 dado
Média de 2 dados
Média de 3 dados
Média de 10 dados
Média de 50 dados
Média de 100 dados
0
500
1000
1500
1
2
3
4
5
6

Exemplo

O último histograma mostra o resultado de 10000 repetições do seguinte experimento: observar a média do lançamento de 100 dados (equivalente a observar a média de 100 lançamentos de um dado).

Exemplo

O último histograma mostra o resultado de 10000 repetições do seguinte experimento: observar a média do lançamento de 100 dados (equivalente a observar a média de 100 lançamentos de um dado).

Repare que conforme o número de dados (tamanho amostral) aumenta, a distribuição da média amostral se aproxima da distribuição normal com média 3.5 e variância cada vez menor (17.5/n).

Teorema do Limite Central (TLC)

Você pode verificar o comportamento de $\bar{X}$ para vários tipos de distribuição de X:

https://nishantsbi.shinyapps.io/CLT_Shiny

https://gallery.shinyapps.io/CLT_mean/

Aproximação da Distribuição Binomial pela Normal

Aproximação da Binomial pela Normal

Consideremos uma população em que a proporção de indivíduos portadores de uma certa característica seja p.

$$X_i = \begin{cases} 1, & \text{se o indivíduo i possui a característica} \\ 0, & \text{caso contrário} \end{cases}$$

Aproximação da Binomial pela Normal

Consideremos uma população em que a proporção de indivíduos portadores de uma certa característica seja p.

$$X_i = \begin{cases} 1, & \text{se o indivíduo i possui a característica} \\ 0, & \text{caso contrário} \end{cases}$$

$\Rightarrow X_i \sim Bernoulli(p);\ i = 1, 2, ..., n.$

Aproximação da Binomial pela Normal

Consideremos uma população em que a proporção de indivíduos portadores de uma certa característica seja p.

$$X_i = \begin{cases} 1, & \text{se o indivíduo i possui a característica} \\ 0, & \text{caso contrário} \end{cases}$$

$\Rightarrow X_i \sim Bernoulli(p);\ i = 1, 2, ..., n.$

Se as observações são independentes:

$$S_n = X_1 + ... + X_n \sim Bin(n, p).$$

Aproximação da Binomial pela Normal

Consideremos uma população em que a proporção de indivíduos portadores de uma certa característica seja p.

$$X_i = \begin{cases} 1, & \text{se o indivíduo i possui a característica} \\ 0, & \text{caso contrário} \end{cases}$$

$\Rightarrow X_i \sim Bernoulli(p);\ i = 1, 2, ..., n.$

Se as observações são independentes:

$$S_n = X_1 + ... + X_n \sim Bin(n, p).$$

Após a coleta de uma amostra aleatória simples de n indivíduos, podemos considerar:

Aproximação da Binomial pela Normal

Consideremos uma população em que a proporção de indivíduos portadores de uma certa característica seja p.

$$X_i = \begin{cases} 1, & \text{se o indivíduo i possui a característica} \\ 0, & \text{caso contrário} \end{cases}$$

$\Rightarrow X_i \sim Bernoulli(p);\ i = 1, 2, ..., n.$

Se as observações são independentes:

$$S_n = X_1 + ... + X_n \sim Bin(n, p).$$

Após a coleta de uma amostra aleatória simples de n indivíduos, podemos considerar:

$\hat{p} = \frac{S_n}{n}$ (média amostral como estimador da média populacional).

Aproximação da Binomial pela Normal

Utilizando a distribuição exata (n pequeno):

$$P\left(\hat{p} = \frac{k}{n}\right) = P\left(\frac{S_n}{n} = \frac{k}{n}\right) = P\left(S_n = k\right) = \left(\begin{array}{c} n \\ k \end{array}\right) p^k \left(1-p\right)^{n-k},$$

para $k = 0, 1, ..., n$.

Aproximação da Binomial pela Normal

Utilizando a distribuição exata (n pequeno):

$$P\left(\hat{p}=\frac{k}{n}\right)=P\left(\frac{S_n}{n}=\frac{k}{n}\right)=P\left(S_n=k\right)=\left(\begin{array}{c} n \\ k \end{array}\right)p^k\left(1-p\right)^{n-k},$$

para $k=0,1,...,n$.

Utilizando a aproximação para a Normal (n grande):

$$\hat{p}\sim N\left(p,\frac{p(1-p)}{n}\right)$$

Exemplo

Se p for a proporção de fumantes no estado de SP, $p = 0.2$ e tivermos coletado uma amostra aleatória simples de 500 indivíduos:

$$X_i = \begin{cases} 1, & \text{se o indivíduo i é fumante} \\ 0, & \text{caso contrário} \end{cases}$$

Exemplo

Se p for a proporção de fumantes no estado de SP, $p = 0.2$ e tivermos coletado uma amostra aleatória simples de 500 indivíduos:

$$X_i = \begin{cases} 1, & \text{se o indivíduo i é fumante} \\ 0, & \text{caso contrário} \end{cases}$$

$\hat{p} = \frac{\sum_{i=1}^{500} X_i}{500}$.

Exemplo

Se p for a proporção de fumantes no estado de SP, $p = 0.2$ e tivermos coletado uma amostra aleatória simples de 500 indivíduos:

$$X_i = \begin{cases} 1, & \text{se o indivíduo i é fumante} \\ 0, & \text{caso contrário} \end{cases}$$

$\hat{p} = \frac{\sum_{i=1}^{500} X_i}{500}$.

$\hat{p} \sim N\left(0.2, \frac{0.2 \times 0.8}{500}\right) = N\left(0.2, 0.00032\right)$

Exemplo

Se p for a proporção de fumantes no estado de SP, $p = 0.2$ e tivermos coletado uma amostra aleatória simples de 500 indivíduos:

$$X_i = \begin{cases} 1, & \text{se o indivíduo i é fumante} \\ 0, & \text{caso contrário} \end{cases}$$

$\hat{p} = \frac{\sum_{i=1}^{500} X_i}{500}.$

$\hat{p} \sim N\left(0.2, \frac{0.2 \times 0.8}{500}\right) = N\left(0.2, 0.00032\right)$

$P\left(\hat{p} \leq 0.25\right) = P\left(Z \leq 2.795\right) = \Phi\left(2.795\right) = 0.9974$

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}.$

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}$.

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}$.

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Qual a distribuição de S_n quando n é grande o suficiente?

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}$.

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Qual a distribuição de S_n quando n é grande o suficiente?

$S_n = X_1 + ... + X_n$

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}$.

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Qual a distribuição de S_n quando n é grande o suficiente?

$S_n = X_1 + ... + X_n$

$\hat{p} = \frac{S_n}{n} \sim N\left(p, \frac{p(1-p)}{n}\right)$

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}.$

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Qual a distribuição de S_n quando n é grande o suficiente?

$S_n = X_1 + ... + X_n$

$\hat{p} = \frac{S_n}{n} \sim N\left(p, \frac{p(1-p)}{n}\right)$

$S_n = n\hat{p} \sim N\left(np, np(1-p)\right)$

Aproximação da Binomial pela Normal

$\hat{p} = \frac{S_n}{n} \Rightarrow S_n = n\hat{p}$.

Quando n é grande o suficiente $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$.

Qual a distribuição de S_n quando n é grande o suficiente?

$S_n = X_1 + ... + X_n$

$\hat{p} = \frac{S_n}{n} \sim N\left(p, \frac{p(1-p)}{n}\right)$

$S_n = n\hat{p} \sim N\left(np, np(1-p)\right)$

Portanto: $Bin(n, p) \approx N\left(np, np(1-p)\right)$ quando n é grande.

Exemplo

$X \sim Bin(100, 0.4)$

Exemplo

$$X \sim Bin(100, 0.4)$$

$$E(X) = 100 \times 0.4 = 40$$

Exemplo

$X \sim Bin(100, 0.4)$

$E(X) = 100 \times 0.4 = 40$

$Var(X) = 100 \times 0.4 \times 0.6 = 24$

Exemplo

$$X \sim Bin(100, 0.4)$$

$$E(X) = 100 \times 0.4 = 40$$

$$Var(X) = 100 \times 0.4 \times 0.6 = 24$$

$$X \approx N(40, 24)$$

Exemplo

$$X \sim Bin(100, 0.4)$$

$$E(X) = 100 \times 0.4 = 40$$

$$Var(X) = 100 \times 0.4 \times 0.6 = 24$$

$$X \approx N(40, 24)$$

$$P\left(X \leq 50\right) = P\left(Z \leq \frac{50-40}{\sqrt{24}}\right) \approx \Phi\left(\frac{10}{\sqrt{24}}\right) = \Phi\left(2.04\right) \approx 0.9793$$

Leituras

- Ross: capítulo 7.
- OpenIntro: seção 4.1
- Magalhães: capítulo 7.