Estatística

IX

M. Cristina Miranda e Anabela Rocha

1. Teorema do Limite Central
2. Aplicações do Teorema do Limite Central
3. Distribuição da média em populações Normais

M. Cristina Miranda e Anabela Rocha

Teorema do Limite Central

Aplicações do Teorema do Limite Central

Distribuição da média em populações Normais

Section 1

Teorema do Limite Central

Teorema do Limite Central (TLC)

Teorema do Limite Central (TLC)

Sejam $X_1, X_2, ..., X_n$, v.a.s i.i.d. com variância finita $V[X_i] = \sigma^2$ e valor esperado $E[X_i] = \mu$. Então,

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

ou, de forma equivalente,

$$\lim_{n\to\infty} \left(\bar{X} \right) = Z \sim N(\mu, \frac{\sigma^2}{n}).$$

Na prática, pode usar-se o TLC quando $n \geq 30$.

Section 2

Aplicações do Teorema do Limite Central

Aproximação da Binomial à Normal

Se X é uma v.a. com distribuição Binomial, $X \sim B(n,p)$, $E[X] = np$ e $V[X] = np(1-p)$, então

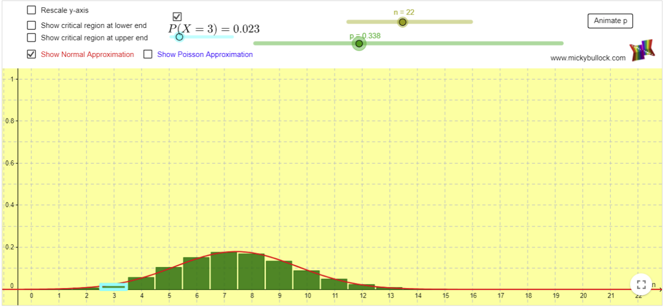
$$\lim_{n\to\infty} \left(\frac{X - np}{\sqrt{np(1-p)}} \right) = Z \sim N(0,1)$$

Binomial$\longrightarrow$ Normal

Ilustração

Author: gesges, Micky Bullock

Topic: Binomial Distribution, Distributions, Hypothesis Testing, Normal Distribution, Poisson Distribution, Probability, Statistics



https://www.geogebra.org/m/zxBQV7te

1) Exemplo

Seja X uma variável aleatória que representa o número de jogadores de basquetebol que conseguem marcar pontos num campeonato, num total de 45 jogadores. Supondo que a probabilidade de encestar é igual para todos e tem o valor de $p = 0.6$, calcule um valor aproximado para a probabilidade de haver entre 20 e 30 jogadores a marcar pontos nesse campeonato.

$X \sim B(45, 0.6)$; $E[X] = 45 \times 0.6 = 27$ e $V[X] = 45 \times 0.6 \times 0.4 = 10.8$. Como n=45, pode aplicar-se o TLC,

$$P(20 < X < 30) = P\left(\frac{20-27}{\sqrt{10.8}} < \frac{X-27}{\sqrt{10.8}} < \frac{30-27}{\sqrt{10.8}}\right)$$

$$\simeq \Phi\left(\frac{30-27}{\sqrt{10.8}}\right) - \Phi\left(\frac{20-27}{\sqrt{10.8}}\right)$$
$$\simeq 0.803.$$

Aproximação da Poisson à Normal

Se $(X_1, X_2, ..., X_n)$ é uma a.a. com distribuição Poisson, $X_i \sim P(\lambda)$, então

$X = \sum_{i=1}^{n} X_i \sim P(n\lambda)$ e

$$\lim_{n\to\infty} \frac{X - n\lambda}{\sqrt{n\lambda}} = Z \sim N(0,1).$$

Poisson $\longrightarrow$ Normal

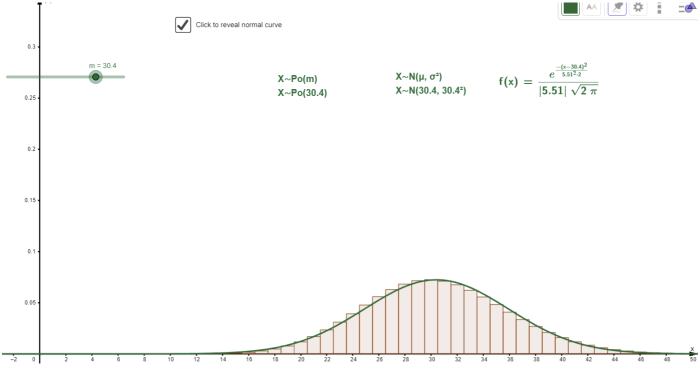
universidade de aveiro

Ilustração

Estatística

M. Cristina Miranda e Anabela Rocha

Teorema do Limite Central
Aplicações do Teorema do Limite Central
Distribuição da média em populações Normais



https://www.geogebra.org/m/F9qSxXwP

2) Exemplo

Para uma população com distribuição Poisson de valor médio 9, calcule um valor aproximado para a probabilidade da média de uma amostra de 32 observações variar entre 8 e 13.

$$X_i \sim P(9) \Leftrightarrow \begin{cases} E[X_i] & = 9 \\ V[X_i] & = 9 \end{cases} \Leftrightarrow \begin{cases} E[\bar{X}] & = 9 \\ V[\bar{X}] & = \frac{9}{32}. \end{cases}$$

Como $n = 32$, vem pelo TLC, $\dfrac{\bar{X} - 9}{\sqrt{9/32}} \rightarrow N(0,1)$.

$$P\left(8 < \bar{X} < 13\right)$$

$$\simeq \Phi\left(\frac{13-9}{\sqrt{9/32}}\right) - \Phi\left(\frac{8-9}{\sqrt{9/32}}\right)$$

$$\simeq 0.970.$$

3) Exemplo

Numa dada região o número de clientes de um determinado operador de telecomunicações, por km^2 é uma v. a. X com distribuição Poisson de média unitária.

3) Exemplo

Numa dada região o número de clientes de um determinado operador de telecomunicações, por km^2 é uma v. a. X com distribuição Poisson de média unitária.

1. Calcule a probabilidade de, num km^2, encontrar pelo menos 1 cliente desse operador.

3) Exemplo

Numa dada região o número de clientes de um determinado operador de telecomunicações, por km^2 é uma v. a. X com distribuição Poisson de média unitária.

1. Calcule a probabilidade de, num km^2, encontrar pelo menos 1 cliente desse operador.
2. Indique, justificando, um valor aproximado para a probabilidade de encontrar entre 440 e 500 clientes desse operador, numa zona com $400 km^2$.

$$X \sim P(1) \Leftrightarrow \begin{cases} E[X] & = 1 \\ V[X] & = 1 \end{cases}$$

1.

$$P(X \geq 1) = 1 - P(X < 1) =$$

$$= 1 - P(X = 0) = 1 - f(0) = 0.632.$$

2. Como $n = 400$, pode aplicar-se o TLC,

$$\frac{\sum_{i=1}^{400} X_i - 400}{20} \to N(0,1)$$

$$P(440 < \sum_{i=1}^{400} X < 500) =$$

$$P\left(\frac{440-400}{20} < \frac{\sum_{i=1}^{400} X_i - 400}{20} < \frac{500-400}{20}\right)$$

$$= \Phi(5) - \Phi(2) \simeq 0.023.$$

Section 3

Distribuição da média em populações Normais

Distribuição da média em populações Normais

Em estatística inferencial, muitas vezes considera-se que a população em estudo segue uma distribuição Normal.

Supondo uma a. a. $(X_1, X_2, ..., X_n)$ de uma população com distribuição Normal, ou seja, considerando que X_i são v.a. independentes e $X_i \sim N(\mu, \sigma^2)$, então tem-se que:

$$\bar{X} \sim N(\mu, \sigma^2/n).$$

Com efeito, por uma propriedade estudada para a distribuição Normal, sabe-se que a v.a. $\bar{X}$ ainda segue uma distribuição Normal, com parâmetros:

$$E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu;$$

e

$$V[\bar{X}] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V[X_i] = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}.$$

4) Exemplo

Duas amostras aleatórias independentes, de tamanhos 10 e 15 respetivamente, foram extraídas de uma população Normal, com valor médio 20 e variância 3.

4) Exemplo

Duas amostras aleatórias independentes, de tamanhos 10 e 15 respetivamente, foram extraídas de uma população Normal, com valor médio 20 e variância 3.

1. Diga, justificando, qual a distribuição da diferença das médias das duas amostras, $\bar{X}_1 - \bar{X}_2$.

4) Exemplo

Duas amostras aleatórias independentes, de tamanhos 10 e 15 respetivamente, foram extraídas de uma população Normal, com valor médio 20 e variância 3.

1. Diga, justificando, qual a distribuição da diferença das médias das duas amostras, $\bar{X}_1 - \bar{X}_2$.
2. Qual a probabilidade do módulo da diferença das médias das duas amostras, ser pelo menos 0.3?

1.

$$A_1 = (X_1, X_2, ..., X_{10}), A_2 = (X_1, X_2, ..., X_{15});$$

$$X_i \sim N(20, 3);$$

$$X_i \sim N(20, 3) \Rightarrow \bar{X}_1 \sim N(20, \frac{3}{10}) \text{ e } \bar{X}_2 \sim N(20, \frac{3}{15}).$$

A combinação linear de v.a. independentes com distribuição Normal é uma v.a. com distribuição Normal. Por isso, a diferença das médias amostrais tem distribuição Normal. Quais os parâmetros?

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = 20 - 20 = 0$$

e

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{3}{10} + \frac{3}{15} = \frac{3+2}{10} = \frac{1}{2};$$

Então,

$$\bar{X}_1 - \bar{X}_2 \sim N(0; 0,5).$$

$$P|\bar{X}_1 - \bar{X}_2| \geq 0,3 = 1 - P|\bar{X}_1 - \bar{X}_2| \leq 0,3 =$$
$$= 1 - P(-0,3 \leq \bar{X}_1 - \bar{X}_2 \leq 0,3) = 1 - F(0,3) + F(-0,3) =$$

$$= 1 - 0.66 + 0.34 = 0.67.$$