Probabilidade e estatística - Aula 16

Estimação pontual de parâmetros - Continuação

Dr. Giannini Italino Alves Vieira

Universidade Federal do Ceará - Campus de Crateús

2024

Recordando

- Na aula passada vimos um método de obtenção de estimadores, chamado de método de momentos.
- Vimos que a ideia da obtenção de estimadores por meio do método de momentos é bem simples e intuitiva, consistindo em obter estimadores igualando-se momentos populacionais aos seus respectivos momentos amostrais.
- Na aula de hoje iremos estudar um outro método de obtenção de estimadores, chamado de método de máxima verossimilhança.
- Estimadores obtidos pelo método de máxima verossimilhança são geralmente preferíveis em relação a estimadores obtidos por meio do método de momentos, pois possuem melhores propriedades de eficiência.

Recordando

- Na aula passada vimos um método de obtenção de estimadores, chamado de método de momentos.
- Vimos que a ideia da obtenção de estimadores por meio do método de momentos é bem simples e intuitiva, consistindo em obter estimadores igualando-se momentos populacionais aos seus respectivos momentos amostrais.
- Na aula de hoje iremos estudar um outro método de obtenção de estimadores, chamado de método de máxima verossimilhança.
- Estimadores obtidos pelo método de máxima verossimilhança são geralmente preferíveis em relação a estimadores obtidos por meio do método de momentos, pois possuem melhores propriedades de eficiência.

Recordando

- Na aula passada vimos um método de obtenção de estimadores, chamado de método de momentos.
- Vimos que a ideia da obtenção de estimadores por meio do método de momentos é bem simples e intuitiva, consistindo em obter estimadores igualando-se momentos populacionais aos seus respectivos momentos amostrais.
- Na aula de hoje iremos estudar um outro método de obtenção de estimadores, chamado de método de máxima verossimilhança.
- Estimadores obtidos pelo método de máxima verossimilhança são geralmente preferíveis em relação a estimadores obtidos por meio do método de momentos, pois possuem melhores propriedades de eficiência.

Recordando

- Na aula passada vimos um método de obtenção de estimadores, chamado de método de momentos.
- Vimos que a ideia da obtenção de estimadores por meio do método de momentos é bem simples e intuitiva, consistindo em obter estimadores igualando-se momentos populacionais aos seus respectivos momentos amostrais.
- Na aula de hoje iremos estudar um outro método de obtenção de estimadores, chamado de método de máxima verossimilhança.
- Estimadores obtidos pelo método de máxima verossimilhança são geralmente preferíveis em relação a estimadores obtidos por meio do método de momentos, pois possuem melhores propriedades de eficiência.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Método de máxima verossimilhança

- Veremos que o estimador de máxima verossimilhança será o valor do parâmetro que maximiza a função de verossimilhança. Dessa forma, é fundamental entender o que é a função de verossimilhança.

Função de verossimilhança

Definição: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n da variável aleatória X com função de probabilidade (ou densidade) $f(x, \theta)$, em que θ seja um único parâmetro desconhecido. Seja $x_1, x_2, \ldots, x_n$ os valores observados da amostra aleatória. Então a função de verossimilhança é definida como

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

- Note que a função de verossimilhança é uma função do parâmetro desconhecido θ.

Estimador de máxima verossimilhança - EMV

Definição: O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}$ que maximiza a função de verossimilhança $L(\theta)$.

Função de verossimilhança - Interpretação

- A fim de interpretar a função de verossimilhança, considere que X é uma variável aleatória discreta.
- Note que a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

  - Em outras palavras, $L(\theta)$ é simplesmente a probabilidade de se obter os valores amostrais $x_1, x_2, \ldots, x_n$.
  - Logo, no caso de X ser discreta, o EMV é simplesmente um valor que maximiza a probabilidade de ocorrência dos valores da amostra.

Função de verossimilhança - Interpretação

- A fim de interpretar a função de verossimilhança, considere que X é uma variável aleatória discreta.
- Note que a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

  - Em outras palavras, $L(\theta)$ é simplesmente a probabilidade de se obter os valores amostrais $x_1, x_2, \ldots, x_n$.
  - Logo, no caso de X ser discreta, o EMV é simplesmente um valor que maximiza a probabilidade de ocorrência dos valores da amostra.

Função de verossimilhança - Interpretação

- A fim de interpretar a função de verossimilhança, considere que X é uma variável aleatória discreta.
- Note que a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

  - Em outras palavras, $L(\theta)$ é simplesmente a probabilidade de se obter os valores amostrais $x_1, x_2, \ldots, x_n$.
  - Logo, no caso de X ser discreta, o EMV é simplesmente um valor que maximiza a probabilidade de ocorrência dos valores da amostra.

Função de verossimilhança - Interpretação

- A fim de interpretar a função de verossimilhança, considere que X é uma variável aleatória discreta.
- Note que a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

  - Em outras palavras, $L(\theta)$ é simplesmente a probabilidade de se obter os valores amostrais $x_1, x_2, \ldots, x_n$.
  - Logo, no caso de X ser discreta, o EMV é simplesmente um valor que maximiza a probabilidade de ocorrência dos valores da amostra.

Função de verossimilhança - Interpretação

- A fim de interpretar a função de verossimilhança, considere que X é uma variável aleatória discreta.
- Note que a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n)$$

  - Em outras palavras, $L(\theta)$ é simplesmente a probabilidade de se obter os valores amostrais $x_1, x_2, \ldots, x_n$.
  - Logo, no caso de X ser discreta, o EMV é simplesmente um valor que maximiza a probabilidade de ocorrência dos valores da amostra.

Função de log-verossimilhança

- O logaritmo natural da função de verossimilhança $L(\theta)$ é chamado de função de log-verossimilhança, isto é

$$l(\theta) = \ln L(\theta)$$

- Não é difícil argumentar que o valor $\hat{\theta}$ que maximiza $L(\theta)$ é o mesmo que maximiza $l(\theta)$, ou seja, o EMV pode ser obtido maximizando a função de log-verossimilhança. De fato, note que se $\hat{\theta}$ é EMV então como a função logaritmo natural é monótona crescente, decorre que

$$L(\hat{\theta}) > L(\theta) \Leftrightarrow \ln L(\hat{\theta}) > \ln L(\theta)$$

para todo θ no espaço paramétrico.

Função de log-verossimilhança

- O logaritmo natural da função de verossimilhança $L(\theta)$ é chamado de função de log-verossimilhança, isto é

$$l(\theta) = \ln L(\theta)$$

- Não é difícil argumentar que o valor $\hat{\theta}$ que maximiza $L(\theta)$ é o mesmo que maximiza $l(\theta)$, ou seja, o EMV pode ser obtido maximizando a função de log-verossimilhança. De fato, note que se $\hat{\theta}$ é EMV então como a função logaritmo natural é monótona crescente, decorre que

$$L(\hat{\theta}) > L(\theta) \Leftrightarrow \ln L(\hat{\theta}) > \ln L(\theta)$$

para todo θ no espaço paramétrico.

Função de log-verossimilhança

- O logaritmo natural da função de verossimilhança $L(\theta)$ é chamado de função de log-verossimilhança, isto é

$$l(\theta) = \ln L(\theta)$$

- Não é difícil argumentar que o valor $\hat{\theta}$ que maximiza $L(\theta)$ é o mesmo que maximiza $l(\theta)$, ou seja, o EMV pode ser obtido maximizando a função de log-verossimilhança. De fato, note que se $\hat{\theta}$ é EMV então como a função logaritmo natural é monótona crescente, decorre que

$$L(\hat{\theta}) > L(\theta) \Leftrightarrow \ln L(\hat{\theta}) > \ln L(\theta)$$

para todo θ no espaço paramétrico.

Função de log-verossimilhança

- O logaritmo natural da função de verossimilhança $L(\theta)$ é chamado de função de log-verossimilhança, isto é

$$l(\theta) = \ln L(\theta)$$

- Não é difícil argumentar que o valor $\hat{\theta}$ que maximiza $L(\theta)$ é o mesmo que maximiza $l(\theta)$, ou seja, o EMV pode ser obtido maximizando a função de log-verossimilhança. De fato, note que se $\hat{\theta}$ é EMV então como a função logaritmo natural é monótona crescente, decorre que

$$L(\hat{\theta}) > L(\theta) \Leftrightarrow \ln L(\hat{\theta}) > \ln L(\theta)$$

para todo θ no espaço paramétrico.

Função de log-verossimilhança

- O logaritmo natural da função de verossimilhança $L(\theta)$ é chamado de função de log-verossimilhança, isto é

$$l(\theta) = \ln L(\theta)$$

- Não é difícil argumentar que o valor $\hat{\theta}$ que maximiza $L(\theta)$ é o mesmo que maximiza $l(\theta)$, ou seja, o EMV pode ser obtido maximizando a função de log-verossimilhança. De fato, note que se $\hat{\theta}$ é EMV então como a função logaritmo natural é monótona crescente, decorre que

$$L(\hat{\theta}) > L(\theta) \Leftrightarrow \ln L(\hat{\theta}) > \ln L(\theta)$$

para todo θ no espaço paramétrico.

Obtendo o EMV

- Portanto, o EMV pode ser obtido como raiz da equação

$$\frac{dl(\theta)}{d\theta} = 0 \tag{1}$$

- Para se concluir que a solução da equação (1) é um ponto de máximo, é necessário verificar se

$$\left.\frac{d^2 l(\theta)}{d\theta^2}\right|_{\theta=\hat{\theta}} < 0 \tag{2}$$

- A equação (1) estabelece que o ponto obtido, $\hat{\theta}$, é um ponto critico e a inequação (2) garante que o ponto obtido em (1) é um ponto de máximo.
- Em algumas situações, a solução de (1) pode ser obtida explicitamente. Em situações mais complexas, a solução de (1) será em geral obtida po meio de procedimentos numéricos.

Obtendo o EMV

- Portanto, o EMV pode ser obtido como raiz da equação

$$\frac{dl(\theta)}{d\theta} = 0 \tag{1}$$

- Para se concluir que a solução da equação (1) é um ponto de máximo, é necessário verificar se

$$\left.\frac{d^2 l(\theta)}{d\theta^2}\right|_{\theta=\hat{\theta}} < 0 \tag{2}$$

- A equação (1) estabelece que o ponto obtido, $\hat{\theta}$, é um ponto critico e a inequação (2) garante que o ponto obtido em (1) é um ponto de máximo.
- Em algumas situações, a solução de (1) pode ser obtida explicitamente. Em situações mais complexas, a solução de (1) será em geral obtida po meio de procedimentos numéricos.

Obtendo o EMV

- Portanto, o EMV pode ser obtido como raiz da equação

$$\frac{dl(\theta)}{d\theta} = 0 \tag{1}$$

- Para se concluir que a solução da equação (1) é um ponto de máximo, é necessário verificar se

$$\left.\frac{d^2 l(\theta)}{d\theta^2}\right|_{\theta=\hat{\theta}} < 0 \tag{2}$$

- A equação (1) estabelece que o ponto obtido, $\hat{\theta}$, é um ponto critico e a inequação (2) garante que o ponto obtido em (1) é um ponto de máximo.
- Em algumas situações, a solução de (1) pode ser obtida explicitamente. Em situações mais complexas, a solução de (1) será em geral obtida po meio de procedimentos numéricos.

Obtendo o EMV

- Portanto, o EMV pode ser obtido como raiz da equação

$$\frac{dl(\theta)}{d\theta} = 0 \tag{1}$$

- Para se concluir que a solução da equação (1) é um ponto de máximo, é necessário verificar se

$$\left.\frac{d^2 l(\theta)}{d\theta^2}\right|_{\theta=\hat{\theta}} < 0 \tag{2}$$

- A equação (1) estabelece que o ponto obtido, $\hat{\theta}$, é um ponto critico e a inequação (2) garante que o ponto obtido em (1) é um ponto de máximo.
- Em algumas situações, a solução de (1) pode ser obtida explicitamente. Em situações mais complexas, a solução de (1) será em geral obtida po meio de procedimentos numéricos.

Obtendo o EMV

- Portanto, o EMV pode ser obtido como raiz da equação

$$\frac{dl(\theta)}{d\theta} = 0 \tag{1}$$

- Para se concluir que a solução da equação (1) é um ponto de máximo, é necessário verificar se

$$\left.\frac{d^2 l(\theta)}{d\theta^2}\right|_{\theta=\hat{\theta}} < 0 \tag{2}$$

- A equação (1) estabelece que o ponto obtido, $\hat{\theta}$, é um ponto critico e a inequação (2) garante que o ponto obtido em (1) é um ponto de máximo.
- Em algumas situações, a solução de (1) pode ser obtida explicitamente. Em situações mais complexas, a solução de (1) será em geral obtida po meio de procedimentos numéricos.

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i - \mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1 - \mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2 - \mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n - \mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i - \mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1 - \mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2 - \mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n - \mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)}$$

Exemplo 1

Exemplo 1: Sejam $X_1, X_2, \dots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \dots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \dots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Exemplo 1: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população normal com média μ e variância 1. Vamos obter o estimador de máxima verossimilhança do parâmetro μ.

- Sol.: Primeiro note que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória proveniente de $X \sim N(\mu, 1)$, então suas densidades são da forma

$$f(x_i, \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

- Logo, a função de verossimilhança é

$$L(\mu) = f(x_1, \mu) \cdot f(x_2, \mu) \cdot \ldots \cdot f(x_n, \mu)$$

ou seja,

$$L(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^{n}2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^{n}(x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^{n}x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^{n}x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left(\left(\sqrt{2\pi}\right)^{-n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^n 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^n (x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^n x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^n x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^n 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^n (x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^n x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^n x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que
$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é
$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos
$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^{n}2(x_i-\mu)(-1) = 0$$
ou seja,
$$\sum_{i=1}^{n}(x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^{n}x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^{n}x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^{n}(x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^{n} x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^{n} x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^{n}2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^{n}(x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^{n}x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^{n}x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^n 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^n (x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^n x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^n x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^n 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^n (x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^n x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^n x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Obtivemos que

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}$$

- Logo, a função de log-verossimilhança é

$$l(\mu) = \ln L(\mu) = \ln\left((\sqrt{2\pi})^{-n} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2}\right) = -n\ln\sqrt{2\pi} - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$$

- Derivando a última equação acima, com respeito a μ, e igualando a zero temos

$$\frac{dl(\mu)}{d\mu} = 0 \Leftrightarrow -\frac{1}{2}\sum_{i=1}^n 2(x_i-\mu)(-1) = 0$$

ou seja,

$$\sum_{i=1}^n (x_i-\mu) = 0 \Leftrightarrow \sum_{i=1}^n x_i - n\mu = 0 \Leftrightarrow \sum_{i=1}^n x_i = n\mu \Leftrightarrow \hat{\mu} = \bar{X}$$

Cont. exemplo 1

- Note ainda que

$$\frac{d^2 l(\mu)}{d\mu^2} = \frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i - \mu)\right) = \frac{d}{d\mu}\left(\sum_{i=1}^{n} x_i - n\mu\right) = -n < 0.$$

- Portanto, temos que o estimador de máxima verossimilhança de μ é $\hat{\mu} = \bar{X}$.

Cont. exemplo 1

- Note ainda que

$$\frac{d^2 l(\mu)}{d\mu^2} = \frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i - \mu)\right) = \frac{d}{d\mu}\left(\sum_{i=1}^{n} x_i - n\mu\right) = -n < 0.$$

- Portanto, temos que o estimador de máxima verossimilhança de μ é $\hat{\mu} = \bar{X}$.

Cont. exemplo 1

- Note ainda que

$$\frac{d^2 l(\mu)}{d\mu^2} = \frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i - \mu)\right) = \frac{d}{d\mu}\left(\sum_{i=1}^{n} x_i - n\mu\right) = -n < 0.$$

- Portanto, temos que o estimador de máxima verossimilhança de μ é $\hat{\mu} = \bar{X}$.

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se } \ x = 0, 1, \\ 0, & \text{se } \ \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se } \ x = 0, 1, \\ 0, & \text{se } \ \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se } \; x = 0, 1, \\ 0, & \text{se } \; \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Exemplo 2

Exemplo 2: Sejam $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com distribuição Bernoulli de parâmetro p, isto é, a função de probabilidade de X é

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{se} \quad x = 0, 1, \\ 0, & \text{se} \quad \text{caso contrário.} \end{cases}$$

Vamos obter o EMV de p.

- Sol. Temos que a função de verossimilhança é

$$L(p) = f(x_1, p) \cdot f(x_2, p) \cdot \ldots \cdot f(x_n, p)$$

ou seja,

$$L(p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

ou seja,

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right) \ln p + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right)\ln p + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right) \ln p + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right) \ln p + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^n x_i\right)\ln p + \left(n - \sum_{i=1}^n x_i\right)\ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^n x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^n x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^n x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^n x_i\right)$$

Cont. exemplo 2

- Obtivemos que

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é

$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right) \ln p + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,

$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Obtivemos que
$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Logo, a função de log-verossimilhança é
$$l(p) = \ln L(p) = \left(\sum_{i=1}^{n} x_i\right)\ln p + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-p)$$

- Derivando a última equação acima, com respeito a p, e igualando a zero temos
$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0$$

ou seja,
$$\frac{dl(p)}{dp} = 0 \Leftrightarrow \frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right)$$

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$.

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$.

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$.

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$.

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$.

Cont. exemplo 2

- Ou ainda

$$\frac{1}{p}\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) \Leftrightarrow \frac{1-p}{p} = \frac{n - \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

simplificando, temos que

- Ou ainda

$$\frac{1}{p} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1 \Leftrightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}.$$

- Não é difícil verificar que $\left.\frac{d^2 l(p)}{dp^2}\right|_{p=\hat{p}} < 0$. Logo, temos que o EMV de p é $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$.

Cont. exemplo 2

- Suponha que, em uma situação prática, estamos querendo estimar a proporção p de peças defeituosas produzida por uma fábrica. Se n itens são selecionados, ao acaso, de uma linha de produção e cada item é considerado como defeituoso ou não defeituoso, então
  - Podemos relacionar o i-ésimo item defeituoso com $x_i = 1$ e o i-ésimo item não defeituoso com $x_i = 0$.
  - Note $\sum_{i=1}^{n} x_i$ é o número de itens defeituosos na amostra. E $\hat{p}$ é a proporção de itens defeituosos na amostra.

Cont. exemplo 2

- Suponha que, em uma situação prática, estamos querendo estimar a proporção p de peças defeituosas produzida por uma fábrica. Se n itens são selecionados, ao acaso, de uma linha de produção e cada item é considerado como defeituoso ou não defeituoso, então
    - Podemos relacionar o i-ésimo item defeituoso com $x_i = 1$ e o i-ésimo item não defeituoso com $x_i = 0$.
    - Note $\sum_{i=1}^{n} x_i$ é o número de itens defeituosos na amostra. E $\hat{p}$ é a proporção de itens defeituosos na amostra.

Cont. exemplo 2

- Suponha que, em uma situação prática, estamos querendo estimar a proporção p de peças defeituosas produzida por uma fábrica. Se n itens são selecionados, ao acaso, de uma linha de produção e cada item é considerado como defeituoso ou não defeituoso, então
  - Podemos relacionar o i-ésimo item defeituoso com $x_i = 1$ e o i-ésimo item não defeituoso com $x_i = 0$.
  - Note $\sum_{i=1}^{n} x_i$ é o número de itens defeituosos na amostra. E $\hat{p}$ é a proporção de itens defeituosos na amostra.

Propriedades do EMV

- O método de máxima verossimilhança é um dos métodos de estimação mais usuais, sendo que os estimadores obtidos por esse métodos possuem boas propriedades estatísticas.
- Se $\hat{\theta}$ for o EMV de θ, então sob condições bem gerais e não restritivas, se n (tamanho da amostra) for grande pode-se mostrar que
  - $\hat{\theta}$ é aproximadamente não viesado para θ, isto é, $E(\hat{\theta}) \approx \theta$;
  - a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador;
  - $\hat{\theta}$ tem uma distribuição normal aproximada.

Propriedades do EMV

- O método de máxima verossimilhança é um dos métodos de estimação mais usuais, sendo que os estimadores obtidos por esse métodos possuem boas propriedades estatísticas.
- Se $\hat{\theta}$ for o EMV de θ, então sob condições bem gerais e não restritivas, se *n* (tamanho da amostra) for grande pode-se mostrar que
  - $\hat{\theta}$ é aproximadamente não viesado para θ, isto é, $E(\hat{\theta}) \approx \theta$;
  - a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador;
  - $\hat{\theta}$ tem uma distribuição normal aproximada.

Propriedades do EMV

- O método de máxima verossimilhança é um dos métodos de estimação mais usuais, sendo que os estimadores obtidos por esse métodos possuem boas propriedades estatísticas.
- Se $\hat{\theta}$ for o EMV de θ, então sob condições bem gerais e não restritivas, se n (tamanho da amostra) for grande pode-se mostrar que
  - $\hat{\theta}$ é aproximadamente não viesado para θ, isto é, $E(\hat{\theta}) \approx \theta$;
  - a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador;
  - $\hat{\theta}$ tem uma distribuição normal aproximada.

Propriedades do EMV

- O método de máxima verossimilhança é um dos métodos de estimação mais usuais, sendo que os estimadores obtidos por esse métodos possuem boas propriedades estatísticas.
- Se $\hat{\theta}$ for o EMV de θ, então sob condições bem gerais e não restritivas, se n (tamanho da amostra) for grande pode-se mostrar que
    - $\hat{\theta}$ é aproximadamente não viesado para θ, isto é, $E(\hat{\theta}) \approx \theta$;
    - a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador;
    - $\hat{\theta}$ tem uma distribuição normal aproximada.

Propriedades do EMV

- O método de máxima verossimilhança é um dos métodos de estimação mais usuais, sendo que os estimadores obtidos por esse métodos possuem boas propriedades estatísticas.
- Se $\hat{\theta}$ for o EMV de θ, então sob condições bem gerais e não restritivas, se n (tamanho da amostra) for grande pode-se mostrar que
  - $\hat{\theta}$ é aproximadamente não viesado para θ, isto é, $E(\hat{\theta}) \approx \theta$;
  - a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida com qualquer outro estimador;
  - $\hat{\theta}$ tem uma distribuição normal aproximada.