Inferência para uma Proporção

Inferência para uma Proporção

- Na prática, os modelos de amostragem Poisson e binomial tem valores para os parâmetros desconhecidos. O método mais utilizado para se estimar parâmetros é o método da máxima verossimilhança.

O Princípio da Máxima Verossimilhança

Afirma que devemos escolher aquele valor do parâmetro desconhecido que ***maximiza*** a probabilidade de obter a amostra observada, ou seja, o valor que torna aquela amostra "***mais provável***" (*verossímil*).

- *Este princípio foi enunciado por Fisher pela primeira vez em 1912.*
- *Em 1922, ele introduziu a expressão "likelihood" quando deu-lhe forma mais completa.*
- *Usa o principio da máxima verossimilhança para obter os estimadores de máxima verossimilhança que, em geral, têm propriedades muito boas.*

Função de Verossimilhança

Seja X uma v.a. com função de probabilidade $p(x;\theta)$ ou função densidade de probabilidade $f(x;\theta)$.

Seja uma amostra aleatória de tamanho ***n*** de X, $(X_1, X_2, \dots, X_n)$, e sejam $(x_1, x_2, \dots, x_n)$ os valores efetivos observados.

A FUNÇÃO DE VEROSSIMILHANÇA é definida por:

$$L(\theta) = L(\theta; x_1, x_2, \dots, x_n) = \begin{cases} \prod_{i=1}^{n} p(x_i;\theta) & se\ X\ v.a.\ discreta \\ \prod_{i=1}^{n} f(x_i;\theta) & se\ X\ v.a.\ contínua \end{cases}$$

o qual é considerada uma função de θ.

Estimação de Máxima Verossimilhança

Seja

$$L(\theta) = L(\theta; x_1, x_2, \ldots, x_n)$$

a função de verossimilhança para $(X_1, X_2, \ldots, X_n)$.

Se

$$\hat{\theta} = g(X_1, X_2, \ldots, X_n)$$

é o valor de θ que maximiza $L(\theta)$, então $\hat{\theta}$ é o estimador de máxima verossimilhança (MV) de θ e

$\hat{\theta} = g(x_1, x_2, \ldots, x_n)$ é a estimativa MV de θ .

Estimação de Máxima Verossimilhança

Sob condições de regularidade da função de verossimilhança, ***o estimador MV*** é a solução da equação:

$$\frac{\partial L(\theta)}{\partial \theta} = 0$$

Temos ainda que $L(\theta)$ e $l(\theta) = \ln L(\theta)$ têm máximo no mesmo valor de θ, e às vezes é mais fácil encontrar o máximo de $l(\theta) = \ln L(\theta)$.

Estimação por Máxima Verossimilhança

- Para um particular modelo de amostragem, podemos substituir os dados da amostra na função de probabilidade e então ver a probabilidade como uma função de valores do parâmetro desconhecido.

- **Exemplo:**

  Em n = 10 ensaios, suponha uma binomial com Y = 0. Da fórmula da binomial com parâmetro π, a probabilidade do resultado é igual a:

$$P(0) = (1-\pi)^{10}$$

Esta probabilidade é definida para valores entre 0 e 1.

Estimação por Máxima Verossimilhança

- A probabilidade dos dados observados, expressa como uma função do parâmetro, é chamada de ***função de verossimilhança***.

  A função de verossimilhança para Y = 0 sucessos em 10 ensaios é dada por:

$$l(\pi) = (1-\pi)^{10}, \quad 0 \le \pi \le 1$$

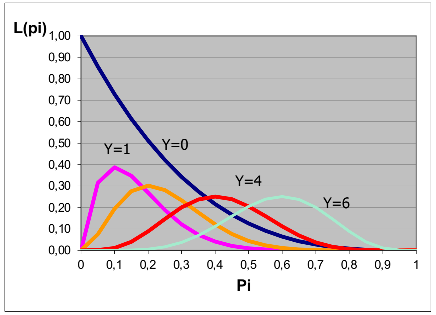
- A estimativa de máxima verossimilhança do parâmetro é definido como sendo o valor na qual a probabilidade dos dados observados assume o seu maior valor. Ou seja, é o valor na qual a função de verossimilhança assume o seu máximo.

Estimação por Máxima Verossimilhança

- *O gráfico de $l(\pi)$ para $n = 10$ e $Y = 0$, assume o máximo no ponto $\pi = 0$.*

- Geralmente, uma binomial com *y* sucessos em *n* ensaios tem estimativa de máxima verossimilhança de π igual a p = y/N. Isto é, a proporção de sucessos da amostra para N ensaios.
- Denotando por 1 o sucesso e 0 a falha. Então a proporção da amostra é igual a média da amostra dos resultados do ensaio.

- *O gráfico de $l(\pi)$ para $n = 10$ e $Y = 0$, assume o máximo no ponto $\pi = 0$.*



- Geralmente, uma Binomial com *y* sucessos em *n* ensaios tem estimador de máxima verossimilhança de π igual a

$$\hat{P} = \frac{Y}{n},$$

isto é, a ***proporção de sucessos da amostra*** para *n* ensaios.

- Denotando por 1 o sucesso e 0 a falha. Então a proporção da amostra é igual a média da amostra dos resultados do ensaio.

Estimação por Máxima Verossimilhança

- A distribuição de Poisson também tem estimativa de máxima verossimilhança igual a média amostral.
- Os estimadores baseados no método da máxima verossimilhança são muito populares porque eles tem propriedades interessantes para grandes amostras:
  - Os estimadores de máxima verossimilhança são aqueles que apresentam os menores erros padrões.
  - Os estimadores de máxima verossimilhança são aproximadamente normais.

Distribuição Amostral de $\hat{P}$

Vamos considerar uma população em que a proporção de elementos portadores de uma certa característica é π.

Assim, a população pode ser considerada como a variável X, tal que:

$$X = \begin{cases} 1 & \text{se o elemento é portador da característica} \\ 0 & \text{se o elemento não é portador da característica} \end{cases}$$

Logo X tem distribuição de Bernoulli e

$$\mu = E(X) = \pi \quad e \quad V(X) = \pi(1-\pi)$$

- Observe que:

$$p = \frac{núm.\,de\ elementos\ da\ população\ com\ a\ característica}{n} = \frac{\sum_{i=1}^{N} X_i}{N}$$

Logo a PROPORÇÃO é uma particular média.

- Retirada uma amostra aleatória simples de tamanho ***n*** dessa população e, se indicarmos por:

$$Y_n = \sum_{i=1}^{n} X_i \rightarrow total\ de\ indivíduos\ com\ a\ característica\ na\ amostra$$

Temos que:

Y: *Binomial* (n,π) e $E(Y_n) = n\pi$ e $V(Y_n) = n\pi(1-\pi)$

- Definindo como $\hat{P}$ a ***proporção de elementos da amostra*** com a característica em estudo, isto é,

$$\hat{P} = \frac{Y_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$$

Logo $\hat{P}$ é também um caso particular de $\overline{X}$ e portanto, são válidos os resultados encontrados para $\overline{X}$

Assim,

$$\mu_{\hat{P}} = E(\hat{P}) = \pi \quad e \quad \sigma^2_{\hat{P}} = V(\hat{P}) = \frac{\pi(1-\pi)}{n}$$

- ***Pequenas Amostras:***

$$P\left(\hat{P} = \frac{y}{n}\right) = P\left(\frac{Y_n}{n} = \frac{y}{n}\right) = P(Y_n = y)$$

$$sendo: \quad Y_n: Binomial\ (n, \pi)$$

- **Grandes Amostras:**
  - Se ***n*** é grande, pelo TEOREMA CENTRAL DO LIMITE (*uma vez que X não é normal*) temos que:

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

Estimação da Proporção Populacional

Deseja-se estimar a proporção de elementos da população (π) que apresentam certa característica.

- ***ESTIMADOR de*** π : $\hat{P}$
  - Distribuição amostral de $\hat{P}$:

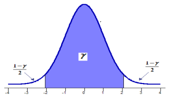
    Se n é "grande" temos que:

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

E consequentemente,

$$Z = \frac{\hat{P} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0,1)$$

Intervalo de Confiança (Υ) para π



Fixada uma confiança Υ tem-se na N(0, 1):

$$P\left(-z < \frac{\widehat{P}-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} < z\right) = \gamma \quad \Rightarrow \quad P\left(\widehat{P} - z\sqrt{\frac{\pi(1-\pi)}{n}} < \pi < \widehat{P} + z\sqrt{\frac{\pi(1-\pi)}{n}}\right) = \gamma$$

Assim:

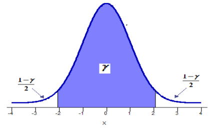
$$\pi \in \left(\widehat{P} - z\sqrt{\frac{\pi(1-\pi)}{n}} \; ; \; \widehat{P} + z\sqrt{\frac{\pi(1-\pi)}{n}}\right)$$

Observe-se, porém, que não se conhece o valore de π e, portanto $\pi(1-\pi)$ é também desconhecido.

- ***Solução:***

$$\pi \in \left(\widehat{P} \mp z\sqrt{\frac{\widehat{p}(1-\widehat{P})}{n}}\right)$$

Intervalo de Confiança (Υ) para p



- ***Solução 1:*** $$p \in \left(\hat{P} - z\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}; \hat{P} + z\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right)$$

- ***Solução 2:*** $$p \in \left(\hat{P} - z\sqrt{\frac{1}{4n}}; \hat{P} + z\sqrt{\frac{1}{4n}}\right)$$

Usando o fato de que
P(1-p)= 1/4