Estatística Computacional

Ronald Targino Nojosa

DEMA-UFC

Notas de Aula 2016
Versão Parcial

Introdução

Na Estatística Computacional abordaremos:

- programa computacional R
- geração de números pseudo-aleatórios
- métodos para geração de variáveis aleatórias
  - inversão, rejeição e composição
- métodos numéricos
  - Monte Carlo, Newton-Raphson, escore de Fisher
- técnicas de reamostragem
  - bootstrap, jackknife

Método de Newton-Raphson

Antes de tratarmos especificamente com o Método de Newton-Raphson vamos abordar o Método da Máxima Verossimilhança.

Método da Máxima Verossimilhança

Método da Máxima Verossimilhança

Método da Máxima Verossimilhança

Motivação 1: Testes de laboratório em amostras de água de um rio

Intesse: Determinar a concentração de bactérias na água

Figura 1 - Bactérias.

Método da Máxima Verossimilhança

Importante: O número de bactérias é determinado para cada volume unitário de amostra de água. São avaliados n volumes (amostras de água) e registrados os números de bactérias para cada volume: $x_1, x_2, \ldots, x_n$.

Figura 2 - Tubos com amostras de água.

Método da Máxima Verossimilhança

Objetivo: Estimar μ, o número médio de bactérias por volume unitário de água do rio.

Suposições: As bactérias são distribuídas aleatoriamente e uniformemente no rio.

Modelo probabilístico a ser empregado: Poisson

Método da Máxima Verossimilhança

Justificativas para o modelo poisson

1. O número de bactérias em um volume v de água do rio assume um dos seguintes valores: 0,1,2,...
2. O número de bactérias nos distintos volumes são independentes.
3. A distribuição do número de bactérias em qualquer volume de água do rio depende apenas do volume e não de que parte do rio o volume foi coletado.
4. A probabilidade de obter duas ou mais bactérias em um volume suficientemente pequeno é desprezível.

Método da Máxima Verossimilhança

Justificativas para o modelo poisson

1. O número de bactérias em um volume v de água do rio assume um dos seguintes valores: 0,1,2,...
2. O número de bactérias nos distintos volumes são independentes.
3. A distribuição do número de bactérias em qualquer volume de água do rio depende apenas do volume e não de que parte do rio o volume foi coletado.
4. A probabilidade de obter duas ou mais bactérias em um volume suficientemente pequeno é desprezível.

Método da Máxima Verossimilhança

Justificativas para o modelo poisson

1. O número de bactérias em um volume v de água do rio assume um dos seguintes valores: 0,1,2,...
2. O número de bactérias nos distintos volumes são independentes.
3. A distribuição do número de bactérias em qualquer volume de água do rio depende apenas do volume e não de que parte do rio o volume foi coletado.
4. A probabilidade de obter duas ou mais bactérias em um volume suficientemente pequeno é desprezível.

Método da Máxima Verossimilhança

Justificativas para o modelo poisson

1. O número de bactérias em um volume v de água do rio assume um dos seguintes valores: 0,1,2,...
2. O número de bactérias nos distintos volumes são independentes.
3. A distribuição do número de bactérias em qualquer volume de água do rio depende apenas do volume e não de que parte do rio o volume foi coletado.
4. A probabilidade de obter duas ou mais bactérias em um volume suficientemente pequeno é desprezível.

Método da Máxima Verossimilhança

Estas condições[1] permitem estabelecer que as probabilidades associadas ao número de bactérias (X) em um volume v de água do rio são regidas pelo modelo

$$f(x) = P(X = x) = \frac{e^{-\lambda v}(\lambda v)^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

em que λ é a densidade esperada de bactérias e é constante para todo o volume. Portanto, X é uma variável aleatória com distribuição de Poisson.

Nota: $\lambda v = E(X) \Rightarrow \lambda = \dfrac{E(X)}{v}$

[1] ver duas outras condições no livro texto

Método da Máxima Verossimilhança

Outros exemplos de variáveis aleatórias modeladas por uma distribuição de poisson:

- Número de chamadas em uma central telefônica durante um período de comprimento t
- Número de glóbulos sanguíneos visíveis a um microscópio de área visível de a unidades quadradas
- Número de estrelas encontradas em um volume v da Via Láctea

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Desde que os volumes disjuntos são independentes, a probabilidade das n contagens observadas é dada por:

$$f(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x!}$$

2. Função de verossimilhança: $L(\mu) = \mu^{\sum x_i} e^{(-n\mu)}$
3. Função de log-verossimilhança: $l(\mu) = logL(\mu) = \sum x_i log(\mu) - n\mu$
4. Função Escore: $U(\mu) = \dfrac{\partial l(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu} - n$
5. Função de Informação: $I(\mu) = -\dfrac{\partial U(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Desde que os volumes disjuntos são independentes, a probabilidade das n contagens observadas é dada por:

$$f(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x!}$$

2. Função de verossimilhança: $L(\mu) = \mu^{\sum x_i} e^{(-n\mu)}$
3. Função de log-verossimilhança: $l(\mu) = logL(\mu) = \sum x_i log(\mu) - n\mu$
4. Função Escore: $U(\mu) = \dfrac{\partial l(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu} - n$
5. Função de Informação: $I(\mu) = -\dfrac{\partial U(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Desde que os volumes disjuntos são independentes, a probabilidade das n contagens observadas é dada por:

$$f(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x!}$$

2. Função de verossimilhança: $L(\mu) = \mu^{\sum x_i} e^{(-n\mu)}$
3. Função de log-verossimilhança: $l(\mu) = logL(\mu) = \sum x_i log(\mu) - n\mu$
4. Função Escore: $U(\mu) = \frac{\partial l(\mu)}{\partial \mu} = \frac{\sum x_i}{\mu} - n$
5. Função de Informação: $I(\mu) = -\frac{\partial U(\mu)}{\partial \mu} = \frac{\sum x_i}{\mu^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Desde que os volumes disjuntos são independentes, a probabilidade das n contagens observadas é dada por:

$$f(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x!}$$

2. Função de verossimilhança: $L(\mu) = \mu^{\sum x_i} e^{(-n\mu)}$
3. Função de log-verossimilhança: $l(\mu) = logL(\mu) = \sum x_i log(\mu) - n\mu$
4. Função Escore: $U(\mu) = \dfrac{\partial l(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu} - n$
5. Função de Informação: $I(\mu) = -\dfrac{\partial U(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Desde que os volumes disjuntos são independentes, a probabilidade das n contagens observadas é dada por:

$$f(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x_i}}{x!}$$

2. Função de verossimilhança: $L(\mu) = \mu^{\sum x_i} e^{(-n\mu)}$
3. Função de log-verossimilhança: $l(\mu) = logL(\mu) = \sum x_i log(\mu) - n\mu$
4. Função Escore: $U(\mu) = \dfrac{\partial l(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu} - n$
5. Função de Informação: $I(\mu) = -\dfrac{\partial U(\mu)}{\partial \mu} = \dfrac{\sum x_i}{\mu^2}$

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{\sum x_i}{\hat{\mu}} - n = 0 \Rightarrow \hat{\mu} = \bar{x}$$

7. Como $I(\hat{\mu}) > 0$, $\hat{\mu} = \bar{x}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(\mu) \to 0$ quando $\mu \to \infty$, $\hat{\mu} = \bar{x}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para μ é $\hat{\mu} = \bar{x}$. Assim, para maximizar a probabilidade da ocorrência dos dados $x_1, x_2, \ldots, x_n$, estimamos a média populacional μ pela média amostral $\bar{x}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{\sum x_i}{\hat{\mu}} - n = 0 \Rightarrow \hat{\mu} = \bar{x}$$

7. Como $I(\hat{\mu}) > 0$, $\hat{\mu} = \bar{x}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(\mu) \to 0$ quando $\mu \to \infty$, $\hat{\mu} = \bar{x}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para μ é $\hat{\mu} = \bar{x}$. Assim, para maximizar a probabilidade da ocorrência dos dados $x_1, x_2, \ldots, x_n$, estimamos a média populacional μ pela média amostral $\bar{x}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{\sum x_i}{\hat{\mu}} - n = 0 \Rightarrow \hat{\mu} = \bar{x}$$

7. Como $I(\hat{\mu}) > 0$, $\hat{\mu} = \bar{x}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(\mu) \to 0$ quando $\mu \to \infty$, $\hat{\mu} = \bar{x}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para μ é $\hat{\mu} = \bar{x}$. Assim, para maximizar a probabilidade da ocorrência dos dados $x_1, x_2, \ldots, x_n$, estimamos a média populacional μ pela média amostral $\bar{x}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{\sum x_i}{\hat{\mu}} - n = 0 \Rightarrow \hat{\mu} = \bar{x}$$

7. Como $I(\hat{\mu}) > 0$, $\hat{\mu} = \bar{x}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(\mu) \to 0$ quando $\mu \to \infty$, $\hat{\mu} = \bar{x}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para μ é $\hat{\mu} = \bar{x}$. Assim, para maximizar a probabilidade da ocorrência dos dados $x_1, x_2, \ldots, x_n$, estimamos a média populacional μ pela média amostral $\bar{x}$.

Método da Máxima Verossimilhança

Motivação 2: Considerando que não é possível contar o número de bactérias em uma amostra de água, mas apenas determinar se elas estão ou não presentes, devemos ter um outro cenário para o experimento.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar n tubos com volume v de água. Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: y tubos dos n, apresentaram resultado negativo.

Método da Máxima Verossimilhança

Motivação 2: Considerando que não é possível contar o número de bactérias em uma amostra de água, mas apenas determinar se elas estão ou não presentes, devemos ter um outro cenário para o experimento.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar n tubos com volume v de água. Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: y tubos dos n, apresentaram resultado negativo.

Método da Máxima Verossimilhança

Motivação 2: Considerando que não é possível contar o número de bactérias em uma amostra de água, mas apenas determinar se elas estão ou não presentes, devemos ter um outro cenário para o experimento.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar n tubos com volume v de água. Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: y tubos dos n, apresentaram resultado negativo.

Método da Máxima Verossimilhança

Motivação 2: Considerando que não é possível contar o número de bactérias em uma amostra de água, mas apenas determinar se elas estão ou não presentes, devemos ter um outro cenário para o experimento.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar n tubos com volume v de água. Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: y tubos dos n, apresentaram resultado negativo.

Método da Máxima Verossimilhança

Pergunta: Qual o número médio de bactérias por volume unitário de água do rio?

Resposta: Será dada pelo estimador de máxima verossimilhança para μ.

Método da Máxima Verossimilhança

Partindo da amostra observada, temos n tubos e, destes, y apresentaram resultado negativo (sem bactérias). Admitindo condições experimentais que atendam ao modelo binomial, definimos a função de probabilidade por

$$f(y) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y},$$

em que p é a probabilidade de um resultado negativo em um tubo.

Método da Máxima Verossimilhança

Com as condições do exemplo 1, a probabilidade de que haja x bactérias em um volume v de água do rio é dado por uma distribuição de Poisson de parâmetro μv. Assim,

$$p = P(X = 0) = e^{-\mu v}, \quad 0 \leq \mu < \infty.$$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança:

1. Função de probabilidade: $f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad p = e^{-\mu v}$
2. Função de verossimilhança: $L(p) = p^y (1-p)^{n-y}$
3. Função de log-verossimilhança: $l(p) = ylog(p) + (n-y)log(1-p)$
4. Função Escore: $U(p) = \frac{\partial l(p)}{\partial p} = \frac{y}{p} - \frac{n-y}{1-p}$
5. Função de Informação: $I(p) = -\frac{\partial U(p)}{\partial p} = \frac{y}{p^2} + \frac{n-y}{(1-p)^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança:

1. Função de probabilidade: $f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad p = e^{-\mu v}$
2. Função de verossimilhança: $L(p) = p^y (1-p)^{n-y}$
3. Função de log-verossimilhança: $l(p) = y log(p) + (n-y) log(1-p)$
4. Função Escore: $U(p) = \frac{\partial l(p)}{\partial p} = \frac{y}{p} - \frac{n-y}{1-p}$
5. Função de Informação: $I(p) = -\frac{\partial U(p)}{\partial p} = \frac{y}{p^2} + \frac{n-y}{(1-p)^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança:

1. Função de probabilidade: $f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad p = e^{-\mu v}$
2. Função de verossimilhança: $L(p) = p^y (1-p)^{n-y}$
3. Função de log-verossimilhança: $l(p) = ylog(p) + (n-y)log(1-p)$
4. Função Escore: $U(p) = \dfrac{\partial l(p)}{\partial p} = \dfrac{y}{p} - \dfrac{n-y}{1-p}$
5. Função de Informação: $I(p) = -\dfrac{\partial U(p)}{\partial p} = \dfrac{y}{p^2} + \dfrac{n-y}{(1-p)^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança:

1. Função de probabilidade: $f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad p = e^{-\mu v}$
2. Função de verossimilhança: $L(p) = p^y (1-p)^{n-y}$
3. Função de log-verossimilhança: $l(p) = y log(p) + (n-y) log(1-p)$
4. Função Escore: $U(p) = \dfrac{\partial l(p)}{\partial p} = \dfrac{y}{p} - \dfrac{n-y}{1-p}$
5. Função de Informação: $I(p) = -\dfrac{\partial U(p)}{\partial p} = \dfrac{y}{p^2} + \dfrac{n-y}{(1-p)^2}$

Método da Máxima Verossimilhança

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança:

1. Função de probabilidade: $f(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad p = e^{-\mu v}$
2. Função de verossimilhança: $L(p) = p^y (1-p)^{n-y}$
3. Função de log-verossimilhança: $l(p) = y log(p) + (n-y) log(1-p)$
4. Função Escore: $U(p) = \dfrac{\partial l(p)}{\partial p} = \dfrac{y}{p} - \dfrac{n-y}{1-p}$
5. Função de Informação: $I(p) = -\dfrac{\partial U(p)}{\partial p} = \dfrac{y}{p^2} + \dfrac{n-y}{(1-p)^2}$

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{y}{\hat{p}} - \frac{n-y}{1-\hat{p}} = 0 \quad \Rightarrow \quad \hat{p} = \frac{y}{n}$$

7. Como $I(\hat{p}) > 0$, $\hat{p} = \dfrac{y}{n}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(1) = 0$, $\hat{p} = \dfrac{y}{n}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para p é $\hat{p} = \dfrac{y}{n}$.
10. Pelo princípio da invariância, o estimador de máxima verossimilhança de μ é dado por $\hat{p} = e^{-\hat{\mu}v}$. Portanto, $\hat{\mu} = -\dfrac{1}{v} log(\hat{p}) = \dfrac{log(n) - log(y)}{v}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{y}{\hat{p}} - \frac{n-y}{1-\hat{p}} = 0 \quad \Rightarrow \quad \hat{p} = \frac{y}{n}$$

7. Como $I(\hat{p}) > 0$, $\hat{p} = \frac{y}{n}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(1) = 0$, $\hat{p} = \frac{y}{n}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para p é $\hat{p} = \frac{y}{n}$.
10. Pelo princípio da invariância, o estimador de máxima verossimilhança de μ é dado por $\hat{p} = e^{-\hat{\mu}v}$. Portanto, $\hat{\mu} = -\frac{1}{v}log(\hat{p}) = \frac{log(n) - log(y)}{v}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{y}{\hat{p}} - \frac{n-y}{1-\hat{p}} = 0 \quad \Rightarrow \quad \hat{p} = \frac{y}{n}$$

7. Como $I(\hat{p}) > 0$, $\hat{p} = \dfrac{y}{n}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(1) = 0$, $\hat{p} = \dfrac{y}{n}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para p é $\hat{p} = \dfrac{y}{n}$.
10. Pelo princípio da invariância, o estimador de máxima verossimilhança de μ é dado por $\hat{p} = e^{-\hat{\mu}v}$. Portanto, $\hat{\mu} = -\dfrac{1}{v}log(\hat{p}) = \dfrac{log(n) - log(y)}{v}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{y}{\hat{p}} - \frac{n-y}{1-\hat{p}} = 0 \quad \Rightarrow \quad \hat{p} = \frac{y}{n}$$

7. Como $I(\hat{p}) > 0$, $\hat{p} = \dfrac{y}{n}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(1) = 0$, $\hat{p} = \dfrac{y}{n}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para p é $\hat{p} = \dfrac{y}{n}$.
10. Pelo princípio da invariância, o estimador de máxima verossimilhança de μ é dado por $\hat{p} = e^{-\hat{\mu}v}$. Portanto, $\hat{\mu} = -\dfrac{1}{v}log(\hat{p}) = \dfrac{log(n) - log(y)}{v}$.

Método da Máxima Verossimilhança

6. Igualando a Função Escore a zero, temos:

$$\frac{y}{\hat{p}} - \frac{n-y}{1-\hat{p}} = 0 \quad \Rightarrow \quad \hat{p} = \frac{y}{n}$$

7. Como $I(\hat{p}) > 0$, $\hat{p} = \dfrac{y}{n}$ é máximo relativo.
8. Sendo $L(0) = 0$ e $L(1) = 0$, $\hat{p} = \dfrac{y}{n}$ é máximo geral.
9. Conclusão: O estimador de máxima verosimilhança para p é $\hat{p} = \dfrac{y}{n}$.
10. Pelo princípio da invariância, o estimador de máxima verossimilhança de μ é dado por $\hat{p} = e^{-\hat{\mu}v}$. Portanto, $\hat{\mu} = -\dfrac{1}{v}log(\hat{p}) = \dfrac{log(n) - log(y)}{v}$.

Método da Máxima Verossimilhança

Para finalizar, suponha que $n = 40$ tubos, com $v = 10$ml cada, são testados. Se 28 são "fracassos" e 12 são "sucessos", teremos

$$\hat{\mu} = \frac{log(40) - log(12)}{10} = 0{,}0357.$$

A concentração de bactérias no rio é estimada em $0{,}0357$ por ml.

Obs.: Se observarmos $y = 0$, teremos $\hat{\mu} = \infty$. Neste caso, seria importante indicar um intervalo de valores para μ. Isto pode ser feito examinando a função de verossimilhança relativa.

Método da Máxima Verossimilhança

Para finalizar, suponha que $n = 40$ tubos, com $v = 10$ml cada, são testados. Se 28 são "fracassos" e 12 são "sucessos", teremos

$$\hat{\mu} = \frac{log(40) - log(12)}{10} = 0{,}0357.$$

A concentração de bactérias no rio é estimada em $0{,}0357$ por ml.

Obs.: Se observarmos $y = 0$, teremos $\hat{\mu} = \infty$. Neste caso, seria importante indicar um intervalo de valores para μ. Isto pode ser feito examinando a função de verossimilhança relativa.

Método de Newton

Método de Newton

Método de Newton

Intesse: Encontrar a raiz $\hat{\theta}$ da equação $g(\theta) = 0$.

Considerações: Seja θ_0 um valor do parâmetro próximo a $\hat{\theta}$ e considere a expansão em série de Taylor de $g(\theta)$ para $\theta = \theta_0$:

$$g(\theta)=g(\theta_0)+\frac{g'(\theta_0)(\theta-\theta_0)^1}{1!}+\frac{g''(\theta_0)(\theta-\theta_0)^2}{2!}+\ldots+\frac{g^n(\theta_0)(\theta-\theta_0)^n}{n!}+\ldots$$

ara $|\theta - \theta_0|$ pequeno, o termo quadrático ou superior da expansão será pequeno. Eliminando estes termos teremos:

$$g(\theta) \approx g(\theta_0) + g'(\theta_0)(\theta - \theta_0).$$

Método de Newton

Intesse: Encontrar a raiz $\hat{\theta}$ da equação $g(\theta) = 0$.

Considerações: Seja θ_0 um valor do parâmetro próximo a $\hat{\theta}$ e considere a expansão em série de Taylor de $g(\theta)$ para $\theta = \theta_0$:

$$g(\theta) = g(\theta_0) + \frac{g'(\theta_0)(\theta - \theta_0)^1}{1!} + \frac{g''(\theta_0)(\theta - \theta_0)^2}{2!} + \ldots + \frac{g^n(\theta_0)(\theta - \theta_0)^n}{n!} + \ldots$$

ara $|\theta - \theta_0|$ pequeno, o termo quadrático ou superior da expansão será pequeno. Eliminando estes termos teremos:

$$g(\theta) \approx g(\theta_0) + g'(\theta_0)(\theta - \theta_0).$$

Método de Newton

Intesse: Encontrar a raiz $\hat{\theta}$ da equação $g(\theta) = 0$.

Considerações: Seja θ_0 um valor do parâmetro próximo a $\hat{\theta}$ e considere a expansão em série de Taylor de $g(\theta)$ para $\theta = \theta_0$:

$$g(\theta) = g(\theta_0) + \frac{g'(\theta_0)(\theta - \theta_0)^1}{1!} + \frac{g''(\theta_0)(\theta - \theta_0)^2}{2!} + \ldots + \frac{g^n(\theta_0)(\theta - \theta_0)^n}{n!} + \ldots$$

Para $|\theta - \theta_0|$ pequeno, o termo quadrático ou superior da expansão será pequeno. Eliminando estes termos teremos:

$$g(\theta) \approx g(\theta_0) + g'(\theta_0)(\theta - \theta_0).$$

Método de Newton

Estamos aproximando $g(\theta)$ por uma função linear de θ que tem o mesmo valor e inclinação de $g(\theta)$ no ponto $\theta = \theta_0$.

esde que $g(\hat{\theta}) = 0$, teremos:

$$
\begin{aligned}
g(\hat{\theta}) &\approx g(\theta_0) + g'(\theta_0)(\hat{\theta} - \theta_0) \Rightarrow \\
0 &\approx g(\theta_0) + g'(\theta_0)(\hat{\theta} - \theta_0) \Rightarrow \\
\hat{\theta} &\approx \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)}
\end{aligned}
$$

Método de Newton

Estamos aproximando $g(\theta)$ por uma função linear de θ que tem o mesmo valor e inclinação de $g(\theta)$ no ponto $\theta = \theta_0$.

Desde que $g(\hat{\theta}) = 0$, teremos:

$$
\begin{aligned}
g(\hat{\theta}) &\approx g(\theta_0) + g'(\theta_0)(\hat{\theta} - \theta_0) \Rightarrow \\
0 &\approx g(\theta_0) + g'(\theta_0)(\hat{\theta} - \theta_0) \Rightarrow \\
\hat{\theta} &\approx \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)}
\end{aligned}
$$

Método de Newton

Note que θ_0 é apenas uma "estimativa qualquer" para a raiz $\hat{\theta}$ da equação $g(\theta)$. A proposta do método é atualizar esse valor proposto da seguinte forma:

1. determinar $\theta_1 = \theta_0 - \dfrac{g(\theta_0)}{g'(\theta_0)}$
2. determinar $\theta_2 = \theta_1 - \dfrac{g(\theta_1)}{g'(\theta_1)}$
3. repetir esse processo até que $\theta_{i+1} \approx \theta_i$
4. θ_i será a raiz da equação.

Método de Newton

Método de Newton - Ilustração

O método de Newton trata com a equação da reta que passa pelo ponto $(\theta_i, g(\theta_i))$ e é tangente à curva nesse ponto:

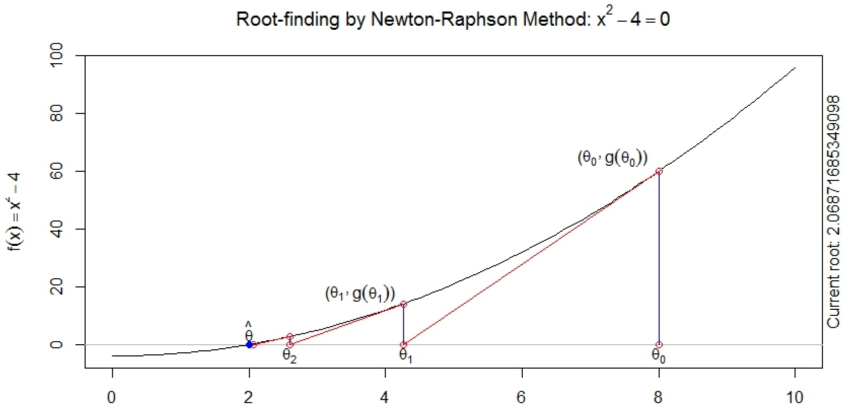
$$g(\theta) - g(\theta_i) = g'(\theta_i)(\theta - \theta_i), \quad i = 0, 1, 2, \ldots$$

em que $g'(\theta_i)$ é a inclinação da reta (derivada). Essa reta passa pelo ponto $(\theta_{i+1}, 0)$. Assim,

$$\begin{aligned} 0 - g(\theta_i) &= g'(\theta_i)(\theta_{i+1} - \theta_i) \Rightarrow \\ \theta_{i+1} &= \theta_i - \frac{g(\theta_i)}{g'(\theta_i)} \end{aligned}$$

Método de Newton

Figura 3 - Método de Newton - Ilustração



Método de Newton

O Método de Newton na Estatística

Método de Newton na Estatística

Considere que $\hat{\theta}$, o estimador de máxima verossimilhança do parâmetro θ, pode ser obtido de $U(\theta) = 0$, ou seja, $\hat{\theta}$ é a raiz dessa equação. Portanto, o procedimento de Newton se estabelece fazendo $g(\theta) = U(\theta)$ e $g'(\theta) = U'(\theta) = -I(\theta)$ e é dado por

$$\theta_{i+1} = \theta_i + \frac{U(\theta_i)}{I(\theta_i)}, \quad i = 0, 1, 2, \dots$$

Quando $\theta_{i+1} \approx \theta_i$, o estimador de máxima verossimilhança será $\hat{\theta} = \theta_{i+1}$.

Método de Newton na Estatística

Observações:

- Com bons chutes iniciais, 3 ou 4 iterações são suficientes para a convergência. Isso se deve ao fato de que para grandes amostras, $U(\theta)$ é próxima da linearidade em θ.
- Se $U(\theta)$ é exatamente linear em θ, o método produz $\hat{\theta}$ em uma única iteração.
- Se $U(\theta) = 0$ tem mais de uma raiz, o método não necessariamente converge para a raiz desejada. Dificuldades também surgem se o máximo ocorre no limite ou próximo a um limite do espaço paramétrico. Sugere-se examinar o gráfico de $l(\theta)$ antes de aplicar o método de Newton.

Método de Newton na Estatística

Motivação 3: No exemplo anterior (motivação 2) é necessário especificar o volume v de água do rio nos tubos. Se v for grande (pequeno), mais provavelmente os testes serão positivos (negativos). Propõe-se, então, usar tubos com dois ou mais diferentes volumes.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar 40 tubos de 10ml (procedimento 1) e 40 tubos de 1ml (procedimento 2). Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: 28 tubos de 10ml e 37 de 1ml apresentaram resultados negativos.

Método de Newton na Estatística

Motivação 3: No exemplo anterior (motivação 2) é necessário especificar o volume v de água do rio nos tubos. Se v for grande (pequeno), mais provavelmente os testes serão positivos (negativos). Propõe-se, então, usar tubos com dois ou mais diferentes volumes.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar 40 tubos de 10ml (procedimento 1) e 40 tubos de 1ml (procedimento 2). Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: 28 tubos de 10ml e 37 de 1ml apresentaram resultados negativos.

Método de Newton na Estatística

Motivação 3: No exemplo anterior (motivação 2) é necessário especificar o volume v de água do rio nos tubos. Se v for grande (pequeno), mais provavelmente os testes serão positivos (negativos). Propõe-se, então, usar tubos com dois ou mais diferentes volumes.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar 40 tubos de 10ml (procedimento 1) e 40 tubos de 1ml (procedimento 2). Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: 28 tubos de 10ml e 37 de 1ml apresentaram resultados negativos.

Método de Newton na Estatística

Motivação 3: No exemplo anterior (motivação 2) é necessário especificar o volume v de água do rio nos tubos. Se v for grande (pequeno), mais provavelmente os testes serão positivos (negativos). Propõe-se, então, usar tubos com dois ou mais diferentes volumes.

Intesse: Determinar se a água está contaminada

Procedimento: Avaliar 40 tubos de 10ml (procedimento 1) e 40 tubos de 1ml (procedimento 2). Para cada tubo registrar:

- teste negativo: se não há bactérias presentes
- teste positivo: se há pelo menos uma bactéria presente

Resultado do procedimento: 28 tubos de 10ml e 37 de 1ml apresentaram resultados negativos.

Método de Newton na Estatística

Pergunta: Qual o número médio de bactérias por volume unitário de água do rio?

Resposta: Será dada pelo estimador de máxima verossimilhança para μ.

Método de Newton na Estatística

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Função de probabilidade:

$$f(y_1, y_2) = f(y_1).f(y_2) = \binom{40}{28} p_1^{28}(1-p_1)^{12}.\binom{40}{37} p_2^{37}(1-p_2)^{3},$$

   com $p_1 = e^{-\mu v_1}$ e $p_2 = e^{-\mu v_2}$. Como os procedimentos são independentes podemos tratá-los separadamente.

2. Funções de verossimilhança:

$$\begin{aligned} L_1(\mu) &= p_1^{28}(1-p_1)^{12}, \quad \text{com} \quad p_1 = e^{-10\mu} \\ L_2(\mu) &= p_2^{37}(1-p_2)^{03}, \quad \text{com} \quad p_2 = e^{-\mu} \end{aligned}$$

   A verossimilhança baseada nos 80 tubos é

$$L(\mu) = L_1(\mu).L_2(\mu).$$

Método de Newton na Estatística

Estimação de μ, o número médio de bactérias por volume unitário de água do rio, pelo método da máxima verossimilhança.

1. Função de probabilidade:

$$f(y_1,y_2)=f(y_1).f(y_2)=\binom{40}{28}p_1^{28}(1-p_1)^{12}.\binom{40}{37}p_2^{37}(1-p_2)^{3},$$

   com $p_1=e^{-\mu v_1}$ e $p_2=e^{-\mu v_2}$. Como os procedimentos são independentes podemos tratá-los separadamente.

2. Funções de verossimilhança:

$$\begin{aligned} L_1(\mu) &= p_1^{28}(1-p_1)^{12}, \quad \text{com} \quad p_1=e^{-10\mu} \\ L_2(\mu) &= p_2^{37}(1-p_2)^{03}, \quad \text{com} \quad p_2=e^{-\mu} \end{aligned}$$

   A verossimilhança baseada nos 80 tubos é

$$L(\mu)=L_1(\mu).L_2(\mu).$$

Método de Newton na Estatística

3. Funções de log-verossimilhança:

$$\begin{aligned} l_1(\mu) &= 28\log p_1 + 12\log(1-p_1), \quad \text{com} \quad p_1 = e^{-10\mu} \\ l_2(\mu) &= 37\log p_2 + 3\log(1-p_2), \quad \text{com} \quad p_2 = e^{-\mu} \end{aligned}$$

Com, $l(\mu) = l_1(\mu) + l_2(\mu)$.

Método de Newton na Estatística

4. Funções Escores:

$$\begin{aligned} U_1(\mu) &= \frac{\partial l_1(\mu)}{\partial \mu} = \frac{\partial l_1(\mu)}{\partial p_1} \times \frac{\partial p_1}{\partial \mu} = \frac{120}{1-p_1} - 400 \\ U_2(\mu) &= \frac{\partial l_2(\mu)}{\partial \mu} = \frac{\partial l_2(\mu)}{\partial p_2} \times \frac{\partial p_2}{\partial \mu} = \frac{3}{1-p_2} - 40 \end{aligned}$$

Com, $U(\mu) = U_1(\mu) + U_2(\mu)$.

Método de Newton na Estatística

5. Funções de Informação:

$$
\begin{aligned}
I_1(\mu) &= -\frac{\partial U_1(\mu)}{\partial \mu} = \frac{\partial U_1(\mu)}{\partial p_1} \times \frac{\partial p_1}{\partial \mu} = \frac{1200 p_1}{(1-p_1)^2} \\
I_2(\mu) &= -\frac{\partial U_2(\mu)}{\partial \mu} = \frac{\partial U_2(\mu)}{\partial p_2} \times \frac{\partial p_2}{\partial \mu} = \frac{3 p_2}{(1-p_2)^2}
\end{aligned}
$$

Com, $I(\mu) = I_1(\mu) + I_2(\mu)$.

Método de Newton na Estatística

Todas as funções foram definidas. Agora é encontrar a raiz $\hat{\mu}$ da equação $U(\mu) = 0$. Entretanto, não há solução algébrica para $U(\mu) = 0$ e, portanto, $\hat{\mu}$ será determinada numericamente.

Uma opção: calcular $U(\mu)$ para vários valorees de μ e encontrar em valor aproximado de μ para o qual $U(\mu) = 0$.

Uma possível solução: usar o método de Newton.

Método de Newton na Estatística

Todas as funções foram definidas. Agora é encontrar a raiz $\hat{\mu}$ da equação $U(\mu) = 0$. Entretanto, não há solução algébrica para $U(\mu) = 0$ e, portanto, $\hat{\mu}$ será determinada numericamente.

Uma opção: calcular $U(\mu)$ para vários valorees de μ e encontrar em valor aproximado de μ para o qual $U(\mu) = 0$.

Uma possível solução: usar o método de Newton.

Método de Newton na Estatística

Todas as funções foram definidas. Agora é encontrar a raiz $\hat{\mu}$ da equação $U(\mu) = 0$. Entretanto, não há solução algébrica para $U(\mu) = 0$ e, portanto, $\hat{\mu}$ será determinada numericamente.

Uma opção: calcular $U(\mu)$ para vários valorees de μ e encontrar em valor aproximado de μ para o qual $U(\mu) = 0$.

Uma possível solução: usar o método de Newton.

Método de Newton na Estatística

Temos:

$$
\begin{aligned}
U(\mu) &= \frac{120}{1-p_1} - 400 + \frac{3}{1-p_2} - 40 \\
I(\mu) &= \frac{1200p_1}{(1-p_1)^2} + \frac{3p_2}{(1-p_2)^2}
\end{aligned}
$$

com $p_1 = e^{-10\mu}$ e $p_2 = e^{-\mu}$.

Para implementar o método de Newton, é necessário um valor inicial μ_0 para μ.

$$
\mu_{i+1} = \mu_i + \frac{U(\mu_i)}{I(\mu_i)}, \quad i = 0, 1, 2, \ldots
$$

Método de Newton na Estatística

Proposta para o valor inicial μ_0: média das estimativas de μ em cada procedimento.

P1: $\hat{\mu} = \dfrac{log(n) - log(y)}{v} = \dfrac{log(40) - log(28)}{10} = 0{,}03567$

P2: $\hat{\mu} = \dfrac{log(n) - log(y)}{v} = \dfrac{log(40) - log(37)}{1} = 0{,}07796$

Portanto, $\mu_0 = 0{,}0568$.

Método de Newton na Estatística

Alguns resultados

Iteração 1:
$U(\mu_0) = -108{,}7516$

$I(\mu_0) = 4550{,}7180$

$$\mu_1 = 0{,}0568 - \frac{108{,}7516}{4550{,}7180} = 0{,}03290$$

Iteração 2:
$U(\mu_1) = 80{,}7194$

$I(\mu_1) = 13758{,}2500$

$$\mu_2 = 0{,}03253 - \frac{80{,}7194}{13758{,}2500} = 0{,}03877$$

O algoritmo converge em 5 iterações, o EMV para μ é $0{,}04005$.

Método de Newton na Estatística

Alguns resultados

Iteração 1:

$U(\mu_0) = -108{,}7516$

$I(\mu_0) = 4550{,}7180$

$$\mu_1 = 0{,}0568 - \frac{108{,}7516}{4550{,}7180} = 0{,}03290$$

Iteração 2:

$U(\mu_1) = 80{,}7194$

$I(\mu_1) = 13758{,}2500$

$$\mu_2 = 0{,}03253 - \frac{80{,}7194}{13758{,}2500} = 0{,}03877$$

O algoritmo converge em 5 iterações, o EMV para μ é $0{,}04005$.

Método de Newton na Estatística

Alguns resultados

Iteração 1:

$U(\mu_0) = -108{,}7516$

$I(\mu_0) = 4550{,}7180$

$$\mu_1 = 0{,}0568 - \frac{108{,}7516}{4550{,}7180} = 0{,}03290$$

Iteração 2:

$U(\mu_1) = 80{,}7194$

$I(\mu_1) = 13758{,}2500$

$$\mu_2 = 0{,}03253 - \frac{80{,}7194}{13758{,}2500} = 0{,}03877$$

O algoritmo converge em 5 iterações, o EMV para μ é $0{,}04005$.

Método de Newton-Raphson

Método de Newton-Raphson

Método Escore de Fisher

Método Escore de Fisher

Referências

1. Ross, Sheldon M. Simulation. 4th ed. San Diego, California: Elsevier, 2006.
2. Efron, B. (1979). Bootstrap methods: another look at the jackknife. The Annals of Statistics 7: 1-25.
3. Efron, B. e Tibshirani, R. J. (1993). An Introducion to the bootstrap. New York: Chapman e Hall.
4. Davidson, A. C. e Hinkley, D. V. (1997). Bootstrap methods e their applications. New York: Cambridge University Press.
5. Chernick, M.R. (1999). Bootstrap methods: A practitioner's guide. New York: Wiley.

Referências

6. Efon, B. (1982) The jackknife, the bootstrap and other resampling plans. Philadelphia: Society for Industrial and Applied Mathematics.
7. Godfrey, L. (2009). Bootstrap tests for regression models. New York: Palgrave MacMillan.
8. Hall, P. (1992). The bootstrap and Edgeworth expansion. New York: Springer–Verlag.
9. Shao, J. & Tu, D. (1995). The jackknife and bootstrap. New York: Springer.
10. Rubin, D.B. (1981). The bayesian bootstrap. Annals of Statistics, 9, 130–134.