# Métodos Computacionais para Inferência Estatística

## Capítulo 1 e 2 - Verossimilhança

Paulo Justiniano Ribeiro Jr.
Wagner Hugo Bonat
Elias Teixeira Krainski
Walmes Marques Zeviani

**LEG**: Laboratório de Estatística e Geoinformação
Universidade Federal do Paraná

20$^o$ SINAPE, 30-31/07/2012

# Comentários

- Agradecimentos
- Recursos
  - Página: `http://www.leg.ufpr.br/mcie`
  - códigos e versões atualizadas !!!
- *Principal público alvo: graduação a início de PG*
- Motivações e Propósitos
- Experiências/exposição das gerações
- Facilidade de recursos computacionais e linguagens
- Uso de rotinas *versus* implementação/teste/ilustração/aprendizado
- Uso crítico e avaliação e apreciação das limitações de rotinas

Laboratório de Estatística e Geoinformação

# Conteúdos

- **texto** $x$ **apresentação**
  1. Verossimilhança - idéias e conceitos básicos e exemplos
  2. Modelos de regressão (GLM, Simplex, Subdisperso, Não Linear, Proc. Poisson não Hom.)
  3. Modelos de efeitos aleatórios (Espaciais, GLMM, Beta Longitudinal, TRI, Linear Dinâmico)
  4. Modelos dinâmicos
  5. (Inferência Bayesiana)
  6. Clonagem de dados
  7. INLA - *Integrated Nested Laplace Aproximation* - fundamentos
  8. INLA - exemplos (regressão dinâmica e modelos espaço-temporais)
- Omissões do material / Omissões no material
- **Foco:** inferência por verossimilhança e alguns métodos numéricos
- incursões eventuais em inferência Bayesiana

# Paradigmas para inferência

- Paradigmas
    - Frequentista (espaço amostral / Newman-Pearson)
    - Verossimilhança
    - Bayesiano

Laboratório de Estatística e Geoinformação

# Paradigmas para inferência

- Paradigmas
  - Frequentista (espaço amostral / Newman-Pearson)
  - Verossimilhança
  - Bayesiano
- Diferentes perguntas:

# Paradigmas para inferência

- Paradigmas
  - Frequentista (espaço amostral / Newman-Pearson)
  - Verossimilhança
  - Bayesiano
- Diferentes perguntas:
  - O que devo fazer?

# Paradigmas para inferência

- Paradigmas
  - Frequentista (espaço amostral / Newman-Pearson)
  - Verossimilhança
  - Bayesiano
- Diferentes perguntas:
  - O que devo fazer?
  - O que os dados dizem?

Laboratório de Estatística e Geoinformação

# Paradigmas para inferência

- Paradigmas
  - Frequentista (espaço amostral / Newman-Pearson)
  - Verossimilhança
  - Bayesiano
- Diferentes perguntas:
  - O que devo fazer?
  - O que os dados dizem?
  - Em que devo acreditar?

Laboratório de Estatística e Geoinformação

# Paradigmas para inferência

- Paradigmas
  - Frequentista (espaço amostral / Newman-Pearson)
  - Verossimilhança
  - Bayesiano
- Diferentes perguntas:
  - O que devo fazer?
  - O que os dados dizem?
  - Em que devo acreditar?
- Uma provocação à reflexão (**Royall, 1997:**)
  *"Fortunatelly we are not forced to choose either of these two evils, the sample-space dependence of the frequentists or the prior distribution of the Bayesians.*
  **Likelihood methods avoid both sources of subjectivity.** "

Laboratório de Estatística e Geoinformação

# Exemplo introdutório

Visitando um exemplo simples:

- População: $X \sim B(\theta)$

# Exemplo introdutório

Visitando um exemplo simples:

- População: $X \sim B(\theta)$
- Amostra: $x_1, \ldots, x_n$

# Exemplo introdutório

Visitando um exemplo simples:

- População: $X \sim B(\theta)$
- Amostra: $x_1, \ldots, x_n$
- O que podemos falar sobre $\theta$?
  - Qual a informação contida na amostra?
  - Consideram-se outras fontes de informação?

Laboratório de Estatística e Geoinformação

# Exemplo introdutório

Visitando um exemplo simples:

- População: $X \sim B(\theta)$
- Amostra: $x_1, \dots, x_n$
- O que podemos falar sobre $\theta$?
    - Qual a informação contida na amostra?
    - Consideram-se outras fontes de informação?
- informação na amostra resumida por $\left(n, y = \frac{\sum_{i=1}^{n} x_i}{n}\right)$ ?

Laboratório de Estatística e Geoinformação

# Espaço do Modelo

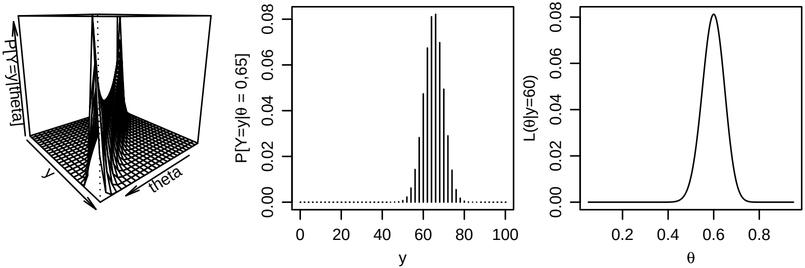
- O espaço definido pelo modelo (3-D)

$$\binom{n}{y}\theta^{y}(1-\theta)^{n-y}$$

# Espaço do Modelo

- O espaço definido pelo modelo (3-D)

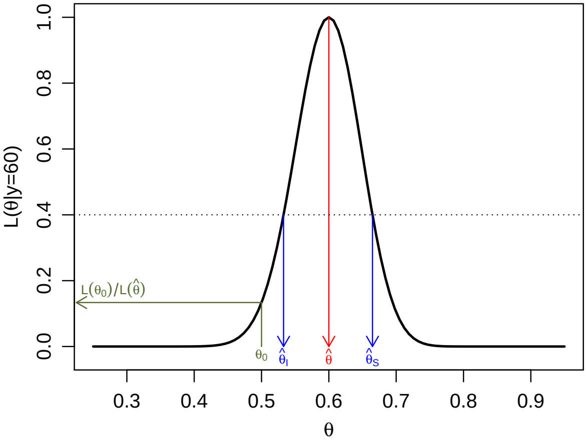
$$\binom{n}{y}\theta^{y}(1-\theta)^{n-y}$$




Laboratório de Estatística e Geoinformação

# Objetivos de Inferência

Objetivos de inferência e a função de verossimilhança



# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$

Laboratório de Estatística e Geoinformação

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador
- conjunto de valores razoavelmente compatíveis com a amostra

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador
- conjunto de valores razoavelmente compatíveis com a amostra
- decidir entre dois valores o mais compatível com a amostra

Laboratório de Estatística e Geoinformação

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador
- conjunto de valores razoavelmente compatíveis com a amostra
- decidir entre dois valores o mais compatível com a amostra
- decidir se a amostra é compatível com certo valor $\theta_0$ de interesse?

Laboratório de Estatística e Geoinformação

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador
- conjunto de valores razoavelmente compatíveis com a amostra
- decidir entre dois valores o mais compatível com a amostra
- decidir se a amostra é compatível com certo valor $\theta_0$ de interesse?
- suposições/pressupostos

Laboratório de Estatística e Geoinformação

# Elementos

- função de verossimilhança
  probabilidade da amostra obtida para diferentes valores de $\theta$
- melhor estimador
- conjunto de valores razoavelmente compatíveis com a amostra
- decidir entre dois valores o mais compatível com a amostra
- decidir se a amostra é compatível com certo valor $\theta_0$ de interesse?
- suposições/pressupostos
- relações e contrastes com outros métodos

Laboratório de Estatística e Geoinformação

# Inferência Bayesiana

- Extensão da definição do modelo

$$
\begin{aligned}
[Y|\theta] &\sim B(n, \theta) \\
[\theta] &\sim Pr(a, b) \ \ (priori)
\end{aligned}
$$

- permite obter (via teorema de Bayes)

$$[\theta|y] \sim \pi(a^*, b^*) \ \ (posteriori)$$

- Analogias diretas para estimação (pontual e intervalar),
- ... mas não para testes de hipótese
(valores na *posteriori* sem analogias diretas com razão de verossimilhanças).

Laboratório de Estatística e Geoinformação

# Função de Verossimilhança

**Definição informal:**

*Dada pela expressão da distribuição conjunta de todas as v.a.'s observáveis ($Y$) o modelo.* ou *Distribuição conjunta de todas as v.a.'s especificadas no modelo, integrada sobre as não observáveis.*

Notação: $[\cdot]$ (função de) distribuição de $\cdot$

Casos de particular interesse:

- Distribuições e modelos de regressão

$$[Y|\theta]$$

Laboratório de Estatística e Geoinformação

# Função de Verossimilhança

**Definição informal:**

*Dada pela expressão da distribuição conjunta de todas as v.a.'s observáveis ($Y$) o modelo.* ou *Distribuição conjunta de todas as v.a.'s especificadas no modelo, integrada sobre as não observáveis.*

Notação: $[\cdot]$ (função de) distribuição de $\cdot$

Casos de particular interesse:

- Distribuições e modelos de regressão

$$[Y|\theta]$$

- Modelos hierárquicos (mistos, efeitos aleatórios, longitudinais, espaciais, etc)

$$[Y|\theta] = \int [Y, b|\theta] d\, b = \int [Y|b, \theta][b|\theta] d\, b$$

Laboratório de Estatística e Geoinformação

# Função de Verossimilhança

**Definição informal:**

*Dada pela expressão da distribuição conjunta de todas as v.a.'s observáveis ($Y$) o modelo.* ou *Distribuição conjunta de todas as v.a.'s especificadas no modelo, integrada sobre as não observáveis.*

Notação: $[\cdot]$ (função de) distribuição de $\cdot$

Casos de particular interesse:

- Distribuições e modelos de regressão

$$[Y|\theta]$$

- Modelos hierárquicos (mistos, efeitos aleatórios, longitudinais, espaciais, etc)

$$[Y|\theta] = \int [Y, b|\theta] d\, b = \int [Y|b, \theta][b|\theta] d\, b$$

- inclui Inf. Bayesiana com especificação de $[\theta]$

# Expressão da Verossimilhança I

**V.A. observável discreta (não há ambiguidade)**

$$L(\theta) \equiv P_\theta[Y = y]$$

**Exemplo:** $Y \sim P(\theta)$

$$L(\theta) = \prod_{i=1}^{n} \frac{\exp\{-\theta\}\theta^{Y_i}}{Y_i!} = \frac{\exp\{-n\theta\}\theta^{\sum_{i=1}^{n} Y_i}}{\prod_{i=1}^{n} Y_i!}$$
$$\propto \exp\{-n\theta\}\theta^{\sum_{i=1}^{n} Y_i}$$

Laboratório de Estatística e Geoinformação

# Expressão da Verossimilhança II

**V.A. contínua: medição à certa precisão** $(y_{iI} \leq y_i \leq y_{iS})$

- Forma mais geral

$$L(\theta) = P_\theta[y_{1I} \leq y_1 \leq y_{1S}, \ldots, y_{nI} \leq y_n \leq y_{nS}]$$

Laboratório de Estatística e Geoinformação

# Expressão da Verossimilhança II

**V.A. contínua: medição à certa precisão** $(y_{iI} \leq y_i \leq y_{iS})$

- Forma mais geral

$$L(\theta) = P_\theta[y_{1I} \leq y_1 \leq y_{1S}, \ldots, y_{nI} \leq y_n \leq y_{nS}]$$

- Sob independência

$$\begin{aligned} L(\theta) &= P_\theta[y_{1I} \leq y_1 \leq y_{1S}] \cdot P_\theta[y_{2I} \leq y_2 \leq y_{2S}] \ldots P_\theta[y_{nI} \leq y_n \leq y_{nS}] \\ &= \prod_{i=1}^{n} P_\theta[y_{iI} \leq y_i \leq y_{iS}] \end{aligned}$$

Laboratório de Estatística e Geoinformação

# Expressão da Verossimilhança II

**V.A. contínua: medição à certa precisão** $(y_{iI} \leq y_i \leq y_{iS})$

- Forma mais geral

$$L(\theta) = P_\theta[y_{1I} \leq y_1 \leq y_{1S}, \ldots, y_{nI} \leq y_n \leq y_{nS}]$$

- Sob independência

$$\begin{aligned} L(\theta) &= P_\theta[y_{1I} \leq y_1 \leq y_{1S}] \cdot P_\theta[y_{2I} \leq y_2 \leq y_{2S}] \ldots P_\theta[y_{nI} \leq y_n \leq y_{nS}] \\ &= \prod_{i=1}^{n} P_\theta[y_{iI} \leq y_i \leq y_{iS}] \end{aligned}$$

- Se grau de precisão comum, $(y_i - \delta/2 \leq Y_i \leq y_i + \delta/2)$;

$$L(\theta) = \prod_{i=1}^{n} P_\theta[y_i - \delta/2 \leq Y_i \leq y_i + \delta/2] = \prod_{i=1}^{n} \int_{y_i-\delta/2}^{y_i+\delta/2} f(y_i, \theta) d(y_i).$$

Laboratório de Estatística e Geoinformação

# Expressão da Verossimilhança II (cont)

- alto grau de precisão ($\delta$ é pequeno em relação a variabilidade dos dados)

$$L(\theta) \approx \left( \prod_{i=1}^{n} f(y_i, \underline{\theta}) \right) \delta^n,$$

Laboratório de Estatística e Geoinformação

# Expressão da Verossimilhança II (cont)

- alto grau de precisão ($\delta$ é pequeno em relação a variabilidade dos dados)

$$L(\theta) \approx \left( \prod_{i=1}^{n} f(y_i, \underline{\theta}) \right) \delta^n,$$

- e se $\delta$ não depende dos valores dos parâmetros

$$L(\theta) \approx \prod_{i=1}^{n} f(y_i, \underline{\theta})$$

# Expressão da Verossimilhança II (cont)

- alto grau de precisão ($\delta$ é pequeno em relação a variabilidade dos dados)

$$L(\theta) \approx \left( \prod_{i=1}^{n} f(y_i, \underline{\theta}) \right) \delta^n,$$

- e se $\delta$ não depende dos valores dos parâmetros

$$L(\theta) \approx \prod_{i=1}^{n} f(y_i, \underline{\theta})$$

- observações não independentes - densidade multivariada:

$$L(\theta) \approx f(\underline{y}, \underline{\theta})$$

Laboratório de Estatística e Geoinformação

# Verossimilhança e Informação

Considere $Y \sim N(\theta, 1)$ e as as seguintes observações.

1. $x = 2.45$
2. $0.9 < x < 4$
3. somente o máximo de uma amostra de tamanho cinco é fornecido $x_{(5)} = 3.5$

Laboratório de Estatística e Geoinformação

# Verossimilhança e Informação

Considere $Y \sim N(\theta, 1)$ e as as seguintes observações.

1. $x = 2.45$
2. $0.9 < x < 4$
3. somente o máximo de uma amostra de tamanho cinco é fornecido $x_{(5)} = 3.5$

Verossimilhança:

$$L(\theta; x) = \phi(x - \theta) \equiv \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(x - \theta)^2\};$$

$$L_1 = L(\theta; x = 2.45) = \phi(x - \theta) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(2.45 - \theta)^2\};$$

$$L_2 = L(\theta; 0{,}9 < x < 4) = \Phi(4 - \theta) - \Phi(0{,}9 - \theta);$$

$$L_3 = L(\theta; x_{(5)} = 3.5) = n\{\Phi(x_{(n)} - \theta)\}^{n-1}\phi(x_{(n)} - \theta).$$

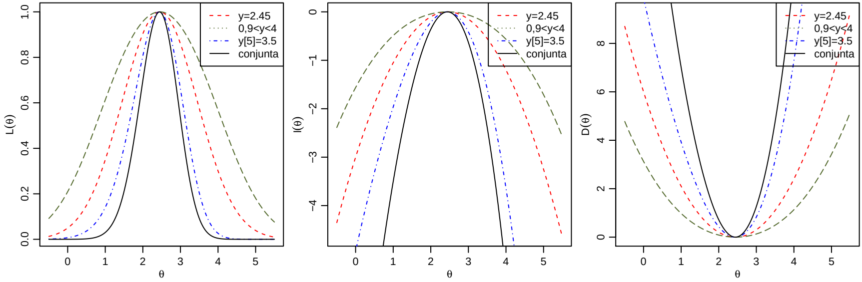
Para última - argumento multinomial e com

$$F(y) = P(X_{\{n\}} \leq y) = P[X_{\{i\}} < y \; \forall i \neq n \text{ e } X_{\{n\}} = y]$$

Laboratório de Estatística e Geoinformação

# Verossimilhança e Informação (cont)

```
L1 <- function(theta) dnorm(2.45, m=theta, sd=1)
L2 <- function(theta)
        pnorm(4,mean=theta,sd=1)-pnorm(0.9,mean=theta,sd=1)
L3 <- function(theta)
        5*pnorm(3.5,m=theta,s=1)^4 * dnorm(3.5,m=theta,s=1)
```



# Formas alternativas

- Verossimilhança:

$$L(\theta)$$

- Verossimilhança Relativa:

$$LR(\theta) = \frac{L(\theta)}{L(\hat{\theta})}$$

- log-Verossimilhança:

$$l(\theta) = \log\{L(\theta)\}$$

- *Deviance*:

$$D(\theta) = -2\log\{LR[\theta]\} = -2\{l[\theta] - l[\hat{\theta}]\}$$

EMV: $\hat{\theta} = \sup_{\Theta} L[\theta]$
Quando possível obtido por: $\hat{\theta} = \max_{\Theta} l[\theta]$

Laboratório de Estatística e Geoinformação

# Exemplo: distribuição Poisson

$$
\begin{aligned}
L[\theta] &= \frac{\exp\{-n\theta\}\theta^{n\overline{Y}}}{\prod_{i=1}^{n} Y_i!} \\
LR[\theta] &= \exp\{-n(\theta - \hat{\theta})\}(\theta/\hat{\theta})^{n\overline{Y}} \\
l[\theta] &= -n(\theta + \overline{Y}\log(\theta) - \overline{\log(Y_i!)}) \\
D(\theta) &= -2n\{(\theta - \hat{\theta}) - \overline{Y}\log(\theta/\hat{\theta})\}
\end{aligned}
$$

Para uma a.a. de observações pontuais:

$$\hat{\theta} = \overline{Y}$$

# Exemplo: Distribuição Poisson

1. Código 1

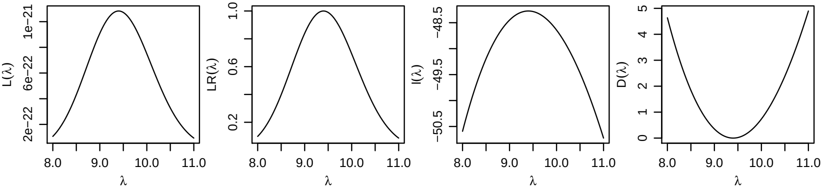
```
veroPois <- function(par, dados, tipo, maxlogL){
   tipo = match.arg(tipo, choices=c("L","LR","logL","dev"))
   ll <- sapply(par, function(p) sum(dpois(dados, lambda=p,
                                              log=TRUE)))
   return(switch(tipo, "L" = exp(ll),
                       "LR" = exp(ll-maxlogL),
                       "logL" = ll,
                       "dev" = 2*(maxlogL-ll)))}
```

2. Código 2

```
veroPois <- function(par, amostra, tipo="logL", maxlogL){
   tipo = match.arg(tipo, choices=c("L","LR","logL","dev"))
   ll <- with(amostra, -n*par + soma * log(par))
   return(switch(tipo, "L" = exp(ll),
                       "LR" = exp(ll-maxlogL),
                       "logL" = ll,
                       "dev" = 2*(maxlogL-ll)))}
```

# Exemplo: Distribuição Poisson (cont)

```
mll <- veroPois(emv, amostra=am, tipo="logL")
...
curve(veroPois(x, amostra=am, tipo="dev", maxlogL=mll), 8, 11,
      ylab=expression(D(lambda)), xlab=expression(lambda))
```



# Funções de Interesse

- Função escore: $U(\theta) = l'(\theta)$
- Hessiano e Informação observada: $I_O(\theta) = -H(\theta) = -l''(\theta)$
- Informação Esperada: $I_E(\theta) = E_Y[I_O(\theta)]$
- Estimadas: $I_O(\hat{\theta})$ e $I_E(\hat{\theta})$

Laboratório de Estatística e Geoinformação

# Funções de Interesse

- Função escore: $U(\theta) = l'(\theta)$
- Hessiano e Informação observada: $I_O(\theta) = -H(\theta) = -l''(\theta)$
- Informação Esperada: $I_E(\theta) = E_Y[I_O(\theta)]$
- Estimadas: $I_O(\hat{\theta})$ e $I_E(\hat{\theta})$

Propriedades assintóticas:

- $\underline{\hat{\theta}} \sim NM_d(\underline{\theta}, I_E(\underline{\theta})^{-1})$
- Assintoticamente equivalentes:

$$\underline{\hat{\theta}} \sim NM_d(\underline{\theta}, I_E(\underline{\hat{\theta}})^{-1})$$

$$\underline{\hat{\theta}} \sim NM_d(\underline{\theta}, I_O(\underline{\theta})^{-1})$$

$$\underline{\hat{\theta}} \sim NM_d(\underline{\theta}, I_O(\underline{\hat{\theta}})^{-1}).$$

- $D(\underline{\theta}) = -2[l(\underline{\theta}) - l(\underline{\hat{\theta}})] \sim \chi^2_d$

# Exemplo: Distribuição Poisson (cont)

$$L(\theta) = \prod_{i=1}^{n} \frac{\exp\{-\theta\}\theta^{Y_i}}{Y_i!} = \frac{\exp\{-n\theta\}\theta^{\sum_{i=1}^{n} Y_i}}{\prod_{i=1}^{n} Y_i!} = \frac{\exp\{-n\theta\}\theta^{n\bar{Y}}}{\prod_{i=1}^{n} Y_i!}$$

$$l(\theta) = -n\theta + (\sum_{i=1}^{n} Y_i)\log(\theta) - \sum_{i=1}^{n} \log Y_i! = -n(\theta - \bar{Y}\log(\theta) - \overline{\log Y_i!})$$

$$U(\theta) = -n + \frac{\sum_{i=1}^{n} Y_i}{\theta} = -n(1 - \frac{\bar{Y}}{\theta})$$

$$U(\hat{\theta}) = 0 \longrightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} Y_i}{n} = \overline{Y}$$

$$I_O(\theta) = \frac{\sum_{i=1}^{n} Y_i}{\theta^2} = \frac{n\bar{Y}}{\theta^2}\ ;\ I_E(\theta) = \frac{n}{\theta}\ ;\ I_O(\hat{\theta}) = I_E(\hat{\theta}) = \frac{n^2}{\sum_{i=1}^{n} Y_i} = \frac{n}{\overline{Y}}$$

$$V(\hat{\theta}) = I_E^{-1}(\theta) \approx I_O^{-1}(\theta) \approx I_O^{-1}(\hat{\theta}) = I_E^{-1}(\hat{\theta})$$

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Poisson (cont)

Função escore:

```
UPois <- function(lambda, amostra){
  return(with(amostra, n - soma/lambda))
}
```

Hessiano (negativo da Informação observada):

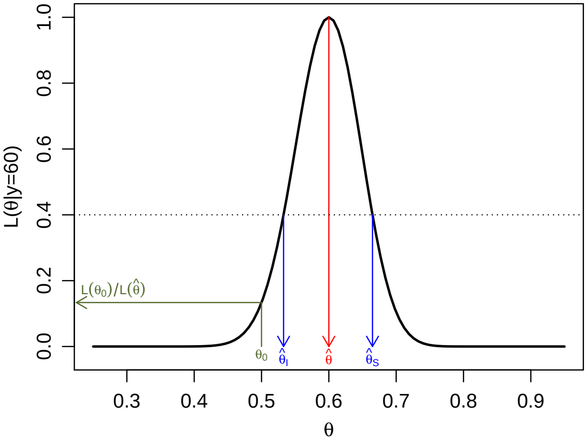
```
HPois <- function(lambda, amostra){
  return(with(amostra, -soma/lambda^2))
}
```

Informação esperada:

```
IePois <- function(lambda, amostra){
  return(with(amostra, n/lambda))
}
```

# Estimação

Objetivos de inferência e a função de verossimilhança



# Obtendo o EMV

- analiticamente:
  estudando comportamento de $l(\theta)$ ou resolvendo $U(\theta) = 0$
- numericamente (otimização/aproximações numéricas)
  - Solução da(s) equação(ões) de estimação (função escore)
    - com uso de derivadas (ex: Newton-Raphson)
    - sem uso de derivadas (ex: Brent)
  - Maximização da função de (log)-verossimilhança
- Outros (ex: EM)
- Simulação (ex: verossimilhança Monte Carlo, *data-cloning*, ...)
- Aproximações da verossimilhança (pseudo-verossimilhanças)

Laboratório de Estatística e Geoinformação

# Estimadores e Inferência

- Análogos para distribuições *posteriori* em Inferência Bayesiana
- maximização numérica: mais comum em EMV
- simulação: mais usual em Inferência Bayesiana

# EMV

Newton Raphson: expansão (Taylor) de 1ª ordem de $U(\theta)$:

$$\lambda^{r+1} = \lambda^{r} - \frac{U(\lambda)}{H(\lambda)}$$

```
maxit <- 100; lambdaNR <- 5; iter <- 0; d <- 1
while(d > 1e-12 & iter <= maxit){
    lambdaNR.new <-
      lambdaNR - UPois(lambdaNR, am)/HPois(lambdaNR, am)
    d <- abs(lambdaNR - lambdaNR.new)
    lambdaNR <- lambdaNR.new ; iter <- iter + 1
}
c(lambdaNR, iter)
```

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Poisson (cont)

- Solução de equação $U(\theta) = 0$:

```
uniroot(UPois, interval=range(y), amostra=am)$root
```

- Maximização da verossimilhança

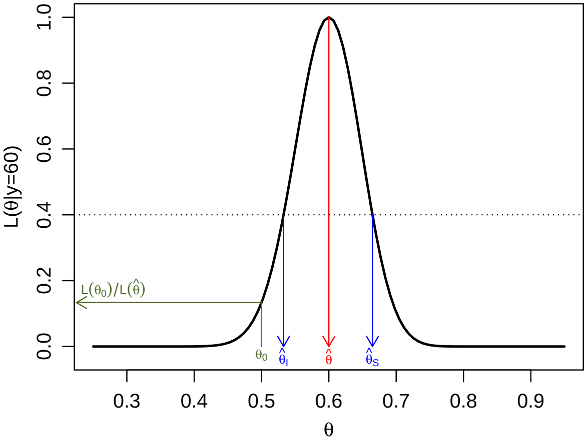
```
optimize(veroPois, interval=range(y), maximum=TRUE, amostra=am)
optim(par = median(y), fn=veroPois, control=list(fnscale=-1),
      amostra=am, hessian = TRUE)
```

  - uso do gradiente: argumento `gr = Upois`
  - pode retornar hessiano estimado ($H(\hat{\theta}) = -I_O(\hat{\theta})$ obtido numericamente)

Laboratório de Estatística e Geoinformação

# Intervalo de Confiança

Objetivos de inferência e a função de verossimilhança



# Estimação por Intervalo

Definição informal:

*Região de valores do parâmetro com compatibilidade aceitável com os dados .*

Definição do "ponto de corte" que define a região por:

a) evidência relativa em $LR(\theta)$:

b) comportamento assintótico $D(\theta) \sim \chi^2_p$:

c) interpretação probabilística direta em Inf. Bayesiana (quantis ou HPD)

Ou seja, evidência avaliada por:

- analogia sobre diferenças em $LR(\theta), l(\theta) ou D(\theta)$
- referência probabilística

OBS: $LR(\theta)$ e $D(\theta)$ são adimensionais.

Laboratório de Estatística e Geoinformação

# Relações entre critérios de corte

- $LR(\theta) \geq r$ na função de verossimilhança relativa
- $l(\hat{\theta}) - l(\theta) \geq c$ na função de log-verossimilhança
- $D(\theta) \leq c^*$ na função deviance

# Relações entre critérios de corte

- $LR(\theta) \geq r$ na função de verossimilhança relativa
- $l(\hat{\theta}) - l(\theta) \geq c$ na função de log-verossimilhança
- $D(\theta) \leq c^*$ na função deviance

$$
\begin{aligned}
LR(\theta) &\geq r \\
l(\hat{\theta}) - l(\theta) &\leq -\log(r) = c \\
D(\theta) = -2[l(\theta) - l(\hat{\theta})] &\leq -2\log(r) = c^* \\
c^* = 2c = -2\log(r) &\longrightarrow r = \mathrm{e}^{-c} = \mathrm{e}^{-c/2}
\end{aligned}
$$

## Relações entre critérios de corte

- $LR(\theta) \geq r$ na função de verossimilhança relativa
- $l(\hat{\theta}) - l(\theta) \geq c$ na função de log-verossimilhança
- $D(\theta) \leq c^*$ na função deviance

$$
\begin{aligned}
LR(\theta) &\geq r \\
l(\hat{\theta}) - l(\theta) &\leq -\log(r) = c \\
D(\theta) = -2[l(\theta) - l(\hat{\theta})] &\leq -2\log(r) = c^* \\
c^* = 2c = -2\log(r) &\longrightarrow r = \mathrm{e}^{-c} = \mathrm{e}^{-c/2}
\end{aligned}
$$

$$
\begin{aligned}
D(\theta) &\approx I_o(\hat{\theta})(\theta - \hat{\theta})^2 \sim \chi^2(1) \\
\sqrt{D[\theta]} &\approx I_o^{1/2}(\hat{\theta})(\theta - \hat{\theta}) \sim \mathrm{N}(0,1)
\end{aligned}
$$

$c^*$ é um quantil da $\chi^2_{(1)}$ ou, equivalentemente,
$\sqrt{c^*}$ é um quantil da $\mathrm{N}(0{,}1)$.

Laboratório de Estatística e Geoinformação

# Relações entre critérios de corte

$$c^* = 2c = -2\log(r) \longrightarrow r = \mathrm{e}^{-c} = \mathrm{e}^{-c/2}$$

Relações entre intervalos baseados no corte na $LR[\theta]$ e por limites de probabilidade.

| $r$ | $c$ | $c^*$ | $z = \sqrt{c^*}$ | $P[\|Z\| < \sqrt{c^*}]$ |
|---|---|---|---|---|
| 50% | 0,693 | 1,386 | 1,177 | 0,761 |
| 26% | 1.347 | 2.694 | 1,641 | 0,899 |
| 15% | 1,897 | 3,794 | 1,948 | 0,949 |
| 3,6% | 3,324 | 6,648 | 2,578 | 0,990 |

Laboratório de Estatística e Geoinformação

# analogia...

# Limites do Intervalo

1. Solução de equação (analítica ou numérica)
    - $LR(\theta) = r$
    - $l(\hat{\theta}) - l(\theta) = c$
    - $D(\theta) = c^*$

Laboratório de Estatística e Geoinformação

# Limites do Intervalo

1. Solução de equação (analítica ou numérica)
   - $LR(\theta) = r$
   - $l(\hat{\theta}) - l(\theta) = c$
   - $D(\theta) = c^*$
2. Aproximação quadrática (Taylor)

$$D(\theta) = -2[l(\theta) - l(\hat{\theta})] =$$
$$= -2\left\{[l(\hat{\theta}) + (\theta - \hat{\theta})l'(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 l''(\hat{\theta})] - l(\hat{\theta})\right\}$$
$$D(\theta) = -(\theta - \hat{\theta})^2 l''(\hat{\theta}) \leq c^*$$
$$\tilde{\theta} = \hat{\theta} \pm \sqrt{\frac{c^*}{-l''(\hat{\theta})}}$$

Laboratório de Estatística e Geoinformação

# Limites do Intervalo

1. Solução de equação (analítica ou numérica)
   - $LR(\theta) = r$
   - $l(\hat{\theta}) - l(\theta) = c$
   - $D(\theta) = c^*$
2. Aproximação quadrática (Taylor)

$$D(\theta) = -2[l(\theta) - l(\hat{\theta})] =$$

$$= -2\left\{[l(\hat{\theta}) + (\theta - \hat{\theta})l'(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 l''(\hat{\theta})] - l(\hat{\theta})\right\}$$

$$D(\theta) = -(\theta - \hat{\theta})^2 l''(\hat{\theta}) \leq c^*$$

$$\tilde{\theta} = \hat{\theta} \pm \sqrt{\frac{c^*}{-l''(\hat{\theta})}}$$

3. equivalência com à Distribuição assintótica: $\underline{\hat{\theta}} \sim NM_d(\underline{\theta}, I_E(\underline{\theta})^{-1})$

# Exemplo: Exponencial (i.i.d.)

$$f(y_i, \theta) = \theta \exp\{-\theta y_i\} \quad y > 0 \; ; \; \theta > 0$$

$$F(y_i, \theta) = 1 - \exp\{-\theta y_i\} \quad y > 0 \; ; \; \theta > 0$$

$$L(\theta) = \theta^n \exp\{-\theta n \bar{y}\}$$

$$l(\theta) = n \log(\theta) - \theta n \bar{y}$$

$$U(\theta) = \frac{n}{\theta} - n\bar{y}$$

$$H(\theta) = -\frac{n}{\theta^2} \text{ (depende do valor de } \theta\text{!!)}$$

$$\hat{\theta} = 1/\bar{y}$$

**Código R**

# Exemplo: Exponencial (cont)

## Intervalos de confiança

1. Corte na deviance: (solução apenas numericamente)

$$D(\theta) = 2n[\log\left(\hat{\theta}/\theta\right) + \overline{y}(\theta - \hat{\theta})] \leq c^*$$

# Exemplo: Exponencial (cont)

## Intervalos de confiança

1. Corte na deviance: (solução apenas numericamente)

$$D(\theta) = 2n[\log\left(\hat{\theta}/\theta\right) + \overline{y}(\theta - \hat{\theta})] \leq c^*$$

2. Aproximação quadrática:

$$D(\theta) \approx n\left(\frac{\theta - \hat{\theta}}{\hat{\theta}}\right)^2$$

$$\left(\hat{\theta}_L \approx \hat{\theta}(1 - \sqrt{c^*/n}) \ , \ \hat{\theta}_U \approx \hat{\theta}(1 + \sqrt{c^*/n})\right)$$

# Exemplo: Exponencial (cont)

## Intervalos de confiança

1. Corte na deviance: (solução apenas numericamente)

$$D(\theta) = 2n[\log\left(\hat{\theta}/\theta\right) + \overline{y}(\theta - \hat{\theta})] \leq c^*$$

2. Aproximação quadrática:

$$D(\theta) \approx n\left(\frac{\theta - \hat{\theta}}{\hat{\theta}}\right)^2$$

$$\left(\hat{\theta}_L \approx \hat{\theta}(1 - \sqrt{c^*/n}) \ , \ \hat{\theta}_U \approx \hat{\theta}(1 + \sqrt{c^*/n})\right)$$

3. Distribuição assintótica: $I_E^{-1}(\theta) \approx I_O^{-1}(\hat{\theta}) = \hat{\theta}^2/n$

$$\hat{\theta} \pm z_{\frac{\alpha}{2}}\sqrt{V(\hat{\theta})}$$

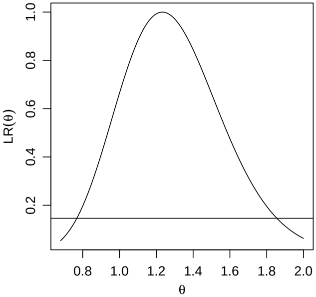
$$\hat{\theta}_L = \hat{\theta} - z_{\frac{\alpha}{2}}\hat{\theta}/\sqrt{n} \quad \text{e} \quad \hat{\theta}_U = \hat{\theta} + z_{\frac{\alpha}{2}}\hat{\theta}/\sqrt{n}$$

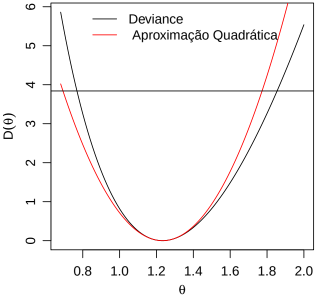
Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Exponencial (cont)

```
ICdevExp <- function(theta, theta.hat, y, nivel=0.95){
  n <- length(y)
  dv <- 2*n*( log( theta.hat/theta) + mean(y)*(theta- theta.hat))
  return(dv - qchisq(nivel,df=1))
}
```

```
require(rootSolve)
uniroot.all(ICdevExp, interval=c(0,10), theta.hat=1/mean(y),y=y)
```



# Reparametrização

Caso univariado (ou reparametrização 1-1)

$$\phi = g(\theta)$$

Como fazer inferências sobre $\phi$?

- estimação pontual
- estimação por intervalo
- testes de hipótese

Um resultado fundamental: invariância da verossimilhança:

$$l(\phi_i) = l(g(\theta_i)) = l(\theta_i)$$

# Reparametrização

Duas alternativas iniciais:

1. Reescrever a função de verossimilhança e "recomeçar" do zero:
   - Pontual: $\hat{\phi} = \text{argmax}\{l(\phi)\}$
   - Intervalar:

$$(\hat{\phi}_I,\hat{\phi}_S) \text{ baseado em } l(\phi) \tag{1}$$

$$(\tilde{\phi}_I,\tilde{\phi}_S) \text{ baseado em } \tilde{l}(\phi) \tag{2}$$

2. "Aproveitar" resultados da inferência já obtida para $\theta$:
   - Pontual: $\hat{\phi} = g(\hat{\theta})$ (invariância)
   - Intervalar:

$$(g(\hat{\theta}_I), g(\hat{\theta}_S)) \text{ transformação dos limites em } l(\theta)(\text{invariância}) \tag{3}$$

$$(g(\tilde{\theta}_I),g(\tilde{\theta}_S)) \text{ transformação dos limites em } \tilde{l}(\theta) \tag{4}$$

Estimativas pontuais: iguais em ambos casos

Estimativas intervalares: (1) = (3) e tem-se então três possibilidades

Laboratório de Estatística e Geoinformação

# Reparametrização

- Resultados exatos baseados na verossimilhança:
  - $\hat{\phi} = g(\hat{\theta})$
  - IC por corte: $(\hat{\phi}_I, \hat{\phi}_S) = (g(\hat{\theta}_I), g(\hat{\theta}_S))$
- Aproximações quadráticas da verossimilhança
  - aproximação $\tilde{l}(\theta)$ : $(g(\tilde{\theta}_I), g(\tilde{\theta}_S))$
  - aproximação $\tilde{l}(\phi)$ : $(\tilde{\phi}_I, \tilde{\phi}_S)$
- Distribuição assintótica do estimador pelo Método *delta*:

$$\mathrm{Var}(\hat{\phi}) = [g'(\theta)]^2 [I_E(\theta)]^{-1} \longrightarrow \left[ \mathrm{se}(\hat{\phi}) = |g'(\theta)| [I_E(\theta)]^{-1/2} \right]$$

Assintoticamente: $\hat{\phi} = g(\hat{\theta}) \sim N(\phi, [g'(\theta)]^2 [I_E(\theta)]^{-1})$

Laboratório de Estatística e Geoinformação

# Método delta

$$\phi = g(\theta)$$

**Teorema**

*(Método delta). Seja $\hat{\theta}$ um estimador de $\theta$ para uma amostra de tamanho $n$ tal que*

$$\sqrt{n}(\hat{\theta} - \theta) \to \mathrm{N}(0, \sigma_\theta^2).$$

*Então, para qualquer função $g(\cdot)$ que é diferenciável ao redor de $\theta$ e $h'(\theta) \neq 0$, tem-se que*

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \to \mathrm{N}(0, \sigma_\theta^2 |h'(\theta)|^2).$$

Em temos do TCL, aplica-se a funções da média amostral.

Produz uma aproximação quadrática da vesossimilhança de $\phi$.

Laboratório de Estatística e Geoinformação

# Reparametrização

- ideal: Resultados exatos baseados na verossimilhança:
- muito utilizado: Resultados baseados no método delta
- conveniente: aproximações baseadas nas aproximações quadráticas das verossimilhanças $\tilde{l}(\theta)$ ou $\tilde{l}(\phi)$
- Se transformação $g(\cdot)$ é não linear, a invariância não é válida para a aproximação quadrática, ou seja:

$$\{g(\tilde{\theta}_I),g(\tilde{\theta}_S)\} = \{g(\hat{\theta} - z_{\alpha/2}[I_E(\hat{\theta})]^{-1/2}),g(\hat{\theta} + z_{\alpha/2}[I_E(\hat{\theta})]^{-1/2})\} \neq$$
$$\{g(\hat{\theta}) - z_{\alpha/2}|g'(\theta)|[I_E(\hat{\theta})]^{-1/2},g(\hat{\theta}) + z_{\alpha/2}|g'(\theta)|[I_E(\hat{\theta})]^{-1/2}\} = (\tilde{\phi}_I,\tilde{\phi}_S)$$

- se $l(\theta)$ é menos assimétrica: usar $(g(\tilde{\theta}_I),g(\tilde{\theta}_S))$
  se $l(\phi)$ é menos assimétrica: usar $(\tilde{\phi}_I,\tilde{\phi}_S)$

# Recomendações

- Melhor abordagem: (mais geral e acurácia)
  IC's baseados verossimilhança/deviance (muitas vezes só obtidos numericamente)
- Intervalos assintóticos (utilizam $se(\hat{\theta})$, obtenção a partir da aproximação quadrática, formas fechadas )
- Escolher parametrização da função que forneça uma boa aproximação quadrática
- IC's para funções dos parâmetros: obtenção pelo método delta ou direta se aproximadamente quadrática

Laboratório de Estatística e Geoinformação

# Reparametrização

Por que reparametrizar (ou considerar resultados ra reparametrização)?

- Interpretabilidade de $\phi$
- "Melhor"formato da função de verossimilhança com melhor comportamento de métodos numéricos.
- Predição!!
  uma predição de um modelo pode ser vista como uma reparametrização!

Laboratório de Estatística e Geoinformação

# Condições de regularidade

- $\Theta$ é finito dimensional e $\theta$ é interior a $\Theta$
- primeiras três derivadas de $l(\theta)$ na vizinhança de $\theta$
- amplitude não depende de $\theta$
- $l(\theta) \approx$ quadrática para $n \to \infty$, passando a depender apenas da posição e curvatura no EMV
- $\ldots I_E(\theta)$ precisa ser inversível

Laboratório de Estatística e Geoinformação

# Exemplo: Exponencial (cont)

## Reparametrização

$$\phi = P[Y \leq u] = 1 - \exp\{-\theta u\}$$

- Obter $se(\hat{\phi})$
- Três intervalos possíveis:

$$(\hat{\phi}_I, \hat{\phi}_S) : (g(\hat{\theta}_I), g(\hat{\theta}_S))$$

$$(\tilde{\phi}_I, \tilde{\phi}_S) : \hat{\phi} \pm z_{\alpha/2} se(\hat{\phi})$$

$$(1 - \exp\{-\tilde{\theta}_S u\}, 1 - \exp\{-\tilde{\theta}_S u\}) : (g(\tilde{\theta}_I), g(\tilde{\theta}_S))$$

- Comparação gráfica das funções e das taxas de cobertura (simulação)

# Exemplo: Distribuição Exponencial (cont)

## Redefinindo

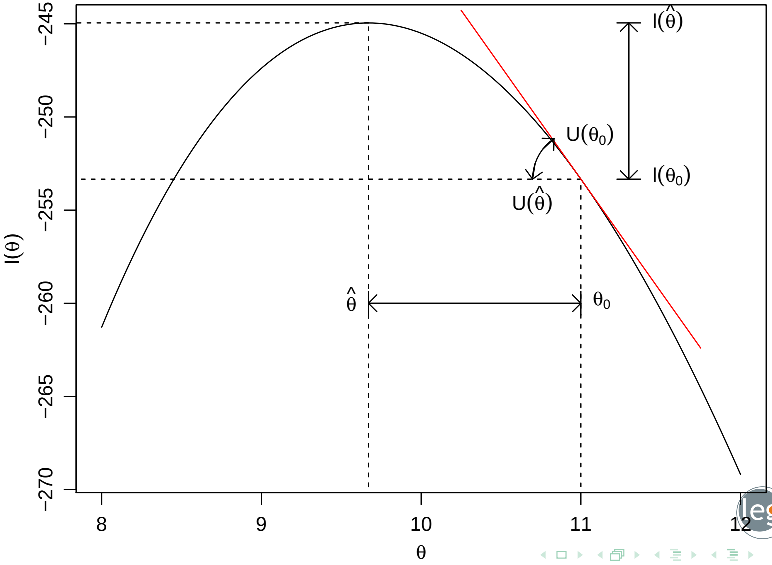
```
ICdevExp <- function(theta, theta.hat, y, nivel=0.95){
  n <- length(y)
  dv <- 2*n*( log( theta.hat/theta) + mean(y)*(theta- theta.hat))
  return(dv - qchisq(nivel,df=1))
}
```

```
require(rootSolve)
uniroot.all(ICdevExp, interval=c(0,10), theta.hat=1/mean(y), y=y)
```

```
ICdevExp <- function(theta, amostra, nivel=0.95){
  ## amostra é um vetor com elementos n e mean(y), nesta ordem
  n <- amostra[1]
  med <- amostra[2]
  dv <- 2*n*(-log(med*theta) + med*theta - 1)
  return(dv - qchisq(nivel, df=1))
}
```

```
am <- c(length(y), mean(y))
uniroot.all(ICdevExp, interval=c(0,10), amostra=am)
```

# Teste de Hipótese



# Teste de Hipótese

- Teste razão de verossimilhança

```
trv <- function(Est, H0, alpha, ...){
  critico <- qchisq(1-alpha, df=1)
  est.calc <- Est(H0, ...)
  print(ifelse(est.calc < critico, "Aceita H0", "Rejeita H0"))
  return(c(est.calc,critico))}
```

- Teste Wald

```
wald <- function(H0, EMV, V.EMV, alpha){
  critico <- qnorm(1-alpha/2)
  Tw <- (EMV - H0)/sqrt(V.EMV)
  print(ifelse(Tw < critico, "Aceita H0", "Rejeita H0"))
  return(c(Tw,critico))
}
```

- Teste Escore

```
escore <- function(H0, U, Ie, alpha, ...){
  critico <- qnorm(1-alpha/2)
  Te <- U(H0,...)/sqrt(Ie(H0,...))
  print(ifelse(Te < critico, "Aceita H0", "Rejeita H0"))
  return(c(Te,critico))
```

# Exemplo: Poisson

TRV

```
Est <- function(H0, x){
   n <- length(x)
   EMV <- mean(x)
   lv <- 2*n*(( H0 - EMV) + EMV*log(EMV/H0))
 return(lv)
}
trv(Est = Est, H0=8, alpha = 0.05, x=x)
```

Wald

```
wald(H0=8, EMV = mean(x), V.EMV = mean(x)/length(x),alpha=0.05)
```

Escore

```
fc.escore <- function(lambda,x){
   n <- length(x)
   esco <- -n + sum(x)/lambda
   return(esco)}
Ie <- function(lambda,x){
   n <- length(x)
   I <- n/lambda
   return(I)}
escore(H0 = 8, U = fc.escore, Ie = Ie, alpha=0.05, x=x)
```

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Normal

log-Verossimilhança

$$l(\mu,\sigma) = -\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2.$$

Escore

$$U(\mu) = \frac{\partial l(\mu,\sigma)}{\partial\mu} = \frac{\sum_{i=1}^{n} y_i}{\sigma^2} - \frac{n\mu}{\sigma^2}$$

$$U(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i - \mu)^2.$$

EMV

$$\hat{\mu} = \frac{\sum_{i=1}^{n} y_i}{n} \quad \text{e} \quad \hat{\sigma^2} = \frac{\sum_{i=1}^{n}(y_i - \mu)^2}{n}.$$

Informação

$$I_O(\hat{\mu},\hat{\sigma}) = \left[\begin{array}{cc} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & \frac{2n}{\hat{\sigma}^2} \end{array}\right].$$

Laboratório de Estatística e Geoinformação

# Intervalos de confiança

## Conjuntos

- corte

$$D(\mu, \sigma) = 2[n \log\left(\frac{\sigma}{\hat{\sigma}}\right) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 - \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(y_i - \hat{\mu})]$$

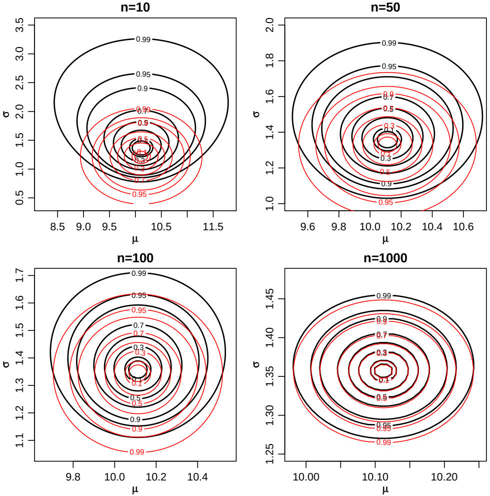
- elipse (aproximação quadrática)

$$D(\mu, \sigma) \approx (\underline{\theta} - \underline{\hat{\theta}})^\top I_o(\underline{\hat{\theta}})(\underline{\theta} - \underline{\hat{\theta}}).$$

- assintótico

$$\begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix} \sim NM_2\left(\begin{bmatrix} \mu \\ \sigma \end{bmatrix}, \begin{bmatrix} \hat{\sigma}^2/n & 0 \\ 0 & \hat{\sigma}^2/2n \end{bmatrix}\right)$$

# Exemplo: Distribuição Normal (cont)



# Intervalos de confiança

Parâmetros de interesse e de inconveniência (*nuisance*): $(\theta, \psi)$

Soluções usuais:

- Condicionando no EMV : $L(\theta) = L(\theta, \hat{\psi}) \equiv [Y|\theta, \hat{\psi}]$
- Verossimilhança Perfilhada : $L(\theta) \equiv L[\theta, \hat{\psi}_\theta]$
- Verossimilhanças marginais integradas Bayesianas : $L(\theta) = \int [Y|\theta, \psi][\psi] d\,\psi$

# Intervalos de confiança

Parâmetros de interesse e de inconveniência (*nuisance*): $(\theta, \psi)$

Soluções usuais:

- Condicionando no EMV : $L(\theta) = L(\theta, \hat{\psi}) \equiv [Y|\theta, \hat{\psi}]$
- Verossimilhança Perfilhada : $L(\theta) \equiv L[\theta, \hat{\psi}_\theta]$
- Verossimilhanças marginais integradas Bayesianas : $L(\theta) = \int [Y|\theta, \psi][\psi] d\,\psi$

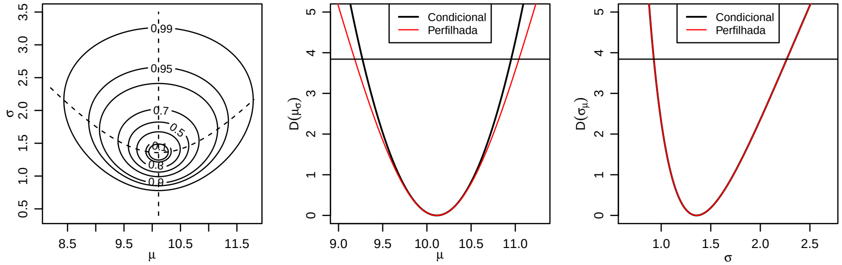
Exemplo Normal: $1/\sigma^2 \sim G(a, b)$

$$f(y|\mu) = \frac{\Gamma(n/2+1)}{\pi^{n/2}\Gamma(a)(\sum_i (y_i - \mu)^2 + 2b)^{n/2+a}}$$

Integrações analíticas e por simulação

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Normal (cont)



# Exemplo: Distribuição Normal (cont)

```
pl.mu <- function(sigma, mu, dados){
    pll <- sum(dnorm(dados, mean=mu, sd=sigma, log=TRUE))
    return(pll)}
##
pl.sigma <- function(mu, sigma, dados){
    pll <- sum(dnorm(dados, mean=mu, sd=sigma, log=TRUE))
    return(pll)}
```

```
grid.mu <- seq(9, 11.3, length=200)
grid.sigma <- seq(0.65, 2.7, length=200)
## Condicionais:
mu.cond <- sapply(grid.mu, pl.sigma, sigma=sqrt(var(y10)*9/10), dados=y10)
sigma.cond <- sapply(grid.sigma, pl.mu, mu=mean(y10), dados=y10)

mu.perf <- matrix(0, nrow=length(mu), ncol=2)
for(i in 1:length(mu)){
mu.perf[i,] <- unlist(optimize(pl.mu,c(0,200),
                     mu=mu[i],dados=y10,maximum=TRUE))}
sigma.perf <- matrix(0, nrow=length(sigma), ncol=2)
for(i in 1:length(sigma)){
sigma.perf[i,] <- unlist(optimize(pl.sigma,c(0,1000),
                sigma=sigma[i],dados=y10,maximum=TRUE))}
```

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Normal (Dados intervalares)

Dados intervalares e parâmetros $\theta = (\mu, \sigma)$

observações "pontuais":

`72,6  81,3  72,4  86,4  79,2  76,7  81,3` ;

observações intervalares:

uma observação com valor acima de 85,
uma observação com valor acima de 80,
quatro observações com valores entre 75 e 80,
seis observações com valores abaixo de 75.

Contribuições para verossimilhança

$$L(\theta) = f(y_i) \text{ para } y_i \text{ pontual},$$
$$L(\theta) = 1 - F(85) \text{ para } y_i > 85,$$
$$L(\theta) = 1 - F(80) \text{ para } y_i > 80,$$
$$L(\theta) = F(80) - F(75) \text{ para } 75 < y_i < 80,$$
$$L(\theta) = F(75) \text{ para } y_i < 75.$$

Laboratório de Estatística e Geoinformação

# Exemplo: Distribuição Normal (Dados intervalares)

Expressão da verossimilhança no exemplo:

$$L(\theta) = \left( \prod_{1=1}^{7} f(y_i) \right) \cdot (1 - F(85)) \cdot (1 - F(80)) \cdot (F(80) - F(75))^4 \cdot (F(75))^6$$

De forma mais geral para $n_p$ dados pontuais e $n_I$ dados intervalares com valores entre $a_i$ e $b_i$:

$$\begin{aligned} L(\theta) &= \prod_{1=1}^{n_p} f(y_i) \cdot \prod_{1=1}^{n_I} (F(b_i) - F(a_i)) \\ &= \prod_{1=1}^{n_p} \phi(\frac{y_i - \mu}{\sigma}) \cdot \prod_{1=1}^{n_I} (\Phi(\frac{a_i - \mu}{\sigma}) - \Phi(\frac{b_i - \mu}{\sigma})) \end{aligned}$$

No exemplo:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | 85 | 80 | 75 | 75 | 75 | 75 | $-\infty$ | $-\infty$ | $-\infty$ | $-\infty$ | $-\infty$ | $-\infty$ |
| b | $\infty$ | $\infty$ | 80 | 80 | 80 | 80 | 75 | 75 | 75 | 75 | 75 | 75 |

# Dados intervalares (cont)

$$l(\theta) = \sum_{1=1}^{n_p} \log\left(\phi(\frac{y_i - \mu}{\sigma})\right) + \sum_{1=1}^{n_I} \log\left((\Phi(\frac{a_i - \mu}{\sigma}) - \Phi(\frac{b_i - \mu}{\sigma}))\right)$$

```
nllnormI <- function(par, xp, XI) {
    ll1 <- sum(dnorm(xp, mean = par[1], sd = par[2], log = T))
    L2 <- pnorm(XI, mean = par[1], sd = par[2])
    ll2 <- sum(log(L2[, 2] - L2[, 1]))
    return(-(ll1 + ll2))
}
```

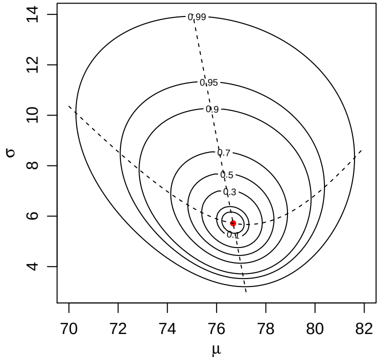
```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
[1,]   85   80   75   75   75   75 -Inf -Inf -Inf  -Inf  -Inf  -Inf
[2,]  Inf  Inf   80   80   80   80   75   75   75    75    75    75
```

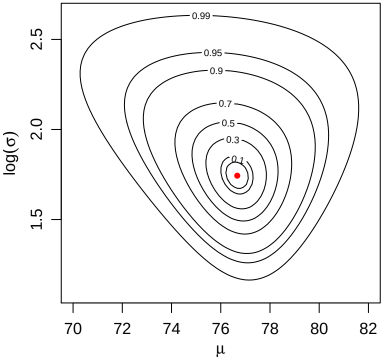
```
ini <- c(mean(y), sd(y))
ests <- optim(, nllnormI, x=y, XI=yI)$par
```

# Dados intervalares (cont)

Função deviance genérica.

```
devFun <- function(theta, est, llFUN, ...){
  return(2 * (llFUN(theta, ...) - llFUN(est, ...)))
}
devSurf <- Vectorize(function(x,y, ...) devFun(c(x,y), ...))
```



# Dados intervalares (cont)

## Código mais geral e cuidadoso

```
nllnormI <- function(par, xp, XI, logsigma=FALSE){
    if(logsigma) par[2] <- exp(par[2])
        ll1 <- ifelse(missing(xp), 0,
                 sum(dnorm(xp, mean=par[1], sd=par[2], log=T)))
    if(missing(XI)) ll2 <- 0
    else{
      if(ncol(XI) != 2 || any(XI[,2] <= XI[,1]))
       stop("XI deve ser matrix com 2 colunas com XI[,2] > XI[,2]"
      L2 <- pnorm(XI, mean=par[1], sd=par[2])
            ll2 <- sum(log(L2[,2] - L2[,1]))
  }
  return(-(ll1 + ll2))
}
```

# Outros exemplos (texto)

1. AR1
2. Outro exemplo de reparametrização
3. Gamma
4. Binomial Negativa
5. Processo de Poisson não-homogêneo
6. Modelo espacial Geoestatístico
7. Códigos Genéricos:
   - `mle` (**stat4**) e `mle2` (**bbmle**)
   - `profile` e `confint`

Laboratório de Estatística e Geoinformação