UNIVERSIDADE FEDERAL DE SERGIPE
DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS
Disciplina: Inferência I
Professor: Sadraque E.F. Lucena

Lista de Exercícios 10

10.1) Sejam $X_1, \ldots, X_n$ uma amostra aleatória. Obtenha o EMV de θ para as seguintes distribuições:

(a) $X \sim$ Poisson(θ);

(b) $X \sim$ Geométrica(θ), onde $P(X = x) = (1-\theta)^{x-1}\theta$, para $x = 1, 2, \ldots$;

(c) $X \sim$ Exponencial(θ), com fdp $f(x) = \frac{1}{\theta} e^{-x/\theta}$, $x > 0$ (parametrização pela média).

10.2) Considere uma amostra aleatória $X_1, \ldots, X_n$ de uma população com densidade dada por:

$$f(x|\theta) = \theta x^{\theta-1}, \quad 0 < x < 1, \quad \theta > 0$$

(a) Encontre o EMV de θ.

(b) Mostre que a estatística $T = -\sum_{i=1}^{n} \ln(X_i)$ é suficiente para θ.

10.3) Suponha que a velocidade de moléculas de um gás siga a distribuição de Maxwell-Boltzmann, com densidade:

$$f(x) = \frac{x^2 e^{-x/(2\theta)}}{2\theta^3}, \quad x > 0, \theta > 0.$$

Obtenha o estimador de máxima verossimilhança de θ.

10.4) Seja $X_1, \ldots, X_n$ uma amostra aleatória da distribuição de Pareto, com densidade:

$$f(x) = \theta k^{\theta} x^{-(\theta+1)}, \quad x \geq k, \quad \theta > 0,$$

onde $k > 0$ é uma constante conhecida. Encontre o EMV de θ.

10.5) Seja $X \sim$ Bernoulli(p). Sabemos que o EMV de p é $\widehat{p} = \overline{X}$.

(a) Um pesquisador deseja estimar a "Razão de Chances" (*Odds Ratio*), definida por $\psi = \frac{p}{1-p}$. Encontre o EMV de ψ.

(b) Encontre o EMV da variância de X, ou seja, de $g(p) = p(1-p)$.

10.6) Seja $X_1, \ldots, X_n$ uma amostra de $N(\mu, \sigma^2)$.

(a) Escreva a função de log-verossimilhança para o vetor $\boldsymbol{\theta} = (\mu, \sigma^2)$.

(b) Derive e iguale a zero para encontrar o sistema de equações de score.

(c) Mostre que $\widehat{\sigma}^2_{MV} = \frac{1}{n} \sum (X_i - \overline{X})^2$.

(d) Compare $\widehat{\sigma}^2_{MV}$ com a variância amostral S^2. Qual deles é viesado?

10.7) Em economia, a distribuição Lognormal é frequentemente utilizada para modelar a renda de populações. Dizemos que $X \sim \text{Lognormal}(\mu, \sigma^2)$ se $Y = \ln(X)$ segue uma distribuição $\text{Normal}(\mu, \sigma^2)$. A densidade de X é dada por:

$$f(x; \mu, \sigma^2) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}, \quad x > 0.$$

Seja $X_1, \ldots, X_n$ uma amostra aleatória dessa população.

(a) Escreva a função de log-verossimilhança $\ell(\mu, \sigma^2)$.

(b) Derive o sistema de equações de score para o vetor de parâmetros $\boldsymbol{\theta} = (\mu, \sigma^2)$.

(c) Resolva o sistema e mostre que os estimadores de máxima verossimilhança são a média e a variância (viesada) dos logaritmos dos dados:

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

10.8) Considere $X \sim \text{Exponencial}(\lambda)$ com média $1/\lambda$. O EMV é $\widehat{\lambda} = 1/\overline{X}$.

(a) Calcule a Informação de Fisher $I_1(\lambda)$.

(b) Determine a distribuição assintótica de $\widehat{\lambda}$.

(c) Um engenheiro quer estimar a confiabilidade no tempo $t = 1$, dada por $R(1) = e^{-\lambda}$. Use o Método Delta para encontrar a distribuição assintótica do estimador de $R(1)$.

10.9) O número de falhas em um sistema de TI segue uma distribuição de Poisson. Em 4 semanas, foram observadas as seguintes quantidades de falhas: 2, 5, 3, 6.

(a) Estime o parâmetro λ (taxa semanal) por MV.

(b) Estime a probabilidade de que na próxima semana não ocorra nenhuma falha ($P(X = 0)$).

10.10) Os tempos entre chegadas de clientes em uma loja (em minutos) foram:

2,5 1,8 4,0 0,5 3,2

Assumindo distribuição Exponencial(λ):

(a) Obtenha a estimativa de MV para a taxa λ e para a média $1/\lambda$.

(b) Estime a probabilidade de um intervalo ser maior que 3 minutos.

10.11) Julgue as afirmações abaixo como Verdadeiras ou Falsas, justificando brevemente:

(a) O Estimador de Máxima Verossimilhança (EMV) é sempre não-viesado para qualquer tamanho de amostra.

(b) Se $\widehat{\theta}$ é o EMV de θ, então $\sqrt{\widehat{\theta}}$ é necessariamente o EMV de $\sqrt{\theta}$.

(c) Sob condições de regularidade, a variância assintótica do EMV atinge o Limite Inferior de Cramér-Rao.

(d) Se $T(\mathbf{X})$ é uma estatística suficiente para θ, então o Estimador de Máxima Verossimilhança deve depender dos dados apenas através de $T(\mathbf{X})$.

(e) A resolução da equação de score ($\frac{\partial \ell}{\partial \theta} = 0$) sempre resulta em uma fórmula explícita para o estimador, dispensando o uso de métodos numéricos iterativos.

(f) Encontrar uma raiz da equação de score é condição suficiente para garantir que encontramos o máximo global da função de verossimilhança.

10.12) A função de verossimilhança $L(\theta; \mathbf{x})$ é frequentemente confundida com a função de densidade conjunta $f(\mathbf{x}; \theta)$. Explique a diferença fundamental entre as duas e por que a integral de $L(\theta; \mathbf{x})$ em relação a θ não é necessariamente igual a 1.

10.13) O gráfico da função de log-verossimilhança $\ell(\theta)$ pode ser "mais pontiagudo"(curvatura acentuada) ou "mais achatado". Qual a relação entre a curvatura dessa função no ponto de máximo e a precisão (variância) do estimador? Relacione sua resposta com o conceito de Informação de Fisher.

10.14) Um atuário está modelando a frequência de sinistros de uma carteira de automóveis utilizando uma distribuição da Família Exponencial. Ele optou pelo Método da Máxima Verossimilhança (MV) para estimar o parâmetro θ de interesse. Com base na teoria da estimação pontual, julgue os itens a seguir como Verdadeiro ou Falso:

(a) Se o atuário precisar estimar a probabilidade de não ocorrência de sinistros ($e^{-\theta}$), basta calcular $e^{-\widehat{\theta}_{MV}}$. Essa propriedade é exclusiva do Método de Máxima Verossimilhança e não é garantida, em geral, pelo Método dos Momentos.

(b) O Estimador de Máxima Verossimilhança é preferido na prática atuarial porque garante, necessariamente, que a estimativa seja não-viesada (ou seja, $E[\widehat{\theta}] = \theta$) para qualquer tamanho de amostra.

(c) Sob condições de regularidade, à medida que o tamanho da amostra (histórico de sinistros) aumenta, a variância do estimador de MV converge para o Limite Inferior de Cramér-Rao, tornando-o assintoticamente eficiente.

10.15) Em uma situação onde o estimador pelo Método dos Momentos (EMM) fornece um resultado diferente do Estimador de Máxima Verossimilhança (EMV), qual dos dois geralmente é preferível para grandes amostras? Cite duas vantagens teóricas do EMV que justificam essa escolha.

Respostas:

10.1) (a) $\widehat{\theta}_{MV} = \overline{X}$

(b) $\widehat{\theta}_{MV} = \frac{1}{\overline{X}}$

(c) $\widehat{\theta}_{MV} = \overline{X}$

10.2) (a) $\widehat{\theta}_{MV} = -\frac{n}{\sum_{i=1}^{n} \ln x_i}$

10.3) $\widehat{\theta}_{MV} = \frac{\overline{X}}{6}$

10.4) $\widehat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} \ln(x_i/k)}$

10.5) (a) $\widehat{\psi}_{MV} = \frac{\overline{X}}{1-\overline{X}}$

(b) $\widehat{Var(X)}_{MV} = \overline{X}(1-\overline{X})$

10.6) (a) $\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$

(b) $-\frac{n}{2\widehat{\sigma}^2} + \frac{1}{2(\widehat{\sigma}^2)^2}\sum_{i=1}^{n}(x_i - \widehat{\mu})^2 = 0$

(c) $\widehat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

(d) $\widehat{\sigma}^2_{MV}$ é viesado, enquanto S^2 não.

10.7) (a) $\ell(\mu, \sigma^2) = -\sum \ln x_i - \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(\ln x_i - \mu)^2$

(b) $\begin{cases} \frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum(\ln x_i - \mu) = 0 \\ \frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(\ln x_i - \mu)^2 = 0 \end{cases}$

(c) $\widehat{\mu}_{MV} = \frac{\sum_{i=1}^{n} \ln x_i}{n}$ e $\widehat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \widehat{\mu})^2$

10.8) (a) $I_1(\lambda) = -E\left[-\frac{1}{\lambda^2}\right] = \frac{1}{\lambda^2}$

(b) $\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda^2)$ ou $\widehat{\lambda} \overset{aprox}{\sim} N\left(\lambda, \frac{\lambda^2}{n}\right)$

(c) $\widehat{R(1)} \overset{aprox}{\sim} N\left(e^{-\lambda}, \frac{\lambda^2 e^{-2\lambda}}{n}\right)$

10.9) (a) $\widehat{\lambda}_{MV} = 4$ falhas por semana.

(b) $\widehat{P}(X = 0) = e^{-4} \approx 0,0183$

10.10) (a) Estimativa da média $(1/\lambda)$: $\widehat{\mu}_{MV} = \overline{x} = 2,4$ minutos
Estimativa da taxa (λ): $\widehat{\lambda}_{MV} = \frac{1}{2,4} \approx 0,4167$ clientes por minuto

(b) $\widehat{P}(X > 3) \approx 0,2865$

10.11) Julgue as afirmações abaixo como Verdadeiras ou Falsas, justificando brevemente:

(a) Verdadeiro. Esta afirmação decorre da Propriedade da Invariância dos estimadores de máxima verossimilhança. Ela estabelece que se $\widehat{\theta}$ é o EMV de θ, então para qualquer função $g(\theta)$ (como a raiz quadrada), o EMV de $g(\theta)$ é dado por $g(\widehat{\theta})$.

(b) Verdadeiro. Esta é a propriedade da Eficiência Assintótica. Sob condições de regularidade, a variância do EMV converge para o inverso da Informação de Fisher ($1/I_n(\theta)$), que é exatamente o Limite Inferior de Cramér-Rao, tornando-o o estimador com a menor variância possível entre os estimadores consistentes para grandes amostras.

(c) Verdadeiro. Pelo Teorema da Fatoração de Neyman-Fisher, a verossimilhança pode ser escrita como $L(\theta) = g(T(\mathbf{x}); \theta)h(\mathbf{x})$. Ao maximizar em relação a θ, o termo $h(\mathbf{x})$ é irrelevante, e a maximização depende dos dados exclusivamente através da estatística suficiente $T(\mathbf{X})$.

(d) Falso. A equação de score frequentemente resulta em expressões transcedentais ou complexas que não possuem solução analítica fechada. Exemplos incluem a estimação dos parâmetros de forma da distribuição Gama ou os coeficientes em uma Regressão Logística. Nestes casos, o uso de métodos numéricos iterativos (como Newton-Raphson ou Fisher-Scoring) é obrigatório.

(e) Falso. A condição $\frac{\partial \ell}{\partial \theta} = 0$ determina apenas um ponto crítico, que pode ser um máximo local, um mínimo local ou um ponto de sela. Para garantir que é um máximo global, é necessário: (1) verificar a segunda derivada (Hessiana negativa definida) e (2) comparar o valor da verossimilhança no ponto crítico com os valores nas fronteiras do espaço paramétrico.

10.12) A diferença reside na definição de variável e parâmetro fixo. Na densidade $f(\mathbf{x}; \theta)$, o parâmetro θ é fixo e $\mathbf{x}$ varia sobre o espaço amostral. Por definição de probabilidade, sua integral em relação a $\mathbf{x}$ é sempre 1. Na verossimilhança $L(\theta; \mathbf{x})$, os dados $\mathbf{x}$ são fixos (observados) e θ varia sobre o espaço paramétrico. Como L mede a plausibilidade e não a probabilidade de θ (que não é aleatório na inferência clássica), não há restrição matemática que faça sua integral em relação a θ ser igual a 1.

10.13) A curvatura da função de log-verossimilhança é inversamente proporcional à variância do estimador. Curva "pontiaguda" (alta curvatura) indica que o máximo é bem definido e os dados fornecem muita informação sobre θ. Isso implica alta Informação de Fisher e, consequentemente, baixa variância (alta precisão). Já a curva "achatada" (baixa curvatura) indica incerteza na localização do máximo. Implica baixa Informação de Fisher e alta variância (baixa precisão). Matematicamente, a variância assintótica é o inverso da curvatura esperada: $Var(\widehat{\theta}) \approx \frac{1}{I(\theta)}$.

10.14) (a) Verdadeiro

(b) Falso. O EMV é frequentemente viesado em pequenas amostras. O que o EMV garante é ser consistente (o viés desaparece com $n \to \infty$).

(c) Verdadeiro

10.15) Geralmente, o Estimador de Máxima Verossimilhança (EMV) é preferível. Uma vantagem teórica é a Eficiência Assintótica (para grandes amostras, o EMV possui a menor variância possível entre os estimadores consistentes – atinge o Limite Inferior de Cramér-Rao), sendo frequentemente mais preciso que o EMM. Outra vantagem é o Princípio da Invariância (o EMV de uma função do parâmetro $g(\theta)$ é dado simplesmente por $g(\widehat{\theta})$). O Método dos Momentos não garante essa propriedade (ex: $E[\overline{X}^2] \neq \mu^2$).