Cálculo das Probabilidades II - Lista 3 - 2020/02

Prof. Hugo Carvalho

02/06/2021

– INSTRUÇÕES – LEIAM ATENTAMENTE! –

- A data limite de entrega da avaliação é sábado 12/06/2021 às 23h59'. Avaliações entregues após esse prazo serão desconsideradas.
- A entrega deve ser feita exclusivamente através do Google Classroom, clicando na caixa "+ Adicionar ou Criar" dentro da postagem dessa lista, para então anexar um arquivo com sua resolução. Após isso, clique em "Entregar" para enviar sua resolução.
  Atenção: Somente anexar a resolução não é suficiente! O envio deve ser feito para que sua resolução de fato seja entregue.
- Você tem a liberdade de escrever sua resolução no computador (usando Word, LaTeX, dentre outros), ou manuscrito e depois escanear ou fotografar a sua resolução. Nesse último caso, tome cuidado para que o documento fique legível. No caso de fotografar, opte por utilizar luz natural e tome cuidado com sombras.
- **Dica**: Se for fotografar sua resolução com um *smartphone* ou *tablet*, utilize o aplicativo próprio da câmera, e evite fotografar através de WhatsApp, Telegram, Messenger, e outros. Os aplicativos de comunicação, ao utilizarem a câmera, fazem uma severa compressão da imagem, incorrendo em uma grande diminuição de sua qualidade. Para transferir a imagem do celular para o computador prefira fazer o envio por e-mail, ou acessando sua galeria de fotos através do Google Photos no computador (caso já utilize esse aplicativo para gerenciar suas fotos no aparelho).
- Independente do modo de escrita, a resolução deve ser entregue em um único documento, no formato PDF, com a resolução em pé (formato retrato). O *layout* da resolução não será levado em consideração na avaliação, porém o texto deve estar legível para ser corrigido.
  Atenção: Resoluções ilegíveis ou fora desse formato não serão corrigidas e serão desconsideradas.
- A troca de conhecimento na realização da avaliação é permitida e encorajada: ciência se faz com colaboração, e devemos seguir esse espírito aqui. Porém, cada aluno deverá ter a sua própria resolução, e cópias ou outras ilegalidades serão severamente punidas com a anulação da avaliação para o(s) aluno(s) suspeito(s).
- Todos os passos de sua resolução devem ser devidamente justificados.
- Ao entregar essa avaliação, você afirma ter lido e estar de acordo com essas regras, comprometendo-se a cumpri-las.

– PROPRIEDADES ASSINTÓTICAS DE ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA –

Seja $X_1, X_2, X_3, \ldots$ uma sequência independente e identicamente distribuída de variáveis aleatórias, todas com a mesma distribuição de uma variável aleatória fixa X, cuja função densidade de probabilidade é dada por $f(x;\theta_\star)$, onde $\theta_\star \in \Theta \subset \mathbb{R}$ é um parâmetro da distribuição – assumido ser um valor fixo porém a nós desconhecido – e o conjunto Θ é chamado de *espaço paramétrico*, denotando o conjunto de valores que tal parâmetro pode assumir. Caso tenhamos interesse em estimar o valor $\theta_\star$ podemos considerar a *função de verossimilhança*, definida como

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} f(x_i;\theta),$$

onde $x_1, x_2, x_3, \ldots$ é uma sequência de observações das variáveis aleatórias $X_1, X_2, X_3, \ldots$. Note que a função de verossimilhança pode ser interpretada como "*algo relacionado à probabilidade dos dados observados assumindo o particular valor θ para o parâmetro da distribuição*". Dessa forma, podemos considerar o *estimador de máxima verossimilhança* (EMV), definido como

$$\widehat{\theta}_n = \underset{\theta\in\Theta}{\operatorname{argmax}}\, \mathcal{L}_n(\theta),$$

interpretado como "*o valor de θ que torna os dados observados os mais prováveis*". Imagina-se, obviamente, que para n suficientemente grande teremos $\widehat{\theta}_n \approx \theta_\star$, porém ao notarmos que $\widehat{\theta}_n$ é também uma variável aleatória, devemos tomar cuidado ao formular tal afirmação. O objetivo desta avaliação é formalizar tais conceitos.

Questão 1: *(Consistência do EMV)* O objetivo desta questão é provar que o EMV é *consistente*, ou seja, sob hipóteses adequadas, temos que $\widehat{\theta}_n \xrightarrow{p} \theta_\star$. Para isso, siga o roteiro abaixo.

a) Mostre que maximizar $\mathcal{L}_n(\theta)$ é equivalente a maximizar $M_n(\theta)$, sendo tal função definida por

$$M_n(\theta) := \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{f(x_i;\theta)}{f(x_i;\theta_\star)}\right).$$

b) Mostre que $M_n(\theta) \xrightarrow{p} -D(\theta_\star,\theta)$, onde a função D é definida como

$$D(\theta_\star,\theta) := \int_{-\infty}^{+\infty} f(x;\theta_\star)\log\left(\frac{f(x;\theta_\star)}{f(x;\theta)}\right)\, dx.$$

c) Vamos assumir, sem prova, dois fatos importantes sobre a função D e a família paramétrica $\{f(\,\cdot\,;\theta) \mid \theta\in\Theta\}$:

   i) $D(\theta_\star,\theta) \geq 0$, para todo $\theta \in \Theta$,

   ii) $D(\theta_\star,\theta) = 0$ se e somente se $\theta = \theta_\star$.

   Com base nisso, argumente que, no limite quando $n \to \infty$, temos que $\theta_\star$ maximiza $\mathcal{L}_n$. Use isso para concluir a consistência do EMV.

   Obs.: Pode ser que neste item a sua argumentação não seja formal. Não há problema algum, estamos jogando algumas coisas para baixo do tapete, que podem ser resolvidas no item abaixo.

d) (BÔNUS) Aqui provaremos formalmente o resultado argumentado no item acima. Para isso, defina $M(\theta) := -D(\theta_\star,\theta)$ e assuma as duas hipóteses a seguir:

   i) $\sup_{\theta\in\Theta} |M_n(\theta) - M(\theta)| \xrightarrow{p} 0$,

   ii) Para todo $\varepsilon > 0$ vale que $\sup_{\{\theta \,\mid\, |\theta-\theta_\star|\geq\varepsilon\}} M(\theta) < M(\theta_\star)$.

   Com base nisso, conclua formalmente que $\widehat{\theta}_n \xrightarrow{p} \theta_\star$.

Questão 2: *(Normalidade assintótica do EMV)* Aqui provaremos que o EMV é *normalmente assintótico*, ou seja, após ser devidamente escalado, a distribuição de $\widehat{\theta}_n$ é aproximadamente normal, se n é suficientemente grande. Para isso, siga o roteiro abaixo.

a) Defina as funções s e S como abaixo:

$$s(\theta; x) := \frac{\partial}{\partial \theta} \log f(x; \theta),$$
$$S_n(\theta; x_1, \ldots, x_n) := \sum_{i=1}^{n} s(\theta; x_i).$$

Mostre que $\mathbb{E}_\theta[s(\theta; X)] = 0$, onde $\mathbb{E}_\theta[\,\cdot\,]$ denota o valor esperado com respeito à densidade $f(x; \theta)$.

Obs.: A função s é conhecida como função score*, extremamente importante em Estatística.*

b) Usando o resultado do item anterior, mostre que $\mathbb{V}_\theta(s(\theta; X)) = -\mathbb{E}_\theta\left[\frac{\partial}{\partial \theta} s(\theta; X)\right]$, onde $\mathbb{V}_\theta(\,\cdot\,)$ denota a variância com respeito à densidade $f(x; \theta)$.

Obs.: Uma razão da grande importância da função score é por ela fazer parte da definição da informação de Fisher*, definida como $I(\theta) := \mathbb{V}_\theta(s(\theta; X))$.*

c) Argumente, usando o Teorema do Valor Médio, que a reta secante ao gráfico de $\theta \mapsto S_n(\theta; x_1, \ldots, x_n)$ que passa pelos pontos $(\theta_\star, S_n(\theta_\star; x_1, \ldots, x_n))$ e $(\widehat{\theta}_n, S_n(\widehat{\theta}_n; x_1, \ldots, x_n))$ tem inclinação aproximadamente igual a $\frac{\partial}{\partial \theta} S_n(\theta_\star; x_1, \ldots, x_n)$.

d) Conclua que $\sqrt{n}(\widehat{\theta}_n - \theta_\star) \approx \dfrac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} s(\theta_\star; x_i)}{-\frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \theta} s(\theta_\star; x_i)}$.

e) Vamos agora lembrar que o numerador e denominador da fração no item acima são funções das amostras, e portanto, variáveis aleatórias. Defina então

$$N_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} s(\theta_\star; X_i),$$
$$D_n := -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \theta} s(\theta_\star; X_i).$$

Argumente que $D_n \xrightarrow{p} I(\theta_\star)$.

f) Na notação do item acima, argumente que $N_n \xrightarrow{d} \mathrm{N}(0, I(\theta_\star))$.

g) Conclua finalmente que $\dfrac{N_n}{D_n} \xrightarrow{d} \mathrm{N}\left(0, \dfrac{1}{I(\theta_\star)}\right)$, e portanto, que o EMV é normalmente assintótico.