

## MAE5911/IME: Fundamentos de Estatística e Machine Learning. Prof.: Alexandre Galvão Patriota

**Questão 01:** Defina formalmente o valor-p (**aproximado**) para testar uma hipótese geral e discuta os problemas de interpretação usuais ao utilizar a sua versão condicional à hipótese nula. Apresente um exemplo numérico e simulações de Monte Carlo, usando um modelo de regressão Poisson, para ilustrar a sua aplicação.

**Questão 02:** Apresente um Teorema da Aproximação Universal que generalize o Teorema de Funahashi, discuta as diferenças entre os resultados por meio de exemplos de redes neurais.

**Questão 03:** Proponha uma função de estimação robusta para estimar os parâmetros de uma regressão Binomial,  $Y|X = x \sim \text{Bin}(m, \mu_\theta(x))$ , em que  $\mu_\theta(\cdot)$  é uma função com imagem em  $[0, 1]$  e  $\theta$  é o vetor de parâmetros. Apresente um exemplo numérico e uma simulação de Monte Carlo, como feito em sala, para ilustrar os resultados. Perturbe a distribuição dos dados e mostre que a sua proposta é de fato robusta contra essa perturbação.

### Resposta para o Questão 01

O conceito do p-valor, tal como proposto originalmente por *Fisher* (1925), é uma medida **condicional sob  $H_0$** , definida por

$$\text{valor-p}(H_0, z_n) = \sup_{\theta \in \Theta_0} P_\theta^{(n)}(T_{H_0}(Z_n) \geq T_{H_0}(z_n)),$$

isto é, a probabilidade de se observar uma estatística tão extrema quanto a obtida, assumindo  $H_0$  verdadeira. O valor-p foi concebido por Fisher como um *instrumento de contraste* entre o dado e hipótese, mas não como uma medida de verdade. Seu objetivo era expressar o grau de incompatibilidade empírica entre o fenômeno observado e as consequências lógicas de  $H_0$ .

Posteriormente, a escola *Neyman–Pearson* reinterpreto esse conceito dentro de uma estrutura de decisão. Nessa transição, o valor-p perdeu seu caráter exploratório e passou a ser tratado como um critério binário, expresso por: “rejeita” ou “não rejeita”  $H_0$ . Esta leitura gerou um equívoco lógico: o valor-p, que é definido dentro do universo em que  $H_0$  é verdadeira, passou a ser interpretado como se pudesse mensurar sobre a probabilidade de  $H_0$  ser verdadeira.

O **valor-p assintótico**, que é o valor usualmente utilizado, na ausência da distribuição exata de  $T_{H_0}$ , é definido por:

$$p^{(a)}(H_0, z_n) = \sup_{\theta \in \Theta_0} P_\theta(W_\theta \geq T_{H_0}(z_n)),$$

em que  $W_\theta$  representa a distribuição-limite de  $T_{H_0}$  quando  $n \rightarrow \infty$ .

O valor-p assintótico não representa um aprimoramento do valor exato, mas a revelação de uma contradição epistemológica. Em tese, se o modelo assintótico descrevesse adequadamente o comportamento empírico, o valor-p aproximado deveria convergir para o valor-p exato à medida que  $n \rightarrow \infty$ . Entretanto, essa convergência é apenas formal: ela ocorre dentro do próprio universo lógico de  $H_0$ , e não em relação ao mundo empírico dos dados. Assim, o valor-p assintótico não corrige o exato, mas o expõe — mostrando que a própria definição do valor-p repousa sobre uma idealização que não pode alcançar o real. O limite teórico que deveria representar a aproximação à verdade empírica termina, paradoxalmente, por evidenciar a distância entre a coerência lógica do modelo e a imperfeição do fenômeno que ele tenta descrever.

Essa versão não é um aprimoramento do valor exato, mas uma *idealização teórica*: o cálculo passa a ocorrer em um mundo em que o tamanho da amostra tende ao infinito e as condições de regularidade são perfeitamente satisfeitas. Nesse sentido, o valor-p aproximado não se aproxima da verdade empírica, mas de uma coerência lógica interna ao modelo.

Assim, o valor-p nunca mede a probabilidade de  $H_0$ , mas apenas a compatibilidade entre dado e modelo, isto é, o quanto o fenômeno observado tensiona o modelo que o descreve. Tratá-lo como probabilidade de  $H_0$  é tentar inferir o mundo empírico a partir de uma suposição epistemológica autorreferente. Por isso, o valor-p não

deve ser entendido como uma medida de crença, mas apenas como originalmente proposto — uma **medida de discrepância entre o real e o teórico** — que opera em um único sentido: como evidência contra a hipótese nula, nunca para confirmá-la.

Em termos das  $\sigma$ -álgebras envolvidas, a construção inferencial pode ser representada como uma sequência de aplicações mensuráveis:

$$(\Omega_n, \mathcal{F}_n, \mathbb{P}_n) \xrightarrow{Z_n} (\mathcal{Z}_n, \mathcal{B}_{Z_n}, \mathbb{P}_{Z_n}) \xrightarrow{T} (\mathcal{T}_n, \mathcal{B}_{T_n}, \mathbb{P}_{T_n}).$$

Cada espaço de probabilidade possui a sua própria  $\sigma$ -álgebra, refletindo níveis distintos de abstração:

- $\mathcal{F}_n$  — descreve o universo empírico de incertezas, isto é, os eventos do mundo observável;
- $\mathcal{B}_{Z_n}$  — corresponde ao espaço amostral modelado, onde as variáveis aleatórias  $Z_n$  são definidas sob  $\mathbb{P}_{Z_n}$ ;
- $\mathcal{B}_{T_n}$  — é a  $\sigma$ -álgebra induzida pela estatística  $T$ , onde vivem as distribuições teóricas dos testes e, em particular, o valor-p.

O ponto crucial é que o valor-p é mensurável apenas em relação à  $\sigma$ -álgebra  $\mathcal{B}_{T_n}$ , isto é, dentro do universo lógico em que a hipótese nula  $H_0$  é assumida como verdadeira. Ele não pertence à  $\sigma$ -álgebra empírica  $\mathcal{F}_n$ , e portanto não pode ser interpretado como uma probabilidade sobre o mundo real ou sobre a veracidade de  $H_0$ . Tratá-lo dessa forma constitui um erro de violação da  $\sigma$ -álgebra — um deslize epistemológico em que se tenta extrair informação empírica de uma construção puramente condicional e autorreferente.

## Resposta para o Questão 01

O valor-p, tal como proposto originalmente por Fisher (1925), é uma medida **condicional sob  $H_0$** , definida por

$$p(H_0, z_n) = P_\theta^{(n)}(T_{H_0}(Z_n) \geq T_{H_0}(z_n) \mid H_0),$$

isto é, a probabilidade de se observar uma estatística tão extrema quanto a obtida, assumindo  $H_0$  verdadeira. O valor-p foi concebido por Fisher como um *instrumento de contraste* entre dado e hipótese, e não como uma medida de verdade. Seu objetivo era expressar o grau de incompatibilidade empírica entre o fenômeno observado e as consequências lógicas de  $H_0$ .

Posteriormente, a escola Neyman–Pearson reinterpretou esse conceito dentro de uma estrutura de decisão. Nessa transição, o valor-p perdeu seu caráter exploratório e passou a ser tratado como um critério binário, expresso por: “rejeita” ou “não rejeita”  $H_0$ . Essa leitura gerou um equívoco lógico: o valor-p, que é definido dentro do universo em que  $H_0$  é verdadeira, passou a ser interpretado como se dissesse algo sobre a probabilidade de  $H_0$  ser verdadeira.

Na ausência da distribuição exata de  $T_{H_0}$ , o **valor-p assintótico** é definido por:

$$p^{(a)}(H_0, z_n) = \sup_{\theta \in \Theta_0} P_\theta(W_\theta \geq T_{H_0}(z_n)),$$

em que  $W_\theta$  representa a distribuição-limite de  $T_{H_0}$  quando  $n \rightarrow \infty$ . Essa versão não é um aprimoramento do valor exato, mas uma *idealização teórica*: o cálculo passa a ocorrer em um mundo em que o tamanho da amostra tende ao infinito e as condições de regularidade são perfeitamente satisfeitas. Nesse sentido, o valor-p aproximado não se aproxima da verdade empírica, mas de uma coerência lógica interna ao modelo.

Assim, o valor-p nunca mede a probabilidade de  $H_0$ , mas apenas a compatibilidade entre dado e modelo, isto é, o quanto o fenômeno observado tensiona o modelo que o descreve. Tratá-lo como probabilidade de  $H_0$  é tentar inferir o mundo empírico a partir de uma suposição epistemológica autorreferente. Por isso, o valor-p não deve ser entendido como uma medida de crença, mas apenas como originalmente proposto, uma **medida de discrepância entre o real e o teórico** — e que apenas tem o poder de ser utilizado em uma única direção e sentido: como uma evidência contra a hipótese nula, mas nunca para confirmá-la.

## Resposta para o Questão 02

## Resposta para o Questão 03

---

Na ausência da distribuição exata de  $T_{H_0}$ , define-se o **valor-p assintótico**:

$$p^{(a)}(H_0, z_n) = \sup_{\theta \in \Theta_0} P_\theta(W_\theta \geq T_{H_0}(z_n)),$$

em que  $W_\theta$  representa a distribuição-limite de  $T_{H_0}$  quando  $n \rightarrow \infty$ . Essa versão não é um aprimoramento do valor exato, mas uma *idealização teórica*: o cálculo passa a ocorrer em um mundo em que o tamanho da amostra tende ao infinito e as condições de regularidade são perfeitamente satisfeitas. Nesse sentido, o valor-p aproximado não se aproxima da verdade empírica, mas de uma coerência lógica interna ao modelo.

O valor-p tal como proposto originalmente por Fisher (1925), é uma medida **condicional sob**  $H_0$ , definida por

$$p(H_0, z_n) = P_\theta^{(n)}(T_{H_0}(Z_n) \geq T_{H_0}(z_n) \mid H_0),$$

isto é, a probabilidade de se observar uma estatística tão extrema quanto a obtida, assumindo  $H_0$  verdadeira. O valor-p foi concebido por Fisher como um *instrumento de contraste* entre dado e hipótese, e não como uma medida de verdade. Seu objetivo era expressar o grau de incompatibilidade empírica entre o fenômeno observado e as consequências lógicas de  $H_0$ .

Posteriormente, a escola *Neyman–Pearson* reinterpretou esse conceito dentro de uma estrutura de decisão formal, introduzindo regras fixas de rejeição, níveis de significância ( $\alpha$ ) e poder do teste. Nessa transição, o valor-p perdeu seu caráter exploratório e passou a ser tratado como um *critério binário de decisão* — “rejeita” ou “não rejeita”  $H_0$ . Essa leitura gerou um equívoco lógico: o valor-p, que é definido dentro do universo em que  $H_0$  é verdadeira, passou a ser interpretado como se dissesse algo sobre a probabilidade de  $H_0$  ser verdadeira.

Assim, o valor-p nunca mede a probabilidade de  $H_0$ , mas apenas a compatibilidade entre dado e modelo, isto é, o quanto o fenômeno observado tensiona o universo hipotético que o contém. Tratá-lo como probabilidade de  $H_0$  é tentar inferir o mundo empírico a partir de uma suposição epistemológica autorreferente e idealizada. Por isso, o valor-p deve ser entendido não como medida de crença, mas como **medida de discrepância entre o real e o teórico** — uma forma de pontuar a coerência entre o fenômeno e o modelo, e não de confirmar a verdade da hipótese nula.