Segunda avaliação (A2)

Disciplina: Inferência Estatística
Professor: Luiz Max Carvalho

29 de Novembro de 2021

- Por favor, entregue um único arquivo PDF;
- O tempo para realização da prova é de 3 horas, mais vinte minutos para upload do documento para o e-class;
- Leia a prova toda com calma antes de começar a responder;
- Responda todas as questões sucintamente;
- Marque a resposta final claramente com um quadrado, círculo ou figura geométrica de sua preferência;
- A prova vale 80 pontos. A pontuação restante é contada como bônus;
- Lembre-se de consultar o catálogo de fórmulas no fim deste documento.

Dicas

- Se $Y \sim \text{Poisson}(\lambda)$ então, para λ grande o suficiente, temos
  - $E[\sqrt{Y} + \sqrt{Y+1}] \approx \sqrt{4\lambda + 1}$;
  - $\text{Var}(\sqrt{Y} + \sqrt{Y+1}) \approx 1$.

  Ademais, temos
  $$\frac{Y - \lambda}{\sqrt{\lambda}} \to \text{Normal}(0, 1),$$
  onde a seta representa convergência em distribuição.

- Se X tem distribuição Poisson com média λ, então
  $$\Pr(X \leq x) = Q(\lfloor x + 1 \rfloor, \lambda),$$
  onde
  $$Q(x, s) = \frac{\Gamma(x, s)}{\Gamma(x)}$$
  é a função Gama regularizada superior e $\lfloor y \rfloor$ é maior inteiro menor ou igual a y – também chamado de *floor*. Ademais, temos
  $$\frac{\partial}{\partial s} Q(x, s) = -\frac{e^{-s} s^{x-1}}{\Gamma(x)},$$
  onde $\Gamma(x) = (x - 1)!$ é a função Gamma.

- Se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória, então
  $$\Pr(\max(X_1, X_2, \ldots, X_n) \leq m) = \prod_{i=1}^{n} \Pr(X_i \leq m).$$

- Em uma regressão linear simples, temos:
  $$\begin{aligned} \hat{\beta_0} &\sim \text{Normal}\left(\beta_0, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{s_x^2}\right)\right), \\ \hat{\beta_1} &\sim \text{Normal}\left(\beta_1, \frac{\sigma^2}{s_x^2}\right), \\ \text{Cov}\left(\hat{\beta_0}, \hat{\beta_1}\right) &= -\frac{\bar{x}\sigma^2}{s_x^2}, \end{aligned}$$
  onde $s_x = \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ e $\hat{\beta_0}$ e $\hat{\beta_1}$ são os estimadores de máxima verossimilhança dos coeficientes.

1. A duck row

Tio Patinhas e seus sobrinhos continuam sua jornada no aprendizado da Estatística e consideram $X_1, X_2, \ldots, X_n \sim \text{Uniforme}(0, \theta)$, $\theta > 0$, uma amostra aleatória com $M := \max(X_1, X_2, \ldots, X_n)$. Várias dúvidas e discordâncias surgiram entre os patinhos, no entanto. No que se segue, você deve utilizar seus conhecimentos de Estatística para esclarecer as coisas.

a) (5 pontos) Huguinho diz que $Y = \theta^2 M$ é pivotal, Zezinho discorda. Mostre que Huguinho está equivocado e encontre uma quantidade pivotal.

b) (5 pontos) Mostre ao Pato Donald como utilizar a quantidade obtida no item anterior para obter um intervalo de confiança unilateral exato de $100\gamma\%$ para θ.

c) (5 pontos) Defina[1] as hipóteses $H_0 : \theta = 1/3$ e $H_1 : \theta > 1/3$. Suponha que propomos um procedimento de teste δ_c que rejeita H_0 quando $M > c$. Deduza a função poder do teste e encontre c de modo que o teste tenha tamanho α_0.

d) (10 pontos) Luisinho cismou que o este teste tem razão de verossimilhanças monótona. Ele está certo? Justifique sua resposta.

e) (10 pontos) Tio Patinhas jura que este teste é uniformemente mais poderoso. Explique para Zezinho o que isso significa e depois discuta se o Tio Patinhas tem razão. Justifique sua resposta.

Conceitos trabalhados: Quantidade pivotal; intervalo de confiança; testes de hipótese; poder; tamanho; UMP.
Nível de dificuldade: médio.
Resolução: Para resolver a) vamos começar lembrando que

$$\begin{aligned} \Pr(M \leq m) &= \prod_{i=1}^{n} \Pr(X_i \leq m), \\ &= \left[\frac{m}{\theta}\right]^n. \end{aligned}$$

Desta forma, temos que

$$\begin{aligned} \Pr(Y \leq y) &= \Pr\left(M \leq \frac{y}{\theta^2}\right), \\ &= \left[\frac{y}{\theta^3}\right]^n, \end{aligned}$$

donde concluímos que $Y = \theta^2 M$ não é pivotal. Os cálculos acima sugerem que $W = M/\theta$ é de fato pivotal:

$$\begin{aligned} \Pr(W \leq w) &= \Pr(M \leq \theta w), \\ &= \left[\frac{\theta w}{\theta}\right]^n = [w]^n, \end{aligned}$$

[1] Preste atenção à formulação das hipóteses!

mostrando que a distribuição de W não depende de θ, como desejado. Para ajudar o Pato Donald em b), vamos primeiro notar que queremos um intervalo da forma $I(\boldsymbol{X}) = (M, b(\boldsymbol{X}))$ de modo que valha

$$\Pr(M \leq \theta \leq b(\boldsymbol{X})) \geq \gamma.$$

Utilizando os cálculos acima, sabemos que

$$\Pr\left(\theta \leq \frac{M}{w}\right) = 1 - w^n,$$

o que sugere que precisamos encontrar w de modo que $1 - w^n = \gamma$, ou seja, $w = (1-\gamma)^{1/n}$. Assim, concluímos que $I(\boldsymbol{X}) = (M, M/(1-\gamma)^{1/n})$ é um intervalo de confiança de $100\gamma\%$ exato para θ. Outras soluções são possíveis. Para resolver c) precisamos lembrar que o poder $\pi(\theta \mid \delta_c)$ é a probabilidade de M cair na região de rejeição:

$$\Pr(M > c \mid \theta) = 1 - \left[\frac{c}{\theta}\right]^n.$$

Dessa forma, pela definição de tamanho de um teste, temos

$$\alpha_0 := \Pr(M > c \mid \theta = 1/3) = 1 - [3c]^n,$$

portanto $c = \frac{(1-\alpha_0)^{1/n}}{3}$ é o valor crítico desejado. Agora vamos atacar d): para começar sabemos que a função de verossimilhança é

$$f_n(\boldsymbol{x} \mid \theta) = \begin{cases} \frac{1}{\theta^n}, \ M \leq \theta, \\ 0, \text{ caso contrário.} \end{cases}$$

Desta forma, queremos mostrar que para $\theta_1, \theta_2 \in (0, \infty)$ com $\theta_1 < \theta_2$ a razão $f_n(\boldsymbol{x} \mid \theta_2)/f_n(\boldsymbol{x} \mid \theta_1)$ depende da amostra apenas através de M e é função monotônica desta estatística. De fato

$$\frac{f_n(\boldsymbol{x} \mid \theta_2)}{f_n(\boldsymbol{x} \mid \theta_1)} = \left[\frac{\theta_2}{\theta_1}\right]^n \mathbb{I}(M \leq \theta_1).$$

Agora vamos descobrir se o teste proposto é UMP. Vamos começar explicando para o Zezinho que um teste UMP δ é aquele que, para um tamanho α_0, satisfaz

$$\pi(\delta' \mid \theta) \leq \pi(\delta \mid \theta),$$

para todo $\theta \in \Omega_1$ e para qualquer outro teste δ'. Ou seja, um teste UMP *domina* todos os outros testes de tamanho α_0 em termos de poder em Ω_1 – e, portanto, tem o menor erro do tipo II. Segundo o Teorema 9.3.1 de DeGroot, provado também no item 3 do Trabalho IV, sabemos que para hipóteses da forma $H_0 : \theta = \theta_0$ e $H_1 : \theta < \theta_0$ ou $H_1 : \theta > \theta_0$ um teste com razão de verossimilhanças monotônica é UMP, confirmando a suspeita do Tio Patinhas. ■

2. The God(damn) particle

O bóson de Higgs[2], que confere massa a partículas que de outra forma não teriam massa, foi a última partícula do chamado Modelo Padrão (*standard model*) da Física Quântica a ter sua existência confirmada experimentalmente.

[2] Assim nomeado em homenagem ao físico inglês Peter Higgs (1929-).

A ideia fundamental da descoberta experimental de novas partículas é a de comparar o que se espera sob um modelo base sem a nova partícula e um modelo proposto que a inclui. Uma maneira de fazer isso é detectar o número X de ocorrências (eventos) de decaimento de bósons e comparar esse número com o que é esperado sob o modelo mais simples. Como o medidor é imperfeito, considere o seguinte modelo:

$$X \sim \text{Poisson}(\beta + \kappa\mu),$$

com $\mu > 0$, onde vamos assumir que a taxa base $\beta > 0$ e a taxa esperada do bóson de Higgs $\kappa > 0$ são conhecidas. Desta forma, o modelo com $\mu = 0$ é o modelo de base e aquele onde $\mu = 1$ é o modelo com o bóson de Higgs.

a) (10 pontos) Encontre o EMV de μ e comente sobre a sua adequação; o EMV sempre produz estimativas válidas?

b) (10 pontos) Encontre um intervalo de confiança bilateral de $100\gamma\%$ **aproximado** para μ;

   Dica: Você pode assumir que β é grande o suficiente. Use as dicas no começo da prova.

c) (10 pontos) Considere um teste em que assumimos que $X \sim \text{Poisson}(\theta)$ e construímos um teste δ_c de modo que, se $X > c$, rejeitamos $H_0 : \theta \leq \theta_0$.

   - Mostre que o teste em questão é não-viesado.
   - Escreva θ_0 em função de β, κ e μ de modo a testar um desvio em relação ao modelo base.

     Dica: Existem várias respostas certas.

d) (10 pontos) Formule um teste para $H_0 : \mu = 0$ vs $H_1 : \mu > 0$ como um teste de razão de verossimilhanças;

Conceitos trabalhados: Estimador de máxima verossimilhança; intervalos de confiança aproximados; teste de razão de verossimilhanças.
Nível de dificuldade: médio.
Resolução: Para encontrar o EMV de μ, vamos começar escrevendo a verossimilhança

$$f_n(X = x \mid \mu) = \frac{(\beta + \kappa\mu)^x \exp(-(\beta + \kappa\mu))}{x!},$$

e o seu logaritmo

$$\log(f_n(x \mid \mu)) =: l(\mu) = x\log(\beta + \kappa\mu) - (\beta + \kappa\mu)),$$

onde, por conveniência, descartamos termos que não dependem dos parâmetros. Agora, temos

- $l'(\mu) = \frac{\kappa x}{\kappa\mu + \beta} - \kappa,$
- $l''(\mu) = -\frac{\kappa^2 x}{(\beta + \kappa\mu)^2},$

que imediatamente nos diz que a função é côncava com respeito a μ de modo que podemos obter

$$\hat{\mu}_{\text{EMV}} = \frac{x - \beta}{\kappa}$$

como nosso maximizador. Note que esse estimador não garante estimativas em $(0, \infty)$; quando a contagem observada é, por acaso, menor que o esperado pelo modelo de base, isto é, menor que a taxa base β, temos estimativas negativas, que não fazem sentido. Uma maneira de corrigir isso seria propor o estimador modificado

$$\hat{\mu}_{\text{ALT}} = \max\left(0, \frac{x - \beta}{\kappa}\right).$$

O interessante é notar que essa deterioração do estimador acontece perto do limite de detecção, isto é, quando a contagem é baixa e perto da taxa base, β. Para obter um intervalo aproximado e responder b), vamos assumir que β é grande o suficiente, de modo que $\beta + \kappa\mu$ se adequa às condições das dicas. Em primeiro lugar, vamos definir $Z_\gamma := \Phi^{-1}((1+\gamma)/2)$, onde Φ é a CDF da Normal padrão. Escrevendo $\sqrt{X} + \sqrt{X+1} =: U$, em seguida afirmamos que

$$\begin{aligned}
\gamma &\approx \Pr\left(-Z_\gamma \leq U - \sqrt{4(\beta + \kappa\mu) + 1} \leq Z_\gamma\right), \\
&= \Pr\left(-Z_\gamma - U \leq -\sqrt{4(\beta + \kappa\mu) + 1} \leq Z_\gamma - U\right), \\
&= \Pr\left(Z_\gamma - U \leq \sqrt{4(\beta + \kappa\mu) + 1} \leq Z_\gamma + U\right), \\
&= \Pr\left(\frac{[Z_\gamma - U]^2 - 1}{4} \leq (\beta + \kappa\mu) \leq \frac{[Z_\gamma + U]^2 - 1}{4}\right), \\
&= \Pr\left(\frac{[Z_\gamma - U]^2 - (1+4\beta)}{4\kappa} \leq \mu \leq \frac{[Z_\gamma + U]^2 - (1+4\beta)}{4\kappa}\right).
\end{aligned}$$

Nossa resposta final é portanto que

$$I(X) = \left(\frac{[Z_\gamma - U]^2 - (1+4\beta)}{4\kappa}, \frac{[Z_\gamma + U]^2 - (1+4\beta)}{4\kappa}\right)$$

é um intervalo bilateral aproximado de $100\gamma\%$ para μ. Para o item c) vamos lembrar que a probabilidade de rejeitar H_0 é:

$$\pi(\theta \mid \delta_c) := \Pr\left(X \geq c \mid \theta\right) = \sum_{k=c}^{\infty} \frac{e^{-\theta}\theta^k}{k!} = 1 - Q(\lfloor c+1 \rfloor, \theta).$$

Utilizando a outra dica dada no começo da prova, vemos que a derivada

$$\begin{aligned}
\frac{d\pi(\theta \mid \delta_c)}{d\theta} &= \frac{d}{d\theta}\left[1 - Q(\lfloor c+1 \rfloor, \theta)\right], \\
&= -\left(-\frac{e^{-\theta}\theta^{\lfloor c+1 \rfloor - 1}}{\Gamma(\lfloor c+1 \rfloor)}\right) = \frac{e^{-\theta}\theta^c}{\Gamma(\lfloor c+1 \rfloor)}
\end{aligned}$$

é não-negativa para todo $\theta \in (0, \infty)$, o que significa que a função poder do teste é monotônica não-decrescente em θ. Desse modo, concluímos que o teste é não-viesado: como a função poder é não-descrescente em θ temos que $\pi(\theta \mid \delta_c) \leq \pi(\theta' \mid \delta_c)$ para todo par (θ, θ') tal que $\theta \leq \theta_0$ e $\theta' > \theta_0 \geq \theta$, isto é, $\theta \in \Omega_0$ e $\theta' \in$

Ω_1. Finalmente, para responder d), está claro que queremos $\theta_0 = \beta + \kappa\mu_0$ com $\mu_0 \geq 0$. Esta formulação dá conta de um modelo com o *background*, β, e mais alguma flutuação além dele. Qualquer escolha de $\mu_0 \in [0, 1)$ está tecnicamente certa, muito embora algumas sejam mais úteis científicamente que outras. Uma possível é $\mu_0 = 1/2$ (van Dyk, 2014). Agora vamos reponder à questão d) nos aproveitando do que já foi feito no item a):

$$\begin{aligned} r_{01}(x) &= \frac{\sup_{\mu\in\Omega_0} f_n(x \mid \mu)}{\sup_{\mu\in\Omega_1} f_n(x \mid \mu)}, \\ &= \frac{f_n(x \mid \mu = 0)}{f_n(x \mid \mu = \hat{\mu}_{\text{EMV}})}, \\ &= \frac{\beta^x \exp(-\beta)}{x^x \exp(-x)}. \end{aligned}$$

Agora podemos formular um teste de razão de verossimilhanças δ_k como

$$\delta_k = \begin{cases} \text{Rejeitar } H_0 \text{ , se } r_{01}(x) < k, \\ \text{Falhar em rejeitar } H_0 \text{ , se } r_{01}(x) \geq k. \end{cases}$$

Comentários adicionais Esta questão foi baseada no artigo de van Dyk (2014) sobre a descoberta do bóson de Higgs e no trabalho de Freeman and Tukey (1950) sobre transformações estabilizadoras da variância para o caso Poisson. ■

3. Gosto muito de você, linearzinho.

O modelo linear (de regressão) é um dos cavalos de batalha da Estatística, sendo aplicado em problemas de Finanças, Medicina e Engenharia. Vamos agora estudar como utilizar as propriedades deste modelo para desenhar experimentos com garantias matemáticas de desempenho e obter estimadores de quantidades de interesse.

a) (5 pontos) Uma prática comum em regressão é a de **centrar** a variável independente (covariável), isto é subtrair a média; isto facilita a interpretação do intercepto e também simplifica alguns cálculos importantes. Mostre que no caso com a covariável centrada, $\hat{\beta}_0$ e $\hat{\beta}_1$ são independentes;

b) (7,5 pontos) Mais uma vez considerando o caso centrado, mostre como obter o número de observações n que faz com que a variância do estimador de máxima verossimilhança do intercepto seja menor que $v > 0$;

c) (5 pontos) Mostre como obter um estimador não-viesado da quantidade $\theta = a\beta_0 + b\beta_1 + c$, com $a, b, c \neq 0$, e encontre o seu erro quadrático médio.

d) (7,5 pontos) Quando $x_{\text{pred}} = \bar{x}$, mostre como obter o número de observações n necessário para que o intervalo de predição de $100(1 - \alpha_0)\%$ para a variável-resposta (Y) tenha largura menor ou igual a $l > 0$ com probabilidade pelo menos γ.

   Dicas:(i) A expressão dependerá *também* da variância dos resíduos, σ^2 e (ii) Você não precisa calcular n, apenas mostrar o procedimento para obtê-lo.

Conceitos trabalhados: Regressão linear; desenho experimental; quantidades derivadas.
Nível de dificuldade: médio.
Resolução: Para resolver a) vamos precisar apenas olhar para as dicas e perceber que quando substituímos a covariável original X por $X' = X - \bar{x}$ temos $\bar{x}' = 0$ e portanto $\operatorname{Cov}\left(\hat{\beta}_0, \hat{\beta}_1\right) = -\frac{\bar{x}'\sigma^2}{s_x^2} = 0$. Para afirmarmos que $\hat{\beta}_0$ e $\hat{\beta}_1$ são independentes é preciso lembrar que estes estimadores têm distribuição conjunta Normal bivariada; quando a covariância é zero, sabemos que são independentes. A resposta de b) vem mais uma vez utilizando a dica dada. Vemos que no caso centrado a variância de $\hat{\beta}_0$ é σ^2/n. Desta forma, precisamos apenas encontrar n tal que $\sigma^2/n < v$, isto é $n > \sigma^2/v$. Como sabemos que os estimadores dos coeficientes são não-viesados (trabalhado em aula, presente nas dicas), podemos encontrar $\hat{\theta} = a\hat{\beta}_0 + b\hat{\beta}_1 + c$ como nosso estimador não-viesado de θ. O EQM de tal estimador é a sua variância:

$$
\begin{aligned}
E[(\hat{\theta} - \theta)^2] = \operatorname{Var}(\hat{\theta}) &= a^2 \operatorname{Var}(\hat{\beta}_0) + b^2 \operatorname{Var}(\hat{\beta}_1) - ab \operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1), \\
&= a^2\sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{s_x^2}\right) + b^2\frac{\sigma^2}{s_x^2} + ab\frac{\bar{x}\sigma^2}{s_x^2}, \\
&= \sigma^2 \left(\frac{a^2}{n} + \frac{a^2\bar{x}^2}{s_x^2} + \frac{b^2}{s_x^2} + \frac{ab\bar{x}}{s_x^2}\right).
\end{aligned}
$$

Por fim, vamos responder d). Note que a expressão necessária aqui é a do intervalo de predição:

$$
\hat{Y} \pm c(n, \alpha_0) \cdot \hat{\sigma}'_r \cdot \sqrt{\left[1 + \frac{1}{n} + \frac{(x_{\text{pred}} - \bar{x})^2}{s_x^2}\right]},
$$

onde

$$
c(n, \alpha_0) := T^{-1}\left(1 - \frac{\alpha_0}{2}; n - 2\right),
$$

e

$$
\hat{\sigma}'_r := \sqrt{\frac{\sum_{i=1}^n \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2}{n-2}}.
$$

Quando $x_{\text{pred}} = \bar{x}$ a expressão se reduz um pouco e podemos deduzir que a largura do intervalo é

$$
\hat{l} = 2 \cdot c(n, \alpha_0) \cdot \hat{\sigma}'_r \sqrt{\left[1 + \frac{1}{n}\right]}.
$$

Desejamos, portanto, encontrar n tal que

$$
\begin{aligned}
\Pr\left(\hat{l} < l\right) &\geq \gamma, \\
\Pr\left(\hat{\sigma}'_r < \frac{l}{2 \cdot c(n, \alpha_0) \cdot \sqrt{\left[1 + \frac{1}{n}\right]}}\right) &\geq \gamma,
\end{aligned}
$$

isto é conseguimos reduzir nossa afirmação probabilística a uma afirmação com respeito à f.d.a. (ou CDF) de $\hat{\sigma}'_r$. Para completar nossos cálculos só precisamos nos lembrar que $n\hat{\sigma}'_r/\sigma^2$ tem distribuição qui-quadrado com $n-2$ graus de liberdade (De Groot, Teorema 11.3.2) e, portanto,

$$\Pr\left(\hat{\sigma}'_r \leq a\right) = F_\chi\left(\frac{\sigma^2}{n}a; n-2\right).$$

Fórmulas úteis

Como usar este catálogo: as fórmulas dadas aqui estão propositalmente privadas do seu contexto. O objetivo desta coleção é ajudar você a lembrar das expressões. Entretanto, saber quais expressões são utilizadas em que contexto é sua tarefa.

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$;
- $\hat{\sigma}' = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2}$;
- $S_X^2 = \sum_{i=1}^{m}(X_i - \bar{X}_m)^2$;
- $S_Y^2 = \sum_{j=1}^{n}(Y_j - \bar{Y}_n)^2$;
- $U = \frac{\sqrt{m+n-2}(\bar{X}_m - \bar{Y}_n)}{\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)(S_X^2+S_Y^2)}}$;
- $V = \frac{S_X^2/(m-1)}{S_Y^2/(n-1)}$;
- $\bar{x} = (1/n)\sum_{i=1}^{n} X_i$;
- $\bar{y} = (1/n)\sum_{i=1}^{n} Y_i$;
- $$E\left[\left(\hat{Y} - Y\right)^2\right] = \sigma^2\left(1 + \frac{1}{n} + \frac{(x_{\text{pred}} - \bar{x})^2}{s_x^2}\right).$$
- $$\hat{\sigma}'_r := \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2}{n-2}}.$$
- $$\hat{\beta}_0 \pm \hat{\sigma}' c\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{s_x^2}} \quad \text{e} \quad \hat{\beta}_1 \pm c\frac{\hat{\sigma}'}{s_x},$$
  $$\hat{\beta}_0 + \hat{\beta}_1 x_{\text{pred}} \pm c\hat{\sigma}'\sqrt{\frac{1}{n} + \frac{(x_{\text{pred}} - \bar{x})^2}{s_x^2}},$$
- $$\hat{Y} \pm c\hat{\sigma}'_r\sqrt{\left[1 + \frac{1}{n} + \frac{(x_{\text{pred}} - \bar{x})^2}{s_x^2}\right]},$$

  onde $c = T^{-1}(1 - \frac{\alpha_0}{2}; n-2)$.

Bibliografia

Freeman, M. F. and Tukey, J. W. (1950). Transformations related to the angular and the square root. *The Annals of Mathematical Statistics*, pages 607–611.

van Dyk, D. A. (2014). The role of statistics in the discovery of a Higgs boson. *Annual Review of Statistics and Its Application*, 1:41–59.