

AValiação DOS MÉTODOS: *PROBIT*, *PROBIT ISOTONIZADO* E *UP AND DOWN* EM DADOS DE SENSIBILIDADE

Edcarlos Miranda de Souza¹, Lucas Monteiro Chaves², Joel Augusto Muniz³

1 INTRODUÇÃO

Em análise de dados de sensibilidade, particularmente em ensaios de dose-resposta, o método mais utilizado é o *probit*, para os casos em que a susceptibilidade de cada indivíduo é considerada como uma variável aleatória com distribuição normal ou *logit* nos casos do uso da função logística. O *probit* foi Proposto por Bliss (1935) e Fisher (1935) e o *logit* por Berkson (1944). Um método de análise alternativa foi proposto por Dixon & Mood (1948) denominado Método *Up and Down* e se utiliza da teoria de passeios aleatórios. Este método alternativo tem a vantagem de ser mais adequado para o caso de amostras pequenas, o que é essencial em ensaios de dose-resposta por razões práticas e éticas. Todos estes procedimentos estimam com eficiência a dose letal de um indivíduo a 50% (DL_{50}). Entretanto, há indícios que para proporções próximas das caldas das distribuições, tais como (DL_{90} , DL_{95} e DL_{99}) estes métodos se tornam menos eficientes. Neste trabalho é proposto o método *probit isotonizado*, que consiste em calcular a regressão isotônica das proporções observadas antes da aplicação do *probit*. Um estudo comparativo dos três métodos é desenvolvido para dados simulados.

2 REFERENCIAL TEÓRICO

2.1 O Método *Probit*

O *probit* é um método para estimar dose críticas em ensaios de dose-resposta, discutido em detalhes em Bliss (1935). Um ensaio do tipo dose-resposta é aquele onde uma determinada droga é administrada em k diferentes doses (níveis), d_1, d_2, \dots, d_k em, respectivamente, m_1, m_2, \dots, m_k indivíduos, obtendo-se como resposta, após um período especificado, y_1, y_2, \dots, y_k indivíduos que mudam de estado (ocorrência de um sucesso, por exemplo, morte).

Neste tipo de ensaio uma amostra com indivíduos de uma mesma espécie é selecionada. Tal

¹Doutorando em Estatística e Experimentação Agropecuária, Bolsista CAPES, Universidade Federal do Acre, profedcarlos@hotmail.com

²Universidade Federal de Lavras, lucas@dex.ufla.br

³Universidade Federal de Lavras, joamuniz@dex.ufla.br

amostra é dividida em k grupos, cada um com m_i indivíduos, $i = 1, 2, \dots, k$. Cada dose é aplicada a cada grupo e o número de sucessos por grupo (morte do indivíduo, por exemplo) é contado. Dessa maneira, a resposta para cada grupo é uma variável aleatória com distribuição binomial com parâmetros m_i e π_i , ou seja, $Y_i \sim Bin(m_i, \pi_i)$.

Uma informação que é de interesse prático em experimentos de dose-resposta, é a medida de *Susceptibilidade* ou *Tolerância* de uma unidade experimental ao tratamento a que ela foi submetida. Esta tolerância pode ser entendida como o valor crítico que o tratamento assume no limiar entre sucesso ou fracasso do ensaio, por exemplo, nos experimentos de dose-resposta, ela pode ser entendida como a menor dose possível para matar um determinado indivíduo, ou seja, a dose letal (*lethal dose - DL*) do indivíduo. Esta Tolerância é uma variável aleatória contínua que não pode ser medida diretamente. O que pesquisador tem em mãos são as proporções de sucessos (ou fracassos) para cada grupo do qual foi submetido o tratamento.

No método *probit*, tal variável será considerada no processo de modelagem como uma variável aleatória com distribuição normal com média μ e variância σ^2 , ou seja, se U é a variável aleatória que representa a Tolerância de cada indivíduo da população, tem-se que $U \sim N(\mu, \sigma^2)$. Se a dose d_i é aplicada à população toda e $f_U(u)$ é a função de densidade de probabilidade para a distribuição de Tolerâncias, todo indivíduo cuja Tolerância é menor do que d_i responderá à droga (tratamento), e a probabilidade de que um indivíduo escolhido ao acaso responda à dose é dada por:

$$\pi_i = P[U \leq d_i] = F(d_i) = \int_{-\infty}^{d_i} f_U(u) du. \quad (1)$$

Assim,

$$\begin{aligned} \pi_i &= P[U \leq d_i] = P\left[\frac{(U - \mu)}{\sigma} \leq \frac{(d_i - \mu)}{\sigma}\right] \\ &= P\left[Z \leq -\frac{\mu}{\sigma} + \frac{1}{\sigma}d_i\right] = P[Z \leq \beta_1 + \beta_2 d_i], \end{aligned}$$

com $\beta_1 = -\frac{\mu}{\sigma}$ e $\beta_2 = \frac{1}{\sigma}$. Considerando $\Phi(x)$ a distribuição acumulada da normal padrão no ponto x , tem-se:

$$\pi_i = \Phi(\beta_1 + \beta_2 d_i) \Rightarrow \Phi^{-1}(\pi_i) = \beta_1 + \beta_2 d_i. \quad (2)$$

O *probit* de π_i é definido como sendo : $probit(\pi_i) = \Phi^{-1}(\pi_i) = \beta_1 + \beta_2 d_i$.

Encontrando-se estimativas para β_1 e β_2 , pode-se estimar os valores de μ e σ . Para a estimação

de dose letais elevadas ou baixas o método sofre de um problema de estabilidade em razão da calda da distribuição. Fisher (1935), sugere um método para contornar este problema: Considerando p_i a proporção de sucessos para a dose i , toma-se todos os pares (d_i, p_i) onde $p_i \neq 0$ e $p_i \neq 1$ e faz-se uma primeira estimativa de β_1 e β_2 por Mínimos Quadrados Ponderados com pesos $w_i = \frac{z_i^2 m_i}{p_i q_i}$. Encontradas estas estimativas, considera-se o *probit* para a dose d_j em que ocorreu $p_j = 0$ por:

$$probit^*(p_j) = \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right) - \frac{\Phi(d_j)}{\phi(d_j)},$$

e para o caso onde ocorreu $p_j = 1$:

$$probit^*(p_j) = \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right) + \frac{1 - \Phi(d_j)}{\phi(d_j)},$$

De posse destes valores, considera-se todos os valores iniciais juntamente com os anteriores e a regressão linear é refeita, calculando-se novas estimativas para β_1 e β_2 , que serão as estimativas finais do processo.

2.2 O Método *Up and Down*

Considere que uma dose inicial d_i é aplicada em um primeiro paciente e defina o seguinte passeio aleatório: se o paciente responde positivamente, para o próximo paciente a dose é diminuída, se a resposta é negativa, a dose é aumentada. Para a análise dos dados é considerado que a probabilidade p_i de um paciente responder positivamente é:

$$\pi_i = \Phi(d_i) = \int_{-\infty}^{d_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt = 1 - \pi_i. \quad (3)$$

Com a simplificação de que o número de fracasso para a dose d_{i-1} é igual ao número de sucessos para dose d_i , é possível obter um sistema tratável de equações para o cálculo dos Estimadores de Máxima Verossimilhança de μ e σ . A característica fundamental deste delineamento é que o passeio rapidamente se concentra em doses próximas da dose média μ , de onde se espera que mesmo para amostras pequenas os estimadores sejam eficientes.

Os Estimadores para μ e σ propostos por Dixon & Mood (1948) são:

$$\hat{\mu} = d' + d \left(\frac{A}{Q} \pm \frac{1}{2} \right), \quad (4)$$

em que $A = \sum in_i$, $Q = \min\{M, N\}$. A análise é feita somente com sucessos ou somente com

fracassos. Ela será feita com sucessos se $Q = N$ e com fracassos se $Q = M$, assim, n_i representa ou a frequência com que ocorreu um sucesso na dose d_i ou a frequência com que ocorreu um fracasso na dose d_i . Além disso, só são considerados na análise, doses que foram visitadas pelo passeio. Na fórmula 4, d' é a menor dose visitada com menor frequência, e o fator 0,5 segue de correção de continuidade, sendo que o sinal é “+” se $Q = M$, ou “-” se $Q = N$. O estimador para o desvio padrão é:

$$s = 1,620d \left(\frac{QB - A^2}{Q^2} + 0,029 \right), \quad (5)$$

em que $B = \sum i^2 n_i$.

2.3 Regressão Isotônica

Um problema que surge com o uso da proporção de sucessos observados $p_i = \frac{y_i}{m_i}$ para estimar π_i é que as estimativas podem não ser crescentes com a dosagem. Nestes casos podem ser usados conceitos de Estimção Isotônica. Uma função f é isotônica nas doses se $f(d_1) \leq f(d_2) \leq \dots \leq f(d_n)$ sempre que $d_1 < d_2 < \dots < d_n$. Uma função isotônica p_i^* sobre um conjunto de doses $D = \{d_1, d_2, \dots, d_n\}$ é a regressão isotônica de $p_i = \frac{y_i}{m_i}$ com pesos w_i se ela minimiza

$$\sum_{i=1}^n (p_i - f_i)^2 w_i \quad (6)$$

na classe de todas as funções isotônicas f sobre D . Barlow et al. (1972) apresentam um algoritmo para o cálculo de p_i^* e demonstram que este é o Estimador de Máxima Verossimilhança com restrição de ordem nas doses, para algumas distribuições da família exponencial.

2.4 Os Estimadores

Seja DL_Γ a dose letal que mata $100\Gamma\%$ da população. Considere os estimadores para DL_Γ :

Estimador Probit - Consiste basicamente no método *probit*:

$$\widehat{DL}_\Gamma^1 = \frac{\Phi^{-1}(\Gamma) - \hat{\beta}_1}{\hat{\beta}_2}, \quad (7)$$

em que $\hat{\beta}_1$ e $\hat{\beta}_2$ são as Estimativas de Mínimos Quadrados Ponderados propostas por Bliss & Fisher (1935).

Estimador Isotonizado - Este estimador se utiliza da mesma metodologia do *probit* usual, só que

no lugar das proporções observadas, considera-se suas estimativas obtidas por regressão isotônica:

$$\widehat{DL}_\Gamma^2 = \frac{\Phi^{-1}(\Gamma) - \hat{\beta}_1}{\hat{\beta}_2}, \quad (8)$$

em que $\hat{\beta}_1$ e $\hat{\beta}_2$ são encontrados sobre a regressão isotônica P_i^* de P_i .

Estimador Up and Down - Considerando $\hat{\mu}$ e s os estimadores apresentado pelo método *Up and Down*, descritos nas Equações 4 e 5, pode-se considerar a Tolerância de cada indivíduo por $U \sim N(\hat{\mu}, s^2)$, e assim, obter o estimador para dose letal como:

$$\widehat{DL}_\Gamma^3 = P[U \leq \Gamma] = \Phi_{\hat{\mu}, s^2}(\Gamma) \quad (9)$$

3 MATERIAL E MÉTODOS

Foi realizado uma simulação considerando que a verdadeira Tolerância dos indivíduos é normal com média igual a 10 e desvio padrão igual a 1,5. O conjunto de doses considerado foi $\{7, 0; 7, 5; \dots; 13\}$ para os três estimadores. Para os estimadores *probit* e *probit isotonizado* foram considerados 5 ensaios em cada dose, totalizando uma amostra com 65 elementos. Para o estimador *Up and Down* foi simulado um passeio aleatório com 40 passos nestas doses, ou seja, uma amostra com 40 elementos.

As doses letais consideradas foram: DL_{50} , DL_{90} , DL_{95} e DL_{99} . Para estimar estas doses foram realizadas 1000 simulações, gerando assim 1000 estimativas para cada uma delas com cada um dos estimadores propostos. Na análise foi considerado a média e o desvio padrão destas estimativas para efeito comparativo das metodologias.

Todas as rotinas foram realizadas em linguagem *R* e efetuadas no software estatístico R Development Core Team versão 2.10.1.

4 RESULTADOS E DISCUSSÃO

A tabela 1 apresenta os resultados para as quatro doses letais de interesse. Pode-se observar que os quatro estimadores funcionam bem para estimação da dose mediana, ou seja, os quatro apresentaram resultados praticamente idênticos para a estimação da dose letal a 50%. Entretanto, para as doses altas, os mesmos apresentaram desempenho diferentes.

O Estimador *Up and Down* e *probit isotonizado* apresentaram desvio médio absoluto em relação ao verdadeiro valor do parâmetro muito semelhantes, entretanto o *Up and Down* foi o que mostrou

maior variabilidade das estimativas para as doses elevadas, sendo esta variabilidade quase o dobro daquela apresentada pelo método do *probit isotonizado*.

TABELA 1: Doses letais, Médias e Erro Padrão das 1000 estimativas obtidas por simulação para cada um dos estimadores. Os valores entre parênteses são os Erros Padrão das Estimativas.

Doses Letais	Parâmetros	\widehat{DL}_Γ^1	\widehat{DL}_Γ^2	\widehat{DL}_Γ^3
DL_{50}	10,00	10,01 (0,28)	10,0 (0,28)	10,10 (0,32)
DL_{90}	11,92	12,23 (0,47)	11,81 (0,44)	11,86 (0,85)
DL_{95}	12,47	12,86 (0,56)	12,32 (0,52)	12,36 (1,06)
DL_{99}	13,49	14,04 (0,75)	13,28 (0,68)	13,29 (1,47)

O Estimador *probit* foi o que apresentou maior desvio absoluto (valor do parâmetro menos estimativa), sendo que este aumenta a medida que a dose também aumenta, e apresentou Erro Padrão menor do que do que o *Up and Down*.

5 CONCLUSÃO

Dentre os Estimadores apresentados, o Estimador *probit isotonizado* foi aquele que apresentou melhor desempenho para a estimação das doses letais altas, sendo que o mesmo apresentou desempenho semelhante aos demais para a estimação da dose mediana.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARLOW, R. E.; BARTHOLOMEW, D. J.; BREMNER, J.M.; BRUNK, H.D. **Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression**. New York: John Wiley & Sons, 1972. 388 p.
- BERKSON, J. Application to the logistic function to bio-assay. **Journal of the American Statistical Association**, v.39, n.227, p. 357-365, Sep. 1944.
- BLISS, C.I. The calculator of the dosage-mortality curve. **Annals of Applied Biology**, v.22, p. 134-167, 1935.
- DIXON, W. J.; MOOD, A. M. A method for obtaining and analyzing sensitivity data. **Journal of the America Statistical Association**, v.43, n. 241, p. 109-126, Mar. 1948.
- FISHER, R. A. Appendix to bliss (1935): The case of zero survivors. **Annals of Applied Biology**, v.22, p. 164-165, 1935.