

UM ESTUDO SOBRE ASPECTOS INFERENCIAIS BAYESIANOS DA DISTRIBUIÇÃO LOGNORMAL SOB UMA NOVA PARAMETRIZAÇÃO



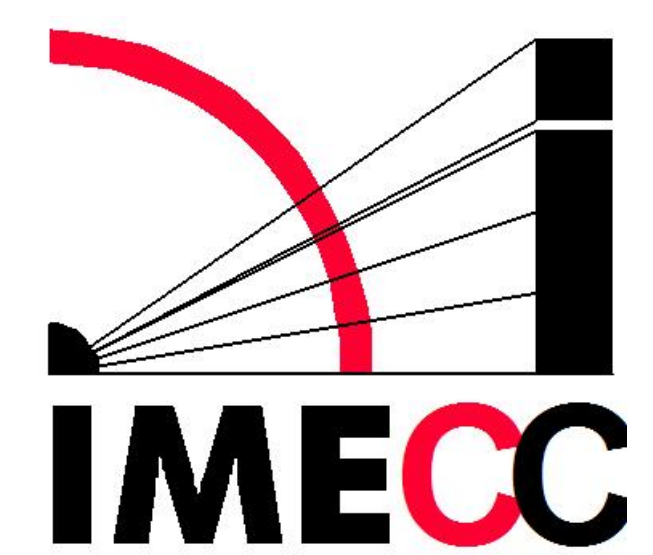
Lígia Silveira Schweller - ligiasilveira@gmail.com

Orientador: Prof. Dr. Caio Lucidius Naberezny Azevedo

DEPARTAMENTO DE ESTATÍSTICA, IMECC - UNICAMP

Agência Financiadora: PIBIC/CNPq

Palavras-Chave: Distribuição Lognormal - Inferência Bayesiana - Priors



Introdução

Em síntese, a distribuição lognormal (DLN) pode ser definida como a distribuição de uma variável aleatória cujo logaritmo natural (\ln) é normalmente distribuído. Por assumir valores positivos e ter comportamento assimétrico, é muito utilizada para modelar dados como renda, peso, pressão arterial, etc ([1]).

A distribuição lognormal é, em geral, apresentada na sua Parametrização Tradicional (PT), em função de parâmetros que não possuem interpretação intuitiva. No presente trabalho, foi considerada uma parametrização em termos da média e variância da referida distribuição. Esta última foi chamada de Parametrização Direta (PD).

Foi estudado o impacto de diferentes priors no processo de estimação. As estimativas pontuais Bayesianas foram comparadas com as frequentistas obtidas através de simulações, para a PT e PD. Por fim, foi feita uma análise de um conjunto de dados reais da área odontológica.

Metodologia

Inicialmente, foram encontradas as distribuições a posteriori a partir de priors informativas e não informativas (priors conjugadas, priors de Jeffreys e priors de Jeffreys sob independência). As posteriores foram encontradas considerando-se o caso em que cada um dos parâmetros é conhecido, e também quando ambos são desconhecidos. Na parte inferencial, considerou-se a situação em que se tem uma amostra aleatória da distribuição lognormal. Foi desenvolvido o processo de estimação pontual, calculando-se esperança, moda e mediana a posteriori. Em relação à estimação intervalar, foram obtidos os intervalos de credibilidade ([2]).

Resultados

Ao reparametrizar a distribuição lognormal, os parâmetros que, na PT, representavam localização e dispersão, expressam na PD o valor esperado e uma medida de dispersão associada à variância. Com tal reparametrização, pode-se interpretar os resultados mais facilmente, e compará-los com os obtidos para a PT.

Inicialmente, foram obtidas as distribuições a posteriori a partir das priors conjugadas, priors de Jeffreys e priors de Jeffreys sob independência, para as duas parametrizações. No caso da PD, a família conjugada existe apenas para o caso em que o parâmetro de dispersão é conhecido. Na PT, as posteriores foram obtidas analiticamente. Já na PD, foram-na de modo numérico.

Em relação aos aspectos inferenciais da distribuição lognormal, sob o ponto de vista Bayesiano, iniciou-se com o processo de estimação dos parâmetros. Foram obtidas a esperança, moda e mediana a posteriori, de acordo com as posteriores obtidas em cada um dos casos. Após a estimação de cada parâmetro, foram estimadas a média e a variância da DLN. A construção dos intervalos de credibilidade de modo analítico só foi possível para as priors conjugadas da PT. Nos demais casos, os intervalos não foram obtidos, pois são necessárias simulações e procedimentos numéricos.

Com o estudo de simulação, pôde-se concluir que a parametrização adotada não tem muito efeito sobre o valor das estimativas, em ambos os casos os resultados são bons. No entanto, as priors de Jeffreys levam a estimativas com menor vício, menor variância e, conseqüentemente, menor erro quadrático médio. As estimativas Bayesianas obtidas foram muito próximas das frequentistas.

Análise de dados reais

Os dados consistem em 20 grupos com diferentes formulações de determinada resina dentária, divididos de acordo com o nível de enchimento do vidro de bário (carga) e a proporção molar (matriz). Em cada um desses grupos, foi medida a tensão que a resina suporta sob as condições a que está submetida, em 7 unidades experimentais. O objetivo é comparar a média da tensão entre os grupos, e verificar como os fatores (carga e matriz) afetam o resultado da tensão ([3]).

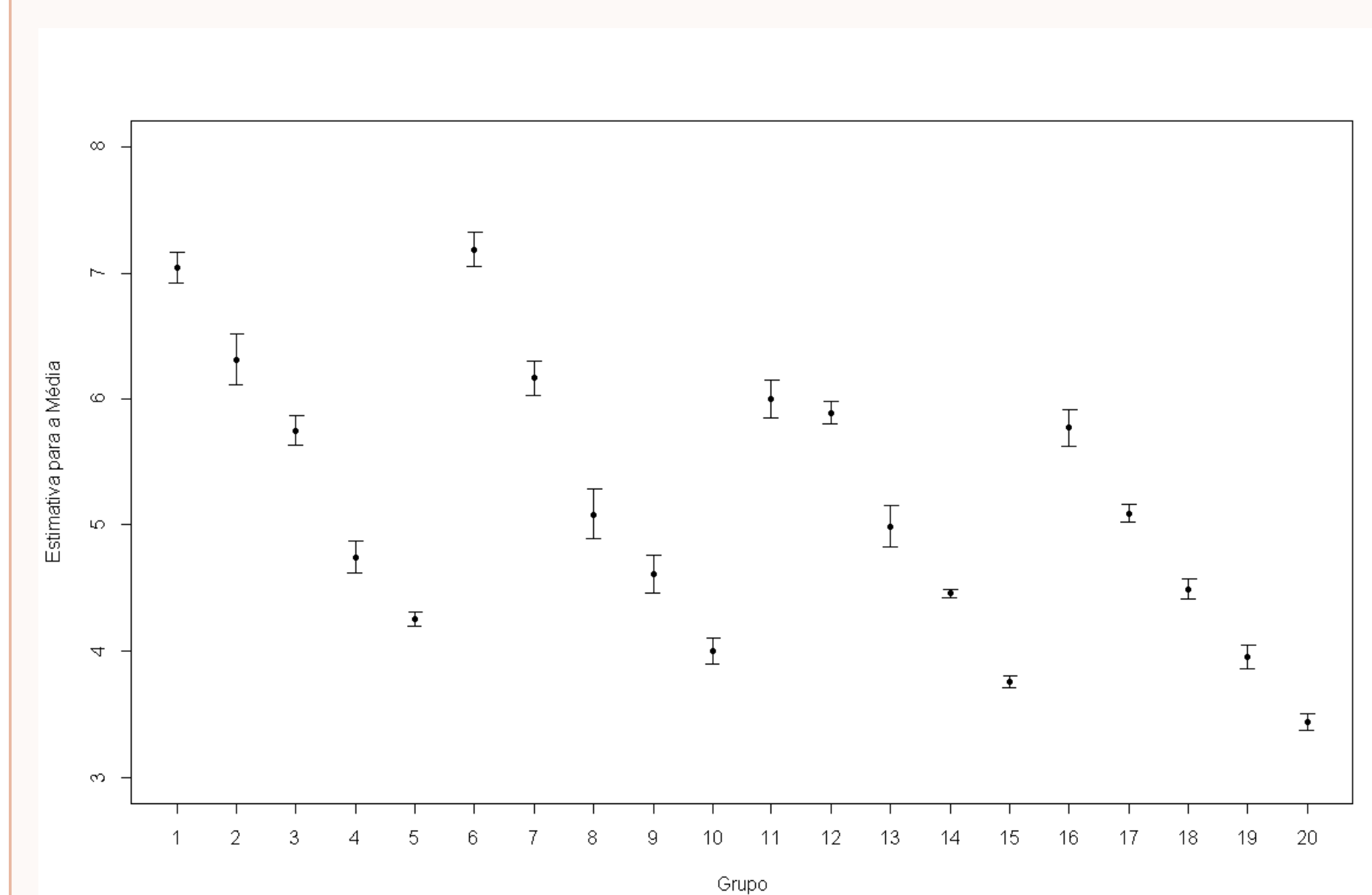


Figura 1: Estimativas pontuais e $IC_B(95\%)$ para a tensão em cada grupo.

Referências

- [1] Crow, E. L. and Shimizu, K. *Lognormal distributions: theory and applications*. Marcel Dekker, Inc, New York, NY, first edition, 1988.
- [2] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. *Bayesian Data Analysis*. Chapman & Hall/CRC, Texts in Statistical Science Series, second edition, 2004.
- [3] Gonçalves, F., Azevedo, C. L. N., Ferracane, J. L., Braga, R. R. *BisGMA/TEGDMA ratio and filler content effects on shrinkage stress*. Dental Materials, Volume 27, 2011.