

UM ESTUDO SOBRE ASPECTOS PROBABILÍSTICOS E INFERENCIAIS FREQUENTISTAS DA DISTRIBUIÇÃO LOG-NORMAL SOB UMA NOVA PARAMETRIZAÇÃO



Lígia Silveira Schweller - ligiasilveiraa@gmail.com
Orientador: Prof. Dr. Caio Lucidius Naberezny Azevedo
DEPARTAMENTO DE ESTATÍSTICA, IMECC - UNICAMP
Agência Financiadora: PIBIC/CNPq



Palavras-Chave: Distribuição lognormal - Estimação pontual - Estimação intervalar

Introdução

Em síntese, a distribuição lognormal pode ser definida como a distribuição de uma variável aleatória cujo logaritmo natural é normalmente distribuído. Por assumir valores positivos e ter comportamento assimétrico, a distribuição lognormal é muito utilizada para modelar dados positivos que apresentam assimetria. Tal distribuição é, em geral, apresentada na sua Parametrização Tradicional (PT). No presente trabalho, consideramos uma nova parametrização, na qual os parâmetros possuem interpretações simples em termos do valor esperado e da variância da referida distribuição. Nomeamos esta última de Parametrização Direta (PD). Apresentamos os principais aspectos probabilísticos, analítica e graficamente, e inferenciais frequentistas desta distribuição, na sua parametrização tradicional e comparamos com os resultados equivalentes obtidos a partir da parametrização direta.

Metodologia

Geramos diversos gráficos, a partir de diferentes combinações de valores dos parâmetros (originais e diretos), para verificar os comportamentos das densidades e das funções de distribuição acumulada geradas e, com isso, avaliar como tais parâmetros afetam o comportamento dessas funções. Na parte inferencial, consideramos a situação em que temos uma amostra aleatória da distribuição lognormal, sob cada uma das parametrizações. Desenvolvemos os métodos de estimação por MV e MM. Os intervalos de confiança foram construídos com base em quantidades pivotais, bem como na distribuição assintótica dos estimadores de MV.

Resultados

Para as duas parametrizações, encontramos as funções densidade e distribuição acumulada. Em ambas as parametrizações, o comportamento que observamos foi equivalente, como esperado. Ilustramos a comparação inferencial (estimação pontual por MM e MV) das duas parametrizações com um estudo de simulação. Concluímos que, para estimar $E(X)$, é indiferente a parametrização adotada, mas o método de MV fornece estimativas melhores, embora tal diferença não seja muito grande. Além disso, para estimar $Var(X)$, a PD se mostrou mais precisa. Neste caso, também foi melhor a estimação por MV. Como era de se esperar, maiores tamanhos de amostra levam a melhores estimativas, e as propriedades assintóticas são verificadas.

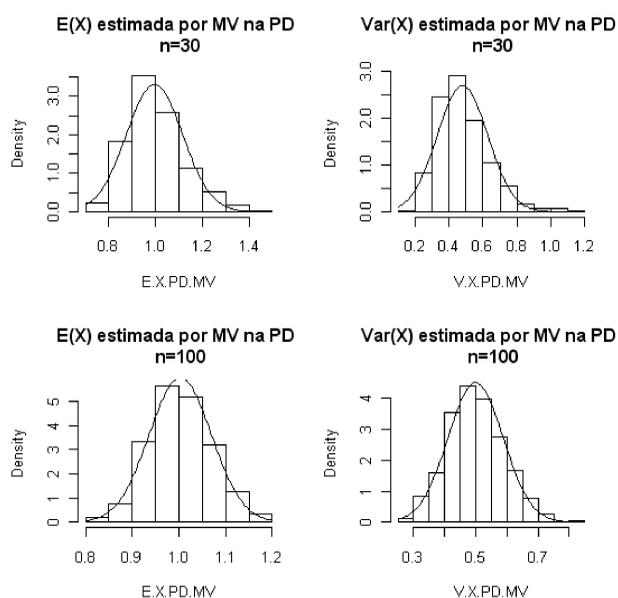


Figura 1: Distribuição de $\widehat{E(X)}_{PD,MV}$ e $\widehat{Var(X)}_{PD,MV}$ para $n = 30$ e $n = 100$.

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. Nosso objetivo é comparar a média da tensão entre os grupos, e verificar como os fatores (carga e matriz) afetam o resultado da tensão.

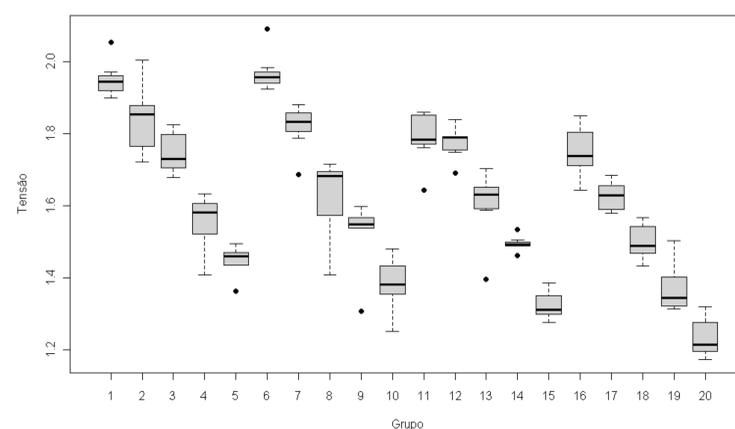


Figura 2: Boxplot da tensão do material por grupo.

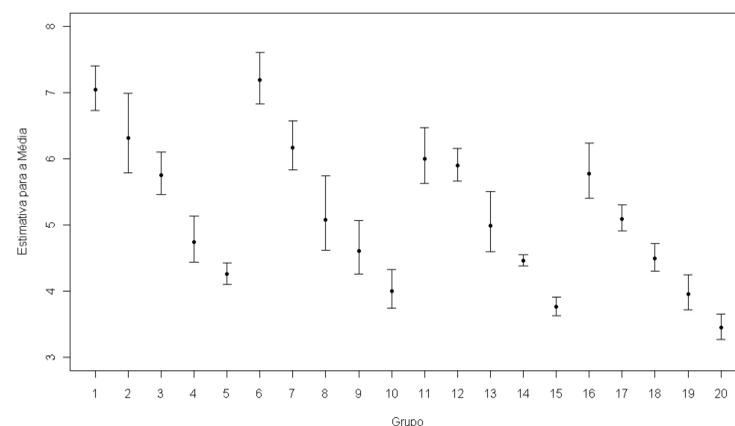


Figura 3: Estimativas pontuais e $IC(95\%)$ para a tensão em cada grupo.