Explanation of the relationship between lognormal and normal models for data

A model for the mean $\mathbf{x}_i'\boldsymbol{\beta}$ of lognormally distributed data y_i implies that

$$y_i = \mathbf{x}_i'\beta e^{\epsilon_i},\ \epsilon_i \sim \text{normal}(0, \sigma^2). \tag{1}$$

Taking the log of both sides, we have

$$\log(y_i) = \log(\mathbf{x}_i'\boldsymbol{\beta}) + \epsilon_i,\ \epsilon_i \sim \text{normal}(0, \sigma^2), \tag{2}$$

which is the same as

$$\log(y_i) \sim \text{normal}(\log(\mathbf{x}_i'\boldsymbol{\beta}), \sigma^2). \tag{3}$$

If the log of a random variable is normally distributed with mean $\log(\mathbf{x}_i'\boldsymbol{\beta})$, and variance σ^2 then the random variable is lognormally distributed with parameters (not moments) $\log(\mathbf{x}_i'\boldsymbol{\beta})$ and σ^2:

$$y_i \sim \text{lognormal}(\underbrace{\log(\mathbf{x}_i'\boldsymbol{\beta})}_{\text{centrality parameter}}, \underbrace{\sigma^2}_{\text{scale parameter}}). \tag{4}$$

The centrality parameter of a lognormal distribution is defined as the mean of the random variable on the log scale and the scale parameter is defined as the variance of the random variable on the log scale as shown in equation 3. **You need to be very careful with coding if you use eq. 3. If this is part of a hierarchical model, you must exponentiate the left hand side if you want the variable to be on the lognormal (i.e., exponentiated normal) scale.**

We can substitute any $g(\boldsymbol{\theta}, \mathbf{x}_i)$ for $\mathbf{x}_i'\boldsymbol{\beta}$ as our model of the mean (or median) and the logic remains the same, but we need to careful about the interpretation of the coefficients. We often see the mean model written as $\exp(\mathbf{x}_i'\boldsymbol{\beta})$ to assure that it produces positive values, however, in this case the slopes specify the *multiplicative* change from the mean (e.g., $e^{\beta_0}e^{\beta_1 x_{1,i}}$) in y per unit change in x. Constraining $\mathbf{x}_i'\boldsymbol{\beta}$ to be positive rather than exponentiating it allows the interpretation of the slopes to be the same as for any linear model, which is the *additive* change in the mean (e.g, $\beta_0 + \beta_1 x_{1,i}$) per unit change in x.