Model

The **inputs** are the real, continuous variable z_x, which standardises the increasing attempt numbers, x, of each individual and the dichotomous variable K_j which indicates whether the individual is a kid:

$$z_x = \frac{x - \bar{x}}{\hat{\sigma}_x}, \qquad K_j = \begin{cases} 1 & \text{if } j \text{ is a kid} \\ 0 & \text{else} \end{cases} \tag{1}$$

where $\bar{x}$ is the sample mean and $\hat{\sigma}_x$ the sample standard deviation of x.

The **outputs** are the reaction times, y_j, for each individual j, given the inputs. For our model, we use the $\log(y_j)$ and work with standardised values, $z_{\log(y_j)}$:

$$z_{\log(y_j)} = \frac{\log(y_j) - \overline{\log(y_j)}}{\hat{\sigma}_{\log(y_j)}} \tag{2}$$

The applied Bayesian **hierarchical model** expresses the probability of the output given the input and 9 parameters by redistributing the probability proportions as follows:

$$\begin{aligned}
&P(z_x, K_j, \theta_{0_j}, \theta_{1_j}, \sigma, \mu_0, \varphi_0, \mu_1, \varphi_1, \tau_0, \tau_1 | z_{\log(y_j)}) \\
&\quad \propto P(z_{\log(y_j)} | z_x, K_j, \theta_{0_j}, \theta_{1_j}, \sigma, \mu_0, \varphi_0, \mu_1, \varphi_1, \tau_0, \tau_1) \cdot P(\theta_{0_j}, \theta_{1_j}, \sigma, \mu_0, \varphi_0, \mu_1, \varphi_1, \tau_0, \tau_1) \\
&\quad \propto P(z_{\log(y_j)} | z_x, \theta_{0_j}, \theta_{1_j}, \sigma) \cdot P(\theta_{0_j} | \mu_0, \varphi_0, \tau_0, K_j) \cdot P(\theta_{1_j} | \mu_1, \varphi_1, \tau_1, K_j) \cdot \\
&\qquad P(\sigma) \cdot P(\mu_0) \cdot P(\mu_1) \cdot P(\varphi_0) \cdot P(\varphi_1) \cdot P(\tau_0) \cdot P(\tau_1)
\end{aligned}$$

K_j and z_x are the outputs and have no priors. The joint posterior over all individuals can then be expressed as:

$$\prod_j^J P(z_x, K_j, \theta_{0_j}, \theta_{1_j}, \cdots | z_{\log(y_j)}) \propto \prod_j^J P(z_{\log(y_j)} | z_x, \theta_{0_j}, \theta_{1_j}, \sigma) \cdot P(\theta_{0_j} | \mu_0, \varphi_0, \tau_0, K_j) \cdot P(\theta_{1_j} | \mu_1, \varphi_1, \tau_1, K_j) \cdots$$

A graphical representation of the hierarchical model can be seen in Figure 2.

The **core** of the model is the distribution of the data (i.e. the likelihood function) which models a regression line for each individual with a certain noise around it:

$$\boxed{z_{\log(y_j)} \sim \mathcal{N}(\theta_{0_j} + \theta_{1_j} \cdot z_x,\ \sigma^2)}$$

The assumed shapes of the intermediate priors and pure (hyper-) priors are as follows:

$$\begin{aligned}
\theta_{0_j} &\sim \mathcal{N}(\mu_0 + \varphi_0 \cdot K_j, \tau_0^2) \\
\theta_{1_j} &\sim \mathcal{N}(\mu_1 + \varphi_1 \cdot K_j, \tau_1^2) \\
\mu_0, \mu_1, \varphi_0, \varphi_1 &\sim \mathcal{U}(-\infty, +\infty) \\
\sigma, \tau_0, \tau_1 &\sim \mathcal{U}\ ]0, +\infty)
\end{aligned}$$

Note, that the priors in the first line above exclude 0.

Transformations

With the random variables $\Xi, \Theta, \Omega \sim \mathcal{N}(0,1)$, we can rewrite the distributions as follows:

$$\begin{aligned} z_{\log(y_j)} &= \theta_{0_j} + \theta_{1_j} \cdot z_x + \sigma \cdot \Xi \\ \theta_{0_j} &= \mu_0 + \varphi_0 \cdot K_j + \tau_0 \cdot \Theta \\ \theta_{1_j} &= \mu_1 + \varphi_1 \cdot K_j + \tau_1 \cdot \Omega \end{aligned}$$

and thus

$$z_{\log(y_j)} = \mu_0 + \varphi_0 \cdot K_j + \tau_0 \cdot \Theta + (\mu_1 + \varphi_1 \cdot K_j + \tau_1 \cdot \Omega) z_x + \sigma \cdot \Xi$$

Assuming independence, we can convolute the addition of the random variables Ξ, Θ and Ω and merge them into a single one which we call again Ξ:

$$z_{\log(y_j)} = \mu_0 + \varphi_0 \cdot K_j + \sqrt{\sigma^2 + \tau_0^2 + \tau_1^2 z_x^2} \cdot \Xi + (\mu_1 + \varphi_1 \cdot K_j) z_x$$

De-standardising from Equation 2 yields:

$$\log(y_j) = \underbrace{\hat{\sigma}_y \mu_0 + \overline{\log(y_j)}}_{\mu_{0,\text{unsc}}} + \underbrace{\hat{\sigma}_y \varphi_0}_{\varphi_{0,\text{unsc}}} \cdot K_j + \sqrt{\underbrace{\hat{\sigma}_y^2 \sigma^2}_{\sigma^2_{\text{unsc}}} + \underbrace{\hat{\sigma}_y^2 \tau_0^2}_{\tau^2_{0,\text{unsc}}} + \underbrace{\hat{\sigma}_y^2 \tau_1^2}_{\tau^2_{1,\text{unsc}}} z_x^2} \cdot \Xi + (\underbrace{\hat{\sigma}_y \mu_1}_{\mu_{1,\text{unsc}}} + \underbrace{\hat{\sigma}_y \varphi_1}_{\varphi_{1,\text{unsc}}} \cdot K_j) z_x$$

and therefore,

$$\log(y_j) \sim \mathcal{N}(\mu_{0,\text{unsc}} + \varphi_{0,\text{unsc}} \cdot K_j + (\mu_{1,\text{unsc}} + \varphi_{1,\text{unsc}} \cdot K_j) z_x, \quad \sigma^2_{\text{unsc}} + \tau^2_{0,\text{unsc}} + \tau^2_{1,\text{unsc}} z_x^2)$$

The distribution of y_j is thus log-normal. The expected value at the **group** level is:

$$\begin{aligned} E[y_j] &= e^{\mu_{0,\text{unsc}} + \varphi_{0,\text{unsc}} \cdot K_j + (\mu_{1,\text{unsc}} + \varphi_{1,\text{unsc}} \cdot K_j) z_x + \frac{\sigma^2_{\text{unsc}}}{2} + \frac{\tau^2_{0,\text{unsc}}}{2} + \frac{\tau^2_{1,\text{unsc}} z_x^2}{2}} \\ &= e^{\mu_{0,\text{unsc}} + \varphi_{0,\text{unsc}} \cdot K_j + \frac{\sigma^2_{\text{unsc}}}{2} + \frac{\tau^2_{0,\text{unsc}}}{2}} \cdot \underbrace{e^{(\mu_{1,\text{unsc}} + \varphi_{1,\text{unsc}} \cdot K_j) z_x + \frac{\tau^2_{1,\text{unsc}} z_x^2}{2}}}_{\text{effect of increasing attempts}} \\ &= \underbrace{e^{\varphi_{0,\text{unsc}} \cdot K_j}}_{\text{child-effect}} \; e^{\mu_{0,\text{unsc}} + \frac{\sigma^2_{\text{unsc}}}{2} + \frac{\tau^2_{0,\text{unsc}}}{2}} \cdot \underbrace{e^{\varphi_{1,\text{unsc}} \cdot K_j z_x + \frac{\tau^2_{1,\text{unsc}} z_x^2}{2}}}_{\text{child-effect}} \; e^{\mu_{1,\text{unsc}} z_x} \end{aligned}$$

At the **individual** level, the hyper-priors just disappear:

$$\begin{aligned} E[y_j] &= e^{\theta_{0_j,\text{unsc}} + \theta_{1_j,\text{unsc}} z_x + \frac{\sigma^2_{\text{unsc}}}{2}} \\ &= \underbrace{e^{\theta_{1_j,\text{unsc}} z_x}}_{\text{effect of increasing attempts}} \; e^{\theta_{0_j,\text{unsc}} + \frac{\sigma^2_{\text{unsc}}}{2}} \end{aligned}$$

where

$$\begin{aligned} \theta_{0_j,\text{unsc}} &= \hat{\sigma}_y \theta_{0_j} + \overline{\log(y_j)} \\ \theta_{1_j,\text{unsc}} &= \hat{\sigma}_y \theta_{1_j} \end{aligned}$$

The multiplicative effect of being a kid acts only on the group level, but the effect of increasing attempts acts on the individual and group level.