# Beta regression – three views

## 1. Index form

*The per-observation reading: what happens for each row $i$ of your data.*

$$\begin{aligned} y_i \mid \mu_i, \, \sigma_i &\sim \mathrm{Beta}(\mu_i \sigma_i, \, (1-\mu_i)\sigma_i) \\ \mathrm{logit}(\mu_i) &= \beta_0 + \beta_1 \, x_i \\ \log(\sigma_i) &= \gamma_0 \end{aligned}$$

***Coefficient reading.*** *On the response scale, $\exp(\hat{\beta})$ is the odds ratio for the mean proportion: a one-unit increase multiplies the odds $\mu/(1-\mu)$ by $\exp(\hat{\beta})$ (logit link).*

## 2. Matrix form

*The same model in matrix notation. The structural contract every textbook past chapter 4 switches to.*

$$\begin{aligned} \mathbf{y} \mid \boldsymbol{\mu}, \, \boldsymbol{\sigma} &\sim \mathrm{Beta}(\boldsymbol{\mu} \odot \boldsymbol{\sigma}, \, (1-\boldsymbol{\mu}) \odot \boldsymbol{\sigma}) \\ \mathrm{logit}(\boldsymbol{\mu}) &= \mathbf{X}\boldsymbol{\beta} \\ \log(\boldsymbol{\sigma}) &= \mathbf{X}_\sigma \boldsymbol{\gamma} \end{aligned}$$

## 3. Worked observation (i = 1)

*The index-form equations evaluated at the first observation of your data, with coefficient estimates plugged in.*

$$\begin{aligned} \hat{\eta}_1 \;&=\; -0.824 - 0.0874 \cdot 1.43 \;\approx\; -0.95 \\ \hat{\mu}_1 \;&=\; \mathrm{logistic}(\hat{\eta}_1) \;=\; \mathrm{logistic}(-0.95) \;\approx\; 0.279 \\ \mathrm{y}_1 \;&\sim\; \mathrm{Beta}(\hat{\mu}_1 \hat{\sigma}_1, \, (1-\hat{\mu}_1)\hat{\sigma}_1) \end{aligned}$$

*Linear predictor $\hat{\eta}_1 = -0.95$ on the link scale; predicted mean $\hat{\mu}_1 = \mathrm{logistic}(\hat{\eta}_1) \approx 0.279$ on the response scale. No additive error term enters the linear predictor – the spread is set by the likelihood above.*

And the $\sigma$ submodel at $i = 1$:

$$\log \hat{\sigma}_1 \;=\; -1.04 \;\Rightarrow\; \hat{\sigma}_1 \;=\; \exp(-1.04) \;\approx\; 0.353$$

*Stacking the same response equation for all $n = 100$ observations:*

$$\underbrace{\begin{bmatrix} -0.95 \\ -0.768 \\ -0.806 \\ -0.79 \\ -0.797 \\ \vdots \\ -0.81 \\ -0.861 \end{bmatrix}}_{\hat{\boldsymbol{\eta}}_{100\times 1} \text{ (linear predictor, link scale)}} = \underbrace{\begin{bmatrix} 1 & 1.43 \\ 1 & -0.651 \\ 1 & -0.207 \\ 1 & -0.393 \\ 1 & -0.32 \\ \vdots & \vdots \\ 1 & -0.164 \\ 1 & 0.421 \end{bmatrix}}_{\mathbf{X}_{100\times 2}} \underbrace{\begin{bmatrix} -0.824 \\ -0.0874 \end{bmatrix}}_{\hat{\boldsymbol{\beta}}_{2\times 1} \text{ (estimated)}}$$

*And the $\sigma$ submodel (no observed counterpart – $\sigma$ here is the family's spread parameter and is not a residual SD; see Tab 1 for its specific role in this distribution):*

$$\log \underbrace{\begin{bmatrix} 0.353 \\ 0.353 \\ 0.353 \\ 0.353 \\ 0.353 \\ \vdots \\ 0.353 \\ 0.353 \end{bmatrix}}_{\boldsymbol{\sigma}_{100\times 1}} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{X}_{\sigma,\,100\times 1}} \underbrace{\begin{bmatrix} -1.04 \end{bmatrix}}_{\boldsymbol{\gamma}_{1\times 1}}$$