Poisson regression – three views

1. Index form

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

$$
\begin{aligned}
y_i \mid \mu_i &\sim \text{Poisson}(\mu_i) \\
\log(\mu_i) &= \beta_0 + \beta_1\, x_i
\end{aligned}
$$

Coefficient reading. *On the response scale,* $\exp(\hat{\beta})$ *is the rate ratio: a one-unit increase in the predictor multiplies the expected count by* $\exp(\hat{\beta})$ *(log link).*

2. Matrix form

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

$$
\begin{aligned}
\mathbf{y} \mid \boldsymbol{\mu} &\sim \text{Poisson}(\boldsymbol{\mu}) \\
\log(\boldsymbol{\mu}) &= \mathbf{X}\boldsymbol{\beta}
\end{aligned}
$$

3. Worked observation (i = 1)

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

$$
\begin{aligned}
\hat{\eta}_1 \;&=\; 0.955 - 0.0438 \cdot 0.45 \;\approx\; 0.936 \\
\hat{\mu}_1 \;&=\; \exp(\hat{\eta}_1) \;=\; \exp(0.936) \;\approx\; 2.55 \\
\mathrm{y}_1 \;&\sim\; \text{Poisson}(\hat{\mu}_1)
\end{aligned}
$$

Linear predictor $\hat{\eta}_1 = 0.936$ *on the link scale; predicted mean* $\hat{\mu}_1 = \exp(\hat{\eta}_1) \approx 2.55$ *on the response scale. No additive error term enters the linear predictor – the spread is set by the likelihood above.*

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

$$
\underbrace{\begin{bmatrix} 0.936 \\ 0.956 \\ 0.969 \\ 0.996 \\ 1.02 \\ \vdots \\ 0.963 \\ 0.911 \end{bmatrix}}_{\hat{\boldsymbol{\eta}}_{100\times 1}\ \text{(linear predictor, link scale)}}
=
\underbrace{\begin{bmatrix} 1 & 0.45 \\ 1 & -0.0186 \\ 1 & -0.318 \\ 1 & -0.929 \\ 1 & -1.49 \\ \vdots & \vdots \\ 1 & -0.166 \\ 1 & 1.02 \end{bmatrix}}_{\mathbf{X}_{100\times 2}}
\underbrace{\begin{bmatrix} 0.955 \\ -0.0438 \end{bmatrix}}_{\hat{\boldsymbol{\beta}}_{2\times 1}\ \text{(estimated)}}
$$