Data-Split Validation: Testing Generalization

Full Fit: $\chi^2 = 13.5$
(all data used for fitting)

Training: Fit on low-z bins
(test bins hidden)

Test: Predict high-z bins
($E_{\mathrm{split}} = 1.4$)

Training
Test (held out)

Test

Larger residuals
= poor generalization

Residual

Redshift z