

Generalized Linear Model

Test name: Generalized Linear Model

Dataset: D:\DATA ANALYSIS\H Regression Tests and Models\Generalized Linear Model\dataset.csv

Rows in original dataset: 649

Columns in original dataset: 33

Rows used in GLM models: 649

Continuous outcome: G3

Binary outcome created: $\text{glm_pass} = 1$ when $G3 \geq 10.00$

Count outcome created: $\text{glm_count} = \text{round}(\text{absences})$ clipped at zero

Numeric predictors: G1, G2, studytime, failures, age, Medu, Fedu

Categorical predictors: school, sex, address

Gaussian formula: $G3 \sim G1 + G2 + \text{studytime} + \text{failures} + \text{age} + \text{Medu} + \text{Fedu} + C(\text{school}) + C(\text{sex}) + C(\text{address})$

Binomial formula: $\text{glm_pass} \sim G1 + G2 + \text{studytime} + \text{failures} + \text{age} + \text{Medu} + \text{Fedu} + C(\text{school}) + C(\text{sex}) + C(\text{address})$

Poisson formula: $\text{glm_count} \sim G1 + G2 + \text{studytime} + \text{failures} + \text{age} + \text{Medu} + \text{Fedu} + C(\text{school}) + C(\text{sex}) + C(\text{address})$

Gaussian GLM: identity link for continuous mean outcome.

Binomial GLM: logit link for pass/event probability; $\exp(\text{coefficient})$ is an odds ratio.

Poisson GLM: log link for count outcome; $\exp(\text{coefficient})$ is a rate ratio.

Decision guide: $p < 0.05$ and confidence interval excluding the null value indicates stronger evidence.

Charts generated: 7

GLM coefficients

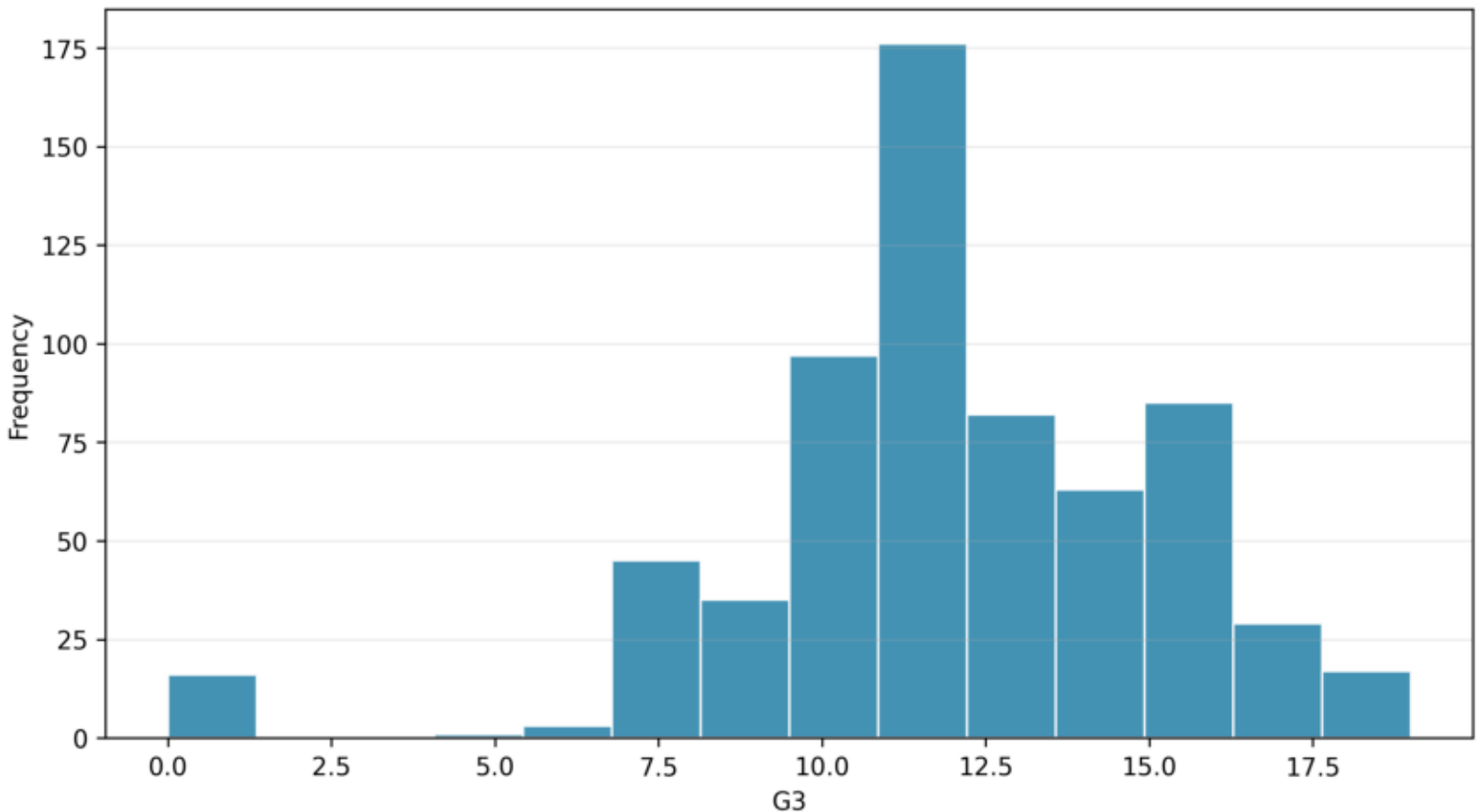
model	term	coefficient	standard_error	z_or_t_value	p_value	ci95_lower	ci95_upper	interpretation	exp_coefficient	exp_ci95_lower	exp_ci95_upper	decision_algorithm
Gaussian id	Intercept	-0.1923	0.7888	-0.2438	0.8074	-1.7382	1.3536	Mean difference	0.8251	0.1758	3.8715	Not statistically significant
Gaussian id	C(school)[T.]	-0.2170	0.1167	-1.8599	0.0629	-0.4456	0.0117	Mean difference	0.8050	0.6404	1.0117	Not statistically significant
Gaussian id	C(sex)[T.M]	-0.1981	0.1046	-1.8936	0.0583	-0.4032	0.0069	Mean difference	0.8203	0.6682	1.0070	Not statistically significant
Gaussian id	C(address)[T.]	0.1132	0.1153	0.9824	0.3259	-0.1127	0.3392	Mean difference	1.1199	0.8934	1.4038	Not statistically significant
Gaussian id	G1	0.1295	0.0368	3.5233	0.0004	0.0575	0.2016	Mean difference	1.1383	1.0592	1.2234	Statistically significant
Gaussian id	G2	0.8804	0.0343	25.6469	0.0000	0.8131	0.9477	Mean difference	2.4119	2.2550	2.5798	Statistically significant
Gaussian id	studytime	0.0539	0.0631	0.8543	0.3929	-0.0697	0.1775	Mean difference	1.0554	0.9326	1.1942	Not statistically significant
Gaussian id	failures	-0.2253	0.0951	-2.3684	0.0179	-0.4118	-0.0389	Mean difference	0.7982	0.6624	0.9619	Statistically significant
Gaussian id	age	0.0319	0.0434	0.7353	0.4621	-0.0531	0.1168	Mean difference	1.0324	0.9483	1.1239	Not statistically significant
Gaussian id	Medu	-0.0526	0.0588	-0.8945	0.3710	-0.1677	0.0626	Mean difference	0.9488	0.8456	1.0646	Not statistically significant
Gaussian id	Fedu	0.0249	0.0591	0.4206	0.6741	-0.0910	0.1408	Mean difference	1.0252	0.9130	1.1512	Not statistically significant
Binomial log	Intercept	-25.2555	4.3307	-5.8317	0.0000	-33.7436	-16.7675	Odds ratio for	0.0000	0.0000	0.0000	Statistically significant

GLM fit statistics

model	family	link	n	df_model	df_resid	deviance	pearson_chi2	scale	aic	log_likelihood
Gaussian iden	Gaussian	identity	649.0000	10.0000	638.0000	1002.1000	1002.1000	1.5707	2145.7210	-1061.8605
Binomial logit	Binomial	logit	649.0000	10.0000	638.0000	176.1896	1327.5772	1.0000	198.1896	-88.0948
Poisson log GL	Poisson	log	649.0000	10.0000	638.0000	3067.1642	3161.2421	1.0000	4462.2974	-2220.1487

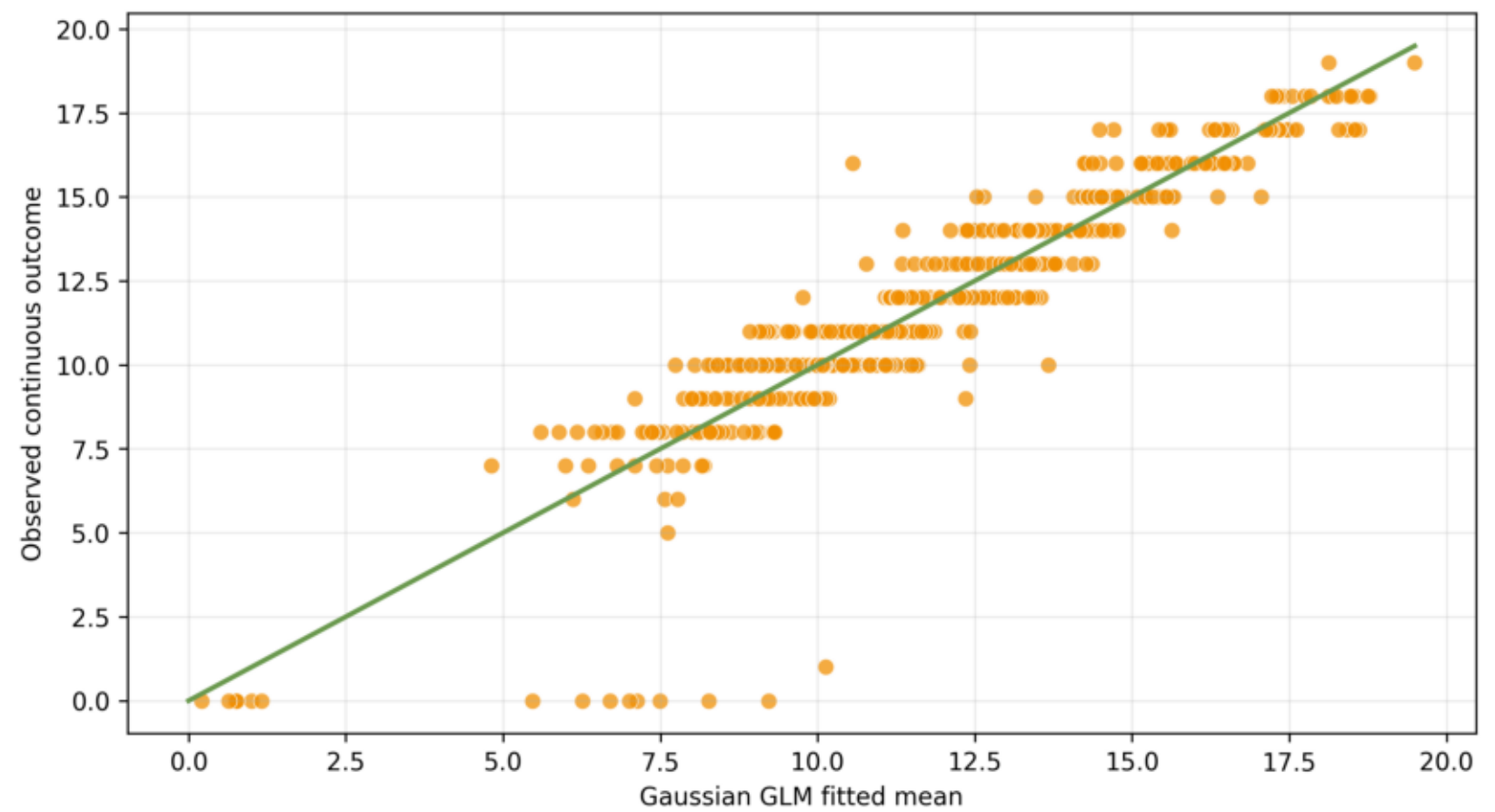
Continuous outcome distribution

Gaussian identity GLM models the expected continuous outcome from selected predictors.



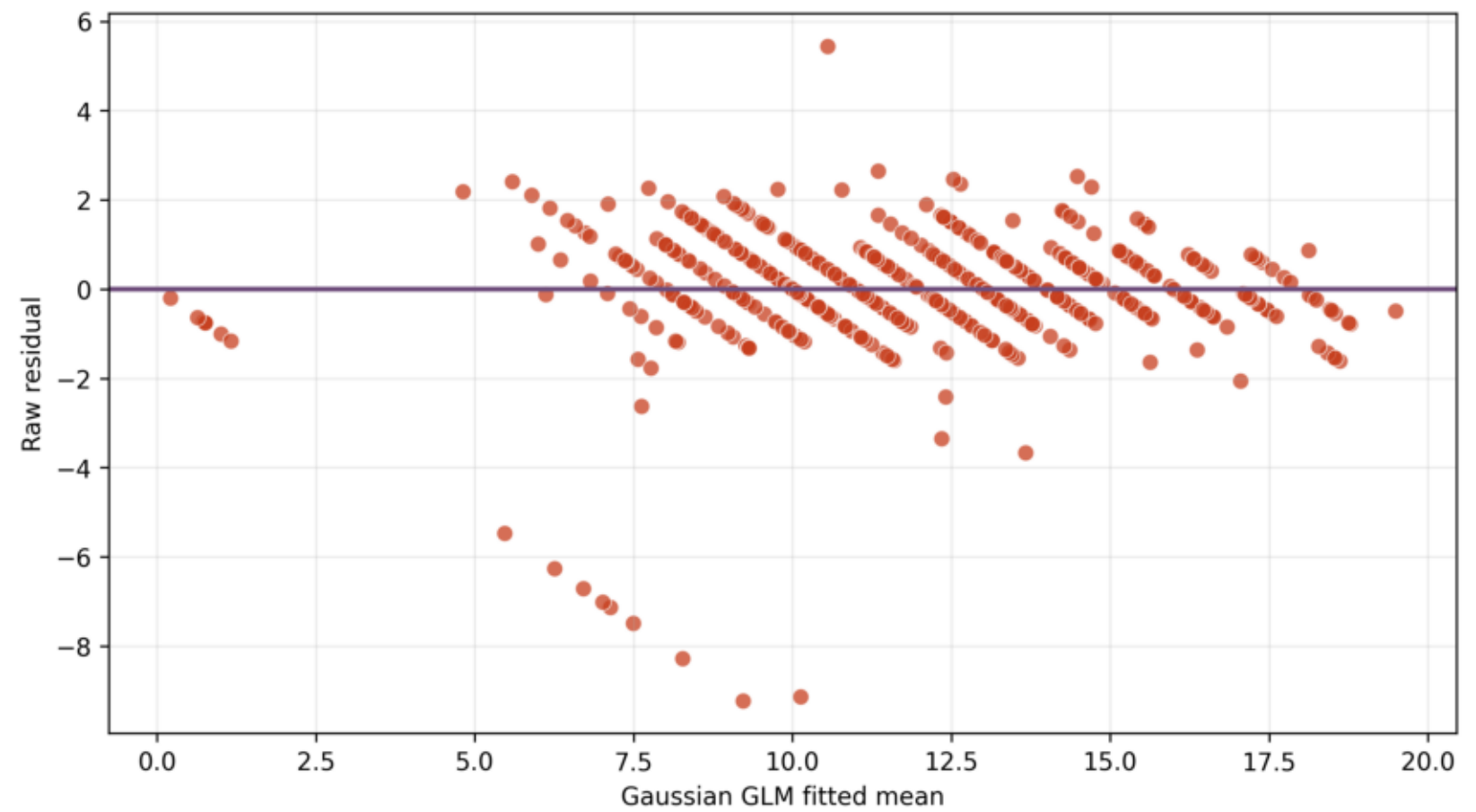
Gaussian GLM observed vs fitted

Points close to the diagonal show stronger agreement between observed and fitted means.



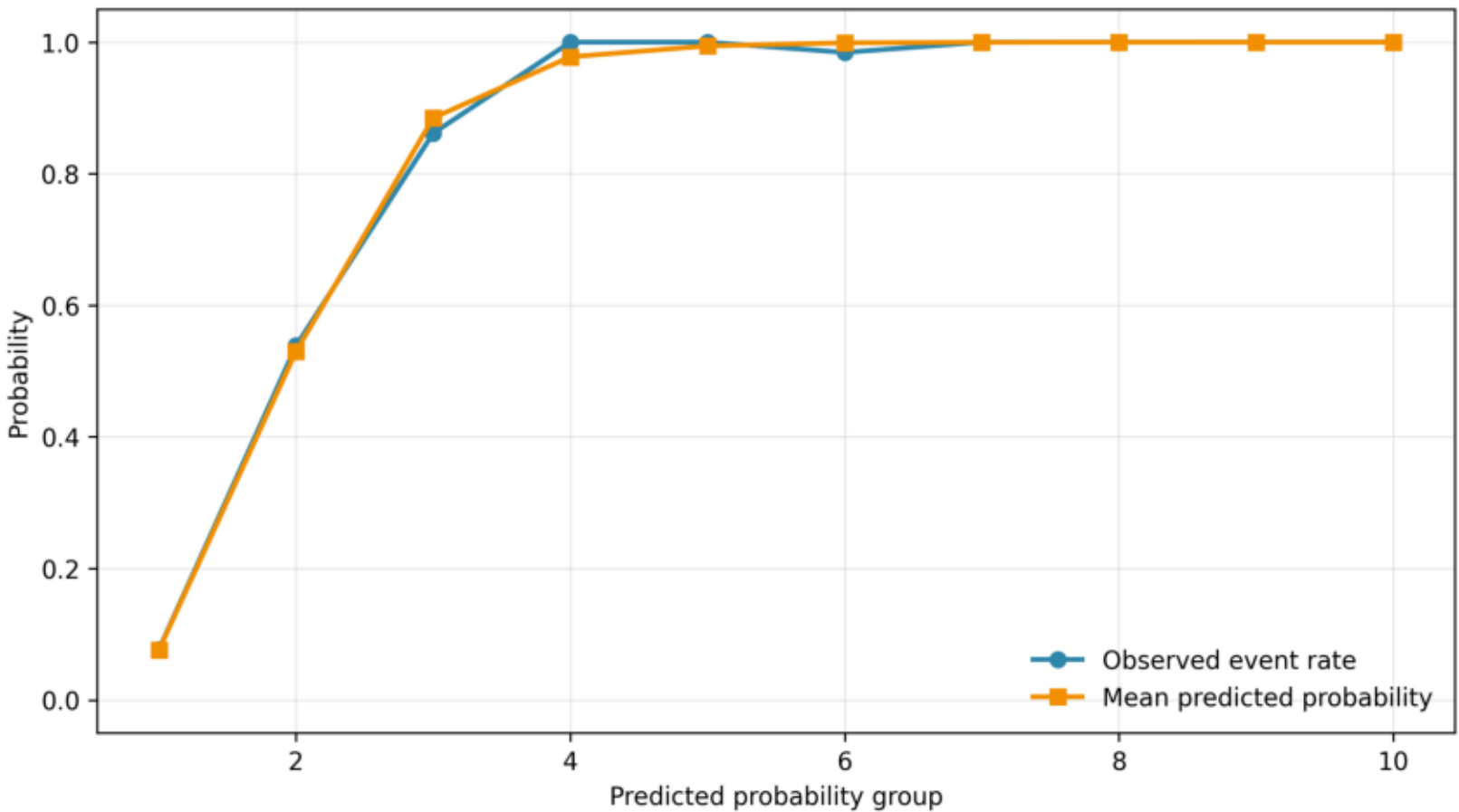
Gaussian GLM residuals vs fitted

A random cloud around zero supports a more stable linear mean specification.



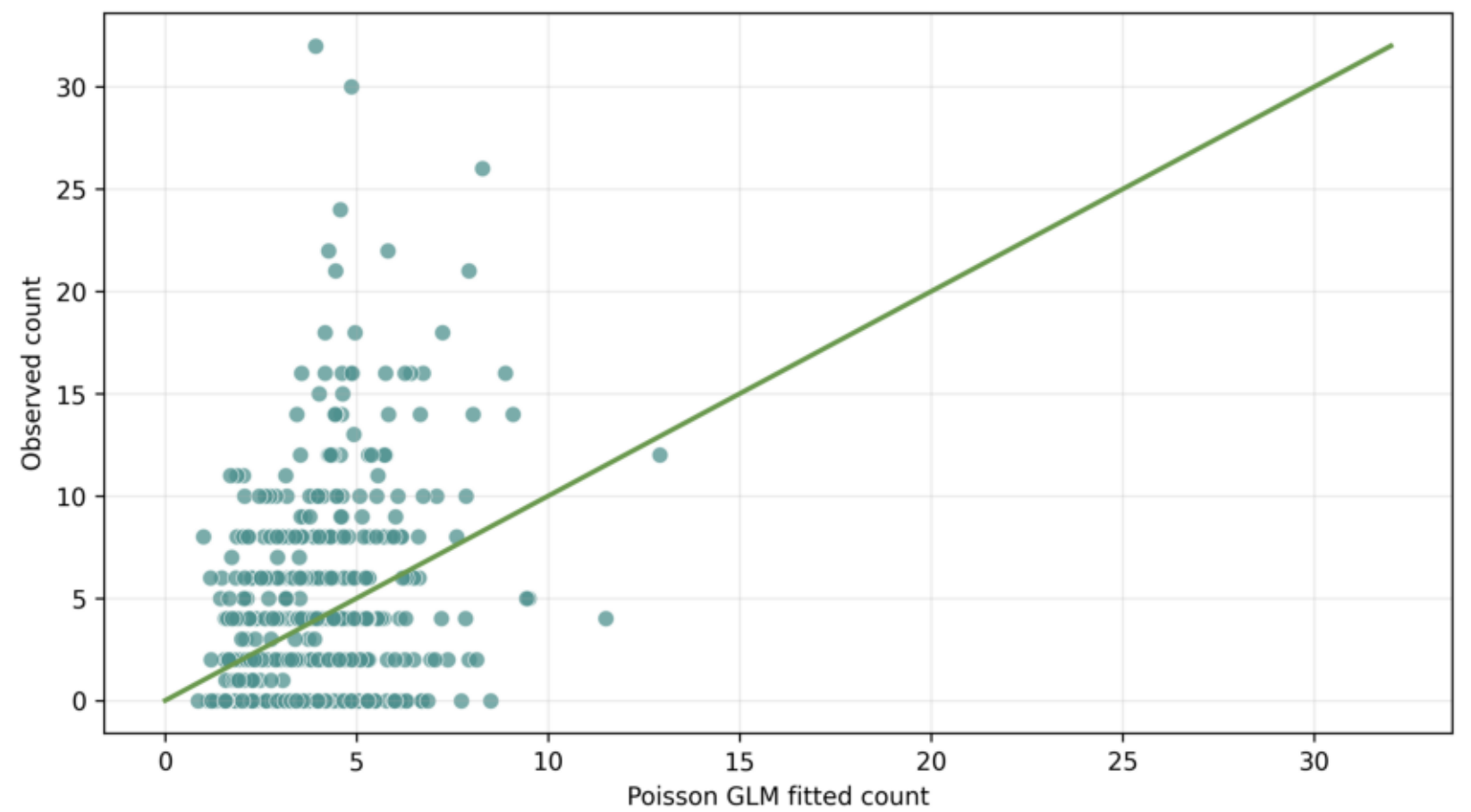
Binomial GLM calibration by predicted probability

Observed event rates should follow predicted probabilities across ordered probability groups.



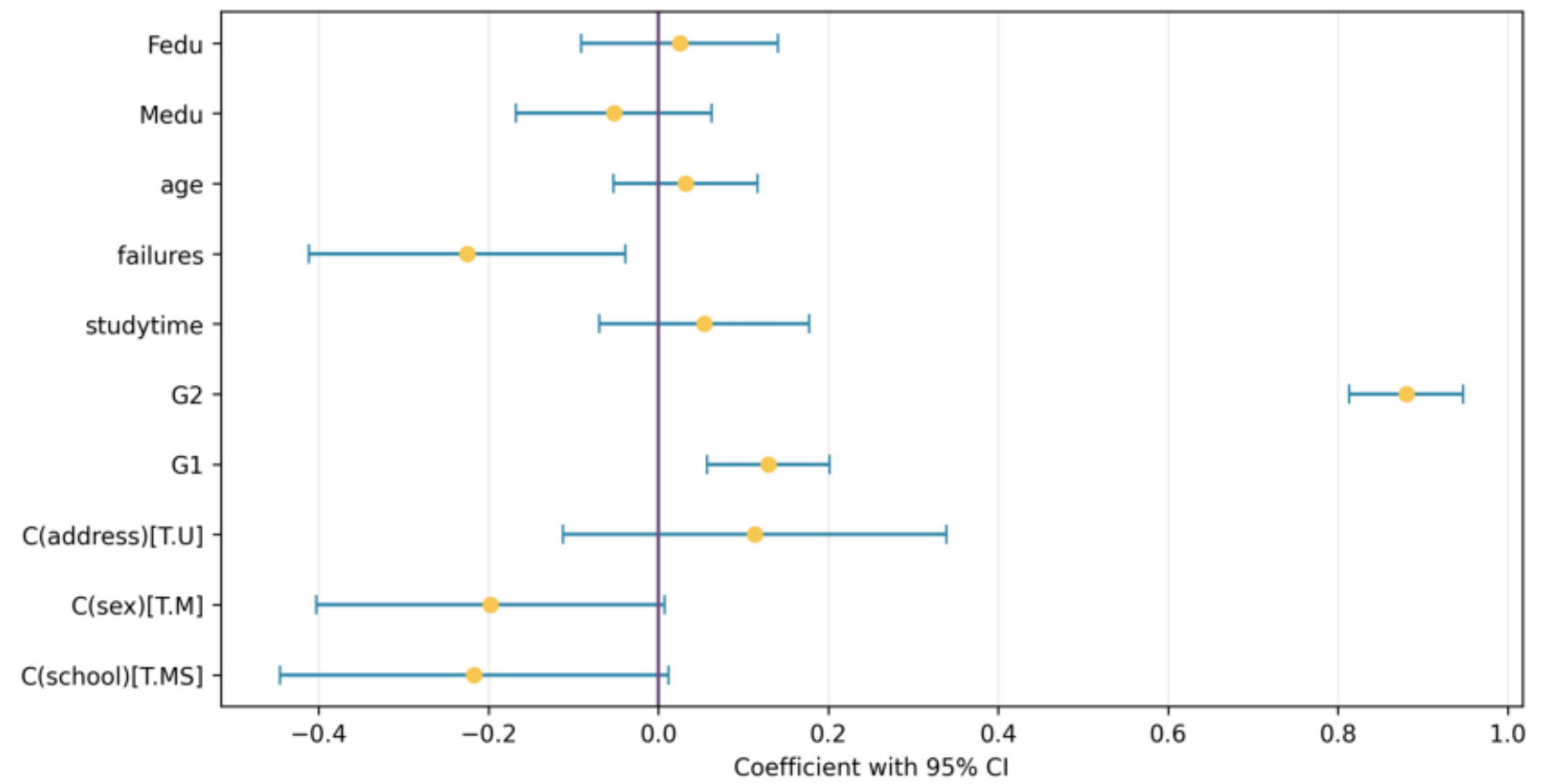
Poisson GLM observed vs fitted counts

The Poisson log GLM models non-negative count outcomes through an expected count.



Gaussian GLM coefficient plot

Intervals not crossing zero indicate stronger evidence for an adjusted mean effect.



GLM model AIC summary

AIC values are displayed for model context; compare only models fitted to the same outcome scale.

