

Hurdle Model

Test name: Hurdle Model

Dataset: D:/DATA ANALYSIS/H Regression Tests and Models/Hurdle Model/dataset.csv

Rows in original dataset: 649

Columns in original dataset: 33

Rows used in hurdle model: 649

Count outcome source: absences

Count outcome created: hurdle_count = round(absences) clipped at zero

Zero count observations: 244

Positive count observations: 405

Numeric predictors: G1, G2, G3, studytime, failures, age, Medu

Categorical predictors: school, sex, address

Hurdle formula: hurdle_positive ~ G1 + G2 + G3 + studytime + failures + age + Medu + school + sex + address

Positive-count formula: hurdle_count ~ G1 + G2 + G3 + studytime + failures + age + Medu + school + sex + address

Hurdle part: Binomial GLM with logit link for $P(\text{count} > 0)$.

Positive-count part: Poisson GLM with log link fitted only on positive counts.

Expected count: predicted probability of positive count multiplied by predicted positive-count mean.

Positive-count Pearson dispersion: 3.0749

Positive-count dispersion differs from 1; consider robust or negative-binomial sensitivity reporting.

Interpretation: hurdle coefficients are odds-ratio effects; positive-count coefficients are rate-ratio effects.

Decision guide: $p < 0.05$ and confidence interval not crossing 1 indicates stronger evidence for a predictor effect.

Charts generated: 8

Hurdle logit coefficients

	model	term	coefficient	standard_error	z_value	p_value	ci95_lower	ci95_upper
	Hurdle binary logit model	(Intercept)	0.99357272	1.34899750	0.7365267	0.4614102079	-1.65041380	3.63755923
	Hurdle binary logit model	G1	-0.24253716	0.06774140	-3.5803389	0.0003431489	-0.37530787	-0.10976646
	Hurdle binary logit model	G2	-0.15624889	0.08780289	-1.7795414	0.0751510363	-0.32833940	0.01584162
	Hurdle binary logit model	G3	0.26682394	0.07422908	3.5946012	0.0003248888	0.12133762	0.41231027
	Hurdle binary logit model	studytime	-0.16569522	0.10624671	-1.5595327	0.1188703622	-0.37393495	0.04254451
	Hurdle binary logit model	failures	0.12497503	0.17856919	0.6998689	0.4840091675	-0.22501416	0.47496422
	Hurdle binary logit model	age	0.08411114	0.07525905	1.1176216	0.2637286279	-0.06339388	0.23161616
	Hurdle binary logit model	Medu	0.08061256	0.08007845	1.0066698	0.3140934956	-0.07633833	0.23756345
	Hurdle binary logit model	schoolMS	-0.62760383	0.20011046	-3.1362870	0.0017110168	-1.01981312	-0.23539454
	Hurdle binary logit model	sexM	-0.20631976	0.17892949	-1.1530786	0.2488781337	-0.55701511	0.14437559
	Hurdle binary logit model	addressU	-0.09881177	0.19769140	-0.4998284	0.6171959355	-0.48627980	0.28865626
odds_ratio	odds_ratio_ci95_lower	odds_ratio_ci95_upper	decision_alpha_0_05					
2.7008667	0.1919705	37.9989768	Not statistically significant					
0.7846346	0.6870777	0.8960434	Statistically significant					
0.8553463	0.7201186	1.0159678	Not statistically significant					
1.3058105	1.1290060	1.5103030	Statistically significant					
0.8473044	0.6880217	1.0434625	Not statistically significant					
1.1331202	0.7985049	1.6079567	Not statistically significant					
1.0877498	0.9385737	1.2606358	Not statistically significant					
1.0839508	0.9265027	1.2681555	Not statistically significant					
0.5338695	0.3606623	0.7902590	Statistically significant					
0.8135729	0.5729166	1.1553180	Not statistically significant					
0.9059132	0.6149097	1.3346329	Not statistically significant					

Positive-count coefficients

	model	term	coefficient	standard_error	z_value	p_value	ci95_lower
Positive-count	Poisson model	(Intercept)	0.871173817	0.32624128	2.67033594	7.577539e-03	0.23175266
Positive-count	Poisson model	G1	0.020524187	0.01624302	1.26356933	2.063846e-01	-0.01131156
Positive-count	Poisson model	G2	0.001602706	0.02525331	0.06346518	9.493961e-01	-0.04789286
Positive-count	Poisson model	G3	-0.071195280	0.02428520	-2.93163266	3.371853e-03	-0.11879339
Positive-count	Poisson model	studytime	-0.102158657	0.02730312	-3.74164705	1.828182e-04	-0.15567180
Positive-count	Poisson model	failures	-0.004495488	0.03474021	-0.12940301	8.970388e-01	-0.07258505
Positive-count	Poisson model	age	0.105044723	0.01749141	6.00550351	1.907390e-09	0.07076219
Positive-count	Poisson model	Medu	-0.014145547	0.02067676	-0.68412769	4.938945e-01	-0.05467126
Positive-count	Poisson model	schoolMS	-0.426506361	0.05275712	-8.08433822	6.250266e-16	-0.52990841
Positive-count	Poisson model	sexM	-0.006357317	0.04396586	-0.14459669	8.850293e-01	-0.09252881
Positive-count	Poisson model	addressU	0.083129884	0.04953112	1.67833638	9.328145e-02	-0.01394933
ci95_upper	rate_ratio	rate_ratio_ci95_lower	rate_ratio_ci95_upper	decision_alpha_0_05			
1.51059498	2.3897143	1.2608078	4.5294249	Statistically significant			
0.05235993	1.0207363	0.9887522	1.0537550	Not statistically significant			
0.05109828	1.0016040	0.9532359	1.0524263	Not statistically significant			
-0.02359717	0.9312800	0.8879912	0.9766791	Statistically significant			
-0.04864552	0.9028863	0.8558400	0.9525187	Statistically significant			
0.06359407	0.9955146	0.9299867	1.0656597	Not statistically significant			
0.13932726	1.1107603	1.0733259	1.1495002	Statistically significant			
0.02638017	0.9859540	0.9467963	1.0267312	Not statistically significant			
-0.32310432	0.6527857	0.5886589	0.7238983	Statistically significant			
0.07981418	0.9936628	0.9116229	1.0830858	Not statistically significant			
0.18020910	1.0866829	0.9861475	1.1974677	Not statistically significant			

Hurdle model fit statistics

model_part	family	link	n	n_zero	n_positive	df_residual	null_deviance	residual_deviance	aic		
Hurdle	binary	part	Binomial	logit	649	244	405	638	859.345	812.4525	834.4525
Positive count	part	Poisson	log	405	0	405	394	1224.883	1049.8428	2444.9760	
pearson_dispersion											note
	NA		Models	whether	count	crosses	the	zero			hurdle.
	3.074874	Poisson	log	GLM	fitted	to	positive	counts			only.















