Linear vs. Logistic Regression

Topic	Linear Regression	Logistic Regression	Comparison
Response variable	Continuous	Binary	Different
Regression function	$\left(\sum_{i=0}^{p} \beta_i X_i\right) + \epsilon_i$	$\left(\sum_{i=0}^{p} \beta_i X_i\right) + \epsilon_i$	Same
Residual distribution	Normal	Bernoulli	Different
Response Transformation	Optional but useful, many choices.	Required: Logit transformation of outcome probability	Different
Fitting method	Least squares (same as maximum likelihood)	Maximum Likelihood (not the same as least squares)	Similar
Measure of variation	Error Sum of Squares	Residual Deviance	Similar
Coefficient inference	t-tests and confidence intervals	t-tests and confidence intervals	Same
Model building	Best subsets, stepwise	Best subsets, stepwise	Same
Residual analysis	Residuals, studentized residuals. Plot vs. fitted or X_i, check for patterns in the points. Q-Q plot.	Deviance residuals (studentized). Plot vs. fitted or X_i, check for shape of Lowess fit. Q-Q plot.	Similar
Fit summary	R^2, R_a^2, residual std. error, overall F test, …	Sensitivity, Specificity, AUC, residual deviance, Likelihood ratio test	Mostly different
Influence diagnostics	Leverage values, Cook's distance, etc.	Leverage values, Cook's distance, etc.	Same
Assess model fit	Plot Y_i vs. $\hat{Y}_i$	2x2 table classifying observed response category vs. predicted category	Different