Analysis	Usage	Assumptions
Canonical Corr. An. (CCA)	Measure correlation between X's and Y's	Linearity of Correlations, Linearity of Relationships, Multivariate Normality
Classification An. KNN	Compute distance from point a to all other points. Assign Group who matches most.	None. Non-parametric test. $K = sqrt(n\_i)$
Cluster Analysis (CA)	Separate items into clusters on the basis of p variables	N objects/cases/rows of data. K clusters/groups
CA - Hierarchical	Cluster data in n steps, joining observations to form clusters.	**Types:** Single Linkage: min, Complete: max, Average: avg
CA - Partitioning	For K, best partition of K to optimize some function.	**Types:** K-Means: Minimize WSS via eucl. Dist. Wards: Mix of Hier. + K-means. K-medoids: nonparam. Dist to median
Mantel-Haenszel	Alt to Pearson Chi-Sq when ordinal factors involved	$M^2 = (n - 1)r^2$. R = corr. $M^2 \sim X^2$, df = 1 under H0
Fisher's Exact Test	Randomization Test when comparing two props.	Independence
Mantel-Haenszel Excess	Residual for cell in RxC table. Conditional Indep. And homogenous assoc for k odds ratios in k 2x2 tables	Odds Ratio same in all tables(Breslow-Day). Sum of expected counts at least 5
Logistic Regression	Binary Responses.	Indep., no multicollinearity, linearity of X's + log odds
Probit Regression	Alternative to logit link for glm. Uses CDF of N(0,1) instead of CDF of logistic fun.	Same as logistic

Sampling Scheme	Known/Not Known	Test Type
Poisson	No Marginal Totals	Homogeneity, Indep.
Multinomial	Same and T (Total) fixed	Homogeneity, Indep.
Prospective Product Bin.	Row Totals Fixed	Homogeneity for odds ratio
Retrospective Product Bin.	Column totals Fixed	Homogeneity for odds ratio
Randomized Bin.	Like P Prod Bin	Homogeneity
Hypergeom Prob.	R + C Totals fixed	Fisher's Exact Test

Chi-square GoF: df = num cells - 1. Chi-square Indep: df = $(r - 1) * (c - 1)$	LRT approx $X^2(v)$ where v = diff(num_param_full, num_param_reduced)	Odds Ratio = $(n_{11} * n_{22}) / (n_{21} * m_{12})$ for 2x2 table of counts
Logistic: $\pi = e^\nu / 1 + e^\nu$. $E(Y) = \pi$. $V(Y) = \pi(1 - \pi)$	LRT/DinD = DevianceFull - DevianceReduced	Deviance = SSRes = $-2 * \log(MaxLikelihoodFunction)$
AIC = deviance + $\log(n) * p$	BIC = deviance + $2 * p$	Logit = $\log(\pi / 1 - \pi)$

Assessment	Applicable Models
Scatterplots	Log-Linear, Negative Bin, Logistic Bin (log-odds vs expl. vars), Logistic
Residuals	Log-Linear, Negative Bin, Logistic Bin
Deviance GoF	Log-Linear, Negative Bin, Logistic Bin
Hosmer-Lemeshow GoF	Logistic
Informal Testing of Terms	Logistic
Hotellings T^2	DFA
Wilks' lambda	DFA