

STAT 102 - Inference Scenarios

	Response variable		Numeric quantity	Assumptions for theory-based inference	Standardized test statistic and its theory-based distribution
1 variable	Numerical		Mean	Independence and at least 30 observations	$t \sim t(df = n - 1)$
	Categorical (Binary)		Proportion	Independence and at least 10 successes and at least 10 failures	$z \sim N(0, 1)$
	Response variable	Explanatory variable	Numeric quantity	Assumptions for theory-based inference	Standardized test statistic and its theory-based distribution
2 variables	Numerical	Categorical (Binary)	Difference in means	Independence and at least 30 observations in each category	$t \sim t(df = \min(n_1, n_2) - 1)$, where n_1 = # of observations in group 1, n_2 = # of observations in group 2
	Categorical (Binary)	Categorical (Binary)	Difference in proportions	Independence and at least 10 successes and at least 10 failures in each category	$z \sim N(0, 1)$
	Numerical	Numerical	Correlation	Independence and at least 30 observations	$t \sim t(df = n - 2)$
	Numerical	Mix	Linear regression coefficient	Linearity, constant variability, independence, and normality	$t \sim t(df = n - p - 1)$, where n is # of observations, p is # of predictors
	Numerical	Categorical (2+)	F-statistic (from ANOVA)	Independence, normality, and constant variability	$F \sim F(df_1, df_2)$, where $df_1 = n_{\text{groups}} - 1$, $df_2 = n_{\text{observations}} - n_{\text{groups}}$
	Categorical (2+)	Categorical (2+)	Chi-squared	Independence and at least 10 observations in each cell	$\chi^2 \sim \chi^2(df = (r - 1)(c - 1))$, where r = # of rows, c = # of columns