Assumptions for Inference *(and the conditions that confirm or override them)*

Proportions (*z*)

- **One Sample**

1. Individuals are independent	1. SRS and < 10% of population
2. Sample is sufficiently large	2. Successes and failures ≥ 10

- **Two Sample**

1. Samples are independent	1. (Think about how the data were collected.)
2. Data in each sample are independent	2. Both SRSs and < 10% of populations OR random allocation
3. Both samples are sufficiently large	3. Successes and failures ≥ 10 for both

Means (*t*)

- **One Sample** (df = $n - 1$)

1. Individuals are independent	1. SRS and < 10% of the population
2. Population has a Normal model	2. Histogram is unimodal and symmetric*****

- **Matched Pairs** (df = $n - 1$)

1. Data are matched	1. (Think about the design.)
2. Individuals are independent	2. SRS and < 10% OR random allocation
3. Population of differences is Normal	3. Histogram of differences is unimodal and symmetric

- **Two Independent Samples** (df from calculator)

1. Samples are independent	1. (Think about the design.)
2. Data in each sample are independent	2. SRSs and < 10% OR random allocation
3. Both populations are Normal	3. Both histograms are unimodal and symmetric*****

Distributions (Chi-Square)

- **Goodness of Fit** (df = #*cells* – 1; one variable, one sample compared to population model)

1. Data are counts	1. (Are they?)
2. Data in sample are independent	2. SRS and < 10% of the population
3. Sample is sufficiently large	3. All expected counts ≥ 5

- **Homogeneity** (df = $(r - 1)(c - 1)$; several groups compared on one variable)

1. Data are counts	1. (Are they?)
2. Data in groups are independent	2. SRSs and < 10% OR random allocation
3. Groups are sufficiently large	3. All expected counts ≥ 5

- **Independence** (df = $(r - 1)(c - 1)$; sample from one population classified on two variables)

1. Data are counts	1. (Are they?)
2. Data are independent	2. SRSs and < 10% of the population
3. Sample is sufficiently large	3. All expected counts ≥ 5

Regression (*t*, df = *n* – 2)

- **Association** between two measurement variables ($\beta = 0$?)

1. Form of relationship is linear	1. Scatterplot looks approx linear
2. Errors are independent	2. No apparent pattern in residuals plot
3. Variability of errors is constant	3. Residuals plot has consistent spread
4. Errors have a Normal model	4. Histogram of residuals is approximately unimodal and symmetric*****

(*****less critical as *n* increases)