Summary of Inference

Starting Point:

1. What are the observational units?
2. What are the variables?
3. Is each variable quantitative or categorical?

What Did You Measure?	Conditions for Inference	Population Parameter	Sample Statistic	R function	CI	Hypothesis Test
A Categorical Variable (or a count of how many in sample were in a certain category)	• Representative sample • Independent observational units • Count of how many in sample were in a certain category	p: proportion of population in a certain category	$\hat{p}$: proportion of sample in a certain category	`binom.test`	from R	from R
A Quantitative Variable	• Representative sample • Independent observational units • Quantitative variable • Mean is a good summary of the center (distribution is approximately unimodal and symmetric, no serious outliers)	μ: population mean	$\bar{x}$: sample mean	`t.test`	$\bar{x} \pm t^*_{n-1} s/\sqrt{n}$	test statistic: $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ If the null hypothesis is true, then $t \sim t_{n-1}$
2 Quantitative Variables (possibly also other variables)	Think of Robert the leprechaun: **R. O'LINE** • **R**epresentative sample • No **O**utliers • **L**inear relationship • **I**ndependent observational units • **N**ormally distributed residuals • **E**qual variance of residuals	β_0: Intercept of line describing population β_1: Slope of line describing population (possibly also other coefficients)	b_0: Intercept of line describing sample b_1: Slope of line describing sample (possibly also other coefficients)	`lm`	$b_0 \pm t^*_{n-k-1} SE(b_0)$ $b_1 \pm t^*_{n-k-1} SE(b_1)$ k = number of explanatory variables (1 for simple linear regression)	test statistic: $t = \frac{b_0-\beta_0^{null}}{SE(b_0)}$ or $t = \frac{b_1-\beta_1^{null}}{SE(b_1)}$ If the null hypothesis is true, then $t \sim t_{n-k-1}$