Inference Summary

How to Organize an Inference Problem: The Four-Step Process		
	Confidence intervals	**Significance tests**
STATE:	State the parameter you want to estimate and the confidence level.	State the hypotheses you want to test and the significance level, and define any parameters you use.
PLAN:	Identify the appropriate inference method and check conditions.	Identify the appropriate inference method and check conditions.
DO:	If the conditions are met, perform calculations.	If the conditions are met, perform calculations. • Give the sample statistic(s). • Calculate the standardized test statistic. • Find the P-value.
CONCLUDE:	Interpret your interval in the context of the problem.	Make a conclusion about the hypotheses in the context of the problem.

$$\text{Confidence interval} = \text{statistic} \pm (\text{critical value}) \cdot (\text{standard error of statistic})$$

$$\text{Standardized test statistic} = \frac{\text{statistic} - \text{parameter}}{\text{standard deviation (error) of statistic}}$$

Inference about	Number of samples/ groups	Interval or test (Section)	Name of procedure (TI-83/84 name) Formula	Conditions
Proportions	1	Interval **(8.2)**	One-sample z interval for p (1-PropZInt) $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	**Random:** The data come from a random sample from the population of interest. ○ **10%:** When sampling without replacement, $n < 0.10N$. **Large Counts:** *Interval:* Both $n\hat{p}$ and $n(1-\hat{p}) \geq 10$. *Test:* Both np_0 and $n(1-p_0) \geq 10$.
		Test **(9.2)**	One-sample z test for p (1-PropZTest) $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	
	2	Interval **(8.3)**	Two-sample z interval for $p_1 - p_2$ (2-PropZInt) $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	**Random:** The data come from two independent random samples or from two groups in a randomized experiment. ○ **10%:** When sampling without replacement, $n_1 < 0.10N_1$ and $n_2 < 0.10N_2$. **Large Counts:** *Interval:* The counts of successes and failures in each sample or group—$n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$ and $n_2(1-\hat{p}_2)$—are all ≥ 10. *Test:* The expected counts of successes and failures in each sample or group—$n_1\hat{p}_C$, $n_1(1-\hat{p}_C)$, $n_2\hat{p}_C$, $n_2(1-\hat{p}_C)$—are all ≥ 10.
		Test **(9.3)**	Two-sample z test for $p_1 - p_2$ (2-PropZTest) $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_1} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_2}}}$ where $\hat{p}_C = \frac{\text{total successes}}{\text{total sample size}} = \frac{X_1 + X_2}{n_1 + n_2}$	

Inference about	Number of samples/ groups	Interval or test (Section)	Name of procedure (TI-83/84 name) Formula	Conditions
Means	1	Interval **(10.1)**	One-sample t interval for μ (TInterval) $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$ df $= n - 1$	**Random:** The data come from a random sample from the population of interest. ○ **10%:** When sampling without replacement, $n < 0.10N$. **Normal/Large Sample:** The population has a Normal distribution or the sample size is large ($n \geq 30$). If the population distribution has unknown shape and $n < 30$, a graph of the sample data shows no strong skewness or outliers.
		Test **(11.1)**	One-sample t test for μ (T-Test) $t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}}$ df $= n - 1$	
	2	Interval **(10.2)**	Two-sample t interval for $\mu_1 - \mu_2$ (2-SampTInt) $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ df from technology or smaller of $n_1 - 1$, $n_2 - 1$	**Random:** The data come from two independent random samples or from two groups in a randomized experiment. ○ **10%:** When sampling without replacement, $n_1 < 0.10N_1$ and $n_2 < 0.10N_2$. **Normal/Large Sample:** For each sample, the corresponding population distribution (or the true distribution of response to the treatment) is Normal or the sample size is large ($n \geq 30$). For each sample, if the population (treatment) distribution has unknown shape and $n < 30$, a graph of the sample data shows no strong skewness or outliers.
		Test **(11.2)**	Two-sample t test for $\mu_1 - \mu_2$ (2-SampTTest) $t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ df from technology or smaller of $n_1 - 1$, $n_2 - 1$	
	Paired data	Interval **(10.2)**	Paired t interval for μ_{diff} (TInterval) $\bar{x}_{diff} \pm t^* \frac{s_{diff}}{\sqrt{n_{diff}}}$ df $= n_{diff} - 1$	**Random:** Paired data come from a random sample from the population of interest or from a randomized experiment. ○ **10%:** When sampling without replacement, $n_{diff} < 0.10N_{diff}$. **Normal/Large Sample:** The population distribution of differences (or the true distribution of differences in response to the treatments) is Normal or the number of differences in the sample is large ($n_{diff} \geq 30$). If the population (treatment) distribution of differences has unknown shape and $n_{diff} < 30$, a graph of the sample differences shows no strong skewness or outliers.
		Test **(11.2)**	Paired t test for μ_{diff} (T-Test) $t = \frac{\bar{x}_{diff} - \mu_0}{\frac{s_{diff}}{\sqrt{n_{diff}}}}$ df $= n_{diff} - 1$	

Inference about	Number of samples/ groups	Interval or test (Section)	Name of procedure (TI-83/84 name) Formula	Conditions
Distribution of a categorical variable	1	Test **(12.1)**	Chi-square test for goodness of fit (χ^2GOF-Test) $\chi^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$ df = number of categories − 1	**Random:** The data come from a random sample from the population of interest. ○ **10%:** When sampling without replacement, $n < 0.10N$. **Large Counts:** All expected counts at least 5
	2 or more	Test **(12.2)**	Chi-square test for homogeneity (χ^2-Test) $\chi^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$ df = (# of rows − 1)(# of columns − 1)	**Random:** Data from independent random samples or from groups in a randomized experiment. ○ **10%:** When sampling without replacement, $n_1 < 0.10N_1$, $n_2 < 0.10N_2$, and so on. **Large Counts:** All expected counts at least 5
Relationship between 2 categorical variables	1	Test **(12.2)**	Chi-square test for independence (χ^2-Test) $\chi^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$ df = (# of rows − 1)(# of columns − 1)	**Random:** The data come from a random sample from the population of interest. ○ **10%:** When sampling without replacement, $n < 0.10N$. **Large Counts:** All expected counts at least 5
Relationship between 2 quantitative variables (slope)	1	Interval **(12.3)**	*t* interval for the slope (LinRegTInt) $b \pm t^* (\text{SE}_b)$ with df $= n - 2$	**Linear:** The actual relationship between x and y is linear. For any particular value of x, the mean response μ_y falls on the population (true) regression line $\mu_y = \alpha + \beta x$. **Independent:** Individual observations are independent of each other. When sampling without replacement, check the *10% condition*, $n < 0.10N$. **Normal:** For any particular value of x, the response y varies according to a Normal distribution. **Equal SD:** The standard deviation of y (call it σ) is the same for all values of x. **Random:** The data come from a random sample from the population of interest or a randomized experiment.
		Test **(12.3)**	*t* test for the slope (LinRegTTest) $t = \frac{b - \beta_0}{\text{SE}_b}$ with df $= n - 2$	