

MATH 203: PRINCIPLES OF STATISTICS 1
FALL 2024
INFERENCE CHECKLIST

Here are the seven key features of an inference problem:

1. **Parameter:** μ or p ?
2. **Sample:** one sample or two sample ?
3. **Paired:** no or yes ?
4. **Sample size:** small sample or large sample ?
5. **Inference:** confidence interval or hypothesis test ?
6. **Level:** value of α ?
7. **Alternative hypothesis:** one-sided or two-sided H_a ?

A summary is given in the table on page 2.

Confidence Intervals: For parameter θ (which is either μ , p , $\mu_1 - \mu_2$ or $p_1 - p_2$), we base the confidence interval construction on the Z -transformed quantity

$$Z = \frac{\hat{\theta} - \theta}{\text{s.e.}(\hat{\theta})}$$

which we assume has either a t-distribution or a Normal distribution depending on the sample size. For these symmetric distributions, we specify that

$$P(-u_{\alpha/2} \leq Z \leq u_{\alpha/2}) = 1 - \alpha$$

for the upper $\alpha/2$ point, $u_{\alpha/2}$, which is equivalent to stating that

$$P\left(\hat{\theta} - u_{\alpha/2} \times \text{s.e.}(\hat{\theta}) \leq \theta \leq \hat{\theta} + u_{\alpha/2} \times \text{s.e.}(\hat{\theta})\right) = 1 - \alpha$$

which yields the confidence interval

$$\left(\hat{\theta} - u_{\alpha/2} \times \text{s.e.}(\hat{\theta}), \hat{\theta} + u_{\alpha/2} \times \text{s.e.}(\hat{\theta})\right)$$

Hypothesis Tests: We form the test statistic

$$T = \frac{\hat{\theta} - \theta_0}{\text{s.e.}(\hat{\theta})}$$

for the hypothesized value θ_0 specified by the null hypothesis. We then construct the appropriate Rejection Region

$$\{T \geq u_{\alpha}\} \quad \text{or} \quad \{T \leq -u_{\alpha}\} \quad \text{or} \quad \{|T| \geq u_{\alpha/2}\}$$

to match the alternative hypothesis.

	Sample	Paired	Sample size	CI	Z-transform	H_0	T	Null distn.
μ	One	No	Small	μ	$\frac{\bar{X} - \mu}{S/\sqrt{n}}$	$\mu = \mu_0$	$\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	t_{n-1}
			Large					$N(0, 1)$
	Two	No	Small	$\mu_1 - \mu_2$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}}$	$\mu_1 = \mu_2$	$\frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}}$	t_{m+n-2}
			Large					$N(0, 1)$
		Yes	Small	$\mu_d = \mu_1 - \mu_2$	$\frac{\bar{D} - \mu_d}{S_d/\sqrt{n}}$	$\mu_1 = \mu_2$	$\frac{\bar{D}}{S_d/\sqrt{n}}$	t_{n-1}
			Large					$N(0, 1)$
p	One	No	Large	p	$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}}$	$p = p_0$	$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$	$N(0, 1)$
	Two	No	Large	$p_1 - p_2$	$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{m} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n}}}$	$p_1 = p_2$	$\frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{\hat{p}(1 - \hat{p})}{m} + \frac{1}{n}}}$	$N(0, 1)$
$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ $S_d^2 = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2$ One sample: $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$ Two sample: $\hat{p}_1 = \frac{1}{m} \sum_{i=1}^m X_i, \hat{p}_2 = \frac{1}{n} \sum_{i=1}^n Y_i$ $\hat{p} = \frac{1}{n+m} \left(\sum_{i=1}^m X_i + \sum_{i=1}^n Y_i \right)$								