CH 10: Hypothesis Testing for Data from Two or More Samples Part 1

1. Case 4: Z-test for difference in Means $(\mu_1 - \mu_2)$ with both σ_1 and σ_2 known.

   (A) Concepts

   (B) The Test Statistic

   eq 10.5 : Test statistic for mean difference $\mu_1 - \mu_2$ (σ_1, σ_2 known): $Z_{cal} = \dfrac{(\overline{X_1} - \overline{X_2}) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

   (C) Assumptions for using this formula: the populations are normally distributed or the samples are large; the two samples are randomly and independently drawn.

   (D) We can draw our conclusion either based on the critical value approach or the p-value approach.

   (E) For two-tailed test: $p - value = 2 * P(Z > |Z_{cal}|)$;
   upper, one-tail test: $p - value = P(Z > Z_{cal})$,
   lower, one-tail test: $p - value = P(Z < Z_{cal})$

   EX 1 Given two independent samples, a sample of size $n_1 = 40$ from a population 1 with known standard deviation $\sigma_1 = 20$ is selected and resulting in a sample mean of $\overline{X_1} = 72$; another sample of size $n_2 = 50$ from population 2 with known standard deviation $\sigma_2 = 10$ is also selected and the sample mean $\overline{X_2} = 66$. Test if the average for population 1 is more than the average for population 2 ($\alpha = 0.025$).

   Step 1: State H_0 vs. H_1.

   Step 2: Compute the test statistic

   Step 3: Make a decision using either the p-value approach or the critical value approach.

2. Case 5: t-test for difference in Means $(\mu_1 - \mu_2)$ with both σ_1 and σ_2 unknown.

   (A) Concepts

   (B) Compute the test statistic

   eq10.8: Test statistic for Mean difference $\mu_1 - \mu_2$ (σ_1, σ_2 unknown): $t_{cal} = \frac{(\overline{X_1} - \overline{X_2}) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

   with eq10.7: $df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{1}{n_1-1}(\frac{s_1^2}{n_1})^2 + \frac{1}{n_2-1}(\frac{s_2^2}{n_2})^2}$

   Note: we round the noninterger degrees of freedom down.

EX 2 Comparing the lifetimes of two brands of batteries, a researcher has randomly selected 20 batteries of brand A with $\overline{X}_A = 22.5$ months and $S_A = 2.5$ months and 30 batteries from brand B with $\overline{X}_B = 20.1$ months and $S_B = 4.8$ months . Test if the means are different ($\alpha = 0.05$)

Step 1: State H_0 vs. H_1.

Step 2: Compute the test statistic and df

Step 3: Make a decision using either the p value approach or the critical value approach.

CH 10: Hypothesis Testing for Data from Two or More Samples Part 2

3. Case 6: t-test for difference in two related samples μ_d

   (A) Basic Concept and Data Structure

   (B) Test Statistic

   eq 10.9 Test statistic for mean difference (related samples): $t_{cal} = \dfrac{\overline{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}$

   (with $(n-1)$ degrees of freedom)

   (C) Hypothesis Testing

   Step 1: State H_0 vs. H_1.

   Step 2: Compute the test statistic

   Step 3: Make a decision using either the p-value approach or the critical value approach.

EX 3 Given a set of matched pair of data, test if the mean has been changed (use $\alpha = 0.05$).

Step 1: State H_0 vs. H_1.

Step 2: Compute the test statistic

Step 3: Make a decision using either the p-value approach or the critical value approach.

4. Case 7: Z-test for the Difference Between Two Proportions $p_1 - p_2$

(A) Basic Concepts

(B) The Test Statistic

eq10.16: Test statistic for the difference between two proportions $Z_{cal} = \frac{(\overline{p}_1 - \overline{p}_2)}{\sqrt{\overline{p}(1-\overline{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$,

where eq 10.15: $\overline{p} = \frac{n_1\overline{p}_1 + n_2\overline{p}_2}{n_1 + n_2}$

EX 4 Auto company suspects that singles have more claims than married policyholders. Let the single policyholder be population 1 and married policyholder be population2. If a random survey indicates that 76 out of 400 single and 90 out of 900 married policyholders did auto claim last year, test the theory with $\alpha = 0.05$.

Step 1: State H_0 vs. H_1.

Step 2: Compute the test statistic

Step 3: Make a decision using either the p-value approach or the critical value approach.