Section 8-6 Inferences about Two Means: Independent and Small Samples

Inferences about Two Means: Independent and Small Samples

Assumptions

1. The two samples are independent.
2. The two samples are simple random samples from normally distributed populations.
3. At least one of the two samples is small ($n \le 30$).

Apr 10-2:48 PM

Three Different Procedures

Case 1: The values of both population variances are known. (This case seldom occurs.)

Case 2: The two populations appear to have equal variances. (That is, $\sigma_1^2 = \sigma_2^2$)

Case 3: The two populations appear to have unequal variances. (That is, $\sigma_1^2 \neq \sigma_2^2$)

Apr 10-2:51 PM

Case 2: The Two Populations Appear to have Equal Variances

- A pooled estimate of the variance σ^2 that is common to both populations, denoted by s_p^2, is calculated.
- s_p^2 is a weighted average of s_1^2 and s_2^2

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

where $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$

degrees of freedom $df = n_1 + n_2 - 2$

Apr 10-2:51 PM

Case 2: Confidence Interval

(Small Independent Samples and Equal Variances)

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

where $E = t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$

and s_p^2 is as given in the test statistic

degrees of freedom $df = n_1 + n_2 - 2$

Apr 10-2:55 PM

Case 3: Test Statistic

(Small Independent Samples and Unequal Variances)

An approximate method

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where df = smaller of $n_1 - 1$ and $n_2 - 1$

Apr 10-2:57 PM

Case 3: Confidence Interval

(Small Independent Samples and Unequal Variances)

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

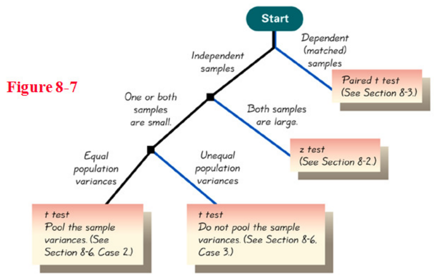
where $E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

and df = smaller of $n_1 - 1$ and $n_2 - 1$

Apr 10-2:59 PM

Inferences about the Means of Two Populations

Figure 8-7



Apr 10-3:02 PM

Ex1)Use a .05 significance level to test the claim that the mean amount of nicotine in filtered king size cigarettes is equal to the mean amount of nicotine in nonfiltered king size cigarettes. Assume that both cigarettes have equal variances.

filtered	nonfiltered
$n_1 = 21$	$n_1 = 8$
$\bar{x}_1 = .94$	$\bar{x}_1 = 1.65$
$s_1 = .31$	$s_1 = .16$

Use Case #2 because variances are equal.

1. $\mu_1 = \mu_2$ (filtered = nonfiltered)
   $\mu_1 \neq \mu_2$

2. Test Statistics: $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$

$$\frac{20(.31)^2 + 7(.16)^2}{20 + 7} = .078$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \qquad \frac{.94 - 1.65}{\sqrt{\frac{.078}{21} + \frac{.078}{8}}} = -6.119$$

3. Critical Value: df = 21 + 8 - 2 = 27

use table A-3 2 tailed $\alpha = .05$ ± 2.052

4. Reject H_0: There is not sufficient evidence to support that they are equal.

Apr 11-8:16 AM

Ex1b) Make a 95% confidence interval to estimate $\mu_1 - \mu_2$.

filtered	nonfiltered
$n_1 = 21$	$n_1 = 8$
$\bar{x}_1 = .94$	$\bar{x}_1 = 1.65$
$s_1 = .31$	$s_1 = .16$

$s_p^2 = .078$

degrees of freedom df = $n_1 + n_2 - 2$

$$E = t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} \qquad E = 2.052\sqrt{\frac{.078}{21} + \frac{.078}{8}}$$

E = .238

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

.94 - 1.65 = -.71

$-.71 - .238 < \mu_1 - \mu_2 < -.71 + .238$

$-.95 < \mu_1 - \mu_2 < -.47$

95% confident that the nonfiltered king size cigarettes have mean nicotine content that exceeds the mean for filtered kings by an amount that is between .47 mg and .95 mg.

Apr 11-9:34 AM

Ex2a)

EXAMPLE Cigarette Filters and Tar

Use a 0.05 significance level to test the claim that the mean amount of tar in filtered king-size cigarettes is *less than* the mean amount of tar in nonfiltered king-size cigarettes. (All measurements are in milligrams, and the data are from the Federal Trade Commission.) Assume unequal variances.

Use Case 3 since told variances aren't equal.

Tar (mg)	
Filtered Kings	Nonfiltered Kings
$n_1 = 21$	$n_2 = 8$
$\bar{x}_1 = 13.3$	$\bar{x}_2 = 24.0$
$s_1 = 3.7$	$s_2 = 1.7$

1. H_0: $\mu_1 \geq \mu_2$ (or $\mu_1 - \mu_2 \geq 0$)
   H_1: $\mu_1 < \mu_2$ (original claim)

2. Test Statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(13.3 - 24.0) - 0}{\sqrt{\frac{3.7^2}{21} + \frac{1.7^2}{8}}} = -10.630$$

3. Critical Value: df = smaller Table A-3

$\alpha = 0.05$ (one tail)
df = 7 $t = -1.895$

4. Reject H_0: There is sufficient evidence that $\mu_1 < \mu_2$ or that filtered cigarettes have less tar than nonfilteredeed.

Apr 11-9:42 AM

2b) **EXAMPLE Cigarette Filters and Tar** Using the data in the preceding example, construct a 95% confidence interval estimate of $\mu_1 - \mu_2$.

df = 7 and use 2 tailed test for critical value:

$$E = t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.365\sqrt{\frac{3.7^2}{21} + \frac{1.7^2}{8}} = 2.381$$

$$\bar{x}_1 - \bar{x}_2 = 13.3 - 24.0 = -10.7$$

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$
$$-10.7 - 2.381 < (\mu_1 - \mu_2) < -10.7 + 2.381$$
$$-13.1 < (\mu_1 - \mu_2) < -8.3$$

INTERPRETATION Again, this final format is somewhat awkward with its negative signs, so we might interpret it by saying that we are 95% confident that the mean tar content for nonfiltered kings exceeds the mean tar content for filtered kings by an amount that is between 8.3 mg and 13.1 mg.

Apr 11-1:51 PM