10.1 Sampling Distribution of Differences between 2 independent sample means

GOALS:

1. Consider how two samples can be compared to determine if they are the same or different or come from the same or different populations.
2. Consider the distribution of the difference of sample means, including:
   - the mean of the difference
   - the standard deviation of the difference

Read Ch. 10.1, Study Key Fact 10.1

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

G: * random samples of 30 males and 30 females
* tested frame of reference, pointed S, error recorded
* table of pointing errors, in degrees, with:

MALE					FEMALE				
13	68	60	22	30	14	78	18	32	80
130	18	5	70	8	8	69	35	35	91
39	3	9	58	20	20	111	111	12	68
33	11	59	3	67	3	3	109	27	66
10	38	5	167	26	138	128	36	8	176
13	23	86	15	19	122	31	27	3	15

$\overline{x}_M = 37.6$ $\overline{x}_F = 55.8$

$s_M = 38.5$ $s_F = 48.3$

F: At the 1% significance level, do the data provide sufficient evidence to conclude that, on average, males have a better sense of direction and, in particular, a better frame of reference than females?

from 10.2

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

How do we start?

Know that mean for males is ~ nd for large samples.
Know that mean for females is ~ nd for large samples.

How do we compare them?

See that: mean for males < mean for females

or $\overline{X}_M < \overline{X}_F$

or $\overline{X}_M - \overline{X}_F < 0$

We need to examine the distribution of

$\overline{X}_M - \overline{X}_F$

In general: the distribution of $\overline{X}_1 - \overline{X}_2$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

the distribution of $\overline{X}_1 - \overline{X}_2$

If x is ~ nd on each of the populations 1 and 2, then $\overline{X}_1 - \overline{X}_2$ is ~nd and:

$$\mu_{\overline{x}_1 - \overline{x}_2} = \mu_1 - \mu_2$$

$$\sigma_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

the distribution of $\overline{X}_1 - \overline{X}_2$

If x is ~ nd on each of the populations 1 and 2, then $\overline{X}_1 - \overline{X}_2$ is ~nd and:

$$\mu_{\overline{x}_1 - \overline{x}_2} = \mu_1 - \mu_2$$

$$\sigma_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

comparison

Mean of the Sample Mean

$$\mu_{\overline{x}} = \mu$$

Standard Deviation of the Sample Mean

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Standard Error (of the Mean)

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}}$$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

When σ is known, can standardize to SNC,

But **rarely have a known σ, so will not consider a z test**

$$z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Instead, either of two *t*-tests: σ_1, σ_2 unknown

- $\sigma_1 = \sigma_2$ --> pooled ***t***
  s_1 and s_2 are estimates of same σ
- $\sigma_1 \neq \sigma_2$ --> non-pooled ***t***
  s_1 and s_2 are not known to estimate same σ

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes

10.1 Sampling Distribution of Differences between 2 independent sample means

When σ is known, can standardize to SNC,

But **rarely have a known σ, so will not consider a z test**

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

No

Yes

Instead, either of two *t*-tests: σ_1, σ_2 unknown

- $\sigma_1 = \sigma_2$ --> pooled ***t***

  s_1 and s_2 are estimates of same σ

- $\sigma_1 \neq \sigma_2$ --> non-pooled ***t***

  s_1 and s_2 are not known to estimate same σ

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page

Class Notes