Module 3

General Sir John Kotelawala Defense University

Department of Mathematics

MA 3102 APPLIED STATISTICS

CONFIDENCE INTERVALS FOR THE DIFFERENCE OF TWO MEANS (σ^2 *Known*)

- Suppose that we are interested in comparing the means of two normal distributions.
- Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ be, respectively, two independent random samples of sizes n and m from the two normal distributions $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$.
- Suppose, for now, that σ_X^2 and σ_X^2 are known.
- The random samples are independent; thus, the respective sample means $\bar{X}$ and $\bar{Y}$ are also independent and have distributions

$$\bar{X} \sim N(\mu_X, {\sigma_X^2}/{n}) \ and \qquad \bar{Y} \sim N(\mu_Y, {\sigma_Y^2}/{m})$$

- *Let's say* $W = \bar{X} - \bar{Y}$
- Consequently, the distribution of

$W = \bar{X} - \bar{Y}$ is $N(\mu_X - \mu_Y, {\sigma_X^2}/{n} + {\sigma_Y^2}/{m})$

$$\sigma_W = \sqrt{{\sigma_X^2}/{n} + {\sigma_Y^2}/{m}}$$

- Once the experiments have been performed and the means of $\bar{x}$ *and* $\bar{y}$ are computed, the confidence intervals are

EXAMPLE

Based on earlier discussion

n=15, m =8, $\bar{x} = 70$, $\bar{y} = 75.3, \sigma_X^2 = 60, \sigma_Y^2 = 40$

Find 90 % confidence interval on difference.

$$\sigma_W = \sqrt{\sigma_X^2/_n + \sigma_Y^2/_m} = \sqrt{\frac{60}{15} + \frac{40}{8}}$$

The CI on $\bar{x}$ *and* $\bar{y}$ are computed as

$$\left[(70 - 75.3) - 1.645\sqrt{\frac{60}{15} + \frac{40}{8}}, (70 - 75.3) + 1.645\sqrt{\frac{60}{15} + \frac{40}{8}}\right]$$

[-10.135, -0.265]

is a 90% confidence interval for $\mu_X - \mu_Y$. Because the confidence interval does not include zero, we suspect that μ_Y is greater than μ_X.

If the sample sizes are large and σ_X and σ_Y are unknown, we can replace σ_X^2 an σ_Y^2 with s_x^2 and s_y^2, where s_x^2 and s_y^2 are the values of the respective unbiased estimate of the variances. This means that

$$\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

serves as an approximate $100(1-\alpha)\%$ confidence interval for $\mu_X - \mu_Y$.

CONFIDENCE INTERVALS FOR THE DIFFERENCE OF TWO MEANS (σ^2 *Unknown*)

- Consider a problem construction confidence interval for difference in means of two normal distributions when **variances are unknown** but the sample sizes are small.

- Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ be, respectively, two independent random samples of sizes n and m from the two normal distributions $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$ respectively. If the sample sizes are not large (Say <30), this problem can be difficult.

- If we can assume common but unknown variances (say $\sigma^2 = \sigma_X^2 = \sigma_Y^2$)
- The $CI\ on\ \mu_X - \mu_Y$ is

$$(\bar{X} - \bar{Y}) \mp t_{\alpha/2,(m+n-2)} \sqrt{\left[\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}\right]\left[\frac{1}{n} + \frac{1}{m}\right]}$$

- The pooled estimator of the common standard deviation $S_{p.}$

$$S_p = \sqrt{\left[\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}\right]}$$

Example

Suppose that scores on a standardized test in mathematics taken by students from large and small high schools are $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$, respectively, where σ^2 is unknown. If a random sample of $n = 9$ students from large high schools yielded $\overline{x} = 81.31$, $s_x^2 = 60.76$, and a random sample of $m = 15$ students from small high schools yielded $\overline{y} = 78.61$, $s_y^2 = 48.24$, then the endpoints for a 95% confidence interval for $\mu_X - \mu_Y$ are given by

$$81.31 - 78.61 \pm 2.074\sqrt{\frac{8(60.76) + 14(48.24)}{22}}\sqrt{\frac{1}{9} + \frac{1}{15}}$$

because $t_{0.025}(22) = 2.074$. The 95% confidence interval is $[-3.65, 9.05]$. ■

Example

7.2-3. Independent random samples of the heights of adult males living in two countries yielded the following results: $n = 12$, $\overline{x} = 65.7$ inches, $s_X = 4$ inches and $m = 15$, $\overline{y} = 68.2$ inches, $s_y = 3$ inches. Find an approximate 98% confidence interval for the difference $\mu_X - \mu_Y$ of the means of the populations of heights. Assume that $\sigma_X^2 = \sigma_Y^2$.

Example

7.2-7. An automotive supplier is considering changing its electrical wire harness to save money. The idea is to replace a current 20-gauge wire with a 22-gauge wire. Since not all wires in the harness can be changed, the new wire must work with the current wire splice process. To determine whether the new wire is compatible, random samples were selected and measured with a pull test. A pull test measures the force required to pull the spliced wires apart. The minimum pull force required by the customer is 20 pounds. Twenty observations of the forces needed for the current wire are

28.8	24.4	30.1	25.6	26.4	23.9	22.1	22.5	27.6	28.1
20.8	27.7	24.4	25.1	24.6	26.3	28.2	22.2	26.3	24.4

Twenty observations of the forces needed for the new wire are

14.1	12.2	14.0	14.6	8.5	12.6	13.7	14.8	14.1	13.2
12.1	11.4	10.1	14.2	13.6	13.1	11.9	14.8	11.1	13.5

(a) Does the current wire meet the customer's specifications?

(b) Find a 90% confidence interval for the difference of the means for these two sets of wire.

(d) What is your recommendation for this company?

TABLE A-3 *t* Distribution: Critical *t* Values

			Area in One Tail		
	0.005	0.01	0.025	0.05	0.10
			Area in Two Tails		
Degrees of Freedom	0.01	0.02	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316