Topic: Confidence interval for the difference of means

Question: A professor is interested in whether exam scores differ between two nearby colleges. He selects a simple random sample of 20 students each from both colleges and finds a mean test score of 350 with a standard deviation of 15 at the first college, and a mean test score of 390 with a standard deviation of 30 at the second college. Assuming exam scores are normally distributed at both colleges, find a 95% confidence interval around the difference in exam scores.

Answer choices:

A $(-55.52, -24.48)$

B $(-55.39, -24.61)$

C $(24.64, 55.36)$

D $(24.61, 55.39)$

Solution: D

We don't know the population standard deviations, and the sample sizes are smaller than 30. The sample variances are $s_1^2 = 30^2 = 900$ and $s_2^2 = 15^2 = 225$. We assign the populations this way so that we end up with a positive result for the difference of means.

Not only were the samples taken from different populations, but 900 is more than double 225, so we can conclude that we're working with unequal population variances.

Therefore, our confidence interval formula will be

$$(a, b) = (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{with df} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Let's start with degrees of freedom.

$$\text{df} = \frac{\left(\frac{900}{20} + \frac{225}{20}\right)^2}{\frac{1}{20 - 1}\left(\frac{900}{20}\right)^2 + \frac{1}{20 - 1}\left(\frac{225}{20}\right)^2}$$

$$\text{df} = \frac{\left(\frac{1{,}125}{20}\right)^2}{\frac{1}{19}\left(\frac{900}{20}\right)^2 + \frac{1}{19}\left(\frac{225}{20}\right)^2}$$

$$\text{df} = \frac{\frac{1{,}265{,}625}{400}}{\frac{810{,}000}{7{,}600} + \frac{50{,}625}{7{,}600}}$$

$$\text{df} = \frac{1{,}265{,}625}{400}\left(\frac{7{,}600}{860{,}625}\right)$$

$$\text{df} \approx 27.94$$

Rounding down to the nearest whole number in order to keep the estimate conservative, we find 27 degrees of freedom. Together with a $95\,\%$ confidence level, the t-table gives $t_{\alpha/2} = 2.052$.

	Upper-tail probability p									
df	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	**2.052**	2.473	2.771	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
	50%	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.8%**	**99.9%**
	Confidence level C									

Then the confidence interval will be

$$(a, b) = (390 - 350) \pm 2.052\sqrt{\frac{30^2}{20} + \frac{15^2}{20}}$$

$$(a, b) = 40 \pm 2.052\sqrt{\frac{900}{20} + \frac{225}{20}}$$

$$(a,b) = 40 \pm 2.052\sqrt{\frac{1{,}125}{20}}$$

$$(a,b) = 40 \pm 2.052(7.5)$$

$$(a,b) = 40 \pm 15.39$$

Therefore, we can say that the confidence interval is

$$(a,b) = (40 - 15.39, 40 + 15.39)$$

$$(a,b) = (24.61, 55.39)$$

So we can say that we're 95% confident that the true difference between mean exam scores is between 24.61 and 55.39.

Topic: Confidence interval for the difference of means

Question: Two college directors want to determine whether there's a difference in the amount that their students spend annually on textbooks. They sample 200 students from college A and 230 from college B and find mean spends of $\bar{x}_A = \$1{,}258$ and $\bar{x}_B = \$1{,}150$. Assuming both populations are normally distributed with $\sigma_A = \$52$ and $\sigma_B = \$64$, find a $99\,\%$ confidence interval around the difference in annual textbook spend.

Answer choices:

A $(-122.44, -93.56)$

B $(97.03, 118.97)$

C $(93.56, 122.44)$

D $(93.35, 122.65)$

Solution: C

Because population standard deviations are known, our confidence interval formula will be

$$(a,b)=(\bar{x}_1-\bar{x}_2)\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

At 99% confidence, we use a critical z-value of $z_{\alpha/2}=2.58$. Then the confidence interval will be

$$(a,b)=(1{,}258-1{,}150)\pm 2.58\sqrt{\frac{52^2}{200}+\frac{64^2}{230}}$$

$$(a,b)=108\pm 2.58\sqrt{\frac{2{,}704}{200}+\frac{4{,}096}{230}}$$

$$(a,b)=108\pm 2.58\sqrt{\frac{338}{25}+\frac{2{,}048}{115}}$$

$$(a,b)=108\pm 2.58\sqrt{\frac{18{,}014}{575}}$$

$$(a,b)\approx 108\pm 2.58(5.597)$$

$$(a,b)\approx 108\pm 14.441$$

Therefore, we can say that the confidence interval is

$$(a,b)\approx(108-14.441,108+14.441)$$

$$(a,b)\approx(93.559,122.441)$$

So we can say that we're 99% confident that the true difference between mean textbook spend is between 93.559 and 122.441.

Topic: Confidence interval for the difference of means

Question: The owners of two restaurants on the same street are interested in whether or not their daily earnings differ. They take simple random samples of earnings over 15 days, and find mean daily earnings of $\$1,365$ with a standard deviation of $\$48$ for the first restaurant, and mean daily earnings of $\$1,230$ with a standard deviation of $\$28$ for the second restaurant. Assuming daily earnings at both restaurants follow a normal distribution, find a $90\,\%$ confidence interval around the difference in daily earnings.

Answer choices:

A $(111.33,158.67)$

B $(134.38,135.62)$

C $(116.16,153.84)$

D $(110.36,159.64)$

Solution: D

Because population standard deviations are unknown, and because we have small samples $n_1, n_2 < 30$, we'll need to use a critical t-value instead of a critical z-value.

Our samples were taken from different populations, and one sample variance is more than twice the other, $48^2 = 2{,}304 > 2(28^2) = 2(784) = 1{,}568$, so we'll say that the population variances are unequal.

Find the number of degrees of freedom.

$$\text{df} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

$$\text{df} = \frac{\left(\frac{48^2}{15} + \frac{28^2}{15}\right)^2}{\frac{1}{15 - 1}\left(\frac{48^2}{15}\right)^2 + \frac{1}{15 - 1}\left(\frac{28^2}{15}\right)^2}$$

$$\text{df} = \frac{\left(\frac{48 \cdot 48 + 28 \cdot 28}{15}\right)^2}{\frac{1}{14}\left(\frac{48 \cdot 48 \cdot 48 \cdot 48}{15 \cdot 15}\right) + \frac{1}{14}\left(\frac{28 \cdot 28 \cdot 28 \cdot 28}{15 \cdot 15}\right)}$$

$$\text{df} = \frac{\frac{(48 \cdot 48 + 28 \cdot 28)(48 \cdot 48 + 28 \cdot 28)}{15 \cdot 15}}{\frac{24 \cdot 48 \cdot 48 \cdot 48}{7 \cdot 15 \cdot 15} + \frac{2 \cdot 28 \cdot 28 \cdot 28}{15 \cdot 15}}$$

$$\text{df} = \frac{\frac{(48 \cdot 48 + 28 \cdot 28)(48 \cdot 48 + 28 \cdot 28)}{15 \cdot 15}}{\frac{24 \cdot 48 \cdot 48 \cdot 48 + 2 \cdot 7 \cdot 28 \cdot 28 \cdot 28}{7 \cdot 15 \cdot 15}}$$

$$\text{df} = \frac{(48 \cdot 48 + 28 \cdot 28)(48 \cdot 48 + 28 \cdot 28)}{15 \cdot 15}\left(\frac{7 \cdot 15 \cdot 15}{24 \cdot 48 \cdot 48 \cdot 48 + 2 \cdot 7 \cdot 28 \cdot 28 \cdot 28}\right)$$

$$\text{df} = \frac{7(48 \cdot 48 + 28 \cdot 28)(48 \cdot 48 + 28 \cdot 28)}{24 \cdot 48 \cdot 48 \cdot 48 + 2 \cdot 7 \cdot 28 \cdot 28 \cdot 28}$$

$$\text{df} = \frac{7(48^4 + 2(28^2 48^2) + 28^4)}{24(48^3) + 2(7)(28^3)}$$

$$\text{df} \approx 22.54$$

Rounding down to the nearest whole number in order to keep the estimate conservative, we find 22 degrees of freedom. Together with a 90% confidence level, the t-table gives $t_{\alpha/2} = 1.717$.

	Upper-tail probability p									
df	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.686	0.858	1.061	1.321	**1.717**	2.074	2.508	2.819	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
	50%	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.8%**	**99.9%**
	Confidence level C									

Then the confidence interval will be

$$(a, b) = (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(a, b) = (1{,}365 - 1{,}230) \pm 1.717\sqrt{\frac{48^2}{15} + \frac{28^2}{15}}$$

$$(a, b) = 135 \pm 1.717\sqrt{\frac{2{,}304}{15} + \frac{784}{15}}$$

$$(a, b) = 135 \pm 1.717\sqrt{\frac{3{,}088}{15}}$$

$$(a, b) \approx 135 \pm 24.64$$

Therefore, we can say that the confidence interval is

$$(a, b) \approx (135 - 24.64, 135 + 24.64)$$

$$(a, b) \approx (110.36, 159.64)$$

Based on the confidence interval, we're 90% confident that the true difference between the mean daily earnings of the two restaurants is between $\$110.36$ and $\$159.64$.