Names:
Math 321 | Differences | April 13, 2016

Often we'll want to compare two populations. For example, "Effects of Fast-Food Consumption on Energy Intake and Diet Quality Among Children in a National Household Study" gives the following summary data on daily calorie intake for a sample of teens divided into two groups based one whether or not they usually eat fast food.

Eat Fast Food?	Sample Size	Sample Mean (Cal)	Sample SD (Cal)
No	663	2258	1519
Yes	413	2637	1138

The question is whether the difference in the means of these two groups is evidence in differences between the population of fast-food eaters and population of fast-food abstainers. Fortunately you know enough math to answer this question.

Theorem. *If X and Y are independent normally distributed random variables, then $X - Y$ is also a normally distributed random variable.*

1. Let $\overline{X}$ be the mean of a random sample of size n_1 from a population with mean μ_1 and variance σ_1^2 and let $\overline{Y}$ be the mean of a random sample of size n_2 from a population with mean μ_2 and variance σ_2^2.

a) What are the the mean and variance of $\overline{X} - \overline{Y}$?

b) If n_1 and n_2 are large, then the Central Limit Theorem tells us that $\overline{X}$ and $\overline{Y}$ are each approximately normally distributed. Use this (and the theorem above) to develop a $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$.

c) Calculate a 99% confidence lower bound for the difference between the two population means in the example above.

In the last problem your analysis was based on a normal distribution (I hope), which is reasonable because sample sizes were large. Smaller samples from normally distributed populations instead give us a t distribution. If the two population are normally distributed and have the same variance, then the following random variable has a t distribution with $n_1 + n_2 - 2$ degrees of freedom:

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$. Note: our book has a different version suitable for use when the populations have different variances.

2. The paper *Effects of Exercise Modality on Insulin Resistance and Functional Limitation in Older Adults* (Lance E. Davidson et al. Arch Intern Med. 2009;169(2):122-131) describes the results of a randomized controlled trial: "136 sedentary, abdominally obese older men and women" were divided randomly into different groups and each group was given a different exercise regimen for 6 months. At the end of the study the 24 members of the control group (no extra exercise) had a mean BMI of 0.10 with a standard error of 0.13 (measurements have been adjusted for baseline value, age, and sex; standard error is $s/\sqrt{n}$). The 30 members of the aerobic exercise group finished with a mean BMI of -0.96 with a standard error of 0.12. Use a t distribution to find a 95% confidence interval for the difference between the two population means associated with this study.

The statistics you have been working with can also be used to test hypotheses. For example, the hypothesis $H_0 : \mu_1 - \mu_2 = \delta_0$ can be tested using the statistic

$$t = \frac{\overline{x} - \overline{y} - \delta_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

with s_p as given on the last page. The complete findings of the study in problem 2 are summarized in this table:

Table 4. Change in Primary and Secondary Outcome Variables

	Study Group, Mean (SE)[a]			
Characteristic	**Control (n=24/ITT=28)**	**Resistance Exercise (n=30/ITT=36)**	**Aerobic Exercise (n=30/ITT=37)**	**Combined Exercise (n=33/ITT=35)**
Anthropometric				
Body weight, kg	0.28 (0.37)	−0.64 (0.37)	−2.77 (0.33)[b,c]	−2.31 (0.33)[b, c]
BMI	0.10 (0.13)	−0.26 (0.12)	−0.96 (0.12)[b,c]	−0.84 (0.12)[b, c]
Waist circumference, cm	−0.28 (0.53)	−3.18 (0.49)[b]	−5.08 (0.46)[b,c]	−4.61 (0.47)[b]
MRI				
Skeletal muscle, kg	−0.01 (0.18)	0.97 (0.20)[b,d]	−0.06 (0.19)	0.62 (0.15)[b, d]
Upper body muscle, kg	−0.07 (0.11)	0.61 (0.16)[b,d]	−0.31 (0.11)	0.38 (0.09)[b, d]
Lower body muscle, kg	0.05 (0.11)	0.39 (0.12)	0.24 (0.12)	0.24 (0.10)
Total fat, kg	−0.52 (0.38)	−1.56 (0.36)	−3.03 (0.39)[b,c]	−3.38 (0.34)[b, c]
Total abdominal fat, kg	−0.05 (0.12)	−0.42 (0.11)	−0.84 (0.13)[b]	−0.76 (0.10)[b]
Visceral fat, kg	0.02 (0.06)	−0.21 (0.06)	−0.43 (0.08)[b]	−0.35 (0.05)[b]
Abdominal subcutaneous fat, kg	−0.04 (0.07)	−0.21 (0.07)	−0.40 (0.07)[b]	−0.40 (0.06)[b]
Fat to muscle ratio	−0.02 (0.02)	−0.12 (0.01)[b]	−0.13 (0.02)[b]	−0.19 (0.02)[b-d]
Metabolic				
Fasting insulin level, μIU/mL	−0.29 (0.63)	−0.97 (0.57)	−1.43 (0.63)	−1.49 (0.53)
Insulin resistance, M/I	0.29 (1.59)	1.84 (1.29)	6.51 (1.27)[b,c]	9.22 (1.33)[b, c]
Functional limitation[e]				
No. of chair stands	0.36 (0.56)	4.3 (0.56)[b]	4.00 (0.50)[b]	5.89 (0.52)[b, d]
No. of arm curls	0.15 (0.84)	6.2 (0.79)[b]	5.11 (0.67)[b]	7.85 (0.69)[b, d]
2-min step test, No. of steps	1.77 (2.35)	19.58 (2.34)[b]	17.00 (1.99)[b]	26.88 (2.02)[b, d]
8-ft up-and-go, s	−0.05 (0.08)	−0.56 (0.08)[b]	−0.45 (0.07)[b]	−0.60 (0.07)[b]
Combined (all 4 tests) improvement, *z* score	−1.01 (0.12)	0.17 (0.12)[b]	−0.01 (0.10)[b]	0.52 (0.10)[b, d]

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); ITT, intent-to-treat; M/I, rate of glucose uptake per unit of insulin per kilogram of skeletal muscle per minute × 100; MRI, magnetic resonance imaging; &Vdot;o$_{2max}$, maximum oxygen consumption.

SI conversion factor: To convert insulin to picomoles per liter, multiply by 6.945.

[a] Adjusted for baseline value, age, and sex. Pairwise comparisons among groups were tested for statistical significance using Tukey studentized range tests.

[b] Significant preintervention vs postintervention treatment differences compared with the control group ($P < .05$).

[c] Significant preintervention vs postintervention treatment differences compared with the resistance exercise group ($P < .05$).

[d] Significant preintervention vs postintervention treatment differences compared with the aerobic exercise group ($P < .05$).

[e] Tests are described in the "Measurement of Functional Limitation and Cardiorespiratory Fitness" subsection of the "Methods" section.

3. Test $H_0 : \mu_1 - \mu_2 = 0$ against $H_1 : \mu_1 - \mu_2 \neq 0$ for at least one of the significant differences (noted with a b, c or d in the table) and calculate the associated P-value.