AP Statistics – Inference with Two Sample Means

THE BASEBALL PROBLEM American League baseball teams play their games with the designated hitter rule, meaning that pitchers do not bat. The league believes that replacing the pitcher, traditionally a weak hitter, with another player in the batting order produces more runs and generates more interest among fans. Below are the average numbers of runs scored in American League and National League stadiums for the 2006 season.

American League:

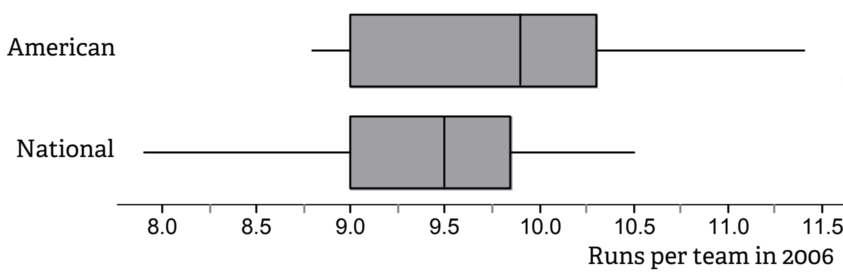
11.4	10.5	10.4	10.3	10.2	10.0	9.9	9.9	9.7	9.1	9.0	9.0	8.9	8.8

Mean: 9.7929 Standard Deviation: 0.7580 Number of teams: 14

National League:

10.5	10.3	10.0	10.0	9.7	9.7	9.6	9.5	9.5	9.4	9.1	9.0	9.0	8.9	8.9	7.9

Mean: 9.4375 Standard Deviation: 0.6386 Number of teams: 16



create BOTH graphs (check for outliers)

a) Construct and interpret a 95% confidence interval for the mean difference in runs per game between the two leagues.

2 – sample t-interval

OPTIONAL

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

$$(9.7929 - 9.4375) \pm t^* \sqrt{\frac{0.758^2}{14} + \frac{0.639^2}{16}}$$

$$(-0.286 \,,\, +0.996)$$

Conditions:

- We must consider the data for # of runs to be representative for each league.
- The two samples are reasonably independent of each other.
- Nearly normal: The boxplots for each league show no major departures from normality. (make sure you sketch both graphs!)

We are 95% confident that the true mean difference in runs per game (American – National) is between -0.286 and 0.996.

b) Carefully interpret the meaning of the 95% confidence level in context.

If we used this method maaaany times, about 95% of the resulting intervals would contain the true mean difference for the # of runs/game between the 2 leagues.

AP Statistics – Inference with Two Sample Means

"EACH DAY I AM GETTING BETTER IN MATH" A subliminal message is below our threshold of awareness but may nonetheless influence us. Can subliminal messages help students learn math? A group of 18 students who had failed the mathematics part of the City University of New York Skills Assessment Test agreed to participate in a study to find out. All received a daily subliminal message, flashed on a screen too rapidly to be consciously read. The treatment group of 10 students (assigned at random) was exposed to "Each day I am getting better in math." The control group of 8 students was exposed to a neutral message, "People are walking on the street." All 18 students participated in the summer program designed to raise their math skills, and all took the assessment test again at the end of the program. The tables below gives data on the each subject's test score improvement:

(You can create dot plots by hand!)

Treatment Group (10 subjects)

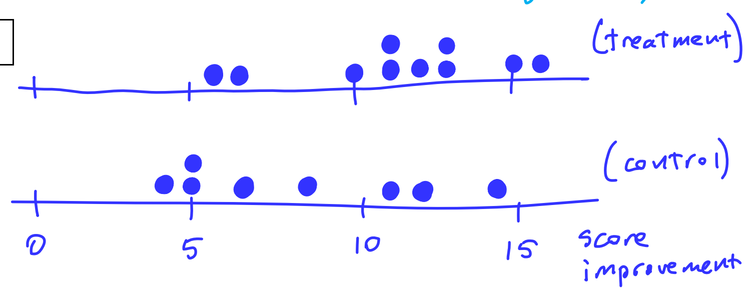
6	7	12	11	15	16	11	13	13	10

Mean: 11.4 Standard Deviation: 3.17

Control Group (8 subjects)

11	5	4	8	14	5	7	12

Mean: 8.25 Standard Deviation: 3.69



At the 5% level of significance, do the data provide evidence that students receiving the positive message would experience a higher mean improvement than students receiving the neutral message?

2-sample t-test

μ_1 = true mean score improvement for students that receive the positive message

μ_2 = " " " " ... " neutral ".

$H_0: \mu_1 = \mu_2$ OR $\mu_1 - \mu_2 = 0$

$H_A: \mu_1 > \mu_2$ $\quad \mu_1 - \mu_2 > 0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(formula is optional)

$t = 1.914$ $\quad df = 13.919$

$p = 0.038$

write these down from calculator

Conditions:

- The 18 students were randomly assigned to one of the two groups. This creates reasonably independent groups.
- NNC: The dotplots for each group show plausible normality.

At $\alpha = 0.05$:

Since $p < \alpha$, we reject Ho. We have evidence that students who receive the positive message experience a higher mean score improvement than students who receive the neutral message.