Agenda

1. Difference between two means
2. ANOVA
3. Multiple Testing

Difference of two means Often the data are *not* naturally paired. In particular, we are often interested in comparing mean from two groups of unequal sizes. For example, the 11 children whose fathers had higher IQs than mothers had a lower average score on the skills test than the 25 children whose mothers had higher IQs than the fathers.

```
library(mosaic)
library(openintro)

gifted <- gifted %>%
  mutate(diff = motheriq - fatheriq, moth_greater = (diff > 0))

favstats(score ~ moth_greater, data = gifted)

##   moth_greater min    Q1 median  Q3 max     mean       sd  n missing
## 1        FALSE 150 152.5    156 161 164 156.5455 4.906397 11       0
## 2         TRUE 154 159.0    160 163 169 160.2800 4.097967 25       0
```

Now the samples are *not* naturally paired. How do we know if the observed difference in means between these two groups is meaningful? Let X be the random variable that gives the analytical skills test score for a gifted child whose father has a higher IQ than her mother, and let Y be the random variable that gives the test score for a gifted child whose mother has a higher IQ. Then we need to understand the sampling distribution of the test statistic $D = \bar{X} - \bar{Y}$.

Just as we did with proportions, the standard error of the difference is a combination of the standard errors of the variables.

$$SE_D = \sqrt{(SE_X)^2 + (SE_Y)^2}$$

If both X and Y meet the conditions for a t-based sampling distribution, then D will meet those conditions as well. We typically use $\min(n_1 - 1, n_2 - 1)$ for the degrees of freedom.

The hypothesis test for a difference of two means constructed in this manner is called the *two-sample t-test*, and it is a commonly applied statistical technique.

1. Use the information above to conduct a two-sample t-test for a difference in mean test score between gifted children whose fathers have higher IQs vs. those whose mothers have higher IQs.

Practice Problem

1. Do customers spend more on a trip to Walmart or Target? Suppose researchers interested in this question collected a systematic sample from 85 Walmart customers and 80 Target customers by asking customers for their purchase amount as they left the stores. Using their data, they calculated that Walmart shoppers spent an average of $45 with a standard deviation of $21, and Target shoppers spent an average of $53 with a standard deviation of $19 dollars. Test an appropriate hypothesis and state your conclusion.

ANOVA We just developed a way to compare differences in means between *two* groups. But what if we have more than two groups? Analysis of Variance (ANOVA) provides a mechanism for simultaneously assessing the differences between multiple groups.

The HELP study was a clinical trial for adult inpatients recruited from a detoxification unit. Patients with no primary care physician were randomized to receive a multidisciplinary assessment and a brief motivational intervention or usual care, with the goal of linking them to primary medical care. We'll consider two variables:

- `cesd`: Center for Epidemiologic Studies Depression measure at baseline (high scores indicate more depressive symptoms)
- `substance`: primary substance of abuse: alcohol, cocaine, or heroin

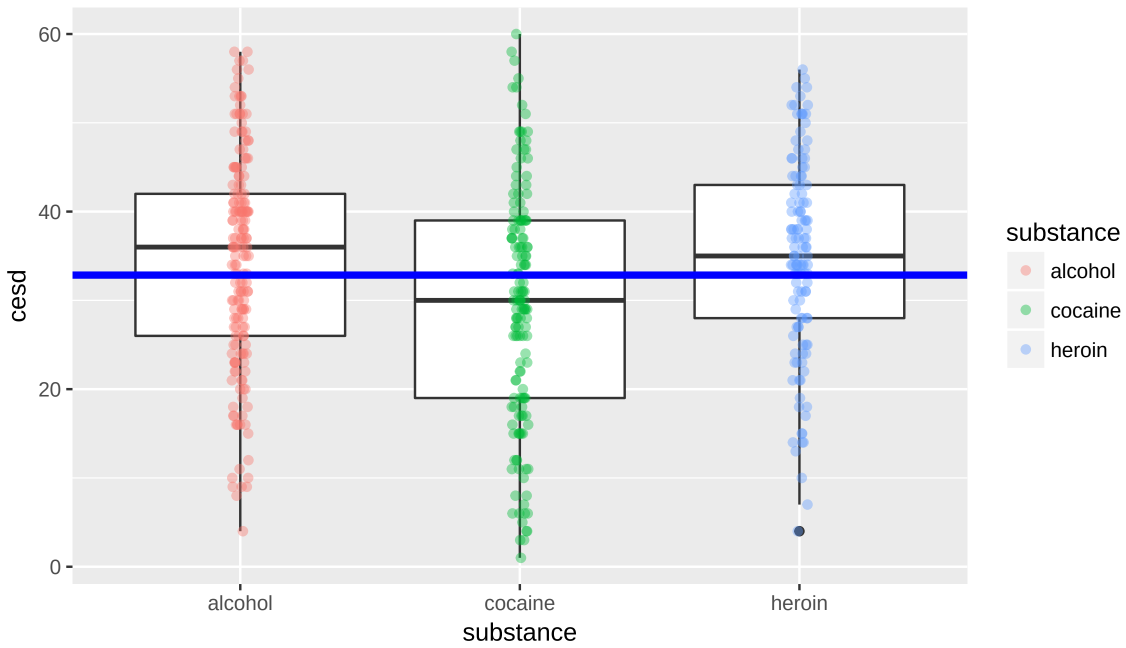
Are there important differences in the depression scores among patients depending on their drug of abuse?

```
favstats(cesd ~ substance, data = HELPrct)

##   substance min Q1 median Q3 max     mean       sd   n missing
## 1   alcohol   4 26     36 42  58 34.37288 12.05041 177       0
## 2   cocaine   1 19     30 39  60 29.42105 13.39740 152       0
## 3    heroin   4 28     35 43  56 34.87097 11.19812 124       0

grand_mean <- mean(~cesd, data = HELPrct)

ggplot(data = HELPrct, aes(y = cesd, x = substance)) +
  geom_boxplot() +
  geom_jitter(height = 0, width = 0.03, alpha = 0.4, aes(color = substance)) +
  geom_hline(yintercept = grand_mean, col = "blue", size = 1.5)
```



```
anova(lm(cesd ~ substance, data = HELPrct))

## Analysis of Variance Table
##
## Response: cesd
##            Df Sum Sq Mean Sq F value    Pr(>F)
## substance   2   2704  1352.1  8.9363 0.0001563 ***
## Residuals 450  68084   151.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

1. Write down the null and alternative hypotheses

2. Check the conditions for ANOVA: is independence reasonable? Is normality reasonable? What about equal variance?

3. Find the value of the test statistic (F) in the ANOVA table. Can you derive it from the other numbers in the table?

4. Draw a picture of the sampling distribution of F. How many degrees of freedom do we have?

5. Find the p-value. [You will need the function `pf()`.]

6. What do you conclude? Write a sentence summarizing your findings.

Multiple Testing Why is the comic on our home page funny?: `http://www.science.smith.edu/~rgarcia/sds201-S17/index.html`

The simplest (and most conservative) way to correct for multiple testing is to use Bonferroni's correction: simply divide the α-level by the number of comparisons that you are making.