2-sample t tests (summary)

- matched samples:
  - 1-sample t test on difference scores
  - compare mean difference to expected difference (of zero)
  - is observed mean unusually large?
  - df = N-1
- independent samples:
  - evaluate difference between two sample means with t test
  - is observed difference between means unusually large?
  - use sampling distribution of difference between means
  - df = n1 + n2 - 2

Evaluating scores & means
z tests & t tests

Are our data unusual?

- Is our observation (test score, group mean, group difference) unusual?
- You cannot answer the question based on the observation/measure alone
- Need information about population from which the observation was drawn
Are our data unusual?

- Is our observation (test score, group mean, group difference) unusual?
- You cannot answer the question based on the observation/measure alone
- Need information about population from which the observation was drawn
- What is a “typical” score? Usually, we use the population mean (μ)
- What is the spread/variability of the scores?
  - Usually, we use the standard deviation (σ)
  - Large variability implies that many scores are far from the mean
  - Small variability implies that most scores are close to the mean
- How are scores distributed in population?
  - Often we assume scores are distributed normally
  - (Central Limit Theorem)

Is our individual test score unusual? (population parameters are known)

- Example: Intra-Ocular Pressure
- Observation: X = 22
- Express observation as a z score:
  - X = 22... z = (22-17)/2.5 = 2
- Score is 2 SD above mean
  - assuming H0 is true (μ=17 & σ=2.5)
  - p(z >= 2) = 0.023
- Only 2.3% of scores in population fall above X=22 when H0 is true

z & t tests

- z & t tests used to evaluate “unusualness” of scores & group means
- “observation” is a test score, group mean, or difference between group means
- express observation as number of standard deviations from mean of the appropriate distribution in the population
  - distribution of test scores (μ, σ)
  - distribution of sample means (μ̄X, σ̄X)
  - distribution of difference between 2 sample means (μ̄X1−X̄2, σ̄X1−X̄2)
Is our group mean unusual?
(known population $\mu$ & $\sigma$)
- Example: Intra-Ocular Pressure
- Mean of $N=100$ individuals with mutation on gene $Y$ is 18
- Express group mean as a $z$ score:
  - $z = (18-17)/2.5/\sqrt{100} = 0.25 = 4$
  - $z$ is 4 SD above mean
- Assuming $H_0$ is true ($\mu=17$ & $\sigma=2.5$)
- $p(z \geq 4) < 0.001$
- Less than 0.1% of group means greater than 18 (when $H_0$ is true)

Is our group mean unusual?
(estimated population $\sigma$)
- Example: tapping test
- Sample consists of $N=5$ individuals
  - mean = 54.2, $s = 6$
- Is mean unusually low?
  - Express group mean as a $t$ score:
    - $t = (54.2 - 59)/6/\sqrt{5} = -1.79$
  - Mean is 1.79 SD below mean
  - Assuming $H_0$ is true ($\mu=59$)
  - $p(t < -1.79) = 0.074$ ($df = N-1 = 4$)
  - When $H_0$ is true, 7.4% of group means ($N=5$) will be less than 54.2

Is our difference between group means unusual?
(estimated population $\sigma$)
- Example: weight gain in adolescent girls with anorexia
  - Therapy Group ($N=17$): mean = 7.26, $s = 7.16$
  - Control Group ($N=26$): mean = -0.45, $sd = 7.99$
- Is difference unusual if true difference is zero?
  - Express group difference as a $t$ score:
    - $t = (7.26 - (-0.45))/6/\sqrt{(7.76^2/17 + 7.99^2/26)} = 3.29$
  - Mean is 3.29 SD below mean
   - Assuming $H_0$ is true ($\mu_1=\mu_2$)
   - $p(t > 3.29) = 0.001$ ($df = n_1 + n_2 - 2 = 41$)
  - When $H_0$ is true, 0.1% of differences between group means will be greater than 7.71

Hypothesis testing summary
- express observations as number of SDs from mean of (sampling) distribution
  - when $\sigma$ is known, statistic follows $z$ distribution
  - when $\sigma$ is estimated, statistic follows $t$ distribution
- consider alternative explanations of “significant” result
  - placebo effect, regression to the mean, etc.
- importance of using a control group
- Null hypothesis testing will lead to errors:
  - Type I error rate (false alarms; alpha; significance level)
  - Type II error rate (misses; beta; statistical power = 1-beta)
  -