Psych 2111
Prof. Matt Jones

Lecture 11: t-test

Hypothesis Test for Population Mean
- Goal: Infer μ from M
- Null hypothesis: $\mu = \mu_0$
  - Usually 0
  - Sometimes another value, e.g. from larger population
- Change scores
  - Memory improvement, weight loss, etc.
- Sub-population within known, larger population
  - IQ of CU undergrads
- Approach:
  - Determine likelihood function for M, using CLT
  - Compare actual sample mean to critical value

Likelihood Function for M
- Probability distribution for M, according to H_0
- Central Limit Theorem:
  - Mean equals population mean: $\mu_M = \mu_0$
  - Standard deviation: $\sigma_M = \frac{\sigma}{\sqrt{n}}$
  - Shape: Normal

Critical Value for M
- Result that has 5% (a) chance of being exceeded, **IF** null hypothesis is true
- Easier to find after standardizing $p(M)$
  - Convert distribution of sample means to z-scores: $z_M = \frac{M-\mu_M}{\sigma_M} = \frac{M-\mu_0}{\sigma/\sqrt{n}}$
- Critical value for z_M always the same
  - Just use `qnorm()`
  - Equals 1.64 for $\alpha = 5\%$

Example: IQ
- Are CU undergrads smarter than population?
  - Sample size $n = 100$, sample mean $M = 103$
- Likelihood function
  - If no difference, what are probabilities of sample means?
  - Null hypothesis: $\mu_0 = 100$
  - CLT:
    - $\mu_M = 100$
    - $\sigma_M = s/\sqrt{n} = 15/10 = 1.5$
- z-score: $(103-100)/1.5 = 2$

Problem: Unknown Variance

$$z_M = \frac{M - \mu_0}{\sigma/\sqrt{n}}$$

- Test statistic depends on population parameter
  - Can only depend on data or values assumed by H_0
- Could include σ in null hypothesis
  - H_0: $\mu = \mu_0$ & $\sigma = \sigma_0$
  - Usually no theoretical basis for choice of σ_0
  - Cannot tell which assumption fails
- Change test statistic
  - Replace population SD with sample SD
  - Depends only on data and μ_0
  - $t = \frac{M - \mu_0}{s/\sqrt{n}}$

t Statistic

- Invented in 1908 by "Student" at Guinness brewery
- $t = \frac{M - \mu_0}{s/\sqrt{n}}$
- Deviation of sample mean divided by estimated standard error
- Depends only on data and μ_0
- Sampling distribution depends only on *n*

t Distribution

- Sampling distribution of *t* statistic
- Derived from ratio of Normal and (modified) χ^2
- Depends only on sample size
  - <u>Degrees of freedom</u>: df = $n - 1$
  - Invariant with respect to μ, σ
- Shaped like Normal, but with fatter tails
  - Reflects uncertainty in sample variance
  - Closer to Normal as *n* increases
- Critical value decreases as *n* increases

α = .05			
df	t_{crit}	df	t_{crit}
1	6.31	5	2.02
2	2.92	10	1.81
3	2.35	30	1.70
4	2.13	∞	1.64

Steps of t-test

1. State clearly the two hypotheses
2. Determine null and alternative hypotheses
   H_0: $\mu = \mu_0$
   H_1: $\mu \neq \mu_0$
3. Compute the test statistic t from the data

$$t = \frac{M - \mu_0}{s/\sqrt{n}}$$

4. Determine likelihood function for test statistic according to H_0
   t distribution with n-1 degrees of freedom
5. Find critical value
   R: `qt(alpha,n-1,lower.tail=FALSE)`
6. Compare actual result to critical value
   $t < t_{crit}$: Retain null hypothesis, $\mu = \mu_0$
   $t > t_{crit}$: Reject null hypothesis, $\mu \neq \mu_0$

Example: Rat Mazes

- Measure maze time on and off drug
  - Difference score: $\text{Time}_{\text{drug}} - \text{Time}_{\text{no drug}}$
  - Data (seconds): 5, 6, -8, -3, 7, -1, 1, 2
- Sample mean: $M = 1.0$
- Sample standard deviation: $s = 5.06$
- Standard error: SE = $s/\sqrt{n}$ = 5.06/2.83 = 1.79
- $t = (M - \mu_0)/\text{SE} = (1.0 - 0)/1.79 = .56$
- Critical value: `qnorm(.05,7,lower.tail=FALSE) = 1.89`
- t does not exceed critical value
  - Cannot reject null hypothesis
  - Assume no effect of drug