Quantitative Research Review: 3-1

- The Scientific Method
- Null Hypotheses, Alternative Hypotheses
- Defining a rejection region based on hypothesis
- T-tests
- Degrees of Freedom
- Error types

Type I, Type II Errors

		True state of nature		
		H_0	H_A	
Our	Reject H_0	Type I error	correct decision	
decision	'Accept' H_0	correct decision	Type II error	

(Orloff & Bloom, 2014)

Type I, Type II Errors

		True state of nature		
Our		H_0	H_A	
	Reject H_0	Type I error	correct decision	
decision	'Accept' H ₀	correct decision	Type II error	
		H_0	H_A (Orloff 8	& Bloom
	Reject H_0	1	H_A	& Bloom
	Reject H_0 'Accept' H_0	P(Reject H ₀ H ₀)	H_A P(Reject $H_0 \mid H_1$)	

Type I, Type II Errors



Power

```
significance level ("p-value") = P(type I error) = P(Reject H_0 | H_0) (probability we are incorrect)

power = 1 - P(type II error) = P(Reject H_0 | H_1) (probability we are correct)
```

$$\begin{array}{c|cccc} & H_0 & H_A \\ \hline \hline \text{Reject H_0} & \text{P(Reject H_0 | H_0)} & \text{P(Reject H_0 | H_1)} \end{array}$$

Power

significance level ("p-value") = P(type I error) = $P(Reject H_0 | H_0)$ (probability we are incorrect)

power = 1 - P(type II error) = $P(Reject H_0 | H_1)$ (probability we are correct)

Formally, a power function of a test with rejection region, *R*, is:

$$\beta(\theta) = P_{\theta}(X \in R)$$

where θ is the parameters of the distribution over which R is defined. (e.g. p, n for a binomial distribution)

Multi-test Correction

If alpha = .05, and I run 40 variables through significance tests, then, by chance, how many are likely to be significant?

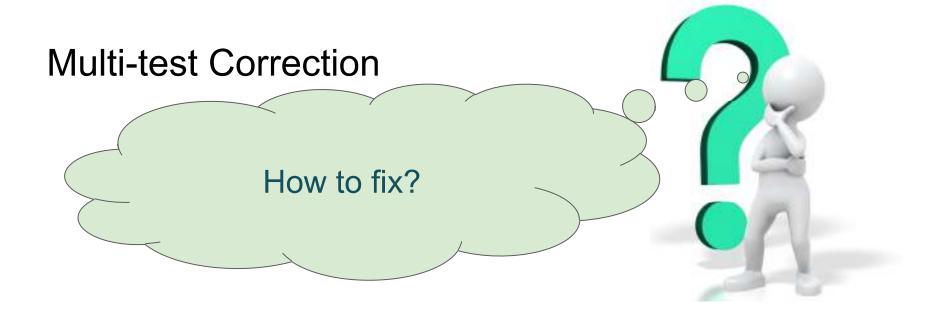


Multi-test Correction

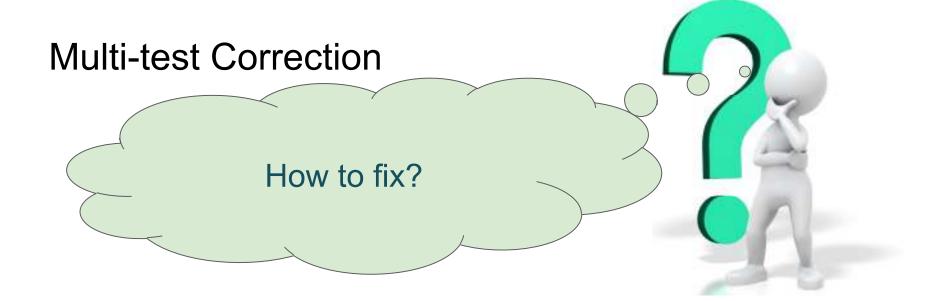
If alpha = .05, and I run 40 variables through significance tests, then, by chance, how many are likely to be significant?



2 (5% any test rejects the null, by chance)



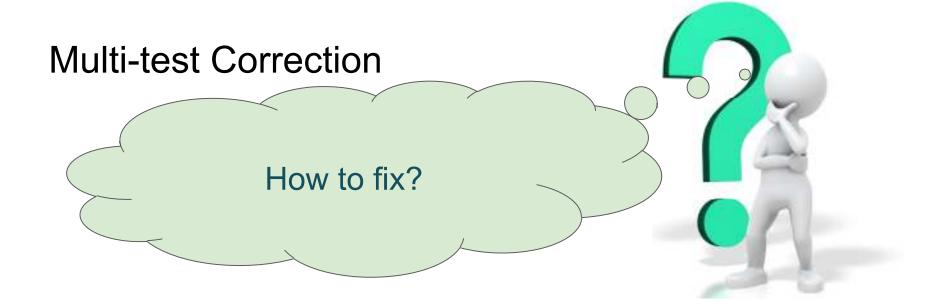
2 (5% any test rejects the null, by chance)



What if all tests are independent?

=> "Bonferroni Correction" (α/m)





What if all tests are independent?

=> "Bonferroni Correction" (α/m)

But this may over-correct.

Multi-test Correction

Benjamini-Hochberg Correction Procedure

- 1. Let $P_{(1)} < ... < P_{(m)}$ denote ordered p-values
- 2. Define: $\ell_i = \frac{ia}{C_m m}, \text{ and } R = max \left\{ i : P_i(i) < \ell \right\}$ where $C_m = 1$ if p-values are independent, $C_m = \sum_{i=1}^m \frac{1}{i}$ otherwise
- 3. Let $T = P_{(R)}$, the "rejection threshold"
- 4. Reject all $H_{(0)}$ for which $P_i \leq T$ (Weiss, 2005)

But this may over-correct.

The Scientific Method

Develop General Theories

General theories must be consistent with most or all available data and with other current theories.

Gather Data to Test Predictions

Relevant data can come from the literature, new observations or formal experiments. Thorough testing requires replication to verify results.

Make Observations

What do I see in nature? This can be from one's own experiences, thoughts or reading.

> Refine, Alter, Expand or Reject Hypotheses

Develop Testable Predictions

If my hypothesis is correct, then I expect a, b, c, ...

Think of Interesting Questions

Why does that pattern occur?

Formulate Hypotheses

What are the general causes of the phenomenon I am wondering about?

The Scientific Method

Which steps are most subjective?

Develop General Theories

General theories must be consistent with most or all available data and with other current theories.

Gather Data to Test Predictions

Relevant data can come from the literature, new observations or formal experiments. Thorough testing requires replication to verify results.

Make Observations

What do I see in nature?
This can be from one's own experiences, thoughts or reading.

Refine, Alter, Expand or Reject Hypotheses

Develop Testable Predictions

If my hypothesis is correct, then I expect a, b, c, ...

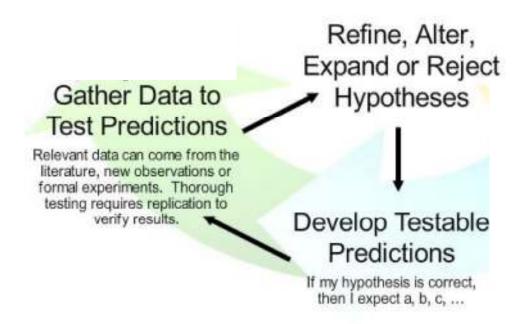
Think of Interesting Questions

Why does that pattern occur?

Formulate Hypotheses

What are the general causes of the phenomenon I am wondering about?

The Scientific Method Potential Effect from Big Data



Hypothesis Testing

Terminology: "tails" -- is the rejection region made up of one or two sides of the rejection region?

Example: Comparing two means:

- two-tailed p-value: P(T > |t| or T < -|t|) = 2*P(T > |t|)?
 (when there is no assumption of direction of difference)
- one-tailed p-value: P(T > t)? (when H_a posits the second mean is greater)
 P(T < t)? (when H_a posits the second mean is less)

Resampling Techniques

"nonparametric" tests

The permutation test:

- t_{obs} = Compute observed score
- passes = 0
- for 1 to *B*:
 - o randomly permute the data, keeping the same sizes per class
 - \circ t_{B} = compute score on permuted data
 - \circ if $t_B > (or <) t_{obs}$: passes+=1
- p_value = passes/B

Application: comparing two distributions, especially when they are unknown.

Finding a linear function based on X to best yield Y.

X = "covariate" = "feature" = "predictor" = "regressor" = "independent variable"

Y = "response variable" = "outcome" = "dependent variable"

Regression: r(x) = E(Y|X=x)

goal: estimate the function r

Finding a linear function based on X to best yield Y.

X = "covariate" = "feature" = "predictor" = "regressor" = "independent variable"

Y = "response variable" = "outcome" = "dependent variable"

Regression:
$$r(x) = E(Y|X = x)$$

goal: estimate the function r

Linear Regression (univariate version): $r(x) = \beta_0 + \beta_1 x$

goal: find β_0 , β_1 such that $r(x) \approx \mathrm{E}(Y|X=x)$

Simple Linear Regression
$$Y_i=\beta_0+\beta_1X_i+\epsilon_i$$
 where $\mathbf{E}(\epsilon_i|X_i)=0$ and $\mathbf{V}(\epsilon_i|X_i)=\sigma^2$

$$r(x) = \beta_0 + \beta_1 x$$

