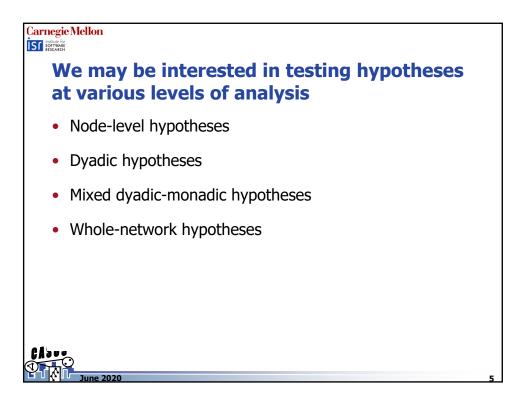
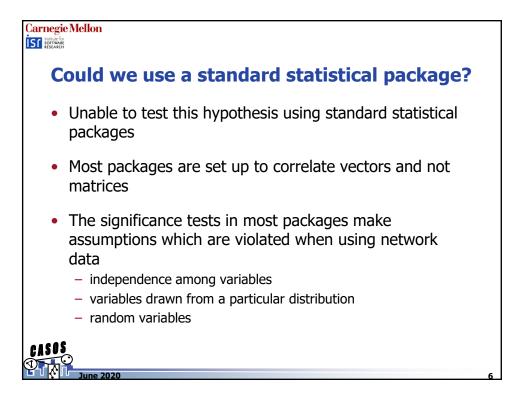


Carnegie Mellon	
	Manufactor
1 300737007	Marriage Business
1 ACCIAIUOL 2 ALBIZZI	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
3 BARBADORI	0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 2 0 0 0 0
4 BISCHERI	0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 1 0 0 1 0
5 CASTELLAN	0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 5 0 0 1 0 0 0 1 0 0 1 0 0 0 0
6 GINORI	
7 GUADAGNI	0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 7 0 0 1 0 0 0 1 0 0 0 0
8 LAMBERTES	0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 8 0 0 0 1 0 0 0 0
9 MEDICI	1 1 1 0 0 0 0 0 0 0 0 1 1 0 1 9 0 0 1 0 0 1 0 0 0 1 0 0 1 0 1
10 PAZZI	0 0 0 0 0 0 0 0 0 0 0 1 0 0 10 0 0 0 0
11 PERUZZI	0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 11 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0
12 PUCCI	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 12 0 0 0 0
13 RIDOLFI	0 0 0 0 0 0 0 1 0 0 0 0 1 1 13 0 0 0 0 0
14 SALVIATI	0 0 0 0 0 0 0 1 1 0 0 0 0 0 14 0 0 0 0 0
15 STROZZI	0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 15 0 0 0 0
16 TORNABUON	0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 6 0 0 0 0
GASOS	
	10







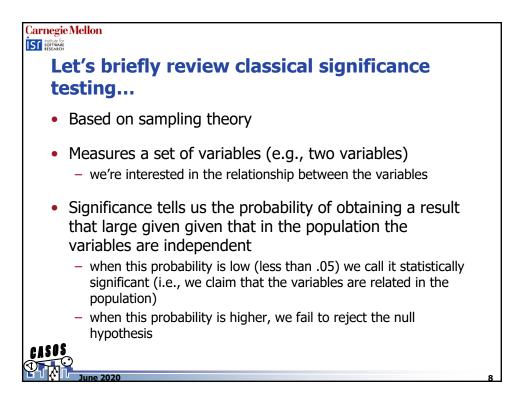


Special Methods for Testing Hypotheses
 Develop statistical models specifically designed for studying the distribution of ties in a network

 Exponential Random Graph Models
 Stochastic Actor-Oriented Longitudinal Models
 Complex models beyond the scope of this presentation

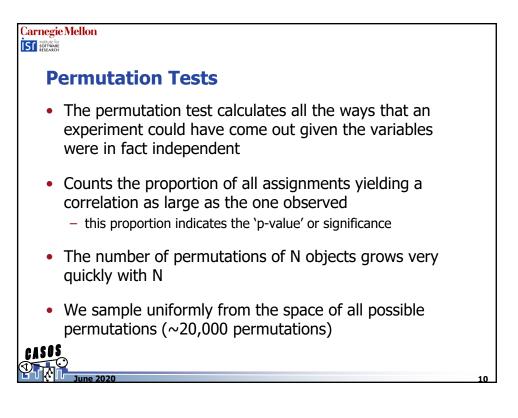
 Permutation Tests

 Easy to use and interpret
 Customizable for different research questions

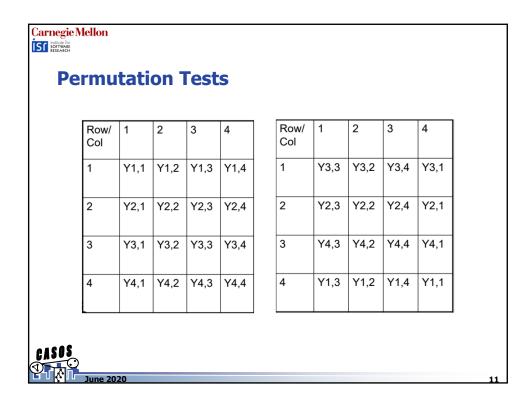


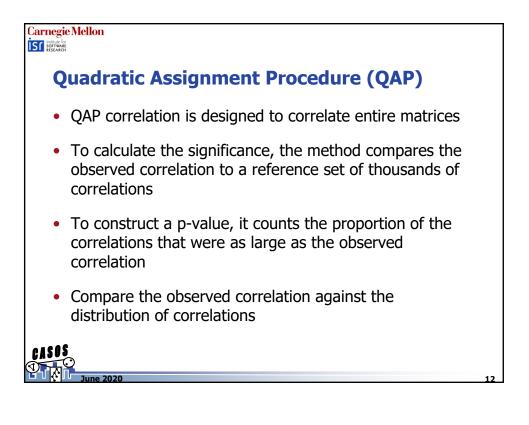


## Correction Mellon The logic of permutation tests differ from standard statistical tests For example, suppose you believe we favor tall people and scores in this course are correlated with height variables of height and score (correlation is .384) Now suppose we write down a set of math scores and have each student draw a score blindly from a hat What proportion of all the ways scores could be pulled would result in a correlation as large as our observation Compare the observed correlation against the distribution of correlations

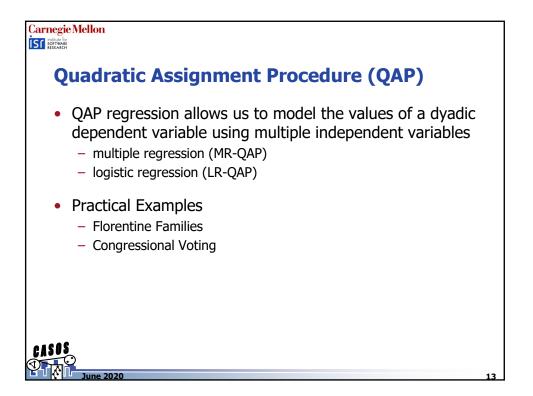






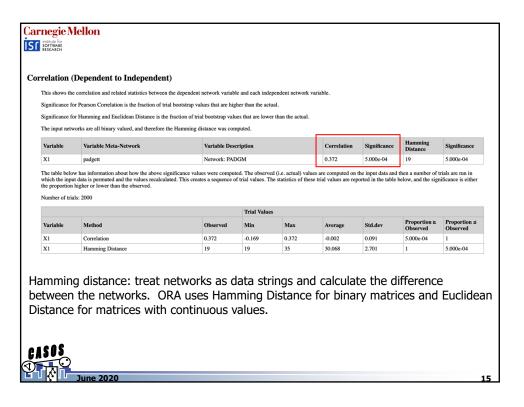






arnegie Mello	n						
institute for SOFTWARE RESEARCH							
RESEARCH							
AP/MROAP	ANALYSIS REPOR	Г					
		•					
Input data: padgett							
Start time: Tue May 19 16	5:35:43 2020						
Data Description							
Parameters							
i ui uniciciti							
Dependent meta-ne	etwork	padgett					
Dependent data		Network: PADGB					
Number of indeper	ident networks	1					
Random seed		0					
Number of permuta			2000				
Diagonal values us		false					
The table below des	cribes how the dependent and indepedent va	riables were constructed. The variable labels Y,	X1, X2, etc. are used consistently throughout the report.				
Variable	Variable Meta-Network		Variable Description				
Y	padgett		Network: PADGB				
X1	padgett		Network: PADGM				
1505							
' Nº ' June	2020						





institute for SOFTWARE RESEARCH							
gression Res	ilte						
Model: b0 + b1	Std.Errors; finally, bootstra	are inree computations for si pped standard errors are repor	rted in column Bootstra	ipped Std.Errors	s.	Errors; heteroskedasticity robus	a standard erfors are reporte
Model Fit							
Model Fit Observations				120			
				120 0.138	-		
Observations	Squares						
Observations R-Squared (R2)				0.138	]		
Observations R-Squared (R2) Residual Sum Of				0.138	]		
Observations R-Squared (R2) Residual Sum Of Total Sum Of Squ		Std. Coef	Std. Errors	0.138 11.310 13.125 0.310	Robust Std.Errors	Bootstrapped Std.Errors	Sig.V-Perm
Observations R-Squared (R2) Residual Sum Of Total Sum Of Squ Standard Error	3035	Std. Coef 0	Std. Errors 0.031	0.138 11.310 13.125 0.310	Robust Std.Errors	Bootstrapped Std.Errors 0.014	Sig.Y-Perm

R-Squared and Standard Error are both goodness-of-fit measures for linear regression models.

R-Squared indicates the percentage of the variance in the dependent variable that the Independent variables explain collectively (~13.8%)

Standard error measures the precision of the model's prediction – the standard distance between the observations and the regression line ( $\sim$ .31%)





**Carnegie Mellon** 

his shows the	e correlation and related statistics between the	e dependent network variab	le and each inde	pendent network v	ariable.				
ignificance fo	or Pearson Correlation is the fraction of trial b	bootstrap values that are hig	ther than the actu	ial.					
ignificance fo	or Hamming and Euclidean Distance is the fra	action of trial bootstrap valu	ues that are lowe	r than the actual.					
At least one in	nput network has non-binary link values, and	therefore the Euclidean dist	ance was compu	ted.					
Variable	Variable Meta-Network	Variable Des	Variable Description		Correlation	Significance	Euclidean Distance	Significance	
X1	Congress	Network: Age	Network: Age Difference		7.002e-04	0.560	188.475	0.560	
X2	Congress	Network: San	ne Committee		0.406	0.026	6.782	0.026	
	-	Network: Sam	Network: Same Gender			1	6.928	1	
X3	Congress	retwork. San	ie Gender		Network: Same Party 0.316 0.045 5.477 0.045				
X4 The table belo which the inpute the proportion	Congress whas information about how the above signi ut data is permuted and the values recalculated higher or lower than the observed.	Network: Sam	ne Party ted. The observe of trial values. Th	e statistics of thes	es are computed on	the input data and	then a number of	rials are run in	
X4 The table belo which the inpute the proportion	Congress whas information about how the above signi ut data is permuted and the values recalculated higher or lower than the observed.	Network: Sam	ne Party ted. The observe	e statistics of thes	es are computed on	the input data and	then a number of below, and the sign	rials are run in hificance is eithe	
X4 The table beloo which the inpute proportion Number of tria	Congress whas information about how the above signi ut data is permuted and the values recalculated higher or lower than the observed.	Network: Sam	ne Party ted. The observe of trial values. Th	e statistics of thes	es are computed on	the input data and	then a number of	rials are run in hificance is eithe	
X4 he table belo hich the inpu- proportion lumber of tria Variable	Congress whas information about how the above signi ut data is permuted and the values recalculated higher or lower than the observed. als: 2000	Network: Sam ificance values were comput d. This creates a sequence of	ne Party ted. The observe of trial values. Th <b>Trial Values</b>	e statistics of thes	es are computed on e trial values are rep	the input data and ported in the table I	then a number of below, and the sign Proportion ≥	rials are run in hificance is eithe Proportion :	
X4 The table belo which the inpute proportion Tumber of tria Variable X1	Congress Whas information about how the above signit at data is permuted and the values recalculate higher or lower than the observed. als: 2000 Method	Network: San dicance values were comput d. This creates a sequence of Observed	ne Party ted. The observe of trial values. Th Trial Values Min	Max	es are computed on e trial values are rep Average	the input data and ported in the table I Std.dev	then a number of below, and the sign Proportion ≥ Observed	rials are run in hificance is either Proportion : Observed	
X4 The table belo thich the inpu- ne proportion fumber of trian Variable X1 X1	Congress whas information about how the above signi ot data is permuted and the values recalculate higher or lower than the observed. als: 2000 Method Correlation	Network: Sam difcance values were computed d. This creates a sequence of Observed 7.002e-04	ted The observe of trial values. The observe <b>Trial Values</b> <b>Min</b> -0.682	Max 0.412	es are computed on e trial values are rep Average 8.368e-04	Std.dev       0.217	then a number of i below, and the sign Proportion ≥ Observed 0.560	Proportion : Observed 0.455	
X4 The table belo which the inpute the proportion Tumber of trian Variable X1 X1 X2	Congress Whas information about how the above signi to data is permuted and the values recalculated higher or lower than the observed. als: 2000 Method Correlation Euclidean Distance	Network: Sam d. This creates a sequence of Observed 7.002e-04 188.475	ted Trial Values Trial Values Min -0.682 187.433	Max 0.412 190.192	es are computed on e trial values are rep Average 8.368e-04 188.474	the input data and poorted in the table I Std.dev 0.217 0.548	Proportion ≥ Observed 0.560 0.455	Proportion : Observed 0.455 0.560	
X4 he table belo which the inpu te proportion aumber of tria Variable X1 X1 X1 X2 X2	Corgress Whas information about how the above signified to the above signified to the solution of the solution	Network: Sam           d. This creates a sequence of           Observed           7.002e-04           188.475           0.406	Trial Values Min -0.682 187.433 -0.279	Max           0.412           190.192           0.482	Average 8.368e-04 188.474 0.004	Std.dev 0.217 0.548 0.144	Proportion ≥ Observed           0.560           0.455           0.026	Proportion : Observed 0.455 0.560 0.998	
X4 The table belo thich the input te proportion Transformed trian Variable X1 X1 X1 X2 X2 X3	Corgress We has information about how the above signified of the abo	Network: Sam           dfcance values         were computed           d. This creates a sequence of         0           0 Doserved         7.002e-04           188.475         0.406           6.782         0	Min         0.682           187.433         -0.279           6.481         -0.481	Max           0.412           190.192           0.482           9.055	Average         8.368-04           188.474         0.004           8.179         8.179	Std.dev         0.217         0.548         0.144         0.470	Proportion ≥ Observed           0.560           0.455           0.026           0.998	Proportion : Observed 0.455 0.560 0.998 0.026	
X4 The table belo which the input	Congress Whas information about how the above signi ot data is permuted and the values recalculate higher or lower than the observed. als: 2000 Kethod Correlation Euclidean Distance Correlation Euclidean Distance Correlation Correlation Euclidean Distance Correlation Euclidean Distance Correlation Euclidean Distance Correlation	Network: Sam           difficance values were computed.           d. This creates a sequence of the	e Party ted. The observe ted. The observe <b>Trial Values</b> . The <b>International Values</b> <b>Min</b> 0.0.682 187.433 0.279 6.481 0.094	Max           0.412           190.192           0.482           9.055           0.756	Average         8.368e-04           188.474         0.004           8.179         -0.004	Std.dev         0.217           0.548         0.144           0.470         0.115	Proportion ≥ Observed           0.560           0.455           0.026           0.998           1	Proportion :           0.455         0.560           0.998         0.026           0.389         0.389	

This shows the c Y X1 X2 X3	orrelation be Y	tween all networ X1 7.002e-04	k variables. X2			
X1 X2			X2			
X1 X2	1			X3	X4	
X2			0.406	-0.094	0.316	
		1	-0.215	0.114	0.136	
			1	0.221	0.056	
X4				1	-0.120 1	
egression Re	sults					
(R2) Residual Sum Of Squares	90 0.280 14.404 20					
Variable	Coef	Std. Coef	Std. Errors	Robust Std.Errors	Bootstrapped Std.Errors	Sig.Y-Perm
	5.747e-04	0	0.107	0	0.169	0.977
	0.003	0.081	0.004	0	0.010	0.408
	0.340	0.447	0.075	0	0.123	0.005
	-0.162	-0.171	0.092	0		0.977
X4 0	0.246	0.260	0.090	0	0.091	0.044



