Empirical Research Methods in Information Science

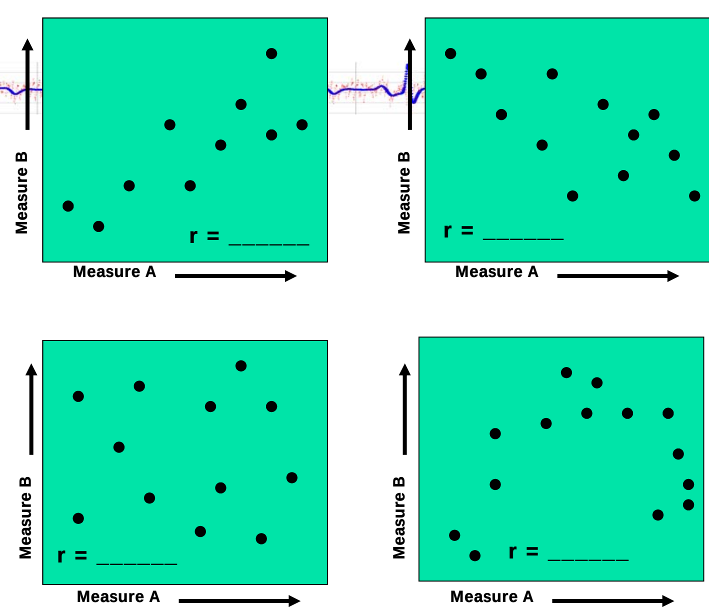
IS 4800 / CS 6350

Lecture 42

Correlation

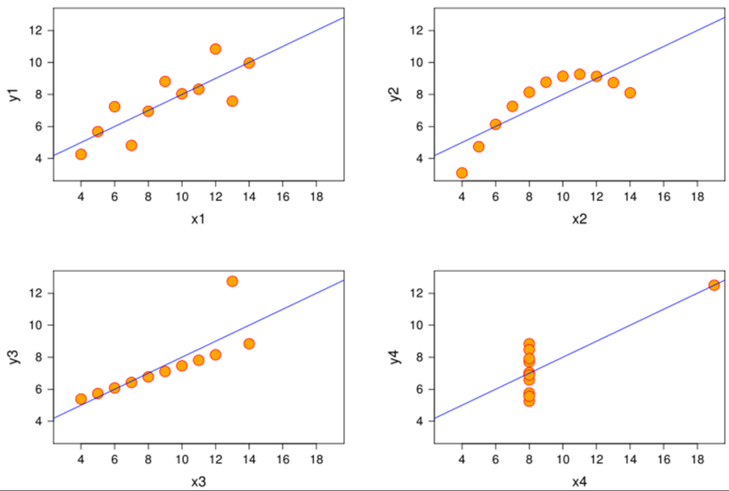
1

Quiz



3

Correlation?



All
r=0.81

4

Predictions Using Correlations

- Predictor vs. Criterion (Dependent) Variable
- Can you assume directionality?
- Depends entirely on your study design.

6

Pearson Correlation Coefficient

- Assumptions
  1. Two interval (or ratio) measures.
  2. Not an obviously curvilinear relationship.
  3. Both populations normally distributed*.

*Unimodal and symmetric frequency distributions.

7

Formula for Pearson correlation

$$r = \frac{\sum Z_X Z_Y}{N - 1}$$

-OR-

$$r = \frac{\sum[(X - M_X)(Y - M_Y)]}{\sqrt{(SS_X)(SS_Y)}}$$

Where

$$SS_X = \sum (X - M_X)^2$$

$$SS_Y = \sum (Y - M_Y)^2$$

8

Procedure for Hypothesis Testing with Correlations

- Further assumptions
  - Two variables are normally distributed.
    (unimodal and symmetric frequency distributions)
  - Equal distribution of each variable at each point of the other variable.

9

Procedure for Hypothesis Testing with Correlations

- Populations being compared:
  - ***Test:*** The population from which the observed sample was drawn.
  - ***Comparison:*** A hypothetical population in which the variables are unrelated, i.e., have a correlation of zero.

10

Procedure for Hypothesis Testing with Correlations

- Form of hypothesis H1?
  - The correlation in the observed population is different from a population in which the correlation is zero.
  - Unlikely we would have obtained a correlation this big if the variables actually were unrelated.
- Form of null hypothesis H0?
  - The correlation in the observed population is the same as a population in which the correlation is zero.

11

Procedure for Hypothesis Testing with Correlations

- Heuristic threshold for α=0.05:

$$r > \frac{2}{\sqrt{N}}$$

- *Exact form given in Aron (t-test).*

12

Procedure for Hypothesis Testing with Correlations

- R:
  - Stacked data format
  - Run cor.test…
  - See if significance<threshold
    - Yes => reject H0
    - No => inconclusive
- Manually:
  - Compute r
  - Is $r > \frac{2}{\sqrt{N}}$
    - If yes => reject H0
    - If no => inconclusive

14

Reporting results

r=*val*, *p<sigthresh*

Where,

- sigthresh = pre-defined significance threshold
  - Note: if p<<sigthresh, can report that as well, e.g., "p<.01", "p=.001"

For example: **r=0.82, p<.05**

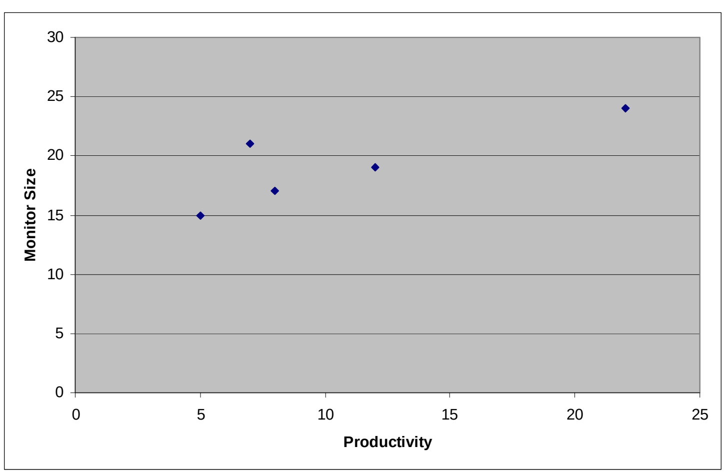
If not significant, than use "n.s." instead of "p<…".

15

Example

Employee	Sue	Sam	Sid	Sal	Sierra
Productivity	8	5	7	12	22
Monitor Size	17	15	21	19	24

Example Continued



Example

Employee	Sue	Sam	Sid	Sal	Sierra
Productivity	8	5	7	12	22
Monitor Size	17	15	21	19	24

P = 10.8 (6.8), M = 19.2 (3.5)

Z scores	-0.414	-0.858	-0.562	0.178	1.657
	-0.63	-1.202	0.515	-0.057	1.374
Z score products	0.260881	1.031666	-0.28968	-0.01016	2.276781

R = sum / (N-1) = +0.82

18

Pearson Correlation coefficient

- Which of the following is it appropriate for?
  - Descriptive study designs
  - Demonstration study designs
  - Correlational study designs
  - Experimental study designs

19

Group Exercise

20

Group Exercise

- For each problem, write
  1. Two populations being compared
  2. Research hypothesis
  3. Null hypothesis
  4. Test criteria
  5. Scatter plot
  6. r (if appropriate)
  7. Hypothesis test results
     - publication format and
     - English

21

Pearson Correlation in R

```
#For vectors v1, v2

#Just ‘r’
cor(v1,v2)

#Hypothesis test (including r)
cor.test(v1,v2)
```

22

Example Correlation Matrix

	Mean	CS	SE	EOU	TP	IC	OV
CS - Customer Support	4.7	1.00	*0.39*	*0.60*	*0.53*	*0.48*	*0.51*
SE – Security	5.0		1.00	*0.30*	*0.34*	*0.36*	*0.32*
EOU - Ease of Use	5.4			1.00	*0.49*	*0.53*	*0.62*
TP - Transactions and Payment	5.0				1.00	*0.58*	*0.49*
IC - Information Content and Innovation	5.0					1.00	*0.64*
OV - Overall Satisfaction	5.4						1.00

Table 6: *Correlation Matrix and Means (all $p < 0.05$)*

23

Example Correlation Matrix

Morrow, et al '96, Medication Instruction Design

Simple Correlations among Instruction Deviation Scores, Age, Vocabulary, Number of Medications Taken, and Health Beliefs for Older and Younger Participants

	External	Need for Information	Chance	Internal	Vocab	# Meds	Instruction Deviation Scores
			Older Group (N = 42)				
Age	.04	−.35*	.07	.15	.04	.08	.13
Ext		.24	−.16	.35*	−.24	−.09	.32*
Info			−.18	−.04	−.16	.14	.10
Ch				−.08	.02	.09	−.07
Int					.26	−.32*	−.22
Voc						−.26	−.59***
Meds							.12
			Younger Group (N = 42)				
Age	−.06	.09	.23	.10	.02	.15	.13
Ext		.07	−.02	.01	.01	.01	−.15
Info			−.43**	.30*	.16	.01	.04
Ch				−.32*	−.04	.12	.16
Int					−.21	−.08	.18
Voc						.21	−.44**
Meds							.06

* $p < .05$, ** $p < .01$, *** $p < .001$.

Comparing r's

- If you want to make statements about how large one correlation is relative to another.
  - e.g. one is twice as large as another
- Don't compare r's directly...
- Compare r^2 ("proportionate reduction in error")

25

Other measures of association

- Point-biserial
  - One numeric & one nominal measure
  - Just dummy code the nominal and use Pearson correlation.
- Spearman Rank Order (rho)
  - Two ordinal measures (or for transformed numeric measures if non-linear)
  - Replace each value with its rank order
  - Compute Pearson correlation with ranks
  - Measures degree of monotonicity

26

Two meanings of 'correlation'

- Correlation statistic   vs.
- Correlational research model

27

Example: Leashes & Attachment

Example: Leashes & Attachment

- You want to see if toddlers who grow up leashed have better attachment scores.
- You recruit 30 parents of toddlers, and randomly give half of them leashes and sign contracts agreeing to leash their toddler every time they leave the house.
- After one year you administer the strange situation protocol to classify the toddler attachment as secure, avoidant, or resistant.

- What kind of study is this?
- What statistic would you use to evaluate results?
- What is df?
- Assuming X^2(df) = 32.4, what would you conclude?
- Assuming X^2(df)=0.2, what would you conclude?

29

Example: Net Latency & Satisfaction

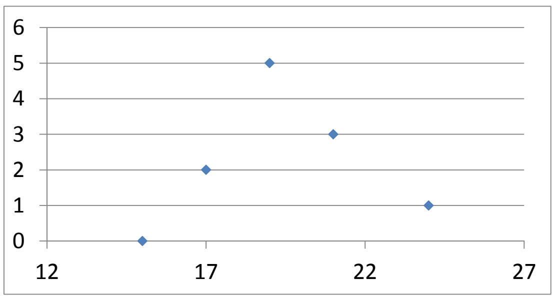
- CCIS wants to save money by switching to slower wireless routers, and wants to assess the impact this will have on student satisfaction. You want to know how slow things have to get before students start complaining.
- You have the crew implement a program that randomly chooses a network latency (between 0s and 10s) every time a student logs into CCIS wireless, then adds that latency to every network access from them. After 10 minutes of use a web form pops up asking students to rate their degree of satisfaction with CCIS systems.
- What kind of study is this?
- What statistic would you use to evaluate results?
- What is df?
- Assuming r = -0.8, p=.021, what would you conclude?
- Assuming r = 0.1, p=.342, what would you conclude?

What do you do your data is clearly not unimodal & symmetric OR this is a clear non-linear relationship?

31

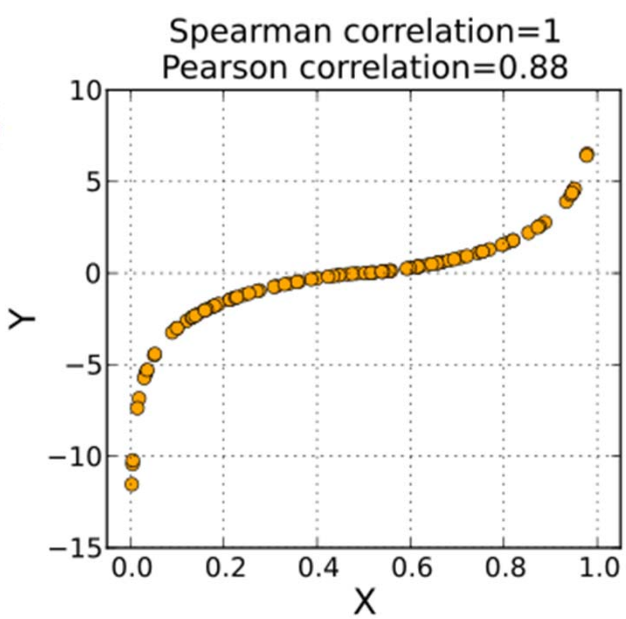
Sample Data for Spearman rho

Monitor Size	21	24	17	19	15
Productivity	3	1	2	5	0



rho = 0.3

32



Spearman measures degree of monotonicity.

33

Parametric vs. Non-parametric Statistics

- non-parametric statistics that do not rely on data belonging to any particular distribution

- E.g., Pearson r is a parametric statistic (assumes underlying distributions are normal – can be described using parameters – mean & stddev)
- E.g., Spearman rho is non-parametric

34

Intro to Power & Effect Size

35

A Brief Note About Power

- The "power" of a statistical test is its ability to detect differences in data that are inconsistent with the null hypothesis.
  - p(rejecting H0|H1)
  - aka Concluding H1, given that H1 is actually true.

Relationship between alpha, beta, and power.

	"The Truth" H1 True	H1 False
Decide to Reject H0 & accept H1	Correct p = power	Type I err p = alpha
Do not Reject H0 & do not accept H1	Type II err p = beta	Correct p = 1-alpha

Effect size

- The *amount* of change in the DVs seen.
- Can have statistically significant test but small effect size.
- *IF* your test yields significance *THEN* you are usually interested in how big the effect is.
  (If not significant, effect size is irrelevant.)
- Different for each statistical test.
- “Cohen” provides qualitative categories
  - small, medium, large
  - Example, for r: 0.1=small, 0.3=medium, 0.5=large

38

Power Analysis

- Power
  - Increases with effect size
  - Increases with sample size
  - Increases with alpha (decreases as you make the criteria more stringent)
- Should determine number of subjects you need ahead of time by doing a ‘power analysis’
- Standard procedure:
  - Fix alpha and beta (power)
  - Estimate effect size from prior studies
  - Determine number of subjects you need

39

Power & Effect Size for Correlation

- Effect size = |r|
- Power, see table 11-7, pg 465 Aron
  - Usually, given
    - Expected effect size
    - Test criteria
      - Desired significance level (usually 0.05)
      - Desired power (usually 0.8)
      - Directionality of test
  - Want to determine how many samples (ie, bodies) you need
  - What happens if you use fewer?

40

Table 11-8, Aron

Approximate number of participants needed for 80% power for a study using the correlation coefficient (r) for testing a hypothesis at the .05 significance level

	Effect size		
	Small (r=0.1)	**Medium (r=0.3)**	**Large (r=0.5)**
Two-tailed......	**783**	**85**	**28**

41

Effect Size & Power for X^2

- Completely different formulas than for Pearson r.
- Dependent on df.
- See Aron Tables 13-9, 13-10

42

Homework

- Read B&A Ch 10 to 303
- Read Aron Ch 3 & 4
  - Or remedial stats: normal curve, 1-sample t-tests
- Homework I10 – Chi^2 & Correlation
  - Work individually on this one. For each problem follow the steps to hypothesis testing we followed in class (statement of hypotheses, specification of testing criteria, etc.). Turn in your documentation for each step. Use R for the calculations, but copy and paste the relevant part of the R output (tables, charts) into your Word document.

43

Homework, cont'd

- Part 1 – Chi-Square Goodness of Fit
  - Translate scores from your "Facebook addiction " I9 survey into an ordinal measure with three values spanning equal parts of your index's range. For example, if your index ranges from 1 to 10, then 1.0 to 3.33 would be labeled "low", 3.33 to 6.66 would be labeled "medium", and 6.66 to 10.0 would be labeled "high". Determine how well your sample matches an expected distribution in which individuals fall equally into the 3 categories..
- Part 2 – Pearson Correlation
  - Determine how well your index of "Facebook addiction " co-varies with your validation measure (final part of I9). Provide scatter plot, calculation of r, and hypothesis test. What does your result tell you about the "computer game addiction " composite measure?

44