Simple Correlation

- Scatterplots & r
  - scatterplots for continuous - binary relationships
  - H0: & RH:
  - Non-linearity
- Interpreting r
- Outcomes vs. RH:
  - Supporting vs. "contrary" results
- Outcomes vs. Population
  - Correct vs. Error results

A scatterplot a graphical depiction of the relationship between two quantitative (or binary) variables

- each participant's x & y values depicted as a point in x-y space

Pearson's correlation coefficient (r value) summarizes the direction and strength of the linear relationship between two quantitative variables into a single number (range from -1.00 to 1.00)

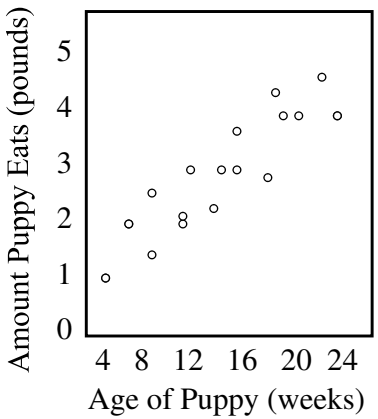
- you should always examine the scatterplot before considering the correlation between two variable
- NHST can be applied to test if the correlation in the sample is sufficiently large to reject H0: of no linear relationship between the variables in the population

A linear regression formula allows us to take advantage of this relationship to estimate or predict the value of one variable (the criterion) from the other (the predictor).

- prediction should only be applied if the relationship between the variables is "linear" and "substantial"

Example of a "scatterplot"

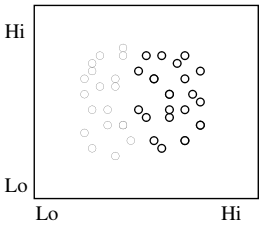
Puppy	Age (x)	Eats (y)
Sam	8	2
Ding	20	4
Ralf	12	2
Pit	4	1
Seff	24	4
…	..	..
Toby	16	3



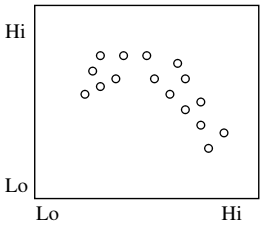
When examining a scatterplot, we look for three things...

- relationship
  - no relationship
  - linear
  - non-linear
- direction (if linear)
  - positive
  - negative
- strength
  - strong
  - moderate
  - weak

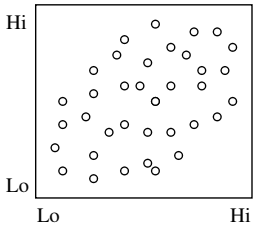
No relationship



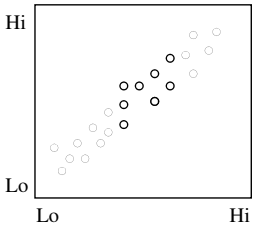
nonlinear, strong



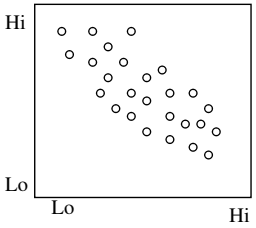
linear, positive, weak



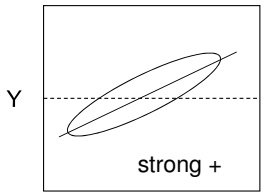
linear, positive, strong

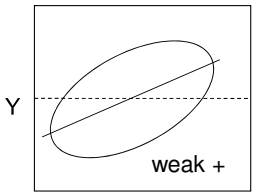


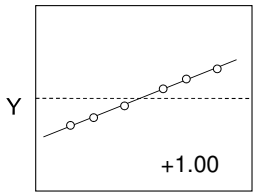
linear, negative, moderate



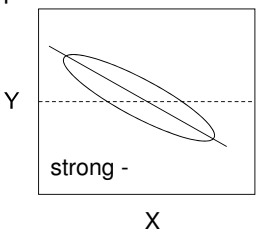
We can use correlation to examine the relationship between a quantitative predictor variable and a quantitative criterion variable.

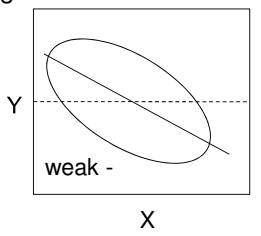


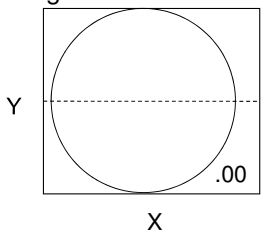




A positive r tells us those higher X values tend to have higher Y values



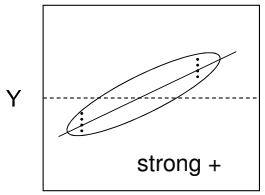


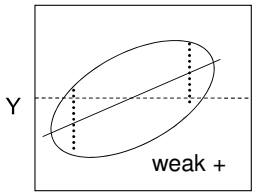


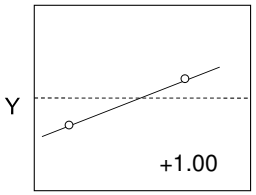
A negative r tells us those with lower X values tend to have higher Y values

A nonsignificant r tells us there is no linear relationship between X & Y

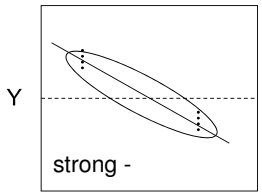
We can also use correlation to examine the relationship between a binary predictor variable and a quantitative criterion variable.

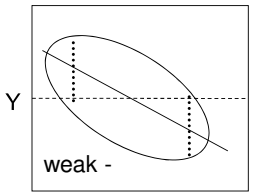


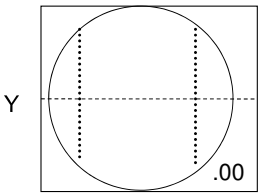




A positive r tells us the group with the higher X code as the higher mean Y





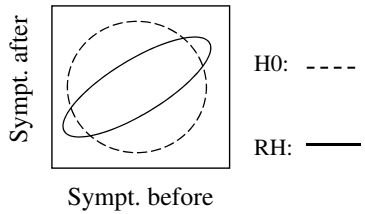


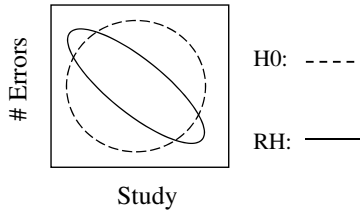
A negative r tells us the group with the lower X code as the higher mean Y

A nonsignificant r tells us the groups have “equivalent” means on Y

For each of the following show the envelope for the H0: and the RH:

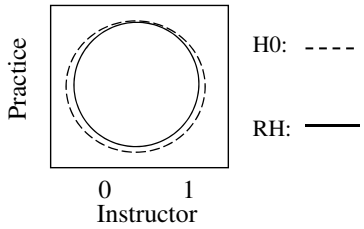
People who have more depressive symptomology before therapy will be those who have more symptomology after therapy.





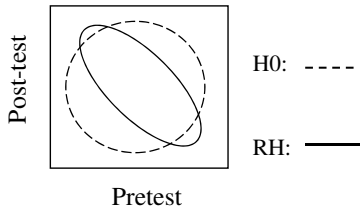
Those who study more have fewer errors on the spelling test

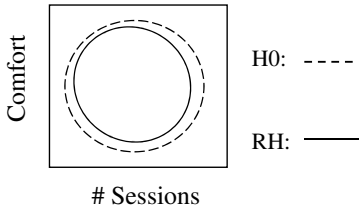
Instructor isn't related to practice.



For each of the following show the envelope for the H0: and the RH:

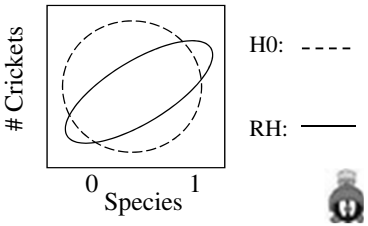
People who score better on the pretest will be those who tend to score worse on the posttest





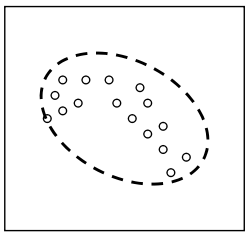
You can't predict patient comfort from the number of therapy sessions

I predict that snapping turtles (coded 1) will eat more crickets than painted turtles (coded 0).

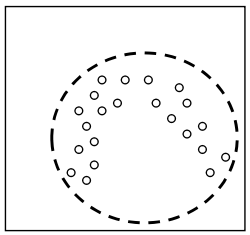


The Pearson's correlation (r) summarizes the direction and strength of the linear relationship shown in the scatterplot

- r has a range from -1.00 to 1.00
  - 1.00 a perfect positive linear relationship
  - 0.00 no linear relationship at all
  - -1.00 a perfect negative linear relationship
- r assumes that the relationship is linear
  - if the relationship is not linear, then the r-value is an underestimate of the strength of the relationship at best and meaningless at worst

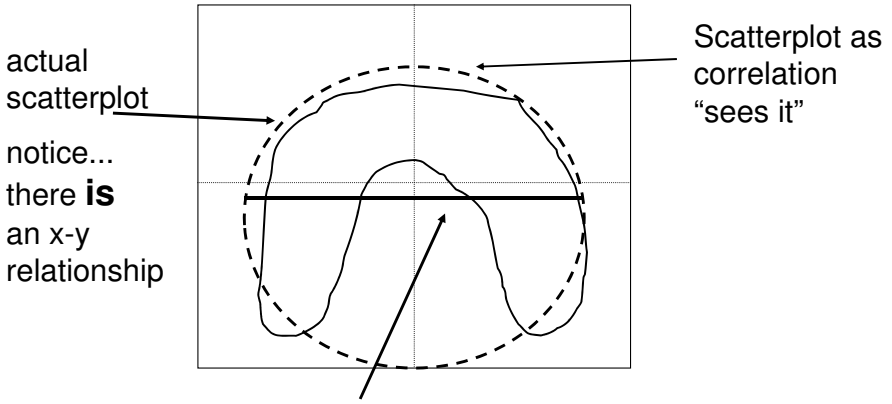


For a non-linear relationship, r will be based on a "rounded out" envelope -- leading to a misrepresentative r



Extreme Non-linear relationship

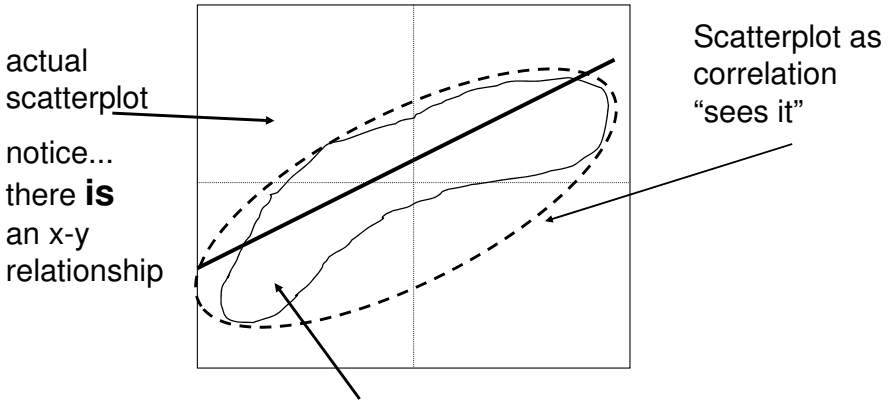
• r value is "misinformative"



regression line has 0 slope & r = 0 -- no linear relationship

Moderate Non-linear relationship

• r value is an underestimate of the strength of the nonlinear relationship



regression line has non-0 slope & r > 0 → but, the regression line not a great representation of the bivariate relationship

NHST testing with r

H0: r = 0.00 is the same as $r^2 = 0.00$

• get used to working with both r (the correlation between the 2 vars) and r^2 (the "variance shared between the 2 vars)"

Performing the significance test

• software will usually provide an exact p-value (use p < .05)

• a general formula is …

$$F = \frac{r^2}{(1 - r^2) / (N - 2)}$$

N = sample size

Find F-critical using df = 1 & N-2

What "retaining H0:" and "Rejecting H0:" means...

- When you retain H0: you're concluding…
  - The linear relationship between these variables in the sample ***is not*** strong enough to allow me to conclude there is a linear relationship between them in the population represented by the sample.
- When you reject H0: you're concluding…
  - The linear relationship between these variables in the sample ***is*** strong enough to allow me to conclude there is a linear relationship between them in the population represented by the sample.

effect significance vs. effect size vs. shared variance

The p-value (value range 1.0 – 0) tells the probability of making a Type I error if you reject the H0: based on the sample data

- e.g., p = .10 means "if we reject H0: based on these data there is a 10% chance that there really is no relationship between the variables in the population represented by the sample"
- The usual "acceptable risk" is less than 5% or $p < .05$

r (range -1.0 – 1.0) tells strength and direction of the bivariate relationship between Y & X

- "large enough to be interesting" value vary across research areas , but a common guideline is .10 = small, .30 = medium and .50 = large

r^2 (range 0 – 1.0) tells how much of the Y variability is "accounted for," "predicted from" or "caused by" X

- e.g., r=.30 means that $.30^2$ (9%) of the Y variability is accounted for by X
- "large enough to be interesting" will vary across research areas , but a common guideline is 1% = small, 10% = medium and 25% = large

Interpret each of the following (significance, strength & direction)

For age & social skills: r = .25, p = .043.

Sig – medium – positive → Older adolescents tend to have higher social skills scores

For practice and performance errors: r = -.52, p = .015

Sig – large – negative → Those who practiced more tended to have fewer errors

For age and performance: r = -.33, p = .231

Nonsig – medium? - negative ? →There is no linear relationship between age and performance??

For age group (<12=1, 12+=2) and social skills: r = .14, p = .004

Sig – small – positive → older group had higher mean on social skills scores

For age group (<12=1, 12+=2) and perf: r = -.31, p = .029

Sig – medium – negative →Younger group had higher mean performance

For age group (<12=1, 12+=2) and practice: r = .11, p = .098

Nonsig – small? – positive? → No mean practice difference between age

Statistical Conclusion Errors

In the population there are only three possibilities...

… and three possible statistical decisions

Outcomes	In the Population: -r	r = 0	+r
-r	Correctly rejected H0:	Type I error	Type III error
r = 0	Type II error	Correctly retained H0:	Type II error
+r	Type III error	Type I error	Correctly rejected H0:

Please note that this is a different question than whether the results "match" the RH: This is about whether the results from the sample are "correct" – whether the results are "represent the population. This is about statistical conclusion validity

The 9 outcomes come in 5 types …

Type I error -- "false alarm" - finding a significant mean difference between the conditions in the study when there really ***isn't*** a difference between the populations

Type II error -- "miss" - finding no difference between the conditions of the study when there really ***is*** a difference between the populations

Type III error -- "misspecification" - finding a difference between the conditions of the study that ***is different from*** the the difference between the populations

Correctly retained H0: -- finding no difference between the conditions of the study when there really ***is no difference*** between the populations

Correctly rejected H0: -- finding a difference between the conditions of the study ***that is the same as*** the the difference between the populations

Practice with statistical decision errors ...

We found that students who did more homework problems tended to have higher exam scores, which is what the other studies have found. — Correct rejection

We found that students who did more homework problems tended to have lower exam scores. All other studies found the opposite effect. — Type III

We found that students who did more homework problems and those who did fewer problems tended to have about the same exam scores, which is what the other studies have found. — Correct retention

We found that students who did more homework problems tended to have lower exam scores. Ours is the only study with this finding. — Can't tell -- what DID the other studies find?

We found that students who did more homework problems tended to have lower exam scores. Ours is the only study with this finding, others find no relationship. — Type I

We found that students who did more homework problems and those who did fewer problems tended to have about the same exam scores. Everybody else has found that homework helps. — Type II

correlation RH: vs. outcomes

There are only three possible Research Hypotheses

… and three possible statistical outcomes

Outcomes	Research Hypotheses: -r	r = 0	+r
-r			? ?  ? ?
r = 0			
+r	? ?  ? ?		

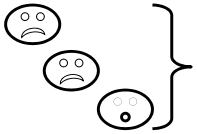
So, there are only 9 possible combinations of RH: & Outcomes …

of 4 types "effect as expected" 

"unexpected null"

"unexpected effect"

"backward effect"



Results contrary to RH:

Keep in mind that rejecting H0: does **not** guarantee support for the research hypothesis?

Why not ???

- The direction of the r might be opposite that of the RH: 
- The RH: might be that's there's no correlation (RH: = H0:) 

? ?

Remember !!!

Our purpose is not to "Reject the H0:" …

nor even to "support our RH:" …

Our real purpose is for our results to represent the relationship between the constructs in the target population !!!!!

A quick focus on the two that are most often confused …

Type III Statistical Decision Error

- When our significant findings have a direction or pattern different from that found in the population
- A difference between "the effect we found" and "the effect we should have found"

"Results contrary to our RH:"

- When our findings have a direction or pattern different from what we had hypothesized
- A difference between "the effect we found" and "the effect we hypothesized"

A result can be BOTH!!!!! (Or neither, or one, or the other !!!) 

RH:, statistical conclusions & statistical decision errors ...

 RH: supported

 Unexpected H0:

 Unexpected effect

 "backward" Results

Statistical Decision	RH: + direction/pattern	RH: H0:	RH: - direction/pattern
+ direction/pattern ($p < .05$)	Correct rejection, Type I or Type III	Correct rejection, Type I or Type III	Correct rejection, Type I or Type III
H0: ($p > .05$)	Correct retention or Type II	Correct retention or Type II	Correct retention or Type II
- direction/pattern ($p < .05$	Correct rejection, Type I or Type III	Correct rejection, Type I or Type III	Correct rejection, Type I or Type III

Lets practice …

Our RH: was that there will be a negative correlation between performance on the GRE and cumulative GPA.

We found r = .47, p = .016.

These results are "contrary to our RH:" -- a significant relationship in the opposite direction from the RH:

A literature review revealed 105 other studies involving these two variables, each of which found a correlation between .43 and .61 (all $p < .05$).

The consistent results of these other studies suggests that our finding was a correct rejection – what we found "does describe the relationship between these variables in the population".

Our RH: was incorrect, not supported, but our results were right!!!

Another …

Our RH: was that there will be a positive correlation between the severity of discomfort at the beginning of therapy and the amount of improvement a patient shows during the first six weeks of therapy.

We found r = .27, p = .085.

These results are "contrary to our RH:" -- a nonsignificant relationship isn't the RH: +r

The 14 studies of these two variables which followed ours each found a correlation between -.33 and -.41 (all $p < .05$).

The consistent findings of these other studies suggests that our finding was a Type II error – what we found "doesn't describe the likely relationship between these variables in the population".

Our RH: was incorrect, not supported & our results were "wrong"!!!

Try this one …

Our RH: was that there will be a positive correlation between social skills and comfort in an unfamiliar social situation.

We found r (82) = .37, p = .016.

These results "support our RH:"
- a significant relationship in the RH: direction

A literature review revealed 22 other studies involving these two variables, each of which found a correlation between -.13 and .11 (all p > .05)

The consistent results of these other studies suggests that our finding was a Type I error – what we found "does not describe the relationship between these variables in the population".

Our RH: was incorrect but supported & our results were wrong !!!

Last one …

Our RH: is that there will be a positive correlation between how much a person likes to compliment people and the number of close friends a person reports.

We found r (58) = .30, p < .05.

These results "support our RH:"
-- a significant, positive relationship, as hypothesized

A literature review revealed 8 other studies of these two variables, each of which found a correlation between .25 and .32 (all p < .05).

Our finding was consistent with earlier research!

The "researchers Trifecta"
RH: is correct & supported and the results are correct 1!!!

Keep in mind … There are 27 combinations of RH: (+ 0 -), Results (+ 0 -) and Population value (+ 0 -).

"Success" depends more on a consistent agreement of the last two than of the first two!