ISS 305:
Evaluating Evidence:
Becoming a Smart Research Consumer

7. Establishing Associations / Relationships

Reminder: Turn on your I<CLICKER

7. Establishing Relationships

I. Example: "Headache Coming On?"
II. Key concepts and tools: Relationship, Contingency table, Perfect relationship, Causal relationship
III. Common errors in judging simple relationships
  A. Single Cell
  B. Single Row
  C. Diagonal
  D. Marginal
IV. Biases in judging simple relationships
V. More complex relationships

Establishing Relationships.

- Example: **Headache coming On?**
- What variables are mentioned here?
- What links are suggested between these variables?

Headache coming on?

☐ Maybe you skipped your coffee this morning. According to new research, people who consume average amounts (one to three cups daily) often suffer debilitating withdrawal symptoms—moderate to severe headaches, depression, fatigue—when they forgo their morning brew. The culprit: caffeine—or, rather, the lack of it. Study findings, reported in the <u>New England Journal of Medicine</u>, explain away the mystery of why some people have headaches only on weekends: They drink less coffee than they do on workdays. Since moderate coffee drinking is considered relatively harmless, the simple antidote to withdrawal symptoms is . . . a cup of java.

Defining relationships

- From univaraiate to bivariate description
  - The more interesting empirical questions are not ones about single variables (e.g. how much coffee do people drink?) but about whether or not pairs of variables "go together" (e.g., does reducing coffee drinking go with headaches?)—whether they're related (and why)
  - We'll focus first on the "whether" question. Later on the "why" question.

Defining relationships

There is a **RELATIONSHIP** between two variables A and B ... (3 versions):

1. When they "go together"
   - as one goes up, the other goes up - **positive relationship**
   - as one goes up, the other goes down - **negative relationship**
     - These are simple, linear relationships
     - Later we'll consider more complex relationships
   - so, the absence or lack of a relationship would mean?
     - **as one goes up (or down), the other <u>stays the same</u>**

2. Whenever (as) different values on variable A are observed, different values on variable B are observed (Anderson, 1971, p. 14)
   - so, a lack of a relationship means?
     - **as different values on Variable A are observed, <u>the same</u> values on Variable B are observed**
     - **(that is, Variable B remains constant)**

3. Knowledge of a person's value on variable A <u>improves our ability to predict</u> that person's value on variable B (relative to how well we could have predicted <u>without</u> knowing A)
   - so, a lack of a relationship means?
     - **knowing A gives us <u>no improved ability to predict B</u> (over not knowing A)**

- Synonyms:
  - association, predictive relationship, probabilistic relationship, stochastic relationship, correlation, relation
  - Examples?

Contingency tables:

- Q: If we let the levels or possible values of Variable A be the rows, and the levels of Variable B be the columns, what's the smallest, simplest table possible to establish a relationship?
- A: 2 x 2 table (i.e., each variable has only 2 values)
- Consider the possible example of a relationship between weekday/weekend and headaches.
- **What's "a" here?**
- **What's "d"?**
- **What's "e"?**
- **What's "g"?**
- **What's "N"?**

Variable A Time of the week	Variable B Had a Headache?		Variable A Totals
	Yes	No	
Thursday	a	b	e=a+b
Saturday	c	d	f=c+d
Variable B Totals	g=a+c	h=b+d	N=a+b+c+d =e+f =g+h

Contingency tables:

Variable A Time of the week	Variable B Had a Headache?		Variable A Totals
	Yes	No	
Thursday	a=30 a/300= 10%	b=270 b/300 =90%	300
Saturday	c=10 c/100= 10%	d=90 d/100 =90%	100
Variable B Totals	40	360	400

Example 1

- For there to be a relationship, as day changes, headache reports must be **different**.
- More specifically, the **relative frequency distribution** (rfd) must be different - the **% or proportion** at each level of Variable B
  – the **base** for the calculation of these %s are the row (or column) totals
- Example 1: Comparing the rfd of B at each level of A (i.e., comparing **row** rfd's)
- It's **not** enough just for the **frequency distribution** to be different (that is, for (a,b)≠(c,d)).
  – For example, the **frequency** distributions are not identical (i.e., (30,270) is not the same as (10,90)),
  – but the **relative frequency** distributions <u>are</u> the same (i.e., (10%,90%)=(10%,90%), so there is **no** relationship in these data

Contingency tables:

Example 2: Comparing the rfd of A at each level of B (i.e., comparing **column** rfd's)

– The observations are the same as in the last table.
– For there to be a relationship, as headache status changes, calling day must be **different**.
– Are they here?
– No. the **relative frequency** distributions <u>are</u> the same (i.e., (75%,25%)=(75%,25%), so there is **no** relationship in these data

Note that you come to the same conclusion, whether you compare row rdf's or column rdf's

Variable A Time of the week	Variable B Had a Headache?		Variable A Totals
	Yes	No	
Thursday	a=30 a/40= 75%	b=270 b/360 =75%	300
Saturday	c=10 c/40= 25%	d=90 d/360 =25%	100
Variable B Totals	40	360	400

Contingency tables:

Variable A Time of the week	Variable B Had a Headache?		Var A Tot
	Yes	No	
Thursday	a=30 Row rf=a/300= 10% Col rf=a/60= 50%	b=270 Row rf=b/300= 90% Col rf=b/340= 79%	300
Saturday **NOTE! → New #s in this row**	c=30 Row rf=c/100= 30% Col rf=c/60= 50%	d=70 Row rf=d/100= 70% Col rf=d/340= 21%	100
Variable B Totals	a+c=60 Row rf=60/400= 15%	b+d=340 Row rf=340/400 =85%	400

- Are the row rdf's different?
  – i.e., (% getting vs. not getting headaches on Thursday) vs. (% getting vs. not getting headaches on Saturday)?
  – 10% headaches on Thursday vs. 30% headaches on Saturday
- Are the column rdf's different?
  – i.e., (% called on Thursday vs. Saturday for those with headaches) vs. (% called on Thursday vs. Saturday for those without headaches)?
  – 50% of headache sufferers called on Thursday vs 79% of those without a headache called on Thursday
- **Moral**:
  – In a 2x2 table, in order for there to be a relationship, all you need to do is show either
    - **that the row <u>relative</u> frequencies differ within either column**
    - **or, that the column <u>relative</u> frequencies differ within either row**
  – You'll come to the same conclusions (about the presence or absence of a relationship) either way.

Q: What's the point of all this?

- A: You need certain information in order to draw a valid conclusion about whether there is or isn't a relationship between two variables.
- Q: What information?
- A: **The information that would permit a comparison of the appropriate relative frequencies.**
- Q: So what?
- A: **It is <u>very</u> common (especially in the media and in informal communication) that people make claims of relationships without having or giving you enough of this information.**

Other important and related concepts

Perfect relationships (or, so close to 1.0 it hurts!)

- A **perfect relationship** between A and B is one where knowing one's standing on Variable A tells you exactly what one's standing on Variable B is
- if the relationship between A and B were held to be "the rule", there are no "exceptions to the rule" for perfect relationships
- Synonyms: deterministic relationship

Other important and related concepts

- What would a perfect relationship look like in a simple 2 x 2 contingency table?
  – Either a & d > 0 and c & b=0
  – or, the other way around, i.e., c & b > 0 and a & d=0
- For example, to make this concrete, what would a perfect relationship between smoking and lung cancer look like?
  – Is either really the case?
- Give me some good examples of perfect empirical relationships (observable rules for which there are no exceptions)

Variable A	Variable B		Variable A Totals
	High/ Yes	Low/ No	
High/Yes	a	b	e=a+b
Low/No	c	d	f=c+d
Variable B Totals	g=a+c	h=b+d	N=a+b+c+d

Smokes?	Die of Lung Cancer?		Variable A Totals
	Yes	No	
Yes	100	0	100
No	0	100	100
Variable B Totals	100	100	200

Smokes?	Die of Lung Cancer?		Variable A Totals
	Yes	No	
Yes	0	100	100
No	100	0	100
Variable B Totals	100	100	200

Other important and related concepts

- Moral?
  - **<u>Almost no</u> relationships in nature (and hence, of scientific interest) are perfect relationships**
    - Why?
    - **Because most things in nature (like lung cancer) have <u>multiple</u> causes**
  - But A relationship does <u>not</u> have to be perfect to be useful
    - Smoking's relationship to lung cancer is useful to know about, even though it is not perfect
    - SAT score's relationship with college success is useful, even though it is not perfect
  - However, as we will see, the most common errors we make in judging relationships **<u>assume exactly the opposite</u>—that <u>all</u> relationships are perfect ones**

Causal Relationships

- When we say that variable A and B are related
  -We are **<u>NOT</u>** saying that A _____________ B

- Many relationships in nature are non-causal:
  - Example: Preventing the Bubonic plague
  - "The pope [Clement VI], in his quarters at Avignon, sat between two large fires. They thought that this would purify the 'bad air' which most blamed for the spread of the plague. Although there was no bad air, the fires actually did prevent the plague…"
  - So it was concluded that bad (cold, damp) air was the cause of the plague
  - The actual cause was a bacteria, carried by fleas that live on rats, but bite humans, too

- Why, if the problem wasn't "bad air" did the fires make any difference?
  - the heat drove away the rats, carrying the fleas that carry the disease
  - but the cause of the disease was not cold, "bad air"
  - This was an example of what some people call "accidental science", or a discovery made from superstition, or by accident.

- Morals:
  - **Not all relationships are causal relationships**
    - air temperature is related to risk of plague, but
    - air temperature does not cause plague
  - **You need more than to show an association between two variables to show that one causes the other.** We'll explain what more you need later.

Errors in judging simple relationships: 1. The "single-cell" error

This comes in many forms. For example, **testimonials**

- The satisfied customer
- Q: what relationship is implied by this testimonial evidence?
  - **A:**
    -
    -
  - What would a contingency table look like for such a relationship?

Variable A: What treatment?	Variable B: How much hair?	
	Same	More
AMG Herbal		
Head & Shoulders		

Errors in judging simple relationships: 1. The "single-cell" error

This comes in many forms. For example, **testimonials**

- The satisfied customer
- What's wrong with such evidence?
  -
  -
  -
  - Advertiser's defense?
  -
  -

More "single-cell" errors

- Single (or multiple) "**person on the street" interviews**
  - Example: "Do Ivy League graduates get higher salaries than equally qualified Big Ten graduates?"
  - What's the implied relationship?
  - What would a contingency table look like?
  - Reporters often solicit opinions or testimonials from isolated people
    - "My friend Bob graduated from Yale and was paid $200,000 to start"
  - What do we know and what's missing?
  - Would it help if others had similar stories?

Variable A: Graduate from	Variable B: Starting salary above national median?	
	Yes	No
Ivy League school		
Big Ten school		

More "single-cell" errors

- **Reasoning from one's own experience**
  - My Dad says, "Don't be a lazy bum. If you work hard, you'll succeed. Look at me. I worked hard all my life and now, I'm the president of my company."
  - What's the implied relationship?
  - What would a contingency table look like?
  - Assume that Dad's absolutely right about his own experience. What's available and what's missing?
  - Would it help if others had similar stories?
    - "All my Dad's friends work very hard and are very successful"

Variable A: Work hard?	Variable B: Financially successful?	
	Yes	No
Yes		
No		

More "single-cell" errors

Case studies

- Case study methods involve an in-depth examination of a single instance or event: **a case**
- e.g., "Lorenzo's Oil"
- What's the implied relationship?
- What would a contingency table look like?
- Assume that Lorenzo really did have a good result. What's available and what's missing?
- Does it really work? Is there really a relationship?
  - Follow up "careful and systematic" observation showed that
  - "The oil doesn't seem to work for people who are already ill - but it does seem to help prevent illness in those whose genes make them vulnerable to developing symptoms."

Variable A: Take Lorenzo's oil?	Variable B: Disease slowed or stopped?	
	Yes	No
Yes		
No		

Another "single-cell" error: Letting an exception disprove the rule

- Several common forms:
- Form 1. **negative testimonials**-- dissatisfied customers
  - "I'll never buy another American built car. This Ford of mine is always in the shop."
    - What's the rule/relationship in question?
    - What do we know and what don't we know?
      - » We know that not all American built cars run well, BUT
      - » We don't know whether there is any general relationship
- Moral:

Variable A: Drive a US built car?	Variable B: Fewer than average rate of repairs?	
	Yes	No
Yes	?	1
No	?	?

Another "single-cell" error: Letting an exception disprove the rule

Form 2: **"person who.." errors**

- e.g., "Smoking can't be that unhealthy. I know a **person who** smoked 4 packs a day and died of old age at 92"
  - What's the rule/relationship in question?
  - assuming that the data is correct, What do we know and what don't we know?
    - we know that not all smokers die of lung cancer BUT
    - we don't know from this **"person who"** evidence whether there is any general relationship

Variable A: Smokes?	Variable B: Die of lung cancer?	
	Yes	No
Yes	?	1
No	?	?

Another "single-cell" error: Letting an exception disprove the rule

Form 3: **Historical counterexamples**

- Example, "People with more experience in government make better presidents."
  - "James Buchanan had lots of experience and was a terrible president, and Abraham Lincoln had very little experience, and was one of our greatest presidents"
- What's the rule/relationship in question?
- What do we know and what don't we know?
  - we know that not all good presidents had lots of experience, and
  - not all bad presidents had little experience, BUT
  - we don't know from these counterexample whether or not there is any general relationship
- Look at similar arguments commonly made by historians for many more examples

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?	
	Good	Bad
Yes	?	1
No	1	?

Roots of the single cell error

- Q: Why do we commit the single cell error?
- A:
  - Consider the relationship between experience and effectiveness of Presidents
  - Consider the first 40 presidents (may be hard to judge effectiveness of 4 most recent presidents)
    - divide them into the more and less experienced halves
    - and the more or less effective halves (assume that we can make this measurement)
  - If every real rule has no exceptions, then we have only 3 possibilities:
    - Experience and effectiveness always go together
    - Inexperience and effectiveness always go together
    - There's no relationship at all
  - What would these look like in contingency tables?

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	20	0	20
No	0	20	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	0	20	20
No	20	0	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	10	10	20
No	10	10	20
Variable B totals	20	20	40

Roots of the single cell error

- Suppose S. Hess' evaluations of Buchanan and Lincoln are correct. What have we learned?
  - -
- Suppose I can find a single President who is both experienced and effective (e.g. George Washington). What more have we learned?
  - -
  - -
  - -

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	20	0	20
No	0	20	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	0	20	20
No	20	0	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	10	10	20
No	10	10	20
Variable B totals	20	20	40

Roots of the single cell error

- Q: What's wrong with this reasoning?
- A1: Nothing--IF our initial assumption—that if there was any relationship at all, it had to be a perfect one—is true.
- A2: Everything--if that assumption is false.
- But, as we've shown, in fact, almost no relationships are perfect
- It is still quite possible that
  - there is a positive relationship—the more the experience, the more effective the president,
  - <u>and</u> that Buchanan and Lincoln are a couple of the rare exceptions
- In a world of imperfect relationships (like the one we live in)
  - **exceptions can <u>only</u> show that a relationship isn't perfect,**
  - **but <u>NOT</u> that there is no relationship**

Variable A: Relatively lots of experience?	Variable B: Good or Bad President? Yes	No	Variable A totals
Yes	18	2	20
	90.00%	10.00%	
No	2	18	20
	10.00%	90.00%	
Variable B totals	20	20	40

Factors that affect making the "Single cell" error

- Q: (From earlier lectures...) If we view this error as a heuristic, a "short cut", when are we more likely to make the error?
  - A1:
  - A2:

Other factors that affect making the "Single cell" error

- The <u>vividness</u> of the "single cell" evidence.
  - the more vivid and attention grabbing the cases/ testimonials/instances are, the more influence they have on our belief that there <u>is</u> a relationship

Shown best by the research of R. Nisbett & colleagues

- e.g. 1, Hamill, Wilson, & Nisbett (1980)
- Consider the following question: "Is being a prison guard related to how humane (i.e., how considerate, compassionate, sensitive) a person is?"
- How would you frame this in a simple 2x2 contingency table?
- Hamill et al.'s Ps saw a filmed interview with "a person who" was allegedly a prison guard
- The person acted in one of 2 ways (in different conditions)
  - 1) very nasty, disagreeable, and inhumane
  - 2) or very friendly, agreeable, and humane.
- Ps then asked their opinion about the humaneness of prison guards (i.e., is there a relationship?)

Factors that affect making the "Single cell" error: The <u>vividness</u> of the "single cell" evidence

- What would you expect?
  - That Ps in condition 2 would think that prison guards were generally more humane than Ps in condition 1
- But there's more.... Hamill et al. also told some Ps that the nasty guard they saw was very <u>typical</u> of all prison guards
  - Would this likely change how they responded? What do we know?
    - Are most people we know (non-prison guards) nasty, inhumane, and disagreeable? How many out of 100?
    - Now, if we know that the prison guard in the interview <u>was</u> inhumane AND we are told that they are very typical
    - So, out of 100 prison guards, how many should be inhumane?
      - at least more than half. Let's say 2/3
  - So, is there a relationship?
    - Yes, and that's what the Ps said.

Variable A: Prison guard?	Variable B: Inhumane person? Yes	No	
Yes	[redacted]	[redacted]	
No	[redacted]	[redacted]	

Factors that affect making the "Single cell" error: The <u>vividness</u> of the "single cell" evidence

- But there's still more.... Hamill et al. also told some other Ps the nasty guard in the interview was very <u>atypical</u> of all prison guards
  - Should this change how they responded?
  - What do we know?
    - We know that the prison guard in the interview <u>was</u> inhumane AND that they are very atypical.
    - So, in 100 prison guards, how many <u>should be</u> humane?
      - at least a majority, say 2/3?
      - If he's "<u>Very atypical</u>", then a large majority, say 95%?
  - So, is there a relationship?
    - No, or <u>at least</u>, a much weaker one.

However, what did Hamill et al. find?

Variable A: Prison guard?	Variable B: Inhumane person? Yes	No	
Yes	[redacted]	[redacted]	
No	[redacted]	[redacted]	

Factors that affect making the "Single cell" error: The <u>vividness</u> of the "single cell" evidence

- Consider the following question: "Is smoking related to lung cancer?"

  -1964 first Surgeon General's report was published.
- However, for at least two decades after that report was published, there was <u>no</u> per capita decrease in cigarette consumption

  -things have improved in the US more recently
- However, <u>one profession</u> did reduce their smoking after the report came out. Can you guess which one?

Variable A: Smoke?	Variable B: Get Lung Cancer Yes	No	
Yes	[redacted]	[redacted]	
No	[redacted]	[redacted]	

Factors that affect making the “Single cell” error

Some implications of the “vividness” effect:

- We’re more likely to pay attention to vivid, testimonial evidence than pallid base rate information
  - E.g. 1. If we’re trying to decide whether or not to buy a Subaru, which evidence will be more convincing:
    - statistics that show that Subaru cars are above average in rate of repairs, OR
    - some **person who** tells us that their Sabaru is always breaking down?
  - E.g. 2. If we’re trying to decide whether it’s safer to fly or drive to New York, which evidence will be more convincing:
    - statistics that show you’re 37 times more likely to be killed in each mile you travel in your car vs. in a commercial airliner, OR
    - pictures of the plane crashes of 9/11?

Factors that affect making the “Single cell” error

Further implications of the “vividness” effect:

- Events which are rare or unusual catch our attention and are more vivid in our memory.
- This can help explain why stereotypes form
- **Stereotypes** are beliefs we hold about members of social groups
  - Sometimes they can contain a kernel of truth
    - “Basketball players are tall”
  - Often they do not, even on average
    - “Women are more easily frightened then men”
      - some evidence that women report more fear
      - but direct measures (e.g., behavior; arousal) shows no mean difference
  - They are almost always overgeneralized
    - They are assumed to apply to everyone in the group, with few or no exceptions
- A major question in social psychology is how such stereotypes form
- Hamilton & Gifford (1976) showed that our tendency to make the single-cell error more with vivid evidence can also lead to stereotypes.

Illusory Correlations: Hamilton & Gifford (1976)

- Ps given 1100 cards describing individuals.
- Each card has 2 bits of information
  - what group the person belongs to
    - a majority group (91%)
    - a minority group (9%)
  - whether the person has committed a crime
    - most people don’t (95%)
    - but a few do (5%)
- H&G’s data can be summarized in a contingency table
- Is there a relationship between group membership and crime in this table?
  -

Group Membership	Commits a crime?		Group Membership Totals
	No	Yes	
Majority	950	50	1000
Minority	95	5	100
"Behavior" Totals	1045	55	1100

Illusory Correlations: Hamilton & Gifford (1976)

- BUT, Hamilton & Gifford’s Ps generally concluded that there was a relationship

-

-Why?

-
-
-

-

Group Membership	Commits a crime?		Group Membership Totals
	No	Yes	
Majority	950	50	1000
Minority	95	5	100
"Behavior" Totals	1045	55	1100

How personal the evidence is

- If “vivid” information is more likely to lead to the “single cell error”, what makes evidence especially vivid to us?
- One thing is when it happens to us, personally
  - For example, do we attach different meaning to coincidences (e.g., getting struck by lightning) that happen to us than when exactly the same thing happens to someone else?
- Falk (1989)
  - Students in 3 classes recorded their birthdays; in Classes B & C, each student was also assigned a “random birthday”
  - All matches in the class of real or random birthdays were determined
  - Everybody in the class was classified in one of 3 categories:
    - Self involved in a “meaningful” match (someone in the class had the same birthday as you did)
    - Self involved, but in a “meaningless” match
      - Class A: the person sitting next to you matched birthdays with someone in the class
      - Classes B & C: your “random” birthday matched someone else’s “random birthday”
    - Everybody else (no matches of any kind)
  - Students then rated how surprising the results were

How personal the evidence is

– Students then rated how surprising the results were

Group/event	Agent of coincidence: Self (involved subjects) — Meaningful	Meaningless	Others (uninvolved subjects)
A/birthdays (74)	5.75 (8)	4.75 (8)	4.41 (58)
B/birthdays (53)	11.50 (6)	7.43 (7)	5.15 (40)
C/birthdays (72)	9.95 (21)	8.00 (7)	6.91 (44)
C/name-sums (72)	9.25 (8)	—	6.66 (64)

Note. ns are in parentheses.
[a]See Guttman (1988).

1 = not at all surprising
20 = very surprising

Remember the Confirmation Bias?

- What is this bias?
  - -
  - Shown dramatically in Wason's rule-discovery task
  - 2-4-6 https://www.youtube.com/watch?v=vKA4w2O61Xo
    - people tend to select only confirming/positive tests of their original rule (e.g., 8-10-12)
    - and never/rarely look at possible disconfirming/negative tests (e.g., 2-1-5)
- When we're looking for a relationship,
  - -
  - -
- So, what if we're trying to determine if there were a relationship between smoking and lung cancer
- So, if we are trying to decide whether A & B are positively related, do we pay special attention to the "present-present" (Yes/Yes; Hi A/Hi B) cell?

Variable A: Smoke?	Variable B: Get Lung Cancer		
	Yes	No	
Yes	[redacted]	[redacted]	
No	[redacted]	[redacted]	

Smedslund (1969)

- One group of nurses were given data from 100 patients and asked to determine whether there was a relationship or correlation between a symptom and a disease.
- If you summarize the data for the 100 cases, it looked like this contingency table
  - Is there a relationship?
    - there is a slight negative relationship here
    - 37 out of 70 or 53% of those who have the symptom have the disease, whereas 17 out of 30 or 57% of those without the symptom have the disease.
    - There's a tendency for the symptom to indicate that one does not have the disease

Variable A: Symptom Present?	Variable B: Disease Present?		Variable A totals
	Yes	No	
Yes	37 52.86%	33 47.14%	70
No	17 56.67%	13 43.33%	30
Variable B totals	54	46	100

Smedslund (1969)

- A second group of nurses were given data from 100 different patients and also asked to determine whether there was a relationship or correlation between a symptom and a disease.
- If you summarize the data for these 100 cases, it looked like the bottom contingency table
  - How is it different?
  - Is there a relationship here?
    - yes, a positive relationship
- **However, most of the nurses in both groups (about 85%) concluded that A and B were positively related**
  - (i.e., that the symptom signaled the disease)
- Why?
  - Follow up analyses showed that the best predictor of Ps judgment of whether or not there was a relationship was the proportion of cases in the Yes/Yes cell.
  - 37 in these two groups, and the same
  - **The nurses (and we) tend to pay most attention to just one cell—the Yes/Yes cell**
- This suggests that our judgments about relationship depend on how we ask or "frame" the question.
  - For example, the nurses would have looked at different data if the question were "Is not having the symptom related to having the disease?"

Variable A: Symptom Present?	Variable B: Disease Present?		Variable A totals
	Yes	No	
Yes	37 52.86%	33 47.14%	70
No	17 56.67%	13 43.33%	30
Variable B totals	54	46	100

Variable A: Symptom Present?	Variable B: Disease Present?		Variable A totals
	Yes	No	
Yes	37 52.86%	33 47.14%	70
No	13 43.33%	17 56.67%	30
Variable B totals	50	50	100

Framing relationships: Crocker (1982)

- Crocker asked her participants to consider 2 questions:
  - Question 1a: Is there a relationship between working out the day before a tennis match and winning that match?
  - Question 2a: Indicate any information needed to answer Question 1a; what information in this contingency table must you have to make an accurate judgment about Question 1a?
- How do you think her participants answered?
  - Write in the % of participants who said each cell was needed
  - -

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	a [redacted]	b [redacted]
No	c [redacted]	d [redacted]

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes		
No	[redacted]	[redacted]

Framing relationships: Crocker (1982)

- Crocker then asked another group to consider 2 questions:
  - Question 1b: Is there a relationship between working out the day before a tennis match and losing that match?
  - Question 2b: Indicate any information needed to answer Question 1; what information in this contingency table must you have to make an accurate judgment about Question 1?
- Note: Question 2b is really the same question as Question 2a
- Now how do you think her participants answered?
  - Write in the % of participants who said each cell was needed
  - -

Moral:

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	a [redacted]	b [redacted]
No	c [redacted]	d [redacted]

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes		
No	[redacted]	[redacted]

Other situational factors

- We tend to make the single cell error more when we receive the data piecemeal than all at once (e.g., summarized in a 2x2 table)
  - Smedslund found that the nurses made the error less when they received the table than when they simply looked through 100 folders
- We tend to make the error more when we are trying to profit from it vs. simply trying to figure out if there is a rule—rewards/hits are especially (and too) "vivid"
  - Gamblers are more likely to (falsely) conclude there is a rule governing when they win (e.g., blow on the dice) than onlookers

Individual differences

- So far, we've focused on <u>when</u> we're likely to make the single cell error
- But <u>who</u> is more likely to make it?
  - are there individual differences in the tendency to make this error?
- There are, including
  - those chronically likely to rely on heuristic thinking
    - low Need for Cognition
  - other factors, including...
    - depression &
    - age

Age & single cell errors: Kuhn et al. (1985)

- Students in different grades were asked
  - Suppose I tell you about 6 car owners (e.g., 6 testimonials) who use an additive and their cars always start right up
  - Knowing this, can you conclude that using the additive helps?
- What is the correct answer?
- To the right is how Kuhn et al.'s Ps answered
- Is there a relationship between grade and getting the answer right?
- Why might this be true?

Variable A: Use the additive?	Variable B: Car starts?	
	Yes	No
Yes	6	
	[redacted]	
No		
	[redacted]	

TABLE 2
Judgments Following Initial (Single-cell) Presentation

	Yes	No	Can't Tell	Total
Grade 4	14	1	2	17
Grade 7	13	0	3	16
Grade 10	10	2	3	15
College	5	0	10	15
(total)	42	[illegible]	18	63

Depression and single cell errors: Alloy & Abramson (1979)

- Studied ability of depressed and non-depressed people to accurately detect presence or absence of a relationship
  - the degree of contingency between their responses (pressing or not pressing a button) and an environmental outcome (onset of a green light).
- Depressed Ps were fairly accurate, whether or not they actually had any control
- Nondepressed Ps
  - Overestimated the degree of contingency between their responses and outcomes when contingent outcomes were frequent and/or desired
  - And underestimated the degree of contingency when contingent outcomes were undesired
  - Saw more control than there was when they experienced or wanted control and saw less control than there was when they didn't like the outcomes.

Single row/column errors

- These are errors that result from having data in one row (or column) of the contingency table, but not the other
- Takes a number of forms, including **superstitious behavior**
- This illustrates another name for the single row/column error: the **"no control group" error**
  - Here, what's the missing control group?

Variable A: Drink the night before an exam?	Variable B: Pass the exam?	
	Yes	No
Yes	18	2
	90.00%	10.00%
No		
	[redacted]	

Single row/column errors

- **Inflation and elderly suicide in Argentina**
  - In 2001, Argentina's currency and economy collapsed
  - NPR report of elderly Argentines, living on now-worthless pensions, committing suicide
- What's wrong with concluding that economic problems have resulted in higher suicide rates among the elderly?

Variable A: After vs. Before Economic Crisis?	Variable B: Death by Suicide?	
	Yes	No
After	40	60
	40.00%	60.00%
Before		
	[redacted]	

Roots of the single row/column error

- Earlier we argued that root of the single cell error was the false assumption that "all relationships are perfect relationships"
- I think the root of the single row/column error is **"the data in all contingency tables are symmetric (and hence, usually indicate a <u>strong</u> relationship)"**
  - if there were symmetry here, what would that imply about those who <u>don't</u> take the medicine?
  - (you'll need to fill in the table)
  - When would this rule NOT imply a relationship?
    - When there was no clear trend in the known row (e.g., 50% who take the medicine get better; 50% do not)
- Of course, not all such data are symmetric, so this heuristic leads to errors

Variable A: Take the tonic?	Variable B: Better in a week?	
	Yes	No
Yes	8	2
	80.00%	20.00%
No		

Factors affecting single row/column errors

- **Restriction of Range**
- Your informal observation suggests that this is odd (taller people tend to be heavier people), so you ask for more detail
  - like what?
- What's the problem?
- What has it to do with "single row/column" error?
- Moral:
  - Whenever someone claims their data shows no relationship, make sure that there is no restriction of range on either variable

Variable A: Height	Variable B: Heavy? >190 lbs	<190 lbs	Variable A totals
Tall	5	5	10
	50.00%	50.00%	
Short	5	5	10
	50.00%	50.00%	
Variable B totals	10	10	20

Factors affecting single row/column errors

- **Restriction of Range: Floor/Ceiling Effects**
- **Ceiling effect**: scores in one condition are as high as they can go; no chance to find a higher score in the other condition
- **Floor effect**: Scores in one condition are as low as they can go; no chance to find a lower score in the other condition.
- Moral:
  - when someone claims no relationship between 2 variables, check to make sure there are no floor or ceiling effects for either variable.

Variable A: % Helium	Variable B: Balloon goes high? Yes	No	Variable A totals
100%	0	10	10
	0.00%	100.00%	
50%	0	10	10
	0.00%	100.00%	
Variable B totals	0	20	20

Factors affecting single row/column errors

- Same as before (for single cell error)
  - low ability or motivation to think carefully
  - "vivid" row/column information
  - individual differences,

Age & single row/column errors: Kuhn et al. (1985)

- Here's how Kuhn et al.'s Ps answered
  - Note a new response: "Sometimes" = not a perfect relationship
- Is there a relationship between grade and getting the answer right?

Variable A: Use the additive?	Variable B: Car starts? Yes	No
Yes	6	2
	[redacted]	
No		
	[redacted]	

TABLE 3
Judgments Following Second (Two-cell) Presentation

	Yes	Sometimes	No	Can't Tell	Total
Grade 4	8	4	3	2	17
Grade 7	7	3	4	2	16
Grade 10	5	3	2	5	15
College	5	1	1	8	15
(total)	25	11	10	17	63

More errors of judging relationships: Diagonal error

An example

- Suppose you're interviewing applicants to be your TV station's weatherperson
- An applicant says, "In my last job, I correctly predicted the weather 82% of the time"
- Can I conclude that they are a good weatherperson—i.e., there's a relationship between this person's forecasts and the actual weather?
- What do we know?
  - That (a+d)/N = 82%
  - That the applicant's success/hit rate is high
  - But does this guarantee that there is a positive relationship?
  - Consider these three sets of data...
  - For each,
    - what's (a+d)/N?
    - is there a relationship?

Variable A: Predicted Weather	Variable B: Actual weather Rain	No Rain	Variable A totals
Rain	81	9	90
	90.00%	10.00%	
No Rain	9	1	10
	90.00%	10.00%	
Variable B totals	90	10	100

Variable A: Predicted Weather	Variable B: Actual weather Rain	No Rain	Variable A totals
Rain	81	18	99
	81.82%	18.18%	
No Rain	0	1	1
	0.00%	100.00%	
Variable B totals	81	19	100

Variable A: Predicted Weather	Variable B: Actual weather Rain	No Rain	Variable A totals
Rain	0	9	9
	0.00%	100.00%	
No Rain	9	82	91
	9.89%	90.11%	
Variable B totals	9	91	100

Origin of the diagonals error

- We suggested that single cell error stemmed from assuming that all relationships were perfect
- What would that mean for diagonal errors?
  - What would the diagonals look like if the relationship were perfect?
  - 100% "hits"
  - 0% "misses"
  - Diagonal error says, "if perfect = 100% hits, then strong positive relationship = "lots of hits"? or "more hits than misses"?
  - Wrong!

Variable A: Predicted Weather	Variable B: Actual weather Rain	No Rain	Variable A totals
Rain	90	0	90
	100.00%	0.00%	
No Rain	0	10	10
	0.00%	100.00%	
Variable B totals	90	10	100

Origin of the diagonals error

- Can even have a positive relationship when there are more "misses" than "hits"
  - here, (50+49)/550= 18% "hits"
  - (1+450)/550= 82% "misses"
  - but there's still a positive relationship
- Moral: **Can't tell if or what kind of relationship there is when all you know is information from one set of diagonal cells (e.g., % hits, entries in diagonal only)**

Variable A: AIDS patient takes AZT?	Variable B: Survive 10 years from AIDS diagnosis? Yes	No	Variable A totals
Yes	50	450	500
	10.00%	90.00%	
No	1	49	50
	2.00%	98.00%	
Variable B totals	51	499	550

More errors of judging relationships: Marginals error

An example

- Suppose you're told, "Although 80% of the dorm residents took Vitamin C last winter, half of them still got colds" (you are not provided a contingency table)
- What's the implication?
  - That Vitamin C must not work if so many people who take it still get colds
- Can I conclude that there is or isn't a relationship between taking Vitamin C and getting a cold?
- What do we know?
- What don't we know?
- Consider the 3 tables to the right
  - what are the marginal distributions?
  - is there a relationship?
- Moral:
  - you cannot tell whether or not there is a relationship simply by knowing the marginal distributions

Variable A: Take Vitamin C?	Variable B: Get a cold? Yes	No	Variable A totals
Yes	250 62.50%	150 37.50%	400
No	0 0.00%	100 100.00%	100
Variable B totals	250	250	500

Variable A: Take Vitamin C?	Variable B: Get a cold? Yes	No	Variable A totals
Yes	200 50.00%	200 50.00%	400
No	50 50.00%	50 50.00%	100
Variable B totals	250	250	500

Variable A: Take Vitamin C?	Variable B: Get a cold? Yes	No	Variable A totals
Yes	150 37.50%	250 62.50%	400
No	100 100.00%	0 0.00%	100
Variable B totals	250	250	500

Origin of Marginals error:

- Q: Why do we make this error?
- A: **Because we assume all relationships are perfect (or nearly perfect) ones**
- Q: What would marginals look like if there were a perfect relationship?
- A: **Marginal distributions would be identical.**
- Hence, if they're NOT identical, can't be a perfect relationship, BUT can still have an imperfect relationship

Variable A: Take Vitamin C?	Variable B: Get a cold? Yes	No	Variable A totals
Yes	400 100.00%	0 0.00%	400
No	0 0.00%	100 100.00%	100
Variable B totals	400	100	500

Other biases in detecting relationships: Our <u>expectations</u> about relationships

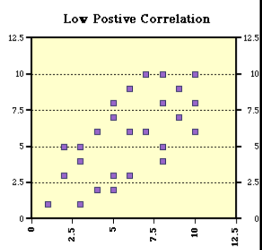
- For example, the use of invalid projective tests, such as the "Draw-a-Person" test or the Rorschach test.
  - Clinicians <u>believe</u> that people with certain disorders (e.g., paranoia) are related to responses (e.g., draw figures with big eyes).
  - Chapman & Chapman (1967) gave clinicians drawings with diagnoses attached and asked to discover what drawing features tended to go with what disorders.
  - Although there was <u>no</u> association in the data, the clinicians saw the relationships that they expected to see (e.g., paranoids tend to draw people with big eyes).
  - Students (without any clinical experience, but with similar expectations) did the same thing
- Why?
  - -
- Implications (e.g., for advertisers)?
  - -

Other biases in detecting relationships: Our expectations about relationships

- Q: What if we already <u>believe</u> that there IS a relationship?
  - $|r| > 0$.
- A:
  - -
  - -
- Q: What if we have no reason to think that there is a relationship?
  - i.e., Expected $r = 0$
- A:
  - -
  - -
- Implications for the last question?

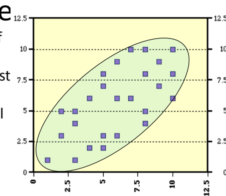
C x C case

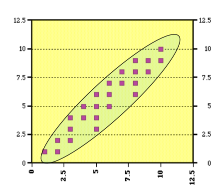
- Use *scatter plot* to display data
  - a 2-dimensional graph with
  - one variable (e.g., A) on one axis, &
  - other variable (e.g., B) on other axis
  - and every person (i.e., pair of scores) represented by a point on the plot
- Use some statistic that summarizes covariation to detect and summarize relationship
  - Usually *correlation coefficient*
  - r varies from -1.0 to 0 to +1.0
  - $r = 0$ means no LINEAR relationship
  - $r > 0$ or $r < 0$ means there is some LINEAR relationship
  - closer r is to -1 or +1, the stronger the relationship
- "Shape" of scatter plot indicates magnitude of correlation



C x C case

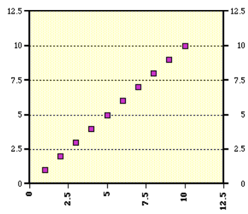
- Shape of scatter plot indicates magnitude of correlation
  - Think of drawing an oval to capture (most of) the points
  - The "thinner" and more "tilted" the oval is, the stronger the relationship
  - Top plot
    - as X goes up, Y tends to go up (positive relation)
    - but at any given X value, fairly wide range of possible Y values, so X cannot predict Y very exactly (weak relationship; low positive r)
  - Lower Plot
    - as X goes up, Y tends to go up (positive relation)
    - but at any given X value, narrower range of possible Y values, so X can predict Y better (strong relationship: high positive r)

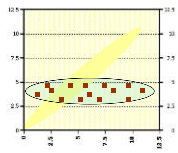


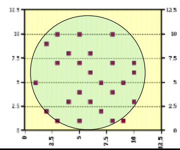


C x C case

- Shape of scatter plot indicates magnitude of correlation



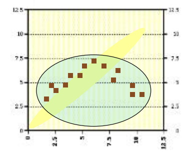




C x C case

- possible for r = 0 (and no <u>LINEAR</u> relationship),
 but there <u>could</u> still be a relationship, but it would be a NONLINEAR relationship

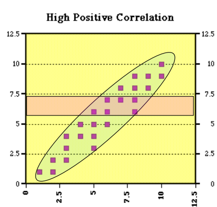
- consider this scatter plot
  - what's the best fitting oval?
    - so, r = 0
    - no LINEAR relationship, but
  - there <u>is</u> a relationship, an "inverted U"



- Moral:
  - **When someone says *r* = 0, it only means that there is no LINEAR relationship**
  - **There might still be a non-linear one**

CxC case: Restriction of range

- Even if there is a strong relationship overall, you won't detect it if you focus on only a thin slice of either variable
- What if we only consider the observations in this slice?
  - what's *r*?
  - Moral: restriction of range➔ low/no association, EVEN WHEN THERE MAY BE A STRONG ONE WITHOUT A RESTRICTION OF RANGE
- Likewise, you may not detect a genuine relationship if there are ceiling or floor effects



What's a critical thinker to do?

<u>Checklist:</u>

A. Is a relationship or lack of relationship between two variables suggested?
- What are the two variables?
- Do you have enough information to conclude that there really is a relationship?
  1. If both variables are dichotomous (or have just a few values), can you compare the row or column relative frequency distributions?
  2. If one variable is continuous and the other categorical, do you have (or can you get) mean values of the former for each level of the latter?
  3. If both variables are continuous
     - Has the data been summarized fully, e.g.
       - with some summary statistic (e.g., Pearson's ***r***, rank-order ***r***, etc.)?
       - with a scatterplot?

B. If there is a claim of <u>no</u> relationship...
- Are restriction of range, ceiling effects, or floor effects likely problems for either variable?
- If the evidence is that some measure of <u>linear</u> relationship (e.g., r) is small, is a non-linear relationship plausible?

Practice in judging relationships from word problems:

- The mean per capita contribution to the United Way from men is $54, the same as the overall mean for both genders.

- What do we know?
- Men mean= $54
- Overall mean= $54
- Woman mean=
  -
- Restriction of range or floor/ceiling effect plausible?

Practice in judging relationships from word problems:

- The rate of dropping out of the Herbert Hoover High School for girls is 12%, twice as high as the rate for boys
- What do we know and not know?

-
-
-
-

Variable A: Drops Out?	Variable B: Sex	
Yes	a	b
	a/g	b/h
No	c	d
	c/g	d/h
Variable B totals	g	h

Practice in judging relationships from word problems:

- Of the 100 heterosexual married couples stopped by the police after midnight on New Year's Eve, only 20 persons passed a breathalyzer test, but 18 of them were wives (and only 2 were husbands). Everyone in the car took a breathalyzer test (200 total people).
- What do we know and not know?
- We know enough to figure out that there IS a relationship

Variable A: Passed Breathalizer Test?	Variable B: Husbands	Wives	Variable A totals
Yes	2	18	20
	2.00%	18.00%	
No	98	82	180
	98.00%	82%	
Variable B totals	100	100	200