Econ 211

Prof. Jeffrey Naecker

Wesleyan University

Borrowing from Psychology

Psychology and Economics

- Behavioral and experimental economics owe much to the psychology literature
- Many of the foundation papers we will read in each section are from major branches of psychology:
  - social psychology
  - cognitive psychology
  - judgement and decision-making
- Many groundbreaking researchers in behavioral economics were trained as psychologist
  - Amos Tversky, PhD in Cognitive Psychology
  - Daniel Kahneman, PhD in Psychology
  - Dan Ariely, PhD in Cognitive Psychology

Discussion Questions

- What ideas and concepts are shared between psychology and economics? Do they have the same names or different?
- What methods to the two areas share? What methods are distinct?
- How is theory in economics different than theory in psychology?

Two Systems Model

- System 1: fast, associative
  - Make connections from similarity
  - Processing occurs automatically
  - Good at: pattern completion, emotional reactions, repetitive tasks

Two Systems Model

- System 1: fast, associative
  - Make connections from similarity
  - Processing occurs automatically
  - Good at: pattern completion, emotional reactions, repetitive tasks
- System 2: slow, deliberative
  - Uses symbols and rules
  - Processing occurs with conscious awareness
  - Good at one-shot learning, formal logic, educational knowledge, abstract theories

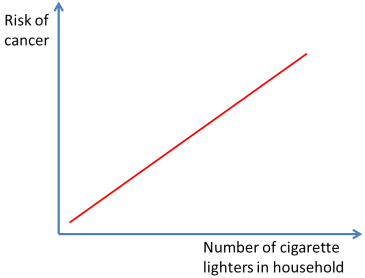
Two Systems Model

- System 1: fast, associative
  - Make connections from similarity
  - Processing occurs automatically
  - Good at: pattern completion, emotional reactions, repetitive tasks
- System 2: slow, deliberative
  - Uses symbols and rules
  - Processing occurs with conscious awareness
  - Good at one-shot learning, formal logic, educational knowledge, abstract theories
- Systems may interact alternatively or simultaneously
- Interaction moderated by mood, energy level, difficulty of problem, type of judgement

Experimental Design

Why Do We Need Experiments?

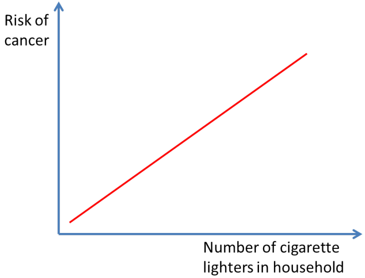
- Suppose we observe the following pattern in observational data:



- Can we conclude that cigarette lighters cause cancer?

Why Do We Need Experiments?

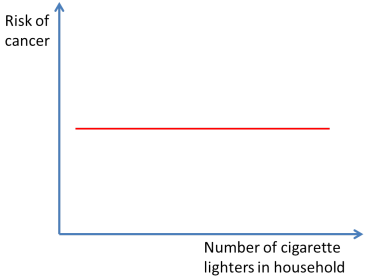
- Suppose we observe the following pattern in observational data:



- Can we conclude that cigarette lighters cause cancer?
  - No, correlation is not causation
  - More likely that there is a third variable (smoking) that causes the other two

Experiments Give Us Control

- Experiments allow the researcher to *control* all the variables (or at least control them more than in observational data)
- Experiments also the researcher to determine causality using *random assignment*
  - For example, if we randomly gave some people more cigarette lighters to have around the house, we would probably see that they have no *causal* relationship with cancer risk



Building Blocks

- Every experiment needs some *participants*, also know as *subjects* or *decision-makers*
- The most basic unit in an experiment is a *task* or *choice*
  - For example, deciding whether or not to buy a food item
- A *treatment* is a series of one or more similar tasks
  - For example, choosing whether or not to buy many food items, one at a time
- An *experiment* consists of one or more treatments
  - For example, could have one treatment where all food items are 25 cents, and another where all food items are at 75 cents

Treatment Variables

- Many experiment consist on only one treatment
- But many others have more than one treatment
  - The parts that differ between the treatments are called *treatment variables*
- A common experimental design is to have two treatment variables that can each take on two levels
  - This is called a *2-by-2* design
  - Example:
    - We are interested in responses to requests for donations to the local animal shelter
    - Vary whether picture of a dog is included, and vary whether a specific amount of money is asked for

	Amount Specified	**Amount Not Specified**
Photo	Treatment A	Treatment B
No Photo	Treatment C	Treatment D

Within- vs Between-Subjects Design

- In a *between-subjects* design, each subject completes only one treatment
- In a *within-subjects* design, each subject completes multiple treatments
- Is one of these designs better than the other?

Within- vs Between-Subjects Design

- In a *between-subjects* design, each subject completes only one treatment
- In a *within-subjects* design, each subject completes multiple treatments
- Is one of these designs better than the other? It depends:
  - A within-subjects design can suffer from *order effects*: the order the subjects do the treatments in can matter
  - However, a within-subjects design needs fewer subjects and gives more control of subject characteristics

Incentives

- Choices in experiments are typically incentivized with some kind of material *incentive* or *payoff*
    - Incentives may be cash, consumption goods, social image
    - Important to calibrate the size of the *stakes* to the task
    - Eg, bad idea to only pay a few cents for correctly solving an entire crossword puzzle
- In some cases, hypothetical stakes may be appropriate

Context

- Generally want to avoid contexts that include unncessary complications or distractions for subjects
- For example, consider food choice experiment
  - Primarily interested in testing the law of demand
  - Avoid confounds such as making decisions publicly observable
  - Unless, of course, this is the treatment variable I'm interested in
  - Controlling context is easier in lab experiments than field experiments
- Related issue: experimenter demand effect
  - Subjects may be influenced by what they think the experimenter wants them to do
  - Avoid using language that implies a value judgement or normative choice

Independence

- A famous (likely apocryphal) story:
  - Graduate student is studying how fast lacerations heal on the skin of mice, depending on whether or not mice have a certain genetic mutation
  - Advisor tells graduate student he should double his sample size for increased statistical power
  - Doubling the number of mice is very expensive and time-consuming
  - Graduate student is very clever: makes a second laceration on each mouse
- Has the number of observations doubled?

Independence

- A famous (likely apocryphal) story:
  - Graduate student is studying how fast lacerations heal on the skin of mice, depending on whether or not mice have a certain genetic mutation
  - Advisor tells graduate student he should double his sample size for increased statistical power
  - Doubling the number of mice is very expensive and time-consuming
  - Graduate student is very clever: makes a second laceration on each mouse
- Has the number of observations doubled?
  - Not really: the two cuts on each mouse are probably not *independent* from another
  - For example, if one cut heals quickly, I can expect the other cut will heal quickly too
  - Thus the second cut does not add new information

Common Pitfalls of Experiment Design

- Changing two treatment variables at the same time
- Poor choice of context
- Order effects not considered
- Independence of observations not considered
- Poor choice of incentives and stakes

Probability

Why Do We Need Probability?

1. Social scientists are interested in making predictions about future behavior
    - Sometimes the best prediction we can give is a likelihood of a certain event of interest happening
2. We need a benchmark to talk about rationality of behavior
    - Probability judgement: Process of assigning a number to an event that represents one's strength of belief that that event will occur
    - The rules of probability come from very general assumptions but still give powerful restrictions on how probability judgement should behave

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- A collection of one or more outcomes is an *event*, eg $E \in \Omega$
  - The event "roll a 3 or greater" is $E = \{3, 4, 5, 6\}$

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- A collection of one or more outcomes is an *event*, eg $E \in \Omega$
  - The event "roll a 3 or greater" is $E = \{3, 4, 5, 6\}$
- A probability function is a function P that assigns numbers between 0 and 1 (inclusive) to every possible event in Ω
  - $P(\{1\}) =$

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- A collection of one or more outcomes is an *event*, eg $E \in \Omega$
  - The event "roll a 3 or greater" is $E = \{3, 4, 5, 6\}$
- A probability function is a function P that assigns numbers between 0 and 1 (inclusive) to every possible event in Ω
  - $P(\{1\}) = \frac{1}{6}$

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- A collection of one or more outcomes is an *event*, eg $E \in \Omega$
  - The event "roll a 3 or greater" is $E = \{3, 4, 5, 6\}$
- A probability function is a function P that assigns numbers between 0 and 1 (inclusive) to every possible event in Ω
  - $P(\{1\}) = \frac{1}{6}$
  - $P(\{3, 4, 5, 6\}) =$

Probability Basics

- Let Ω be the space of all possible outcomes
  - Eg if we are rolling a single die, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- A collection of one or more outcomes is an *event*, eg $E \in \Omega$
  - The event "roll a 3 or greater" is $E = \{3, 4, 5, 6\}$
- A probability function is a function P that assigns numbers between 0 and 1 (inclusive) to every possible event in Ω
  - $P(\{1\}) = \frac{1}{6}$
  - $P(\{3, 4, 5, 6\}) = \frac{2}{3}$

Axioms of Probability

- An axiom is a mathematical rule that is assumed to be true
- We have just two axioms in probability theory:
  1. $P(\Omega) = 1$
     - You can think of this as "something is guaranteed to happen"
     - Eg $P(\{1, 2, 3, 4, 5, 6\}) = 1$ in dice example

Axioms of Probability

- An axiom is a mathematical rule that is assumed to be true
- We have just two axioms in probability theory:
  1. $P(\Omega) = 1$
     - You can think of this as "something is guaranteed to happen"
     - Eg $P(\{1, 2, 3, 4, 5, 6\}) = 1$ in dice example
  2. $P(A \text{ or } B) = P(A) + P(B)$ for any two mutually exclusive events A and B
     - Called the *addition axiom*
     - Eg $P(\{1\}) + P(\{3, 4, 5, 6\}) = \frac{1}{6} + \frac{2}{3} = \frac{5}{6}$

Where Do These Probabilities Come From?

- Frequentist perspective
  - Probabilities represent *long run* averages
  - Eg, dice: $P(4) = \frac{1}{6}$ because if I roll a die a very large number of times, $\frac{1}{6}$ of the time I will roll a 4

Where Do These Probabilities Come From?

- Frequentist perspective
  - Probabilities represent *long run* averages
  - Eg, dice: $P(4) = \frac{1}{6}$ because if I roll a die a very large number of times, $\frac{1}{6}$ of the time I will roll a 4
- Counting perspective
  - Break event space into small, distinct pieces that they are be equally likely to happen
  - Probabilities of more complex events can then be built by addition rule, since pieces are mutually exclusive
  - Example: what is the probability of flipping two heads in a row?
    - Break down into 4 equally likely events: HH, HT, TH, TT
    - Only 1 of these has two heads, so $P(HH) = \frac{1}{4}$

Conditional Probabilities

- We can define *conditional probability* $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$
  - This is the probability of A occurring given that we know B has occurred.
  - Example: $P(\text{winter}) = \frac{1}{4}$ but $P(\text{winter} \mid \text{snowed last weekend}) > \frac{1}{4}$

Conditional Probabilities

- We can define *conditional probability* $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$
  - This is the probability of A occurring given that we know B has occurred.
  - Example: $P(\text{winter}) = \frac{1}{4}$ but $P(\text{winter} \mid \text{snowed last weekend}) > \frac{1}{4}$
- We say two events A and B are *independent* if $P(A|B) = P(A)$

Conditional Probabilities

- We can define *conditional probability* $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$
  - This is the probability of A occurring given that we know B has occurred.
  - Example: $P(\text{winter}) = \frac{1}{4}$ but $P(\text{winter} \mid \text{snowed last weekend}) > \frac{1}{4}$
- We say two events A and B are *independent* if $P(A|B) = P(A)$
  - Dice example: Probability of rolling a six is independent of what you rolled previously

Conditional Probabilities

- We can define *conditional probability* $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$
  - This is the probability of A occurring given that we know B has occurred.
  - Example: $P(\text{winter}) = \frac{1}{4}$ but $P(\text{winter} \mid \text{snowed last weekend}) > \frac{1}{4}$
- We say two events A and B are *independent* if $P(A|B) = P(A)$
  - Dice example: Probability of rolling a six is independent of what you rolled previously
- Multiplication rule: For any two events A and B then $P(A \text{ and } B) = P(A|B)P(B)$
  - Comes from rearranging definition of conditional probability
  - If $P(A|B) = P(A)$, we say A and B are *independent*
  - Note that if A and B are independent, $P(A \text{ and } B) = P(A)P(B)$

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability
- Question: What is the probability that a patient has cancer given that you see a positive mammogram result?

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability
- Question: What is the probability that a patient has cancer given that you see a positive mammogram result?
  - Consider 100 women, all of whom we test for breast cancer

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability
- Question: What is the probability that a patient has cancer given that you see a positive mammogram result?
  - Consider 100 women, all of whom we test for breast cancer
  - 10 will have cancer (from the baseline rate)
    - Of these, 9 will return positive, 1 will return negative
  - 90 will not have cancer

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability
- Question: What is the probability that a patient has cancer given that you see a positive mammogram result?
  - Consider 100 women, all of whom we test for breast cancer
  - 10 will have cancer (from the baseline rate)
    - Of these, 9 will return positive, 1 will return negative
  - 90 will not have cancer
    - Of these 72 (90×0.8) will return negative, 18 will return positive

Bayes' Rule: Intuition

- Suppose you have the following information:
  - The baseline breast cancer rate in women is 10%
  - If a patient has breast cancer, a mammogram will return positive with 90% probability
  - If a patient does not have breast cancer, a mammogram will return positive with 20% probability
- Question: What is the probability that a patient has cancer given that you see a positive mammogram result?
  - Consider 100 women, all of whom we test for breast cancer
  - 10 will have cancer (from the baseline rate)
    - Of these, 9 will return positive, 1 will return negative
  - 90 will not have cancer
    - Of these 72 (90×0.8) will return negative, 18 will return positive
  - So among the positive tests, only $\frac{9}{27} = \frac{1}{3}$ are true positives for cancer

Bayes' Rule: Formalization

- Suppose we are considering two events A_1 and B
- From the definition of conditional probability, we have
  - $P(A_1|B) = \frac{P(A_1 \text{ and } B)}{P(B)}$
  - $P(B|A_1) = \frac{P(B \text{ and } A_1)}{P(A_1)} = \frac{P(A_1 \text{ and } B)}{P(A_1)}$

Bayes' Rule: Formalization

- Suppose we are considering two events A_1 and B
- From the definition of conditional probability, we have
  - $P(A_1|B) = \frac{P(A_1 \text{ and } B)}{P(B)}$
  - $P(B|A_1) = \frac{P(B \text{ and } A_1)}{P(A_1)} = \frac{P(A_1 \text{ and } B)}{P(A_1)}$
- Substituting one into the other:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B)}$$

- This is the most basic statement of Bayes' rule

Bayes' Rule: Alternate Formulations

- That $P(B)$ on the bottom is not very useful
- Suppose A_2 is the even that A_1 does not happen, ie $A_1 + A_2 = \Omega$
- Then

$$\begin{aligned} P(B) &= P(B \text{ and } A_1) + P(B \text{ and } A_2) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \end{aligned}$$

- This give us a more useful version of Bayes' rule:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

Bayes' Rule: Alternate Formulations

- That $P(B)$ on the bottom is not very useful
- Suppose A_2 is the even that A_1 does not happen, ie $A_1 + A_2 = \Omega$
- Then

$$\begin{aligned} P(B) &= P(B \text{ and } A_1) + P(B \text{ and } A_2) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \end{aligned}$$

- This give us a more useful version of Bayes' rule:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

- Or more generally:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\Sigma_i P(B|A_i)P(A_i)}$$