GV300 Quantitative Political Analysis

Week 3
Probability Theory – Part 1

Dominik Duell (University of Essex)

October 16, 2019

Housekeeping

- Problem set 1 due on Thursday October 23, 9:45 via Faser, will be on Moodle by the end of the day
- More refresher material on R/Stata online
- Support session next week, more R/Stata (loops, functions, programs), some algebra, some set theory
- No office hours on this Thursday

What are we talking about today?

Probability

- Uncertainty plays a huge role in the social world
- humans constantly make descriptive inferences about available facts and causal inference about how the world works under uncertainty
- Statistics is nothing else then the study of such judgments
- We need a precise language to talk about judgment under uncertainty and all starts with **probability theory** – probability is the formal language of uncertainty!

What are we talking about today?

How is probability used in this regression table? For what purpose?

	Dependent variable:	
	Social Distance	
	(1)	(2)
Ideology	0.314***	0.276***
	(0.055)	(0.052)
Religiosity (baseline: religious)		
Secular	−0.425*	−0.353
	(0.256)	(0.246)
Traditional	0.236	0.287
	(0.285)	(0.272)
Ultra-Orthodox	0.650***	0.495**
	(0.241)	(0.232)
Education (baseline: graduate)		
Primary school	1.013**	0.634
	(0.466)	(0.458)
High school	0.368	0.250
	(0.337)	(0.321)
Undergrad	0.570*	0.430
	(0.340)	(0.325)
Income (baseline: average)		
Very low income	−0.002	0.048
	(0.202)	(0.195)
Low income	−0.096	−0.180
	(0.216)	(0.208)
High income	−0.474*	−0.384
	(0.270)	(0.260)
Very high income	0.151	−0.108
	(0.402)	(0.387)
Age	−0.003	−0.004
	(0.005)	(0.005)
Foreign Born	−0.326*	−0.246
	(0.197)	(0.188)
Male	−0.519***	−0.376***
	(0.149)	(0.144)
Ethnicity (baseline: Ashkenazy)		

What are we talking about today?

```
. reg voteLabour income;

      Source |       SS       df       MS              Number of obs =   13500
-------------+------------------------------           F(  1, 13498) = 2214.01
       Model |  475.32687     1  475.32687             Prob > F      =  0.0000
    Residual |  2897.8935 13498  .214690584            R-squared     =  0.1409
-------------+------------------------------           Adj R-squared =  0.1408
       Total |  3373.22037 13499  .249886686           Root MSE      =  .46335

------------------------------------------------------------------------------
  voteLabour |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      income |  -.0092478   .0001965   -47.05   0.000     -.009633   -.0088625
       _cons |   .9457103    .010503    90.04   0.000      .925123    .9662976
------------------------------------------------------------------------------
```

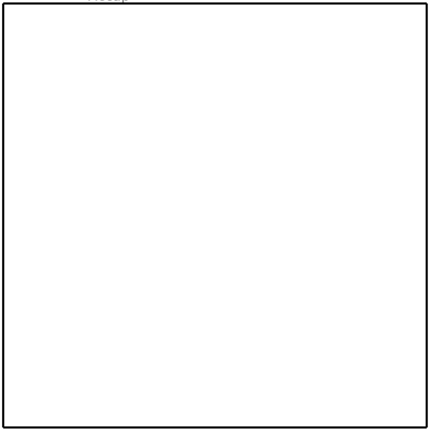
Plan for week 3 and 4 (and 5)

Probability Theory

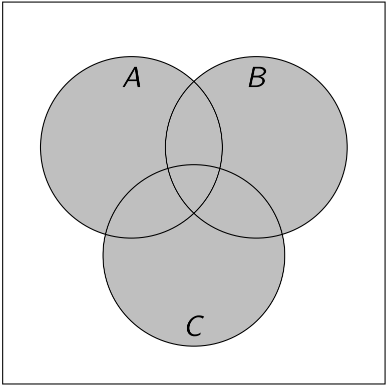
1. Week 3: Outcomes, events, samples, and the definition of probability; conditional probability and independence
2. Week 4 (maybe 5): Bayes Theorem, random variables, probability mass functions

- **Sets**: bounded collections defined by its contents: set of EPL teams {Manchester City, Manchester United, Tottenham Hotspurs, ...}
- **Elements**: contained in a set
- **Experiment**: specific snapshot of the world that can be repeated many times
- **Outcome**: anything that may happen in a given experiment

- **Sample space** of a given experiment: set of all possible outcomes of an experiment
- **Event**: any collection of possible outcomes of an experiment – simple events cannot be broken down further into constituting outcomes
- **Event space**: any mutually exclusive, collectively exhaustive collection of events of an experiment
- **Compound events**: composed of two or more simple events – either **independent** or **conditional** on one another



- Sample space or **universal set** of an **experiment**: set of all possible outcomes of an **experiment**



- **Events**: any collection of possible outcomes of an **experiment**

Example

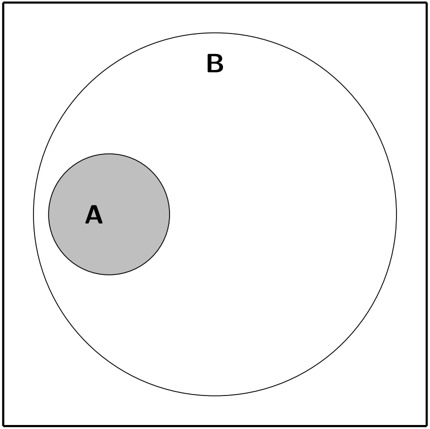
- Roll 6-sided die once:
  - Sample space: 1,2,3,4,5,6
  - Events: Roll 1, roll 3, roll # larger than 4, ...
- Example: Toss coin 3 times:
  - Sample space: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH
  - Events: Toss 3 T, Toss 1 T on 1st and 1 H on 3rd, ...

- **Countable set**: elements can be placed in one-to-one correspondence with positive integers
- **Finite set**: contains non-infinite numbers of elements
- **Cardinality** of a set: number of contained elements
- **Empty set**: contains no element, $\emptyset$

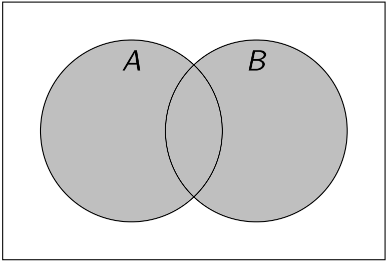
Operations on sets

- **Complement**: $A' = \{X : X \notin A\}$
- **Subset**: A is a subset of B if every element of A is also in B: $A \subset B \Leftrightarrow \forall X \, X \in A, X \in B$
- **Equal sets**: $A = B \Leftrightarrow A \subset B, B \subset A$
- **Union of sets A, B**: $A \cup B = \{X : X \in A \text{ or } X \in B\}$
- $A_1 \cup A_2 \cup \dots A_n = \bigcup_{i \leq n} A_i$
- **Intersection of sets A,B**: $A \cap B = \{X : X \in A \text{ and } X \in B\}$
- $A_1 \cap A_2 \cap \dots A_n = \bigcap_{i \leq n} A_i$

A subset of B

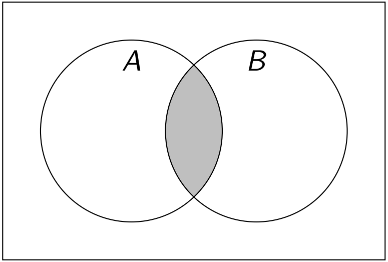


Union of A and B



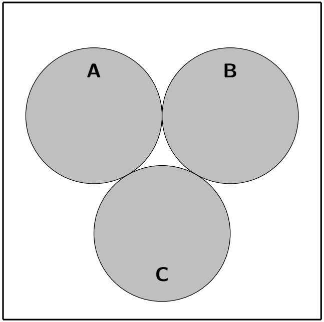
- $A \cup B$ or $A + B$

Intersection of A and B



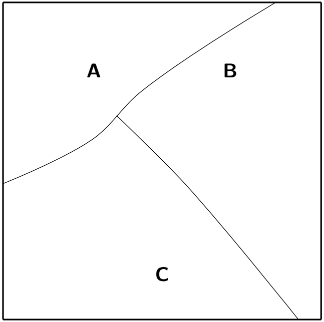
- $A \cap B$ or AB
- Generally, if you are confused about probabilities, draw such pictures

Mutually exclusive



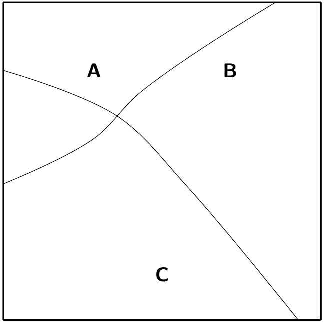
- A, B, and C are **mutually exclusive**
- k sets $A_1, A_2, \ldots, A_k$ are mutually exclusive iff $A_1 \cap A_j = \emptyset$ $\forall i \neq j$
- Also called *pairwise disjoint*

Collectively exhaustive



- A, B, and C are **collectively exhaustive**
- k sets $A_1, A_2, \ldots, A_k$ are collectively exhaustive iff $\bigcup_{k=1}^{K} A_k = S$
- This one is also mutually exclusive

Collectively exhaustive but not mutually exclusive



Summarizing

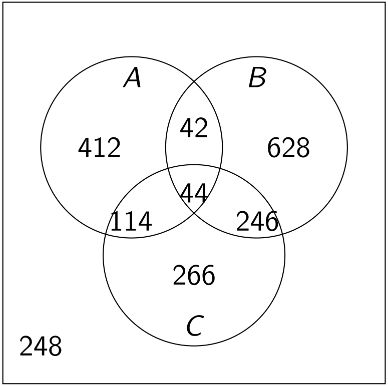
- **Outcome** is anything that may happen in an experiment
- **Event** is any collection of possible outcomes of an experiment
- **Event space** is any mutually exclusive, collectively exhaustive collection of events of an **experiment**
- **Sample space**: finest grained mutually exclusive, collectively exhaustive set of all possible outcomes of an experiment – finest grained? Means it cannot be divided any further

Example

- Experiment: We ask 2000 people about their smoking and drinking habits, and their age
- Here is a series of events:
  - 612 Smokers
  - 960 Drinkers
  - 670 Older than 25
  - 86 Drink and smoke
  - 290 Drink and are older than 25
  - 158 Smoke and are older than 25
  - 44 Drink, smoke, and are older than 25
  - 248 don't drink, don't smoke, and are younger than 25
- Let's represent these events in a Venn diagram

Example

Let's define Event A: Smoke, Event B: Drink, Event C: Older than 25. Fill all intersections of events and that part of an event that is not intersected with any other event with the number of people representing for which the event is true (that is, how many people smoke but don't drink, how many people drink, smoke, and are older than 25, etc.):



- To make probability statements we need to assign each point in the event space a probability

  Note, discrete world, we use sums not integrals
- Consider an experiment with sample space S, a real-valued function $\mathbb{R}$ on the event space is called probability measure
- Loosely speaking:

$$\text{Prob(Event)} = \frac{\#\text{ of ways event could happen}}{\text{Total }\#\text{ of possible outcomes}}$$

- This is a numerical measure of the likelihood of an event
- Probability function is a mapping from a defined event(s) onto a metric bounded by zero and one

$$\text{Prob(Event)} = \frac{\#\text{ of ways event could happen}}{\text{Total }\#\text{ of possible outcomes}}$$

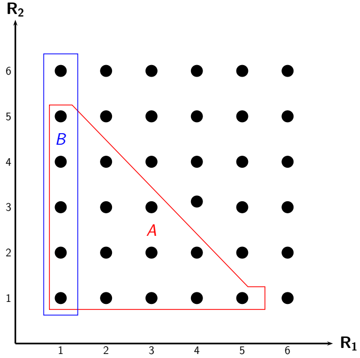
- Note, this is a theoretical construct
- We could think of probability as empirical construct as well:

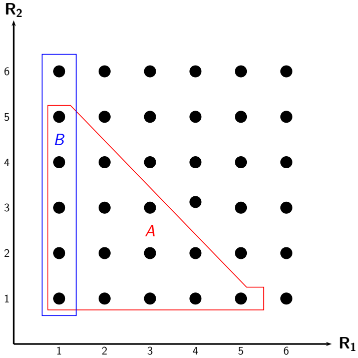
$$\text{Prob(Event)} = \frac{\#\text{ of times a given outcome occurs}}{\#\text{ of times any outcome occurs}}$$

- Example: Toss coin once, what is the probability that it comes up "Heads?" $p(E) = \frac{1}{2}$
- Example: Toss coin 3 times, what is the probability that it comes up "Heads" at 2nd Toss?
    - 8 possible outcomes: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH
    - $p(E) = 4/8 = 1/2$

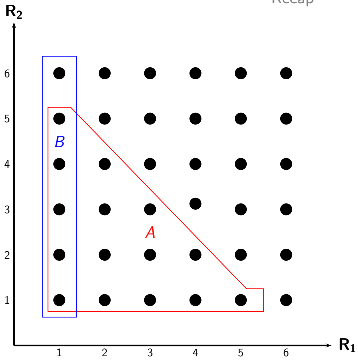
- Rules a probability measure needs to satisfy:
  - non-negative #: for any event A, $p(A) \geq 0$
  - Probability of S is 1: $p(S) = 1$
  - If there are two outcomes that cannot happen at same time, then the probability that either outcome occurs is the sum of probability of individual outcomes:
    If $AB = \emptyset$, $P(A + B) = p(A) + p(B)$
  - **Joint probability** is the probability of a compound event
  - Compound events are either **independent** or **conditional** on each other

- An event has occurred how does that affect the probability of another event?
- Example
  - Roll 2 dice
  - Event A: $R_1 + R_2 < 7$
  - Event B: $R_1 = 1$





- Probability of each point? $1/36$
- $P(A)$? $P(A) = \frac{15}{36} = .42$
- $P(B)$? $P(B) = \frac{6}{36} = \frac{1}{6}$
- What is the probability of B given A? What is $P(B|A)$



- Count it!
  $P(B|A) = 5/15 = 1/3$

- Note, relative probability within A is not changed
- We learned, A happened but do not have any more information about any outcome not in A – Event A becomes new total number of expected outcomes

- The conditional probability $p(B|A)$ includes other information (Event A) when specifying the probability that Event B occurs.
- How to scale probability so that total probability is 1?

$$P(B|A) = \begin{cases} P(B)/P(A) & \text{if } B \subset A \\ 0 & \text{if } B \subset A' \end{cases}$$

- two events are independent if knowing about one tells you nothing about the other
- sometimes it is trivial to see:
  - Flip 2 coins
  - Event A: first heads
  - Event B: second tails
  - Unless coin not damaged in first flip, certainly independent events

- A and B are **independent** iff $P(A|B) = P(A)$
- Generalizing to N events:

  N events $A_1, \ldots, A_N$ are mutually independent iff $P(A_i|A_j, A_k, \ldots, A_p) = P(A_i) \, \forall \, i \neq j, k, \ldots, p$ with $1 \geq i, j, k, \ldots, p \geq N$

- A and B are **conditionally independent** iff $P(AB|C) = P(A|C)P(B|C)$
- conditional independence matters, for example:
  - we assume that error term in regression model is conditionally independent of X (aka independent variables)
  - people could have independent health outcome but not conditionally independent given hospitalization
  - advanced regression and experimental tools in spring are about when we can make the conditional independence assumption

What can we do with knowing the probability of a particular event?

- Say we got the electoral returns of a Northern English district for the last parliamentary election
- 250612 voted Tory, 198027 voted Labour, 45312 voted LibDem
- What is a good guess of the probability that an arbitrary voter chosen at random from the district votes LibDem?
- What type of probability measure is this?

Conditional probability, example:

- Flip coin twice
- Event A: At least 1 head
- Event B: Two heads
- What's $P(A)$, $P(B)$, $P(B|A)$, $P(A|B)$?

- Most events of interest in political science are not simple – we face **compound events** that often are
- not mutually exclusive
- or one is conditional on other(s):
  - say the decision to vote, it is dependent already on the weather!
  - $prob(turnout > .5|weather) = prob(y > .5|\beta, x) = prob((\beta x + \epsilon)|\beta, x)$
- We need independence and/or conditional probability to determine the probability of **compound events**

Remember this handy notation:

- $P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$
- $P(B|A) = \frac{P(A \cap B)}{P(A)}$ with $P(A)$ non-zero
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ with $P(B)$ non-zero
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Summarizing

- Probability function: mapping from an event(s) onto a metric bounded by 0 and 1

  allows us to discuss various degrees of likelihood of occurrence of events
- Probability measure:

$$\text{Prob(Event)} = \frac{\text{\# of ways event could happen}}{\text{Total \# of possible outcomes}}$$

- Events may occur as compound events so we need to determine conditional probability and/or independence of those events