Day 5: Probability

Unsolicited advice

- Figure out how you're going to manage your files now.
- Stop wasting time doing stupid things
  - Stupid = you don't like it + it's not benefitting you
  - Stupid != rest and leisure
- If you're going to be in academia, you need to really love your job
  - If you don't, that doesn't mean you need to quit! Our Ph.D. program is really good at preparing people for industry jobs.
- Your research isn't that important.
- The impact you have on the people around you is important.

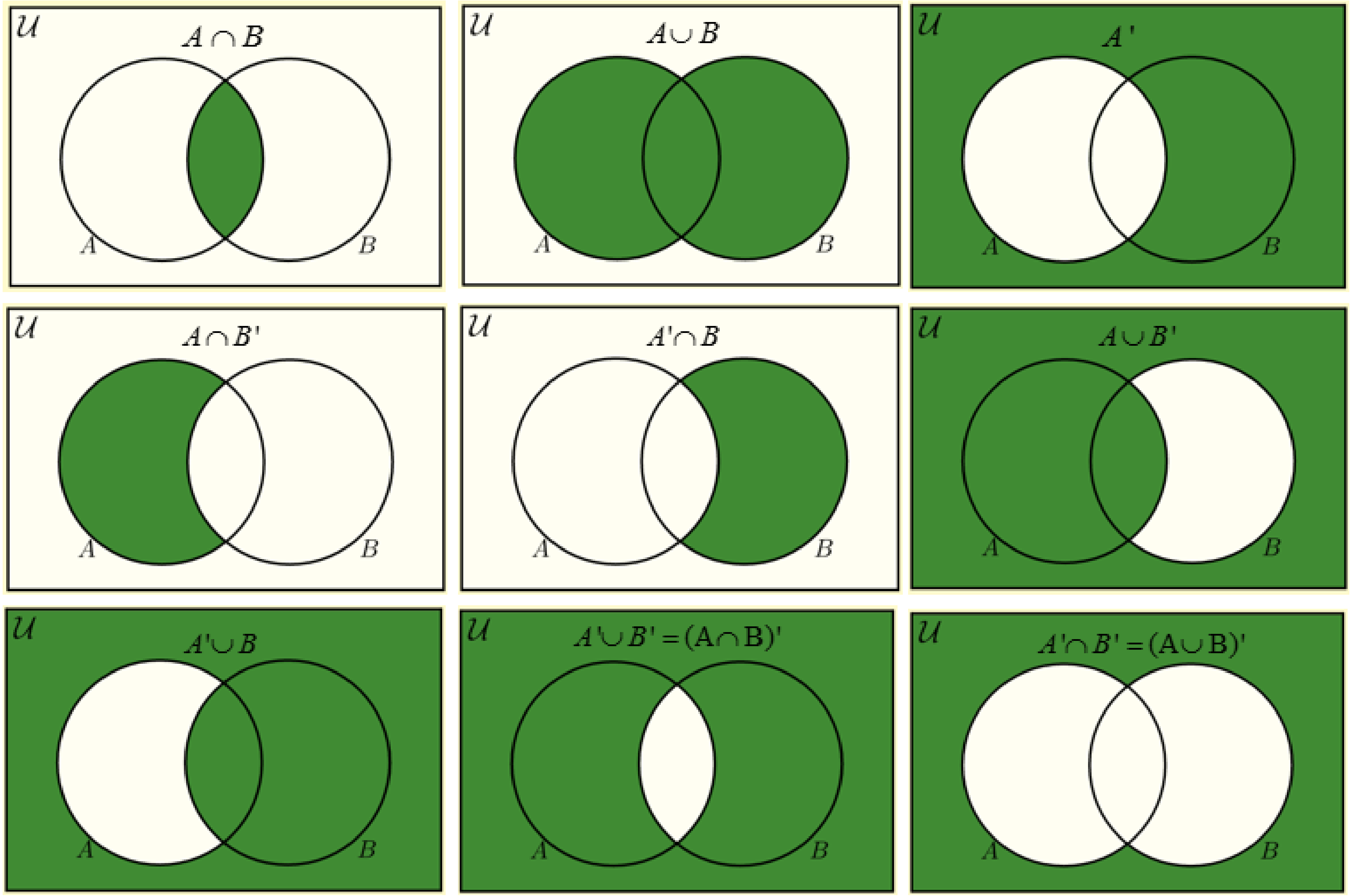
Sets and Set Notation

Set notation

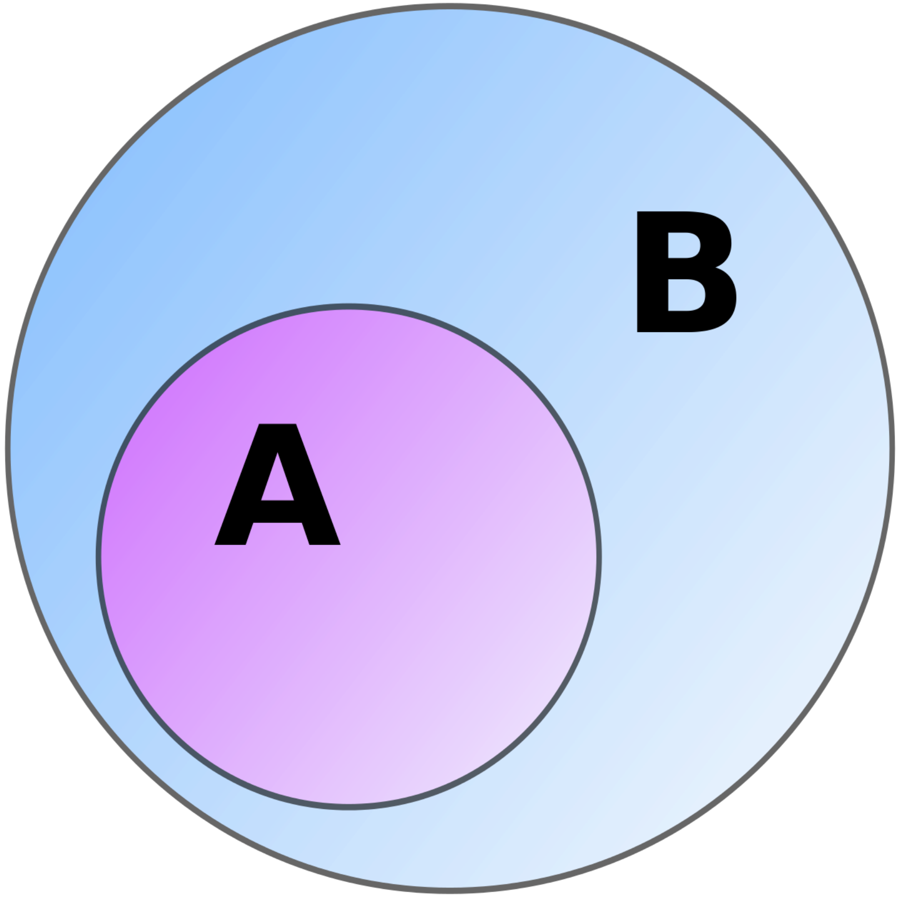
Set Theory Symbols

Symbol	Name	Example	Explanation
{ }	Set	$A = \{1,3\}$ $B = \{2,3,9\}$ $C = \{3,9\}$	Collection of objects
$\cap$	Intersect	$A \cap B = \{3\}$	Belong to both set A and set B
$\cup$	Union	$A \cup B = \{1,2,3,9\}$	Belong to set A or set B
$\subset$	Proper Subset	$\{1\} \subset A$ $C \subset B$	A set that is contained in another set
$\subseteq$	Subset	$\{1\} \subseteq A$ $\{1,3\} \subseteq A$	A set that is contained in or equal to another set
$\not\subset$	Not a Proper Subset	$\{1.3\} \not\subset A$	A set that is not contained in another set
$\supset$	Superset	$B \supset C$	Set B includes set C
$\in$	Is a member	$3 \in A$	3 is an element in set A
$\notin$	Is not a member	$4 \notin A$	4 is not an element in set A

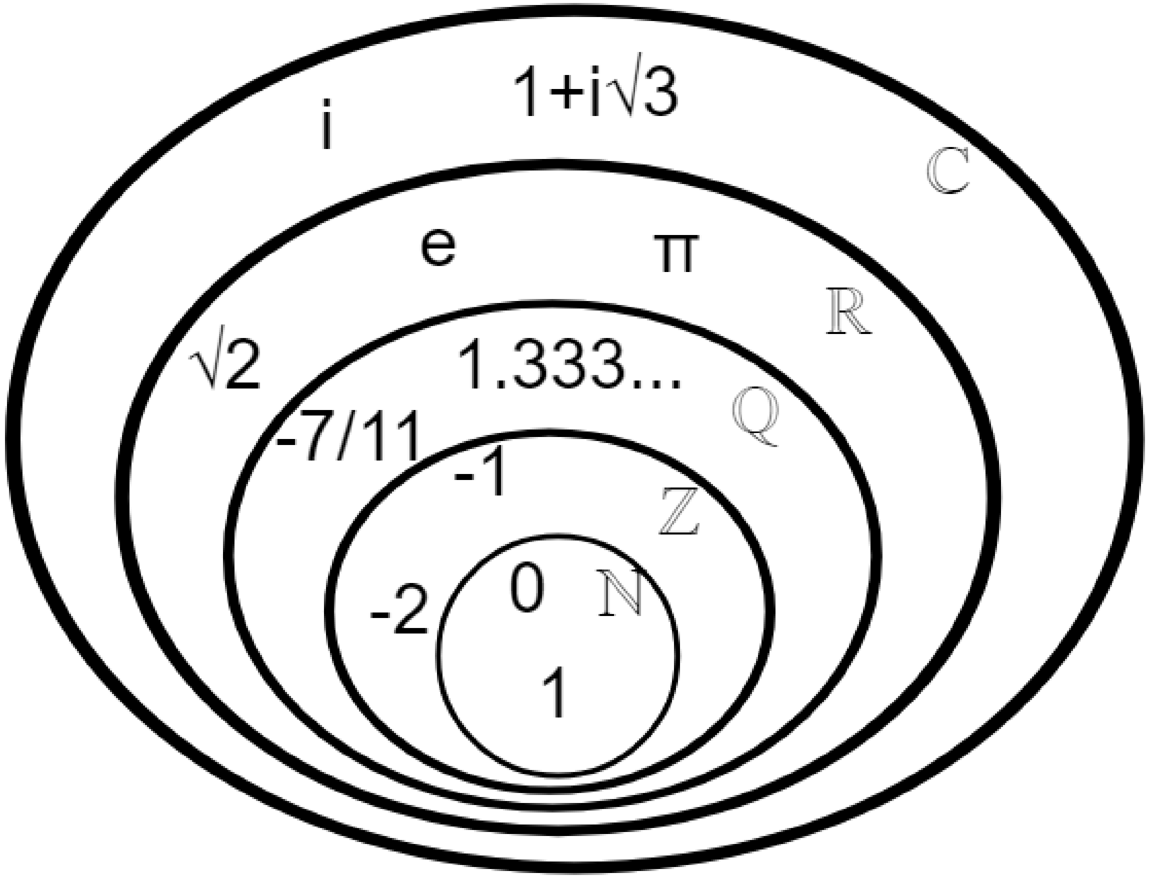
Basic sets as Venn diagrams



$$A \subset B, B \supset A$$



Defined sets



Probability

Probability in set notation

- $P(A)$: Probability of A
- $P(A \cup B)$: Probability of A or B
- $P(A \cap B)$: Probability of A and B

Reasoning with set notation

$$P(A \cup B) =$$

Reasoning with set notation

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Reasoning with set notation

$$P(A') =$$

Reasoning with set notation

$$P(A') = 1 - P(A)$$

Conditional Probability

Conditional probabilities

- Conditional probabilities allow us to update the probability of an event occurring based on additional information.

Example: Suppose we have *y*, a vector of binary values for 1,000 voters indicating if they voted in the last election. We have:

  - A{y: y = 1}
  - B{y: y=0}

We select a random voter. What is the probability that they voted in the last election?

Conditional probabilities

- Conditional probabilities allow us to update the probability of an event occurring based on additional information.

Example: Suppose we have y, a vector of binary values for 100 voters indicating if they voted in the last election. We have:

- $A\{y: y = 1\}$
- $B\{y: y=0\}$

We select a random voter: y_i. What is the probability that they voted in the last election?

$$\bar{y}$$

Conditional probabilities

- Now assume that in addition to y, we have c – the travel time of the voter to their nearest polling station.
- I tell you $\bar{c} = 30$ and that $c_i = \bar{c} + 15$
- How do you adjust your estimate?

Conditional probabilities

Generically expressed as:

$$P(A|B)$$

"The probability of A given B"

Conditional probabilities

Generically expressed as:

$$P(A|B)$$

"The probability of A given B"

It can also be expanded:

$$P(A|B, C, D, X, Y, Z)$$

$$P(A \cap Q|B, C, D, X)$$

Random events

Random events

- Random events are events that are non-deterministic
- Random does not imply even probability

Bernoulli trials

- A Bernoulli trial is a single event where the outcome is either a 1 or 0, True or False
- E.g. A coin flip or whether or not someone voted.

$$p(x) = p$$
$$p(x') = 1 - p$$

Binomial distribution

- A series of Bernoulli trials is ***binomially distributed***
- e.g. if we observer three elections, and count how many times people voted across those elections, that count is ***binomially distributed***

X_i = if someone voted in election i

Y = the count of elections someone voted in:

$$Y = \sum_{i=1}^{n} X_i$$

Binomial Distribution cont.

In our example:

$$p(Y = y) = \binom{3}{y} p^y 1 - p^{3-y}$$

More generally:

$$p(Y = y|n, p) = \binom{n}{y} p^y 1 - p^{n-y}$$

If something is binomially distributed we can express it as:

$$B(n, p)$$

View this as you would any other defined function. Don't pay attention to the specifics of the math right now, just know what the function is communication about the data and then plug and play.

Distribution notation

- Binomial distribution: $B(n, p)$
- Normal distribution: $N(\mu, \sigma^2)$
- Poisson distribution: $Pois(\lambda)$
- Chi-squared distribution: $\chi^2(k)$
- Etc.

$$X \sim N(\mu, \sigma^2)$$

"X is normally distributed around with a mean of μ and a variance of σ^2"

More Review and Practice

Example 1:

- How did we get from the first equation to the second?
- What happened to u_i?
- How can we explain what the second equation means in plain English?

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Example 2:

- What is the relationship between the second two equations and the first equation?
- Why would a model like this be used? What is it expressing about the relationship between the variables and Y?

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 X_i D_i + u_i$$

$$\begin{aligned} \mathrm{E}(Y|X, D = 0) &= \beta_0 + \beta_1 X \\ \mathrm{E}(Y|X, D = 1) &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X \end{aligned}$$

Example 3:

$$P\left(y_i = 1 \middle| \alpha_j, \beta_i, \gamma, \theta_i, \phi_j\right) = \; logit^{-1}\left(\alpha_j + \beta_i \; - \gamma \left|\theta_i - \phi_j\right|^2\right)$$

Example 3:

$$P\left(y_i = 1 \middle| \alpha_j, \beta_i, \gamma, \theta_i, \phi_j\right) = \ logit^{-1}\left(\alpha_j + \beta_i \ - \gamma\left|\theta_i - \phi_j\right|^2\right)$$

- α: The popularity of politician j
- β: How interested person i is in politics
- θ: The ideology of person i
- ϕ: The ideology of politician j
- γ: Normalizing constant (ignore this for now)

Some more notation

E(x)	Expected value of x
ρ	correlation
$\forall$	“For all”
$\propto$	“Proportional to”
$\equiv$	Identical to, true for all values, true by definition
iff	“If and only if”
IID	“Independent and identically distributed”
p	Probability of success
q	Probability of failure
$A \perp B$	A and B are independent events