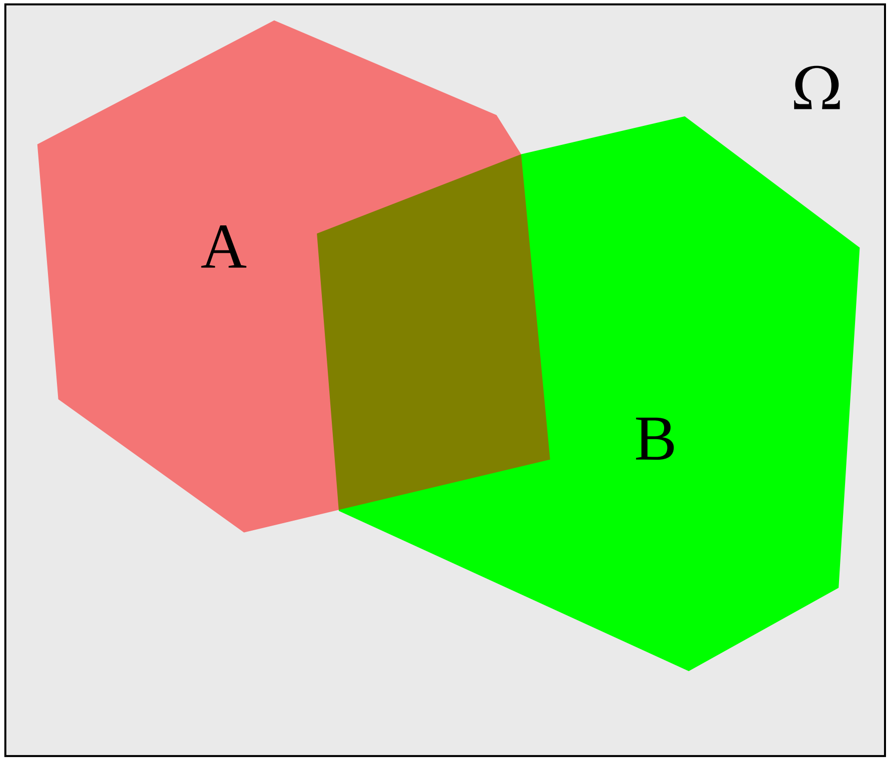
HST.582J / 6.555J / 16.456J Biomedical Signal and Image Processing
Spring 2007

Harvard-MIT Division of Health Sciences and Technology
HST.582J: Biomedical Signal and Image Processing, Spring 2007
Course Director: Dr. Julie Greenberg

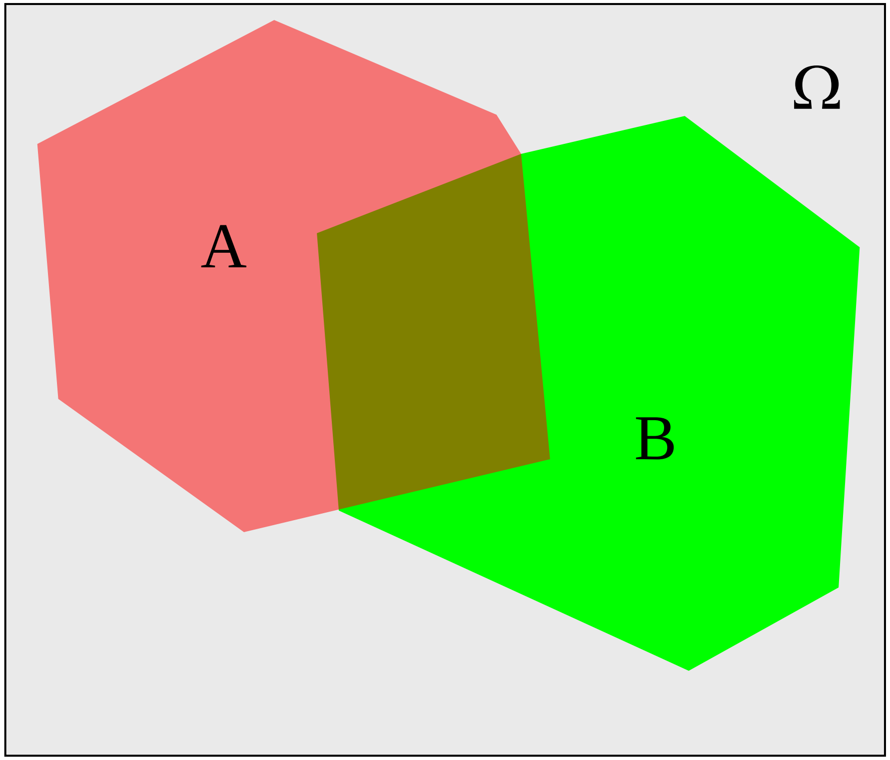
Probability with Venn Diagrams

Venn Diagrams and Probability



- Venn diagrams are a graphical representations of sets, which we can use to get probabilities.
- Ω - the set of all possible outcomes, the "certain event" (i.e. everything in the gray box).
- A, B - subsets of Ω.
- The event "A" (or "B") has occurred when an experimental outcome from the region "A" (or "B") is observed.
- A^c (or B^c) - the complement of A (or B), i.e. the region outside A (or B), but inside Ω.

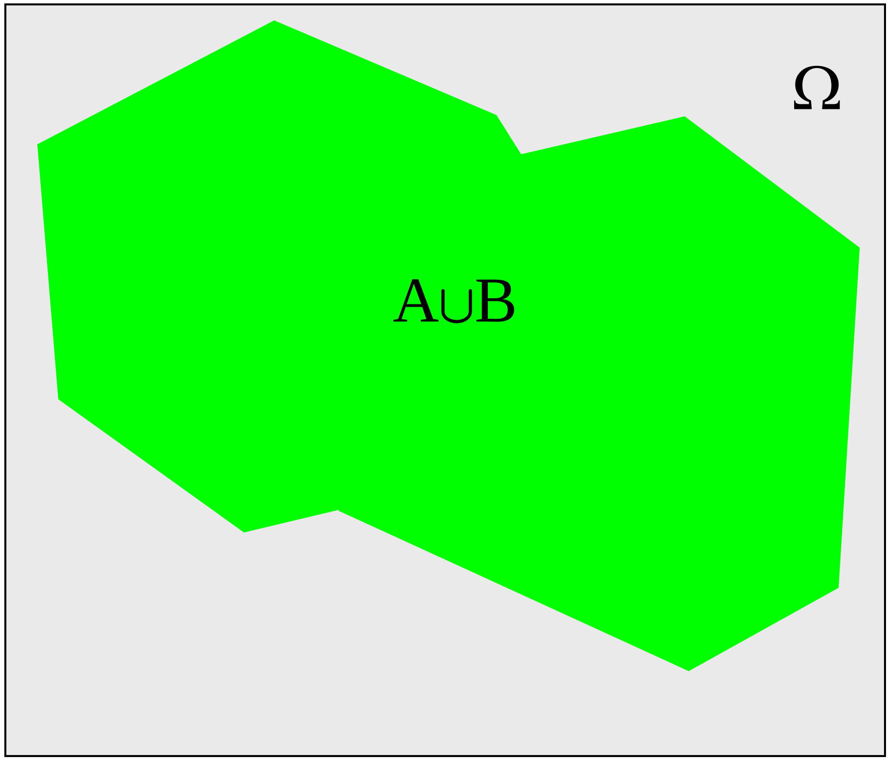
Venn Diagrams and Probability



To be concrete

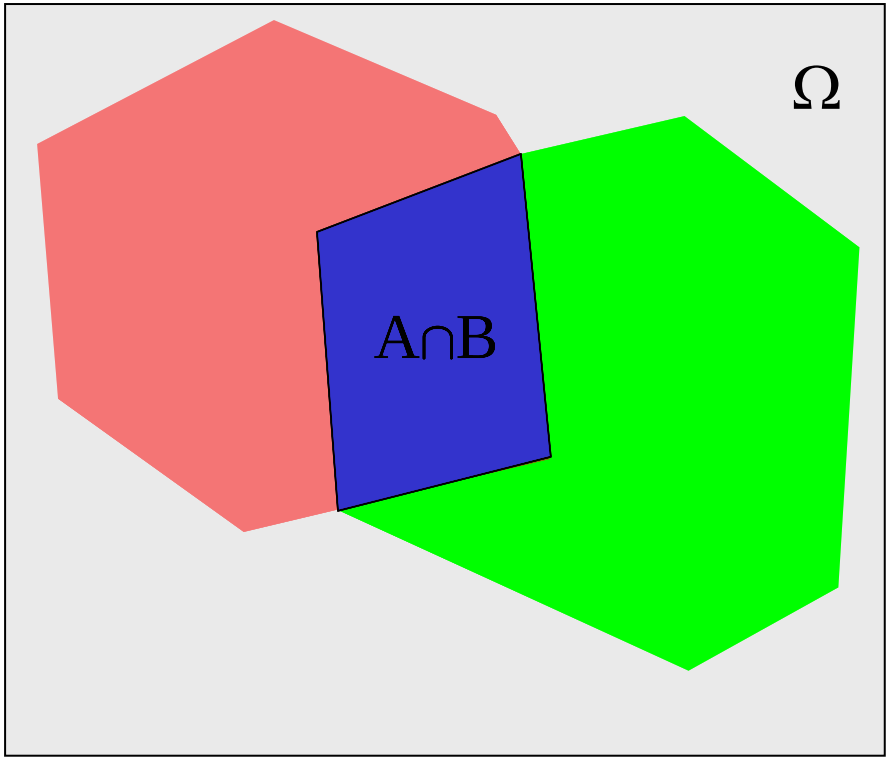
- Ω ={set of all possible results of flipping a coin 100 times}.
- A ={the event that at least 30 heads are observed}.
- B ={the event that at least 30 tails are observed}.
- These are not mutuall exclusive because it is possible to have a result that has 30+ heads and 30+ tails (the region of overlap).
- What 30 were changed to 51 in the definition of both A and B?

Venn Diagrams (Union)



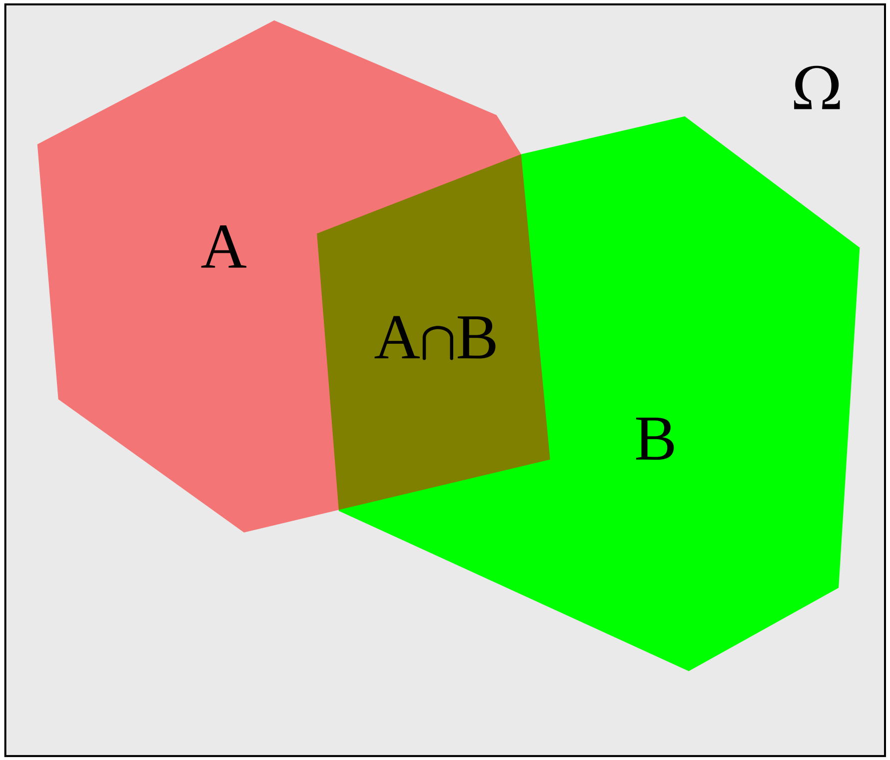
- The event "A or B", denoted $A \cup B$ or $A + B$
  - 30+ heads OR 30+ tails were observed
- It's complement $(A \cup B)^c$.
  - less that 30 heads AND less than 30 tails were observed

Venn Diagrams (Intersection)



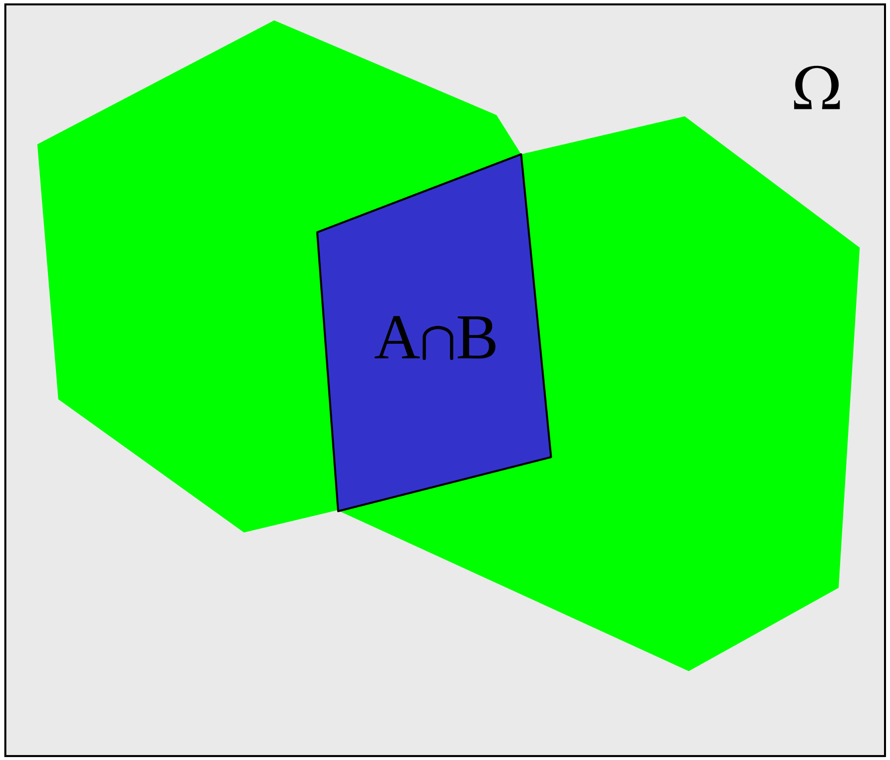
- The event "A and B", denoted $A \cap B$ or AB
  - 30+ heads AND 30+ tails were observed
- It's complement $(A \cap B)^c$.
  - (30+ heads, but less then 30 tails) OR (30+ tails, but less than 30 heads) OR (less than 30 heads AND less than 30 tails).

Venn Diagrams (probability from area)



- Equate the sizes of the regions (relative to Ω) as the chance that the event occurs.
- The probability of any event is equal the size of its area when the area of Ω is set to unity.
  - $P(\Omega) = 1$
  - $P(A) =$area of "A"
  - $P(A^c) = 1 - P(A)$
  - $P(B) =$area of "B"
  - $P(A \cup B) =$area of "A or B"
  - $P(A \cap B) =$area of "A and B"
- This is equivalent to saying all experimental outcomes (i.e. sample space events) are equally likely.

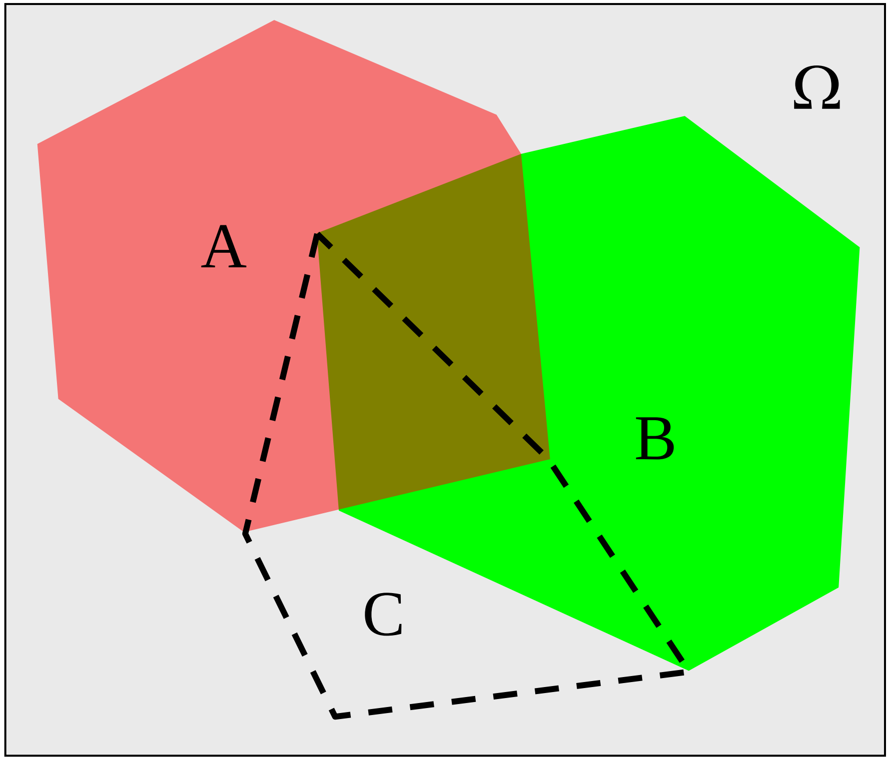
Venn Diagrams (probability relationship)



$$P(A+B) = P(A) + P(B) - P(AB)$$

- Using the area argument, one can compute $P(A \cup B)$ as a function of $P(A)$, $P(B)$, $P(A|B)$
- It is the sum of the areas of events A and B minus their overlap $A \cap B$ (to avoid double counting)
- If A and B are mutually exclusive events, then they will not overlap and $A \cap B$ will have zero area.

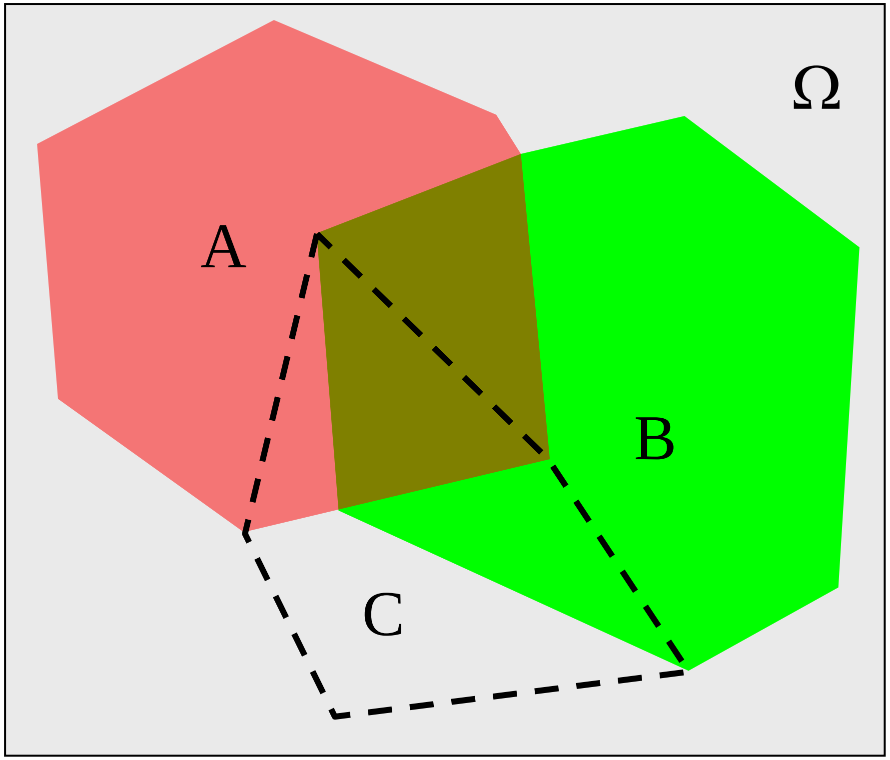
Venn Diagrams (combining multiple events)



- Suppose we add a third event, C (the region inside the dashed line) and we want to compute the $P(A \cup B \cup C)$ using the area argument.
- We could do something like: $P(A \cup B \cup C) = P(A) + P(B) + P(C)$ minus "the areas which were counted multiple times".
- However, it is easier to group $A \cup B$ as one event and use the previous result.
- This leaves one term, $P((A \cup B)C)$, which is neither simply a union nor an intersection of events.

$$
\begin{aligned}
P(A+B+C) &= P((A+B)+C) \\
&= P(A+B) + P(C) - P((A+B)C) \\
&= P(A) + P(B) - P(AB) + P(C) - \\
&\quad P((A+B)C) \\
&= P(A) + P(B) + P(C) \\
&\quad -P(AB) - P((A+B)C)
\end{aligned}
$$

Venn Diagrams (combining multiple events)



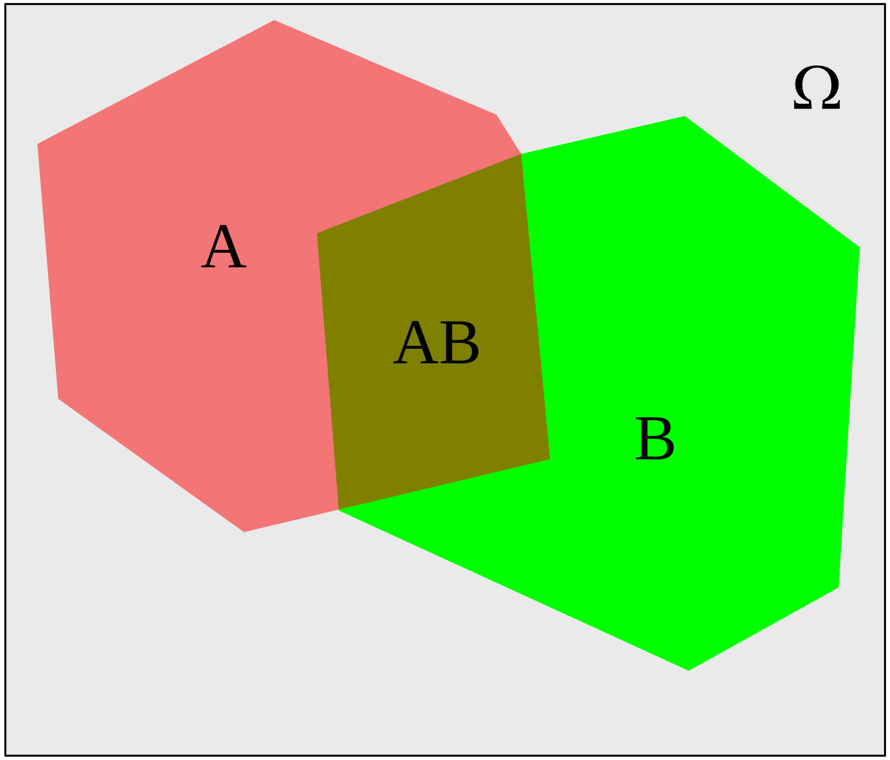
- $P((A \cup B)C)$ corresponds to the areas AC plus BC minus ABC

$$
\begin{aligned}
((A \cup B)\,C) = & P\,(AC) \\
& + P\,(BC) - P\,(ABC)
\end{aligned}
$$

- Substitute this into the equation and with some rearranging you get the result below.
- The main point here is that the probability of the union of events can be computed recursively by combining events one at a time.

$$
\begin{aligned}
P\,(A+B+C) &= P\,(A)+P\,(B)+P\,(C) \\
&\quad -P\,(AB) - P\,((A+B)\,C) \\
&= P\,(A)+P\,(B)+P\,(C) \\
&\quad -(P\,(AB)+P\,(AC)+P\,(BC)) \\
&\quad +P\,(ABC)
\end{aligned}
$$

Venn Diagrams (conditional probabilities)



$$P(A|B) = P(AB)/P(B)$$
$$P(B|A) = P(AB)/P(A)$$

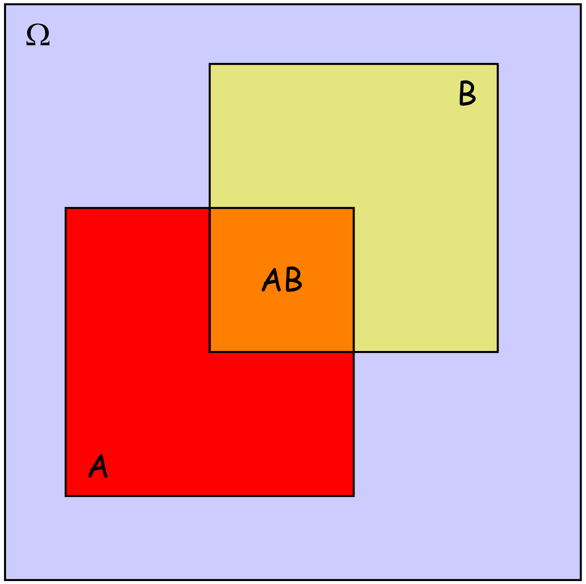
Back to heads/tails example

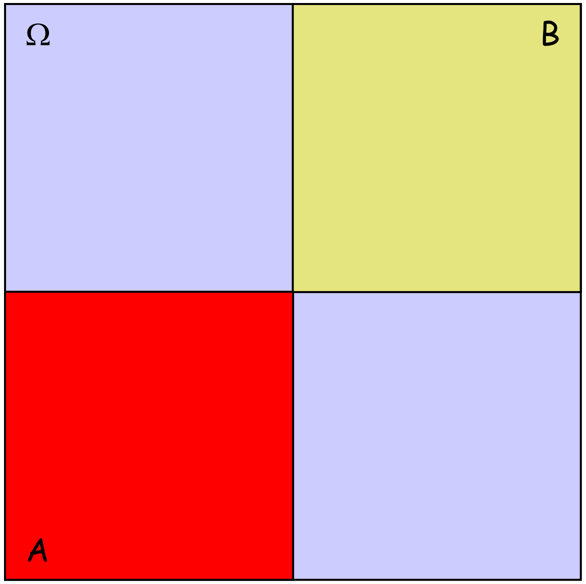
- What is the probability that 30+ heads (event A) **will** occur conditioned on the observation that 30+ tails (event B) **did** occur.
- This can be computed by looking at the relative size of the set AB to B. That is, looking at the proportion of the set B which also are in set A.
- $P(A|B)$ denotes the probability of A conditioned on B.
- $P(B|A)$ is computed similarly.

Statistical Independence (for events)

- Two Events A and B are statistically independent if their **joint** probability is the product of their **marginal** probabilities:

$$\Pr\{AB\} = \Pr\{A\}\Pr\{B\} \quad \Leftrightarrow \quad \Pr\{A|B\} = \frac{\Pr\{AB\}}{\Pr\{B\}} = \Pr\{A\}$$





Which Venn diagram depicts *independent* events?