Events, Probabilities, and Bayes' Rule

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

Outline

Experiments, Events, and Probabilities

Properties of Probabilities

Approaches to Probability

Conditional Probability and Bayes' Rule

Discussion

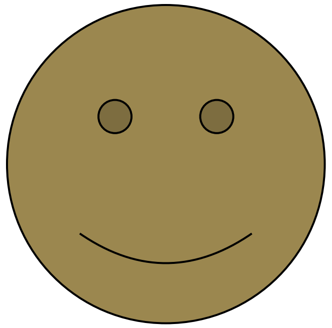
Homework

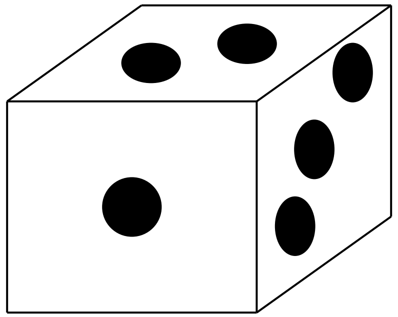
Experiments, Events, and Probabilities

Let's talk about **experiments**, **events**, and **probabilities**.

An **experiment** is a process by which a measurement is taken or observations is made.

i.e. flip coin or roll die





Experiments, Events, and Probabilities

An **outcome** is the result of an experiment. i.e. *Heads*, or 3

Coin*:* $O_1=H$, $O_2=T$

Die: $O_1=1$, $O_2=2$, $O_3=3$, $O_4=4$, $O_5=5$, $O_6=6$

Sample space is a listing of possible outcomes.

$S=\{O_1,O_2\}$ or $S=\{O_1,O_2,O_3,O_4,O_5,O_6\}$

Coin: $S=\{H,T\}$

Die: $S=\{1,2,3,4,5,6\}$

Experiments, Events, and Probabilities

An **event** A is an outcome or a combination of outcomes.

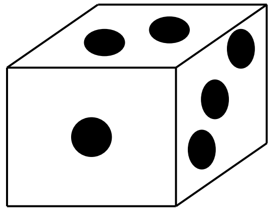
i.e. A=*even number when rolling a die*={2,4,6}

The probability of an event A is written $P(A)$.

i.e. $P(A)=P(\textit{even number when rolling a die})$

$=P(\{2,4,6\})$

$=3/6$



Properties of Events

Property 1

In words:

"A probability is always a numerical value between 0 and 1."

In algebra:

A is an event

$$0 \le \text{each } P(A) \le 1$$

P(Heads on coin flip),
or P(3 on roll of die)

If the event A can never occur, then $P(A)=0$.

If the event A is sure to occur, then $P(A)=1$.

Properties of Events

Property 2

In words:

"The sum of probabilities for all outcomes of an experiment is equal to exactly 1."

In algebra:

O_i are nonoverlapping outcomes that include all possibilities

$$\sum_{i=1}^{n} P(O_i) = 1 \qquad i = 1, \ldots, n$$

Approaches to Probability

Now that we talked about events and probabilities,
how do we get probabilities of events?

Probability of a event: The relative frequency with which that event can be expected to occur.

There are three different approaches to probability.

Empirical (AKA experimental)
Theoretical (AKA classical or equally likely)
Subjective (AKA expression of belief, generally not discussed)

Approaches to Probability-Empirical

Empirical (Observed) Probability: $P(A)$

In words:

$$\text{empirical probability of } A = \frac{\text{number of times } A \text{ occured}}{\text{number of trials}}$$

In algebra:

$$P(A) = \frac{n(A)}{n}$$

Approaches to Probability-Empirical

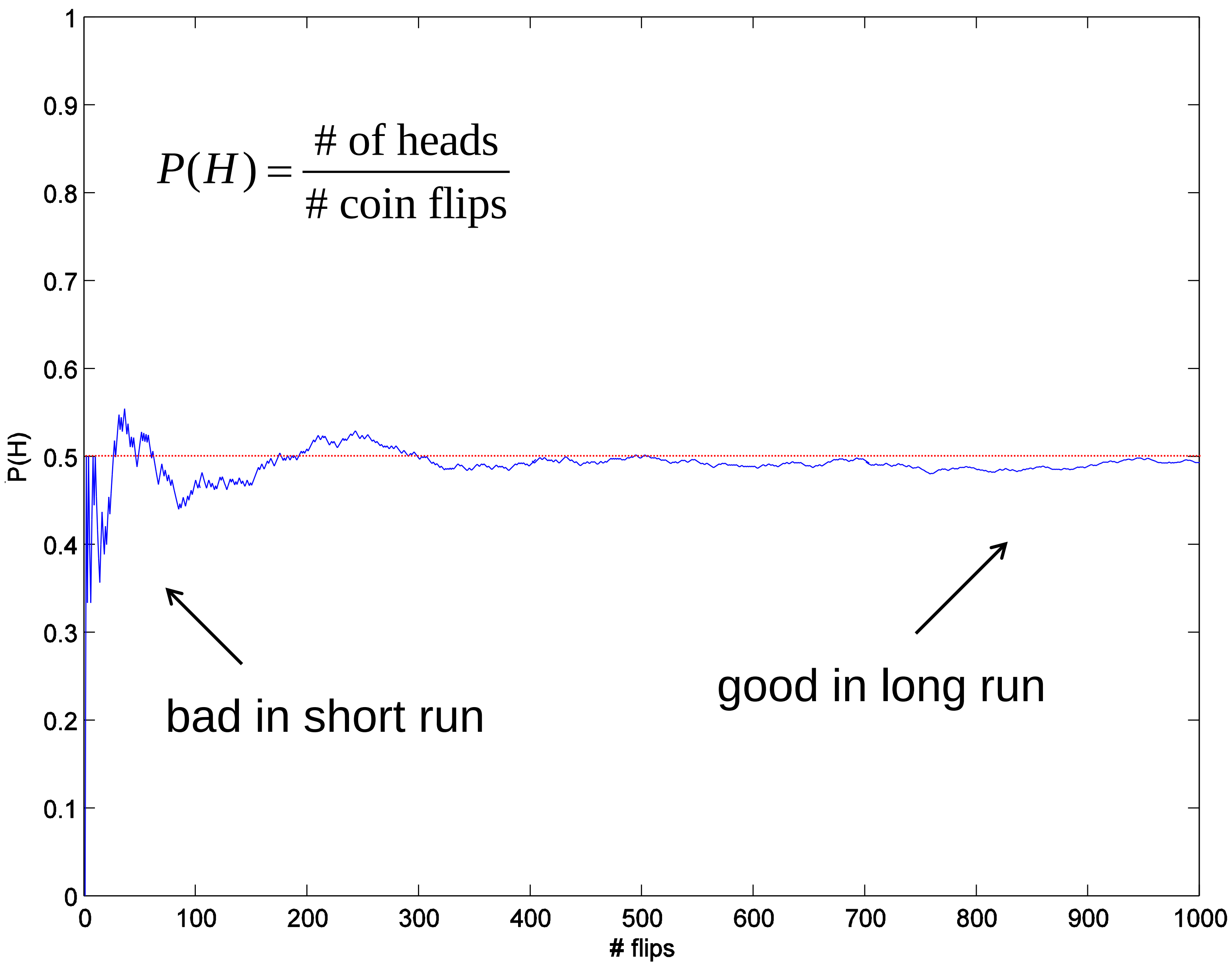
```
m=1;, p=1/2;, n=1000;
x=binornd(m,p,n,1);
cumprob=(cumsum(x)./(1:n)');
figure;
plot((1:n),cumprob)
```

Had computer flip a single coin 1000 times.

Flip # on x axis

$P(H)$ on y axis.

This shows convergence to true value of 1/2.



Approaches to Probability-Empirical

In the empirical method you actually have to perform the experiment of flipping the coin.

The empirical approach may be off in the short run.

Suppose you get on a streak and out of 10 flips all 10 are heads?

By the empirical method we would say that $P(H)=1$.

Approaches to Probability-Theoretical

Theoretical (Expected) Probability: $P(A)$

In words:

$$\text{theoretical probability of } A = \frac{\text{number of times } A \text{ occus in sample space}}{\text{number of elements in the sample space}}$$

In algebra:

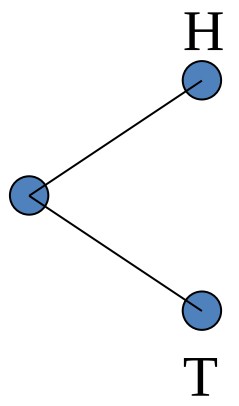
$$P(A) = \frac{n(A)}{n(S)}$$

Approaches to Probability-Theoretical

Flip once.

Sample space:
listing of outcomes
for 1 flip

$$S = \{H, T\}$$



$$P(H) = \frac{\#\text{ times } H \text{ occurs in } S}{\#\text{ elements in } S} = \frac{1}{2}$$

Approaches to Probability-Subjective

Generally results from personal judgement.

Often assessed before having any data.

Quantify your beliefs or the beliefs of a group of experts.

By the subjective method we would say that I feel that or from my lengthy experience with coin flipping that $P(H) = 0.52$.

We will come back to this and use subjective probability for the entire course.

Experiments, Events, and Probabilities

What is the probability that the Professor will put an exam question on topic x?

What is the probability that the Professor will put an exam question on topic x given that he covered topic x in class?

Let A = Professor will put an exam question on topic x

B = s/he covered topic x in class

$P(A)$ vs. $P(A|B)$

Experiments, Events, and Probabilities

We use conditional probability in our daily lives and sometimes we do not realize it.

What is the probability that the Professor will put an exam question on topic x?

What is the probability that the Professor will put an exam question on topic x given that he covered topic x in class?

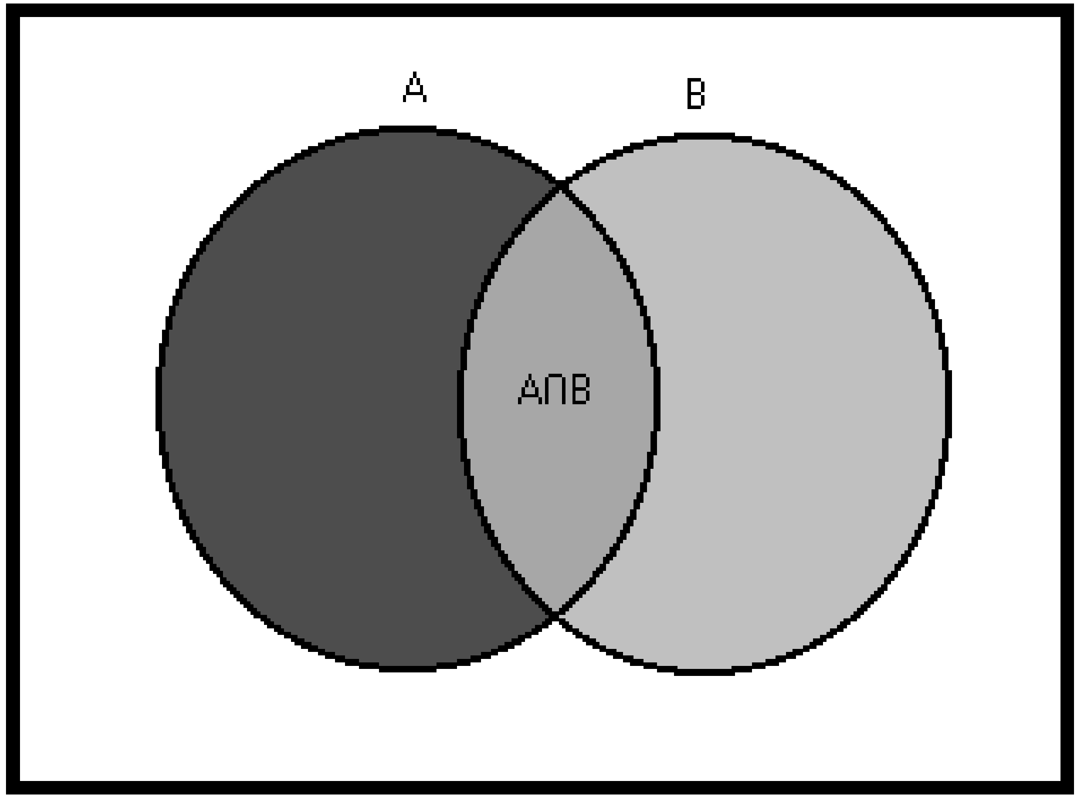
Conditional Probability and Bayes' Rule

We learned about the conditional probability of B given A.

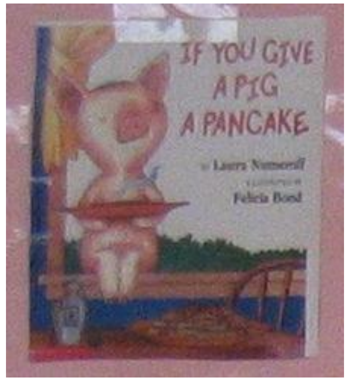
If A and B are events in S, and $P(A)>0$, then the *conditional probability of* B *given* A is,

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$

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



Conditional Probability and Bayes' Rule

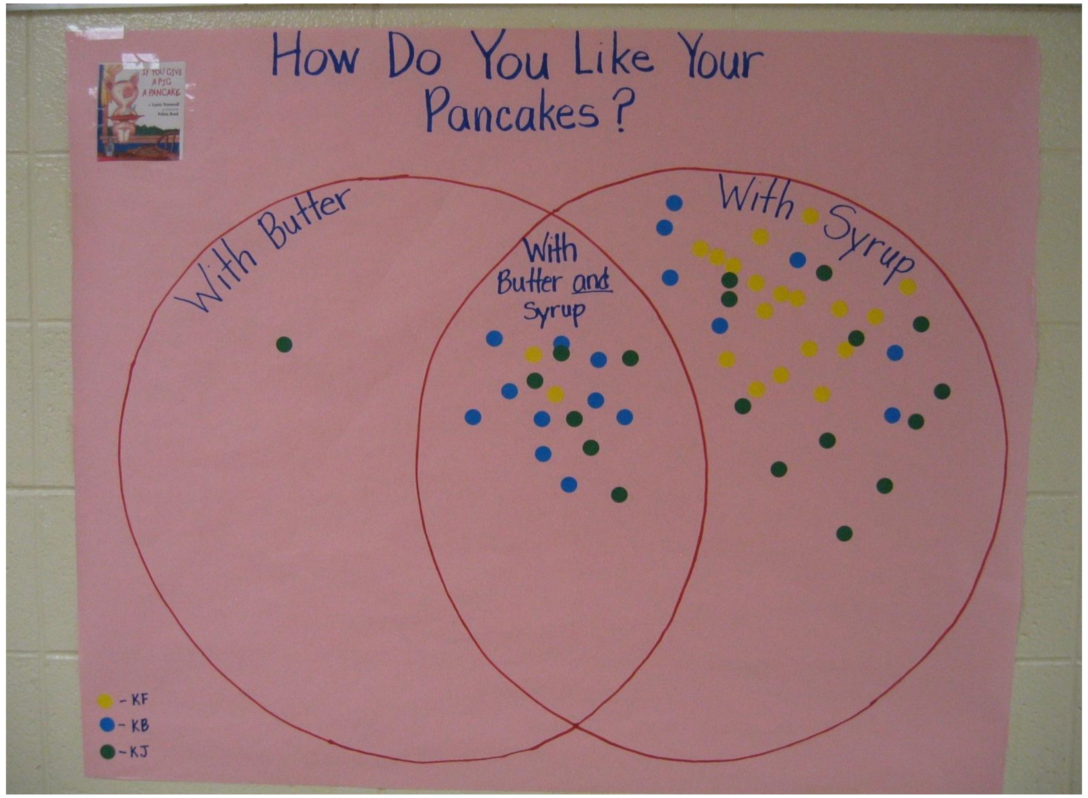


We learned about the conditional probability of B given A.

If A and B are events in S, and $P(A)>0$, then the *conditional probability of* B *given* A is,

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$

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



Conditional Probability and Bayes' Rule

Conditional probability an event will occur: A conditional probability is the relative frequency with which an event can be expected to occur under the condition that that additional preexisting information is known about some other event.

$P(A|B)$, the "|" is spoken as "given" or "knowing"

Conditional Probability and Bayes' Rule

$$P(A)=\frac{n(A)}{n(S)}$$

$$P(B)=\frac{n(B)}{n(S)}$$

Example: Roll two die.
Let A be that 10 is the sum of the two die.
$P(A)$=

Let B that the first die is a 4.
$P(B)$=

What is $\text{P}(A|B)$?.
$P(A|B)$=

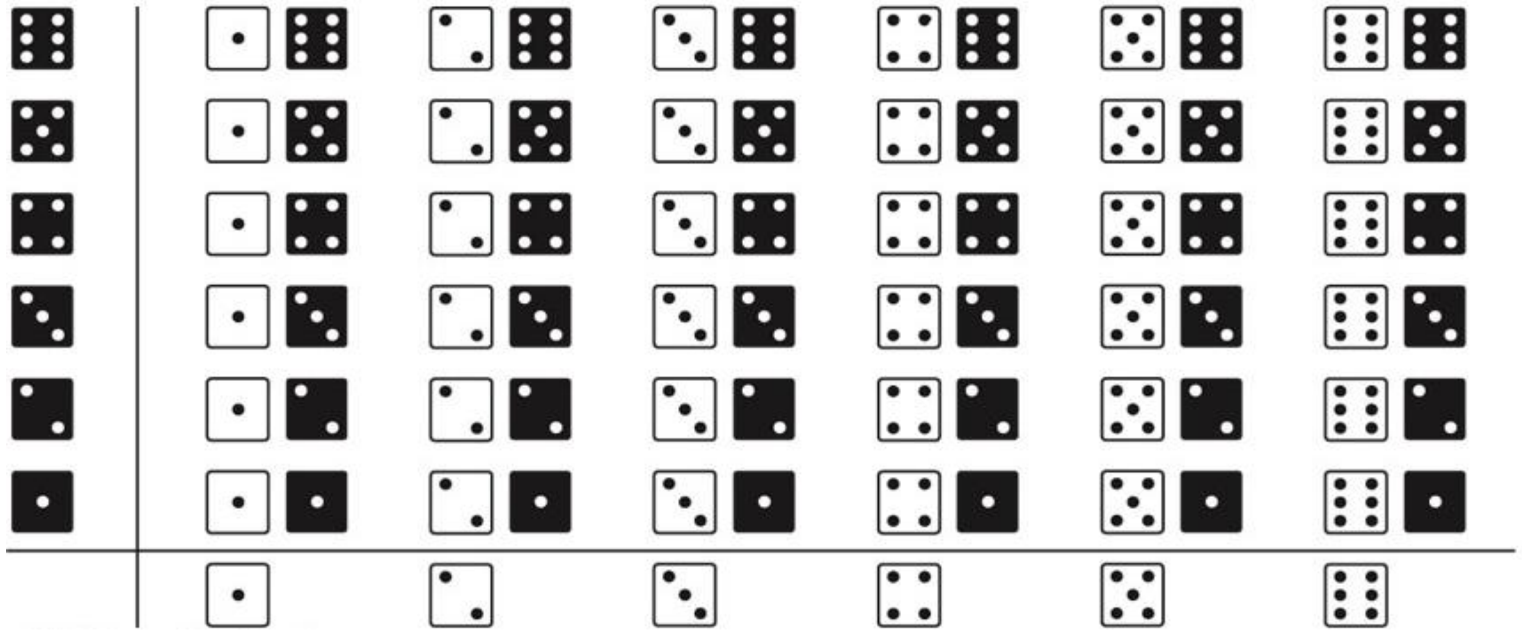


Figure from Johnson & Kuby, 2012.

Conditional Probability and Bayes' Rule

Example: Roll two die.
Let A be that 10 is the sum of the two die.
$P(A)$=3/36

$$P(A)=\frac{n(A)}{n(S)}$$

$$P(B)=\frac{n(B)}{n(S)}$$

Let B that the first die is a 4.
$P(B)$=

What is $\mathrm{P}(A|B)$?.
$P(A|B)$=

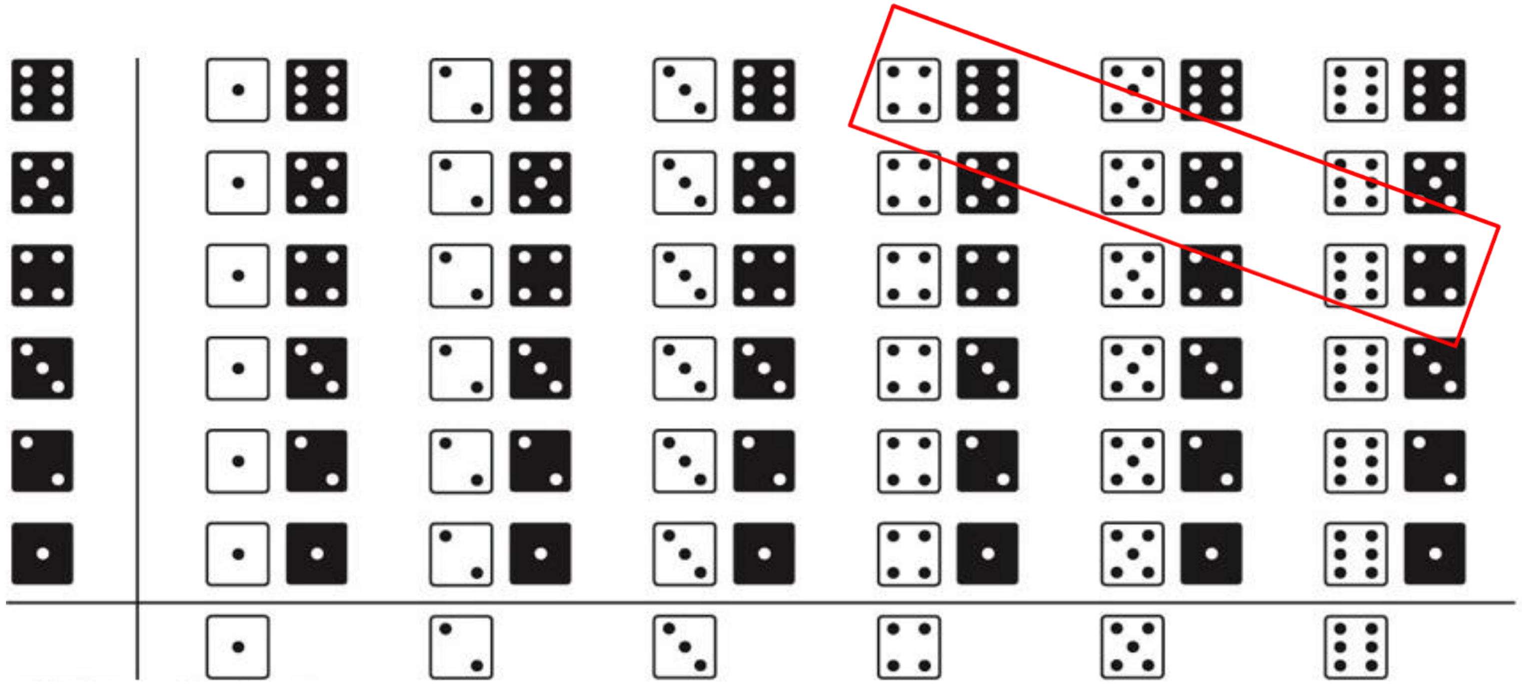


Figure from Johnson & Kuby, 2012.

Conditional Probability and Bayes' Rule

$$P(A)=\frac{n(A)}{n(S)}$$

$$P(B)=\frac{n(B)}{n(S)}$$

Example: Roll two die.
Let A be that 10 is the sum of the two die.
$P(A)$=3/36

Let B that the first die is a 4.
$P(B)$=6/36

What is $\mathrm{P}(A|B)$?.
$P(A|B)$=1/6

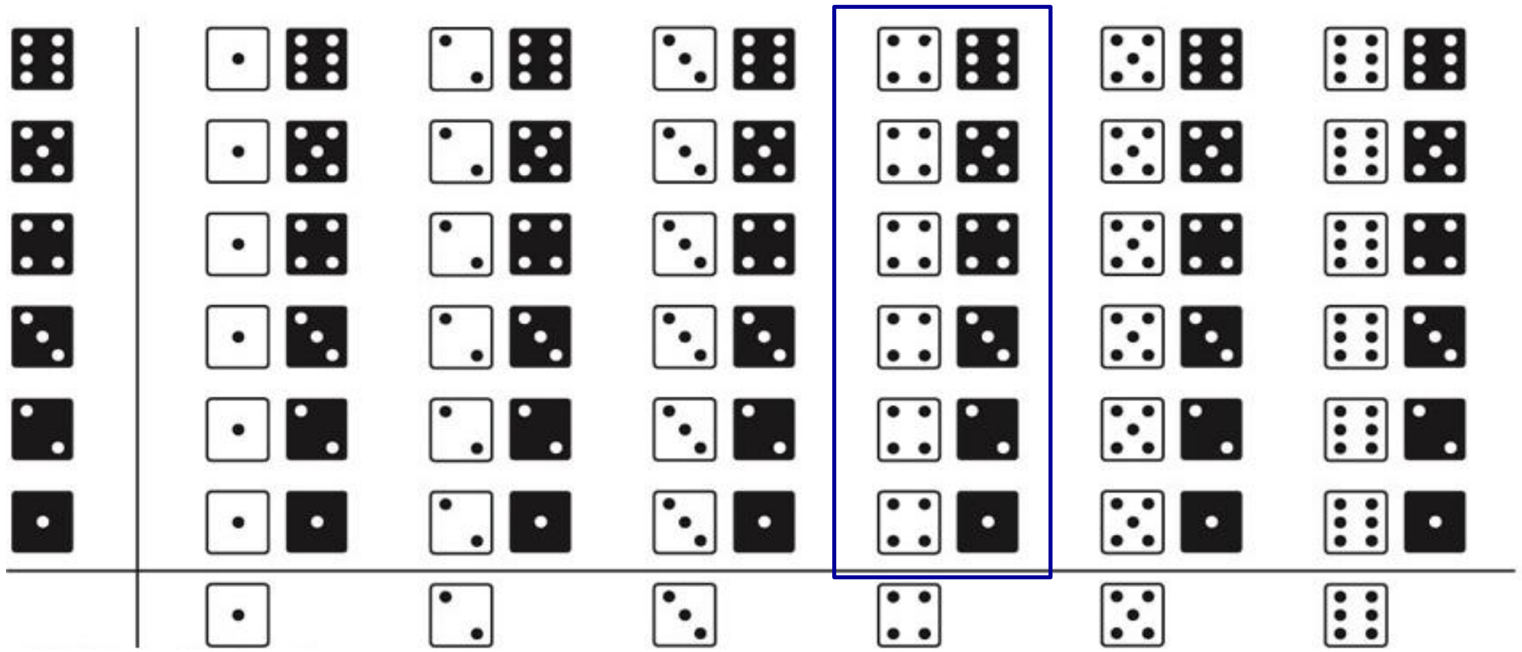


Figure from Johnson & Kuby, 2012.

Conditional Probability and Bayes' Rule

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(B) = \frac{n(B)}{n(S)}$$

Example: Roll two die.

Let A be that 10 is the sum of the two die.

$P(A)$=

Let B that the first die is a 4.

$P(B)$=

What is $\text{P}(A|B)$?.

$P(A|B)$=

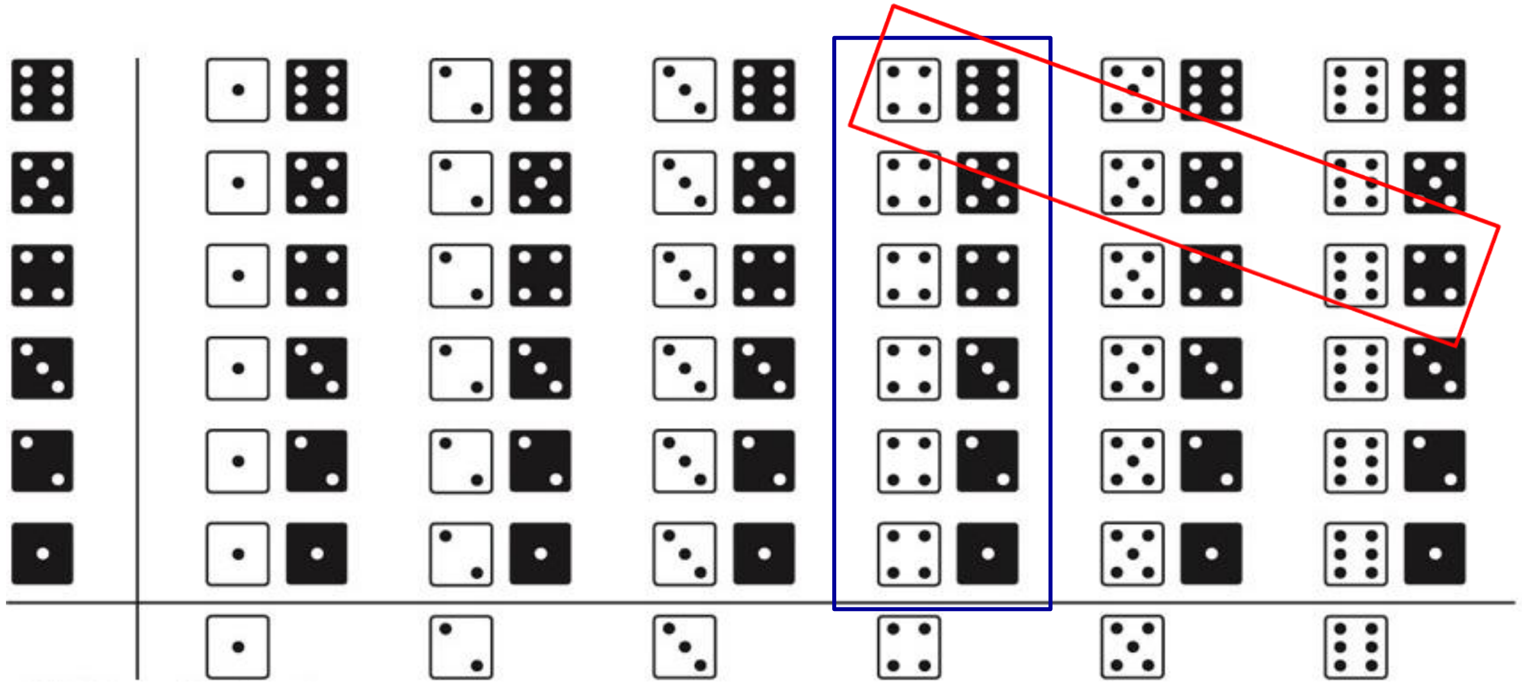


Figure from Johnson & Kuby, 2012.

Conditional Probability and Bayes' Rule

We extended to more A events, $A_1, A_2, \ldots$

Let $B_1, B_2, \ldots$ be a partition of the sample space, and

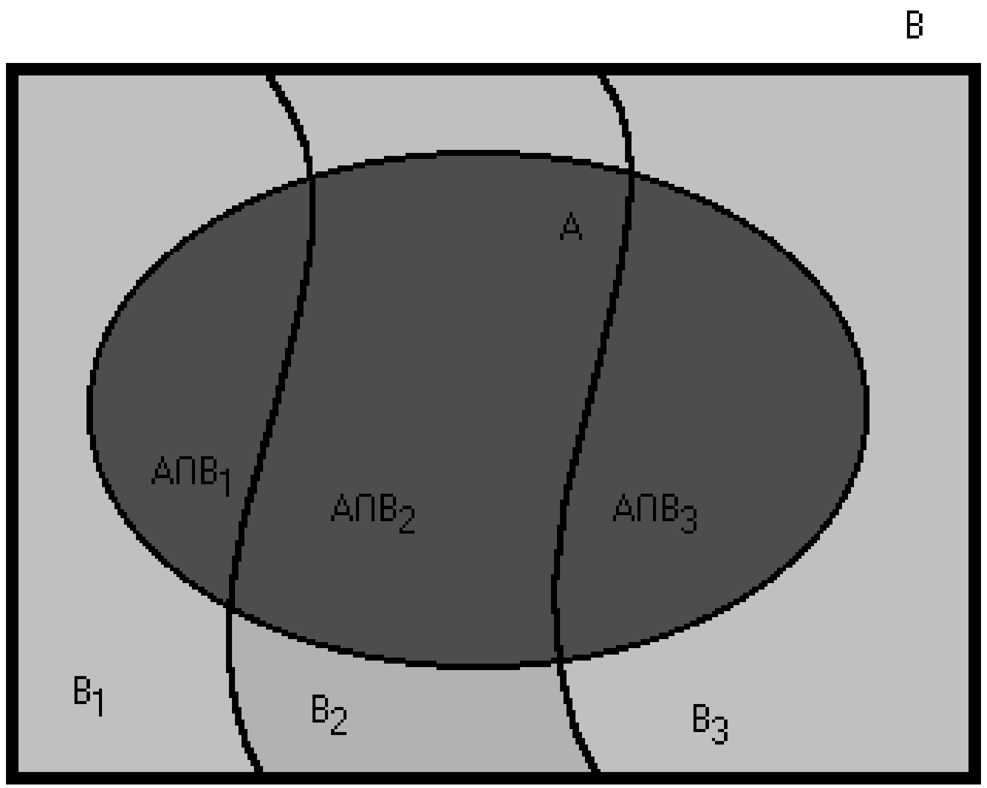
let B be any set.

Then for each $i=1,2,\ldots,$

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A)}$$

$$P(A) = \sum_{i=1}^{\infty} P(A \mid B_i)P(B_i)$$

S



Conditional Probability and Bayes' Rule

Example: Medical Test. P(have disease|test positive).

$T+$: The event that the test is positive.

$T-$: The event that the test is negative.

$D+$: The event that the person truly has disease.

$D-$: The event that the person truly does not has disease.

The sensitivity of test is $P(T+|D-)=0.99$.

The specificity of test is $P(T\text{-}|D-)=0.99$.

If the proportion of population that truly has disease is 10^{-6}.

$$P(D-|T+)=\frac{P(T+|D-)P(D-)}{P(T+)}=0.99990101$$

$$P(T+)=P(T+|D+)P(D+)+P(T+|D-)P(D-)$$

Conditional Probability and Bayes' Rule

We will extend this idea for events A and B

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

to random variables/parameters x and θ

$$f_{\Theta|X}(\theta|x) = \frac{f_{X|\Theta}(x|\theta)\, f_{\Theta}(\theta)}{f_X(x)} \quad ,$$

We will in general not include PDF subscripts.

where

$$f_X(x) = \int f_{X|\Theta}(x|\theta)\, f_{\Theta}(\theta)\; d\theta \quad .$$

Conditional Probability and Bayes' Rule

So we can calculate posterior quantities such as

$$E(\Theta \mid X) = \int \theta f(\theta \mid x)\ d\theta$$

$$\text{var}(\Theta \mid X) = \int [\theta - E(\Theta \mid X)]^2 f(\theta \mid x)\ d\theta$$

$$\text{mode}(\Theta \mid X) = \underset{\theta}{\arg\max}\ f(\theta \mid x)$$

etc…

Discussion

Questions?

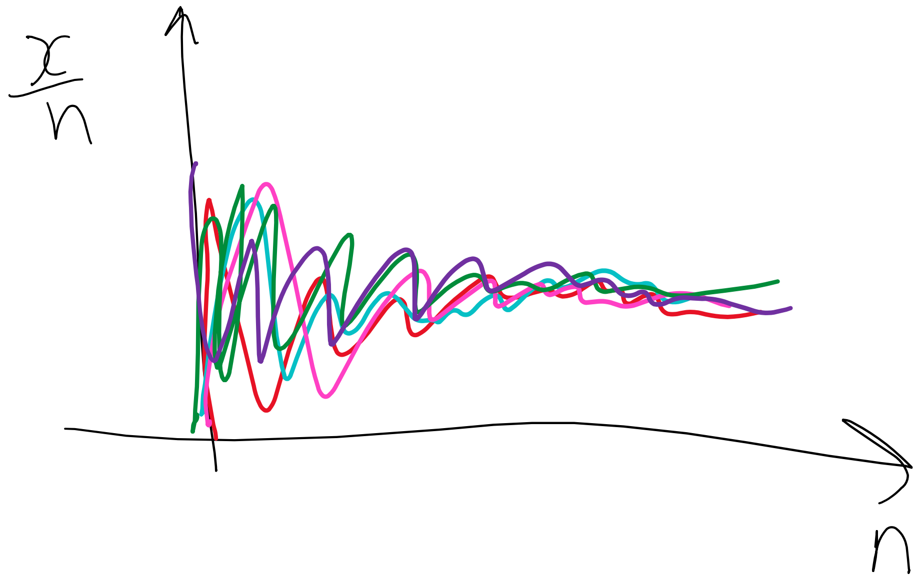
Homework 1b

1. Simulate n=1000 random flips of a coin with probability p=1/2.

   Make a plot of the cumulative probability of heads after each flip.

   $$\hat{p}_n = \frac{x}{n}$$

   Use “hold on” after first plot command and add more simulated flip traces.



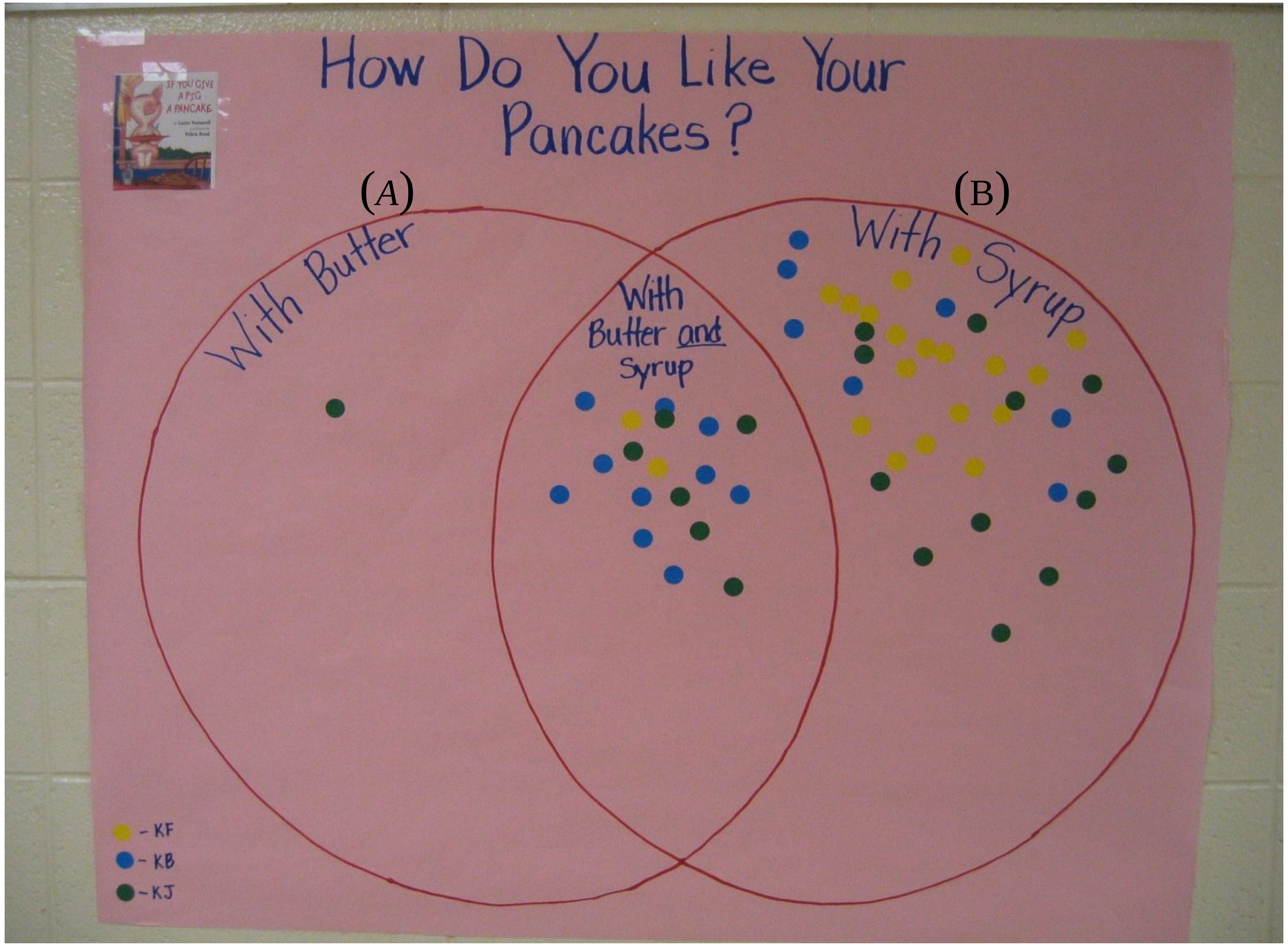
Homework 1b

2. What is the probability that a randomly selected kindergartener likes syrup (B) on their pancakes given that they like butter (A) on their pancakes?

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$

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

i.e. $$P(A)=\frac{n(A)}{n}$$



(each dot is a kindergartner, treat all colors the same)