# ASEN 5044, Fall 2024

## Statistical Estimation for Dynamical Systems

## Lecture 09:
## Joint Probabilities, Marginal Probabilities, Conditional Probabilities, Dependence/Independence

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Tuesday 09/24/2024

Ann and H.J. Smead
Aerospace Engineering Sciences
UNIVERSITY OF COLORADO **BOULDER**

# Announcements

- **HW 3 Due Fri 09/27**
- **HW 4: Out Thurs 09/26, Due Thurs 10/03 [1 week]**
- **Quiz 4: out this Fri 09/27, due Tues 10/01**
- **First advanced topic lecture: this Fri 9/27 (pre-recorded, to be posted)**
  - ***Optional:*** **focused on Bayesian estimation and related topics**
  - **Will post blank + written slides**
- **Midterm 1: out next Thurs 10/03, due Thurs 10/10**
  - One week long take home exam posted to Canvas
  - Due Thurs 10/10/2024 via Gradescope
  - Open book/notes – honor code applies (must complete by yourself)
  - Will cover HW 1-4 (solns will be posted) + Quizzes 1-4 + associated lectures
  - No office hours for TFs or Prof. Ahmed from 10/07-10/10 (private questions via email/Piazza only)
  - Effort: ~5.5-8 hrs (including sanity checking/proof reading time)

# Overview

Last time: Intro to Probability (pre-recorded lectures and notes)
- Motivation
- Formal definitions: sample spaces, event spaces, axioms

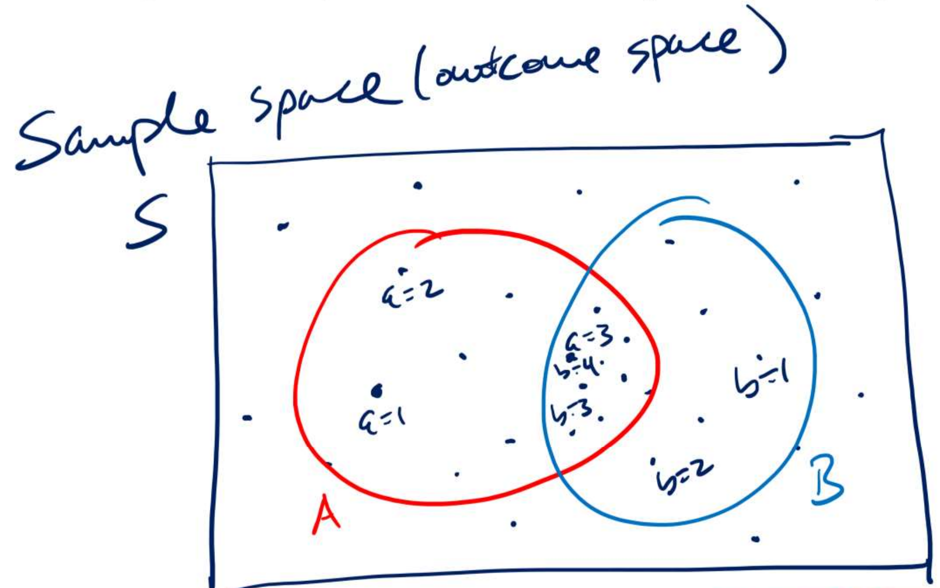
Today: other important fundamental concepts
- Joint probabilities
- Marginal probabilities
- Conditional probabilities
- Bayes' Rule
- Dependent/independent probabilities

- Next time:
  - Random variables (RVs)
  - Probability distributions for RVs (discrete/continuous)
  - Probability density functions (pdfs) for continuous RVs

**READ: Chapter 2.4 in Simon book**

# Joint Distributions

- The joint probability P(A,B) is the probability that events A and B occur *simultaneously*




**Definition:** $P(A=a \text{ AND } B=b) = P(A=a \land B=b)$

$= P(A=a \,\&\, B=b)$

$= P(A=a, B=b)$

Can be extended to arbitrary number of events:

$P(A=a \,\&\, B=b \,\&\, C=c \,\&\, D=d \,\&\, \ldots)$

$=$ probability that $\{A=a, B=b, C=c, D=d, \ldots\}$ occur simultaneously

(*) Note generally: $P(A,B) \neq P(A) \cdot P(B)$ (could be true in some cases but NOT ALWAYS!)

*ALSO NOTE: Must have (from probability axioms):

$$\sum_a \sum_b P(A=a, B=b) = 1$$

(i.e. joint prob must sum to 1 for all possible values $a$ for $A$ and $b$ for $B$)

**Example:** Draw card from standard 52 card playing deck

2 separate events
- A : suit of card (hearts $a=1$, spades $a=2$, clubs $a=3$, diamonds $a=4$)
- B = #/rank of card — 13 different outcomes $b_i$ ($i=1:13$)

$P(A=\text{hearts}, B=10) = \frac{1}{52}$

$C$ = card odd $\rightarrow c=1$ / card even $\rightarrow c=0$

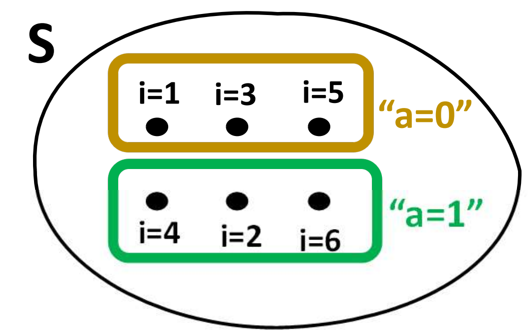
$D$ = face card $\rightarrow d=1$ / # card $\rightarrow d=0$

$P(C=1 \,\&\, D=0) = \frac{\#\text{ cards odd } \& \text{ \# cards}}{\text{total \# cards}}$

# Example: Joint Distribution on Die Roll Events

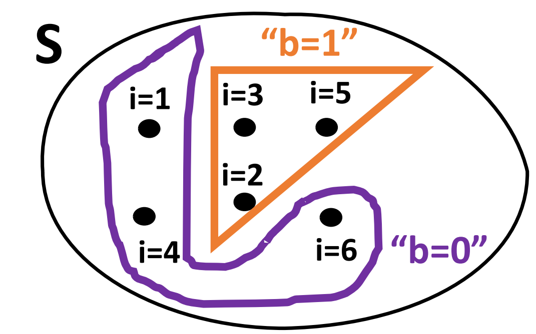
- Consider rolling of a fair 6-sided die with outcome i
- Let event A=1 if roll is even (i = 2,4,6); A=0 otherwise
- Let event B=1 if roll is prime (i = 2,3,5); B=0 otherwise
- Here: S (outcome space/sample space) is {i=1,…,i=6}

**Event space $A$ on $S$**




'$A = a$': event $a \in \{0, 1\}$ occurs in space of event $A$

**Event space $B$ on $S$**




'$B = b$': event $b \in \{0, 1\}$ occurs in space of event $B$

**Examine possible joint event outcomes:**

| $A$ | $B$ | occurs if… |
|---|---|---|
| 0 | 0 | i=1: not even and not prime → Probability = 1/6 |
| 0 | 1 | i=3,5: not even and is prime → Probability = 2/6 |
| 1 | 0 | i=4,6: is even and not prime → Probability = 2/6 |
| 1 | 1 | i=2: is even and prime → Probability = 1/6 |

**Joint Probability Table (JPT):**

| $P(A \wedge B)$ | $B:$ $b = 0$ | $b = 1$ |
|---|---|---|
| $A:$ $a = 0$ | $\frac{1}{6}$ | $\frac{2}{6}$ |
| $a = 1$ | $\frac{2}{6}$ | $\frac{1}{6}$ |

**Sum over all rows and columns = 1:**

$$\sum_a \sum_b P(A = a, B = b) = 1$$

Note: $P(A = 1) = P(i \text{ is even}) = \frac{1}{2}$

$P(B = 1) = P(i \text{ is prime}) = \frac{1}{2}$

$P(A = 1) \cdot P(B = 1)$

$= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{1}{6}$!!!

recall: in general, $P(A) \cdot P(B) \neq P(A, B)$

# Marginal Probabilities

- How are P(A) and P(B) related to P(A,B)?
- Turns out we can always recover P(A) and P(B) from the joint probability P(A,B), (even though **in general** $P(A)P(B) \neq P(A,B)$)

Given joint probabilities $P(A=a, B=b)$, define:

$$\text{Marginal distribution of } A\text{: } P(A=a) = \sum_b P(A=a, B=b)$$

$$\text{Marginal distribution of } B\text{: } P(B=b) = \sum_a P(A=a, B=b)$$

Also holds for more than 2 events, e.g. if 3rd event C with $P(A, B, C)$, then

$$P(A=a) = \sum_b \sum_c P(A=a, B=b, C=c),$$
$$P(B=b) = \sum_a \sum_c P(A=a, B=b, C=c),$$
$$P(C=c) = \sum_a \sum_b P(A=a, B=b, C=c)$$

Also possible:

$$P(A=a, C=c) = \sum_b P(A=a, B=b, C=c)$$

# Examples of Marginal Probabilities

- 6-sided die example again: A: roll is even # (1=yes, 0 = no)

  B: roll is prime # (1=yes, 0 = no)

| P(A & B) | B=0 | B=1 |
|---|---|---|
| A=0 | 1/6 | 2/6 |
| A=1 | 2/6 | 1/6 |

$1/2 = P(A=0)$

$1/2 = P(A=1)$

$P(B=0) = 1/2$ $P(B=1) = 1/2$

Marginal Prob of $A = 0$: $P(A=0) = \sum_b P(A=0, B=b)$
$$= P(A=0, B=0) + P(A=0, B=1) = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}$$

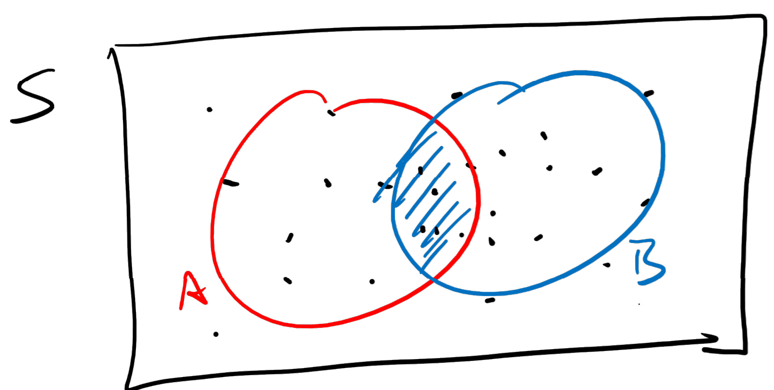
Marginal Prob of $A = 1$: $P(A=1) = \sum_b P(A=1, B=b)$
$$= P(A=1, B=0) + P(A=1, B=1) = \frac{2}{6} + \frac{1}{6} = \frac{1}{2}$$

Likewise, can show that:

$$P(B=0) = \sum_a P(A=a, B=0) = \frac{1}{2} \qquad P(B=1) = \sum_a P(A=a, B=1) = \frac{1}{2}$$

# Conditional Probabilities

- Are events A and B related, such that knowing whether/not B occurs alters prob. of A?




P(A = a GIVEN B=b is KNOWN/SUPPOSED to occur)

$$= P(A=a \mid B=b) \triangleq \frac{P(A=a \;\&\; B=b)}{P(B=b)} = \frac{P(A=a \;\&\; B=b)}{\sum_a P(A=a \;\&\; B=b)}$$

"GIVEN" or "conditioned on" the event(s)

Basically: restricting sample/outcome space S such that we only care about the part of S where the event B=b must occur (renormalizing joint probability to be consistent with GIVEN statement B=b)

→ For more than 2 events, e.g. $P(A, B, C)$:

$$P(A \mid B, C) = \frac{P(A,B,C)}{P(B,C)} = \frac{P(A,B,C)}{\sum_a P(A=a,B,C)}$$

$$P(A, C \mid B) = \frac{P(A,B,C)}{P(B)} = \frac{P(A,B,C)}{\sum_a \sum_c P(A=a,B,C=c)}$$

# Example: Conditional Probabilities

- 6-sided die again:

A: roll is even # (1=yes, 0 = no)

B: roll is prime # (1=yes, 0 = no)

| P(A & B) | B=0 | B=1 |
|---|---|---|
| A=0 | 1/6 | 2/6 |
| A=1 | 2/6 | 1/6 |

$$P(B=0|A=1) = \frac{P(A{=}1 \ \& \ B{=}0)}{P(A{=}1)} = \frac{(2/6)}{(1/2)} = \frac{2}{3}$$

$$P(B=1|A=1) = \frac{P(A{=}1 \ \& \ B{=}1)}{P(A{=}1)} = \frac{(1/6)}{(1/2)} = \frac{1}{3} \ [= 1 - P(B=0|A=1)]$$

$$P(B=0|A=0) = \frac{P(A{=}0 \ \& \ B{=}0)}{P(A{=}0)} = \frac{(1/6)}{(1/2)} = \frac{1}{3}$$

$$P(A=0|B=0) = \frac{P(A{=}0 \ \& \ B{=}0)}{P(B{=}0)} = \frac{(1/6)}{(1/2)} = \frac{1}{3}$$

# Consequences of Conditioning

- FACT #1: P(A,B) can always be **conditionally factored** in two ways

(chain rule of probability)

$$P(A,B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$$

$$\left( P\cancel{(A)} \cdot \frac{P(A,\ B)}{P\cancel{(A)}} = P\cancel{(B)} \cdot \frac{P(A,\ B)}{P\cancel{(B)}} \right)$$

- FACT #2: **the law of total probability**: because of FACT #1 and definition of marginal distributions, we have

$$P(A=a) = \sum_b P(A=a, B=b) = \sum_b P(B=b) \cdot P(A=a|B=b)$$

$$P(B=b) = \sum_a P(A=a, B=b) = \sum_a P(A=a) \cdot P(B=b|A=a)$$

# Bayes' Rule for Reverse Conditioning

- Very handy for "**inverse problems**", where we see the "effects" B=b (evidence) and want to infer the "cause" A (explanation), based only on knowing P(A) and P(B=b|A)
  - i.e. useful in cases where P(A) & P(B|A) are easy to specify, but P(A|B) is not…
- Allows us to update P(A) [prior belief in A] given new data (observations) B=b
  - P(A) [a priori, old belief before data] → P(A|B=b) [a posteriori, new belief given observed data]
- **Derivation:** start with FACT #1 from previous slide:

$$P(A,B) = P(A)\cdot P(B|A) = P(B)\cdot P(A|B)$$

Now, since $P(A)\cdot P(B|A) = P(B)\cdot P(A|B)$, re-arrange and solve for

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

$\rightarrow$ but: $P(B=b) = \sum_a P(A=a, B=b) = \sum_a P(A=a)\cdot P(B=b|A=a)$, so:

(from FACT #2)

$$\boxed{P(A|B=b) = \frac{P(A)\,P(B{=}b|A)}{\sum_a P(A{=}a)P(B{=}b|A{=}a)}}$$

likelihood

prior

Bayes' Rule ⊛ (or Bayes' Theorem)

# Bayes' Rule Example: "Bayesian Inference"

- What is the probability that Prof. Ahmed is in his home given that lights are on?

A: Ahmed is at home (0 = no, 1 = yes)

B: Lights are on in his window (0 = no, 1 = yes)

| P(A=0) | P(A=1) |
|---|---|
| 0.5 | 0.5 |

| P(B=0\|A=0) | P(B=1\|A=0) |
|---|---|
| 0.8 | 0.2 |

| P(B=0\|A=1) | P(B=1\|A=1) |
|---|---|
| 0.1 | 0.9 |

Want to use this data to find $P(A=1|B=1)$.

From Bayes' rule, we get

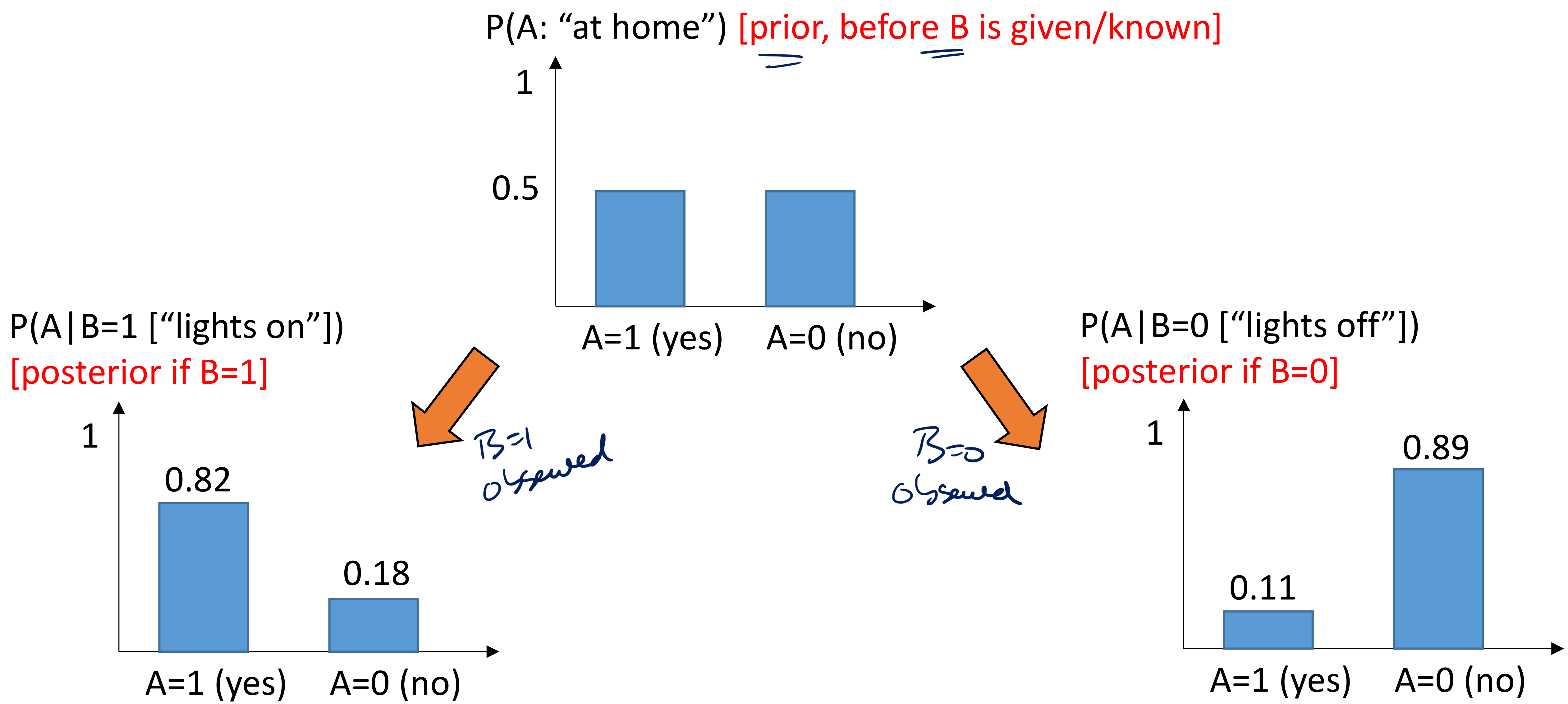
$$P(A=1|B=1) = \frac{P(A{=}1)\cdot P(B{=}1|A{=}1)}{\sum_a P(A{=}a)\cdot P(B{=}1|A{=}a)} = \frac{P(A{=}1)\cdot P(B{=}1|A{=}1)}{P(A{=}0)\cdot P(B{=}1|A{=}0)+P(A{=}1)\cdot P(B{=}1|A{=}1)}$$

$$= \frac{0.5\cdot 0.9}{(0.5\cdot 0.2)+(0.5\cdot 0.9)} = \frac{0.9}{0.2+0.9} \approx 0.82$$

$P(B=1)$ = prob. lights on
No matter what A is
$= (0.5)(0.2)+(0.5)(0.9)$
$= 0.55$

# Bayes' Rule Example

- Can also apply Bayes' rule to compute full posterior distribution of A given B=0, or given B=1
- Compare posterior (probs of A after Bayes' rule) to prior (probs of A before Bayes' rule, i.e. NOT given B)



# Independent Events and Independent Probabilities

- If knowledge of the occurrence of event B <u>never</u> alters P(A), then we say that A and B are **independent events**

if $A \perp\!\!\!\perp B$, then $P(A=a \,\&\, B=b) = P(A=a) \cdot P(B=b) \;\; \forall \; \underline{\underline{a \,\&\, b}}$ (for all)

"independent" [product of marginals]

i.e. events A & B have <u>no</u> <u>information</u> about each other

$\longrightarrow$ so if $A \perp\!\!\!\perp B$, then $P(A|B) = \dfrac{P(A \,\&\, B)}{P(B)} = \dfrac{P(A) \cdot P(B)}{P(B)} = P(A)$

$$\Rightarrow \boxed{\text{if } A \perp\!\!\!\perp B, \text{ then } P(A|B) = P(A) \;\; \& \;\; P(B|A) = P(B)}$$

(*) (*) This must hold for every possible configuration $A=a \,\&\, B=b$!

<u>Ex.</u>:
- prob. of getting "heads" on 1st & 2nd coin flip are $\perp\!\!\!\perp$
- prob. of getting prime & even die rolls are <u>NOT</u> $\perp\!\!\!\perp$