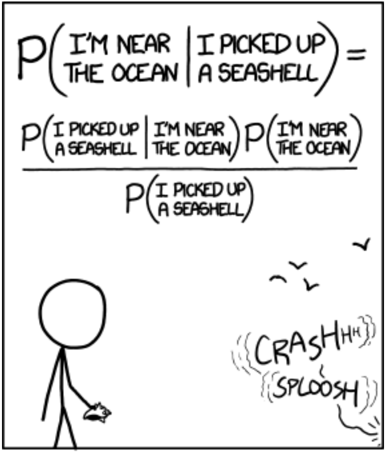
Bayesian Updating: Discrete Priors: 18.05 Spring 2014



STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND *DON'T* HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.

http://xkcd.com/1236/

Learning from experience

Which treatment would you choose?

1. Treatment 1: cured 100% of patients in a trial.
2. Treatment 2: cured 95% of patients in a trial.
3. Treatment 3: cured 90% of patients in a trial.

Which treatment would you choose?

1. Treatment 1: cured 3 out of 3 patients in a trial.
2. Treatment 2: cured 19 out of 20 patients treated in a trial.

3. Standard treatment: cured 90000 out of 100000 patients in clinical practice.

Which die is it?

- I have a bag containing dice of two types: 4-sided and 10-sided.
- Suppose I pick a die at random and roll it.
- Based on what I rolled which type would you guess I picked?

• Suppose you find out that the bag contained one 4-sided die and one 10-sided die. Does this change your guess?

• Suppose you find out that the bag contained one 4-sided die and 100 10-sided dice. Does this change your guess?

Board Question: learning from data

- A certain disease has a prevalence of 0.005.
- A screening test has 2% false positives an 1% false negatives.

Suppose a patient is screened and has a positive test.

1. Represent this information with a tree and use Bayes' theorem to compute the probabilities the patient does and doesn't have the disease.
2. Identify the data, hypotheses, likelihoods, prior probabilities and posterior probabilities.
3. Make a full likelihood table containing all hypotheses and possible test data.
4. Redo the computation using a Bayesian update table. Match the terms in your table to the terms in your previous calculation.

Solution on next slides.

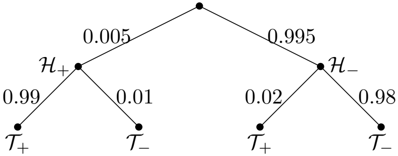
Solution

1. Tree based Bayes computation

Let $\mathcal{H}_+$ mean the patient has the disease and $\mathcal{H}_-$ they don't.

Let $\mathcal{T}_+$: they test positive and $\mathcal{T}_-$ they test negative.

We can organize this in a tree:



Bayes' theorem says $P(\mathcal{H}_+ \,|\, \mathcal{T}_+) = \dfrac{P(\mathcal{T}_+ \,|\, \mathcal{H}_+)P(\mathcal{H}_+)}{P(\mathcal{T}_+)}$.

Using the tree, the total probability

$$\begin{aligned} P(\mathcal{T}_+) &= P(\mathcal{T}_+ \,|\, \mathcal{H}_+)P(\mathcal{H}_+) + P(\mathcal{T}_+ \,|\, \mathcal{H}_-)P(\mathcal{H}_-) \\ &= 0.99 \cdot 0.005 + 0.02 \cdot 0.995 = 0.02485 \end{aligned}$$

Solution continued on next slide.

Solution continued

So,

$$P(\mathcal{H}_+ \,|\, \mathcal{T}_+) = \frac{P(\mathcal{T}_+ \,|\, \mathcal{H}_+)P(\mathcal{H}_+)}{P(\mathcal{T}_+)} = \frac{0.99 \cdot 0.005}{0.02485} = 0.199$$

$$P(\mathcal{H}_- \,|\, \mathcal{T}_+) = \frac{P(\mathcal{T}_+ \,|\, \mathcal{H}_-)P(\mathcal{H}_-)}{P(\mathcal{T}_+)} = \frac{0.02 \cdot 0.995}{0.02485} = 0.801$$

The positive test greatly increases the probability of $\mathcal{H}_+$, but it is still much less probable than $\mathcal{H}_-$.

Solution continued on next slide.

Solution continued

2. Terminology

Data: The data are the results of the experiment. In this case, the positive test.

Hypotheses: The hypotheses are the possible answers to the question being asked. In this case they are $\mathcal{H}_+$ the patient has the disease; $\mathcal{H}_-$ they don't.

Likelihoods: The likelihood given a hypothesis is the probability of the data given that hypothesis. In this case there are two likelihoods, one for each hypothesis

$$P(\mathcal{T}_+ \,|\, \mathcal{H}_+) = 0.99 \qquad \text{and} \qquad P(\mathcal{T}_+ \,|\, \mathcal{H}_-) = 0.02.$$

We repeat: the likelihood is a probability **given** the hypothesis, **not** a probability of the hypothesis.

Continued on next slide.

Solution continued

Prior probabilities of the hypotheses: The priors are the probabilities of the hypotheses prior to collecting data. In this case,

$$P(\mathcal{H}_+) = 0.005 \qquad \text{and} \qquad P(\mathcal{H}_-) = 0.995$$

Posterior probabilities of the hypotheses: The posteriors are the probabilities of the hypotheses **given** the data. In this case

$$P(\mathcal{H}_+ \,|\, \mathcal{T}_+) = 0.199 \qquad \text{and} \qquad P(\mathcal{H}_- \,|\, \mathcal{T}_+) = 0.801.$$

Posterior ↓ Likelihood ↓ Prior ↓

$$P(\mathcal{H}_+ \,|\, \mathcal{T}_+) = \frac{P(\mathcal{T}_+ \,|\, \mathcal{H}_+) \cdot P(\mathcal{H}_+)}{P(\mathcal{T}_+)}$$

↗ Total probability of the data

Solution continued

3. Full likelihood table

The table holds likelihoods $P(\mathcal{D}|\mathcal{H})$ for every possible hypothesis and data combination.

hypothesis $\mathcal{H}$	likelihood $P(\mathcal{D}\|\mathcal{H})$	
disease state	$P(\mathcal{T}_+\|\mathcal{H})$	$P(\mathcal{T}_-\|\mathcal{H})$
$\mathcal{H}_+$	0.99	0.01
$\mathcal{H}_-$	0.02	0.98

Notice in the next slide that the $P(\mathcal{T}_+ \mid \mathcal{H})$ column is exactly the likelihood column in the Bayesian update table.

Solution continued

4. Calculation using a Bayesian update table

$\mathcal{H}$ = hypothesis: $\mathcal{H}_+$ (patient has disease); $\mathcal{H}_-$ (they don't).

Data: $\mathcal{T}_+$ (positive screening test).

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{T}_+\|\mathcal{H})$	$P(\mathcal{T}_+\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{T}_+)$
$\mathcal{H}_+$	0.005	0.99	0.00495	0.199
$\mathcal{H}_-$	0.995	0.02	0.0199	0.801
total	1	NO SUM	$P(\mathcal{T}_+) = 0.02485$	1

Data $\mathcal{D} = \mathcal{T}_+$

Total probability: $P(\mathcal{T}_+)$ = sum of Bayes numerator column = 0.02485

Bayes' theorem: $P(\mathcal{H}|\mathcal{T}_+) = \dfrac{P(\mathcal{T}_+|\mathcal{H})P(\mathcal{H})}{P(\mathcal{T}_+)} = \dfrac{\text{likelihood} \times \text{prior}}{\text{total prob. of data}}$

Board Question: Dice

Five dice: 4-sided, 6-sided, 8-sided, 12-sided, 20-sided.

Suppose I picked one at random and, without showing it to you, rolled it and reported a 13.

1. Make the full likelihood table (be smart about identical columns).

2. Make a Bayesian update table and compute the posterior probabilities that the chosen die is each of the five dice.

3. Same question if I rolled a 5.

4. Same question if I rolled a 9.

(Keep the tables for 5 and 9 handy! Do not erase!)

Tabular solution

$\mathcal{D}$ = 'rolled a 13'

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	$1/5$	0	0	0
$\mathcal{H}_6$	$1/5$	0	0	0
$\mathcal{H}_8$	$1/5$	0	0	0
$\mathcal{H}_{12}$	$1/5$	0	0	0
$\mathcal{H}_{20}$	$1/5$	$1/20$	$1/100$	1
total	1		$1/100$	1

Tabular solution

$\mathcal{D}$ = 'rolled a 5'

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	$1/5$	0	0	0
$\mathcal{H}_6$	$1/5$	$1/6$	$1/30$	0.392
$\mathcal{H}_8$	$1/5$	$1/8$	$1/40$	0.294
$\mathcal{H}_{12}$	$1/5$	$1/12$	$1/60$	0.196
$\mathcal{H}_{20}$	$1/5$	$1/20$	$1/100$	0.118
total	1		0.085	1

Tabular solution

$\mathcal{D}$ = 'rolled a 9'

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	$1/5$	0	0	0
$\mathcal{H}_6$	$1/5$	0	0	0
$\mathcal{H}_8$	$1/5$	0	0	0
$\mathcal{H}_{12}$	$1/5$	$1/12$	$1/60$	0.625
$\mathcal{H}_{20}$	$1/5$	$1/20$	$1/100$	0.375
total	1		.0267	1

Iterated Updates

Suppose I rolled a 5 and then a 9.

Update in two steps:

First for the 5

Then update the update for the 9.

Tabular solution

$\mathcal{D}_1$ = 'rolled a 5'
$\mathcal{D}_2$ = 'rolled a 9'
Bayes numerator$_1$ = likelihood$_1 \times$ prior.
Bayes numerator$_2$ = likelihood$_2 \times$ Bayes numerator$_1$

hyp.	prior	likel. 1	Bayes num. 1	likel. 2	Bayes num. 2	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}_1\|\mathcal{H})$	$***$	$P(\mathcal{D}_2\|\mathcal{H})$	$***$	$P(\mathcal{H}\|\mathcal{D}_1, \mathcal{D}_2)$
$\mathcal{H}_4$	1/5	0	0	0	0	0
$\mathcal{H}_6$	1/5	1/6	1/30	0	0	0
$\mathcal{H}_8$	1/5	1/8	1/40	0	0	0
$\mathcal{H}_{12}$	1/5	1/12	1/60	1/12	1/720	0.735
$\mathcal{H}_{20}$	1/5	1/20	1/100	1/20	1/2000	0.265
total	1				0.0019	1

Board Question

Suppose I rolled a 9 and then a 5.

1. Do the Bayesian update in two steps:
   First update for the 9.
   Then update the update for the 5.

2. Do the Bayesian update in one step
   The data is $\mathcal{D} =$ '9 followed by 5'

Tabular solution: two steps

$\mathcal{D}_1$ = 'rolled a 9'
$\mathcal{D}_2$ = 'rolled a 5'
Bayes $\text{numerator}_1 = \text{likelihood}_1 \times$ prior.
Bayes $\text{numerator}_2 = \text{likelihood}_2 \times$ Bayes numerator_1

hyp.	prior	likel. 1	Bayes num. 1	likel. 2	Bayes num. 2	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}_1 \mid \mathcal{H})$	$* * *$	$P(\mathcal{D}_2 \mid \mathcal{H})$	$* * *$	$P(\mathcal{H} \mid \mathcal{D}_1, \mathcal{D}_2)$
$\mathcal{H}_4$	$1/5$	0	0	0	0	0
$\mathcal{H}_6$	$1/5$	0	0	$1/6$	0	0
$\mathcal{H}_8$	$1/5$	0	0	$1/8$	0	0
$\mathcal{H}_{12}$	$1/5$	$1/12$	$1/60$	$1/12$	$1/720$	0.735
$\mathcal{H}_{20}$	$1/5$	$1/20$	$1/100$	$1/20$	$1/2000$	0.265
total	1				0.0019	1

Tabular solution: one step

$\mathcal{D}$ = 'rolled a 9 then a 5'

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\mid\mathcal{H})$	$P(\mathcal{D}\mid\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\mid\mathcal{D})$
$\mathcal{H}_4$	$1/5$	0	0	0
$\mathcal{H}_6$	$1/5$	0	0	0
$\mathcal{H}_8$	$1/5$	0	0	0
$\mathcal{H}_{12}$	$1/5$	$1/144$	$1/720$	0.735
$\mathcal{H}_{20}$	$1/5$	$1/400$	$1/2000$	0.265
total	1		0.0019	1

Board Question: probabilistic prediction

Along with finding posterior probabilities of hypotheses. We might want to make posterior predictions about the next roll.

With the same setup as before let:
$\mathcal{D}_1 =$ result of first roll
$\mathcal{D}_2 =$ result of second roll

(a) Find $P(\mathcal{D}_1 = 5)$.

(b) Find $P(\mathcal{D}_2 = 4 | \mathcal{D}_1 = 5)$.

Solution

$\mathcal{D}_1$ = 'rolled a 5'

$\mathcal{D}_2$ = 'rolled a 4'

hyp.	prior	likel. 1	Bayes num. 1	post. 1	likel. 2	post. 1 × likel. 2
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}_1\|\mathcal{H})$	$***$	$P(\mathcal{H}\|\mathcal{D}_1)$	$P(\mathcal{D}_2\|\mathcal{H},\mathcal{D}_1)$	$P(\mathcal{D}_2\|\mathcal{H},\mathcal{D}_1)P(\mathcal{H}\|\mathcal{D}_1)$
$\mathcal{H}_4$	1/5	0	0	0	$*$	0
$\mathcal{H}_6$	1/5	1/6	1/30	0.392	1/6	$0.392 \cdot 1/6$
$\mathcal{H}_8$	1/5	1/8	1/40	0.294	1/8	$0.294 \cdot 1/40$
$\mathcal{H}_{12}$	1/5	1/12	1/60	0.196	1/12	$0.196 \cdot 1/12$
$\mathcal{H}_{20}$	1/5	1/20	1/100	0.118	1/20	$0.118 \cdot 1/20$
total	1		0.085	1		0.124

The law of total probability tells us $P(\mathcal{D}_1)$ is the sum of the Bayes numerator 1 column in the table: $\boxed{P(\mathcal{D}_1) = 0.085}$.

The law of total probability tells us $P(\mathcal{D}_2|\mathcal{D}_1)$ is the sum of the last column in the table: $\boxed{P(\mathcal{D}_2|\mathcal{D}_1) = 0.124}$