Bayes’ Theorem

Chengbo ZHENG

Bayes' Theorem

- **Theorem**
  Suppose that *E* and *F* are events from a sample space *S* such that $p(E) \neq 0$ and $p(F) \neq 0$. Then:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Proof:

Bayes' Theorem.

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\bar{F})p(\bar{F})}$$

E, F are events from a sample space S

$$p(E|F)p(F) = \frac{p(E \cap F)}{p(F)} p(F) = p(E \cap F)$$

$$p(E|\bar{F})p(\bar{F}) = p(E \cap \bar{F})$$

$$p(E \cap F) + p(E \cap \bar{F}) = p(E) \rightarrow \text{Denominator}$$

Numerator: $p(E \cap F)$

By Definition: $p(F|E) = \dfrac{p(E \cap F)}{p(E)}$

$$\underbrace{p(F|E)}_{\text{Posterior}} = \frac{\underbrace{p(E|F)}_{\text{Likelihood}}\;\underbrace{p(F)}_{\text{Prior}}}{\underbrace{p(E)}_{\text{Evidence}}}$$

- Naive Bayes Classifier.
- Bayes Belief Network......

Example

- **Example**: We have two boxes.
  - Box 1 contains 2 green balls and 7 red balls.
  - Box 2 contains 4 green balls and 3 red balls.

  Bob first picks one of the boxes at random. Then he selects a ball from that box at random. If he has a red ball, what is the probability that he picked box 1 at first.

Example: • Box 1 : 2 green balls, 7 red balls.
• Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

① pick the box.

② Select the ball.

Problem: If has a Red Ball, what's the probability the picked Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Example

- **Example**: We have two boxes.
  - Box 1 contains 2 green balls and 7 red balls.
  - Box 2 contains 4 green balls and 3 red balls.

  Bob first picks one of the boxes at random. Then he selects a ball from that box at random. If he has a red ball, what is the probability that he picked box 1 at first.

Example: • Box 1 : 2 green balls, 7 red balls.

• Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

① pick the box.

② Select the ball.

Problem: If has a Red Ball,

what's the probability the picked

Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Goal ? $P(F|E)$

Example: • Box 1 : 2 green balls, 7 red balls.

• Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

1. pick the box.
2. Select the ball.

Problem: If has a Red Ball,

what's the probability the picked

Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Goal ? $P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F)+P(E|\bar{F})P(\bar{F})}$

What we know?

$P(E|F) = \boxed{}$

$P(E|\bar{F}) = \boxed{}$

$P(F) = \boxed{}$

Example: • Box 1 : 2 green balls, 7 red balls.
• Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

①. pick the box.

②. Select the ball.

Problem: If has a Red Ball,
what's the probability the picked Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Goal ? $P(F|E) = \dfrac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}$

What we know?

$P(E|F) = \frac{7}{9}$

$P(E|\bar{F}) = \boxed{}$

$P(F) = \boxed{}$

Example: • Box 1 : 2 green balls, 7 red balls.
• Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

①. pick the box.

②. Select the ball.

Problem: If has a Red Ball,
what's the probability the picked Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Goal ? $P(F|E) = \dfrac{P(E|F)P(F)}{P(E|F)P(F)+P(E|\bar{F})P(\bar{F})}$

What we know?

$P(E|F) = \frac{7}{9}$

$P(E|\bar{F}) = \frac{3}{7}$

$P(F) = \boxed{}$

Example:
- Box 1 : 2 green balls, 7 red balls.
- Box 2 : 4 green balls, 3 red balls.

Bob do 2-step ~~things~~.

① pick the box.

② Select the ball.

Problem: If has a Red Ball, what's the probability the picked Box 1 first.

E : Bob has chosen a red Ball

F : Bob has chose Box 1

Goal? $P(F|E) = \dfrac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}$

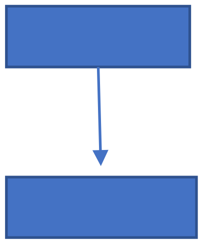
What we know?

$P(E|F) = \frac{7}{9}$

$P(E|\bar{F}) = \frac{3}{7}$

$P(F) = \frac{1}{2}$

$$\therefore P(F|E) = \frac{\frac{7}{9} \times \frac{1}{2}}{\frac{7}{9} \times \frac{1}{2} + \frac{3}{7} \times \frac{1}{2}}$$



```
In [1]: (7/18)/(7/18+3/14)
Out[1]: 0.6447368421052632
```

Example

- **Example**: There is a test for a particular disease.
  - The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
  - The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
  - On average, one person out of 100,000 has the disease.
- Question:
  Should someone who tests positive be worried?

Example

- **Example**: There is a test for a particular disease.
  - The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
  - The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
  - On average, one person out of 100,000 has the disease.
- Question:
  Should someone who tests positive be worried?

D: the person has the disease.

E: this person tests positive

Goal ?

Example

- **Example**: There is a test for a particular disease.
  - The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
  - The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
  - On average, one person out of 100,000 has the disease.
- Question:
  Should someone who tests positive be worried?

D: the person has the disease.

E: this person tests positive

Goal: $P(D|E)$

$$= \frac{P(E|D)P(D)}{P(E|D)P(D)+P(E|\bar{D})P(\bar{D})}.$$

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

false positive	true positive
false negative	true negative.

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

false positive. $P(E \mid \bar{D})$	true positive
false negative $P(\bar{E} \mid D)$	true negative.

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

false positive. $P(E\mid\bar{D})$	true positive $P(E\mid D)$
false negative $P(\bar{E}\mid D)$	true negative. $P(\bar{E}\mid\bar{D})$

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

false positive. $P(E \mid \bar{D})$ 0.5%	true positive $P(E \mid D)$
false negative $P(\bar{E} \mid D)$ 1%	true negative. $P(\bar{E} \mid \bar{D})$

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

false positive. $P(E \mid \bar{D})$ 0.5%	true positive $P(E \mid D)$ $1 - 1\%$
false negative $P(\bar{E} \mid D)$ 1%	true negative. $P(\bar{E} \mid \bar{D})$ $1 - 0.5\%$

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

$$p(D|E) = \frac{p(E|D)p(D)}{p(E|D)p(D) + p(E|\overline{D})p(\overline{D})}$$

$$= \frac{(0.99)(0.00001)}{(0.99)(0.00001) + (0.005)(0.99999)}$$

$$\approx 0.002$$

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

$$p(D|E) = \frac{p(E|D)p(D)}{p(E|D)p(D) + p(E|\overline{D})p(\overline{D})}$$
$$= \frac{(0.99)(0.00001)}{(0.99)(0.00001) + (0.005)(0.99999)}$$
$$\approx 0.002$$

- Why? Why testing positive does not imply high probability of getting disease?

- The test's false negative rate is 1%, i.e., it gives a negative result with prob. 1% when given to someone with the disease.
- The test's false positive rate is 0.5%, i.e., it gives a positive result with prob. 0.5% when given to someone without the disease.
- On average, one person out of 100,000 has the disease.

D: the person has the disease.

E: this person tests positive

- Why? Why testing positive does not imply high probability of getting disease?

$$p(D|E) = \frac{p(E|D)p(D)}{p(E|D)p(D) + p(E|\overline{D})p(\overline{D})}$$

$$= \frac{(0.99)(0.00001)}{(0.99)(0.00001) + (0.005)(0.99999)}$$

$$\approx 0.002$$

False positive rate is much higher than the probability of getting disease.

If the false positive rate is 0.00001, p(D | E)=0.4975

- What if the result is negative?

Goal?

- What if the result is negative?

Goal? $p(D|\bar{E})$

$$p(\overline{D}|\overline{E}) = \frac{p(\overline{E}|\overline{D})p(\overline{D})}{p(\overline{E}|\overline{D})p(\overline{D}) + p(\overline{E}|D)p(D)}$$

$$= \frac{(0.995)(0.99999)}{(0.995)(0.99999) + (0.01)(0.00001)}$$

$$\approx 0.9999999$$

$$p(D|\overline{E})$$

$$\approx 1 - 0.9999999$$

$$= 0.0000001.$$

- What if the result is negative?

Goal? $p(D|\bar{E})$

false positive $P(E\|\bar{D})$ 0.5%	true positive $P(E\|D)$ 1-1%
false negative $P(\bar{E}\|D)$ 1%	true negative $P(\bar{E}\|\bar{D})$ 1-0.5%

$$p(D|\bar{E}) = \frac{p(\bar{E}|D)\,p(D)}{p(\bar{E}|D)p(D) + p(\bar{E}|\bar{D})p(\bar{D})}$$

$$= \frac{0.01 \times 0.00001}{0.01 \times 0.00001 + 0.995 \times 0.99999}$$

=

```
In [2]: 0.01*0.00001/(0.01*0.00001+0.995*0.99999)
Out[2]: 1.0050350749703402e-07
```

Bayesian Spam Filter

- Given:
  - A set B of spam messages
  - A set G of non-spam messages
- Goal: Compute the probability that a new email message is spam

- Given:
  - $|B| = 2000$
  - $|G| = 1000$
  - The word “Rolex” occurs in 250 spam messages and 5 good messages

△ Filter Based on "spam words"....

eg: "on sale", "discount", "Apple", "Rolex"....

Sometimes we call these selected words as the features of the emails used by spam filter

Notations

- B: spam messages
- G: non-spam messages
- w: the particular word, i.e., “Rolex” in our case
- $n_B(w)$: The number of messages in B that w occurs
- $n_G(w)$: The number of messages in G that w occurs

Goal: $p(B|W)$

$$= \frac{p(W|B)\,p(B)}{p(W|B)\,p(B) + p(W|G)\,p(G)}$$

$p(W|B) =$

$p(W|G) =$

$p(B) =$

$p(G) =$

Goal: $p(B|w)$

$$= \frac{p(w|B)\,p(B)}{p(w|B)\,p(B) + p(w|G)\,p(G)}.$$

$$p(w|B) = \frac{n_B(w)}{|B|} = \frac{250}{2000}$$

$$p(w|G) = \frac{n_G(w)}{|G|} = \frac{5}{1000}$$

$$p(B) =$$

$$p(G) =$$

Goal: $p(B|w)$

$$= \frac{p(w|B)\,p(B)}{p(w|B)\,p(B) + p(w|G)\,p(G)} \approx 0.98$$

$$p(w|B) = \frac{n_B(w)}{|B|} = \frac{250}{2000}$$

$$p(w|G) = \frac{n_G(w)}{|G|} = \frac{5}{1000}$$

$$p(B) = \frac{|B|}{|G|+|B|} = \frac{2000}{3000}$$

$$p(G) = \frac{|G|}{|G|+|B|} = \frac{1000}{3000}$$

Tutorial: Binomial Coefficients

Chengbo Zheng

EP3-26. Prove the identity

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$$

for $1 \leq k < n$. (Use both combinatorial and algebraic proofs)

2 ways of proofs.

$\begin{cases} \text{combinatorial proof.} \\ \text{algebraic poof} \end{cases}$

EP3-26. Prove the identity

$$\binom{n}{k} = \frac{n}{k}\boxed{\binom{n-1}{k-1}}$$

for $1 \leq k < n$. (Use both combinatorial and algebraic proofs)

① Algebraic proof:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n}{k}\,\frac{(n-1)!}{(k-1)!\,(n-k)!}$$

$$\frac{(n-1)!}{(k-1)!\,(n-k)!} = \frac{(n-1)!}{(k-1)!\,((n-1)-(k-1))!}$$

$$= \binom{n-1}{k-1}$$

EP3-26. Prove the identity

$$\binom{n}{k} = \frac{n}{k}\boxed{\binom{n-1}{k-1}}$$

for $1 \le k < n$. (Use both combinatorial and algebraic proofs)

② Combinatorial proof:

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

Consider the counting problem, select one leader and $(k-1)$ workers from n people.

① first way of counting:

step 1: Select k people from n

step 2: Select 1 leader from k people.

$$k \times \binom{n}{k}$$

② Second way of counting:

step 1: Select 1 leader from n people.

step 2: Select $k-1$ workers from $(n-1)$ people.

$$n \times \binom{n-1}{k-1}$$

EP3-27. Prove the identity

$$\binom{n}{k} = \frac{n}{n-k}\binom{n-1}{k}$$

for $0 \leq k < n$. (Use both combinatorial and algebraic proofs)

EP3-35. Prove the identity (use both combinatorial and algebraic proofs)

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$

① Algebraic Proof :: $2\sum_{k=1}^{n} k\binom{n}{k} = n2^{n}$

① Algebraic Proof :

$$2\sum_{k=1}^{n} k\binom{n}{k} = n2^{n}$$

$$\sum_{k=1}^{n}\left(k\binom{n}{k}\right) = \sum_{k=0}^{n}\left(k\binom{n}{k}\right) \qquad \underline{k=0,1,2,\cdots,} \; n$$

① Algebraic Proof :

$$2\sum_{k=1}^{n} k\binom{n}{k} = n2^{n}$$

$$\sum_{k=1}^{n}\left(k\binom{n}{k}\right) = \sum_{k=0}^{n}\left(k\binom{n}{k}\right) \qquad \underline{k = 0, 1, 2, \cdots, n}$$

$$= \sum_{k=0}^{n}\left((n-k)\binom{n}{n-k}\right) \qquad \underline{n-k = n, n-1, \cdots, 1, 0}$$

① Algebraic Proof :

$$2\sum_{k=1}^{n} k\binom{n}{k} = n2^{n}$$

$$\sum_{k=1}^{n}\left(k\binom{n}{k}\right) = \sum_{k=0}^{n}\left(k\binom{n}{k}\right) \qquad \underline{k = 0, 1, 2, \cdots, n}$$

$$= \sum_{k=0}^{n}\left((n-k)\binom{n}{n-k}\right) \qquad \underline{n-k = n, n-1, \cdots, 1, 0}$$

$$= \sum_{k=0}^{n}(n-k)\binom{n}{k}$$

① Algebraic Proof :: $2\sum_{k=1}^{n} k\binom{n}{k} = n2^n$

$$\sum_{k=1}^{n}\left(k\binom{n}{k}\right) = \sum_{k=0}^{n}\left(k\binom{n}{k}\right) \qquad \underline{k = 0, 1, 2, \cdots, n}$$

$$= \sum_{k=0}^{n}\left((n-k)\binom{n}{n-k}\right) \qquad \underline{n-k = n, n-1, \cdots, 1, 0}$$

$$= \sum_{k=0}^{n}(n-k)\binom{n}{k}$$

$$2\sum_{k=1}^{n} k\binom{n}{k} = \sum_{k=0}^{n}\left(k\binom{n}{k} + (n-k)\binom{n}{k}\right)$$

$$= \sum_{k=0}^{n} n\binom{n}{k}$$

$$= n\sum_{k=0}^{n}\binom{n}{k}$$

$$= n2^n$$

Combinatorial Proof

- The counting problem: choose 1 leader and any number of workers from n people
- First way of counting
  - We first select 1 leader+0 workers, 1 leader + 1 worker, and then 1 leader + 2 workers, ⋯, 1 leader + (n-1) workers
  - Every time, select k people and choose 1 leader from them.

$$\sum_{k=1}^{n} k\binom{n}{k}$$

Combinatorial Proof

- The counting problem: choose 1 leader and any number of workers from n people
- Second way of counting
  - We first choose 1 leader
  - Then for the rest (n-1) people, each of them has two choices

$$n2^{n-1}$$

EP3-37. Use the binomial theorem to prove that $2^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}3^{n-k}$

Solution

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}3^{n-k} = (3-1)^n = 2^n$$

QB5-4. Arrange the following running times in order of increasing asymptotic complexity. Just give the answer; no explanation is needed.

$$n^3, \sqrt{2n}, n+10, \log(n^4), 20^n, 2^n, n^2\log n$$

Note that you must write function $f(n)$ before function $g(n)$ if $f(n) = O(g(n))$.

Solution

$$\log(n^4), \sqrt{2n}, n+10, n^2\log n, n^3, 2^n, 20^n$$