條件機率 Conditional Probability & Bayes' Rule

new information

- **Q**: Should the following probabilities be different?
  - Event = a pitcher wins at least 16 games in a season (same event)
    - P=?? in the beginning of the season
    - P=?? in the middle of the season
    - P=?? close to the end of the season

    changed
  - Event = rain tomorrow (same event) **Q**: how to evaluate them?
    - P=?? if no information about where you are staying
    - P=?? if you are staying in a desert
    - P=?? if a typhoon will hit the place you stay tomorrow

    changed
- **Q**: What causes the differences?
  - For an event, *new information* (i.e., some other event has occurred) could change its probability
  - We call the altered probability a conditional probability

Recall subjective probability

- Mathematical Definition: If A and B are two events in a sample space Ω and $P(A)>0$, then

event occurred   event of interest

reasonable to define for A with P(A)=0?

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

is called the *conditional probability* of B given A.

conditioned on continuous random variables (future lecture)

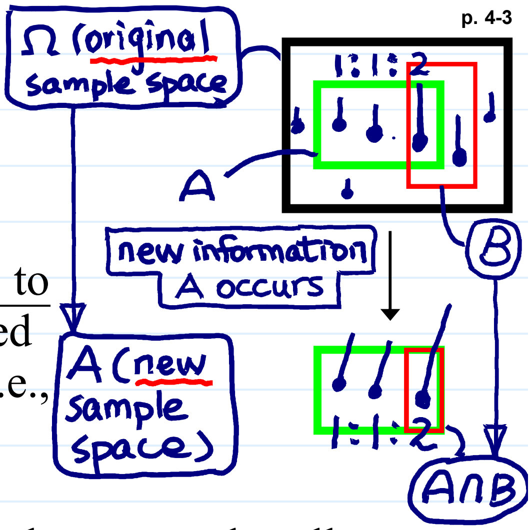
- In the classical probability, (symmetric outcomes in Ω)

  $P(A) = \#A/\#\Omega$ and $P(A \cap B)=\#(A\cap B)/\#\Omega$ ~~$A\cap B$~~

  $$\Rightarrow P(B|A) = \frac{\#(A\cap B)/\cancel{\#\Omega}}{\#A/\cancel{\#\Omega}} = \frac{\#(A\cap B)}{\#A} \quad \text{cf.} \quad P(B) = \frac{\#B}{\#\Omega}$$

  $P(B) \xrightarrow{\text{update}} P(B|A)$

  - Example: A family is known to have 2 children, *at least one of whom is a girl* (A). **Q**: Probability that the other is a boy=?? ($P(B|A)$)
    - $\Omega = \{bb, bg, gb, gg\}$ symmetric outcomes
    - $A=\{bg, gb, gg\}$ and $B=\{bb, bg, gb\}$ (at least one boy) ← $P(B)=3/4$
    - $P(B|A) = \#(A\cap B)/\#A = 2/3$ ← update
  - Note: $\#\Omega$ is reduced to $\#A$.



- In effect by conditioning,
  - we are restricting the sample space from Ω to A, i.e.,

$$\Omega \rightarrow A,$$

  - and, for an arbitrary event B (in Ω) to occur when A has occurred, we need that both A and B occur together, i.e.,

$$B \rightarrow B \cap A.$$

- The division by $P(A)$ in the definition above rescales all probabilities from the entire sample space Ω to being relative to the new sample space A ($\because$ P(sample space) = 1)

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

- $P(B|A)$ is a probability measure defined on B, $\boxed{\text{but not } A.}$ —(*)

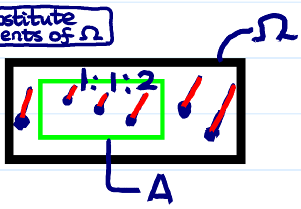
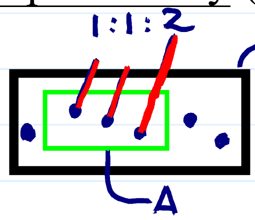
a subset of Ω

It can be regarded as a probability measure defined on 2^Ω

  - $P(\cdot|A)$ satisfies the 3 axioms of probability (exercise)

can substitute any events of Ω

check textbook Sec. 3.5 for details

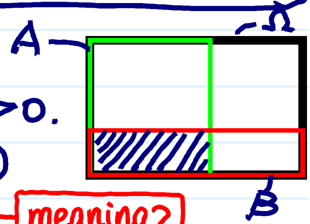
$P(\cdot)$  $P(\cdot|A)$ 

  - Any propositions developed in Chapter 2 for probability measures can be applied on $P(\cdot|A)$. (For example, $P(B^c|A)=1-P(B|A)$.)

cf. $P(B|A)+P(B|A^c) \neq 1$ in general ((*) in LN p. 4-3)

$\because$ $P(B|A)$ & $P(B|A^c)$ might be easier to evaluate

- 3 Useful Formulas for Calculating Probabilities (for 2-events case)



order not matter $= P(B)P(A|B)$ if $P(B)>0$.

1. If $P(A)>0$, then $P(A\cap B) = P(A)\ P(B|A)$. —①

to garantee the cond. prob. well defined

proof. By the definition of cond. prob., $P(B|A) = P(A\cap B)/P(A)$.

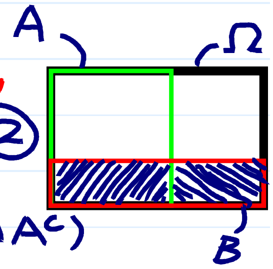
meaning?

2. If $0<P(A)<1$, then

weighed average

$P(B|A)$ $P(B)$ $P(B|A^c)$

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c). \quad \text{—②}$$



proof. $B = B\cap\Omega$
$= B\cap(A\cup A^c)$
$= (B\cap A)\cup(B\cap A^c)$

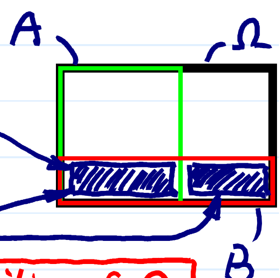
mutually exclusive

$P(B\cap A)$ by ① $P(B\cap A^c)$

3. If $0<P(A)<1$ and $P(B)>0$, then

$P(\cdot) \xrightarrow[\text{B occurs}]{\text{update}} P(\cdot|B)$

③ $$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^c)P(B|A^c)}.$$



proof. $\frac{P(A\cap B) \text{ —by ① → 分子}}{P(B) \text{ —by ② → 分母}}$

Note. In ② & ③, A & A^c : a partition of Ω

➢ Example (Urn Problem) ← prototype of many applications

- The Story. n balls sequentially and randomly chosen, *without replacement*, from an urn containing R red and $N-R$ (total # of balls) white balls ($n \le N$). **Q**: Given that k of the n balls are red ($k \le R$), probability that the 1st ball chosen is red = ??

do not put the drawn balls back into the urn

#A∩B

- Let $A=\{k$ of the n balls are red$\}$ (#A; symmetric outcomes)

  $B=\{$1st ball chosen is red$\}$

#Ω ← symmetric outcomes

ball 1 (r), ball R (r), ball R+1 (w), ball N (w); 1st draw, ..., nth draw

~~1st draw~~ 2nd draw ... nth draw; k-1 positions for red balls

1st draw ... nth draw; k positions for red balls

- Method 1: $P(B|A) = \binom{n-1}{k-1} \Big/ \binom{n}{k} = k/n$ — meaning? — **Q**: Why different #A?

- Method 2: $\dfrac{\#A\cap B}{\#A}$

hyper-geometric distribution

$$P(A) = \left[\binom{R}{k} \times \binom{N-R}{n-k}\right] \Big/ \binom{N}{n}$$

"#A/#Ω" $= \binom{n}{k} \times [(R)_k \times (N-R)_{n-k}] / (N)_n$

$$= \frac{\binom{n}{k} \times \binom{R}{k} \cancel{k!} \times \binom{N-R}{n-k} \cancel{(n-k)!}}{\binom{N}{n} \cancel{n!}}$$

#A: 1st draw, 2nd draw, ..., nth draw; k positions for red balls; ball 1 (r) ... ball R (r): r1th draw ... rkth draw; ball R+1 (w) ... ball N (w): w1th draw ... wn-kth draw

#Ω: symmetric outcomes; ball 1 (r), ball R (r), ball R+1 (w), ball N (w); n positions for n balls

#A: ball 1 (r) ... ball R (r); ball R+1 (w) ... ball N (w); k positions, n-k positions

- $P(A\cap B) = P(B)P(A|B) = \dfrac{R}{N} \times \dfrac{\binom{R-1}{k-1} \times \binom{N-R}{n-k}}{\binom{N-1}{n-1}}$ (by ①) (easier to evaluate) (urn containg R-1 ⓡ, N-R ⓦ, n-1 balls drawn, k-1 ⓡ, n-k ⓦ)

- $P(B|A) = P(A\cap B)/P(A) = k/n$

➢ Example (Sampling Experiments): An urn contains R red balls and $N-R$ white balls. Sample 2 balls from the urn.

\# of balls in the urn

- $A = \{$red on the first draw$\}$

  $B = \{$red on the second draw$\}$

ball 1 (r), ball R (r), ball R+1 (w), ball N (w); 1st draw, 2nd draw; Ω ← symmetric outcomes

(r,w)	(w,w)
(r,r)	(w,r)

A, Ω, B

#Ω = N(N-1)

- Sampling *Without Replacement*:

$$P(A\cap\Omega) = P(A) = \frac{R(N-1)}{N(N-1)} = \frac{R}{N}, \quad P(A\cap B) = \frac{R(R-1)}{N(N-1)} = P(A)P(B|A)$$

$P(A\cap\Omega) = P(A)P(\Omega|A)$; by ①

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}.$$

easier to evaluate

Similarly, $P(B|A^c) = \dfrac{R}{N-1}$. (exercise) ⇒ Note: $P(B|A)+P(B|A^c) = \frac{R-1}{N-1} + \frac{R}{N-1} \neq 1$ in general

meaning

$P(B|A) \xleftarrow{\text{update}}$ $P(B|A^c) \leftarrow$ $P(B) = P(A)P(B|A) + P(A^c)P(B|A^c)$ (by ②)

check ② in LNp.4-4: $P(B|A) < P(B) < P(B|A^c)$

$$= \frac{R}{N}\cdot\frac{R-1}{N-1} + \frac{N-R}{N}\cdot\frac{R}{N-1}$$

$$= \frac{R^2 - R + NR - R^2}{N(N-1)} = \frac{R(N-1)}{N(N-1)} = \frac{R}{N}. = P(A)$$

fair

($\boxed{r/w}$, $\boxed{r}$)

1st draw, 2nd draw

$$P(B|A) = P(A|B) \overset{\text{by ③}}{=} \frac{P(A\cap B)}{P(B)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}.$$

- Notes:

  also, same for the 3rd, 4th, ⋯ draw

  - The probabilities are proportional to # of red balls left

9/12

  - $P(A|B) = P(B|A) \Rightarrow$ Symmetry.

    $P(A\cap B)/P(B)$ " $= P(A\cap B)/P(A)$

- Sampling *With Replacement*: put the ball back into the urn after draw

  cf. without replacement

$$P(A) = \frac{\boxed{R}\cdot\boxed{N}}{\boxed{N}\cdot\boxed{N}} = \frac{R}{N}, \quad P(B) = \frac{\boxed{N}\cdot\boxed{R}}{\boxed{N}\cdot\boxed{N}} = \frac{R}{N},$$

symmetric outcomes

$$P(A\cap B) = \frac{\boxed{R}\cdot\boxed{R}}{\boxed{N}\cdot\boxed{N}} = \frac{R^2}{N^2},$$

independent (LNp.4-19)

Ω: (ball 1 (r), ⋮, ball R (r), ball R+1 (w), ⋮, ball N (w) [1st draw] , ball 1 (r), ⋮, ball R (r), ball R+1 (w), ⋮, ball N (w) [2nd draw])

cf. Ω in LNp.4-6

$\#\Omega = N^2$

$P(B|A) = P(B) = P(B|A^c)$

$$P(B|A) \overset{\text{update}}{=} \frac{R^2/N^2}{R/N} = \frac{R}{N} = P(B).$$