# 條件机率 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 $\Omega$ and $P(A)>0$, then

  event occurred / event of interest

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

  original probability space $(\Omega, \mathcal{F}, P)$

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

  is called the *conditional probability* of $B$ given $A$.
  - In the classical probability, (symmetric outcomes in $\Omega$)

    conditioned on continuous random variables (future lecture)

    $$P(A) = \#A/\#\Omega \quad \text{and} \quad P(A \cap B) = \#(A \cap B)/\#\Omega$$

    $$\Rightarrow P(B|A) = \frac{\#(A \cap B)/\#\Omega}{\#A/\#\Omega} = \frac{\#(A \cap B)}{\#A}$$

    cf. $P(B) = \frac{\#B}{\#\Omega}$

    P(B) update → P(B|A)
  - Example: A family is known to have 2 children, *at least one of whom is a girl*. Q: Probability that the other is a boy=??

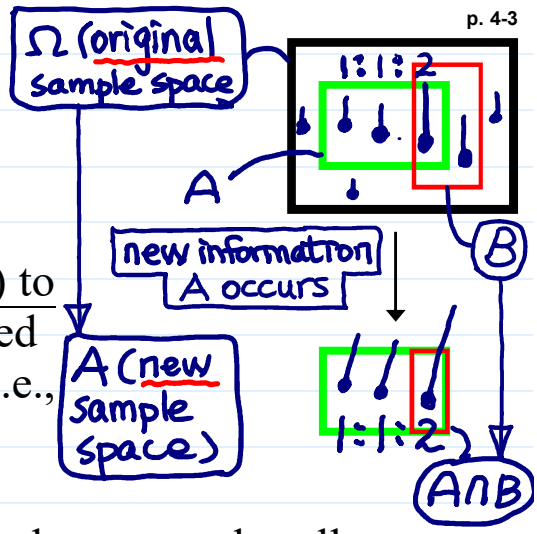
    A
    - $\Omega = \{bb, bg, gb, gg\}$ (symmetric outcomes; at least one boy)
    - $A=\{bg, gb, gg\}$ and $B=\{bb, bg, gb\}$ ← $P(B) = 3/4$
    - $P(B|A) = \#(A \cap B)/\#A = 2/3$ ← update

    NO → 1/2 (intuition?)

    P(B|A)

    Change A to be 1st is a girl
  - Note: $\#\Omega$ is reduced to $\#A$.

- In effect by conditioning,
  - we are restricting the sample space from $\Omega$ to $A$, i.e.,
    $$\Omega \to A,$$
  - and, for an arbitrary event $B$ (in $\Omega$) to occur when $A$ has occurred, we need that both $A$ and $B$ occur together, i.e.,
    $$B \to B \cap A.$$



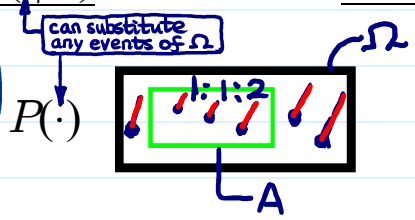

- The division by $P(A)$ in the definition above rescales all probabilities from the entire sample space $\Omega$ to being relative to the new sample space $A$ ($\because P(\text{sample space})=1$)

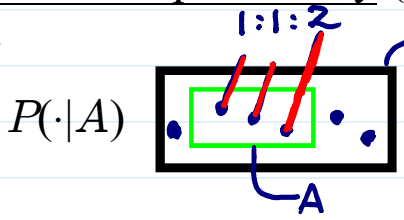
$\frac{P(B\cap A)}{P(A)} =$

- $P(B|A)$ is a probability measure defined on $B$, but not $A$. (*)
  - $P(\cdot|A)$ satisfies the 3 axioms of probability (exercise)

It can be regarded as a probability measure defined on $2^{\Omega}$

can substitute any events of Ω




check textbook sec. 3.5 for details



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

∵ P(B|A) & P(B|A^c) might be easier to evaluate

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

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

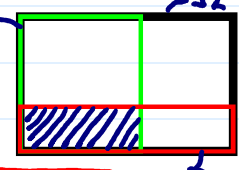
1. If $P(A)>0$, then $P(A\cap B) = P(A)\,P(B|A)$. (1)

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

to guarantee the cond. prob. well defined

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

meaning?

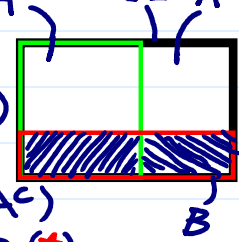



2. If $0<P(A)<1$, then (2)

weighed average

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

$\frac{P(B|A)}{P(B|A^c)}$ $\frac{P(B)\,P(B|A^c)}{P(B|A)}$




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

mutually exclusive

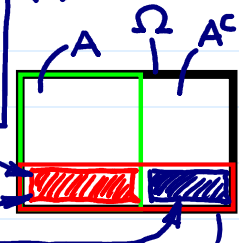
$P(B\cap A)$ (by ①) $P(B\cap A^c)$

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

$P(\cdot)$ update B occurs $P(\cdot|B)$

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

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

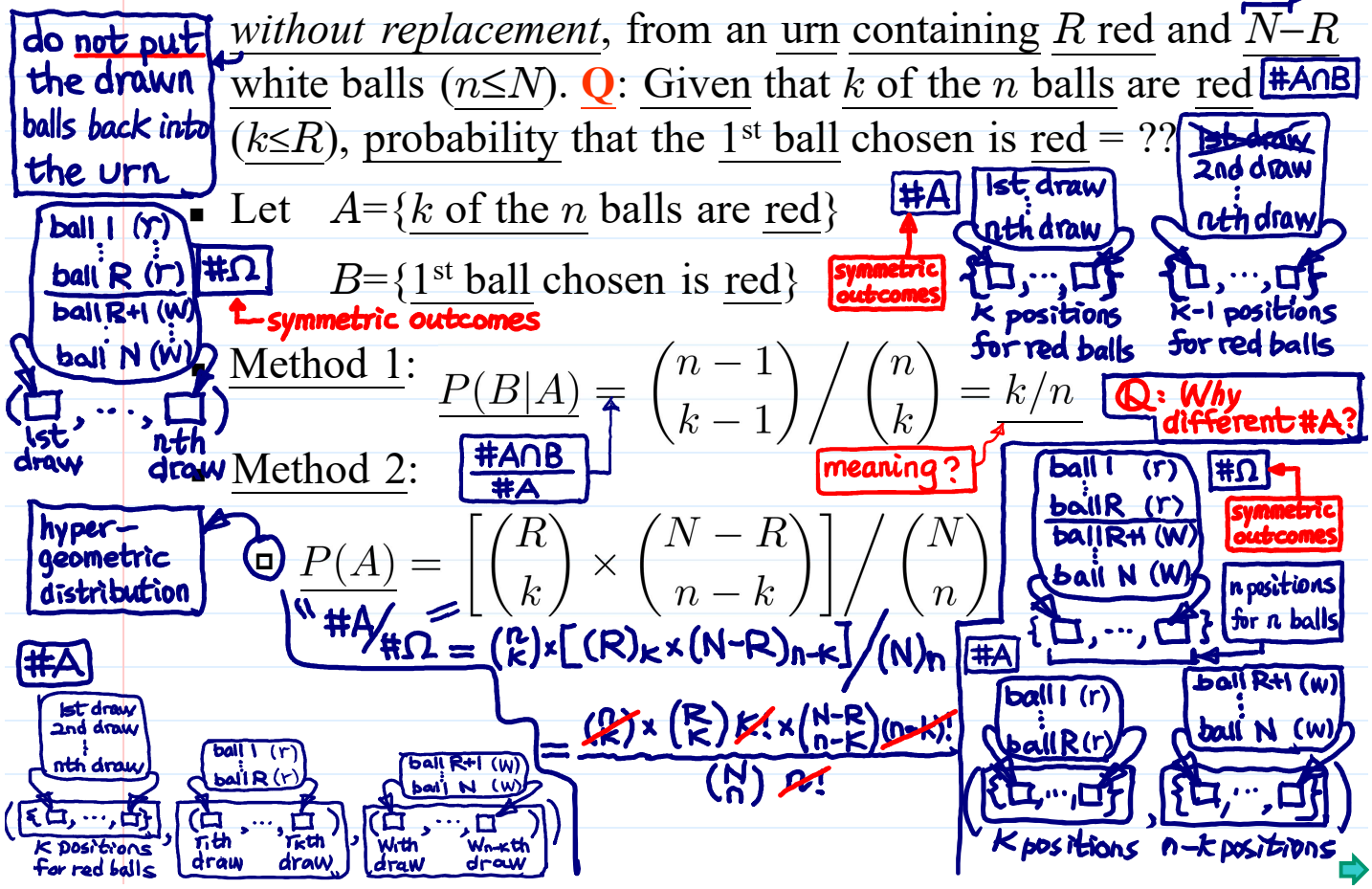



proof, $\frac{P(A\cap B)}{P(B)}$ — By ① → 分子; By ② → 分母

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 $N-R$ 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 = ??




- Method 1: $P(B|A) = \binom{n-1}{k-1} \Big/ \binom{n}{k} = k/n$
- Method 2:
  - $P(A) = \left[\binom{R}{k} \times \binom{N-R}{n-k}\right] \Big/ \binom{N}{n}$



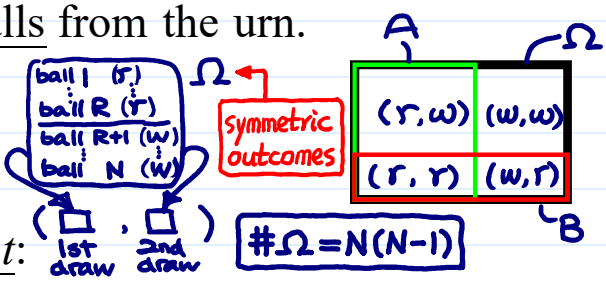
- $P(A \cap B) = P(B)P(A|B) = \dfrac{R}{N} \times \dfrac{\binom{R-1}{k-1} \times \binom{N-R}{n-k}}{\binom{N-1}{n-1}}$ (by ①; easier to evaluate)
- $P(B|A) = P(A \cap B)/P(A) = k/n$

urn containing R-1 (r), N-R (w), n-1 balls drawn, k-1 (r), n-k (w)

## 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}




- Sampling *Without Replacement*:

$P(A \cap \Omega) = P(A) = P(A)P(\Omega|A)$

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

easier to evaluate

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

meaning

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

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

update $P(B|A)$, $P(B|A^c)$ (by ②)

check ② in LNp.4-4 p. 4-7

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

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

(r/w, r): 1st draw, 2nd draw

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

fair

by ③

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

also, same for the 3rd, 4th, ··· draw (exercise)

- Notes:
  - The probabilities are proportional to # of red balls left
  - $P(A|B) = P(B|A) \Rightarrow$ Symmetry.

9/17

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

- Sampling *With Replacement*:

put the ball back into the urn after draw

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

cf. without replacement

symmetric outcomes

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

independent (LNp.4-19)

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

update

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

Ω: ball 1 (r), ball R (r), ball R+1 (w), ball N (w)

cf. Ω in LNp.4-6

1st draw, 2nd draw

$\#\Omega = N^2$