生存分析

2020春季本科课程
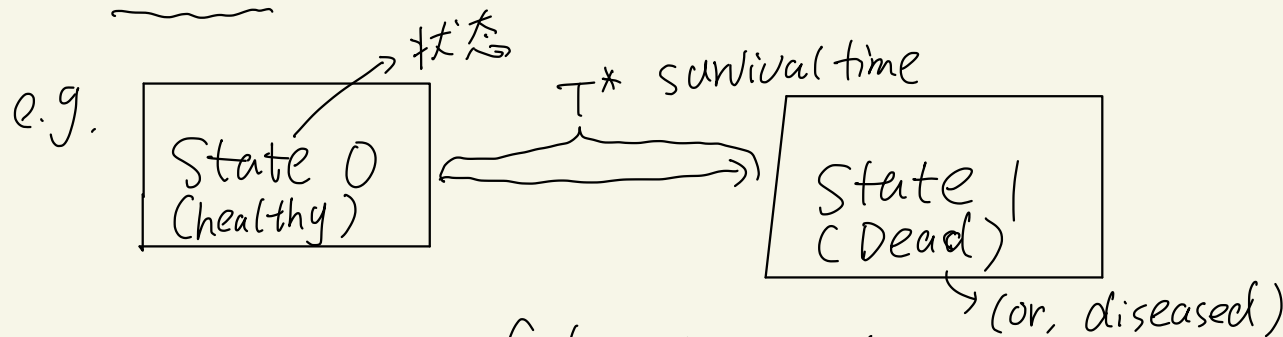
严颖

Survival Analysis

Chp 1. Introduction

§1.1 Definition
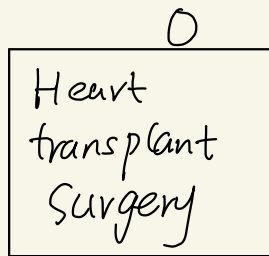
Def. <u>Survival Analysis</u> is a collection of statistical procedure for which the outcome of interest is <u>time until an event</u> <u>occurs.</u>

e.g.

T^* survival time

State 0 (healthy) → (状态)

State 1 (Dead) (or, diseased)

T^*: survival time (failure time; lifetime): <u>from beginning of followup</u> of an individual until an event occurs.

e.g.

医学：


```
    0
┌──────────────┐              T*        ┌──────────────┐ 1
│  Heart       │ ─────────────────────> │              │
│  transplant  │                        │  Dead        │
│  Surgery     │                        │              │
└──────────────┘                        └──────────────┘
  心脏移植手术
```
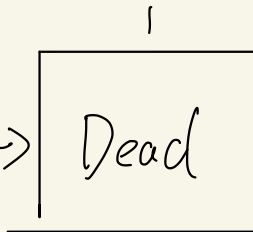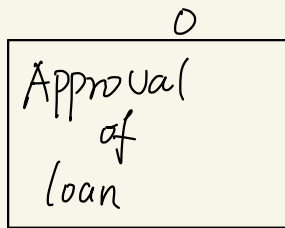
银行：

```
          0
┌──────────────┐              T*        ┌──────────────┐ 1
│  Approval    │ ─────────────────────> │              │
│     of       │                        │ Bankcruptcy  │
│  loan        │                        │              │
└──────────────┘                        └──────────────┘
                                           破产
```
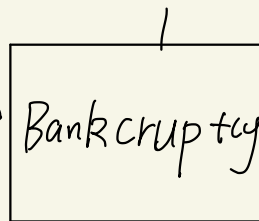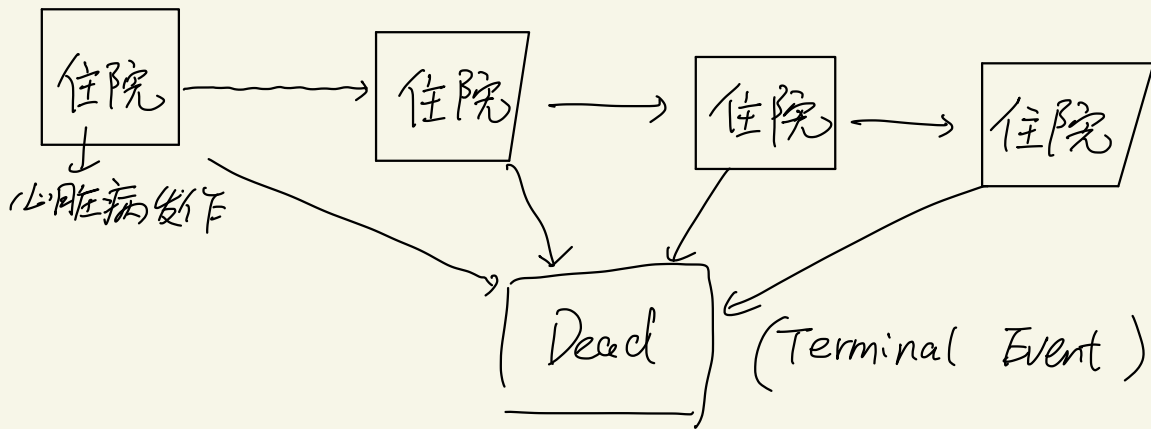
Def. Event History Analysis : We are interested in multiple events
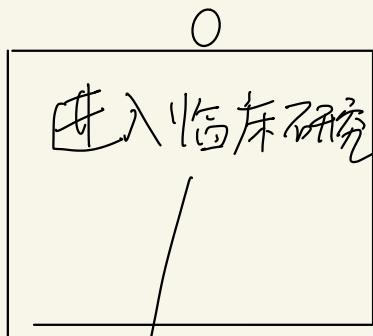事件历史分析

① Recurrent event analysis. (复发事件)

e.g.

住院 → 住院 → 住院 → 住院

心脏病发作

Dead (Terminal Event)

② Competing risk Analysis (竞争风险)

e.g.

```
              1
          ┌──────────────┐
          │ Dead because │
          │ of liver     │      (Cause 1)
          │ failure      │
          └──────────────┘
    0
┌──────────┐          2
│          │     ┌──────────────┐
│  In ICU  │────▶│ Heart        │      (Cause 2)
│          │     │ disease      │
└──────────┘     └──────────────┘

              3
          ┌──────────────┐
          │ Infection    │      (Cause 3)
          └──────────────┘
```
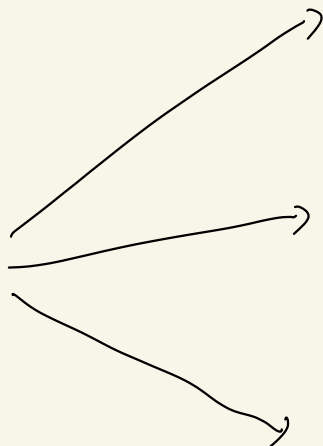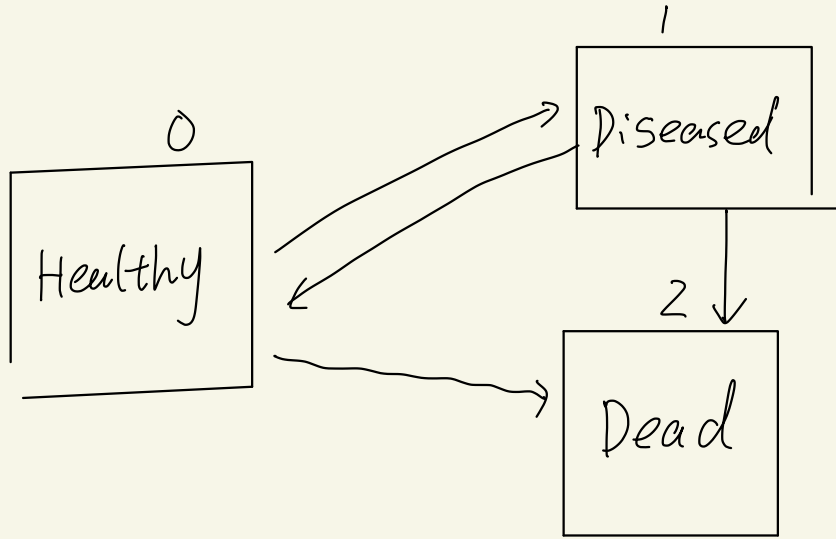
临床研究

e.g.

进入临床研究 O

分配 { 新药
      安慰剂
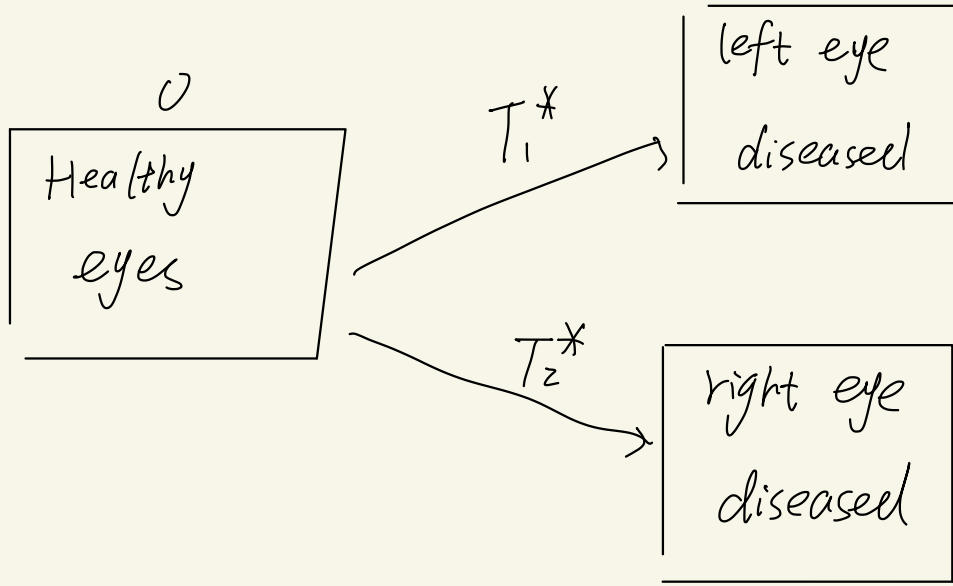
治愈 1

感染 2
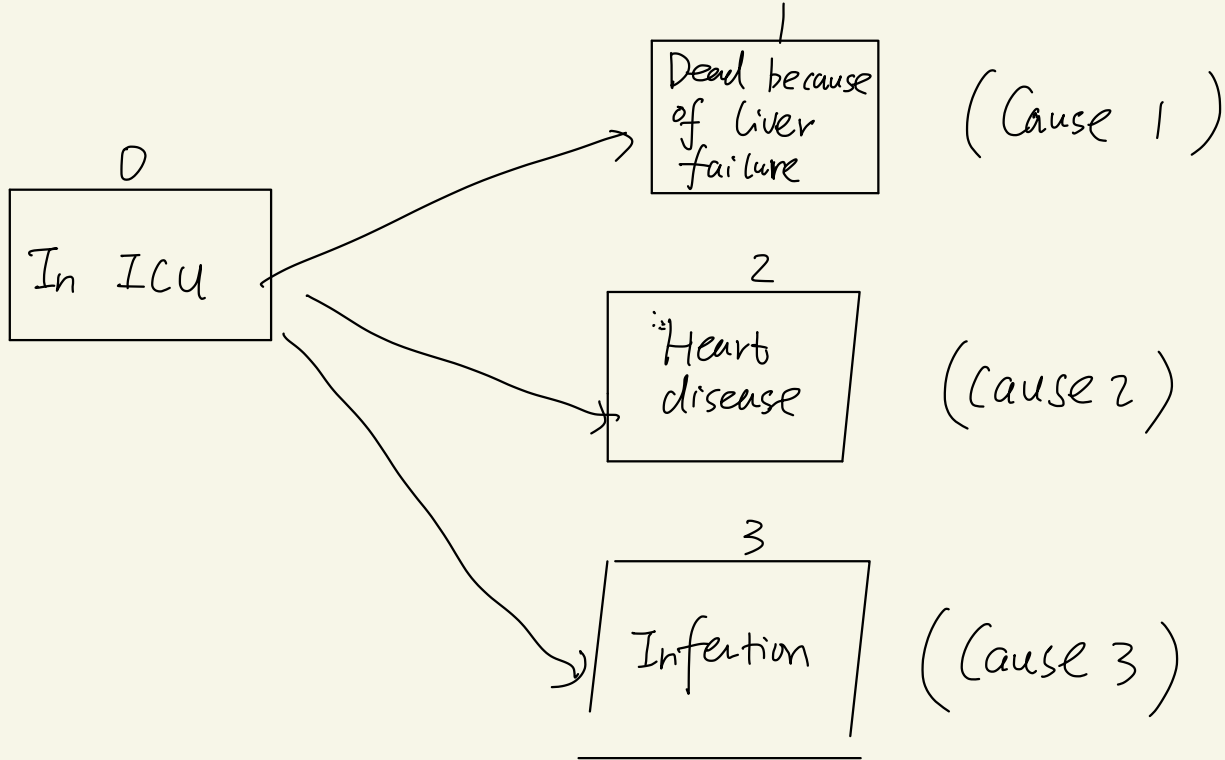
恶化/死亡 3

③ Multistate model （多状态模型）

e.g.

illness-death model

Note: Survival, recurrent event, competing risk are
special cases of multistate models.

④ Multivariate Survival Analysis (多元生存分析)

e.g.

$$O$$

| Healthy eyes |

T_1^* →

| left eye diseased |

T_2^* →

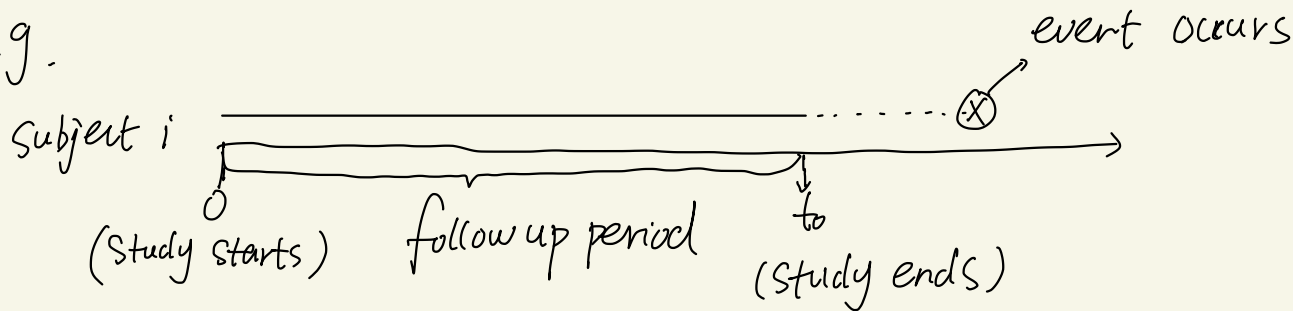| right eye diseased |

$$T^* = \begin{pmatrix} T_1^* \\ T_2^* \end{pmatrix}$$

Note: it's different from Competing risk.

§1.2 Censoring (删失) and Truncation (截断)

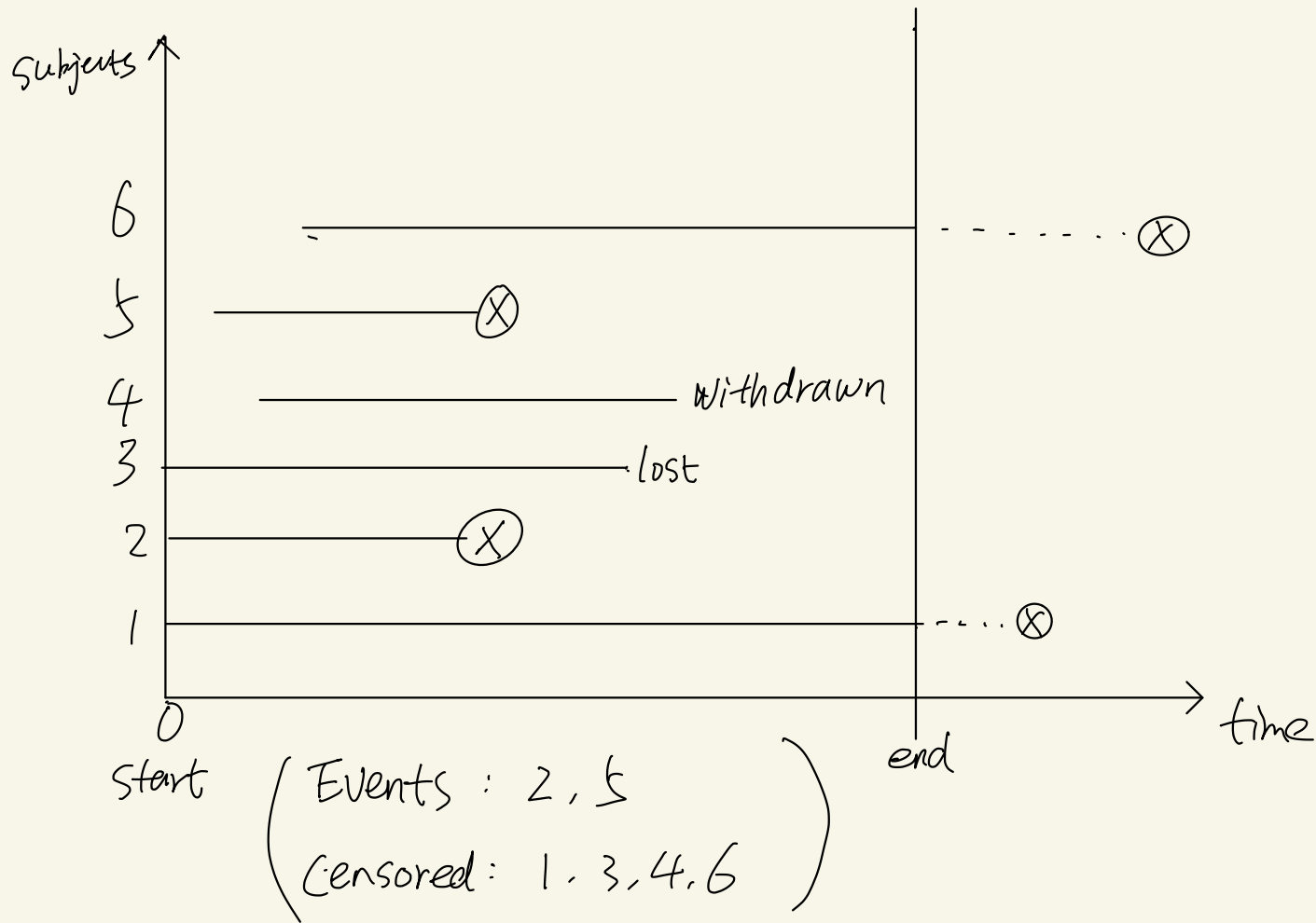Def. Censoring occurs when we have some information of survival time, but we don't know the survival time exactly.

e.g.

Subject i

(study starts) — follow up period — t_0 (study ends) — event occurs

we know $T_i^* > t_0$ only.

missing data

Subjects (y-axis), time (x-axis)

- 6: withdrawn line extending past end, ⊗
- 5: ⊗ (event)
- 4: withdrawn
- 3: lost
- 2: X (event)
- 1: ⊗

0 Start — end

$$\left(\begin{array}{l} \text{Events} : 2, 5 \\ \text{Censored} : 1, 3, 4, 6 \end{array} \right)$$

Def.   Right-censored (右删失):   Survival time > observed time

left-censored (左删失) :   <

e.g.

$$\Downarrow$$

observed time

survival time
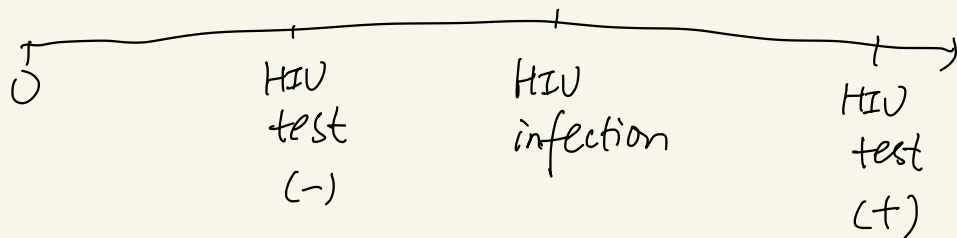
HIV infection

HIV test is positive (t)
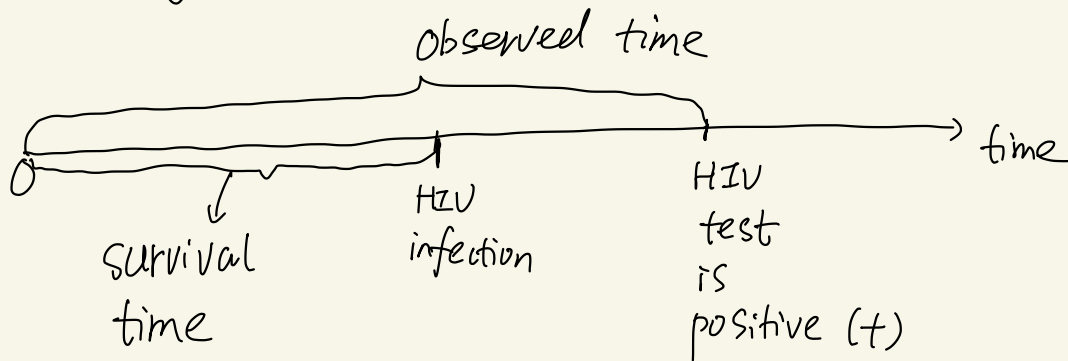
time

椭圆发检测

Interval-censored (区间删失): Survival time is unknown, but we know it's <u>within a time interval</u>.
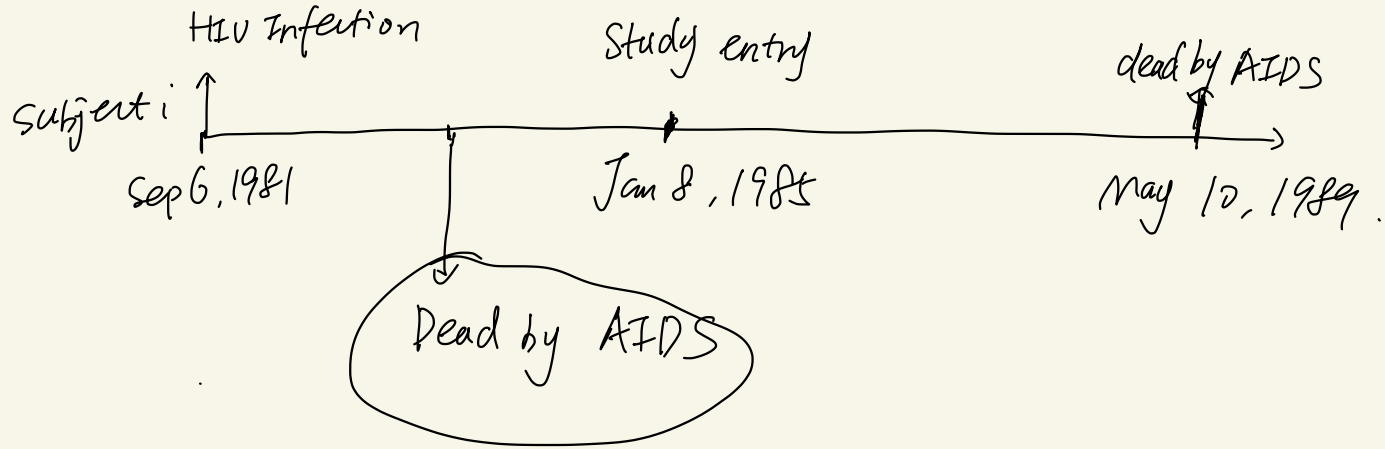
e.g.

Note: left & right censoring are special cases of interval censoring.

Left truncation (左截断):

e.g. First test of HIV: 1984年

Subject i

HIV Infection — Sep 6, 1981

Study entry — Jan 8, 1985

dead by AIDS — May 10, 1989.

Dead by AIDS

<u>Left trunction</u>: Subjects that become HIV infected <u>and</u> have a short time to Dead are likely to be missed by the study. Those who are <u>missed</u> are called <u>left - truncated</u>.

任果: [biased sample]
(有偏样本)

survivor bias
(幸存者偏差)

population

Sample

随机抽样 (无偏样本)