March 20

"All models are wrong but some are useful!"
– George Box"

Approaches (model)

Parametric

↓

we assume that the mathematical (form) model is

Non parametric

↓

The model is not known.

known.

Eg: If we assume that a certain data follows $N(\mu, \sigma^2)$, μ & σ^2 may be unknown.

Then the related ~~this~~ approach is parametric.

Approaches (goals)

Estimation

Hypothesis Testing

we try to approximate unknown pop. parameter

Test some belief.

Estimation (Parametric)

let $X_1, X_2, \ldots, X_n$ are IID observations from a population with known CDF (PDF/PMF) form but some unknown parameters.

Further let an unknown parameter is β and we want to "estimate" (approximate) β.

In statistics.

$\boxed{\text{approximation of unknown parameter based on sample}}$

estimation

Let $T(X_1, \ldots, X_n)$ is an approximation of β.

$T(X_1, \ldots, X_n)$ β

$\underbrace{T(x_1, \ldots)}$

statistic

parameter

we say that

(i) $T(x_1, \ldots, x_n)$ estimates β.

(ii) $T(x_1, \ldots, x_n)$ is an estimator of β

"Estimator" is a statistic and targets (estimates) one parameter..

"Estimate" is the value for a given

of estimator for θ

sample.

Eg: let $X_1, X_2, \ldots, X_n$ are IID observations from Poisson(λ) distribution, λ is unknown(parameter).

An estimator for λ is

$$\bar{X}_n = \frac{X_1 + \ldots + X_n}{n}$$

$\downarrow$

will change sample to sample

1st sample

1, 2, 3, 4, 5

$$\bar{X}_{n(1)} = \frac{1+2+3+4+5}{5}$$

2nd sample

5, 3, 2, 4, 6

$$\bar{X}_{n(2)} = \frac{5+3+2+4+6}{5}$$

Eg: Tossing of a coin with unknown prob. of getting head, denote it by p.

parameter - pop. mean - $\overset{E(X)}{\ }$ p

=

$X_1, \ldots, X_{100}$

$$\bar{X}_{100} = \frac{X_1 + \cdots + X_{100}}{100}$$

$\downarrow$

Estimator

value estimate $\frac{60}{100}$

$\bar{X}_{100}(1)$

$\vdots$ value estimate

$\bar{X}_{100}$

Eg: Value of π?

Properties of a good estimator:

① Unbiasedness:

An estimator $T(X_1, \dots, X_n)$ is said to be unbiased for parameter β if

$$\mathbb{E}\left(T(X_1, \dots, X_n)\right) = \beta.$$

Eg: Coin toss example.

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

$$\mathbb{E}(\bar{X}_n) = p$$

$\bar{X}_n$ is unbiased for p.

$$\bar{} \quad X_1 + \dots + X_n + 1$$

$$T_n = \frac{\quad}{n}$$

$$\mathbb{E}(T_n) = \mathbb{E}\left(\bar{X}_n + \frac{1}{n}\right)$$

$$= p + \frac{1}{n}$$

$$\mathbb{E}(T_n) \neq p$$

---

(2) Consistency:

An estimator $T(X_1, \ldots, X_n)$ is said to be consistent for β, if

$$T(X_1, \ldots, X_n) \xrightarrow{p} \beta.$$