Lecture — 15

Parameter Estimation

E.g. About 54% of Indians have developed antibodies for Covid-19.

E.g. Election in W.B.
opinion poll
BJP:
TMC:
Congress: 0 seats

E.g. Height of an average adult Indian male is 165 cms.

E.g: Let's say that you have a biased coin. Assume that you don't know the probability of getting a head. You toss this coin 100 times. You get 70 heads. You will assume that the probability of getting a head is 0.7

Binomial distribution

(n, p)

$n = 100$, $p = ?$

Maximum likelihood estimators

There is some quantity/parameter that we want to estimate.

E.g. Suppose that n independent trials, each of which is a success with probability p are performed. What is the maximum likelihood estimator of p?

$$x_i = \begin{cases} 1 & \text{is a success} \\ 0 & \text{is a failure} \end{cases}$$

$$P(x_i = 1) = p$$

$$P(x_i = 0) = 1 - p$$

$$P(x_i = x) = p^x (1-p)^{1-x},$$

where $x = 0$ or 1

Assumption is that the trials are independent

likelihood, joint probability mass function

$$f(x_1, x_2, \dots, x_n | p)$$

NOT given p

f is a function of x_i and p.

$$= P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | p\}$$

$$= P\{X_1 = x_1\} P\{X_2 = x_2\} \cdots P\{X_n = x_n\}$$

$$= p^{x_1}(1-p)^{1-x_1} p^{x_2}(1-p)^{1-x_2} \cdots p^{x_n}(1-p)^{1-x_n}$$

$$= p^{\sum x_i}(1-p)^{\sum(1-x_i)}$$

$\rightarrow$ maximize this over p

$x_i = 0, 1$, $i = 1, \dots, n$

Find the value of p

that maximizes this

likelihood. Take log first

Taking logs

$$\log f(x_1, x_2, \dots, x_n | p) =$$

$$\sum x_i \log p + \sum (1 - x_i) \log(1-p)$$

$$\frac{d}{dp}(\,) = 0$$

$$\frac{1}{p} \sum x_i \quad \frac{-1}{1-p} \sum (1 - x_i) = 0$$

$$\frac{\sum x_i}{p} = \frac{\sum 1 - \sum x_i}{1-p} = \frac{n - \sum x_i}{1-p}$$

$$(1-p)\sum x_i = p\left(n-\sum x_i\right)$$

$$\sum x_i \cancel{-p\cdot\sum x_i} = np \cancel{-p\sum x_i}$$

$$\sum x_i = np$$

$$\boxed{p = \frac{\sum x_i}{n}}$$

$$p = \frac{\text{no. of successes}}{\text{no. of trials}}$$

this is an estimate for the probability of success e.g.: Two proofreaders who go through a book.

Proofreader 1 finds 300 errors in the book.

Proofreader 2 finds 400 errors in the book.

There 150 errors that both of them find.

Estimate the total no. of errors in the book.

$= 550$?

$300 + 400 - 150$

550 is the no. of errors that have been found by either of the proof readers.

$$|E_1 \cup E_2| = \boxed{550 \; 800}$$

n_1, n_2, $n_{1,2}$, Estimate

found by both

total no. of errors

Results of the proofreaders are independent: assumption

p_1 = probability that an error is found by proofreader 1

p_2 = prob. that an error is found by proofreader 2

$$p_1 = \frac{n_{1,2}}{n_2}$$

n_2 errors that were found by proofreader 2

$$p_1 = \frac{n_1}{N}$$

$$\frac{n_{1,2}}{n_2} = \frac{n_1}{N}$$

$$\Rightarrow \boxed{N = \frac{n_1 \quad n_2}{n_{1,2}}}$$

$$= \frac{\cancel{300}^{2} \times 400}{\cancel{150}}$$

$$= 800$$

$$p_2 = \frac{n_{1,2}}{n_1}$$

$$= \frac{n_2}{N}$$