ch 5 Estimation

Probability Assume model $Y \sim N(0,1)$ $\rightarrow$ calculate $P(|Y| \leq 3) = 0.997$

Statistics

is it plausible that $Y \sim N(0,1)$ $\Leftarrow$ observe Y data point

$Y = 4.5$

Data $\underbrace{X_1, \ldots, X_n}_{\text{data}}$ iid from some $\underbrace{\text{cdf}}_{\text{model}}$ F

cases ① **Parametric model** $\Rightarrow$ we know F except for possibly a finite-dimensional **parameter** θ

F is $\exp(\theta)$

F is $N(\mu, \sigma^2)$ $\rightarrow$ $\theta = (\mu, \sigma^2)$

F is Binomial(p) $\theta = p$

② **Nonparametric model** $\Rightarrow$ can only make qualitative assumptions about F

e.g. F has smooth density.

Data $X_1, \ldots, X_n \sim$ iid from F

unknown parameter of interest θ

Def An Estimator (or a statistic) is a function of the data, e.g. $g(X_1, \ldots, X_n)$, that is used to estimate/approximate the unknown parameter θ.

$\hat{\theta}_n = g(X_1, \ldots, X_n)$

$\boxed{\text{r.v.}}$ has a distribution

$\boxed{\text{statistics display their variability across samples.}}$

example $\theta = EX_1$

$\hat{\theta}_n = \frac{1}{n} \sum_1^n X_i$

using data $X_1 = 3$
$X_2 = -4$
$X_3 = 5$
$\vdots$
$X_{10} = -1$

$\Rightarrow \hat{\theta}_{10} = -0.57$

$\boxed{\text{an estimate of } \theta}$

example $X_1, \ldots, X_n \sim$ iid from F with finite variance.

$\theta = EX_i$

$\hat{\theta}_n = \frac{1}{n} \sum_1^n X_i$ (motivated by LLN & CLT)

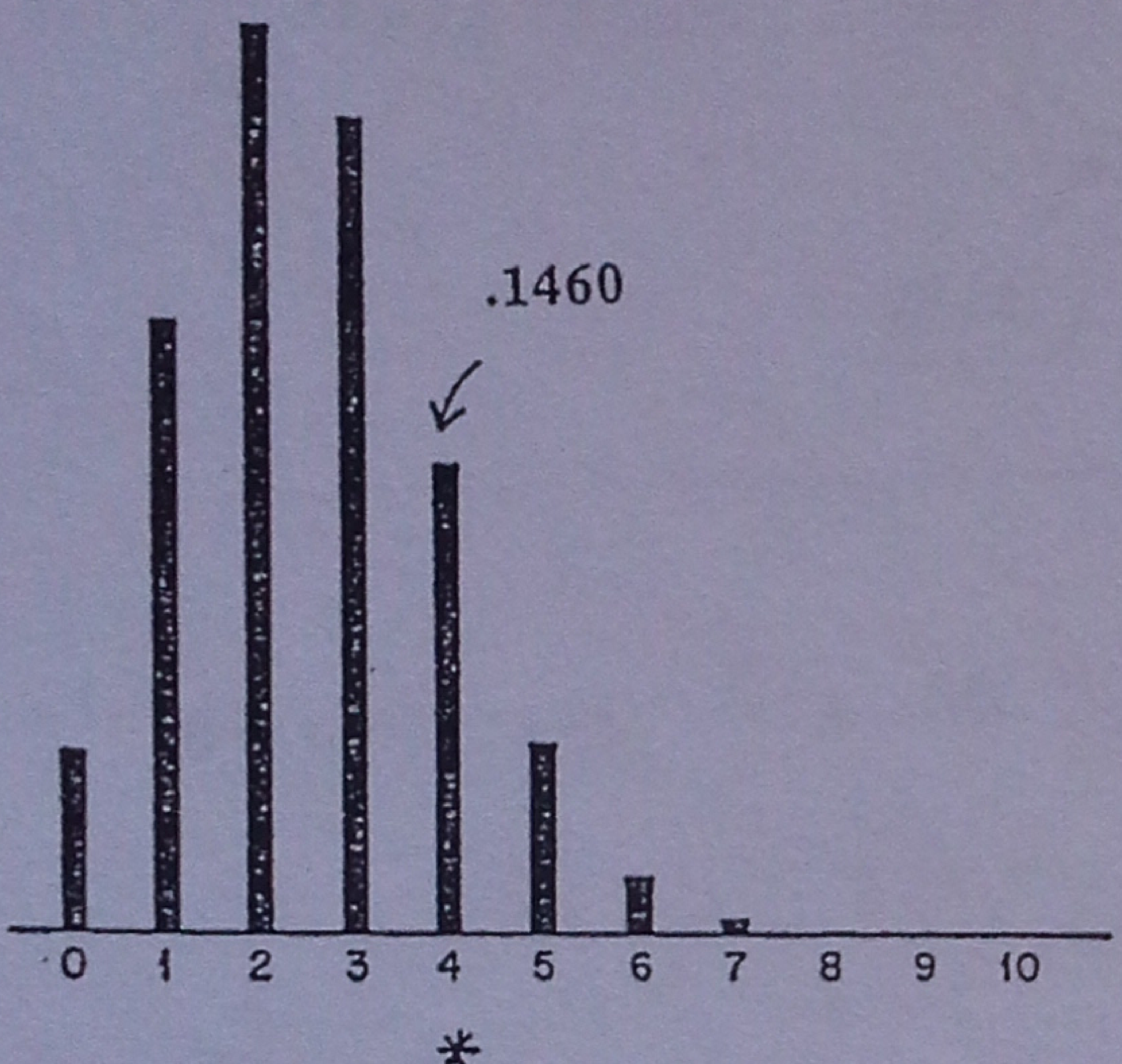
ex. 2 $\boxed{\theta = EX_i^2}$ Define $Y_i = X_i^2$ so look at $Y_1, \ldots, Y_n$

METHOD OF MOMENTS ESTIMATION

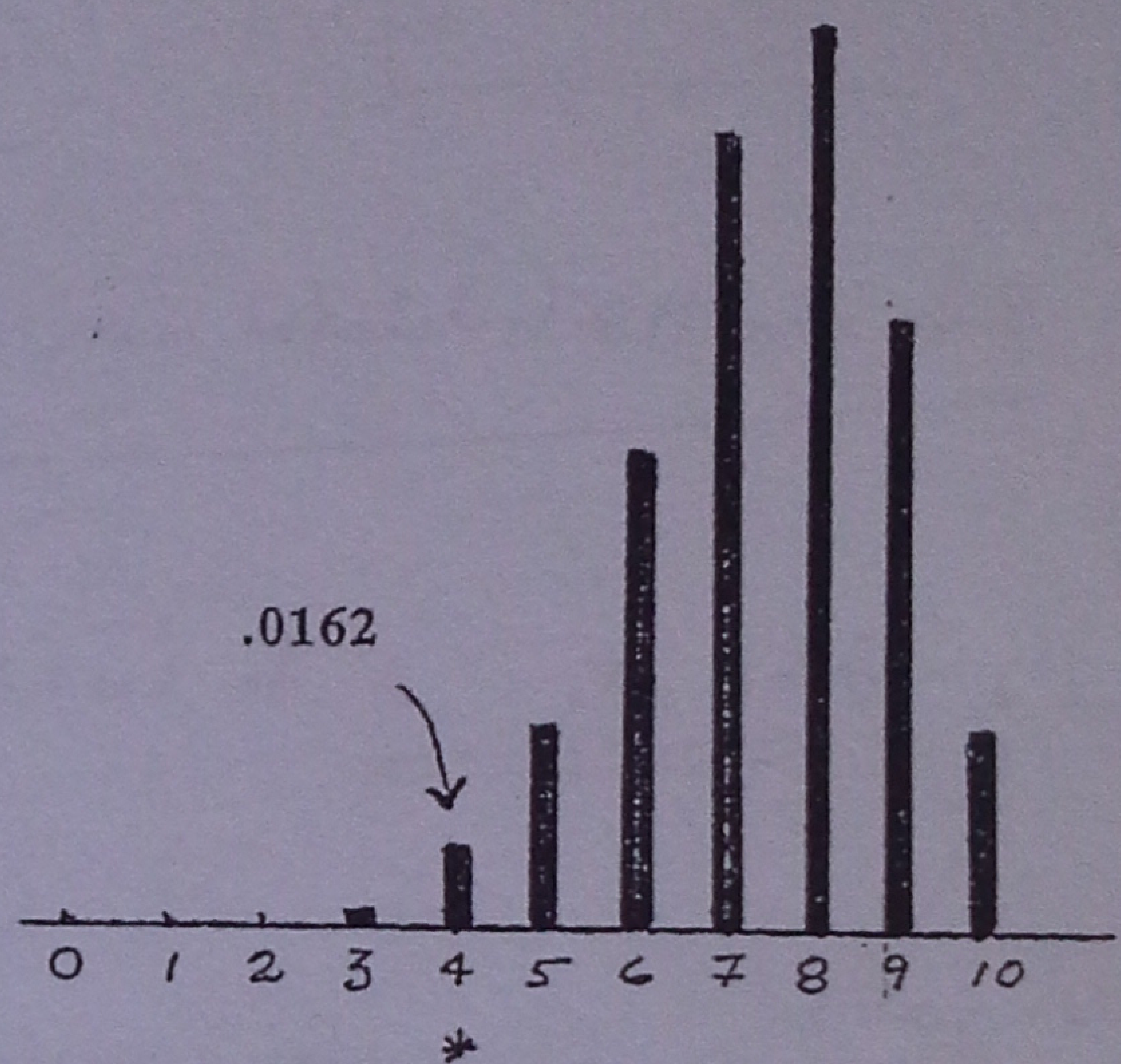
$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n Y_i = \frac{1}{n} \sum_1^n X_i^2$

Binomial Distributions

(θ = Probability of "heads"; $\theta = .25$ or $\theta = .75$;
with probabilities depicted by bar lengths)



$n = 10$; $\theta = .25$



$n = 10$; $\theta = .75$

Data : # of Heads = 4

The maximum likelihood idea: Suppose we know that $\theta = .25$ or $\theta = .75$; but the value of θ is unknown, otherwise. If we observe exactly 4 "heads", and then we are required to choose a value of θ, which choice is reasonable ?

$$P(\#\text{ of Heads} = 4) = \underbrace{\binom{10}{4}\theta^4(1-\theta)^6}_{= L(\theta)}$$ ← Likelihood function

$=$ Prob. of your data w.r.t. data fixed to what you observed & viewed as a function of θ.

Choose θ that maximizes $L(\theta)$

$$\frac{\partial}{\partial\theta}L(\theta) = 0 \Rightarrow \cdots \Rightarrow \theta = \frac{4}{10}$$

MLE) Maximum Likelihood Estimate of θ

In general you form the likelihood of $X_1, \ldots, X_n$ iid from pdf $f_\theta(\cdot)$

$$L(\theta) = \prod_{i=1}^{n} f_\theta(X_i)$$

observed values

either discrete or continuous density.

view it as a function of θ with X_i fixed to observed values.

$n=1 \quad X_1 \sim Bin(n, \theta)$

unknown

$X_1, \ldots, X_{10} \sim$ iid Bernoulli(θ)

$P(X_i = 1) = \theta$

$P(X_i = 0) = (1-\theta)$

observed: 0010011100

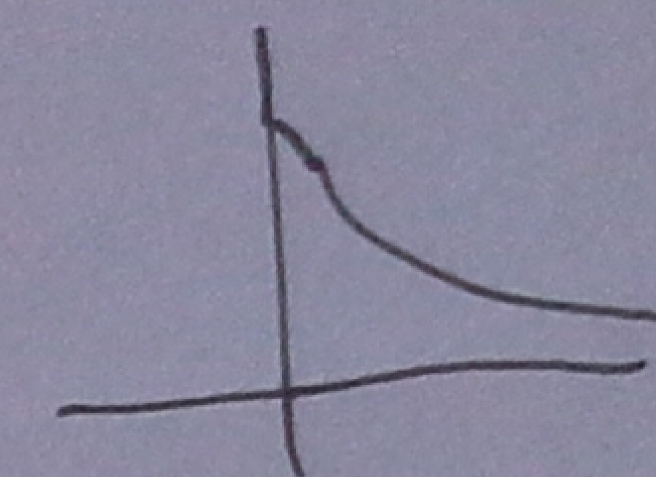
$$L(\theta) = (1-\theta)(1-\theta)\theta(1-\theta)\cdots = \theta^4(1-\theta)^6$$

admits a sufficient statistic $g(X_1, \ldots, X_n) = \sum_1^{10} X_i$

$$\frac{d}{d\theta} L(\theta) = 0 \iff \frac{d}{d\theta} \log L(\theta) = 0$$

$$L(\theta) = \prod_{i=1}^{n} f_\theta(X_i) \qquad \underbrace{\log L(\theta) = \sum_{i=1}^{n} \log f_\theta(X_i)}$$

<u>EX 5.2.2</u> $X_1, \dots, X_n \sim \text{iid } f_\theta(x) = \frac{1}{\theta} \exp\{-x/\theta\}$ for $x \geq 0$ & $\theta \geq 0$



<u>MOM</u> estimator of θ is $\bar{X} = \frac{1}{n} \Sigma X_i$

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \exp\left\{-\frac{X_i}{\theta}\right\} = \frac{1}{\theta^n} \exp\left\{-\frac{\Sigma X_i}{\theta}\right\}$$

$$\log L(\theta) = -n \log \theta - \frac{\Sigma X_i}{\theta}$$

$$\frac{d}{\partial\theta} \log L(\theta) = 0 \Rightarrow \dots \Rightarrow \hat{\theta}_{MLE} = \frac{1}{n} \Sigma X_i$$

Range of X_i does not depend on θ.