STAT 231

October 13, 2016

Roadmap

INTERVAL ESTIMATION

- LIKELIHOOD INTERVALS
- CONFIDENCE INTERVALS using sampling distributions

Objective: Data set: $\{y_1, \dots, y_n\}$ indep

MODEL: $Y_i \sim f(y_i; \theta)$

f = distn function of Y.

__Objective__: To estimate θ = unknown parameter.

Instead of estimating the "most likely" value of θ, we want to find the "reasonable" values θ can take. We want to construct an interval based on our data

$$[a \,,\, b]$$

θ

which will contain θ with a high degree of confidence.

Application: How are Margin of Errors calculated for polling data?

METHOD I: USING THE RELATIVE LIKELIHOOD FUNCTION.

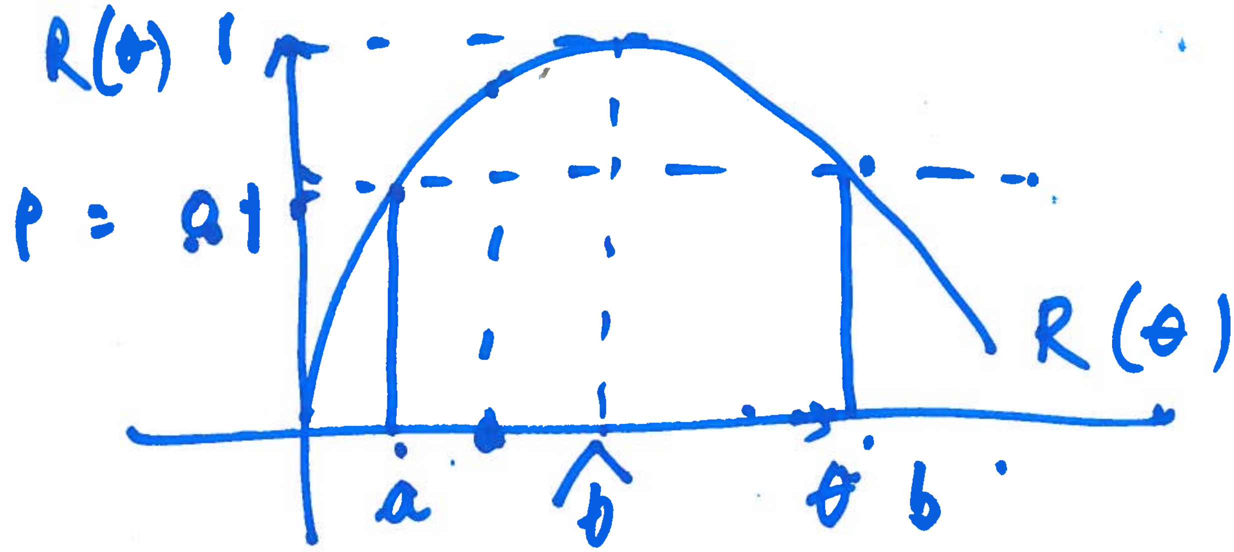
Definition: Take $p \in (0,1)$. A $100p\%$ likelihood interval is the set of all θ: $\{\theta : R(\theta) \geq p\}$

$$R(\theta) = L(\theta) / L(\hat{\theta})$$

$$\underline{\underline{L(\theta)}} \geqslant p \cdot L(\hat{\theta})$$

$p = 0.1$

The value of the likelihood function evaluated at θ is at least $100p\%$ the value of L evaluated at $\hat{\theta}$.



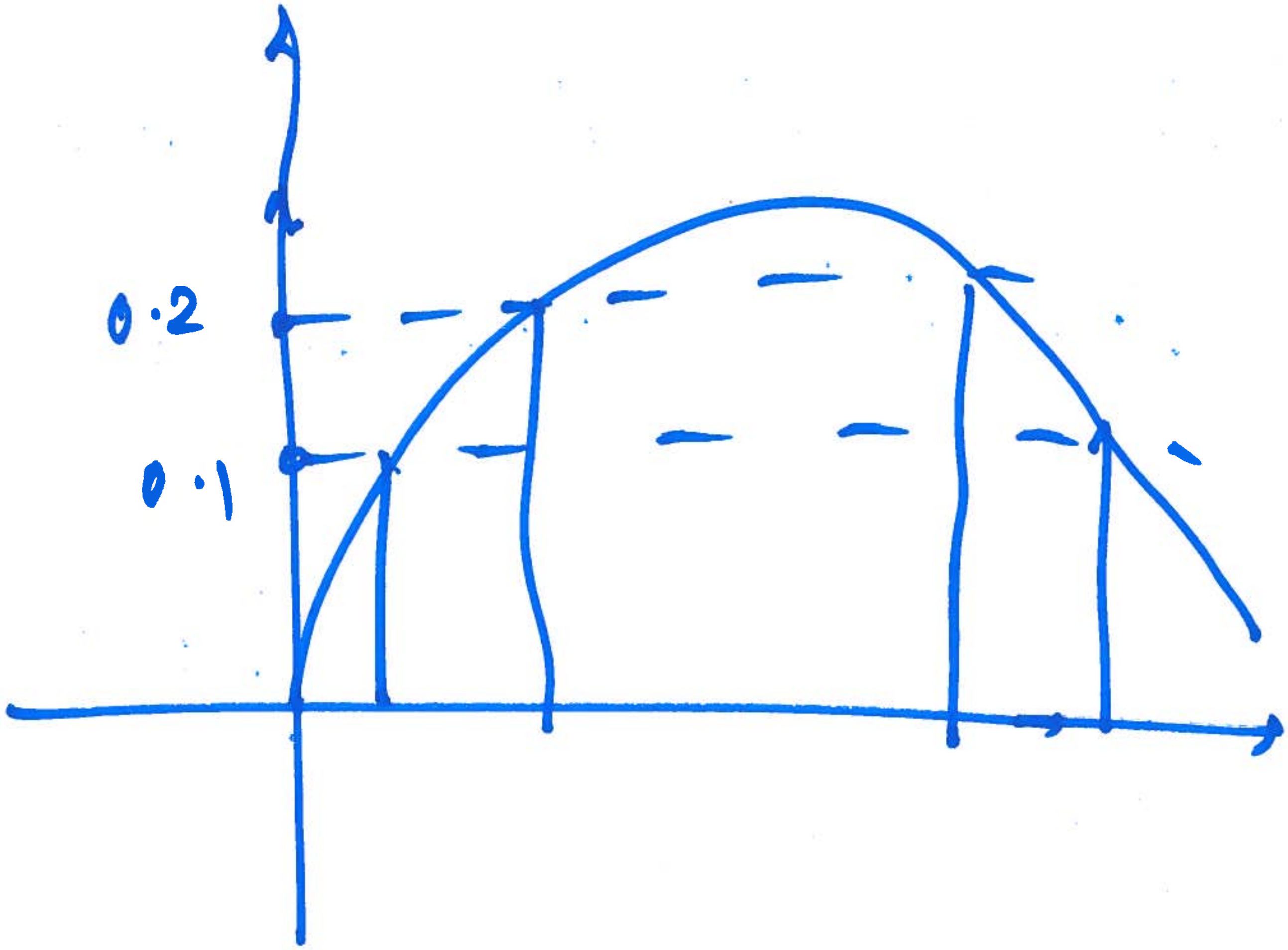
Geometrically, the Relative Likelihood function is easier to interpret.

Question 1

If θ belongs to the 20% likelihood interval, it **must** belong to the 10% likelihood interval.

(a) TRUE 73%

(b) FALSE 27%



$20\% \Rightarrow \quad R(\theta) \geq 0.2$

≥ 0.1

Suppose we are given the log relative likelihood function

$$r(\theta) = \log R(\theta)$$

$$R(\theta) \geqslant p \iff r(\theta) \geq \log p$$

Example: Suppose $Y \sim Bin(n, \theta)$

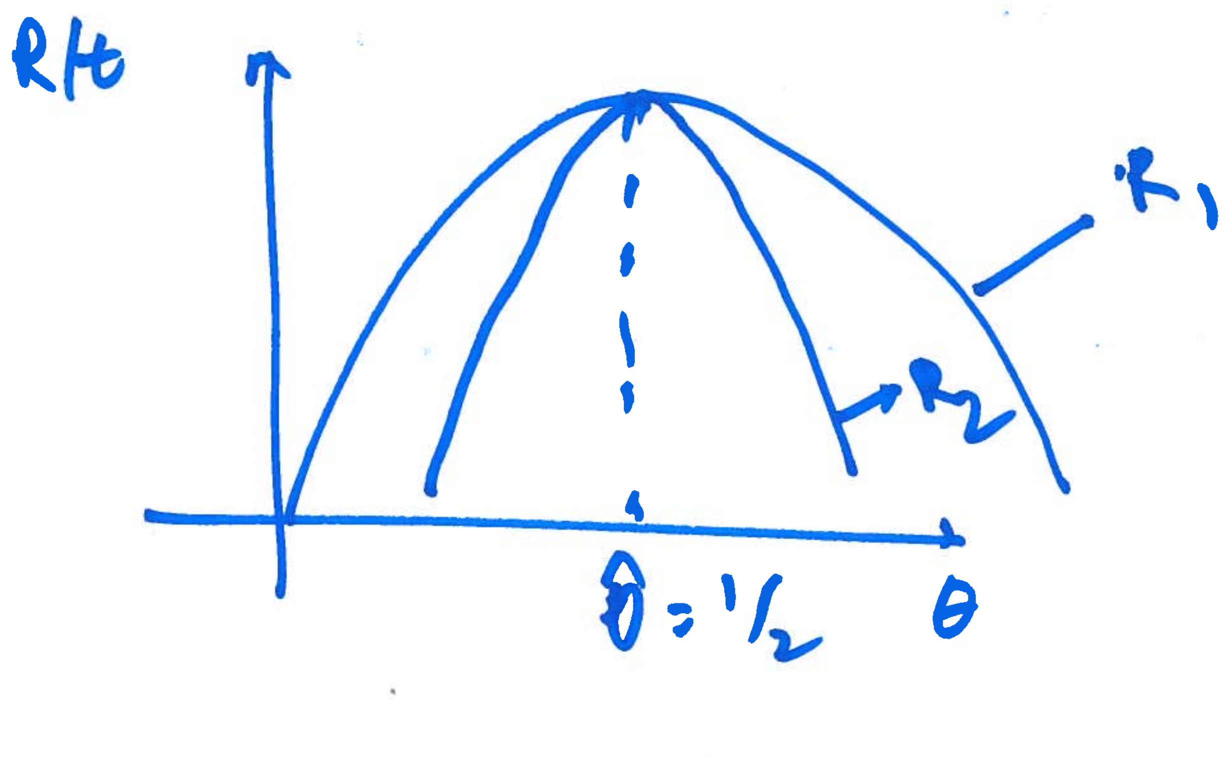
θ = prob. of success.

EXPERIMENT 1

$n = 100$ $\quad y_1 = 50$ $\quad \hat{\theta} = 1/2$

EXPERIMENT 2

$n = 500$ $\quad y = 250$ $\quad \hat{\theta} = 1/2$



Question

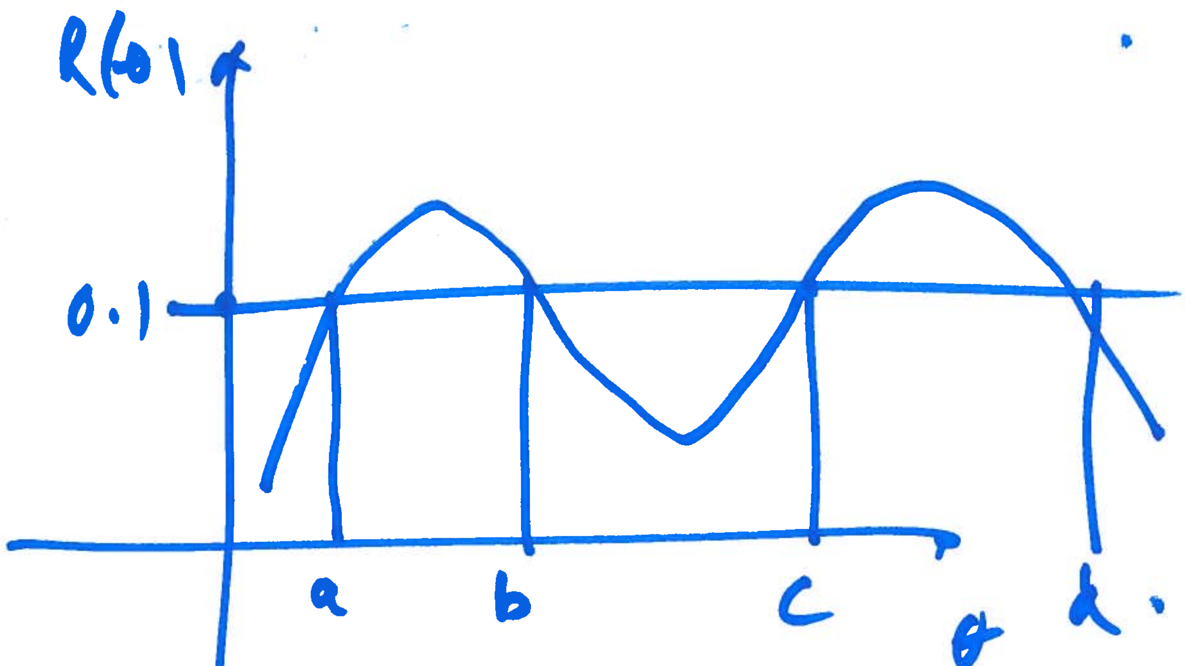
For the graph R_1, $n = 500$

(a) TRUE 50% A

(b) FALSE 50% B

Notes

- If $R(\theta)$ is bimodal, then the likelihood interval might be the union of two disjoint intervals



10% L.I $\longrightarrow [a,b] \cup [c,d]$

Typically, p is chosen to be between 0.1 and 0.15

CONVENTIONS

If $R(\theta) \geq 0.5$, i.e θ lies in the 50% likelihood interval $\Rightarrow$

VERY PLAUSIBLE

$\cancel{0.5 \quad R(\theta)}$

If $R(\theta) \geq 0.1$ but less than 0.5

PLAUSIBLE.

If $R(\theta) \leq 0.1$, but > 0.01

$\Rightarrow$ IMPLAUSIBLE

If $R(\theta) \leq 0.01$

$\Rightarrow$ VERY "

Downside to Likelihood Intervals

$\Rightarrow$ difficult interpretation

We want to estimate the probability that ~~the~~ interval that we construct would contain θ.
this

METHOD OF SAMPLING DISTRIBUTIONS

All numerical measures that we construct using our data set can be thought of as outcomes of a r.v.

@ μ = POPULATION MEAN

$\bar{y}$ = sample mean.

$\bar{Y}$ = r.v. from which $\bar{y}$ is drawn.

Notation

$\hat{\theta}$ = estimate = #

$\dot{\theta}$ = unknown parameter

$\tilde{\theta}$ = random variable of

which θ is an outcome.