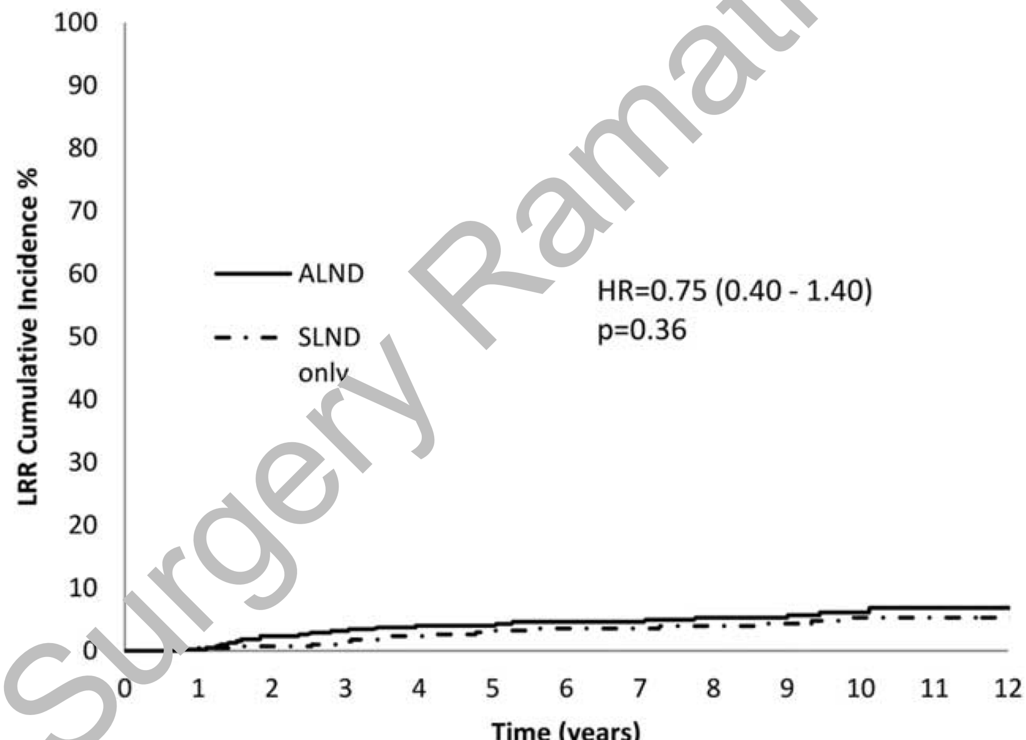
Confidence Intervals

Panuwat Lertsithichai

25 Feb 2019

Locoregional Recurrence After Sentinel Lymph Node Dissection With or Without Axillary Dissection in Patients With Sentinel Lymph Node Metastases

Long-term Follow-up From the American College of Surgeons Oncology Group (Alliance) ACOSOG Z0011 Randomized Trial



Ann Surg 2016;264:413-20.

What is Interval Estimation?

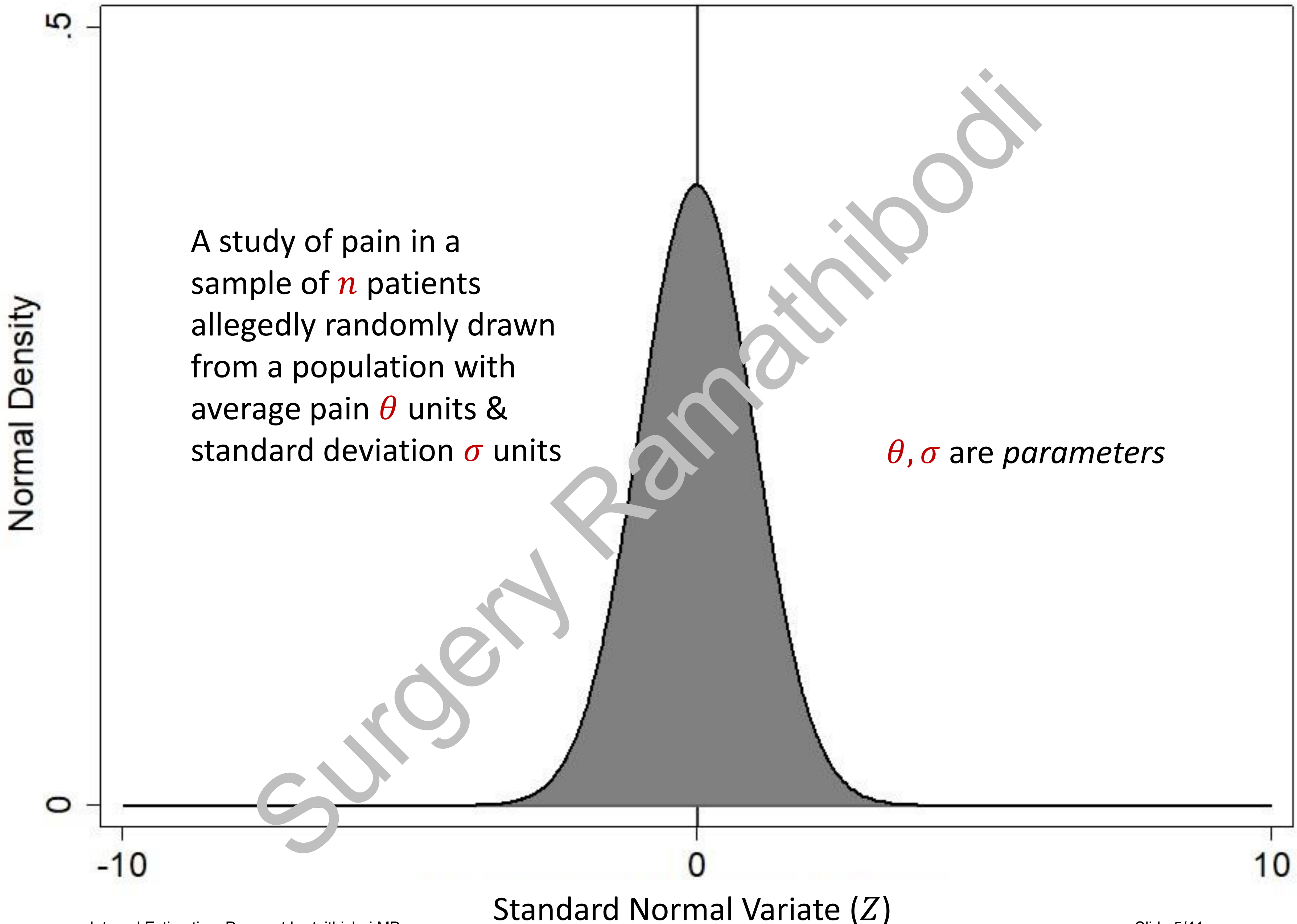
Statistics are often concerned with

- Hypothesis testing: **test of given quantities**
- Point estimation: **best guess of quantities**
- Interval estimation, e.g. Confidence Intervals (Frequentist statistics): **uncertainty in quantities**

- Hypothesis testing & interval estimation are intimately related

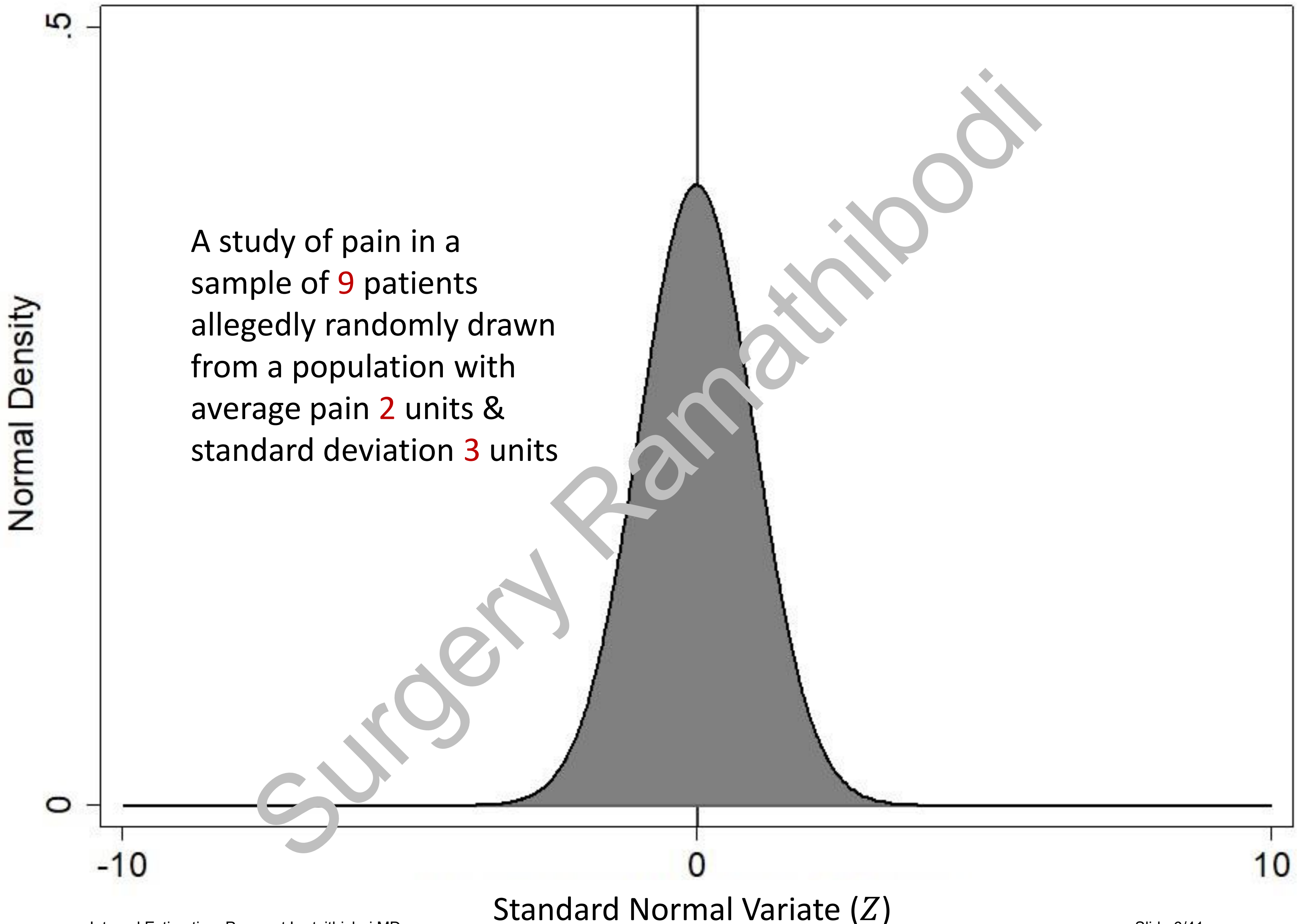
The Standard Normal Distribution

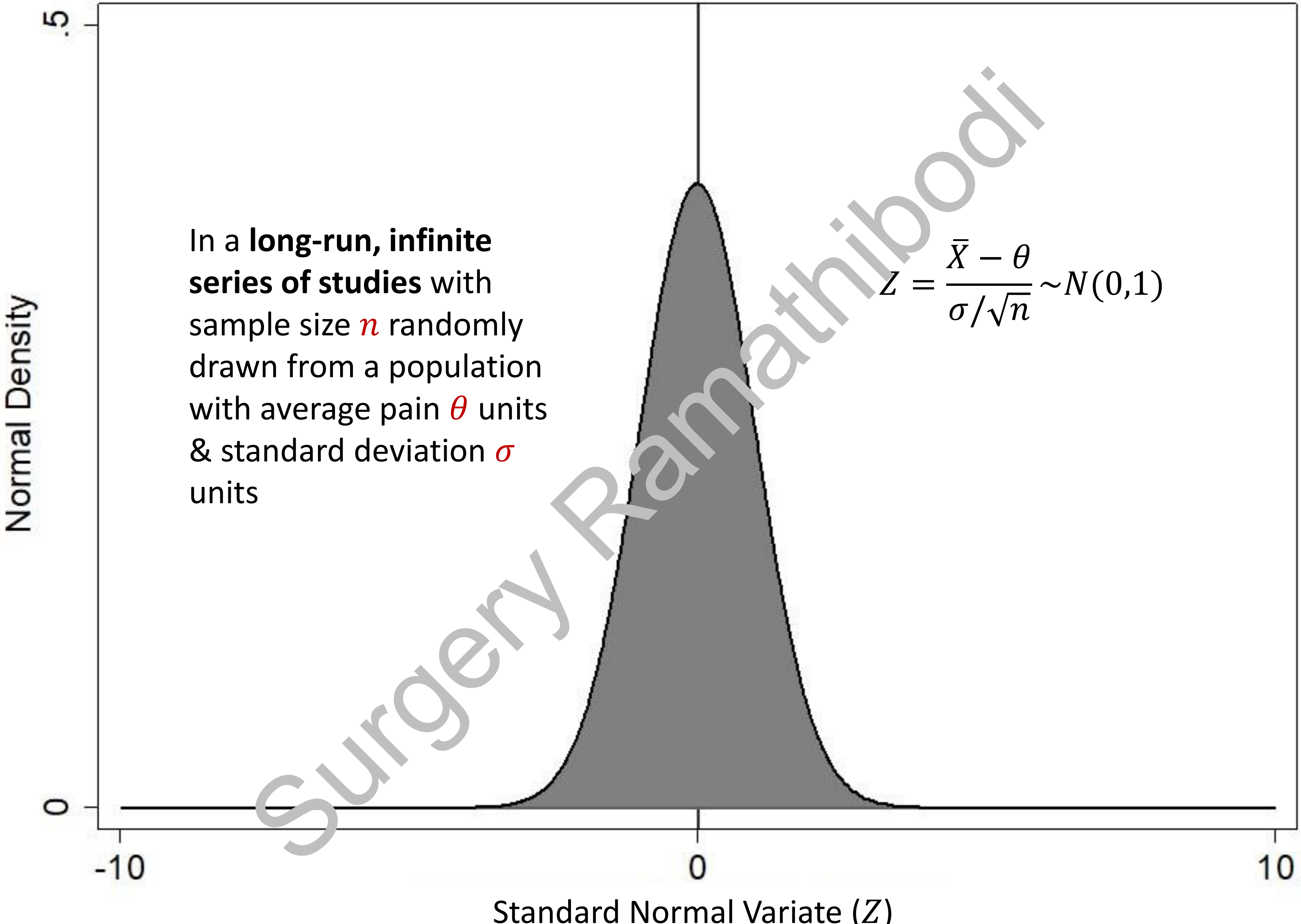
- A study with a continuous outcome measure
- Assumed to have a Normal Distribution in a population
- Can be standardized to have a **Standard** Normal Distribution
- In an *imagined* **long-run** series of experiments/ clinical studies



A study of pain in a sample of n patients allegedly randomly drawn from a population with average pain θ units & standard deviation σ units

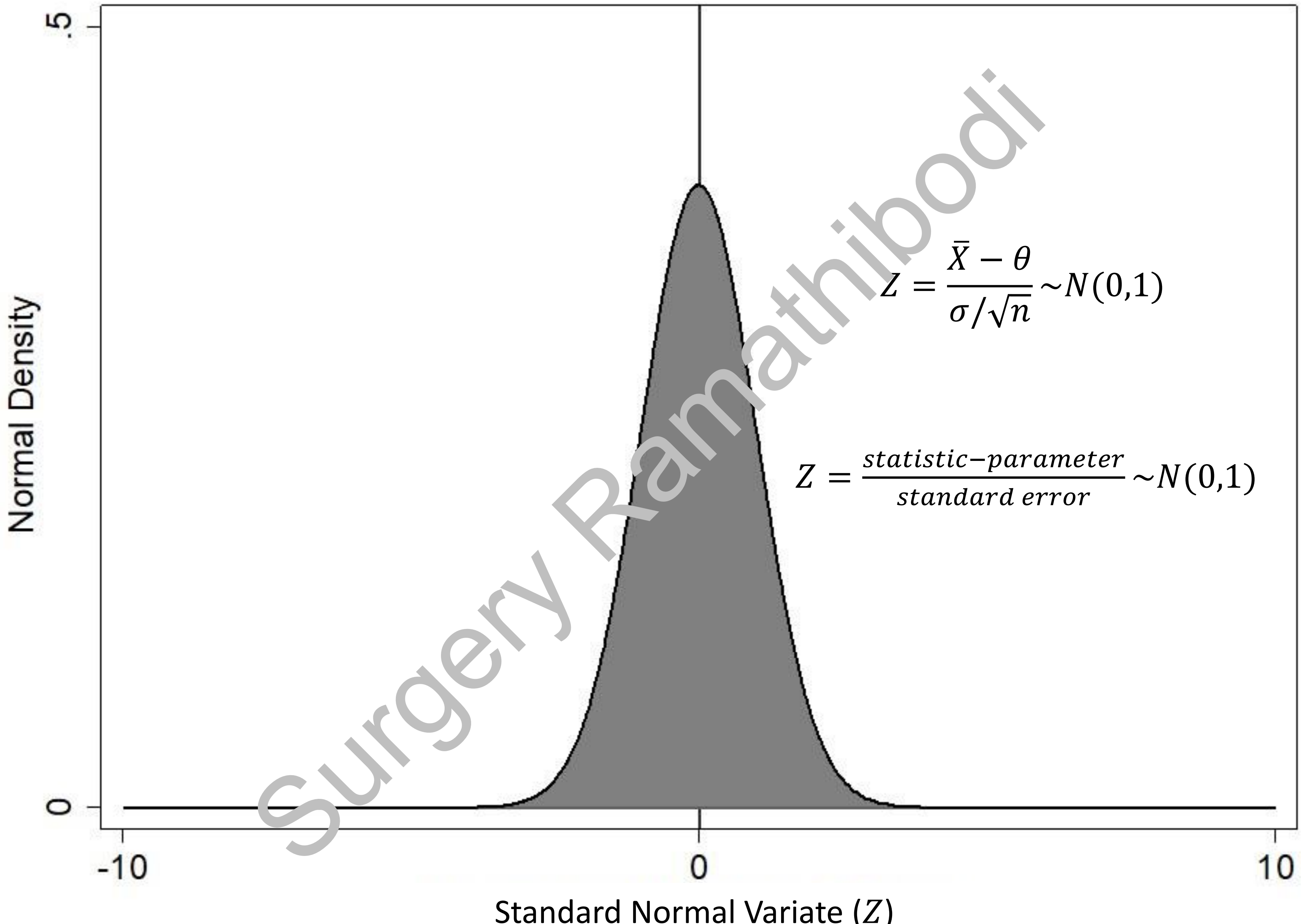
θ, σ are *parameters*


.5
0
Normal Density
-10
0
10
A study of pain in a sample of 9 patients allegedly randomly drawn from a population with average pain 2 units & standard deviation 3 units
Standard Normal Variate (Z)



In a **long-run, infinite series of studies** with sample size n randomly drawn from a population with average pain θ units & standard deviation σ units

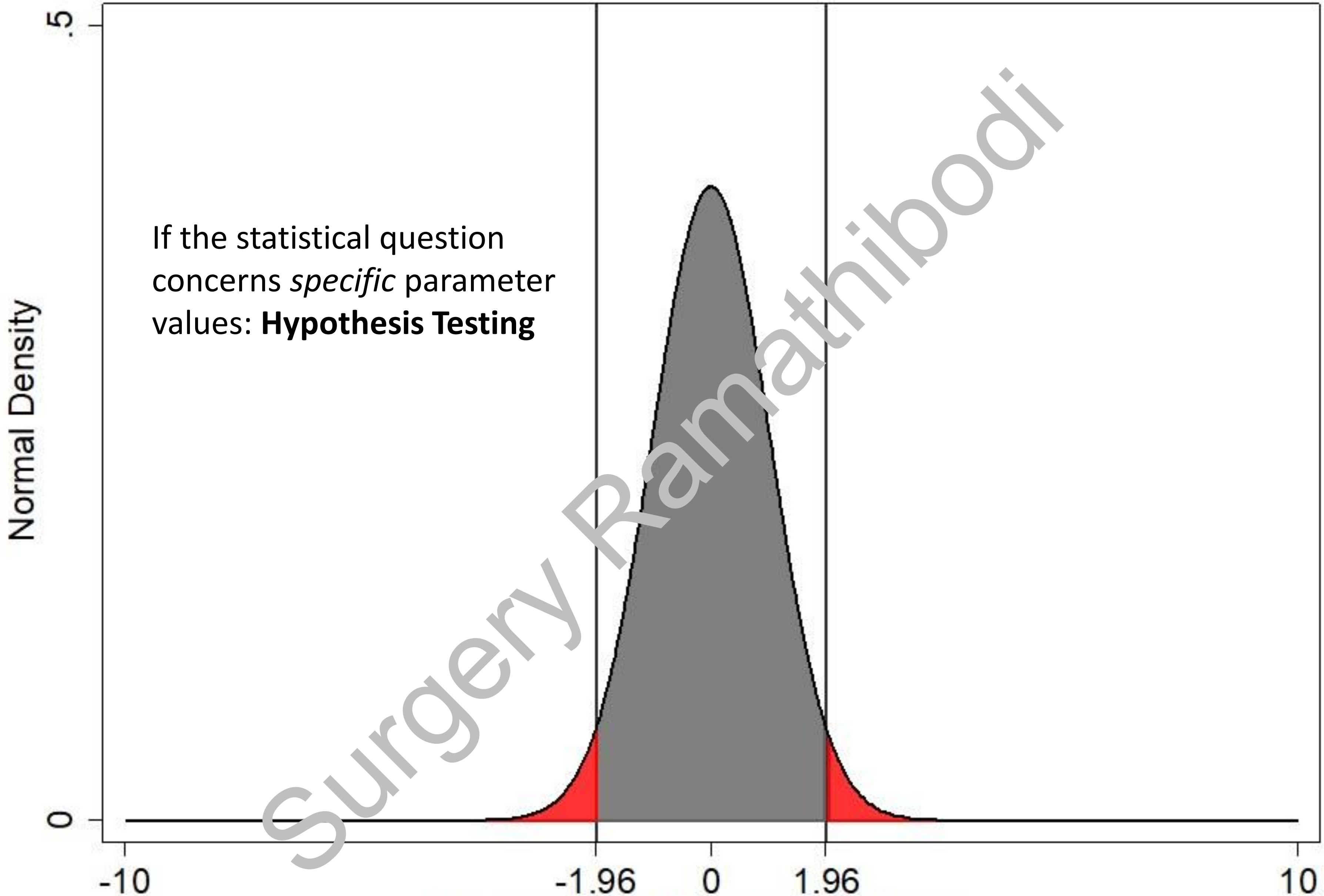
$$Z = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \sim N(0,1)$$



$$Z = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \sim N(0,1)$$

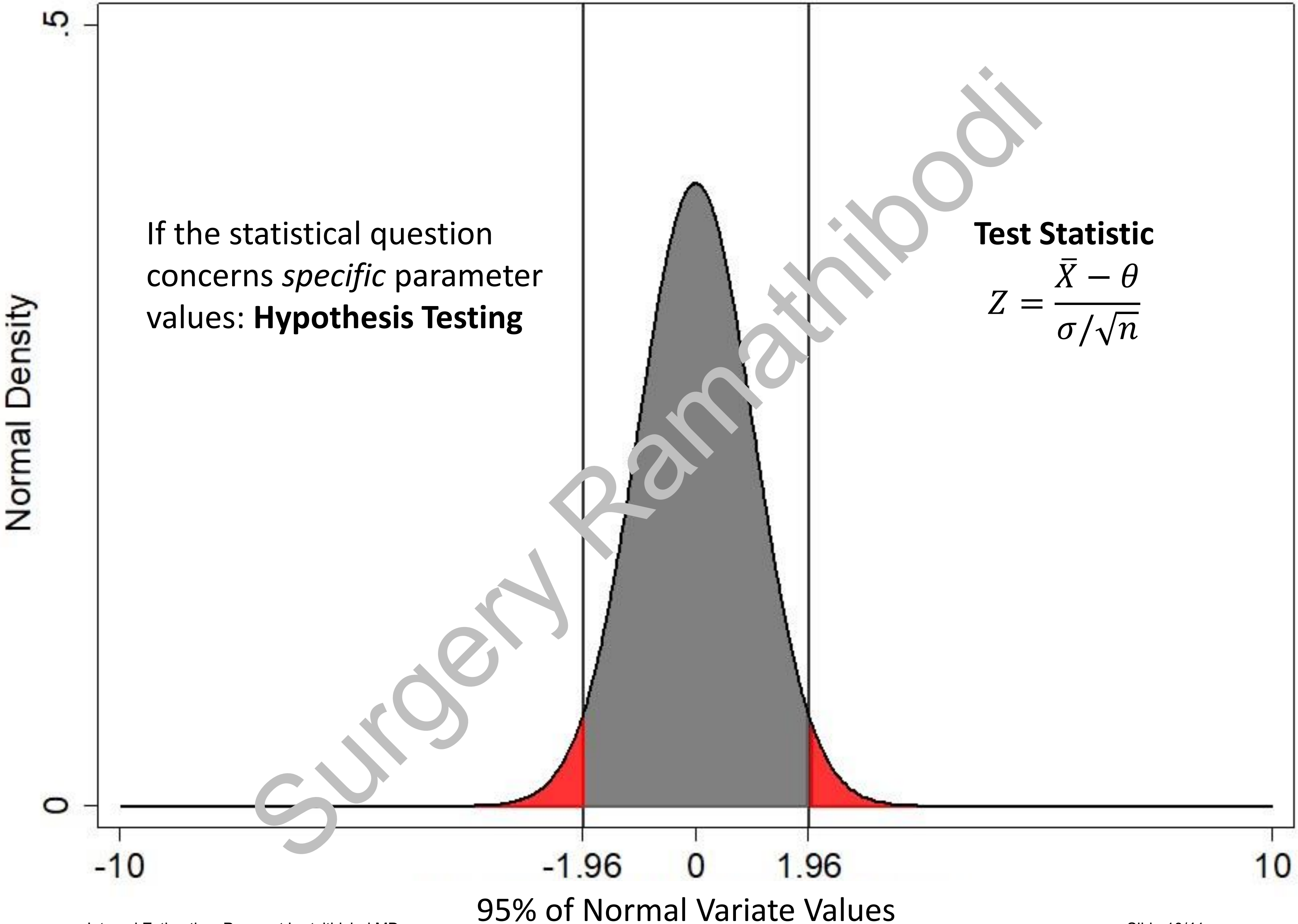
$$Z = \frac{statistic - parameter}{standard\ error} \sim N(0,1)$$

Hypothesis Testing Under Null: mean $= \theta$ vs. Alternative: mean $\neq \theta$



If the statistical question concerns *specific* parameter values: **Hypothesis Testing**

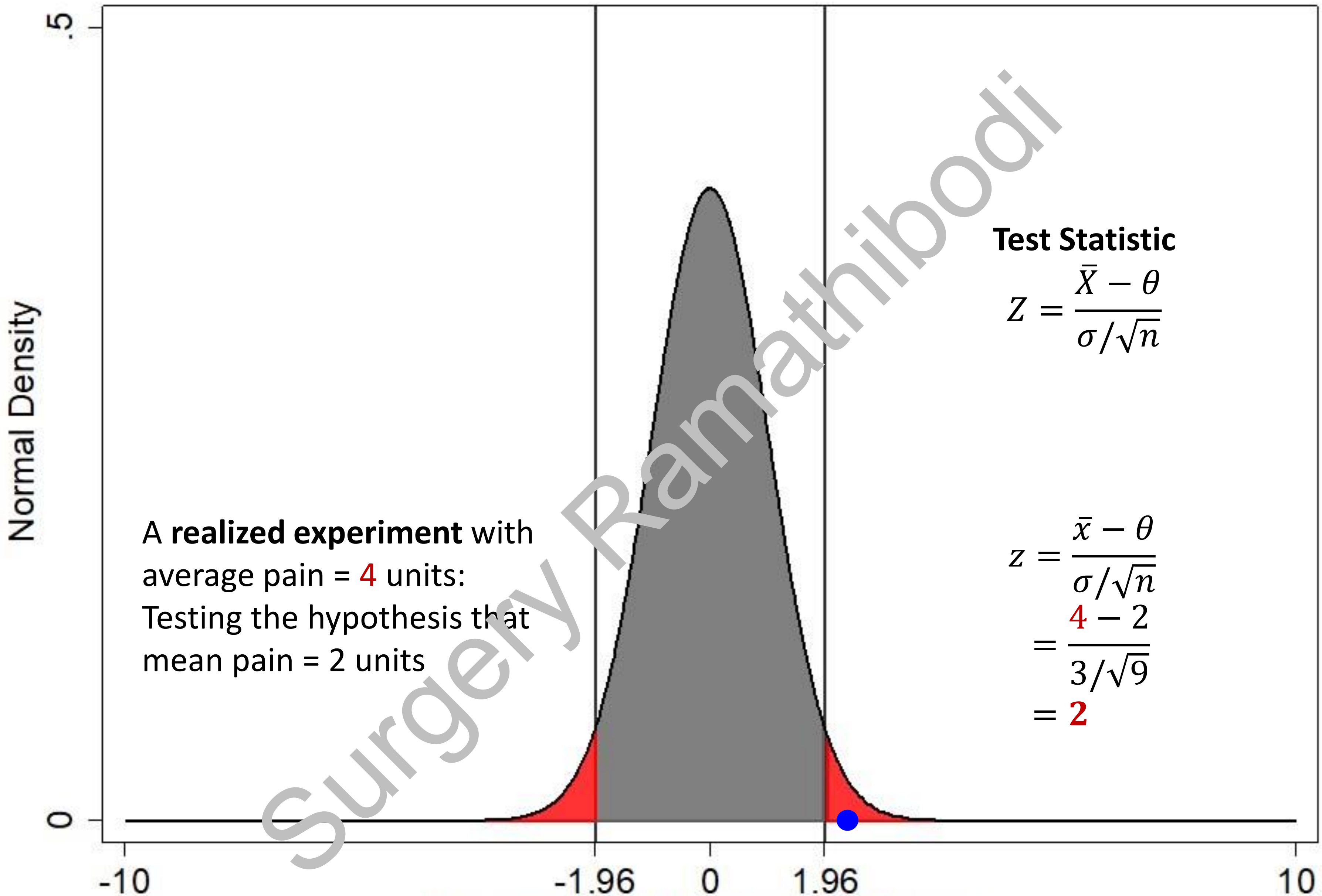
Hypothesis Testing Under Null: mean $= \theta$ vs. Alternative: mean $\neq \theta$



If the statistical question concerns *specific* parameter values: **Hypothesis Testing**

Test Statistic

$$Z = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}}$$



95% of Normal Variate Values

Test Statistic

$$Z = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}}$$

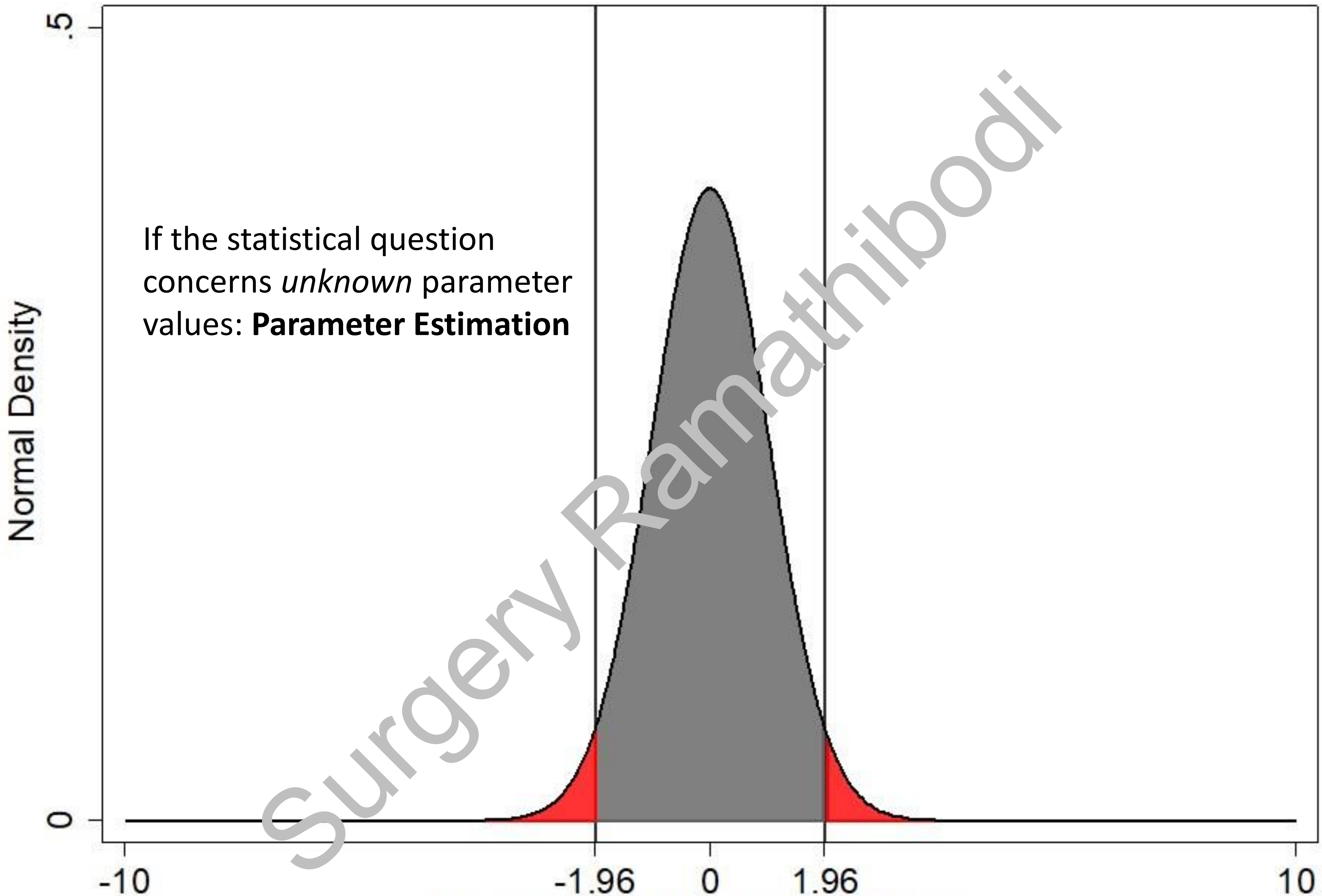
A **realized experiment** with average pain = 4 units: Testing the hypothesis that mean pain = 2 units

$$z = \frac{\bar{x} - \theta}{\sigma/\sqrt{n}} = \frac{4 - 2}{3/\sqrt{9}} = \mathbf{2}$$

Random Variable & Realized Value

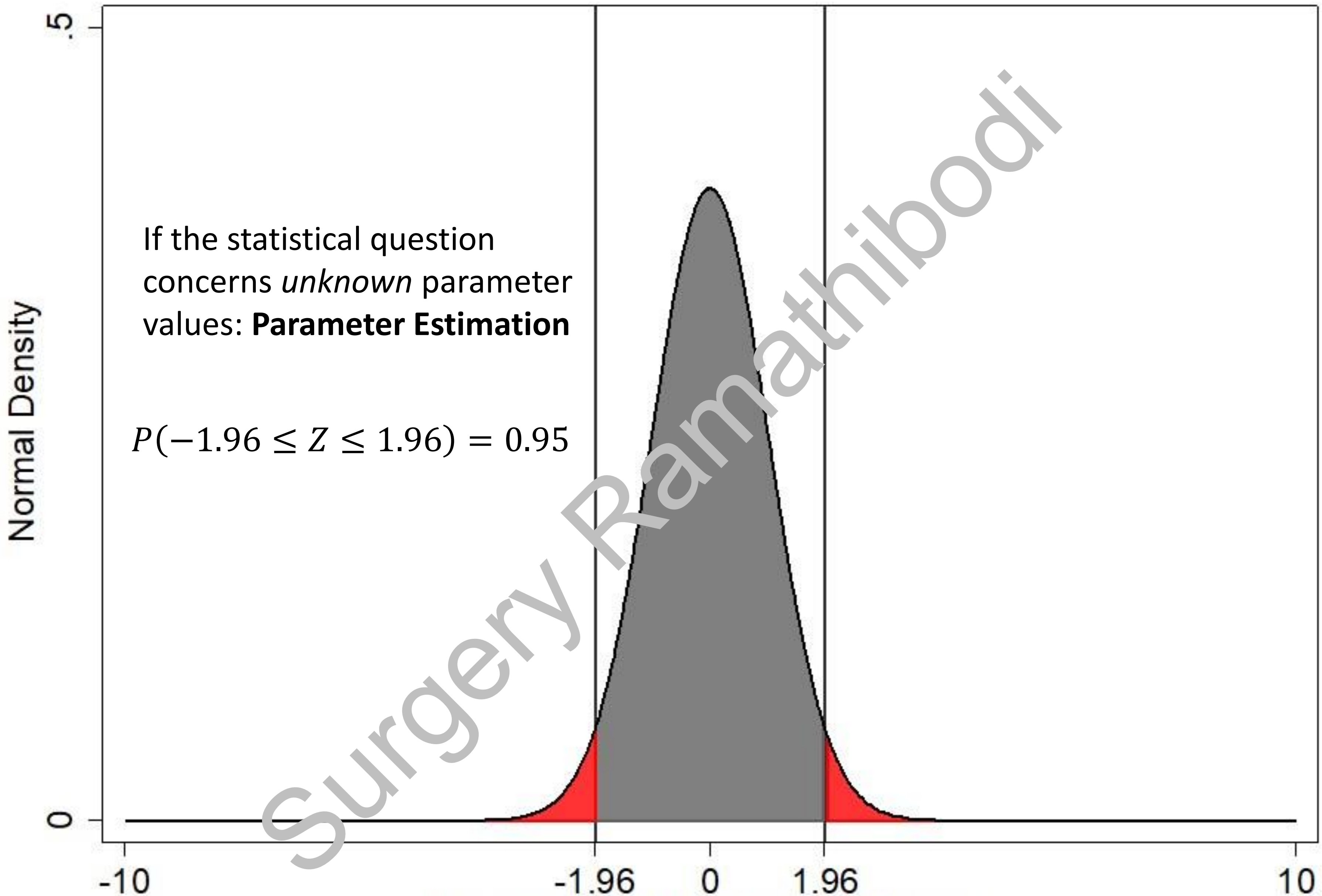
- Throwing a fair coin
- The outcome X, with *possible values* $X = H$ “Heads” and $X = T$ “Tails”, is a random variable
- Thus, $P(X = H) = 0.5$
- The *realized outcome* of a completed throw $x = H$ is no longer random: it is “Heads” with probability 1 and “Tails” with probability 0
- Random experiment: outcome not known, but with a known (class of) probability distribution(s)
- Realized experiment: outcome has occurred & known

If the statistical question concerns *unknown* parameter values: **Parameter Estimation**

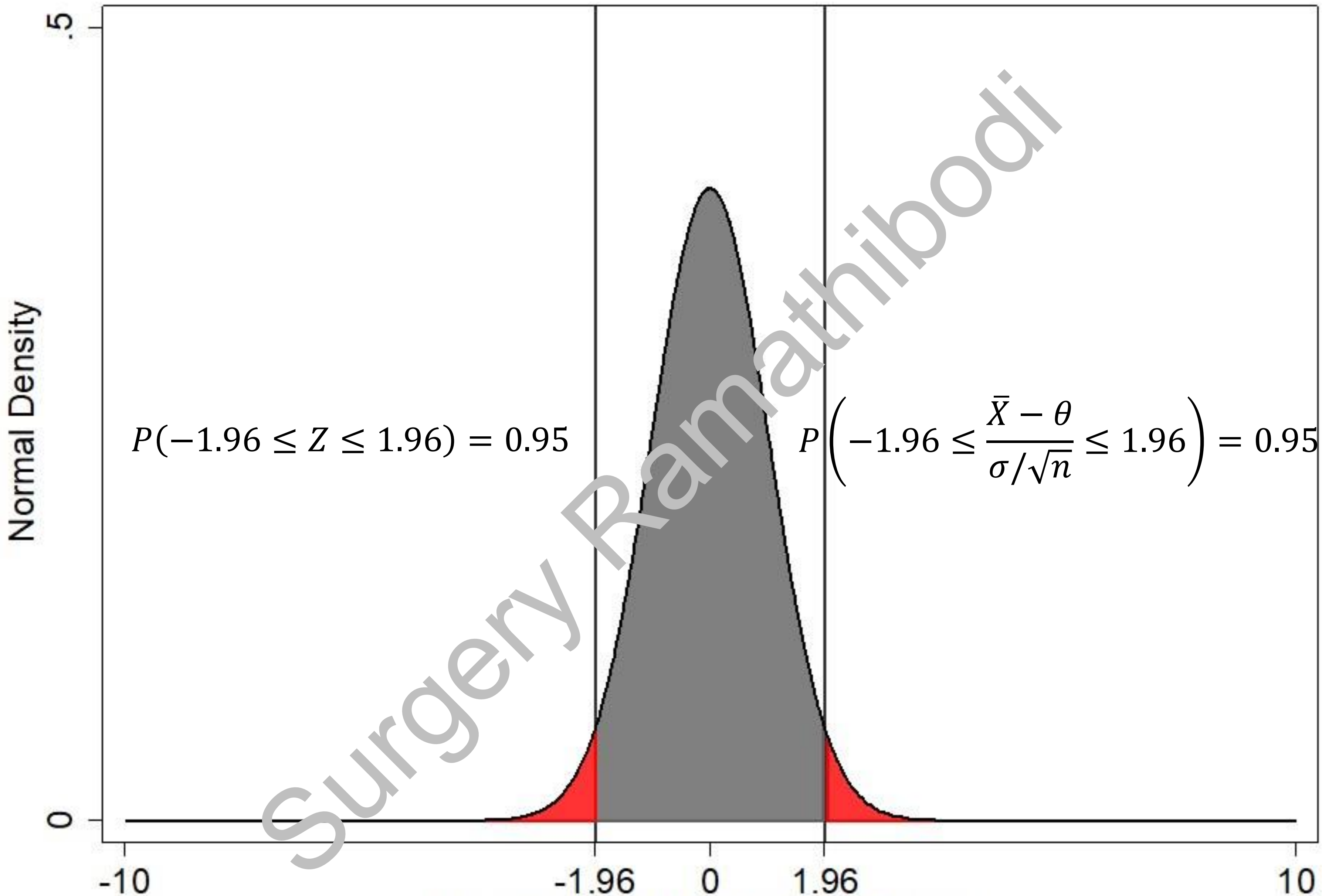


If the statistical question concerns *unknown* parameter values: **Parameter Estimation**

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$



95% of Normal Variate Values


.5
0
Normal Density
$P(-1.96 \leq Z \leq 1.96) = 0.95$
$P\left(-1.96 \leq \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \leq 1.96\right) = 0.95$
-10
-1.96
0
1.96
10
95% of Normal Variate Values

Inversion of Test Statistics

- $P\left(-1.96 \leq \frac{\bar{X}-\theta}{\sigma/\sqrt{n}} \leq 1.96\right) = 0.95$
- $P\left(-1.96\frac{\sigma}{\sqrt{n}} \leq \bar{X} - \theta \leq 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$
- $P\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} \leq \theta \leq \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$

Thus, the 95% upper & lower bounds are $\bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}}$

Null Hypothesis of mean $= \theta$ against Alternative Hypotheses of mean $\neq \theta$ implies 2-sided interval estimates

Confidence Interval (CI)

The (95%) Confidence Interval is

- $\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$

If is $\boldsymbol{\sigma}$ unknown, and if n is sufficiently large, then the approximate 95% CI is

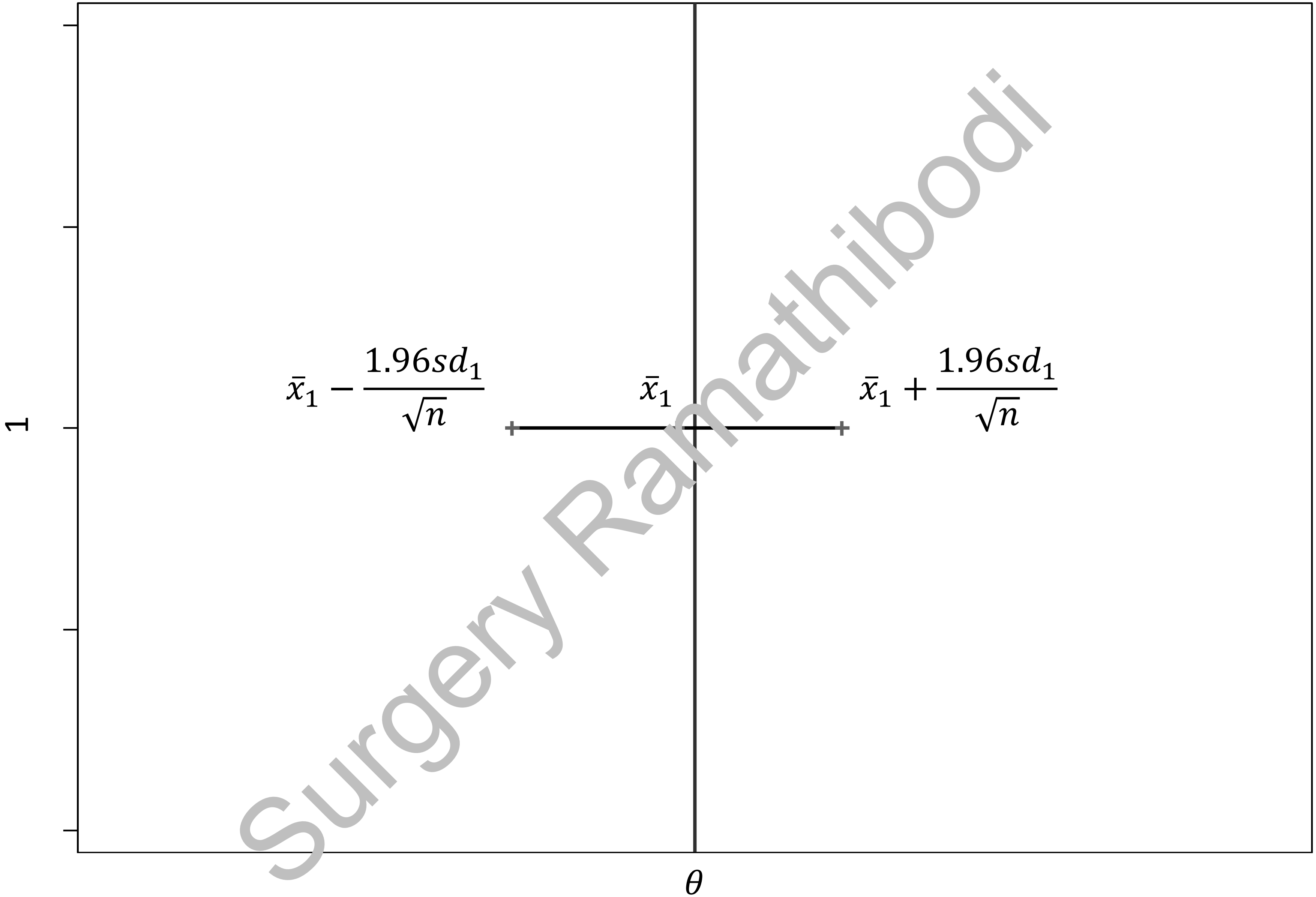
- $\left[\bar{X} - 1.96\frac{SD}{\sqrt{n}}, \bar{X} + 1.96\frac{SD}{\sqrt{n}}\right]$

Where SD is the standard deviation of the data

The Meaning of *Random* CI's

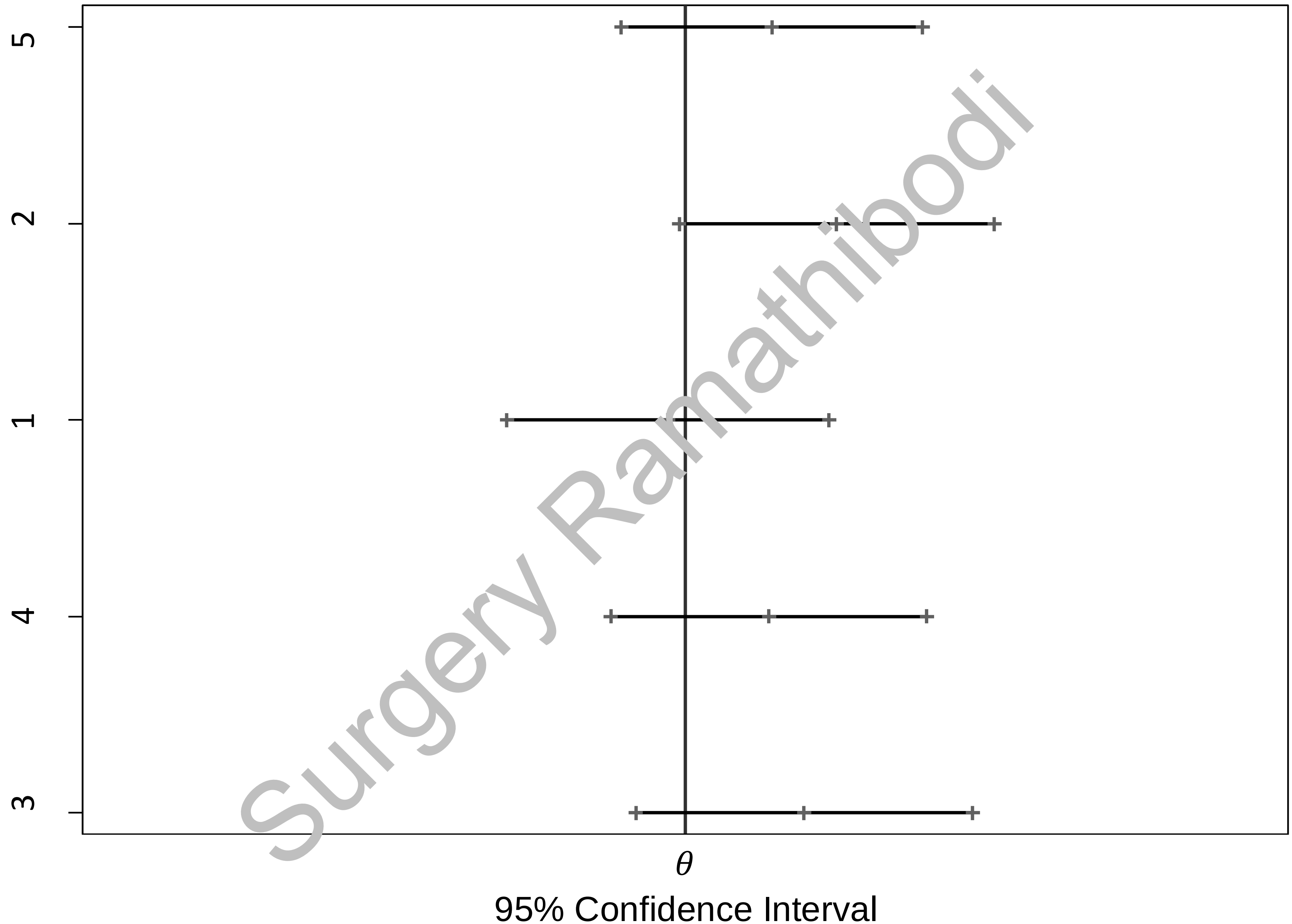
- Experimental outcomes are *random*, so $\bar{X}$ (**sample mean/average**) are *random variables*
- In the Frequentist paradigm, θ (***population* mean/ average**) is *fixed* and the probabilities are *long-run relative frequencies*
- And so the interpretation of the **random** Confidence Interval is …

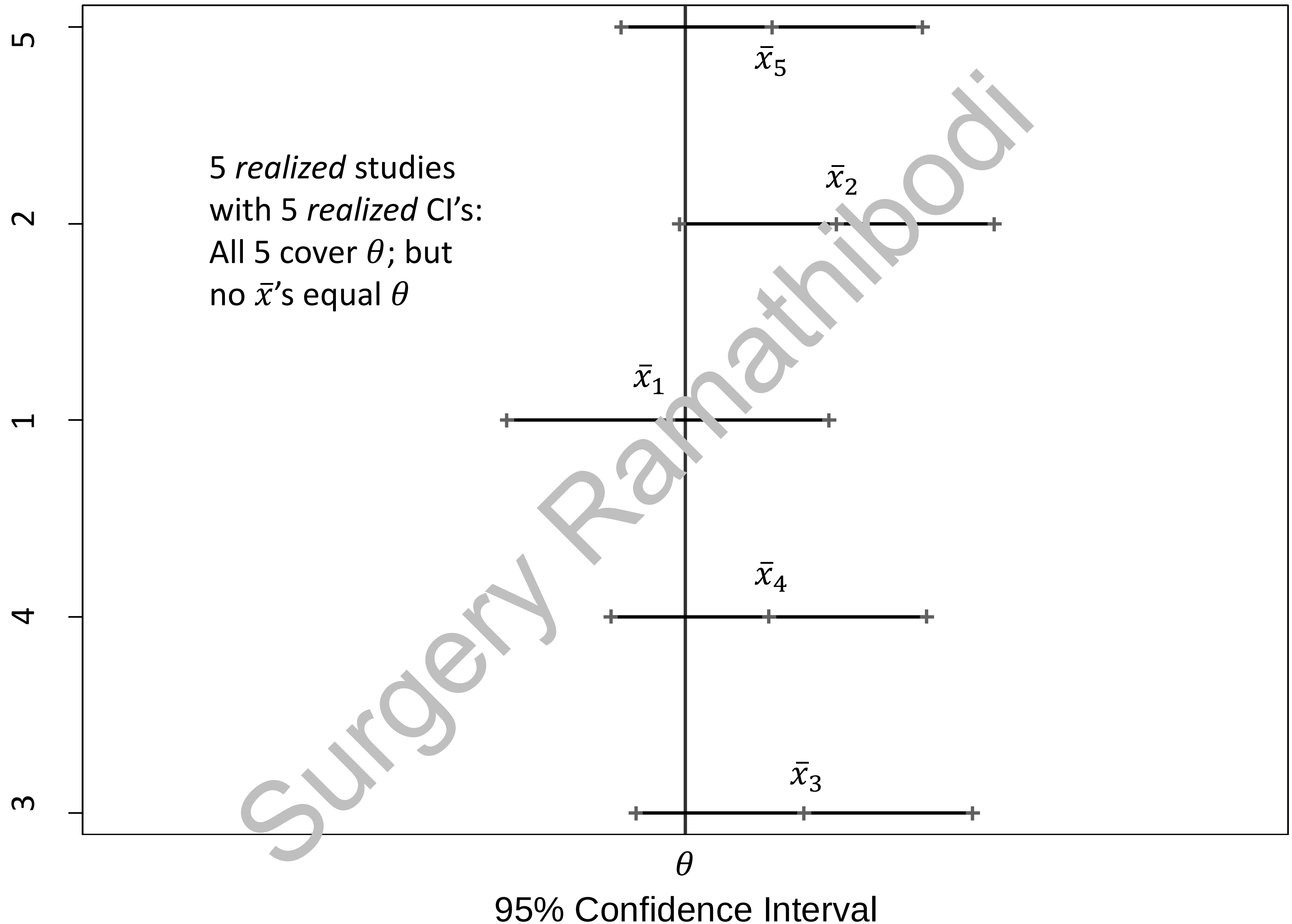

A *realized* 95% CI: study 1
1
θ
95% Confidence Interval

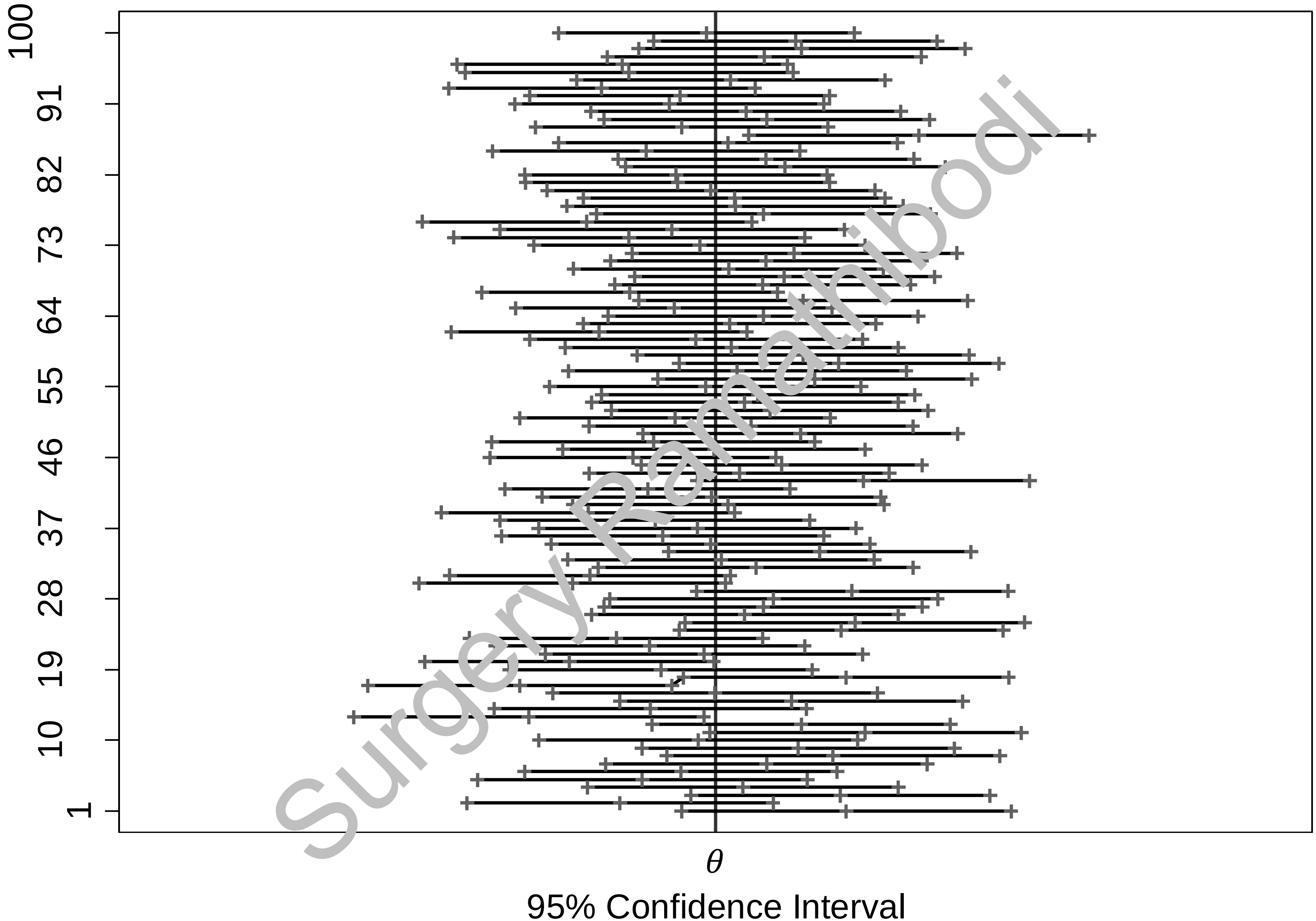


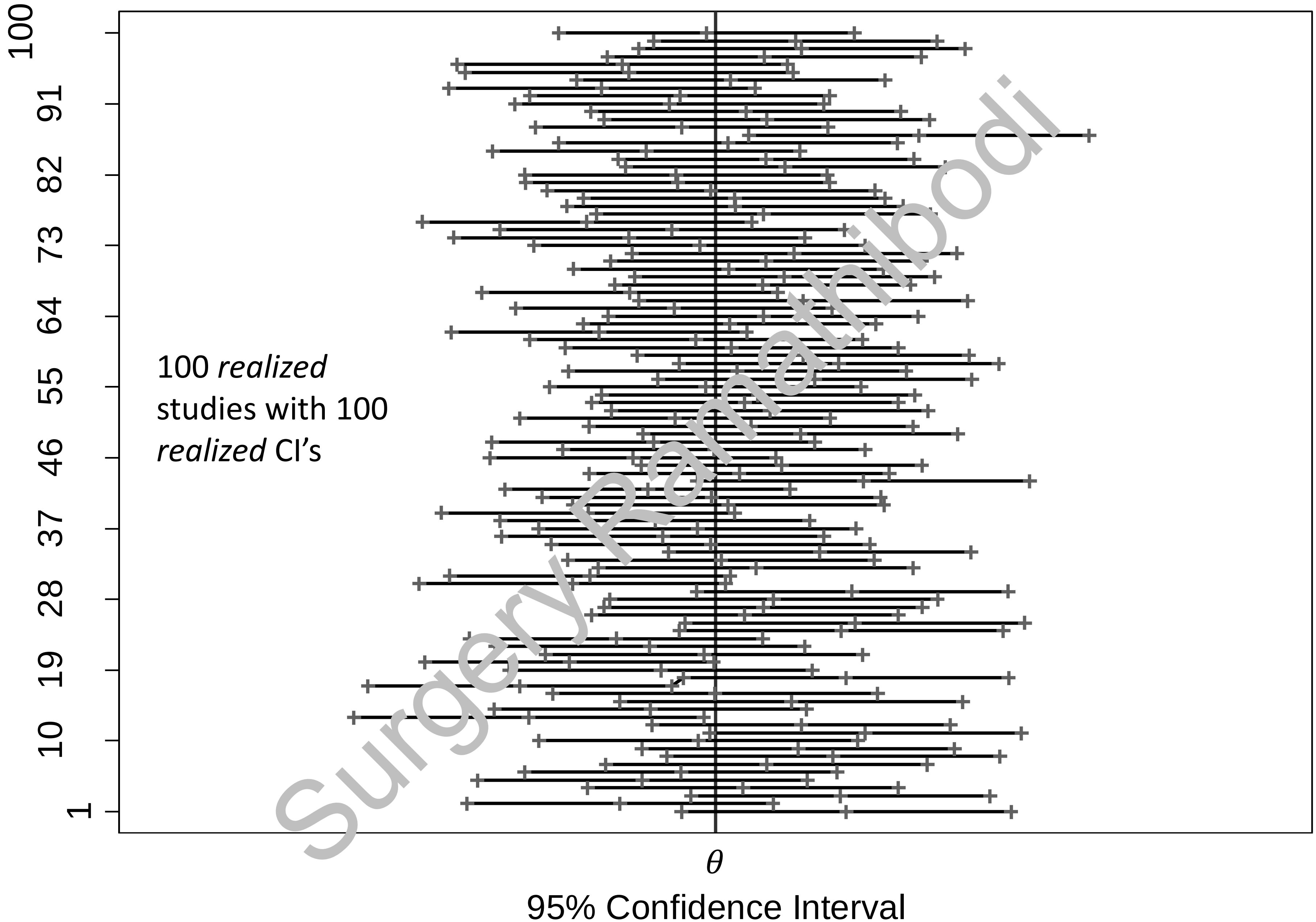
95% Confidence Interval

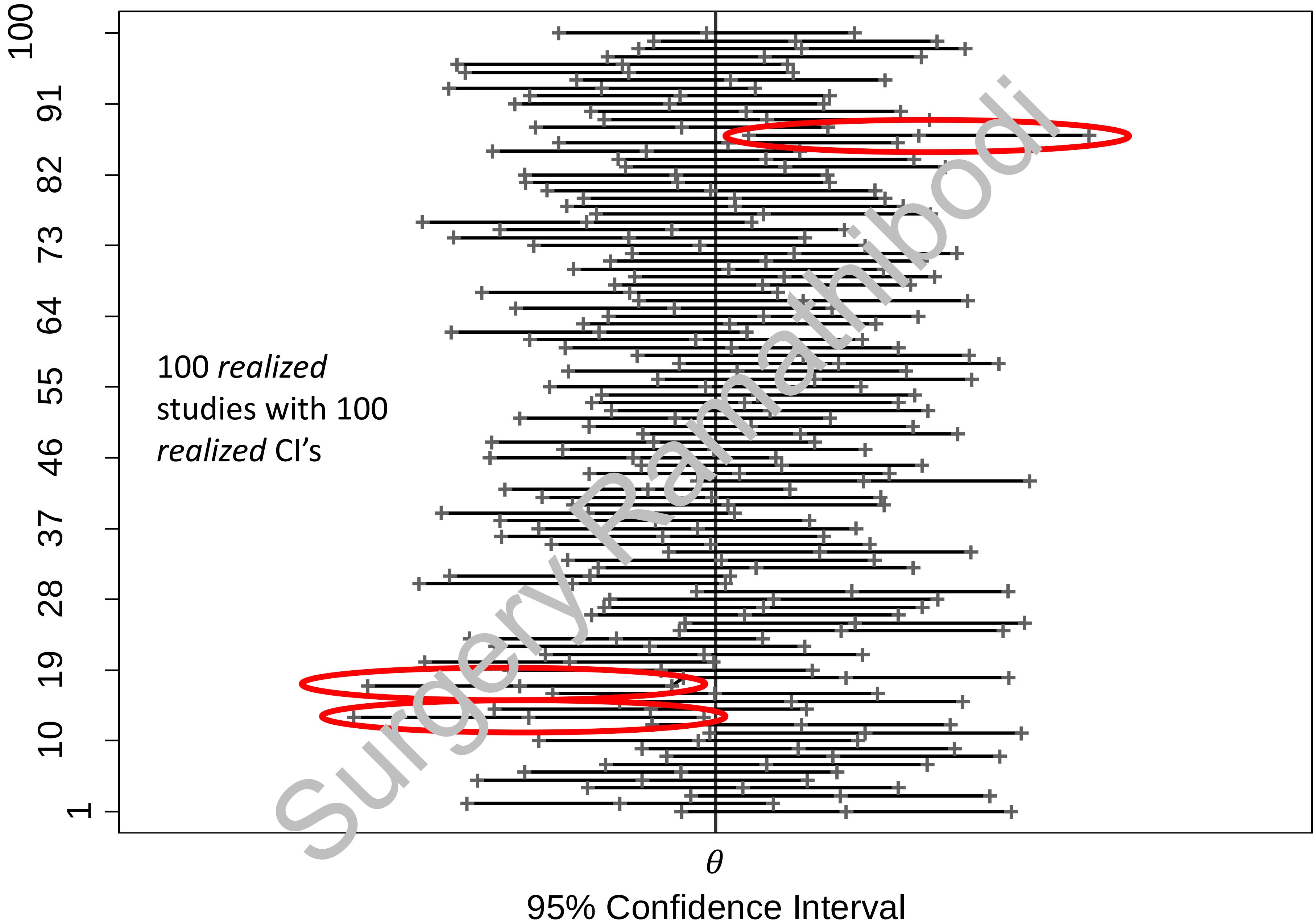
$\bar{x}_2 - \frac{1.96sd_2}{\sqrt{n}}$
$\bar{x}_2$
$\bar{x}_2 + \frac{1.96sd_2}{\sqrt{n}}$
2
1
θ
95% Confidence Interval


5
2
1
4
3
θ
95% Confidence Interval

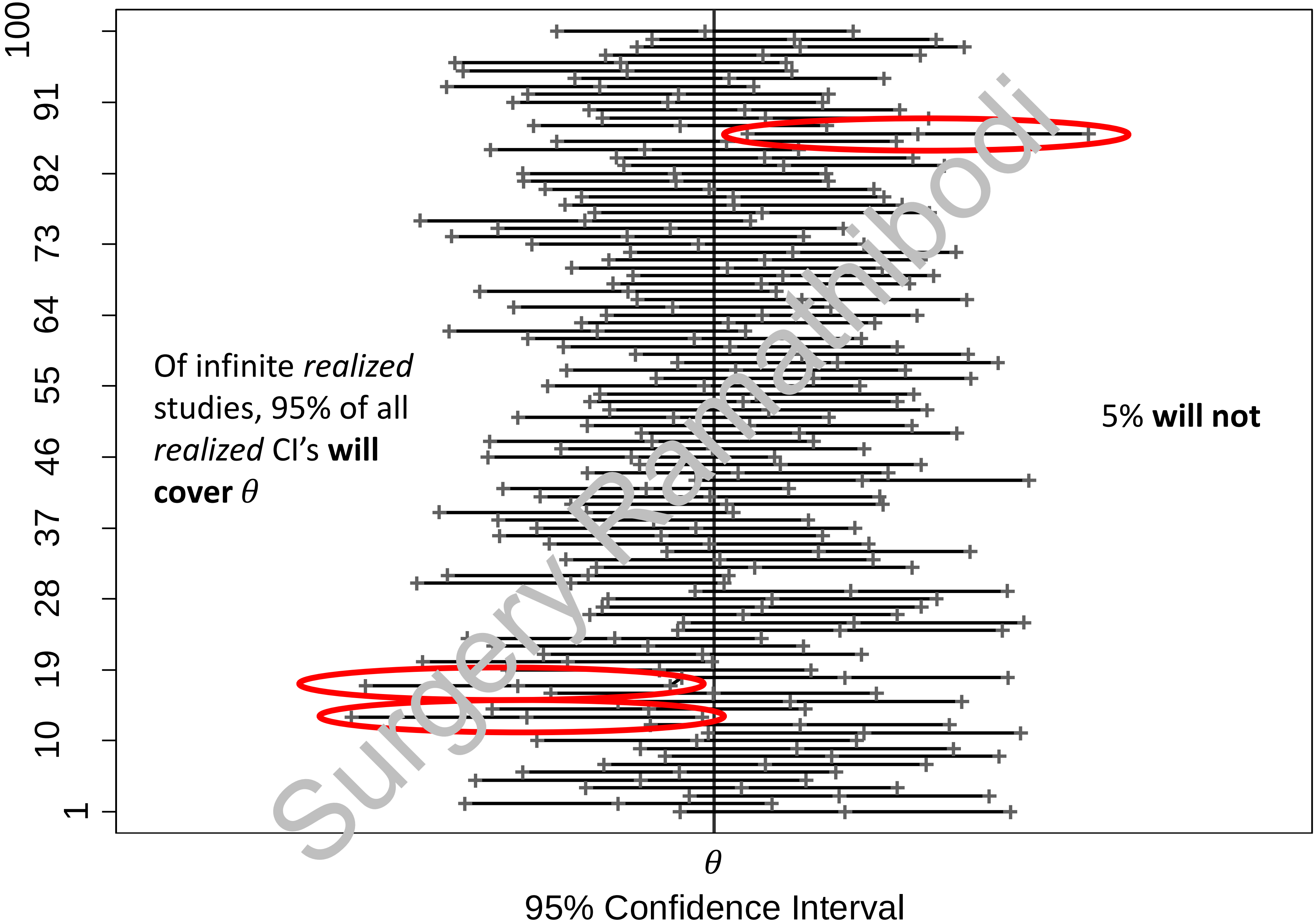

5 *realized* studies
with 5 *realized* CI's:
All 5 cover θ; but
no $\bar{x}$'s equal θ
$\bar{x}_5$
$\bar{x}_2$
$\bar{x}_1$
$\bar{x}_4$
$\bar{x}_3$
5
2
1
4
3
θ
95% Confidence Interval


100
91
82
73
64
55
46
37
28
19
10
1
θ
95% Confidence Interval

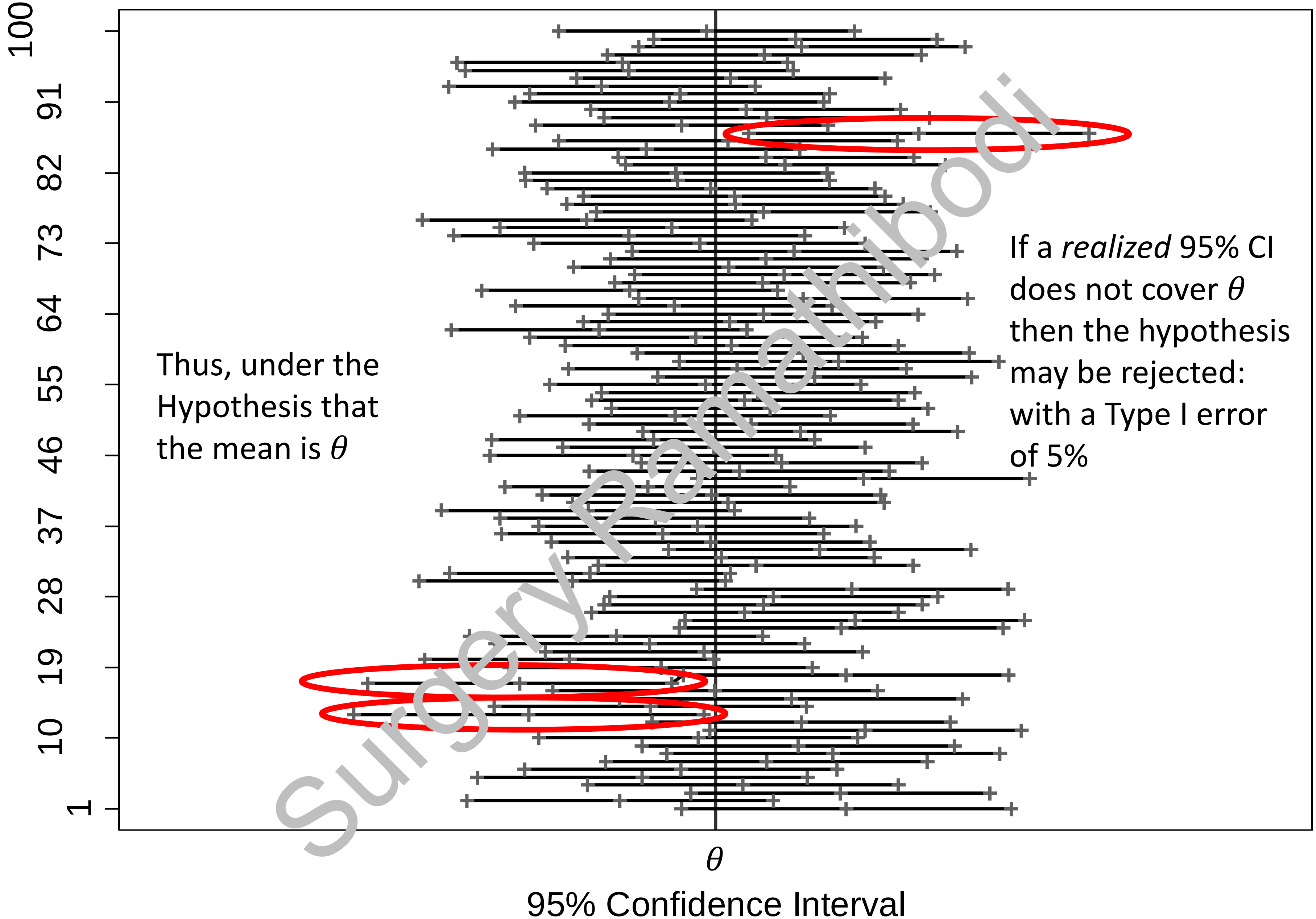

100
91
82
73
64
55
46
37
28
19
10
1
100 *realized* studies with 100 *realized* CI's
θ
95% Confidence Interval


100
91
82
73
64
55
46
37
28
19
10
1
100 *realized* studies with 100 *realized* CI's
θ
95% Confidence Interval

The Meaning of the **Random** 95% Confidence Interval



Use 95% Confidence Interval for **Hypothesis Testing**?



Thus, under the Hypothesis that the mean is θ

If a *realized* 95% CI does not cover θ then the hypothesis may be rejected: with a Type I error of 5%

The Meaning of *Realized* CI's

The *probability* of the **random** 95% CI

- $\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$ covering θ is 95%

However, a *specific*, **realized** 95% CI

- $\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$ e.g. [2,6] either *does or does not* cover θ

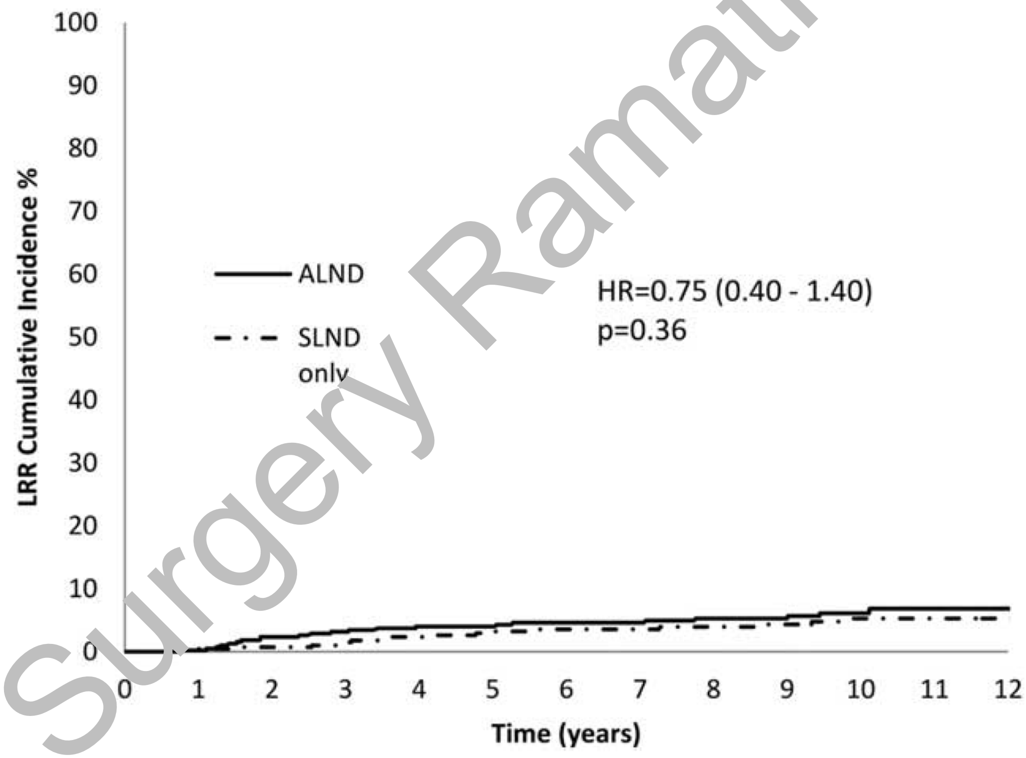
We thus use the phrase, we are 95% *confident* that [2,6] will cover θ

Examples of Estimation Problems

- What is the likely average pain of the operation, in the population, and what is the uncertainty?
- From our study of 9 patients, the average pain is 4 units, the 95% CI is [2,6] (realized experiment)
- Estimate the Odds Ratio of 10-yr Breast Cancer Recurrence if given RT vs. no RT in BCS
- Estimate the Hazard Ratio of Disease-free Survival at 5 years for Stage III Breast Cancer vs. Stage I

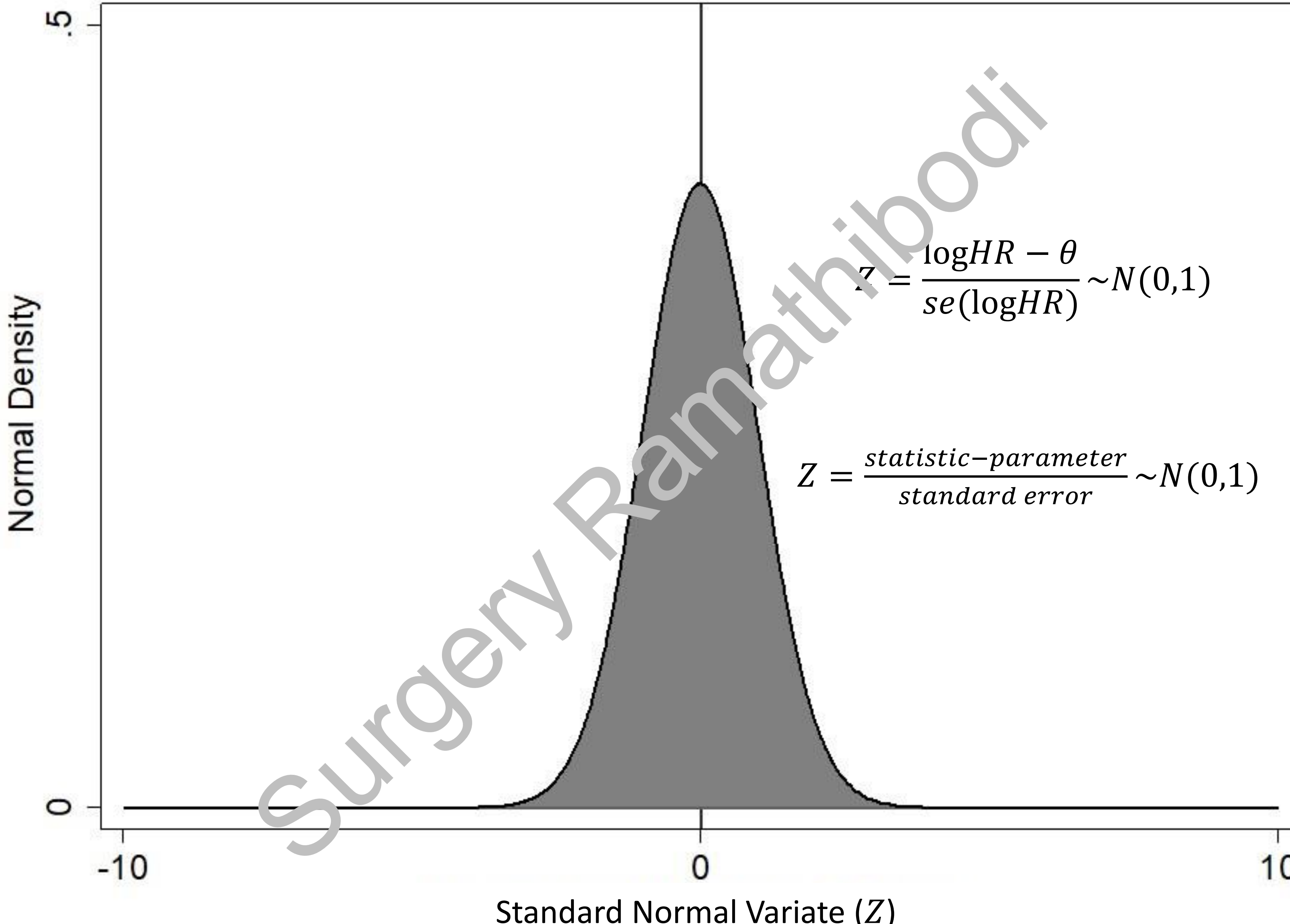
Locoregional Recurrence After Sentinel Lymph Node Dissection With or Without Axillary Dissection in Patients With Sentinel Lymph Node Metastases

Long-term Follow-up From the American College of Surgeons Oncology Group (Alliance) ACOSOG Z0011 Randomized Trial



Ann Surg 2016;264:413-20.

Log(HR) has an approximate (asymptotic) Normal Distribution!



$$Z = \frac{\log HR - \theta}{se(\log HR)} \sim N(0,1)$$

$$Z = \frac{statistic - parameter}{standard\ error} \sim N(0,1)$$

The Hazard Ratio (HR)

The (95%) Confidence Interval is, in the log scale

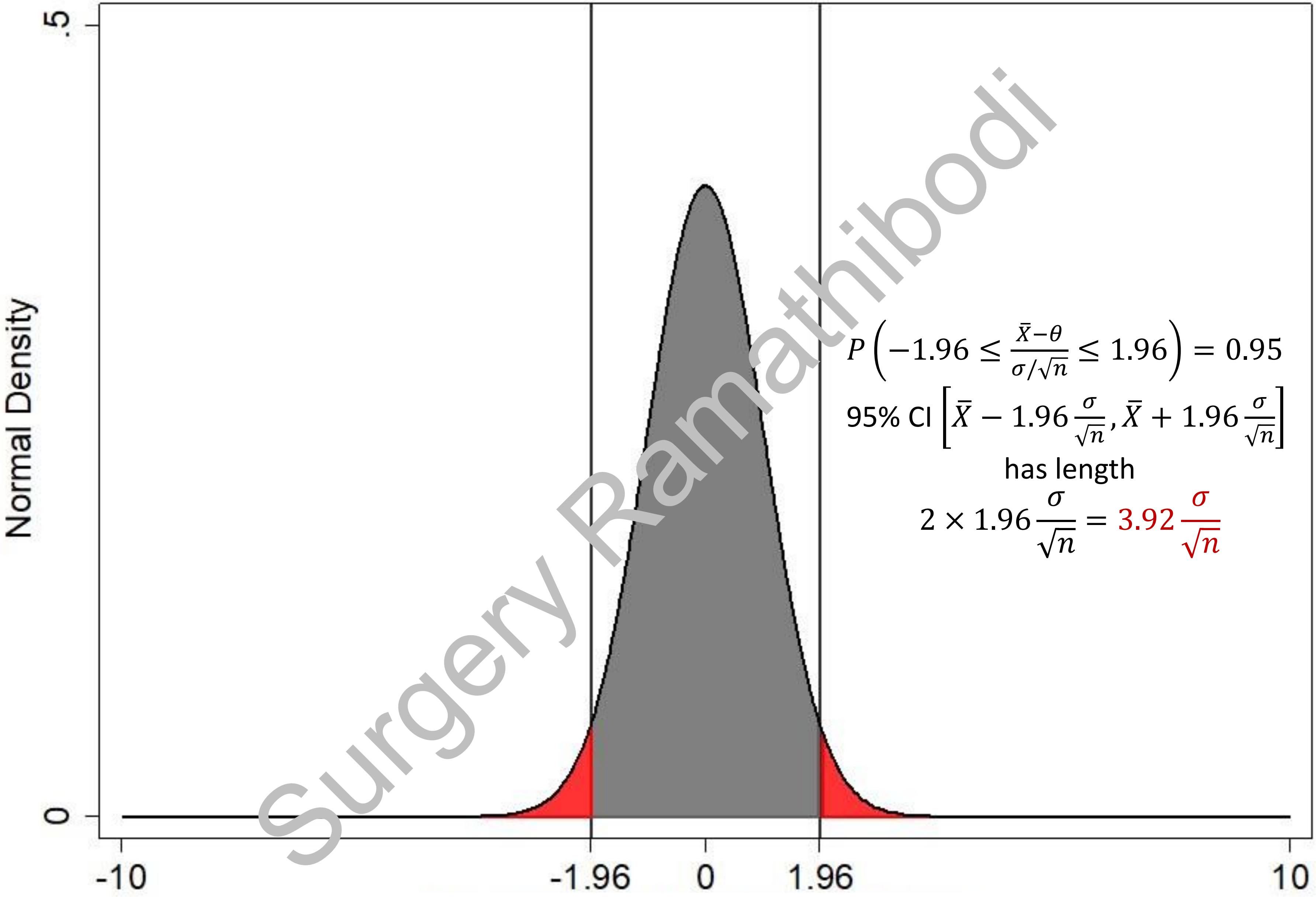
- $[\log HR - 1.96se(\log HR), \log HR + 1.96se(\log HR)]$

Or, in the HR scale

- $\left[HRe^{-1.96se(\log HR)}, HRe^{1.96se(\log HR)}\right]$
- [0.40, 1.40] (interval estimate)
- With HR = 0.75 (point estimate)

“Optimal” Interval Estimate?

- There are many interval estimates for a given problem
- Which one is the “best” in some sense?



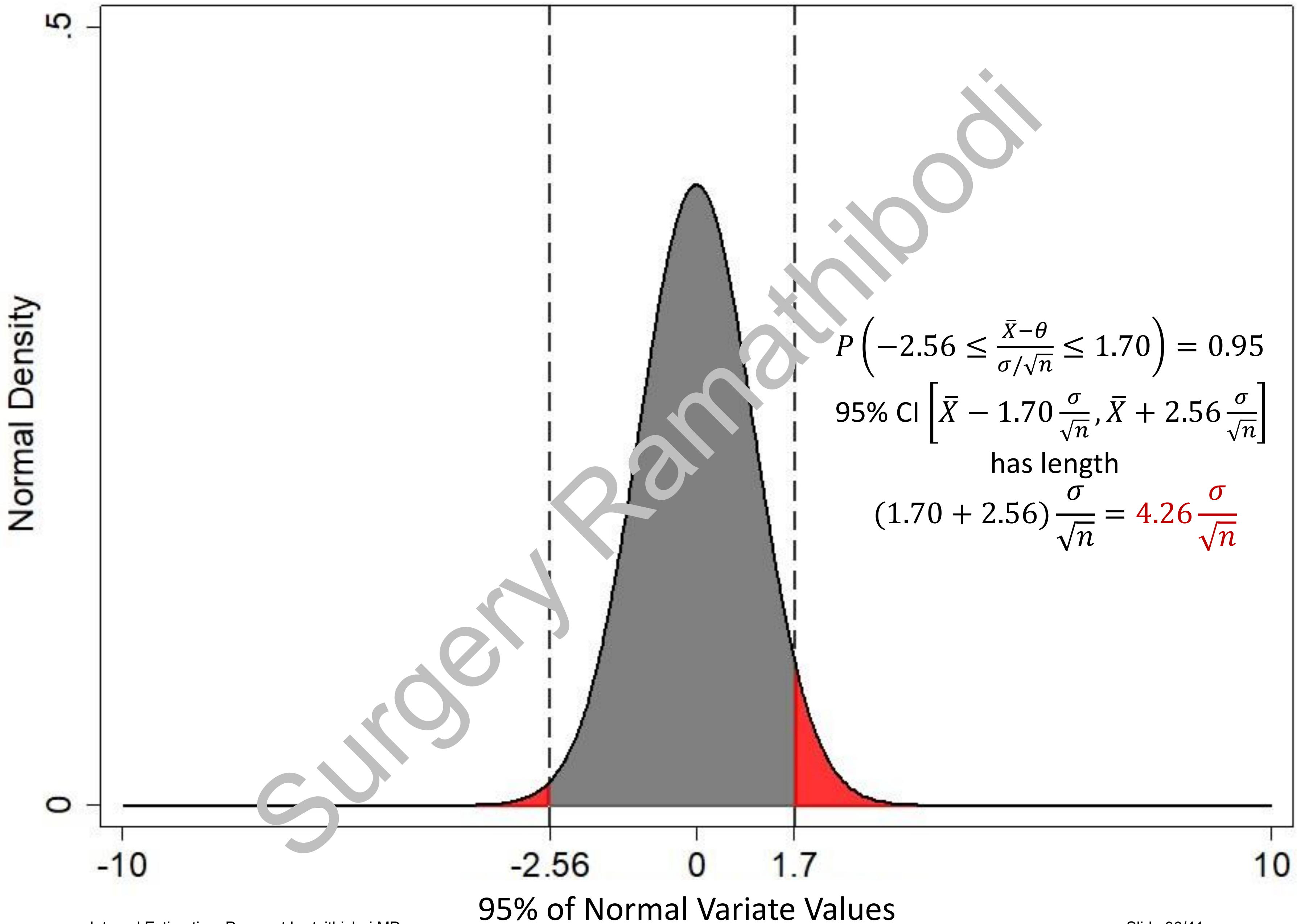
$$P\left(-1.96 \le \frac{\bar{X}-\theta}{\sigma/\sqrt{n}} \le 1.96\right) = 0.95$$

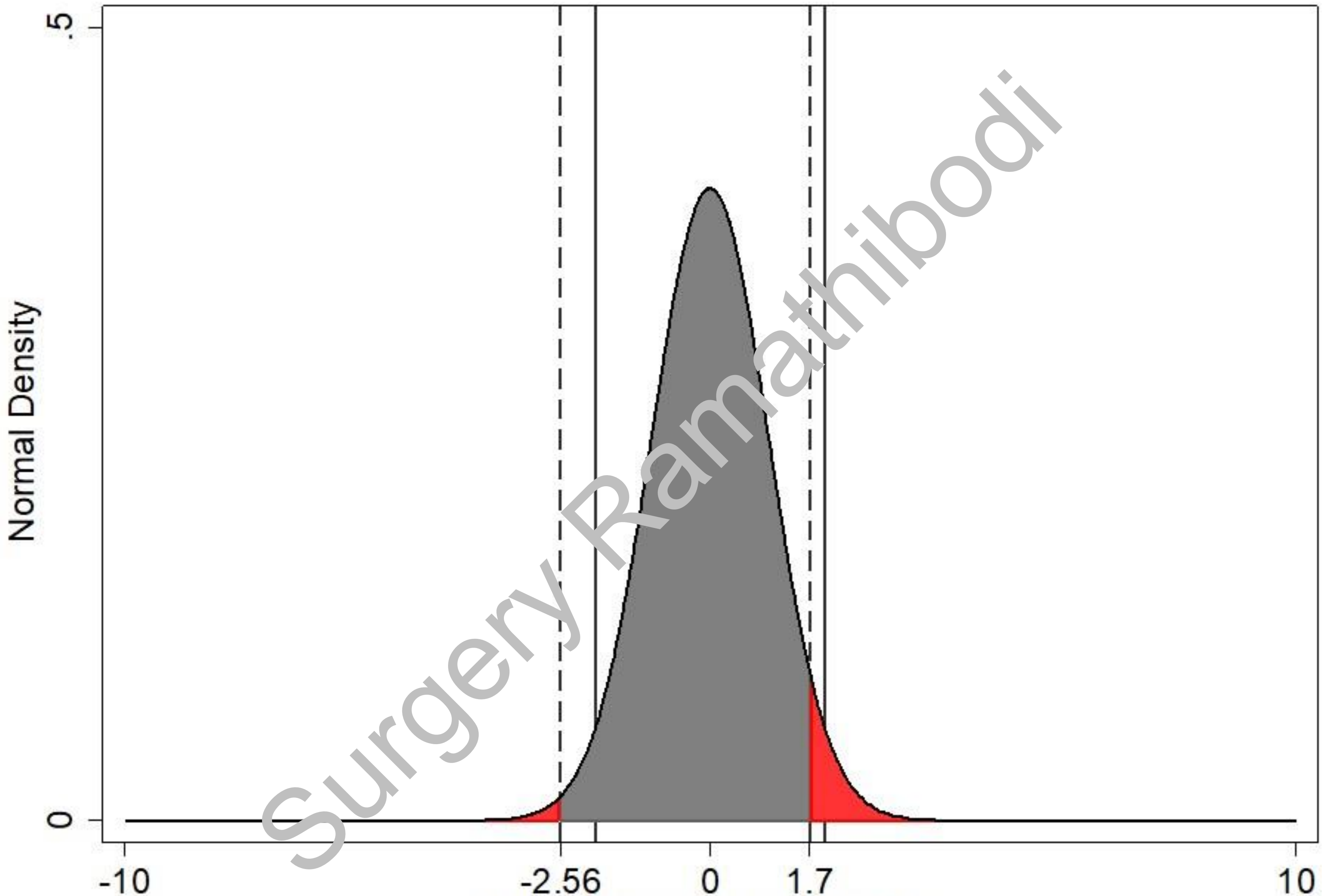
95% CI $\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$

has length

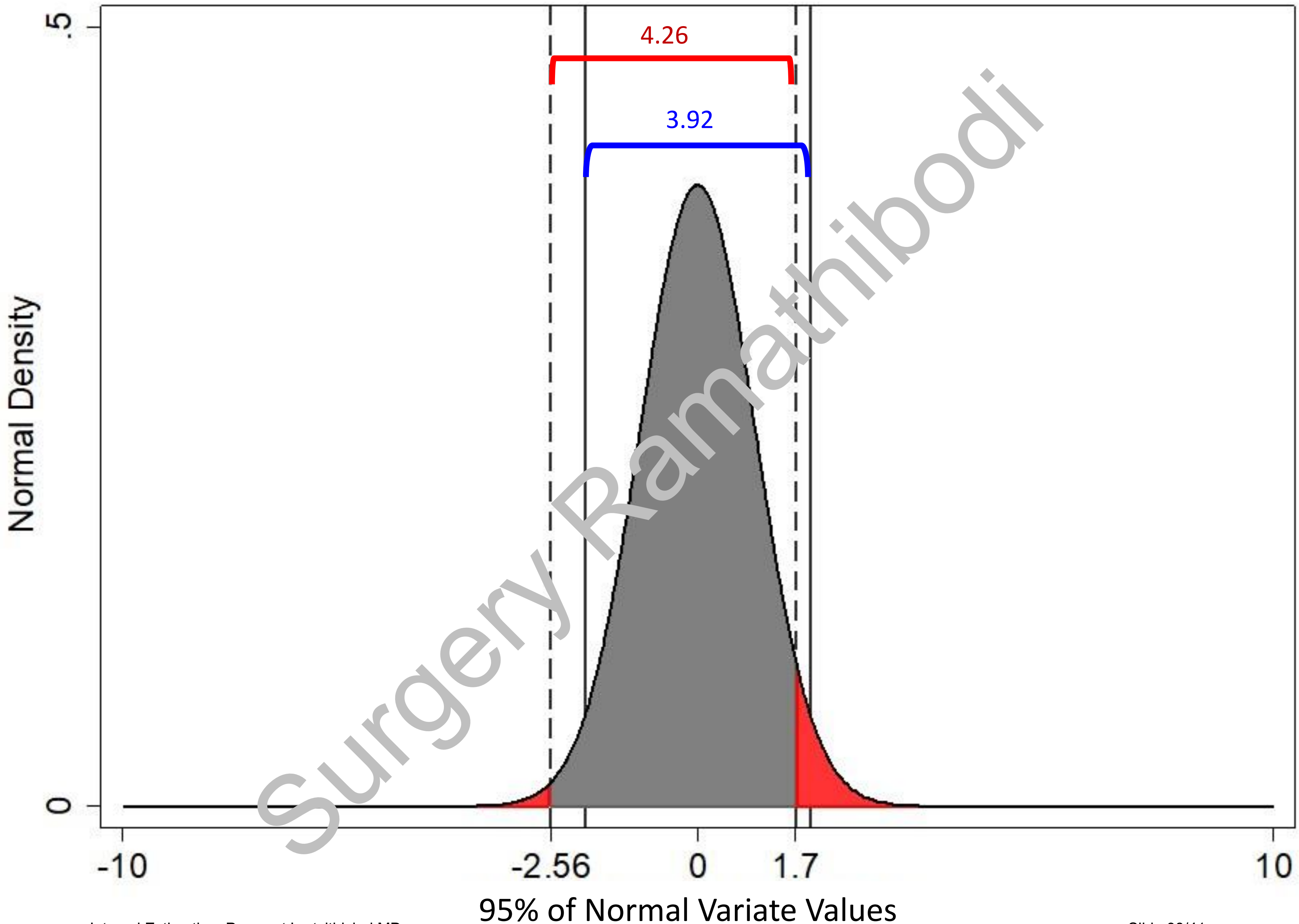
$$2 \times 1.96\frac{\sigma}{\sqrt{n}} = 3.92\frac{\sigma}{\sqrt{n}}$$

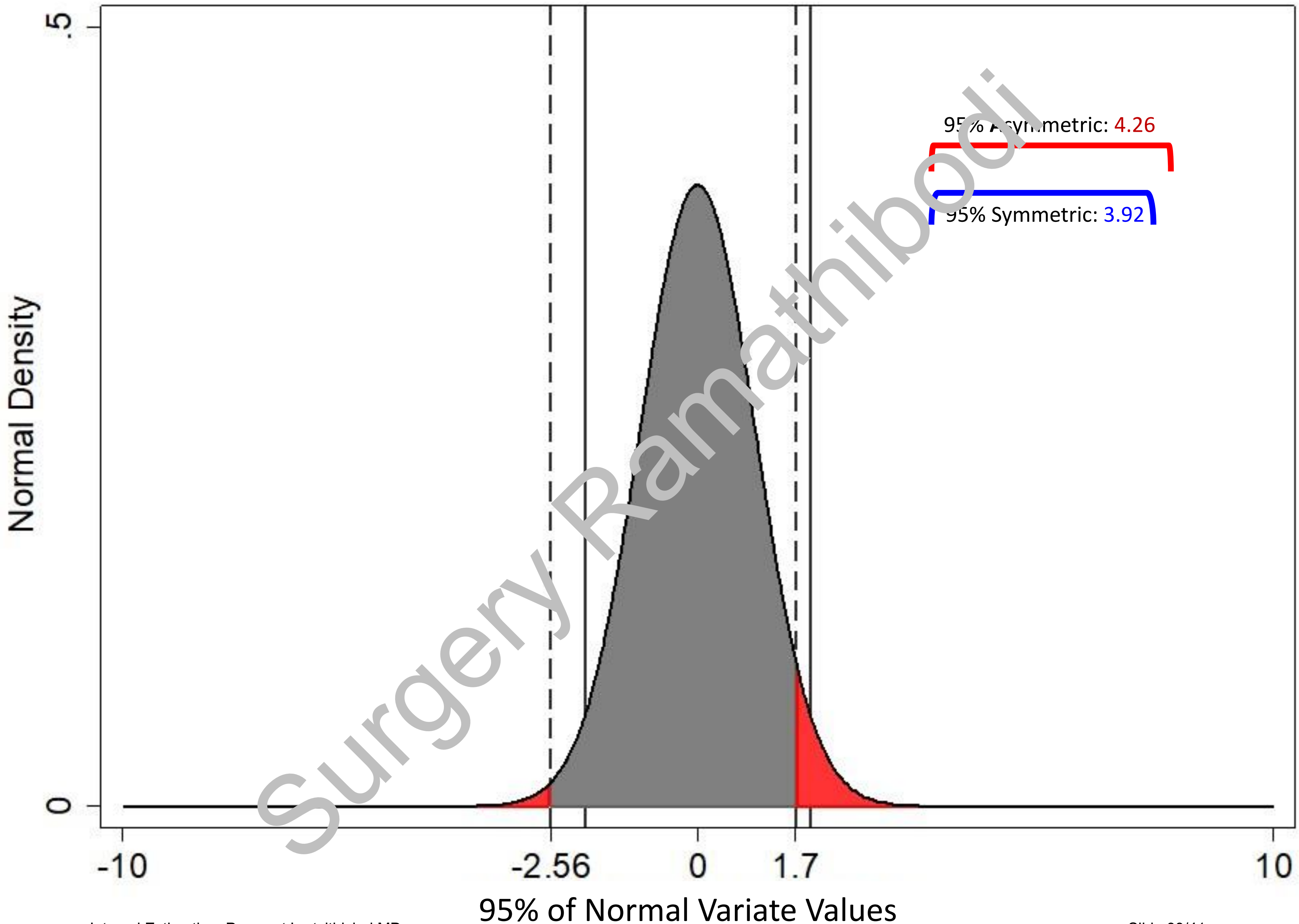
95% of Normal Variate Values


.5
0
Normal Density
-10
-2.56
0
1.7
10
95% of Normal Variate Values
$P\left(-2.56 \le \frac{\bar{X}-\theta}{\sigma/\sqrt{n}} \le 1.70\right) = 0.95$
95% CI $\left[\bar{X} - 1.70\frac{\sigma}{\sqrt{n}}, \bar{X} + 2.56\frac{\sigma}{\sqrt{n}}\right]$
has length
$(1.70 + 2.56)\frac{\sigma}{\sqrt{n}} = 4.26\frac{\sigma}{\sqrt{n}}$



95% of Normal Variate Values


4.26
3.92
.5
0
Normal Density
-10
-2.56
0
1.7
10
95% of Normal Variate Values



.5
0
Normal Density
95% Asymmetric: 4.26
95% Symmetric: 3.92
-10
-2.56
0
1.7
10
95% of Normal Variate Values

Is There an “Optimum” CI?

- Shortest true coverage probability?
- Minimizing *false* coverage probability? (e.g. Uniformly Most Accurate – UMA confidence interval) – by **inverting (UMP) test statistics**
- Etc.
- For certain situations and from certain points of view – Yes! – But there may be many and different “optimum” intervals!

Casella & Berger. Statistical Inference, 2nd ed. 2002

Some Notes (after lecture)

- A **realized confidence interval** is a collection (set) of *parameter values* within the realized boundaries $\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$
- This is the *Confidence Set** $C(\boldsymbol{x})$ in **parameter space** for the realized $\boldsymbol{x}$
- This either does or does not contain θ
- On the other hand we have the acceptance region $A(\theta)$, where, for each θ, a set of $\boldsymbol{x}$'s are considered "compatible"

*Casella & Berger. Statistical Inference, 2nd ed. 2002