ECO 602 Analysis of Environmental Data

FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Today’s Agenda

1. Building up to confidence Intervals
2. Group activity/quiz
3. Confidence intervals
4. Hypothesis testing and confidence intervals revisited

Confidence Intervals: Key Terminology

1. Sampling distribution
2. Distribution of sample
3. Sample standard deviation
4. Standard error
5. Standard error of the mean
6. Alpha, beta, statistical power

Let’s review parameters and statistics!

In the frequentist paradigm:

- A parameter is?
- A statistic is?

Let’s review parameters and statistics!

In the frequentist paradigm:

- A parameter is?
- A statistic is?

Key distinction: one is an unknowable property of a **population**, the other is a known quantity we calculate from a **sample**.

Sampling Distributions and the Distribution of a Sample

Sampling distribution is the theoretical distribution of a sample statistic, if we were able to take many samples.

- If we could perform 100 sampling realizations, how would the distribution of means look?

Distribution of the sample is a description of the results of one sampling realization.

The central limit theorem

The central limit theorem states that when you perform repeated, independent sampling, the distribution of the mean of your samples approaches a normal distribution regardless of the true, unknown/unknowable, distribution of the underlying population.

This is actually very profound and very weird! And also useful.

Standard Errors

A measure of the **spread of a statistic**.

Not a measure of **spread of the data** itself.

We can **exactly** calculate the SE if we know the true **population** standard deviation.

We can **estimate** the SE from a **sample**.

A standard error is a property of a sampling distribution, not the distribution of one sample.

Standard Errors

A measure of the **spread of a statistic**.

Standard errors can be calculated for any sample statistic.

Most commonly we calculate the **standard error of the mean**

It is possible to calculate a **standard error** of the **standard deviation**.

Standard error of the mean

We estimate it from a sample by dividing the sample standard deviation by the square root of the sample size.

Possible confusing trio of concepts:

- Population standard deviation
- Sample standard deviation
- Standard error of the mean

Standard deviation vs error

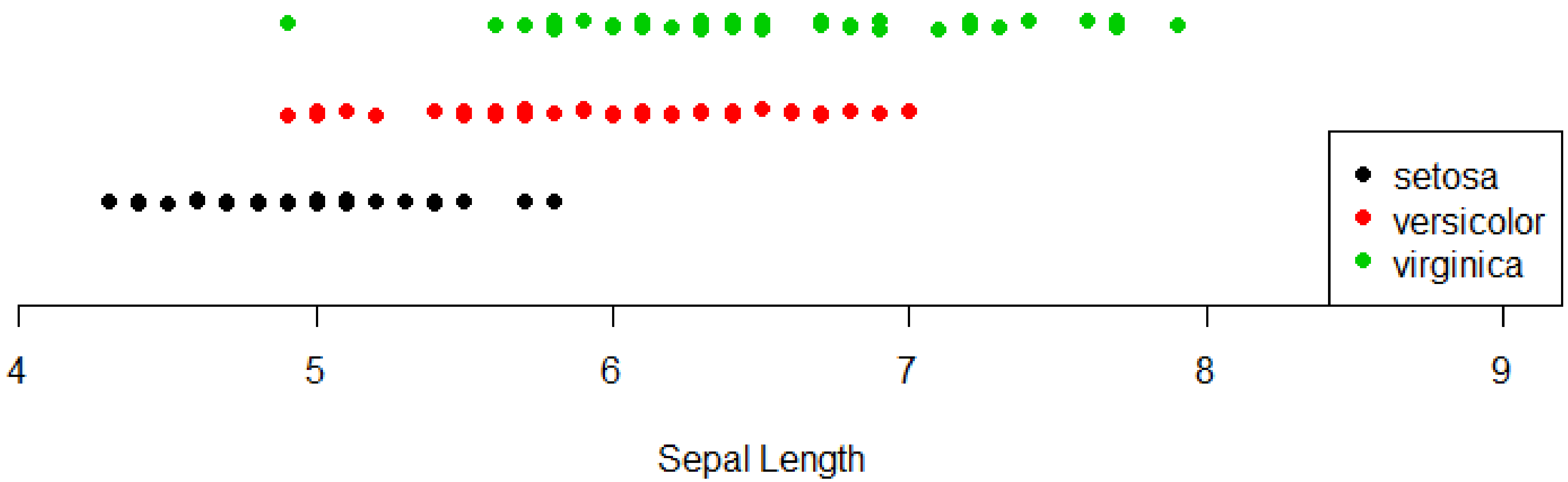
Standard deviation describes the variability in a population.

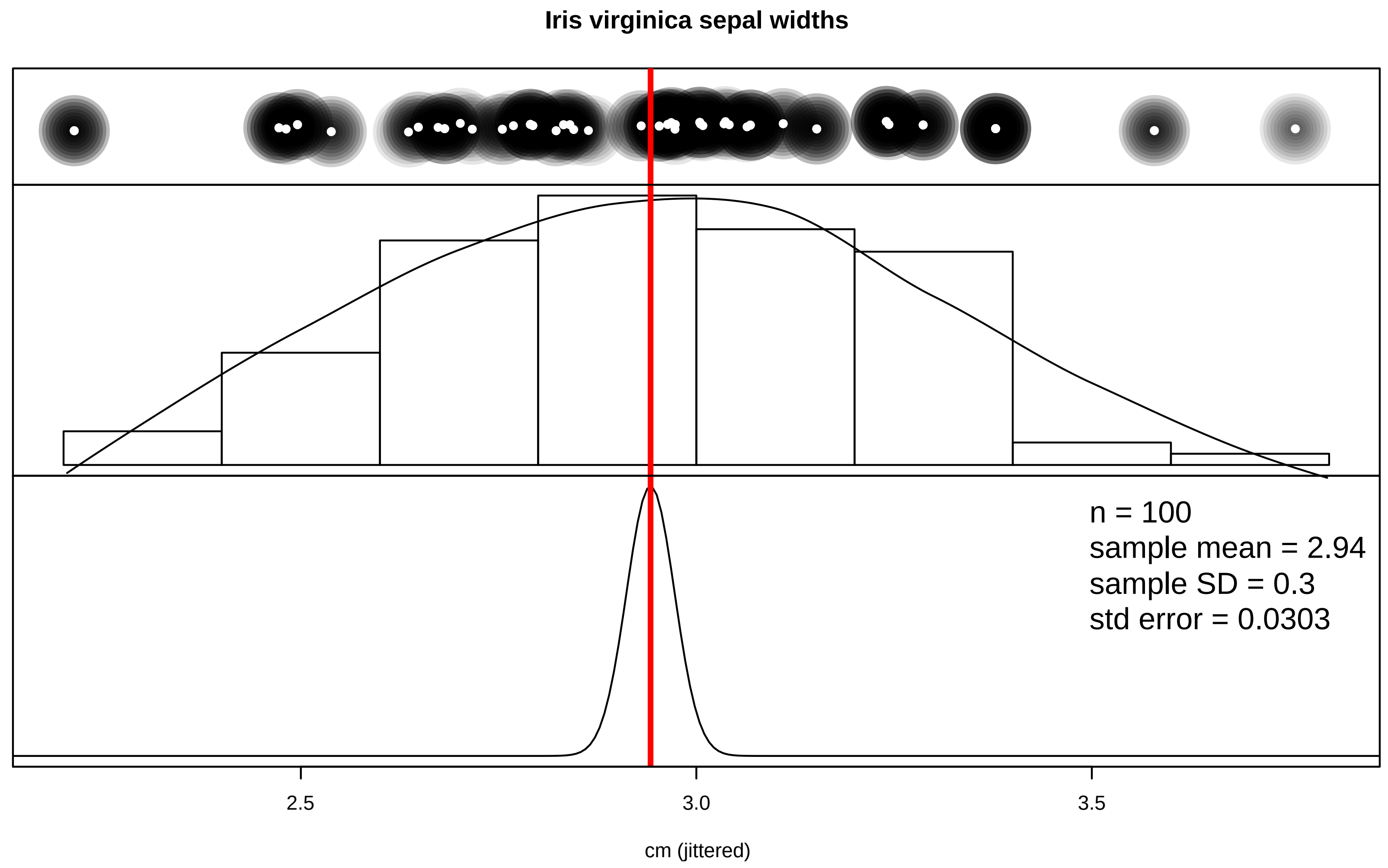
Standard error describes the variability of a statistic calculated from a sample of the population.

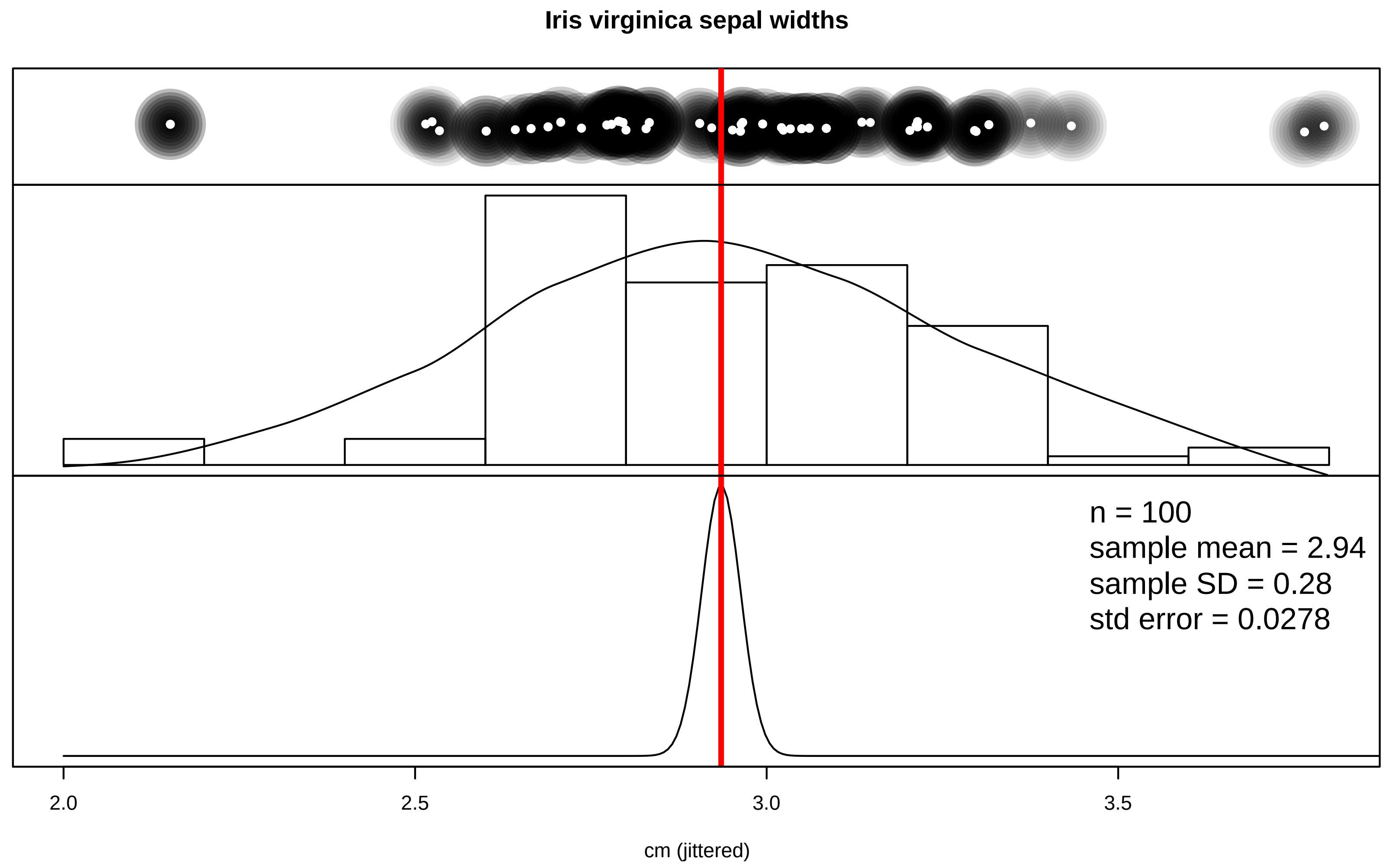
Standard errors decrease with increasing sample size.

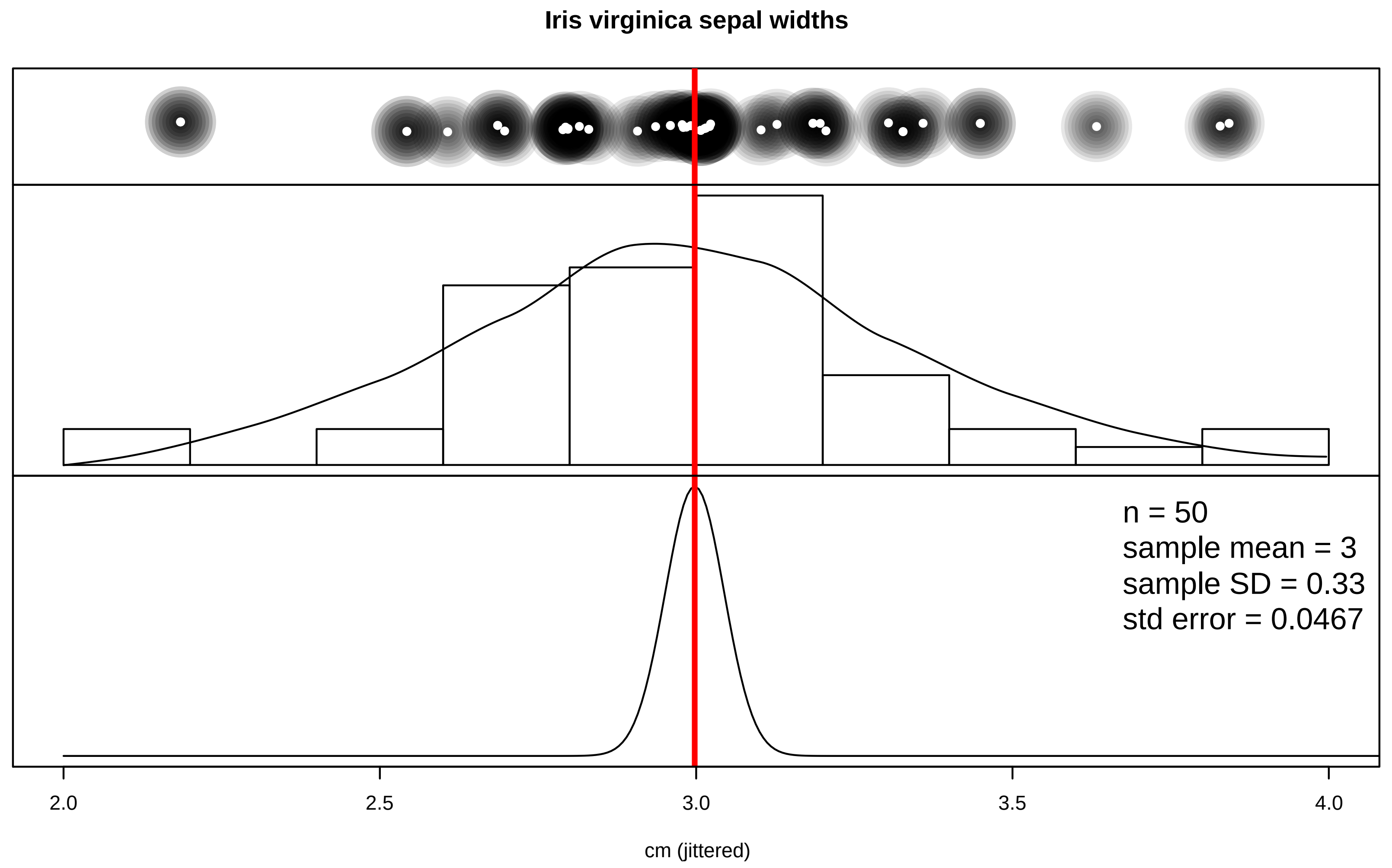
Standard error, standard deviation graphical intuition

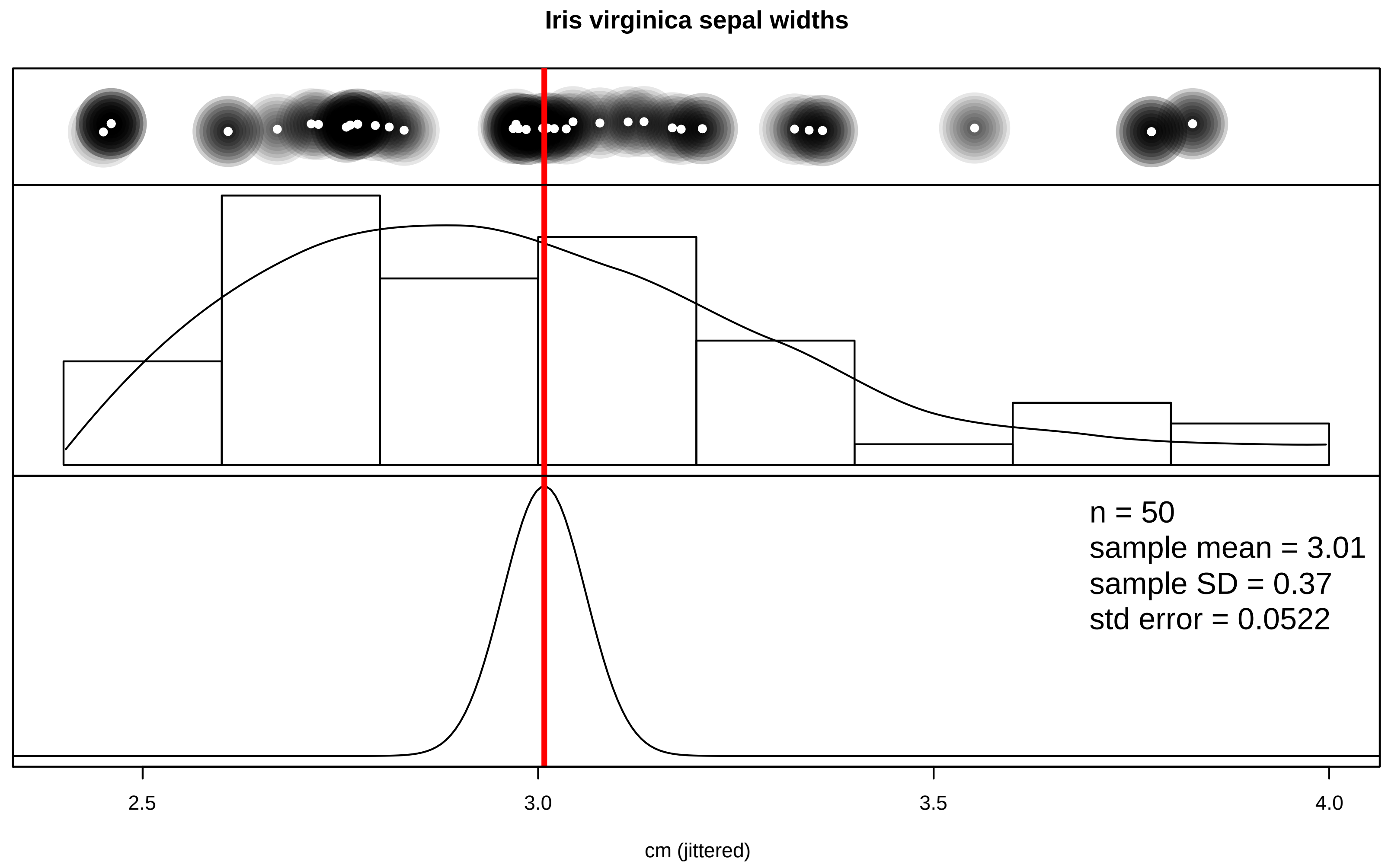
Iris sepals:

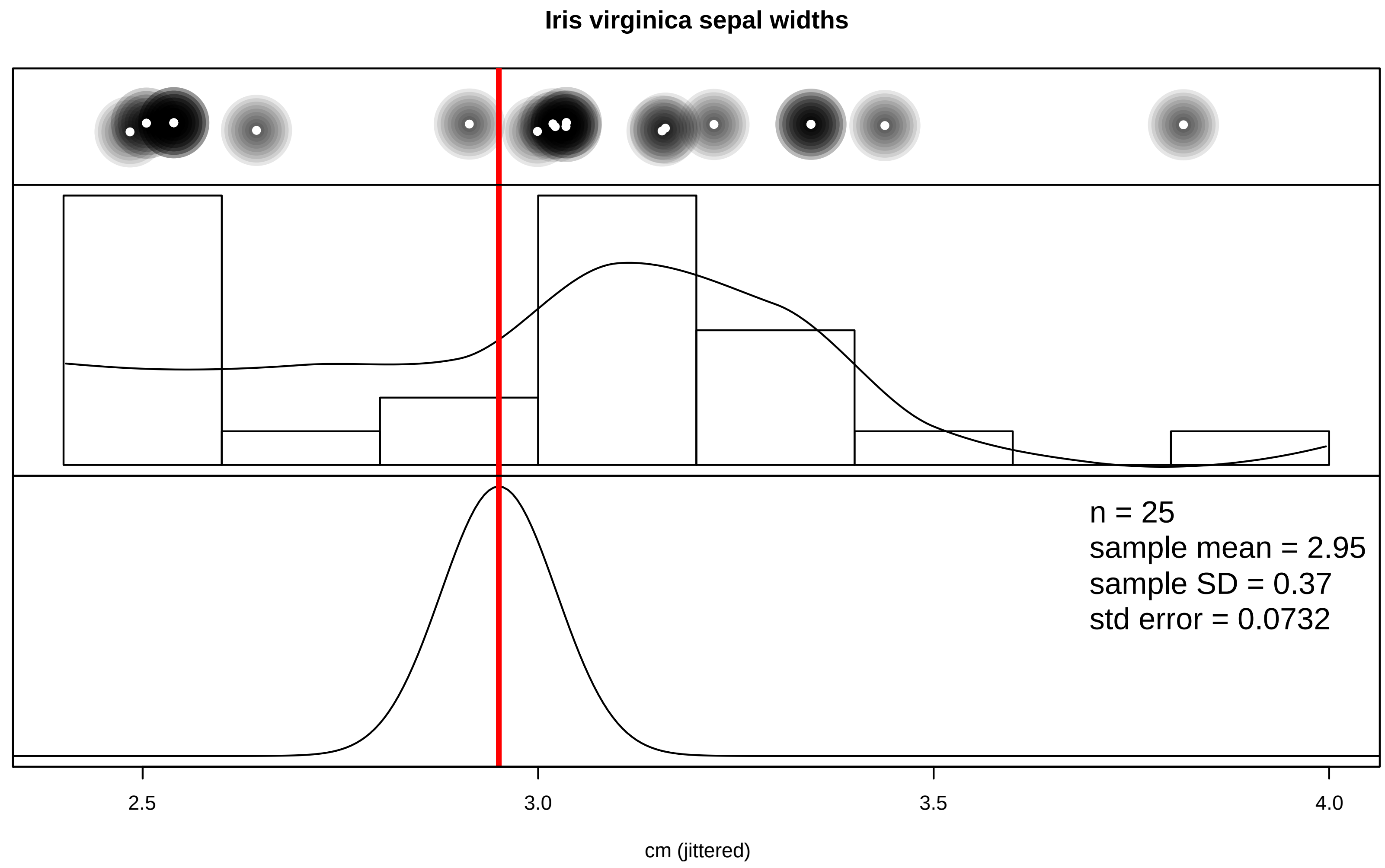


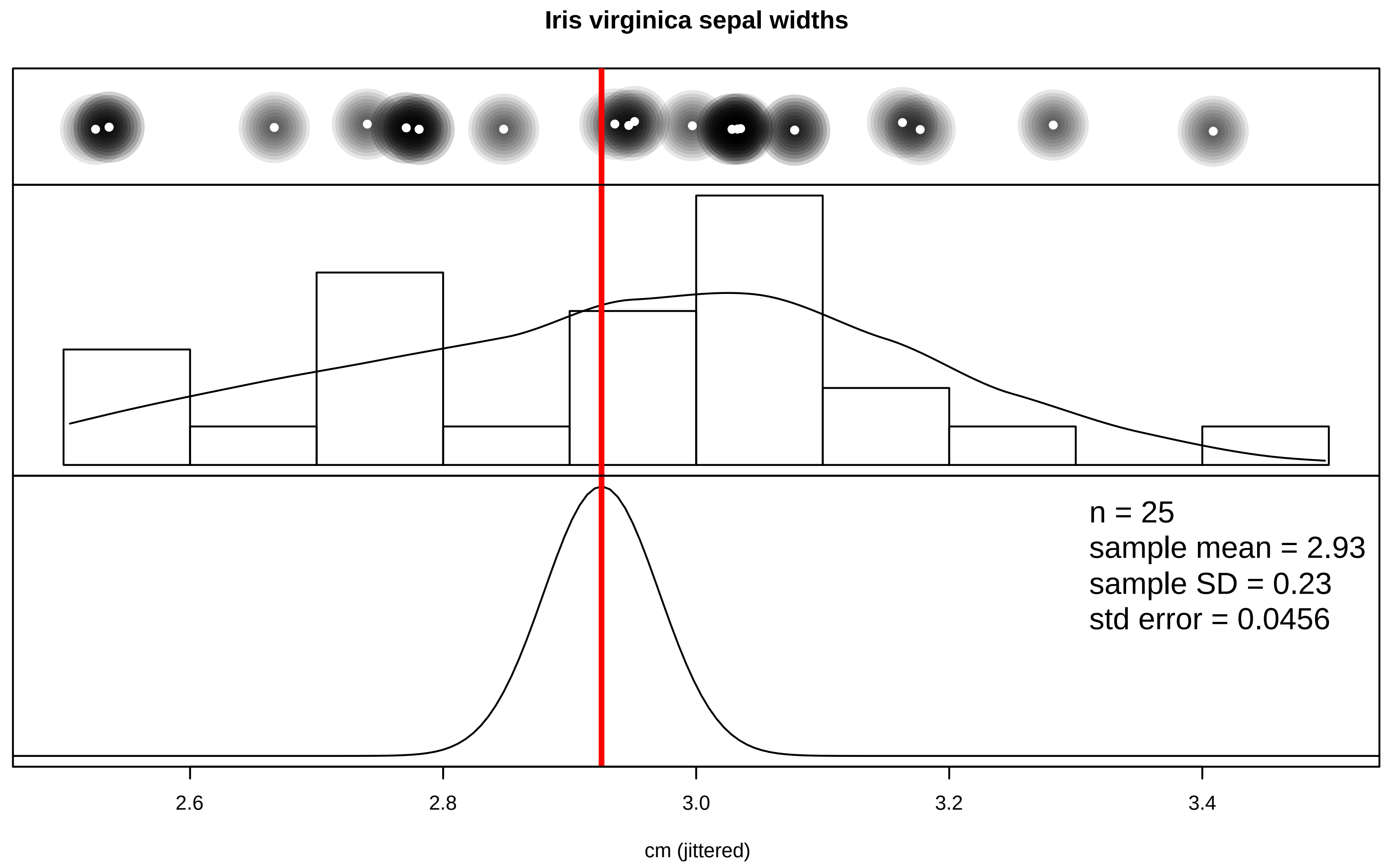

Iris virginica sepal widths
n = 100
sample mean = 2.94
sample SD = 0.3
std error = 0.0303
2.5
3.0
3.5
cm (jittered)

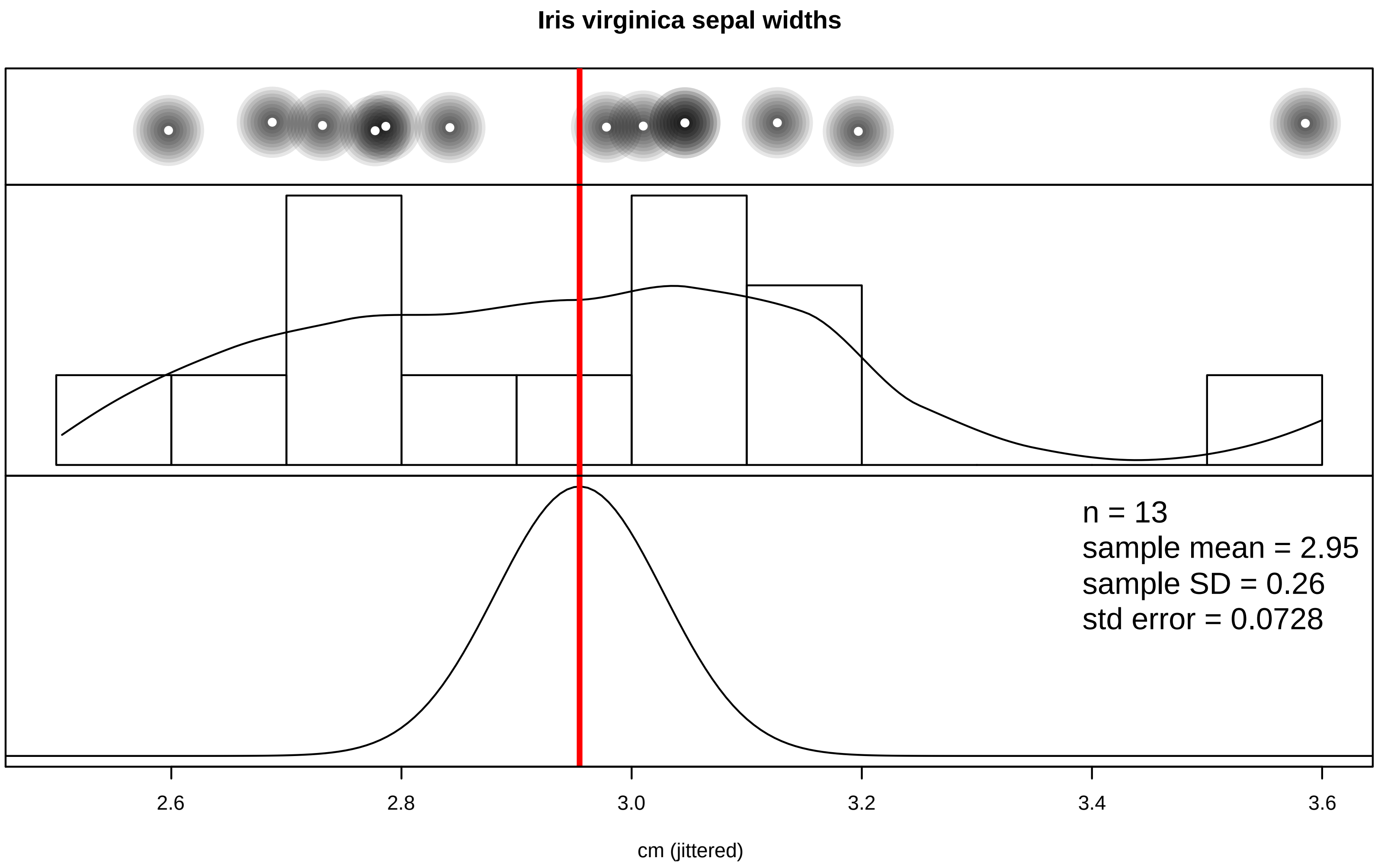

Iris virginica sepal widths
n = 100
sample mean = 2.94
sample SD = 0.28
std error = 0.0278
2.0
2.5
3.0
3.5
cm (jittered)

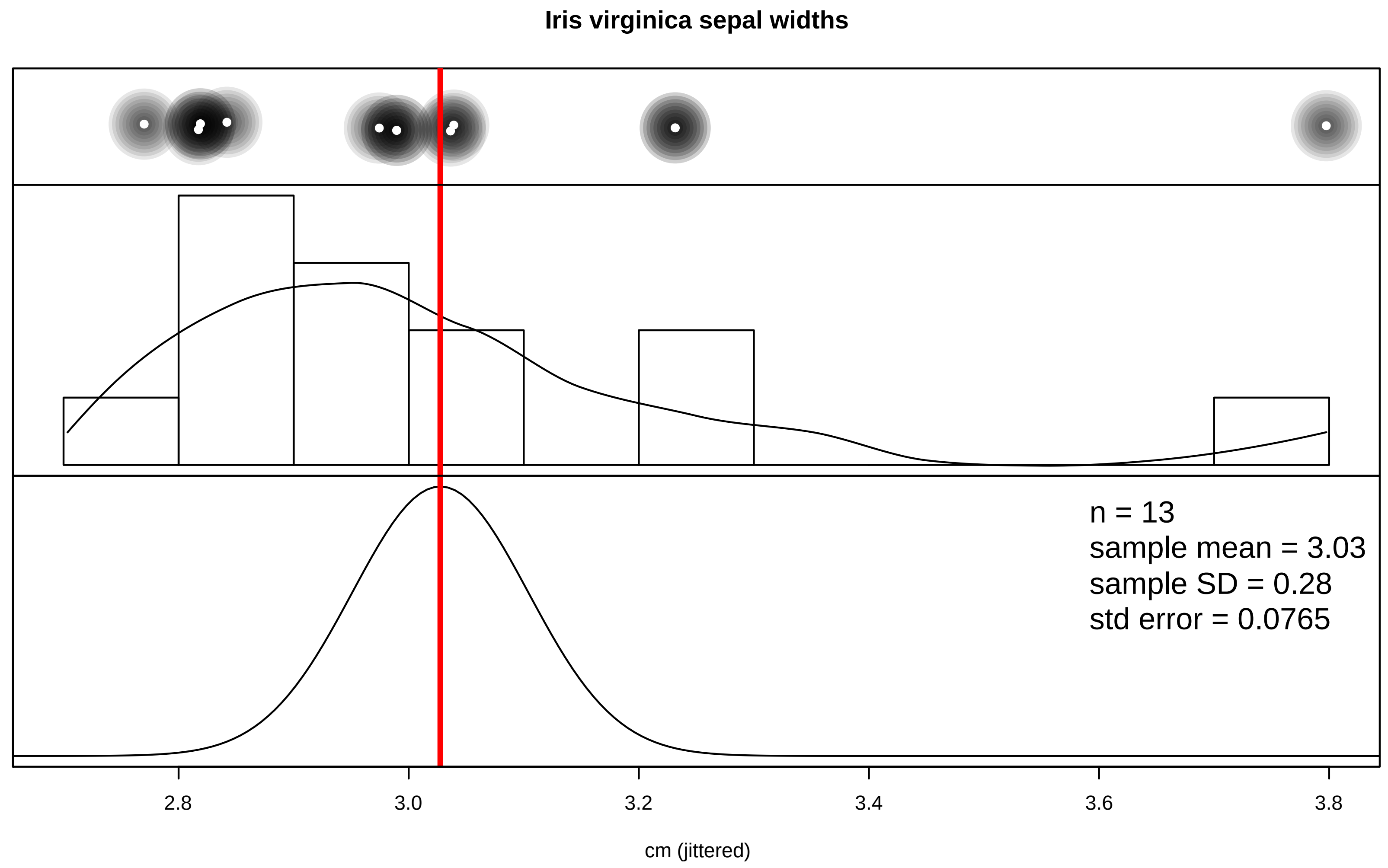

Iris virginica sepal widths
n = 50
sample mean = 3
sample SD = 0.33
std error = 0.0467
2.0
2.5
3.0
3.5
4.0
cm (jittered)


Iris virginica sepal widths
n = 50
sample mean = 3.01
sample SD = 0.37
std error = 0.0522
2.5
3.0
3.5
4.0
cm (jittered)


Iris virginica sepal widths
n = 25
sample mean = 2.95
sample SD = 0.37
std error = 0.0732
2.5
3.0
3.5
4.0
cm (jittered)


Iris virginica sepal widths
n = 25
sample mean = 2.93
sample SD = 0.23
std error = 0.0456
2.6
2.8
3.0
3.2
3.4
cm (jittered)


Iris virginica sepal widths
n = 13
sample mean = 2.95
sample SD = 0.26
std error = 0.0728
2.6
2.8
3.0
3.2
3.4
3.6
cm (jittered)


Iris virginica sepal widths
n = 13
sample mean = 3.03
sample SD = 0.28
std error = 0.0765
2.8
3.0
3.2
3.4
3.6
3.8
cm (jittered)

Things to note:

1. Standard error decreases as sample sizes increase.
2. Estimators of population parameters (sample mean, sample variance) stabilize with bigger samples.

Group activity: hypothesis testing

1. Type I and II error rates
2. Alpha, beta
3. Compare the null hypothesis distribution with the alternative hypothesis distributions.
4. In a t-test, we are comparing sampling distributions of the mean.

Group activity: hypothesis testing

Question 1: Are Iris setosa petals longer than 0 cm, on average? We observed from our sample:

- Mean petal length = 4.2 cm
- Standard error of the mean petal length is 0.2 cm.
- Critical petal length for alpha = 0.05 is 1.3 cm.

State null and alternative hypotheses for a t-test.

Sketch null and alternative sampling distribution curves, indicating the critical value, alpha, beta, and statistical power.

Group activity: hypothesis testing

Question 1: Are Iris setosa petals longer than 3.5 cm, on average? We observed from our sample:

- Mean petal length = 4.2 cm
- Standard error of the mean petal length is 0.2 cm.
- Critical petal length for alpha = 0.05 is 4.3 cm.

State null and alternative hypotheses for a t-test.

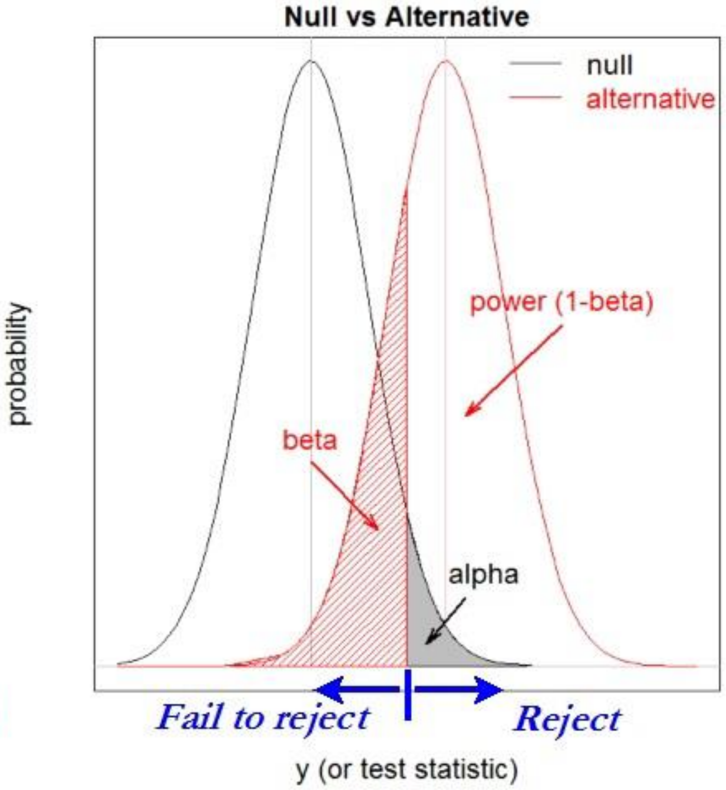
Sketch null and alternative sampling distribution curves, indicating the critical value, alpha, beta, and statistical power.

Hypothesis Testing Concepts

Neyman-Pearson decision framework

- *alpha* = probability of wrongly rejecting the null hypothesis (Type I error)
- *beta* = probability of wrongly accepting the null hypothesis (Type II error)
- *power* = probability of correctly rejecting the null hypothesis

alpha is under the null; *beta* and *power* are under the alternative



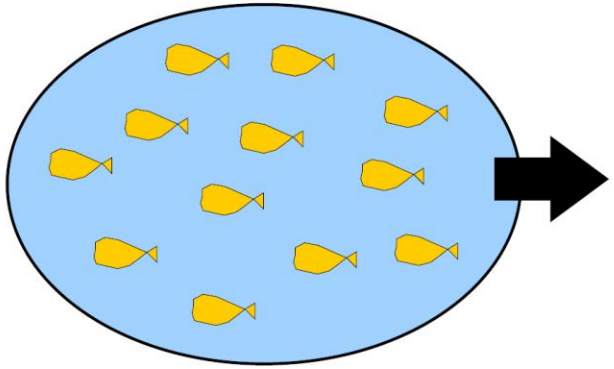
Take a picture of your quiz question answers and email them to me.

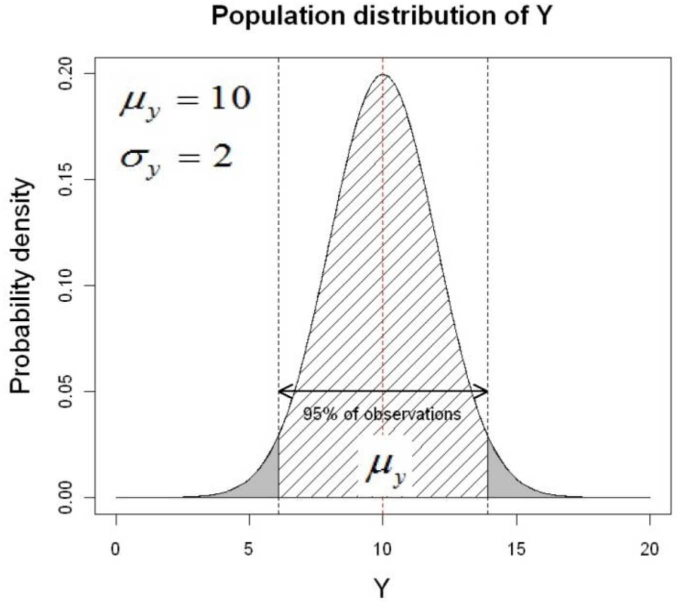
Now we’re ready to talk about confidence intervals!

Now we’re ready to talk about confidence intervals!

Primer on confidence intervals and more...

Population distribution of a random variable





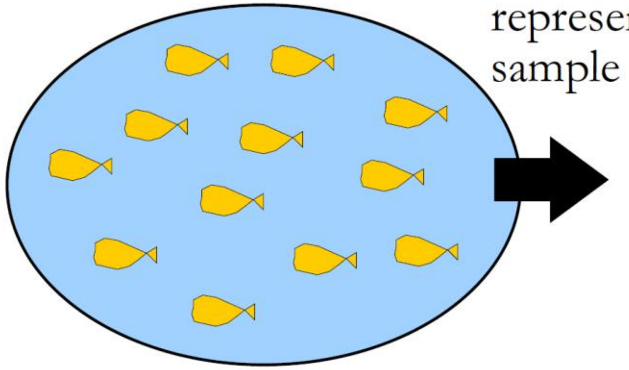
$$\Pr\left\{\mu_y - 1.96\sigma_y \leq Y \leq \mu_y + 1.96\sigma_y\right\} = 0.95$$

This is <u>not</u> a confidence interval!

Primer on confidence intervals and more...

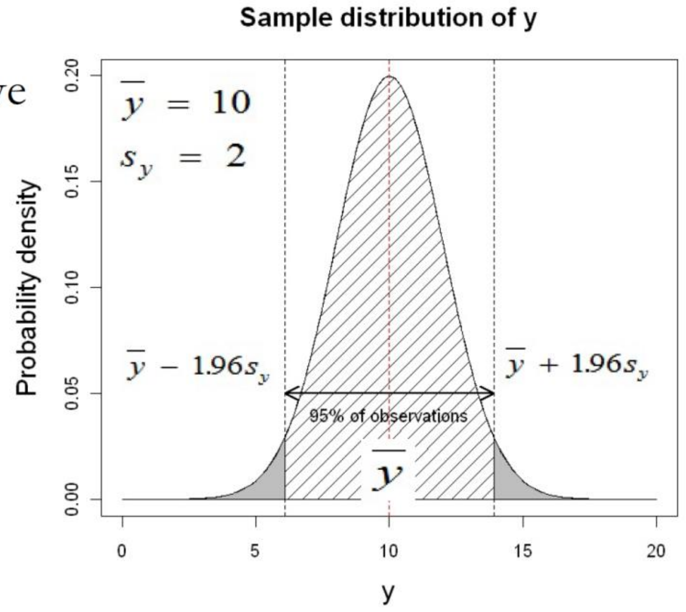
Sample distribution of a random variable

Sample of fish



A perfectly representative sample

Y = fish size



$$\Pr\left\{\bar{y} - 1.96 s_y \le y \le \bar{y} + 1.96 s_y\right\} = 0.95$$

This is not a confidence interval!

Primer on confidence intervals and more…

Confidence interval for the sample estimate of population parameter

Confidence interval for the mean:

- Convert the distribution of *sample means* into a standard normal distribution via the *z*-score standardization

σ_y = population standard deviation

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} \qquad z = \frac{\bar{y} - \mu_y}{\sigma_{\bar{y}}}$$

$$\Pr\left\{\bar{y} - 1.96\sigma_{\bar{y}} \leq \mu_y \leq \bar{y} + 1.96\sigma_{\bar{y}}\right\} = 0.95$$

This <u>is</u> a confidence interval!

Was that a letdown?

- Confidence interval: The ‘confidence’ refers to the interval, not the population parameter
- The width of the confidence interval depends on:
  - Alpha
  - Population variability
  - Sample size

Confidence interval is a very frequentist concept.

- Based on hypothetical repeated sampling.
- With alpha = 0.05:
- “If we repeated our sampling scheme many times, around 95% of our confidence intervals would bracket the true population mean.”

Confidence interval is a very frequentist concept.

- We can’t say that we are 95% sure a particular CI contains the true population mean.
- A CI either contains the true mean, or it doesn’t… But we cannot tell a particular CI because the true population mean is unknowable.

T-test null and alternative hypotheses

- 1-sample:
- 2-sample:
- Using iris data:
  - 3 species: setosa, virginica, versicolor
  - What are some possible 1-sample hypotheses?
  - What are some possible 2-sample hypotheses?

For next time:

Read McGarigal chapter 8.

Read the assignment 3 questions. This set is very technical, and we'll most likely need to spend some class time working through them.