

Hypothesis Testing Framework

- 1) Setup the Null and Alternate Hypothesis
- 2) Choose the right test statistic
- 3) Left-tailed Vs Right-tailed Vs Two-tailed
- 4) Compute p-value
- 5) If p-value is less than alpha, then reject the null hypothesis

Recap central limit theorem

Average height is 65 inches with std dev 2.5

We take a sample of 50 people

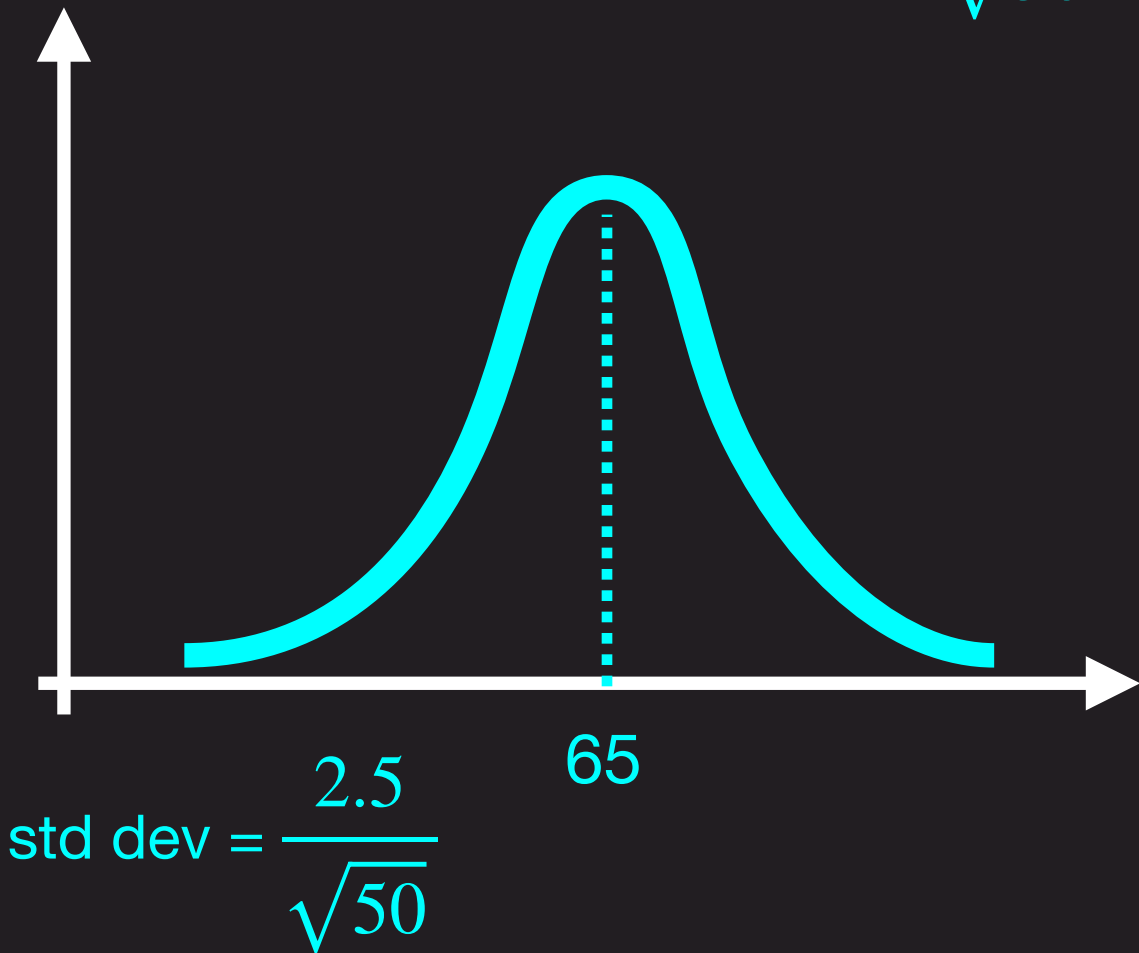
Let m represent the sample mean

Is m a random variable? Yes

What is its distribution? Gaussian (From CLT)

What is $E[m]$? 65

What is the std dev of m ? $\frac{2.5}{\sqrt{50}}$



We take a sample of 5 people

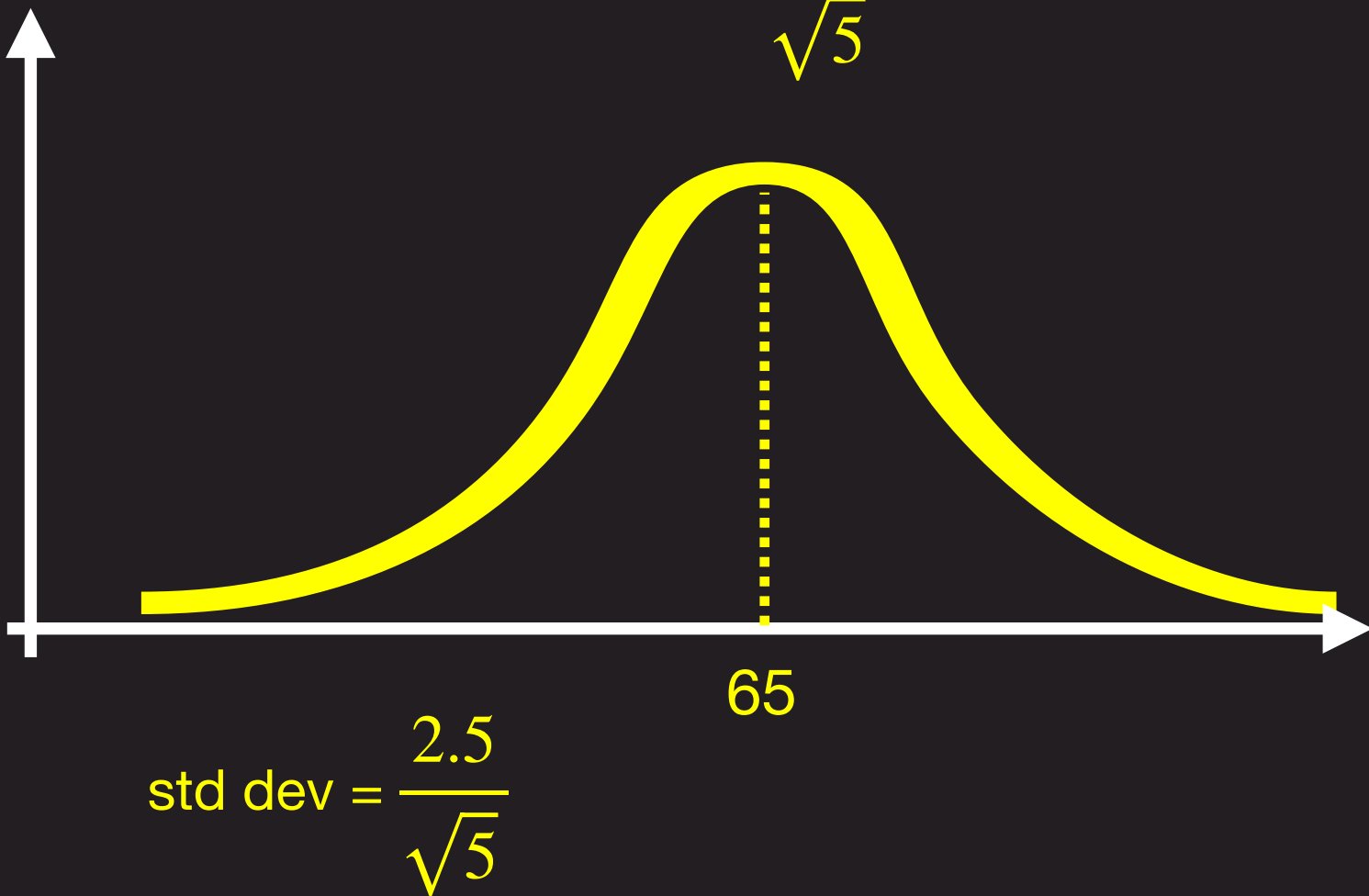
Let m represent the sample mean

Is m a random variable? Yes


What is its distribution? Gaussian (From CLT)

What is $E[m]$? 65

What is the std dev of m ? $\frac{2.5}{\sqrt{5}}$



Supply chain example


$$\mu = 1800$$
$$\sigma = 100$$

A retailer has 2000 stores in the country

Historical data tells us that weekly sales of shampoo bottles has an average of 1800, with a standard deviation of 100

Sales team wants to improve sales by hiring a marketing team

Hiring a marketing team can be expensive, so we need to be very sure that they will improve sales

Before deploying their strategy for all 2000 stores, they are tested in 50 stores

On the 50 stores, their average sales for that week was 1850

You are the data scientist who should tell your sales team whether this is statistically significant

Sales team has said that we will hire only if we are 99 % confident $\alpha = 0.01$

Another marketing team is also being considered

They are tested on 5 stores

On the 5 stores, their average sales for that week was 1900

Would you say this team is better than the first one?

Between the “blue team” and the “yellow team”, whom will you choose?

Supply chain example

50 stores with average of 1850

$H_0 : \mu_b = 1800$

$H_a : \mu_b > 1800$

$$m_b = \frac{x_1 + x_2 + \cdots + x_{50}}{50}$$

Is m_b a random variable?

What is its distribution?

What is $E[m_b]$?

What is the std dev of m_b ?

What is the observed value of m_b ?

Right or Left tailed?

How to compute p -value?

Is the p -value less than α ?

We reject the null hypothesis

$\alpha = 0.01$

Let x_1, x_2, \cdots, x_{50} denote the sales

m_b is the sample mean

Yes

Gaussian (From CLT)

1800

$$\frac{100}{\sqrt{50}}$$

1850

Right tailed

H_a says “greater than”

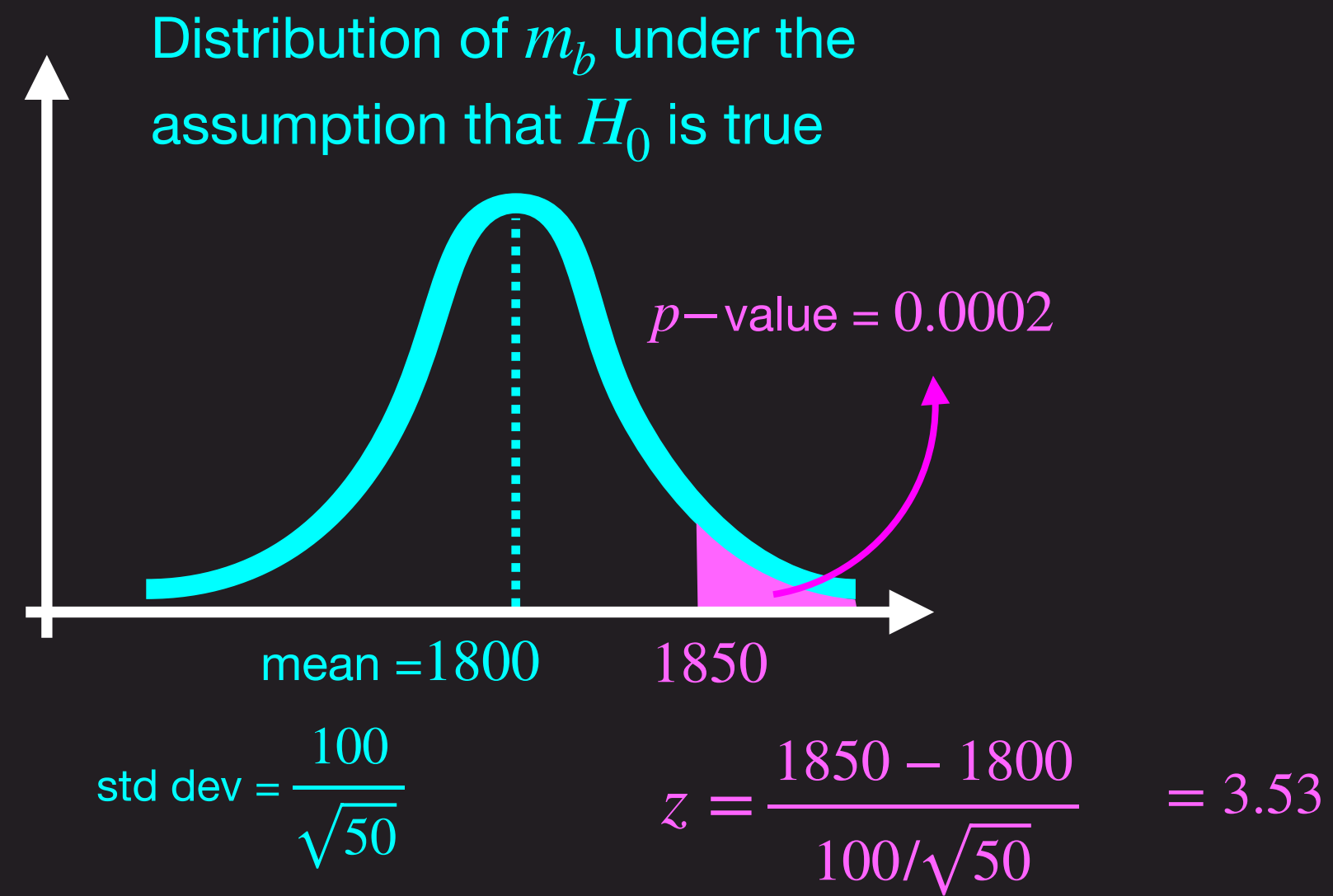
$$P \left[m_b \geq 1850 \mid H_0 \text{ is true} \right] = 1 - \text{norm.cdf}(3.53) = 0.0002$$

Yes

This means the marketing team had a positive effect on the sales



$\mu = 1800$
 $\sigma = 100$



Supply chain example

5 stores with average of 1900

$H_0 : \mu_y = 1800$

$H_a : \mu_y > 1800$

$$m_y = \frac{x_1 + x_2 + \dots + x_5}{5}$$

Is m_y a random variable?

What is its distribution?

What is $E[m_y]$?

What is the std dev of m_y ?

What is the observed value of m_y ?

Right or Left tailed?

How to compute p -value?

Is the p -value less than α ?

We fail to reject the null hypothesis

$\alpha = 0.01$

Let x_1, x_2, x_3, x_4, x_5 denote the sales

m_y is the sample mean

Yes

Gaussian (From CLT)

1800

$$\frac{100}{\sqrt{5}}$$

1900


Right tailed

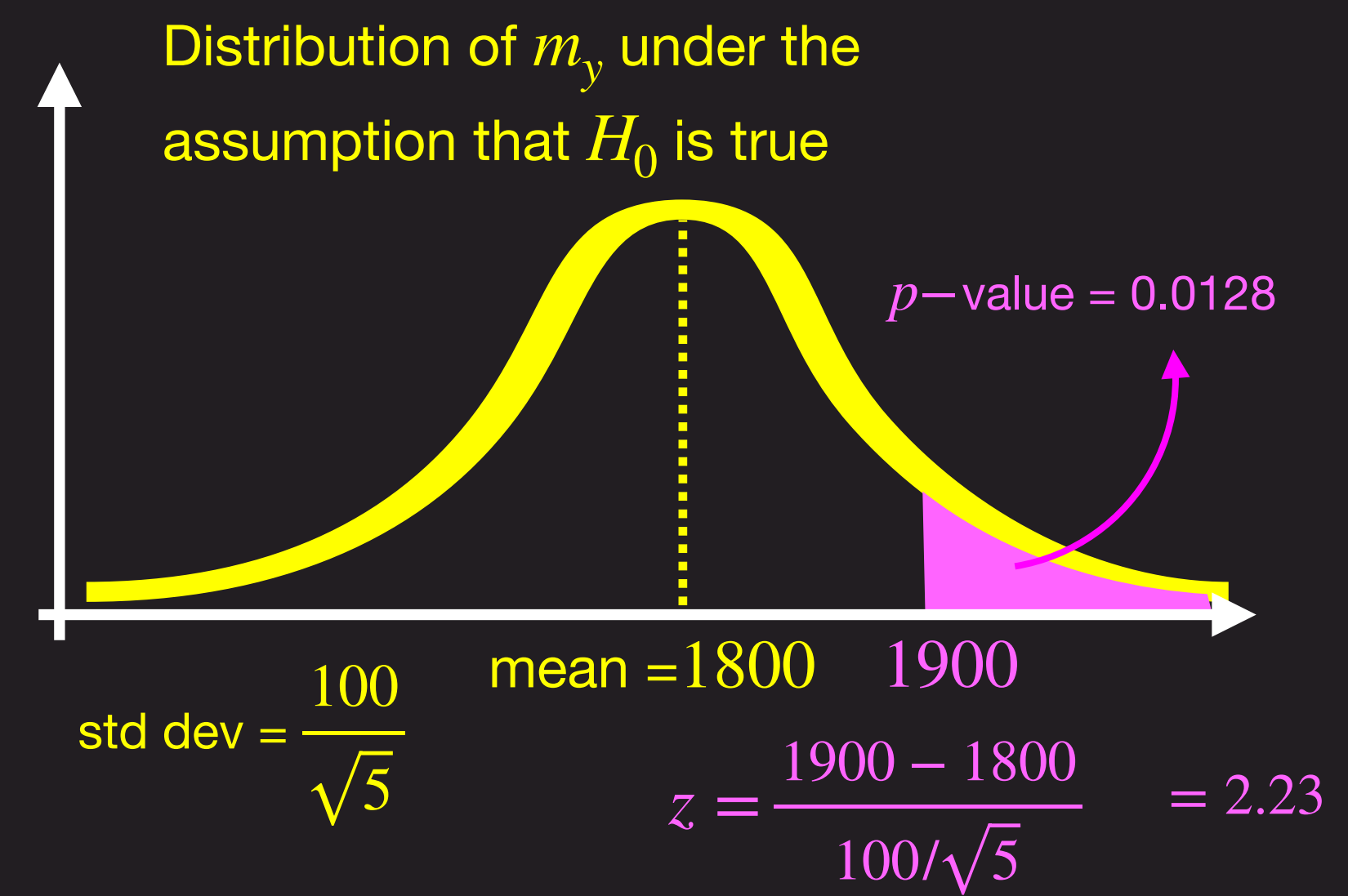
H_a says “greater than”

$$P \left[m_y \geq 1900 \mid H_0 \text{ is true} \right] = 1 - \text{norm.cdf}(2.23) = 0.0128$$

No

The effect of marketing was not statistically significant

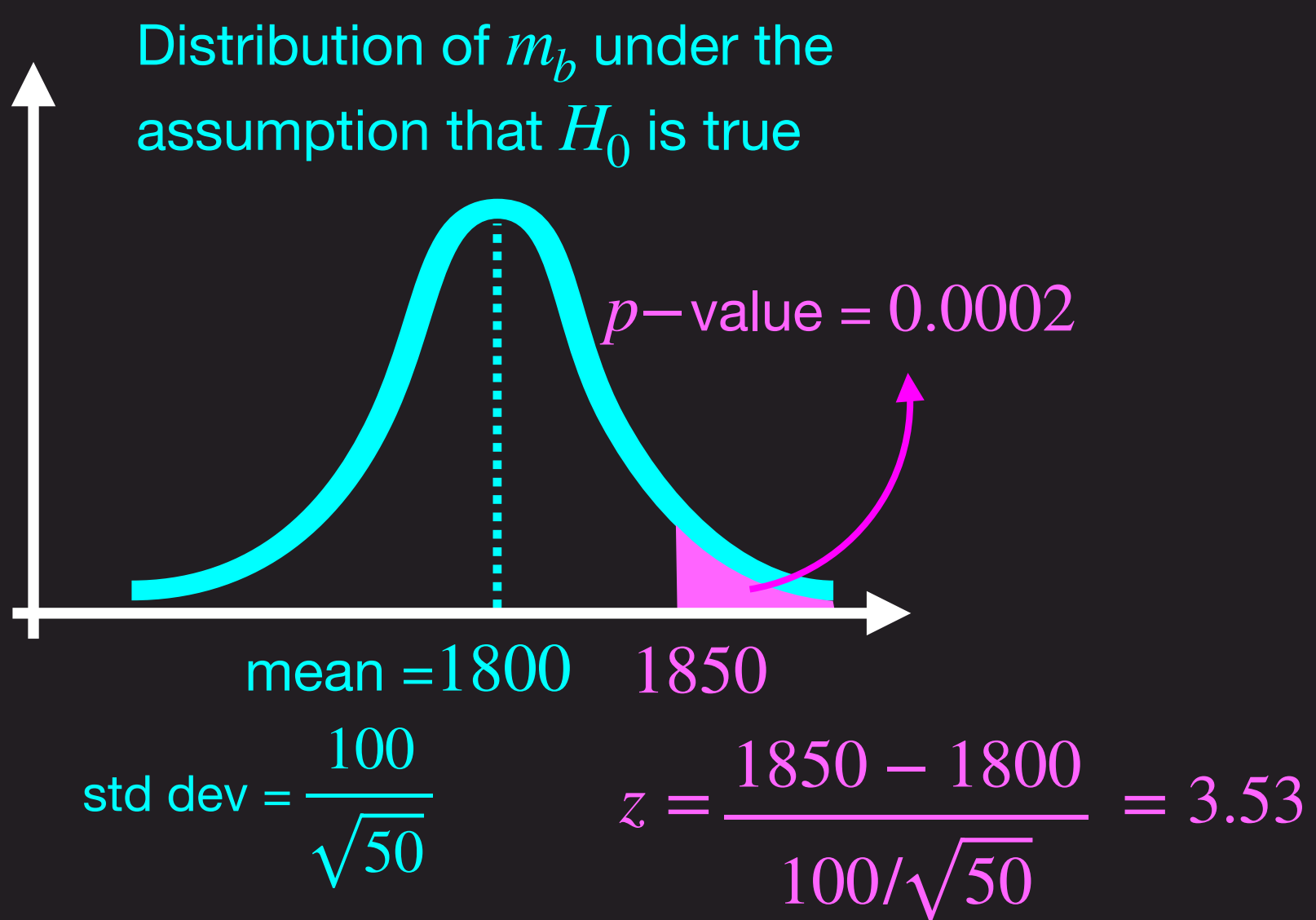
 $\mu = 1800$
 $\sigma = 100$



50 stores with average of 1850

$H_0 : \mu_b = 1800$

$H_a : \mu_b > 1800$

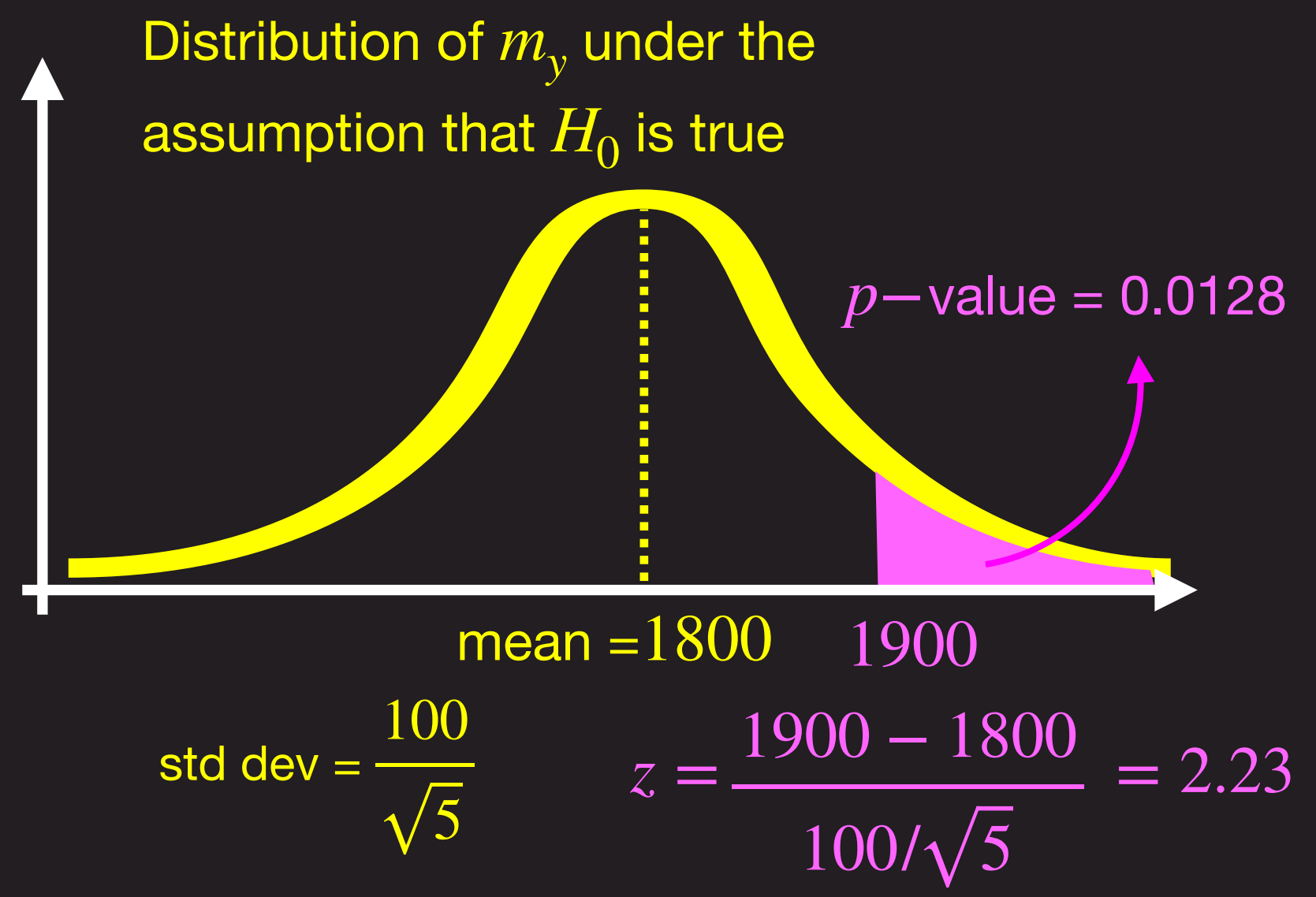


Reject H_0

5 stores with average of 1900

$H_0 : \mu_y = 1800$

$H_a : \mu_y > 1800$

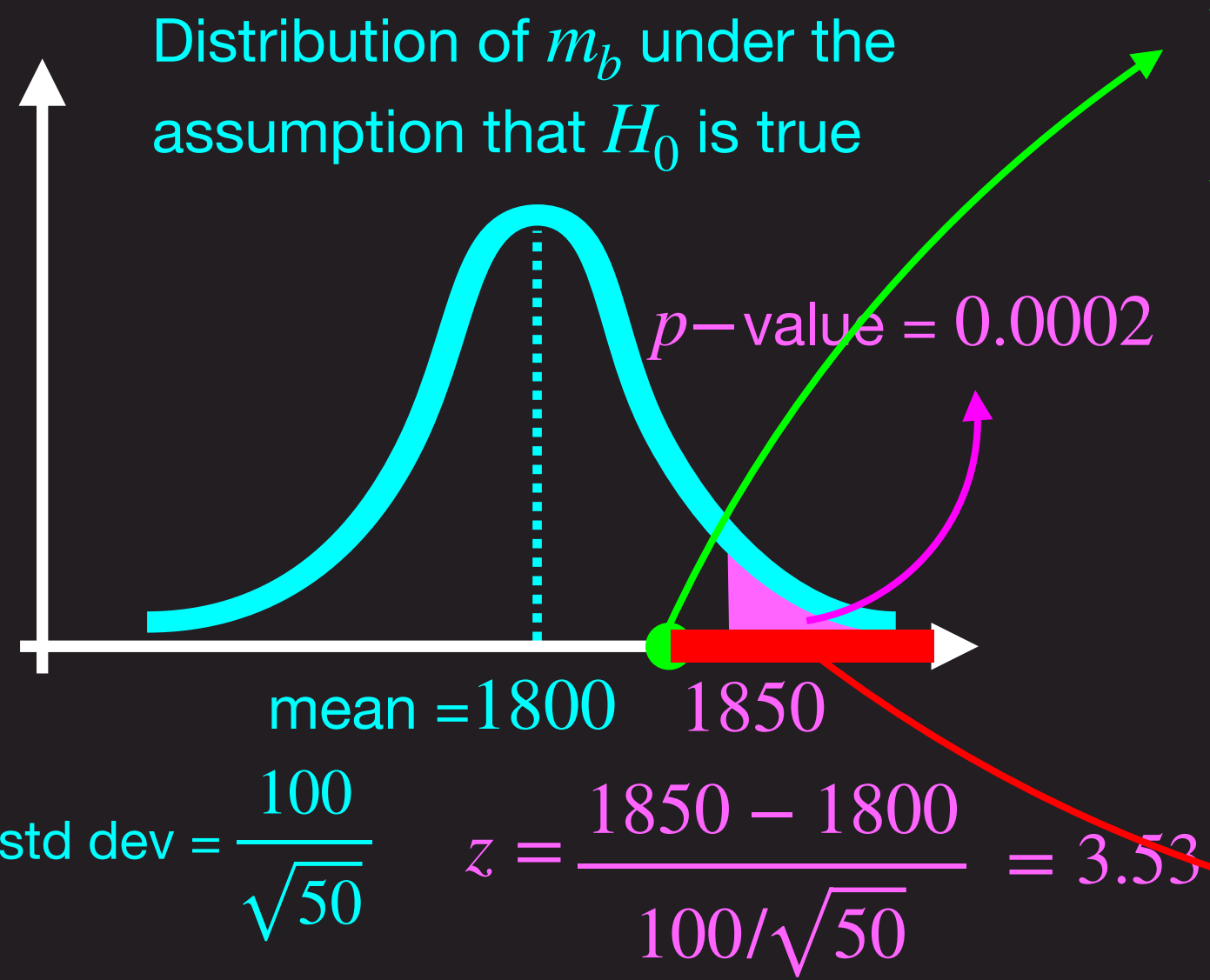


Fail to reject H_0

50 stores with average of 1850

$H_0 : \mu_b = 1800$

$H_a : \mu_b > 1800$



What should be the z-score such that we can reject if mean is larger, and accept if mean is lesser?

We want only 1% area to the right

Upper critical value = `norm.ppf(0.99)` = 2.32


$z = 2.32 \quad x = 1800 + 2.32 * \frac{100}{\sqrt{50}} = 1832.8$

To summarise, if we are testing for 50 samples, we can reject the null hypothesis only if the average sales is greater than 1832.8

This region is called the “critical region”

Supply chain example

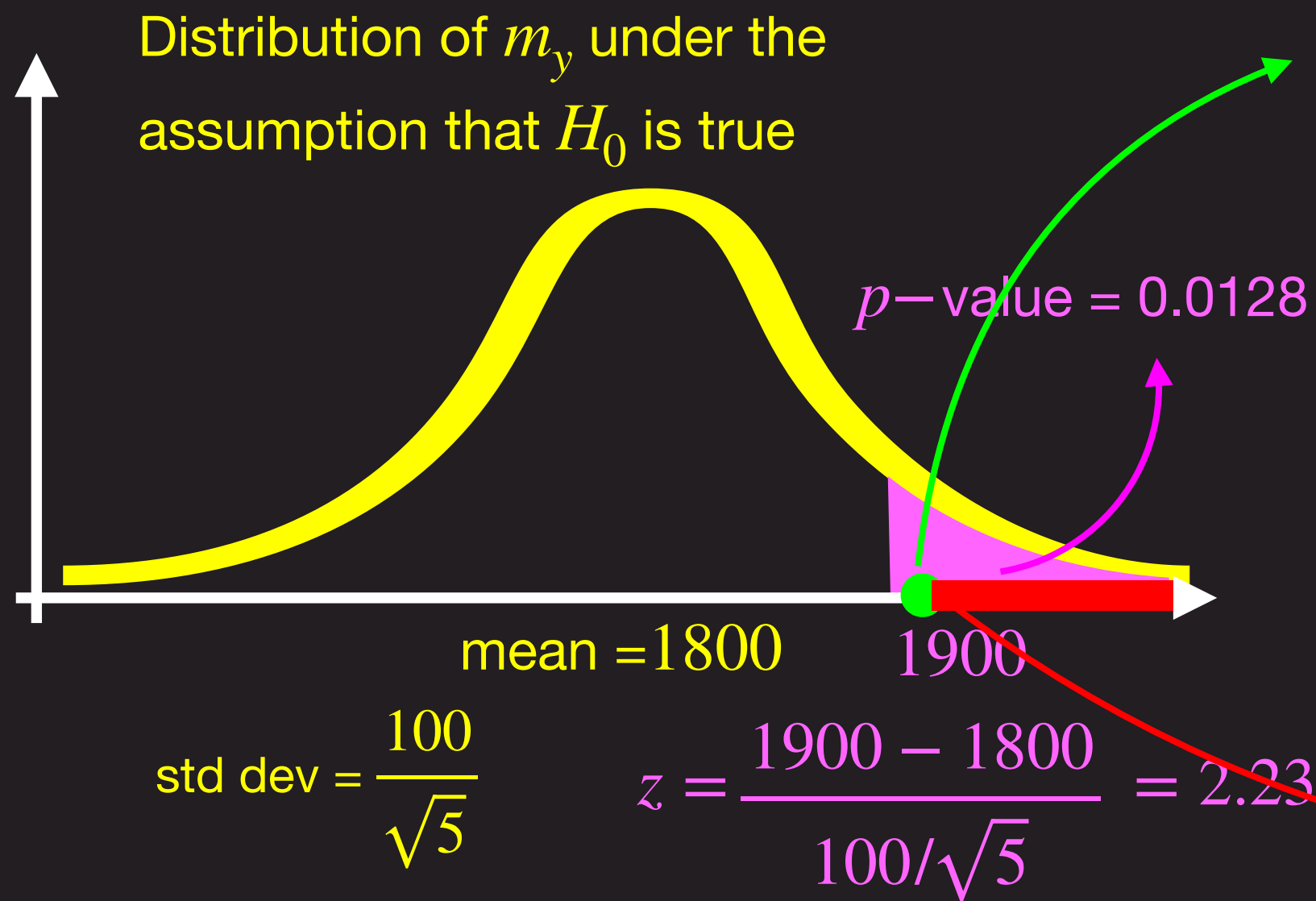
$\alpha = 0.01$

 $\mu = 1800$
 $\sigma = 100$

5 stores with average of 1900

$H_0 : \mu_y = 1800$

$H_a : \mu_y > 1800$



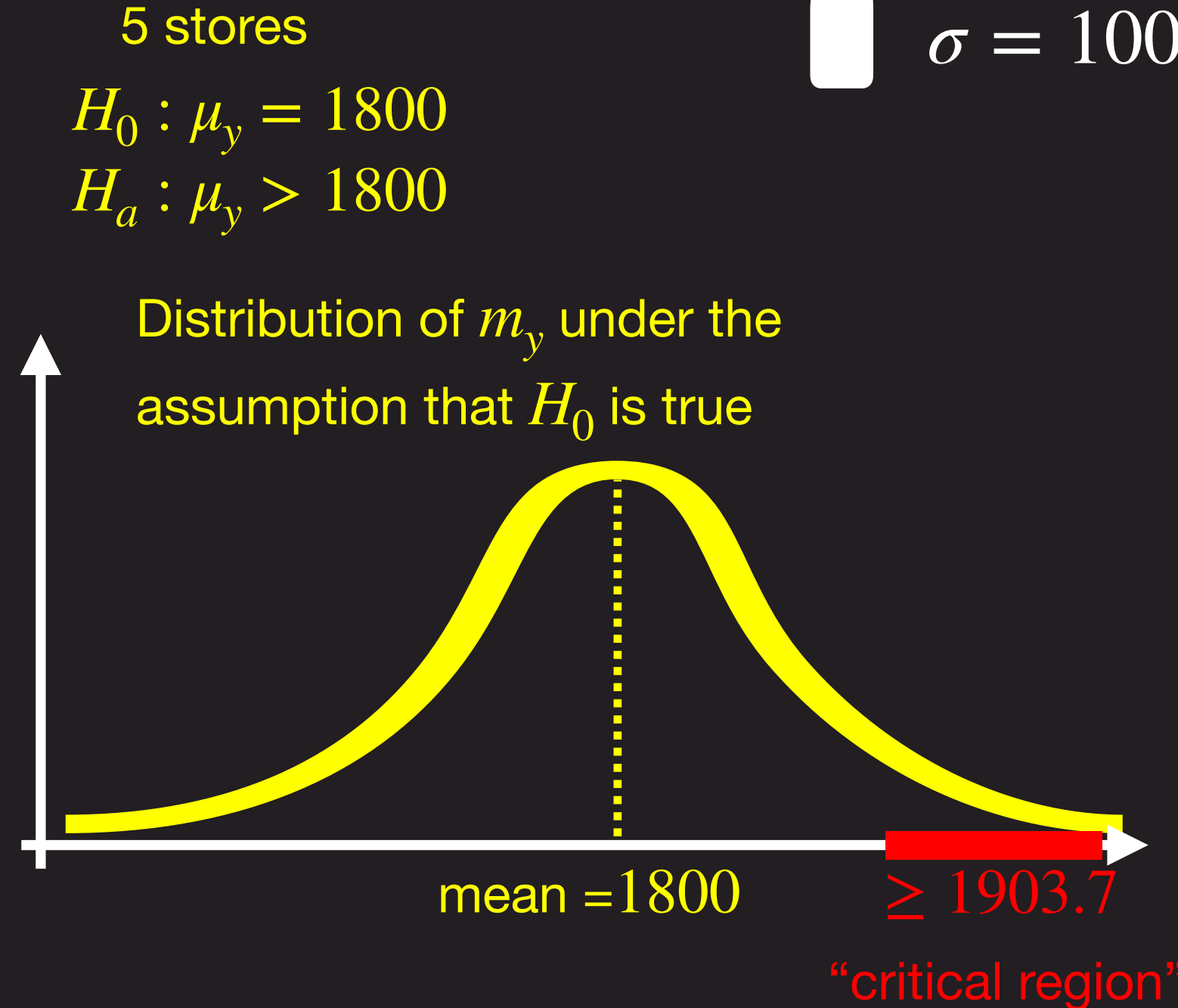
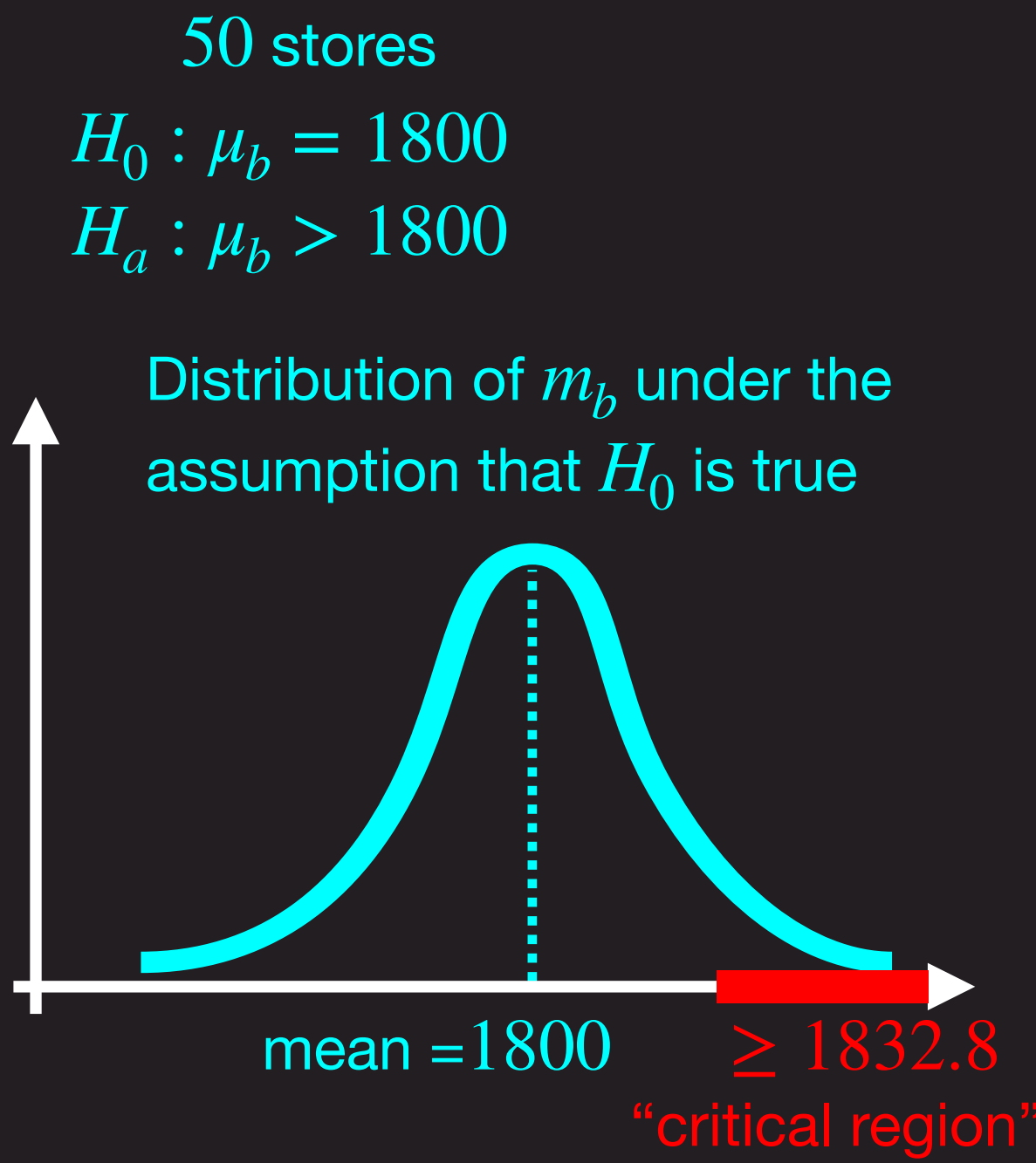
What should be the z-score such that we can reject if mean is larger, and accept if mean is lesser?
We want only 1% area to the right

Upper critical value = `norm.ppf(0.99)` = 2.32

If $z = 2.32$, then $x = 1800 + 2.32 * \frac{100}{\sqrt{5}} = 1903.7$

To summarise, if we are testing for 5 samples, we can reject the null hypothesis only if the average sales is greater than 1903.7

This region is called the “critical region”



Note: For right-tailed test, the critical region is on the right

The probability associated with critical region is α

The rule to reject is very simple: If the observed test statistic is in the critical region, then reject the null hypothesis

Hypothesis Testing Framework

- 1) Setup the Null and Alternate Hypothesis
- 2) Choose the right test statistic
- 3) Left-tailed Vs Right-tailed Vs Two-tailed
- 4) Compute p-value (Or compute the critical region)
- 5) If p-value is less than alpha, then reject the null hypothesis
(Or check if observed test statistic is in the critical region. If so, reject the null hypothesis)

Premature Children

Average IQ of all people is 100, with a standard deviation of 15

Medical researches want to know if prematurely born children have similar IQ or not

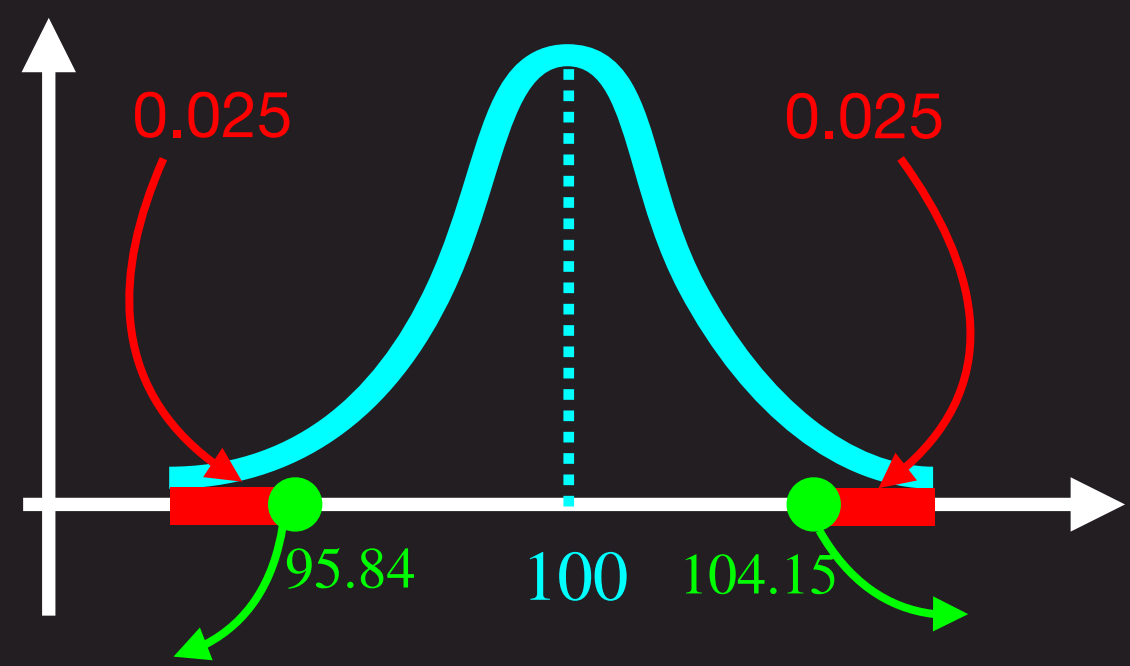
They sampled 50 such children and did an IQ test

In what range should the sample mean be to say they have normal IQ with 95% confidence?

$H_0 : \mu = 100$	What is α ?	What is the test statistic?	Right-tailed/Left-tailed/Two-tailed?
$H_a : \mu \neq 100$	$\alpha = 0.05$	Sample mean	Two-tailed

For two-tailed, we have two critical regions

The 0.05 gets split into two pieces



When do we reject H_0 ?

If sample mean is below 95.84 or above 104.15

Z-score for 2.5%
 $\text{norm.ppf}(0.025) = -1.96$

Lower critical value = $100 - 1.96 * \frac{15}{\sqrt{50}} = 95.84$

std dev = $\frac{15}{\sqrt{50}}$

Z-score for 97.5%
 $\text{norm.ppf}(0.975) = 1.96$

Upper critical value = $100 + 1.96 * \frac{15}{\sqrt{50}} = 104.15$