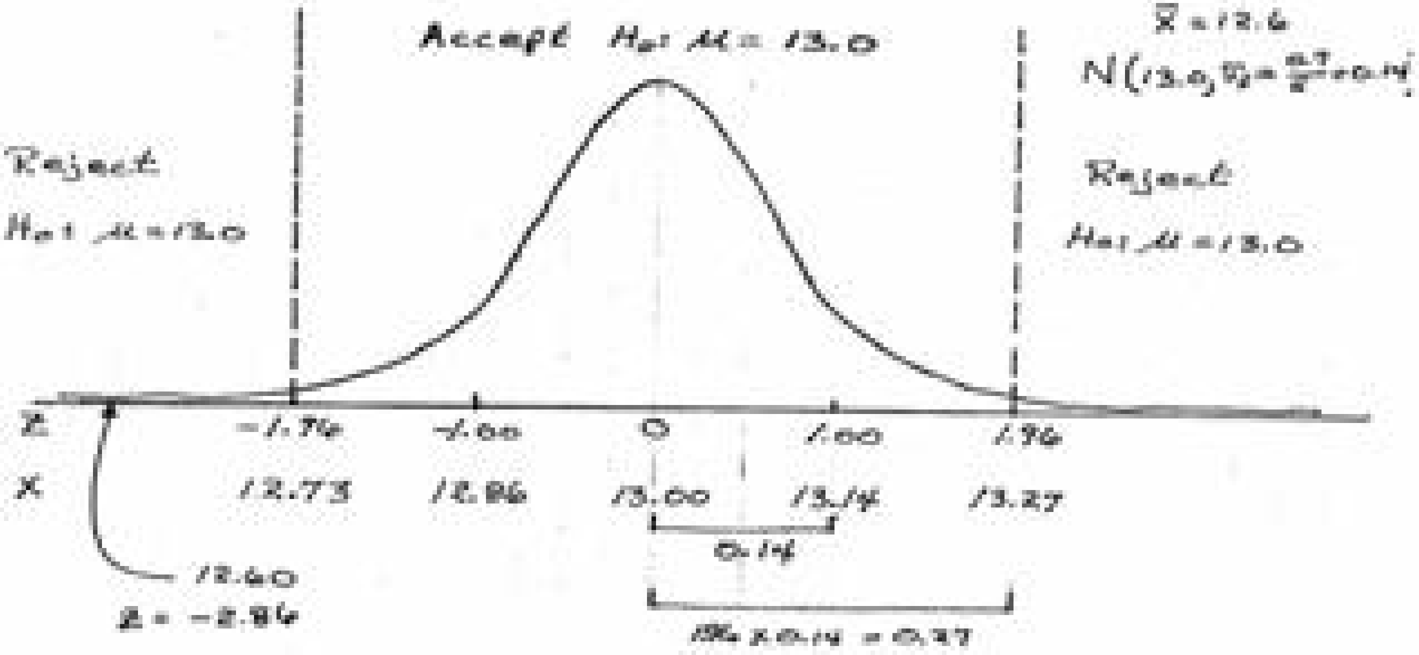


Under $H_0: \mu = 13$

$\mu = 13.00$

$\sigma^2 = 0.49$

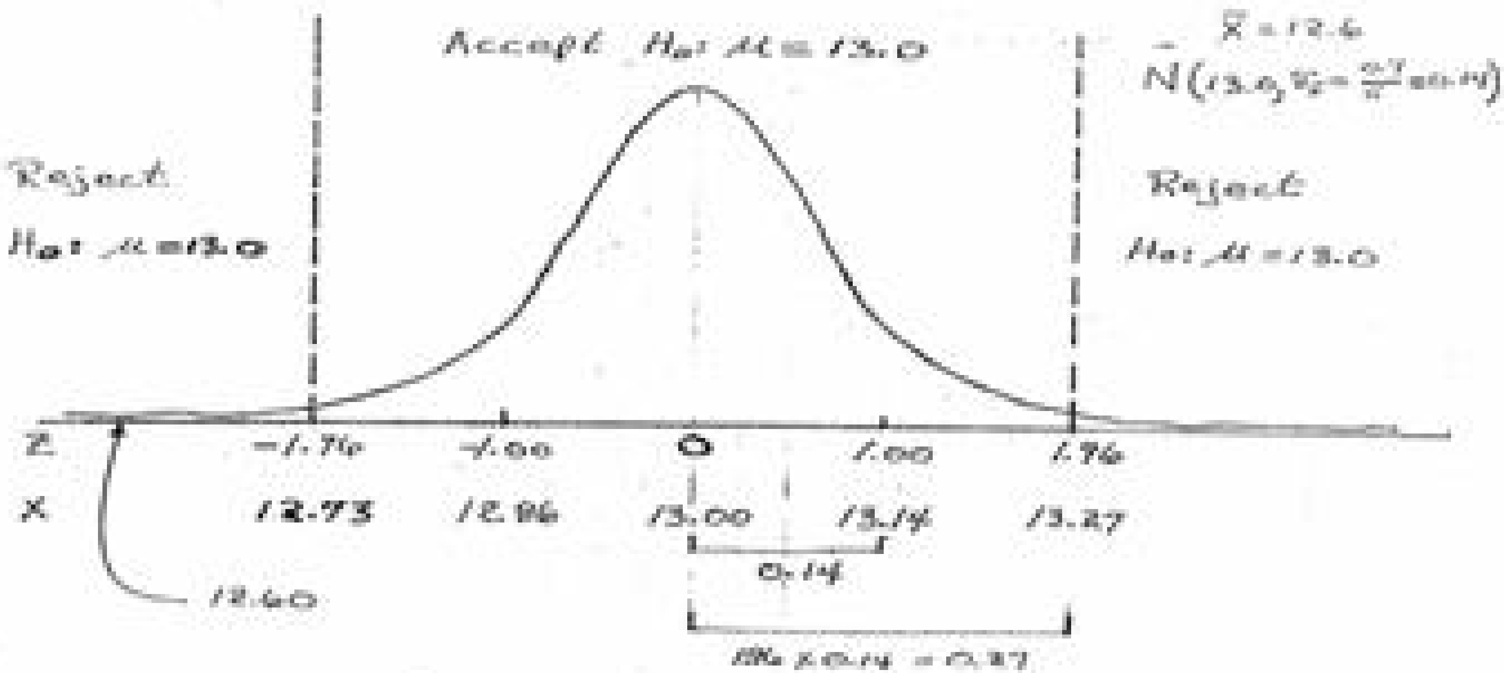
$\sigma = 0.70$

For the sample mean, with $n = 25$

$\mu_{\bar{x}} = 13.00$

$\sigma_{\bar{x}}^2 = 0.49/25 = 0.02$

$\sigma_{\bar{x}} = 0.7/5 = 0.14$



Area below $z = -2.86 = 0.00212$

Area above $z = +2.86 = 0.00212$

Total 0.00424

$P < 0.005$

Hypothesis Testing: $\bar{x} = 13.5$

A random sample of n = 25 measurements of chest circumferences from a population of newborns having σ = 0.7 inches provides a sample mean of $\bar{x}$ = **13.5** in. Is it likely that the population mean has the value μ = 13.0 in.?

1. The hypothesis: H_0: $\mu = 13.0$ versus H_1: $\mu \neq 13.0$
2. The assumptions: Random sample from a normal distribution with σ = 0.7 inches
3. The α-level: $\alpha = 0.05$

4. The test statistic:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

5. The critical region: Reject H_0: $\mu = 13.0$ if the value calculated for z is not between ±1.96

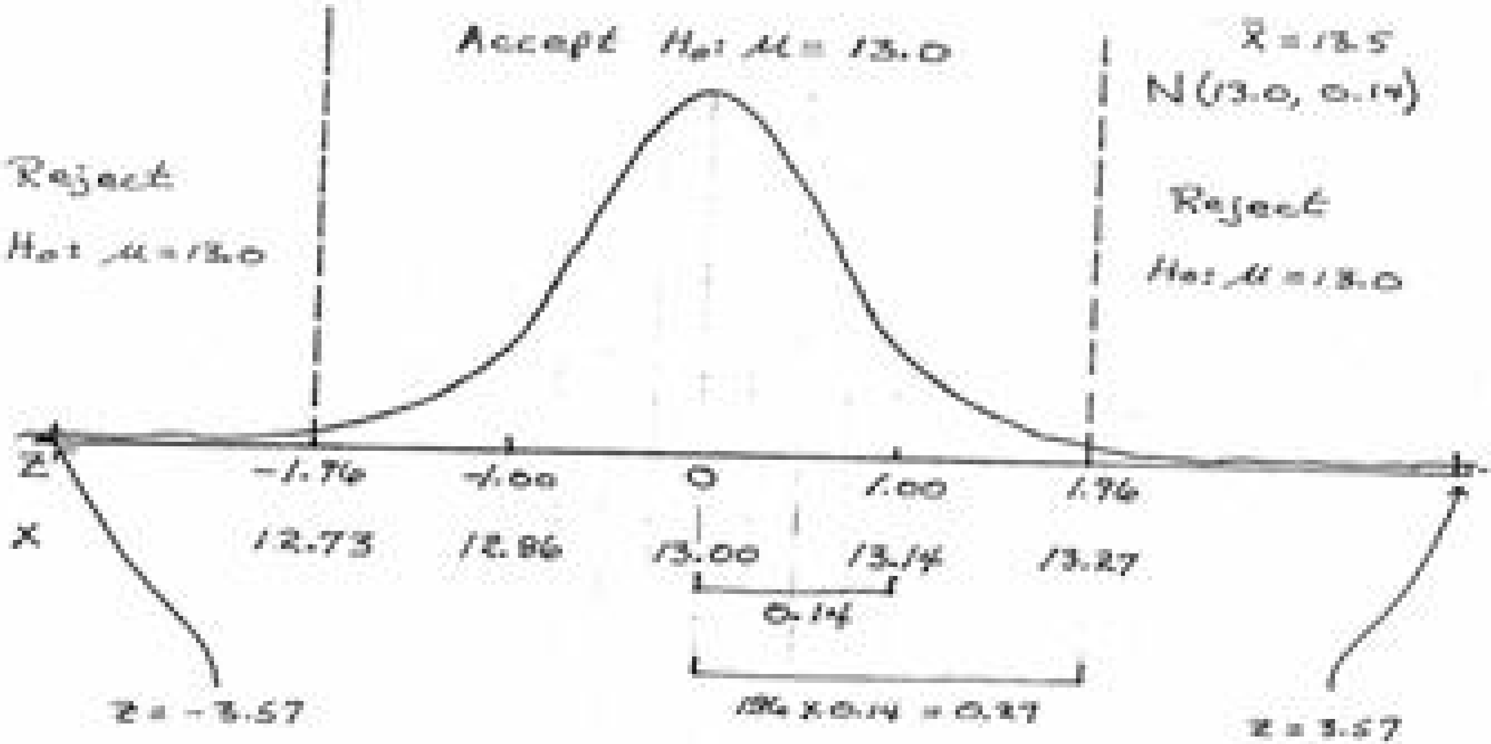
6. The result:

$$z = \frac{13.5 - 13.0}{0.7/\sqrt{25}} = \frac{0.5}{0.14} = 3.57$$

7. The conclusion: Reject H_0: $\mu = 13.0$ since the value calculated for z is not between ±1.96

This test was performed under the assumption that $\mu = 13.0$. Our conclusion is that our sample mean $\bar{x} = 13.5$ is so far away from $\mu = 13.0$ that we find it hard to believe that $\mu = 13.0$. That is, our observed value of $\bar{x} = 13.5$ for the sample mean is too rare for us to believe that $\mu = 13.0$.

How rare is $\bar{x} = 13.5$ under the assumption that $\mu = 13.0$?



Area below −3.57 = 0.00018

Area above +3.57 = 0.00018

Total 0.00036

$P < .0005$

Hypothesis Testing: $\bar{x} = 13.1$

A random sample of n = 25 measurements of chest circumferences from a population of newborns having $\sigma = 0.7$ inches provides a sample mean of $\bar{x} = \mathbf{13.1}$ in. Is it likely that the population mean has the value $\mu = 13.0$ in.?

1. The hypothesis: H_0: $\mu = 13.0$ versus H_1: $\mu \neq 13.0$

2. The assumptions: Random sample from a normal distribution with $\sigma = 0.7$ inches

3. The α-level: $\alpha = 0.05$

4. The test statistic:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

5. The critical region: Reject H_0: $\mu = 13.0$ if the value calculated for z is not between ± 1.96

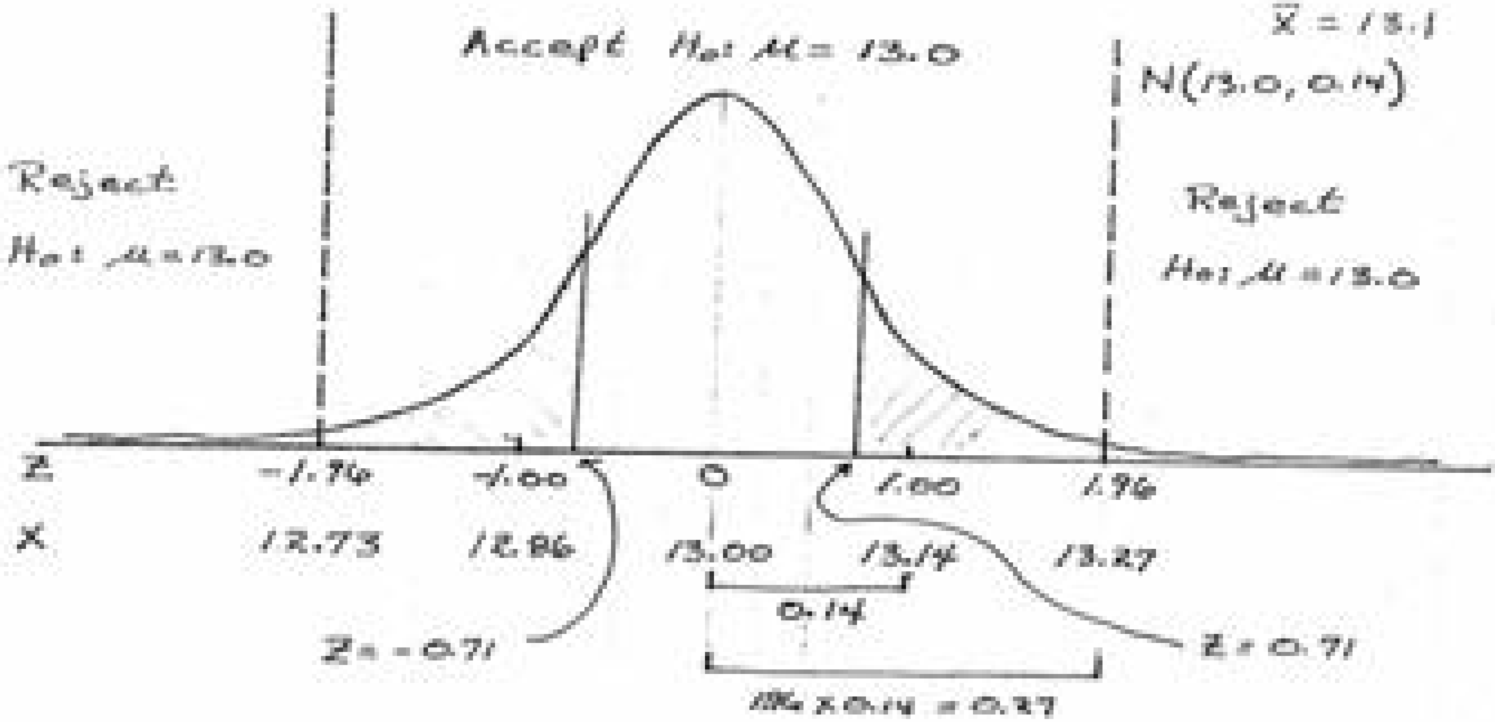
6. The result:

$$z = \frac{13.1 - 13.0}{0.7/\sqrt{25}} = \frac{0.1}{0.14} = 0.71$$

7. The conclusion: Accept H_0: $\mu = 13.0$ since the value calculated for z *is* between ± 1.96

This test was performed under the assumption that $\mu = 13.0$. Our conclusion is that our sample mean $\bar{x} = 13.1$ is *not* so far away from $\mu = 13.0$ that we find it hard to believe that $\mu = 13.0$. That is, our observed value of $\bar{x} = 13.1$ for the sample mean is not so rare and it could be that $\mu = 13.0$.

How rare is $\bar{x} = 13.1$ under the assumption that $\mu = 13.0$?

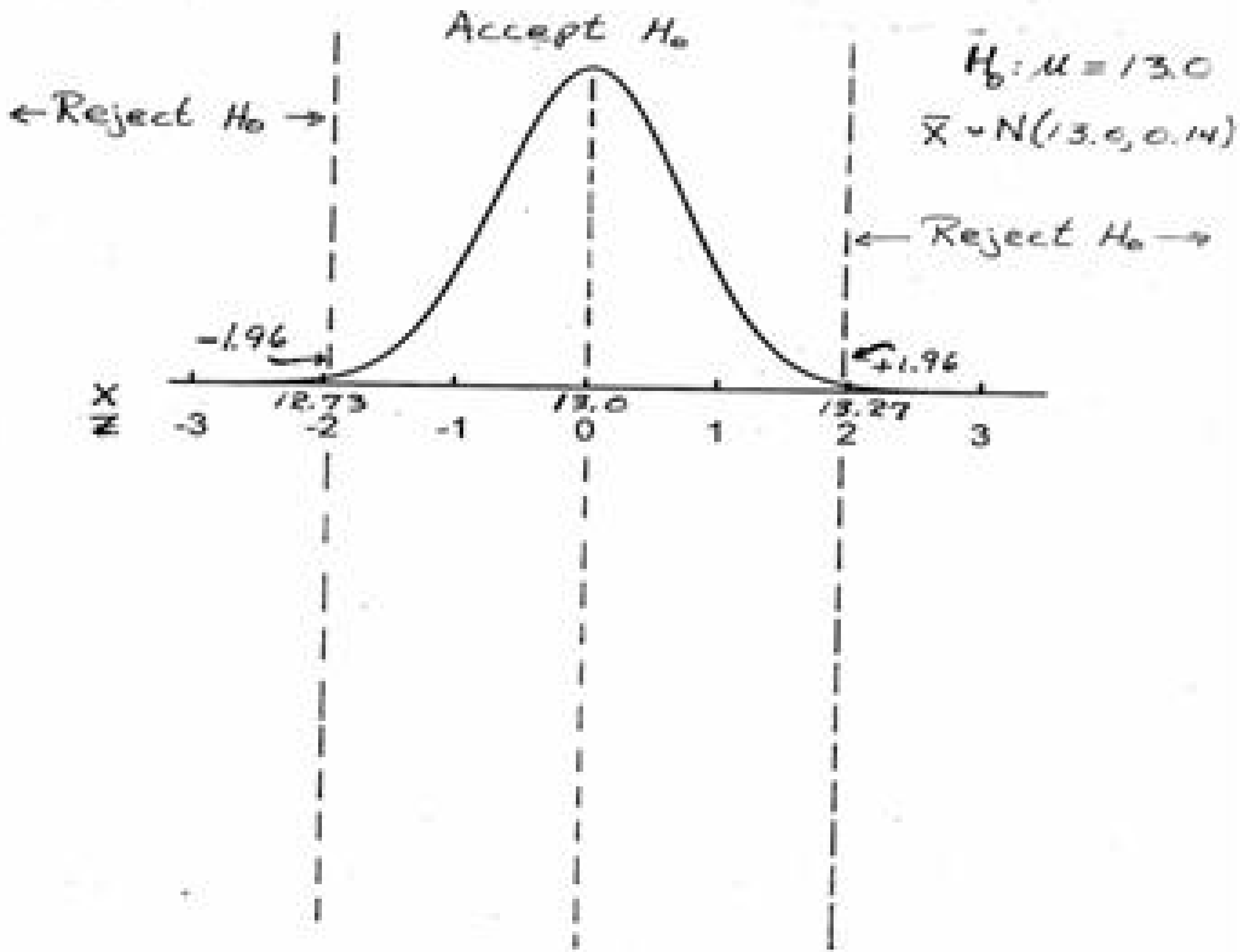


Area below −0.71 = 0.2389

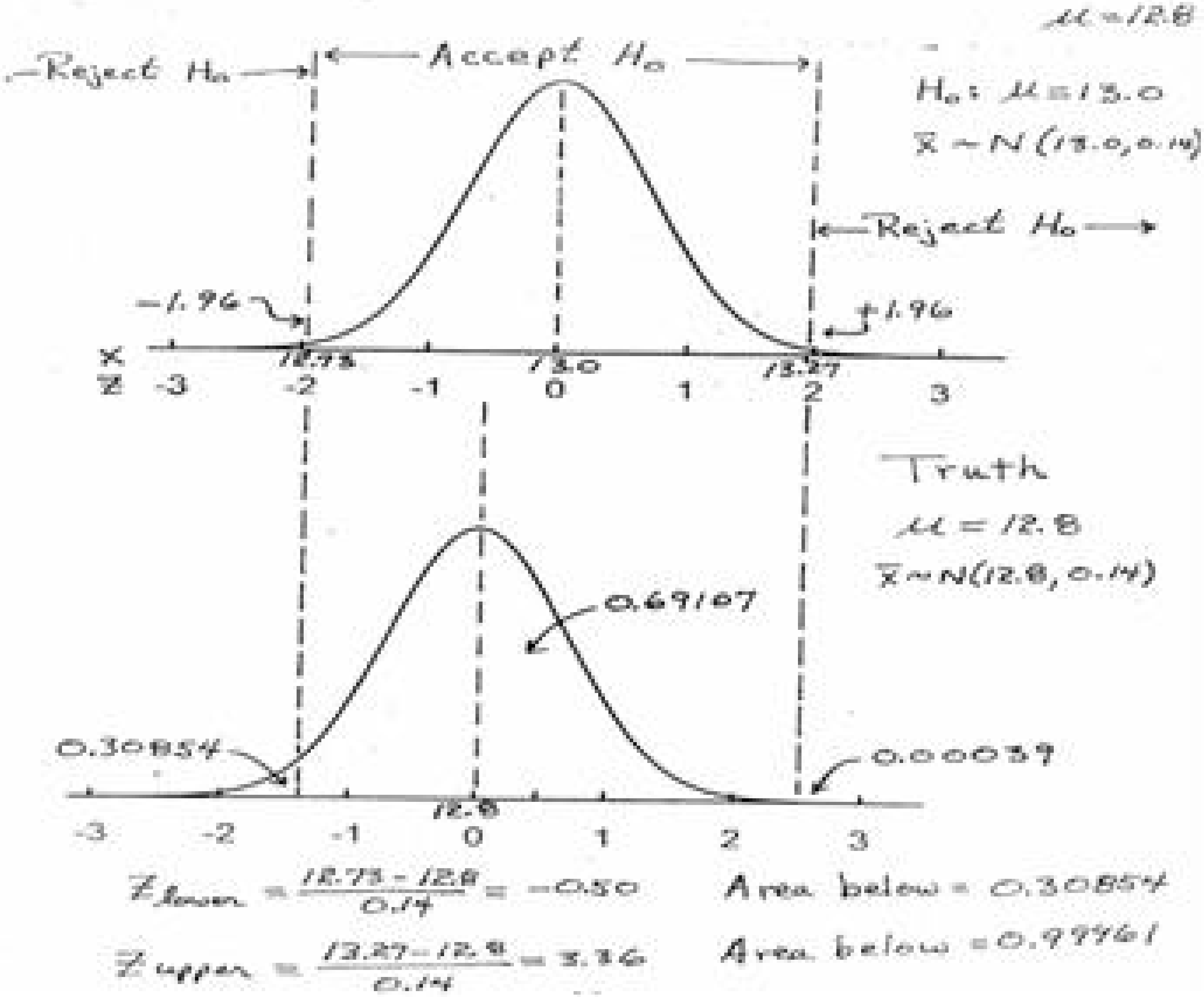
Area above +0.71 = 0.2389

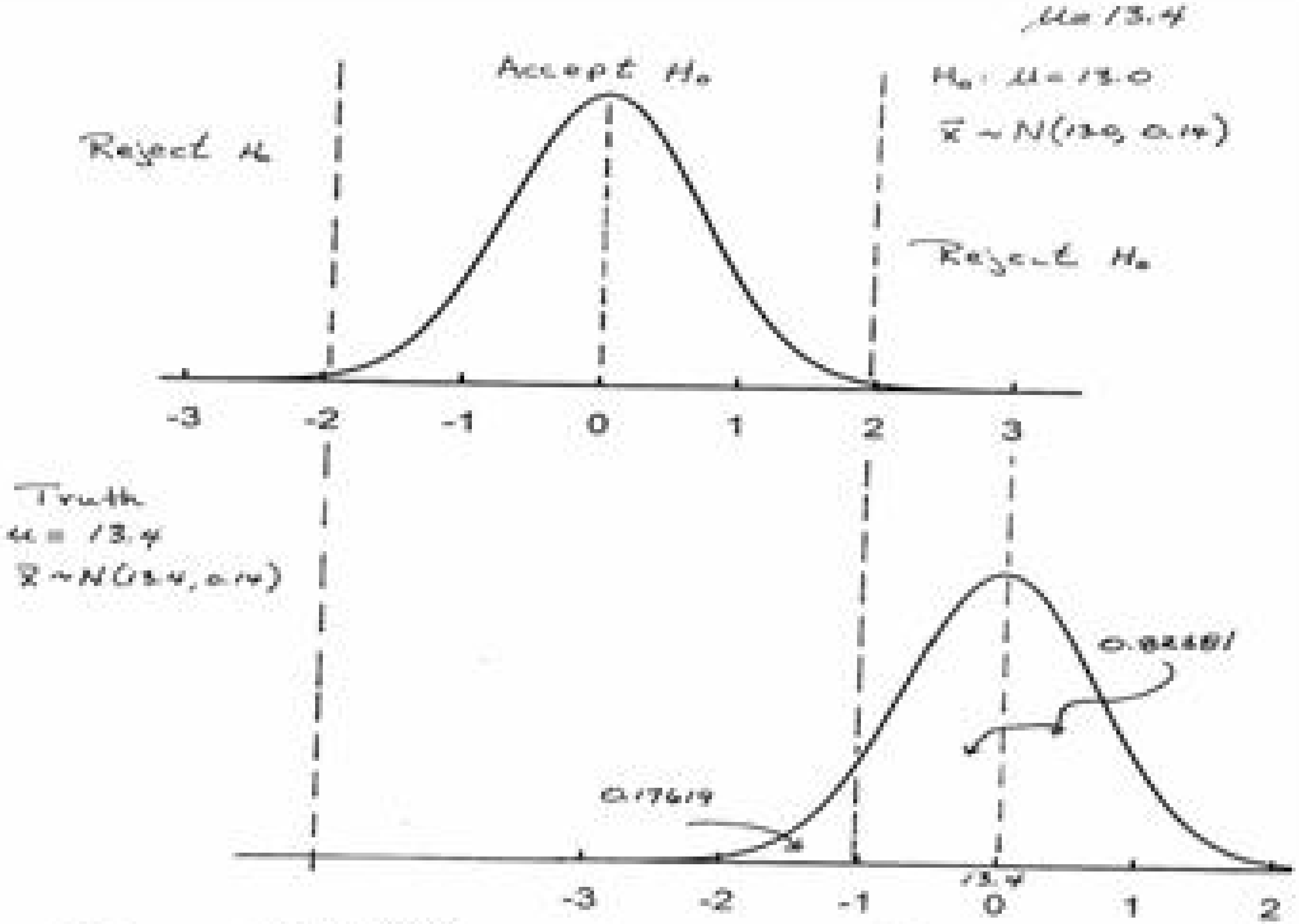
Total 0.4778

$P > 0.05$



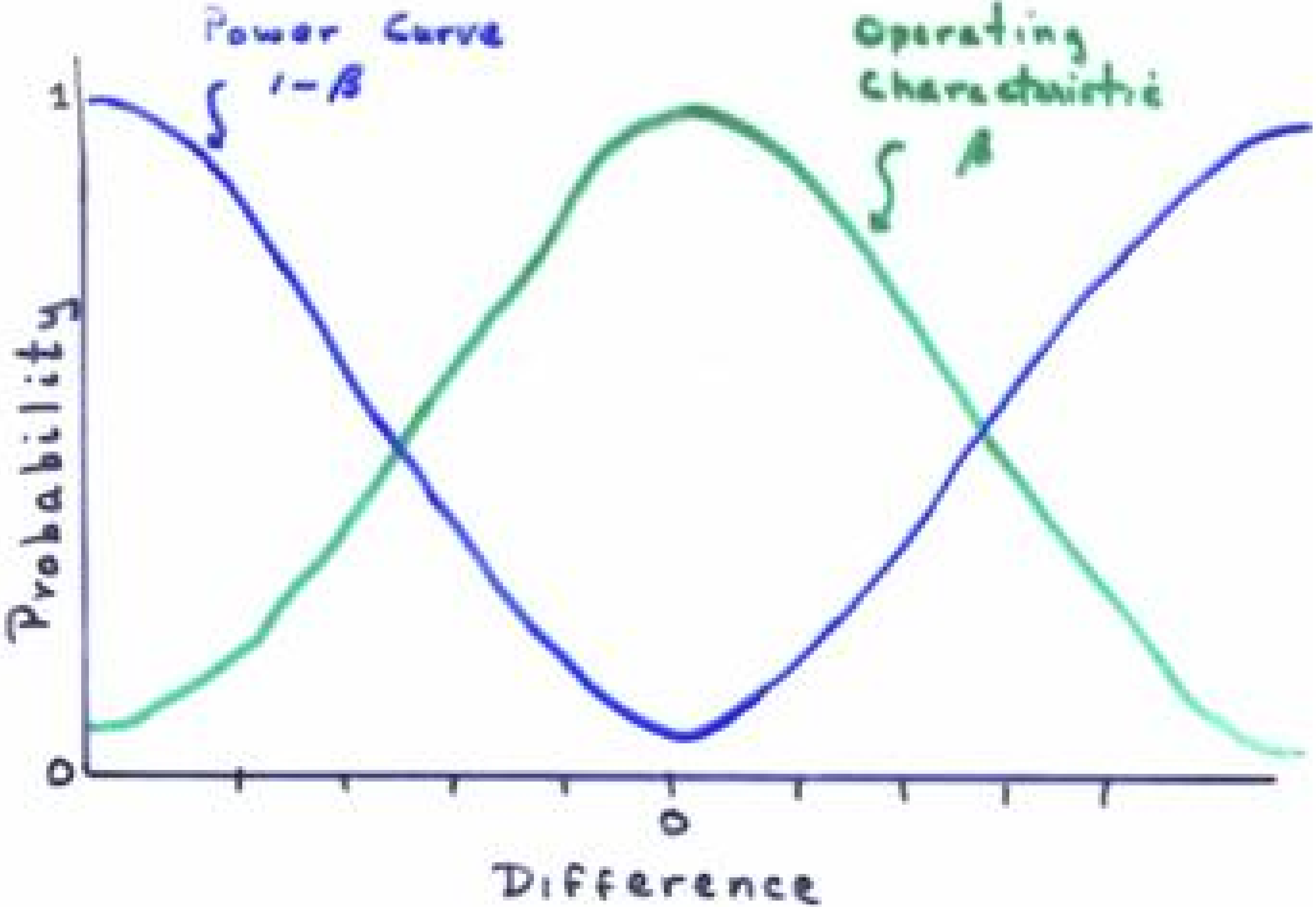
Accept H_0
← Reject H_0 →
$H_0: \mu = 13.0$
$\bar{X} \sim N(13.0, 0.14)$
← Reject H_0 →
-1.96
+1.96
X
Z
12.73
13.0
13.27
-3
-2
-1
0
1
2
3


μ=12.8
Reject H₀
Accept H₀
H₀: μ=13.0
X̄ ~ N(13.0, 0.14)
Reject H₀
−1.96
+1.96
X
Z
12.73
13.0
13.27
-3
-2
-1
0
1
2
3
Truth
μ = 12.8
X̄ ~ N(12.8, 0.14)
0.69107
0.30854
0.00039
12.8
-3
-2
-1
0
1
2
3
Z lower = (12.73 − 12.8)/0.14 = −0.50
Area below = 0.30854
Z upper = (13.27 − 12.8)/0.14 = 3.36
Area below = 0.99961



$Z_{lower} = \frac{12.73 - 13.4}{0.14} = -4.79$ Area below ≈ 0

$Z_{upper} = \frac{13.27 - 13.4}{0.14} = -0.93$ Area below = 0.17619



Power Curve
1 - β
Operating
Characteristic
β
1
0
Probability
0
Difference

The Process of Testing Hypotheses

“ The null hypothesis is never proved or established, but is possibly disproved in the course of experimentation.”

“Every experiment may be said to exist only to give the facts a chance of disproving the null hypothesis.”

R. A.Fisher

Design of experiments

Alpha (α) and Beta (β) Errors

	Truth	
Test Result	H_0 Correct	H_0 Wrong
Accept H_0	OK	Type II or β Error
Reject H_0	Type I or α Error	OK

Hypothesis Testing Key Concepts

p-value: For a specific test of a hypothesis, the likelihood or probability of observing, ***under the assumption that the null hypothesis is true***, an outcome as far away or further from the null hypothesis as the one observed.

The p-value measures the rareness of an observed outcome, under the assumption that the null hypothesis is true. If the p-value is small, typically $p < 0.05$, then it is often judged that the null hypothesis is unlikely to be true because, if it were, one would not expect to have observed so unlikely an outcome.

The p-value and the α-level are related concepts, but there are important differences between them. The α-level should be determined as a part of the process of setting up the hypothesis test. That is, it is a fundamental component of the ground rules established in setting up the hypothesis test and ***it should be set before the test is actually conducted***. Setting the α-level at, say, $\alpha = 0.05$, is the way we implement the decision as to how willing we are to reject the null hypothesis as being true when it is actually true, that is, our willingness to make a Type I or α-error.

The p-value, on the other hand, can be determined only after the test statistic is calculated. They are related in the sense that, if $\alpha = 0.05$, then we will reject the null hypothesis as being true if the value we calculate for p is $p < 0.05$.

α-error or Type I error: The error made when we reject a null hypothesis as being true when it is, if fact, true. The number we select for α, typically $\alpha = 0.05$, represents our willingness to make an α-error. We control the frequency of making this error when we select the α-level as an initial part of setting up the hypothesis test. When we use this level, we are saying that we are willing to run a 5% risk of rejecting the null hypothesis as true when it is actually true.

β-error or Type II error: The error made when we accept as true a null hypothesis that is false. The value of β is the likelihood of making a β-error. Controlling this error is more complicated than controlling the α-error. It usually involves selecting an appropriate sample size to detect a difference of a specified magnitude.

Power or 1-β: The probability of rejecting the null hypothesis when it is false. This is often thought of in the context of the **power** to detect a difference of a certain magnitude.

Module 15: Hypothesis Testing

This modules discusses the concepts of hypothesis testing, including α-level, P-values, and statistical power.

Example

Suppose we have a random sample of $n = 25$ measurements of chest circumference from a population of newborns with $\sigma = 0.7$ in.

The sample mean is $\bar{x} = 12.6$ in.

A 95% confidence interval for μ is:

$$C[\bar{x} - 1.96\,\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 1.96\,\sigma/\sqrt{n}\,] = 0.95$$

$$C[12.6 - 1.96(0.7/5) \leq \mu \leq 12.6 + 1.96(0.7/5)] = 0.95$$

$$C[12.6 - 0.27 \leq \mu \leq 12.6 + 0.27] = 0.95$$

$$C[12.33 \leq \mu \leq 12.87] = 0.95$$

Is it possible that this sample with $\bar{x} = 12.6$ came from a population with $\mu = 13.0$?

Is it likely that this sample with $\bar{x} = 12.6$ came from a population with $\mu = 13.0$?

Hypothesis Testing: $\bar{x} = 12.6$

A random sample of $n = 25$ measurements of chest circumferences from a population of newborns having $\sigma = 0.7$ inches provides a sample mean of **$\bar{x} = 12.6$ in.** Is it likely that the population mean has the value $\mu = 13.0$ in.?

1. The hypothesis:	H_0: $\mu = 13.0$ versus H_1: $\mu \neq 13.0$
2. The assumptions:	Random sample from a normal distribution with $\sigma = 0.7$ inches
3. The α-level:	$\alpha = 0.05$

4. The test statistic:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

5. The critical region: Reject H_0: $\mu = 13.0$ if the value calculated for z is not between ± 1.96

6. The result:

$$z = \frac{12.6 - 13.0}{0.7/\sqrt{25}} = \frac{-0.4}{0.14} = -2.86$$

7. The conclusion: Reject H_0: $\mu = 13.0$ since the value calculated for z is not between ± 1.96

This test was performed under the assumption that $\mu = 13.0$. Our conclusion is that our sample mean $\bar{x} = 12.6$ is so far away from $\mu = 13.0$ that we find it hard to believe that $\mu = 13.0$.

That is, our observed value of $\bar{x} = 12.6$ for the sample mean is too rare for us to believe that $\mu = 13.0$.

How rare is $\bar{x} = 12.6$ under the assumption that $\mu = 13.0$?