ISyE 3770, Spring 2024
Statistics and Applications

Hypothesis Testing

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

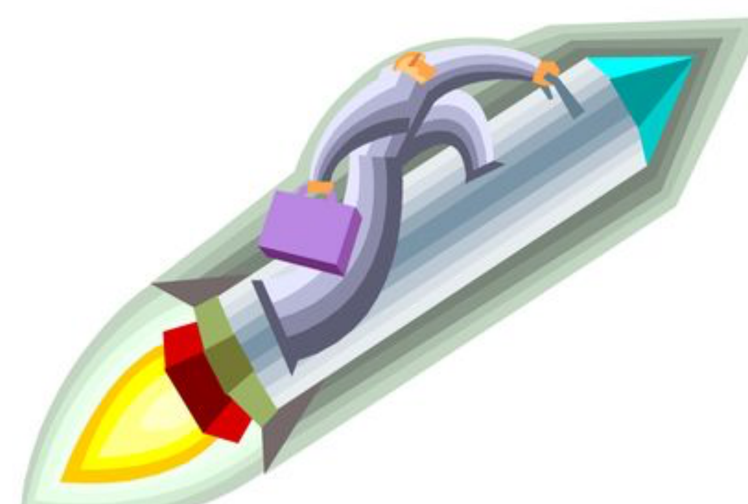
jwang3163@gatech.edu
Office: ISyE Main 447

Aircrew Escape System

- Aircrew escape systems are powered by a solid propellant. Rocket motor contains a propellant.
- To reject seat properly, specification require that the mean burn rate must be 50cm/s. Burning too slow or too fast are both not safe.
- 10 samples are tested to determine

$H_0: \mu = 50$

$H_1: \mu \neq 50.$



1

Statistical hypothesis test

- Statistical hypothesis testing of parameters are the fundamental methods used at the data analysis stage of a comparative experiment
- Many types of decision-making problems, tests, or experiments in engineering world can be formulated as hypothesis testing problems

$N(\mu, 1)$

$H_0: \mu = 0$

$H_1: \mu \neq 0$

Statistical Hypothesis

A **statistical hypothesis** is a statement about the parameters of one or more populations.

Hypothesis Test

- Interested in burning rate of a solid propellant
- Burning rate is a random variable
- Deciding whether or not the mean burning rate is 50 cm/s

Null hypothesis	H_0: $\mu = 50$ centimeters per second
Alternative hypothesis	H_1: $\mu \neq 50$ centimeters per second

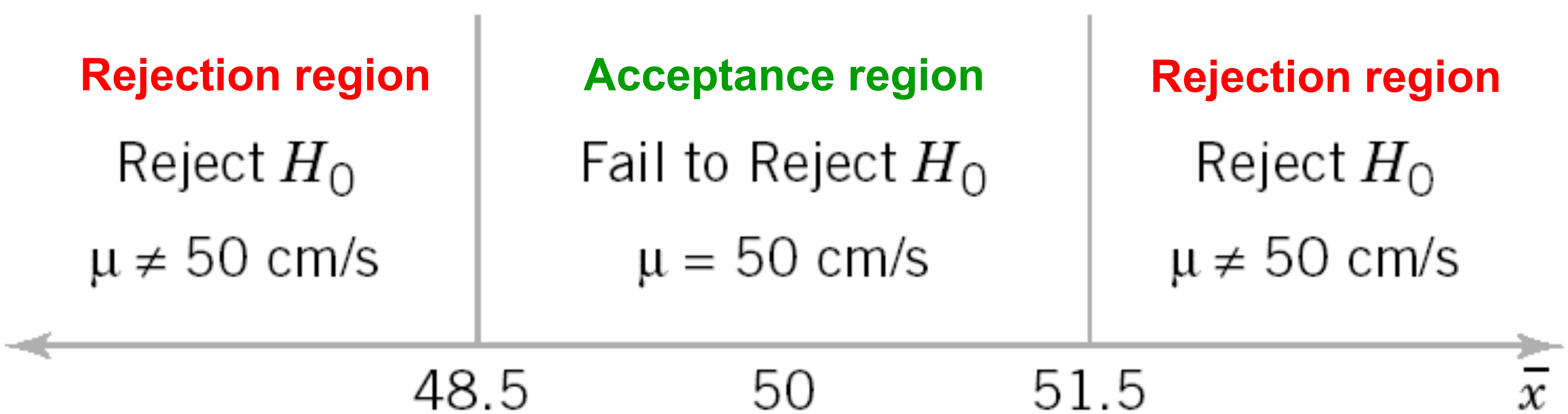
- A procedure leading to a decision about a particular hypothesis using <u>data</u> is called a test of a hypothesis

Testing a hypothesis

- Testing a hypothesis involves taking a random sample, computing a test statistic from the sample data, and then use the test statistic to make a decision about the null hypothesis
- E.g. Take 10 samples

$\bar{X}$: test statistic

$$H_0: \mu = 50 \text{ centimeters per second}$$
$$H_1: \mu \neq 50 \text{ centimeters per second}$$



Statistical Hypothesis Testing

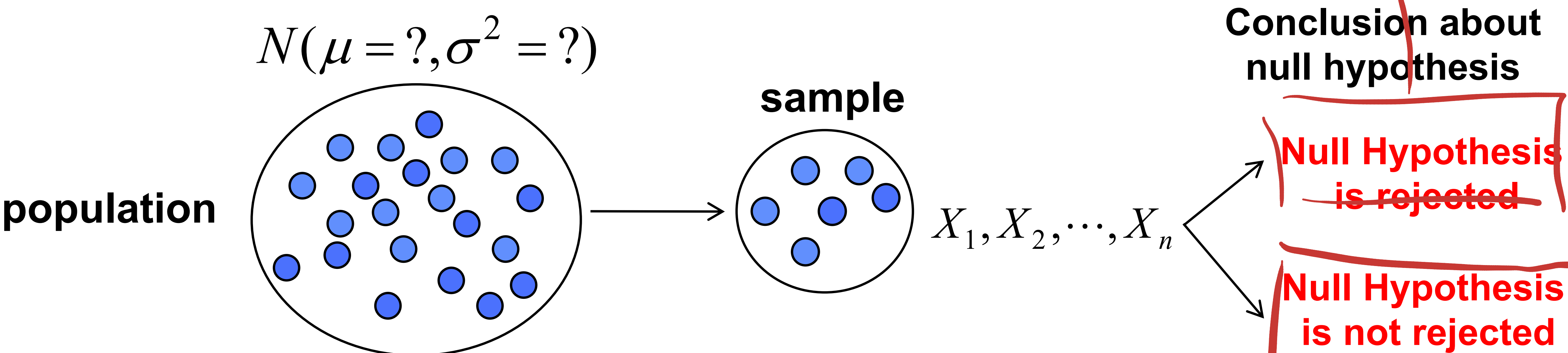
Hypothesis-testing procedures rely on using the information in a random sample from the population of interest.

If this information is *consistent* with the hypothesis, then we will conclude that the hypothesis is true; if this information is *inconsistent* with the hypothesis, we will conclude that the hypothesis is false.

Statistical Hypothesis Testing

- For example, comparing the mean of a population to a specified value

(H0 is false
H1 is true)



(H0 is true
H1 is false)

Null Hypothesis	$H_0: \mu = \mu_0$	$H_0: \sigma^2 = \sigma_0^2$
Alternative Hypothesis	$H_1: \mu \neq \mu_0$	$H_1: \sigma^2 \neq \sigma_0^2$

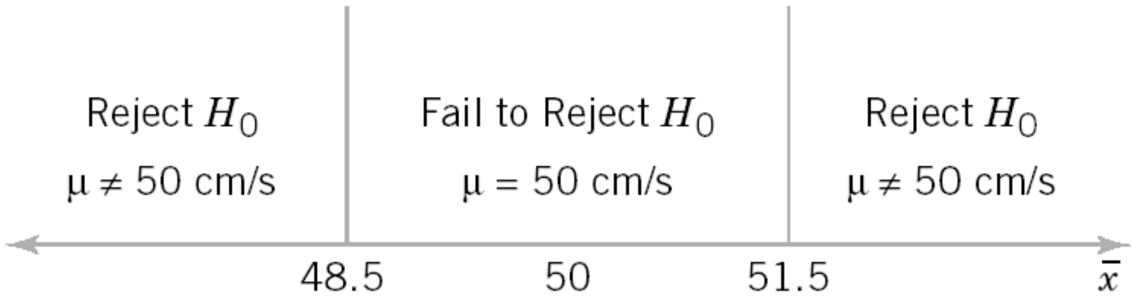
> A **statistical hypothesis** is a statement about the parameters of one or more populations.

Key Questions

- **How to set-up the hypothesis test?**
- **How to make decision?**

testing statistics

  - **How to choose detection statistic?**
  - **How to determine threshold?**



- **Concepts**
  - **Test statistic**
  - **Decision: "Rejection" "acceptance"**
  - **Significance level**
  - **p-value**

Class Activity

(1) Is the coin fair? P(tails)=P(heads)

$P(tail) = \frac{1}{2}$

H_0: ??? H_1: ??? $P(tail) \neq \frac{1}{2}$.

(2) A machine produces product (X) with mean μ, variance σ^2

(2a) Is the variability under control by σ_0^2?

$\sigma^2 \leq \sigma_0^2$

H_0: ??? H_1: ??? $\sigma^2 > \sigma_0^2$

(2b) Do we support the hypothesis that the machine in average produce an item of a size larger than a known μ_0?

H_0: ??? H_1: ???

$\mu > \mu_0$ $\mu \leq \mu_0$.

Type of Hypothesis

Simple Hypothesis

Testing two possible values of the parameter

$$H_0: \mu = 12$$ **null hypothesis**

$$H_1: \mu = 24$$ **alternative hypothesis**

Composite Hypotheses

Testing a range of values

$$H_0: \mu = 10$$

$$H_1: \mu < 10$$

X: customers' waiting time in a bank

$$H_0: \mu = 50$$

$$H_1: \mu \neq 50$$

Average diameter of screw

Errors in Hypothesis Test

Type I Error

Rejecting the null hypothesis H_0 when it is true is defined as a **type I error.**

Probability of Type I Error

$$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true}) \qquad (9\text{-}3)$$

Type II Error

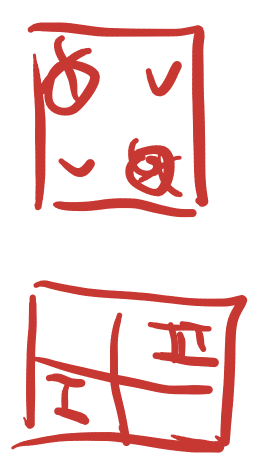
Failing to reject the null hypothesis when it is false is defined as a **type II error.**

Probability of Type II Error

$$\beta = P(\text{type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false}) \qquad (9\text{-}4)$$

Table 9-1 Decisions in Hypothesis Testing

Decision	H_0 Is True	H_0 Is False
Fail to reject H_0	no error	type II error
Reject H_0	type I error	no error



Court room decision

Suppose you are the prosecutor in a courtroom trial. The defendant is either guilty or not. The jury will either convict or not.

		truth	
		H_0: Not Guilty	H_1: Guilty
decision	Free-of-guilt	Right decision	Wrong decision Type-II error
	Convict	Wrong decision Type-I error	Right decision

Class Activity 2

In 1999, a study on the weight of students at GT provided an average weight of $\mu = 160$ lbs. We would like to test our belief that the GT student weight average did not increase in 2020 (compare to 1999)

1. What is the alternative hypothesis?

$H_0: \mu = 160$

A. $H_1: \mu = 160$ B. $H_1: \mu > 160$ C. $H_1: \mu < 160$

2. Test $H_0: \mu = 160$ vs. $H_1: \mu > 160$. What is P(Reject H_0 | $\mu = 160$)?

A. Type I error B. Type II error C. ~~Power~~

3. Test $H_0: \mu = 160$ vs. $H_1: \mu > 160$. What is P(Accept H_0 | $\mu > 160$)?

A. Type I error B. Type II error C. Power

Example

$$\alpha = P(\bar{X} < 48.5 \text{ when } \mu = 50) + P(\bar{X} > 51.5 \text{ when } \mu = 50)$$

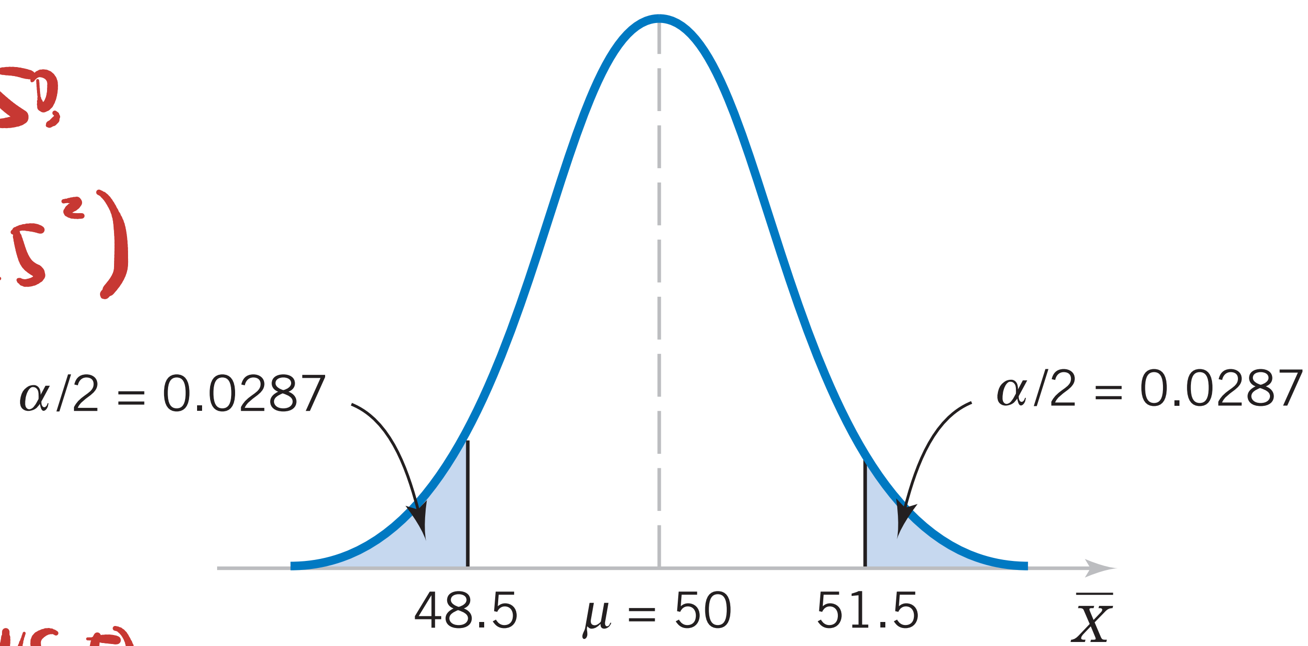
Type I error

• under H_0: $\mu = 50$

$\sigma = 2.5$

$X \sim N(50, 2.5^2)$

$\bar{X} \sim N(50, \frac{2.5^2}{10})$



$\alpha = P(N(50, \frac{2.5^2}{10}) < 48.5)$

$+ P(N(50, \frac{2.5^2}{10}) > 51.5)$

Figure 9-2 The critical region for H_0: $\mu = 50$ versus H_1: $\mu \neq 50$ and $n = 10$.

$= \text{pnorm}(48.5, 50, \sqrt{\frac{2.5^2}{10}})$

$+ 1 - \text{pnorm}(51.5, \ldots)$

$= 0.057$

Rejection region	Acceptance region	Rejection region
Reject H_0	Fail to Reject H_0	Reject H_0
$\mu \neq 50$ cm/s	$\mu = 50$ cm/s	$\mu \neq 50$ cm/s

48.5 50 51.5 $\bar{x}$

Suppose that the standard deviation of burning rate is $\sigma = 2.5$ centimeters per second

$$\alpha = P(\overline{X} < 48.5 \text{ when } \mu = 50) + P(\overline{X} > 51.5 \text{ when } \mu = 50)$$

n = 10 samples

$$\sigma/\sqrt{n} = 2.5/\sqrt{10} = 0.79.$$

$$z_1 = \frac{48.5 - 50}{0.79} = -1.90 \qquad \text{and} \qquad z_2 = \frac{51.5 - 50}{0.79} = 1.90$$

$$\alpha = P(Z < -1.90) + P(Z > 1.90) = 0.0287 + 0.0287 = 0.0574$$

$$H_1: \mu = 52$$

$\bar{X} \sim N(52, \frac{2.5^2}{10})$

$\beta = P(48.5 \le \bar{X} \le 51.5 \text{ when } \mu = 52)$

$= N(52, 0.625)$

The z-values corresponding to 48.5 and 51.5 when $\mu = 52$ are $\frac{\bar{X} - 52}{\sqrt{0.625}}$

$$z_1 = \frac{48.5 - 52}{0.79} = -4.43 \qquad \text{and} \qquad z_2 = \frac{51.5 - 52}{0.79} = -0.63$$

$\beta = \text{pnorm}(51.5, 52, \sqrt{\frac{2.5^2}{10}}) - \text{pnorm}(48.5, 52, \sqrt{\frac{2.5^2}{10}})$

$$\beta = P(-4.43 \le Z \le -0.63) = P(Z \le -0.63) - P(Z \le -4.43)$$
$$= 0.2643 - 0.0000 = 0.2643$$

Significance Level

- Generally, the statistician <u>controls type I error</u> probability α when she selects the critical values
- Set type I error probability at a <u>desired level</u> called significance level
- Deal with the corresponding type II error β resulted from this

significance level

A widely used procedure in hypothesis testing is to use a type 1 error or significance level of $\alpha = 0.05$. This value has evolved through experience, and may not be appropriate for all situations.

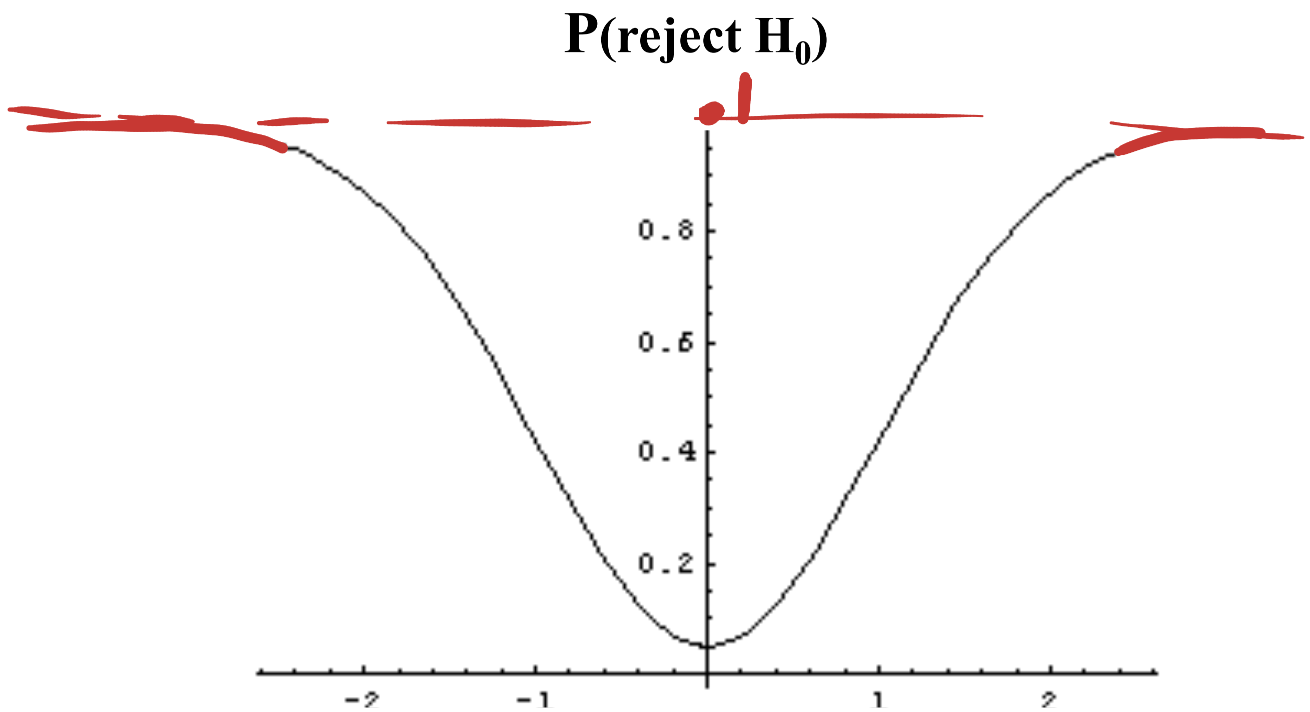
In general we control P(type-I error) $\leq$ 0.05 !!

Typical value for significance level: 0.1, 0.05, 0.001

Statistical Power

Power = 1 – β = P(reject H_0 when H_1 is true)

Example: power function for the test



$$H_0 : \mu = 0$$
$$H_1 : \mu \neq 0$$

Power

The **power** of a statistical test is the probability of rejecting the null hypothesis H_0 when the alternative hypothesis is true.

p-value

- Fixed significance level report the results of a hypothesis test

type- I $\leq \alpha$

- Significance level is an upper bound by α

decision

- But this way we have no idea how strong the evidence is

0.08

- Also report **p-value** of the statistic

if p-value is smaller, evidence is stronger

0.005
0.001

p-value

- $H_0: \mu = 0$ $H_1: \mu \neq 0$
- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ • Realization of $\bar{X}$ is 2

• p-value $= P(\bar{X} \geq 2 \mid H_0 \text{ is true})$
$= P(N(\mu, \frac{\sigma^2}{n}) \geq 2 \mid \mu = 0)$

The P-value is the probability that the test statistic will take on a value that is at least as extreme as the observed value of the statistic when the null hypothesis H_0 is true.

$= P(N(0, \frac{\sigma^2}{n}) \geq 2)$

- **Given a significance level** $\alpha = 0.05$
- **If the p-value is smaller than** α $\Rightarrow$ **Reject** H_0 (p-value $= 0.0001$)
- **If the p-value is larger than** α $\Rightarrow$ (p-value $= 0.06$)

there is not enough evidence to reject H_0

Computing p-value

- Let $\bar{X}$ be the testing statistic, and x_0 be its realized value.

- p-value $= \begin{cases} Pr_{H_0}(\bar{X} \geq x_0), & \text{if } H_0: \mu = \mu_0, \; H_1: \mu > \mu_0 \\ Pr_{H_0}(\bar{X} \leq x_0), & \text{if } H_0: \mu = \mu_0, \; H_1: \mu < \mu_0 \\ 2\min\{Pr_{H_0}(\bar{X} \geq x_0), Pr_{H_0}(\bar{X} \leq x_0)\} & \text{if } H_0: \mu = \mu_0, \; H_1: \mu \neq \mu_0 \end{cases}$

- Smaller p-value indicates stronger evidence to reject H_0.

Example

9-20. A manufacturer is interested in the output voltage of a power supply used in a PC. Output voltage is assumed to be normally distributed, with standard deviation 0.25 volt, and the manufacturer wishes to test $H_0: \mu = 5$ volts against $H_1: \mu \neq 5$ volts, using $n = 8$ units.

(a) The acceptance region is $4.85 \leq \bar{x} \leq 5.15$. Find the value of α.

(b) Find the power of the test for detecting a true mean output voltage of 5.1 volts.

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) = N(5.1, \frac{0.25^2}{8})$

9-23. In Exercise 9-20, calculate the *P*-value if the observed statistic is

(a) $\bar{x} = 5.2$ (b) $\bar{x} = 4.7$ (c) $\bar{x} = 5.1$

power $= 1 - \beta = 1 - P(4.85 \leq \bar{X} \leq 5.15 \mid \mu = 5.1) = 1 - P(N(5.1, \frac{0.25^2}{8}) \in [4.85, 5.15])$

- Normal population $\sigma = 0.25$
- $H_0 : \mu = 5$ $H_1 : \mu \neq 5$ $n = 8$

(a) $\boxed{4.85 \leq \bar{x} \leq 5.15} \longrightarrow$ accept H_0

$\alpha = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

$= P(\bar{X} < 4.85 \text{ or } \bar{X} > 5.15 \mid \mu = 5)$

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$\bar{X} \sim N(5, \frac{0.25^2}{8})$

$= P(N(5, \frac{0.25^2}{8}) < 4.85 \text{ or } N(5, \frac{0.25^2}{8}) > 5.15)$

$= P(N(5, \frac{0.25^2}{8}) < 4.85) + P(N(5, \frac{0.25^2}{8}) > 5.15)$

$= \text{pnorm}(4.85, 5, \sqrt{\frac{0.25^2}{8}}) + 1 - \text{pnorm}(5.15, 5, \sqrt{\frac{0.25^2}{8}}) = ?$

$n = 8$ $\qquad\qquad$ $H_0 : \mu = 5$

$\sigma = 0.25$ $\qquad\qquad$ $H_1 : \mu \neq 5$

$4.85 \leq \bar{x} \leq 5.15$

$$\alpha = 1 - \mathbb{P}(4.85 \leq \bar{x} \leq 5.15 \mid \mu = 5) \qquad \bar{x} \sim N\left(5, \frac{0.25^2}{8}\right)$$

$$= 1 - \mathbb{P}\left(\frac{4.85-5}{0.25/\sqrt{8}} \leq \frac{\bar{x}-5}{0.25/\sqrt{8}} \leq \frac{5.15-5}{0.25/\sqrt{8}}\right)$$

$$= 1 - [\Phi(1.697) - \Phi(-1.697)]$$

$$= 0.0897$$

Use R-command pnorm(x,0,1) to find $\Phi(x)$.

① $\beta = P(4.85 \leq \bar{x} \leq 5.15 \mid \mu = 5.2)$

$$= 0.2857$$

② $\beta = \mathbb{P}(4.85 \leq \bar{x} \leq 5.15 \mid \mu = 4.7)$

$$= \Phi\left(\frac{5.15-4.7}{0.25/\sqrt{8}}\right) - \Phi\left(\frac{4.85-4.7}{0.25/\sqrt{8}}\right) = 0.0448$$

③ $\beta = \mathbb{P}(4.85 \leq \bar{x} \leq 5.15 \mid \mu = 5.1)$

$$= \Phi\left(\frac{5.15-5.1}{0.25/\sqrt{8}}\right) - \Phi\left(\frac{4.85-5.1}{0.25/\sqrt{8}}\right) = 0.7118$$

Solution to 9.23

Assume $H_0: \mu = 5$ $H_1: \mu \neq 5$.

(a) When $\bar{X} = 5.2$, under H_0, $\bar{X} \sim N(5, \frac{0.25^2}{8})$

$$p\text{-value} = 2\min\{Pr_{H_0}(\bar{X} \geq 5.2), Pr_{H_0}(\bar{X} \leq 5.2)\}$$
$$= 2 Pr_{H_0}(\bar{X} \geq 5.2) = 0.024$$

(b) When $\bar{X} = 4.7$,

$$p\text{-value} = 2\min\{Pr_{H_0}(\bar{X} \geq 4.7), Pr_{H_0}(\bar{X} \leq 4.7)\}$$
$$= 2 Pr_{H_0}(\bar{X} \leq 4.7) = 0.0006$$

(c) When $\bar{X} = 5.1$, $p\text{-value} = 2 Pr_{H_0}(\bar{X} \geq 5.1) = 0.2579$.

Additional Example:

Suppose $\bar{X} \sim N(\mu, \frac{0.25^2}{8})$. $H_0: \mu = 5$, $H_1: \mu > 5$.

When realization value $\bar{x} = 5.1$,

$$p\text{-value} = Pr_{H_0}(\bar{X} \geq 5.1) = Pr(N(5, \frac{0.25^2}{8}) \geq 5.1)$$

$$= 0.1289$$

Additional Example:

Suppose $\bar{X} \sim N\left(\mu, \frac{0.25^2}{8}\right)$. $H_0: \mu = 5$. $H_1: \mu < 5$

When realized value $\bar{x} = 4.8$,

$$\text{p-value} = Pr_{H_0}(\bar{X} \leq 4.8)$$

$$= 0.012$$

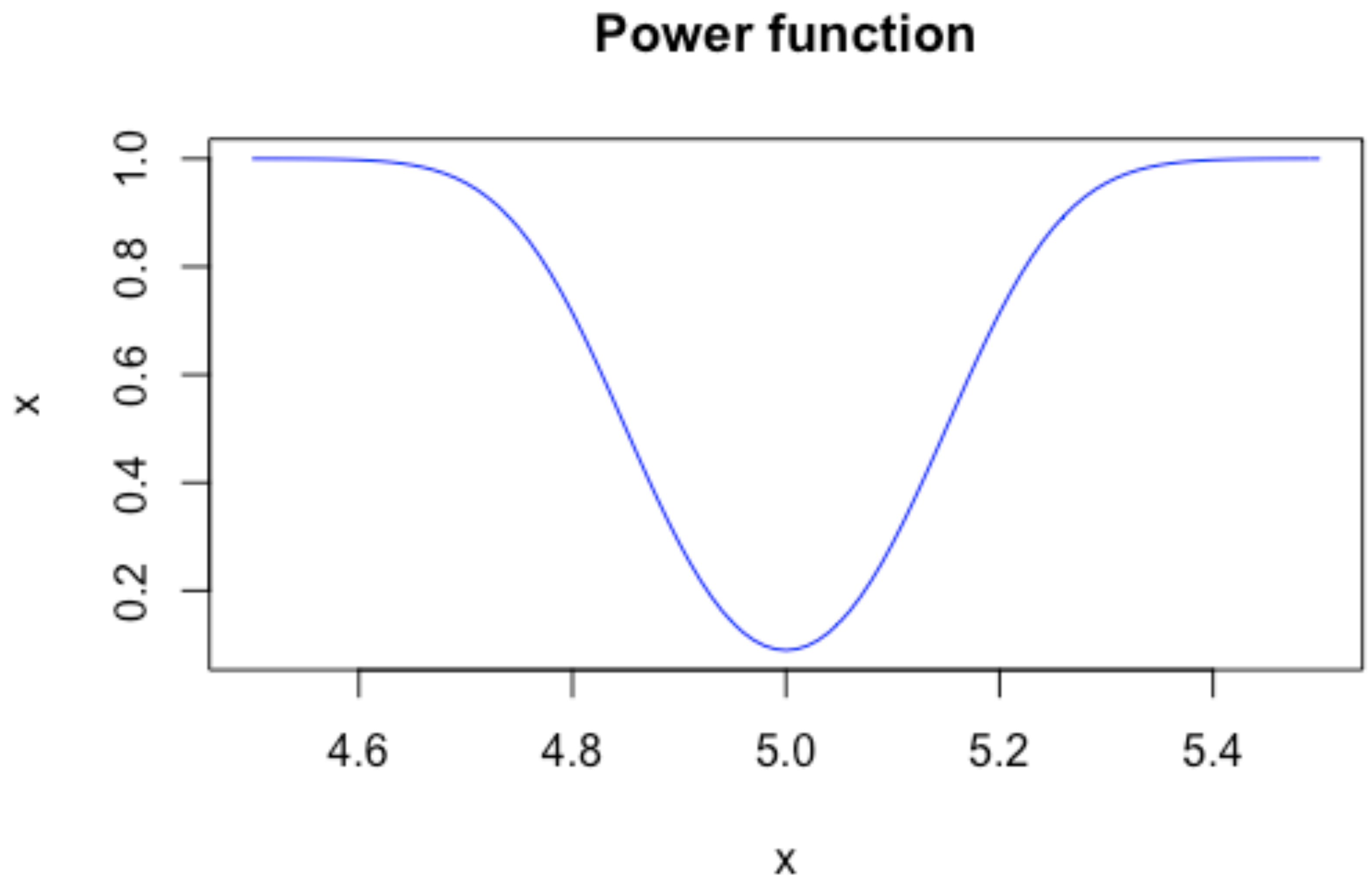
```
# clear space
rm(list = ls())

# initialize
n <- 50
x <- (450:550)/100
y <- rep(0, length(x))
len <- length(x)
ind <- 1:len
CR_upper <- 5.15
CR_lower <- 4.85

#CR_upper = 10
#CR_lower = 3

for (i in ind){
y[i] <- pnorm((CR_upper-x[i])/(0.25/sqrt(8)), 0, 1)-
pnorm((CR_lower-x[i])/(0.25/sqrt(8)), 0, 1)
print(1-y[i])
}

plot(x, 1-y, type="l", main="Power function", ylab = "x",
col = "blue")
```



Hypothesis Testing Procedures

Example: $H_0: \mu = 0$ & $H_1: \mu \neq 0$

1. **Parameter of interest:** From the problem context, identify the parameter of interest.
2. **Null hypothesis, H_0:** State the null hypothesis, H_0.
3. **Alternative hypothesis, H_1:** Specify an appropriate alternative hypothesis, H_1.
4. **Test statistic:** Determine an appropriate test statistic.
5. **Reject H_0 if:** State the rejection criteria for the null hypothesis.
6. **Computations:** Compute any necessary sample quantities, substitute these into the equation for the test statistic, and compute that value.
7. **Draw conclusions:** Decide whether or not H_0 should be rejected and report that in the problem context.

H_0 is rejected

There is no evidence to reject H_0 $\Leftrightarrow$ H_0 is accepted

8. **Report p-values**

Inference on the Mean of a Normal Population – Known Variance

Null Hypothesis

$$H_0 : \mu = \mu_0$$

Test Statistic

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

Distribution under H0

$$Z_0 \sim N(0,1)$$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : \mu \neq \mu_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$
$H_1 : \mu > \mu_0$	$Z_0 > Z_\alpha$
$H_1 : \mu < \mu_0$	$Z_0 < -Z_\alpha$

Threshold: Critical values

H_0 is accepted

$-Z_{\alpha/2}$ $+Z_{\alpha/2}$

H_0 is rejected

values of Z_0

Accept H_0 Z_α Reject H_0

values of Z_0

Reject H_0 $-Z_\alpha$ Accept H_0

Why this gives the desired significance level?

- **Proof** : e.g. $H_1 : \mu \neq \mu_0$

$$\alpha = \mathbb{P}(\text{reject } H_0 \text{ when } H_0 \text{ is true})$$

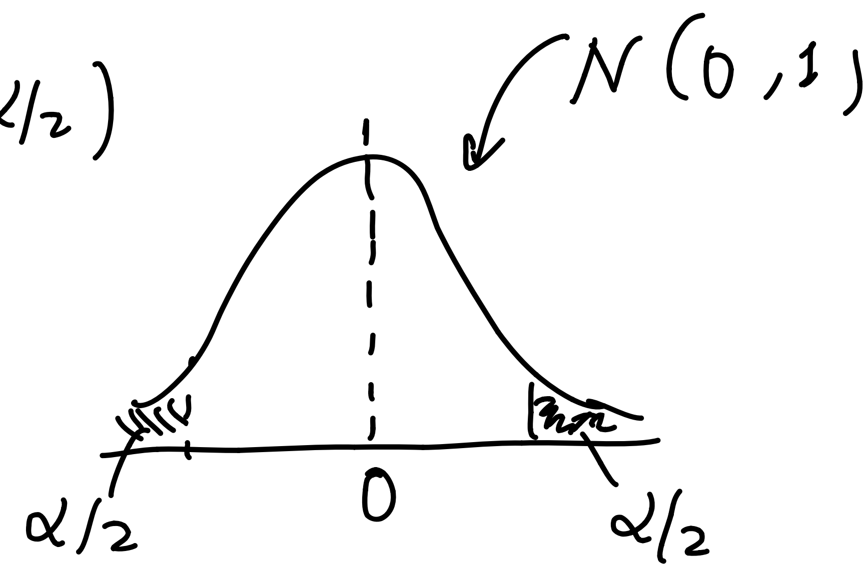
$$= \mathbb{P}\left(|Z| \geq z_{\alpha/2} \mid \mu = \mu_0 \right)$$

$$X_1 \cdots X_n \sim N(\mu_0, \sigma^2)$$

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

(from definition of $z_{\alpha/2}$)

$$\alpha = \alpha/2 \times 2 = \alpha$$



Example: Battery life

- The life in hours of a battery is known to be approximately normally distributed with standard deviation 1.25 hours
- A random sample of 10 batteries has a mean life of 40.5 hours
- Is there an evidence to support the claim that battery life exceeds 40 hours? Use. $\alpha = 0.05$

1. Parameter:

true mean parameter μ $n=10$ $\alpha=0.05$

population $\sim N(\mu, \sigma^2 = 1.25^2)$

2/3: Hypothesis: $H_0: \mu = 40$ $H_1: \mu > 40$ $\mu_0 = 40$

4. Testing statistic: $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{\bar{X} - 40}{1.25/\sqrt{10}}$

5. Rejection region:

Reject H_0 if $Z_0 > Z_{0.05}$

6. Computation: $Z_0 = \dfrac{40.5 - 40}{1.25/\sqrt{10}} = 1.26$ $Z_{0.05} = \text{qnorm}(1-0.05) = 1.64$

7. Conclusion:
H_0 is not rejected

Solution

1. Parameter: true battery life μ, $n = 10$

$X_1, \dots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, $\sigma = 1.25$

2. null hypothesis: $H_0: \mu = 40$
3. Alternative: $H_1: \mu > 40$, $(\mu_0 = 40)$
4. test statistic: standardized sample mean $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$
5. Reject H_0 if: $Z_0 > Z_\alpha$, $\alpha = 0.05$
6. Computation:

$$Z_0 = \frac{40.5 - 40}{1.25 / \sqrt{10}} = 1.265$$

$$Z_\alpha = Z_{0.05} = 1.645$$

7. Draw conclusion: Z_0 is not greater than Z_α (critical value). So we cannot reject H_0. There's not enough evidence to support H_1.

p-value: observed statistic value = 1.265

we reject H_0 when μ is large.

Therefore, $p\text{-value} = \mathbb{P}(\overset{N(0,1)}{\text{test stat}} > 1.265)$

$= 1 - \Phi(1.265)$

$= 0.103 > \alpha = 0.05$

So indeed, we don't have enough evidence to reject H_0

More Examples and Case Study

1. Ice Hockey Player: Variance Parameter

9-77. The data from *Medicine and Science in Sports and Exercise* described in Exercise 8-48 considered ice hockey player performance after electrostimulation training. In summary, there were 17 players and the sample standard deviation of performance was 0.09 seconds.

(a) Is there strong evidence to conclude that the standard deviation of performance time exceeds the historical value of **0.07** seconds? Use $\alpha = 0.05$. Find the *P*-value for this test.

1. parameter σ

2/3. Null / Alternative Hypothesis

$H_0: \sigma = 0.07$

$H_1: \sigma > 0.07$

Inference on the Variance of a Normal Population

sample variance

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \sigma^2 = \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 \sim \chi^2(n-1)$

$\chi^2_{\alpha, n-1}$ = qchisq(1-α, n-1)

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : \sigma^2 \neq \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha/2, n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2, n-1}$
$H_1 : \sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha, n-1}$
$H_1 : \sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha, n-1}$

Critical values

Accept H0

$\chi^2_{\alpha/2, n-1}$ $\chi^2_{1-\alpha/2, n-1}$

Reject H0

values of χ_0^2

$\chi^2_{\alpha, n-1}$

values of χ_0^2

$\chi^2_{1-\alpha, n-1}$

Hypothesis Testing Using Confidence Intervals - Variance of a Normal Population

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0 : \sigma^2 = \sigma_0^2 \qquad H_1 : \sigma^2 \neq \sigma_0^2 \quad \Rightarrow \quad \text{CI} : \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}} \right]$$

$$H_0 : \sigma^2 = \sigma_0^2 \qquad H_1 : \sigma^2 > \sigma_0^2 \quad \Rightarrow \quad \text{CI} : \left[\frac{(n-1)s^2}{\chi^2_{\alpha, n-1}}, +\infty \right)$$

Lower CI

$$H_0 : \sigma^2 = \sigma_0^2 \qquad H_1 : \sigma^2 < \sigma_0^2 \quad \Rightarrow \quad \text{CI} : \left(0, \frac{(n-1)s^2}{\chi^2_{1-\alpha, n-1}} \right]$$

Upper CI

- **If the Confidence Interval does NOT include σ_0^2, then Reject H0**

Solution

1. parameter of interest: σ

2. $\begin{cases} H_0: \sigma = 0.07 \\ H_1: \boxed{\sigma > 0.07} \end{cases}$ $\sigma_0 = 0.07$
3.

4. Test stats: $\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

5. Reject when $\chi_0^2 > \chi^2_{\alpha, n-1}$

6. Compute: $\chi_0^2 = \dfrac{(17-1) \times 0.09^2}{0.07^2} = 26.45$

$n = 17$

$\sigma_0 = 0.75$ $\chi^2_{\alpha, n-1} = \chi^2_{0.05, 16} = 26.3$

$s = 0.09$ (R: qchisq (0.05, 16, lower.tail = F))

$\hookrightarrow$ qchisq(1- 0.05, 16)

7. Conclusion: reject H_0.

p-value: $\mathbb{P}(\chi^2(n-1) > 26.45) = 0.048$

$\hookrightarrow$ 1-pchisq(26.45, 16)

Can also solve using C.I.

2. Engineering Higher Education: Sample Proportion

9-92. An article in *Fortune* (September 21, 1992) claimed that nearly one-half of all engineers continue academic studies beyond the B.S. degree, ultimately receiving either an M.S. or a Ph.D. degree. Data from an article in *Engineering Horizons* (Spring 1990) indicated that 117 of 484 new engineering graduates were planning graduate study.

(a) Are the data from *Engineering Horizons* consistent with the claim reported by *Fortune*? Use $\alpha = 0.05$ in reaching your conclusions. Find the *P*-value for this test.

(b) Discuss how you could have answered the question in part (a) by constructing a two-sided confidence interval on p.

1. parameter: p.

2/3: H_0: $p = \frac{1}{2}$.

H_1: $p \neq \frac{1}{2}$.

Inference on a Population Proportion

Null Hypothesis

$$H_0 : p = p_0$$

Test Statistic

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Asymptotic Distribution under H0

$$Z_0 \sim N(0,1)$$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : p \neq p_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$
$H_1 : p > p_0$	$Z_0 > Z_{\alpha}$
$H_1 : p < p_0$	$Z_0 < -Z_{\alpha}$

Critical values

Hypothesis Testing Using Confidence Intervals a Population Proportion

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0: p = p_0 \qquad H_1: p \neq p_0 \quad \Rightarrow \text{CI}: \left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$$H_0: p = p_0 \qquad H_1: p > p_0 \quad \Rightarrow \text{CI}: \left[\hat{p} - Z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, +\infty\right)$$

Lower CI

$$H_0: p = p_0 \qquad H_1: p < p_0 \quad \Rightarrow \text{CI}: \left(-\infty, \hat{p} + Z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Upper CI

- **If the Confidence Interval does NOT include p_0, then Reject H0**

Solution

1. Parameter-of-interest: p $\quad p \neq 0.5$

2. $H_0: p = 0.5$ $\quad (p_0 = 0.5, \quad n = 484$
3. $H_1: p < 0.5$ $\quad$ data: $\hat{p} = 117/484 = 0.242$

4. Test statistic: $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{0.242 - 0.5}{\sqrt{\dfrac{0.5 \times 0.5}{484}}}$

5. Reject H_0 when $z_0 < -z_\alpha$ $\Rightarrow$ $|z_0| > z_{\alpha/2}$

6. Compute: $z_0 = -11.352$

$z_\alpha = z_{0.05} = 1.645$

7. reject H_0 since $z_0 < -z_\alpha$

8. p-value: $\mathbb{P}(Z < z_0) = \Phi(-11.352)$

$= 3.6 \times 10^{-30}$

3. Comparing two paints

A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested; formulation 1 is the standard chemistry, and formulation 2 has a new drying ingredient that should reduce the drying time. From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient. Ten specimens are painted with formulation 1, and another 10 specimens are painted with formulation 2; the 20 specimens are painted in random order. The two sample average drying times are $\bar{x}_1 = 121$ minutes and $\bar{x}_2 = 112$ minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using $\alpha = 0.05$?

$\Rightarrow \sigma_1 = \sigma_2 = 8$

$n_1 = n_2 = 10$

$\bar{x}_1 = 121 \quad \bar{x}_2 = 112$

population distributions are assumed to be normal.

Inference for Differences in Means of Two Normal Distributions (known variances)

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \mu_1 - \mu_2 = \Delta_0$	$Z_0 = \dfrac{(\overline{X} - \overline{Y}) - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$Z_0 \sim N(0,1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$Z_0 > Z_\alpha$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$Z_0 < -Z_\alpha$

Critical values

Hypothesis Testing Using Confidence Intervals Mean of Two Normal Populations – Known Variance

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$
$$H_1 : \mu_1 - \mu_2 \neq \Delta_0$$
$$\Rightarrow (\bar{X} - \bar{Y}) - Z_{\alpha/2} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{X} - \bar{Y}) + Z_{\alpha/2} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$
$$H_1 : \mu_1 - \mu_2 > \Delta_0$$
$$\Rightarrow \left[(\bar{X} - \bar{Y}) - Z_{\alpha} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, +\infty \right)$$

Lower CI

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$
$$H_1 : \mu_1 - \mu_2 < \Delta_0$$
$$\Rightarrow \left(-\infty, (\bar{X} - \bar{Y}) + Z_{\alpha} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right]$$

Upper CI

- **If the Confidence Interval does NOT include Δ_0, then Reject H0**

Solution

1. parameter-of-interest: $\mu_1 - \mu_2$

2. $H_0: \mu_1 - \mu_2 = 0$ $\quad \Delta_0 = 0$,
3. $H_1: \mu_1 - \mu_2 > 0$

4. Test stat: $Z_0 = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$, under H_0

5. Reject when $Z_0 > Z_\alpha$

6. Compute: $Z_0 = \dfrac{121 - 112}{\sqrt{\dfrac{8^2}{10} \times 2}} = 2.5156$

$\sigma_1^2 = \sigma_2^2 = 8^2$

$n_1 = n_2 = 10$

$Z_\alpha = Z_{0.05} = 1.645$

7. reject H_0 since $Z_0 > Z_\alpha$
8. p-value: $\mathbb{P}(Z > Z_0) = 1 - \Phi(2.5156) = 0.0059$

Question 3 of HW 7

use R software to perform hypothesis testing, Avoid Step 4-8 to have conclusions.

4: comparing catalysts

EXAMPLE 10-5 Yield from a Catalyst
Two catalysts are being analyzed to determine how they affect the mean yield of a chemical process. Specifically, catalyst 1 is currently in use, but catalyst 2 is acceptable. Since catalyst 2 is cheaper, it should be adopted, providing it does not change the process yield. A test is run in the pilot plant and results in the data shown in Table 10-1. Is there any difference between the mean yields? Use $\alpha = 0.05$, and assume equal variances.

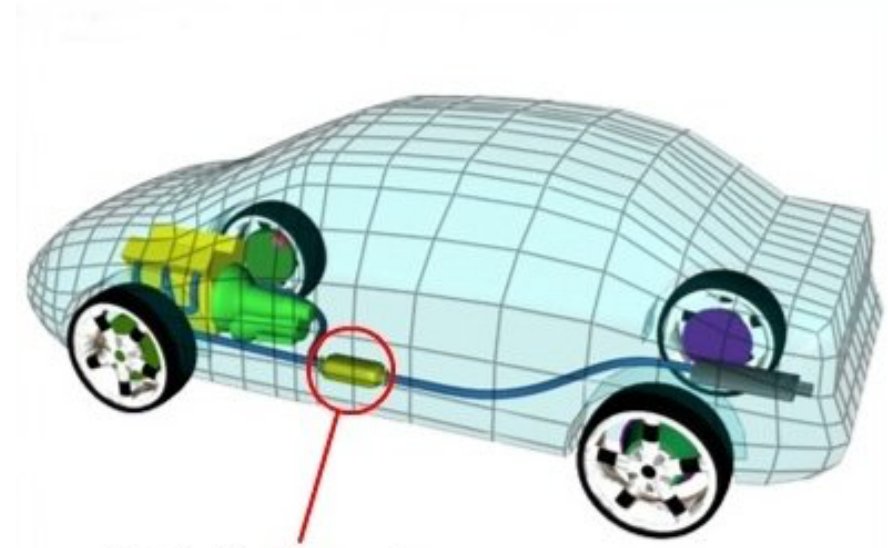
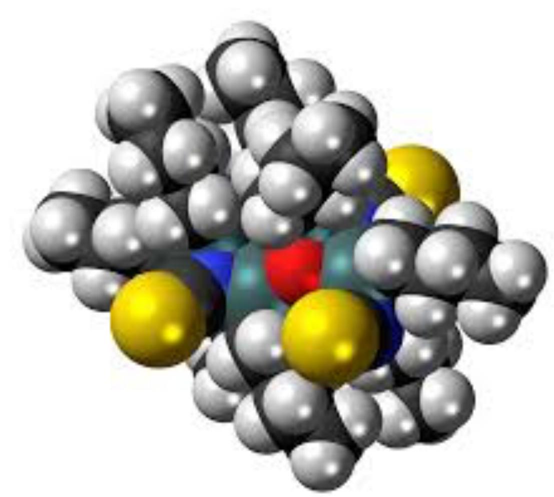


Table 10-1 Catalyst Yield Data, Example 10-5

30 33: Statistics courses in advance

Observation Number	Catalyst 1	Catalyst 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$



Comparing by descriptive statistics

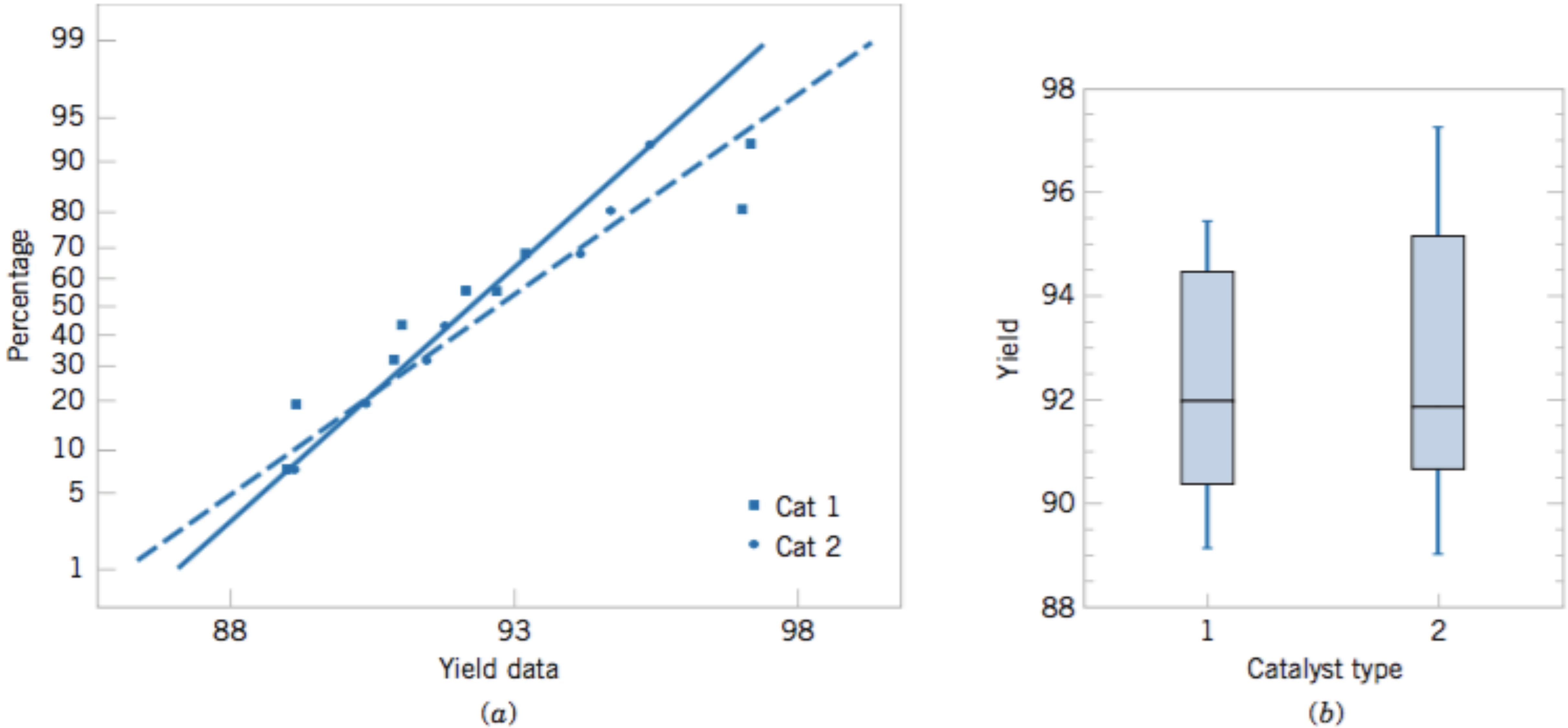


Figure 10-2 Normal probability plot and comparative box plot for the catalyst yield data in Example 10-5. (a) Normal probability plot, (b) Box plots.

Summary

- General procedure for hypothesis test
  - Direct method: define rejection region
  - Using confidence interval
  - Compute p-value

- For several parameters of interest
  - Mean: when variance is known and unknown
  - Variance
  - Sample proportion
  - Comparing two populations

Additional Examples

Quiz

9-47. Medical researchers have developed a new artificial heart constructed primarily of titanium and plastic. The heart will last and operate almost indefinitely once it is implanted in the patient's body, but the battery pack needs to be recharged about every four hours. A random sample of 50 battery packs is selected and subjected to a life test. The average life of these batteries is 4.05 hours. Assume that battery life is normally distributed with standard deviation $\sigma = 0.2$ hour.

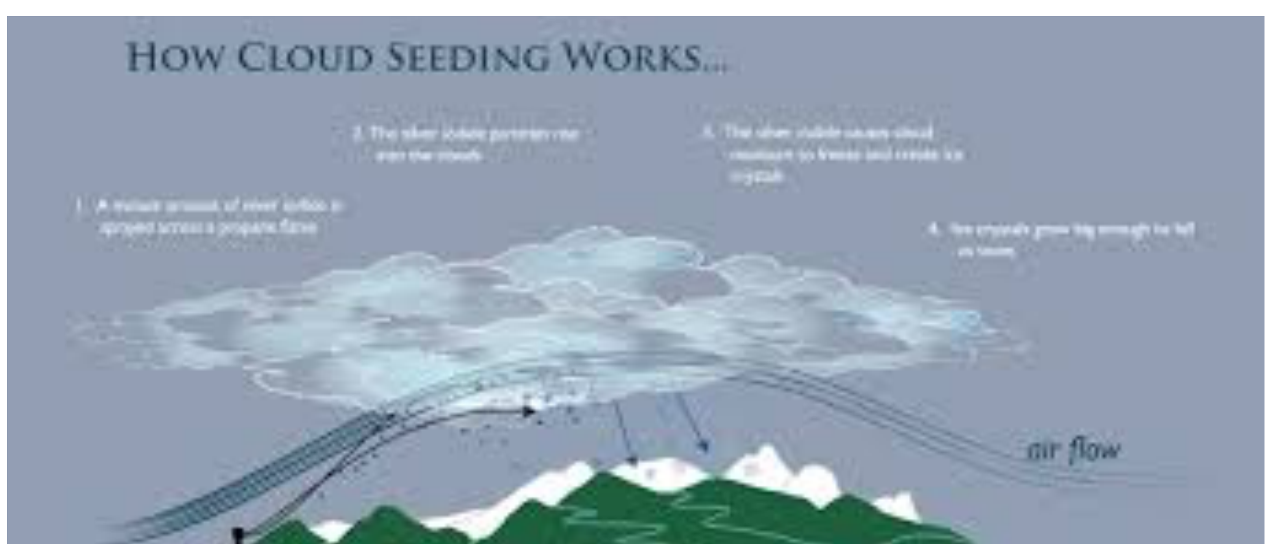
(a) Is there evidence to support the claim that mean battery life exceeds 4 hours? Use $\alpha = 0.05$.

(b) What is the P-value for the test in part (a)?

A1. Cloud seeding

9-60. Cloud seeding has been studied for many decades as a weather modification procedure (for an interesting study of this subject, see the article in *Technometrics*, "A Bayesian Analysis of a Multiplicative Treatment Effect in Weather Modification," Vol. 17, pp. 161–166). The rainfall in acre-feet from 20 clouds that were selected at random and seeded with silver nitrate follows: 18.0, 30.7, 19.8, 27.1, 22.3, 18.8, 31.8, 23.4, 21.2, 27.9, 31.9, 27.1, 25.0, 24.7, 26.9, 21.8, 29.2, 34.8, 26.7, and 31.6. **Assume the true standard deviation is 4.**

(a) Can you support a claim that mean rainfall from seeded clouds exceeds 25 acre-feet?



Solution

Parameter of interest μ

Hypotheses $H_0: \mu = 25$

H1 $: \mu > 25$

Test statistic $Z_0 = \frac{\bar{X} - \mu 0}{\sigma/\sqrt{n}} = \frac{26.035 - 25}{4/\sqrt{20}}$ **=1.157,** $\bar{X}$ **= 26.035,** σ **= 4, n = 20**

Rejection H0 if $Z_0 > Z_\alpha = Z_{0.05}$ **= 1.645**

Fail to reject H0. So H0 is true.

A2. Network response time

Example The response time of a distributed computer system is an important quality characteristic. The system manager wants to know whether the mean response time to a specific type of command exceeds 75 millisec. From past experience, he knows that the standard deviation of response time is 8 millisec.

If the command is executed 25 times and the response time for each trial is recorded. The sample average response time is 79.25 millisec. Formulate an appropriate hypothesis and test the hypothesis.

Solution

Parameter of interest μ

Hypotheses $H_0: \mu = 75$

H1 $: \mu > 75$

$\bar{X}$ = 79.25, σ = 8, n = 25

Test statistic $Z_0 = \frac{\bar{X}-\mu 0}{\sigma/\sqrt{n}} = \frac{79.25-75}{8/\sqrt{25}}$ =2.656

Rejection H0 if $Z_0 > Z_\alpha = Z_{0.05}$ = 1.645

Reject H0. So H1 is true.

A3. Engine controller

EXAMPLE 9-10 Automobile Engine Controller

A semiconductor manufacturer produces controllers used in automobile engine applications. The customer requires that the process fallout or fraction defective at a critical manufacturing step not exceed 0.05 and that the manufacturer demonstrate process capability at this level of quality using $\alpha = 0.05$. The semiconductor manufacturer takes a random sample of 200 devices and finds that four of them are defective. Can the manufacturer demonstrate process capability for the customer?

We may solve this problem using the seven-step hypothesis-testing procedure as follows:

1. **Parameter of Interest:** The parameter of interest is the process fraction defective p.
2. **Null hypothesis:** $H_0\colon p = 0.05$
3. **Alternative hypothesis:** $H_1\colon p < 0.05$
   This formulation of the problem will allow the manufacturer to make a strong claim about process capa bility if the null hypothesis $H_0\colon p = 0.05$ is rejected.
4. The test statistic is (from Equation 9-40)

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}}$$

   where $x = 4$, $n = 200$, and $p_0 = 0.05$.
5. **Reject H_0 if:** Reject $H_0\colon p = 0.05$ if the p-value is less than 0.05.
6. **Computations:** The test statistic is

$$z_0 = \frac{4 - 200(0.05)}{\sqrt{200(0.05)(0.95)}} = -1.95$$

7. **Conclusions:** Since $z_0 = -1.95$, the P-value is $\Phi(-1.95) = 0.0256$, so we reject H_0 and conclude that the process fraction defective p is less than 0.05.

Practical Interpretation: We conclude that the process is capable.