ISE 315: Engineering Statistics

Lecture 10: Hypothesis Testing for a Single Sample (cont'd)

Instructor: Mansur M. Arief, PhD
Industrial and Systems Engineering, KFUPM

Office: 22-219 — Email: mansur.arief@kfupm.edu.sa

Based on Montgomery & Runger, Applied Statistics and Probability for Engineers, 6th Ed.

Lecture 10

Hypothesis Testing: χ^2-Test, Proportion Test, Type II Error, and Practice

Lecture 10 Outline

- Rule review: 1) test statistic vs critical value and 2) P-value vs α
- Chi-square (χ^2) test for the variance σ^2
- Z-test for a proportion p
- Type II error (β) and the α–β tradeoff
- Practice problems (Sets A–D)

Choosing the Right Test

Procedure 4.5; see also Table 4.9

Parameter	Condition	Test Statistic	Distribution	Book Eq.
μ	σ known	$Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Standard Normal	(4.3)

Textbook: Chapter 4 ((4.7) and (4.8))

Choosing the Right Test

Procedure 4.5; see also Table 4.9

Parameter	Condition	Test Statistic	Distribution	Book Eq.
μ	σ known	$Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Standard Normal	(4.3)
μ	σ unknown	$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1}	(4.6)

Textbook: Chapter 4 ((4.7) and (4.8))

Choosing the Right Test

Procedure 4.5; see also Table 4.9

Parameter	Condition	Test Statistic	Distribution	Book Eq.
μ	σ known	$Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Standard Normal	(4.3)
μ	σ unknown	$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1}	(4.6)
σ^2	Normal pop.	$\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$	χ_{n-1}^2	(4.7)

Textbook: Chapter 4 ((4.7) and (4.8))

Choosing the Right Test

Procedure 4.5; see also Table 4.9

Parameter	Condition	Test Statistic	Distribution	Book Eq.
μ	σ known	$Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Standard Normal	(4.3)
μ	σ unknown	$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1}	(4.6)
σ^2	Normal pop.	$\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$	χ_{n-1}^2	(4.7)
p	Large n	$Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$	Standard Normal	(4.8)

Textbook: Chapter 4 ((4.7) and (4.8))

Choosing the Right Test

Procedure 4.5; see also Table 4.9

Parameter	Condition	Test Statistic	Distribution	Book Eq.
μ	σ known	$Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	Standard Normal	(4.3)
μ	σ unknown	$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1}	(4.6)
σ^2	Normal pop.	$\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$	χ_{n-1}^2	(4.7)
p	Large n	$Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$	Standard Normal	(4.8)

Today: the χ^2 row and the p row in detail, plus the α–β tradeoff.

Textbook: Chapter 4 ((4.7) and (4.8))

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α
- **Rules:**

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α
- **Rules:**
  - One-tail (upper): Reject H_0 if test statistic $>$ critical value

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α
- **Rules:**
  - One-tail (upper): Reject H_0 if test statistic $>$ critical value
  - One-tail (lower): Reject H_0 if test statistic $<$ critical value

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α
- **Rules:**
  - One-tail (upper): Reject H_0 if test statistic $>$ critical value
  - One-tail (lower): Reject H_0 if test statistic $<$ critical value
  - Two-tail: Reject H_0 if test statistic is outside both critical values (left or right tail)

Decision Rule: Approach 1 - Rejection Region (Critical Value)

(1) Rejection Region (Critical Value) Approach:

- Compute the test statistic (Z_0, t_0, or χ_0^2)
- Find critical value(s) from the table at significance level α
- **Rules:**
  - One-tail (upper): Reject H_0 if test statistic $>$ critical value
  - One-tail (lower): Reject H_0 if test statistic $<$ critical value
  - Two-tail: Reject H_0 if test statistic is outside both critical values (left or right tail)
- *For χ^2 tests, remember: the distribution is not symmetric.*

Decision Rule: Approach 2 - P-value Approach (Procedure 4.1)

(2) P-value Approach:

Decision Rule: Approach 2 - P-value Approach (Procedure 4.1)

(2) P-value Approach:

- Calculate the P-value (4.4): probability of obtaining your result or more extreme if H_0 is true

Decision Rule: Approach 2 - P-value Approach (Procedure 4.1)

(2) P-value Approach:

- Calculate the P-value (4.4): probability of obtaining your result or more extreme if H_0 is true
- Compare to α

Decision Rule: Approach 2 - P-value Approach (Procedure 4.1)

(2) P-value Approach:

- Calculate the P-value (4.4): probability of obtaining your result or more extreme if H_0 is true
- Compare to α
- **Rule:** Reject H_0 if **P-value** $< \alpha$

Decision Rule: Approach 2 - P-value Approach (Procedure 4.1)

(2) P-value Approach:

- Calculate the P-value (4.4): probability of obtaining your result or more extreme if H_0 is true
- Compare to α
- **Rule:** Reject H_0 if **P-value** $< \alpha$

Both methods lead to the same conclusion!

Chi-Square Test for the Variance

When to use: Testing claims about σ^2 (or σ). The population **must be Normal**.

Chi-Square Test for the Variance

When to use: Testing claims about σ^2 (or σ). The population **must be Normal**.

Test statistic:

$$\chi_0^2 = \frac{(n-1)\,S^2}{\sigma_0^2} \tag{4.7}$$

Under H_0, this follows χ^2_{n-1}.

Chi-Square Test for the Variance

When to use: Testing claims about σ^2 (or σ). The population **must be Normal**.

Test statistic:

$$\chi_0^2 = \frac{(n-1)\,S^2}{\sigma_0^2} \tag{4.7}$$

Under H_0, this follows χ^2_{n-1}.

Key difference from Z and t: The chi-square distribution is *not symmetric*, so critical values for left and right tails are different.

χ^2-Test Decision Rules

Collected in Table 4.7

Test Type	H_1	**Reject H_0 if**
Two-tailed	$\sigma^2 \neq \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha/2,\, n-1}$ or $\chi_0^2 > \chi^2_{\alpha/2,\, n-1}$

χ^2-Test Decision Rules

Collected in Table 4.7

Test Type	H_1	**Reject H_0 if**
Two-tailed	$\sigma^2 \neq \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha/2,\,n-1}$ or $\chi_0^2 > \chi^2_{\alpha/2,\,n-1}$
Right-tailed	$\sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha,\,n-1}$

χ^2-Test Decision Rules

Collected in Table 4.7

Test Type	H_1	**Reject H_0 if**
Two-tailed	$\sigma^2 \neq \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha/2,\, n-1}$ or $\chi_0^2 > \chi^2_{\alpha/2,\, n-1}$
Right-tailed	$\sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha,\, n-1}$
Left-tailed	$\sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha,\, n-1}$

χ^2-Test Decision Rules

Collected in Table 4.7

Test Type	H_1	**Reject H_0 if**
Two-tailed	$\sigma^2 \neq \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha/2,\, n-1}$ or $\chi_0^2 > \chi^2_{\alpha/2,\, n-1}$
Right-tailed	$\sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha,\, n-1}$
Left-tailed	$\sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha,\, n-1}$

Important: $\chi^2_{\alpha,\nu}$ is the value such that $P(\chi^2_\nu > \chi^2_{\alpha,\nu}) = \alpha$.

So χ^2 gives $P(\chi^2 > \chi^2) = 0.95$, i.e., only 5% is in the *left* tail

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Hypotheses: $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed)

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Hypotheses: $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed)

Test statistic (4.7): $\chi_0^2 = \dfrac{(20-1)(0.14)}{0.25} = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Hypotheses: $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed)

Test statistic (4.7): $\chi_0^2 = \dfrac{(20-1)(0.14)}{0.25} = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$

Critical value: $\chi^2_{0.95,\,19} = 10.117$ (left-tail cutoff)

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Hypotheses: $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed)

Test statistic (4.7): $\chi_0^2 = \dfrac{(20-1)(0.14)}{0.25} = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$

Critical value: $\chi^2_{0.95,\,19} = 10.117$ (left-tail cutoff)

Decision: $\chi_0^2 = 10.64 > 10.117$, so it does *not* fall in the rejection region.

Example: Sensor-Fusion Latency Variance

AI Safety Testing Context

An AV team wants to verify that the variance of sensor-fusion latency has been **reduced below** the legacy value $\sigma^2 = 0.25$ ms^2 after a firmware patch. Data: $n = 20$ drives, $s^2 = 0.14$ ms^2. Test at $\alpha = 0.05$.

Hypotheses: $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed)

Test statistic (4.7): $\chi_0^2 = \dfrac{(20-1)(0.14)}{0.25} = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$

Critical value: $\chi^2_{0.95,\,19} = 10.117$ (left-tail cutoff)

Decision: $\chi_0^2 = 10.64 > 10.117$, so it does *not* fall in the rejection region.

Conclusion: **Fail to reject H_0**. Not enough evidence at 5% that the latency variance has been reduced below 0.25 ms^2.

Z-Test for a Proportion p

When to use: Testing a claimed rate p_0 (defect rate, late-delivery rate, conversion rate) with x successes in n trials, $\hat{p} = x/n$.

Z-Test for a Proportion p

When to use: Testing a claimed rate p_0 (defect rate, late-delivery rate, conversion rate) with x successes in n trials, $\hat{p} = x/n$.

Test statistic:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \tag{4.8}$$

Under H_0, approximately standard Normal.

Z-Test for a Proportion p

When to use: Testing a claimed rate p_0 (defect rate, late-delivery rate, conversion rate) with x successes in n trials, $\hat{p} = x/n$.

Test statistic:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \tag{4.8}$$

Under H_0, approximately standard Normal.

Condition: Normal approximation needs $np_0 \geq 5$ and $n(1 - p_0) \geq 5$.

Z-Test for a Proportion p

When to use: Testing a claimed rate p_0 (defect rate, late-delivery rate, conversion rate) with x successes in n trials, $\hat{p} = x/n$.

Test statistic:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \tag{4.8}$$

Under H_0, approximately standard Normal.

Condition: Normal approximation needs $np_0 \geq 5$ and $n(1 - p_0) \geq 5$.

Note: the standard error uses p_0 (the H_0 value), *not* $\hat{p}$ — unlike the CI (3.7).

Type II Error (β) and The α–β Tradeoff

Recap: Two Types of Errors

Definitions in (4.2)

	H_0 **True**	H_0 **False**
Reject H_0	Type I error (α)	Correct! (Power)

Recap: Two Types of Errors

Definitions in (4.2)

	H_0 **True**	H_0 **False**
Reject H_0	Type I error (α)	Correct! (Power)
Fail to reject H_0	Correct!	Type II error (β)

Recap: Two Types of Errors

Definitions in (4.2)

	H_0 **True**	H_0 **False**
Reject H_0	Type I error (α)	Correct! (Power)
Fail to reject H_0	Correct!	Type II error (β)

- $\alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$ – **false alarm**
- $\beta = P(\text{fail to reject } H_0 \mid H_0 \text{ is false})$ – **missed detection**
- **Power** $= 1 - \beta = P(\text{reject } H_0 \mid H_0 \text{ is false})$

The α–β Tradeoff

For a **fixed sample size** n:

The α–β Tradeoff

For a **fixed sample size** *n*:

- Decreasing α (stricter) $\Rightarrow$ β **increases** (harder to detect effects)

The α–β Tradeoff

For a **fixed sample size** *n*:

- Decreasing α (stricter) $\Rightarrow$ β **increases** (harder to detect effects)
- Increasing α (more lenient) $\Rightarrow$ β **decreases** (more false alarms)

The α–β Tradeoff

For a **fixed sample size** *n*:

- Decreasing α (stricter) $\Rightarrow$ β **increases** (harder to detect effects)
- Increasing α (more lenient) $\Rightarrow$ β **decreases** (more false alarms)

How to reduce both α and β?

The α–β Tradeoff

For a **fixed sample size** n:

- Decreasing α (stricter) $\Rightarrow$ β **increases** (harder to detect effects)
- Increasing α (more lenient) $\Rightarrow$ β **decreases** (more false alarms)

How to reduce both α and β?

Get more evidence (i.e. increase the sample size n)

Example: Effect of Changing α

Supply Chain Context

Delivery route time test (Set A, later today): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, P-value $= 0.0455$

Example: Effect of Changing α

Supply Chain Context

Delivery route time test (Set A, later today): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, P-value $= 0.0455$

At $\alpha = 0.05$: P-value $= 0.0455 < 0.05 \Rightarrow$ **Reject H_0**

Example: Effect of Changing α

Supply Chain Context

Delivery route time test (Set A, later today): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, P-value $= 0.0455$

At $\alpha = 0.05$: P-value $= 0.0455 < 0.05 \Rightarrow$ **Reject** H_0

At $\alpha = 0.01$: P-value $= 0.0455 > 0.01 \Rightarrow$ **Fail to reject** H_0

Example: Effect of Changing α

Supply Chain Context

Delivery route time test (Set A, later today): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, P-value $= 0.0455$

At $\alpha = 0.05$: P-value $= 0.0455 < 0.05 \Rightarrow$ **Reject** H_0

At $\alpha = 0.01$: P-value $= 0.0455 > 0.01 \Rightarrow$ **Fail to reject** H_0

Same data, same test statistic, **different conclusion**!

The stricter $\alpha = 0.01$ means we need stronger evidence to reject.
We claim β is now larger – but can we actually *calculate* β?

Example: Effect of Changing α

Supply Chain Context

Delivery route time test (Set A, later today): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, P-value $= 0.0455$

At $\alpha = 0.05$: P-value $= 0.0455 < 0.05 \Rightarrow$ **Reject** H_0

At $\alpha = 0.01$: P-value $= 0.0455 > 0.01 \Rightarrow$ **Fail to reject** H_0

Same data, same test statistic, **different conclusion**!

The stricter $\alpha = 0.01$ means we need stronger evidence to reject.
We claim β is now larger – but can we actually *calculate* β?

Yes! But we need to specify a particular "true" alternative value μ_{true}.
$\beta(\mu_{\text{true}}) = P\big(\text{fail to reject } H_0 \mid \mu = \mu_{\text{true}}\big)$

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

We have $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{76 - 71}{16/\sqrt{64}} = 2.50$ (4.3) and P-value $= 0.0124$.

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

We have $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{76 - 71}{16/\sqrt{64}} = 2.50$ (4.3) and P-value $= 0.0124$.

At $\alpha = 0.05$: P-value $= 0.0124 < 0.05 \Rightarrow$ **Reject** H_0

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

We have $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{76 - 71}{16/\sqrt{64}} = 2.50$ (4.3) and P-value $= 0.0124$.

At $\alpha = 0.05$: P-value $= 0.0124 < 0.05 \Rightarrow$ **Reject** H_0

At $\alpha = 0.01$: P-value $= 0.0124 > 0.01 \Rightarrow$ **Fail to reject** H_0

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

We have $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{76 - 71}{16/\sqrt{64}} = 2.50$ (4.3) and P-value $= 0.0124$.

At $\alpha = 0.05$: P-value $= 0.0124 < 0.05 \Rightarrow$ **Reject H_0**

At $\alpha = 0.01$: P-value $= 0.0124 > 0.01 \Rightarrow$ **Fail to reject H_0**

The stricter $\alpha = 0.01$ means we need stronger evidence to reject.

Example: Effect of Changing α (cont'd)

Container Dwell Time from Lectures 8–9

Container dwell time (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 71$, $\bar{x} = 76$

We have $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{76 - 71}{16/\sqrt{64}} = 2.50$ (4.3) and P-value $= 0.0124$.

At $\alpha = 0.05$: P-value $= 0.0124 < 0.05 \Rightarrow$ **Reject H_0**

At $\alpha = 0.01$: P-value $= 0.0124 > 0.01 \Rightarrow$ **Fail to reject H_0**

The stricter $\alpha = 0.01$ means we need stronger evidence to reject.

We claim β is now larger. To see this, we need a particular "true" value μ_{true}.
$\beta(\mu_{\text{true}}) = P\big(\text{fail to reject } H_0 \mid \mu = \mu_{\text{true}}\big)$

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Scenario	**Parameter Values**	**Interpretation**
H_0 (Null hypothesis)	$\mu = \mu_0 = 71$ h	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 75$ h	What actually holds

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Scenario	Parameter Values	Interpretation
H_0 (Null hypothesis)	$\mu = \mu_0 = 71$ h	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 75$ h	What actually holds

The fail-to-reject interval for H_0 (in terms of $\bar{X}$) is the book's acceptance region:

$$\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{X} \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{4.9}$$

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Scenario	Parameter Values	Interpretation
H_0 (Null hypothesis)	$\mu = \mu_0 = 71$ h	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 75$ h	What actually holds

The fail-to-reject interval for H_0 (in terms of $\bar{X}$) is the book's acceptance region:

$$\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{4.9}$$

At $\alpha = 0.05$: $71 - 1.96 \times 2 \leq \bar{X} \leq 71 + 1.96 \times 2 \Rightarrow 67.08 \leq \bar{X} \leq 74.92$

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Scenario	Parameter Values	Interpretation
H_0 (Null hypothesis)	$\mu = \mu_0 = 71$ h	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 75$ h	What actually holds

The fail-to-reject interval for H_0 (in terms of $\bar{X}$) is the book's acceptance region:

$$\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{X} \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{4.9}$$

At $\alpha = 0.05$: $71 - 1.96 \times 2 \le \bar{X} \le 71 + 1.96 \times 2 \Rightarrow 67.08 \le \bar{X} \le 74.92$

Computing β for a Given μ_{true}

Procedure 4.4, Step 1: the Fail-to-Reject Interval

Recall: $n = 64$, known $\sigma = 16$ h, the operator claims $\mu_0 = 71$ h, and we observed $\bar{x} = 76$ h. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ h.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 75$ h.

Scenario	Parameter Values	Interpretation
H_0 (Null hypothesis)	$\mu = \mu_0 = 71$ h	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 75$ h	What actually holds

The fail-to-reject interval for H_0 (in terms of $\bar{X}$) is the book's acceptance region:

$$\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{X} \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{4.9}$$

At $\alpha = 0.05$: $71 - 1.96 \times 2 \le \bar{X} \le 71 + 1.96 \times 2 \Rightarrow 67.08 \le \bar{X} \le 74.92$

Computing β for a Given μ_{true} (cont'd)

Procedure 4.4, Step 2: Probability under the Truth, via (4.10)

If $\mu_{\text{true}} = 75$ h, then $\beta(\mu_{\text{true}} = 75) = \overbrace{P\left(\bar{X} \in I_{H_0,\alpha} \mid \mu = 75\right)}^{\text{prob. } \bar{X} \text{ in the fail-to-reject region}}$

Computing β for a Given μ_{true} (cont'd)

Procedure 4.4, Step 2: Probability under the Truth, via (4.10)

$$\text{If } \mu_{\text{true}} = 75 \text{ h, then } \beta(\mu_{\text{true}} = 75) = \overbrace{P\left(\bar{X} \in I_{H_0,\alpha} \mid \mu = 75\right)}^{\text{prob. } \bar{X} \text{ in the fail-to-reject region}}$$

- For $\alpha = 0.05$: $I_{H_0,\alpha=0.05} = [67.08,\ 74.92]$

$$\beta(\mu_{\text{true}} = 75) = P\left(67.08 \le \bar{X} \le 74.92 \mid \mu = 75\right) = P\left(\frac{67.08 - 75}{2} \le Z \le \frac{74.92 - 75}{2}\right)$$

$$= P(-3.96 \le Z \le -0.04) = \Phi(-0.04) - \Phi(-3.96) = 0.4840 - 0.0000 = \boxed{0.484}$$

Computing β for a Given μ_{true} (cont'd)

Procedure 4.4, Step 2: Probability under the Truth, via (4.10)

$$\text{If } \mu_{\text{true}} = 75 \text{ h, then } \beta(\mu_{\text{true}} = 75) = \overbrace{P\left(\bar{X} \in I_{H_0,\alpha} \mid \mu = 75\right)}^{\text{prob. } \bar{X} \text{ in the fail-to-reject region}}$$

- For $\alpha = 0.05$: $I_{H_0,\alpha=0.05} = [67.08,\ 74.92]$

$$\beta(\mu_{\text{true}} = 75) = P\left(67.08 \le \bar{X} \le 74.92 \mid \mu = 75\right) = P\left(\frac{67.08 - 75}{2} \le Z \le \frac{74.92 - 75}{2}\right)$$

$$= P(-3.96 \le Z \le -0.04) = \Phi(-0.04) - \Phi(-3.96) = 0.4840 - 0.0000 = \boxed{0.484}$$

- For $\alpha = 0.01$: $I_{H_0,\alpha=0.01} = [65.85,\ 76.15]$

$$\beta(\mu_{\text{true}} = 75) = P\left(65.85 \le \bar{X} \le 76.15 \mid \mu = 75\right) = P\left(\frac{65.85 - 75}{2} \le Z \le \frac{76.15 - 75}{2}\right)$$

$$= P(-4.58 \le Z \le 0.58) = \Phi(0.58) - 0 = \boxed{0.719}$$

Computing β for a Given μ_{true} (cont'd)

Procedure 4.4, Step 2: Probability under the Truth, via (4.10)

$$\text{If } \mu_{\text{true}} = 75 \text{ h, then } \beta(\mu_{\text{true}} = 75) = \overbrace{P\left(\bar{X} \in I_{H_0,\alpha} \mid \mu = 75\right)}^{\text{prob. } \bar{X} \text{ in the fail-to-reject region}}$$

- For $\alpha = 0.05$: $I_{H_0,\alpha=0.05} = [67.08,\ 74.92]$

$$\beta(\mu_{\text{true}} = 75) = P(67.08 \le \bar{X} \le 74.92 \mid \mu = 75) = P\left(\frac{67.08 - 75}{2} \le Z \le \frac{74.92 - 75}{2}\right)$$

$$= P(-3.96 \le Z \le -0.04) = \Phi(-0.04) - \Phi(-3.96) = 0.4840 - 0.0000 = \boxed{0.484}$$

- For $\alpha = 0.01$: $I_{H_0,\alpha=0.01} = [65.85,\ 76.15]$

$$\beta(\mu_{\text{true}} = 75) = P(65.85 \le \bar{X} \le 76.15 \mid \mu = 75) = P\left(\frac{65.85 - 75}{2} \le Z \le \frac{76.15 - 75}{2}\right)$$

$$= P(-4.58 \le Z \le 0.58) = \Phi(0.58) - 0 = \boxed{0.719}$$

$$\text{By } \underbrace{\text{reducing } \alpha\ (0.05 \to 0.01)}_{\text{fewer false alarms}},\ \underbrace{\beta \text{ increases } 0.484 \to 0.719}_{\text{more failures to reject}}.$$

What We Have So Far

- **Four test types:** Z-test (4.3), t-test (4.6), χ^2-test (4.7), proportion Z-test (4.8)

What We Have So Far

- **Four test types:** Z-test (4.3), t-test (4.6), χ^2-test (4.7), proportion Z-test (4.8)
- **Two decision approaches:** Compare test statistic to critical value, *or* compare P-value to α (Procedure 4.1)

What We Have So Far

- **Four test types:** Z-test (4.3), t-test (4.6), χ^2-test (4.7), proportion Z-test (4.8)
- **Two decision approaches:** Compare test statistic to critical value, *or* compare P-value to α (Procedure 4.1)
- α–β **tradeoff:** Lowering α increases β (and vice versa) for fixed n

What We Have So Far

- **Four test types:** Z-test (4.3), t-test (4.6), χ^2-test (4.7), proportion Z-test (4.8)
- **Two decision approaches:** Compare test statistic to critical value, *or* compare P-value to α (Procedure 4.1)
- α–β **tradeoff:** Lowering α increases β (and vice versa) for fixed *n*
- **Computing** β**:** Specify μ_{true}, find $P(\bar{X} \in \text{fail-to-reject region} \mid \mu = \mu_{\text{true}})$ — (4.9) and (4.10)

What We Have So Far

- **Four test types:** Z-test (4.3), t-test (4.6), χ^2-test (4.7), proportion Z-test (4.8)
- **Two decision approaches:** Compare test statistic to critical value, *or* compare P-value to α (Procedure 4.1)
- α–β **tradeoff:** Lowering α increases β (and vice versa) for fixed *n*
- **Computing** β**:** Specify μ_{true}, find $P(\bar{X} \in \text{fail-to-reject region} \mid \mu = \mu_{\text{true}})$ — (4.9) and (4.10)

Now let's put it all together with practice problems.

Practice Problems

Sets A–D, each with the seven steps of Procedure 4.2

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Steps 4–5: Z-test (4.3); reject if $|Z_0| > z_{0.025} = 1.96$.

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Steps 4–5: Z-test (4.3); reject if $|Z_0| > z_{0.025} = 1.96$.

Step 6: $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{154 - 150}{14/\sqrt{49}} = \dfrac{4}{2} = 2.00$; P-value (4.5) $= 2(1 - \Phi(2.00)) = 2(0.0228) = 0.0455$

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Steps 4–5: Z-test (4.3); reject if $|Z_0| > z_{0.025} = 1.96$.

Step 6: $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{154 - 150}{14/\sqrt{49}} = \dfrac{4}{2} = 2.00$; P-value (4.5) $= 2(1-\Phi(2.00)) = 2(0.0228) = 0.0455$

Step 7: $|2.00| = 2.00 > 1.96 \Rightarrow$ **Reject** H_0.
Sufficient evidence at 5% that the mean route time differs from 150 minutes.

Set A: Z-Test, Two-Tailed (Delivery Route Time)

Supply Chain Context

A logistics planner tests whether the mean delivery route time differs from the scheduled 150 minutes. $n = 49$ routes, $\bar{x} = 154$ min, $\sigma = 14$ min, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Steps 4–5: Z-test (4.3); reject if $|Z_0| > z_{0.025} = 1.96$.

Step 6: $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{154 - 150}{14/\sqrt{49}} = \dfrac{4}{2} = 2.00$; P-value (4.5) $= 2(1 - \Phi(2.00)) = 2(0.0228) = 0.0455$

Step 7: $|2.00| = 2.00 > 1.96 \Rightarrow$ **Reject H_0**.
Sufficient evidence at 5% that the mean route time differs from 150 minutes.

α–β: At $\alpha = 0.01$: P-value $= 0.0455 > 0.01 \Rightarrow$ fail to reject. β increases.

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ unknown. $H_0 : \mu = 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ unknown. $H_0 : \mu = 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Steps 4–5: t-test (4.6), df $= 15$; reject if $t_0 > t_{0.05,\,15} = 1.753$.

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ unknown. $H_0 : \mu = 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Steps 4–5: t-test (4.6), df $= 15$; reject if $t_0 > t_{0.05,\,15} = 1.753$.

Step 6: $t_0 = \dfrac{10.8 - 10.0}{1.2/\sqrt{16}} = \dfrac{0.8}{0.3} = 2.67$; P-value bounds (Procedure 4.3): $t_{0.01,\,15} = 2.602 < 2.67 < 2.947 = t_{0.005,\,15}$, so $0.005 < \text{P} < 0.01$ (≈ 0.0088)

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ unknown. $H_0 : \mu = 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Steps 4–5: t-test (4.6), df $= 15$; reject if $t_0 > t_{0.05,\,15} = 1.753$.

Step 6: $t_0 = \dfrac{10.8 - 10.0}{1.2/\sqrt{16}} = \dfrac{0.8}{0.3} = 2.67$; P-value bounds (Procedure 4.3): $t_{0.01,\,15} = 2.602 < 2.67 < 2.947 = t_{0.005,\,15}$, so $0.005 < \mathrm{P} < 0.01$ (≈ 0.0088)

Step 7: $2.67 > 1.753 \Rightarrow$ **Reject H_0**.
Sufficient evidence at 5% that the mean activations exceed 10.0 per user.

Set B: t-Test, Right-Tailed (Feature Activations)

Startup Context

A startup founder claims the new onboarding flow pushes the mean number of first-week feature activations above 10.0 per user. $n = 16$ pilot users, $\bar{x} = 10.8$, $s = 1.2$, $\alpha = 0.05$.

Steps 1–3: Parameter: μ, σ unknown. $H_0 : \mu = 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Steps 4–5: t-test (4.6), df $= 15$; reject if $t_0 > t_{0.05,15} = 1.753$.

Step 6: $t_0 = \dfrac{10.8 - 10.0}{1.2/\sqrt{16}} = \dfrac{0.8}{0.3} = 2.67$; P-value bounds (Procedure 4.3): $t_{0.01,15} = 2.602 < 2.67 < 2.947 = t_{0.005,15}$, so $0.005 < \text{P} < 0.01$ (≈ 0.0088)

Step 7: $2.67 > 1.753 \Rightarrow$ **Reject H_0**.
Sufficient evidence at 5% that the mean activations exceed 10.0 per user.

α–β: At $\alpha = 0.01$: $t_{0.01,15} = 2.602$. Since $2.67 > 2.602$, still reject. β increases but the conclusion holds.

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Steps 1–3: Parameter: σ^2, Normal population. $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Steps 1–3: Parameter: σ^2, Normal population. $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Steps 4–5: χ^2-test (4.7), df $= 19$; reject if $\chi^2_0 < \chi^2_{0.95,\,19} = 10.117$.

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Steps 1–3: Parameter: σ^2, Normal population. $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Steps 4–5: χ^2-test (4.7), df $= 19$; reject if $\chi_0^2 < \chi^2_{0.95,\,19} = 10.117$.

Step 6: $\chi_0^2 = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$; P-value $= P(\chi^2_{19} < 10.64) \approx 0.065$

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Steps 1–3: Parameter: σ^2, Normal population. $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Steps 4–5: χ^2-test (4.7), df $= 19$; reject if $\chi_0^2 < \chi^2_{0.95,\,19} = 10.117$.

Step 6: $\chi_0^2 = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$; P-value $= P(\chi^2_{19} < 10.64) \approx 0.065$

Step 7: $10.64 > 10.117$ (not in rejection region) $\Rightarrow$ **Fail to reject H_0**.
Not enough evidence at 5% that the variance is below 0.25 ms^2.

Set C: χ^2-Test, Left-Tailed (Sensor-Fusion Latency Variance)

AI Safety Testing Context

The AV team tests whether the latency variance has been reduced below 0.25 ms^2 after the firmware patch. $n = 20$ drives, $s^2 = 0.14$ ms^2, $\alpha = 0.05$.

Steps 1–3: Parameter: σ^2, Normal population. $H_0 : \sigma^2 = 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Steps 4–5: χ^2-test (4.7), df $= 19$; reject if $\chi_0^2 < \chi^2_{0.95,\,19} = 10.117$.

Step 6: $\chi_0^2 = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$; P-value $= P(\chi^2_{19} < 10.64) \approx 0.065$

Step 7: $10.64 > 10.117$ (not in rejection region) $\Rightarrow$ **Fail to reject H_0**.
Not enough evidence at 5% that the variance is below 0.25 ms^2.

α–β: At $\alpha = 0.10$: $\chi^2_{0.90,\,19} = 11.651$. Since $10.64 < 11.651$, now reject. β decreases.

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Steps 1–3: Parameter: p. $H_0 : p = 0.05$ vs $H_1 : p > 0.05$ (right-tailed). Check: $np_0 = 20 \geq 5$, $n(1 - p_0) = 380 \geq 5$. ✓

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Steps 1–3: Parameter: p. $H_0 : p = 0.05$ vs $H_1 : p > 0.05$ (right-tailed). Check: $np_0 = 20 \geq 5$, $n(1 - p_0) = 380 \geq 5$. ✓

Steps 4–5: Proportion Z-test (4.8); reject if $Z_0 > z_{0.05} = 1.645$.

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Steps 1–3: Parameter: p. $H_0 : p = 0.05$ vs $H_1 : p > 0.05$ (right-tailed). Check: $np_0 = 20 \geq 5$, $n(1 - p_0) = 380 \geq 5$. ✓

Steps 4–5: Proportion Z-test (4.8); reject if $Z_0 > z_{0.05} = 1.645$.

Step 6: $Z_0 = \dfrac{0.07 - 0.05}{\sqrt{0.05 \times 0.95/400}} = \dfrac{0.02}{0.01090} = 1.84$; P-value (4.5) $= 1 - \Phi(1.84) = 0.0329$

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Steps 1–3: Parameter: p. $H_0 : p = 0.05$ vs $H_1 : p > 0.05$ (right-tailed). Check: $np_0 = 20 \geq 5$, $n(1 - p_0) = 380 \geq 5$. ✓

Steps 4–5: Proportion Z-test (4.8); reject if $Z_0 > z_{0.05} = 1.645$.

Step 6: $Z_0 = \dfrac{0.07 - 0.05}{\sqrt{0.05 \times 0.95/400}} = \dfrac{0.02}{0.01090} = 1.84$; P-value (4.5) $= 1 - \Phi(1.84) = 0.0329$

Step 7: $1.84 > 1.645 \Rightarrow$ **Reject H_0**.
Sufficient evidence at 5% that the late-delivery rate exceeds the 5% contract level.

Set D: Proportion Z-Test, Right-Tailed (Late Deliveries)

Supply Chain Context

A carrier's contract requires at most 5% late arrivals. An audit samples $n = 400$ shipments and finds 28 late, so $\hat{p} = 28/400 = 0.07$. Is the late rate **above** the contract level? $\alpha = 0.05$.

Steps 1–3: Parameter: p. $H_0 : p = 0.05$ vs $H_1 : p > 0.05$ (right-tailed). Check: $np_0 = 20 \geq 5$, $n(1 - p_0) = 380 \geq 5$. ✓

Steps 4–5: Proportion Z-test (4.8); reject if $Z_0 > z_{0.05} = 1.645$.

Step 6: $Z_0 = \dfrac{0.07 - 0.05}{\sqrt{0.05 \times 0.95/400}} = \dfrac{0.02}{0.01090} = 1.84$; P-value (4.5) $= 1 - \Phi(1.84) = 0.0329$

Step 7: $1.84 > 1.645 \Rightarrow$ **Reject H_0.**
Sufficient evidence at 5% that the late-delivery rate exceeds the 5% contract level.

α–β: At $\alpha = 0.01$: $z_{0.01} = 2.326$. Since $1.84 < 2.326$, fail to reject. β increases.

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A
Test	Z (4.3)	t (4.6)	χ^2 (4.7)	Z (4.8)

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A
Test	Z (4.3)	t (4.6)	χ^2 (4.7)	Z (4.8)
Tail	Two	Right	Left	Right

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A
Test	Z (4.3)	t (4.6)	χ^2 (4.7)	Z (4.8)
Tail	Two	Right	Left	Right
Decision	Reject H_0	Reject H_0	Fail to reject	Reject H_0

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A
Test	Z (4.3)	t (4.6)	χ^2 (4.7)	Z (4.8)
Tail	Two	Right	Left	Right
Decision	Reject H_0	Reject H_0	Fail to reject	Reject H_0
P-value	0.0455	0.0088	0.065	0.0329

Summary: Comparing the Four Sets

See Table 4.10 in the textbook

	Set A	**Set B**	**Set C**	**Set D**
Parameter	μ	μ	σ^2	p
σ known?	Yes	No	N/A	N/A
Test	Z (4.3)	t (4.6)	χ^2 (4.7)	Z (4.8)
Tail	Two	Right	Left	Right
Decision	Reject H_0	Reject H_0	Fail to reject	Reject H_0
P-value	0.0455	0.0088	0.065	0.0329

Same seven-step procedure (Procedure 4.2), only the test statistic and distribution change!

Lecture 10 Summary

- The χ^2-test (4.7) is used for testing variance (σ^2). Requires normality.

Lecture 10 Summary

- The χ^2-test (4.7) is used for testing variance (σ^2). Requires normality.
- The proportion Z-test (4.8) handles claimed rates; check np_0 and $n(1-p_0)$.

Lecture 10 Summary

- The χ^2-test (4.7) is used for testing variance (σ^2). Requires normality.
- The proportion Z-test (4.8) handles claimed rates; check np_0 and $n(1 - p_0)$.
- Type I error (α) and Type II error (β) have an inverse relationship for fixed n. Increasing n reduces both.

Lecture 10 Summary

- The χ^2-test (4.7) is used for testing variance (σ^2). Requires normality.
- The proportion Z-test (4.8) handles claimed rates; check np_0 and $n(1 - p_0)$.
- Type I error (α) and Type II error (β) have an inverse relationship for fixed n. Increasing n reduces both.
- The same seven steps (Procedure 4.2) work for every test:
  1. Parameter of interest
  2. Null hypothesis H_0
  3. Alternative hypothesis H_1
  4. Test statistic
  5. Rejection criterion for the chosen α
  6. Computations
  7. Conclusion in context

Lecture 10 Summary

- The χ^2-test (4.7) is used for testing variance (σ^2). Requires normality.
- The proportion Z-test (4.8) handles claimed rates; check np_0 and $n(1 - p_0)$.
- Type I error (α) and Type II error (β) have an inverse relationship for fixed n. Increasing n reduces both.
- The same seven steps (Procedure 4.2) work for every test:
  1. Parameter of interest
  2. Null hypothesis H_0
  3. Alternative hypothesis H_1
  4. Test statistic
  5. Rejection criterion for the chosen α
  6. Computations
  7. Conclusion in context
- **Next:** Chapter 5 – Hypothesis testing for two samples

Announcements

- Quiz #3 on Chapter 4 next Tuesday

Announcements

- Quiz #3 on Chapter 4 next Tuesday
- HW #3 due next week, HW #4 on Chapter 4 due in two weeks

Announcements

- Quiz #3 on Chapter 4 next Tuesday
- HW #3 due next week, HW #4 on Chapter 4 due in two weeks
- Mid-semester evaluation is out (please fill it out by next week)

Announcements

- Quiz #3 on Chapter 4 next Tuesday
- HW #3 due next week, HW #4 on Chapter 4 due in two weeks
- Mid-semester evaluation is out (please fill it out by next week)
- Make-up quiz and homeworks will be posted by Thursday EOD

Announcements

- Quiz #3 on Chapter 4 next Tuesday
- HW #3 due next week, HW #4 on Chapter 4 due in two weeks
- Mid-semester evaluation is out (please fill it out by next week)
- Make-up quiz and homeworks will be posted by Thursday EOD
- Recording materials (up to Lecture 10) will be made available by Thursday EOD