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, 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
- Type II error (β) and the α–β tradeoff
- Practice problems

Choosing the Right Test

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

Today: We cover the χ^2 row in detail and discuss the α–β tradeoff.

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: distribution is not symmetric.*

Decision Rule: Approach 2 - P-value Approach

(2) P-value Approach:

- Calculate the P-value (probability of obtaining your result or more extreme if H_0 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**.

Test statistic:

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

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

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 $\chi^2_{0.95,\,19}$ gives $P(\chi^2_{19} > \chi^2_{0.95,\,19}) = 0.95$, i.e., only 5% is in the *left* tail.

Example: Pipe Wall Thickness Variance

A quality engineer wants to verify that the variance of pipe wall thickness has been **reduced below** the historical value $\sigma^2 = 0.25$ mm^2. Data: $n = 20$, $s^2 = 0.14$ mm^2. Test at $\alpha = 0.05$.

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

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

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 variance has been reduced below 0.25 mm^2.

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

Recap: Two Types of Errors

	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*:

- 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

Pipeline pressure test (Set A): $n = 49$, $\sigma = 14$, $\mu_0 = 150$, $Z_0 = 2.00$, 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_1.
$\beta(\mu_1) = P\big(\text{fail to reject } H_0 \mid \mu = \mu_1\big)$

Example: Effect of Changing α

Crude oil viscosity (Lectures 8–9): $n = 64$, $\sigma = 16$, $\mu_0 = 115$, $\bar{x} = 120$

We have $Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{120 - 115}{16/\sqrt{64}} = 2.50$ 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}

Recall: $n = 64$, known $\sigma = 16$ cP, supplier claims $\mu_0 = 115$ cP, and we observed $\bar{x} = 120$ cP. The standard error is $\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{16}{8} = 2$ cP.

Assume (for illustration) that the true mean is $\mu_{\text{true}} = 119$ cP.

Scenario	Parameter Values	Interpretation
H_0 (Null hypothesis)	$\mu = \mu_0 = 115$ cP	Assumed for testing
Truth (unknown in reality)	$\mu = \mu_{\text{true}} = 119$ cP	What actually holds

Calculate the "fail-to-reject" interval for H_0 (in terms of $\bar{X}$) : $\mu_0 \pm z_{\alpha/2} \cdot \text{SE}$

At $\alpha = 0.05 : 115 - 1.96 \times 2 \leq \bar{X} \leq 115 + 1.96 \times 2 \Rightarrow 111.08 \leq \bar{X} \leq 118.92$

At $\alpha = 0.01 : 115 - 2.576 \times 2 \leq \bar{X} \leq 115 + 2.576 \times 2 \Rightarrow 109.85 \leq \bar{X} \leq 120.15$

Let's call these intervals $I_{H_0,\alpha=0.05}$ and $I_{H_0,\alpha=0.01}$.

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

$$\text{If } \mu_{\text{true}} = 119 \text{ cP, then } \beta(\mu_{\text{true}} = 119) = \overbrace{P\left(\bar{X} \in I_{H_0,\alpha} \mid \mu = 119\right)}^{\text{prob. } \bar{X} \text{ in the fail-to-reject region}}$$

- For $\alpha = 0.05$, $I_{H_0,\alpha=0.05} = [111.08,\ 118.92]$

$$\beta(\mu_{\text{true}} = 119) = P\left(111.08 \leq \bar{X} \leq 118.92 \mid \mu = 119\right) = P\left(\frac{111.08 - 119}{2} \leq Z \leq \frac{118.92 - 119}{2}\right)$$
$$= P(-3.96 \leq Z \leq -0.04) = \Phi(-0.04) - \Phi(-3.96) = 0.4840 - 0.0000 = \boxed{0.484}$$

- For $\alpha = 0.01$: $I_{H_0,\alpha=0.01} = [109.85,\ 120.15]$

$$\beta(\mu_{\text{true}} = 119) = P\left(109.85 \leq \bar{X} \leq 120.15 \mid \mu = 119\right) = P\left(\frac{109.85 - 119}{2} \leq Z \leq \frac{120.15 - 119}{2}\right)$$
$$= P(-4.58 \leq Z \leq 0.58) = \Phi(0.58) - 0 = \boxed{0.718}$$

$$\text{By } \underbrace{\text{reducing } \alpha\ (0.05 \to 0.01)}_{\text{fewer false alarms}},\ \underbrace{\beta \text{ increases } 0.484 \to 0.718}_{\text{more failures to reject}}.$$

What We Have So Far

- **Three test types:** Z-test (σ known or proportion), t-test (σ unknown), χ^2-test (variance)
- **Two decision approaches:** Compare test statistic to critical value, *or* compare P-value to α
- α–β **tradeoff:** Lowering α increases β (and vice versa) for fixed n
- **Computing** β**:** Specify a true alternative μ_{true}, find $P(\bar{X} \in \text{fail-to-reject region} \mid \mu = \mu_{\text{true}})$

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

Practice Problems

Set A: Z-Test, Two-Tailed (Pipeline Pressure)

A petroleum engineer tests whether the mean operating pressure differs from 150 psi. $n = 49$, $\bar{x} = 154$ psi, $\sigma = 14$ psi, $\alpha = 0.05$.

Step 1–2: Parameter: μ, σ known. $H_0 : \mu = 150$ vs $H_1 : \mu \neq 150$ (two-tailed).

Step 3: Z-test. $Z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{154 - 150}{14/\sqrt{49}} = \dfrac{4}{2} = 2.00$

Step 4: Reject if $|Z_0| > z_{0.025} = 1.96$. P-value $= 2(1 - \Phi(2.00)) = 2(0.0228) = 0.0455$

Step 5–6: $|2.00| = 2.00 > 1.96 \Rightarrow$ **Reject H_0**.
Sufficient evidence at 5% that the mean pressure differs from 150 psi.

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

Set B: t-Test, Right-Tailed (Drilling Fluid Density)

A drilling engineer claims the mean fluid density exceeds 10.0 lb/gal. $n = 16$, $\bar{x} = 10.8$ lb/gal, $s = 1.2$ lb/gal, $\alpha = 0.05$.

Step 1–2: Parameter: μ, σ unknown. $H_0 : \mu \leq 10.0$ vs $H_1 : \mu > 10.0$ (right-tailed).

Step 3: t-test, df $= 15$. $t_0 = \dfrac{10.8 - 10.0}{1.2/\sqrt{16}} = \dfrac{0.8}{0.3} = 2.67$

Step 4: Reject if $t_0 > t_{0.05,\,15} = 1.753$. P-value $= P(t_{15} > 2.67) \approx 0.0088$

Step 5–6: $2.67 > 1.753 \Rightarrow$ **Reject** H_0.
Sufficient evidence at 5% that the mean density exceeds 10.0 lb/gal.

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

Set C: χ^2-Test, Left-Tailed (Pipe Thickness Variance)

A quality engineer tests whether the variance has been reduced below 0.25 mm^2. $n = 20$, $s^2 = 0.14$ mm^2, $\alpha = 0.05$.

Step 1–2: Parameter: σ^2. $H_0 : \sigma^2 \geq 0.25$ vs $H_1 : \sigma^2 < 0.25$ (left-tailed).

Step 3: χ^2-test, df $= 19$. $\chi_0^2 = \dfrac{19 \times 0.14}{0.25} = \dfrac{2.66}{0.25} = 10.64$

Step 4: Reject if $\chi_0^2 < \chi^2_{0.95,\,19} = 10.117$. P-value $= P(\chi^2_{19} < 10.64) \approx 0.0646$

Step 5–6: $10.64 > 10.117$ (not in rejection region) $\Rightarrow$ **Fail to reject H_0**.
Not enough evidence at 5% that the variance is below 0.25 mm^2.

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

Summary: Comparing the Three Sets

	Set A	**Set B**	**Set C**
Parameter	μ	μ	σ^2
σ known?	Yes	No	N/A
Test	Z-test	t-test	χ^2-test
Tail	Two	Right	Left
Decision	Reject H_0	Reject H_0	Fail to reject
P-value	0.0455	0.0088	0.0646

Same 6-step procedure, only the test statistic and distribution change!

Lecture 10 Summary

- The χ^2-test is used for testing variance (σ^2). Requires normality.
- Type I error (α) and Type II error (β) have an inverse relationship for fixed n.
- Decreasing $\alpha \Rightarrow \beta$ increases. Increasing n reduces both.
- The same 6-step procedure works for every test:
  1. Identify parameter and conditions
  2. State H_0, H_1, α
  3. Choose test statistic (Z_0, t_0, or χ_0^2)
  4. Find rejection region or P-value
  5. Compute the test statistic
  6. Decide and state conclusion in context
- **Next:** Chapter 10 – Hypothesis testing for two samples

Announcements

- Quiz #3 on Chapter 9 next Tuesday
- HW #3 due next week, HW #4 on Chapter 9 due next 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