Technical University of Košice

Faculty of Electrical Engineering and Informatics

Applied Statistics

Exercise Session 5

Worked Solutions in R

Jana Fortes

One-Sample Parametric Tests

Exercise 5.1 Normal Mean: Variance Known

Problem statement (textbook, p. 62)

The required length of a manufactured part is 10 cm. Nine randomly selected parts have the following lengths [cm]: 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. Assume a normal population, with known variance $\sigma^2 = 0.0016$.

Test whether the population mean length:

a) differs from 10 cm, at $\alpha = 0.05$;
b) is less than 10 cm, at $\alpha = 0.01$;
c) is greater than 10 cm, at $\alpha = 0.01$.

R solution: one-sample Z-test

```
x = c(10.24,10.12,9.91,10.19,9.78,10.14,9.86,10.17,10.05)
mu0 = 10
population_variance = 0.0016
sample_mean = mean(x)
sigma = sqrt(population_variance)
n = length(x)
statistic = (sample_mean - mu0)*sqrt(n)/sigma  # Test statistic
```

a) H0: mu = 10; H1: mu != 10

```
alpha1 = 0.05  # Significance level
critical = qnorm(1 - alpha1/2)
cat("Critical region: (-Inf,", -critical, ") union (",
    critical, ", Inf)\n")

# Decision using the critical region
if (abs(statistic) > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using TeachingDemos

```
test_result = TeachingDemos::z.test(x, mu = mu0, sd = sigma,
    alternative = "two.sided", conf.level = 1 - alpha1)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha1) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.1 (continued)

b) H0: mu = 10; H1: mu < 10

```
alpha2 = 0.01  # Significance level
critical = qnorm(1 - alpha2)
cat("Critical region: (-Inf,", -critical, ")\n")

# Decision using the critical region
if (statistic < -critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using TeachingDemos

```
test_result = TeachingDemos::z.test(x, mu = mu0, sd = sigma,
    alternative = "less", conf.level = 1 - alpha2)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha2) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

c) H0: mu = 10; H1: mu > 10

```
alpha3 = 0.01  # Significance level
critical = qnorm(1 - alpha3)
cat("Critical region: (", critical, ", Inf)\n")

# Decision using the critical region
if (statistic > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using TeachingDemos

```
test_result = TeachingDemos::z.test(x, mu = mu0, sd = sigma,
    alternative = "greater", conf.level = 1 - alpha3)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha3) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.2 Normal Mean: Variance Unknown

Problem statement (textbook, p. 64)

The required length of a manufactured part is 10 cm. Nine randomly selected parts have the following lengths [cm]: 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. Assume a normal population.

Test whether the population mean length:

a) differs from 10 cm, at $\alpha = 0.05$;
b) is less than 10 cm, at $\alpha = 0.01$;
c) is greater than 10 cm, at $\alpha = 0.01$.

R solution: one-sample t-test

```
x = c(10.24,10.12,9.91,10.19,9.78,10.14,9.86,10.17,10.05)
mu0 = 10
sample_mean = mean(x)
sample_sd = sd(x)
n = length(x)
statistic = (sample_mean - mu0)*sqrt(n)/sample_sd
```

a) H0: mu = 10; H1: mu != 10

```
alpha1 = 0.05  # Significance level
critical = qt(1 - alpha1/2, n - 1)
cat("Critical region: (-Inf,", -critical, ") union (",
    critical, ", Inf)\n")

# Decision using the critical region
if (abs(statistic) > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using stats

```
test_result = stats::t.test(x, mu = mu0,
    alternative = "two.sided", conf.level = 1 - alpha1)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha1) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.2 (continued)

b) H0: mu = 10; H1: mu < 10

```
alpha2 = 0.01  # Significance level
critical = qt(1 - alpha2, n - 1)
cat("Critical region: (-Inf,", -critical, ")\n")

# Decision using the critical region
if (statistic < -critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using stats

```
test_result = stats::t.test(x, mu = mu0,
    alternative = "less", conf.level = 1 - alpha2)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha2) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

c) H0: mu = 10; H1: mu > 10

```
alpha3 = 0.01  # Significance level
critical = qt(1 - alpha3, n - 1)
cat("Critical region: (", critical, ", Inf)\n")

# Decision using the critical region
if (statistic > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using stats

```
test_result = stats::t.test(x, mu = mu0,
    alternative = "greater", conf.level = 1 - alpha3)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha3) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.3 Normal Variance: Mean Known

Problem statement (textbook, p. 67)

The required part length is 10 cm. Nine parts were randomly selected. Their lengths [cm] are: 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. Assume a normal population. The stated variability corresponds to a standard deviation $\sigma_0 = 0.1$ cm (variance $\sigma_0^2 = 0.01$ cm^2).

For this exercise, take the population mean $\mu = 10$ cm as known.

At $\alpha = 0.05$, test whether the population variance:
a) differs from the stated variance;
b) is smaller than the stated variance;
c) is larger than the stated variance.

R solution: chi-square test with n degrees of freedom

```
x = c(10.24,10.12,9.91,10.19,9.78,10.14,9.86,10.17,10.05)
mu_known = 10
variance0 = 0.01
alpha = 0.05
n = length(x)
variance_known_mean = sum((x - mu_known)^2)/n
statistic = n*variance_known_mean/variance0
```

a) H0: sigma^2 = 0.01; H1: sigma^2 != 0.01

```
critical_lower = qchisq(alpha/2, n)
critical_upper = qchisq(1 - alpha/2, n)
cat("Critical region: (0,", critical_lower, ") union (",
    critical_upper, ", Inf)\n")

# Decision using the critical region
if (statistic < critical_lower || statistic > critical_upper) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

b) H0: sigma^2 = 0.01; H1: sigma^2 < 0.01

```
critical = qchisq(alpha, n)
cat("Critical region: (0,", critical, ")\n")

# Decision using the critical region
if (statistic < critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

c) H0: sigma^2 = 0.01; H1: sigma^2 > 0.01

```
critical = qchisq(1 - alpha, n)
cat("Critical region: (", critical, ", Inf)\n")

# Decision using the critical region
if (statistic > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Exercise 5.4 Normal Variance: Mean Unknown

Problem statement (textbook, p. 69)

Nine parts were randomly selected from a box. Their lengths [cm] are: 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. Assume a normal population. The stated variability corresponds to a standard deviation $\sigma_0 = 0.1$ cm (variance $\sigma_0^2 = 0.01$ cm^2).

At $\alpha = 0.05$, test whether the population variance:
a) differs from the stated variance;
b) is smaller than the stated variance;
c) is larger than the stated variance.

R solution: chi-square test with n-1 degrees of freedom

```
x = c(10.24,10.12,9.91,10.19,9.78,10.14,9.86,10.17,10.05)
variance0 = 0.01
alpha = 0.05
n = length(x)
sample_variance = var(x)
statistic = (n - 1)*sample_variance/variance0
```

a) H0: sigma^2 = 0.01; H1: sigma^2 != 0.01

```
critical_lower = qchisq(alpha/2, n - 1)
critical_upper = qchisq(1 - alpha/2, n - 1)
cat("Critical region: (0,", critical_lower, ") union (",
    critical_upper, ", Inf)\n")

# Decision using the critical region
if (statistic < critical_lower || statistic > critical_upper) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using DescTools

```
test_result = DescTools::VarTest(x, sigma.squared = variance0,
    alternative = "two.sided", conf.level = 1 - alpha)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.4 (continued)

b) H0: sigma^2 = 0.01; H1: sigma^2 < 0.01

```
critical = qchisq(alpha, n - 1)
cat("Critical region: (0,", critical, ")\n")

# Decision using the critical region
if (statistic < critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using DescTools

```
test_result = DescTools::VarTest(x, sigma.squared = variance0,
    alternative = "less", conf.level = 1 - alpha)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

c) H0: sigma^2 = 0.01; H1: sigma^2 > 0.01

```
critical = qchisq(1 - alpha, n - 1)
cat("Critical region: (", critical, ", Inf)\n")

# Decision using the critical region
if (statistic > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

Alternative calculation using DescTools

```
test_result = DescTools::VarTest(x, sigma.squared = variance0,
    alternative = "greater", conf.level = 1 - alpha)
test_result
p_value = test_result$p.value

# Decision using the p-value
if (p_value > alpha) {
  cat("p > alpha: do not reject H0.\n")
} else {
  cat("p <= alpha: reject H0 in favour of H1.\n")
}
```

Exercise 5.5 Exponential Mean

Problem statement (textbook, p. 72)

Bus waiting times were recorded for 100 randomly selected passengers:

Waiting time [min]	0.5	2.5	4.5	6.5	8.5
Number of passengers	60	20	11	5	4

Assume an exponential distribution with mean parameter λ. The carrier claims that the mean waiting time is 2 minutes. Test whether the population mean waiting time:
a) differs from 2 minutes, at α = 0.05;
b) is greater than 2 minutes, at α = 0.01;
c) is less than 2 minutes, at α = 0.01.

R solution: one-sample test of the exponential mean

```
values = c(0.5,2.5,4.5,6.5,8.5)
frequencies = c(60,20,11,5,4)
lambda0 = 2
n = sum(frequencies)
# Repeat each value according to its frequency.
x = rep(values, frequencies)
sample_mean = mean(x)
statistic = 2*n*sample_mean/lambda0
```

a) H0: lambda = 2; H1: lambda != 2

```
alpha = 0.05
critical_lower = qchisq(alpha/2, 2*n)
critical_upper = qchisq(1 - alpha/2, 2*n)
cat("Critical region: (0,", critical_lower, ") union (",
    critical_upper, ", Inf)\n")

# Decision using the critical region
if (statistic < critical_lower || statistic > critical_upper) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

b) H0: lambda = 2; H1: lambda > 2

```
alpha1 = 0.01
critical = qchisq(1 - alpha1, 2*n)
cat("Critical region: (", critical, ", Inf)\n")

# Decision using the critical region
if (statistic > critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```

c) H0: lambda = 2; H1: lambda < 2

```
critical = qchisq(alpha1, 2*n)
cat("Critical region: (0,", critical, ")\n")

# Decision using the critical region
if (statistic < critical) {
  cat("Statistic in critical region: reject H0 in favour of H1.\n")
} else {
  cat("Statistic outside critical region: do not reject H0.\n")
}
```