Course Code 009959 (Spring 2018)

Probability and Statistics with Programming

Hypotheses Testing in R

Prof. S. M. Riazul Islam, Dept. of Computer Engineering, Sejong University, Korea

E-mail: riaz@sejong.ac.kr

Hypotheses Testing in R

- Problem Solving on Hypotheses Testing using R

Hypotheses Testing in R

Problem

Suppose the manufacturer claims that the mean lifetime of a light bulb is more than 10,000 hours. In a sample of 30 light bulbs, it was found that they only last 9,900 hours on average. Assume the population standard deviation is 120 hours. At .05 significance level, can we reject the claim by the manufacturer?

Solution

The null hypothesis is that $\mu \geq 10000$. We begin with computing the test statistic.

```
> xbar = 9900            # sample mean
> mu0 = 10000            # hypothesized value
> sigma = 120            # population standard deviation
> n = 30                 # sample size
> z = (xbar-mu0)/(sigma/sqrt(n))
> z                      # test statistic
[1] -4.5644
```

We then compute the critical value at .05 significance level.

```
> alpha = .05
> z.alpha = qnorm(1-alpha)
> -z.alpha               # critical value
[1] -1.6449
```

// qnorm(0.05)

REJECT

Alternative Approach

```
> pval = pnorm(z)
> pval                   # lower tail p-value
[1] 2.5052e-06
```

Hypotheses Testing in R

Problem

Suppose the mean weight of King Penguins found in an Antarctic colony last year was 15.4 kg. In a sample of 35 penguins same time this year in the same colony, the mean penguin weight is 14.6 kg. Assume the population standard deviation is 2.5 kg. At .05 significance level, can we reject the null hypothesis that the mean penguin weight does not differ from last year?

Solution

The null hypothesis is that $\mu = 15.4$. We begin with computing the test statistic.

```
> xbar = 14.6             # sample mean
> mu0 = 15.4              # hypothesized value
> sigma = 2.5             # population standard deviation
> n = 35                  # sample size
> z = (xbar-mu0)/(sigma/sqrt(n))
> z                       # test statistic
[1] -1.8931
```

We then compute the critical values at .05 significance level.

```
> alpha = .05
> z.half.alpha = qnorm(1-alpha/2)
> c(-z.half.alpha, z.half.alpha)
[1] -1.9600  1.9600
```

// qnorm(0.05/2)

Do not REJECT

Alternative

```
> pval = 2 * pnorm(z)     # lower tail
> pval                    # two-tailed p-value
[1] 0.058339
```

Hypotheses Testing in R

Problem

Suppose the food label on a cookie bag states that there is at most 2 grams of saturated fat in a single cookie. In a sample of 35 cookies, it is found that the mean amount of saturated fat per cookie is 2.1 grams. Assume that the sample standard deviation is 0.3 gram. At .05 significance level, can we reject the claim on food label?

Solution

The null hypothesis is that $\mu \le 2$. We begin with computing the test statistic.

```
> xbar = 2.1                # sample mean
> mu0 = 2                   # hypothesized value
> s = 0.3                   # sample standard deviation
> n = 35                    # sample size
> t = (xbar−mu0)/(s/sqrt(n))
> t                         # test statistic
[1] 1.9720
```

We then compute the critical value at .05 significance level.

```
> alpha = .05
> t.alpha = qt(1−alpha, df=n−1)
> t.alpha                   # critical value
[1] 1.6991
```

// qt(0.05, lower.tail=FALSE)

Alternative

```
> pval = pt(t, df=n−1, lower.tail=FALSE)
> pval                      # upper tail p−value
[1] 0.028393
```

REJECT

Hypotheses Testing in R

Problem

Let μ denote the mean reaction time to a certain stimulus. For a large-sample z test of H_0: $\mu = 5$ versus H_a: $\mu > 5$, find the P-value associated with each of the given values of the z test statistic.

a. 1.42 **b.** .90 **c.** 1.96 **d.** 2.48 **e.** $-.11$

a0.0778, b0.1841, c0.0250, d0.0066, e0.5438

Hypotheses Testing in R

Problem

Give as much information as you can about the *P*-value of a *t* test in each of the following situations:

a. Upper-tailed test, df $= 8$, $t = 2.0$
b. Lower-tailed test, df $= 11$, $t = -2.4$
c. Two-tailed test, df $= 15$, $t = -1.6$
d. Upper-tailed test, df $= 19$, $t = -.4$
e. Upper-tailed test, df $= 5$, $t = 5.0$
f. Two-tailed test, df $= 40$, $t = -4.8$

a. .040 **b.** .018 **c.** .130 **d.** .653
e. $< .005$ **f.** $\approx .000$

Q&A