PES University, Bangalore

(Established under Karnataka Act No. 16 of 2013)

UE19CS203 – STATISTICS FOR DATA SCIENCE

Unit-5 - Power of Test and Simple Linear Regression

QUESTION BANK-SOLVED

Power of a Test

Exercises for section 6.13: [Text Book Exercise 6.13– Pg. No. [485 – 487]]

1. A copper smelting process is supposed to reduce the arsenic content of the copper to less than 1000 ppm. Let μ denote the mean arsenic content for copper treated by this process, and assume that the standard deviation of arsenic content is σ = 100 ppm. The sample mean arsenic content X of 75 copper specimens will be computed, and the null hypothesis $H_0 : \geq 1000$ will be tested against the alternate $H_1 : \mu < 1000$.

   a. A decision is made to reject H_0 if $X \leq 980$. Find the level of this test.
   b. Find the power of the test in part (a) if the true mean content is 965 ppm.
   c. For what values of X should H_0 be rejected so that the power of the test will be 0.95 when the true mean content is 965?
   d. For what values of X should H_0be rejected so that the level of the test will be 5%?
   e. What is the power of a 5% level test if the true mean content is 965 ppm?
   f. How large a sample is needed so that a 5% level test has power 0.95 when the true mean content is 965 ppm?

[Text Book Exercise – Section 6.13 – Q. No.6 – Pg. No. 486]

Solution:

Given n= sample size = 75

True mean = 965

σ = 100

α = 5% = 0.05

$H_{0:}: \ \mu \geq 1000$

$H_1: \ \mu < 1000$

(a) $\bar{X} \leq 980$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{980 - 1000}{100/\sqrt{75}} \approx -1.73$$

The level of the test is the probability of rejecting the null hypothesis.
Level = P(Z < - 1.73) = 0.0418

(b) The power is the probability of rejecting the null hypothesis when the alternative hypothesis is true

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{980 - 965}{100/\sqrt{75}} \approx 1.30$$

The level of the test is the probability of not rejecting the null hypothesis.

Power= P(Z < 1.30)=0.9032

(c) Power = 0.95

Determine the z-score corresponding with a probability of 0.95 using the normal probability table

z = 1.645

The sampling distribution of the sample mean has mean μ and standard deviation $\sigma/\sqrt{n}$

The corresponding sample mean is

$$\bar{x} = \mu + z\frac{\sigma}{\sqrt{n}} = 965 + 1.645\frac{100}{\sqrt{75}} \approx 983.9948$$

Thus, H_0 should be reject when $\bar{X} \ \leq 983.9948$

(d) α = 5% = 0.05

Determine the z-score corresponding with a probability of α = 5% = 0.05 using the normal probability table

z = - 1.645

The sampling distribution of the sample mean has mean μ and standard deviation $\sigma/\sqrt{n}$

The corresponding sample mean is

$$\bar{x} = \mu + z\frac{\sigma}{\sqrt{n}} = 1000 - 1.645\frac{100}{\sqrt{75}} \approx 981.0052$$

Thus, H_0 should be reject when $\bar{X} \ \leq 981.0052$

(e) α = 5% = 0.05

Determine the z-score corresponding with a probability of α = 5% = 0.05 using the normal probability table

z = - 1.645

The sampling distribution of the sample mean has mean μ and standard deviation $\sigma/\sqrt{n}$

The corresponding sample mean is

$$\bar{x} = \mu + z\frac{\sigma}{\sqrt{n}} = 1000 - 1.645\frac{100}{\sqrt{75}} \approx 981.0052$$

The z-value is

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{981.0052 - 965}{100/\sqrt{75}} \approx 1.39$$

Determine the probability of rejecting the null hypothesis using the normal probability table

$$\text{Power} = P(z < 1.39) = 0.9177$$

(f) Power=0.95

α = 5% = 0.05

Determine the z-score corresponding with a probability of α = 5% = 0.05 using the normal probability table

$$z = -1.645$$

The sampling distribution of the sample mean has mean μ and standard deviation $\sigma/\sqrt{n}$

The corresponding sample mean is

$$\bar{x} = \mu + z\frac{\sigma}{\sqrt{n}} = 1000 - 1.645\frac{100}{\sqrt{n}}$$

Determine the z-score corresponding with a probability of 0.95 using the normal probability table

$$z = 1.645$$

The sampling distribution of the sample mean has mean μ and standard deviation $\sigma/\sqrt{n}$

The corresponding sample mean is

$$\bar{x} = \mu + z\frac{\sigma}{\sqrt{n}} = 965 + 1.645\frac{100}{\sqrt{n}}$$

The two found samples means need to be equal:

$$1000 - 1.645\frac{100}{\sqrt{n}} = 965 + 1.645\frac{100}{\sqrt{n}}$$

$$35 - 1.645\frac{100}{\sqrt{n}} = 1.645\frac{100}{\sqrt{n}}$$

Multiply each side of the equation by $\sqrt{n}$:

$$35\sqrt{n} - 1.645(100) = 1.645(100)$$

$$35\sqrt{n} = 1.645(100) + 1.645(100)$$

$$\sqrt{n} = \frac{329}{35} = 9.4$$

Therefore, $n = 9.4^2 = 88.26 \approx 89$

2. A new process for producing silicon wafers for integrated circuits is supposed to reduce the proportion of defectives to 10%. A sample of 250 wafers will be tested. Let X represent the number of defectives in the sample. Let p represent the population proportion of defectives produced by the new process. A test will be made of $H_0 : p \geq 0.10$ versus $H_1 : p < 0.10$. Assume the true value of p is actually 0.06.

   a. It is decided to reject H_0 if $X \leq 18$. Find the level of this test.
   b. It is decided to reject H_0 if $X \leq 18$. Find the power of this test.
   c. Should you use the same standard deviation for X to compute both the power and the level? Explain.
   d. How many wafers should be sampled so that the power is 0.90 if the test is made at the 5% level?

[Text Book Exercise – Section 6.13 – Q. No.8 – Pg. No. 486]

Solution:

Given $n = 250, \quad True\ proportion = 0.06, \ \alpha = 0.05$

$$H_0 : p \geq 0.10$$

$$H_1 : p < 0.10$$

(a) The sample proportion is the number of successes divided by the sample size:

$$\hat{p} = \frac{x}{n} = \frac{18}{250} = 0.072$$

Determine the value of the test-statistic (using the proportion from the null hypothesis):

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.072 - 0.10}{\sqrt{\frac{0.10(1-0.10)}{250}}} \approx -1.48$$

The level of the test is the probability of rejecting the null hypothesis.
Level= P(Z < -1.48) = 0.0694

(b) The sample proportion is the number of successes divided by the sample size:

$$\hat{p} = \frac{x}{n} = \frac{18}{250} = 0.072$$

Determine the value of the test-statistic (using the proportion from the null hypothesis):

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.072 - 0.06}{\sqrt{\frac{0.06(1-0.06)}{250}}} \approx 0.80$$

The power of the test is the probability of rejecting the null hypothesis.
Power= P(Z < 0.80)=0.7881

(c) The standard deviation of the sampling distribution of the sample proportion is given by the formula $\sqrt{\frac{p_0(1-p_0)}{n}}$.

When we determine the level of the test, then we need to use the proportion of the null hypothesis as an estimate for the proportion p_0 and thus we need to use the proportion of the null hypothesis in the calculation of the standard deviation as well.

When we determine the power of the test, then we need to use the true proportion as an estimate for the proportion p_0 and thus we need to use the true proportion in the calculation of the standard deviation as well.

Since we use difference proportions in the calculation of the standard deviation when determining the level and the power of the test, we **do not use** the same standard deviation for x to compute both the power and the level.

(d) Power=0.90

α= Significance level= 5%=0.05

Determine the z- score corresponding with a probability of α= 5%=0.05 using the normal probability

z = - 1.645

The sampling distribution of the sample proportion has mean p_0 and the standard deviation $\sqrt{\frac{p_0(1-p_0)}{n}}$.

The corresponding sample proportion is

$$\hat{p} = p_0 + z\sqrt{\frac{p_0(1-p_0)}{n}} = 0.10 - 1.645\sqrt{\frac{0.10(1-0.10)}{n}}$$

Determine the z-score corresponding with a probability of 0.90 using the normal probability table

Z=1.28

The sampling distribution of the sample proportion has mean p_0 and the standard deviation $\sqrt{\frac{p_0(1-p_0)}{n}}$.

The corresponding sample proportion is

$$\hat{p} = p_0 + z\sqrt{\frac{p_0(1-p_0)}{n}} = 0.06 + 1.28\sqrt{\frac{0.06(1-0.06)}{n}}$$

The two found samples proportions needs to be equal

$$0.10 - 1.645\sqrt{\frac{0.10(1-0.10)}{n}} = 0.06 + 1.28\sqrt{\frac{0.06(1-0.06)}{n}}$$

$$0.04 - 1.645\sqrt{\frac{0.10(1-0.10)}{n}} = 1.28\sqrt{\frac{0.06(1-0.06)}{n}}$$

Multiply each side of the equation by $\sqrt{n}$,

$$0.04\sqrt{n} - 1.645\sqrt{0.10(1-0.10)} = 1.28\sqrt{0.06(1-0.06)}$$

$$0.04\sqrt{n} = 1.645\sqrt{0.10(1-0.10)} + 1.28\sqrt{0.06(1-0.06)}$$

$$\sqrt{n} = \frac{1.645\sqrt{0.10(1-0.10)}+1.28\sqrt{0.06(1-0.06)}}{0.04}$$

$$Therefore, n \approx 398$$