Linear Least Squares Regression

Here there are only five pairs of numbers so we can enter them in manually. Each of the five pairs consists of a year and the mean interest rate:

```
> year <- c(2000 ,  2001  , 2002  , 2003 ,  2004)
> rate <- c(9.34 ,  8.50  , 7.62  , 6.93  , 6.60)
```

We find the correlation between the year and the mean interest rates:

```
> cor(year,rate)
[1] -0.9880813
```

The command to perform the least square regression is the *lm* command. Since we specified that the interest rate is the response variable and the year is the explanatory variable this means that the regression line can be written in slope-intercept form:

rate=(slope)year+(intercept)

rate=(slope)year+(intercept)


```
> fit <- lm(rate ~ year)

> fit

Call:

lm(formula = rate ~ year)

Coefficients:

(Intercept)       year

  1419.208      -0.705
```

Lower Tail Test of Population Mean with Known Variance

The null hypothesis of the **lower tail test of the population mean** can be expressed as follows:

$$\mu \geq \mu_0$$

where μ_0 is a hypothesized lower bound of the true population mean μ.

Let us define the test statistic z in terms of the sample mean, the sample size and the population standard deviation σ:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Then the null hypothesis of the lower tail test is to be *rejected* if $z \leq -z_\alpha$, where z_α is the $100(1 - \alpha)$ percentile of the standard normal distribution.

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

$\le -z_\alpha$, where z_α is the $100(1 - \alpha)$ percentile of the standard normal distribution.

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 $\mu \ge 10000$. 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.

```
[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
```

Answer

The test statistic -4.5644 is less than the critical value of -1.6449. Hence, at .05 significance level, we reject the claim that mean lifetime of a light bulb is above 10,000 hours.

Alternative Solution

Instead of using the critical value, we apply the pnorm function to compute the lower tail **p-value** of the test statistic. As it turns out to be less than the .05 significance level, we reject the null hypothesis that $\mu \geq 10000$.

```
> pval = pnorm(z)
> pval                      # lower tail p-value
```

Upper Tail Test of Population Mean with Known Variance

The null hypothesis of the upper tail test of the population mean

Upper Tail Test of Population Mean with Known Variance

The null hypothesis of the **upper tail test of the population mean** can be expressed as follows:

$$\mu \leq \mu_0$$

where μ_0 is a hypothesized upper bound of the true population mean μ.

Let us define the test statistic z in terms of the sample mean, the sample size and the population standard deviation σ :

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Then the null hypothesis of the upper tail test is to be *rejected* if $z \geq z_\alpha$, where z_α is the $100(1 - \alpha)$ percentile of the standard normal distribution.

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 population standard

$\geq z_\alpha$, where z_α is the 100(1 − α) percentile of the standard normal distribution.

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 population standard deviation is 0.25 grams. At .05 significance level, can we reject the claim on food label?

Solution

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

```
> xbar = 2.1                 # sample mean
> mu0 = 2                    # hypothesized value
> sigma = 0.25               # population standard deviation
> n = 35                     # sample size
> z = (xbar-mu0)/(sigma/sqrt(n))
> z                          # test statistic
[illegible]
```

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

-p × +

TjVVNDFZSdbXtNtRFvgIWRfZxFRQ?projector=1&messagePartId=0.1

Answer

The test statistic 2.3664 is greater than the critical value of 1.6449. Hence, at .05 significance level, we reject the claim that there is at most 2 grams of saturated fat in a cookie.

Two-Tailed Test of Population Mean with Known Variance

The null hypothesis of the **two-tailed test of the population mean** can be expressed as follows:

$$\mu = \mu_0$$

where μ_0 is a hypothesized value of the true population mean μ.

Let us define the test statistic z in terms of the sample mean, the sample size and the population standard deviation σ :

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Then the null hypothesis of the two-tailed test is to be *rejected* if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$, where $z_{\alpha/2}$ is the 100$(1 - \alpha/2)$ percentile of the standard normal distribution.

Page 4 / 7

Problem

the standard normal distribution.

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.

jVVNDFZSdbXtNtRFvgIWRfZxFRQ?projector=1&messagePartId=0.1

```
[1] -1.9600  1.9600
```

Answer

The test statistic -1.8931 lies between the critical values -1.9600 and 1.9600. Hence, at .05 significance level, we do *not* reject the null hypothesis that the mean penguin weight does not differ from last year.

Lower Tail Test of Population Mean with Unknown Variance

The null hypothesis of the lower tail test of the population mean can be expressed as follows:

$$\mu \geq \mu_0$$

where μ_0 is a hypothesized lower bound of the true population mean μ.

Let us define the test statistic t in terms of the sample mean, the sample size and the sample standard deviation s:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

sample size and the sample standard deviation s :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Then the null hypothesis of the lower tail test is to be rejected if $t \le -t_\alpha$, where t_α is the $100(1 - \alpha)$ percentile of the Student t distribution with $n - 1$ degrees of freedom.

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 sample standard deviation is 125 hours. At .05 significance level, can we reject the claim by the manufacturer?

Solution

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

```
> xbar = 9900            # sample mean
> mu0 = 10000            # hypothesized value
> s = 125                # sample standard deviation
> n = 30                 # sample size
> t = (xbar-mu0)/(s/sqrt(n))
> t                      # test statistic
[1] -4.3818
```

Patil - p × +

cJHsTjVVNDFZSdbXtNtRFvgIWRfZxFRQ?projector=1&messagePartId=0.1

test statistic.

```
> xbar = 9900            # sample mean
> mu0 = 10000            # hypothesized value
> s = 125                # sample standard deviation
> n = 30                 # sample size
> t = (xbar−mu0)/(s/sqrt(n))
> t                      # test statistic
[1] −4.3818
```

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
```

Answer

The test statistic -4.3818 is less than the critical value of -1.6991.

Hence, at .05 significance level, we can reject the claim that mean lifetime of a light bulb is above 10,000 hours.

Upper Tail Test of Population Mean with Unknown Variance

Sea

The null hypothesis of the **upper tail test of the population mean**

Upper Tail Test of Population Mean with Unknown Variance

The null hypothesis of the **upper tail test of the population mean** can be expressed as follows:

$$\mu \le \mu_0$$

where μ_0 is a hypothesized upper bound of the true population mean μ.

Let us define the test statistic t in terms of the sample mean, the sample size and the sample standard deviation s :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Then the null hypothesis of the upper tail test is to be *rejected* if $t \ge t_\alpha$, where t_α is the $100(1 - \alpha)$ percentile of the Student t distribution with $n - 1$ degrees of freedom.

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 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 $\mu \leq 2$. 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
```

Answer

The test statistic 1.9720 is greater than the critical value of 1.6991. Hence, at .05 significance level, we can reject the claim that there is at most 2 grams of saturated fat in a cookie.

Two-Tailed Test of Population Mean with

Page 6 / 7

Two-Tailed Test of Population Mean with Unknown Variance

The null hypothesis of the **two-tailed test of the population mean** can be expressed as follows:

$$\mu = \mu_0$$

where μ_0 is a hypothesized value of the true population mean μ.

Let us define the test statistic t in terms of the sample mean, the sample size and the sample standard deviation s :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Then the null hypothesis of the two-tailed test is to be *rejected* if $t \leq -t_{\alpha/2}$ or $t \geq t_{\alpha/2}$, where $t_{\alpha/2}$ is the $100(1 - \alpha)$ percentile of the Student t distribution with $n - 1$ degrees of freedom.

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 sample standard deviation is 2.5 kg. At .05

14.6 kg. Assume the sample 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
> s = 2.5                      # sample standard deviation
> n = 35                       # sample size
> t = (xbar-mu0)/(s/sqrt(n))
> t                            # test statistic
[1] -1.8931
```

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

```
> alpha = .05
> t.half.alpha = qt(1-alpha/2, df=n-1)
> c(-t.half.alpha, t.half.alpha)
[1] -2.0322  2.0322
```

Answer

The test statistic -1.8931 lies between the critical values -2.0322, and 2.0322. Hence, at .05 significance level, we do *not* reject the null hypothesis that the mean penguin weight does not differ from last year.