**Chapter 2**

We assume that the normal error regression model is applicable. This model is:

$$Y\_{i}=β\_{0}+β\_{1}X\_{i}+ε\_{i}$$

where:

$β\_{0}$ and $β\_{1}$, are parameters

$X\_{i}$ are known constants

$ε\_{i}$ are independent $N (0, σ^{2})$

$$E\left(Y\_{i}\right)=β\_{0}+β\_{1}X\_{i}$$

**Sampling Distribution of** $\hat{β\_{1}}$

$$\hat{β\_{1}}=b\_{1}=\sum\_{i=1}^{n}\frac{\left(X\_{i}-\overbar{X}\right)\left(Y\_{i}-\overbar{Y}\right)}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}$$

$$E\left(\hat{β\_{1}}\right)=β\_{1}$$

$$σ^{2}\left(\hat{β\_{1}}\right)=\frac{σ^{2}}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}$$

$$s^{2}\left(\hat{β\_{1}}\right)=\frac{MSE}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}$$

$$\frac{b\_{1}-β\_{1}}{s\left(b\_{1}\right)}\~t\_{\left(n-2\right)}$$

**Confidence Interval for** $β\_{1}$

$$P\left[b\_{1}-t\_{\left(1-\frac{α}{2},n-2\right)}s\left(b\_{1}\right)\leq β\_{1}\leq b\_{1}+t\_{\left(1-α/2,n-2\right)}s\left(b\_{1}\right)\right]=1-α$$

C.I $\left(1-α\right)\%$ for $β\_{1}$

$$b\_{1}-t\_{\left(1-\frac{α}{2},n-2\right)}s\left(b\_{1}\right)\leq β\_{1}\leq b\_{1}+t\_{\left(1-α/2,n-2\right)}s\left(b\_{1}\right)$$

**Tests Concerning** $β\_{1}$

|  |
| --- |
| 1. Hypothesis  |
| $$H\_{0}:β\_{1}=β\_{10} $$$$H\_{1}:β\_{1}\ne β\_{10}$$ | $$H\_{0}:β\_{1}=β\_{10} $$$$H\_{1}:β\_{1}>β\_{10}$$ | $$H\_{0}:β\_{1}=β\_{10} $$$$H\_{1}:β\_{1}<β\_{10}$$ |
| 2. Test statistic |
| $$T\_{0}=\frac{b\_{1}-β\_{10}}{s\left(b\_{1}\right)}$$ |
| 3. Decision: Reject $H\_{0}$ if |
| $$\left|T\_{0}\right|>t\_{\left(1-\frac{α}{2},n-2\right)}$$ | $$T\_{0}>t\_{\left(1-α,n-2\right)}$$ | $$T\_{0}<t\_{\left(α,n-2\right)}$$ |
| P-value: Reject $H\_{0}$ if $p-value<α$ |
| p-value=$2P\left(t\_{\left(n-2\right)}>\left|T\_{0}\right|\right)$ | p-value$=P\left(t\_{\left(n-2\right)}>T\_{0}\right)$ | $$p-value=P\left(t\_{\left(n-2\right)}<T\_{0}\right)$$ |

**Sampling Distribution of** $\hat{β\_{0}}$

$$\hat{β\_{0}}=b\_{0}=\overbar{Y}-b\_{1}\overbar{X}$$

$$E\left(\hat{β\_{0}}\right)=β\_{0}$$

$$σ^{2}\left(\hat{β\_{0}}\right)=σ^{2}\left(\frac{1}{n}+\frac{\overbar{X}^{2}}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}\right)$$

$$s^{2}\left(\hat{β\_{0}}\right)=MSE\left(\frac{1}{n}+\frac{\overbar{X}^{2}}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}\right)$$

$$\frac{b\_{0}-β\_{0}}{s\left(b\_{0}\right)}\~t\_{\left(n-2\right)}$$

**Confidence Interval for** $β\_{1}$

$$P\left[b\_{0}-t\_{\left(1-\frac{α}{2},n-2\right)}s\left(b\_{0}\right)\leq β\_{0}\leq b\_{0}+t\_{\left(1-α/2,n-2\right)}s\left(b\_{0}\right)\right]=1-α$$

C.I $\left(1-α\right)\%$ for $β\_{0}$

$$b\_{0}-t\_{\left(1-\frac{α}{2},n-2\right)}s\left(b\_{0}\right)\leq β\_{0}\leq b\_{0}+t\_{\left(1-α/2,n-2\right)}s\left(b\_{0}\right)$$

**Tests Concerning** $β\_{1}$

|  |
| --- |
| 1. Hypothesis  |
| $$H\_{0}:β\_{0}=β\_{00} $$$$H\_{1}:β\_{0}\ne β\_{00}$$ | $$H\_{0}:β\_{0}=β\_{00} $$$$H\_{1}:β\_{0}>β\_{00}$$ | $$H\_{0}:β\_{1}=β\_{00} $$$$H\_{1}:β\_{1}<β\_{00}$$ |
| 2. Test statistic |
| $$T\_{0}=\frac{b\_{0}-β\_{00}}{s\left(b\_{0}\right)}$$ |
| 3. Decision: Reject $H\_{0}$ if |
| $$\left|T\_{0}\right|>t\_{\left(1-\frac{α}{2},n-2\right)}$$ | $$T\_{0}>t\_{\left(1-α,n-2\right)}$$ | $$T\_{0}<t\_{\left(α,n-2\right)}$$ |
| P-value: Reject $H\_{0}$ if $p-value<α$ |
| p-value=$2P\left(t\_{\left(n-2\right)}>\left|T\_{0}\right|\right)$ | p-value$=P\left(t\_{\left(n-2\right)}>T\_{0}\right)$ | $$p-value=P\left(t\_{\left(n-2\right)}<T\_{0}\right)$$ |

$$Y\_{h}=b\_{0}+b\_{1}X\_{h}$$

ANOVA TABLE

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | d.f | SS | MS | F | p-value |
| Regression | 1 | SSR=$\sum\_{}^{}\left(\hat{Y\_{i}}-\overbar{Y}\right)^{2}$ | $$MSR=\frac{SSR}{1}$$ | $$\frac{MSR}{MSE}$$ |  |
| Error | n-2 | SSE=$\sum\_{}^{}\left(Y\_{i}-\hat{Y\_{i}}\right)^{2}$ | $$MSE=\frac{SSE}{n-2}$$ |  |  |
| Total | n-1 | SSTo= $\sum\_{}^{}\left(Y\_{i}-\overbar{Y}\right)^{2}$ |  |  |  |

1. Hypothesis

$H\_{0}:β\_{1}=0 $(Non liner)

$$H\_{1}:β\_{1}\ne 0$$

1. Test statistic

$$F^{\*}=\frac{MSR}{MSE}$$

1. Decision: Reject $H\_{0}$ if

$$F>F\_{\left(1-α, 1,n-2\right)}$$

P-value: Reject $H\_{0}$ if $p-value<α$

$$p-value=P\left(F\_{\left(1,n-2\right)}>F^{\*}\right)$$

Q2.6**. Refer to Airfreight breakage Problem 1.21.**

$\overbar{X}=1, \overbar{Y}=14.2$,

$$\sum\_{i=1}^{n=10}\left(X\_{i}-\overbar{X}\right)\left(Y\_{i}-\overbar{Y}\right)=40, \sum\_{i=1}^{10}\left(X\_{i}-\overbar{X}\right)^{2}=10$$

$$\sum\_{i=1}^{n=10}\left(Y\_{i}-\overbar{Y}\right)^{2}=177.6, MSE=2.2$$

$$b\_{0}=10.2, b\_{1}=4$$

d) A consultant has suggested, on the basis of previous experience, that the mean number of broken ampules should not exceed 9.0 when no transfers are made. Conduct an appropriate test, using $α=0.025$. State the alternatives, decision rule, and conclusion. What is the P-value of the test?

$$α=0.025$$

1. Hypothesis

$$H\_{0}:β\_{0}\leq 9 $$

$$H\_{1}:β\_{0}>9$$

2. Test statistic

$$T\_{0}=\frac{b\_{0}-β\_{00}}{s\left(b\_{0}\right)}=\frac{10.2-9}{0.6633}=1.809$$

$$s^{2}\left(\hat{β\_{0}}\right)=MSE\left(\frac{1}{n}+\frac{\overbar{X}^{2}}{\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}\right)=2.2\left(\frac{1}{10}+\frac{1^{2}}{10}\right)=0.44$$

$$s\left(b\_{0}\right)=0.6633$$

3. Decision: Reject $H\_{0}$ if $ T\_{0}>t\_{\left(1-α,n-2\right)}$,

 $ 1.809≯t\_{\left(0.975,8\right)}=2.306$

Then not reject $H\_{0}$

p-value=$P\left(t\_{\left(n-2\right)}>T\_{0}\right)=\left(1-P\left(t\_{\left(n-2\right)}<1.809\right)\right)=\left(1-0.945\right)=0.055≮0.025$

, then we not reject $H\_{0}$.

at $α=0.05$

$$b\_{0}-t\_{\left(1-\frac{α}{2},n-2\right)}s\left(b\_{0}\right)\leq β\_{0}\leq b\_{0}+t\_{\left(1-α/2,n-2\right)}s\left(b\_{0}\right)$$

$$t\_{\left(1-\frac{α}{2},n-2\right)}=t\_{\left(0.975,8\right)}=2.306$$

$$10.2-2.306\*0.6633\leq β\_{0}\leq 10.2+2.306\*0.6633$$

$$8.76\leq β\_{0}\leq 11.728$$

**Analysis of Variance**

**Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value**

**Regression 1 160.000 90.09% 160.000 160.000 72.73 0.000**

 **Xi 1 160.000 90.09% 160.000 160.000**

**Error 8 17.600 9.91% 17.600 2.200**

**Total 9 177.600 100.00%**

**Coefficients**

**Term Coef SE Coef 95% CI T-Value P-Value VIF**

**Constant 10.200 0.663 (8.670, 11.730) 15.38 0.000**

**Xi 4.000 0.469 (2.918, 5.082) 8.53 0.000 1.00**

**Regression Equation**

**Yi = 10.200 + 4.000 Xi**

Q2.25. Refer to Airfreight breakage Problem 1.21.

a. Set up the ANOVA table. Which elements are additive?

b. Conduct an *F* test to decide whether or not there is a linear association between the number of times a carton is transferred and the number of broken ampules; control the $α$risk at 0.05. State the alternatives, decision rule, and conclusion.

c. Obtain the *t\** statistic for the test in part (b) and demonstrate numerically its equivalence to the *F\** statistic obtained in part (b).

$\overbar{X}=1, \overbar{Y}=14.2$, $\sum\_{i=1}^{n=10}\left(X\_{i}-\overbar{X}\right)\left(Y\_{i}-\overbar{Y}\right)=40, \sum\_{i=1}^{10}\left(X\_{i}-\overbar{X}\right)^{2}=10$

$$\sum\_{i=1}^{n=10}\left(Y\_{i}-\overbar{Y}\right)^{2}=177.6, MSE=2.2, b\_{0}=10.2, b\_{1}=4$$

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| $$X\_{i}$$ | $$Y\_{i}$$ | $$\left(X\_{i}-\overbar{X}\right)$$ | $$\left(Y\_{i}-\overbar{Y}\right)$$ | $$\left(X\_{i}-\overbar{X}\right)^{2}$$ | $$\left(X\_{i}-\overbar{X}\right)\*\left(Y\_{i}-\overbar{Y}\right)$$ | $$\left(Y\_{i}-\overbar{Y}\right)^{2}$$ | $$\hat{Y\_{i}}$$ | $$\left(Y\_{i}-\hat{Y\_{i}}\right)^{2}$$ |
| 1 | 16 | 0 | 1.8 | 0 | 0 | 3.24 | 14.2 | 3.24 |
| 0 | 9 | -1 | -5.2 | 5.2 | 1 | 27.04 | 10.2 | 1.44 |
| 2 | 17 | 1 | 2.8 | 2.8 | 1 | 7.84 | 18.2 | 1.44 |
| 0 | 12 | -1 | -2.2 | 2.2 | 1 | 4.84 | 10.2 | 3.24 |
| 3 | 22 | 2 | 7.8 | 15.6 | 4 | 60.84 | 22.2 | 0.04 |
| 1 | 13 | 0 | -1.2 | 0 | 0 | 1.44 | 14.2 | 1.44 |
| 0 | 8 | -1 | -6.2 | 6.2 | 1 | 38.44 | 10.2 | 4.84 |
| 1 | 15 | 0 | 0.8 | 0 | 0 | 0.64 | 14.2 | 0.64 |
| 2 | 19 | 1 | 4.8 | 4.8 | 1 | 23.04 | 18.2 | 0.64 |
| 0 | 11 | -1 | -3.2 | 3.2 | 1 | 10.24 | 10.2 | 0.64 |
| 10 | 142 | 0 | 0 | 40 | 10 | 177.6 | 142 | 17.6 |

$$\sum\_{}^{}\left(Y\_{i}-\hat{Y\_{i}}\right)^{2}=17.6$$

ANOVA TABLE

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | d.f | SS | MS | F | p-value |
| Regression | 1 | SSR=$177.6-17.6=160$ | $$MSR=160$$ | $$\frac{160}{2.2}=72.72$$ | 0.00 |
| Error | 8 | SSE=$17.6$ | $$MSE=\frac{17.6}{8}=2.2$$ |  |  |
| Total | 9 | SSTo= 177.6 |  |  |  |

$$α=0.05$$

1. Hypothesis

$$H\_{0}:β\_{1}=0 $$

$$H\_{1}:β\_{1}\ne 0$$

2. Test statistic

$$F^{\*}=72.72$$

3. Decision: Reject $H\_{0}$ if $ F^{\*}>F\_{\left(1-α,1,n-2\right)}$, $ 72.72>F\_{\left(0.95,1,8\right)}=5.31$

Then reject $H\_{0}$

p-value=$P\left(F\_{\left(1,n-2\right)}>F^{\*}\right)=\left(1-P\left(F\_{\left(1,8\right)}<72.72\right)\right)=\left(1-0.9999\right)=0.0001<0.05$

, then we reject $H\_{0}$.

**Analysis of Variance**

**Source DF Adj SS Adj MS F-Value P-Value**

**Regression 1 160.000 160.000 72.73 0.000**

 **Xi 1 160.000 160.000 72.73 0.000**

**Error 8 17.600 2.200**

 **Lack-of-Fit 2 0.933 0.467 0.17 0.849**

 **Pure Error 6 16.667 2.778**

**Total 9 177.600**

$t^{\*}=8.528$, $\left(t^{\*}\right)^{2}=\left(8.528\right)^{2}=72.72=F^{\*}$

Q2.26. Refer to Plastic hardness Problem 1.22.

a. Set up the ANOVA table.

b. Test by means of an *F* test whether or not there is a linear association between the hardness of the plastic and the elapsed time. Use *a* = .01. State the alternatives, decision rule, and conclusion.

**Analysis of Variance**

**Source DF Adj SS Adj MS F-Value P-Value**

**Regression 1 5297.51 5297.51 506.51 0.000**

 **Xi 1 5297.51 5297.51 506.51 0.000**

**Error 14 146.43 10.46**

 **Lack-of-Fit 2 17.67 8.84 0.82 0.462**

 **Pure Error 12 128.75 10.73**

**Total 15 5443.94**

$$α=0.01$$

1. Hypothesis

$$H\_{0}:β\_{1}=0 $$

$$H\_{1}:β\_{1}\ne 0$$

2. Test statistic

$$F^{\*}=506.51$$

3. Decision:

p-value=0.000<0.01

, then we reject $H\_{0}$.