1. A sample of customers have difficulty in operating printer's owner manual. As the given data is a sample data by using point estimation. We calculate the difficulty faced by customers in operating printer manual and we project it on entire users or population. By doing so we Calculated the confidence interval or range for estimated or predicted value for the entire population by using R-graph or R-plot.

Number of customers (N)=10

Number of customers facing difficulty in operating printers manual (X)=7

P(probability)=(X/N)

=7/10

=0.7.

As the given values are categorical variable but not time variables. We use proportion test not t.test more over fiber does not change with respect to time. This are responsive in proportion or percentages. So, we use proportion test to determine confidential interval and estimate a value.

> prop.test(7,10,conf.level = 0.95,correct = TRUE)

1-sample proportions test with continuity correction

data:  7 out of 10, null probability 0.5

X-squared = 0.9, df = 1, p-value = 0.3428

alternative hypothesis: true p is not equal to 0.5

95 percent confidence interval:

 0.3536707 0.9190522

sample estimates:

  p

0.7

So the confidence interval is 0.3536707 to 0.9190522

This sample test or point estimation is projected on entire population. So the estimated population have the difficulty in operating printers will be in the range of 0.3536707 to 0.9190522. Minimum value is 35% approximately and maximum value is 91% approximately.

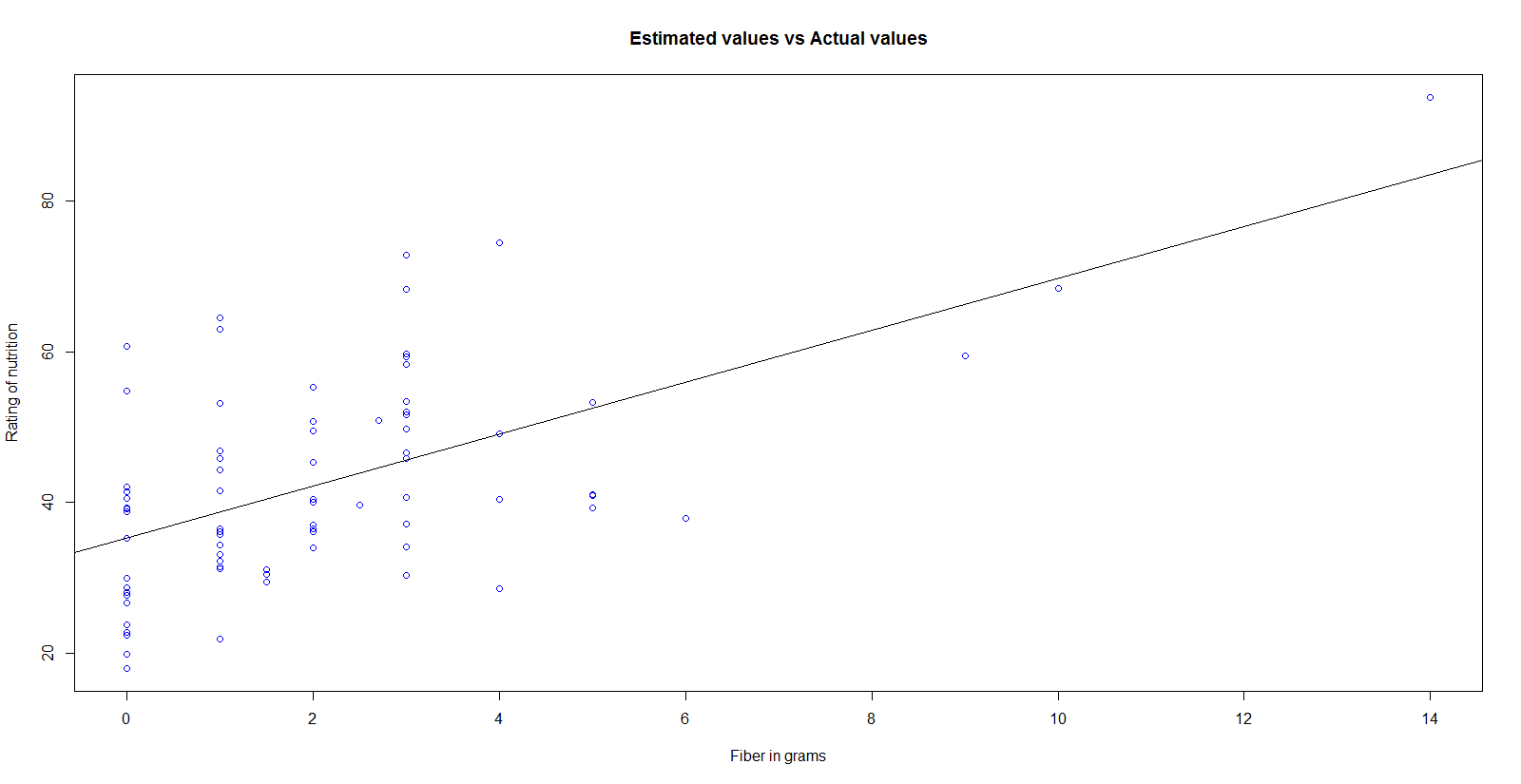
Finally, by this estimation we can clearly say that within the range of 35% to 92%(approximation) percent people will have difficulty in operating printer manual for entire users(population). Estimated value will be in this range for entire users or population.

The minimum percentage of population face difficulty in operating printer by using owner manual on the whole population is 0.3536707 approximately (35.3 percent).

2. By using regression analysis we are going to determine estimated value or predicted value and we calculate the prediction error or estimated error to see the accuracy of our estimate. we should maintain prediction error minimum to obtain tangible estimate. The dataset we are using is cereal gives information about all the ingredients, rating, weight etc. All these are the variables names or attributes. These values are given to the different products as shown in the dataset. Fiber in grams and Nutrition Rating are taken as the reference variables. In order to calculate the fiber of 3 grams effects on nutritional rating of cereals.

The regression line is written in the form y^=b0+b1x, is called the regression equation, where:

* y^ is the estimated value of the response variable.
* b0 is the y-intercept of the regression line.
* b1 is the slope of the regression line.
* b0 and b1, together, are called the regression coefficients.

plot(cereals$Fiber,cereals$Rating,xlab="Fiber in grams",ylab="Rating of nutrition",main="Estimated values vs Actual values",col="blue")

> abline(lm(cereals$Rating~cereals$Fiber))

**From graph points represents Actual Values and line represents Estimated value**

#for calculating slope:

lm(formula=cereals$Rating~cereals$Fiber)

(Intercept) cereals$Fiber or slope

35.257 3.443

The slope of the regression line indicates the estimated change in y per unit increase in x.

Here, we calculate the slope from regression line and intercept for Nutrition Rating in order to calculate the estimated value of nutrition rating for cereals with 3 grams of fiber. Actual value of nutrition rating for cereals is given in the data for 3 grams.

So, here we compare the two variables such as fiber and nutrition. Nutrition rating changes with respect to time. So, we take it on y-axis and fiber on x-axis.

Rating as y-axis because which is responsive with respect to time.

Rating = (Slope)fiber+(intercept)

y^=b1x+b0

y^= 3.443(fiber in grams)+35.357

The estimated rating is 35.357+3.443(fiber in grams). So, this regression equation used in estimation and prediction.

Here as given fiber value is equal to 3 grams.

=35,257+3.443(3.0)

=35.25+10.329

=45.579

This is estimated value for all the products with fiber content of 3 grams.

Note that this estimated value for the nutritional rating lies directly on the regression line, at the location .In fact, for any given value of x (fiber content), the estimated value for y (nutritional rating) lies precisely on the regression line.

Now we calculate the prediction error or estimation error for products such as Great-Grain-Pecan, Total-Whole-Grain, Wheat-Chex.

For 3-grams of Fiber, product Great-Grain-Pecan Actual Nutrition Rating = 45.81172

For 3-grams of Fiber, product Total-Whole-Grain Actual Nutrition Rating = 46.6588

For 3-grams of Fiber, product Wheat-Chex Actual Nutrition Rating = 49.78744

Difference between actual value and estimated value gives us predicted error or estimated error. keep error as long as small to obtain adequate result.

Calculation of predicted error (or) estimated error for three products:

Predicted error=Estimated Value-Actual Value.

1.Total-Whole-Grain: 46.65884-45.579=1.07984

2.Great-Grain-Pecan: 45.81172-45.579=0.23272

3.Wheat-chex: 49.78744-45.579=4.20844

1.07984,0.23272,4.20844 by maintaining prediction error minimal. We can obtain concise estimated value or predicted value.