Department of Industrial Engineering & Operations Research

IEOR 165 (Spring 2017)

Homework 1

Due: Friday, Feb 9

NOTE: The solutions to this HW will be posted in one week. Please do not write your solutions in red ink as this HW should be self graded or peer graded in red ink. Grades will awarded for the completion of the problems and having graded the solutions fairly. Correctness or incorrectness of the solutions will not be considered for the overall HW grade.

Question 1. Let $X_1, \dots, X_n$ be iid from the pdf

$$f(x) = \frac{2}{\pi\theta^2}\sqrt{\theta^2 - x^2}, \quad -\theta \le x \le \theta,\ 0 < \theta < \infty$$

Find the method of moments estimator of θ.

Solution. Note that the first moment is:

$$\mathbb{E}[X] = \int_{-\theta}^{\theta} \frac{2}{\pi\theta^2} x\sqrt{\theta^2 - x^2}dx = \int_0^0 \frac{2}{\pi\theta^2} u^2 du = 0 \tag{1}$$

Where we used u–substitution with $u = \sqrt{\theta^2 - x^2}$. This is not helpful so instead we consider the second moment:

$$\mathbb{E}[X^2] = \int_{-\theta}^{\theta} \frac{2}{\pi\theta^2} x^2\sqrt{\theta^2 - x^2}dx = \frac{2}{\pi\theta^2}\int_0^{\pi} \sin^2(\gamma)\cos^2(\gamma)d\gamma = \frac{2}{\pi\theta^2}\int_0^{\pi} \frac{1-\cos(4\gamma)}{8}d\gamma = \frac{\theta^2}{4} \tag{2}$$

Here we used trigonometric substitution with $\sin(\gamma) = \frac{x}{\theta}$. Solving for the second moment we obtain that our estimator is:

$$\hat{\theta} = 2\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2} \tag{3}$$

Question 2. Let $X_1, \dots, X_n$ be iid from the binomial distribution with (n, p), where integer $n > 0$ and $0 \le p \le 1$. Use the method of moments to estimate n and p. (Note: the p.m.f. of a binomial distribution is given by $f(x) = \binom{n}{x}p^x(1-p)^{n-x}$)

Solution. This distribution has two unknown parameters so we must consider both the first and second moments of the distribution:

$$\mathbb{E}[X] = np, \quad \mathbb{E}[X^2] = np(1-p) + n^2p^2 = np(1 - p + np) \tag{4}$$

Solving the first equation for n and then substituting into the following equation gives that:

$$\mathbb{E}[X^2] = \mathbb{E}[X](1 - p + \mathbb{E}[X]) \implies p = \mathbb{E}[X] + 1 - \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]} = \frac{\mathbb{E}[X] + (\mathbb{E}[X])^2 - \mathbb{E}[X^2]}{\mathbb{E}[X]} \tag{5}$$

And hence: $n = \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X]+(\mathbb{E}[X])^2-\mathbb{E}[X^2]}$. Substituting in the empirical first and second moments of the data gives the desired estimates.

Question 3. The following data set specifies the number of units of a good ordered and the price of the good at six different locations. Suppose we would like to build a linear model that predicts the number ordered as a function of price.

Number Ordered	88	101	59	48	43	61
Price	47	53	22	18	19	21

a) State the predictor and the response.

Solution. The response is Number ordered and the predictor is Price.

b) What is the linear model?

Solution. The linear model is

$$\text{Number Ordered} = \beta_0 + \beta_1 \text{Price} + \epsilon$$

c) Estimate the parameters of the linear model using least squares.

Solution. Let x be the price and y be the number ordered, from the formula of least squares we have

$$\hat{\beta}_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2} \approx 1.42$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \approx 23.94$$

Question 4. The corrosion of a certain metallic substance has been studied in dry oxygen at 500 degrees Centigrade. In this experiment, the gain in weight after various periods of exposure was used as a measure of the amount of oxygen that had reacted with the sample. The data can be found in the table below.

Hour	Percent Gain
1.0	0.035
2.0	0.024
2.5	0.030
3.0	0.030
3.5	0.055
4.0	0.029

Suppose we would like to build a linear model that predicts the percent weight gain as a function of time of exposure.

a) State the predictor and the response.

Solution. The response is Percent Gain and the predictor is Hour.

b) What is the linear model?

Solution. The linear model is

$$\text{Percent Gain} = \beta_0 + \beta_1 \text{Hour} + \epsilon$$

c) Estimate the parameters of the linear model using least squares.

Solution. Let x be the hour and y be the percent gain, from the formula of least squares we have

$$\hat{\beta}_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2} = 0.0026$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 0.0269$$

d) Predict the percent weight gain when the metal is exposed for 4.5 hours.

Solution. From the linear model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot 4.5 = 0.0386$$

So the percent weight gain after 4.5 hours is predicted as 0.0386.

Question 5. Imagine you are an election consultant for a Berkeley city council campaign and the following data is available for citizens of the district:

- Property value
- Age
- Voting Record
- Registered Party Affiliation $\in \{\text{Democratic}, \text{Republican}, \text{Unafilliated}, \text{Other}\}$

Suppose you would like to build a linear model to predict the voting record based on age, property value, and party affiliation information. Please specify the response and predictor variables.

Solution. The response: voting record. The predictors: Age, property value, registered party affiliation. And for the registered party affiliation variable, we need to code it by dummy variables. In total, we need 4 dummy variables.