

Statistics

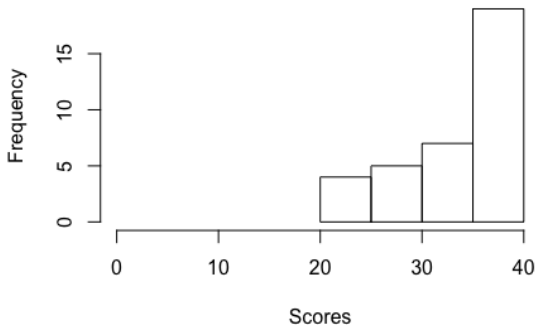
Week 13 Recitation

ESD, SUTD

Term 5, 2017

Homework 3

HW3 Score Distribution



Grading details are available on eDimension if you submitted your homework online.

Homework 3

Question 2. While trying to find the least square regression line for some data points (x_i, y_i) , a drunk statistician used the points (y_i, x_i) instead.

- (a) Does the correlation coefficient of the resulting regression line agree with the correct r ?
- (b) What about the slope of the resulting regression line: is it the same as the correct $\hat{\beta}_1$, or is it $\hat{\beta}_1$ flipped around the line $y = x$, or does something else happen?

Question 8 & 9.

Permutation and bootstrapping, find p-value and CI.

Question - Maximum Likelihood

Let $x_1, x_2, x_3 \dots x_n$ be n iid observations from a Bernoulli distribution.

The probability mass function for a Bernoulli random variable with probability of success θ may be written as

$$f(x|\theta) = \theta^x (1 - \theta)^{1-x},$$

where x is either 0 or 1.

Find the maximum likelihood estimate for θ .
Is the estimator unbiased?

Question - Maximum Likelihood (Solution)

$$\begin{aligned}L(\theta) &= \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}, \\ \log L(\theta) &= \sum_i x_i \log \theta + (n - \sum_i x_i) \log(1 - \theta), \\ \frac{d}{d\theta} \log L(\theta) &= \frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1 - \theta}.\end{aligned}$$

Setting the last equation to 0, we obtain

$$\sum_i x_i - \theta \sum_i x_i = \theta n - \theta \sum_i x_i,$$

which gives $\hat{\theta} = \bar{x}$ upon simplifying.

$$E[\hat{\theta}] = E[\bar{x}] = E[x_i] = \theta$$

Thus the estimator is unbiased.

Please complete the course survey!!!