

Data Analysis Course

Problem Set 2

Due date March 22, 2017

1. Let X be a random variable uniformly distributed in the interval $[-1, 1]$. For each $x \in [-1, 1]$ generate Y uniformly from the interval $[0, x^2]$.

a) Compute the correlation of X and Y .

b) Compute mutual information $I(X, Y)$.

2. Generate $N(0, 1)$ samples of size $N = 100$. Make $S = 10000$ such samples of size 100 each. Make maximum likelihood estimates for mean and variance for each of the 10000 samples. Plot the bivariate distribution (histogram) of $(\hat{\mu}, \hat{\sigma}^2)$.

3. If you know the mean of a normal distribution to be zero but ~~have~~ have estimate the variance, what would your maximum likelihood estimate be?

a) Is this estimator of variance unbiased?

b) " " " " " Consistent?

c) What is the variance of this estimator?

d) How does it compare with the Cramer-Rao bound?

4. A multinomial distribution with parameters $\theta_1, \dots, \theta_k$ produces random vector $(x_1, \dots, x_k)$ where $x_1 + \dots + x_k = n$. If we have N samples $(x_1^{(1)}, \dots, x_k^{(1)}), \dots, (x_1^{(N)}, \dots, x_k^{(N)})$, what would the MLE for $\theta_1, \dots, \theta_k$ be?