COS424
Homework #1
Due Thursday, February 22, 2007

See the course website for important information about collaboration and late policies, as well as where and when to turn in assignments.

Question 1. An *indicator function* is equal to one when its argument is true, and zero otherwise. Specifically,

$$\mathbf{1}(x) = \begin{cases} 1 & \text{when } x \text{ is true.} \\ 0 & \text{otherwise} \end{cases}$$

Let, X be the outcome of a (fair) six-sided die. Compute

1. The probability that $X > 2$.
2. $\mathrm{E}[\mathbf{1}(X > 2)]$

Question 2. Let $(X_1, X_2, \ldots)$ denote a sequence of IID coin flips where

$$\begin{aligned} P(X_n = H) &= \pi \\ P(X_n = T) &= 1 - \pi. \end{aligned}$$

Suppose we flip the coin until we see two heads in a row. Use iterated expectation to compute the expected number of flips we will make.

Question 3. Let $C_1, C_2, and C_3$ each be coins such that

$$\begin{aligned} P(C_1 = H) &= 0.5 \\ P(C_2 = H) &= 0.7 \\ P(C_3 = H) &= 0.5 \end{aligned}$$

Suppose we first flip C_1 and record its outcome Z. If it is heads, then we flip C_2 twice and record the outcomes (X, Y). If it is tails, then we flip C_3 twice and record the outcomes (X, Y).

- Compute the joint probability distribution of the outcomes of the flips $P(X, Y)$.
- Compute $P(Z \mid X = H)$.
- Compute $P(X \mid Y = H)$.

Question 4. Precisely describe three random variables, X, Y, and Z, such that $X \perp\!\!\!\perp Y$ and $X \not\perp\!\!\!\perp Y \mid Z$. Prove that your answer is correct using the definitions of marginal independence and conditional independence.

Question 5. Let $(x_1, \ldots, x_N)$ be binary values assumed IID Bernoulli with parameter π. Write down the log likelihood function. Derive the maximum likelihood estimate of π.

Question 6. Let $(x_1, \ldots, x_N)$ be continuous observations assumed IID Gaussian with mean μ and variance σ^2. Write down the log likelihood function. Derive the maximum likelihood estimate of μ. Derive the maximum likelihood estimate of σ^2 given an estimate of $\mu = \hat{\mu}$.

The following two questions have an R programming component and a write-up component. The write-up should be clear, concise, and thoughtful and include relevant output from the programming component, such as plots. Your grade for these questions will be determined by the completeness, correctness, and clarity of your programming and writeup.

Please submit the write-up for these questions along with the written portion of your assignment, and the programming portion of these questions through moodle. After each question, we have made clear what the electronic submission entails.

Question 7 (R). Recall the Gaussian density,

$$p(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}$$

Collect 25 continuous univariate data points from the real world. Write a function that takes the data as input and plots the following on a single plot:

- the data
- the Gaussian density fit to the data
- a vertical line at the sample mean $\hat{\mu}$
- a horizontal line with length equal to twice the square root of the sample variance $2\sqrt{\hat{\sigma}^2}$. It should be centered at the mean and at height $p(\hat{\mu} - \sqrt{\hat{\sigma}^2} \mid \hat{\mu}, \hat{\sigma}^2)$
- 25 additional points drawn from the same distribution and plotted to be distinguished from the original data

What do you notice about the data you collected compared with the fit Gaussian?

Electronic submissions for this question:

- *A data file containing your 25 data points.*
- *A file named Question7.R containing a single entry-point function named "make.plot" that will read in your data and generate the plots.*

```
make.plot <- function() {
# your code here
}
```

Question 8 (R). Write a function that takes a binary vector of data as input and returns the Bernoulli log likelihood *function* $\mathcal{L}(\pi)$ given that data. Specifically, the returned function takes one argument $\pi \in (0,1)$ and returns the log likelihood of the data given that parameter.

Sample 100 data points $(x_1, \ldots, x_{100})$ from a Bernoulli with $\pi^* = 0.5$. For various values of N, plot the log likelihood function over $\pi \in (0,1)$ using data $(x_1, \ldots, x_N)$ on a single plot.

How does the log likelihood function change for different values of N?

Electronic submissions for this question:

- *A file named Question8.R containing a single entry-point function named "make.plot" that will generate the plots, and a function named "Bernoulli.log.likelihood.function." This function takes a data set and returns a function. The returned function takes a parameter and returns the Bernoulli log likelihood evaluated for the data and parameter.*

```
Bernoulli.log.likelihood.function <- function(x) {
  # your code here
}

make.plot <- function() {
 # your code here
}
```