

Assignment #3 — Due Wednesday, February 11, 2008, by 5:00 P.M.

Turn in homework in lecture, discussion, or your TA's mailbox. Indicate the discussion section in which you expect to attend to pick up this assignment on the assignment.

311: Monday 1:20–2:10

312: Monday 12:05–12:55

For each problem, you may use a combination of paper/pencil calculations and or calculations with R. If you use R, include the commands you use for your solution. You do not need to include all of the output, but you can cut and paste if you wish. Graphing functions require the `lattice` library.

1. This problem asks you to use R for some graphics and data summaries.

- (a) Write R code to generate a random sample of size 1000 from a Normal(3, 2) distribution (using `rnorm()`) and save the data. (Note that `rnorm()` parameterizes by the mean and sd, not the mean and variance.)

```
> x = rnorm(1000,3,sqrt(2))
```

- (b) Plot the data using a density histogram over the range $(-5, 10)$ using intervals of length 1 and using intervals of length 0.1. Use `histogram()`. Compare the plots.

- (c) Plot the data using `densityplot()`. Compare this method for graphing continuous quantitative variables with the previous histograms.

```
> print( densityplot(~x) )
```

- (d) Plot the data with a box-and-whisker plot using `bwplot()`. Compare this method for graphing continuous quantitative variables with the previous plots.

```
> print( bwplot(~x) )
```

- (e) Use `fivenum()` to find the five-number summary of the data (minimum, lower quartile, median, upper quartile, maximum). How are these numerical values represented in the box-and-whisker plot?

- (f) Compute the sample mean, standard deviation, variance, and median using `mean()`, `sd()`, `var()`, and `median()`. Compare the sample estimates with the expected values from a normal distribution.

2. Do Exercise 5.5.5 using R.

3. For a statistical model $\{f_\theta(s) : \theta \in \Omega\}$ with sufficient statistic $T(s)$, prove that for any two parameter values $\theta_1, \theta_2 \in \Omega$ and for any two data samples s_1 and s_2 such that $T(s_1) = T(s_2)$ that

$$\frac{L(\theta_1 | s_1)}{L(\theta_2 | s_1)} = \frac{L(\theta_1 | s_2)}{L(\theta_2 | s_2)} .$$

In other words, if a data set is swapped for another with the same sufficient statistic, there is no change in the likelihood ratio to compare to arbitrary parameter values.

4. Let X_1, \dots, X_n be an i.i.d. sample of size n from a Uniform $(-\theta, \theta)$ distribution.

- (a) Find a minimal sufficient statistic for θ .

- (b) Find a maximum likelihood estimate of θ .

5. Let α_0 be a known fixed constant and let X_1, \dots, X_n be an i.i.d. sample from a Gamma distribution with density for a single observation

$$f_\theta(x) = \frac{\theta^{\alpha_0}}{\Gamma(\alpha_0)} x^{\alpha_0-1} e^{-\theta x}, \quad \text{for } x > 0, \theta > 0.$$

- (a) Find a minimal sufficient statistic for θ .
(b) Find a maximum likelihood estimate of θ .
6. Let X_1, \dots, X_n be an i.i.d. sample from a Beta($\theta, 1$) distribution with density for a single observation

$$f_\theta(x) = \frac{\Gamma(\theta + 1)}{\Gamma(\theta)\Gamma(1)} x^{\theta-1}, \quad \text{for } 0 < x < 1, \theta > 0.$$

- (a) Find a minimal sufficient statistic for θ .
(b) Find a maximum likelihood estimate of θ .

Work to do, but not turn in.

- Read section 6.3.
-