# STA261: Assignment 3

*Alex Stringer*

*January, 2018*

This assignment is not for credit. Complete the questions as preparation for quizzes and tests.

Suggested reading: Textbook sections 8.5 and 8.8. Note we didn't do all of 8.5 this week.

1. *Sufficiency*: Prove that any one-to-one function of a sufficient statistic is also sufficient.

2. *Sufficiency*: Show the following estimators are sufficient for their respective population parameters, for the following independent random samples and corresponding distributions. If you use the *factorization theorem*, be sure to state the functions $g(\hat{\theta}, \theta)$ and $h(\mathbf{x})$.

   (a) $X_i \sim Gamma(\alpha, \beta)$ with density $f_{x_i}(x_i) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}}$, $(\hat{\alpha}, \hat{\beta}) = \left(\prod_{i=1}^n x_i, \sum_{i=1}^n x_i\right)$

   (b) $X_i \sim N(\mu, \sigma)$, $(\hat{\mu}, \hat{\sigma}) = \left(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2\right)$

   (c) $X_i \sim N(\mu, \sigma)$, $(\hat{\mu}, \hat{\sigma}) = \left(\bar{x}, \bar{x^2}\right)$

   (d) $X_i \sim Beta(\alpha, \beta)$ with $f_{x_i} = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} x_i^{\alpha-1} (1-x_i)^{\beta-1}$, $\left(\hat{\alpha}, \hat{\beta}\right) = \left(\prod_{i=1}^n x_i, \prod_{i=1}^n (1-x_i)\right)$

   (e) $X_i \sim Beta(\alpha, \beta)$ as before, $\left(\hat{\alpha}, \hat{\beta}\right) = \left(\sum_{i=1}^n \log x_i, \sum_{i=1}^n \log(1-x_i)\right)$

3. *Sufficiency*. For $X_i \sim Unif(a, b)$, the *continuous* uniform distribution on $(a, b)$, find a sufficient statistic for $(a, b)$. *Hint*: the density is only defined over a certain subset of $\mathbb{R}$, what is it? Make sure to include the corresponding *indicator function* of the support when you write out the density, i.e.

$$f_{x_i}(x_i) = \frac{1}{b-a} \times I(support)$$

   It's actually good form to *always* do this, even if I often don't do it for you.

4. Show that the following two statistics are sufficient for any parameter from any distribution:

   (a) The full dataset, $\mathbf{x} = (x_1, \ldots, x_n)$

   (b) The *order statistics*, which are just the ordered sample values $(x_{(1)}, \ldots, x_{(n)})$ with $x_{(1)} \leq \ldots \leq x_{(n)}$

5. State and prove the *factorization theorem*.

6. State and prove the *Rao-Blackwell theorem*.

7. Suppose we flip a coin repeatedly so that $X_i \sim Bern(p), i = 1 \ldots n$ as usual, and we want to estimate $p$. I am a very stubborn person- I say that we only base our estimate $\hat{p}$ off of the first $m < n$ flips,

$$\hat{p} = \frac{1}{m} \sum_{i=1}^m x_i$$

(a) Use the Rao-Blackwell theorem to find an estimator $\tilde{p}$ of $p$ with lower variance than my estimator.

(b) Verify both estimates' respective variance for data $\mathbf{x} = (1, 1, 1, 0, 0, 1, 0, 0, 1)$, with $m = 4$.

8. *Likelihood*: Let $X_i \sim Bern(p)$ be a sequence of independent coin flips. Find the likelihood function for $\mathbf{x} = (x_1, \ldots, x_n)$. Compare your answer to the *binomial* probability mass function, and explain why they are slightly different.

9. *Likelihood*: Suppose we observe a sequence of random variables following a *Markov Process*, such that

$$\begin{aligned} X_0 \sim & N(0, \sigma) \\ X_j | X_{j-1} = x_{j-1} \sim & N(x_{j-1}, \sigma), j = 1 \ldots n \\ X_j | X_{j-1} \perp & (X_{j-2}, \ldots, X_0) \end{aligned}$$

This notation means that each random variable $X_j$ is dependent only on the previous value $X_{j-1}$, and is independent of the rest of the sample. Find the likelihood function for $\sigma$. Don't worry about the normalization constant. *Hint* if you're stuck, it helps to try this out for a few values of *j*, e.g. $j = 1, 2, \ldots$ and then try to generalize your answer. You don't need to simplify your answer; just write down the correct formula for the likelihood.

10. *Maximum Likelihood*: Show that the maximum likelihood estimator can depend on the data only through a function of a sufficient statistic.

11. *Maximum Likelihood*: Let $X_i \sim N(\mu, \sigma)$ independently. Find the maximum likelihood estimator for $(\mu, \sigma)$. As mentioned in lecture, you don't need to do the second derivative test.