STA 711 Homework 8

Due: Monday, April 21, 10:00pm on Canvas.

Instructions: Submit your work as a single PDF. You may choose to either hand-write your work and submit a PDF scan, or type your work using LaTeX and submit the resulting PDF. See the course website for a homework template file and instructions on getting started with LaTeX and Overleaf.

Confidence intervals

1. Suppose $X_1, ..., X_n \overset{iid}{\sim} N(\theta, \theta)$, where $\theta > 0$. Find a pivotal quantity $Q(X_1, ..., X_n, \theta)$, and use the quantity to create a $1 - \alpha$ confidence interval for θ.

2. Suppose $X_1 \overset{iid}{\sim} Uniform[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$. Find a $1 - \alpha$ confidence interval for θ, using the single observation X_1.

3. Suppose that $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

   (a) If σ^2 is known, the interval for μ is $\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, and the *width* of the interval is $2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Find the minimum value of n so that a 95% confidence interval for μ will have a length of at most $\sigma/4$.

   (b) If σ^2 is unknown, the interval for μ is $\overline{X} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$, where $s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$. Find the minimum value of n such that, with probability 0.9, a 95% confidence interval for μ will have a length of at most $\sigma/4$.

4. Let $X_1, ..., X_n$ be an iid sample from the *inverse Gaussian* distribution, with pdf

$$f(x|\mu) = \frac{1}{\sqrt{2\pi x^3}} \exp\left\{-\frac{(x-\mu)^2}{2\mu^2 x}\right\} \qquad x > 0, \mu > 0.$$

   On Exam 2, you showed that $\sqrt{n}(\overline{X} - \mu) \xrightarrow{d} N(0, \mu^3)$.

   Using delta method, find a variance stabilizing transformation g such that the (asymptotic) variance of $g(\overline{X})$ does not depend on μ.

Sufficient statistics and minimal sufficiency

In class, we discussed *sufficient statistics*. Informally, a sufficient statistic captures all the information about a parameter of interest. In that sense, a sufficient statistic is a form of data reduction. However, there are many different possible sufficient statistics, and we might ask which reduction is the "most efficient".

A **minimal sufficient statistic** $T(X_1, ..., X_n)$ is one which achieves the greatest possible data reduction. That is, if $T'(X_1, ..., X_n)$ is another sufficient statistic, then $T(X_1, ..., X_n)$ is a function of $T'(X_1, ..., X_n)$ (see definition 6.2.11 in Casella and Berger).

Theorem 6.2.13 in Casella and Berger tells us how to find a minimal sufficient statistic. Let $f(x_1, ..., x_n|\theta)$ be the joint probability function, and suppose that the ratio $f(x_1, ..., x_n|\theta)/f(y_1, ..., y_n|\theta)$ does not depend on θ if and only if $T(x_1, ..., x_n) = T(y_1, ..., y_n)$. Then, T is a minimal sufficient statistic.

In the following questions, you will practice finding minimal sufficient statistics. I recommend reading section 6.2.1, including the definition and theorem mentioned here and examples 6.2.14 and 6.2.15.

5. Suppose that $X_1, ..., X_n \overset{iid}{\sim} Geometric(p)$. Using the information above, find a minimal sufficient statistic for p.

6. Suppose that $X_1, ..., X_n \overset{iid}{\sim} Uniform(a, b)$. Find a minimal sufficient statistic for (a, b). *Hint:* Like in the normal example in class and in 6.2.14, your minimal sufficient statistic will be a vector.