MATH/STAT 355: Problem Set 5

Prof. Taylor Okonek

Due: April 18, 2025

Hypothesis Testing

1. Suppose we observe a random sample $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$, where σ^2 is known, and recall that the MLE for μ is given by $\hat{\mu}_{MLE} = \overline{X}$. Let our null and alternative hypotheses be specified by $H_0 : \mu = 0$, $H_1 : \mu \neq 0$.

   (a) Derive the Wald test statistic, and show that the asymptotic distribution of the Wald test statistic is χ^2_1 when H_0 is true.

   (b) Derive the LRT statistic, and show that the asymptotic distribution of the LRT statistic is χ^2_1 when H_0 is true.

   (c) Derive the score test statistic, and show that the asymptotic distribution of the score test statistic is χ^2_1 when H_0 is true.

   *Note: These three test statistics are asymptotically equivalent for more than just normally distributed random variables, we're just using it as one example here!

2. Suppose we have iid observations $X_1, \dots, X_n$, and an MLE given by $\overline{X}$, where $E[X_i] = \theta$, and $Var[X_i] = \sigma^2$, where σ^2 is known. For this question, let $\sigma^2 = 10$, $n = 5$, $\{X_1, X_2, X_3, X_4, X_5\} = \{3.5, 7, 8, 2.75, 6\}$, and assume $\theta > 0$.

   One downside to the Wald test statistic is that is *non-invariant to reparameterizations.* This means that the test statistic may have a different value if we reparameterize our hypotheses, even if they are technically testing the same thing. Demonstrate this property, by showing that the following two sets of hypotheses lead to different Wald test statistics in this scenario:

   i. $H_0 : \theta = 1$, $H_1 : \theta \neq 1$

   ii. $H_0 : \log(\theta) = 0$, $H_1 : \log(\theta) \neq 0$

3. Suppose $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$, where both μ *and* σ are unknown. Recall that we have previously shown

$$\sqrt{n}\left(\frac{\overline{X} - \mu}{s_n}\right) \xrightarrow{d} N(0, 1)$$

   where s_n is the sample standard deviation, given by $s_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2}$.

   (a) Write $\sqrt{n}\left(\frac{\overline{X}-\mu}{s_n}\right)$ as $\frac{\sqrt{n-1}Y}{\sqrt{Z}}$, noting what distributions Y and Z have.

   (b) Let $W = \frac{\sqrt{n-1}Y}{\sqrt{Z}}$. Use the result from part (a) to show that, in *finite* samples (i.e. no asymptotics, no convergence, etc.), $W \sim t_{n-1}$.*

*This is why, when you fit a linear regression model in `R`, you get a t-statistic instead of a z-statistic! When we don't assume the variance of our observations is known and instead *estimate* it using the sample variance (more realistic), our test statistic follows a t-distribution rather than a normal distribution, in finite samples.

Bayesian Statistics

1. Suppose we have a random sample $Y_1, \ldots, Y_n \overset{iid}{\sim} Exponential(\lambda)$ where our prior distribution for λ is Gamma(α, β). Find the posterior distribution for λ.

2. Suppose we have a random sample $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$ with *known* mean μ, and an Inverse-Gamma(α, β) prior on σ^2. The pdf of a random variable $W \sim$ Inverse-Gamma(α, β) is given by

$$\pi(w \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{w}\right)^{\alpha+1} e^{-\beta/w}.$$

   Find the posterior distribution for σ^2.

3. The purpose of this problem is to demonstrate the idea of sequentially updating our beliefs using a Bayesian model. Prof. Kristin Heysse and Taylor often flip their favorite coins in their spare time (ha!), and decide to compete to see who can get the most "heads" in 50 coin flips. Taylor gets 23 heads, and Kristin gets 38 heads, which causes Taylor to suspect that Kristin's coin isn't fair!

   You will need to use `R` for this question. Just paste the few lines of code you need into your LaTeXdocument for your problem set submission. You may also use Wolfram Alpha, or something similar, to solve a system of equations in parts (b) and (c).

   (a) Derive the posterior distribution for θ, the probability of heads, using a $Beta(\alpha, \beta)$ (plugging in actual numbers for coin flips, etc., as you go).

   (b) Calculate the posterior probability that Kristin would have flipped *at least* 38 out of 50 heads, given a prior belief that her coin was fair (i.e., 50% chance of heads) *and* that the variation in her coin flips was only 10% (i.e., $Var(X) = 0.1$, where X denotes a single coin flip). Assume a Beta prior distribution.

   (c) Calculate the posterior probability that Kristin would have flipped between 20 and 30 heads, given a prior belief that her coin has a 76% chance of heads, and the variation in her coin flips was again 10%.* Again, assume a Beta prior distribution.

   (d) Suppose Kristin (frustrated with Taylor's disbelief) flips her favorite coin 30 more times, and this time observes 18 heads. Update your posterior probability that Kristin would have originally flipped *at least* 38 out of 50 heads, in light of this new information. (Hint: treat the posterior probability calculated in Part b as a new *prior* distribution).

   *Note: This is called an *empirical* Bayesian approach, where the prior beliefs we specify are calculated from the data we observe. A fun fact: empirical Bayesian hierarchical models are mathematically equivalent to frequentist mixed effects models!)

4. This problem will walk you through one way to construct a Frequentist confidence interval in a Bayesian framework.

   Suppose we have data $Y_1, \ldots, Y_n \sim f(\mathbf{y} \mid \theta)$ for some probability density function f. Let $C(\mathbf{y})$ denote a $1 - \alpha$ level confidence region for θ, such that

$$\Pr(\theta \in C(\mathbf{y})) = \int I\{\theta \in C\} f(\mathbf{y} \mid \theta) d\mathbf{y} \geq 1 - \alpha.$$

   Additionally, suppose we have a prior for θ, $f(\theta)$. Define the prior-posterior ratio as

$$R(\theta) = \frac{f(\theta)}{f(\theta \mid \mathbf{y})}.$$

   Show that the confidence region $C_R(\mathbf{y}) = \{\theta \mid R(\theta) \leq 1/\alpha\}$ is a $1 - \alpha$ level confidence region. (Hint: Markov's inequality)