STAT120C Assignment 1

Due: April 8th by 5pm in the STAT120C (Pluta) Dropbox in DBH

Problem 1. Suppose $X_1, \cdots, X_n$ are i.i.d. $N(\mu, \sigma^2)$.

(a) Compute $E[\bar{X}]$ and $Var[\bar{X}]$, where $\bar{X}$ is defined as $\frac{1}{n}\sum_{i=1}^{n} X_i$

(b) What is the distribution of $X_i - \mu$? What is the distribution of $(X_i - \mu)/\sigma$? What is the distribution of $\bar{X} - \mu$?

(c) If μ and σ^2 are known, how can you construct a χ^2-distributed random variable using $X_1, \cdots, X_n$? Provide the degrees of freedom.

(d) How can you construct a t-distributed random variable using $X_1, \cdots, X_n$? Provide the degrees of freedom for your constructed random variable.

Problem 2. Suppose $X_1, \cdots, X_n$ are i.i.d. $N(\mu, \sigma^2)$, with both mean μ and variance σ^2 unknown.

(a) Find the maximum likelihood estimates of μ and σ^2.

(b) Derive and simplify the likelihood ratio Λ for $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$.

(c) What distribution does $-2log(\Lambda)$ follow under the null hypothesis?

Problem 3. (a) Show that the MLE $\hat{\sigma}^2$ derived in 2(a) is biased for σ^2. Does the MLE overestimate or underestimate σ^2 on average?

(b) What effect might this bias have on testing results when $\hat{\sigma}^2$ is used? Will the test reject or fail to reject H_0 more often than if an unbiased estimate is used?

Problem 4. Suppose $X_1, \cdots, X_n$ are i.i.d. $N(\mu, \sigma^2)$. To test the null hypothesis $H_0 : \mu = \mu_0$, the t-test is often used:

$$t = \frac{\bar{X} - \mu_0}{s_{\bar{X}}}$$

Under H_0, t follows a t distribution with $n-1$ df. The null hypothesis is rejected when t^2 is large. Show that in the likelihood ratio test, the null hypothesis is also rejected for large values of t^2.

Hints:

(a) You might want to use results you obtain from **Problem 2 (b)**.

(b) The following equation is also helpful. Prove it before you use it (it is not very difficult).

$$\sum_{i=1}^{n}(X_i - \mu_0)^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(\bar{X} - \mu_0)^2$$

Problem 5. (a) Write the definition of a P-value in the Null Hypothesis Significance Testing framework (NHST).

(b) Carefullly explain why the P-value is **not** the "probability that H_0 is true".

Problem 6. Suppose $X_1, \cdots, X_n$ are i.i.d. $N(\mu, \sigma^2)$, and that we want to test the null hypothesis $H_0 : \mu = 0$.

(a) Write the formula for the test statistic for a one-sample t-test from this sample.

(b) What is the reference distribution for the statistic in (a)? Draw a picture or produce a plot of the reference distribution and shade the region corresponding to the P-value for a right-tailed test.

(c) How will the P-values compare for the two possible alternative hypotheses $H_1 : \mu = 1$, and $H_1 : \mu = 100$? Explain your answer.

(d) How will the P-values compare for a left-tailed, right-tailed, and two-tailed test?