STAT120C Assignment 1

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

(a) $E[\bar{X}] = \mu$, $Var[\bar{X}] = \frac{\sigma^2}{n}$

(b) $X_i - \mu \sim \mathcal{N}(0, \sigma^2),\ (X_i - \mu)/\sigma \sim \mathcal{N}(0,1),\ \bar{X} - \mu \sim \mathcal{N}(0, \sigma^2/n)$?

(c) Two examples:

$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma^2}\right)^2 \sim \chi^2_{n-1}$

$\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2_1$

(d) $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$, where s^2 is the sample variance. See derivation in Lecture 1.

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.

$$\begin{aligned}
\mathcal{L}(\mu, \sigma^2 | X) &= (2\pi\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \right\} \\
\ell(\mu, \sigma^2) &= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \\
\frac{\partial \ell}{\partial \mu} &= -\frac{1}{2\sigma^2}\sum -2(X_i - \mu) \\
0 &= \sum (X_i - \hat{\mu}) \Rightarrow \hat{\mu} = \bar{X} \\
\frac{\partial \ell}{\partial \sigma^2} &= -\frac{n}{2} \cdot \frac{1}{\sigma^2} - \frac{1}{(\sigma^2)^2} \sum (X_i - \mu)^2 \\
0 &= \frac{n}{\hat{\sigma}^2} - \frac{1}{(\hat{\sigma}^2)^2}\sum (X_i - \hat{\mu})^2 \\
\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2
\end{aligned}$$

(b) Let $\hat{\sigma}^2$ be the MLE of σ^2 and $\hat{\sigma}_0^2$ be the MLE of σ^2 when $\mu = \mu_0$.

$$\begin{aligned}
\frac{\max_{\Omega_0} \mathcal{L}(\mu, \sigma^2|X)}{\max_{\Omega} \mathcal{L}(\mu, \sigma^2|X)} &= \frac{(2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left\{\frac{-1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(X_i - \mu_0)^2\right\}}{(2\pi\hat{\sigma}^2)^{-n/2}\exp\left\{\frac{-1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(X_i - \hat{\mu})^2\right\}} \\
&= \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2}
\end{aligned}$$

Problem 3. (a) Recall that $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$, from which it follows that $E[s^2] = \sigma^2$. We can write $\hat{\sigma}^2 = \frac{n-1}{n}s^2$, implying $E[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2$.

(b) Using $\hat{\sigma}^2$ instead of s^2 could inflate the type I error, i.e. H_0 will be rejected more often than desired due to underestimating the variance.

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.

Solution: The LRT statistic is equivalent to rejecting H_0 when $\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}$ is large. This is equivalent to rejecting H_0 when $\frac{(\bar{X}-\mu_0)^2}{\frac{1}{n}\sum(X_i-\bar{X})^2}$ is large. Multiplying the denominator by $\frac{1}{n-1}$ gives the t^2 statistic.

Problem 5. (a) The P-value of testing a hypothesis H_0 from a test statistic $T(X)$ is the probability of a sample having statistic as extreme or more extreme than the observed data, assuming H_0 is true.

(b) In the NHST framework, popluation parameters are assumed to be fixed, rather than random, so H_0 is true with probability 0 or 1.

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) & (b) See Lecture 1.

(c) P-values will be equal in the NHST framework, since the P-value only depends on the null hypothesis and the observed data.

(d) $P_{right} = 1 - P_{left}$, $P_{two} = 2 * min(P_{right}, P_{left})$.