Problem Set #6

1. From Lecture 13b, we have

$$L(\mu|X) = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^n \exp\left(\frac{-\sum_{i=1}^n (X_i - \mu)^2}{2\sigma^2}\right)$$

The likelihood ratio statistic:

$$\lambda(X) = \frac{L(\hat{\mu}_{MLE}|\, X)}{L(\mu_0|\, X)} = \exp\left(\frac{\sum_{i=1}^n (X_i - \mu_0)^2 - \sum_{i=1}^n (X_i - \bar{X}_n)^2}{2\sigma^2}\right)$$

Therefore, we have two cases. When $\bar{X}_n < 0$, our best estimate for μ_0 will be $\bar{X}_n$ and

$$\lambda(X) = \frac{L(\hat{\mu}_{MLE}|\, X)}{L(\mu_0|\, X)} = 1 > c$$

Otherwise, if $\bar{X}_n \geq 0$, our closest estimate for $\mu_0$will be $\mu_0 = 0$:

$$\lambda(X) = \exp\left(\frac{\sum_{i=1}^n (X_i - 0)^2 - \sum_{i=1}^n (X_i - \bar{X}_n)^2}{2\sigma^2}\right) = \exp\left(\frac{n(\bar{X}_n)^2}{2\sigma^2}\right)$$

We reject when

$$\left|\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}}\right| > c' = \sqrt{2\log(c)}$$

Making a test size:

$$P\left(\left|\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}}\right| > c'\right) = \alpha$$

According to the notes

$$\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Therefore,

$$P\left(\left|\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}}\right| > c'\right) = 1 - P\left(\left|\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}}\right| \leq c'\right) = 1 - \phi(c') = \alpha$$

$$c' = \phi^{-1}(1-\alpha)$$

Solving,

$$\left|\frac{\bar{X}_n}{\frac{\sigma}{\sqrt{n}}}\right| > \phi^{-1}(1-\alpha)$$

$$\bar{X}_n > \frac{\sigma\phi^{-1}(1-\alpha)}{\sqrt{n}}$$

Which is the same.

2. Similar to the previous question, we have the likelihood function with a subtle difference:

$$L(\mu|X) = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^n \exp\left(\frac{-\sum_{i=1}^{n}(X_i-\mu)^2}{2\sigma^2}\right)$$

However, we must do an MLE for the variance and the mean. We know that

$$\hat{\sigma}_{MLE}{}^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}$$

$$\hat{\mu}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$$

Therefore, we can set up our ratio:

$$\lambda(X) = \frac{L(\hat{\mu}_{MLE}|\,X)}{L(\mu_0|\,X)} = \frac{\left(\frac{1}{\sqrt{2\pi\hat{\sigma}_{MLE}}}\right)^n \exp\left(\frac{-\sum_{i=1}^{n}(X_i-\hat{\mu}_{MLE})^2}{2\hat{\sigma}_{MLE}{}^2}\right)}{\left(\frac{1}{\sqrt{2\pi\sigma_0}}\right)^n \exp\left(\frac{-\sum_{i=1}^{n}(X_i-\mu_0)^2}{2\sigma_0^2}\right)}$$

With

$$\sigma_0^2 = \frac{\sum_{i=1}^{n}(X_i-\mu_0)^2}{n}$$

$$\lambda(X) = \frac{\left(\frac{1}{\sqrt{2\pi\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n}}}}\right)^n \exp\left(\frac{-\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{2\frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n}}\right)}{\left(\frac{1}{\sqrt{2\pi\sqrt{\frac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{n}}}}\right)^n \exp\left(\frac{-\sum_{i=1}^{n}(X_i - \mu_0)^2}{2\frac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{n}}\right)}$$

$$= \left(\sqrt{\frac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}}\right)^n > c$$

$$= \frac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2} > c^{\frac{2}{n}}$$

Using the property from Piazza,

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

$$\lambda(X) = \frac{n(\bar{X}_n - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2} > c^{\frac{2}{n}} - 1$$

To get the sample variance, we need to have $\frac{1}{n-1}$:

$$= \frac{n(\bar{X}_n - \mu_0)^2}{(n-1)s_n^2} > c^{\frac{2}{n}} - 1$$

$$= \left|\frac{\bar{X}_n - \mu_0}{\frac{s_n}{\sqrt{n}}}\right| > \sqrt{\left(c^{\frac{2}{n}} - 1\right)(n-1)}$$

This takes the form of a t-statistic, so we reject the null hypothesis when

$$\left|\frac{\bar{X}_n - \mu_0}{\frac{s_n}{\sqrt{n}}}\right| > t_{n-1,1-\frac{\alpha}{2}}$$

3. We have

$$Y_i = \beta X_i + e_i$$

$$||e_i|| = ||Y_i - \beta X_i||$$

Therefore, we want

$$\min_{\beta} ||Y - \beta X||^2$$

$$= \min_{\beta} \sum_{i=1}^{n} (Y_i - \beta X_i)^2$$

We differentiate and set equal to zero

$$\frac{d}{d\beta}\left(\sum_{i=1}^{n} (Y_i - \beta X_i)^2\right) = -2\sum_{i=1}^{n}\left(X_i * (Y_i - \beta X_i)\right) = 0$$

$$\sum_{i=1}^{n} X_i Y_i - \beta \sum_{i=1}^{n} X_i^2 = 0$$

$$\beta = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$

Showing this is a minimum:

$$\frac{d^2}{d\beta^2}\left(\sum_{i=1}^{n} (Y_i - \beta X_i)^2\right) = \sum_{i=1}^{n} 2X_i^2$$

This is always positive, assuming we don't have the case where all $X = 0$, since anything squared is positive. Therefore, since the second derivate is positive when the first derivative is zero, this is the minimum.