Worksheet 10 (Solutions)

1. (Ratio Test) Let $X_1, \ldots, X_n \overset{iid}{\sim} Exp(\lambda)$. What is the test statistic Λ for the corresponding likelihood ratio test for the null hypothesis $H_0 : \lambda = 1$.

Solution: In general for the exponential distribution the likelihood ratio test statistic will be:

$$\begin{aligned}
\Lambda &= 2 \cdot \sum_i \left[l(\hat{\lambda}) - l(\lambda_0) \right] \\
&= 2 \cdot \sum_i \left[\log(\hat{\lambda}) - \hat{\lambda} x_i - \log(\lambda_0) + \lambda_0 x_i \right] \\
&= 2 \cdot \left[\sum_i \log(\hat{\lambda}) - \hat{\lambda} \sum_i x_i - \sum_i \log(\lambda_0) + \sum_i \lambda_0 x_i \right] \\
&= 2 \cdot \left[n \log(\hat{\lambda}) - \hat{\lambda} n \cdot \bar{x} - n \log(\lambda_0) + \lambda_0 \cdot n \bar{x} \right] \\
&= 2 \cdot n \left[\log(\hat{\lambda}) - \hat{\lambda} \bar{x} - \log(\lambda_0) + \lambda_0 \bar{x} \right]
\end{aligned}$$

We know from last time that $\hat{\lambda} = \bar{x}^{-1}$ and have that $\lambda_0 = 1$ from the null-hypothesis. So:

$$\begin{aligned}
\Lambda &= 2 \cdot n \left[\log(\bar{x}^{-1}) - \bar{x}^{-1} \bar{x} - \log(1) + 1 \cdot \bar{x} \right] \\
&= 2 \cdot n \left[- \log(\bar{x}) - 1 - 0 + \bar{x} \right] \\
&= 2 \cdot n \left[\bar{x} - \log(\bar{x}) - 1 \right].
\end{aligned}$$

And that's as much as we can simplify.

2. (Ratio Test) Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. What is the test statistic Λ for the corresponding likelihood ratio test for the null hypothesis $H_0 : \lambda = 1$.

Solution: In general for the Poisson distribution, the likelihood ratio test statistic will be:

$$\begin{aligned}
\Lambda &= 2 \cdot \sum_i \left[l(\hat{\lambda}) - l(\lambda_0) \right] \\
&= 2 \cdot \sum_i \left[x_i \log(\hat{\lambda}) - \hat{\lambda} - \log(x_i!) - x_i \log(\lambda_0) + \lambda_0 + \log(x_i!) \right] \\
&= 2 \cdot n \left[\bar{x} \log(\hat{\lambda}) - \hat{\lambda} - \bar{x} \log(\lambda_0) + \lambda_0 \right]
\end{aligned}$$

We know that $\hat{\lambda} = \bar{x}$ for a Poisson MLE and have $\lambda_0 = 1$ from the null-hypothesis, so:

$$\begin{aligned}\Lambda &= 2 \cdot n \left[\bar{x}\log(\bar{x}) - \bar{x} - \bar{x}\log(1) + \lambda_0\right] \\ &= 2 \cdot n \left[\bar{x}\log(\bar{x}) - \bar{x} + \lambda_0\right] \\ &= 2 \cdot n \left[\bar{x}\log(\bar{x}) - \bar{x} + 1\right]\end{aligned}$$

And that's about all that we can do.

3. **(Ratio Test)** Let $X_1, \dots, X_n \overset{iid}{\sim} Bernoulli(p)$. What is the test statistic Λ for the corresponding likelihood ratio test for the null hypothesis $H_0 : p = 0.2$.

Solution: One more, same idea. In general for the Bernouilli distribution, the likelihood ratio test statistic will be:

$$\begin{aligned}\Lambda &= 2 \cdot \sum_i \left[l(\hat{p}) - l(p_0)\right] \\ &= 2 \cdot \sum_i \left[x_i \log(\hat{p}) + (1 - x_i)\log(1 - \hat{p}) - x_i \log(p_0) - (1 - x_i)\log(1 - p_0)\right] \\ &= 2 \cdot n \left[\bar{x}\log(\hat{p}) + (1 - \bar{x})\log(1 - \hat{p}) - \bar{x}\log(p_0) - (1 - \bar{x})\log(1 - p_0)\right]\end{aligned}$$

Plugging in the MLE ($\hat{p} = \bar{x}$) and the null hypothesis ($p_0 = 0.2$), we have:

$$\Lambda = 2 \cdot n \left[\bar{x}\log(\bar{x}) + (1 - \bar{x})\log(1 - \bar{x}) - \bar{x}\log(0.2) - (1 - \bar{x})\log(1 - 0.2)\right]$$

And again, there's not much more to simplify here. We could easily compute the value of Λ for a particular dataset, and then compare the a chi-squared distribution.

4. **(Ratio Test)** Let $X \sim Bin(n, p_1)$ and $Y \sim Bin(n, p_2)$ be independent random variables, assuming that n is a known quantity. We want to test the hypothesis that $H_0 : p_1 = p_2$. What are the corresponding Θ and Θ_0 in our updated formulation of hypothesis testing?[1] If we use a Likelihood Ratio Test for this hypothesis, how many degrees of freedom should Λ have?

[1] We will derive the actual test itself in a more general form next class.

Solution: We will write values of the parameter $\theta = (p_1, p_2)$. These can be any values between 0 and 1, so $\Theta = [0, 1] \times [0, 1] \subset \mathbb{R}^2$. The null-hypothesis is then the subset of this where the two values are equal: $\Theta_0 = \{(x, y) | x = y, x \in [0, 1]\} \subset \Theta$. Visually, this is a line of values between the origin and the point $(1, 1)$. The difference in dimensionality is $2 - 1 = 1$, so $\Lambda \sim \chi^2(1)$.

5. **(Ratio Test)** Recall that we used the one-sample ANOVA test with the null-hypothesis that the means of K samples are all the same.

Write down and describe the values of Θ and Θ_0 that correspond to this test. If we use a Likelihood Ratio Test for this hypothesis, how many degrees of freedom should Λ have?

Solution: Here, we will let $\theta = (\mu_1, \ldots, \mu_k, \sigma^2) \in \mathbb{R}^{k+1}$. For the parameter space we have:

$$\begin{aligned}\Theta &= \mathbb{R}^k \times (0, \infty) \\ &= \{(x_1, \ldots, x_k, x_{k+1}) | x_{k+1} > 0\}\end{aligned}$$

The null hypothesis is the set where the first k elements are equal:

$$\Theta_0 = \{(x_1, \ldots, x_k, x_{k+1}) | x_i = x_j (i <= k, j <= k), x_{k+1} > 0\}$$

6. (MLE Practice) Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, a)$. Find the MLE estimator for a. Note: You cannot do this using the derivative. Just think about it!

Solution: The density $f(x)$ of the uniform distribution from 0 to a will be $1/a$ if $x \in [0, a]$ and zero otherwise. So, to maximize the likelihood, clearly we need to have $a \leq max_i\{x_i\}$ (otherwise the likelihood is zero). But, as long as the maximum is no larger than a, the likelihood decreases with with a larger a. So, $\hat{a} = max_i\{x_i\}$. If that's not clear, ask me to draw you a picture in class!