Some of my white board writing in weeks 8&9

1) In week 8, I first discussed the examples on p. 45-47 of the notes. Looking at the amount of material provided in the notes, I believe that you should be able to reconstruct the details of these examples by yourself.

In week 9, I started with recalling the Neyman-Pearson Lemma.

2) Proof of the Neyman-Pearson Lemma: Again, I did complete derivation on the white board but looking at the content that is put on p. 50-51 of the notes, I believe that you should be able to reconstruct the details of the proof yourself. Then I did:

3.) Example about uniformly most powerful (UMP) α-test for the normal distribution:

$\mathbb{X} = (X_1, X_2, \dots, X_n)$ i.i.d. $N(\theta, 1)$. Consider $H_0: \theta = \theta_0 \in \mathbb{R}^1$ versus a composite $H_1: \theta > \theta_0$

all these are alternative

θ_0

We are looking for UMP α-test φ^* which means: if we take any competitor $\varphi \in \Phi_\alpha = \{$set of all tests φ such that $E_{\theta_0}\varphi \le \alpha\}$ then we claim that $E_\theta \varphi^* \ge E_\theta \varphi$ for all $\theta > \theta_0$

We first simplify the problem by considering testing a simple $H_0: \theta = \theta_0$ versus simple $H_1: \theta = \theta_1$ for a fixed $\theta_1 > \theta_0$. Because this is a Neyman-Pearson lemma-type problem, for it we have the most powerful α test and it is given by

$$\varphi^* = \begin{cases} 1 & \text{if } L(\mathbb{X}, \theta_1)/L(\mathbb{X}, \theta_0) > C \\ 0 & \text{if } L(\mathbb{X}, \theta_1)/L(\mathbb{X}, \theta_0) \le C \end{cases}$$

Notice that

$$\frac{L(X, \theta_1)}{L(X, \theta_0)} = \exp\left((\theta_1 - \theta_0)\sum_{i=1}^{n} X_i + \frac{n}{2}(\theta_0^2 - \theta_1^2)\right)$$

Since $\theta_1 - \theta_0 > 0$, $\frac{L(X, \theta_1)}{L(X, \theta_0)}$ is <u>monotonically increasing</u>

in $T = \sum_{i=1}^{n} X_i$ and $\frac{L(X, \theta_1)}{L(X, \theta_0)} > C$ is equivalent to

$\sum_{i=1}^{n} X_i > C_1$ or, by renaming constants, to $\bar{X} > \bar{\bar{C}}$.

To find $\bar{\bar{C}}$, we must "exhaust the given level α" which means $E_{\theta_0}\varphi^* = 1 * P_{\theta_0}(\bar{X} > \bar{\bar{C}}) = \alpha$ must hold (see the statement of the NP Lemma).

But $E_{\theta_0}\varphi^* = P_{\theta_0}(\bar{X} > \bar{\bar{C}}) = P_{\theta_0}\left(\frac{\sqrt{n}(\bar{X} - \theta_0)}{1} > \frac{\sqrt{n}(\bar{\bar{C}} - \theta_0)}{1}\right)$

$= P\left(Z > \frac{\sqrt{n}(\bar{\bar{C}} - \theta_0)}{1}\right) = \alpha$ where $Z \sim N(0,1)$.

This implies that $\frac{\sqrt{n}(\bar{\bar{C}} - \theta_0)}{1} = z_\alpha$ must hold where z_α is the upper $\alpha * 100\%$ point of the $N(0,1)$.

Then $\bar{\bar{C}} = \theta_0 + \frac{z_\alpha}{\sqrt{n}}$ and φ^* becomes:

$$\varphi^*(X) = \begin{cases} 1 & \text{if } \bar{X} > \theta_0 + \frac{z_\alpha}{\sqrt{n}} \\ 0 & \text{if } \bar{X} \le \theta_0 + \frac{z_\alpha}{\sqrt{n}} \end{cases}$$

<u>NOW WE NOTICE</u> that the resulting $\varphi^*(X)$ above, although having been constructed for a particular $H_1: \theta = \theta_1$, <u>DOES NOT</u> involve this $\theta_1 > \theta_0$ in **its** shape. Hence the <u>SAME test</u> φ^* will be the most powerful α-test for <u>any</u> chosen $\theta_1 > \theta_0$! Therefore, $\varphi^*(X)$ will be the <u>UNIFORMLY</u> most powerful for testing also $H_0: \theta = \theta_0$ versus $H_1: \theta > \theta_0$.

Notice that we used the monotonicity of the Likelihood ratio in our argument.

This example was generalized in the Blackwell-Girshick (BG) Theorem (p.52 of the notes).

I also gave an example of applying the BG theorem to derive UMP α tests.

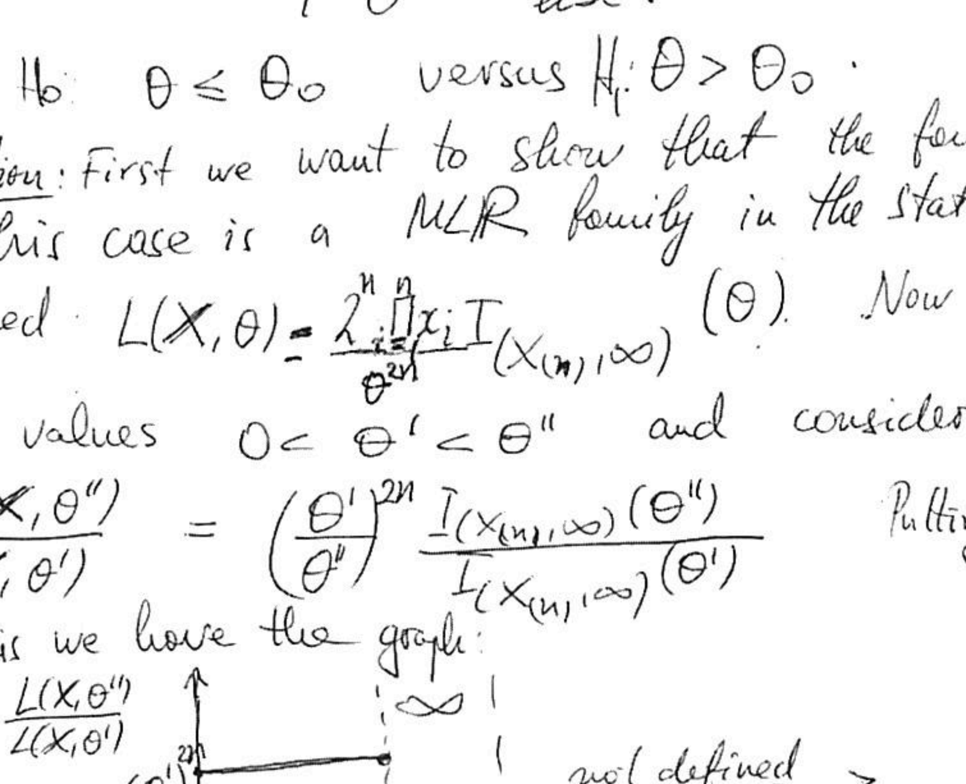
Example: Assume that $\mathbb{X} = (X_1, X_2, \dots, X_n)$ are i.i.d. from $f(x,\theta) = \begin{cases} 2x/\theta^2, & 0 < x < \theta \\ 0 & \text{else}. \end{cases}$ Construct a UMP α-test of $H_0: \theta \le \theta_0$ versus $H_1: \theta > \theta_0$.

Solution: First we want to show that the family $L(X,\theta)$ in this case is a MLR family in the statistic $T = X_{(n)}$.

Indeed $L(X,\theta) = \dfrac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} I_{(X_{(n)},\infty)}(\theta)$. Now take two values $0 < \theta' < \theta''$ and consider

$$\frac{L(X,\theta'')}{L(X,\theta')} = \left(\frac{\theta'}{\theta''}\right)^{2n} \frac{I_{(X_{(n)},\infty)}(\theta'')}{I_{(X_{(n)},\infty)}(\theta')}$$

Putting $X_{(n)}$ on the OX axis we have the graph:



Hence we have MLR property in $T = X_{(n)}$. Then BG theorem tells us that UMP α test of H_0 vs H_1 exists and is given by

$$\varphi^* = \begin{cases} 1 & \text{if } X_{(n)} > K \\ 0 & \text{if } X_{(n)} \le K. \end{cases}$$

To find K we need to exhaust the level, i.e. must satisfy $E_{\theta_0}\varphi^* = \alpha$. However $E_{\theta_0}\varphi^* = P_{\theta_0}(X_{(n)} > K) = 1 - P_{\theta_0}(X_{(n)} \le K)$
$= 1 - [P_{\theta_0}(X_1 \le K)]^n = 1 - \left(\frac{K}{\theta_0}\right)^{2n} = \alpha \Rightarrow \boxed{K = \theta_0 (1-\alpha)^{\frac{1}{2n}}}$ and φ^* is completely determined.

Some of my white board writing in weeks 8&9 continued

I finished the discussion of the example:

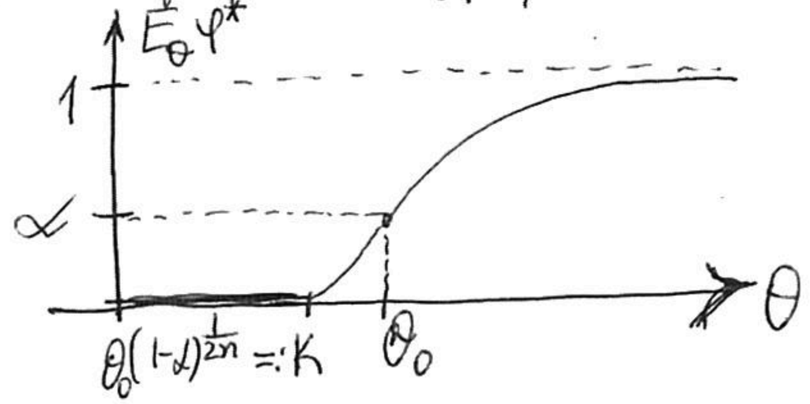
$\mathbf{X}=(X_1,X_2,\ldots,X_n)$ are i.i.d from $f(x,\theta)=\begin{cases}2x/\theta^2, & 0<x<\theta\\ 0 & \text{else}\end{cases}$

We constructed the UMP α-test of $H_0:\theta\le\theta_0$ vs $H_1:\theta>\theta_0$. We ended up with

$$\varphi^*=\begin{cases}1 & \text{if } X_{(n)}>K=\theta_0(1-\alpha)^{\frac{1}{2n}}\\ 0 & \text{if } X_{(n)}\le K=\theta_0(1-\alpha)^{\frac{1}{2n}}\end{cases}$$

I discussed the graph of the resulting power function: Since $E_\theta\varphi^*=\begin{cases}1-\left(\frac{\theta_0}{\theta}\right)^{2n}(1-\alpha), & 0<K<\theta\\ 0 & K\ge\theta\end{cases}$

we can graph this function $E_\theta\varphi^*,\ \theta>0$ and the graph looks like:



Next, I discussed the construction of UMP α test in the case of DISCRETE data where randomization was necessary.

Problem: $n=25$ i.i.d. observations from Bernoulli (θ).
$\alpha=0.01$ is chosen for level of significance.
Construct the UMP α test of $H_0:\theta\le .15=\theta_0$ vs $H_1:\theta>.15$

Solution:
First of all, UMP α test exists because of the Blackwell-Girshick theorem since,

$f(x,\theta)=\theta^x(1-\theta)^{1-x}=(1-\theta)\cdot e^{x\ln\left(\frac{\theta}{1-\theta}\right)}$ is an one-parameter exponential family with

$c(\theta) = \ln\left(\frac{\theta}{1-\theta}\right)$ monotone increasing in θ, and $d(x) = x$.
Hence we have the MLR property in the statistic $T = \sum_{i=1}^{n} X_i$ and Blackwell-Girshick's theorem tells us that UMP α test exists and has the structure

$$\varphi^* = \begin{cases} 1 & \text{if } T > K \\ \gamma & \text{if } T = K \\ 0 & \text{if } T < K \end{cases}$$

$\left(T = \sum_{i=1}^{25} X_i \text{ here}\right)$

To determine K and γ we have to find the smallest natural number for which still $P_{\theta_0}(T > K) < \alpha$.
Since $T \sim Bin(25, .15)$ under the borderline $\theta_0 = .15$ value, we can find any of the probabilities $P_{\theta_0}(T = t) = \binom{25}{t}(.15)^t(.85)^{25-t}$, $t = 0, 1, 2, \ldots, 25$ and tabulate the whole cdf (or "ask" the computer to do it for us)

An extract from the distribution of T follows:

x	...	7	8	9	...
$P(T \le x)$		.974532	.99207	.99786	

We need $P(T > x) < 0.01$ and x should be the smallest with this property $\Rightarrow K = x = 8$. Then

$$\gamma = \frac{\alpha - P_{\theta_0}(S)}{P_{\theta_0}(R)} = \frac{0.01 - (1 - .99207)}{.99207 - .9745} = .114$$

Hence the UMP 0.01-test is completely determined:

$$\varphi^* = \begin{cases} 1 & \text{if } \sum_{i=1}^{25} X_i > 8 \\ .114 & \text{if } \sum_{i=1}^{25} X_i = 8 \\ 0 & \text{if } \sum_{i=1}^{25} X_i < 8 \end{cases}$$

I then discussed Q 2 from tutorial sheet 3 but because a complete solution to it is given on moodle I will not reproduce it here (see it there)

I then moved over to discuss the notion of an uniformly most powerful <u>unbiased</u> (UMPU) α-test. This discussion is well and thoroughly represented in the notes (p. 53–54 there) and I am not reproducing it here again.

I also went through the two examples in Section 6.8 (p. 54-55). Both are related to constructing UMPU α tests of $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$.

- the first example treats the case of sample of size $n = 1$ from the exponential distribution
- the second example treats the testing of the same hypothesis for a sample of size n from $N(\theta, 1)$ distribution. For this second problem, it is known that statistical folklore advises to use the so-called "Z-test", i.e.

$$\varphi^*(\mathbf{X}) = \begin{cases} 1 & \text{if } \sqrt{n}\,|\bar{X} - \theta_0| \geq Z_{\frac{\alpha}{2}} \\ 0 & \text{if } \sqrt{n}\,|\bar{X} - \theta_0| < Z_{\frac{\alpha}{2}} \end{cases}$$

I explained <u>why</u> this test $\varphi^*(\mathbf{X})$ was in fact the UMPU α test for the testing problem $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$

Pages 54-55 contain the detailed argument and I abstain from reproducing it here again.