Some of my white board writing in week 9

1) I finished the discussion of the example from last week: $\mathbf{X}=(X_1,X_2,\dots,X_n)$ are i.i.d from $f(x,\theta)=\begin{cases} 2x/\theta^2, & 0<x<\theta \\ 0 & \text{else}\end{cases}$

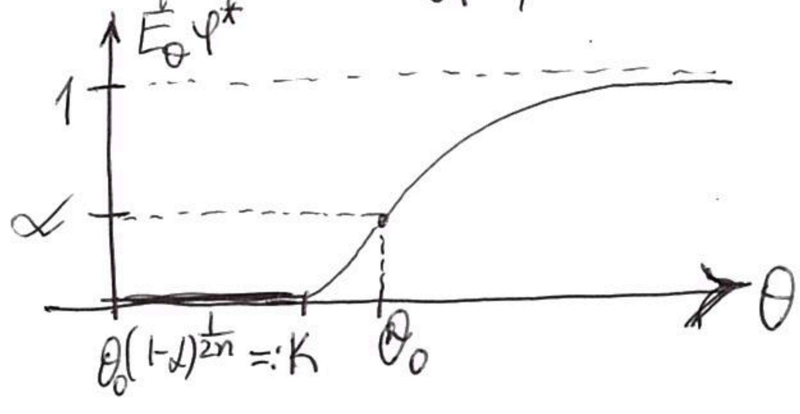
Last week we constructed the UMP α-test of $H_0: \theta\le\theta_0$ vs $H_1: \theta>\theta_0$. We ended up with

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

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:



2.) 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 $T \sim \text{Bin}(25, .15)$ under the borderline $\theta_0 = .15$ value, we can find any of the probabilities $P_{\theta_0}(T = t) = \binom{25}{t}(.15)^t(.85)^{25-t}$, $t = 0, 1, 2, \dots, 25$ 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:

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

3) I then discussed Q2 from tutorial sheet 3 but because a <u>complete solution to it</u> is given on moodle I will not reproduce it here (see it <u>there</u>)

4) I also discussed the generalized likelihood ratio test for two examples related to the normal distribution:

a) Testing $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ for a sample of n i.i.d. $N(\mu, \sigma^2)$ (σ^2 assumed known)

In this case

$$-2\left[\ln L(X, H_0) - \ln L(X, \bar{x})\right] = -2\left[-\frac{\sum_{i=1}^{n}(X_i-\mu_0)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{2\sigma^2}\right] =$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\bar{X})^2\right] = \frac{n(\bar{X}-\mu_0)^2}{\sigma^2}$$

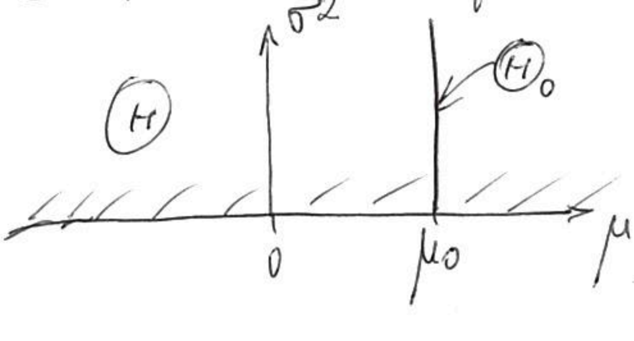
This should be $\sim \chi^2_1$ asymptotically but in this case, because of dealing with normal, the result is precise (not only asymptotic). Indeed, we know that under H_0: $\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n}) \Rightarrow \sqrt{n}\frac{(\bar{X}-\mu_0)}{\sigma} \sim N(0,1) \Rightarrow \frac{n(\bar{X}-\mu_0)^2}{\sigma^2} \sim \chi^2_1$

The GLRT is: $\varphi^* = \begin{cases} 1 & \text{if } \frac{n(\bar{X}-\mu_0)^2}{\sigma^2} > \chi^2_{\alpha,1} \\ 0 & \text{if } \frac{n(\bar{X}-\mu_0)^2}{\sigma^2} \leq \chi^2_{\alpha,1} \end{cases}$ and

is equivalent to the standard Z-test in this case.

b) Testing $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ again but when σ^2 is unknown. Hence we are testing in effect:

$H_0: \begin{cases} \mu = \mu_0 \\ \sigma^2 > 0 \end{cases}$ vs $H_1: \begin{cases} \mu \neq \mu_0 \\ \sigma^2 > 0. \end{cases}$ In terms of the notation of GLRT: we have a parameter vector $\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$ and



H_0 is one-dimensional subspace as sketched, whereas H is "anything" above the Ox axis.

The dimensions K, r, s, as discussed in Section 6.10.1 of Lecture 6, p. 56, are $K=2$, $r=1$, $s=1$.
To perform the GLRT we need to maximize $L(\mathbf{X}, \theta)$ under the Hypothesis and under the alternative.
i) under the hypothesis: $\mu=\mu_0$, so we need to optimize w.r. to σ only. $\ln L = -\frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu_0)^2 + \text{const}$

$$\frac{\partial}{\partial\sigma^2}\ln L = 0 = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (X_i-\mu_0)^2 \quad \text{implies}$$

$$\hat{\sigma}^2_{H_0} = \frac{1}{n}\sum_{i=1}^n (X_i-\mu_0)^2$$

and $$\sup L|H_0 = \frac{1}{(\sqrt{2\pi})^n (\hat{\sigma}^2_{H_0})^{n/2}} \exp\left(-\frac{n}{2}\right)$$

ii) without the restriction of H_0, we have to maximise $\ln L$ w.r. to both μ and σ^2, i.e. solve the system

$$\begin{cases} \frac{\partial}{\partial\mu}\ln L = 0 \\ \frac{\partial}{\partial\sigma^2}\ln L = 0 \end{cases}$$

This leads to $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2$

When plugging-in, we get the $\sup L$ without any restriction and it is $$\sup L = \frac{1}{(\sqrt{2\pi})^n (\hat{\sigma}^2)^{\frac{n}{2}}} \exp\left(-\frac{n}{2}\right)$$

Hence $$-2\log\Lambda = -2\log\left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2_{H_0}}\right)^{\frac{n}{2}} = n\log\left(\frac{\hat{\sigma}^2_{H_0}}{\hat{\sigma}^2}\right)$$ and the GLRT is

$$\varphi^* = \begin{cases} 1 & \text{if } n\log\left(\frac{\hat{\sigma}^2_{H_0}}{\hat{\sigma}^2}\right) > \chi^2_{\alpha,1} \\ 0 & \text{if } n\log\left(\frac{\hat{\sigma}^2_{H_0}}{\hat{\sigma}^2}\right) \le \chi^2_{\alpha,1} \end{cases}$$

(the degrees of freedom are $=1$ since in this case $r=1=K-s$ $(K=2, s=1)$).
Note that now the convergence of $-2\log\Lambda$ to the limiting χ^2_1 is only <u>asymptotic</u> (not precise as in case a)).
But $-2\log\Lambda = n\log\left(1+\frac{(\bar{X}-\mu_0)^2}{\hat{\sigma}^2}\right) \approx \frac{n(\bar{X}-\mu_0)^2}{S^2}$, so it is "almost" equivalent to the standard t-test for $\mu=\mu_0$ when σ^2 is unknown.