Generalized Likelihood Ratio Test Example

a.k.a. Too much for in class… but certainly worth making sure you can do each step!

Consider testing H_0: $\mu \le \mu_0$ vs. H_A: $\mu > \mu_0$ for an random sample form a population that is normally distributed (where σ^2 is unknown).

One way to do this is to construct the likelihood ratio test where $P(\Lambda \le \lambda | H_0 \text{ is true}) = \alpha$. Using the definition of the likelihood ratio test on page 308, and plugging in the normal p.d.f., we get:

$$\Lambda = \frac{\max\limits_{H_0} f(x_1,\ldots x_n \mid \mu,\sigma)}{\max\limits_{H_0 \cup H_A} f(x_1,\ldots x_n \mid \mu,\sigma)} = \frac{\max\limits_{H_0} \prod_{i=1}^{n} f(x_i \mid \mu,\sigma)}{\max\limits_{H_0 \cup H_A} \prod_{i=1}^{n} f(x_i \mid \mu,\sigma)} = \frac{\max\limits_{H_0} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}{\max\limits_{H_0 \cup H_A} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}$$

The next step is to find what the maximum is in both the numerator and denominator. The denominator is the easy place to start because it is over the set of all possible μ and σ values. In other words we just need to find the mle. For the numerator we have to deal with the restriction on μ.

<u>Denominator:</u> To find the MLE we need to take the derivative with respect to both of the parameters, so…

$$lik = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$\log lik = \sum_{i=1}^{n} \log\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)\right) = \sum_{i=1}^{n}\left[\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$= -n\log\sigma - \frac{n}{2}\log(2\pi) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial}{\partial\mu} = -\sum_{i=1}^{n} 2\frac{(x_i-\mu)}{2\sigma^2}(-1) = \sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2} = 0 \Rightarrow \sum_{i=1}^{n} x_i = n\mu \Rightarrow \hat{\mu}_{mle} = \bar{x}$$

$$\frac{\partial}{\partial\sigma} = -\frac{n}{\sigma} - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^3}(-2) = -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^3} \Rightarrow \hat{\sigma}_{mle} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\hat{\mu}_{mle})^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}$$

Numerator: Notice in the numerator we need to find the maximum… subject to the null hypothesis being true. That is, since μ must be $\le\mu_0$, $\hat{\mu}_{H_0 mle}$ must be $\le\mu_0$.

So we do the math just like we did for the denominator… and get that $\bar{x}$ is the H_0mle as long as it is $\le\mu_0$. If it is > then the maximum must happen at the boundary, and be either μ_0 or $-\infty$ (which would mean the maximum would never be reached). Notice what happens if we plug in $-\infty$ for μ; the likelihood is zero. Therefore the H_0mle is μ_0 if $\bar{x} > \mu_0$. To get the H_0mle for σ we simply follow the math above and plug in the value of H_0mle for μ when needed. This gives:

If $\bar{x} \le \mu_0$, $\hat{\mu}_{H_0 mle} = \bar{x}$ and $\hat{\sigma}_{H_0 mle} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$

If $\bar{x} > \mu_0$, $\hat{\mu}_{H_0 mle} = \mu_0$ and $\hat{\sigma}_{H_0 mle} = \dfrac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{n}$

Plugging the values in: First notice that $\Lambda=1$ (and we accept) if $\bar{x} \le \mu_0$ because the maximums plugged into the numerator and denominator are the same. The tricky case then is if $\bar{x} > \mu_0$. So lets carry through the product and just plug the values in:

$$\Lambda = \frac{\max\limits_{H_0} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}{\max\limits_{H_0 \cup H_A} \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)} = \frac{\max\limits_{H_0}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right)}{\max\limits_{H_0 \cup H_A}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right)}$$

$$= \frac{\left(\frac{1}{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{n}}\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{2\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{n}}\right)}{\left(\frac{1}{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{2\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}\right)} = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2}\right)^{\frac{n}{2}} \frac{\exp\left(\frac{-n}{2}\right)}{\exp\left(\frac{-n}{2}\right)} = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2}\right)^{\frac{n}{2}}$$

Now we need to find the λ so that

$$P\left(\Lambda = \left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \mu_0)^2}\right)^{\frac{n}{2}} \le \lambda\right) = \alpha$$

This doesn't look friendly at all… but if we remember that the x's are normal we might be able to change this into a distribution we can work with (t? χ^2? F?). Notice that the numerator looks very close to the standard deviation… what you would find in the denominator of a t-test. Also notice that the denominator is at least suggestive of the $\bar{x} - \mu_0$ you would find in the numerator of the t-test statistic. Since this is are only lead, lets see if we can work it out…

$$\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \mu_0)^2}\right)^{\frac{n}{2}} \le \lambda \quad \Rightarrow \quad \left(\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{\frac{n}{2}} \ge \lambda' \quad \Rightarrow \quad \frac{\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda''$$

$$\Rightarrow \quad \frac{\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda'' \quad \Rightarrow \quad \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + 2\sum_{i=1}^{n}(x_i - \bar{x})(\bar{x} - \mu_0)^2 + \sum_{i=1}^{n}(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda''$$

Note that $\bar{x}$ and μ_0 are constant with respect to n, so that the second term is zero, and the sum turns into a multiplier of n in the third term:

$$' \quad \Rightarrow \quad \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda'' \quad \Rightarrow \quad 1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda''$$

$$\Rightarrow \quad \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \ge \lambda''' \quad \Rightarrow \quad \frac{\sqrt{n}(\bar{x} - \mu_0)^2}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \ge \lambda^{1V} \quad \Rightarrow \quad \frac{(\bar{x} - \mu_0)}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}} \ge \lambda^{1V}$$

$$\Rightarrow \quad \frac{(\bar{x} - \mu_0)}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 / n-1}{n}}} \ge \lambda^{V} \quad \Rightarrow \quad \frac{(\bar{x} - \mu_0)}{s/\sqrt{n}} \ge \lambda^{V}$$

This is just the classic t-statistic!! (And recall that if $\bar{x} \le \mu_0$ then we would accept.)

<u>Finally!</u> So the test for H_0: $\mu \le \mu_0$ vs. H_A: $\mu > \mu_0$ is reject if

$$\frac{(\bar{x}-\mu_0)}{s/\sqrt{n}} \ge t_{1-\frac{\alpha}{2}}$$

<u>Question:</u> How would the above math change if we had used H_0: $\mu=\mu_0$ vs. H_A: $\mu \ne \mu_0$?

<u>What about Theorem A?</u> Notice that theorem A on page 310 says that $-2\log\Lambda$ should be chi-squared with degrees of freedom equal to $\dim\Omega - \dim\omega_o$. Does that work here?

Notice that it probably shouldn't for two different reasons!!!!

Reason 1: Ω = possible values of μ and σ is two-dimensional… but so is ω_o because it is half of a plane (a line is one-dimensional). So the degrees of freedom should be 2-2=0 which makes no sense!!

Reason 2: If $\bar{x} \le \mu_0$ we get that $\Lambda=1$ so that , so that $-2\log\Lambda = 0$. There is thus some probability that $-2\log\Lambda = 0$. But a chi-squared distribution is continuous and so has to have 0 probability of taking any given value!

We get something called a chi-bar distribution that is a mixture of a discrete distribution at zero and part of a chi-squared distribution. Yuck!

However, if we had used the hypotheses H_0: $\mu=\mu_0$ vs. H_A: $\mu \ne \mu_0$ it would have worked out just fine. In this case Ω is two-dimensional, but ω_o has a fixed value of μ so that it is only one-dimensional. If everything works out we should get that $-2\log\Lambda$ will be chi-squared with one degree of freedom.

$$\Lambda = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2} \right)^{\frac{n}{2}}$$

So, using some of the math we saw earlier…

$$-2\log\Lambda = -2\log\left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2} \right)^{\frac{n}{2}} = \log\left(\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \right)^{n} = \log\left(1+\frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \right)^{n}$$

$$= \log\left(1+\left(\frac{(\bar{x}-\mu_0)^2}{\left(\frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}{\frac{n}{n-1}}\right)}\right)\right)^n = \log\left(1+\left(\frac{\frac{(\bar{x}-\mu_0)^2}{s^2/n}}{n-1}\right)\right)^n = \log\left(1+\frac{\left(\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)^2}{n-1}\right)^n = \log\left(1+\frac{t^2}{n-1}\right)^n$$

where t has a t distribution with n-1 degrees of freedom.

As n goes to infinity the t goes to a standard normal and (by one of the definitions of the exponential function from calculus) we get that the distribution of –2logΛ is $\chi \log(\exp(z^2)) = z^2 = \chi^2_{df=1}$.