Statistical Inference

Lecture 10b

ANU - RSFAS

Last Updated: Wed May 10 14:21:43 2017

Neyman-Pearson Set-up

- Consider simple hypotheses - those which consist of only a single parameter value.
- We will examine the case of a statistical test for which both the null and alternative hypotheses are simple.
- Suppose that $X_1, \ldots, X_n$ are a sample from a population characterized by a probability model with density function $f(x|\theta)$ for $\theta \in \Theta$ where $\Theta = \{\theta_0, \theta_1\}$.
- We shall focus on:

$$\begin{aligned} H_0 : \quad & \theta = \theta_0 \\ H_1 : \quad & \theta = \theta_1 \end{aligned}$$

Neyman-Pearson Lemma

- Consider the likelihood-ratio:

$$\lambda(\boldsymbol{x}) = \frac{L(\theta_0|\boldsymbol{x})}{L(\theta_1|\boldsymbol{x})}$$

- The test we shall define has a critical region of the form

$$R = \{\lambda(\boldsymbol{x}) \leq k\}$$

- The ratio of the likelihood for any given sample at each of the two possible parameter values is precisely a relative measure of how plausible the two hypotheses are.
- In other words, when $\lambda(\boldsymbol{x})$ is very small, this is strong evidence that the observations arose from the alternative hypothesis rather than the null hypothesis.

Neyman-Pearson Set-up

- It should seem intuitively reasonable that the likelihood ratio is a good method of distinguishing between samples which support the null hypothesis versus samples which support the alternative hypothesis.
- From what we have done, we know for a given α we could compare the power $\beta(\theta)$.
- We would like to find a uniformily most powerful test . . .

$$\beta(\theta) \geq \beta(\theta^*)$$

- It turns out that N-P tests lead to UMP tests.

Example Suppose that $X_1, \ldots, X_n$ are a random sample from a normal distribution with mean μ and unit variance. Further, suppose that we know $\mu \in \{0, 1\}$. We wish to test:

$$\begin{aligned} H_0 : & \quad \mu = 0 \\ H_1 : & \quad \mu = 1 \end{aligned}$$

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{exp\left(-\frac{1}{2}\sum_{i=1}^{n} X_i^2\right)}{exp\left(-\frac{1}{2}\sum_{i=1}^{n}(X_i-1)^2\right)} \\
&= exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left[X_i^2-(X_i-1)^2\right]\right) \\
&= exp\left(\frac{n}{2}-\sum_{i=1}^{n} X_i\right)
\end{aligned}
$$

- So we get the rejection region:

$$
\begin{aligned}
R &= \left\{ exp\left(\frac{n}{2} - \sum_{i=1}^{n} X_i\right) \leq k \right\} \\
&= \left\{ \frac{n}{2} - \sum_{i=1}^{n} X_i \leq log(k) \right\} \\
&= \left\{ -\sum_{i=1}^{n} X_i \leq log(k) - \frac{n}{2} \right\} \\
&= \left\{ \sum_{i=1}^{n} X_i \geq -log(k) + \frac{n}{2} \right\} \\
&= \left\{ \bar{X} \geq -log(k)/n + \frac{1}{2} \right\} \\
&= \left\{ \bar{X} \geq c^* \right\}
\end{aligned}
$$

$$
\begin{aligned}
P_{H_0}(R) &= P_{H_0}\left(\bar{X} \geq c^*\right) = \alpha \\
&= P_{H_0}\left(\frac{\bar{X} - 0}{1/\sqrt{n}} \geq c^{**}\right) = \alpha \\
&= P_{H_0}\left(Z \geq c^{**}\right) = \alpha
\end{aligned}
$$

- If $\alpha = 0.05$ then c^{**} is 1.644854.

```
qnorm(0.95)
```

```
## [1] 1.644854
```

- This is a UMP test!

Example Suppose that $X_1, \ldots, X_{10}$ are a random sample from a Bernoulli distribution with parameter θ. Further, suppose that we wish to test:

$$\begin{aligned} H_0 : &\quad \theta = 0.5 \\ H_1 : &\quad \theta = 0.2 \end{aligned}$$

- Let's get the likelihood:

$$L(\theta|\boldsymbol{x}) = \theta^{\sum x_i}(1-\theta)^{10-\sum x_i} = \theta^{10\bar{x}}(1-\theta)^{10-10\bar{x}}$$

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{0.5^{10\bar{x}}(1-0.5)^{10-10\bar{x}}}{0.2^{10\bar{x}}(1-0.2)^{10-10\bar{x}}} \\
&= \left(\frac{5}{8}\right)^{10} 4^{10\bar{x}}
\end{aligned}
$$

- So we get the rejection region:

$$
\begin{aligned}
R &= \left\{\left(\frac{5}{8}\right)^{10} 4^{10\bar{x}} \leq k\right\} \\
&= \left\{10\bar{x} \leq log_4 \left[\left(\frac{8}{5}\right)^{10} k\right]\right\} \\
&= \{10\bar{x} \leq c^*\}
\end{aligned}
$$

- Let's get a UMP test for $\alpha = 0.01$.

$$
\begin{aligned}
P_{H_0}(R) &= P(10\bar{X} \leq c^*) = 0.01 \\
&= P\left(\sum_{i=1}^{n} X_i \leq c^*\right) = 0.01
\end{aligned}
$$

- Recall that under H_0: $\sum_{i=1}^{n} X_i \sim \mathrm{binomial}(n = 10, p = 0.5)$.
- Due to the discreteness, we can't find a c^* such that we achieve $\alpha = 0.01$.

```r
qbinom(0.01, 10, 0.5)
```

```
## [1] 1
```

- The closest we can find is $c^* = 1$.

```r
pbinom(1, 10, 0.5)
```

```
## [1] 0.01074219
```

- So we have a UMP test of size $\alpha = 0.01074$, which is close to $\alpha = 0.01$.

$$P_{H_0}(R) = P(10\bar{X} \leq 1) = 0.01074$$

Neyman-Pearson Lemma

Rice Section 9.2:

- Suppose that H_0 and H_1 are simple hypotheses and that the test that rejects H_0 whenever the likelihood ratio is less than k has significance level α.
- Then any other test for which the significance level is less than or equal to α has power less than or equal to that of the likelihood ratio test.

Neyman-Pearson Lemma

Proof:

- Consider any other test of size $\alpha^* \leq \alpha$.
- We need to show that $P_{\theta_1}(R) \geq P_{\theta_1}(R^*)$.
- Consider:

$$
\begin{aligned}
P_{\theta_1}(R) &= P_{\theta_1}(R \cap R^*) + P_{\theta_1}(R \cap R^{*c}) \\
P_{\theta_1}(R^*) &= P_{\theta_1}(R^* \cap R) + P_{\theta_1}(R^* \cap R^c)
\end{aligned}
$$

Neyman-Pearson Lemma

- As $P_{\theta_1}(R \cap R^*) = P_{\theta_1}(R^* \cap R)$, substitute into $P_{\theta_1}(R)$:

$$
\begin{aligned}
P_{\theta_1}(R) &= \{P_{\theta_1}(R^*) - P_{\theta_1}(R^* \cap R^c)\} + P_{\theta_1}(R \cap R^{*c}) \\
&= P_{\theta_1}(R^*) + \{P_{\theta_1}(R \cap R^{*c}) - P_{\theta_1}(R^* \cap R^c)\} \\
P_{\theta_1}(R) - P_{\theta_1}(R^*) &= P_{\theta_1}(R \cap R^{*c}) - P_{\theta_1}(R^* \cap R^c)
\end{aligned}
$$

- So we need to show: $P_{\theta_1}(R \cap R^{*c}) - P_{\theta_1}(R^* \cap R^c) \geq 0$.

Neyman-Pearson Lemma

- Note that for event $E \subseteq R$ we have:

$$
\begin{aligned}
P_{\theta_1}(E) &= \int_E L(\theta_1|\boldsymbol{x})d\boldsymbol{x} \\
&= \int_E L(\theta_1|\boldsymbol{x})\frac{L(\theta_0|\boldsymbol{x})}{L(\theta_0|\boldsymbol{x})}d\boldsymbol{x} \\
&= \int_E \frac{1}{\lambda(\boldsymbol{x})}L(\theta_0|\boldsymbol{x})d\boldsymbol{x}
\end{aligned}
$$

Neyman-Pearson Lemma

- By the definition of R, we have $\lambda(\boldsymbol{x}) \leq k$ for any $\boldsymbol{x} \in E \subseteq R$:

$$
\begin{aligned}
P_{\theta_1}(E) &= \int_E \frac{1}{\lambda(\boldsymbol{x})} L(\theta_0|\boldsymbol{x}) d\boldsymbol{x} \\
&\geq \frac{1}{k} \int_E L(\theta_0|\boldsymbol{x}) d\boldsymbol{x} \\
&= \frac{1}{k} Pr_{\theta_0}(E)
\end{aligned}
$$

- A similar argument shows that for any event $F \subseteq R^c$, then $P_{\theta_1}(F) \leq \frac{1}{k} P_{\theta_0}(F)$.

Neyman-Pearson Lemma

- Now let $E = [R \cap R^{*c}] \subseteq R$ and $F = [R^* \cap R^c] \subseteq R^c$, we have

$$
\begin{aligned}
P_{\theta_1}(R \cap R^{*c}) - P_{\theta_1}(R^* \cap R^c) \quad \geq \quad & \frac{1}{k} P_{\theta_0}(R \cap R^{*c}) - \frac{1}{k} P_{\theta_0}(R^* \cap R^c) \\
&= \frac{1}{k} \left[P_{\theta_0}(R \cap R^{*c}) - P_{\theta_0}(R^* \cap R^c) \right] \\
&= \frac{1}{k} [P_{\theta_0}(R \cap R^{*c}) + P_{\theta_0}(R \cap R^*) \\
&- P_{\theta_0}(R^* \cap R) - P_{\theta_0}(R^* \cap R^c)] \\
&= \frac{1}{k} \left[P_{\theta_0}(R) - P_{\theta_0}(R^*) \right] \\
&= \frac{1}{k} (\alpha - \alpha^*) \geq 0
\end{aligned}
$$

Neyman-Pearson Lemma

- On the surface, it seems the N-P Lemma is too simple to be of any real use. Can we push the result a bit? Let's consider the example.

Example: Suppose that $X_1, \ldots, X_n$ are a random sample from a normal distribution with mean μ and unit variance. Consider testing:

$$\begin{aligned} H_0 : \quad & \mu = \mu_0 = 0 \\ H_1 : \quad & \mu = \mu_1 \end{aligned}$$

Where $\mu_1 > 0 = \mu_0$.

Neyman-Pearson Lemma

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{exp\left(-\frac{1}{2}\sum_{i=1}^{n} X_i^2\right)}{exp\left(-\frac{1}{2}\sum_{i=1}^{n}(X_i - \mu_1)^2\right)} \\
&= exp\left(\frac{n\mu_1^2}{2} - n\mu_1\bar{X}\right)
\end{aligned}
$$

Neyman-Pearson Lemma

- So we get the rejection region:

$$
\begin{aligned}
R &= \left\{ exp\left(\frac{n\mu_1^2}{2} - n\mu_1 \bar{X} \right) \le k \right\} \\
&= \left\{ \left(\frac{n\mu_1^2}{2} - n\mu_1 \bar{X} \right) \le log(k) \right\} \\
&= \left\{ \bar{X} > \frac{\mu_1}{2} - \frac{1}{n\mu_1} log(k) \right\} \\
&= \left\{ \bar{X} > c^* \right\} \\
&= \left\{ \frac{\bar{X} - 0}{1/\sqrt{n}} \ge c^{**} \right\} = \{ Z \ge c^{**} \}
\end{aligned}
$$

- We assumed that $\mu_1 > 0$, so we get the sign switch.
- If $\alpha = 0.05$ then $c^{**} = 1.64$.

Neyman-Pearson Lemma

- The UMP test has the same rejection region as our previous example: $H_0 : \mu = 0$ vs $H_1 : \mu = 1$.
- This test is actually UMP for $H_0 : \mu = 0$ vs $H_1 : \mu > 0$.
- It can also be shown that the test is UMP for $H_0 : \mu \leq 0$ vs $H_1 : \mu > 0$.

Neyman-Pearson Lemma

- What if we wanted to test: $H_0 : \mu = 0$ vs $H_1 : \mu < 0$?
- We get a UMP test with rejection region:
- So we get the rejection region:

$$
\begin{aligned}
R &= \left\{ exp\left(\frac{n\mu_1^2}{2} - n\mu_1\bar{X}\right) \leq k \right\} \\
&= \left\{ \left(\frac{n\mu_1^2}{2} - n\mu_1\bar{X}\right) \leq log(k) \right\} \\
&= \left\{ \bar{X} \leq \frac{\mu_1}{2} - \frac{1}{n\mu_1} log(k) \right\} \\
&= \left\{ \bar{X} \leq c^* \right\} \\
&= \left\{ \frac{\bar{X} - 0}{1/\sqrt{n}} \leq c^{**} \right\} = \{Z \leq c^{**}\}
\end{aligned}
$$

Neyman-Pearson Lemma

* How does that compare to a Maximum Likleihood Ratio Test (Generalized Likelihood Ratio Test) [an extension we will discuss shortly]? For:

$$\begin{aligned} H_0 : \quad & \mu = \mu_0 \\ H_1 : \quad & \mu \neq \mu_0 \end{aligned}$$

- Let's have $\mu_0 = 0$. We will show the rejection region is:

$$\left\{ |Z| > \sqrt{n}\sqrt{[-2log(c)]/n} \right\} = \{|Z| > c^*\}$$

- So we will reject H_0 if:

$$\left\{ \left| \frac{(\bar{x} - 0)}{1/\sqrt{n}} \right| > 1.96 \right\}$$

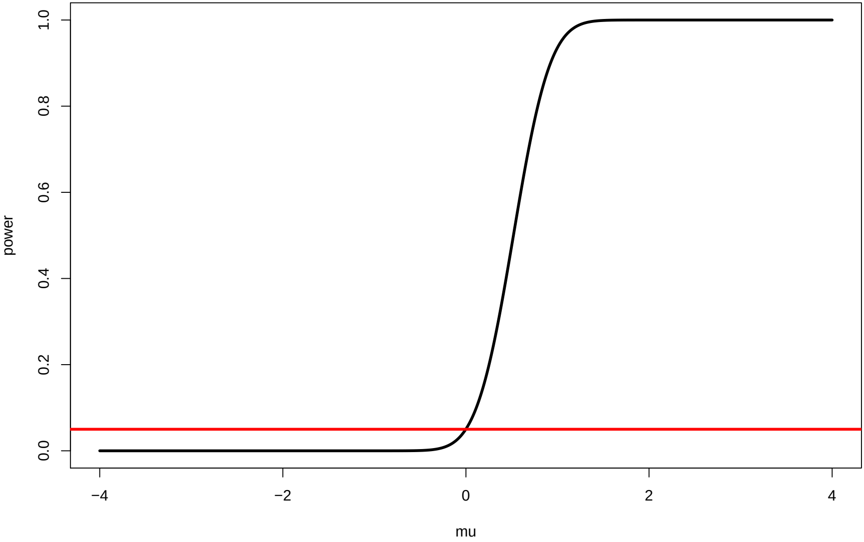
Neyman-Pearson Lemma

- Let's plot the power for the three tests for $n = 10, \mu_0 = 0, \alpha = 0.05$:

1. $H_0 : \mu = 0$ vs $H_1 : \mu > 0$

$$
\begin{aligned}
\beta(\mu) &= P\left(\frac{\bar{X} - \mu_0}{1/\sqrt{n}} \geq 1.64\right) \\
&= P\left(\frac{\bar{X} - \mu}{1/\sqrt{n}} - \frac{\mu_0 - \mu}{1/\sqrt{n}} \geq 1.64\right) \\
&= P\left(\frac{\bar{X} - \mu}{1/\sqrt{n}} \geq 1.64 + \frac{\mu_0 - \mu}{1/\sqrt{n}}\right) \\
&= P\left(Z \geq 1.64 - \frac{\mu}{1/\sqrt{n}}\right) = 1 - P\left(Z < 1.64 - \sqrt{n}\mu\right)
\end{aligned}
$$

```r
mu <- seq(-4,4, by=0.01)
n <- 10
power <- 1 - pnorm(1.64 - sqrt(n)*mu)
plot(mu, power, lwd=3, type="l")
abline(h=0.05, lwd=3, col="red")
```



1.0
0.8
0.6
0.4
0.2
0.0
power
−4
−2
0
2
4
mu

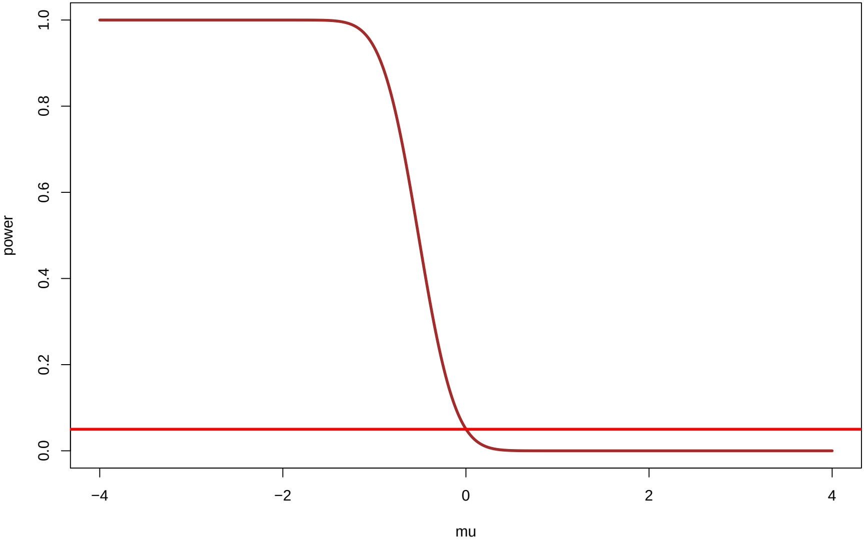
Neyman-Pearson Lemma

- Let's plot the power for the three tests for $n = 10, \mu_0 = 0, \alpha = 0.05$:

2. $H_0 : \mu = 0$ vs $H_1 : \mu < 0$

$$
\begin{aligned}
\beta(\mu) &= P\left(\frac{\bar{X} - \mu_0}{1/\sqrt{n}} \leq -1.64\right) \\
&= P\left(\frac{\bar{X} - \mu}{1/\sqrt{n}} - \frac{\mu_0 - \mu}{1/\sqrt{n}} \leq -1.64\right) \\
&= P\left(\frac{\bar{X} - \mu}{1/\sqrt{n}} \leq -1.64 + \frac{\mu_0 - \mu}{1/\sqrt{n}}\right) \\
&= P\left(Z \leq -1.64 - \sqrt{n}\mu\right)
\end{aligned}
$$

```r
mu <- seq(-4,4, by=0.01)
n <- 10
power <- pnorm(-1.64 - sqrt(n)*mu)
plot(mu, power, lwd=3, type="l", col="brown")
abline(h=0.05, lwd=3, col="red")
```

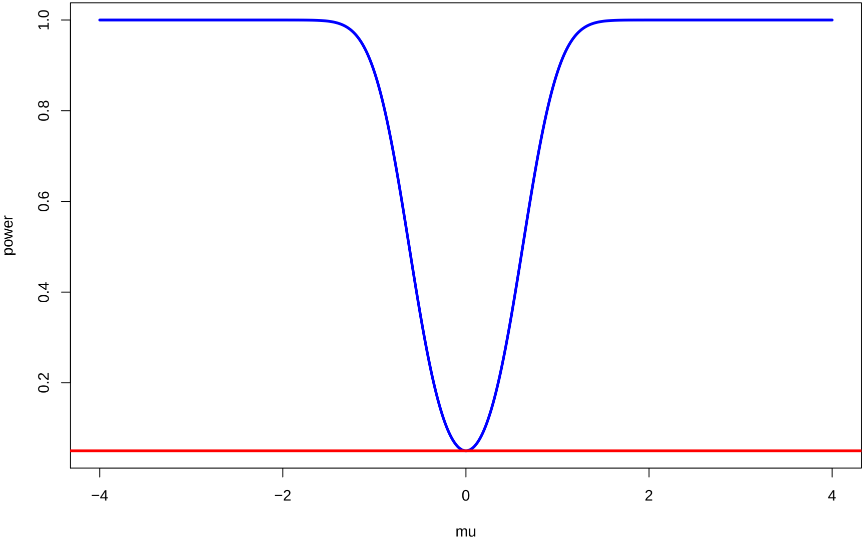

power
0.0
0.2
0.4
0.6
0.8
1.0
−4
−2
0
2
4
mu

Neyman-Pearson Lemma

3. $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$

$$
\begin{aligned}
\beta(\mu) &= P\left(\frac{\bar{X}-\mu_0}{1/\sqrt{n}} \geq 1.96\right) + P\left(\frac{\bar{X}-\mu_0}{1/\sqrt{n}} \leq -1.96\right) \\
&= P\left(\frac{\bar{X}-\mu}{1/\sqrt{n}} \geq 1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) + P\left(\frac{\bar{X}-\mu}{1/\sqrt{n}} \leq -1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) \\
&= P\left(Z \geq 1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) + P\left(Z \leq -1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) \\
&= 1 - P\left(Z < 1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) + P\left(Z \leq -1.96 + \frac{\mu_0-\mu}{1/\sqrt{n}}\right) \\
&= 1 - P\left(Z < 1.96 - \sqrt{n}\mu\right) + P\left(Z \leq -1.96 - \sqrt{n}\mu\right)
\end{aligned}
$$

```r
mu <- seq(-4,4, by=0.01)
n <- 10
power <- 1 - pnorm(1.96 - sqrt(n)*mu) +
  pnorm(-1.96 - sqrt(n)*mu)
plot(mu, power, lwd=3, type="l", col="blue")
abline(h=0.05, lwd=3, col="red")
```


1.0
0.8
0.6
0.4
0.2
power
−4
−2
0
2
4
mu

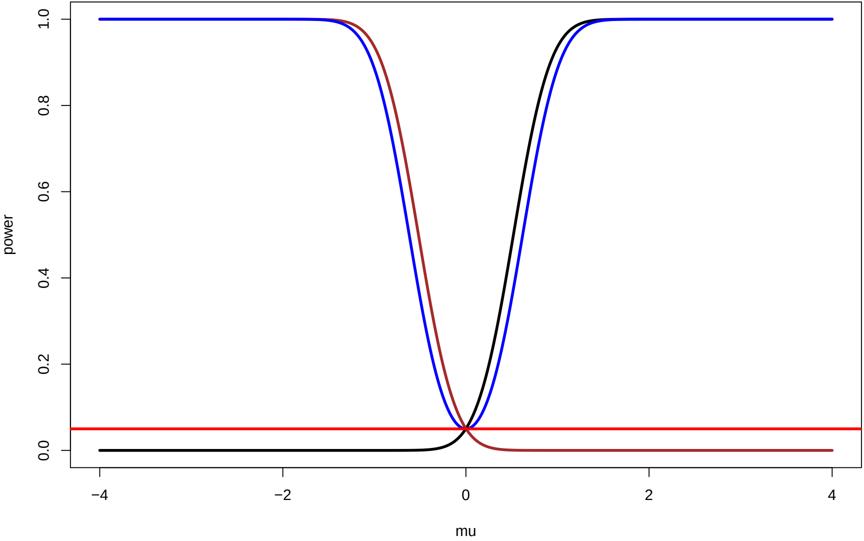
All Together

```
mu <- seq(-4,4, by=0.01)
n <- 10

##
power.1 <- 1 - pnorm(1.64 - sqrt(n)*mu)
power.2 <- pnorm(-1.64 - sqrt(n)*mu)
power.3 <- 1 - pnorm(1.96 - sqrt(n)*mu) +
  pnorm(-1.96 - sqrt(n)*mu)

##
plot(mu, power.1, lwd=3, type="l", ylab="power")
lines(mu, power.2, col="brown", lwd=3)
lines(mu, power.3, col="brown", lwd=3)

#
abline(h=0.05, lwd=3, col="red")
```



- Test 1 (black): $H_1 : \mu > 0$, Test 2 (brown): $H_1 : \mu < 0$, Test 3 (blue): $H_1 : \mu \neq 0$.

N-P Lemma

- From the plot, we see that Test 1 is UMP for $H_1 : \mu > 0$.
- From the plot, we see that Test 2 is UMP for $H_1 : \mu < 0$.

- Test 3 (maximum likelihood ratio test) is not UMP!
- Fortunately, it turns out that even when the maximum likelihood ratio test is not UMP (and many times it is), it typically has excellent properties (in particular, it can be shown to have nearly the largest possible power as the sample size increases towards infinity). As such, we tend to use the maximum likelihood ratio test in most complex testing situations where no other specific UMP test is available.

Likelihood Ratio Tests

Rice Section 9.4: The likelihood ratio test for testing $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_0^c = \Theta_1$ is:

$$\lambda(\boldsymbol{x}) = \frac{\underset{\Theta_0}{sup} L(\theta|\boldsymbol{x})}{\underset{\Theta}{sup} L(\theta|\boldsymbol{x})}$$

- The test has a rejection of the form $R = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq c\}$.
- Where $0 \leq c \leq 1$.
- Note:
  - $\underset{\Theta_0}{sup} L(\theta|\boldsymbol{x})$ is a restricted maximization.
  - $\underset{\Theta}{sup} L(\theta|\boldsymbol{x})$ is a unrestricted maximization.

Likelihood Ratio Tests

Example: Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{normal}(\theta, 1)$.

- Test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.
- θ_0 is a number fixed by the experimenter prior to the experiment.

$$\sup_{\Theta_0} L(\theta|\boldsymbol{x}) = L(\theta_0|\boldsymbol{x})$$

$$\sup_{\Theta} L(\theta|\boldsymbol{x}) = L(\hat{\theta}|\boldsymbol{x}) \Rightarrow \hat{\theta} = \bar{X}$$

Likelihood Ratio Tests

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{(2\pi)^{-n/2} exp[-\sum(x_i-\theta_0)^2/2]}{(2\pi)^{-n/2} exp[-\sum(x_i-\bar{x})^2/2]} \\
&= exp\left[\left(-\sum(x_i-\theta_0)^2+\sum(x_i-\bar{x})^2\right)/2\right] \\
&= exp\left[\left(-\left[\sum(x_i-\bar{x})^2+n(\bar{x}-\theta_0)^2\right]+\sum(x_i-\bar{x})^2\right)/2\right] \\
&= exp\left[-n(\bar{x}-\theta_0)^2/2\right]
\end{aligned}
$$

Likelihood Ratio Tests

$$
\begin{aligned}
R &= \{\lambda(\boldsymbol{x}) \leq c\} \\
&= \left\{exp\left[-n(\bar{x}-\theta_0)^2/2\right] \leq c\right\} \\
&= \left\{-n(\bar{x}-\theta_0)^2/2 \leq log(c)\right\} \\
&= \left\{(\bar{x}-\theta_0)^2 > [-2log(c)]/n\right\} \\
&\Rightarrow \left\{|\bar{x}-\theta_0| > \sqrt{[-2log(c)]/n}\right\} \\
&\Rightarrow \left\{\frac{|\bar{x}-\theta_0|}{1/\sqrt{n}} > \frac{\sqrt{[-2log(c)]/n}}{1/\sqrt{n}}\right\} \\
&= \left\{|Z| > \frac{\sqrt{[-2log(c)]/n}}{1/\sqrt{n}}\right\}
\end{aligned}
$$

Likelihood Ratio Tests

- Now we have:

$$R = \left\{ |Z| > \sqrt{n}\sqrt{[-2log(c)]/n} \right\} = \{|Z| > c^*\}$$

- Under the null hypothesis $\theta = \theta_0$. So $Z \sim \mathrm{normal}(0, 1)$.

$$\begin{aligned} P(|Z| > c^*) &= P(Z > c^*) + P(Z < -c^*) = \alpha \\ &= 2P(Z < -c^*) = \alpha \\ &= P(Z < -c^*) = \alpha/2 \\ &= P(Z < c^{**}) = \alpha/2 \end{aligned}$$

Likelihood Ratio Tests

- Suppose $\alpha = 0.05$, then $c^{**} = -1.96$

```r
qnorm(0.05/2)
```

```
## [1] -1.959964
```

- So we will reject H_0 if:

$$\left\{ \left| \frac{(\bar{x} - \theta_0)}{1/\sqrt{n}} \right| > 1.96 \right\}$$

Likelihood Ratio Tests

Eg. Let $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\theta)$.

- Test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.
- θ_0 is a number fixed by the experimenter prior to the experiment.

$$\underset{\Theta_0}{sup} L(\theta|\boldsymbol{x}) = L(\theta_0|\boldsymbol{x})$$

$$\underset{\Theta}{sup} L(\theta|\boldsymbol{x}) = L(\hat{\theta}|\boldsymbol{x}) \Rightarrow \hat{\theta} = \bar{X}$$

Likelihood Ratio Tests

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{\frac{exp(-n\theta_0)\theta_0^{\sum x_i}}{\prod x_i!}}{\frac{exp(-n\hat{\theta})\hat{\theta}^{\sum x_i}}{\prod x_i!}} \\
&= \frac{exp(-n\theta_0)\theta_0^{\sum x_i}}{exp(-n\hat{\theta})\hat{\theta}^{\sum x_i}} \\
&= exp(-n(\theta_0 - \hat{\theta}))\left(\frac{\theta_0}{\hat{\theta}}\right)^{\sum x_i} \\
&= exp(-n(\theta_0 - \hat{\theta}))\left(\frac{\theta_0}{\hat{\theta}}\right)^{n\bar{x}} \\
&= exp(-n(\theta_0 - \bar{x}))\left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}}
\end{aligned}
$$

Likelihood Ratio Tests

- The rejection region is of the form:

$$R = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq c\} \quad = \quad \left\{ exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}} \leq c \right\}$$

- Notice again that this is based on a sufficient statistic.
- If we could determine the distribution of $\lambda(\boldsymbol{X})$ we could then determine c for a given α!
- Looks a bit tricky here!!

Likelihood Ratio Tests - Asymptotics

Theorem A: For testing $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_0^c = \Theta_1$,

- suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x|\theta)$ and $\hat{\theta}$ is the MLE of θ and $f(x|\theta)$ satisfies the regularity conditions (smoothness).
- Then under H_0, as $n \to \infty$,

$$-2log[\lambda(\boldsymbol{x})] \xrightarrow{D} \chi_1^2$$

Likelihood Ratio Tests - Asymptotics

Proof:

- Do a two-step Taylor series expansion of $\ell(\theta|\boldsymbol{x})$ around $\hat{\theta}$:

$$\ell(\theta|\boldsymbol{x}) = \ell(\hat{\theta}|\boldsymbol{x}) + \ell'(\hat{\theta}|\boldsymbol{x})(\theta - \hat{\theta}) + \ell''(\hat{\theta}|\boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2} + \cdots$$

- $\ell'(\hat{\theta}|\boldsymbol{x}) = 0$ and dropping $(\cdots)$, we have:

$$\ell(\theta|\boldsymbol{x}) = \ell(\hat{\theta}|\boldsymbol{x}) + \ell''(\hat{\theta}|\boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2}$$

Likelihood Ratio Tests - Asymptotics

- Now consider:

$$-2log(\lambda) = -2[\ell(\theta_0|\boldsymbol{x}) - \ell(\hat{\theta}|\boldsymbol{x})]$$

- Substitute Taylor's approximation for $\ell(\theta_0|\boldsymbol{x})$:

$$\begin{aligned}
-2log(\lambda) &= -2\ell(\theta_0|\boldsymbol{x}) + 2\ell(\hat{\theta}|\boldsymbol{x}) \\
&= -2\left[\ell(\hat{\theta}|\boldsymbol{x}) + \ell''(\hat{\theta}|\boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2}\right] + 2\ell(\hat{\theta}|\boldsymbol{x}) \\
&= -\ell''(\hat{\theta}|\boldsymbol{x})(\theta - \hat{\theta})^2
\end{aligned}$$

Likelihood Ratio Tests - Asymptotics

- Now, $-\frac{1}{n}\ell''(\hat{\theta}|\boldsymbol{x}) \stackrel{LLN}{\rightarrow} i(\theta_0)$.

$$
\begin{aligned}
-2log(\lambda) &= -\ell''(\hat{\theta}|\boldsymbol{x})(\theta - \hat{\theta})^2 \\
&= ni(\theta)(\hat{\theta} - \theta)^2 \\
&= \left[\sqrt{ni(\theta)}(\hat{\theta} - \theta)\right]^2 \\
&= \left[\frac{\sqrt{n}(\hat{\theta} - \theta)}{1/\sqrt{i(\theta)}}\right]^2
\end{aligned}
$$

- We showed:

$$\sqrt{n}(\hat{\theta} - \theta) \stackrel{D}{\rightarrow} \mathrm{normal}(0, i(\theta)^{-1})$$

Likelihood Ratio Tests - Asymptotics

- So:

$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{1/\sqrt{i(\theta)}} = Z \xrightarrow{D} \mathrm{normal}(0, 1)$$

- Thus:

$$-2log(\lambda) = Z^2 \xrightarrow{D} \chi^2_1$$

Likelihood Ratio Tests - Asymptotics

- Back to our Poisson example:

$$R = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq c\} \quad = \quad \left\{exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}} \leq c\right\}$$

- Consider the asymptotic distribution:

$$\begin{aligned} -2log(\lambda) &= -2log\left[exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}}\right] \\ &= 2n\left[(\bar{x} - \theta_0) + \bar{x}\,log\left(\frac{\theta_0}{\bar{x}}\right)\right] \sim \chi^2_1 \end{aligned}$$

Likelihood Ratio Tests - Asymptotics

- If we reject when $\{\lambda \leq c\}$, then we reject when

$$\{-2log(\lambda) > -2log(c)\} = \{-2log(\lambda) > c^*\}$$

- What value of c^* should we pick so that $\alpha = 0.05$?

$$P(-2log(\lambda) > c^*) = 0.05$$

```r
qchisq(0.95, 1)
```

```
## [1] 3.841459
```

```r
1-pchisq(3.841,1)
```

```
## [1] 0.05001368
```

Likelihood Ratio Tests - Asymptotics

Theorem A: This theorem extends the previous one to allow for more parameters. It can be shown:

$$-2log(\lambda) \xrightarrow{D} \chi^2_\nu$$

where $\nu = \#\text{number of constraints set in } H_0$.

- Another way to think about it is: Let p be the number of parameters estimated (are free) under H_1. And let p_0 be the the number of parameters estimated (are free) under H_0.
- Then $\nu = p - p_0$.