1 Hypothesis Testing

Consider the family $\{f(\underset{\sim}{x} \mid \theta), \theta \in \Theta, \underset{\sim}{x} \in \mathbb{R}^n\}$. Data $\underset{\sim}{X} \sim f(\underset{\sim}{x} \mid \theta)$. We are interested in testing the hypothesis:

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1$$

for $\Theta_0, \Theta_1 \subset \Theta$ and $\Theta_0 \cap \Theta_1 = \phi$. Often $\Theta_1 = \Theta_0^c$.

1.1 Procedure

Definition 1. *We define rejection (critical region) as a subset $R \subset \mathbb{R}^n$ such that if $\underset{\sim}{X} \in R$, we reject H_0.*

Simplest situation: $H_0 : \theta = \theta_0$ (simple null) vs. $H_1 : \theta = \theta_1$ (simple alternative).
Terminalogy: size: $P_{\theta_0}(\underset{\sim}{X} \in R)$, power $= P_{\theta_1}(\underset{\sim}{X} \in R)$. In the design of tests: we fix the size in advance and choose R to maximize power.

1.2 Neyman-Pearson Lemma

The most powerful tests of $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ are based on the likelihood ratio

$$\text{LR}(\underset{\sim}{x}) = \frac{f(\underset{\sim}{X} \mid \theta_1)}{f(\underset{\sim}{X} \mid \theta_0)}.$$

and the rejection region is given by

$$\{\underset{\sim}{x} : \text{LR}(\underset{\sim}{x}) > k\}.$$

Lemma 1. *Assume both $f(\underset{\sim}{x} \mid \theta_0)$ and $f(\underset{\sim}{x} \mid \theta_1)$ to be densities or pmf's. Suppose R satisfies (for some $k \geq 0$),*

1. *If $LR(\underset{\sim}{x}) > k$, then $\underset{\sim}{x} \in R$ (equivalently $f(\underset{\sim}{x} \mid \theta_1) > kf(\underset{\sim}{x} \mid \theta_0)$)*

2. If $LR(\underset{\sim}{x}) < k$, then $\underset{\sim}{x} \in R^c$ (equivalently $f(\underset{\sim}{x} \mid \theta_1) < kf(\underset{\sim}{x} \mid \theta_0)$)

and

$$P_{\theta_0}(\underset{\sim}{X} \in R) = \alpha, \quad P_{\theta_1}(\underset{\sim}{X} \in R) = \beta.$$

Then for any other test R', if $P_{\theta_0}(\underset{\sim}{X} \in R') \leq \alpha$, then $P_{\theta_1}(\underset{\sim}{X} \in R') \leq \beta$.

Proof. Assume $f(\underset{\sim}{x} \mid \theta_0)$ and $f(\underset{\sim}{x} \mid \theta_1)$ are densities. Define

$$\phi(\underset{\sim}{x}) = I_R(\underset{\sim}{x}), \quad \phi'(\underset{\sim}{x}) = I_{R'}(\underset{\sim}{x})$$

Note that

$$P_\theta(\underset{\sim}{X} \in R) = \int \phi(\underset{\sim}{x}) f(\underset{\sim}{x} \mid \theta) d\underset{\sim}{x}.$$

Observe that

$$\begin{aligned} & 0 \leq \int [\phi(\underset{\sim}{x}) - \phi'(\underset{\sim}{x})][f(\underset{\sim}{x} \mid \theta_1) - kf(\underset{\sim}{x} \mid \theta_0)] d\underset{\sim}{x} \\ = \quad & P_{\theta_1}(\underset{\sim}{X} \in R) - P_{\theta_1}(\underset{\sim}{X} \in R') - k\{P_{\theta_0}(\underset{\sim}{X} \in R) - P_{\theta_0}(\underset{\sim}{X} \in R')\}. \end{aligned}$$

Since $P_{\theta_0}(\underset{\sim}{X} \in R) \geq P_{\theta_0}(\underset{\sim}{X} \in R')$, we have $P_{\theta_1}(\underset{\sim}{X} \in R) \geq P_{\theta_1}(\underset{\sim}{X} \in R')$. □

<u>Example</u>

```
f0 and f1 are mass functions
assigning probs to values x from 1 to 12.

LR is the likelihood ratio: LR = f1(x)/f0(x)


       1     2     3      4     5     6     7     8     9    10    11    12
f0 0.123 0.016 0.125  0.006 0.047 0.149 0.011 0.144 0.028 0.034 0.103 0.214
f1 0.087 0.032 0.103  0.120 0.008 0.015 0.109 0.071 0.082 0.150 0.054 0.169
LR 0.707 2.000 0.824 20.000 0.170 0.101 9.909 0.493 2.929 4.412 0.524 0.790


Reordering x so that LR is decreasing

        4     7    10     9     2     3    12     1    11     8     5     6
f0  0.006 0.011 0.034 0.028 0.016 0.125 0.214 0.123 0.103 0.144 0.047 0.149
f1  0.120 0.109 0.150 0.082 0.032 0.103 0.169 0.087 0.054 0.071 0.008 0.015
LR 20.000 9.909 4.412 2.929 2.000 0.824 0.790 0.707 0.524 0.493 0.170 0.101


Various likelihood ratio tests: reject when LR > const

                              const alpha typeII power
{}                              Inf 0.000  1.000 0.000
{4}                          14.955 0.006  0.880 0.120
{4,7}                         7.160 0.017  0.771 0.229
{4,7,10}                      3.670 0.051  0.621 0.379
{4,7,10,9}                    2.464 0.079  0.539 0.461
{4,7,10,9,2}                  1.412 0.095  0.507 0.493
{4,7,10,9,2,3}                0.807 0.220  0.404 0.596
{4,7,10,9,2,3,12}             0.749 0.434  0.235 0.765
{4,7,10,9,2,3,12,1}           0.616 0.557  0.148 0.852
{4,7,10,9,2,3,12,1,11}        0.509 0.660  0.094 0.906
{4,7,10,9,2,3,12,1,11,8}      0.332 0.804  0.023 0.977
{4,7,10,9,2,3,12,1,11,8,5}    0.135 0.851  0.015 0.985
{4,7,10,9,2,3,12,1,11,8,5,6} -1.000 1.000  0.000 1.000


Calculation of alpha (= size), power, and Prob(type II error)
For two different rejection regions:

rejection region: {4, 7}

alpha = f0(4) + f0(7) = 0.006 + 0.011 = 0.017
power = f1(4) + f1(7) = 0.12 + 0.109 = 0.229
typeII = 1 - power = 0.771


rejection region: {4, 7, 10, 9}

alpha = f0(4) + f0(7) + f0(10) + f0(9) = 0.006 + 0.011 + 0.034 + 0.028 = 0.079
power = f1(4) + f1(7) + f1(10) + f1(9) = 0.12 + 0.109 + 0.15 + 0.082 = 0.461
typeII = 1 - power = 0.539
```

Figure 1: Likelihood ratio test for a discrete problem

Example: Let $X \sim \mathrm{N}(\theta, 1)$ (assume $\theta_1 > \theta_0$). Test $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$. Then

$$\begin{aligned} \mathrm{LR}(x) = \frac{f(x \mid \theta_1)}{f(x \mid \theta_0)} &= \exp\Big\{ -\frac{1}{2}(x-\theta_1)^2 + \frac{1}{2}(x-\theta_0)^2 \Big\} \\ &= \exp\Big\{ (\theta_1 - \theta_0)x + \frac{1}{2}(\theta_0^2 - \theta_1^2) \Big\} \end{aligned}$$

LR(x) is strictly increasing (strictly decreasing when $\theta_1 < \theta_0$). By NP Lemma, most powerful tests have rejection revisions of the form

$$R = \{x : \text{LR}(x) > k\} = \{x : x > k^*\}.$$

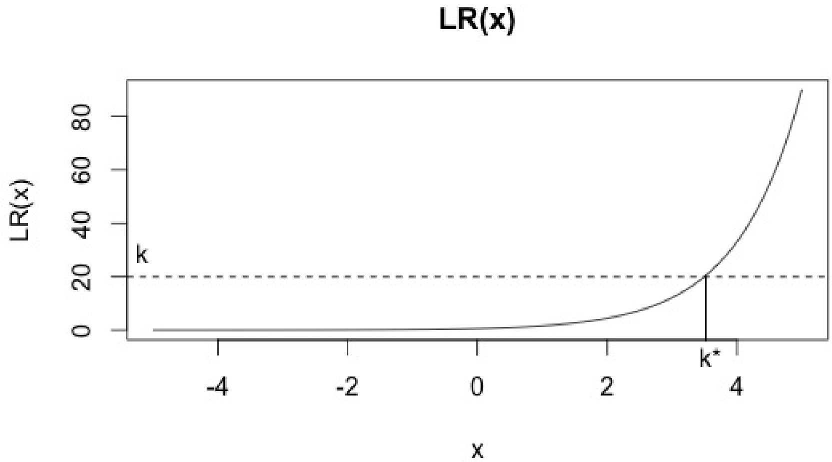

Figure 2: Likelihood ratio for normal problem as a function of x

(Suppose $\theta_0 = 0$.) To get a size α test, choose k^* such that

$$P_{\theta_0}(X > k^*) = \alpha.$$

Note: k^* does not depend on θ_1, so long as $\theta_1 > \theta_0$.
Example: $X \sim \text{Cauchy}(\theta)$ where

$$f(x \mid \theta) = \frac{1}{\pi(1 + (x - \theta)^2)}.$$

Test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ (where $\theta_1 > \theta_0$).

$$\text{LR}(x) = \frac{1 + (x - \theta_0)^2}{1 + (x - \theta_1)^2}.$$

LR not monotonic.

$$\lim_{x \to \pm\infty} \text{LR}(x) = 1.$$

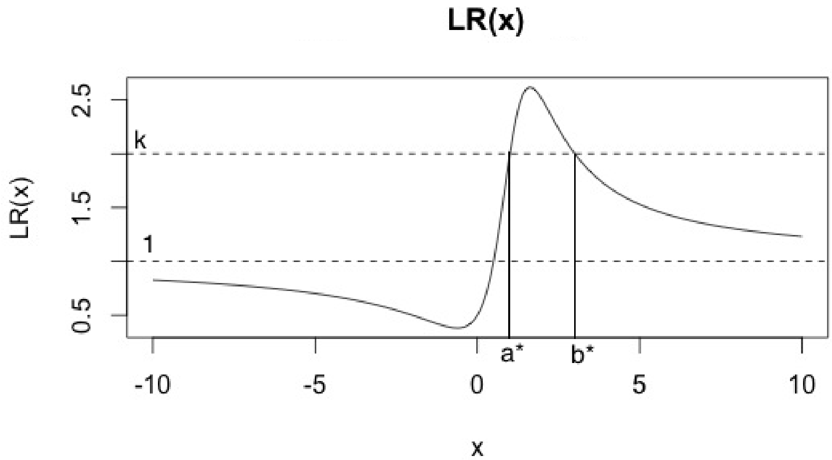

Figure 3: Likelihood ratio for Cauchy problem as a function of x

By NP Lemma, most powerful tests have rejection regions of the form

$$R = \{x : \text{LR}(x) > k\} = \{x : a^* < x < b^*\}.$$

To get size α test: Choose k so $P_{\theta_0}(\text{LR}(X) > k) = \alpha$. This gives a^* and b^* which depend on both θ_0 and θ_1.

1.3 Sufficient Statistics and Testing

If $T(X)$ is sufficient for θ, then testing can be done in terms of $T = T(\underset{\sim}{X})$.

$$\begin{aligned}
\text{LR}_{\underset{\sim}{X}}(\underset{\sim}{x}) &= \frac{f(\underset{\sim}{x} \mid \theta_1)}{f(\underset{\sim}{x} \mid \theta_0)} = \frac{P_{\theta_1}(\underset{\sim}{X} = \underset{\sim}{x})}{P_{\theta_0}(\underset{\sim}{X} = \underset{\sim}{x})} \\
&= \frac{P_{\theta_1}(T(\underset{\sim}{X} = T(\underset{\sim}{x}))P_{\theta_1}(\underset{\sim}{X} = \underset{\sim}{x} \mid T = t)}{P_{\theta}(T = t)P_{\theta_0}(\underset{\sim}{X} = \underset{\sim}{x} \mid T = t)} \\
&= \frac{P_{\theta_1}(T = t)}{P_{\theta_0}(T = t)} = \frac{f_T(t \mid \theta_1)}{f_T(t \mid \theta_0)} = \text{LR}_T(t).
\end{aligned}$$

Thus $\text{LR}_X(x) > k$ iff $\text{LR}_T(t) > k$. Most powerful tests based on X or T are equivalent. (Always produce same outcome, accept or reject.)

1.4 Power function, size and level

Let $X \sim P_\theta, \theta \in \Theta$. Let R be the rejection region for a test of $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$.

Definition 2. *(Power function)* $P_\theta(X \in R) = \beta(\theta)$.

Note: For $\theta \in \Theta_0$, $\beta(\theta)$ = prob of Type I error. For $\theta \in \Theta_1$, $1 - \beta(\theta)$ = prob of Type II error. Performance of a test is judged by power function. Tests are compared via power functions.
For $0 \leq \alpha \leq 1$, a test is size α if

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha.$$

and level α if

$$\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha.$$

2 Likelihood Ratio Test (LRT)

Suppose $\mathbf{X} \sim P_\theta, \theta \in \Theta$, with oint pdf (or pmf) $f(\mathbf{x} \mid \theta)$. For observed data $\mathbf{x}$, the likelihood function is $L(\theta \mid \mathbf{x}) \equiv f(\mathbf{x} \mid \theta)$. The LRT statistic for testing

$$H_0 : \theta \in \Theta_0, \quad \text{versus} \quad H_1 : \theta \in \Theta_0^c$$

is given by

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta \mid \mathbf{x})}{\sup_{\theta \in \Theta} L(\theta \mid \mathbf{x})} = \frac{L(\hat{\theta}_0 \mid \mathbf{x})}{L(\hat{\theta} \mid \mathbf{x})}$$

where $\hat{\theta}_0 = \text{argmax}_{\theta \in \Theta_0} L(\theta \mid \mathbf{x})$ and $\hat{\theta} = \text{argmax}_{\theta \in \Theta} L(\theta \mid \mathbf{x})$.
The LRT rejects for small values of $\lambda(\mathbf{x})$; the test has rejection region (critical region) given by

$$R = \{\mathbf{x} : \lambda(\mathbf{x}) \leq c\}$$

where c is chosen so that

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(\mathbf{X}) \leq c) = \alpha \quad \text{(or failing that, } \leq \alpha).$$

for some pre-specified value α (say 0.05 or 0.01). Sometimes the exact distribution of $\lambda(\mathbf{X})$ can be obtained and then used to find c giving an exact size α test. But often this cannot be done, and we have to rely on the following asymptotic approximation.

3 Asymptotic Distribution of LRT Statistic

Consider a sequence of successively larger data sets

$$\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$$

and let $\lambda_n(\mathbf{X}_n)$ be the LRT statistic based on $\mathbf{X}_n$.

Theorem 1. *If $\theta \in \Theta_0$, then (under regularity conditions)*

$$-2\log\lambda_n(\mathbf{X}_n) \xrightarrow{d} \chi^2_k, \quad as \quad n \to \infty$$

where $k \equiv (dim(\Theta) - (dim\Theta_0))$.

Thus, if c^* satisfies $P(\chi^2_k \geq c^*) = \alpha$, then the rejection region $R = \{\mathbf{x} : -2\log\lambda(\mathbf{x}) \geq c^*\}$ gives an approximate size α test for larger sample sizes.
Comment: The test statistics $\lambda(\mathbf{x})$ and $-2\log\lambda(\mathbf{x})$ are equivalent since

$$\lambda(\mathbf{x}) \leq c, \quad \text{iff} -2\log\lambda(\mathbf{x}) \geq c^*$$

where $c^* \equiv -2\log c$.
It is often convenient to replace the LRT statistic $\lambda(\mathbf{x})$ by an equivalent statistic obtained by applying a strictly monotone transformation.
Comment on the regularity conditions: Conditions are required on both the family of distributions $f(\mathbf{x} \mid \theta)$ and the set Θ_0. The family $f(\mathbf{x} \mid \theta)$ must satisfy conditions like those required for the consistency and asymptotic normality of the MLE (and the validity of the Fisher information). The set Θ_0 must be a lower dimensional subspace (or manifold) of Θ.

Lemma 2. *Let $T, n > 0$. Define*

$$H(\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-T/(2\sigma^2)}, \quad \sigma^2 > 0.$$

Then

$$\begin{aligned} argmax_{\sigma^2>0} H(\sigma^2) &= T/n \equiv \hat{\sigma}^2, and \\ \sup_{\sigma^2>0} H(\sigma^2) &= H(\hat{\sigma}^2) = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}. \end{aligned}$$

Example: Observe $X_1, \ldots, X_n$ iid $\text{N}(0, \sigma^2)$. Find LRT of

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 \neq \sigma_0^2.$$

Here:

$$
\begin{aligned}
\Theta &= (0,\infty) \quad \text{and}\, \Theta_0 = \{\sigma_0^2\}. \\
L(\sigma^2) &= (2\pi\sigma^2)^{-n/2}\exp\{-(2\sigma^2)^{-1}\sum_i x_i^2\}. \\
\text{argmax}_\Theta L(\sigma^2) &= n^{-1}\sum_i x_i^2 \equiv \hat{\sigma}^2 \quad (\text{so that } \sum_i x_i^2 = n\hat{\sigma}^2) \\
\lambda(\mathbf{x}) &= \frac{L(\sigma_0^2)}{L(\hat{\sigma}^2)} = \frac{(2\pi\sigma_0^2)^{-n/2}\exp\{-(2\sigma_0^2)^{-1}n\hat{\sigma}^2\}}{(2\pi\hat{\sigma}^2)^{-n/2}\exp\{-n/2\}} \\
&= e^{n/2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2}\exp\left[-\frac{n}{2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)\right] \\
&= \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)
\end{aligned}
$$

where $\psi(u) \equiv e^{n/2}u^{n/2}e^{-(n/2)u}$.

With $X_1, X_2, \ldots, X_n$ iid $\mathrm{N}(0,\sigma^2)$, the LRT of

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs.} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

rejects in the region

$$R = \{x : \lambda(x) \leq c\}$$

where

$$\lambda(x) = \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right), \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n x_i^2, \quad \psi(u) = e^{n/2}u^{n/2}e^{-(n/2)u}.$$

The function ψ is maximized at $u = 1$, refer to the following figure. Find a and b which are functions of c such that $\psi(a) = \psi(b) = c$.

Thus

$$R = \left\{x : \frac{\hat{\sigma}^2}{\sigma_0^2} \leq a(c) \quad \text{or} \quad \frac{\hat{\sigma}^2}{\sigma_0^2} \geq b(c)\right\}.$$

We reject when $\frac{\hat{\sigma}^2}{\sigma_0^2}$ departs far enough from 1.

<u>Obtaining an exact level α test:</u>

$$
\begin{aligned}
P_{\sigma_0^2}(X \in R) &= 1 - P_{\sigma_0^2}\left(a(c) < \frac{\hat{\sigma}^2}{\sigma_0^2} < b(c)\right) \\
&= 1 - P_{\sigma_0^2}\left(na(c) < \frac{n\hat{\sigma}^2}{\sigma_0^2} < nb(c)\right).
\end{aligned}
$$

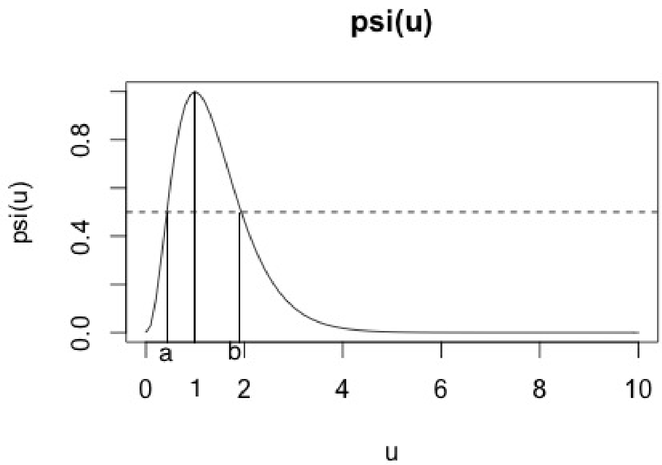

Figure 4: Plot of $\psi(u)$

Note that

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} = \sum_{i=1}^{n} \left(\frac{X_i}{\sigma_0}\right)^2 \sim \chi_n^2$$

under H_0. An exact level α test is thus obtained by chosen c so that

$$P(na(c) < \chi_n^2 < nb(c)) = 1 - \alpha.$$

Finding c requires computation. An easier approach is to reject H_0 when

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} \le \chi_n^2(\alpha/2) \quad \text{or} \quad \frac{n\hat{\sigma}^2}{\sigma_0^2} \ge \chi_n^2(1-\alpha/2)$$

where $\chi_n^2(\alpha/2)$ and $\chi_n^2(1-\alpha/2)$ are the values which cut off probability $\alpha/2$ in the left and right tails of the χ_n^2 distribution.
Example continued: (A variation) Find the LRT with size α of

$$H_0 : \sigma^2 \le \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 > \sigma_0^2$$

Now we have: $\Theta_0 \subset (0, \sigma_0^2]$. Note that

$$\text{argmax}_{\sigma^2 \in \Theta_0} L(\sigma^2) \equiv \hat{\sigma}_0^2 = \begin{cases} \hat{\sigma}^2 & \text{if} \quad \hat{\sigma}^2 \le \sigma_0^2 \\ \sigma_0^2 & \text{if} \quad \hat{\sigma}^2 > \sigma_0^2. \end{cases}$$

since the likelihood function falls away monotonically on each size of $\hat{\sigma}^2$. Hence

$$\lambda(\mathbf{x}) = \frac{L(\sigma_0^2)}{L(\hat{\sigma}^2)} = \begin{cases} 1 & \text{if} \quad \hat{\sigma}^2 \le \sigma_0^2 \\ L(\sigma_0^2)/L(\hat{\sigma}^2) & \text{if} \quad \hat{\sigma}^2 > \sigma_0^2 \end{cases}$$
$$= \begin{cases} 1 & \text{if} \quad \hat{\sigma}^2 \le \sigma_0^2 \\ \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) & \text{if} \quad \hat{\sigma}^2 > \sigma_0^2 \end{cases}$$

Since $\psi(u)$ decreases for $u \ge 1$, we have $\lambda(\mathbf{x}) \le c$ iff $\hat{\sigma}^2/\sigma_0^2 \ge c^*$ iff $S \equiv \sum_{i=1}^n X_i^2/\sigma_0^2 \ge c'$ where c' is chosen to give size α.

$$\sup_{\sigma^2 \in \Theta_0} P_{\sigma^2}(S \ge c') = P_{\sigma_0^2}(S \ge c') = \alpha$$

if we choose c' such that $P(\chi_n^2 \ge c') = \alpha$.
Example continued: Another variation Observe that $X_1, X_2, \ldots, X_n$ iid $\mathrm{N}(\mu, \sigma^2)$. Fin the LRT of

$$H_0 : \sigma^2 = \sigma_0^2, \mu \in \mathbb{R} \quad \text{vs.} H_1 : \sigma^2 \ne \sigma_0^2, \mu \in \mathbb{R}.$$

Now we have:

$$\begin{aligned} \Theta &= \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\} \\ \Theta_0 &= \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 = \sigma_0^2\} \\ L(\mu, \sigma^2) &= (2\pi\sigma^2)^{n/2} \exp\left(-(2\sigma^2)^{-1} \sum_i (x_i - \mu)^2 \right) \\ \text{argmax}_{(\mu,\sigma^2)\in\Theta} L(\mu, \sigma^2) &= (\bar{x}, \hat{\sigma}^2), \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2 \\ \text{argmax}_{(\mu,\sigma^2)\in\Theta_0} L(\mu, \sigma^2) &= (\bar{x}, \sigma_0^2). \end{aligned}$$

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{L(\bar{x}, \sigma_0^2)}{L(\bar{x}, \hat{\sigma}^2)} = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left(-(2\sigma_0^2)^{-1} n\hat{\sigma}^2 \right)}{(2\pi\hat{\sigma}^2)^{-n/2} \exp\left(-n/2 \right)} \\ &= e^{n/2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{n}{2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)\right] = \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right). \end{aligned}$$

where $\psi(u) = e^{n/2}u^{n/2}e^{(-n/2)u}$. Just like before but with a different definition of $\hat{\sigma}^2$. Now determine critical values using $n\hat{\sigma}^2/\sigma_0^2 \sim \chi_{n-1}^2$ under H_0.

4 Uniformly most powerful test for composite hypothesis

Consider a family of distributions $f(\underset{\sim}{x} \mid \theta)$ with

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

and let T be sufficient for θ and $\mathrm{LR}_T(t)$ is non-decreasing.

Theorem 2. *(Karlin-Rubin) If any test with rejection region $R = \{T > c\}$ satisfies $\sup_{\theta \leq \theta_0} P_\theta(\underset{\sim}{X} \in R) = \alpha$, then it is an uniformly most powerful test in the sense that its power function dominates the power function of any α-level test for all points in the alternative hypothesis.*