# Statistical Inference

**Lecture 09a**

ANU - RSFAS

Last Updated: Sun Apr 29 15:24:50 2018

# Some Fun with MLRTs

$E(X) = \frac{1}{\theta}$

- **Eg:** Suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x;\theta) = \theta exp(-\theta x)$.
- Consider testing:

$$H_0: \ \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta \neq \theta_0.$$

$L(\theta) = \prod \theta \exp(-\theta x)$

$= \theta^n \exp(-\theta \sum x_i)$

$$
\begin{aligned}
\lambda &= \frac{\max_{\theta\in\theta_0} L(\theta)}{\max_{\theta\in\Omega} L(\theta)} \\
&= \frac{\theta_0^n exp(-\theta_0 \sum_{i=1}^n x_i)}{\hat{\theta}_n exp(-\hat{\theta}_1 \sum_{i=1}^n x_i)} \\
&= \frac{\theta_0^n exp(-\theta_0 \sum_{i=1}^n x_i)}{\left(\frac{1}{\bar{x}}\right)^n exp(-\left(\frac{1}{\bar{x}}\right) \sum_{i=1}^n x_i)} \\
&= \frac{\theta_0^n\ \bar{x}^n\ exp(-\theta_0 n\bar{x})}{exp(-n)}
\end{aligned}
$$

$$
\begin{aligned}
C &= \{\lambda \le k\} \\
&= \left\{ \frac{\theta_0^n \ \bar{x}^n \ exp(-\theta_0 n\bar{x})}{exp(-n)} \le k \right\} \\
&= \{\theta_0^n \ \bar{x}^n \ exp(-\theta_0 n\bar{x}) \le k^*\} \\
&= \{\bar{x}^n \ exp(-\theta_0 n\bar{x}) \le k^{**}\} \\
&= \{[\bar{x} \ exp(-\theta_0 \bar{x})]^n \le k^{**}\} \\
&= \{\bar{x} \ exp(-\theta_0 \bar{x}) \le k^{***}\}
\end{aligned}
$$

- This does not have a nice form.
- We could rely on the asymptotic result. Or let's see what we can do.

- Consider testing:

$$H_0: \ \theta = 1 \quad \text{vs.} \quad H_1: \theta \neq 1.$$

$$C \ = \ \{\bar{x} \ exp(-1 \ \bar{x}) \leq k^{***}\}$$

$P(C) = P(\lambda \leq k) = \alpha$
$= 0.05$

- Notice we have a function of the form:

$$f(a) = a \ exp(-a)$$

- Let's see what this looks like:

$$f'(a) = (1 - a)\ exp(-a)$$

- $f'(a)$ is positive for $a \ \in (0, 1)$
- $f'(a)$ is negative for $a \ \in (1, \infty)$
- So f(a) is increasing from $(0, 1)$ and decreasing from $(1, \infty)$.
- This suggests we may find:

$$C = \{\underset{\text{why } \bar{x}?}{\overset{\alpha}{\bar{x}}} \leq x_0\} \cup \{\overset{\alpha}{\bar{x}} \geq x_1\}$$

can figure out the dist'n with CLT

- We need to know the distribution of $\bar{x}$ under $H_0$.

$$MGF_{x} = \left(\frac{1}{1 - t/\theta}\right)$$

$$MGF_{\sum_{i=1}^{n} x_i} = \left(\frac{1}{1 - t/\theta}\right) \times \cdots \times \left(\frac{1}{1 - t/\theta}\right) = \left(\frac{1}{1 - t/\theta}\right)^{n}$$

- We can see that $Y = \sum_{i=1}^{n} X_i \sim \mathrm{gamma}(n, 1)$.

- Then let $W = Y/n$:

change of variable method.

$Y = nW$

$\frac{\partial Y}{\partial W} = n$

$$
\begin{aligned}
f_W(w) &= f_Y(y = wn)|n| \\
&= \frac{1^n}{\Gamma(n)} y^{n-1} exp(-y)|n| \\
&= \frac{1^n}{\Gamma(n)} (wn)^{n-1} exp(-(wn))|n| \\
&= \frac{n^n}{\Gamma(n)} w^{\overset{\alpha}{n}-1} exp(-w\overset{\beta}{n})
\end{aligned}
$$

- We can see that $W = Y/n = \bar{X} \sim \text{gamma}(n, n)$.

$$
\begin{aligned}
P(C) &= P(\{\bar{x}\ exp(-\theta_0 \bar{x}) \leq k\}) = \alpha \\
&= P(\{\bar{x} \leq x_0(k)\} \cup \{\bar{x} \geq x_1(k)\}) = \alpha \\
&= P(\{\bar{x} \leq x_0(k)\}) + P(\{\bar{x} \geq x_1(k)\}) = \alpha \\
&= F(x_0(k)) + (1 - F(x_1(k)))
\end{aligned}
$$

- Computationally:

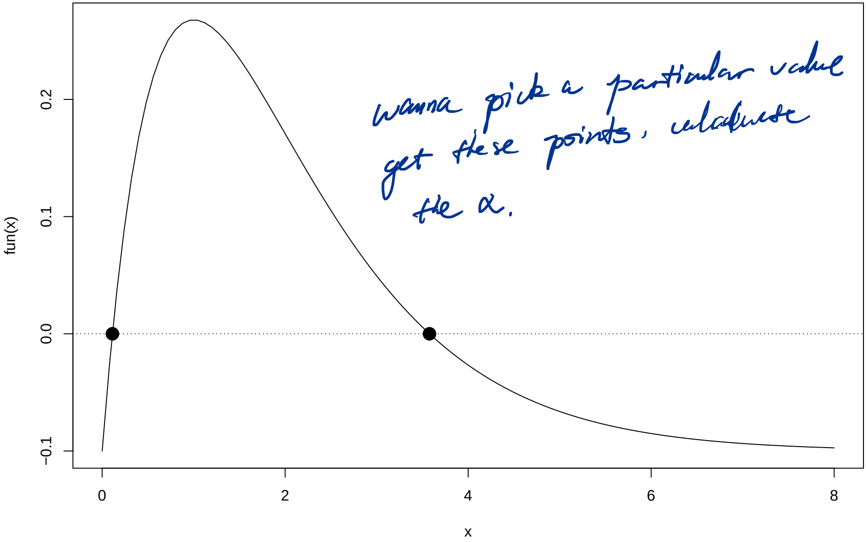
1. Try a value of $k$.
2. Compute $x_0, x_1$.
3. Compute $F(x_0(k)) + (1 - F(x_1(k)))$ and get the close to 0.05.

# Suppose $n = 10$

```
n <- 10

##
library(rootSolve)
k <- 0.10
fun <- function(x.bar){x.bar* exp(-x.bar) - k}
curve(fun(x), 0, 8)
abline(h = 0, lty = 3)
All <- uniroot.all(fun, c(0, 8))
points(All, y = rep(0, length(All)), pch = 16, cex = 2)

##    CDF
pgamma(All[1], n, rate=n) + 1 - pgamma(All[2], n, rate=n)
```




wanna pick a particular value
get these points, calculate
the α.

```
## [1] 4.080553e-07
```

- We can range $k$ from 0 to $\max(x * exp(-x))$. Note the maximum is at $\bar{x} = 1$.

```
1* exp(-1)
```

```
## [1] 0.3678794
```

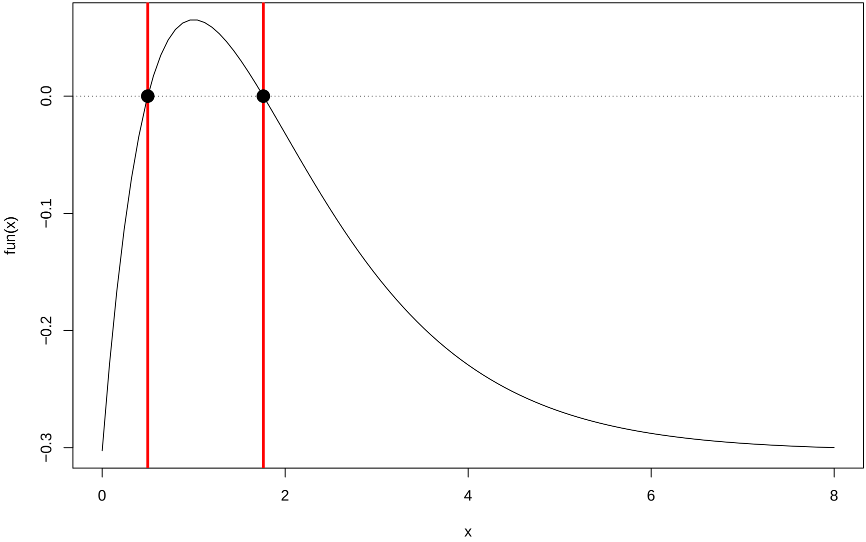
- Now let's use the `while` loop in R:

```
alpha <- 0.9
fun.k <- sort(seq(0.01, 0.366, by=0.00001), decreasing=TRUE)
i <- 1

while(alpha > 0.05){
  k <- fun.k[i]
  fun <- function(x.bar){x.bar* exp(-x.bar) - k}
  All <- uniroot.all(fun, c(0, 8))
  ##
  alpha <- pgamma(All[1], n, rate=n) +
    1 - pgamma(All[2], n, rate=n)
  i <- i+1
}
```

tried diff values of k

```r
curve(fun(x), 0, 8)
abline(h = 0, lty = 3)
points(All, y = rep(0, length(All)), pch = 16, cex = 2)
```



fun(x)
0.0
−0.1
−0.2
−0.3
0
2
4
6
8
x

- Let's examine quantities of interest:

```r
k
```

```
## [1] 0.30262
```

```r
All
```

$x_0$ $x_1$

```
## [1] 0.4978634 1.7613539
```

```r
round(alpha,2)
```

```
## [1] 0.05
```

- Reject $H_0$ if $\bar{x} \leq 0.498$ or $\bar{x} \geq 1.761$.

# Another Computational Approach

$\lambda$

$$P(C) = P[\bar{x}\ exp(-\theta_0 \bar{x}) \leq k] = \alpha = 0.05$$

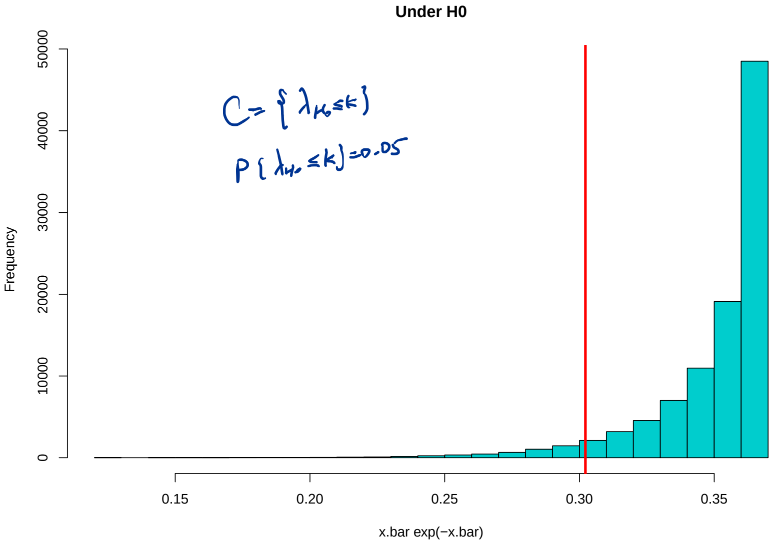
$\theta_0 = 1$

- Let's examine the statistic $\bar{X}exp(-\bar{X})$ under $H_0$.
- Generate repeated samples of size $n = 10$ from an exponential distribution with $\theta_0 = 1$.

```
set.seed(1001)
S <- 100000
out <- rep(0, S)
n <- 10

for(s in 1:S){
    x <- rexp(10, rate=1)
    x.bar <- mean(x)
    out[s] <- x.bar*exp(-x.bar)
}
```

```r
hist(out, col="cyan3", main="Under H0", xlab="x.bar exp(-x.bar)")
k <- quantile(out, 0.05)
abline(v=k, col="red", lwd=3)
```




- Reject $H_0$ if $(\bar{x}\ exp(-\bar{x})) \le 0.302$.

# Properties of Maximum Likelihood Ratio Tests

- The MLRT
  1. is asymptotically most powerful unbiased;
  2. is asymptotically similar;
  3. is asymptotically efficient.

# Unbiased Tests

**Definition 4.6:** Suppose that we wish to test:

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

A test of size $\alpha$ is said to be **unbiased** if

$$\eta(\boldsymbol{\theta}) \geq \alpha \text{ for all } \boldsymbol{\theta} \in \Theta_1.$$

# Similar Tests

**Definition 4.7:** Suppose that we wish to test:

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

A test of size $\alpha$ is said to be **similar** if

$$\eta(\boldsymbol{\theta}) = \alpha \text{ for all } \boldsymbol{\theta} \in \Theta_0.$$

# Efficiency

$\alpha = 0.05$
$\eta = 0.80$

**Definition 4.9:** Suppose that we have to possible tests of $H_0$ vs. $H_1$, where both tests are simple.

- If $n_1$ and $n_2$ are the minimum possible sample sizes for tests 1 and 2 for which we can achieve a size $\alpha$ and power $\geq \eta$, then the **relative efficiency** of test 1 compared to test 2 is:

$$n_2/n_1.$$

If $n_1 < n_2$
then I like
test 1

## Two Other Tests

- **The Score Test**

$$\boldsymbol{u}(\boldsymbol{\theta}) = \left(\frac{\partial \ell}{\partial \theta_1}, \frac{\partial \ell}{\partial \theta_2}, \cdots, \frac{\partial \ell}{\partial \theta_k}\right)^T$$

- Suppose that we wish to test:

$$H_0 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{vs.} \quad H_1 : \Omega - \{\boldsymbol{\theta}_0\}.$$

- Test statistic:

$$\boldsymbol{u}(\boldsymbol{\theta})^T \ \boldsymbol{I}_{\boldsymbol{\theta}_0}^{-1} \boldsymbol{u}(\boldsymbol{\theta}) \ \dot{\sim} \ \chi^2_{df=k}$$

- Note: We don't have to determine the MLEs!

- **The Wald Test**
- Suppose that we wish to test:

$$H_0: \ \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{vs.} \quad H_1 : \Omega - \{\boldsymbol{\theta}_0\}.$$

- Test statistic:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \ \boldsymbol{I}_{\hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \ \dot{\sim} \ \chi^2_{df=k}$$

- The MLRT, the Score Test, and the Wald Test are asymptotically equivalent!