Homework 12

Chen Xupeng

2017/12/25

1

(i) construct the test as below:

$$H_0 : \mu = \mu_0 \quad H_1 : \mu = \mu_1 \ (\mu_1 > \mu_0)$$

we can use NP lemma to have $\lambda(x)$

$$\lambda(x) = \frac{f_1(x)}{f_0(x)} = exp\{n\mu_1\bar{x} - n\mu_2\bar{x} + \frac{1}{2}(n\mu_0^2 - n\mu_1^2)\}$$

$$\lambda(x) \text{ does not involve } \mu_1, \text{ the test we construct is } H_0 : \mu = \mu_0 \quad H_1 : \mu \geq \mu_0$$

$$\lambda(x) \text{ is strictly increasing for } \bar{X}, \text{so the UMPT's rejection region is}\{X : \sqrt{n}(\bar{X} - \mu_0) > c\}$$

$$E(\phi(x)) = \alpha, \text{ we can have UMPT}\phi(x) = \begin{cases} 1, \ \sqrt{n}(\bar{x} - \mu_0)/\sigma > \mu_{\alpha/2} \\ 0, \ \sqrt{n}(\bar{x} - \mu_0)/\sigma \leq \mu_{\alpha/2} \end{cases}$$

(ii)

$$\sqrt{n}(\bar{x} - \mu_0)/\sigma = 1 < 1.96, \text{ we do not reject null hypothesis}$$

$$\text{we reject the null hypothesis when } \bar{x} > 3.392$$

$$\bar{X} \sim N(3.5, 0.04). \text{ So the power function at } \mu = 3.5 \text{ is } 1 - \Phi(\frac{3.392 - 3.5}{0.04}) = 0.9965$$

2

(i)

$$\lambda(x) = \frac{f_1(x)}{f_0(x)} = (\frac{\beta_2}{\beta_1})^{n\alpha} exp\{n(\beta_1 - \beta_2)\bar{x}\}$$

$$\textit{It is plain to see that } \lambda(x) \textit{ increases as } -\bar{x} \textit{ increases, so we let } T = -\bar{X}$$

$\lambda(x)$ increases as $-\bar{x}$ increases. Set T $= -\bar{X}$

$$\phi(x) = \begin{cases} 1, \ T > c \\ 0, \ T \leq c \end{cases}$$

(ii)

$$\text{the MGF of gamma distribution is: } M(t) = (1 - \frac{t}{\beta})^{-\alpha}$$

the sum of n i.i.d. variables in the form of Gamma distribution:

$$M(t) = (1 - \frac{t}{\beta})^{-n\alpha} \sim Gamma(n\alpha, \beta)$$

(iii) using CLT we can have:

$$\frac{\beta_1\sqrt{n}(T+\frac{\alpha}{\beta_1})}{\sqrt{\alpha}}$$

is approximately distributed as standard normal distribution

$$\text{when} z=\frac{\beta_1\sqrt{n}(T+\frac{\alpha}{\beta_1})}{\sqrt{\alpha}}>u_\alpha \text{we reject the null hypothesis}$$

The power function is:

$$\pi(\theta)=\Phi(\frac{\beta_1\sqrt{n}(T+\frac{\alpha}{\beta_1})}{\sqrt{\alpha}})$$

3

$$(i) T=\sum_{i=1}^n X_i \text{test function of UMPT is } \phi(x)=\begin{cases}1,\ T>C\\ r,\ T=C\\ 0,\ T<C\end{cases}$$

(ii)

$$Let P_\theta(T>C)+rP_\theta(T=C)=\alpha$$
$$\text{so we have: } C=5,\ r=0.52$$

(iii) Rcode:

```
### power at 0.375
1-pbinom(5,size=10,prob=0.375)+0.519*(pbinom(5,size=10,prob=0.375)-pbinom(4,size=10,prob=0.375))
```

```
## [1] 0.2200042
```

```
### power at 0.5
1-pbinom(5,size=10,prob=0.5)+0.519*(pbinom(5,size=10,prob=0.5)-pbinom(4,size=10,prob=0.5))
```

```
## [1] 0.5046758
```

(iv)

$$For\ \theta>0.5,\ \pi(\theta)=P_\theta(X>C)+rP_\theta(X=C)=P_{1-\theta}(X\le n-C-1)+rP_{1-\theta}(X=n-c) Since\ P_{1-\theta}(X=n-C)=P_\theta($$

And $P_{1-\theta}(X\le n-C)-P_{1-\theta}(X\le n-C+1)=P_{1-\theta}(X\le n-C+1)-P_{1-\theta}(X\le n-C+1)+P_{1-\theta}(X=n-C)=P_{1-\theta}(X=n-$

(v)

$$\pi(0.625)=0.787\ \ \pi(0.875)=0.998$$

(vi)

$$n=62 \text{ by using CLT}$$

4

$$(i)\ \ C(\theta)=\frac{1}{\theta},\ Q(\theta)=-\frac{1}{\theta}$$

supporting set does not rely on θ so it's exponential family

(ii)

$$when\ \sum_{i=1}^n x_i<C,\ where\ P_{\theta_0}(\sum_{i=1}^n x_i<C)=\alpha$$

we reject the null hypothesis

(iii)

$$M_Y(t) = \frac{1}{(1-2t)^{\frac{2n}{2}}} Y \sim \chi^2_{2n}$$

(iv)

$$(iv) \;\; C = \frac{\theta_0}{2}\chi^2_{2n;1-\alpha} \quad \pi(\theta) = P(Y < \frac{2C}{\theta})$$

(v)

$$n = 23$$

5

(i)

$$Let\ Z_i = Y_i - X_i, \text{ so } Z_i \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$Z_i \text{ is distributed as normal and we use the exponential family's property:}$$

$$T = \sum_{i=1}^{4} Z_i\ as\ the\ test\ statistics,\ and\ \phi(z) = \begin{cases} 1, & C_1 < T < C_2 \\ 0, & otherwise \end{cases}$$

to determine C_1, C_2

$$\alpha = 0.05 = \Phi(\frac{C_2 - n\theta}{\sqrt{n(\sigma_1^2+\sigma_2^2)}}) - \Phi(\frac{C_1 - n\theta}{\sqrt{n(\sigma_1^2+\sigma_2^2)}})\theta = 1\ and\ -1C_1 = -0.6793,\ C_2 = 0.6793$$

we do not reject null hypothesis

6

$$Let\ Z_i = Y_i - X_i, \text{ suppose } Z \sim N(\mu, \sigma), \text{ the test is : } H_0 : \mu = 0\ H_1 : \mu \neq 0$$

the reject region:

$$|T_z| = |\frac{\sqrt{n}\bar{Z}}{S_z}| > t_{n-1}(\alpha/2)$$

$$|T| = 0.4159 < t_7(0.025) = 2.365 \text{accept } H_0$$

7

(i) Rcode
(ii) Rcode

```r
data<-read.csv("data.csv")
ols<-lm(SalePrice~ Gr_Liv_Area+Central_Air,data=data)
summary(ols)
```

```
##
## Call:
## lm(formula = SalePrice ~ Gr_Liv_Area + Central_Air, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -200275  -30225   -3438   22292  328643
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -37583.819   5837.338  -6.439 1.59e-10 ***
## Gr_Liv_Area     113.736      2.747  41.402  < 2e-16 ***
## Central_AirY  51726.059   5031.497  10.280  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 52940 on 1595 degrees of freedom
## Multiple R-squared:  0.5569, Adjusted R-squared:  0.5564
## F-statistic:  1002 on 2 and 1595 DF,  p-value: < 2.2e-16
```

p value are smaller than 0.01. so we reject null hypothesis in (i) and (ii)

(iii) use the previous bootstrapping result, the 95% confidence interval is $[2.612 \times 10^9, 3.001 \times 10^9]$, σ so we do not reject H_0