实验物理中的统计方法 作业12

1.

习题 9.2. 随机变量 x 服从均值为 ξ 的指数分布，考虑对 x 的 n 次观测。参数 ξ 的最大似然估计量 (见 *Statistical Data Analysis (6.6)* 式) 由下式给出

$$\hat{\xi} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{9.2}$$

并且 $\hat{\xi}$ 的概率密度函数（参见 *Statistical Data Analysis (10.25)* 式）为

$$g(\hat{\xi};\xi) = \frac{n^n}{(n-1)!}\frac{\hat{\xi}^{n-1}}{\xi^n}e^{-n\hat{\xi}/\xi}. \tag{9.3}$$

(a) 证明：定义置信带的曲线 $u_\alpha(\xi)$ 和 $v_\beta(\xi)$ 为

$$\begin{aligned} u_\alpha(\xi) &= \frac{\xi}{2n}F_{\chi^2}^{-1}(1-\alpha;2n), \\ v_\beta(\xi) &= \frac{\xi}{2n}F_{\chi^2}^{-1}(\beta;2n), \end{aligned} \tag{9.4}$$

其中 $F_{\chi^2}^{-1}$ 为 χ^2 分布的分位数。根据习题2.6，χ^2 分布的累积分布可以与不完全伽马函数 $P(x,n)$ 联系起来：

$$F_{\chi^2}(2x;2n) = P(x,n) \equiv \frac{1}{\Gamma(n)}\int_0^x e^{-t}t^{n-1}dt. \tag{9.5}$$

取 $\alpha = \beta = 0.159$，$n = 5$，画出 $u_\alpha(\xi)$ 和 $v_\beta(\xi)$。*(* χ^2 分布的分位数可以从标准分布表中查出，或者在 *ROOT* 中调用 `TMath::ChisquareQuantile(Double_t p, Double_t ndf)` 函数得到。*)*

(b) 求出置信区间 $[a,b]$ 作为估计值 $\hat{\xi}$、样本容量 n 以及置信水平 α 和 β 的函数。假设估计值为 $\hat{\xi} = 1.0$，在 $u_\alpha(\xi)$ 和 $v_\beta(\xi)$ 的图上画出该估计量的值。取 $n = 5$，$\alpha = \beta = 0.159$，计算 a 和 b。将计算结果与估计值加减一倍标准差得到的区间进行比较。

解:

(9.2)

(a)
$$\alpha = \int_{u_\alpha}^{\infty} \frac{n^n}{(n-1)!} \frac{\hat{\xi}^{n-1}}{\xi^n} e^{-n\hat{\xi}/\xi} d\hat{\xi}$$

$$1-\alpha = \frac{1}{(n-1)!} \int_0^{u_\alpha} e^{-n\hat{\xi}/\xi} \left(\frac{n\hat{\xi}}{\xi}\right)^{n-1} d\left(\frac{n\hat{\xi}}{\xi}\right)$$

$$= \frac{1}{\Gamma(n)} \int_0^{\frac{n u_\alpha}{\xi}} e^{-t} t^{n-1} dt$$

$$= F_{\chi^2}\left(\frac{2n u_\alpha}{\xi}; 2n\right)$$

$$\Rightarrow u_\alpha = \frac{\xi}{2n} F_{\chi^2}^{-1}(1-\alpha; 2n)$$

$$\beta = \int_0^{u_\beta} \frac{n^n}{(n-1)!} \frac{\hat{\xi}^{n-1}}{\xi^n} e^{-n\hat{\xi}/\xi} d\hat{\xi}$$

同理，$u_\beta = \frac{\xi}{2n} F_{\chi^2}^{-1}(\beta; 2n)$

画图附后.

(b) $u_\alpha(a) = \hat{\xi} \Rightarrow a = 2n\hat{\xi} / F_{\chi^2}^{-1}(1-\alpha; 2n)$

$u_\beta(b) = \hat{\xi} \Rightarrow b = 2n\hat{\xi} / F_{\chi^2}^{-1}(\beta; 2n)$

置信区间: $\left[2n\hat{\xi} / F_{\chi^2}^{-1}(1-\alpha; 2n),\ 2n\hat{\xi} / F_{\chi^2}^{-1}(\beta; 2n)\right]$

画图与计算附后

```
from scipy.stats import chi2
import matplotlib.pyplot as plt
import numpy as np

n = 5
xi_hat = 1.0
alpha = 0.159
```

```python
beta = 0.159

def u_alpha(x):
    return x/(2*n)*chi2.ppf(1-alpha, 10)

def u_beta(x):
    return x/(2*n)*chi2.ppf(beta, 10)

x = np.linspace(0.01, 2, 1000)

plt.plot(x, u_alpha(x), label = 'u_alpha(xi)', color = 'blue')
plt.plot(x, u_beta(x), label =  'u_beta(xi)', color = 'orange')
plt.xlabel('xi')
plt.ylabel('u(xi)')
plt.axhline(1, label=f'xi_hat = {xi_hat}', color = 'green')
a = 2*n*xi_hat/chi2.ppf(1-alpha, 10)
b =  2*n*xi_hat/chi2.ppf(beta, 10)
plt.scatter([a, b], [1, 1], color='red', zorder=5)
plt.text(a, 1.02, f'a = {a:.3f}', ha='center', color='red')
plt.text(b, 1.02, f'b = {b:.3f}', ha='center', color='red')
plt.legend()
plt.savefig('9.2.png')
print(f'a = {a}, b = {b}')
```

结果:

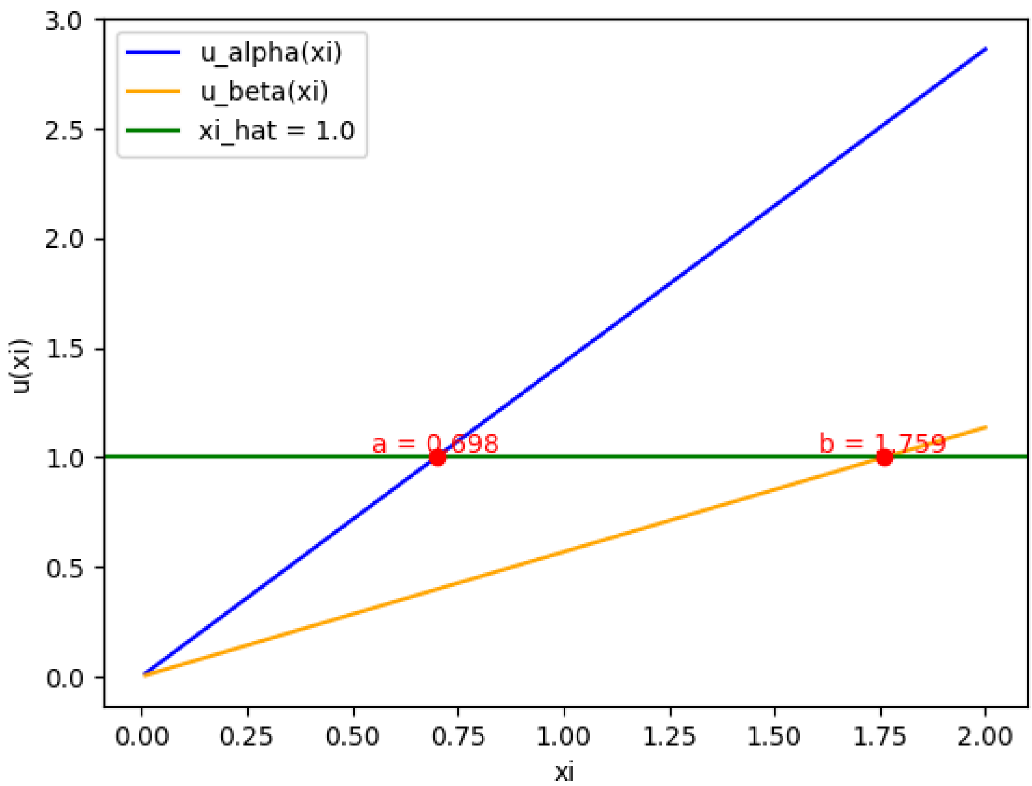

Figure 9.2

结果$[a, b]$在一倍标准差$[0.0, 2.0]$之内。

2.

习题 9.3. 证明二项分布的参数 p 的上限和下限为

$$\begin{aligned} p_{lo} &= \frac{nF_F^{-1}[\alpha; 2n, 2(N-n+1)]}{N-n+1+nF_F^{-1}[\alpha; 2n, 2(N-n+1)]} \\ p_{up} &= \frac{(n+1)F_F^{-1}[1-\beta; 2(n+1), 2(N-n)]}{(N-n)+(n+1)F_F^{-1}[1-\beta; 2(n+1), 2(N-n)]}. \end{aligned} \tag{9.6}$$

其中上下限的置信水平分别为 $1-\alpha$ 和 $1-\beta$，n 为 N 次试验中成功的次数，F_F^{-1} 为 F 分布的分位数，由 F 分布定义:

$$f(x; n_1, n_2) = \left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{\Gamma(\frac{1}{2}(n_1+n_2))}{\Gamma(\frac{1}{2}n_1)\Gamma(\frac{1}{2}n_2)} x^{n_1/2-1} \left(1+\frac{n_1}{n_2}x\right)^{-(n_1+n_2)/2}, \tag{9.7}$$

其中 $x>0$，参数 n_1 和 n_2 为整数 (自由度)。利用二项分布累积分布函数与自由度为 $n_1=2(n+1)$ 和 $n_2=2(N-n)$ 的累积分布函数 $F_F(x)$ 的关系[1]

$$\sum_{k=0}^{n} \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} = 1 - F_F\left[\frac{(N-n)p}{(n+1)(1-p)}; 2(n+1), 2(N-n)\right]. \tag{9.8}$$

F 分布的分位数可以从标准分布表中查得，或者调用 *ROOT* 中的函数计算。

解:

9.3

二项分布:

$$P(n=k;N,p)=\frac{N!}{k!(N-k)!}p^k(1-p)^{N-k}$$

依题:

$$\alpha=\sum_{k=n}^{N}\frac{N!}{k!(N-k)!}p_{lo}^k(1-p_{lo})^k$$

$$\Rightarrow 1-\alpha=\sum_{k=0}^{n-1}\frac{N!}{k!(N-k)!}p_{lo}^k(1-p_{lo})^{N-k}$$

$$=1-F_F\left[\frac{(N-n+1)p_{lo}}{n(1-p_{lo})};2n,2(N-n+1)\right]$$

$$\Rightarrow \frac{(N-n+1)p_{lo}}{n(1-p_{lo})}=F_F^{-1}[1-\alpha;2n,2(N-n+1)]$$

$$\Rightarrow p_{lo}=\frac{nF_F^{-1}[1-\alpha;2n,2(N-n+1)]}{(N-n+1)+nF_F^{-1}[1-\alpha;2n,2(N-n+1)]}$$

同理:

$$\beta=\sum_{k=0}^{n}\frac{N!}{k!(N-k)!}p_{up}^k(1-p_{up})^{N-k}$$

$$=1-F_F\left[\frac{(N-n)p_{up}}{(n+1)(1-p_{up})};2(n+1),2(N-n)\right]$$

$$\Rightarrow p_{up}=\frac{(n+1)F_F^{-1}[1-\beta;2(n+1),2(N-n)]}{N-n+(n+1)F_F^{-1}[1-\beta;2(n+1),2(N-n)]}$$