IDSSMA

zhou ziliang

2019/11/20

IDSSM

1. a. $\because P(y;\pi) = \pi^y(1-\pi)^{1-y}$

$\therefore$ Pdf $f(y) = \begin{cases} \pi & y=1 \\ 1-\pi & y=0. \end{cases}$

$E(Y) = 1\times\pi + 0\times(1-\pi) = \pi$

$Var(Y) = (0-E(Y))^2\cdot P_0 + (1-E(Y))^2\cdot P_1$

$= \pi^2(1-\pi) + (1-\pi)^2\cdot\pi$

$= \pi^2 - \pi^3 + \pi - 2\pi^2 + \pi^3$

$= \pi - \pi^2.$

$\therefore E(Y) = \pi \quad Var(Y) = \pi - \pi^2 = \pi(1-\pi)$

b. according to a, we have

$E(Y) = \pi.$

$\therefore \hat{\pi} = \bar{Y}$

$E(\hat{\pi}) = E(\bar{Y}) = E\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right) = \frac{1}{n}\times E(\sum Y_i) = \frac{n\pi}{n} = \pi.$

So. π is unbiased estimate

c. the method-of-monents estimate of π is $\bar{Y}$

and $MSE(\hat{\pi}) = Var(\hat{\pi}) + Bias(\hat{\pi}, \pi)^2$

Because the $\hat{\pi}$ is unbiased.

$\therefore MSE(\hat{\pi}) = Var(\hat{\pi}) = Var(\bar{Y})$

$= \frac{1}{n^2} \sum_{i=1}^{n} Var(Y_i) = \frac{\pi(1-\pi)}{n}$

$\therefore MSE = \frac{\pi(1-\pi)}{n}$

d. estimator: $Var(Y) = \pi(1-\pi)$

$E[Var(Y)] = E(\pi - \pi^2)$

$= E(\pi) - E(\pi^2)$

$= E(\pi) - [Var(\pi) + E^2(\pi)]$

as ~~above~~ before, we know.

$E(\pi) = \pi$, $Var(\pi) = \frac{\pi(1-\pi)}{n}$

$\therefore E[Var(Y)] = \pi - \frac{\pi(1-\pi)}{n} - \pi^2$

$= \frac{n-1}{n} \cdot \pi(1-\pi) \neq \pi(1-\pi)$

this is ~~not~~ ~~un~~ biased. in tend to make it unbiased.

$\frac{n}{n-1} \cdot E[Var(Y)] = \pi(1-\pi)$.

the unbiased estimator is $\frac{n}{n-1} Var(Y)$.

is $\frac{n}{n-1} \cdot (1-\pi) \cdot \pi = \frac{n}{n-1} \bar{Y}(1-\bar{Y})$

```r
library(tidyverse)
titanic <- read.csv("titanic.csv")
```

```r
titanic <- select(titanic, survived, gender, pclass)

male <- filter(titanic, gender == "male")

female <- filter(titanic, gender == "female")

firstclass.male <- filter(male, pclass == 1)
firstclass.female <- filter(female,pclass == 1)

msur <- sum(male$survived)/length(male$survived);msur
```

```
## [1] 0.1909846
```

```r
fsur <- sum(female$survived)/length(female$survived);fsur
```

```
## [1] 0.7274678
```

```r
firstclass.msur <- sum(firstclass.male$survived)/length(firstclass.male$survived)
firstclass.fsur <- sum(firstclass.female$survived)/length(firstclass.female$survived)
firstclass.fsur
```

```
## [1] 0.9652778
```

```r
firstclass.msur
```

```
## [1] 0.3407821
```

The male survived rate is 0.19,female survived rate is 0.73. In 1st class male survived rate is 0.34, female survived rate is 0.96

2. the posson distribution's pmf: $\frac{\lambda^k e^{-\lambda}}{k!}$

so we get likehood function:

$$L_k(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{k_i} \cdot e^{-\lambda}}{k_i!}$$

$$\ln L_k(\lambda) = \sum_{i=1}^{n} \ln \lambda^{k_i} + \sum_{i=1}^{n} \ln e^{-\lambda} - \sum_{i=1}^{n} \ln(k_i!)$$

$$\ln L_k(\lambda) = \ln\lambda \cdot \sum_{i=1}^{n} k_i + (0 - n\lambda) - \sum_{i=1}^{n} \ln(k_i!)$$

$$\left(\ln L_k(\lambda)\right)' = \frac{\sum_{i=1}^{n} k_i}{\lambda} - n$$

Let $\left(\ln L_k(\lambda)\right)' = 0$.

$$\lambda = \frac{\sum_{i=1}^{n} k_i}{n} = \overline{k}$$

$\therefore$ the maximum likehood estimate of λ is

$$\lambda = \overline{k} = 1.5$$