

# exercise2

巫小珍

## 目录

<b>1</b>	<b>exercise2</b>	<b>1</b>
1.1	Q1:kenya、sweden 与 world 的 CBR 粗出生率 . . . . .	2
1.2	Q2:kenya、sweden 与 world 的年龄生育率 ASFR . . . . .	3
1.3	Q3:kenya、sweden 与 world 的妇女生育平均子女数 TFR . . . .	8
1.4	Q4:kenya、sweden 与 world 的粗出生率 CDR . . . . .	10
1.5	Q5:2005-2010 年 kenya、sweden 的各年龄组的死亡率 ASDR	11
1.6	Q6:2005-2010 年 kenya 的反事实 CDR . . . . .	13

## 1 exercise2

导入数据

```
kenya <- read.csv("../Datasets/Kenya.csv") # 读取肯尼亚的表格数据
```

```
sweden <- read.csv("../Datasets/Sweden.csv") # 读取瑞典的表格数据
```

```
world <- data.frame(read.csv("../Datasets/World.csv")) # 读取世界的表格数据
```

```
summary(kenya)
```

```
summary(sweden)
```

```
summary(world)
```

```
> summary(kenya)
  country      period      age      births      deaths      py.men      py.women
Length:30    Length:30    Length:30    Min. : 0.0    Min. : 18.51    Min. : 39.85    Min. : 82.24
Class :character    Class :character    Class :character    1st Qu.: 0.0    1st Qu.: 21.24    1st Qu.: 840.92    1st Qu.: 733.89
Mode :character      Mode :character      Mode :character      Median : 0.0    Median : 55.13    Median : 1712.56    Median : 1800.20
                                Mean : 298.3    Mean : 91.06    Mean : 3681.89    Mean : 3684.72
                                3rd Qu.: 267.5    3rd Qu.:103.75    3rd Qu.: 4506.66    3rd Qu.: 4417.88
                                Max. :2285.7    Max. :661.25    Max. :15932.59    Max. :15674.83

  l_x
Min. : 9457
1st Qu.: 53537
Median : 66138
Mean : 64688
3rd Qu.: 81411
Max. :100000
> summary(sweden)
  country      period      age      births      deaths      py.men      py.women
Length:30    Length:30    Length:30    Min. : 0.00    Min. : 0.191    Min. : 250.7    Min. : 320.6
Class :character    Class :character    Class :character    1st Qu.: 0.00    1st Qu.: 1.522    1st Qu.:1217.2    1st Qu.:1230.0
Mode :character      Mode :character      Mode :character      Median : 0.00    Median : 4.157    Median :1383.4    Median :1358.1
                                Mean : 36.48    Mean : 26.908    Mean :1350.0    Mean :1364.1
                                3rd Qu.: 58.80    3rd Qu.: 14.406    3rd Qu.:1508.7    3rd Qu.:1499.5
                                Max. :193.19    Max. :271.644    Max. :2612.9    Max. :2635.8

  l_x
Min. : 35709
1st Qu.: 93116
Median : 96880
Mean : 91839
3rd Qu.: 98904
Max. :100000
> summary(world)
  country      period      age      births      deaths      py.men      py.women
Length:30    Length:30    Length:30    Min. : 0    Min. : 3248    Min. : 30527    Min. : 47262
Class :character    Class :character    Class :character    1st Qu.: 0    1st Qu.: 6263    1st Qu.: 380398    1st Qu.: 388215
Mode :character      Mode :character      Mode :character      Median : 0    Median : 8104    Median : 673948    Median : 665470
                                Mean : 38782    Mean : 17517    Mean : 778510    Mean : 770349
                                3rd Qu.: 55748    3rd Qu.: 14409    3rd Qu.:1184047    3rd Qu.:1153630
                                Max. :219277    Max. :101090    Max. :1619802    Max. :1512274
```

## 1.1 Q1:kenya、sweden 与 world 的 CBR 粗出生率

### 1.1.1 计算方法

CBR 的计算公式为：

CBR = 出生总人数 / 总寿命数（人-年）

需要定义一个函数用于计算 1950-1955 年和 2005-2010 年的 CBR。

### 1.1.2 代码实现

定义函数

```
CBRCalculate <- function(x){
# 计算 1950-1955 年和 2005-2010 年的 CBR

# Args:

# x: data.frame

# Returns:
```

```

# 返回 1950-1955 年和 2005-2010 年的 CBR 数值

x1 <- x$py.men[1:15] + x$py.women[1:15] # 计算 1950-1955 年
的寿命数 (人-年)

x1CBR <- sum(x$births[1:15]) / sum(x1) # 计算 1950-1955 年的
粗出生率

x2 <- x$py.men[16:30] + x$py.women[16:30] # 计算 2005-2010 年
的寿命数 (人-年)

x2CBR <- sum(x$births[16:30]) / sum(x2) # 计算 2005-2010 年
的粗出生率

xCBR <- c(x1CBR, x2CBR)

return(xCBR)
}

调用函数

kenCBR <- CBRCalculate(kenya)

sweCBR <- CBRCalculate(sweden)

worCBR <- CBRCalculate(world)

```

### 1.1.3 结果

```

> CBR
      year      kenCBR      sweCBR      worCBR
1 1950-1955 0.05209490 0.01539614 0.03732863
2 2005-2010 0.03851507 0.01192554 0.02021593

```

## 1.2 Q2:kenya、sweden 与 world 的年龄生育率 ASFR

### 1.2.1 计算方法

ASFR 的计算公式:

$ASFR[x, x+ ) = [x, x+ ) \text{ 年龄段女性人数} / [x, x+ ) \text{ 年龄段女性寿命数 (人-年)}$

定义一个函数用于计算 ASFR, 定义一个函数用于画两条数据的折线图。

### 1.2.2 代码实现

定义计算 ASFR 函数

```
ASFRCalculate <- function(x){  
  # 计算 ASFR  
  # Args:  
  #   x: data.frame  
  # Returns:  
  #   返回 numeric 表示 ASFR  
  #  
  return(x$births/ x$py.women)  
}
```

定义画图函数

```
PlotDraw <- function(num1, num2, num3, y1, y2, cha1, cha2,cha3,  
  cha4){  
  # 画两条折线图  
  # Args:  
  #   num1: numeric , 表示 x 轴数据的个数  
  #   num2、num3:y 轴取值范围  
  #   y1:vector, y1 的取值  
  #   y2:vector, y2 的取值  
  #   cha1:character, y1 的名称  
  #   cha2:character, y2 的名称  
  #   cha3:character, y 轴的名称  
  #   cha4:character, 图名称  
  # Returns:
```

# 没有

```
ID <- c(1:num1)

my_frame <- data.frame(ID, y1, y2)

plot(y1~ID, pch=15, col="DarkTurquoise", ylim=c(num2, num3),
     ylab=cha3, main = cha4)

points(ID, y2, pch=16, col="DeepPink", cex=1)

lines(y1, col="DarkTurquoise", lty=1)

lines(y2, col="DeepPink", lty=2)

legend("topleft", c(cha1, cha2), col=c("DarkTurquoise", "DeepPink"),
     text.col=c("DarkTurquoise","DeepPink"), pch=c(15,16),lty=c(1,2))

}
```

函数调用

```
kenASFR <- ASFRCalculate(kenya) # 肯尼亚的 ASFR

sweASFR <- ASFRCalculate(sweden) # 瑞典的 ASFR

worASFR <- ASFRCalculate(world) # 世界的 ASFR

PlotDraw(7, 0, 0.4, kenASFR[4:10], kenASFR[19:25], "1950-1955",
     "2005-2010", "ASFR", "kenASFR" )

PlotDraw(7, 0, 0.15, sweASFR[4:10], sweASFR[19:25], "1950-1955",
     "2005-2010", "ASFR", "sweASFR" )

PlotDraw(7, 0, 0.3, worASFR[4:10], worASFR[19:25], "1950-1955",
     "2005-2010", "ASFR", "worASFR" )
```

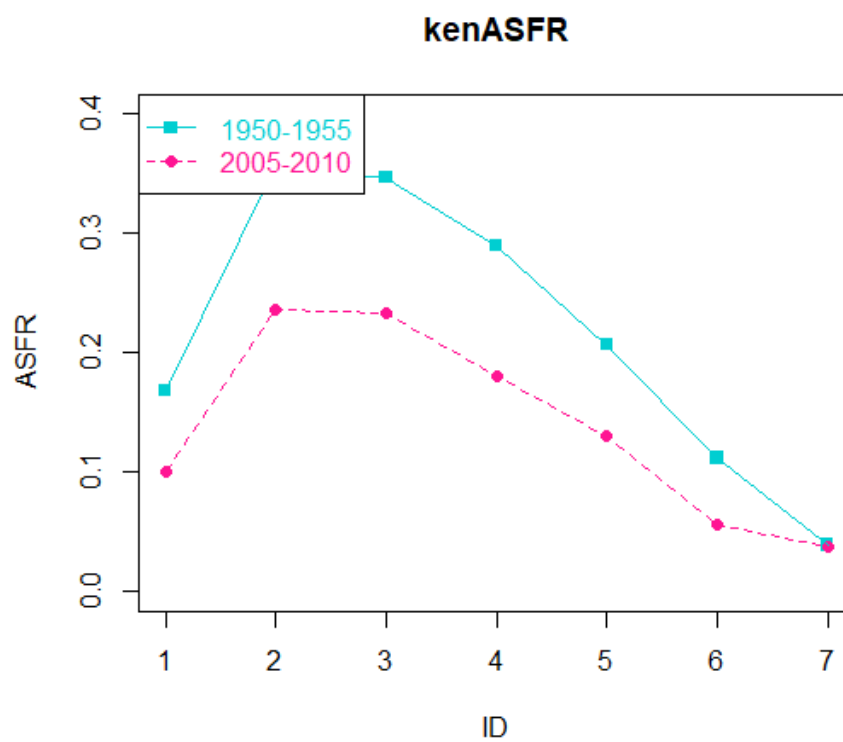
## 1.2.3 结果

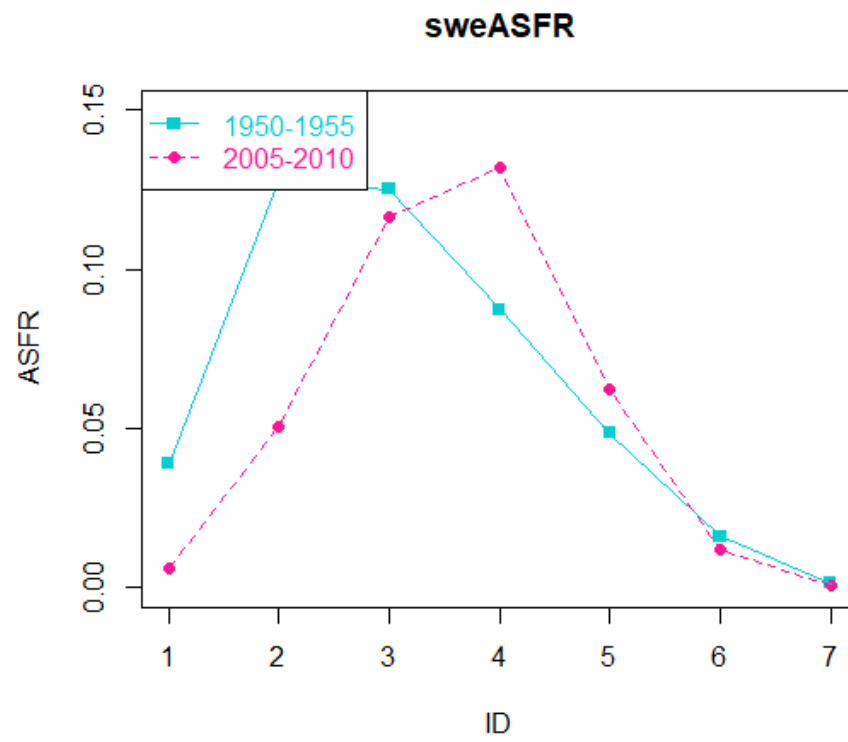
```

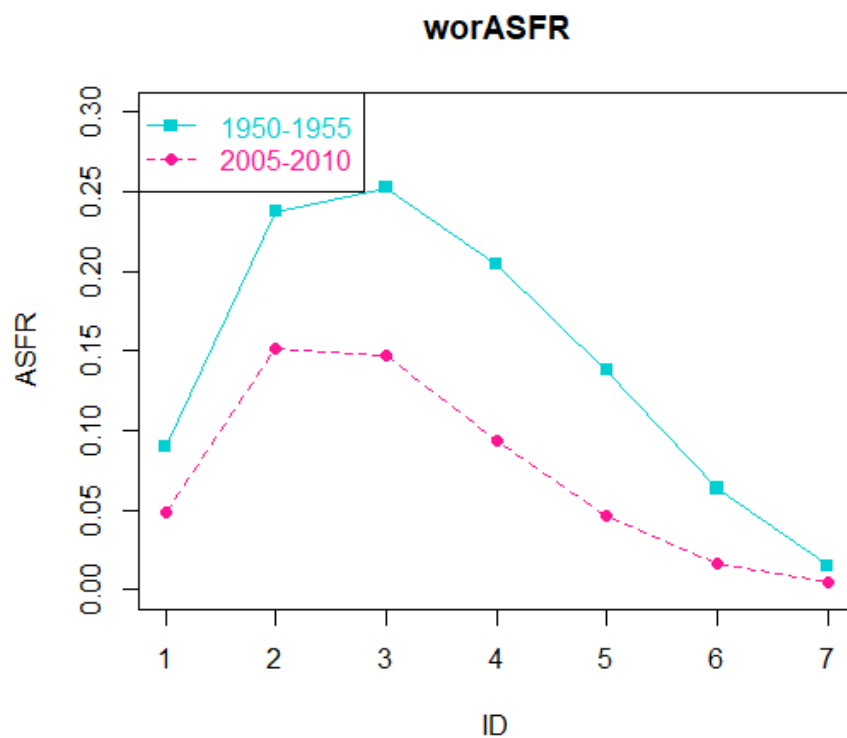
$`1950-1955年ASFR`
age      kenya      sweden      world
1 15-19 0.16884585 0.038908952 0.09029521
2 20-24 0.35596942 0.127710883 0.23763370
3 25-29 0.34657814 0.125243665 0.25245229
4 30-34 0.28946367 0.087364159 0.20416410
5 35-39 0.20644016 0.048603771 0.13810534
6 40-44 0.11193267 0.016210186 0.06360832
7 45-49 0.03905205 0.001341829 0.01519064

$`2005-2010年ASFR`
age      kenya      sweden      world
1 15-19 0.10057087 0.0059709097 0.048489719
2 20-24 0.23583536 0.0507320271 0.151971307
3 25-29 0.23294721 0.1162085625 0.146980966
4 30-34 0.18087964 0.1322744621 0.093813813
5 35-39 0.13126805 0.0625923991 0.046689639
6 40-44 0.05626214 0.0121600765 0.016268995
7 45-49 0.03815044 0.0006143942 0.004510245

```







### 1.3 Q3:kenya、sweden 与 world 的妇女生育平均子女数 TFR

#### 1.3.1 计算方法

ASFR 的计算公式:

$$\text{TFR} = \text{ASFR}[15, 20) * 5 + \text{ASFR}[20, 25) * 5 + \text{ASFR}[25, 30) * 5 + \text{ASFR}[30, 35) * 5 + \text{ASFR}[35, 40) * 5 + \text{ASFR}[40, 45) * 5 + \text{ASFR}[45, 50) * 5$$

定义一个函数用于计算 TFR。

计算全世界妇女数量，1950-1955 年、2005-2010 年全世界总出生人数。



### 1.3.2 代码实现

定义计算 TFR 的函数

```
TFRCalculate <- function(x){
# 计算 TFR
# Args:
#   x: data.frame
# Returns:
#   返回 1950-1955 年和 2005-2010 年的 TFR 数值
xTFR <- c(sum(x[4:10] * 5), sum(x[19:25] * 5))
return(xTFR)
}
```

函数调用计算 kenya、sweden、world 的 TFR

```
kenTFR <- TFRCalculate(kenASFR)
sweTFR <- TFRCalculate(sweASFR)
worTFR <- TFRCalculate(worASFR)
```

全世界女性数量的变化

```
worWomenChange = sum(world$py.women[16:30]) - sum(world$py.women[1:15])
```

全世界出生总人数

```
sum(world$births[1:15]) # 1950-1955 年
sum(world$births[16:30]) # 2005-2010 年
```

### 1.3.3 结果

	year	kenTFR	sweTFR	worTFR
1	1950-1955	7.591410	2.226917	5.007248
2	2005-2010	4.879568	1.902764	2.543623

```

> worWomenChange # 增加了9999096 (以千计)
[1] 9999096

> sum(world$births[1:15]) # 1950-1955年出生总人数为488891.5 (以千计)
[1] 488891.5
>
> sum(world$births[16:30]) # 2005-2010年出生总人数为674581.3 (以千计)
图 1: [1] 674581.3

```

## 1.4 Q4:kenya、sweden 与 world 的粗出生率 CDR

### 1.4.1 计算方法

CDR 的计算公式为:

$CDR = \text{死亡总人数} / \text{总寿命数 (人-年)}$

定义一个计算 CDR 的函数，并调用函数进行计算。

### 1.4.2 代码实现

定义计算 CDR 的函数

```

CDRCalculate <- function(x){
# 计算 CDR
# Args:
#   x: data.frame
# Returns:
#   返回 1950-1955 年和 2005-2010 年的 CDR 数值
x1 <- x$py.men[1:15] + x$py.women[1:15] # 计算 1950-1955 年
的寿命数 (人-年)
x1CDR <- sum(x$deaths[1:15]) / sum(x1) # 计算 1950-1955 年的
粗死亡率
x2 <- x$py.men[16:30] + x$py.women[16:30] # 计算 2005-2010 年
的寿命数 (人-年)

```

```
x2CDR <- sum(x$deaths[16:30]) / sum(x2)  # 计算 2005-2010 年
的粗死亡率

xCDR <- c(x1CDR, x2CDR)

return(xCDR)

}
```

调用函数计算 kenya、sweden 和 world 的 CDR

```
kenCDR <- CDRCalculate(kenya)

sweCDR <- CDRCalculate(sweden)

worCDR <- CDRCalculate(world)
```

### 1.4.3 结果

	year	kenCDR	sweCDR	worCDR
1	1950-1955	0.02396254	0.009844842	0.019318929
2	2005-2010	0.01038914	0.009968455	0.008166083

## 1.5 Q5:2005-2010 年 kenya、sweden 的各年龄组的死亡率 ASDR

### 1.5.1 计算方法

ASDR 的计算公式为：

$ASDR[x, x+] = [x, x+] \text{ 年龄段死亡人数} / [x, x+] \text{ 年龄段寿命数 (人-年)}$

定义一个函数计算 2005-2010 年的 ASDR。

调用 Q2 定义的画图函数进行画图。

### 1.5.2 代码实现

定义计算 2005-2010 年 ASDR 的函数

```
ASDRCalculate <- function(x){

# 计算 ASDR
```

```

# Args:

#   x: data.frame

# Returns:

#   返回 2005-2010 年的 ASDR 数值

x2 <- x$py.men[16:30] + x$py.women[16:30] # 计算 2005-2010 年
的寿命数 (人-年)

xASDR <- x$deaths[16:30] / x2 # 计算 2005-2010 年的 ASDR

return(xASDR)

}

```

函数调用计算 2005-2010 年 kenya 和 sweden 的 ASDR

```

kenASDR <- ASDRCalculate(kenya)

sweASDR <- ASDRCalculate(sweden)

```

调用画图函数画图

```

PlotDraw(15, 0, 0.18, kenASDR, sweASDR, "kenASDR", "sweASDR",
"ASDR", "ken-swe-ASDR" )

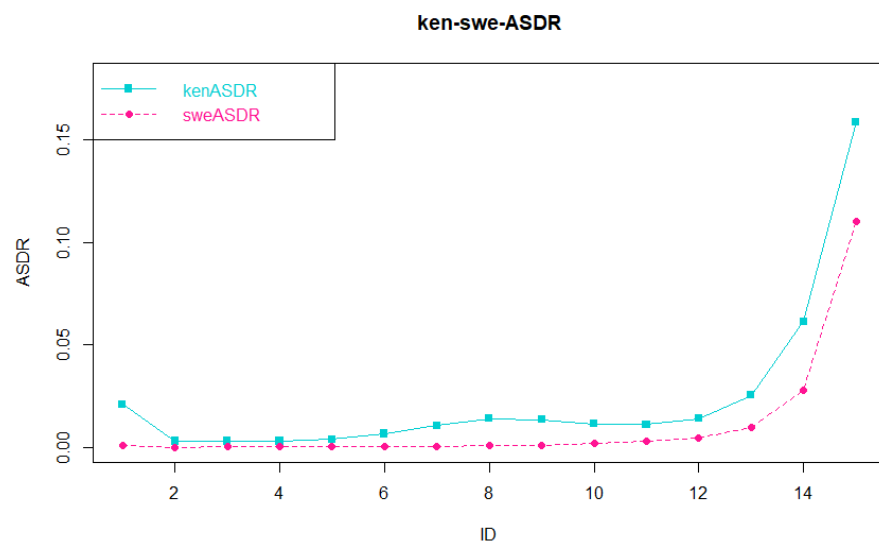
```

### 1.5.3 结果

```

      kenASDR      sweASDR
1 0.020920755 6.790712e-04
2 0.002911301 8.138094e-05
3 0.002918895 1.135496e-04
4 0.002942986 2.687775e-04
5 0.003885368 4.697344e-04
6 0.006558131 4.941440e-04
7 0.010603913 5.057066e-04
8 0.013881062 6.689578e-04
9 0.013474598 1.039256e-03
10 0.011288057 1.769621e-03
11 0.011152339 2.988715e-03
12 0.013898334 4.709913e-03
13 0.023395531 9.828772e-03
14 0.061261551 2.803963e-02
15 0.158620510 1.098892e-01

```



## 1.6 Q6:2005-2010 年 kenya 的反事实 CDR

### 1.6.1 计算方法

用瑞典的人口分布比例计算肯尼亚的反事实。

$[x, x + )$  年龄段的人口分布比例的计算公式为：

$$P[x, x + ) = [x, x + ) \text{ 年龄段人数} / \text{总人数}$$

反事实 CDR 的计算公式为：

$$\begin{aligned} \text{CDR} = & \text{ASDR}[0, 5) * P[0, 5) + \text{ASDR}[5, 10) * P[5, 10) + \text{ASDR}[10, 15) \\ & * P[10, 15) + \text{ASDR}[15, 20) * P[15, 20) + \dots \end{aligned}$$

定义一个函数计算 2005-2010 年的人口分布比例。

通过反事实 CDR 公式计算肯尼亚的反事实。

### 1.6.2 代码实现

定义计算人口分布比例的函数

```

PCalculate <- function(x){
# 计算 2005-2010 年的 P 人口分布比例
# Args:
#   x: data.frame
# Returns:
#   返回 2005-2010 年的 P 数值
P <- (x$py.men[16:30] + x$py.women[16:30]) / sum(x$py.men[16:30]
+ x$py.women[16:30])
return(P)
}

```

调用函数计算

```
sweP <- PCalculate(sweden)
```

```
kenP <- PCalculate(kenya)
```

肯尼亚的反事实 CDR

```
kenSweCDR <- sum(kenASDR * sweP)
```

### 1.6.3 结果

	age	sweP	kenP
1	0-4	0.05787711	0.167329267
2	5-9	0.05121759	0.139518775
3	10-14	0.06149968	0.119952403
4	15-19	0.06755209	0.112321935
5	20-24	0.06123102	0.102701807
6	25-29	0.05979620	0.085055681
7	30-34	0.06485878	0.066238790
8	35-39	0.07030027	0.050043424
9	40-44	0.07005059	0.038855296
10	45-49	0.06548210	0.031800814
11	50-54	0.06272159	0.026205201
12	55-59	0.06702610	0.019889495
13	60-69	0.11454018	0.023300771
14	70-79	0.07190144	0.012919992
15	80+	0.05394527	0.003866347

```

> kenSweCDR
[1] 0.02321646

```