**实验报告1.1-129**

**学号：** 117060400129 **姓名**： 李娟 **班级：**应用统计学一班  **指导老师：** 林卫中

**实验名称**： 习题练习

**实验要求：** 熟悉程序并加以习题练习

**实验题目：习题6.4-6.6**

**算法实现：6.4**

from operator import itemgetter

def analyseText(text):

D = {}

tlist = list(text)

for e in tlist:

if e in D:

D[e] = D[e] + 1

else:

D[e] = 1

return D

text=input("请输入一段文本")

D = analyseText(text)

D = sorted(D.items(),key=itemgetter(1), reverse=True)

for d in D:

ch, count = d[0],d[1]

print("{0:<10}{1:>5}".format(ch,count))

**6.5**

def generateSamples1(n:int):

birthdays = []

days = [31,29,31,30,31,30,31,31,30,31,30,31]

for i in range(n):

month = randint(1,12)

day = randint(1,days[month-1])

someday = (month,day)

birthdays.append(someday)

return birthdays

def generateSamples2(n:int):

birthdays = []

days = [31,28,31,30,31,30,31,31,30,31,30,31]

year = randint(1950,2000)

for i in range(n):

month = randint(1,12)

if (year%400==0) and (year%4==0 and year%100 != 0):

days[1] = 29

else:

days[1] = 28

day = randint(1,days[month-1])

someday = (month,day)

birthdays.append(someday)

return birthdays

def calSameBirthdayProb(birthdays:list, n:int):

num = 0

for i in range(n):

people = sample(birthdays,23)

pset = set(people)

if len(pset) != len(people):

num += 1

return num/n

def main():

while True:

n = int(input("输入一个整数："))

if n < 23:

break

birthdays = generateSamples2(n)

print("{}个随机样本数量下，23个人中至少有两人生日相同的概率是：{}".format(n, calSameBirthdayProb(birthdays, 100000)))

main()

**6.6**

import jieba

excludes = {'什么','一个','我们','你们','如今','说道','老太太','知道','姑娘','起来',

'这里','出来','众人','那里','奶奶','自己','太太','一面','只见','两个',

'没有','怎么','不是','这个','听见','这样','进来','咱们','就是','不知',

'东西','告诉','回来','只是','大家','老爷','只得','丫头','这些','他们',

'不敢','出去','所以','不过','不好','姐姐','的话','一时','过来'

}

txt = open('红楼梦.txt','r', encoding='GB18030').read()

words = jieba.lcut\_for\_search(txt)

counts = {}

for word in words:

if len(word) == 1:

continue

elif word == '凤姐' or word == '熙凤':

rword = '凤姐'

elif word == '元春' or word == '贵妃':

rword = '元春'

else:

rword = word

counts[rword] = counts.get(rword,0) + 1

for word in excludes:

del(counts[word])

items = list(counts.items())

items.sort(key=lambda x: x[1], reverse = True)

for i in range(50):

word, count = items[i]

print("{0:<10}{1:>5}".format(word, count))

**实验结果：**