

گزارش سوال اول تمرین کامپیوتری دوم آمار و احتمال مهندسی

امیرمرتضی رضائی – 810101429

ابتدا با استفاده از تابع `random.normal` موجود در کتابخانه `numpy`، 1000 عدد با توزیع نرمال و `p` و `std` مطلوب تولید کرده و در آرایه‌ی `the_array` که از قبل آنرا ایجاد کرده‌ایم ذخیره می‌کنیم.

```
p = 5
std = 10
the_array = []
the_array = np.random.normal(p, std, 1000)
```

حال نیاز است تا 10000 بار از این مجموعه نمونه‌ی 5 عضوی بگیریم. برای تکرار 10000 باره‌ی `sample` گیری، از حلقه‌ی `for` بهره می‌گیریم. اکنون برای انتخاب 5 نمونه‌ی تصادفی از مجموعه‌مان، در هر مرحله با استفاده از تابع `randint`، 5 عدد تصادفی بین 0 و 999 ایجاد می‌کنیم. در واقع این 5 عدد تولید شده را به عنوان `index` هایی از مجموعه `the_array` در نظر گرفته و عدد موجود در هر یک از آن `index` ها، یکی از آن پنج `sample` ما خواهد بود. پس از تولید `index` ها، یک متغیر به نام `sum_five` و یک متغیر با نام `average_five` با مقدار اولیه صفر ایجاد کرده و در 5 بار تکرار، هر یک از `sample` ها را استخراج کرده و با مقدار قبلی متغیر `sum_five` جمع می‌کنیم. در نهایت نیز مقدار میانگین این 5 عدد تولید شده را که در متغیر `average_five` ذخیره کرده‌ایم، به آرایه‌ی `list_five` که در ابتدا ایجاد کرده‌ایم اضافه می‌کنیم.

```
list_five = []
for i in range(0, 10000):
    random_indexes_five = [random.randint(0, 999) for _ in range(5)]
    sum_five = 0
    average_five = 0
    for j in range(0, 5):
        sum_five = sum_five + the_array[random_indexes_five[j]]
    average_five = sum_five / 5
    list_five.append(average_five)
```

حال پس از 10000 بار نمونه‌گیری 5 تایی، آرایه‌ی `list_five` حاوی 10000 درایه است که هر یک از آنها در واقع برابر با میانگین هر گروه از نمونه‌گیری‌های 5 تایی هستند. در نهایت میانگین درایه‌های این آرایه را محاسبه کرده و آنرا نمایش می‌دهیم:

```
total_average_five = np.mean(list_five)
print("average for sample size 5 : ", total_average_five)
```

همین روند را برای نمونه‌گیری 10 و 20 و 40 عضوی پیش می‌گیریم.

نتایج بدست آمده به صورت زیر می باشد:

average for sample size 5 : 5.211919266078634

average for sample size 10 : 5.256232133030447

average for sample size 20 : 5.277174905741612

average for sample size 40 : 5.268064443220075

باتوجه به قضیه ی حد مرکزی می دانیم اگر توزیع اولیه نرمال باشد، هر نمونه گیری خود یک توزیع نرمال با میانگین نمونه ی اولیه می باشد. در اینجا نیز مشاهده می شود میانگین نمونه ها حدودا برابر با 5 است. دلیل اختلاف اندک نیز آن است که امکان دارد اعضای یکی از نمونه ها با یکی دیگر از نمونه ها برابر باشد و برخی داده ها را چند بار حساب کرده باشیم.