

1- فرض کنید $X_1, \dots, X_n \sim P(4)$.

- الف- ابتدا تعداد 2000 نمونه از این توزیع تولید کنید و سپس هیستوگرام داده های تولید شده را رسم نمایید.
 ب- نشان دهید که برای n بزرگ تقریباً دارای توزیع نرمال است (قضیه حد مرکزی). برای این منظور از تمام روش های زیر استفاده کنید:
- 1- رسم هیستوگرام داده ها،
 - 2- استفاده از دستور q-q plot موجود در SAS.
 - 3- انجام آزمون فرض با $(\alpha=0.05)$.

- 2- کارخانه داری برای بررسی مقاومت دو نوع لاستیک A و B به یکی از دو طریق زیر آزمایش را انجام داده است:
- الف- 10 دوچرخه انتخاب کرده و برای آنها لاستیک نوع A و برای 10 دوچرخه دیگر لاستیک نوع B را مورد استفاده قرار داده است.
 ب- 10 دوچرخه انتخاب کرده و برای هر دوچرخه به طور تصادفی برای یکی از چرخ ها لاستیک نوع A و برای چرخ دیگر لاستیک نوع B را مورد استفاده قرار داده است.
- آیا می توان ادعا کرد که متوسط مقاومت لاستیک نوع A، حداقل 2 واحد بیشتر از متوسط مقاومت لاستیک نوع B است. مسأله را یکبار با فرض الف و یکبار با فرض ب حل کنید. $(\alpha=0.05)$
- تمام مراحل باید توسط نرم افزار SAS انجام گیرد و هیچ کمیتی به طور دستی محاسبه نشود.

مقاومت لاستیک نوع A	120	125	122	118	120	121	119	125	124	123
مقاومت لاستیک نوع B	119	121	121	117	116	115	120	121	122	122

- 3- داده های موجود در فایل های Patient1-4 را در نظر بگیرید. این داده ها اطلاعات فشار خون (BP)، وزن (Weight)، سن (age) و جنسیت (gender) چهار بیمار است که در طی چندین سال ثبت شده است. این داده ها را به فرمت زیر تبدیل کنید. که در آن متغیر iter نشان دهنده تعداد دفعات مراجعه فرد برای ثبت اطلاعاتش و متغیر ave_age متوسط سن فرد در دوره زمانی مراجعاتش است.
- تمام مراحل باید توسط نرم افزار SAS انجام گیرد و هیچ داده ای به طور دستی وارد نشود.

سوال امتیازدار: اگر تعداد دفعات ثبت اطلاعات برای افراد مختلف متفاوت باشد، به چه صورت می توان به فرمت مورد نظر دست یافت.

توضیح: چنین داده هایی را داده های پانلی یا طولی گویند و معمولاً تعداد واحدهای آزمایشی (تعداد بیماران در این مثال) زیاد است.

THE REQUIRED FORMAT

```
Iter BP weight ave_age
1 130 80.0 33
2 132 82.1 33
3 132 81.5 33
4 134 85.0 33
5 134 84.2 33
1 140 65.2 45
2 142 66.3 45
3 142 66.5 45
4 143 70.0 45
5 144 72.5 45
.
.
.
```