

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



درس آمار و احتمال

تمرین کامپیوتری شماره 2

آذر ماه ۱۳۹۹

بخش اول – ماتریس همبستگی و داده‌های گم‌شده

در این بخش داده‌های مربوط به شاخص‌های کلی کشورها در اختیار شما قرار گرفته است.

1. داده‌ها در فایل اکسل countries.csv ذخیره شده‌اند. ابتدا داده‌ها را در یک dataframe ذخیره کنید. دقت کنید که به طور معمول در داده‌های واقعی بعضی از متغیرها حاوی نمونه‌هایی با مقادیر نامشخص هستند. داده‌ها را به گونه‌ای ذخیره کنید که به جای مقادیر نامشخص <NA> قرار بگیرد.
2. همان طور که اشاره شد وجود داده‌های گم‌شده اجتناب ناپذیر است، با جستجو در منابع اینترنتی درباره راهکارهای حل این مشکل جستجو کنید و به صورت خلاصه در گزارش کار شرح دهید.
3. با ماتریس کواریانس در درس آشنا شده‌اید، تفاوت عمده ماتریس همبستگی با ماتریس کواریانس در این است که درایه‌های آن حاوی همبستگی دوجه‌دوی متغیرها را به جای کواریانس نشان است. به کمک ابزارهای موجود در R ماتریس همبستگی داده‌ها را بدست آورید.
4. همان طور که در ماتریس همبستگی مشاهده می‌کنید، مقدار همبستگی بین نرخ زادو ولد و کشاورزی مقداری نزدیک به یک دارد. نمودار scatter plot مربوط به این دو متغیر را رسم کنید.
5. در نمودار نیز رابطه خطی معناداری بین مقادیر دو متغیر مشاهده می‌کنید. حال با پیاده سازی الگوریتم مناسبی مقادیر گم شده کشاورزی را پر کنید.

بخش دوم – شبیه سازی متغیر تصادفی

اگر X یک متغیر تصادفی باشد که مقدار ۱ را با احتمال p و مقدار ۰ را با احتمال $1-p$ بگیرد، آنگاه X را یک متغیر تصادفی با توزیع برنولی می‌نامند. p پارامتر توزیع برنولی نامیده می‌شود که مقداری بین صفر و یک است.

1. با استفاده از تابع $\text{runif}()$ یک متغیر تصادفی با توزیع یکنواخت استاندارد بسازید.
2. به کمک دستور بالا یک متغیر تصادفی با توزیع برنولی، $p=0.6$ بسازید.
3. حال با استفاده از متغیر تصادفی برنولی که ساختید، یک متغیر تصادفی دیگر با توزیع دوجمله‌ای، $n=10$ بسازید. در ادامه 100 نمونه از این متغیر تصادفی بگیرید و میانگین و واریانس نمونه‌ها را محاسبه کرده و نتایج بدست آمده را با روابط تئوری مقایسه کنید.

بخش سوم – تبدیل معکوس

یکی روش‌های عمومی برای شبیه‌سازی متغیرهای تصادفی، روش تبدیل معکوس می‌باشد. اگر $U \sim \text{Uniform}[0,1]$ و $F_X(x)$ تابع توزیع تجمعی X باشد، خواهیم داشت: $X = F^{-1}(U)$.

1. درستی این رابطه را نشان دهید.
2. نشان دهید که اگر $X \sim \exp(\lambda)$ باشد، خواهیم داشت: $X = -\frac{1}{\lambda} \ln(U)$. حال به کمک این رابطه و با نمونه‌گیری از توزیع یکنواخت استاندارد، 1000 نمونه از متغیر تصادفی X استخراج کرده و توزیع فراوانی آنها را رسم کنید.

نحوه تحویل

فایل یا فایل‌های R. حاوی کدها و فایل pdf گزارش را در یک فایل زیپ با نام CA#2-STD_NUM.zip قرار داده و بارگذاری کنید.

توجه داشته باشید که تمامی خروجی‌های خواسته شده اعم از اعداد و نمودارها در تمامی بخش‌های سوالات باید در گزارش به طور کامل ذکر شوند تا مورد ارزیابی قرار گیرند.