



سوال ۱) با فرض نرمال بودن داده های  $\{x_1, x_2, \dots, x_n\}$  زیر و با کمک درست نمایی بیشینه پارامتر های  $(\mu, \sigma^2)$  را مشخص نمایید.

سوال ۲) جدول احتمالی سه طرفه را برای ویژگی های  $X, Y, Z$  نشان داده شده در جدول زیر در نظر بگیرید.

|       | $Z=f$ |       | $Z=g$ |       |
|-------|-------|-------|-------|-------|
|       | $Y=d$ | $Y=e$ | $Y=d$ | $Y=e$ |
| $X=a$ | ۵     | ۲     | ۶     | ۹     |
| $X=b$ | ۵     | ۲     | ۴     | ۳     |
| $X=c$ | ۱۵    | ۲     | ۱۷    | ۷     |

مقدار  $\chi^2$  را برای همبستگی بین  $X$  و  $Z$  محاسبه کنید. در سطح معناداری ۱۰ درصد بررسی کنید که آیا آنها وابسته هستند یا مستقل؟ برای یافتن مقادیر بحرانی از جدول ۳-۱۰ کتاب [مرجع](#) استفاده نمایید.

سوال ۳) ماتریس کوواریانس را با جزییات برای مثال زیر محاسبه نمایید:

|         | $x_1$ | $x_2$ | $x_3$ |
|---------|-------|-------|-------|
| $X_1^T$ | ۲     | ۱     | ۲     |
| $X_3^T$ | ۰     | ۳     | ۲     |
| $X_3^T$ | ۱     | ۴     | ۳     |

نکته: سطر ها نشان دهنده نمونه ها و ستون ها مشخصه ها هستند.

سوال ۴: نمونه داده های زیر را در نظر بگیرید:

| age | sex    | bmi  | children | smoker | region | charges |
|-----|--------|------|----------|--------|--------|---------|
| 19  | female | 16   | 0        | no     | south  | 16884   |
| 18  | male   | 31   | 0        | yes    | south  | 26874   |
| 32  | female | 27.9 | 3        | yes    | east   | 21984   |
| 25  | female | 22.3 | 0        | no     | north  | 1725    |
| 34  | male   | 31.4 | 1        | yes    | east   | 22584   |
| 42  | male   | 37   | 2        | no     | north  | 35284   |

الف: ستون های عددی را با روش z\_score یا min-max نرمال کنید.

ب: ستون های غیر عددی را به روش مناسب عددی کنید.

ج: ستون age را به سه دسته تقسیم کنید.