Shawn Malik
DS3001

A2: EDA

Q1.

1. Show that $m(a+bX) = a + b \times m(X)$

1. $m(X) = \frac{1}{N}\sum_{i=1}^{N} x_i$

$\downarrow$

$$m(a+bX) = \frac{1}{N}\sum_{i=1}^{N}(a+bx_i)$$

$\downarrow$

$$m(a+bX) = \frac{1}{N}\left(\sum_{i=1}^{N} a + \sum_{i=1}^{N} bx_i\right)$$

$\downarrow$

$$m(a+bX) = \frac{1}{N}\left(Na + b\sum_{i=1}^{N} x_i\right)$$

$\downarrow$

$m(a+bX) = a \cdot b \cdot \frac{1}{N}\sum_{i=1}^{N} x_i$, by definition: $\frac{1}{N}\sum_{i=1}^{N} x_i = m(X)$

$\downarrow$

$$\boxed{m(a+bX) = a + b\,m(X)}$$

2. Show that $cov(X, a+bY) = b * cov(X, Y)$

We know that $cov(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))(y_i - m(Y))$, substituting $X=X$ and $Y = a+bY$, we get:

$$cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))((a+by_i - m(a+bY)))$$

$\downarrow$

Knowing $m(a+bY) = a+bm(Y)$, substitute:

$$cov(X, a+bY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))\,b(y_i - m(Y)) = b\frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))(y_i - m(Y))$$

$\downarrow$

$$\boxed{cov(X, a+bY) = b\,cov(X,Y)}$$

3. Show that $cov(a+bX, a+bX) = b^2 cov(X,X)$, in particular that $cov(X,X) = s^2$

From 2, we know that $cov(X, a+bY) = b\, cov(X,Y)$

With $X = a+bX$ and $Y = X$, we get $cov(a+bX, a+bX) =$ $b\, cov(a+bX, X)$

applying it again to the remaining covariance, we get

$cov(a+bX, X) = b\, cov(X,X)$

Multiplying the b's, we get $\boxed{cov(a+bX, a+bX) = b^2 cov(X,X)}$

Finally, ~~by definition~~ of covariance,

$$\boxed{cov(X,X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))(x_i - m(X)) = \frac{1}{N}\sum_{i=1}^{N}(x_i - m(X))^2 = s^2}$$

4. Yes, if g is non-decreasing, then the median of $g(X)$ is just $g(\text{median}(X))$. More generally, any quantile transforms in the same way: $Q_{g(X)}(p) = g(Q_X(p))$. This means that the median (and other quantiles) are preserved under monotone transformations. However, the interquantile range (IQR) and the range do not in general transform so simply, as they become the difference of the transformed quartiles or extremes: $IQR(g(X)) = g(Q_X(0.75)) - g(Q_X(0.25))$ and range $(g(X)) = g(\max X) - g(\min X)$. These only scale nicely when g is affine ($a+bX$ with $b \geq 0$).

5. It is generally not true that $m(g(X)) = g(m(X))$. Let X take values -1 and 1 with equal probability, then $m(X) = 0$. If we take $g(X) = X^2$, then $m(g(X)) = m(X^2) = 1$, while $g(m(X)) = g(0) = 0$, and $1 \neq 0$.