DS 3001: Wrangling and EDA

Q6: Linear transformation

1) $m(a+bX) = a+bm(X)$

Let $y = a+bx$

$m(a+bx) = m(y)$

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

$= \frac{1}{N}\left(\sum a + b\sum x\right) = a+b\left(\frac{1}{N}\sum x_i\right)$

$= a+bm(x)$

2) $Cov(X,X) = s^2$

$= \frac{1}{N}\sum_{i=1}^{N}(x-m(X))(x-m(X))$

$= \frac{1}{N}\sum_{i=1}^{N}(x-m(X))^2 = s^2$

3) $Cov(X, a+bY) = bCov(X,Y)$

$m(a+bY) = a+bm(Y)$ * part 1

$Cov(X, a+bY) = \frac{1}{N}\sum (x_i - m(X))((a+by) - m(a+bY))$

sub $m(a+bY) = a+bm(Y)$,

$(a+by) - (a+bm(Y)) = b(y-m(Y))$

then,

$cov(X, a+bY) = \frac{1}{N}\sum (x-m(X)) \cdot b(y-m(Y))$

$= b\left[\frac{1}{N}\sum (x_i - m(X))(y_i - m(Y))\right] = bCov(X,Y)$

4) $cov(a+bX, a+bY) = b^2 cov(X,Y)$

let $U = a+bX$, $V = a+bY$ } $m(U) = a+bm(X)$, $m(V) = a+bm(Y)$

$U_i - m(U) = (a+bx_i) - (a+bm(X)) = b(x_i - m(X))$

$V_i - m(V) = (a+by) - (a+bm(Y)) = b(y - m(Y))$

Then,

$cov(U,V) = \frac{1}{N}\sum (b(x_i - m(x)))(b(y - m(Y))) = b^2 cov(X,Y)$

If $X = Y$,

$cov(bX, bX) = b^2 cov(X,X) = b^2 s^2$

5. $med(a+bX) = a + b\,med(X)$?

Yes (for $b>0$)

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

b/c transform $g(x) = a+bx$ is increasing, so middle values remain the same.

IQR is $Q_3 - Q_1$, for $b>0$, so:

$Q_p(a+bX) = a + bQ_p(X)$

$IQR(a+bX) = Q_3(a+bX) - Q(a+bX)$

$= b(Q_3(X) - Q_1(X)) = b\,IQR(X)$

correct statement:

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

6. Squares examples where $X = \{0,2\}$ & $N=2$

$m(X) = \frac{0+2}{2} = 1$, $(m(X))^2 = 1^2 = 1$

$m(X^2) = 2 \neq 1 = (m(X))^2$

Square root ex where $X = \{1,9\}$

$m(X) = 1+9/2 = 5$, $\sqrt{5} \approx 2.236$

$\sqrt{X} = \{1,3\}$, $m(\sqrt{X}) = 1+3/2 = 2$

$m(\sqrt{X}) = 2 \neq \sqrt{5} = \sqrt{m(X)}$