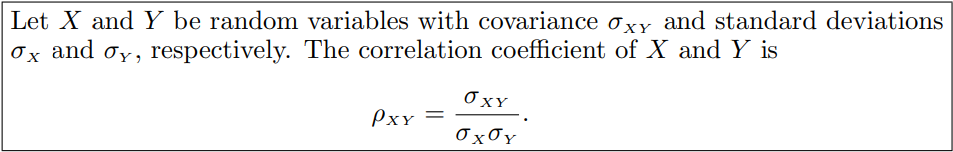
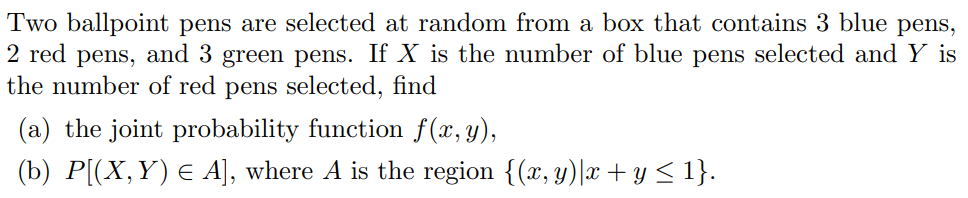
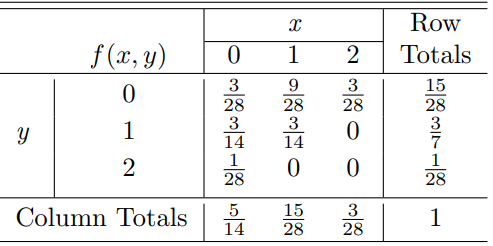
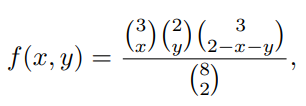
**UNIT-6: Correlation and Regression**

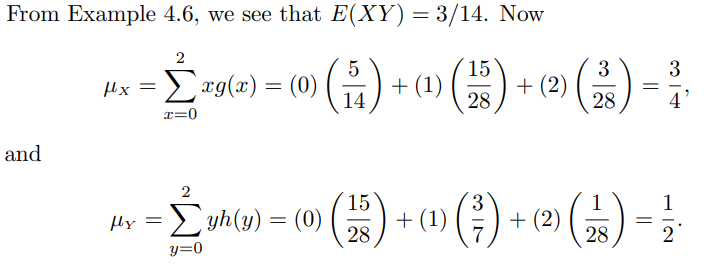


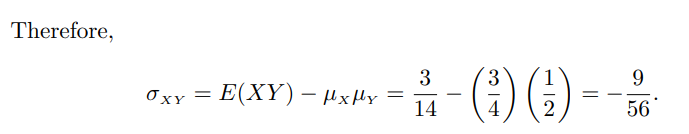
**Example:**

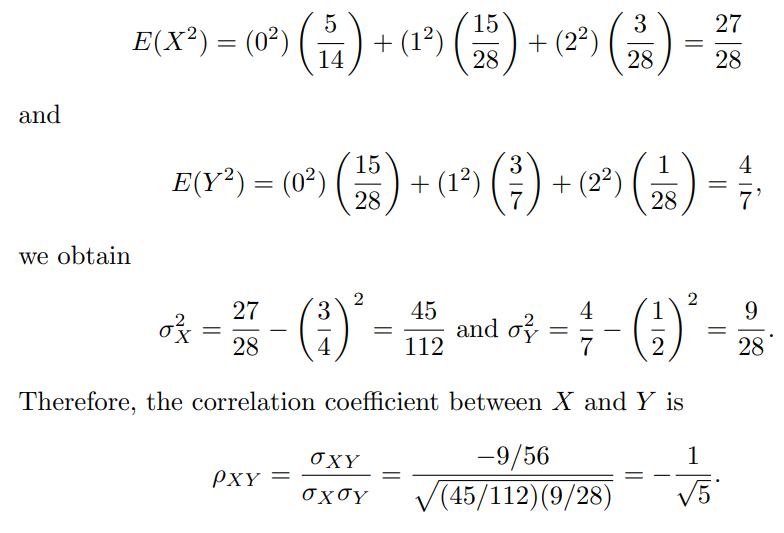


**(c) Find the correlation coefficient between X and Y.**

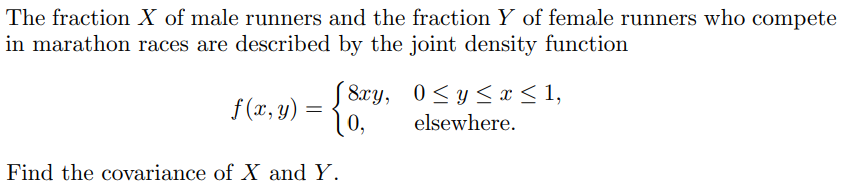
 



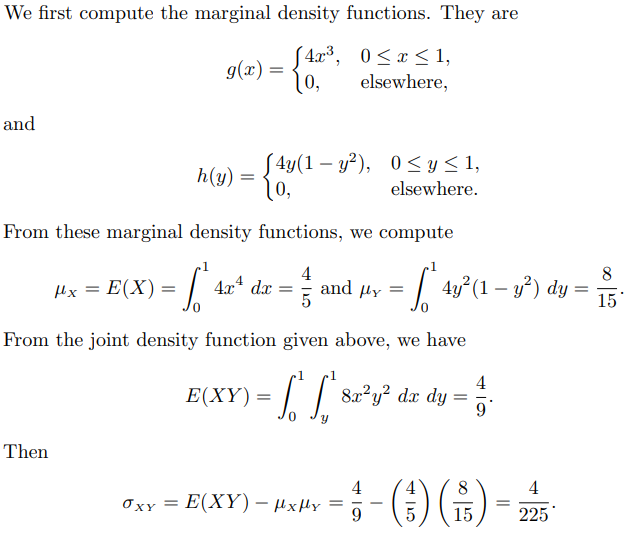


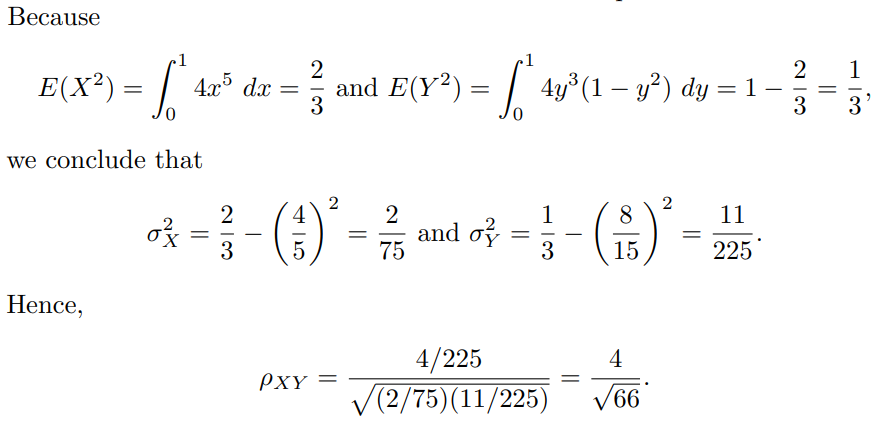
****

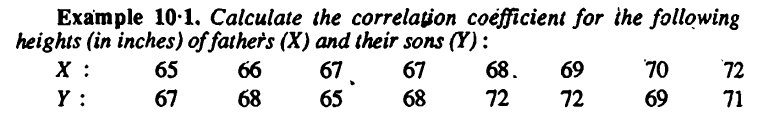
**Example:**

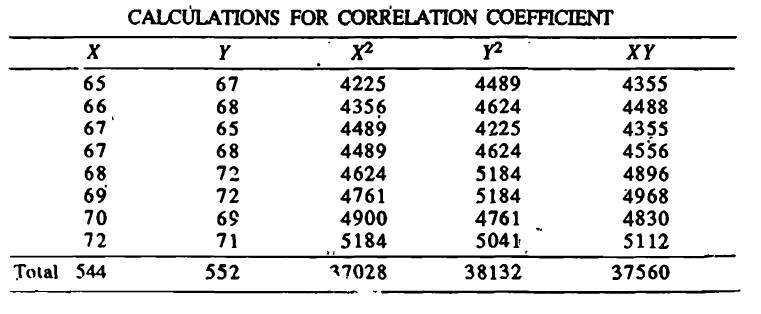


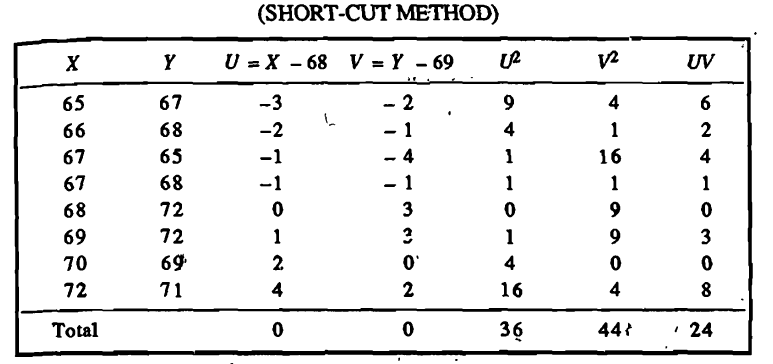
**Find the correlation coefficient between X and Y.**

****

****

****

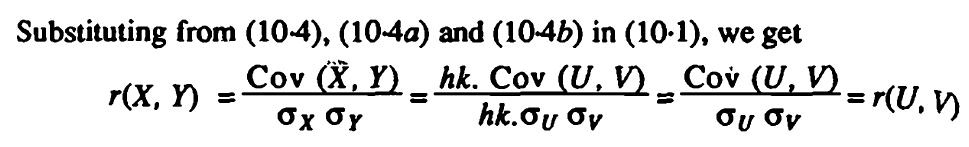
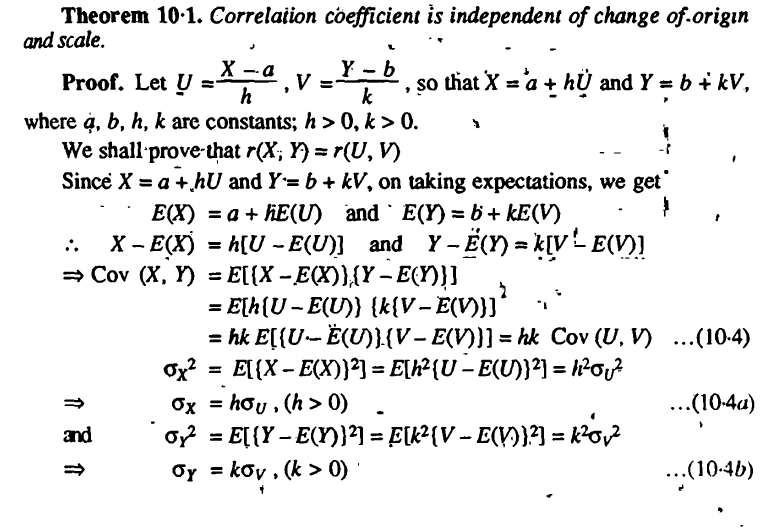
****

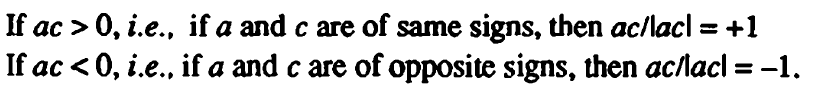
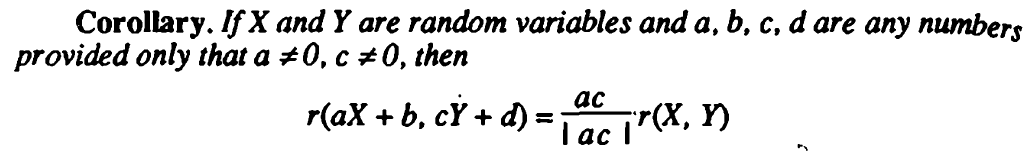
****

Poll Que: If then which of the following options is correct?

(a) (b) (c) (d)

**Theorem: *Correlation coefficient is independent of change of origin and scale.***



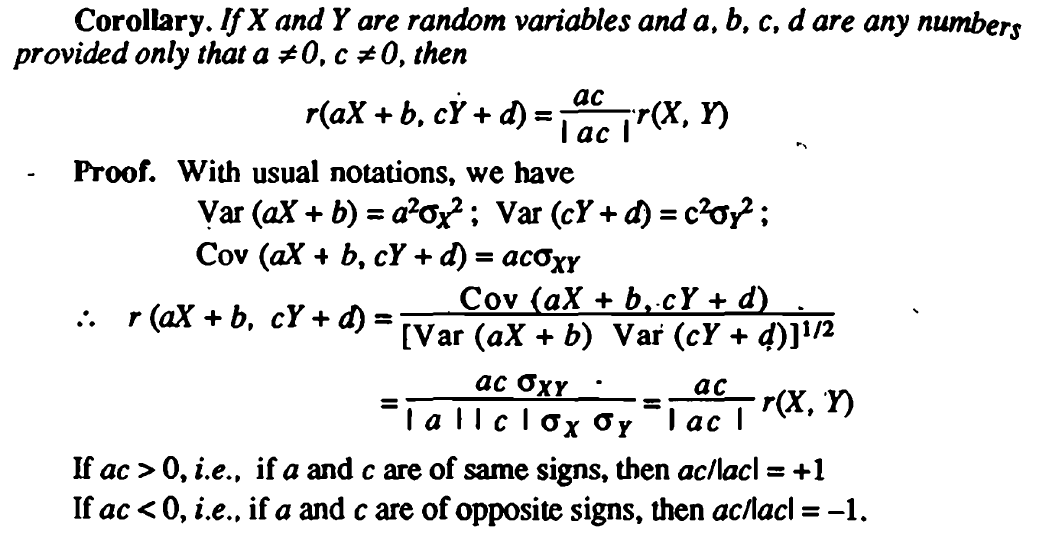


Poll Que: If and are random variables and are any numbers provided

that , then which of the following options is correct?

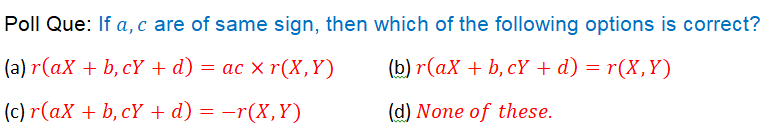
(a) (b)

(c) (d)



**Result:**

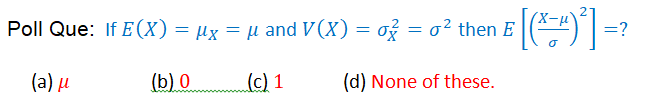


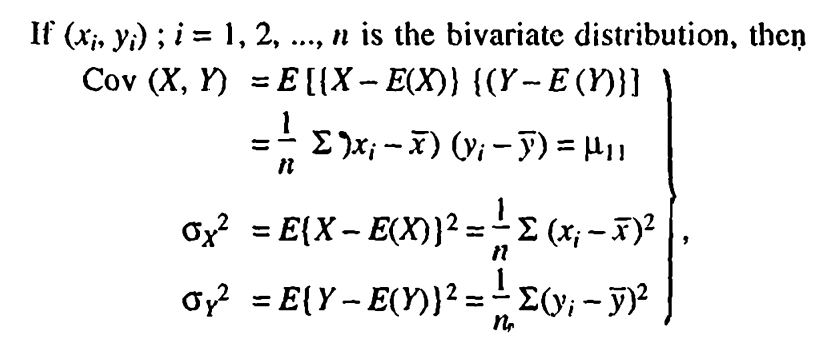
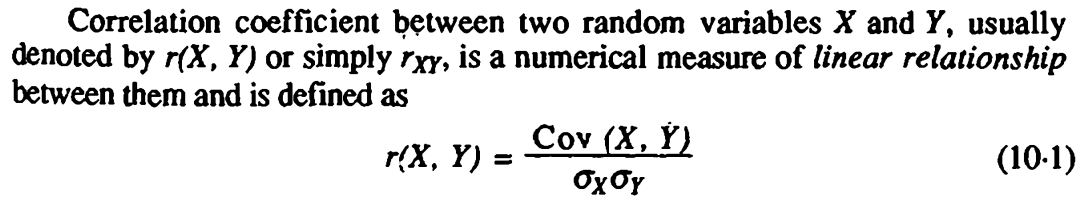


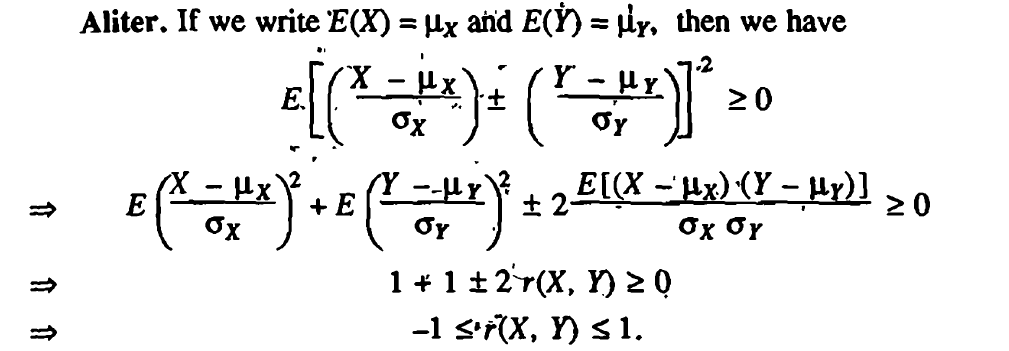
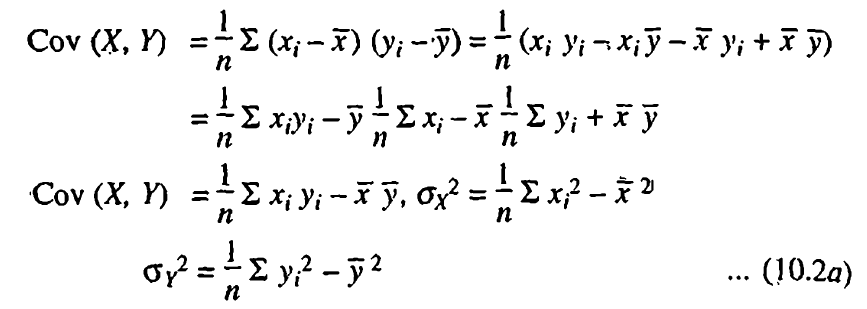
Poll Que: If are of same sign, then which of the following options is correct?

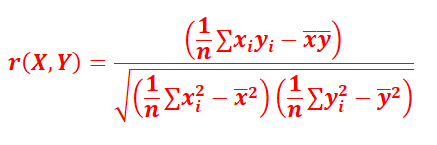
(a) (b)

(c) (d)









Poll Que: Which of the following options is correct?

(a) (b)

(c) (d)

Poll Que: Which of the following options is correct?

(a) (b)

(c) (d)

Q. If and then which of the following options is correct?

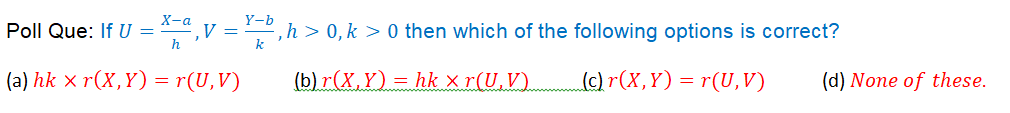
(Here Cov.(X,Y) is Covariance between X and Y)

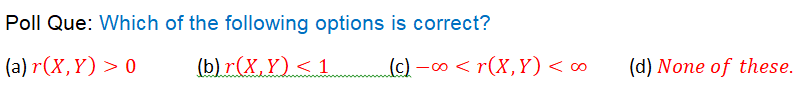
Q. If and then which of the following options is correct?

(Here r(X,Y) is the Karl Pearson’s correlation coefficient)

Q. If and then which of the following options is correct?

(b(X,Y) is regression coefficient)



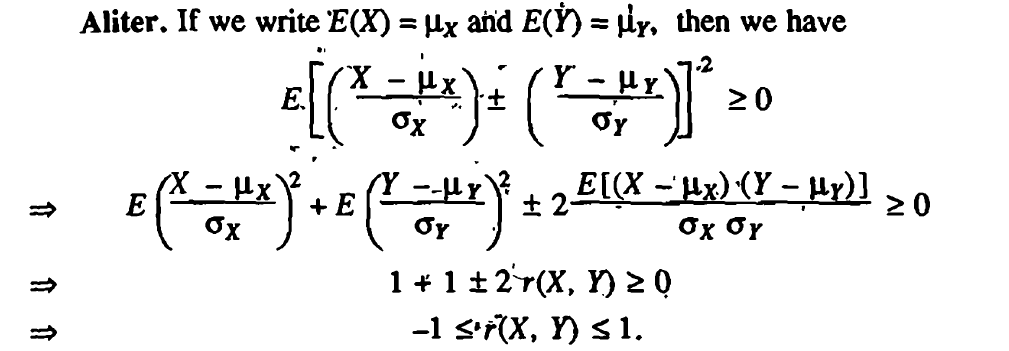


Poll Que:If and then

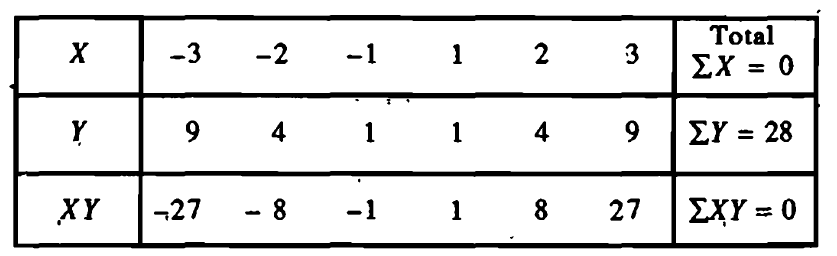
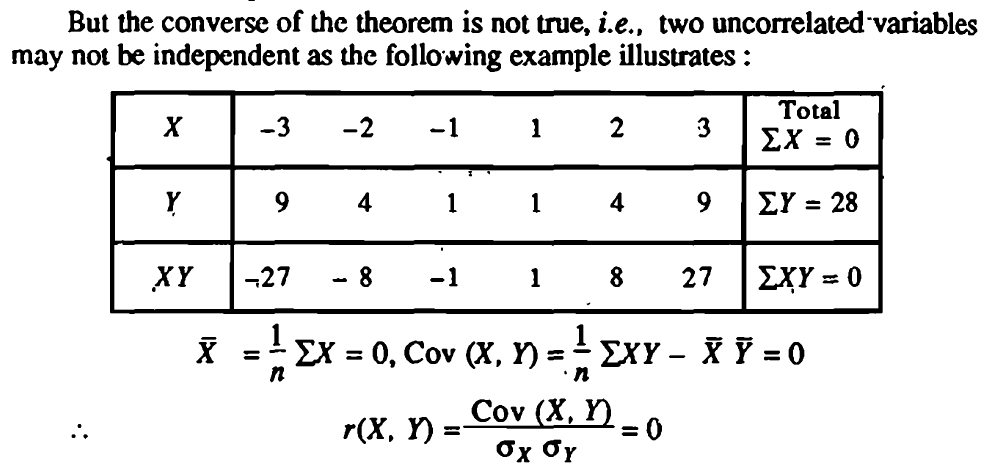
(a) (b) (c) (d) None of these.

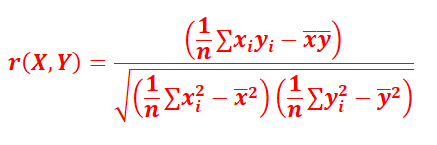
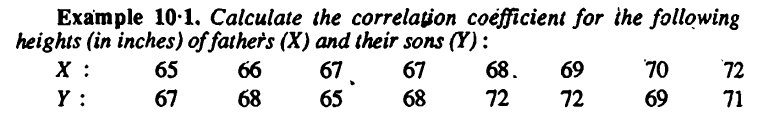
Poll Que: Which of the following options is correct?

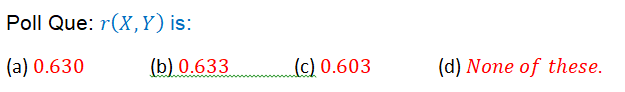
(a) (b) (c) (d)



**Theorem: *Two independent variables are uncorrelated.***

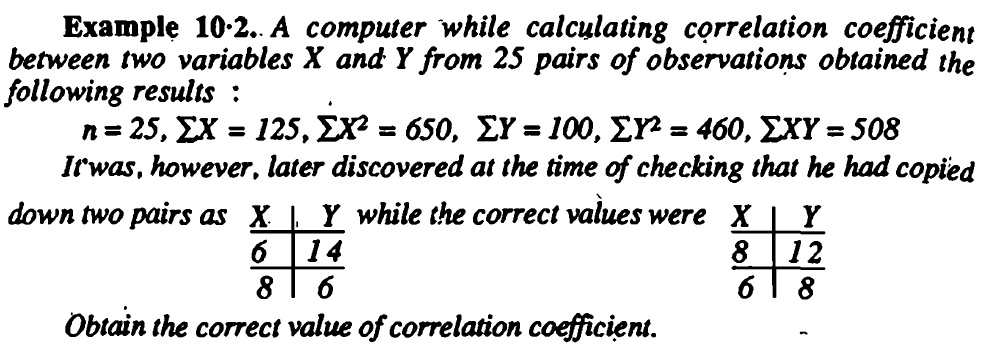


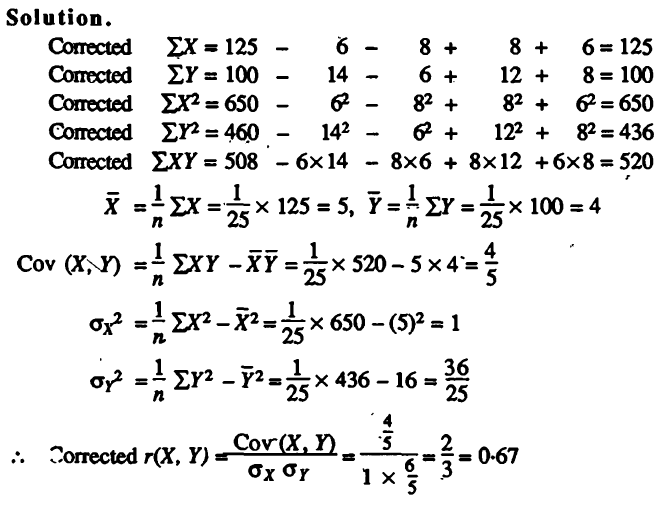


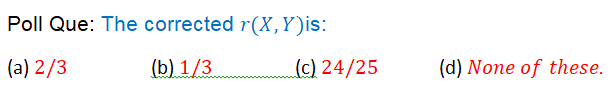


Poll Que:is:

(a) (b) (c) (d)

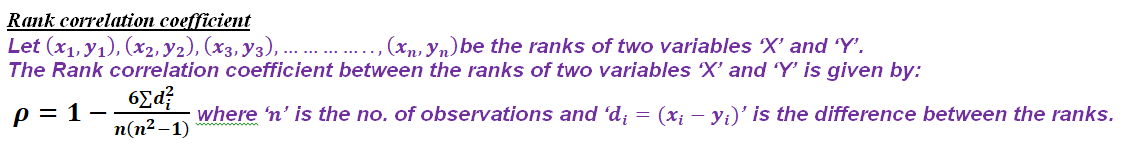






Poll Que:The corrected is:

(a) (b) (c) (d)



***Rank correlation coefficient***

***Let be the ranks of two variables ‘X’ and ‘Y’.***

***The Rank correlation coefficient between the ranks of two variables ‘X’ and ‘Y’ is given by:***

***where ‘’ is the no. of observations and ‘’ is the difference between the ranks.***

**Example:***The ranks of same* 16 *students in Mathematics and Physics are as follows. Two numbers within brackets denote the rank of the*

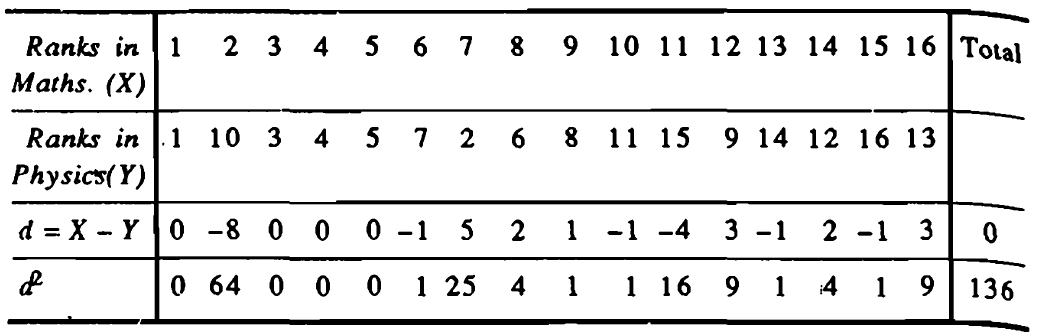
*students in Mathematics and Physics.*

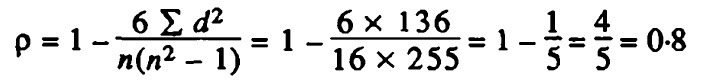
(1,1) *(2,10)* (3,3) (4,4) (5,5) (6,7) (7,2) (8,6) (9,8) *(10,11)* (11,15) (12,9) (13,14) (14,12) (15,16) (16,13).

*Calculate the rank correlation coefficient for proficiencies of this group in Mathematics and Physics.*



(a) (b) (c) (d)





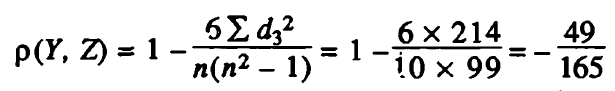
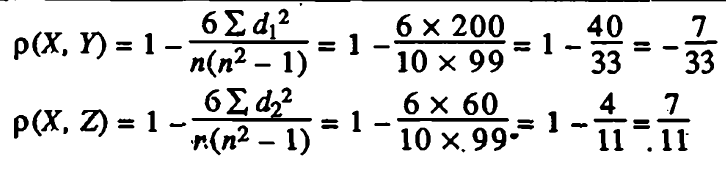
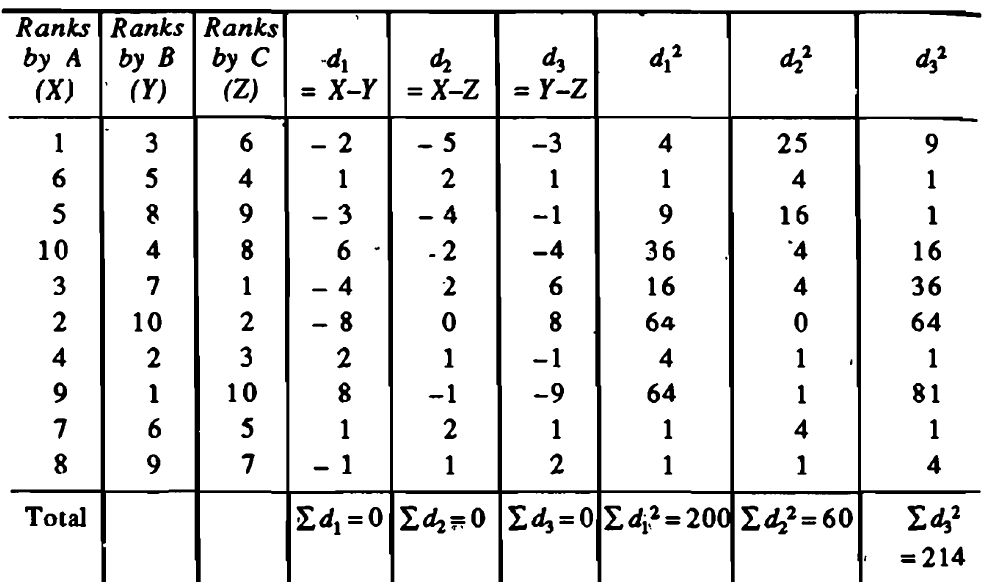
**Example:** *Ten competitors in a musical test were ranked by the three judges A,B and* C *in the following order:*

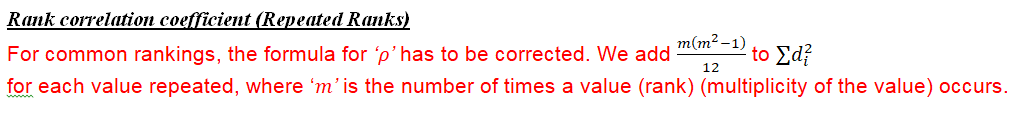
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Ranks by A:*** | ***1*** | ***6*** | ***5*** | ***10*** | ***3*** | ***2*** | ***4*** | ***9*** | ***7*** | ***8*** |
| ***Ranks by B*:** | ***3*** | ***5*** | ***8*** | ***4*** | ***7*** | ***10*** | ***2*** | ***1*** | ***6*** | ***9*** |
| ***Ranks by* C:** | ***6*** | ***4*** | ***9*** | ***8*** | ***1*** | ***2*** | ***3*** | ***10*** | ***5*** | ***7*** |

*Using rank correlation method, discuss which pair of judges has the nearest approach to common likings in music.*



(a) (b) (c) (d)





***Rank correlation coefficient (Repeated Ranks)***

For common rankings, the formula for *‘’*has to be corrected. We add to

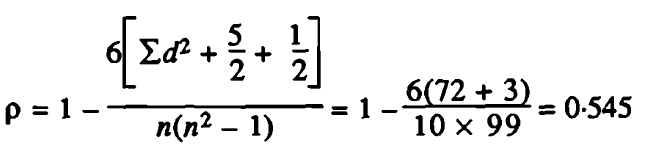
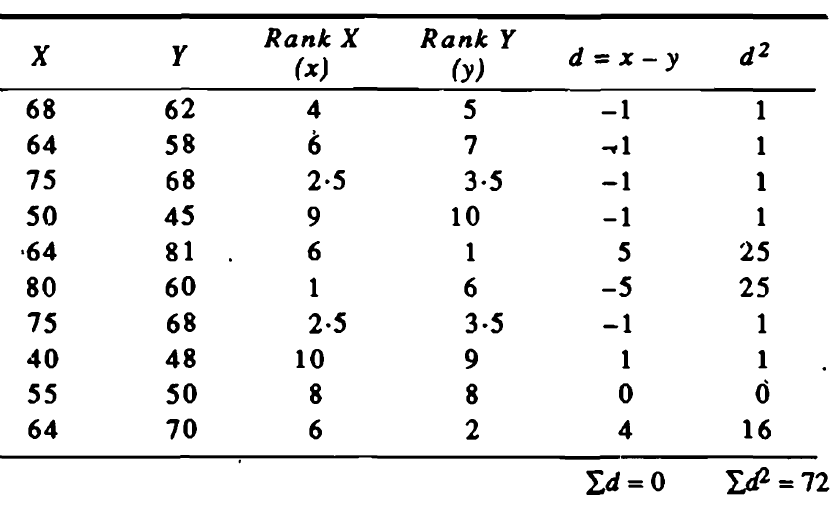
for each value repeated, where ‘*’*is the number of times a value (rank) (multiplicity of the value)occurs.

**Example:** Obtain the rank correlation coefficient for the following data:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X: | 68 | 64 | 75 | 50 | 64 | 80 | 75 | 40 | 55 | 64 |
| Y: | 62 | 58 | 68 | 45 | 81 | 60 | 68 | 48 | 50 | 70 |

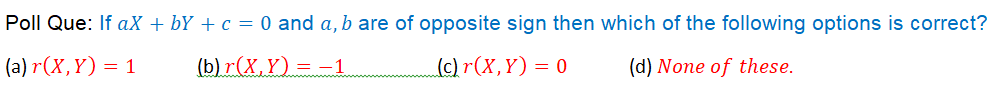
80,75,75,68,64,64,64,55,50,40

81,70,68,68,62,60,58,50,48,45



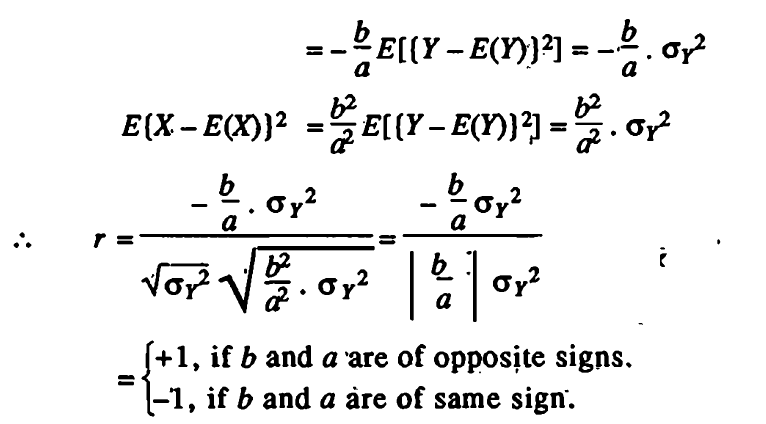
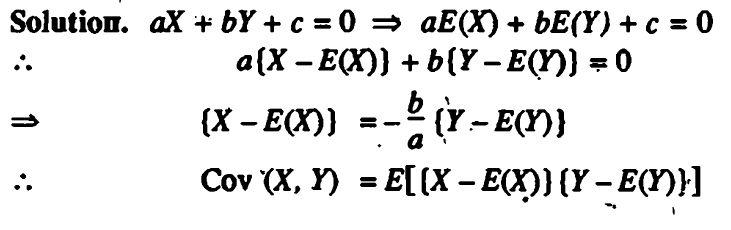
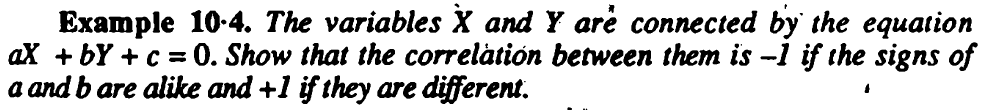


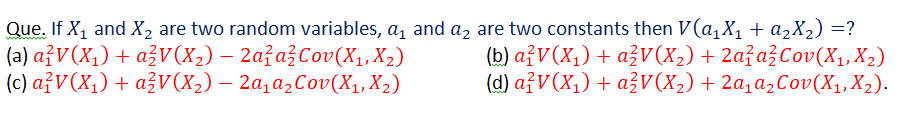
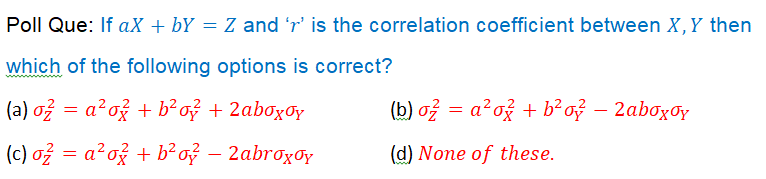
(a) (b) (c) (d)



Poll Que:If and are of opposite sign then which of the following options is correct?

(a) (b) (c) (d)



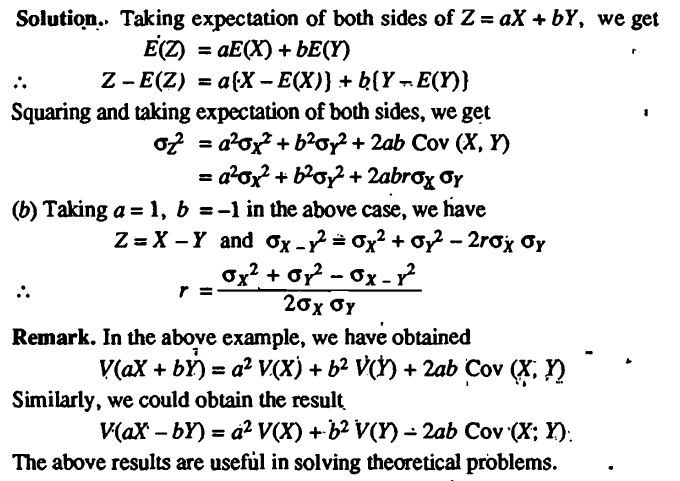
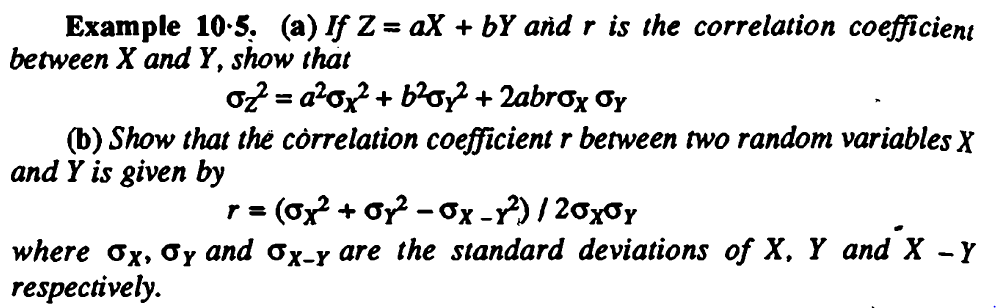


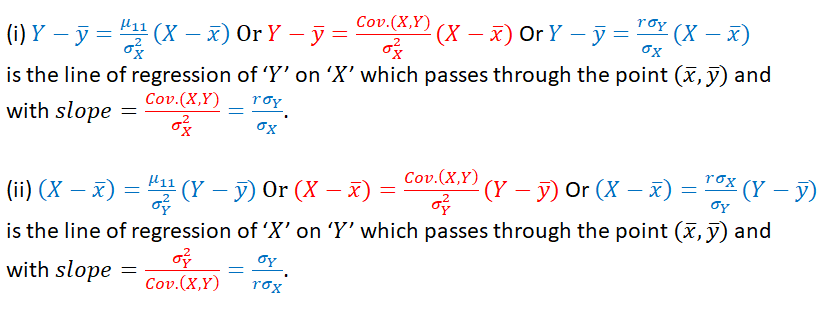
Poll Que:If and ‘’ is the correlation coefficient between then

which of the following options is correct?

(a) (b)

(c) (d)





**Regression**

(i) Or

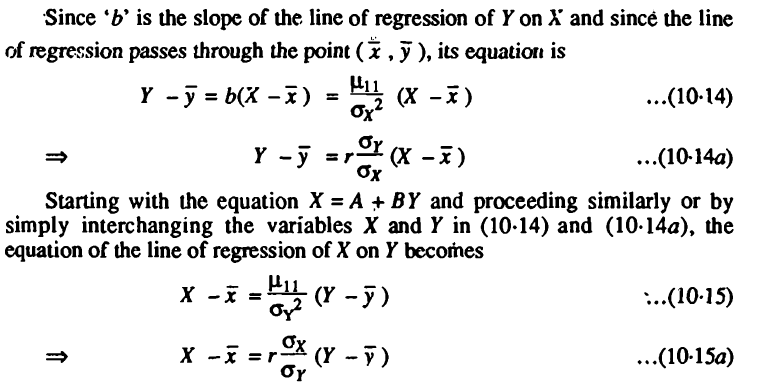
is the line of regression of ‘’ on ‘’ which passes through the point and

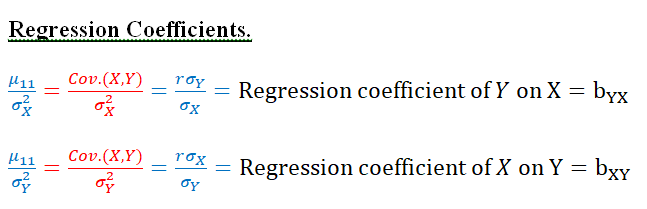
with .

(ii) Or

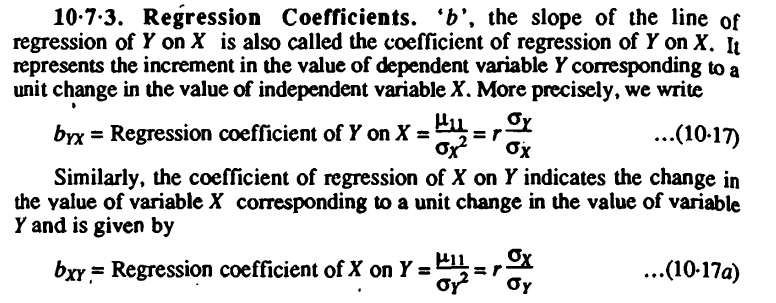
is the line of regression of ‘’ on ‘’ which passes through the point and

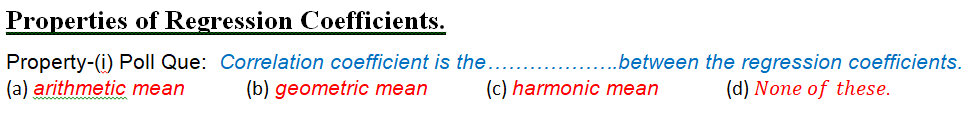
with .





**Regression Coefficients.**

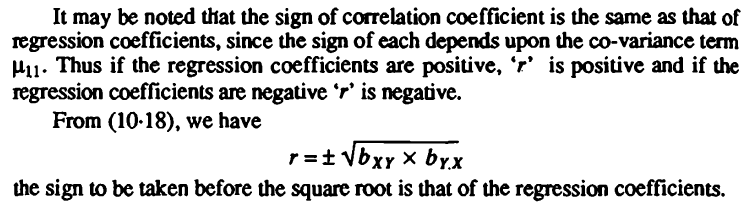
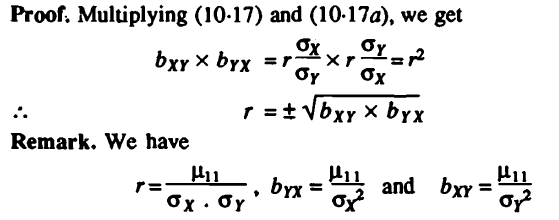


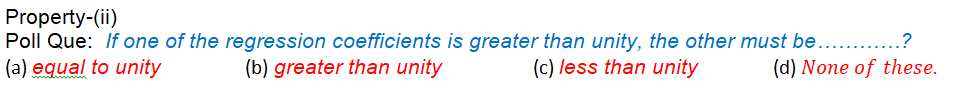


**Properties of Regression Coefficients.**

Property-(i) Poll Que:*Correlation coefficient is the……………….between the regression coefficients.*

(a) *arithmetic mean* (b)*geometric mean* (c) *harmonic mean* (d)



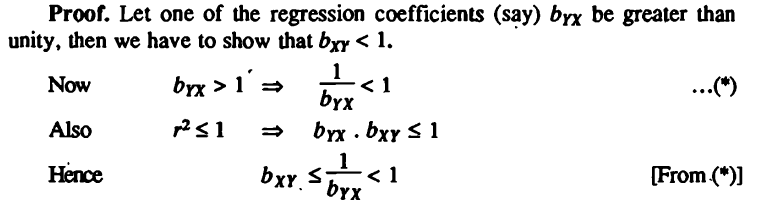


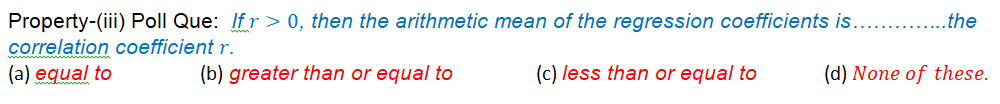
Property-(ii)

Poll Que: *If one of the regression coefficients is greater than unity, the other must be…………?*

(a) *equal to unity* (b)*greater than unity* (c) *less than unity* (d)

**Remark:**The converse of the above property may not be true i.e. if one of the regression coefficients is less than unity then the other regression coefficient may be greater than or less than unity. In other words, both the regression coefficients can be less than unity, but both cannot be greater than unity.

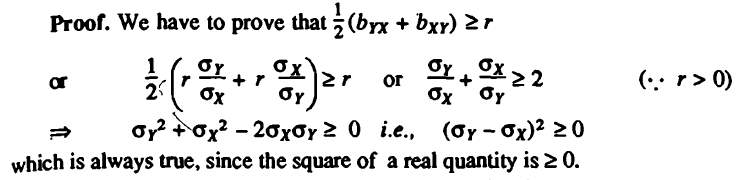




Property-(iii) Poll Que: *If , then the arithmetic mean ofthe regression coefficients is…………..the*

*correlation coefficient .*

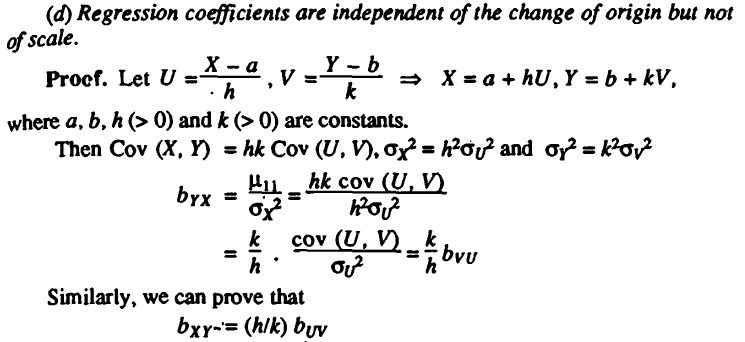
(a) *equal to* (b)*greater than or equal to* (c) *less than or equal to* (d)None of these.

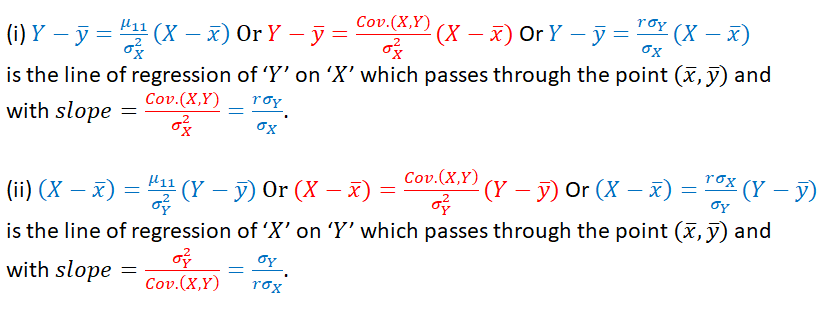
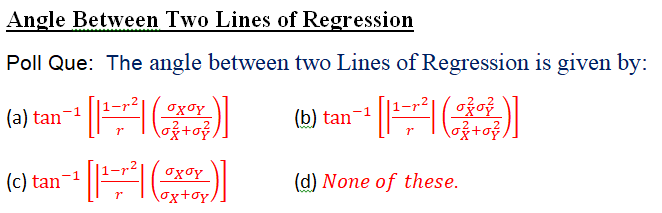


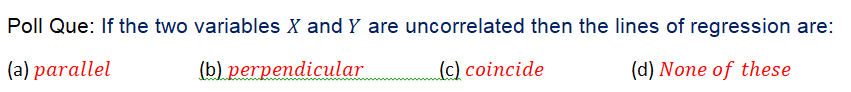


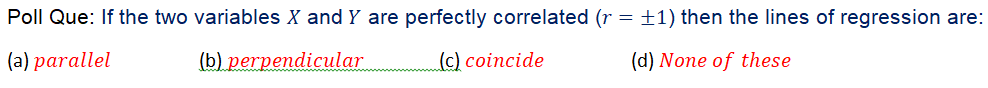
Property-(iv) Poll Que: *Regression coefficients are independent of:*

(a) *change of origin and scale* (b)*change of scale only* (c) *change of origin only* (d)









**Angle Between Two Lines of Regression**

Poll Que: The angle between two Lines of Regression is given by:

(a) (b)

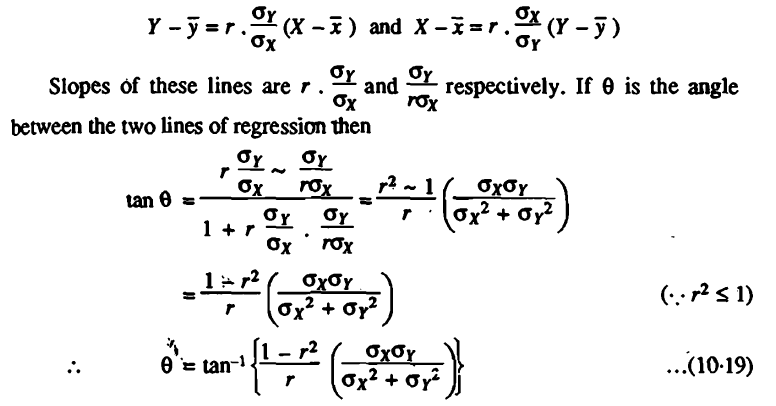
(c) (d)

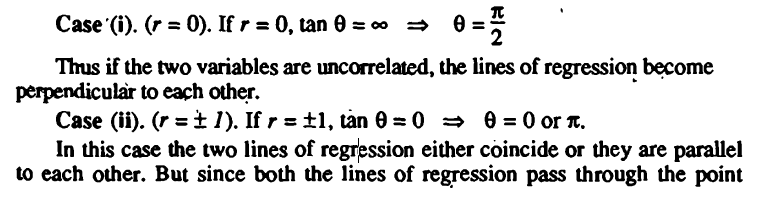
Poll Que: If the two variables and are uncorrelated then the lines of regression are:

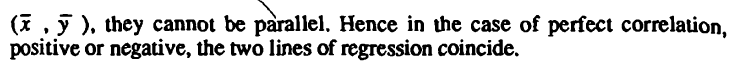
(a) (b) (c) (d)

Poll Que: If the two variables and are perfectly correlated () then the lines of regression are:

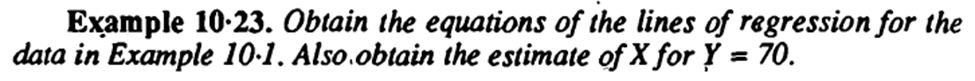
(a) (b) (c) (d)

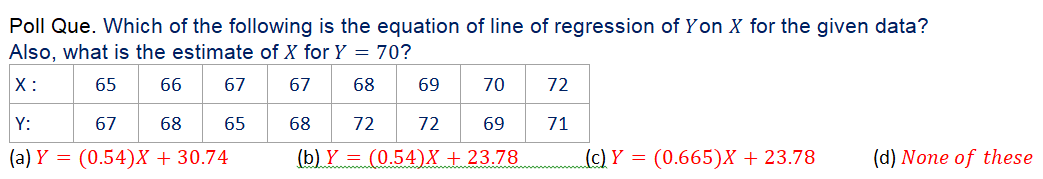


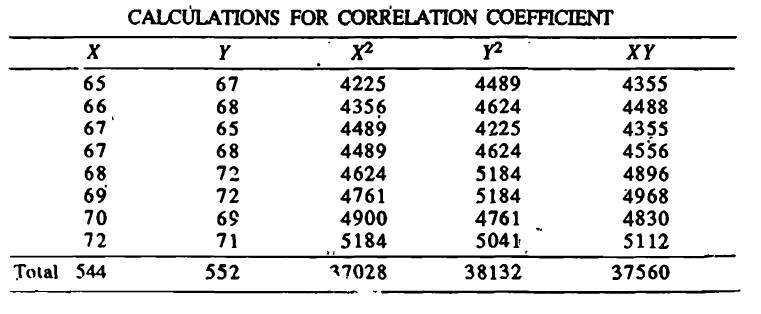










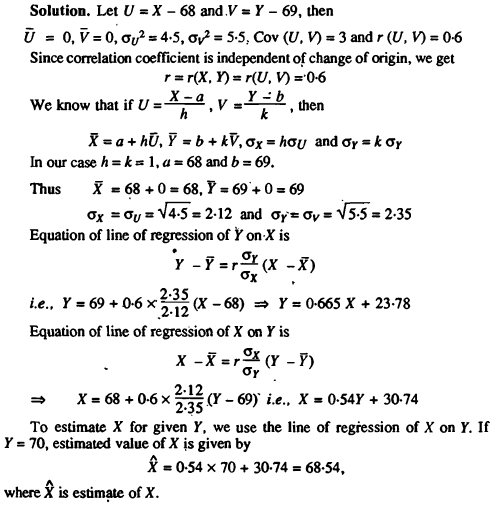
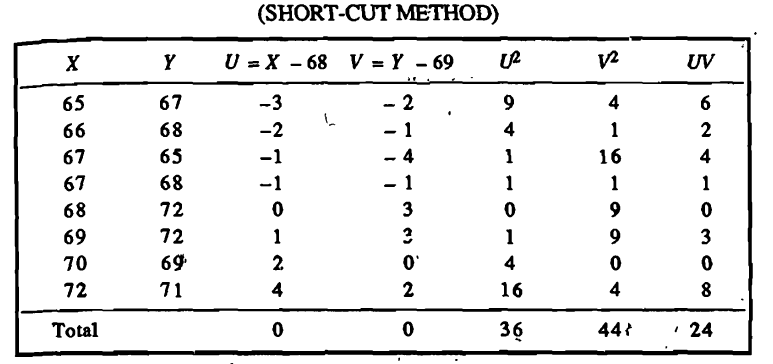


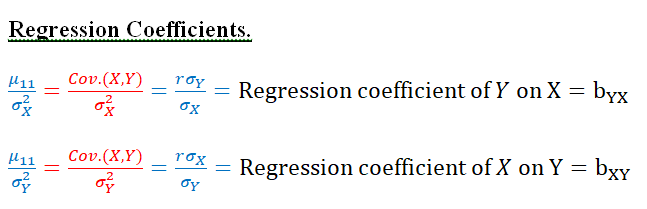
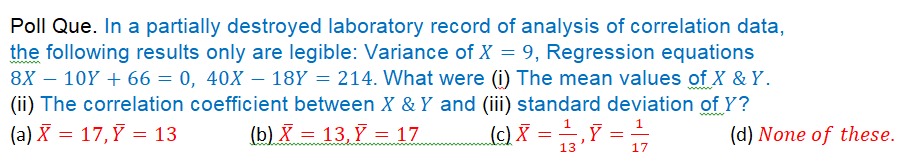
Poll Que. Which of the following is the equation of line of regression of on for the given data?

Also, what is the estimate of for?

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X : | 65 | 66 | 67 | 67 | 68 | 69 | 70 | 72 |
| Y: | 67 | 68 | 65 | 68 | 72 | 72 | 69 | 71 |

(a) (b) (c) (d)









Poll Que. In a partially destroyed laboratory record of analysis of correlation data,

the following results only are legible: Variance of , Regression equations

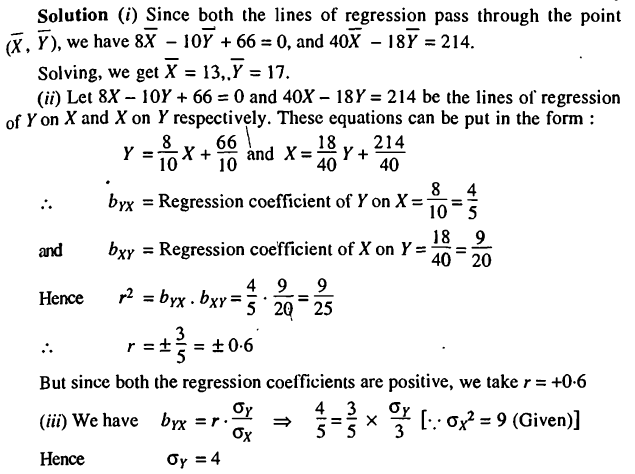
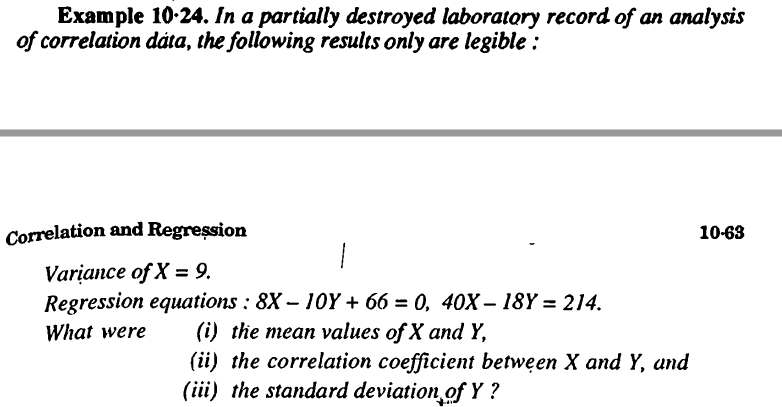
What were(i)The mean values of .

(ii)The correlation coefficient between and (iii) standard deviation of ?

(a) (b) (c) (d).

(a) (b) (c) (d).

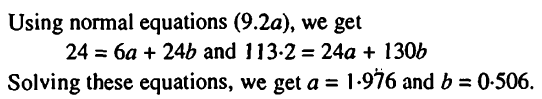
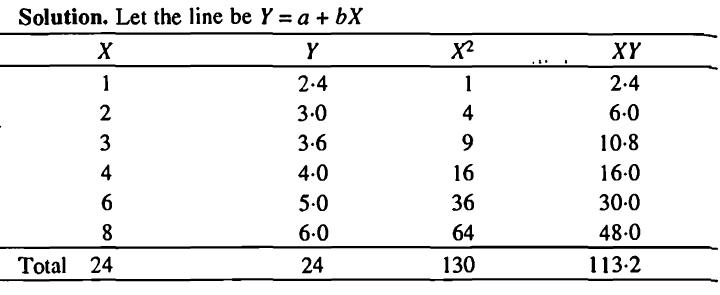
(a) (b) (c) (d).

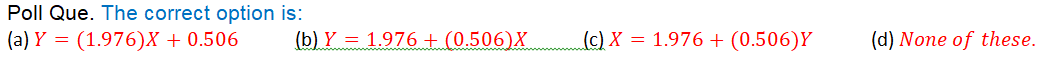


**Curve Fitting**

Example: Fit a straight line to the following data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| X: | 1 | 2 | 3 | 4 | 6 | 8 |
| Y: | 2·4 | 3 | 3·6 | 4 | 5 | 6 |





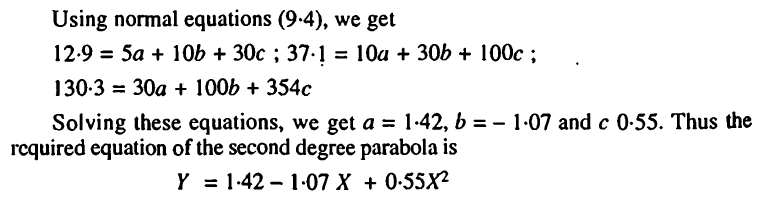
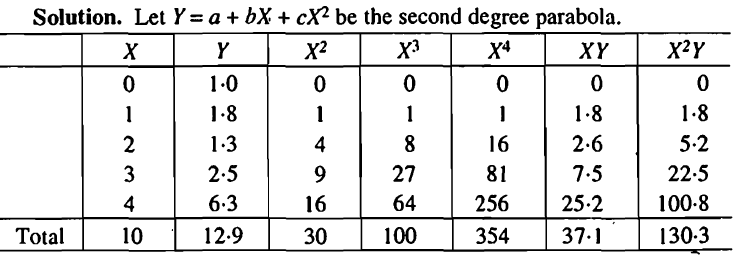
[Solve For a Variable Calculator - Symbolab](https://www.symbolab.com/solver/solve-for-equation-calculator)

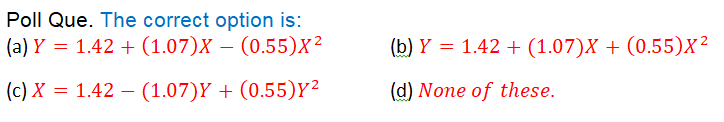
Poll Que. The correct option is:

(a) (b) (c) (d).

Example: Fit a parabola of second degree to the following data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X: | 0 | 1 | 2 | 3 | 4 |
| Y: | 1 | 1·8 | 1·3 | 2·5 | 6·3 |





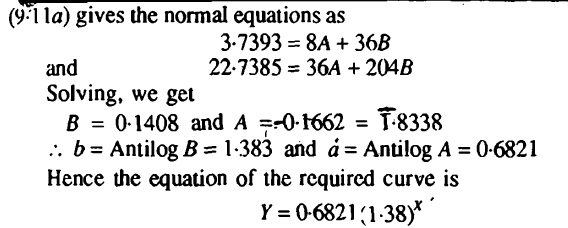
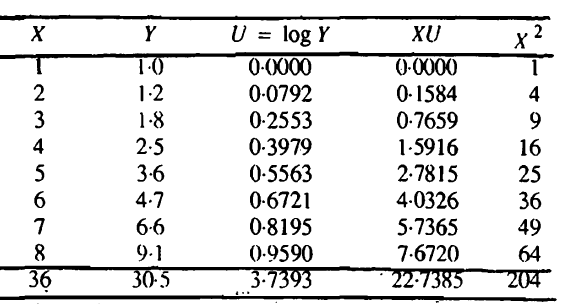
Poll Que. The correct option is:

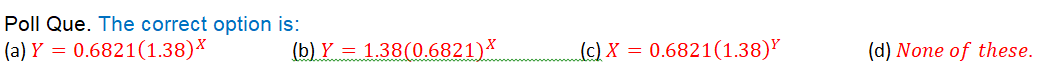
(a) (b)

(c) (d).

Example: Fit an exponential curve of the form to thefollowing data:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Y: | 1·0 | 1·2 | 1·8 | 2·5 | 3·6 | 4·7 | 6·6 | 9·1 |





Poll Que. The correct option is:

(a) (b) (c) (d).