13.02.22

S.no	Sp	Size	Age
1	✓	✓	✓
2	✓	✓	✓
3		Given	
.	.		
n			

Sp = 1 var.

Size & Age = 2 var.

1. Sp no relation w association with Size & Age.
2. Sp having association with Size only.
3. Sp having association with Age only.
4. Sp having association with Size & Age.

Sol^n to Q1 — Sp = Constant [no Size no Age]

Sol^n to Q2 — Sp = Constant + Coeff × Size.

Sol^n to Q3 — Sp = Constant + Coeff × Age

Sol^n to Q4 — Sp = Const + Coeff 1 × Size + Coeff 2 × Age + Coeff 3 × Size × Age.

Sp = y. Size = x_1 Age = x_2

Q1 $\boxed{y = \beta_0}$. [β_0 = Constant]

Q2 $y = \beta_0 + \beta_1 x_1$.

Q3 $y = \beta_0 + \beta_2 x_2$

Q4 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$.

Objective: To forecast or To predict Dv for some given values of Iv.

Sp	Size
y	x_1

Two variables.

1 Dv   1 Iv

Simple Regression

Sp	Size	Age
y	x_1	x_2

1 Dv. more than 1 Iv.

y depending on x_1 & x_2

⇒ Multiple Regression

Why ? we go for MR ✓ instead of SR.

Three types of Correlation Coefficient

✓ 1. ✓ Simple Correlation Coefficient

✓ 2. ✓ Multiple " "

✓ 3. Cannonical "

4. Partial Correlation

5. Part "

Dv	Iv	
1	1	SC.
1	>1	MC. ✓
>1	>1	Cannonical Co.

	y	x_1	x_2
y	1	r_{yx_1}	r_{yx_2}
x_1	r_{yx_1}	1	$r_{x_1x_2}$
x_2	r_{yx_2}	$r_{x_1x_2}$	1

$r_{yx_1} = r_{x_1y}$

$r_{yx_2} = r_{x_2y}$

$r_{x_1x_2} = r_{x_2x_1}$

$R_{y.x_1x_2}$ = Correlation between 1 DV & more than 1 IV.

→ Multiple Correlation

DV. IV.

$y = \beta_0 + \beta_1$ size → SR. ⇒ r, r^2

linear relation between DV & IV.

r^2 is very low.

$r^2 \to 0$

No

$r^2 \to 1$ ⇒ High linear association occurs between DV & IV.

$r^2 = 0.10$

$r^2 = 0.661$

$r^2 > 0.5$ | $r^2 < 0.5$

SR. → MR

$r^2 = 0.91$

$r^2 = 0.661$

r = Correlation Coefficient

$y = a + bx$

$r^2 = 0.90$

10%

r^2 = explained

$1 - r^2$ = unexplained

$$y = \beta_0 + \beta_1 x_1 .$$

$$y = 2 + 3x .$$

$\beta_0 = 2 \qquad \beta_1 = 3 .$

y	x_1

Given values

y	x	$(y-\bar{y})$	$(x-\bar{x})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
2	1	-1.5	-1.5	2.25	2.25	2.25
3	2	-0.5	-0.5	0.25	0.25	0.25
4	3	0.5	0.5	0.25	0.25	0.25
5	4	1.5	1.5	2.25	2.25	2.25

$\bar{y} = \frac{14}{4} = 3.5 \qquad \bar{x} = 10/4 = 2.5$

$\sum (x-\bar{x})^2 = 5 \qquad \sum (y-\bar{y})^2 = 5 \qquad \sum (x-\bar{x})(y-\bar{y}) = 5$

$$r_{xy} = \frac{Cov(x,y)}{SD \text{ of } x \times SD \text{ of } y}$$

$$Cov(x,y) = \frac{1}{n-1} \sum (x-\bar{x})(y-\bar{y}) = \frac{5}{3} = 1.66 .$$

$$S_x = \sqrt{\frac{1}{n-1} \sum (x-\bar{x})^2} = \sqrt{5/3} = \sqrt{1.66}$$

$$S_y = \sqrt{\frac{1}{n-1} \sum (y-\bar{y})^2} = \sqrt{5/3} = \sqrt{1.66}$$

$$r_{xy} = \frac{1.66}{\sqrt{1.66}\,\sqrt{1.66}} = 1$$

$\beta_0 \qquad \beta_1$

The reg. eq^n of y on x [y depends on x]

$$y = \bar{y} + r \frac{S_y}{S_x} (x - \bar{x})$$

$$y = 3.5 + 1 \cdot \frac{\sqrt{1.66}}{\sqrt{1.66}} (x - 2.5)$$

$$= 3.5 + (x - 2.5) = x + 1$$

Predicting model. $\hat{y} = x + 1$.

$\beta_0 = 1$

$\beta_1 = 1$

$r = 1$ $r^2 = 1$.

$E = 0$

Observed.

Predicted

$\hat{y} = x + 1$

y	x	$\hat{y}$	$y - \hat{y}$	$(y-\hat{y})^2$
2	1	2	0	0
3	2	3	0	0
4	3	4	0	0
5	4	5	0	0

$\sum (y - \hat{y})^2 = 0$

Mean sum of square due to error

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2$$

Given

y	x_1	x_2

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

$y = f(x_1, x_2, x_3)$

1 Dv.

R^2

Ivs ⇒ 1) No of Iv. is very large

2) No of Iv combination is large

$p = 20$

$2^{20} = 10,48,576$

$x_1 \quad x_2 \quad x_3$

Case 1: No dep. with $x_1\ x_2\ x_3$ ⇒ β_0

2: Dep with x_1 only ⇒ $\beta_0 + \beta_1 x_1$

3: Dep with x_2 only ⇒ $\beta_0 + \beta_2 x_2$

4: Dep with x_3 only ⇒ $\beta_0 + \beta_3 x_3$

5: " with $x_1\ x_2$ ⇒ $\beta_0 + \beta_4 x_1 x_2$

6: Dep. with $x_1\ x_3$ ⇒ $\beta_0 + \beta_5 x_1 x_3$

7: " " $x_2\ x_3$

8: " " $x_1\ x_2\ x_3$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_6 x_2 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_7 x_1 x_2 x_3$$

$2^3 = 8$

$\beta_4 \to 0 \quad \beta_5 \to 0 \quad \beta_6 \to 0$

$\beta_7 \to 0$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Sp

Sp

Advertising = f(TV, Radio, News).

$y = f(x_1, x_2, x_3)$

R^2

β

R^2

$R^2 = 0.9$

$y = f(x_1, x_3)$

$R^2 = 0.89$