RES-075

APCCMI2025 BANGKOK

Statistical Properties of A Least Squares Method for Linear Regression Analysis

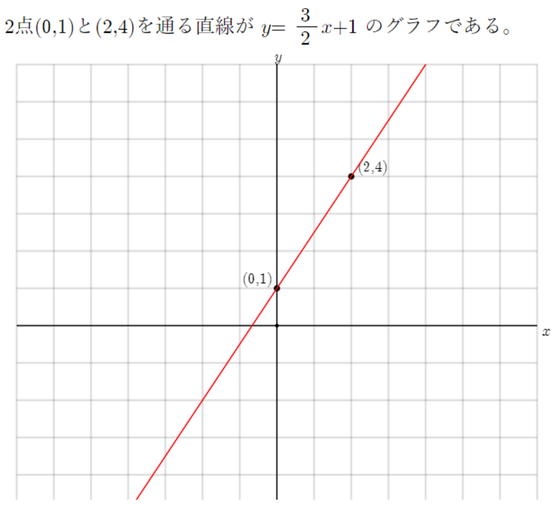
Soichi Takeishi MD, Tatsuo Inoue MD, Koichi Miyamura MD, PhD

Internal medicine, Inuyama Chuo General Hospital

Introduction

- Partial regression coefficients (β) indicate the number of changes in response variables per an increase in covariates by a unit quantity, which is 1, for linear regression analysis.

2点(0,1)と(2,4)を通る直線が $y=\frac{3}{2}x+1$ のグラフである。



The origin of partial regression coefficients

Slope of linear function

↓

The number of changes of y per an increase in x by a unit quantity, which is 1, is calculated from the given 2 coordinates.

↓

Slope of linear function depends on the difference in x and y in the 2 coordinates.

A least squares method is used to calculate partial regression coefficients.

Regression line: $\hat{y} = \hat{\alpha} + \hat{\beta}\, x$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \qquad \hat{\beta} = \frac{S_{xy}}{S_x^2} \qquad S_x^2 = \sum (x - \bar{x})^2,\; S_{xy} = \sum (x - \bar{x})(y - \bar{y})$$

Partial regression coefficients depend on differences

1. between covariates and a mean covariate

and

2. between response variables and a mean response variable

For example, when "the covariates" are replaced with "the covariates – 10", "the mean covariates" are replaced with "the mean covariates – 10".

$(x - 10) - (\bar{x} - 10) = (x - \bar{x})$

The calculated partial regression coefficients are identical. We examined this statistical properties using simulation data.

Method

Case 1

A covariate: "duration from an admission date (a date of onset) to 'an enforcement date of PCR for COVID-19 detection' (PCR date)" (Admission → PCR)

A response variable: threshold cycle values (CT values) of the PCR

Case 2

A covariate: "duration from a ward transfer permission date (defined as ten days after admission) to PCR date" (Transfer permission → PCR)

A response variable: CT values of the PCR

We calculated partial regression coefficients for the two cases using a least squares method (n = 10).

Results

Case 1	Admission → PCR (x_1)	$\bar{x}_1$	$(x_1 - \bar{x}_1)^2$	CT value of PCR (y)	$\bar{y}$	$(x_1 - \bar{x}_1)(y - \bar{y})$	$(x_2 - \bar{x}_2)(y - \bar{y})$
	11		20.25	31		20.25	20.25
	12		12.25	33		8.75	8.75
	13		6.25	34		3.75	3.75
	14		2.25	32		5.25	5.25
	15	15.5	0.25	35	35.5	0.25	0.25
	16		0.25	36		0.25	0.25
	17		2.25	37		2.25	2.25
	18		6.25	39		8.75	8.75
	19		12.25	38		8.75	8.75
	20		20.25	40		20.25	20.25
Total			Sx_1^2: 82.5			Sx_1y: 78.5	Sx_2y: 78.5
Case 2	Transfer permission → PCR (x_2)	$\bar{x}_2$	$(x_2 - \bar{x}_2)^2$	$\hat{\beta}_1$: Sx_1y/Sx_1^2	$\hat{\alpha}_1$	$\hat{\beta}_2$: Sx_2y/Sx_2^2	$\hat{\alpha}_2$
	1		20.25	0.95	20.75	0.95	30.27
	2		12.25				
	3		6.25				
	4		2.25				
	5	5.5	0.25				
	6		0.25				
	7		2.25				
	8		6.25				
	9		12.25				
	10		20.25				
Total			Sx_2^2: 82.5				

The covariates in Case 1 were all "10 + the covariates in Case 2".

The mean of covariates in Case 1 was "10 + the mean of covariates in Case 2".

The differences between the covariates and the mean covariate, as well as the partial regression coefficients, were identical in Cases 1 and 2.

The constant term (α) in Cases 1 and 2 were different.

Conclusion

Partial regression coefficients are calculated by apprehending the differences between the given covariates arranged in ascending order as the number of changes.