

AUTHOR BY:

Sana Ur Rehman Arain

LEARNING JOURNAL 7

University of the People

2025

INSTRUCTOR: ANSON XUAN

MATH 1281-01 Statistical Inference – AY2025-T3

a. Regression Output in JASP

Results

Linear Regression

Model Summary - days

Model	R	R ²	Adjusted R ²	RMSE	R ² Change	df1	df2	p
M ₀	0.000	0.000	0.000	16.253	0.000	0	145	
M ₁	0.299	0.089	0.070	15.673	0.089	3	142	0.004

Note. M₁ includes eth, sex, lrm

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
M ₁	Regression	3421.776	3	1140.592	4.643	0.004
	Residual	34882.477	142	245.651		
	Total	38304.253	145			

Note. M₁ includes eth, sex, lrm

Note. The intercept model is omitted, as no meaningful information can be shown.

Coefficients

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
M ₀	(Intercept)	16.459	1.345		12.236	< .001	13.800	19.117
M ₁	(Intercept)	18.932	2.570		7.365	< .001	13.851	24.013
	eth (N)	-9.112	2.599		-3.506	< .001	14.250	-3.975

Coefficients



Model	Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
						Lower	Upper
sex (M)	3.104	2.637		1.177	0.241	-2.109	8.317
lnr (SL)	2.154	2.651		0.813	0.418	-3.085	7.394

^a Standardized coefficients can only be computed for continuous predictors.

Residuals Statistics

	Minimum	Maximum	Mean	SD	N
Predicted Value	9.820	24.190	16.459	4.858	146
Residual	-22.190	59.914	1.775×10 ⁻¹⁵	15.510	146
Std. Predicted Value	-1.367	1.592	8.766×10 ⁻¹⁷	1.000	146
Std. Residual	-1.442	3.876	-1.079×10 ⁻⁵	1.004	146

b. Regression Equation

The regression equation based on the coefficients is:

$$\text{days}^{\wedge} = 18.932 - 9.112(\text{eth}) + 3.104(\text{sex}) + 2.154(\text{lnr})$$

Where:

- **eth** = 0 (Aboriginal), 1 (Not Aboriginal)
- **sex** = 0 (Female), 1 (Male)
- **lnr** = 0 (Average learner), 1 (Slow learner)

c. Interpretation of Slopes and Statistical Significance

1. **eth** (-9.112, $p < 0.001$)

- Interpretation: Being **not Aboriginal (eth = 1)** reduces the predicted absenteeism by **9.112 days** compared to Aboriginal students.
 - **Statistically significant** at $\alpha = 0.05$ since $p < 0.001$.
2. **sex (3.104, p = 0.241)**
- Interpretation: Being **male (sex = 1)** increases absenteeism by **3.104 days** compared to females.
 - **Not statistically significant** since $p = 0.241 > 0.05$.
3. **lnr (2.154, p = 0.418)**
- Interpretation: Being a **slow learner (lnr = 1)** increases absenteeism by **2.154 days** compared to average learners.
 - **Not statistically significant** since $p = 0.418 > 0.05$.

d. Calculate the Residual for the Second Observation

Given Regression Equation:

$$\hat{y} = 18.932 - 9.112(\text{eth}) + 3.104(\text{sex}) + 2.514(\text{lnr})$$

Given Values for the Second Observation:

- **eth = 0 (Aboriginal)**
- **sex = 1 (Male)**
- **lnr = 1 (Slow learner)**
- **Actual absenteeism (y) = 11 days**

Calculate Predicted Value (\hat{y})

$$\hat{y} = 18.932 - 9.112(0) + 3.104(1) + 2.514(1)$$

$$\hat{y} = 18.932 + 3.104 + 2.514$$

$$\hat{y} = 24.55$$

Calculate Residual:

Residual = Actual Value – Predicted Value

$$\text{Residual} = 11 - 24.55$$

$$\text{Residual} = -13.55$$

e. Model Adjusted R-Squared and Interpretation

- **Adjusted $R^2 = 0.070$**
 - **Interpretation:** The model explains **7%** of the variation in absenteeism. This means that **ethnicity, sex, and learner type only explain a small proportion of absenteeism, suggesting other unaccounted factors influence absenteeism more significantly.**
-
-
-