

## EXAMPLE OF WRITTEN EXAMINATION – 3<sup>rd</sup> APRIL 2023

What does the incorrelation between  $X$  and  $Y$  suggest about model  $Y = \beta_0 + \beta_1 X + \varepsilon$ ? (3 points; 1 correct answer; penalty 33.3%; no answer: 0 points)

Scegli un'alternativa:

- ☐ a. the estimate of  $\beta_0$  is 0
- ☒ b. the  $p$ -value associated to the significance test for  $\beta_1$  is close to 1
- ☐ c.  $R^2 = 1$
- ☐ d. the residual standard error is 0.5
- ☐ e. no answer

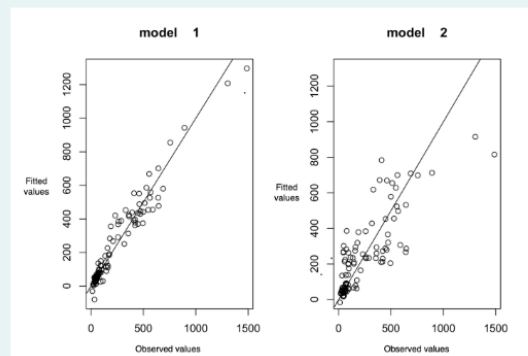
[Annulla la scelta](#)

Which one of the following is an assumption about the response variable  $Y$  in the linear regression model?(3 points; 1 correct answer; penalty 33.3%; no answer: 0 points)

Scegli un'alternativa:

- ☐ a. minimum distance from residuals
- ☐ b. variance that depends on the covariates
- ☐ c. mean equal to zero
- ☐ d. Normal distribution
- ☐ e. no answer

The following graph compares the observed values and the fitted values from two linear regression models. Model 2 is a sub-model of model 1 (nested model). The line in both the graphs is the bisector of the first and third quadrants. What can we conclude? (3 points; 1 correct answer; penalty 33.3%; no answer: 0 points)



Scegli un'alternativa:

- ☐ a. model comparison based on test F suggests to reject model 2, at significance level 0.05
- ☐ b. the explained deviance for model 2 is larger than explained deviance for model 1
- ☐ c. model comparison based on test F suggests not to reject model 2, at significance level 0.05
- ☐ d.  $R^2$  is larger for model 2 than for model 1
- ☐ e. no answer

What kind of information does a numerical interval with confidence level equal to 95% suggest? What can we expect from the interval in case of big datasets? (8 points)

Consider the analysis of the features of 120 houses for sale. The following output refers to a model relating the logarithm of the sale price (in thousands of dollars, Lprice) to the size of the house (in square feet, Size) and the State (CA, NJ, NY, PA, State).

```
Call:
lm(formula = Lprice ~ Size + State)

Residuals:
    Min       1Q   Median       3Q      Max
-1.77915 -0.27947 -0.03962  0.29892  1.69701

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.84306    0.13469   35.957 < 2e-16 ***
Size         0.52416    0.04409   11.888 < 2e-16 ***
StateNJ      -0.13825    0.13835  -0.999 0.319735
StateNY      -0.01309    0.13842  -0.095 0.924802
StatePA      -0.47680    0.13860  -3.440 0.000811 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5358 on 115 degrees of freedom
Multiple R-squared:  0.5864, Adjusted R-squared:  0.572
F-statistic: 40.76 on 4 and 115 DF, p-value: < 2.2e-16
```

Describe how variable State is treated and how the sale price changes with respect to State. (8 points)

Model houses1 is extended to include other information. Consider the following models

houses2: with covariates Size, Baths, Beds

houses3: with covariates Size, State, Beds

houses4: with covariates Size, State, Baths

where Baths and Beds indicate the number of bathrooms and the number of bedrooms in the houses, respectively. Consider the following comparisons between models.

```
> anova(houses1, houses2)
Analysis of Variance Table

Model 1: Lprice ~ Size + Baths + Beds
Model 2: Lprice ~ Size + State
  Res.Df  RSS Df Sum of Sq    F Pr(>F)
1    116 36.245
2    115 33.009  1    3.2362 11.274 0.001066 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(houses1, houses3)
Analysis of Variance Table

Model 1: Lprice ~ Size + State
Model 2: Lprice ~ Size + State + Beds
  Res.Df  RSS Df Sum of Sq    F Pr(>F)
1    115 33.009
2    114 33.009  1 0.00050993 0.0018 0.9666
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(houses1, houses4)
Analysis of Variance Table

Model 1: Lprice ~ Size + State
Model 2: Lprice ~ Size + State + Baths
  Res.Df  RSS Df Sum of Sq    F Pr(>F)
1    115 33.009
2    114 31.515  1    1.4947  5.407 0.02183 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```