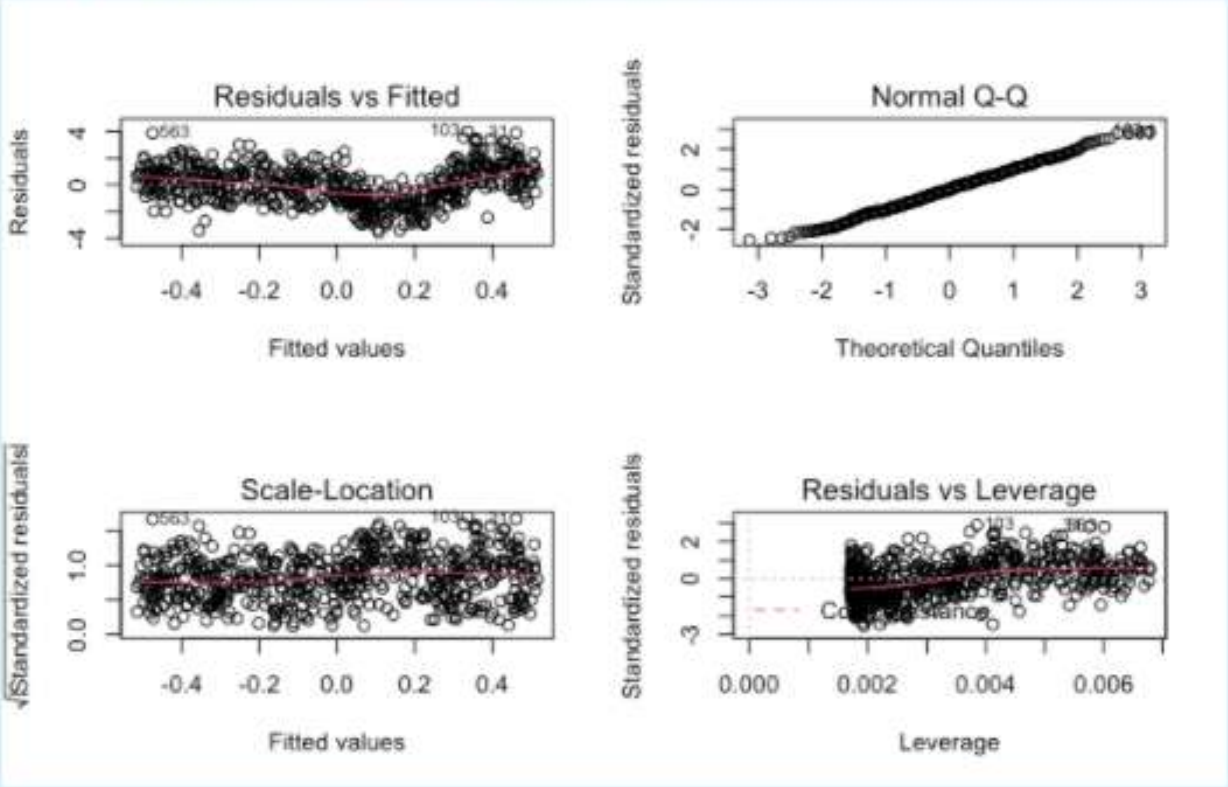
Domanda 5

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

From the following analysis of the residuals in a linear model, we could conclude that one of the model's assumptions is more questionable than the others. Which one?



Scegli un'alternativa:

- a. Linearity ✓
- b. Normality
- c. Homoschedasticity
- d. The random components have zero mean.

La risposta corretta è: Linearity

Domanda 1

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

The dataset `airquality` reports Daily air quality measurements in New York, May to September 1973. The response variable is the Ozone concentration in ppb `Ozone`. Suppose as possible covariates: the wind speed in mph `Wind`, the solar activity in lang `Solar.R`, the temperature in Fahrenheit degrees `Temp` and the month `Month`. Look at the following model summary:

```
Call:
lm(formula = Ozone ~ Solar.R + Wind * Temp + Month, data = airquality)

Residuals:
    Min      1Q  Median      3Q     Max
-36.534 -13.270  -1.179   9.119  96.335

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -235.52336   46.53648  -5.061 1.79e-06 ***
Solar.R        0.05665    0.02180   2.598 0.010714 *
Wind          14.01795    4.08595   3.431 0.000862 ***
Temp           4.06195    0.57038   7.121 1.38e-10 ***
Month         -2.69378    1.40674  -1.915 0.058225 .
Wind:Temp     -0.22304    0.05201  -4.289 4.00e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.37 on 105 degrees of freedom
  (42 observations deleted due to missingness)
Multiple R-squared:  0.6765,    Adjusted R-squared:  0.6611
F-statistic: 43.92 on 5 and 105 DF,  p-value: < 2.2e-16
```

Which of the following sentences is false:

Scegli un'alternativa:

- a. The comparison of the fitted model with the null model is described by the F statistics, whose observed value is 43.92 with $p - p_0$ and $n - p$ degrees of freedom.
- b. The interaction between `Wind` and `Temp` is statistically significant.
- c. At level $\alpha = 0.01$ the variable `Month` is not statistically significant.
- d. The estimated model for the i-th unit is $y_i = -235.52 + 0.057\text{Solar. R}_i + 14.02\text{Wind}_i + 4.06\text{Temp}_i - 2.69\text{Month}_i$. ✓

La risposta corretta è: The estimated model for the i-th unit is $y_i = -235.52 + 0.057\text{Solar. R}_i + 14.02\text{Wind}_i + 4.06\text{Temp}_i - 2.69\text{Month}_i$.

Domanda 3

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

Assume you want to study the proportion of patients that positively react to a drug. In a previous test the drug was effective on the 80% of patients. If you want to include this information in your Bayesian analysis of the data, which prior do you think is most appropriate among the following ones?

Scegli un'alternativa:

a. `curve(dunif(x, 0, 100), from = 0, to = 100)`

b. `curve(dnorm(x, 0.8, 2), from = -5, to = 6)`

c. `curve(dgamma(x, 1, 1))`

d. `curve(dbeta(x, 5, 2))` ✓

La risposta corretta è: `curve(dbeta(x, 5, 2))`

Domanda 4

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

Suppose from a linear model we get that the log-likelihood evaluated in the MLE $\hat{\theta}$ is $l(\hat{\theta}) = -25.1$. Suppose then to have Akaike information criterion, AIC = 74.2. Which of the following sentences is true?

Scegli un'alternativa:

a. A model with 7 parameters and the same value of $l(\hat{\theta})$ is worse than the current one.

b. The number of parameters of the model is $p = 24$.

c. The value $l(\hat{\theta}) = -25.1$ alone is indicative of a satisfactory goodness of fit.

d. A model with 13 parameters and the same value of $l(\hat{\theta})$ is worse than the current one. ✓

La risposta corretta è: A model with 13 parameters and the same value of $l(\hat{\theta})$ is worse than the current one.

Domanda **2**

Risposta corretta

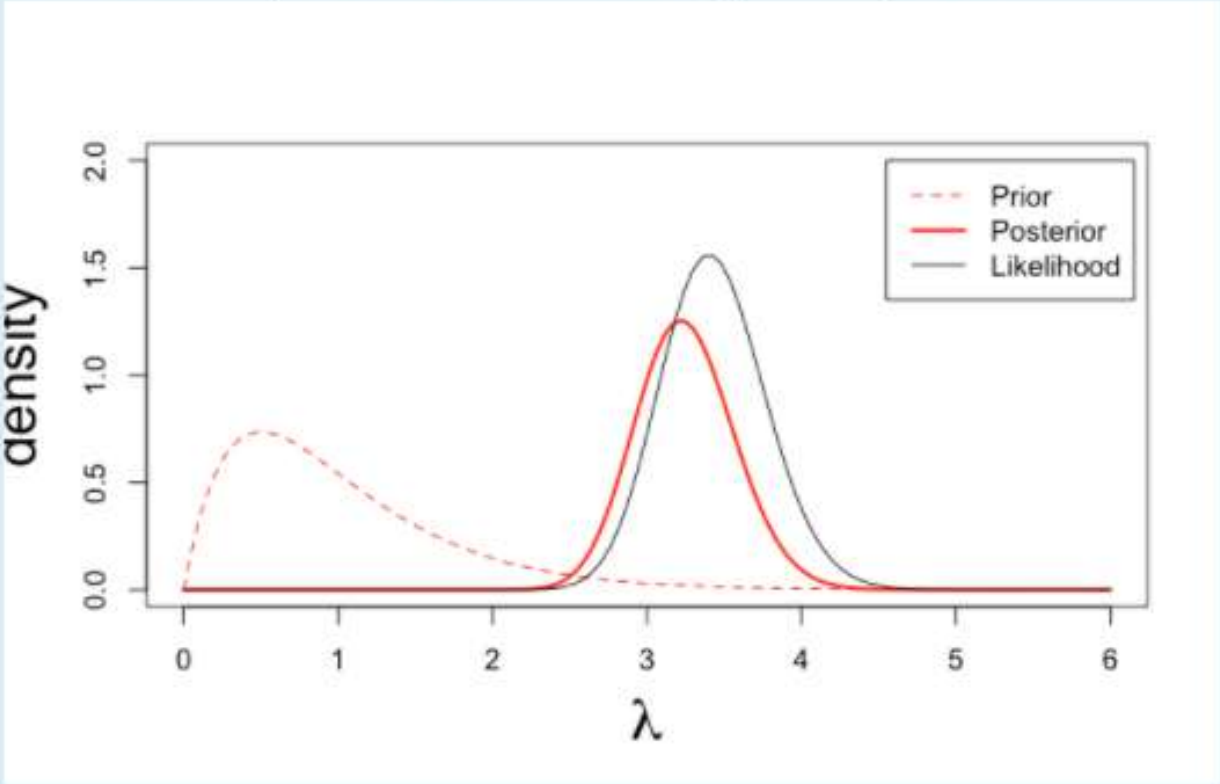
Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

Suppose $y_1, \ldots, y_{30} \sim i.i.d.\ Po(\lambda)$ and the following prior density is proposed on the parameter λ

$\pi(\lambda) = \frac{4}{\Gamma(2)} \lambda e^{-2\lambda}, \quad \lambda > 0$

We are interested in the posterior distribution. From the following plot what can you conclude?



Scegli un'alternativa:

- a. The posterior is a $Po(\lambda + \sum y_i)$.
- b. The prior mean is greater than the posterior mean.
- c. The posterior mode is closer to the prior mode than to the MLE.
- d. The posterior mean is $(2 + \sum y_i)/32$ ✓

La risposta corretta è: The posterior mean is $(2 + \sum y_i)/32$

Domanda 6

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Contrassegna domanda

Suppose a statistical test $H_0 : \lambda = 1$ against $H_1 : \lambda = 2$ for a sample $(y_1, \ldots, y_n)$ from an exponential model of the form $f(Y; \lambda) = \lambda e^{-\lambda y}$ is posed. For a specified threshold k_α, the rejection region of the Neyman-Pearson likelihood ratio-test is then:

Scegli un'alternativa:

a. $\lambda(\boldsymbol{Y}) = 2^n e^{-\sum_{i=1}^n (y_i - 2)^2} \geq k_\alpha$.

b. $\lambda(\boldsymbol{Y}) = 2^n e^{-\sum_{i=1}^n y_i} \geq k_\alpha$.

c. $\lambda(\boldsymbol{Y}) = \frac{1}{2} e^{-\sum_{i=1}^n y_i} \leq k_\alpha$.

d. $\lambda(\boldsymbol{Y}) = e^{\sum_{i=1}^n y_i^2} \geq k_\alpha$.

La risposta corretta è: $\lambda(\boldsymbol{Y}) = 2^n e^{-\sum_{i=1}^n y_i} \geq k_\alpha$.