EXAM OF STATISTICS
(DESCRIPTIVE STATISTICS AND REGRESSION)

2nd Physiotherapy | Version A | **June, 06 2022**

Duration: 1 hour.

(5 pts.) 1. The patients of a physiotherapy clinic were asked to assess their satisfaction in a scale from 0 to 10. The assessments are summarized in the table below.

Assessment	Patients
$0-2$	3
$2-4$	12
$4-6$	9
$6-8$	18
$8-10$	22

(a) Compute the interquartile range of the assessment and interpret it.

(b) If it is required an assessment greater than 5 in more than 50% of patients for the clinic to remain open, will the clinic remain open?

(c) Is the assessment mean representative?

(d) Compute the coefficient of kurtosis of the assessment and interpret it. Is the kurtosis normal?

(e) If the assessment mean of another clinic is 6.8 and the standard deviation is 2.6, which assessment is relatively higher 6 in the first clinic or 6.2 in the second?

Use the following sums for the computations:
$\sum x_i n_i = 408$, $\sum x_i^2 n_i = 3000$, $\sum (x_i - \bar{x})^3 n_i = -548.25$ and $\sum (x_i - \bar{x})^4 n_i = 5140.45$.

Solution

Let X be the patient assessment.

(a) $Q_1 = 4.4444$, $Q_3 = 9.0907$ and $IQR = 4.6463$, so the central dispersion is moderate.

(b) $F(5) = 0.2695$, and the percentage of patients with an assessment greater than 5 is 73.05%.

(c) $\bar{x} = 6.375$, $s_x^2 = 6.2344$, $s_x = 2.4969$ and $cv = 0.3917$, thus the representativity of the mean is moderate.

(d) $g_2 = -0.9335$ and the distribution is flatter than a Gauss bell, but normal, as g_2 is between -2 and 2.

(e) First clinic: $z(6) = -0.1502$
Second clinic: $z(6.2) = -0.3077$.
Thus, an assessment of 6 in the first clinic is relatively higher as its standard score is greater.

(5 pts.) 2. A study tries to determine the effectiveness a training program to increase the grip strength. The table below shows the grip strength in Kg in some weeks of the training program.

Week	1	3	6	9	14	17	21	24
Grip strength	15	22	29	34	36	39	40	41

(a) Compute the regression coefficient of the grip strength on the weeks and interpret it.

(b) According to the logarithmic regression model, what is the expected grip strength after 5 and 25 weeks. Are these predictions reliable? Would these predictions be more reliable with the linear regression model?

(c) According to the exponential regression model, how many weeks are required to have a grip strength of 25 Kg?

(d) What percentage of the total variability of the weeks is explained by the exponential model?

Use the following sums (X=Weeks and Y=Grip strength):
$\sum x_i = 95$, $\sum \log(x_i) = 16.7824$, $\sum y_j = 256$, $\sum \log(y_j) = 27.3423$,
$\sum x_i^2 = 1629$, $\sum \log(x_i)^2 = 43.606$, $\sum y_j^2 = 8804$, $\sum \log(y_j)^2 = 94.3237$,
$\sum x_i y_j = 3552$, $\sum x_i \log(y_j) = 342.9642$, $\sum \log(x_i) y_j = 608.4186$, $\sum \log(x_i)\log(y_j) = 60.047$.

Solution

(a) $\overline{x} = 11.875$ weeks, $s_x^2 = 62.6094$ weeks^2.
$\bar{y} = 32$ Kg, $s_y^2 = 76.5$ Kg^2.
$s_{xy} = 64$ weeks·Kg.
Regression coefficient of Y on X: $b_{yx} = 1.0222$ Kg/weeek. The grip strength increases 1.0222 Kg per week.

(b) $\overline{\ln(x)} = 2.0978$ ln(weeks), $s_{\ln(x)}^2 = 1.05$ ln(weeks)^2 and $s_{\ln(x)y} = 8.9226$ ln(weeks)Kg.
Logarithmic regression model of Y on X: $y = 14.1729 + 8.498\ln(x)$.
Predictions: $y(5) = 27.8499$ Kg and $y(25) = 41.5268$ Kg.
Logarithmic coefficient of determination: $r^2 = 0.9912$. The predictions are not reliable because the sample size is small.
Linear coefficient of determination: $r^2 = 0.8552$.
As the linear coefficient of determination is less than the logarithmic one, the predictions with the logarithmic model are more reliable.

(c) Exponential regression model of X on Y: $x = e^{-1.6345+0.1166y}$.
Prediction: $x(25) = 3.6015$ Weeks.

(d) As $r^2 = 0.9912$, the exponential models explains 99.12% of the variabilidty of the weeks.
