EXAM OF STATISTICS
(DESCRIPTIVE STATISTICS AND REGRESSION)

Pharmacy/Biotechnology 1st year — Version A — **February, 8 2021**

Duration: 1 hour.

(5 pts.) 1. Laetisaric acid is a compound that can be used to control fungal diseases in some plantations. The table below shows the radial growth (in cm) of the fungus Pleurotus ostreatus exposed to different concentrations of laetisaric acid (ing mg/l).

Laetisaric acid (mg/l)	0.0	0	3.0	3.0	6	6	10.0	10.0	20.0	20.0	30.0	30
Fungus growth (cm)	33.3	31	29.8	27.8	28	29	25.5	23.8	18.3	15.5	11.7	10

(a) Compute the equation of the regression line that best explains fungal growth as a function of acid concentration.

(b) Compute the linear correlation and determination coefficients and interpret them.

(c) How much does the fungal population increase or decrease for each one mg/dl increment in acid concentration?

(d) What are the expected fungal growths for acid concentrations of 20 and 200 mg/l? Which of these predictions is more reliable?

Use the following sums for the computations: $\sum x_i = 138$ mg/l, $\sum y_i = 283.7$ cm, $\sum x_i^2 = 2890$ (mg/l)^2, $\sum y_i^2 = 7384.49$ cm^2 and $\sum x_i y_j = 2334.8$ mg/l$\cdot$ cm.

Solution

(a) $\bar{x} = 11.5$ mg/dl, $s_x^2 = 108.5833$ (mg/dl)^2,
$\bar{y} = 23.6417$ cm, $s_y^2 = 56.4458$ cm^2,
$s_{xy} = -77.3125$ mg/dl $\cdot$ cm.
Regression line of fungal growth on acid concentration: $y = 31.8298 - 0.712x$.

(b) $r^2 = 0.9752$ and $r = -0.9875$. As the linear correlation coefficient is close to -1 there is a strong inverse relation between the acid concentration and the fungal growth.

(c) The fungi will decrease 0.712 for each one mg/l increment in the acid concentration.

(d) $y(20) = 17.5896$ cm and $y(200) = -110.5724$ cm. The prediction for 20 mg/dl is more reliable because 20 is in the range of values in the sample.

(5 pts.) 2. The table below shows the blood uric acid concentration of 8 men and 10 women in mg/dl.

Men (X)	4.7	3.6	5.2	6.8	9.5	4.8	5.6	5.4		
Women (Y)	3.2	4.5	5.4	2.1	6.7	5.2	3.8	4.3	7.2	2.6

Se pide:

(a) In which group, men or women, is the mean more representative?

(b) In which group, men or women, is the uric acid distribucion more simmetric?

(c) In which group, men or women, is the uric acid distribution flatter?

(d) Can we assume that the uric acid sample from women comes from a normal population?

(e) What uric acid concentration must a woman have to be relative higher than 6 mg/dl in a man?

(f) If all the values of the variable X are multiplied by a number, what must that number be to make the mean of the new variable as representative as the mean of Y?

Use the following sums for the computations:
Men: $\sum x_i = 45.6$ mg/dl, $\sum x_i^2 = 282.14$ (mg/dl)^2, $\sum (x_i - \bar{x})^3 = 45.06$ (mg/dl)^3 and $\sum (x_i - \bar{x})^4 = 231.15$ (mg/dl)^4.
Women: $\sum y_i = 45$ mg/dl, $\sum y_i^2 = 227.52$ (mg/dl)^2, $\sum (y_i - \bar{y})^3 = 8.17$ (mg/dl)^3 and $\sum (y_i - \bar{y})^4 = 126.77$ (mg/dl)^4.

Solution

(a) Men: $\bar{x} = 5.7$ mg/dl, $s^2 = 2.7775$ (mg/dl)^2, $s = 1.6666$ mg/dl and $cv = 0.2924$.
Women: $\bar{y} = 4.5$ mg/dl, $s^2 = 2.502$ (mg/dl)^2, $s = 1.5818$ mg/dl and $cv = 0.3515$.
Thus, the mean of uric acid is more representative in men than in women as the coefficient of variation is smaller.

(b) $g_{1x} = 1.2168$ and $g_{1y} = 0.2065$. Thus, the distribution of uric acid in men is more symmetric as the coefficient of skewness is closer to 0.

(c) $g_{2x} = 0.7454$ and $g_{2y} = -0.9749$. Thus, the distribution of uric acid in men is flatter as the coefficient of kurtosis is smaller.

(d) As the coefficient of skewness and the coefficient of kurtosis are between -2 and 2 we can assume that the uric acid sample of women comes from a normal population.

(e) The uric acid must be at least 4.7847.
