< Statistics >

1. [30 points] As soon as a woman is pregnant, her body begins producing the pregnancy hormone called human Chorionic gonadotropin(hCG). A home pregnancy test is designed to detect hCG in urine and produce a positive result for pregnancy. Although many home pregnancy tests claim high accuracy, they are not perfect. Suppose that a pregnancy test correctly detects pregnancy 90% of the time for women who are pregnant. Also, the pregnancy test mistakenly detects pregnancy 20% of the time for women who are not pregnant. Suppose that a newly wed woman has a 10% probability of being pregnant.

(a) Suppose that a newly wed women takes 10 independent pregnancy tests. Compute the probability that at least one of the 10 pregnancy test has a positive result.

(b) Compute the probability that a newly wed woman is pregnant given that a pregnancy test result is positive.

(c) Compute the probability that a newly wed woman is pregnant given that two independent pregnancy test results are both positive.

2. [70 points]

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$

b_0, b_1 : coefficients of regression

$\hat{y}_i = b_0 + b_1 x_i$: fitted line

y_i : response

x_i : covariate

(a) Derive b_0, b_1

(b) Derive $Var(b_1) = \frac{\sigma^2}{\sum_{j=1}^{n}(x_j - \bar{x})^2}$

(c) Derive $Var(b_0)$

(d) Construct a $(1-\alpha)100\%$ C.I. for the mean response with covariate x_0.

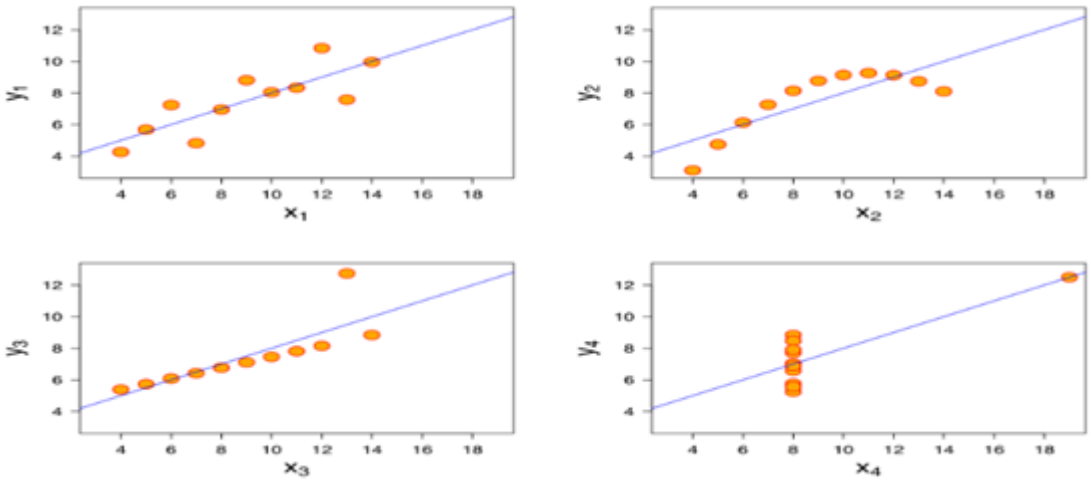
Hint : Let i be a student t random variable with ν degrees of freedom. Then the t-value such that the area in the right tail is p is denoted by $t_{\beta,\nu}$ with $P(T > t_{\beta,\nu}) = \beta$

(e) The leverant h_i of the i^{th} observation measures how far the value of the observation is from the mean of the values of the other observations and is defined as leverage $h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{j=1}^{n}(x_j - \bar{x})^2}$.

Prove that $Var(r_i) = \sigma^2 h_i$ for the residual $r_i = y_i - \hat{y}_i$.

(f) Explain why the variance of the i^{th} residual and the variance of the fitted value of the response with covariate x_i in part (e) justify the use of h_i as a measure of leverage from the i^{th} observation. Is it possible to have $h_i = 1$? If possible, when? If not, why?

(g) Below is the Anscomle's quarter that comprises four data sets with nearly identical simple descriptive statistics but with very different shapes when graphed. The closed circles represent data points, while the dashed lines represent fitted linear regression lines when an influential point is defined as a data point with high leverage that greatly affects the slope of a regression line, choose a figure that contains an influential point, identify the point in the figure and explain why, by redistributing the other points in the chosen figure and drawing a sketch. Provide an example where a high leverage point is not an influential point.



< Mathematical Economics >

1. [70 points] Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 5 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 4 & 0 \\ 1 & 2 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$. Compute.

(a) arithmetic mean of the eigenvalues of B^{-1}

(b) geometic mean of the eigenvalues of B^{-1}

(c) harmonic mean of the eigenvalues of B^{-1}

(d) $|CAC'|$

(e) $|A \otimes B|$

(f) Let $E = C(C'C)^{-1}C'$. Compute the difference between maximum and minimum of the eigenvalues of E.

2. [30 points] Let $f(x_1, x_2) = |x_1| + |x_2|, f: \mathbb{R}^2 \to \mathbb{R}$.

(a) Find the maximum of f subject to $x_1^2 + x_2^2 \leq 1$. (Explain with reasons)

(b) Find the minimum of f subject to $x_1^2 + x_2^2 \leq 1$. (Explain with reasons)

(c) Find the maximum of f subject to $x_1 \geq -1, x_2 \geq -1, 2x_1 + x_2 \geq 1$. (Explain with reasons)