Quiz 2 for Statistics 301
Elementary Statistical Methods - Spring 1999
Material Covered: Chapter 4 of notes; Chapter 3 of text
For: 12th February

Name (please print): _____.
last first

---

1. Consider the following data relating number of patients per 1000 who were diagnosed with colon cancer, x, and the quantity of vegetables and fruit consumed, y.

x	12	24	17	28	24	36	20
y	43	36	25	23	32	17	24

(a) [1] Quantity of vegetables and fruit consumed is the (check none, one or more) **independent variable** / **dependent variable** / **response variable** / **explanatory variable**

(b) [1] Number of patients with colon cancer is (circle one) **positively** / **not** / **negatively** associated with quantity of vegetables and fruit consumed.

(c) [1] The scatter plot could be described as a (circle one) **positive moderately linear** / **negative moderately linear** / **negative strongly linear** / **positive strongly linear**

(d) [1] **True** / **False** The data shows that an increase in the consumption of vegetables and fruits causes a decrease in the number of patients with colon cancer.

(e) [1] The linear regression is given by $\hat{y} = a + bx$,

where $a =$ _____.

(f) [1] The residual at point $(17, 25)$ is (circle one) $\mathbf{-10.2}$ / $\mathbf{-8.4}$ / $\mathbf{-3.2}$ / $\mathbf{5.8}$ / $\mathbf{12.4}$.

(g) [1] If the scatter plot could be fit in a rectangular area which was four times long as it was wide, then the correlation coefficient could be estimated to be

$r \approx 1 - \frac{1}{k^2} =$ _____.

(a) [1] Quantity of vegetables and fruit consumed is the **independent variable**, **explanatory variable**

(b) [1] Number of patients with colon cancer is **negatively** associated with quantity of vegetables and fruit consumed.

(c) [1] The scatter plot could be described as a **negative moderately linear**

(d) [1] **False** The data shows that an increase in the consumption of vegetables and fruits causes (does not *cause* necessarily, but is associated with) a decrease in the number of patients with colon cancer.

(e) [1] The linear regression is given by $\hat{y} = a + bx$, where $a = 47.06$.

(f) [1] The residual at point $(17, 25)$ is $\boldsymbol{y - \hat{y} = 25 - [47.06 - 0.80(17)] \approx -8.4}$

(g) [1] If the scatter plot could be fit in a rectangular area which was four times long as it was wide, then the correlation coefficient could be estimated to be $r \approx 1 - \frac{1}{k^2} = 0.9375$.