Chapter 15 Inference about Regression

(11) $\hat{y} = a + bx$ $\quad \hat{y} = -3.6596 + 1.1969x$

a) length of humerus $= -3.6596 + 1.1969$ (length of femur)

b) $t = \dfrac{b}{SE_b}$ $\quad t = \dfrac{1.1969}{.0751} = 15.94$

c) $df = n - 2 = 5 - 2 = 3$

d) $15.94 > 12.92$ $\quad$ p-value $< .0005$

e) $df = 3$ $\quad t^* = 5.841$ $\quad$ 99% conf int for β true slope

$$b \pm t^* SE_b$$

$1.1969 \pm 5.841(.0751)$

$1.1969 \pm .4389$ $\quad (.7582, 1.6356)$

We are 99% confident that for every extra cm in femur length, the humerus length will increase on average between 0.7582 cm and 1.6356 cm.

(12) a) $r^2 = 99.8\%$ very close to $\underline{1}$; almost perfect linear relationship

b) the slope parameter β $\quad \hat{y} = 1.76608 + .080284x$

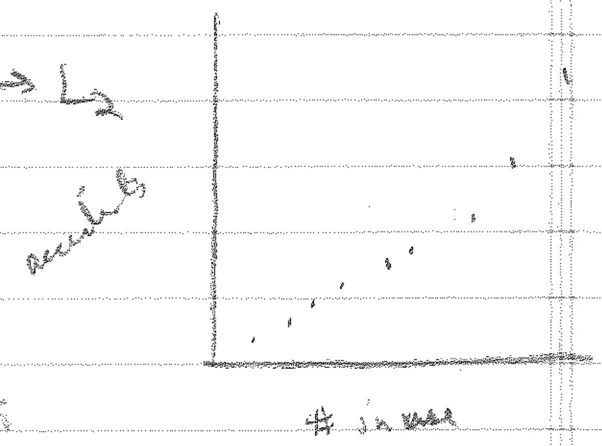
99% conf int for β $\quad 0.080284 \pm 4.032(.0016)$

$t^* = 4.032$ $\quad .080284 \pm .00645$

$b \pm t^* SE_b$ $\quad (.0738, .0867)$

We are 99% confident that the true rate at which steps per second increase as running speed increases by 1 f/s is on average between 0.0738 and 0.0867.

(13) a) # in use → L_1, accidents → L_2



Strong positive linear association between # of jet skis in use and # of accidents

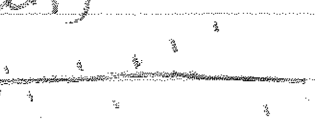
b) $H_0: \beta = 0$ there is not association between # of jet skis in use and # of accidents

$H_a: \beta > 0$ there is a positive association between number of jet skis in use and # of accidents

c) • Independence –

• Linear relationship exists between # of jet skis in use and # of accidents (see scatterplot)

• Standard deviation



• Normal – histogram of residuals

d) Linear Regression t test: $t = 21.08$ $df = 8$

p-value = .00000013

There is strong evidence to reject the null hypothesis and conclude there is a significant positive linear association between # of jet skis in use and # of accidents

e) 98% conf int for β: $.0048 \pm 2.896(SE_b) = .0048 \pm 2.896(.00023)$

$df = 8$ $t^* = 2.896$

$.0048 \pm .0007$

$t = \frac{b}{SE_b}$ $21.08 = \frac{.0048}{SE_b}$

$SE_b = .00023$

$(.0041, .0055)$

p 909

(14) a) Parameter of interest: β = the average rate of change in heart disease death rate

Test: Inference for regression

Conditions: 1) Independence

2) Constant Variance

3) Linear Relationship

4) Normality

Must post both L1 + L2

$H_0: \beta = 0$ No association between consumption of wine and deaths from heart disease

$H_a: \beta < 0$ there is a negative association between consumption of wine and deaths from heart disease

Test statistic: $t = \frac{b}{SE_b}$ $b = -22.969$

$SE_b = \frac{-22.969}{-6.457} = 3.557$

LinRegTTest: $t = -6.457$

$df = 17$

$P(t < -6.457) = 0.000003$ Small p-value.

Reject Null. Conclude there is a negative association...

b) 95% conf int for β: $b \pm t^* SE_b$

$t^*_{.95}(17) = 2.110$ $-22.969 \pm 2.110(3.557)$

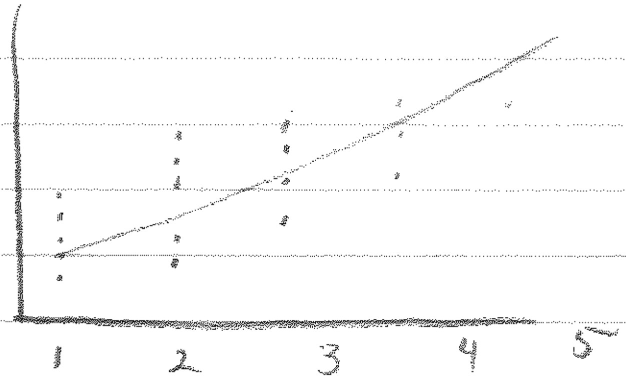
-22.969 ± 7.505

$(-30.474, -15.464)$ (per 100,000)

We are 95% conf that number of deaths from heart disease decreases on average between 15.4 and 30.5 for each additional liter of wine consumed per person

15.

a)



moderately strong positive linear association between . . .

b) $\hat{y} = -1.286 + 11.894x$

# of beetle larvae = −1.286 + 11.894 (# of beaver-caused stumps)

$r^2 = 83.9\%$

$H_0: \beta = 0$ there is no association between

$H_a: \beta \neq 0$ there is an association

Conditions:

c) $t = 10.47 = \dfrac{11.894}{1.136} = \dfrac{b}{SE_b}$

$2P(t > 10.47) = 0.0000$

$df = 21$

Significant

Beaver stump counts explain beetle larvae counts.