

سوالات تمرین سری اول :

4.2) a \rightarrow Height = 65 inches \rightarrow weight = $99.41 + 3.94 \times 65 = 156.69$ pounds.

b Height = 70 inches \rightarrow weight = $99.41 + 3.94 \times 70 = 176.39$ Pounds

Height = 74 inches \rightarrow weight = $99.41 + 3.94 \times 74 = 192.15$ Pounds.

c \rightarrow 1.5 اینچ به قدر مرد اضافه شده، وزن $3.94 \times 1.5 = 5.91$ Pounds زیاد شده.

c \rightarrow

4.5) a \rightarrow چون در مدل تنها دایره‌های سطح زمان موجود برای استنتاج بر روی مدل قرار می‌گیرد و ورودی هر مرد در نظر گرفته شده، عواملی مانند توانایی مردی (سطح آمادگی) دانشجو برای امتحان، اعتماد بنفس، سلامت روحی و جسمی و... لحاظ نشده است. پس هر یک از این عوامل در جمله a یا b همان خط خطیه است. علت تفاوت طول بالاها بین دانشجویان مصافحه نشده هم است.

ب - زیرا اگر مدل رگرسیونی ما خواهد بر اساس خوبی از داده‌های ما باشد باید دو شرط زیر را داشته باشد -

تقریباً برای X متغیری دارد - ii) $E(u|X) = E(u)$ یعنی میانگین مربعات متغیرها مشاهده شده $E(u) = 0$

با اجماع در مورد گفته خواهیم داشت: $E(u|X) = E(u) = 0$

ک. این مال به صورتی دیگر از سوال (ب) می‌آید -

ب - $E(u_i | X_i)$ مقدار خط در نمونه i از X_i نشان می‌دهد (مقادیر مختلف خط X_i مقدار مختلف X_i)

خبرایب تخمین زده شده نادرست نیستند زیرا مقادیر مختلف در خط X_i در مقدار X_i و Y_i متغیر است مانند توانایی فردی، سطح مدارس، هوش، استعداد و ...

ج - این است که در نمونه i از مدل X_i گرفته شده می‌تواند تخمین بزند و جواب با کیفیت نداشته باشد و ...

د - i) $X_1 = 60 \text{ mins}$ so $Y_1 = 55 + 0.17 \times 60 = 65.2$ $X_2 = 75 \text{ mins}$ so $Y_2 = 55 + 0.17 \times 75 = 67.75$

$X_3 = 90 \text{ mins}$ so $Y_3 = 55 + 0.17 \times 90 = 70.3$

Estimated gain = $0.17 \times 5 = 0.85$

4.6)

تساوی $E(u_i | x_i) = 0$ نشان می‌دهد که میانگین متغیر خطای برای مقدار ثابت x برابر با صفر است.
در نتیجه برای هر x_i ، $E(u_i) = 0$ و در نتیجه مجموع خطاها برای هر x برابر با صفر است.
و در نتیجه $E(u_i) = 0$

$$Y_i = \beta_0 + \beta_1 x_i + u_i$$

حال مدل رگرسیون خطی را به شکل دوم داریم

$$E(Y_i) = E(\beta_0) + E(\beta_1 x_i) + E(u_i) \Rightarrow E(Y_i | x_i) = \beta_0 + \beta_1 x_i$$

پس β_0 و β_1 ضرایب رگرسیون خطی هستند. Y_i و x_i متغیرهای وابسته و مستقل هستند.
خواهیم داشت

$$Y_i = \beta_0 + \beta_1 x_i + u_i$$

اگر تعداد نمونه‌ها n باشد داریم: $Y_1 = \beta_0 + \beta_1 x_1 + u_1, \dots, Y_n = \beta_0 + \beta_1 x_n + u_n$
ما جمع هر دو طرف و تقسیم آن‌ها بر n خواهیم داشت: $\sum_{i=1}^n Y_i / n = \sum_{i=1}^n (\beta_0 + \beta_1 x_i + u_i) / n = \bar{Y} = n\beta_0 / n + n\beta_1 \sum_{i=1}^n x_i / n$
چون متغیر رگرسیون صفر است

$$\frac{\sum_{i=1}^n u_i}{n} = \beta_0 + \beta_1 \bar{x} + E(u) = \beta_0 + \beta_1 \bar{x} \Rightarrow \bar{Y} = \beta_0 + \beta_1 \bar{x} \Rightarrow (\bar{x}, \bar{Y}) \text{ نقطه‌ای}$$

خط برازش شده قرار دارد (خطی وجود ندارد)

(2)

$$4.2) a \rightarrow \text{Height} = 64 \text{ in} \Rightarrow \text{Weight} = -79.24 + 4.16 \times 64 = 187.24 \text{ pounds (lbs)}$$

$$(ii) \text{Height} = 68 \text{ in} \Rightarrow \text{Weight} = -79.24 + 4.16 \times 68 = 203.64 \text{ (lbs)}$$

$$(iii) \text{Height} = 72 \text{ in} \Rightarrow \text{Weight} = -79.24 + 4.16 \times 72 = 220.28$$

b \rightarrow چون میزان افزایش وزن مشخص است دیگر ضریب ثابت (-79.24) مطابقت نمی‌شود.

$$\Delta \text{Weight} = 2 \cdot 4.6 = 9.2 \text{ lbs}$$

$$c \rightarrow (i) 64 \text{ in} = 162.56 \text{ cm} \quad \text{از نقاط نمونه / داده‌ها استفاده کرده و آن‌ها را تبدیل کرده ایم.}$$

$$187 \text{ lbs} = 84.8218 \text{ kgs}$$

$$(ii) 68 \text{ in} = 172.72 \text{ cm}, 203.64 \text{ lbs} = 92.3695502 \text{ kgs}$$

در نتیجه می‌توان ضریب β_1 در معادله رگرسیونی $Y = \beta_0 + \beta_1 x$ درست آورد.

$$\beta_1 = \frac{W_{(ii)} - W_{(i)}}{H_{(ii)} - H_{(i)}} = \frac{92.3695502 - 84.8218}{172.72 - 162.56} = 0.743$$

با قرار دادن β_1 در معادله رگرسیونی β_1 نیز بدست می‌آید.

$$84.8218 = \beta_0 + 0.743 \times 162.56 \Rightarrow \beta_0 = -35.96$$

نیت

✓

از آن جا که بر مبنای واحد اندازه گیری را تفسیر داریم و نفع بر از سن ما تفسیر می آید. مقدار R^2 به برابر داده ها توضیح داده شده با خطا به کل داده ها است. تفسیر می کنند پس $R^2 = 0.72$ اما مقدار SER حول وابسته به SSR یا همان مجموع مربعات باقی مانده ها است. $(SER = \frac{SSR}{n-2})$ یا تبدیل واحد تفسیر می کنند. از آن جا که هر متغیر مستقل وزن است پس این مقدار باید در نسبت $\frac{1}{n-2}$ یکپارچه ضرب شود. یعنی $SER = 12.6 \times 0.45 = 5.67$ به صورت