

علیرضا درویشی - تمرین 6 - سنجی دوره فنی - 96109674

$$\arg \min_{\beta_0, \beta_1, \dots, \beta_n} \sum_i (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_n x_{in})^2 = \hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_n$$

$$\arg \min_{\tilde{\beta}_0, \tilde{\beta}_1, \dots, \tilde{\beta}_n} \sum_i (c_0 y_i - \tilde{\beta}_0 - \tilde{\beta}_1 c_1 x_{i1} - \dots - \tilde{\beta}_n c_n x_{in})^2$$

$$= \arg \min_{\tilde{\beta}_0, \tilde{\beta}_1, \dots} \sum_i c_0^2 (y_i - \frac{\tilde{\beta}_0}{c_0} - \frac{\tilde{\beta}_1 c_1}{c_0} x_{i1} - \dots - \frac{\tilde{\beta}_n c_n}{c_0} x_{in})^2$$

$$= \arg \min_{\tilde{\beta}_0, \tilde{\beta}_1, \dots, \tilde{\beta}_n} \sum_i (y_i - \frac{\tilde{\beta}_0}{c_0} - \frac{\tilde{\beta}_1 c_1}{c_0} x_{i1} - \dots - \frac{\tilde{\beta}_n c_n}{c_0} x_{in})^2$$

$$\arg \min_{\gamma_0, \gamma_1, \dots, \gamma_n} \sum_i (y_i - \gamma_0 - \gamma_1 x_{i1} - \dots - \gamma_n x_{in})^2 = \hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_n$$

$$\Rightarrow \gamma_0 = \hat{\beta}_0 \Rightarrow \tilde{\beta}_0 = c_0 \hat{\beta}_0$$

$$\gamma_1 = \hat{\beta}_1 \Rightarrow \tilde{\beta}_1 = \frac{c_0 \hat{\beta}_1}{c_1}, \gamma_2 = \hat{\beta}_2 \Rightarrow \tilde{\beta}_2 = \frac{c_0 \hat{\beta}_2}{c_2}$$

$$\gamma_n = \hat{\beta}_n \Rightarrow \tilde{\beta}_n = \frac{c_0 \hat{\beta}_n}{c_n}$$

-3

$$0.0003 - 0.00000007 \text{ Sale} \times 28 \leq 0$$

(i)

$$\Rightarrow 21428 \leq \text{Sale}$$

(22) خیر چون به صورت فیزی این مدل می تواند نتایج خیلی کم یا حتی منفی

داشته باشد. از طرفی ضریب توان دوم هم با سطح اطمینان 95٪ نمی تواند
رضی منفر بودن

رد شود. پس بهتر است مدل لگاریتمی غایت شد.

(222)

$$rdintens = 2.613 + 0.3 \text{ Sales} - 0.007 \text{ Sales}^2$$

(0.429)

(0.14)

(0.0037)

$$n = 32, R^2 = 0.1484$$

(23) معادله پیش قبل راحت تر است. چون ضرایب و ضرایب استاندارد

دارند هر چند منفی هر معادله یکی است.

نحل 7:

$$t\text{-score} = \frac{2.19}{0.53} = -4.13$$

(i) -30

ضریب $hsize^2$ با اطمینان بیش از 99 درصد منفی است $\Rightarrow p\text{-value} < 0.01$

$$19.3 - 2 \times 2.19 hsize = 0 \Rightarrow hsize = \underset{opt}{4.41}$$

$$\text{nonblack-female: } \hat{sat} = C - 45.09 - 0 + 0$$

(ii)

$$\text{nonblack-male: } \hat{sat} = C - 0 - 0 + 0$$

$$\Rightarrow \text{diff} = -45.09 \rightarrow t\text{-score} = \frac{-45.09}{4.29} = -10.51$$

فصلی مذکور با سطح اطمینان 99.9٪ منفی است. $\Rightarrow p\text{-value} < 0.001$

$$\text{black-male: } \hat{sat} = C - 0 - 169.81 + 0$$

(iii)

$$\text{nonblack-male: } \hat{sat} = C - 0 - 0 + 0$$

$$\text{diff} = 169.81 \rightarrow t\text{-score} = \frac{169.81}{12.71} = 13.36$$

فصلی مذکور با سطح اطمینان 99.9٪ مثبت است. $\Rightarrow p\text{-value} < 0.001$

JAHAN NAMA

(iv)

$$\text{black-female: } \hat{S}_{at} = C - 45.09 - 169.81 + 62.31$$

$$\text{nonblack-female: } \hat{S}_{at} = C - 45.09 - 169.81 - 62.31 - 0 + 0$$

$$\Rightarrow \text{diff} = 107.5$$

برای حساب کردن $SE(\text{diff})$ نیاز به $Cov(d_2, d_3)$ داریم:

$$SE(d_1 + d_2) = \sqrt{\text{Var}(d_1) + \text{Var}(d_2) + 2Cov(d_1, d_2)}$$

(i) 4

$$\text{utility: } \log(\hat{\text{Salary}}) = C + 0 + 0 - 0.283$$

$$\text{trans: } \log(\hat{\text{Salary}}) = C + 0 + 0 - 0$$

$$\log \text{diff} = -0.283$$

\Rightarrow اختلاف حقیقی حدود (-28.3) درصد است.

$$t\text{-score} = \frac{-283}{99} = -2.86 \Rightarrow P\text{-value} < 0.01$$

فرض صفر، منجر به اختلاف، با اطمینان 99٪ رد می شود.

JAHAN NAMA

$$100 (e^{-0.283} - 1) = -24.6$$

(ii)

این عدد، دقیق‌تر است. در واقع عدد بخش قبلی، تقریب خطی است.

$$\text{Cons: } \log(\hat{\text{Salary}}) = C + 0 + 0.181 - 0$$

(22)

$$\text{Fin: } \log(\hat{\text{Salary}}) = C + 0.158 + 0 + 0$$

اختلاف در این حدود 23% است.

$$\text{diff} = 0.023 \Rightarrow$$

برای معاسری دقت آماری داریم،

$$SE(d_{\text{Cons}} - d_{\text{Fin}}) = \sqrt{\text{Var}(d_{\text{Cons}}) + \text{Var}(d_{\text{Fin}}) - 2\text{Cov}(d_{\text{Cons}}, d_{\text{Fin}})}$$

$$t\text{-score} = \frac{\text{diff}}{SE(\text{diff})} \rightarrow \text{معاسری P مقدار}$$

: سپس

(i) - 8

$$\hat{\log(\text{salary})} = \alpha_0 + \alpha_1 \text{Smoke} + \alpha_2 \text{ed} + \alpha_3 \text{exp} + \alpha_4 \text{female}$$

(ii) با آزمون فرض: $H_0: \alpha_5 = 0$ می توان تفاوت بین زن و مرد رسید.

$$\hat{\log(\text{salary})} = \alpha_0 + \alpha_1 \text{Smoke} + \alpha_2 \text{ed} + \alpha_3 \text{exp} + \alpha_4 \text{female} + \alpha_5 \text{female} \cdot \text{Smoke}$$



(iii)

$$\hat{\log(\text{salary})} = \alpha_0 + \alpha_1 \text{moderate} + \alpha_2 \text{Light} + \alpha_3 \text{heavy} + \alpha_4 \text{ed} + \alpha_5 \text{exp} + \alpha_6 \text{female}$$

(iv) با انجام آزمون F برای ضرایب $\alpha_1, \alpha_2, \alpha_3$ می توان فهمید که

آیا مصرف وید روی حقوق تأثیر دارد یا خیر

(v)

(v) شاید افرادی که وید مصرف می کنند، تمایل به  خدا افکاری کمتری داشته باشند.باشند. شاید افراد با ed بیشتر  خدا افکاری واقعی تر نسبت به

افراد با ed کمتر باشند یا افراد با salary بیشتر.

(i) -9

$$f_0(z) = \beta_0 + \beta_1 z$$

$$f_1(z) = (\beta_0 + \gamma_0) + (\beta_1 + \gamma_1) z$$

$$f_0(z) = f_1(z) \Rightarrow \gamma_0 + \gamma_1 z^* = 0 \Rightarrow z^* = -\frac{\gamma_0}{\gamma_1} \quad (ii)$$

$$\gamma_0 \gamma_1 < 0 \Rightarrow \frac{\gamma_0}{\gamma_1} < 0 \Rightarrow z^* > 0 \quad \checkmark$$

$$Total^* = -\frac{-0.357}{0.03} = 11.9 \quad (iii)$$

(iv) خیر واقعاً امکان ۱۱.۹ سال اضافه تحصیل نسبت به مردان وجود ندارد

(i) ۱۵

اگر انتخاب افراد برای طرح اتفاقی بوده باشد، جواب منفی است.

(ii)

اگر پارامترهای بیشتری را در مسئله لایتم، اولین که ضریب مورد نیاز مسئله

با دست بیشتری تعیین می شود و در من که خطای ناشی از انتخاب رندم

افراد کمتر می شود.

JAHAN NAMA

1 (الف) برای درست آوردن ضرایب با دقت بیشتر و با خطای کمتر

2
3 اگر تعداد داده‌های دانش آموزان n زیاد باشد طوری که خطای اتفاقی

4
5 پارامترهای کمتر و اطمینان زیاد شود، مشکلی پیش نمی‌آید.