4.

Table 2.10 Regression Output When Y is Regressed on X for Labor Force Participation Rate of Women

Variable	Coefficient	s.e.	t-Test	p-value
Constant	0.203311	0.0976	2.08	0.0526
X	0.656040	0.1961	3.35	< 0.0038
$n = 19$	$R^2 = 0.397$	$R_a^2 = 0.362$	$\hat{\sigma} = 0.0566$	df = 17

Let Y and X denote the labor force participation rate of women in 1972 and 1968, respectively, in each of 19 cities in the United States. The regression output for this data set is shown in Table 2.10. It was also found that SSR = 0.0358 and SSE = 0.0544. Suppose that the model $Y = \beta_0 + \beta_1 X + \varepsilon$ satisfies the usual regression assumptions.

(a) Compute $\mathrm{Var}(Y)$ and $\mathrm{Cor}(Y, X)$.

(b) Suppose that the participation rate of women in 1968 in a given city is 45%. What is the estimated participation rate of women in 1972 for the same city?

(a) $\mathrm{Var}(Y) = \sqrt{\dfrac{SSE}{N-2}}^{\,2} = \sqrt{\dfrac{0.0544}{19-2}}^{\,2} = 0.05657^2 = 0.0032$

$\mathrm{Cov}(X, Y) = \sqrt{\dfrac{SSR}{SST}} = \sqrt{\dfrac{0.0358}{0.0358+0.0544}} = 0.629997$

$\mathrm{Cor}^2 = R^2$

$\sqrt{R^2} = \sqrt{\dfrac{SSR}{SST}}$

(b) $\hat{Y} = 0.656040X + 0.203311$

x가 45% 주어졌을때 계산

$0.656040 \cdot 45/100 + 0.203311 = 49.85\%$

$\mu_x = 0.5 \quad \sigma_x^2 = 0.005 = S_{xx}$

$s.e.(\hat{y}_0) = s\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}}$

(c) $\hat{y} \pm t_{0.025,17} \times \hat{\sigma}\sqrt{\dfrac{1}{n} + \dfrac{(x-\bar{x})^2}{\mathrm{Var}(x)}}$

$= 0.4985 \pm 2.11 \times 0.0566\sqrt{\dfrac{1}{19} + \dfrac{(0.45-0.5)^2}{0.005}}$

$= 0.4985 \pm 2.11 \times 0.0421 \quad \therefore (0.410, 0.587)$

4.

Table 2.10 Regression Output When Y is Regressed on X for Labor Force Participation Rate of Women

Variable	Coefficient	s.e.	t-Test	p-value
Constant	0.203311	0.0976	2.08	0.0526
X	0.656040	0.1961	3.35	< 0.0038
$n = 19$	$R^2 = 0.397$	$R_a^2 = 0.362$	$\hat{\sigma} = 0.0566$	df = 17

Let Y and X denote the labor force participation rate of women in 1972 and 1968, respectively, in each of 19 cities in the United States. The regression output for this data set is shown in Table 2.10. It was also found that SSR = 0.0358 and SSE = 0.0544. Suppose that the model $Y = \beta_0 + \beta_1 X + \varepsilon$ satisfies the usual regression assumptions.

(d) Construct the 95% confidence interval for the slope of the true regression line, β_1.

(e) Test the hypothesis: $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 > 1$ at the 5% significance level.

(f) If Y and X were reversed in the above regression, what would you expect R^2 to be?

(d) $\left(\hat{\beta}_1 - t_{\frac{\alpha}{2}}(n-2)\cdot s.e.(\hat{\beta}_1),\ \hat{\beta}_1 + t_{\frac{\alpha}{2}}(n-2)\cdot s.e.(\hat{\beta}_1)\right)$

$(0.656040 \pm (2.11 \cdot 0.1961))$

(e) $t = \dfrac{\hat{\beta}_1 - 1}{s.e.(\hat{\beta}_1)} = \dfrac{0.656040 - 1}{0.1961} = -1.75$ p-value : $P(t_{17} < -1.75) = 0.0487$ < 0.05

$\therefore H_0$ is not rejected

(f) Same!

단순선형회귀모델에서의 결정계수는 종속변수와 독립변수가 뒤바뀌어도 유지.

SST = SSR + SSE = 0.0902