[STATS 413] HW5

Huy Le

February 2026

1 Blood pressure prediction

The fitted model is

$$\widehat{\text{BloodPressure}} = 2.20531 + 1.20093 \cdot \text{Weight}.$$

(a) For an individual who weighs = 90 kg, the predicted blood pressure is

$$\widehat{\text{BloodPressure}} = 2.20531 + 1.20093(90) = 110.28901 \approx 110.29.$$

(b) We know that $R^2 = r^2$, where r is the sample correlation between Y and X. Since the slope is positive, $r > 0$, so we have

$$r = \sqrt{0.9026} \approx 0.9501.$$

(c) $\text{Weight}_{lb} = 2.204\,\text{Weight}_{kg}$. From definition, rescaling X by a constant does not change R^2.. The estimated regression line does change in terms of the slope coefficient, but the predictor's coefficient:

$$\widehat{\text{BloodPressure}} = 2.20531 + \frac{1.20093}{2.20462}\,\text{Weight}_{lb} \approx 2.20531 + 0.5448\,\text{Weight}_{lb}.$$

(d) The slope estimate 1.20093 means that for each additional 1 kg of weight, the predicted blood pressure increases by about 1.20093 units (on average).

(e) A 90% CI for the true slope β_1 is

$$\begin{aligned}\hat{\beta}_1 \pm t_{0.95,18}\,\text{SE}(\hat{\beta}_1) &= 1.20093 \pm t_{0.95,18}(0.09297)\\ &= 1.20093 \pm 1.734(0.09297)\\ &= 1.20093 \pm 0.1612.\end{aligned}$$

so the 90% CI is approximately $(1.0397,\ 1.3621)$.

(f) For the test $H_0 : \beta_0 = 0$ vs. $H_A : \beta_0 \neq 0$, the output gives $p = 0.802$. So at 0.05 α levels, we fail to reject H_0, so there is no evidence that the intercept differs from 0.

(g) This is not necessarily. A non-significant intercept does not, by itself, indicate model assumption violations. To check the stronger regression assumptions, we can examine the residuals and look for patterns that would suggest violations. We can also compare nested models with ANOVA F-tests to assess whether a more flexible mean structure is needed.

The new multiple regression model is

$$\widehat{\text{BloodPressure}} = -16.57937 + 1.03296 \cdot \text{Weight} + 0.70825 \cdot \text{Age}.$$

(h) We test

$$H_0 : \beta_{\text{Weight}} = \beta_{\text{Age}} = 0 \quad \text{vs.} \quad H_A : \text{at least one is nonzero}$$

From the output, the overall F-statistic is $F = 978.2$ on $(2, 17)$ df with p-value $< 2.2 \times 10^{-16}$. Since $p < 0.05$, we reject H_0 at $\alpha = 0.05$. Hence, the covariates collectively improve predictive performance.

(i) To test whether Age adds predictive power after controlling for Weight, we test

$$H_0 : \beta_{\text{Age}} = 0 \quad \text{vs.} \quad H_A : \beta_{\text{Age}} \neq 0$$

From the coefficient table, Age has $t = 13.235$ with p-value 2.22×10^{-10}. Since $p < 0.001$, we reject H_0 at $\alpha = 0.001$. We conclude that age provides additional predictive power beyond Weight.

(j) Test $H_0 : \beta_{\text{Weight}} = 1.2$ vs. $H_A : \beta_{\text{Weight}} \neq 1.2$ at $\alpha = 0.1$. We have $\hat{\beta}_{\text{Weight}} = 1.03296$ and $\text{SE}(\hat{\beta}_{\text{Weight}}) = 0.03116$, so

$$t = \frac{1.03296 - 1.2}{0.03116} \approx -5.36,$$

with 17 df. Since $|t|$ is far larger than $t_{0.95,17} \approx 1.740$, we reject H_0 at $\alpha = 0.1$.

(k) In (e), Weight has a slope of 1.2 in the simple regression of BloodPressure. But in (j), we are testing the multiple-regression coefficient on Weight, which is the effect of Weight holding Age fixed. If Age is correlated with Weight and is itself predictive of blood pressure, the simple-regression slope can differ from the multiple-regression slope. Specifically, controlling for Age reduces the estimated Weight slope from about 1.20 to about 1.03. So therefore, there are no contradiction.

2 Unbiased estimates of σ_ε^2

(a) We have

$$y - \hat{y} = y - Hy = (I - H)y$$

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = (y-\hat{y})^\top(y-\hat{y}) = \big((I-H)y\big)^\top\big((I-H)y\big) = y^\top(I-H)^\top(I-H)y$$

We now justify that H is symmetric and idempotent. Indeed, we have

$$\begin{aligned} H^\top &= \big(X(X^\top X)^{-1}X^\top\big)^\top \\ &= X\big((X^\top X)^{-1}\big)^\top X^\top. \end{aligned}$$

Since $X^\top X$ is symmetric, its inverse is also symmetric, so $\big((X^\top X)^{-1}\big)^\top = (X^\top X)^{-1}$. Hence $H^\top = X(X^\top X)^{-1}X^\top = H$.

$$\begin{aligned} H^2 &= X(X^\top X)^{-1}X^\top\, X(X^\top X)^{-1}X^\top \\ &= X(X^\top X)^{-1}(X^\top X)(X^\top X)^{-1}X^\top \\ &= X(X^\top X)^{-1}X^\top \\ &= H \end{aligned}$$

Therefore $(I-H)^\top(I-H) = (I-H)(I-H) = I-H$, so

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = y^\top(I-H)y$$

(b) We have $E(y) = X\beta$. Then

$$\{E(y)\}^\top(I-H)E(y) = (X\beta)^\top(I-H)(X\beta)$$

But $HX = X$ (since H projects onto the column space of X), so $(I-H)X = 0$. Hence $(I-H)(X\beta) = 0$ and the expression equals 0.

(c) Since $\mathrm{Var}(y) = \mathrm{Var}(\varepsilon) = \sigma_\varepsilon^2 I$,

$$\begin{aligned} \mathrm{tr}\{(I-H)\mathrm{Var}(y)\} &= \mathrm{tr}\{(I-H)\sigma_\varepsilon^2 I\} \\ &= \sigma_\varepsilon^2\,\mathrm{tr}(I-H) \\ &= \sigma_\varepsilon^2\big(\mathrm{tr}(I)-\mathrm{tr}(H)\big). \end{aligned}$$

$\mathrm{tr}(I) = n$, and note that H is idempotent, so $\mathrm{tr}(H) = \mathrm{rank}(H) = \mathrm{rank}(X) = p+1$ since X has $p+1$ columns and full column rank. Therefore,

$$\mathrm{tr}\{(I-H)\mathrm{Var}(y)\} = \sigma_\varepsilon^2(n-p-1).$$

(d) With $E(y) = X\beta$ and $\mathrm{Var}(y) = \sigma_\varepsilon^2 I$,

$$\begin{aligned} E\big(y^\top(I-H)y\big) &= \{E(y)\}^\top(I-H)E(y) + \mathrm{tr}\big((I-H)\mathrm{Var}(y)\big) \\ &= 0 + \sigma_\varepsilon^2(n-p-1) \end{aligned}$$

by parts (b) and (c). Thus,

$$E\big(\hat{\sigma}_\varepsilon^2\big) = E\left(\frac{1}{n-p-1}\sum_{i=1}^n (y_i - \hat{y}_i)^2\right) = \frac{1}{n-p-1} E\big(y^\top (I-H)y\big) = \sigma_\varepsilon^2$$

So $\hat{\sigma}_\varepsilon^2$ is unbiased for σ_ε^2.

(e) This is not true. Even if $\hat{\sigma}_\varepsilon^2$ is unbiased for σ_ε^2, it does not follow that $\hat{\sigma}_\varepsilon = \sqrt{\hat{\sigma}_\varepsilon^2}$ is unbiased for σ_ε. In general, expection is not preserved under square root transformation.

(f) Let $Z = (Z_1, \ldots, Z_n)^\top$. Then $E(Z) = \theta\mathbf{1}$ and $\mathrm{Var}(Z) = \sigma^2 I$. Let $\bar{Z} = \frac{1}{n}\mathbf{1}^\top Z$ and note that $\hat{Z} = \bar{Z}\,\mathbf{1} = HZ$ where

$$X = \mathbf{1} \in R^{n\times 1}, \qquad H = X(X^\top X)^{-1}X^\top = \frac{1}{n}\mathbf{1}\mathbf{1}^\top$$

Then

$$\sum_{i=1}^n (Z_i - \bar{Z})^2 = (Z-\hat{Z})^\top (Z-\hat{Z}) = Z^\top (I-H)Z.$$

Applying part (d) with $p = 0$, we get

$$E\left(\frac{1}{n-1}\sum_{i=1}^n (Z_i - \bar{Z})^2\right) = \sigma^2,$$

so the sample variance $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \bar{Z})^2$ is unbiased for σ^2.