Next week: No recitation, maybe extra OHs

Today: Non-empirical practice problems

$$\left.\begin{array}{l} 14.1 \\ 14.2 \\ 14.3 \text{ (Extra)} \\ 14.5 \\ 14.8 \text{ (Extra)} \end{array}\right\} \text{"Big data"}$$

Empirical practice problems (see R notes)

16.1 Dynamic Causal Effects

14.1 A researcher is interested in predicting average test scores for elementary schools in Arizona. She collects data on three variables from 200 randomly chosen Arizona elementary schools: average test scores (*TestScore*) on a standardized test, the fraction of students who qualify for reduced-priced meals (*RPM*), and the average years of teaching experience for the school's teachers (*TExp*). The table below shows the sample means and standard deviations from her sample.

Variable	Sample Mean	Sample Standard Deviation
TestScore	750.1	65.9
RPM	0.60	0.28
TExp	13.2	3.8

After standardizing *RPM* and *TEXP* and subtracting the sample mean from *TestScore*, she estimates the following regression:

$$\widehat{TestScore} = -48.7 \times RPM + 8.7 \times TExp,\ SER = 44.0$$

a. You are interested in using the estimated regression to predict average test scores for an out-of-sample school with $RPM = 0.52$ and $TEXP = 11.1$.

i. Compute the transformed (standardized) values of *RPM* and *TEXP* for this school; that is, compute the X^{oos} values from the X^{*oos} values, as discussed preceding Equation (14.2).

ii. Compute the predicted value of average test scores for this school.

(Y, X_1, X_2) Raw : X_i

$\bar Y, \bar X_1, \bar X_2$ Averages : $\bar X = \frac{\sum_i X_i}{n}$

$\tilde Y, \tilde X_1, \tilde X_2$ Standardized : $\tilde X_i = \frac{X_i - \bar X}{\sigma_X}$, $\tilde Y_i = Y_i - \bar Y$

$Y_i = \beta_0 + \beta_1 X_1 + u_i$: for a one unit increase in X, we expect Y to increase by $\hat\beta_1$

$\tilde Y_i = \beta_0 + \beta_1 \tilde X_1 + u_i$: for a one standard deviation increase in X, we expect Y to increase by $\hat\beta_1$

a. $X_1^{oos} = 0.52$, $X_2^{oos} = 11.1$

i. $\tilde X_1^{oos} = \frac{X_1^{oos} - \bar X_1}{\sigma_{X_1}} = \frac{0.52 - 0.60}{0.28} = -0.289$

$\tilde X_2^{oos} = \frac{X_2^{oos} - \bar X_2}{\sigma_{X_2}} = \frac{11.1 - 13.2}{3.8} = -0.553$

ii. $\hat{\tilde Y}^{oos}(\tilde X_1^{oos}, \tilde X_2^{oos})$

Raw prediction

$\hat{\tilde Y}^{oos} = -48.7\, \tilde X_1^{oos} + 8.7\, \tilde X_2^{oos}$

$= -48.7 \cdot (-0.289) + 8.7(-0.553)$

$= 9.2632$

$\hat{\tilde Y}^{oos} = \hat Y^{oos} - \bar Y \iff \hat Y^{oos} = \hat{\tilde Y}^{oos} + \bar Y$

$= 9.2632 + 750.1$

$= 759.3632$

b. The actual average test score for the school is 775.3. Compute the error for your prediction.

b. $Y^{oos} = 775.3$

$e^{oos} = \hat Y^{oos}(X^{oos}) - Y^{oos}$

$= 775.3 - 759.3632$

$= 15.9368$

c. The regression shown above was estimated using the standardized regressors and the demeaned value of *TestScore*. Suppose the regression had been estimated using the raw data for *TestScore*, *RMP*, and *TExp*. Calculate the values of the regression intercept and slope coefficients for this regression.

c. $\tilde Y^{oos} = \tilde\beta_1 \tilde X_1 + \tilde\beta_2 \tilde X_2 \Rightarrow Y^{oos} = \hat\beta_0 + \hat\beta_1 X_1 + \hat\beta_2 X_2$

$\tilde Y_i = Y_i - \bar Y$, $\tilde X_i = \frac{X_i - \bar X}{\sigma_X}$

$\tilde Y = \tilde\beta_1 \tilde X_1 + \tilde\beta_2 \tilde X_2$

$Y_i - \bar Y = \tilde\beta_1 \left(\frac{X_1 - \bar X_1}{\sigma_{X_1}}\right) + \tilde\beta_2 \left(\frac{X_2 - \bar X_2}{\sigma_{X_2}}\right)$

$\Rightarrow Y_i = \underbrace{\left[\bar Y + \tilde\beta_1 \left(\frac{-\bar X_1}{\sigma_{X_1}}\right) + \tilde\beta_2 \left(\frac{-\bar X_2}{\sigma_{X_2}}\right)\right]}_{\hat\beta_0} + \underbrace{\left(\frac{\tilde\beta_1}{\sigma_{X_1}}\right)}_{\hat\beta_1} X_1 + \underbrace{\left(\frac{\tilde\beta_2}{\sigma_{X_2}}\right)}_{\hat\beta_2} X_2$

$\hat\beta_0 = 750.1 + (-48.7)\left(\frac{-0.60}{0.28}\right) + 8.7\left(\frac{-13.2}{3.8}\right)$

$= 824.23609$

$\hat\beta_1 = \frac{-48.7}{0.28}$

$= -173.92857$

$\hat\beta_2 = \frac{8.7}{3.8}$

$= 2.2894737$

$\hat Y_i = 824.23609 - 173.92857\, X_1 + 2.2894737\, X_2$

d. Use the regression coefficients that you computed in (c) to predict average test scores for an out-of-sample school with $RPM = 0.52$ and $TExp = 11.1$. Verify that the prediction is identical to the prediction you computed in (a.ii).

d. $\hat Y_i(X_1^{oos} = 0.52,\ X_2^{oos} = 11.1) = 759.20639$

14.3 Describe the relationship, if any, between the standard error of a regression and the square root of the MSPE of the regression's out-of-sample predictions.

SE of an (in-sample) regression

$\sqrt{MSPE}$: st. dev. of forecast error

In-Sample Regression : $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_1 + u_i^{IS}$ SE = standard deviation of in-sample error

Data-generating process : $Y_i^{oos} = \beta_0 + \beta_1 X_i^{oos} + u_i^{oos}$

$$\text{Forecast error} := Y^{oos} - \hat{Y}(X^{oos})$$
$$= \beta_0 + \beta_1 X_i^{oos} + u_i^{oos} - \hat{\beta}_0 - \hat{\beta}_1 X_i^{oos}$$
$$= (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1) X_i^{oos} + u_i^{oos}$$

$$MSPE = Var\left[(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1) X_i^{oos} + u_i^{oos}\right]$$
$$= \underbrace{Var\left((\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1) X_i^{oos}\right)}_{\text{a thing}} + \underbrace{Var\left(u_i^{oos}\right)}$$

$$Var(u_i^{is})$$
$$E[(u_i^{is} - 0)^2]$$
$$= SE_i^2$$

14.5 Y is a random variable with mean $\mu = 2$ and variance $\sigma^2 = 25$.

a. Suppose you know the value of μ.

i. What is the best (lowest MSPE) prediction of the value of Y? That is, what is the oracle prediction of Y?

ii. What is the MSPE of this prediction?

b. Suppose you don't know the value of μ but you have access to a random sample of size $n = 10$ from the same population. Let $\overline{Y}$ denote the sample mean from this random sample. You predict the value of Y using $\overline{Y}$.

i. Show that the prediction error can be decomposed as $Y - \overline{Y} = (Y - \mu) - (\overline{Y} - \mu)$, where $(Y - \mu)$ is the prediction error of the oracle predictor and $(\mu - \overline{Y})$ is the error associated with using $\overline{Y}$ as an estimate of μ.

ii. Show that $(Y - \mu)$ has a mean of 0, that $(\overline{Y} - \mu)$ has a mean of 0, and that $Y - \overline{Y}$ has a mean of 0.

iii. Show that $(Y - \mu)$ and $(\overline{Y} - \mu)$ are uncorrelated.

iv. Show that the MSPE of $\overline{Y}$ is MSPE $= E(Y - \mu)^2 + E(\overline{Y} - \mu)^2 = \text{var}(Y) + \text{var}(\overline{Y})$.

v. Show that MSPE $= 25(1 + 1/10) = 27.5$.

$Y \sim (\mu = 2, \sigma^2 = 25)$

a. μ is known

i) $\mathbb{E}[Y^{oos} \mid X = X^{oos}] = \mathbb{E}[\mathbb{E}[Y^{oos} \mid X = X^{oos}]]$

$= \mathbb{E}[Y^{oos}]$ by Law of Iterated Expectations

$= \mu = 2$

ii) $\text{MSPE} = \sigma_u^2 + \underbrace{\mathbb{E}[(\hat{\beta}_1 - \beta_1) X_1^{oos} + (\hat{\beta}_2 - \beta_2) X_2^{oos} + \dots]^2}_{0}$

$= 25$

b. Sample $n = 10$ $\quad \bar{Y}_n = \frac{1}{10} \sum Y_i$

i) $Y - \bar{Y} = Y - \mu + \mu - \bar{Y}$

$= \underbrace{(Y - \mu)}_{\text{Oracle prediction error}} - \underbrace{(\bar{Y} - \mu)}_{\text{error associated w/ using } \bar{Y}}$

ii) $\mathbb{E}[Y - \mu] = \mathbb{E}[Y] - \mathbb{E}[\mu]$

$= \mu - \mu = 0$

$\mathbb{E}[\bar{Y} - \mu] = \mathbb{E}[\bar{Y}] - \mathbb{E}[\mu]$

$= \mathbb{E}[\frac{1}{10} \sum_{i=1}^{10} Y_i] - \mu$

$= \frac{1}{10} \cdot 10\, \mathbb{E}[Y_i] - \mu$

$= \mu - \mu = 0$

$\mathbb{E}[Y - \bar{Y}] = \mathbb{E}[(Y - \mu) - (\bar{Y} - \mu)]$

$= 0 - 0 = 0$

iii) $\bar{Y}$ comes on iid sample
Y comes from OOS $\Big\}$ iid $\Rightarrow (\bar{Y}, Y)$ uncorrelated

iv) $\text{MSPE} = \mathbb{E}[(Y - \bar{Y})^2]$

$= \mathbb{E}[\{(Y - \mu) - (\bar{Y} - \mu)\}^2]$

$= \underbrace{\mathbb{E}[(Y - \mu)^2]}_{Var(Y)} + \underbrace{2\mathbb{E}[(Y - \mu)(\bar{Y} - \mu)]}_{0} + \underbrace{\mathbb{E}[(\bar{Y} - \mu)^2]}_{Var(\bar{Y})}$

v) $\text{MSPE} = 25 + Var(\bar{Y})$

$= Var(\frac{1}{10}(Y_1 + \dots + Y_{10}))$

$= \frac{1}{100} Var(Y_1 + \dots + Y_{10})$

$= \frac{1}{100} \cdot 10\, Var(Y)$

$= \frac{1}{10} Var(Y)$

$= \frac{1}{10} \cdot 25$

$= 25 + \frac{25}{10}$

$= 25 + 2.5$

$= 27.5$

14.8 Let X and Y be two random variables. Denote the mean of Y given $X = x$ by $\mu(x)$ and the variance of Y by $\sigma^2(x)$.

a. Show that the best (minimum MSPE) prediction of Y given $X = x$ is $\mu(x)$ and the resulting MSPE is $\sigma^2(x)$. (*Hint:* Review Appendix 2.2.)

$\mu(x) := \mathbb{E}[Y \mid X = x]$

$\sigma^2(x) := Var(Y \mid X = x)$

a. $MSPE = \mathbb{E}[(Y - g(x))^2]$ $g(x)$: prediction function of $X = x$

Assume $Y_i = \begin{cases} Y_1 & \text{w/ probability } P_1 \\ Y_2 & P_2 \\ \vdots & \vdots \\ Y_K & \text{w/ probability } P_K \end{cases}$ $P_i(x) = \mathbb{P}(Y = Y_i \mid X = x)$, $\sum_{i=1}^{K} P_i(x) = 1$

$$= \sum_{i=1}^{K} (Y_i - g(x))^2 P_i(x)$$

Take derivative wrt $g(x)$

$$0 = \frac{d}{dg(x)} \sum_{i=1}^{K} (Y_i - g(x))^2 P_i(x)$$

$$= \cancel{2(-1)} \sum_{i=1}^{K} (Y_i - g(x)) P_i(x)$$

$$= \sum_{i=1}^{K} Y_i P_i(x) - \sum_{i=1}^{K} g(x) P_i(x)$$

$$\Rightarrow \sum_{i=1}^{K} g(x) P_i(x) = \sum_{i=1}^{K} Y_i P_i(x)$$

$$\hat{g}(x) = \mathbb{E}[Y \mid X = x] \equiv \mu(x)$$

b. Suppose X is chosen at random. Use the result in (a) to show that the best prediction of Y is $\mu(X)$ and the resulting MSPE is $E[Y - \mu(X)]^2 = E[\sigma^2(X)]$.

b. X chosen at random

$$\mathbb{E}[(Y - \mu(X))^2] = \mathbb{E}[\mathbb{E}[(Y - \underset{\mathbb{E}[Y \mid X]}{\mu(X)})^2 \mid X]]$$

by Law of Iterated Expectations

$$= \mathbb{E}[\sigma^2(X)]$$