Economics 102_A01-A04: Analysis of Economic Data
Cameron Fall 2021
Department of Economics, U.C.-Davis
Answers to Final Exam (Versions A and B)

Version A

1.(a) `kdensity hp` or `hist hp, kdensity`.

(b) `detail`

(c) Expect 95% within two standard deviations of mean: $\bar{x} \pm 2s = 161.68 \pm 2 \times 57.69 = (46.30, 277.06)$.

(d) $H_0 : \mu = 150$ versus $H_a : \mu \neq 150$ where claim is the alternative hypothesis.
$t = (\bar{x} - \mu_0)/(s/\sqrt{n}) = (161.68 - 150)/(57.69/\sqrt{95}) = 11.68/5.92 = 1.973$.
Reject H_0 at level .05 if $|t| > t_{.025;94} = 1.986$. So do not reject H_0 as $1.973 < 1.986$.
Conclude that (at sig. level 0.05) population mean horsepower differs from 150.

(e) `display 2*ttail(94,1.973)`

(f) 0.05.

(g) Regressing `hp` on `torque` will best explain `hp` since from the correlation output it is the variable that is the most highly correlated with `hp`.

(h) R^2 for `hp` on `curbwt` equals squared correlation coefficient $= (0.8354)^2 = 0.698$.

2.(a) The slope coefficient is -.0709. A standard deviation of `hp` is 57.69.
So `mpg` changes by $-.0709 \times 57.69 = -4.09$. A decrease of 4.09 miles per gallon.

(b) This is directly given in regression output.
A 95% confidence interval for $\beta_{\texttt{hp}}$ is $(-.083, -.059)$.

(c) A 99% confidence interval for $\beta_{\texttt{hp}}$ is $b_2 \pm t_{.005;93} \times se(b_2) = -0.07086 \pm 2.630 \times 0.00605 = -0.07086 \pm 0.0159 = (-.0867, -.0549)$.

(d) $H_0 : \beta_2 = -.05$ against $H_a : \beta_2 \neq -.05$.
$t = (b_2 - -.05)/se(b_{2b}) = (-0.07086 + .05)/0.00605 = -0.02086/0.00605 = -3.45$.
$|t| = 3.45 > t_{.025;93} = 1.986$.
Reject H_0 at level 0.05. Conclude that the slope coefficient differs from $-.05$.

(e) Prediction is $\widehat{y} = b_1 + b_2 x^* = 40.26 - .07086 \times 161.68 = 28.80$.

3.(a) A one horsepower increase is associated with a 0.00717 increase in miles per gallon (holding the other variables constant).

(b) Cars manufactured by European countries have 1.65 lower miles per gallon than those manufactured by U.S. companies (the omitted category) (holding the other variables constant).

(c) No. Then we have the dummy variable trap since dUS, dEUR and dASIA are a complete set of mutually exclusive indicator variables.

(d) `curbwt` and `dEUR` are statistically significant at 5%.

(e) $H_0 : \beta_2 = 0, ..., \beta_7 = 0$ against H_a : At least one of $\beta_2, ..., \beta_7 \neq 0$.
From the regression output $F = 67.99$ has p value of $0.0000 < .05$.
Reject H_0. Conclude that the regressors are jointly significant at level 0.05.

(f) `test curbwt torque accel dEUR dASIA`

(g) The computer output does not give $\bar{R}^2$.
So use $s_e = \sqrt{\frac{1}{n-k} \sum_{i=1}^{n} (y_i - \widehat{y}_i)}$ which also has a penalty for model size.
Favors larger model as RMSE $2.1329 < 2.998$.

4.(a) Miles per gallon is a proportion .0842 lower, or 8.42% lower for each one second increase in acceleration.

(b) This is now an elasticity as now log-log. A one percent increase in curb weight is associated with a 0.040 percent decrease in miles per gallon.

(c) The error terms may be likely to be correlated for cars from the same manufacturer.
So obtain cluster-robust standard errors with clustering on manufacturer.

(d) Yes, because the dependent variable is a log. We cannot simply use the model to get $\widehat{\ln y}$ and then let $\widehat{y} = \exp(\widehat{\ln y})$ due to retransformation bias.

(e) Yes. OLS is biased and inconsistent if the error is correlated with one of the regressors.

(f) No. The reason we included it as a regressor is because it is correlated with `lmpg`.

(g) `generate ltorque = ln(torque)`

5.(a) $\mathrm{E}[X] = 0.7 \times 10 + 0.3 \times 20 = 7 + 6 = 13$.
$\mathrm{Var}[X] = \mathrm{E}[(X-\mu)^2] = 0.7 \times (10-13)^2 + 0.3 \times (20-13)^2 = 0.7 \times (-3)^2 + 0.3 \times 7^2 = 6.3 + 14.7 = 21$.

(b) Annual growth is $\Delta \ln y / \Delta x = (26.9 - 25.4)/(2012 - 2002) = 1.5/10 = 0.15$.
Annual growth is 15% per year.

(c) It doubles in $72/4 = 18$ years.

(d) For women $\widehat{\texttt{wage}} = b_1 + b_2 \times 0 = b_1$ so $b_1 = 21$ (the average for women).
For men $\widehat{\texttt{wage}} = b_1 + b_2 \times 1 = b_1 + b_2$ so $b_1 + b_2 = 24$ (the average for men).
It follows that $b_1 = 21$ and $b_2 = 24 - 21 = 3$.

(e)(i) $E[y|x] = 10 + 3 \times 3 = 19$
(ii) $u = y - E[y|x] = 30 - 19 = 11$
(iii) $\widehat{y} = 8 + 5 \times 3 = 23$.
(iv) $e = y - \widehat{y} = 30 - 23 = 7$.

Multiple Choice Version A:
1. c **2. d** **3. a** **4. c** **5. c** **6. d** **7. c** **8. b** **9. a** **10. d**

Scores out of	60	
75*th* percentile	43	(72%)
Median	34.5	(58%)
25*th* percentile	27	(45%)

A	43 and above
A-	39 and above
B+	36 and above
B	33 and above
B-	30 and above
C+	28 and above
C	25 and above
C-	23 and above
D+	22 and above
D	21 and above
D-	20 and above