Solutions Autumn 2018

November 14, 2018

All questions are worth 10 points. Bullet points below each question give general grading instructions.

Exercise 1

1a) The test:

$$H_0 : \frac{Var(Offer2006)}{Var(Offer2008)} = 1$$

$$H_1 : \frac{Var(Offer2006)}{Var(Offer2008)} > 1$$

The test statistic is:

$$F = \frac{0.85^2}{0.12^2} = 50.17$$

which is F-distributed with 17,17 degrees of freedom. The cutoff value is $F_{17,17,0.01} = 3.24$. We reject the null of equality of variances.

Note: The larger variance goes in the numerator of the F-statistic. If the larger variance is in the denominator, then the student has to transform the cutoff value by taking the inverse and picking the cut-off value itself with flipped df - $\frac{1}{F_{17,17}} \approx \frac{1}{2.19} = 0.31$

The test indicates that in 2008 the variance in bank interest rate offers was lower than in 2006.

When we have a cartel, involved banks will behave in a similar way. When submitting interest rate offers they would submit similar ones, resulting in a low variance. This would move the mean towards the cartel interest rate offers, reducing the distance between the mean and the cartel offer. The variance is the squared deviations from the mean for every observation, so these deviations will be smaller for the cartel members, resulting overall in a smaller variance.

- Correct test, but with calculation mistake -5
- Puts the smaller variance on top, without correcting the critical value -5
- Correct test, but wrong conclusion -7
- Correct test, but no conclusion -5
- Long argument with no clear intuition -8

1b) The students should pick the positive correlation coefficients from this table. Given that colluding banks submit *identical* interest rates, then there would be a perfect correlation of 1 btw their offers. An additional requirement is that all banks are in it together, so that if H & K have a high correlation btw one another, and K has a high correlation with A, then H & A should also have a high correlation.

If they pick 1, then the offending banks are: D, E & K

Students can demonstrate independence and pick a lower threshold of 0.9, then the offending banks are: D, E, K & B

1c) This question should be solved with T-test for unequal variances:

$$H_0 : Mean(Libor2006) = Mean(Libor2008)$$

$$H_1 : Mean(Libor2006) < Mean(Libor2008)$$

The test statistic is:

$$T = \frac{3.09 - 2.00}{\sqrt{\frac{0.56^2+0.06^2}{10}}} \approx 6.12$$

$$df = \frac{(\frac{0.56^2+0.06^2}{10})^2}{\frac{(0.56^2/10)^2+(0.06^2/10)^2}{9}} = \frac{0.001006158}{0.0001092866} \approx 9$$

The cutoff value is $t_{9,0.01} = 2.82$, which means we reject the null hypothesis of no difference in means in favor of the alternative

The test indicates that the Libor in 2006 is lower than the Libor in 2008

- Correct test, but does not calculate the correct df -3
- Correct test, but with calculation mistake -5
- Correct test, but wrong conclusion -7
- Correct test, but no conclusion -5

1d) Observing the 10 out of 10 positive differences, it is obvious that the two populations have different locations.

H_0 : The two population locations are the same

H_1 : The two populations have different locations

The Sign Test gets the following test statistic:

$$z = \frac{10 - .5 * 10}{.5 * \sqrt{10}} = \frac{5}{1.58} = 3.16 > 2.32 = z_{0.01}$$

We reject the null hypothesis. In other words, all values for 2008 are below the values for 2006.

Libor 2006	Libor 2008	Difference	Sign
2.29	1.91	0.38	+
2.52	1.93	0.58	+
2.60	1.96	0.64	+
2.71	1.98	0.74	+
2.80	2.01	0.79	+
3.34	2.02	1.32	+
3.58	2.03	1.55	+
3.67	2.04	1.63	+
3.67	2.07	1.60	+
3.77	2.09	1.68	+
		SUM:	10 + out of 10

- Correct test, but with calculation mistake -5
- Correct test, but wrong conclusion -7
- Correct test, but no conclusion -5

1e) Interpretation: We reject the null hypothesis of equality in means. The two groups of banks are significantly different at the 1 percent level.

df=1540, given in Table 4.

Confidence interval:

$$(20000 - 19000) \pm 2.576 \times \sqrt{(\frac{500^2}{630} + \frac{1000^2}{990})}$$

$$1000 \pm 96.62$$

The true difference in means belongs in the interval $[903.38, 1096.62]$

1f) Test: $H_0 : f_1 = e_1, f_2 = e_2, ...f_{10} = e_{10}$, where f is the empirical frequency and e is the expected frequency

H_1 : At least one f is not equal to its specified value

Second Digit	Expected frequency (Benfords Law)	Empirical frequency	e	f	f-e	Summation Component
0	0.12	0.038	12	3.8	-8.2	5.603
1	0.114	0.122	11	12	0.8	0.056
2	0.109	0.015	11	1.5	-9.4	8.106
3	0.104	0.374	10	37	27	70.096
4	0.1	0.008	10	0.8	-9.2	8.464
5	0.097	0.107	9.7	11	1	0.103
6	0.093	0.115	9.3	12	2.2	0.520
7	0.09	0.107	9	11	1.7	0.321
8	0.088	0.099	8.8	9.9	1.1	0.138
9	0.085	0.015	8.5	1.5	-7	5.765
Obs.	100	100			Sum:	99.173

$\chi^2 = 99.173 > \chi^2_{0.01,9} = 21.7$

We reject the null hypothesis at the 1 percent significance level. Therefore, the observed frequencies are significantly different than the expected frequencies at the 1 percent level.

- Correct test, but does not calculate the correct df -3
- Correct test, but with calculation mistake -5
- Correct test, but wrong conclusion -7
- Correct test, but no conclusion -5

1g) In Table 6 we observe a two-way ANOVA.

At the 10 percent significant with 6 percent p-value we can't reject the alternative hypothesis of risk exposure affecting bank cartel participation. We can reject the alternative hypothesis of the factor year and for the interaction between risk and year affecting participation. Given that the degrees of freedom on year are 2, and the formula for df is the treatments-1, then we observe 2+1=3 years

Exercise 2

2a) The regression equation looks like this:

$$Conflict = 0.445 - 0.0351 \times GDPCapita$$

The two point predictions:

$$0.445 - 0.0351 * 7.453 = .1833$$

$$0.445 - 0.0351 * 7.453 * \frac{101}{100} = .1807837$$

A 1 percent increase in GDP capita is associated with a .26 percent decrease in the probability of conflict

2b) 3 point predictions:

The probability of conflict at the mean for GDP capita: $0.434 - 0.629 * 7.453 + 0.597 * 7.453 = 0.195$

10% increase this year: $0.434 - 0.629 * 7.453 * 1.3 + 0.597 * 7.453 = -0.273$. The answer makes no sense, because we arrive at a point prediction below 0, which is not possible for a probability. This is a limitation of using OLS for modeling probabilities.

10% increase the previous year: $0.434 - 0.629 * 7.453 + 0.597 * 7.453 * 1.3 = 0.64$. The 10 percent increase in log of GDP capita is associated with a $0.64 - 0.195 = 44$ percent increase in the probability of conflict this year.

2c) 2 point predictions:

The mean: $0.174 - 3.792 \times 7.453 + 3.569 \times 7.453 = -1.488$. The corresponding probability is $p = \frac{e^{-1.488}}{1+e^{-1.488}} = \frac{0.226}{1.226} \approx 0.18$

The 10 percent increase: $0.174 - 3.792 \times 7.453 \times 1.1 + 3.569 \times 7.453 = -4.314$. The corresponding probability is $p = \frac{e^{-4.314}}{1+e^{-4.314}} = \frac{0.013}{1.013} \approx 0.01$.

This means that increasing GDP capita by 10 percent leads to 17 percent decline in the probability of conflict.

Logit predicts a probability within the bounds of (0,1), whereas OLS did not predict correctly the probability.

2d) There seems to be a severe problem of multicollinearity. Variables with a high correlation coefficient are likely to be multicollinear. Using a cutoff of 0.7 these seem to be:

1) Log of GDP capita has a high correlation coefficient with it's lag, with the shares of young, old and urban populations. Same goes for the lag of the variable.

2) Share of young is highly collinear with the share of old and urban populations

2e) Deriving the formula with short names for the variables:

$$\hat{\beta} = \frac{cov(Conflict, GDP)}{Var(GDP)} = \frac{cov(\beta GDP + \alpha ED + \epsilon, GDP)}{Var(GDP)} =$$

$$\frac{\beta cov(GDP, GDP) + \alpha cov(ED, GDP) + cov(\epsilon, GDP)}{Var(GDP)} = \beta + \alpha \frac{cov(ED, GDP)}{Var(GDP)}$$

2f) If $\hat{\beta} < \beta$, then

$$\beta + \alpha \frac{cov(ED, GDP)}{Var(GDP)} < \beta$$

$$\alpha \frac{cov(ED, GDP)}{Var(GDP)} < 0$$

divide both sides by positive α

$$\frac{cov(ED, GDP)}{Var(GDP)} < 0$$

Covariance and correlation have the same sign, so the correlation will be negative.

If $\hat{\beta} < \beta$, then the true coefficient is higher than what we have estimated in column (1).

2g) Ethnic diversity seems to be a variable that has mainly variation between countries, which means that it's effect will be absorbed by including country fixed effects. Fixed effects absorb country specific differences that are fixed over time such as ethnic diversity.

$\beta = -0.880$

$\hat{\beta} = -0.0351$

$Var(GDPXCapita) = SD(GDPxCapita)^2 = 1.572 * 1.571 = 2.469$

$$Cov(GDP, ED) = (\hat{\beta} - \beta) * Var(GDPxCapita)/\alpha = (-0.0357 + 0.880) * 2.469/\alpha = \frac{2.084}{\alpha}$$

Therefore, the sign of the covariance is positive.

The smart students can also note that the two two estimates are opposite tol the expectations from the previous points, meaning that the covariance has to be positive.

- No intuition on fixed effects - 6
- No intuition on the sign of the covariance - 6
- Overly long response with no clear idea, - 8

2h) By using different modeling approaches we find that increases in the GDP predict a decrease in the likelihood of conflict in a given year. We find evidence that increases in the GDP capita are also associated with increases in the probability of conflict next year. In column 4 we observe that fixed effects models provide the best fit of the data. We do not claim causality. (67 words)

- Not coherent - 10
- Implies causality -5
- Exceeds maximum -6