Final Exam, Fall 2001 Answer Key

Question A

1. A: $GPA = \beta_0 + \beta_1 HSGPA + \varepsilon$
   C: $GPA = \beta_0 + \varepsilon$

2. $\square^2$ □ .154 because no adjustment is needed for the enormous sample size of 78,000 in the California analysis. Interpolating between the sample sizes needed for .1 and .2, the necessary sample size is between 45 and 100, probably about 75.

3. A: $GPA = \beta_0 + \beta_1 HSGPA + \beta_2 ACH + \varepsilon$
   C: $GPA = \beta_0 + \varepsilon$

4. A: $GPA = \beta_0 + \beta_1 HSGPA + \beta_2 ACH + \beta_3 SAT + \varepsilon$
   C: $GPA = \beta_0 + \beta_1 HSGPA + \beta_2 ACH + \varepsilon$

5. A: $GPA = \beta_0 + \beta_1 HSGPA + \beta_2 ACH + \beta_3 SAT + \beta_4 QUAL + \varepsilon$
   C: $GPA = \beta_0 + \beta_1 HSGPA + \beta_2 ACH + \beta_4 QUAL + \varepsilon$

6. A: $GPA = \beta_0 + \beta_1 SAT + \beta_2 QUAL + \beta_3 (SAT \text{ X } QUAL) + \varepsilon$
   C: $GPA = \beta_0 + \beta_1 SAT + \beta_2 QUAL + \varepsilon$

7. ACHdev = ACH – mean(ACH) = ACH - 115
   HSGPAdev = HSGPA – mean(HSGPA) = HSGPA – 3.22
   A: $GPA = \beta_0 + \beta_1 HSGPAdev + \beta_2 ACHdev + \varepsilon$
   C: $GPA = 3.02 + \beta_1 HSGPAdev + \beta_2 ACHdev + \varepsilon$

8. ** This was a zinger question – we haven't covered this, yet **
   A: $GPA = \beta_0 + \beta_1 SAT + \beta_2 SAT^2 + \varepsilon$
   C: $GPA = \beta_0 + \beta_1 SAT + \varepsilon$

9. ** This was a zinger question **
   SATdev = SAT – mean(SAT) = SAT – 110
   A: $GPA = \beta_0 + \beta_1 SATdev + \beta_2 SATdev^2 + \varepsilon$
   C: $GPA = \beta_0 + \beta_2 SATdev^2 + \varepsilon$

10.

Var	White	Af-Amer	As-Amer	Hisp
X1	3	-1	-1	-1
X2	0	2	-1	-1
X3	0	0	1	-1

A: $GPA = \beta_0 + \beta_1 X1 + \beta_2 X2 + \beta_3 X3 + \varepsilon$
C: $GPA = \beta_0 + \varepsilon$

11. A: $GPA = \beta_0 + \beta_1 X1 + \beta_2 X2 + \beta_3 X3 + \beta_4 SAT + \beta_5 (SAT \times X1) + \varepsilon$
C: $GPA = \beta_0 + \beta_1 X1 + \beta_2 X2 + \beta_3 X3 + \beta_4 SAT + \varepsilon$

12. This is a rare instance when dummy codes are more appropriate than contrast codes. Using dummy codes for all the other groups means that the simple slope for SAT applies to the African American students.

Var	White	Af-Amer	As-Amer	Hisp
WH	1	0	0	0
ASA	0	0	1	0
HSP	0	0	0	1

A: $GPA = \beta_0 + \beta_1 WH + \beta_2 ASA + \beta_3 HSP + \beta_4 SAT + \beta_5 (SAT \times WH) + \beta_6 (SAT \times ASA) + \beta_7 (SAT \times HSP) + \varepsilon$
C: $GPA = \beta_0 + \beta_1 WH + \beta_2 ASA + \beta_3 HSP + \beta_5 (SAT \times WH) + \beta_6 (SAT \times ASA) + \beta_7 (SAT \times HSP) + \varepsilon$

Note that for the African-American students, WH, ASA, and HSP are all equal to 0, so Models A and C reduce to:

A: $GPA = \beta_0 + \beta_1 SAT + \varepsilon$
C: $GPA = \beta_0 + \varepsilon$

So the test of $\beta_4 = 0$ is the full models provides the appropriate test of whether SAT scores predict GPA scores for African-American students.

Question B

1.

Var	1 Pint	2 Pints	3 Pints
X1	-1	0	1
X2	1	-2	1

2. $bx_1 = (-11.5 + 6.7)/2 = -2.4$
$bx_2 = [11.5 - 2(9.3) + 6.7]/6 = -0.067$

3. $SSR_{x1} = (-11.5 + 6.7)^2/(2/8) = 92.16$
$SSR_{x2} = [11.5 - 2(9.3) + 6.7]^2/(6/8) = 0.21$

4. They conclude that there are no significant differences between the groups. Their conclusion is problematic because the omnibus test does not specifically test their hypothesis about the linear relationship between beer consumption and math performance.

5. A: $Math = \beta_0 + \beta_1 X1 + \beta_2 X2 + \varepsilon$ PA=3
C: $Math = \beta_0 + \beta_2 X2 + \varepsilon$ PC=2
$H_0: \beta_1 = 0$ PA-PC = 1 n-PA = 21

6. $F^*_{(2,21)} = [SSR/(PA-PC)]/[SSE(A)/(n-PA)]$
$2.59 = [92.16 + 0.21)/2]/[SSE(A)/21]$
$SSE(A) = 374.47$

$F^*_{(1,21)} = 92.16/(374.47/21) = 5.17$
$PRE = 92.16/(374.47+92.16) = 0.198$

The researchers can conclude that there is a significant linear relationship between alcohol consumption and math performance, such that increased alcohol consumption is associated with decreased performance.

Question C (you should write out the models for these tests)

1. Yes, as density in a state goes up, fatality rates come down.
PRE = .3648; F(1,48) = 27.57

2. Yes, the negative effect of density on fatality rates becomes less strong at higher levels of density. PRE = .08; F(1,47) = 4.16

3. In Model 1, the slope for density is the effect of that variable assuming linearity. In Model 2, the slope for density is the simple slope for that variable when density equals zero, and allowing the density-fatality relationship to be nonlinear. The simple slope at density equals zero is so much more negative, because the nonlinearity is such that the simple slope becomes less negative as density increases.

4. a) -.0042 (= -.00525 + (2)(.00000349)(144.72))
b) Model A: $Fat = \beta_0 + \beta_1 densd + \beta_2 densd^2 + \varepsilon$
Model C: $Fat = \beta_0 + \beta_2 densd^2 + \varepsilon$

5. From model 3, there is a marginally significant effect of temp on fatrate, controlling for density: as min temp goes up, fatrate goes up. From model 4, there is a nonsignificant simple effect of temp on fatrate when density equals zero, allowing the effect of temp on fatrate to depend on density. From model 5, there is a significant effect of temp on fatrate, controlling for linear and quadratic effects of density on fatrate. As temp goes up, fatrate goes up. Overall I would conclude that temp is predictive, once we control for density and allow density to have a nonlinear effect on fatrate.

6. No, there is no evidence that the effect of temp on fatrate depends on density.
PRE = .0026; F(1,46) = 0.12

7. No, this is not a case of too much power. The magnitude of the slope for the squared item depends upon the metric of the predictor and dependent variables. Here the metric of the predictor is very small: number of people per square mile. The slope is meaningful and significant.