Summative Quiz 2 (Multiple Logistic Regression) Solutions:

1. A random sample of 200 patients admitted to an adult intensive care unit (ICU) was collected to examine factors associated with death during hospital stay for ICU patients. Data was also collected on patient's age (in years), race, whether the patient had an infection at the time of ICU admission, and whether the patient had CPR administered prior to the hospital admission. The following table presents the unadjusted and adjusted slopes (**these have not been exponentiated**), and standard errors from logistic regressions, as well as the intercept for the multiple logistic regression model:

Table 1: Logistic Regression Results for Predictors of ICU Mortality
Regression Coefficents (SE)

Predictor	Unadjusted	Adjusted
Infection		
No	ref	ref
Yes	0.92 (0.36)	0.68 (0.38)
CPR Performed		
No	ref	ref
Yes	1.70 (0.60)	1.62 (0.62)
Age Quartile	p=0.04	
< 47 years	ref	ref
[47, 63) year	1.21 (0.62)	0.93 (0.64)
[64, 72) year	1.06 (0.63)	0.99 (0.65)
>= 72 years	1.58 (0.61)	1.47 (0.64)
Race		
White	ref	ref
Non-white	-0.68 (0.19)	-0.39 (0.67)
Intercept		-2.76

What is the adjusted ln(odds ratio of death) and 95% CI for those who received CPR prior to ICU admission, compared to those who did not receive CPR prior to ICU admission?

Answer: 1.62 (0.38, 2.86)

Reasoning: This adjusted *ln(odds ratio)* is simply the slope for "CPR performed" of 1.62. The 95% CI is given by $\hat{\beta}_{CPR} \pm 2\widehat{SE}(\hat{\beta}_{CPR})$, or 1.68±2(0.62)→ (0.38, 2.86).

2. (this item references the same logistic regression results as item #1)

What is the adjusted odds ratio of death and 95% CI for those who received CPR prior to ICU admission, compared to those who did not receive CPR prior to ICU admission?

Answer: $5.05\ (1.46, 17.46)$

Reasoning: For this, one just needs to exponentiate the results from part a: the estimated odds ratio is $e^{\hat{\beta}} = e^{1.62} \approx 5.05$. For the 95% CI for the odds ratio, exponentiate the endpoints of the 95% CI for the ln(odds ratio): ($e^{0.38}$, $e^{2.86}$) →(1.46, 17.46).

3. (this item references the same logistic regression results as item #1)

Does it appear that the association between death and receiving CPR was confounded by infection status and/or age and/or race? Why or why not?

Answer: No, because the adjusted odds ratio of death for those who received CPR (compared to those who did not receive CPR) is similar in value to than unadjusted odds ratio, and the 95% CIs overlap substantially.

4. (this item references the same logistic regression results as item #1)

Based on the multiple logistic regression results, what is the estimated probability of death in the ICU for 45 year old White patients who did not receive CPR and did not have an infection at the time of ICU admission?

Answer: 0.059

Reasoning: The multiple regression described by the above table ("Adjusted" column) is:

$$\begin{aligned}\ln(ODDs\ of\ death\ in\ ICU) \\ &= -2.76 + 1.62x_1 + 0.68x_2 + 0.93x_3 + 0.99x_4 + 1.47x_5 \\ &+ -0.39x_6\end{aligned}$$

Where x_1 = 1 if patient received CPR, x_2 = 1 if patient had infection, x_3=1 if patient's age in quartile 2, x_4=1 if patient's age in quartile 3, x_5=1 if patient's age in quartile 4 and x_6 =1 if patient is Non-White.

This group of patients x's are 0 for all xs, x_1-x_6: So the resulting $\ln(ODDs\ of\ death\ in\ ICU) = -2.76$. Hence, $\widehat{ODDS} = e^{-2.76} \approx 0.063$ and $\hat{p} = \frac{\widehat{ODDS}}{1+\widehat{ODDS}} = \frac{0.063}{1.063} \approx 0.059, or\ 5.9\%.$

5. (this item references the same logistic regression results as item #1)

Based on the multiple logistic regression results, what is the estimated probability of death in the ICU for 75 year old White patients received CPR and had an infection at the time of ICU admission?

Answer: 0.73

Reasoning: $\ln(ODDs\ of\ death\ in\ ICU) = -2.76 + 1.62 + 0.68 + 1.47 = 1.01$

Hence, $\widehat{ODDs} = e^{1.01} = 2.75$ and $\hat{p} = \frac{\widehat{ODDs}}{1+\widehat{ODDs}} = \frac{2.75}{3.75} \approx 0.733$ or 73.3 %.

6. (this item references the same logistic regression results as item #1)

Based on the multiple logistic regression results, what is the estimated odds ratio of death in the ICU for 75 year old White patients received CPR and had an infection at the time of ICU admission compared to 45 year old white patients who did not receive CPR and did not have an infection at the time of ICU admission?

Answer: 43.7

Reasoning: Reasoning: Based on the computations for items 4 and 5, the estimated odds of death for the first group in the comparison is 2.75, and 0.063 for the second group. Hence the estimated odds ratio of death for the first group compared to the second group is $\widehat{OR} = \frac{2.75}{0.063} \approx 43.7.$

7. (this item references the same logistic regression results as item #1)

   Based on the multiple logistic regression results, what is the estimated relative risk of death in the ICU for 75 year old White patients received CPR and had an infection at the time of ICU admission compared to 45 year old white patients who did not receive CPR and did not have an infection at the time of ICU admission ?

Answer: 12.4

Reasoning: Reasoning: Based on the computations for items 4 and 5, the estimated risks of death for the first group in the comparison is 0.73, and 0.059 for the second group. Hence the estimated odds ratio of death for the first group compared to the second group is
$\widehat{RR} = \frac{0.73}{0.059} \approx 12.4.$

8. (this item makes a reference to the logistic regression results from item #1)

Model 2 was extended to include an interaction term between infection (yes/no) and race (non-white vs. white). The results are as follows:

$\ln(odds\ of\ death) = -2.80 + 0.73x_1 + 0.18x_2 + 0.54x_3 +$ *(slopes and xs for CPR and age quartiles)*

where x_1 = 1 if patient had infection; x_2= 1 for non-white, 0 white ; x_3=x_1*x_2;

Based on this model, what is the estimated (CPR and age adjusted) odds ratio of death for white patients who had an infection compared to white patients who did not have an infection?

Answer: 2.08

Reasoning: For those who identify as white, x_2=0, and hence x_3= x_1*0 =0. The resulting equation reduces to:

$\ln(odds\ of\ death) = -2.8 + 0.73x_1 + -0.18(0) + -0.54(0) +$
$(slopes\ and\ xs\ for\ CPR\ and\ age\ quartiles)$ =

$\ln(odds\ of\ death) = -2.8 + 0.73x_1 + slopes\ and\ xs\ for\ CPR\ and\ age\ quartiles)$

So the slope for infection status for those who identify as white is simply 0.73, and the corresponding adjusted odds ratio is $e^{0.73} \approx 2.08$.

9. (this item makes a reference to the logistic regression results from item #1)

Model 2 was extended to include an interaction term between infection (yes/no) and race (non-white vs. white). The results are as follows:

$\ln(odds\ of\ death) = -2.80 + 0.73x_1 + 0.18x_2 + 0.54x_3 +$ *(slopes and xs for CPR and age quartiles)*

where $x_1 = 1$ if patient had infection; $x_2= 1$ for non-white, 0 white ; $x_3=x_1*x_2$;

Based on this model, what is the estimated (CPR and age adjusted) odds ratio of death for non-white patients who had an infection compared to non-white patients who did not have an infection?

Answer: 1.21

Reasoning: For those who identify as non-white, x_2=0, and hence $x_3= x_1*1= x_1$.

The resulting equation reduces to:
$\ln(odds\ of\ death) = -2.8 + 0.73x_1 + -0.18(1) + -0.54(x_1 \times 1) + (slopes\ and\ xs\ for\ CPR\ and\ age\ quartiles)$ =

$\ln(odds\ of\ death) = -2.8 + -0.18 + (0.73 + -0.54)x_1 + slopes\ and\ xs\ for\ CPR\ and\ age\ quartiles)$

So the slope for infection status for those who identify as white is 0.73+-0.54 = -0.19, and the corresponding adjusted odds ratio is $e^{0.19} \approx 1.21$.

Notice that $e^{0.19} = e^{0.73+\ -0.54} = (e^{0.73})(e^{-0.54}) = (2.08)(0.58)$. So $e^{\hat{\beta}_3}$ is what?

10. (this item makes a reference to the logistic regression results from item #1)

Model 2 was extended to include an interaction term between infection (yes/no) and race (non-white vs. white). The results are as follows:

$\ln(odds\ of\ death) = -2.80 + 0.73x_1 + 0.18x_2 + 0.54x_3 +$ *(slopes and xs for CPR and age quartiles)*

where x_1 = 1 if patient had infection; x_2= 1 for non-white, 0 white ; x_3=x_1*x_2;

The standard error for the slope of the interaction term (x) is 1.4. Is the (CPR and age adjusted) relationship between death in the ICU and infection statistically significantly modified by race?

Answer: No

Reasoning: The 95% CI for the true value of β_3 is -0.54±2(1.4) → (-3.34, 2.26). As this confidence interval includes the null value of 0, the resulting p-value for testing $H_o: \beta_3 = 0$ will be greater than 0.05.

11. Data from the National Epidemiologic Survey on Alcohol and Related Conditions were analyzed to measure the association between first time homelessness, poverty and substance abuse . The abstract from this research is as follows:

Objectives. We examined whether substance-use disorders and poverty predicted first-time homelessness over 3 years.

Methods. We analyzed longitudinal data from waves 1 (2001–2002) and 2 (2004–2005) of the National Epidemiologic Survey on Alcohol and Related Conditions to determine the main and interactive effects of wave 1 substance use disorders and poverty on first-time homelessness by wave 2, among those who were never homeless at wave 1 (n = 30 558). First-time homelessness was defined as having no regular place to live or having to live with others for 1 month or more as a result of having no place of one's own since wave 1.

Results. Alcohol-use disorders (adjusted odds ratio [AOR] = 1.34), drug-use disorders (AOR = 2.51), and poverty (AOR = 1.34) independently increased prospective risk for first-time homelessness, after adjustment for ecological variables. Substance-use disorders and poverty interacted to differentially influence risk for first-time homelessness ($P < .05$), before, but not after, adjustment for controls.

Conclusions. This study reinforces the importance of both substance-use disorders and poverty in the risk for first-time homelessness, and can serve as a benchmark for future studies. Substance abuse treatment should address financial status and risk of future homelessness. (*Am J Public Health.* 2013;103: S282–S288. doi:10.2105/AJPH.2013.301302)

The authors present the following results (unadjusted and adjusted odds ratios) from simple and multiple logistic regression models that relate first time homelessness to poverty, substance abuse and other predictors:

TABLE 3—Associations Between First-time Homelessness at Wave 2 and Poverty, Substance-Use Disorders, and Control Variables at Wave 1: National Epidemiologic Survey on Alcohol and Related Conditions, United States, 2001–2005

Variable	Unadjusted OR (95% CI)	Adjusted[a] OR (95% CI)
Main predictors		
Poverty	2.31 (1.94, 2.75)	1.34 (1.09, 1.64)
Alcohol- and drug-use disorders (Ref = neither disorder)[b]		
Alcohol-use disorder only	2.23 (1.80, 2.77)	1.33 (1.06, 1.67)
Drug-use disorder only	5.39 (3.44, 8.43)	2.51 (1.53, 4.11)
Both alcohol- and drug-use disorders	4.78 (2.89, 7.91)	1.55 (0.87, 2.79)
Control variables		
Age (Ref = ≥ 50), y		
18-29	8.53 (6.92, 10.51)	6.40 (5.08, 8.07)
30-39	3.39 (2.68, 4.28)	3.53 (2.79, 4.48)
40-49	1.97 (1.49, 2.59)	2.09 (1.58, 2.76)
Race (Ref = Non-Hispanic White)		
Non-Hispanic Black	1.74 (1.43, 2.12)	1.12 (0.90, 1.39)
Native American	0.98 (0.58, 1.64)	0.80 (0.47, 1.37)
Asian/Pacific Islander	0.69 (0.47, 1.01)	0.52 (0.36, 0.77)
Hispanic	1.25 (0.99, 1.57)	0.66 (0.53, 0.84)
Gender (Ref = male)		
Female	0.90 (0.77, 1.04)	0.99 (0.84, 1.17)
Education (Ref = at least some college)		
< high school	1.54 (1.25, 1.89)	1.70 (1.35, 2.14)
High school graduate	1.16 (0.99, 1.36)	1.21 (1.02, 1.42)
Married or live as married	0.33 (0.28, 0.38)	0.56 (0.47, 0.67)
Live in urban area	1.14 (0.94, 1.38)	1.09 (0.89, 1.34)
State cost of living above average	0.95 (0.82, 1.10)	1.10 (0.90, 1.34)
Region (Ref = Northeast)		
Midwest	1.15 (0.93, 1.42)	1.11 (0.89, 1.37)
South	1.42 (1.17, 1.72)	1.44 (1.13, 1.84)
West	1.49 (1.20, 1.86)	1.58 (1.25, 2.00)
Any psychiatric disorder	2.77 (2.38, 3.23)	2.08 (1.77, 2.44)

Note. CI = confidence interval; OR = odds ratio.
[a]Model simultaneously controls for all variables in the table.
[b]Raw sample sizes in each category: only alcohol-use disorder (n = 2364), only drug-use disorder (n = 282), both alcohol- and drug-use disorder (n = 330), neither disorder (n = 27 582).

Generally speaking, how do the adjusted results for substance abuse disorders compare to the unadjusted results?

Answer: While the three substance abuse disorders are still positively and statistically significantly associated (with the exception of alcohol and drug use disorder) with the risk of homelessness, the magnitudes of these associations attenuated (decreased) after adjustment

12. (this item refers to the same results presented in item #11)

Based on the multiple logistic regression results (the "adjusted column") what is the estimated odds ratio of homelessness for :20 year olds living in poverty who have a drug use disorder compared to 55 year olds not in poverty with no substance abuse disorder, where both groups are otherwise comparable on the other adjustment variables?

Answer: 21.5.

Reasoning: 1.34*2.51*6.4=21.5

These differences are multiplicative on the odds ratio scale. (If you were to go back to the logistic regression scale, the difference would be additive:

The ln(odds ratio) would be $\hat{\beta}_{age\ 18-19} + \hat{\beta}_{poverty} + \hat{\beta}_{drug\ use\ disorder}$. Exponentiating this yields $e^{\hat{\beta}_{age\ 18-19} + \hat{\beta}_{poverty} + \hat{\beta}_{drug\ use\ disorder}} =$
$\left(e^{\hat{\beta}_{age\ 18-19}}\right)\left(e^{\hat{\beta}_{poverty}}\right)\left(e^{\hat{\beta}_{drug\ use\ disorder}}\right) =$
$(\widehat{AOR}_{age\ 18-19})(\widehat{AOR}_{poverty})(\widehat{AOR}_{drug\ use\ disorder})$, ie: the product of the 3 adjusted odds ratios.