Pharm 3011 - Fall 2019 - Assignment 2 Solutions
(out of 50 points)

1. Following is a partial computer output for a regression of body length of newborn on gestational age, birthweight, mom's age and toxemia. There were 85 observations in the data set.

   Let y denote body length in cm, x_1 denote gestational age in weeks, x_2 denote birthweight in grams, x_3 denote mom's age in years and x_4 denote presence ($x_4 = 1$) or absence ($x_4 = 0$) of toxemia. The regression equation, formally, is

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$$

   where the errors ϵ are assumed to be normal with mean 0 and standard deviation σ.

```
The regression equation is

length = 1.8 + 0.36 gestage + 0.01 birthwt -0.03 momage - 1.00 toxemia

Predictor    Coef   SE Coef     T
Constant     1.80     3.0      0.6
gestage      0.36    0.12      3.0
birthwt      0.01    0.001    10.0
momage      -0.03    0.03     -1.0
toxemia     -1.00    4.00     -0.25

Analysis of Variance

Source          DF    SS     MS      F      P
Regression       4   900  225.0   75.0  0.000
Residual Error  80   240    3.0
Total           84  1140
```

   For this data set, the error sum of squares SSE=240.

   (a) What is the predicted body length of a baby of gestational age of 33 weeks and birthweight 1250 grams, whose mother is 37 years of age and was NOT toxemic?

$$1.8 + .36(33) + .01(1250) - .03(37) - 1.00(0) = 25.07$$

(3)

   (b) All other things being equal what is the mean difference in body length of a baby of a toxemic mother, as compared to a non-toxemic mother?

(3) Average length is shorter by 1.00 for a toxemic mother.

   (deduct 1 point if the answer is just 1 or -1, without indicating which group is shorter or longer on average)

(c) What is the estimate of σ, the standard deviation of the errors in the regression model?

(3) $\sqrt{3.0} \approx 1.73$

(d) Of the four predictor variables, which is the least useful for predicting body length, given that the other three are already included in the regression model? Why?

(3) toxemia. It has the smallest absolute value of the T statistic.

It's also OK to calculate p-values, and report that toxemia has the largest p-value, but that's not necessary.

(e) What proportion of the variation in body length is explained by the linear relationship of length with the four predictor variables?

(3) $R^2 = 900/1140 \approx .79$, or 79%

(f) What is the least squares estimate of β_3?

(2) -.03

(g) Construct a 90% confidence interval for β_3. Note that $t_{.05,80} = 1.664$.

(3) $-.03 \pm 1.664(.03)$, or $-.03 \pm .0499$ or (-.0799, .0199)

(subtract 1 point for each error)

2. **There were n=24 observations from a similar study at a different hospital.** The following is a partial computer output using those data, and carrying out a regression of body length y on gestational age (x_1), toxemia (x_2), and the interaction between toxemia and gestational age ($x_1 \times x_2$).

Formally, the regression equation is

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon$$

```
The estimated regression equation is

length = 2.0 + 1.0 gestage - 3.0 toxemia - 0.5 gestage*toxemia
```

Predictor	Coef	SE Coef	T
Constant	2.0		2.0
gestage	1.0		10.0
toxemia	-3.0		-2.0
interaction	-0.5	0.25	

(a) What is the difference in predicted body length for two babies of gestational age x, where one of the mothers was toxemic, and the other was not? (Note: the answer should be a formula, not a number.)

(4) $\pm\{(2.0 + 1.0x - 3.0 - 0.5x) - (2.0 + 1.0x)\} or \pm \{-3.0 - 0.5x\}$

(subtract 1 point for each error. The $\pm$ means that either the expression as given, or its negative, are correct.)

(b) In testing for the effect of toxemia $H_0 : \beta_2 = 0$, against the two sided alternative, the observed value of the test statistic is -2.0.

(2) i. What are the degrees of freedom? $n - 1 - 3 = 24 - 4 = 20$

(2) ii. Bound the p-value as closely as possible. $2P(t_{20} > |-2.0|) = 2P(t_{20} > 2.0)$
Using the class tables, the p-value is in the interval (.05,.10).

(c) What is the observed value of the test statistic used to test for the significance of the interaction?

(2) $t_{obs} = -0.5/.25 = -2.0$

3. In the meta-analysis of prophylactic lidocaine as a treatment for acute myocardial infarction, suppose we focus on the probability of death in the control group. The data are

Study	randomized	dead				
i	n_i	x_i	$\hat{p}_i$	s_i^2	W_i	$\frac{W_i \hat{p}_i}{\sum W_i}$
1.	43	1	.023	.0005	1893	.0026
2.	44	4	$\hat{p}_2$	s_2^2	W_2	.0028
3.	110	4	.036	.0003	3140	.0067
4.	100	5	.050	.0005	2105	.0062
5.	106	3	.028	.0003	3854	.0064
6.	146	4	.027	.0002	5479	.0088
Total	549	21			17003	

where $\hat{p}_i = \frac{x_i}{n_i}$, $s_i^2 = \frac{\hat{p}_i(1-\hat{p}_i)}{n_i}$, and $W_i = 1/s_i^2$.

(a) Calculate a 95% confidence interval for the probability p of death for patient in the control group, **using only the data from the first study**. (ie $\hat{p}_1 \pm z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1}$)

(5) $.023 \pm 1.96\sqrt{.0005}$, or $.023 \pm 1.96(.023)$, which is $.023 \pm .045$, or (-.022, .068).

(5 points. subtract one point for each error.)

(b) What is the value of $\hat{p}_2$?

(1) $\hat{p}_2 = 4/44 \approx .091$

(c) What is the value of s_2^2?

(2) $s_2^2 = \frac{(4/44)(1-4/44)}{44} \approx .0019$

(d) What is the value of W_2? $W_2 = 1/s_2^2 \approx 532.4$ (OK to round to 532, to be compatible with other tabulated entries.If a rounded value reported in part (c) was used, there will be substantial error in the weight, for example $1/.002 = 500$. Subtract 1 point if this

(2) was the case.)

(e) Carry out a meta analysis to construct a 95% CI for the probability p of death for patient in the control group.

$\hat{p} = .0026 + .0028 + .0067 + .0062 + .0064 + .0088 = 0.0335$

Confidence interval is $.0335 \pm 1.96\sqrt{1/17003}$, or, approximately, $.0335 \pm 1.96(.0077)$.

(10) The 95% confidence interval is $.0335 \pm .0150$, or (.0185, .0485). (10 points. Give full marks for reporting the interval in either of these two ways. Subtract 1 point for each error. It's OK to report to 3 decimal digits, eg [.018, .049])