Homework 9

Due: 11:59pm, Thursday, April 28

Instruction: Please scan or typeset your solutions and upload them as a single pdf file to Canvas. Do not just take a picture of your solutions.

0. Readings: Sections 9.3, 9.4, 9.5, 10.1 and Notes 11 and 12.

1. Let $Y_1, Y_2, \ldots, Y_n \sim_{iid} N(\theta, \sigma^2)$ where σ^2 is known. Consider

$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0$$

(a) What would be Type I Error? What would be Type II Error?

(b) In HW8, we show that the likelihood ratio test procedure is to reject H_0 if

$$z = \frac{\bar{y} - \theta_0}{\sigma/\sqrt{n}} \geq k_1 = \sqrt{-2 \log k}.$$

The power function of this test is (shown in the class, and see notes 11)

$$\gamma(\theta) = P(N(0,1) \geq k_1 + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}).$$

Let $\theta_0 = 105, \sigma = 10, n = 100$ and $k_1 = 1.8$, plot this function, and comment on your plot.

(c) For this test, what is the probability of Type I Error when $\theta = 105$?

(d) For this test, what is the probability of Type II Error when $\theta = 110$? What is the power of rejecting H_0 when $\theta = 110$?

(e) If we set the significance level $\alpha = 0.05$, what is k_1?

(f) For this test procedure with $\alpha = 0.05$, what sample size n is necessary to ensure that the power of rejecting H_0 at $\theta = 108$ is at least 80%?

2. For the carprice example, see `Notes 7`, we have the following R '`lm`' output:

```
##
## Call:
## lm(formula = y ~ x1 + x2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.364  -5.243   1.028   5.926  11.545
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 183.0352    11.3476  16.130 2.19e-07 ***
## x1           -9.5043     3.8742  -2.453   0.0397 *
## x2           -0.8215     0.2552  -3.219   0.0123 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.805 on 8 degrees of freedom
## Multiple R-squared:  0.9361, Adjusted R-squared:  0.9201
## F-statistic: 58.61 on 2 and 8 DF,  p-value: 1.666e-05
```

(a) What is the $\hat{\beta}_1$? How do you interpret this number?

(b) To test $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, what is the P-value? What is your conclusion?

(c) Based on this output, what is the prediction of the average car price for a 3-year-old car with mileage of $25,000$?

3. For the Default example, see `Notes 8`, we have the following R '`glm`' output

```
##
## Call:
## glm(formula = default ~ student + balance + income, family = "binomial",
##     data = Default)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4691  -0.1418  -0.0557  -0.0203   3.7383
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.087e+01  4.923e-01 -22.080  < 2e-16 ***
## studentYes  -6.468e-01  2.363e-01  -2.738  0.00619 **
## balance      5.737e-03  2.319e-04  24.738  < 2e-16 ***
## income       3.033e-06  8.203e-06   0.370  0.71152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2920.6  on 9999  degrees of freedom
## Residual deviance: 1571.5  on 9996  degrees of freedom
## AIC: 1579.5
##
## Number of Fisher Scoring iterations: 8
```

(a) What is the $\hat{\beta}_1$? How do you interpret this number?

(b) What is your prediction of Default for someone who is a student, with balance of 800 and income of $15,000$?

(c) To test $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, what is the p-value? What is your conclusion?