MTH 326 - Spring 2022
Assignment #8
Due: Wednesday, April 13 2022 (11:59pm)

1. Suppose the following represents a random sample of points (x, y):

x	-2	-1	0	1	2
y	3	2	1	1	0.5

(a) Find the 90% confidence interval for $\mathrm{E}(Y)$ when $x^* = 0$ and again when $x^* = 2$.

Solution: For a 90% CI then $\alpha = 1 - 0.9 = 0.1$ and a two sided is $\alpha/2 = 0.05$. There are $n-2$ degrees of freedom, hence df $= 5 - 2 = 3$. Therefore $t_{\alpha/2}(\mathrm{df}) = t_{0.05}(3) = 2.353$. We want to use the following formula,

$$\left(\widehat{\beta_0} + \widehat{\beta_1} x^*\right) \pm t_{\alpha/2}(\mathrm{df}) S \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

so we need to compute the remaining unknows. It is trivial that $\bar{x} = 0$ and $\bar{y} = 1.5$. Then

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 = 4 + 1 + 0 + 1 + 4 = 10$$

$$S_{yy} = \sum (y_i - \bar{y})^2 = \sum (y_i - 1.5)^2 = (1.5)^2 + (0.5)^2 + 2(-0.5)^2 + (-1)^2 = 4$$

and

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i(y_i - 1.5) = -2(1.5) - (0.5) + (-0.5) + 2(-1) = -6.$$

Then

$$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}} = \frac{-6}{10} = -0.6$$

and

$$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x} = 1.5.$$

Then

$$S = \sqrt{\frac{\mathrm{SSE}}{n-2}} = \sqrt{\frac{S_{yy} - \widehat{\beta_1} S_{xy}}{n-2}} = \sqrt{\frac{4 - (-0.6 \cdot -6)}{3}} = \sqrt{\frac{0.4}{3}} \approx 0.365148.$$

Therefore our inverval is

$$(1.5 - 0.6x^*) \pm 2.353 \cdot 0.365148 \sqrt{\frac{1}{5} + \frac{(x^*)^2}{10}}$$

so for $x^* = 0$ then

$$1.5 \pm 2.353 \cdot 0.365148 \sqrt{\frac{1}{5}} \equiv (1.116, 1.884)$$

and for $x^* = 2$ then

$$(1.5 - 0.6(2)) \pm 2.353 \cdot 0.365148 \sqrt{\frac{1}{5} + \frac{(2)^2}{10}} \equiv (-0.366, 0.966).$$

(b) Find the 90% confidence interval for Y^* when $x^* = 0$ and again when $x^* = 2$.

Solution: We adjust the interval to

$$(1.5 - 0.6x^*) \pm 2.353 \cdot 0.365148\sqrt{1 + \frac{1}{5} + \frac{(x^*)^2}{10}}$$

so for $x^* = 0$ then

$$1.5 \pm 2.353 \cdot 0.365148\sqrt{1 + \frac{1}{5}} \equiv (0.5588, 2.4412)$$

and for $x^* = 2$ then

$$(1.5 - 0.6(2)) \pm 2.353 \cdot 0.365148\sqrt{1 + \frac{1}{5} + \frac{(2)^2}{10}} \equiv (-0.7868, 1.3868).$$

(c) Are the intervals for $\mathrm{E}(Y)$ or the intervals for Y^* wider? How can this be explained?

Solution: The intervals for Y^* are wider than $\mathrm{E}(Y)$ because the $(\widehat{\beta}_0 + \widehat{\beta}_1 x^*)$ is held constant as the center of the interval, then the scalar $t_{\alpha/2}(\mathrm{df})S$ is a fixed constant. Lastly, the scalar value under the squareroot is positive since $\frac{1}{n} > 0$, $(x^* - \bar{x})^2 > 0\,\forall\, x \in \mathbb{R}$, and $S_{xx} > 0$ since its also squares. Adding positive values together is stilll positive. Hence adding another positive number (+1) to this value results in a larger positive value. Therefore the squareroot of the new positive value is also larger. Hence the scalar around the center of the interval is larger ("wider").

$$\overbrace{\left(\widehat{\beta}_0 + \widehat{\beta}_1 x^*\right) \pm t_{\alpha/2}(\mathrm{df})S\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}}^{\mathrm{E}(Y)} < \overbrace{\left(\widehat{\beta}_0 + \widehat{\beta}_1 x^*\right) \pm t_{\alpha/2}(\mathrm{df})S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}}^{Y^*}$$

$$\iff t_{\alpha/2}(\mathrm{df})S\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} < t_{\alpha/2}(\mathrm{df})S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

$$\iff \sqrt{\underbrace{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}_{>0}} < \sqrt{1 + \underbrace{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}_{>0}}$$

The following table contains dietary data (calories and the content of fat, sodium, carbohydrate, and protein) in some standard hamburgers that can be found at local fast food restaurants.

	cal	fat (g)	sodium (mg)	carbs (g)	protein (g)
BK Jr.	310	18	390	27	13
Wendy's Jr.	250	11	420	25	13
McDonald's	250	9	480	31	12
Culvers	390	17	480	38	20
Steak-n-Shake	320	14	830	32	15
Sonic Jr.	330	16	610	32	15

2. Assume the relationship between defining calories is a linear one. That is,

$$\text{cal} = A \cdot (\text{fat}) + B \cdot (\text{carbs}) + C \cdot (\text{protein}).$$

(a) Determine the best-fit hyperplane that predicts calories of a burger based upon the fat, carbohydrate, and protein content of the burger. Explicitly state the terms in the normal forms: X, $X^T X$, $(X^T X)^{-1}$, and $X^T Y$.

Solution: Let the fat, carbs, and protein be the column vectors of X and calories be the column vector of Y. Then

$$X = \begin{bmatrix} 18 & 27 & 13 \\ 11 & 25 & 13 \\ 9 & 31 & 12 \\ 17 & 38 & 20 \\ 14 & 32 & 15 \\ 16 & 32 & 15 \end{bmatrix} \qquad \text{and} \qquad Y = \begin{bmatrix} 310 \\ 250 \\ 250 \\ 390 \\ 320 \\ 330 \end{bmatrix}.$$

Then

$$X^T X = \begin{bmatrix} 18 & 11 & 9 & 17 & 14 & 16 \\ 27 & 25 & 31 & 38 & 32 & 32 \\ 13 & 13 & 12 & 20 & 15 & 15 \end{bmatrix} \begin{bmatrix} 18 & 27 & 13 \\ 11 & 25 & 13 \\ 9 & 31 & 12 \\ 17 & 38 & 20 \\ 14 & 32 & 15 \\ 16 & 32 & 15 \end{bmatrix} = \begin{bmatrix} 1267 & 2646 & 1275 \\ 2646 & 5807 & 2768 \\ 1275 & 2768 & 1332 \end{bmatrix}.$$

Expanding across the first row,

$$\begin{aligned} \begin{vmatrix} 1267 & 2646 & 1275 \\ 2646 & 5807 & 2768 \\ 1275 & 2768 & 1332 \end{vmatrix} &= 1267 \begin{vmatrix} 5807 & 2768 \\ 2768 & 1332 \end{vmatrix} + 2646 \begin{vmatrix} 2646 & 2768 \\ 1275 & 1332 \end{vmatrix} + 1275 \begin{vmatrix} 2646 & 5807 \\ 1275 & 2768 \end{vmatrix} \\ &= 92617700 - (-12510288) - 101741175 \\ &= 3386813. \end{aligned}$$

Since $\det\left(X^T X\right) \neq 0$ then the inverse exists.

$$\left[X^T X \mid I_3\right] = \left[\begin{array}{ccc|ccc} 1267 & 2646 & 1275 & 1 & 0 & 0 \\ 2646 & 5807 & 2768 & 0 & 1 & 0 \\ 1275 & 2768 & 1332 & 0 & 0 & 1 \end{array}\right]$$

This row reduces to $\left[I_3 \mid (X^TX)^{-1}\right]$ in which

$$(X^TX)^{-1} = \frac{1}{3{,}386{,}813}\begin{bmatrix} 73100 & 4728 & -79797 \\ 4728 & 62019 & -133406 \\ -79797 & -133406 & 356153 \end{bmatrix}.$$

Lastly,

$$X^TY = \begin{bmatrix} 18 & 11 & 9 & 17 & 14 & 16 \\ 27 & 25 & 31 & 38 & 32 & 32 \\ 13 & 13 & 12 & 20 & 15 & 15 \end{bmatrix}\begin{bmatrix} 310 \\ 250 \\ 250 \\ 390 \\ 320 \\ 330 \end{bmatrix} = \begin{bmatrix} 26970 \\ 57990 \\ 27830 \end{bmatrix}.$$

Solving the normal equation,

$$\begin{aligned} \widehat{\beta} &= (X^TX)^{-1}X^TY \\ &= \frac{1}{3{,}386{,}813}\begin{bmatrix} 73100 & 4728 & -79797 \\ 4728 & 62019 & -133406 \\ -79797 & -133406 & 356153 \end{bmatrix}\begin{bmatrix} 26970 \\ 57990 \\ 27830 \end{bmatrix} \\ &= \frac{1}{3{,}386{,}813}\begin{bmatrix} 24{,}933{,}210 \\ 11{,}306{,}990 \\ 23{,}398{,}960 \end{bmatrix} \\ &\approx \begin{bmatrix} 7.36185021 \\ 3.3385339 \\ 6.90884321 \end{bmatrix}. \end{aligned}$$

Therefore the hyperplane of best fit is

$$\widehat{y} = 7.36185021 \cdot (\text{fat}) + 3.3385339 \cdot (\text{carbs}) + 6.90884321 \cdot (\text{protein}).$$

Testing this on our data:

```
BK Jr:
    Actual:   310.0
    Estimate: 312.46868081
Wendy's Jr:
    Actual:   250.0
    Estimate: 254.25866154
McDonalds:
    Actual:   250.0
    Estimate: 252.65732131
Culvers:
    Actual:   390.0
    Estimate: 390.19260596999993
Steak-n-shake:
    Actual:   320.0
    Estimate: 313.53163588999996
Sonic Jr:
    Actual:   330.0
    Estimate: 328.25533630999996
```

(b) Based upon your least-squares regression analysis, how many calories are expected in a burger made with 10 grams of fat, 20 grams of carbohydrates, and 15 grams of protein.

Solution:

$$\widehat{y}(10, 20, 15) = 7.36185021 \cdot (10) + 3.3385339 \cdot (20) + 6.90884321 \cdot (15) \approx 244.02182825 \text{ cal}$$

(c) Find SSE and S.

Solution: We need $\text{SSE} = Y^TY - \widehat{\beta}^TX^TY$ and using the matrices from part (a) we obtain $Y^TY = 584{,}500$ and

$$\widehat{\beta}^TX^TY = 7.36185021 \cdot 26970 + 3.3385339 \cdot 57990 + 6.90884321 \cdot 27830 = 584423.788.$$

Hence $\text{SSE} = 76.212$.

Now using $S = \sqrt{\frac{\text{SSE}}{\text{df}}}$, since we aren't using a constant, we only have 3 variables, so $\text{df} = n - (k+1) = 6 - (2+1) = 3$. Therefore

$$S = \sqrt{\frac{76.212}{3}} \approx 5.03722807$$

(d) Find a 95% confidence interval for the amount of calories in a burger made with 10 grams of fat, 20 grams of carbohydrates, and 15 grams of protein.

Solution: Let a vector $a \in \mathbb{R}^3$ be defined as $a = \begin{pmatrix} 10 \\ 20 \\ 15 \end{pmatrix}$. By duality we will treat this as a matrix. As shown in (c), we have 3 degrees of freedom, so for $\alpha = 1 - 0.95 = 0.05$ then $t_{\alpha/2}(\text{df}) = t_{0.05}(3) = 3.182$. Then using the formula

$$a^T\widehat{\beta} \pm t_{\alpha/2}S\sqrt{a^T(X^TX)^{-1}a},$$

we have $a^T\widehat{\beta} = 244.02182825$ from part (b). Next,

$$\begin{aligned}
a^T(X^TX)^{-1}a &= (10, 20, 15) \cdot \frac{1}{3{,}386{,}813}\begin{bmatrix} 73100 & 4728 & -79797 \\ 4728 & 62019 & -133406 \\ -79797 & -133406 & 356153 \end{bmatrix} \cdot \begin{pmatrix} 10 \\ 20 \\ 15 \end{pmatrix} \\
&= (10, 20, 15) \cdot \frac{1}{3{,}386{,}813}\begin{pmatrix} -371395 \\ -713430 \\ 1876205 \end{pmatrix} \\
&= \frac{1}{3{,}386{,}813}\left[10 \cdot -371395 + 20 \cdot -713430 + 15 \cdot 1876205\right] \\
&= \frac{10{,}160{,}525}{3{,}386{,}813} \\
&\approx 3.00002539.
\end{aligned}$$

Therefore

$$t_{\alpha/2}S\sqrt{a^T(X^TX)^{-1}a} = 3.182 \cdot 5.03722807\sqrt{3.00002539} \approx 27.76222$$

and

$$I \equiv 244.02182825 \pm 27.76222 \equiv (216.259608,\ 271.784048).$$

3. A family kept track of its natural gas usage for two heating seasons and the accompanying outdoor temperatures. Gas usage is measured in hundred of cubic feet, and temperature is average temperature is Fahrenheit.

month:	Oct 1	Nov 1	Dec 1	Jan 1	Feb 1	Mar 1
temp (°F):	53	41	14	22	32	39
gas:	3.0	6.2	12.6	9.2	7.5	5.5
month:	Oct 2	Nov 2	Dec 2	Jan 2	Feb 2	Mar 2
temp (°F):	56	36	33	13	35	43
gas:	2.0	6.5	7.3	12.5	6.9	5.3

(a) Fit the model $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$ where X is the temperature and Y is the gas usage.

Solution: We need to construct the X and Y matrices first, so

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 53 & 41 & 14 & 22 & 32 & 39 & 56 & 36 & 33 & 13 & 35 & 43 \\ 53^2 & 41^2 & 14^2 & 22^2 & 32^2 & 39^2 & 56^2 & 36^2 & 33^2 & 13^2 & 35^2 & 43^2 \end{bmatrix}^T$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 53 & 41 & 14 & 22 & 32 & 39 & 56 & 36 & 33 & 13 & 35 & 43 \\ 2809 & 1681 & 196 & 484 & 1024 & 1521 & 3136 & 1296 & 1089 & 169 & 1225 & 1849 \end{bmatrix}^T$$

and

$$Y = \begin{bmatrix} 3 & 6.2 & 12.6 & 9.2 & 7.5 & 5.5 & 2 & 6.5 & 7.3 & 12.5 & 6.9 & 5.3 \end{bmatrix}^T.$$

Computing $X^T X$, which exceeds Symbolab's space and WolframAlpha's "maximum number of characters", but my old java code works for 2147483641 x 2147483641 matrices, (assuming you have 32 EB (exabytes) of RAM (and a lot of patience)), hence

$$X^T X = \begin{bmatrix} 12 & 417 & 16479 \\ 417 & 16479 & 706065 \\ 16479 & 706065 & 31998951 \end{bmatrix}.$$

Now WolframAlpha can handle this, and $\det X^T X = 9{,}928{,}933{,}560 \neq 0$. Hence, it is invertible and bijective. Then, according to WolframAlpha (after setting it to decimal approximations),

$$(X^T X)^{-1} \approx \begin{bmatrix} 2.89889 & -0.172054 & 0.00230354 \\ -0.172054 & 0.0113235 & -0.00016125 \\ 0.00230354 & -0.00016125 & 2.40298 \cdot 10^{-6} \end{bmatrix}.$$

Then using my code with rational precision,

$$X^T Y = \begin{bmatrix} 169/2 \\ 24653/10 \\ 169655/2 \end{bmatrix} = \begin{bmatrix} 84.5 \\ 2465.3 \\ 84827.5 \end{bmatrix}.$$

Then using my code with floating point precision,

$$\begin{aligned}\widehat{\beta} &= (X^T X)^{-1} X^T Y \\ &= \begin{bmatrix} 16.195018149999925 \\ -0.30117282499999476 \\ 0.0009574426749999754 \end{bmatrix}\end{aligned}$$

Therefore $\widehat{y} \approx 16.195018 - 0.301172825x + 0.000957442675x^2$.

(b) Test if the quadratic term is really necessary. That is, test $H_0 := \beta_2 = 0$.

Solution: We need to compute $\mathrm{Var}(\widehat{\beta}_2) = S^2 c_{2\,2}$. From $(X^T X)^{-1}$ we already have $c_{2\,2} = 0.00000240297$, so we're left with computing S^2. This requires the SSE and $\mathrm{SSE} = Y^T Y - \widehat{\beta}^T X^T Y$. Hence, computing the components (with code) $Y^T Y = 708.83$ and $\widehat{\beta}^T X^T Y \approx 707.2151367160669$. Then SSE ≈ 1.61486328.

Then we have $\mathrm{df} = n - (k+1) = 12 - (2+1) = 9$. As such,

$$S^2 = \frac{\mathrm{SSE}}{\mathrm{df}} = \frac{1.61486328}{9} \approx 0.179429253.$$

Therefore $\mathrm{Var}(\widehat{\beta}_2) = 0.179429253 \cdot 0.00000240297 = 0.000000431163112$. Computing our t-value,

$$t = \frac{\widehat{\beta}_2 - \beta_2}{\sqrt{\mathrm{Var}(\widehat{\beta}_2)}} = \frac{0.0009574426749999754 - 0}{\sqrt{0.000000431163112}} = \frac{0.0009574426749999754}{0.000656630118} \approx 1.45811569.$$

Then $p = \Pr(|T| > 1.45811569) = 0.178804$ (by Webassign Student's t-distribution with two-tailed, the t-value of 1.45811569 and 9 degrees of freedom). Since $p > \alpha$ there is **not** sufficient evidence that the quadratic term is necessary.