Advanced Data Analysis HW5

Ao Liu, al3472

1.

(a)

Answer:

```
data = read.csv("Shuttle.csv", header = TRUE)
glm(ThermalDistress~Temperature, data = data, family = binomial("logit"))
```

```
Call:  glm(formula = ThermalDistress ~ Temperature, family = binomial("logit"),
    data = data)

Coefficients:
(Intercept)  Temperature
    15.0429      -0.2322

Degrees of Freedom: 22 Total (i.e. Null);  21 Residual
Null Deviance:      28.27
Residual Deviance: 20.32  AIC: 24.32
```

(b)

Answer:

According to the result that we got in (a), our estimation of β_1, the effect of temperature on the probability of thermal distress is

$$-0.2322$$

This implies that when we increase the temperature by 1 degree, the odds of having Thermal Distress changes by a multiplicative factor of $e^{-0.2322}$

(c)

Answer:

```
confint(glm(ThermalDistress~Temperature, data = data, family = binomial("logit")))
```

```
                2.5\%       97.5\%
(Intercept)  3.3305848   34.34215133
Temperature -0.5154718   -0.06082076
```

According to the results in R, the 95% confidence interval for β_1 is

$$(-0.515718, -0.06082076)$$

so the the 95% confidence interval for e^{β_1} is

$$(0.597071743167396, 0.940991888047314)$$

This indicates that we are 95% confident that when we increase the temperature by 1 degree, the odds of having Thermal Distress changes by a multiplicative factor between 0.597071743167396 and 0.940991888047314.

(d)

Answer: