

Name: Rahul Gupta
ID: 100405295

Activity 4, 20 Marks
DANA 4820

Fall 2023, Due: Dec 1

Please include your RStudio output to support your answers

Question

the Credit data file at the text website, shows data for a sample of 100 adults randomly selected for an Italian study on the relation between x = annual income and y = whether you have a travel credit card (1 = yes, 0 = no). At each level of x (in thousands of euros), the table indicates the number of subjects in the sample and the number of those having at least one travel credit card.

- a) Report the prediction equation and interpret the sign of $\hat{\beta}$.

```
> model1 <- glm(cards/n ~ income, family=binomial, weights=n, data = Credit)
> summary(model1)
```

Call:

```
glm(formula = cards/n ~ income, family = binomial, data = Credit,
     weights = n)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3.51795	0.71034	-4.953	7.33e-07 ***
income	0.10541	0.02616	4.030	5.58e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 65.870 on 23 degrees of freedom
Residual deviance: 39.276 on 22 degrees of freedom
AIC: 58.833

Number of Fisher Scoring iterations: 5

->Logit[P(Y=1)] = -3.51 + 0.10*Income(x)

For Income:

The estimated odds for having a travel credit card for 1 thousand euros increase in annual income is increased by a factor of $\exp(\hat{\beta}(0.10541))$ equals 1.11. With positive sign the relationship between travel credit card and income will increase.

b) When $P(Y = 1) = 0.50$, show that the estimated logit value is 0. Based on this, for these data explain why the estimated probability of a travel credit card is 0.50 at income = 33.4 thousand euros.

-> For $P(Y = 1) = 0.50$,

-> $\text{Logit}(0.50) = \log(0.50/1-0.50)$

-> $\text{Logit}(0.50) = \log(0.50/0.50) = \log(1) = 0$

Hence, the estimated logit value at $P(Y = 1) = 0.50$ is 0.

-> when $\pi(x) = 0.50$

-> $\log(0.5/1-0.5) = \log(1) = 0 = \alpha + \beta x$

-> $x = -\alpha/\beta = -(-3.51/0.1054) = 33.4$ thousand euros.

c) Use the logistic output to interpret the effect of income on the odds of possessing a travel credit card, and conduct a significance test and a confidence interval for that effect.

Interpretation:

For Income:

The estimated odds for having a travel credit card for 1 thousand euros increase in annual income is increased by a factor of $\exp(\beta(0.10541))$ equals 1.11.

Significance test:

H_0 (Null hypothesis): $\beta_0 = 0$

H_a (Alternate hypothesis): $\beta_0 \neq 0$

```
> Anova(model1)
```

```
Analysis of Deviance Table (Type II tests)
```

```
Response: cards/n
```

```
      LR Chisq Df Pr(>Chisq)
income  26.594  1  2.51e-07 ***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With chisq statistic of 26.594 and Df=1, chisq p-value of $2.51e-07 < 0.05$, at 5% level of significance, we reject null hypothesis and estimate that income is significant in predicting travel credit card from the model.

Confidence Interval:

```

> confint(model1)
Waiting for profiling to be done...
              2.5 %      97.5 %
(Intercept) -5.06127577 -2.2513245
income       0.05958583  0.1625808
>

```

For the model above, we are 95% confident that the interval for income on predicting travel credit card is between (0.059,0.162).