SMM634 Answers Case Study - Model for Binary Data

Data on the 23 space shuttle flights that occurred before the Challenger mission in 1986 are given in the following table. For each of the 23 missions, data on the temperature, in ^oF, at the time of flight (Temp.), and whether at least one primary O-ring suffered thermal distress (TD) were recorded.

Temp	TD	Temp	TD	Temp	TD
66	0	57	1	70	0
70	1	63	1	81	0
69	0	70	1	76	0
68	0	78	0	79	0
67	0	67	0	75	1
72	0	53	1	76	0
73	0	67	0	58	1
70	0	75	0		

We fit a logistic regression model between TD and Temp.

```
td <- c(0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1)
temp <- c(66, 70, 69, 68, 67, 72, 73, 70, 57, 63, 70, 78, 67, 53, 67, 75, 70,
          81, 76, 79, 75, 76, 58)
shuttle <- data.frame(td, temp)
shuttle.glm <- glm(td ~ temp, family = binomial, data = shuttle)
summary(shuttle.glm)
```

Assuming, for the sake of argument, that the model does provide a good fit, then we see that

$$\widehat{\pi} = \frac{\exp(\widehat{\beta}_1 + \widehat{\beta}_2 x)}{1 + \exp(\widehat{\beta}_1 + \widehat{\beta}_2 x)} \tag{1}$$

where $\widehat{\beta}_1 = 15.043$, $\widehat{\beta}_2 = -0.232$, x is the temperature at the time of flight, and $\widehat{\pi}$ is the (corresponding) estimated probability of thermal distress of at least one primary O-ring.

Question 1: What is the predicted probability of thermal distress at 31^oF (supposedly the temperature at the time of the Challenger flight)?

Setting $x = 31$ in (1) yields

$$\widehat{\pi} = \frac{\exp(15.043 - 0.232 \times 31)}{1 + \exp(15.043 - 0.232 \times 31)} = 0.999$$

i.e. it is predicted that thermal distress is almost certain at that temperature. Note however, that we are predicting a long way outside the range of the observed data, always a dangerous thing to do!

Question 2: At which temperature is it estimated that there is a 50% chance that thermal distress occurs?

Substituting for $\widehat{\pi} = 0.5$ into (1), we find that the corresponding x satisfies

$$\widehat{\beta}_1 + \widehat{\beta}_2 x = 0$$

which implies that

$$x = -\frac{\widehat{\beta}_1}{\widehat{\beta}_2} = -\frac{15.043}{-0.232} = 64.794$$

i.e. about $65^o F$.

See below for a plot of the observed data and fitted model.

```
fittemp <- seq(20, 85, length = 100)
etashuttle <- coef(shuttle.glm)[1] + coef(shuttle.glm)[2] * fittemp
fitshuttle <- exp(etashuttle)/(1 + exp(etashuttle))
ylim <- c(0, 1)
plot(fittemp, fitshuttle, xlab = "Temperature (degrees F)", ylab =
    "Probability of Thermal Distress", type = "l",
     ylim = ylim, xlim = c(20, 85),lty = 1)
points(temp, td, cex=0.5)
```