1. Explainatory variable: quantitative \((x)\)  
Response variable: categorical with two categories \((y)\)

2. Examples:

3. The Challenger disaster.

4. Model: the probability of success \(p\) for a given \(x\) is given by

\[
\ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x
\]

5. How to choose \(b_0\) and \(b_1\) – not least squares. Maximum likelihood.

**Homework**

1. Read Chapter 27, pages 690–694.

2. To turn in on Friday, April 23, do the following problem. A dataset that you can find at the Mathematics 243 website called pima contains data on 768 women of Pima descent living in Arizona. (This is the same dataset used as the example in the book.) The variable test codes whether the person has signs of diabetes and the variable bmi contains the body mass index. Use logistic regression to predict the probability that a women has diabetes if she has a bmi of 20 and 40. (The data is available using get243() or at the course website. It is not in the course package.)
> orings[1:3,]
  temp damage
1 53 3
2 57 1
3 58 1
> plot(damage~temp,data=orings)
> y = cbind( orings$damage, 6 - orings$damage)  # a matrix with two columns, successes and failures
> morings=glm( y ~ temp, data=orings, family=binomial)  # y can also be a vector of ones and zeros
> morings
Call: glm(formula = y ~ temp, family = binomial, data = orings)

Coefficients:
(Intercept) temp
   6.7518  -0.1397

Degrees of Freedom: 22 Total (i.e. Null); 21 Residual
Null Deviance: 28.76
Residual Deviance: 19.09   AIC: 36.76
> x = data.frame(temp=31)
> predict(morings, x, type='response')
1 0.9184025
> summary(morings)
Call: glm(formula = y ~ temp, family = binomial, data = orings)

Deviance Residuals:
        Min       1Q   Median       3Q      Max
-0.9876  -0.7798  -0.4987  -0.2975   2.7483

Coefficients:  Estimate  Std. Error   z value Pr(>|z|)
(Intercept)   6.75183    2.97989     2.266   0.0234 *
 temp        -0.13971    0.04647    -3.007   0.0026 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1  1