Wednesday, Apr. 5

Class #34

The Maximum Likelihood Principle
(This is an advanced topic not covered on tests/quizzes)

Maximum likelihood is a way of estimating parameters in a model.

The ML principle states that the parameters chosen for a model should be those that yield the highest probability of observing precisely what was observed.

Consider the simple logistic regression models of Chapter 20.

We assume \( Y_i \) equals 0 or 1 for each \( i \).
\[
P(Y_i = 1) = \pi_i \quad \text{and} \quad P(Y_i = 0) = 1 - \pi_i.
\]

We may write this more succinctly as \( P(Y_i = y_i) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i} \) where \( y_i \) is a constant equal to either 0 or 1.

If \( y_i \) denotes the observed value of \( Y_i \), the function \( \pi_i^{y_i}(1 - \pi_i)^{1-y_i} \) is called the likelihood function for the \( i \)th observation. We multiply the likelihoods for each observation together to obtain the overall likelihood function:
\[
\prod_{i=1}^{n} \pi_i^{y_i}(1 - \pi_i)^{1-y_i}
\] (1)

The ML principle says we should find the values of \( \pi_1, \pi_2, \ldots \) that give the largest possible value of the likelihood function after we know what the \( y_i \) values are.
If \(\pi_1, \pi_2, \ldots\) maximize the likelihood, they also maximize the log of the likelihood. Thus, we often consider the log-likelihood instead of the likelihood because of mathematical convenience. Taking the logarithm of (1) gives the function

\[
\sum_{i=1}^{n} y_i \log(\pi_i) + (1 - y_i) \log(1 - \pi_i).
\]  

(2)

According to the logistic regression model,

\[
\text{logit}(\pi_i) = \beta_0 + \beta_1 x_i.
\]

Rearranging this using simple algebra gives

\[
\pi_i = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}.
\]  

(3)

Substituting (3) into (2) gives the following function of \(\beta_0\) and \(\beta_1\):

\[
L(\beta_0, \beta_1) = \sum_{i=1}^{n} y_i \log \left( \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \right) + (1 - y_i) \log \left( 1 - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \right)
\]

This can be simplified:

\[
L(\beta_0, \beta_1) = \sum_{i=1}^{n} y_i (\beta_0 + \beta_1 x_i) - \log (1 + \exp(\beta_0 + \beta_1 x_i))
\]

When you fit a logistic regression model in Minitab, it finds the values \(\hat{\beta}_0\) and \(\hat{\beta}_1\) that maximize \(L(\beta_0, \beta_1)\).