15.4 Refer to the previous exercise. One of the WAIS subtests, called Picture completion, asks questions about 20 pictures that have one vital detail missing. It IS considered a test of attention to fine detail. The observations for 20 subjects on (x, y), where x = picture completion score and y = symptoms of senility (1 = yes), are
(7,1), (5,1), (3,1), (8,1) (1,1), (2,1) (9,1), (3,1),(6,1),(4,1),(6,0),(9,0),(7,0),(7,0),(10,0),(12,0),(14,0),(8,0),(8,0),(11,0).

(a) Using software, estimate the logistic regression equation.

- **Logistic Regression**

```
. bysort senility: su picture

-> senility = no

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     picture |        10         9.2    2.529822          6         14

-> senility = yes

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     picture |        10         4.8     2.65832          1          9

. keep if !missing(picture, senility)
(0 observations deleted)
```

Table 1: Picture Completion and Senility

```
. logit senility picture

Iteration 0:   log likelihood = -13.862944
Iteration 1:   log likelihood = -8.0259967
Iteration 2:   log likelihood =  -8.024565
Iteration 3:   log likelihood = -8.0245642

Logistic regression                               Number of obs   =         20
                                                  LR chi2(1)      =      11.68
                                                  Prob > chi2     =     0.0006
Log likelihood = -8.0245642                       Pseudo R2       =     0.4212

------------------------------------------------------------------------------
    senility |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     picture |  -.7721148    .3578407    -2.16   0.031     -1.47347   -.0707599
       _cons |   5.426567    2.627976     2.06   0.039     .2758298     10.5773
------------------------------------------------------------------------------
```

Table 2: Logistic Regression for Picture Completion and Senility

x = picture completion score
y = symptoms of senility

$$logit\ [p(y=1)] = a + \beta x$$
$$logit\ [p(y=1)] = 5.4 - 0.77x$$

Since the estimate -0.77 of β is negative, the estimated probabilities of having senility increasing at lower levels of picture completion score.

➢ Graph of Regression

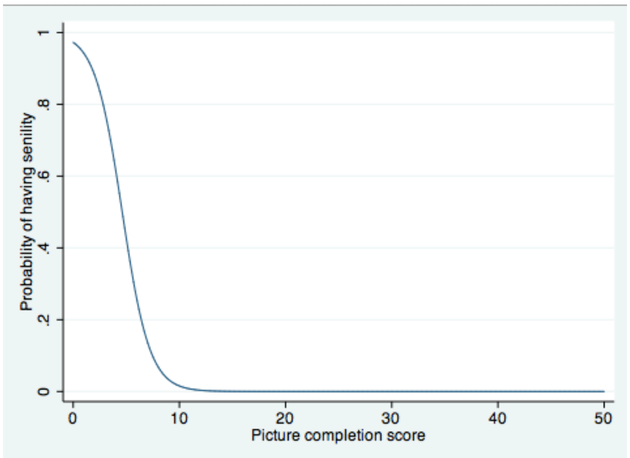

Figure 1：Graph of logistic regression

For the estimates in table 1, a person with picture completion x score has estimated probability of having senility equal to:

$$\hat{P}(y=1) = \frac{e^{5.4-0.77x}}{1 + e^{5.4-0.77x}}$$

➢ **x (6,14)**

For subjects with picture x = 6, the lowest picture completion score, the estimated probability equals:

$$P\ (y=1) = 0.69$$

For subjects with picture x = 14 > 6, the highest picture completion score, the estimated probability equals:

$$P\ (y=1) = 0.0046 < 0.69$$

As shown in figure 1, the higher score of picture completion, the lower probability of having senility.

(b) Estimate the probability that symptoms of senility are present when (i) x = 0, (ii) x = 20.

➢ **x = 0**

For subjects with picture completion x = 0, the estimated probability equals:

$$\hat{P}(y=1) = \frac{e^{5.4-0.77*0}}{1 + e^{5.4-0.77*0}}$$

$$P\ (y=1) = 0.97$$

As shown in Software:

```
. di _b[_cons] + 0* _b[picture]
5.4265673

. di exp( _b[_cons] + 0* _b[picture]) / (1+ _b[_cons] + exp( _b[_cons] + 0* _b[picture]))
.97251184
```

When picture completion score is 0, the probability of having senility is 0.97.

- **x = 20**

For subjects with picture completion x = 20, the estimated probability equals:

$$\hat{P}(y=1) = \frac{e^{5.4-0.77*20}}{1 + e^{5.4-0.77*20}}$$

$$P\ (y=1) = 0.000045$$

As shown in Software:

```
. di _b[_cons] + 20* _b[picture]
-10.015728

. di exp( _b[_cons] + 20* _b[picture]) / (1+ _b[_cons] + exp( _b[_cons] + 20* _b[picture]))
6.954e-06
```

When picture completion score is 20, the probability of having senility is 0.000045.

- **Logistic Regression Graph**

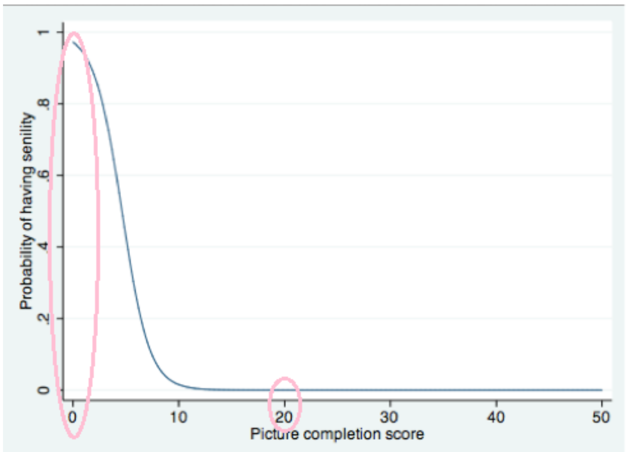

Figure 2：Logistic regression for x=0 and x=20.

As indicated before, since the estimate -0.77 of β is negative, the estimated probabilities of having senility increasing and the score of picture completion is negative connected:

$$x = 0 < x = 20$$
$$P = 0.97 > P = 0.000045$$

(c) Over what range of x-score is the estimated probability of senility greater than 0.5.

- **P = 0.5**

$$x = -a/\beta = 5.4/(-0.77) = 7.01$$
$$P(y = 1) = 0.50 \text{ when } x = 7.01$$

That is to say, the estimated probability equals 0.50 at x equal 7.01.

- **P(y = 1) > 0.50**

As indicated above, β = -0.77, there is negative relations between x and y.
Thus, the estimated probability of having senility is below 0.50 for Picture completion score above 7.01; and above 0.50 for Picture score blow this level.

$$P(y = 1) < 0.50 \text{ when } x > 7.01$$
$$P(y = 1) > 0.50 \text{ when } x < 7.01$$

- **Line Drawn Tangent to a Logistic Regression**

A straight line drawn tangent to the curve has slope βP (y=1)[1-P(y=1)]. Where P(y=1) is the probability at that point.
The slope is greatest when P(y=1) = 1/2, where β(1/2)(1/2) = β/4.

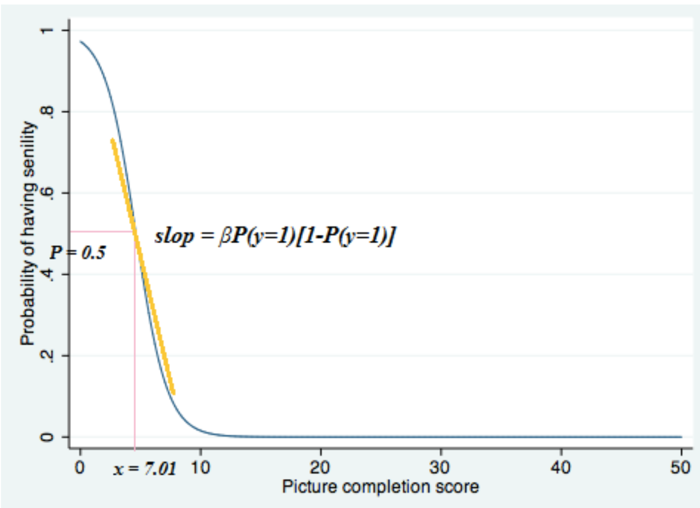

Figure 3：Graph of Logistic regression, slope = -0.77/4 = -0.1295.

(d) Estimate the effect of a one-unit increase in x on the odds that senility symptoms exist.

- **Using the Odds and Odds Ratio**

$$\frac{P(y = 1)}{1 - P(y = 1)} = e^{\alpha+\beta x} = e^{\alpha}(e^{\beta})^{x}.$$

This exponential relationship implies that every unit increase in x has a multiplicative effect of e^{β} on the odds:

$$e^{\beta} = e^{-0.77} = 0.46$$

- **Interpretation**

When annual picture completion score increases by one-unit, the estimated odds of having senility symptoms multiple 0.46.
For instance, when x = 6:

$$\textit{Estimated odds} = P(y=1) / (1-P(y=1)) = e^{5.4-0.77*6} = 0.46$$

Whereas when x = 7:

$$\textit{Estimated odds} = P(y=1) / (1-P(y=1)) = e^{5.4-0.77*7} = 0.99$$

$$0.46/0.99 = 0.46 = e^{-0.77}$$

The probability of x = 7 is 0.46 times of x =6. In other words, $e^{-0.77}$ =0.46 is an estimated odds ratio, equaling the estimated odds at x = n divided by the estimated odds at x = n+1.