

**NOTE: Give numerical answers with three decimals**

### Exercise 1

a. The estimated coefficient for agerec of the linear model is  $0.90653$

The standard error is  $0.04416$

The residual standard deviation ( $\sigma$ )  $6.789$

b. The estimated coefficient for agerec of the generalized linear model is  $0.90653$  ✓

The standard error is  $0.04416$  ✓

The residual standard deviation ( $\sigma$ ) can be calculated by  $\sqrt{46.097} = 6.789$  ✓

c. The link function used is  $\text{logit function: } \log\left(\frac{\pi_i}{1-\pi_i}\right) = X_i\beta$  ✓

d. The model used is  $\pi_i = \beta_0 + \beta_1 X_{i1}$  ✓

Interpretation of the regression coefficient for agerec:  $\text{An increase of 1 year in agerec will make } P(Y=1) \text{ increase } 0.0076.$  ✓

### Exercise 2

a. Total number of women with thrombosis for second generation pill is  $1367$  ✓  
 $5730208$

Total number of women with thrombosis for third generation pill is  $1514930$   $343$  ✓

Total number of "person-months" for second generation pill is  $30033222$  ✓

Total number of "person-months" for third generation pill is  $4579452$  ✓

b. expected number of thrombosis cases per person-month for second generation pill is  $4.55 \times 10^{-5}$  ✓

expected number of thrombosis cases per person-month for third generation pill is  $7.49 \times 10^{-5}$  ✓ and third generation pill.

c. The model is  $\log(\mu) = \beta_0 + \beta_1 \text{type pill} + \log(\text{Users})$  ✓

The rate of thromboses cases per person per month for the second generation pill is  $e^{-9.997} = 4.55 \times 10^{-5}$  ✓

The rate of thromboses cases per person per month for the third generation pill is  $e^{-9.997+0.498} = 7.49 \times 10^{-5}$  ✓

The rate ratio for a third generation pill versus a second generation pill is:  $e^{0.498} = 1.646$  ✓

d. The rate ratio for a third generation pill versus a second generation pill is .....  $e^{0.175} = 1.191$  ✓

e. The results of the goodness of fit test were

P-value = 0.09 > 0.05, not reject  $H_0$ , model fits well. ✓

f. The estimated overdispersion parameter is ..... 1.134 ✓

Conclusion ..... there is no substantial overdispersion ✓