Practice 4 Solutions

1. (a) i. Test not valid, as one model is not nested within the other.

   ii. Test is OK ($M(3)$ nested in $M(1)$). Test of dose as a factor with 6 levels versus dose as a variable.

   iii. Test is OK ($M(7)$ nested in $M(1)$). Test of dose (or dosage) as a factor with 6 levels versus dosage as a variable.

   (b) All but models 2 and 3 provide an adequate fit to the data. (Residual deviance $< \chi^2_{0.95}(df)$ in these cases where $\chi^2_{0.95}(df)$ is the 0.95 quantile of a $\chi^2(df)$ distribution).
   With dose, model 5 is significantly better than model 4 ($\Delta D = 14.84 - 8.18 = 6.66 > 3.84 = \chi^2_{0.95}(1)$), and not significantly worse than model 1 ($\Delta D = 8.18 - 5.01 = 3.17 < 5.99 = \chi^2_{0.95}(2)$).
   With dosage, model 7 is not significantly worse than models 6, 8 and 9 ($\Delta D = 6.76-4.99 = 1.77$, $6.76-5.85 = 0.91$ and $6.76-5.01 = 1.75$, all $< 3.84 = \chi^2_{0.95}(1)$). Model 7 is not significantly worse than model 1 either ($\Delta D = 1.75$ on 4 df).
   So choice is between model 5 (bud.5) and model 7 (bud.7). Direct comparison is not possible, but model 7 is simpler and has a smaller residual deviance, hence prefer model 7.

   (c) i. Yes. $z = -3.09$ and $|z| \geq 1.96$.

   ii. For a given dosage, odds of males being killed is greater than those for females by a factor $e^{1.1} = 3.0$.
   For either sex, odds of death increase by a factor of $e^{1.064} = 2.9$ for each increase in dosage of 1 unit (i.e. as dose doubles).

   iii. For males, $\widehat{\text{logit}}(\theta) = -3.473+1.1007+1.0642\times\texttt{dosage} = 0$ when `dosage`=2.23. So LD(50) =2.23 for males.
   For females, $\widehat{\text{logit}}(\theta) = -3.473 + 1.0642 \times \texttt{dosage} = 0$ when `dosage`=3.26. So LD(50) =3.26 for females.

   iv. For females, $\widehat{\text{logit}}(\theta) = -3.473 + 1.064 \times 4 = 0.784$ when dosage = 4.
   So the estimated probability is $\hat{\theta} = \dfrac{e^{0.784}}{1 + e^{0.784}} = 0.6865$.

   (d) Parameter estimates and standard errors would stay the same. Changes in residual deviances and changes in df for nested models would stay the same. But residual deviances and their df would be different.