1. Je 1 Punkt.

1) c
2) a
3) d
4) a
5) d
6) c
7) d

2. **a)** `x1,x2` not obviously correlated with `y`. Interaction between `x3` and `x4`: Visible change in the effect of `x3` on `y` depedent on the state of `x4`. **(2 Points)**, each wrong answer: **-0.5 Points**

b) `y` no trsf., `x1`: trsf.: log + shift, `x2`: trsf.: squareroot, `x3`: no trsf.; **(2 Points)**, each wrong answer: **-0.5 Points**

c) Excluding more than one predictor at once from the model is not wise. Better perform a stepwise variable selection. **(1 Point)**

d) The multiple R-squared should be larger for this model since more predictors are incorporated. **(1 Point)**

e) n = 30, p = 10, q = 3, $F_{part} = \frac{1}{p-q} \frac{F_p \cdot p \cdot \hat{\sigma}^2_{\epsilon,p} - F_q \cdot q \cdot \hat{\sigma}^2_{\epsilon,q}}{\hat{\sigma}^2_{\epsilon,p}} = 0.6299385$.**(2 Points)**. The Nullhypothesis (all variables of the full model not contained in the nested/small model do not have a significant influense on the response) cannot be rejected. (**1 Point**), **(total: 3 Points)**

3. **a)** i. False. Leaving out observations from a regression analysis always needs close examination. Leverage points which are in addition outliers have to be handeled with care and can be excluded from the analysis if their corresponding Cooks distance is at least 0.5. **(1 Point)**

ii. True. Confidence intervals can be directly converted from one metric to the other. Reason: interval estimates are percentiles of a distribution. These are not affected by monotone transformations. **(1 Point)**

iii. False. Individual parameter test for dummy variable coefficients are not meaningful. Either exclude the whole categorial variable or not. **(1 Point)**

iv. True. Due to the highly significant global F-test we suspect both predictors to be correlated. In this case, variable selection should be applied to choose the predictor which best represents the fluctuations in the response. **(1 Point)**

b) 21 parameters. Follows from $(j-1) * p^*$ with $j-1 = 3$ and $p^* = 7$. See scriptum p. 102 for comparison. **(1 Point)**. We need at least 105 observations for fitting. **(1 Point)**, **(total: 2 Points)**

c) This is a partial F-test. The corresponding test statistic follows an F-distribution. **(1 Point)**

d) There are $p = 8$ parameters in the model. An intercept, all single parameters, all 2-way interactions and one 3-way interaction. **(1 Point)**

4. **a)** (1 Punkt)

$$Y \sim Ber(p)(0.5 Punkte), \quad \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 (0.5 Punkte)$$

b) (1 Punkt) Das relevante Quantil für einen zweiseitigen z-Test mit Niveau 5% ist 1.96. Daher ist der Koeffizient von X_3 signifikant von 0 verschieden. Doch das ist nicht so relevant da die z-Werte von X_2 und X_3 sich nur marginal unterscheiden.

c) (1 Punkt) $n = df_{Resid} + p = 20$

d) (1 Punkt) Die gesuchte Odds Ratio beträgt $\exp(\hat{\beta}_2) = 6.463$.

e) (2 Punkte) $\text{logit}(\hat{p}) = -0.7129 + 0.0931 \cdot 3 + 1.8661 \cdot 2 - 2.8236 \cdot 1 = 0.4750956$
$\hat{p} = \frac{\exp(0.4750956)}{1+\exp(0.4750956)} = 0.6165891$ (1P)
Daher prognostiziert man $\hat{y} = 1$ (1P).

f) (2 Punkt) Damit die Wahrscheinlichkeit, dass $Y = 1$, genau 50% (also $p = 0.5$) beträgt, müssen die Log-Odds $\eta = \log(p/(1-p) = \log(1) = 0$ sein. Bei $X_1 = 3$ und $X_2 = 5$ erhalten wir aus der Modellgleichung

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

die Gleichung

$$0 = -0.7129 + 0.0931 \cdot 3 + 1.8661 \cdot 5 - 2.8236 \cdot X_3 \ ,$$

(1P) was sich auflösen lässt zu $X_3 = 3.150967$. (1P)

g) (1 Punkt) Nein das grössere Model ist nicht signifikant besser denn $D^{(S)} - D^{(B)} = 13.979 - 13.979 = 0$. Das heisst die Nullhypothese $H_0: \ \beta_1 = 0$ wird nicht verworfen.

5. **a)** (1 Point)

$$Y \sim Pois(\lambda), \quad \log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

b) (1 Point) The model fits very well since the residual deviance is nearly equal to the degrees of freedom ($16.068 \approx 16$).

c) (1 Point) $\lambda = e^{1.7719+1.6350 \cdot 3+1.0897 \cdot 3-4.1656 \cdot 1} = 323.8887$

d) (1 Point) By looking at the plot we see that the variance is clearly bigger then the mean i.e. we have overdispersion.

e) (1 Point) $\phi = \frac{\text{sum of squared Pearson residuals}}{\text{residual degrees of freedom}} = \frac{211.9841}{17} = 12.46965$

f) (2 Points) It leads to smaller confidence intervals (1P) and less significant test results for the individual hypothesis tests (1P).