(1, 2pts) The four varibles are mutually independent.

(2, 2pts) The df is $2^4 - 1 - 4 \times 1 = 11$. We reject the hypothesis of mutual independence because 973 is very large for 11.

(3, 2pts) The Pearson Chi-square statistic is the sum of the squared Pearson residuals and it is expected to distribute as a Chi-square with 11 degrees of freedom if the main-effect-only model fits. When the model fits, we would therefore expect Pearson residuals to be about one in absolute value. Some will be larger and other smaller, but 14.6 is clearly very large in this context.

(4, 2pts) Among the 3 2-factor interactions, only comps:smoking is not significant. Removing this interaction yields a model corresponding to conditional independence. We may conclude that having gynaecological problems during pregnancy and being a smoker are conditionally independent given birth weight status.

(5, 2pts) For mothers with heart problems, children with birth weights less than 1250g ($62 = 10 + 25 + 12 + 15$) are roughly as many as those with birth weights at least 1250g ($53 = 7 + 5 + 22 + 19$). However, among mothers without heart problems, normal birth weight is much more common (429 vs. 117). Thus, the result does not generalize to the full dataset.

(6, 2pts) A binomial GLM with the birth weight as the response and the other 3 variables as the explanatory variables.

(7, 2pts) The residuals is obtained from a model corresponding to independence. If a Poisson GLM with the counts in the 2-way table as the response is adopted, the corresponding model is the main-effect-only model:

$$count \sim opinion + food.$$

If a multinomial GLM with food (or opinion) as the response and the other variable as the predictor is adopted, the corresponding model is the intercept-only model:

$$food \sim 1 \qquad \text{or} \qquad opinion \sim 1.$$

(8, 2pts) Bacon and eggs (or perhaps cereal), because it has large absolute value on X-axis (and/or Y-axis).

(9, 2pts) Bacon and eggs (or perhaps stewed fruit), because it and quickeasy would associate with a large negative residual.

(10, 2pts) The food proportion in the opinion “tasteless” is similar to the marginal proportion of food. We can say that the opinion “tasteless” is approximately proportionately represented across all foods and it is not associated with any particular food.

(11, 2pts) In a proportional odds model, we assume there is a latent (unobserved) continuous variable underlying the observed ordinal categories. The intercept values represent the cutpoints (thresholds) on this latent continuous scale that separate the latent variable into intervals corresponding to each observed ordered category.

(12, 2pts) The two relevant cutpoints, 5.979 and 6.244, are only 0.265 units apart. This difference is small compared to their associated standard errors and is relatively smaller than the differences between other adjacent cutpoints when standardized by their standard errors. This suggests that the model cannot clearly distinguish between the "2–3 years" and "more than 3 years" categories on the latent scale, supporting the idea that these two categories can be reasonably combined.

(13, 2pts) To answer this question, we can reparametrize the model by taking Musical revues as the reference level, instead of Musicals. In this parameterization, we compare both Musicals and Plays relative to Musical revues. From the original output, the estimated coefficient for Musical revues is 0.340005, and for Plays it is -1.595308, with Musicals as the reference. Thus, if we switch the reference category to Musical revues, the relative coefficient for Musicals would be: $0 - 0.340005 = -0.340005$, and for Plays: $-1.595308 - 0.340005 = -1.935313$. This shows that the difference between Plays and Musical revues is much larger than that between Musicals and Musical revues. Moreover, the smaller difference between Musicals and Musical revues is not statistically significant relative to the standard error (SE $= 0.853489$), while the difference between Plays and Musical revues is clearly substantial. Therefore, musical revues resemble musicals more than plays in terms of the model-estimated probabilities across the five outcome categories.

(14, 2pts) The correct answer is No. In this case, the coefficient for RevivalYes is -0.718951. In the proportional odds model, a negative coefficient means that the $\eta_{\mathbf{x},j} = \beta_{0,j} - \mathbf{x}^\top \boldsymbol{\beta}$ becomes larger when Revival changes from No to Yes, which increases the cumulative probability $P(Y_{\mathbf{x}} \le j)$. In other words, the probability mass is shifted toward lower response categories. This implies that revival shows are more likely to fall into lower-duration categories (e.g., "less than 6 months") than non-revival shows. Therefore, the model suggests that revival shows are *more likely*, not less likely, to fall into the lowest category.

(15, 2pts) In a multinomial logit model, we need to estimate one set of parameters (including an intercept and coefficients for all explanatory terms) for each non-reference category. There are 4 explanatory terms and 1 intercept, giving 5 parameters per category. Since there are 5 response categories, and one must be used as the reference category, we need to estimate parameters for $5 - 1 = 4$ categories. Thus, the total number of parameters is $(4 + 1) \times (5 - 1) = 20$.

(16, 2pts) Because this is a proportional odds model, the link function for the cumulative probability is the logit function. To compute the probability that the show lasts less than six months (i.e., falls into the lowest category), we use the cumulative logit corresponding to that cutpoint:

$$\eta = 3.618 + 0.718951 - 92 \times 0.049819 = -0.246397.$$

Therefore, the predicted cumulative probability is:

$$\text{logistic}(-0.246397) = \frac{e^{-0.246397}}{1 + e^{-0.246397}} \approx 0.4387105.$$

(17, 4pts) We may recognize that the data form a 3-way contingency table with the following variables: (i) A: Rocking group (Yes = experimental, No = control); (ii) B: Crying outcome (Yes or No); (iii) C: Day (1 to 18). Each cell in this $2 \times 2 \times 18$ table corresponds to the count of babies for a given combination of treatment, crying status, and day. The following two approaches are both acceptable.

Approach 1: Binomial GLM. We model the number y of babies not crying (successes) out of the total number n of babies in each group defined by the combination of the predictors A and C using a binomial distribution:

$$y_{A,C} \sim \text{binomial}(n_{A,C}, p_{A,C}),$$

for $2 \times 18 = 36$ combinations of (A, C). Here, A is a binary indicator variable and C is treated as a qualitative block factor with 18 levels (thus requiring 17 dummy variables). The effects of C account for nuisance variation (e.g., temperature or time of day), and the primary parameter of interest is the effect of A, which reflects the impact of rocking on crying. We can express $p_{A,C}$ as a function of a linear combination of the effects of A and C using a logit link, and fit the data using a binomial GLM.

Approach 2: Poisson GLM. Alternatively, we model the count $y_{A,B,C}$ in each cell of the 3-way table using a Poisson distribution:

$$y_{A,B,C} \sim \text{Poisson}(\mu_{A,B,C}),$$

for $2 \times 2 \times 18 = 72$ combinations of the predictors A, B, and C. Here, A and B are binary indicator variables, and C is treated as a qualitative block factor with 18 levels (thus requiring 17 dummy variables). The effects of C account for nuisance variation (e.g., temperature or time of day), while the interaction between A and B is of primary interest (rather than the main effects of A and B), as it reflects whether the crying outcome depends on rocking, adjusted for day C. We can express $\mu_{A,B,C}$ as a function of a linear combination of the effects of A, B, and C using a log link and fit the data using a Poisson GLM.

(18, 2pts) Since C is treated as a block factor, we typically assume the effect of A on crying is constant across levels of the block factor.

Approach 1: Binomial GLM. The null model includes only the block effect (day), with no treatment effect:

$$y \sim 1 + C$$

The alternative model adds the treatment effect (A):

$$y \sim 1 + A + C$$

Approach 2: Poisson GLM. The null model assumes that A and B are conditionally independent given C:

$$y \sim 1 + A + B + C + A : C + B : C$$

The alternative model is the uniform association model, which adds the $A : B$ interaction, allowing the association between rocking and crying to differ from independence, while adjusting for the block factor C:

$$y \sim 1 + A + B + C + A : B + A : C + B : C$$

(19, 2pts) We should take the difference in the deviances of these two models and compare it to Chi-square with *one* degree of freedom. If the difference is sufficiently large we may conclude that there is an experimental effect. The size and direction of the effect may be expressed in terms of odds using the estimated parameter representing the experimental indicator. Note that the alternative model is not a saturate model so that we cannot use the goodness-of-fit test based on the residual deviance of the null model.