Logistic Regression with SAS - Test 2, max 15 points
Subject number: 223481-0131 – Full time Master's studies, Summer Semester 2018/2019

1. The table below shows the counts observed in a study the aim of which was to assess the relationship between income and job satisfaction. Job satisfaction is to be used as response variable in cumulative logit model. Assess if the proportional odds assumption is met. Provide the formulas and calculate the relevant measures to prove your answer. Interpret the result. [5p]

Income in USD	Very dissatisfied or little satisfied	Moderately satisfied	Very satisfied
15K or less	14	35	7
More than 15K	4	28	16

Source: Agresti, 2002, p. 288.

2. A study was conducted in Czech Republic (CZ), France (FR), Great Britain (GB) and Poland (PL) on how happy the citizens are? Happiness is measured on a scale with three categories: unhappy (1 – reference category in the model), moderately happy (2), very happy (3). A multinomial logistic model was estimated with happiness as response and country as explanatory variable. Is there a statistically significant difference in the chance of being very happy as compared to unhappy between Great Britain and Poland? Provide necessary formulas and calculate the relevant measures to prove the verification of this hypothesis. [5p]

Analysis of Maximum Likelihood Estimates

Parameter		Happy?	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept		Moderately	1	3.2572	0.1427	521.0828	<.0001	25.978
Intercept		Very	1	2.8559	0.1440	393.3899	<.0001	17.391
Country	FR	Moderately	1	0.1836	0.2294	0.6409	0.4234	1.202
Country	FR	Very	1	0.6223	0.2301	7.3137	0.0068	1.863
Country	GB	Moderately	1	-0.3829	0.2178	3.0905	0.0787	0.682
Country	GB	Very	1	0.5934	0.2172	7.4623	0.0063	1.810
Country	PL	Moderately	1	-0.0615	0.2399	0.0657	0.7977	0.940
Country	PL	Very	1	0.6588	0.2398	7.5469	0.0060	1.932

Logistic Regression with SAS - Test 1, max 15 points
Subject number: 223481-1380 – Full time Master's studies, Summer Semester 2018/2019

1. Model (1) shows the estimates of the binary logistic regression model assessing the effect of a categorical variable X on the response Y using reference parametrization (see Table 1). [10 p]

$$\text{logit}[\pi(x_i)] = -2 + 0.8x_{1i} + 2x_{2i} + 1.6x_{3i},\ i=1,2,\ldots,n. \quad (1)$$

Table 1. Class level information				
Variable X	A	1	0	0
	B	0	1	0
	C	0	0	1
	D	0	0	0

1.1. Calculate the estimate of odds ratio for A vs. D providing intermediate formulas.
1.2. Interpret the estimate of odds ratio from (1.1).

2. Assess the statistical significance of the two variables included in the model as shown in Table 2. Formulate the hypotheses and make the decision using provided results. [5 p]

Table 2. Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
acc_status	3	107.2777	<.0001
age	1	5.2307	0.0222