Name:

USC ID:

Notes:

- Write your name and ID number in the solution you submit.
- No books, cell phones or other notes are permitted. Only one letter size cheat sheet (back and front) and a calculator are allowed.
- Problems are not sorted in terms of difficulty. Please avoid guess work and long and irrelevant answers.
- Show all your work and your final answer. Simplify your answer as much as you can.
- Open your exam only when you are instructed to do so.
- The exam has 5 questions, 9 pages, and 13 points extra credit.

Problem	Score	Earned
1	22	
2	25	
3	22	
4	22	
5	22	
Total	113	

1. Assume that we built a linear regression model with $n = 22$ observations and $p = 5$ predictors. Determine the minimum value of R^2 for which at least one of the predictors is statistically significant when $\alpha = 0.01$.

2. Choose either T (True) or F (False) (no need to explain why):

   (a) When the assumption of conditional independence of features holds, the Naïve Bayes' classifier provides the best accuracy among all possible classifiers. T F

   (b) The F1 score is not an appropriate measure for evaluating binary classifiers when data are not imbalanced. T F

   (c) Leave-One-Out Cross Validation has less bias in estimating the error of a classifier for a large data set than 5 fold cross validation. T F

   (d) When classifying imbalanced data into two classes, we can decrease the threshold on class conditional probability $\Pr(Y = k|X_1 = x_1, \ldots, X_p = x_p]$ to increase the true positive rate at the expense of increasing the false negative rate. T F

   (e) Logistic regression assumes that the conditional odds of the outcome Y given the features, $\mathbb{O}[Y = k|X_1 = x_1, \ldots, X_p = x_p]$, is a logistic function of the features. T F

3. Assume that in a binary classification problem with one feature X, the distribution of X in class $k = 1$ is

$$f_1(x) = \frac{x}{\sigma_1^2} \exp\left(\frac{-x^2}{2\sigma_1^2}\right), x \geq 0$$

and the distribution of X in class $k = 2$ is

$$f_2(x) = \frac{1}{x\sqrt{2\pi}\sigma_2} \exp\left(\frac{-(\ln x - \mu_2)^2}{2\sigma_2^2}\right), x \geq 0$$

(a) Are there any conditions under which the discriminant function is a linear function of x?

(b) If $\sigma_1 = \sigma_2 = 1$, $\mu_2 = 10$, and $\pi_1 = \pi_2 = 0.5$, in what class will $x = 10$ be classified?

4. Consider multinomial regression for multiclass classification with three features $\mathbf{X} = (X_1, X_2, X_3)$, formulated by

$$p_k(\mathbf{X}) = \frac{e^{\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\beta_{3k}X_3}}{1 + e^{\beta_{01}+\beta_{11}X_1+\beta_{21}X_2+\beta_{31}X_3} + e^{\beta_{02}+\beta_{12}X_1+\beta_{22}X_2+\beta_{32}X_3}}, \ k \in \{1, 2\}$$

where the classes are determined by $k \in \{1, 2, 3\}$

Assume that using a data set of 210 observations from three classes, we obtained the following results:

Coefficient	Value
β_{01}	1
β_{11}	-2
β_{21}	-1
β_{31}	1
β_{02}	0
β_{12}	0
β_{22}	1
β_{32}	1

Assume that the coefficients are all statistically significant.

(a) In what class will the classifier classify $\mathbf{X}^* = (1, 0, -1)$?

(b) Explain why despite having three classes, we formulated multinomial regression using ONLY TWO sets of parameters, $(\beta_{01}, \beta_{11}, \beta_{21}, \beta_{31})$ and $(\beta_{02}, \beta_{12}, \beta_{22}, \beta_{32})$, specific to classes 1 and 2, respectively?

5. In a weird simulated world, the net worth n_i of everyone who has a spouse is *uniformly* distributed between $10n_m^3$ and $0.2n_s^2$, where n_m is the net worth of their mother and n_s is the net worth of their spouse.

   (a) What is the estimate with minimum mean squared error of the net worth of Fej Zebos, whose mother's net worth is 10 Dollars and whose spouse's net worth is 500 Dollars?

   (b) What type of supervised learning problem are you solving in this question? Explain.

Scratch paper

Name:

USC ID:

Scratch paper

Name:

USC ID:

F - Distribution ($\alpha = 0.01$ in the Right Tail)

Numerator Degrees of Freedom (df_1) / Denominator Degrees of Freedom (df_2)

df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971
19	8.1849	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981
22	7.9454	5.7190	4.8166	4.3134	3.9880	3.7583	3.5867	3.4530	3.3458
23	7.8811	5.6637	4.7649	4.2636	3.9392	3.7102	3.5390	3.4057	3.2986
24	7.8229	5.6136	4.7181	4.2184	3.8951	3.6667	3.4959	3.3629	3.2560
25	7.7698	5.5680	4.6755	4.1774	3.8550	3.6272	3.4568	3.3239	3.2172
26	7.7213	5.5263	4.6366	4.1400	3.8183	3.5911	3.4210	3.2884	3.1818
27	7.6767	5.4881	4.6009	4.1056	3.7848	3.5580	3.3882	3.2558	3.1494
28	7.6356	5.4529	4.5681	4.0740	3.7539	3.5276	3.3581	3.2259	3.1195
29	7.5977	5.4204	4.5378	4.0449	3.7254	3.4995	3.3303	3.1982	3.0920
30	7.5625	5.3903	4.5097	4.0179	3.6990	3.4735	3.3045	3.1726	3.0665
40	7.3141	5.1785	4.3126	3.8283	3.5138	3.2910	3.1238	2.9930	2.8876
60	7.0771	4.9774	4.1259	3.6490	3.3389	3.1187	2.9530	2.8233	2.7185
120	6.8509	4.7865	3.9491	3.4795	3.1735	2.9559	2.7918	2.6629	2.5586
∞	6.6349	4.6052	3.7816	3.3192	3.0173	2.8020	2.6393	2.5113	2.4073