Exam Practice (MTH499/522)

Your name in **BLOCK** letters: _____

You have 75 minutes for this exam. There are **?** pages with **?** questions for a total of 80 points.

Instructions:

1. **UMassD HAS ZERO TOLERANCE TO ANY VIOLATIONS OF THE ACADEMIC HONOR CODE.**
2. This is a CLOSED-BOOK and CLOSED-NOTES exam.
3. Show **ALL** your work. An answer (even if correct) without an explanation for how you obtained it will receive **LITTLE** credit.
4. The use of iPhone or computer devices is **STRICTLY** prohibited during the exam.

---

1. (20 points). Short answering questions.
   (a) (2 points) Please describe the simple linear model. What are the assumptions?

   (b) (2 point) What are the properties of a linear regression estimate?

   (c) (2 points) What is a convolutional neural network?

   (d) (2 point) Please explain type I error and type II error in hypothesis testing.

   (e) (2 point) What are the leverage and influence of a data point in regression diagnosis?

(f) (2 points) What is R^2 in linear regression? Please name one application.

(g) (2 point) Please explain how the logistic regression model can be understood via the generalized linear model.

(h) (2 points) What is a statistical model?

(i) (2 points) Please name two performance measures for machine learning algorithms.

(j) (2 points) Please explain the regression phenomenon.

2. (10 points). Please describe deep learning. Explain challenges of deep neural network and why it may do well in certain applications.

3. (10 points) Please describe tasks in regression diagnosis in detail.

4. (10 points). Data scientist Alex has a coin which is suspected to be biased towards 'T'. To see if this is truly the case, he flipped this coin for 100 times. The outcome were observed as follows.

   H, T, T, H, ..., T, T, H

   Based on the above observation, we can compute the proportion of 'T' to be 0.48. Answer the following questions. Is it true that die of interest is biased towards 'T'? Perform hypothesis testing at 1% significance level.

(a) (2 point) Formulate the null and alternative hypothesis.

(b) (2 points) State the test statistic and its approximate distribution.

(c) (2 points) Determine the rejection region for $\alpha = .01$ and state the decision rule.

(d) (1 point) Calculate the observed value of the test statistic from the sample.

(e) (1 point) State whether H_0 is rejected and tell why.

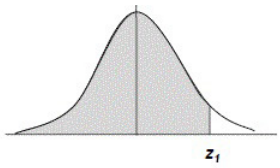
(f) (2 points) What are the interpretation of Type I error and the Type II error *in the context of the problem*? Please describe in plain English.

5. (10 points). The following is a regression analysis output. Please answer questions.

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-12.595  -2.730  -0.518   1.777  21.199

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.646e+00  5.103e-01   7.144 3.28e-12 ***
x1          -1.080e-01  3.286e-02  -3.287 0.001087 **
x2           4.642e-02  1.373e-02   3.382 0.000778 ***
x3           2.056e-02  6.150e-02   0.334 0.738288
```



$$p(z \le z_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_1} e^{-\frac{1}{2}z^2} dz$$

z_1	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

```
x4           2.687e+00  8.616e-01   3.118 0.001925 **
x5          -1.777e+01  3.820e+00  -4.651 4.25e-06 ***
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1

Residual standard error: 3.745 on 500 degrees of freedom
Multiple R-squared:  0.8406,Adjusted R-squared:  0.8338
F-statistic: 108.1 on 5 and 500 DF,  p-value: < 2.2e-16
```

a (2 points). Please write the fitted linear model.

b (2 points). Which variable(s) are statistically significant? Assume the significance level is $\alpha = 0.01$.

c (2 points). Is the linear model used appropriate? Why?

d (2 points). Suppose one drops variable 'x3' and re-do the regression analysis, will the resulting R^2 increase? Why?

e (2 points). Please explain individual items shown in the line for variable 'x2'.

6. (10 points). There is a three-layer neural network with 8, 3, and 5 nodes in its input, hidden and output layer, respectively.

a (2 points). Please sketch the neural network.

b (2 points). How many parameters are there for such a neural network?

a (2 points). Please describe two potential activation functions that you could use.

c (2 points). What would be the dimension and the number of categories of data suitable for such a neural network?

d (2 points). What algorithm will you use to fit the neural network?

7. (10 points). Assume a linear model specified as

$$Y = X\beta + \epsilon, \quad \mathbb{E}\epsilon = 0,\ Var(\epsilon) = \sigma^2,$$

and the data $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ are expressed in matrix notation as $X_{n\times p}$ and $Y_{n\times 1}$. Suppose we wish to find an estimator, $\hat{\beta}$, for β such that it minimizes the following

$$(Y - X\beta)^T (Y - X\beta).$$

a (5 points). Please derive the expression for $\hat{\beta}$ in terms of X and Y.

b (2 points). Is $\hat{\beta}$ an unbiased estimator for β? Why?

b (3 points). What is the distribution of $\hat{\beta}$ if ϵ follows a normal distribution with the mean and variance stated above?