Practice Final Examination

Final exam: Monday, June 6th, 3:30pm-6:30pm
Location: CEMEX Auditorium

This is an open note, closed calculator/computer exam. The last page of the exam is a Standard Normal Table, in case you need it.

You have 3 hours (180 minutes) to take the exam. The exam is 180 points, meant to roughly correspond to one point per minute of the exam. You may want to use the point allocation for each problem as an indicator for pacing yourself on the exam.

In the event of an incorrect answer, any explanation you provide of how you obtained your answer can potentially allow us to give you partial credit for a problem. For example, describe the distributions and parameter values you used, where appropriate. It is fine for your answers to include summations, products, factorials, combinations, and exponentials (but not integrals), unless the question asks for a numeric quantity or closed form. Where numeric answers are required, the use of fractions is fine.

Problem	Score
1 (15 pts)	
2 (20 pts)	
3 (20 pts)	
4 (20 pts)	
5 (25 pts)	
6 (25 pts)	
7 (20 pts)	
8 (15 pts)	
9 (20 pts)	
Total (180 pts)	

THE STANFORD UNIVERSITY HONOR CODE

A. The Honor Code is an undertaking of the students, individually and collectively:

(1) that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;

(2) that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

B. The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid as far as practicable, academic procedures that create temptations to violate the Honor Code.

C. While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to create optimal conditions for honorable academic work.

I acknowledge and accept the letter and spirit of the honor code:

Signature: _____

NAME (print): _____

1. (15 points) Say we have a standard 52 card deck of cards, where the cards have the usual ordering of ranks: 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King, and Ace. Assume each card in the deck is equally likely to be drawn. Two cards are then drawn sequentially (without replacement) from the deck.

   What is the probability that the second card drawn has a rank **greater than** the rank of the first card drawn? Show the derivation you used to compute the probability and also provide your final answer in closed form (if a closed form exists).

2. (20 points) You decide to buy one share of stock in "Happy Computing Inc." (abbreviated: HCI) for $10. Each day, the price of HCI either increases $1 (with probability p) or decreases $1 (with probability $1 - p$). You decide that you will only sell your stock at some point if <u>either</u> it gains $2 (i.e., reaches a price of $12) or loses $2 (i.e., reaches a price of $8).

    a. (7 points) What is the probability that you will sell your share of HCI (for either a gain or a loss of $2) *exactly* 4 days after you buy it?

    b. (13 points) Assuming you just bought a share of HCI, so you don't know how long it will be until you sell it, what is the probability (in terms of p) that when you eventually sell it, you will sell it for a gain (i.e., sell it for $12)? Give a closed form answer.

3. (20 points) Say you are monitoring energy usage for two different types of computers (type W and type X) in an office. There are 10 machines of type W and 10 machines of type X in the office, and the energy usage of all 20 machines are independent.

Let W_i = the number of kilowatt-hours used by machine i of type W in a day.
Let X_j = the number of kilowatt-hours used by machine j of type X in a day.

Say that $W_i \sim \text{Poi}(4)$ for $1 \le i \le 10$ and $X_j \sim N(5, 3)$ for $1 \le j \le 10$.

Now say that you have a separate machine that (probabilistically) monitors the energy usage of the 10 type W and 10 type X machines (call these 20 machines the machine pool). Each day, a new "watch list" is generated, where each machine in the machine pool independently has a 0.2 probability of being placed on the "watch list". Let Y = the total energy usage (in kilowatt-hours) of all the machines on the "watch list".

a. (4 points) What is $E[Y]$?

b. (8 points) Say that three type W (and no type X) machines are on the "watch list" on a particular day. What is $P(Y \ge 20)$ that day?

c. (8 points) Now say that three type X (and no type W) machines are on the "watch list" on a particular day. What is $P(Y \ge 20)$ that day? Give a numeric answer, if possible.

4. (20 points) We have a random number generator that generates random <u>integers</u> independently uniformly distributed from 1 to 5, inclusive. This random number generator is used to generate a sequence on n independent random integers.

Let X = number of 1's generated in the sequence of n integers.

Let Y = number of 5's generated in the sequence of n integers.

(Hint: it might be useful to define indicator variables denoting whether the i-th integer generated is a 1, and likewise for whether the i-th integer generated is a 5.)

a. (13 points) What is Cov(X, Y)?

b. (7 points) What is the correlation ρ(X, Y)?

5. (25 points) Say that our favorite manufacturer of computers produces machines whose lifetimes (independently) each have a mean of 100 weeks with a variance of 25 weeks^2. We buy and use a computer built by this manufacturer until it dies. When the machine dies, we immediately buy a new machine to use by the same manufacturer, and repeat this process every time the machine we are using dies. Let X = the total number of weeks that we use a machine built by our favorite manufacturer.

What is the minimal number of machines we would have to buy so that the total number of weeks we use a machine built by our favorite manufacturer is greater than 2000 with 95% probability? Justify your answer mathematically.

(Hint: derive an expression for the number of machines that will be necessary to purchase in order to have $P(X > 2000) \geq 0.95$, and use that to determine the answer).

6. (25 points) Consider the following recursive functions:

```
int Near() {
   int b = randomInteger(1, 4); // equally likely to be 1, 2, 3 or 4
   if (b == 1) return 2;
   else if (b == 2) return 4;
   else if (b == 3) return (6 + Near());
   else return (8 + Near());
}

int Far() {
   int a = randomInteger(1, 3); // equally likely to be 1, 2 or 3
   if (a == 1) return 2;
   else if (a == 2) return (2 + Near());
   else return (4 + Far());
}
```

Let Y = the value returned by `Far()`.

a. (10 points) What is E[Y]? Give a numeric answer.

b. (15 points) What is Var(Y)? Give a numeric answer.

7. (20 points) Consider a hash table with n buckets. Assume that all buckets of the hash table are equally likely to get hashed to. Let X = the number of strings that are hashed up to and including the first time a string is hashed to a non-empty bucket (i.e., a "collision" occurs).

   a. (6 points) Say the number of buckets in the hash table $n = 4$. Calculate $P(X = k)$ for $1 \le k \le 5$. In other words, calculate the probability that _exactly_ k strings are hashed at the time that the first string is hashed to a non-empty bucket.

     $P(X = 1) =$

     $P(X = 2) =$

     $P(X = 3) =$

     $P(X = 4) =$

     $P(X = 5) =$

   b. (7 points) Still considering the case where $n = 4$, calculate $E[X]$.

   c. (7 points) Now considering the general case of a hash table with n buckets, give a well-defined mathematical expression to compute $E[X]$ in terms of n (for arbitrary positivex integer values of n). Of course, it's fine for your expression to include summations, products, etc. as long as it is well-defined.

8. (15 points) Consider a sample of independently and identically distributed (I.I.D.) random variables X_1, X_2, …, X_n, that each have Geometric distributions. In other words, $X_i \sim \text{Geo}(p)$ for all $1 \le i \le n$.

   a. (11 points) Derive the Maximum Likelihood Estimate for the parameter p of the Geometric distribution. Give a numeric answer. Explicitly show all the steps in your derivation. (For this problem, the derivation is worth many more points than the final answer, so please be clear how you derived your final answer).

   b. (4 points) Say that we have a sample of five such I.I.D. Geometric variables with the following values: $X_1 = 4$, $X_2 = 3$, $X_3 = 4$, $X_4 = 2$, $X_5 = 7$. What value of p in the Geometric distribution would maximize the likelihood of these observations? Give a numeric answer.

9. (20 points) In a particular domain, we are able to observe three real-valued input variables X_1, X_2, and X_3 and want to predict a single binary output variable Y (which can have values 0 or 1). We know the functional forms for the input variables are all uniform distributions, namely: $X_1 \sim$ Uni(a_1, b_1), $X_2 \sim$ Uni(a_2, b_2), and $X_3 \sim$ Uni(a_3, b_3), but we are not given the values of the parameters a_1, b_1, a_2, b_2, a_3 or b_3. We are, however, given the following data set of 8 training instances:

X_1	X_2	X_3	Y
0.1	0.8	0.4	0
0.7	0.6	0.1	0
0.3	0.7	0.2	0
0.4	0.4	0.6	0
0.8	0.2	0.5	1
0.5	0.7	0.8	1
0.9	0.4	0.7	1
0.6	0.6	0.4	1

a. (10 points) Use Maximum Likelihood Estimators to estimate the parameters a_1, b_1, a_2, b_2, a_3 and b_3 in the case where Y = 0 as well as the case Y = 1. (I.e., estimate the distribution P(X_i | Y) for i = 1, 2, and 3). Note that the parameter values for a_1, b_1, a_2, b_2, a_3 and b_3 may be different when Y = 0 versus when Y = 1.

b. (10 points) You are given the following 3 testing instances, numbered 1, 2 and 3. (Note that the testing instances do not have output variable Y specified).

	X_1	X_2	X_3
test instance 1	0.5	0.6	0.4
test instance 2	0.7	0.7	0.7
test instance 3	0.5	0.4	0.3

Using the Naive Bayes assumption and your probability estimates from part (a), predict the output variable Y for *each* instance (you should have 3 predictions). Show how you derived the prediction by showing the computations you made for each test instance.

Standard Normal Table

Note: An entry in the table is the area under the curve to the left of z, $P(Z \le z) = \Phi(z)$

Z	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	0.7703	0.7734	0.7764	0.7793	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	0.8906	0.8925	0.8943	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
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998