Math 201: Introduction to Probability

Sample Final Exam

June 20th, 2023

NAME (please print legibly): ________________________________

Your University ID Number: ________________________________

- The exam will be 180 minutes long. You will get extra time in the end to upload the exam to Gradescope.
- There are 14 pages.
- A sheet with values of $\Phi(x)$ is provided.
- You may use any formulas from class without proof as long as you state it accurately.
- No calculators, phones, electronic devices, books, notes are allowed during the exam. The only materials you are allowed to use are are pen/pencil and paper. In particular, you are NOT allowed to take the exam on a tablet.
- You are allowed to use a phone or tablet to take photographs of your answer sheet once the exam is over. If you finish early, you must take permission before taking photographs. Once you start taking photographs, you are not allowed to write.
- **Show all work and justify all answers as much as possible.** You may not receive full credit for a correct answer if insufficient work is shown or insufficient justification is given.
- You do not need to simplify complicated expressions such as $\binom{200}{15}$ or 500!.

QUESTION	VALUE	SCORE
1	0	
2	15(A)	
3	15(A)	
4	10(A)	
5	20(A)	
6	30(B)	
7	30(B)	
8	20(B)	
9	20(B)	
10	20(B)	
11	20(B)	
TOTAL	200	

Value Table of $\Phi(x)$

	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
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

1. (0 points) Copy the following honesty pledge on to your answer sheet. Remember to sign and date it.

Pledge of Honesty

I affirm that I will not give or receive any unauthorized help on this exam, and that all work will be my own.

Signature: ______________________________

(A) **2. (15(A) points)**

Let X be a random variable with probability density function

$$f_X(x) = \begin{cases} e^{-x} + 2cx & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

where c is a constant.

(a) Find c.

(b) Find $\mathbb{E}[X]$.

(c) Find the cumulative distribution function (c.d.f.) of X.

(A) **3. (15(A) points)** Scientists are trying to do an experiment. They know the outcome of the experiment X is random variable with distribution $\mathrm{Ber}(1/2)$. However, due to inaccuracies in their measuring equipment, there is a noise parameter $G \sim \mathrm{Exp}(1/3)$, and what actually gets measured by the equipment is $M = X + G$. Assuming that X and G are independent, what is the probability that $X = 1$ if the measurement M is at least 4?

(A) **4. (10(A) points)** In the city of Gotham, anyone who likes drinking tea doesn't like drinking anything else. If everyone likes at least one of the beverages tea, coffee, or soda, 10% of people like tea, 50% people like soda, and 60% of people like coffee, then how many people like both coffee and soda?

(A) **5. (20(A) points)**

You want to find out how popular pineapple is on pizzas. You randomly call $90,000$ people around the US and among them $42,000$ said pineapple on a pizza is unacceptable.

(a) Give a 95% confidence interval for the true proportion who find pineapple on a pizza unacceptable.

(b) A national vote was held about pineapple on a pizza. The result says 40% of the population finds pineapple on a pizza unacceptable. Alfredo's sells pizzas with pineapple on them at \$9 per pizza, and pizzas without pineapple on them at \$10. Let X_n be the money Alfredo's makes from selling n pizzas. Express X_n in terms of a distribution you know.

(c) Find

$$\lim_{n\to\infty} \mathbb{P}(X_n > 9.5n).$$

(B) **6. (30(B) points)**

Suppose (X, Y) are uniformly distributed on the triangular region

$$D = \{(x, y) : x + y \leq 1, x \geq 0, y \geq 0\}$$

(a) Find the marginal density functions f_X and f_Y.

(b) Find $M_X(t)$.

(c) Let A be the area of the square bounded by the points $(0, 0)$, $(X, 0)$, $(0, Y)$, and (X, Y). Compute $\mathbb{E}[A]$.

(d) Using your computations in (b) and (c) or otherwise, find the correlation coefficient $\text{Corr}[X, Y]$.

(e) Are X and Y independent? Remember to justify your answer.

(B) **7. (30(B) points)**

Suppose X and Y are independent random variables with the following moment generating functions (m.g.f.),

$$M_X(t) = \frac{1}{4}e^{-t} + \frac{1}{4} + \frac{1}{2}e^{t},$$
$$M_Y(t) = \frac{2}{3}e^{-t} + \frac{1}{3}e^{2t}.$$

(a) Find the probability mass functions (p.m.f.) p_X and p_Y.

(b) Find the joint p.m.f., $p_{X,Y}(x, y)$ and express it as a table.

(c) Let $Z = X + Y$. What is $M_Z(t)$?

(d) Compute the p.m.f. p_Z.

(e) Compute $\mathrm{Var}[Z^2]$.

(B) **8. (20(B) points)**

Recall that NegBin(k, p) for $k \in \mathbb{N}$ and $0 < p < 1$ is the negative binomial distribution and it is defined as the number of independent Ber(p) trials before one sees k successes.

The p.m.f. of $X \sim$ NegBin(k, p) is given by

$$p_X(n) = P(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}.$$

In the following parts, it may be useful to recall that if $Y \sim$ Geom(p), then

$$E[Y] = \frac{1}{p} \qquad \mathrm{Var}[Y] = \frac{1-p}{p^2}.$$

(a) Find a formula for $\mathbb{E}[X]$ and Var$[X]$. (**Hint:** do not compute this directly! Instead, use the relationship between NegBin and Geom.)

(b) In class, we showed using the central limit theorem that when n is a large positive integer, then Poisson(n) behaves like a normal variable. Mimic this argument to show that when k is a large positive integer, then NegBin(k, p) behaves like a normal variable. What are its parameters?

(B) **9. (20(B) points)**

On average Watson and Holmes have to wait 20 days between visits by Lestrade. Lestrade just left after a visit, and suppose X is the number of days from now he will visit again.

(a) If you know nothing else about the distribution of X, what upper bound can you provide for the probability that $X > 50$? Clearly state any inequality you use, and explain why the hypotheses apply.

(b) Suppose X is a continuous random variable. You can assume that X is memoryless – if Lestrade hasn't visited after t days, then probability he will take at least s more days is the same as the probability that he would have taken s days in the first place. That is,

$$P(X > s + t | X > t) = P(X > s).$$

Can you identify the distribution of X?

(c) Compute $P(X > 50)$ exactly.

(d) Compute an upper bound on $P(X > 50)$ using Chebyshev's inequality.

(B) **10. (20(B) points)**

Two-Face rolls a standard die repeatedly. Let A_j be the indicator random variable for the event that the jth roll is 1, and B_j be the indicator random variable for the event that the jth roll is 2.

(a) Find $\mathbb{E}[A_j B_k]$. (**Hint:** consider the cases $j = k$ and $j \neq k$ separately)

(b) Let X be the number of 1s and Y be the number of 2s that show up in n rolls of the die. Find $\mathbb{E}[XY]$.

(c) Compute $\text{Cov}[X, Y]$.

(B) **11. (20(B) points)** No Hobbit in Middle-Earth is taller than 5 feet. You decide to go around randomly asking Hobbits their height to get a good estimate for their average height. Using the law of large numbers, estimate how many Hobbits you need to ask before there is a 99% chance that the average height you sampled differs from the actual average by at most 1 inch? [**Note:** 12 inches is 1 feet.]