Probability and Statistics I

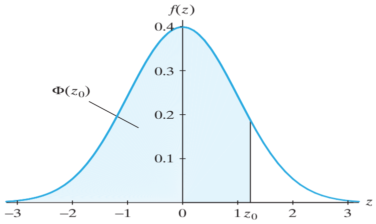
Mid-term Examination
SDS, CUHK(SZ)

March 20, 2021

Name: ____________________ Student ID: ____________________

Answer the multiple choice questions (Section I) in the Answer Card, and answer the regular questions (Section II) in the Answer Book.

Table Va The Standard Normal Distribution Function



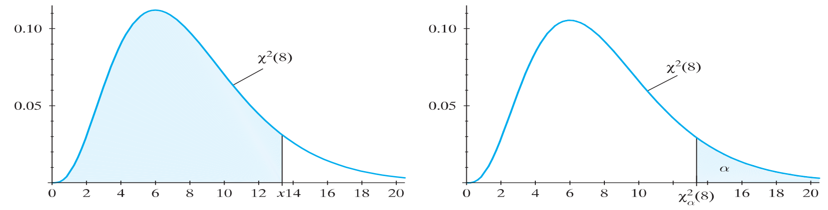
$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

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	0.7794	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

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.326	2.576	2.807	3.291

Table IV The Chi-Square Distribution



$$P(X \le x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2} dw$$

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58
16	5.812	6.908	7.962	9.312	23.54	26.30	28.84	32.00
17	6.408	7.564	8.672	10.08	24.77	27.59	30.19	33.41
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.80
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	35.56	38.88	41.92	45.64
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
40	22.16	24.43	26.51	29.05	51.80	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4
80	53.34	57.15	60.39	64.28	96.58	101.9	106.6	112.3

This table is abridged and adapted from Table III in *Biometrika Tables for Statisticians*, edited by E.S.Pearson and H.O.Hartley.

I Multiple Choices (72 points)

- **3 points for each correct answer; -1 point for each incorrect answer; 0 points for no answer**
- **For each question, only choose (at most) one out of four given choices (A,B,C and D). If you choose more than one choice in one question, your answer will be incorrect and 1 point will be deduced.**

1. Let A and B be two events. Suppose $P(A) = 0.4$, $P(B) = 0.5$.
   (a) $P(A \cap B') = 0.2$ if $P(A \cap B) = 0.2$
   (b) $P(A \mid B) = 0.4$ if A and B are mutually exclusive
   (c) $P(A \cup B) = 0.7$ if A and B are independent
   (d) $P(A \cap B \mid A \cup B) = 0.5$ if $P(A \mid B) = 0.6$.

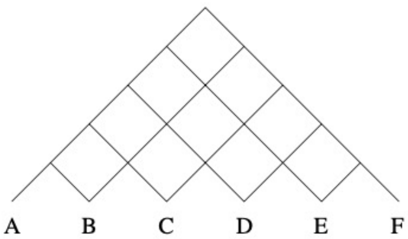
   Find which one of the following statements is correct.
   A. (a)(c)(d)
   B. (c)(d)

C. (b)(d)
D. (a)(c)

Solution: (A)
(a) $P(A \cap B') = P(A) - P(A \cap B) = 0.2$.
(b) $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = 0$.
(c) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = 0.7$.
(d)

$$\begin{aligned} P(A \cap B \mid A \cup B) = \frac{P((A \cap B) \cap (A \cup B))}{P(A \cup B)} &= \frac{P(A \cap B)}{P(A \cup B)} \\ &= \frac{P(A \mid B)P(B)}{P(A) + P(B) - P(A \mid B)P(B)} = 0.5. \end{aligned}$$

2. Skiers at the top of the mountain have a variety of choices as they head down the trails. Assume that at each intersection, a skier is equally likely to go left or right. Find the percent of skiers who end up at B and C, respectively.



A. (a) $\frac{5}{32}$, (b) $\frac{6}{16}$
B. (a) $\frac{5}{32}$, (b) $\frac{10}{32}$
C. (a) $\frac{4}{16}$, (b) $\frac{10}{32}$
D. (a) $\frac{4}{16}$, (b) $\frac{6}{16}$

Solution: (B)
At each turn there is the choice of going left or right. The total number of paths down the mountain is 2^5. To end up at point B the skier must make 1 left turns and 4 right turns, and to end up at point C the skier must make 2 left turns and 3 right turns. The proportion of skiers who end up at B is then

$$P(B) = \frac{\binom{5}{1}}{2^5} = \frac{5}{32},$$

and the proportion of skiers who end up at C is

$$P(C) = \frac{\binom{5}{2}}{2^5} = \frac{10}{32}.$$

3. In an urn are 7 blue and 3 red marbles.

(a) If you draw 3 marbles without replacement, what is the probability that less than 2 will be red?

(b) If you draw the marbles one by one without replacement, and STOP when only one color of marbles are left, what is the probability that the first drawn marble is red and the last drawn marble is blue?

Find which one of the following statements is correct.

A. (a) $\frac{98}{120}$ (b) $\frac{1}{10}$
B. (a) $\frac{88}{120}$ (b) $\frac{1}{10}$
C. (a) $\frac{98}{120}$ (b) $\frac{1}{15}$
D. (a) $\frac{88}{120}$ (b) $\frac{1}{15}$

Solution: (C)
(a) The probability of no red at all is $P(A) = \frac{\binom{7}{3}}{\binom{10}{3}}$, and the probability of one red is $P(B) = \frac{\binom{3}{1}\binom{7}{2}}{\binom{10}{3}}$. Hence $P(A \cup B) = P(A) + P(B) = \frac{\binom{7}{3}+\binom{3}{1}\binom{7}{2}}{\binom{10}{3}} = \frac{98}{120}$.
(b) The probability that the first drawn marble is red is $P(R) = \frac{3}{10}$. Given that the first one is red, the probability that the last drawn marble is blue can be computed as the following: supposing all the 10 marbles are drawn one by one, compute the probability that the last (10th) marble is red. The probability is given by $P(B \mid R) = \frac{2}{9}$. Hence $P(R \cap B) = \frac{3}{10}\frac{2}{9} = \frac{1}{15}$.

4. Let A, B and C be three events.
(a) If the events A and B are mutually exclusive, then A and B are NOT independent.
(b) If $A \subseteq B$, then A and B are NOT independent.
(c) If $P(A) = 0.4$, $P(B) = 0.6$ and $P(A \cup B) = 0.76$, then A and B are independent.
(d) Suppose A, B, C are pairwise independent. In addition, suppose A and $B \cup C$ are independent. Then A, B, C are mutually independent.

How many of the above four statements are TRUE?
A. 1
B. 2
C. 3
D. 4.

Solution: (B)
(a) False. Suppose $P(A) = 0$ and $B = A'$. Then $P(A \cap B) = 0 = P(A)P(B)$.
(b) False. Suppose $P(A) = 0$ and $B = S$. Then $P(A \cap B) = 0 = P(A)P(B)$.
(c) True. $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.24 = P(A)P(B)$.
(d) True. Since A and $B \cup C$ are independent, we have $P(A \cap (B \cup C)) = P(A)P(B \cup C)$ (*). Noting that $B \cup C = B \cup (B' \cap C)$, then $A \cap (B \cup C) = (A \cap B) \cup (A \cap B' \cap C)$. Hence $P(A \cap (B \cup C)) = P(A \cap B) + P(A \cap B' \cap C)$. We

then have from (*) that

$$P(A \cap B) + P(A \cap B' \cap C) = P(A)(P(B) + P(C) - P(B)P(C)).$$

It follows that

$$\begin{aligned} &P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\ &\qquad = P(A)P(B) + P(A)P(C) - P(A)P(B)P(C). \end{aligned}$$

Then we get $P(A \cap B \cap C) = P(A)P(B)P(C)$.

5. I have in my pocket five coins. Four of them are ordinary coins with equal chances of coming up head and tail when tossed, and the fifth has two heads.
(a) If I take one of the coins at random from my pocket and toss it, what is the probability that it comes up head?
(b) If I toss a randomly taken coin and it comes up head, what is the probability that it is the coin with two heads?

Find which one of the following statements is correct.

A. (a) $\frac{3}{5}$, (b) $\frac{1}{3}$
B. (a) $\frac{3}{5}$, (b) $\frac{1}{4}$
C. (a) $\frac{4}{7}$, (b) $\frac{1}{3}$
D. (a) $\frac{4}{7}$, (b) $\frac{1}{4}$

Solution: (A)

(a) Denote by D the event that the coin is the one with two heads, and denote by H the event that we get a head when we toss the coin. We have

$$\begin{aligned} P(H) =& P(H \cap D) + P(H \cap D') \\ =& P(H|D)P(D) + P(H|D')P(D'). \end{aligned}$$

Since $P(H|D) = 1$, $P(D) = \frac{1}{5}$, $P(H|D') = \frac{1}{2}$, and $P(D') = \frac{4}{5}$,

$$P(H) = \frac{3}{5}.$$

(b) Then we want to find $P(D|H)$. By Bayes' theorem, we have

$$P(D|H) = \frac{P(H|D)P(D)}{P(H)} = \frac{1}{3}.$$

6. Suppose that for three dice of the standard type all 216 outcomes of a throw are equally likely. Denote the obtained scores of the three dice by X_1, X_2 and X_3, respectively.
(a) $P(X_1 + X_2 + X_3 \leq 5) = 10/216$,

(b) $P(\min\{X_1, X_2, X_3\} \geq 2) = 125/216$,
(c) $P(X_1 + X_2 < (X_3)^2) = 137/216$.

Find which one of the following statements is correct.
A. (a)(b)
B. (a)(c)
C. (b)(c)
D. (a)(b)(c)

Solution: (D)

(a) There are 216 equally likely triples and of these only 10 have a sum ≤ 5 so $P(X_1 + X_2 + X_3 \leq 5) = 10/216$.

(b) The smallest of three numbers is bigger or equal to 2 only when all three are so $P(\min\{X_1, X_2, X_3\} \geq 2) = P(X_1 \geq 2)P(X_2 \geq 2)P(X_3 \geq 2) = 125/216$.

(c) Of 36 triples with $X_3 = 2$ only 3 have $X_1 + X_2 < 4$ and of 36 triples with $X_3 = 3$, 26 have $X_1+X_2 < 9$ so that $P(X_1+X_2 < (X_3)^2) = \sum_j P(X_1+X_2 < j^2, X_3 = j) = 137/216$.

7. A small plane went down and was missing, and the search was organized into three regions. Starting with the likeliest, they are:

Region	Initial chance the plane is there	Chance of being overlooked in the search
Mountains	0.5	0.3
Prairie	0.3	0.2
Sea	0.2	0.9

The last column gives the chance that if the plane is there, it will not be found. For example, if it went down at sea, there is 90% chance it will have disappeared, or otherwise not be found. Since the pilot is not equipped to long survive a crash in the mountains, it is particularly important to determine the chance that the plane went down in the mountains.
(a) Before any search is started, what is the chance that the plane is in the mountains?
(b) The initial search was in the mountains, and the plane was not found. Now what is the chance the plane is nevertheless in the mountains?
(c) The search was continued over the other two regions, and unfortunately the plane was not found anywhere. Finally now what is the chance that the plane is in the mountains?

Find which one of the following statements is correct.
A. (a) 0.50, (b) 0.23, (c) 0.38
B. (a) 0.50, (b) 0.28, (c) 0.30
C. (a) 0.20, (b) 0.23, (c) 0.38
D. (a) 0.20, (b) 0.28, (c) 0.30

Solution: (A)
Let M, P, S be the events that the plane went down in the mountains, prairie, sea respectively. Let OM, OP, OS be the events that the plane is not found in mountains, prairie, sea respectively. Then we have:

$$\begin{array}{rr} \mathrm{P}(M) = 0.5, & \mathrm{P}(P) = 0.3, \quad \mathrm{P}(S) = 0.2 \\ \mathrm{P}(OM \mid M) = 0.3, & \mathrm{P}(OP \mid M) = \mathrm{P}(OS \mid M) = 1 \\ \mathrm{P}(OP \mid P) = 0.2, & \mathrm{P}(OM \mid P) = \mathrm{P}(OS \mid P) = 1 \\ \mathrm{P}(OS \mid S) = 0.9, & \mathrm{P}(OM \mid S) = \mathrm{P}(OP \mid S) = 1 \end{array}$$

(a) $P(M) = 0.5$

(b) Consider

$$\begin{aligned} \mathrm{P}(M \mid OM) &= \frac{\mathrm{P}(OM \mid M)P(M)}{\mathrm{P}(OM \mid M)\,\mathrm{P}(M) + \mathrm{P}(OM \mid P)\,\mathrm{P}(P) + \mathrm{P}(OM \mid S)\,\mathrm{P}(S)} \\ &= \frac{(0.3)(0.5)}{(0.3)(0.5) + 0.3 + 0.2} = 0.2308 \end{aligned}$$

(c)
$$\begin{aligned} &\mathrm{P}(OM, OP, OS) \\ =&\mathrm{P}(OM, OP, OS \mid M)\,\mathrm{P}(M) + \mathrm{P}(OM, OP, OS \mid P)\,\mathrm{P}(P) + \mathrm{P}(OM, OP, OS \mid S)\,\mathrm{P}(S) \\ =&\mathrm{P}(OM \mid M)\,\mathrm{P}(M) + \mathrm{P}(OP \mid P)\,\mathrm{P}(P) + \mathrm{P}(OM \mid S)\,\mathrm{P}(S) \\ =&(0.3)(0.5) + (0.2)(0.3) + (0.9)(0.2) = 0.39 \\ &\therefore \quad \mathrm{P}(M \mid OM, OP, OS) = \frac{\mathrm{P}(OM \mid M)P(M)}{\mathrm{P}(OM, OP, OS)} = \frac{(0.3)(0.5)}{0.39} = 0.3846 \end{aligned}$$

8. Two balls are chosen randomly from an urn containing 7 white, 4 black, and 1 orange balls. Suppose that we win \$1 for each white ball drawn and we lose \$1 for each orange ball drawn. Denote $\$X$ as the amount that we can win. Determine the probability mass function $p(x)$ for (a) $x = 0$ and (b) $x = 2$, and determine the cumulative distribution function $F(x)$ for (c) $x = 0$ and (d) $x = 2$.

   Find which one of the following statements is correct.
   A. (a) $\frac{13}{66}$, (b) $\frac{7}{22}$, (c) $\frac{17}{66}$, (d) 1
   B. (a) $\frac{13}{66}$, (b) $\frac{7}{22}$, (c) $\frac{2}{33}$, (d) $\frac{15}{22}$
   C. (a) $\frac{2}{33}$, (b) $\frac{14}{33}$, (c) $\frac{15}{66}$, (d) 1
   D. (a) $\frac{2}{33}$, (b) $\frac{14}{33}$, (c) $\frac{2}{33}$, (d) $\frac{15}{22}$

Solution: (A)
Sample space: $\Omega = \{WW, WB, WO, BB, BO\}$ Note that the outcomes are not equally likely with associated Probabilities: $\mathrm{P}(\{WW\}) = \dfrac{\binom{7}{2}}{\binom{12}{2}} = \frac{7}{22}$, $\quad \mathrm{P}(\{WB\}) =$

$\frac{7\times4}{66} = \frac{14}{33}$, $\mathrm{P}(\{WO\}) = \frac{7\times1}{66} = \frac{7}{66}$ $\mathrm{P}(\{BB\}) = \frac{\binom{4}{2}}{66} = \frac{1}{11}$, $\mathrm{P}(\{BO\}) = \frac{4\times1}{66} = \frac{2}{33}$.

First analyze that $\mathrm{P}(X=-1) = \mathrm{P}(\{BO\}) = \frac{2}{33}, \mathrm{P}(X=0) = \mathrm{P}(\{WO, BB\}) = \frac{7}{66} + \frac{1}{11} = \frac{13}{66}, \mathrm{P}(X=1) = \mathrm{P}(\{WB\}) = \frac{14}{33}, \mathrm{P}(X=2) = \mathrm{P}(\{WW\}) = \frac{7}{22}$. The probability mass function (pmf) is given by

$$p(x) = \Pr(X = x) = \begin{cases} \frac{2}{33}, & \text{for } x = -1; \\ \frac{13}{66}, & \text{for } x = 0 \\ \frac{14}{33}, & \text{for } x = 1 \\ \frac{7}{22}, & \text{for } x = 2 \\ 0, & \text{otherwise} \end{cases}$$

The cumulative distribution function (cdf) is given by

$$F(x) = \mathrm{P}(X \le x) = \begin{cases} 0 & \text{for } x < -1; \\ p(-1) = \frac{2}{33}, & \text{for } -1 \le x < 0; \\ p(-1) + p(0) = \frac{17}{66}, & \text{for } 0 \le x < 1; \\ p(-1) + p(0) + p(1) = \frac{15}{22}, & \text{for } 1 \le x < 2; \\ 1, & \text{for } x \ge 2. \end{cases}$$

9. Consider a random experiment of "tossing a coin (not necessarily a fair one) indefinitely until it turns out a head, and the number of dollars you will win is equal to the number of tosses it takes to see the first head turning out." You are interested in how many dollars you will win after the experiment. Assume $\mathrm{P}(\{H\}) = p$.
(a) What is the expected amount of money you will win?
(b) What is the probability that you will win at least 3 dollars?

Find which one of the following statements is correct.
A. (a) $\frac{1}{p}$, (b) $1 - p$
B. (a) $\frac{1}{1-p}$, (b) p^2
C. (a) p, (b) p^2
D. (a) $\frac{1}{p}$, (b) $(1-p)^2$.

Solution: (D)
Define a random variable Z. With $\mathrm{P}(\{H\}) = p$, the pmf is

$$\mathrm{P}(Z = k) = (1-p)^{k-1}p = q^{k-1}p$$

for $k = 1, 2, 3, \ldots$, where $q = 1 - p$.

(a) Hence, the expected amount of money (in dollars) you will win is

$$
\begin{aligned}
\mathrm{E}(Z) &= \sum_{k=1}^{\infty} k q^{k-1} p \\
&= p \sum_{k=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}q} q^k \\
&= p \frac{\mathrm{d}}{\mathrm{d}q} \sum_{k=1}^{\infty} q^k \\
&= p \frac{\mathrm{d}}{\mathrm{d}q} \left(\frac{q}{1-q} \right) \\
&= p \frac{\mathrm{d}}{\mathrm{d}q} \left(\frac{1}{1-q} - 1 \right) \\
&= \frac{p}{(1-q)^2} \\
&= \frac{1}{p}.
\end{aligned}
$$

(b) The probability that you win at least 3 dollars is

$$
\begin{aligned}
\mathrm{P}(Z \geq 3) &= 1 - \mathrm{P}(Z = 1) - \mathrm{P}(Z = 2) \\
&= 1 - p - qp \\
&= q^2
\end{aligned}
$$

10. Consider a two-engine plane and a four-engine plane. Suppose that each engine on each plane will fail independently with the same probability p (where $0 < p < 1$) and that each plane will make a safe flight if at least half of the engines remain operative.
(a) Flying with a four-engine plane is safer than flying with a two-engine plane;
(b) Flying with a two-engine plane is safer than flying with a four-engine plane;
(c) Flying with a two-engine plane may be safer than flying with a four-engine plane;
(d) Flying with a two-engine plane is the same as flying with a four-engine plane.

Find which one of the following statements is correct.
A. (a)
B. (d)
C. (b) and (d)
D. (c) .

Solution: (D)
P(safe flight for a two-engine plane) $= 1 - p^2$.

$$\mathrm{P}(\text{ safe flight for a four-engine plane }) = 1 - p^4 - 4p^3(1-p)^1 \ .$$

Consider P(safe flight for a two-engine plane) $\geq$ P(safe flight for a four-engine plane)

$$
\begin{aligned}
1-p^2 &\geq 1-p^4-4p^3(1-p)^1 \\
p^4+4p^3(1-p) &\geq p^2 \\
-3p^4+4p^3-p^2 &\geq 0 \\
p^2\left(3p^2-4p+1\right) &\leq 0 \\
(3p-1)(p-1) &\leq 0 \quad \because p^2>0 \text{ since } p>0
\end{aligned}
$$

Therefore, $\frac{1}{3} \leq p \leq 1$, with $p<1$, the solution found for the condition is $\frac{1}{3} \leq p < 1$. That means, it is possible that flying with a two-engine plane is safer than (or the same as) flying with a four-engine plane.

11. Let X follow a discrete uniform distribution on $\{a, \ldots, b\}$, where a and b are integers with $a \leq b$. The pmf of X is

$$
\mathrm{P}(X=x)=p(x)=\begin{cases} \frac{1}{b-a+1}, & \text{for } x \in \{a, \ldots, b\} \\ 0, & \text{otherwise} \end{cases}
$$

Let $E(X)$, $Var(X)$ and $M_X(t)$ denote the mean, variance and moment generating function of X, respectively.

Find which one of the following statements is correct.

A. (a) $E(X)=\frac{a+b}{2}$, (b) $Var(X)=\frac{(b-a+1)^2+1}{12}$, (c) $M_X(1)=\frac{e^a-e^b}{(b-a+1)(1-e)}$
B. (a) $E(X)=\frac{b-a}{2}$, (b) $Var(X)=\frac{(b-a+1)^2-1}{12}$, (c) $M_X(1)=\frac{e^a-e^{(b-1)}}{(b-a+1)}$
C. (a) $E(X)=\frac{a+b}{2}$, (b) $Var(X)=\frac{(b-a+1)^2-1}{12}$, (c) $M_X(1)=\frac{e^a-e^{(b+1)}}{(b-a+1)(1-e)}$
D. (a) $E(X)=\frac{b-a}{2}$, (b) $Var(X)=\frac{(b-a+1)^2+1}{12}$, (c) $M_X(1)=\frac{e^a-e^b}{(b-a+1)}$

Solution: (C)

(a) The expected value of X is $\mathrm{E}(X)=\sum_{x=a}^{b} \frac{x}{b-a+1}=\frac{a+(a+1)+\cdots+b}{b-a+1}=\frac{1}{b-a+1} \cdot \frac{(a+b)(b-a+1)}{2}=\frac{a+b}{2}$.

(b) To find the variance of $X, \mathrm{Var}(X)$, we can make use of a newly defined r.v. $Y=X-a+1$. Write $n=b-a+1$. Then Y is discrete uniform on $\{1, \ldots, n\}$. Also, $\mathrm{Var}(X)=\mathrm{Var}(Y)$. Hence, we need only find $\mathrm{Var}(Y)$.

$$
\begin{aligned}
\mathrm{E}(Y) &= \sum_{x=1}^{n} \frac{y}{n}=\frac{1}{n} \cdot(1+2+\cdots+n)=\frac{1}{n} \cdot \frac{n(n+1)}{2}=\frac{n+1}{2} \\
\mathrm{E}\left(Y^2\right) &= \sum_{x=1}^{n} \frac{y^2}{n}=\frac{1}{n} \cdot\left(1^2+2^2+\cdots+n^2\right)=\frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6}=\frac{(n+1)(2n+1)}{6}
\end{aligned}
$$

Therefore.

$$\begin{aligned}
\operatorname{Var}(X) = \operatorname{Var}(Y) = \mathrm{E}\left(Y^2\right) - [\mathrm{E}(Y)]^2 &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{2n^2+3n+1}{6} - \frac{n^2+2n+1}{4} \\
&= \frac{4n^2+6n+2-3n^2-6n-3}{12} \\
&= \frac{n^2-1}{12} \\
&= \frac{(b-a+1)^2-1}{12}.
\end{aligned}$$

(c) The mgf of X is

$$M_X(t) = \mathrm{E}\left(e^{tX}\right) = \sum_{x=a}^{b} \frac{e^{tx}}{b-a+1} = \frac{e^{ta}}{b-a+1}\left[1+e^t+e^{2t}+\cdots+e^{(b-a)t}\right]$$

$$= \begin{cases} \frac{e^{at}\left(1-e^{(b-a+1)t}\right)}{(b-a+1)(1-e^t)}, & \text{for } t \neq 0 \\ 1, & \text{for } t = 0 \end{cases}$$

$$= \begin{cases} \frac{e^{at}-e^{(b+1)t}}{(b-a+1)(1-e^t)}, & \text{for } t \neq 0 \\ 1, & \text{for } t = 0 \end{cases}$$

12. The number of times that an individual contracts a cold in a given year is a Poisson random variable with mean $\theta = 6$. Suppose a new wonder drug (based on large quantities of vitamin C) has just been marketed that reduces the Poisson mean to $\theta = 4$ for 60 percent of the population. For the other 40 percent of the population the drug has no appreciable effect on colds. If an individual tries the drug for a year and has 3 colds in that time, how likely is it that the drug is beneficial for him/her?

A. 0.82
B. 0.77
C. 0.56
D. 0.69 .

Solution: (B)
Let X be the no. of colds the individual contracts within a year. Then $X \sim \mathrm{Po}(4)$ if the drug is effective and $X \sim \mathrm{Po}(6)$ if the drug is not effective. Denote E as the event that the drug is effective, then

$$\begin{aligned}
\mathrm{P}(X=3 \mid E) &= \frac{e^{-4}4^3}{3!} = \frac{32e^{-4}}{3}, \quad \mathrm{P}(E) = 0.6 \\
\mathrm{P}\left(X=3 \mid E^c\right) &= \frac{e^{-6}6^3}{3!} = 36e^{-6}, \quad \mathrm{P}\left(E^c\right) = 0.4
\end{aligned}$$

Hence using law of total Probability and Bayes' theorem,

$$\mathrm{P}(X=3) = 0.6 \times \frac{32e^{-4}}{3} + 0.4 \times 36e^{-6} = 6.4e^{-4} + 14.4e^{-6}$$
$$\mathrm{P}(E \mid X=3) = \frac{\mathrm{P}(X=3 \mid E)\,\mathrm{P}(E)}{\mathrm{P}(X=3)} = \frac{6.4e^{-4}}{6.4e^{-4} + 14.4e^{-6}} = 0.7666$$

13. Ten percent of all trucks undergoing a certain inspection will fail the inspection. Assume that trucks are independently undergoing this inspection one at a time. The expected number of trucks inspected before a truck fails inspection is

A. 1 B. 5 C. 10 D. 20

Solution: (C) Since this follows a geometric distribution with a probability of success 0.1, simply use the formula for the mean (or expected value) of a geometric distribution is $1/0.1 = 10$.

14. A discrete random variable X has a pmf such that $P(X = k+1) = aP(X = k)$, where $1 > a > 0$ and $k = 0, 1, 2 \ldots$. Compute the probability $P(X \geq 11)$.

A. a^{11} B. $1 - a^{11}$ C. a^{10} D. a^{12}

Solution: (A)

$$P(X = k+1) = aP(X = k) = a^2 P(X = k-1) = \cdots = a^{k+1} P(X = 0)$$

Normalization:

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} a^k P(X = 0) = P(X = 0) \left(\frac{1}{1-a} \right) = 1$$

Therefore, $P(X = 0) = 1 - a$.

$$P(X \geq 11) = 1 - P(X \leq 10) = 1 - \sum_{k=0}^{10} P(X = k)$$
$$= 1 - \sum_{k=0}^{10} a^k P(X = 0) = 1 - (1-a) \left(\frac{1 - a^{11}}{1-a} \right) = a^{11}$$

15. Consider a sequence of N Bernoulli trials with probability of success being equal to

p. It is observed that two successes occurred in these N trials, where $2 < N$ and N is even. What is the probability that one success occurred in the first $N/2$ trials?

A. $\frac{N/2}{\binom{N}{2}p(1-p)^{(N-2)}}$ B. $\frac{N}{2(N-1)}$ C. $\frac{(N/2)p^{(N/2)}}{\binom{N}{2}p(1-p)^{(N-2)}}$ D. $\frac{(N/2)p^{(N-2)}(1-p)^{(N/2-1)}}{\binom{N}{2}p(1-p)^{(N-2)}}$

Solution: (B)

Define the events:

A : 2 successes occurred in N trials

B : 1 success occurred in first $N/2$ trials

$$\begin{aligned}
P(B \mid A) &= \frac{P(A \cap B)}{P(A)} \\
&= \frac{P\left(\left\{1 \text{ success in 1st } \frac{N}{2} \text{ trials}\right\} \cap \left\{1 \text{ success in 2nd } \frac{N}{2} \text{ trials}\right\}\right)}{P(A)} \\
&= \frac{(N/2)p(1-p)^{(N/2-1)}(N/2)p(1-p)^{(N/2-1)}}{\binom{N}{2}p^2(1-p)^{(N-2)}} \\
&= \frac{N}{2(N-1)}
\end{aligned}$$

Note that the events $\{1$ success in 1 st $\frac{N}{2}$ trials$\}$ and $\{1$ success in 2 nd $\frac{N}{2}$ trials$\}$ are independent. The final answer can also be obtained by noting that out of the total of $\binom{N}{2}$ ways of picking 2 successes from the N Bernoulli trials, there are $\left(\frac{N}{2}\right)$ ways of choosing one from the first $\left(\frac{N}{2}\right)$ trials and another $\left(\frac{N}{2}\right)$ ways of choosing another from the second $\left(\frac{N}{2}\right)$ trials.

16. A telephone company employs 5 operators who receive requests independently of one another. The number of the requests received by each operator has a Poisson distribution, and each operator receives on average 1 request every 30 minutes. What is the probability that during a given 2 hour period, exactly 4 of the 5 operators receive no requests?

    A. $5(e^{-8} - e^{-10})$.
    B. $4(e^{-10} - e^{-16})$.
    C. $5(e^{-16} - e^{-20})$.
    D. $5(e^{-10} - e^{-16})$.

Solution: (C)
Let X be the number of requests per hour. Then $X \sim$ Poisson(4). The probability that an operator gets no requests within an hour is $P(X = 0) = \frac{4^0 e^{-4}}{0!} = e^{-4}$. Hence the probability that 4 out of 5 operators receive no request is $P(Y = 4) = \binom{5}{4}(e^{-4})^4(1 - e^{-4})^{5-4} = 5(e^{-16} - e^{-20})$.

17. Suppose that X is a continuous random variable with the pdf

$$f(x) = \begin{cases} 1+x, & -1 \le x < 0, \\ 1-x, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Define a random variable $Y = X^2 + 1$. The question is what is the value of $P(\frac{5}{4} < Y \le \frac{7}{4})$.

A. $\sqrt{3} - \dfrac{3}{2}$ B. $2\sqrt{3} - 3$. C. $2 - \sqrt{3}$. D. $3 - 2\sqrt{2}$.

Solution: (A)
We have $F_Y(y) = P(Y \le y) = P(X^2 + 1 \le y)$. Apparently, when $y \le 1, F_Y(y) = 0$; when $y \ge 2, F_Y(y) = 1$. When $1 < y < 2$,

$$\begin{aligned} F_Y(y) &= P(-\sqrt{y-1} \le X \le \sqrt{y-1}) \\ &= \int_{-\sqrt{y-1}}^{\sqrt{y-1}} f_X(x)dx \\ &= \int_{-\sqrt{y-1}}^{0} (1+x)dx + \int_{0}^{\sqrt{y-1}} (1-x)dx \\ &= 2\sqrt{y-1} - y + 1. \end{aligned}$$

Hence,

$$F_Y(y) = \begin{cases} 0, & y \le 1, \\ 2\sqrt{y-1} - y + 1, & 1 < y < 2, \\ 1, & y \ge 2. \end{cases}$$

Then $P(\frac{5}{4} < Y \le \frac{7}{4}) = F_Y(\frac{7}{4}) - F_Y(\frac{5}{4}) = \sqrt{3} - \frac{3}{2}$.

18. Suppose that $f_1(x)$ is the pdf of the standard normal distribution, and $f_2(x)$ is the pdf of the uniform distribution over $[-1, 3]$.Let $f(x)$ be a function defined as

$$f(x) = \begin{cases} af_1(x), & x \le 0, \\ bf_2(x), & x > 0, \end{cases}$$

where $a > 0$ and $b > 0$. If $f(x)$ is a pdf, then which one of the following equalities must hold?

A. $2a + 3b = 4$.

B. $3a + 2b = 4$.

C. $a + b = 1$.

D. $a + b = 2$.

Solution: (A)
According to definition of probability density function:

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Since

$$f_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

and

$$f_2(x) = \begin{cases} \frac{1}{4}, & -1 \le x \le 3 \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= a\int_{-\infty}^{0} f_1(x)dx + b\int_{0}^{\infty} f_2(x)dx \\ &= a\frac{1}{2} + b\int_0^3 \frac{1}{4}dx \\ &= \frac{a}{2} + \frac{3}{4}b \\ &= 1. \end{aligned}$$

So $2a + 3b = 4$.

19. Let the random variable X have a distribution with probability density function

$$f(x) = \frac{1}{\theta}e^{-(x-\delta)/\theta}, \quad \delta < x < \infty.$$

What is the mean and variance of X?
A. $E(X) = \frac{\delta}{\theta}, Var(X) = (\frac{\delta}{\theta})^2$.
B. $E(X) = \theta + \delta, Var(X) = (\theta + \delta)^2$.
C. $E(X) = \theta - \delta, Var(X) = \theta^2$.
D. $E(X) = \theta + \delta, Var(X) = \theta^2$.

Solution: (D)
$M(t) = E[e^{tx}] = \int_\delta^\infty \frac{1}{\theta}e^{tx}e^{-(x-\delta)/\theta}dx = \int_\delta^\infty \frac{1}{\theta}e^{(t-1/\theta)x}e^{\delta/\theta}dx = \frac{1}{1-\theta t}$,
$M'(t) = \frac{\theta e^{t\delta}}{(1-\theta t)^2} + \delta e^{t\delta}\frac{1}{1-\theta t} \quad E(x) = M'(0) = \theta + \delta.$
$M''(t) = \frac{\theta\delta e^{t\delta}(1-\theta t)^2 + 2\theta^2(1-\theta t)e^{t\delta}}{(1-\theta t)^4} + \frac{\delta^2 e^{t\delta}(1-\theta t) + \theta\delta e^{t\delta}}{(1-\theta t)^2}$
$Var(x) = M''(0) - [M'(0)]^2 = 2\theta^2 + \delta^2 + 2\theta\delta - (\theta+\delta)^2 = \theta^2.$
or let $Y = X - \delta, f_Y(y) = \frac{1}{\theta}e^{-y/\theta}, \quad 0 < y < \infty$,
Y has an exponential distribution. $E(Y) = \theta$,
$X = Y + \delta, E(X) = E(Y + \delta) = E(Y) + \delta = \theta + \delta.$
$Var(X) = Var(Y + \delta) = Var(Y) = \theta^2.$

20. Consider the nonnegative random variable X with the pdf

$$f_X(x) = \alpha x e^{-x^2} + \beta I(0,1)$$

Here, α and β are constants to be determined. $I(0,1)$ is the indicator function given by

$$I(0,1) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find α and β such that the 80^{th} percentile is at the point $\pi_{0.8} = 1$.

A. $\alpha = 0.4e, \beta = 1 - 0.2e$
B. $\alpha = 1 - 0.2e, \beta = 0.4e$
C. $\alpha = 1 - 0.4e, \beta = 1 - 0.2e$
D. $\alpha = 0.4e, \beta = 0.2e$

Solution: (A)
Normalization:

$$1 = \int_0^\infty \left(\alpha x e^{-x^2} + \beta I(0,1)\right) dx = \frac{\alpha}{2} \int_0^\infty e^{-x^2} d\left(x^2\right) + \beta = \frac{\alpha}{2} + \beta$$

80^{th} percentile implies

$$\begin{aligned} 0.8 &= \int_0^1 \left(\alpha x e^{-x^2} + \beta I(0,1)\right) dx \\ &= \frac{\alpha}{2} \int_0^{x=1} e^{-x^2} d\left(x^2\right) + \beta \\ &= \frac{\alpha}{2}\left(1 - e^{-1}\right) + \beta \end{aligned}$$

Solving the two equations gives $\frac{\alpha}{2} e^{-1} = 0.2$, or $\alpha = 0.4e$ while $\beta = (1 - 0.2e) > 0$.

21. For a class of students, consider the probability that at least two of them have their birthdays on the same day. What's the minimum size of class, i.e., the minimum number of students in the class, such that the probability is larger than 0.5? (For simplicity, suppose each year has 365 days.)

Hint: The probability of a binomial distribution $b(n,p)$ can be approximated by the probability of a Poisson distribution with $\lambda = np$ for $n \ge 20$ and $p \le 0.05$, and moreover, $\ln 2 = 0.693$.

A. 22 B. 23 C. 24 D. 25

Solution: (B)
Let N be the number of students in a class, there are $n = \binom{N}{2} = \frac{N(N-1)}{2}$ pairs of students sharing birthdays in this class. For each pair, both students are born on the same day with probability $p = 1/365$. Each pair is a Bernoulli trial because

the two birthdays either match or not match. Besides, matches in two different pairs are independent. Therefore, X, the number of pairs sharing birthdays, is binomial distribution.

P(there are two students sharing birthday)=$1 - P$(no matches)
$=1 - P(X = 0) = 1 - (1 - \frac{1}{365})^n$.
We note that the probability of a binomial distribution $b(n, p)$ can be approximated by that of a Poisson distribution with $\lambda = np$ for $n \geq 20, p \leq 0.05$. So we can use Poisson approximation with $(1 - \frac{1}{365})^n \approx e^{-\lambda}, \lambda = np = N(N-1)/730$, then $1 - P(X = 0) = 1 - (1 - \frac{1}{365})^n \approx 1 - e^{-\lambda} \approx 1 - e^{-N^2/730}$.
$1 - e^{-N^2/730} > 0.5$, we obtain $N > \sqrt{730 ln 2} \approx 22.5$. Therefore, the minimum class size should be 23.

22. In a coin tossing game, the gambler has to pay the casino \$$X$ for each toss of the coin which has a probability p of a head and probability $q = 1 - p$ of a tail. The game stops when either a head shows up at which point he gets a reward of \$$Y$, or no head occurs after K tosses at which point he gets no reward. Find the expected net gain of the casino from the game.

A. $KX(1-p)^{K+1} + X\left(\frac{1}{p}\right)\left((1-q^K) - (K+1)pq^K\right) - Y$
B. $KX(1-p)^K + X\left((1-q^{K+1}) - Kpq^K\right) - Y$
C. $(K+1)X(1-p)^{K+1} + X\left(\frac{1}{p}\right)\left((1-q^K) - (K+1)pq^K\right) - Y$
D. $KX(1-p)^K + X\left(\frac{q}{p}\right)(1-q^K) + X(1-(K+1)q^K) - Y$

Solution: (D)
Suppose no head shows up after K tosses. The casino gains \$$KX$ and this event has a probability of $(1-p)^K$.
Let Z be the number of tosses to get to the first head where $Z \leq K$.
Net gain for $Z \leq K$ is $G = XZ - Y$.

$$\begin{aligned}
E[G] &= XE[Z] - Y \\
&= X\sum_{n=1}^{K} n(1-p)^{n-1}p - Y \\
&= Xp\frac{d}{dq}\sum_{n=1}^{K} q^n - Y \\
&= X\left(\frac{q}{p}\right)(1-q^K) + X(1-(K+1)q^K) - Y
\end{aligned}$$

where $q = 1 - p$. Therefore, the net gain from the game is

$$KX(1-p)^K + X\left(\frac{q}{p}\right)(1-q^K) + X(1-(K+1)q^K) - Y$$

23. Consider the following statements:

(i) Let $X \sim N(\mu, 1)$ with $\mu \geq 0$. The largest possible value of c such that the inequality $P(|X| \leq c) \leq 0.8064$ holds is $c = 1.3$.

(ii) Let $Z \sim N(0,1)$, then the probability that the quadratic equation $0.1x^2 + Zx + 0.04 = 0$ has real roots is 0.9.

(iii) If $X \sim N(0,3)$, then $\frac{E(X^6)}{E(X^4)} = 3$.

Find which one of the following statements is correct.

A. All statements are ture. B. Only (i) and (iii) are ture. C. Only (ii) and (iii) are ture. D. Only (i) and (ii) are ture.

Solution: (D)

(i) We can write

$$\begin{aligned} P(|X| \leq c) &= P(-c \leq X \leq c) \\ &= P\left(\frac{-c-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{c-\mu}{\sigma}\right) \\ &= P\left(\frac{-c-\mu}{\sigma} \leq Z \leq \frac{c-\mu}{\sigma}\right) \\ &= P\left(-c-\mu \leq Z \leq c-\mu\right) \\ &\leq P\left(-c \leq Z \leq c\right) \end{aligned}$$

where the last inequality is due to $\mu \geq 0$ and it will be an equality if $\mu = 0$.

Therefore, to find the largest possible value of c we only need to find c such that $P(-c \leq Z \leq c) = 0.8064$. So, we solve

$$P(-c \leq Z \leq c) = \Phi(c) - \Phi(-c) = 2\Phi(c) - 1 = 0.8064$$

and we obtain $c = 1.3$ by checking the table.

(ii) To have real roots, the coefficients of the equations must satisfy $b^2 - 4ac \geq 0$, which is equivalent to $Z^2 - 4 \times 0.1 \times 0.04 \geq 0$ in this question.

As $Z \sim N(0,1)$, using the mgf technique we may show that $Z^2 \sim \chi^2(1)$. Hence, we obtain $P(Z^2 \geq 0.016) = 1 - P(\chi^2(1) < 0.016) = 0.9$, where we obtain the result by checking the chi-square table.

(iii) Use the mgf of a normal random variable, we can show that

$$\begin{aligned} E\left[X^4\right] &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \\ E\left[X^6\right] &= \mu^6 + 15\mu^4\sigma^2 + 45\mu^2\sigma^4 + 15\sigma^6 \end{aligned}$$

Plug in the numbers, then we have $E\left[X^6\right] \div E\left[X^4\right] = 15$.

24. The number of cars passing a speed camera follows a Poisson distribution, and the average number of cars passing the camera in 1 minute is 2. Suppose current time is $t = 0$ and the unit is minute, and let the random variable Y be the time that the 6th

car passes the speed camera and let the random variable Z be the time that the 1st car passes the speed camera.

Consider the following statements:

(i) The probability of at least 3 cars passing the speed camera in the first 2 minute is 0.7619 approximately.

(ii) $E(Y) = 3$, $\text{Var}(Y) = 1.5$, $E(Y^4) = 189$

(iii) for any $t, s > 0$, $P(Z > s + t|Z > t) = P(Z > s)$.

Find which one of the following statements is correct.

A. Only (i) is true. B. Only (i) and (ii) are true.
C. Only (ii) and (iii) are true. D. All statements are true.

Solution: (D)

(i) The mean number of car passing the speed camera in 2 minute is 4. Let X be the number of car passing the speed camera in the first 2 minute, so $X \sim \text{Poisson}(4)$, and then

$$\begin{aligned} P(X \geq 3) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - \frac{e^{-4}(4)^0}{0!} - \frac{e^{-4}(4)^1}{1!} - \frac{e^{-4}(4)^2}{2!} \\ &= 0.76189 \end{aligned}$$

(ii) Note that Y follows a Gamma distribution with $\alpha = 6$ and $\theta = 1/2 = 0.5$.

$$E(Y) = \alpha\theta = 3, \quad \text{Var}(Y) = \alpha\theta^2 = 1.5$$

$$E(Y^4) = \alpha(\alpha + 1)(\alpha + 2)(\alpha + 3)\theta^4 = 189$$

(iii) This is the memoryless property of the exponential distribution as shown in the assignment.

II Regular Questions (28 points)

25. (12 points) Consider a sequence of n Bernoulli trials over the interval $[0, 1]$ which is divided into a large number n of subintervals each of width $1/n$ such that each trial occurs within a subinterval. The probability of a success in each trial is p, and the probability of two or more successes in each subinterval is practically zero. **Note: all details should be included!**

    (a) (2 points) Supposing X is the number of successes in the interval $[0, 1]$, write an expression for $P(X = k)$ where $k = 0, 1, 2, \ldots, n$ is an integer.

    (b) (1 point) Derive an expression for the ratio $\frac{P(X=k+1)}{P(X=k)}$.

    (c) (4 points) Let $n \to \infty$ and the probability of success $p \to 0$ in such a way that the product np remains a constant μ. Derive an approximate expression for the limiting value of this ratio $\frac{P(X=k+1)}{P(X=k)}$ for a given fixed value of k. You must present your reasons in full.

(d) (4 points) Using the result in (c), derive the expression for $P(X = 0)$.

(e) (1 point) Finally, derive the expression for $P(X = k)$ for an arbitrary integer k.

Solution:

(a) X is a Binomial random variable

$$P(X = k) = \binom{n}{k}(p)^k(1-p)^{n-k},\ k = 0, 1, 2, \ldots, n$$

(b) $\frac{P(X=k+1)}{P(X=k)} = \frac{\left(1-\frac{k}{n}\right)np}{(k+1)(1-p)}$, after simplification

(c) Let $n \to \infty$ and $p \to 0$ such that np remains a constant μ.
Then, $\left(1 - \frac{k}{n}\right) \to 1$ for a fixed k, and $(1-p) \to 1$, and we have

$$\frac{P(X = k+1)}{P(X = k)} \approx \frac{\mu}{(k+1)}$$

(d) $P(X = k+1) = \frac{\mu}{(k+1)}P(X = k)$
This gives, for $k = 0, 1, 2, \ldots,$

$$\begin{aligned} P(X = 1) &= \frac{\mu}{1!}P(X = 0) \\ P(X = 2) &= \frac{\mu^2}{2!}P(X = 0) \\ &\vdots \\ P(X = k) &= \frac{\mu^k}{k!}P(X = 0) \end{aligned}$$

Normalization:

$$1 = \sum_{k=0}^{\infty} P(X = k) = P(X = 0)\sum_{k=0}^{\infty}\frac{\mu^k}{k!} = e^{\mu}P(X = 0)$$

This yields $P(X = 0) = e^{-\mu}$.

(e) $P(X = k) = \frac{\mu^k}{k!}P(X = 0) = \frac{\mu^k}{k!}e^{-\mu}, \quad k = 0, 1, 2, \ldots$
This is the Poisson pmf.

26. (16 points) Let X be a continuous random variable with probability density function (pdf) defined as follows

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, x \in (-\infty, \infty).$$

Note: all details should be included!

(a) (4 points) Derive the moment generating function for X and show that $\mathrm{E}(X) = 0$, $\mathrm{Var}(X) = 1$.

(b) (3 points) Define a random variable $Y = aX + b$. What is the distribution of Y? Show the proof in details.

(c) (2 points) What is $E((Y - b)^{177})$?

(d) (7 points) Prove that $\mathrm{Var}(X) = 1$ without using the moment generating function.

Solution:

(a) The derivation and the proof for $\mathrm{E}(X) = 0$, $\mathrm{Var}(X) = 1$ can be found in the end of lecture note 11.

(b) Using the moment generating function technique, we have

$$\begin{aligned}\mathrm{E}(\exp(tY)) &= \mathrm{E}(\exp(taX + tb)) = \mathrm{E}(\exp(taX)\exp(tb)) \\ &= \mathrm{E}(\exp(taX)\exp(tb)) = \exp(\frac{1}{2}a^2t^2)\exp(tb) \\ &= \exp(tb + \frac{1}{2}a^2t^2)\end{aligned}$$

Therefore, $Y \sim N(b, a)$.

(c) Since $Y - b = aX$, $\mathrm{E}((Y - b)^{177}) = \mathrm{E}(b^{177}X^{177}) = b^{177}\mathrm{E}(X^{177})$.

$$\mathrm{E}(X^{177}) = \int_{-\infty}^{\infty} x^{177}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx = -\int_{-\infty}^{\infty} z^{177}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz = -\mathrm{E}(X^{177})$$

Therefore, $\mathrm{E}((Y - b)^{177}) = 0$.

(d)

$$\begin{aligned}\mathrm{Var}(X) = E(X^2) &= \int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx \\ &= 2\int_{0}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx\end{aligned}$$

Let $z = x^2$. Then we have

$$\mathrm{Var}(X) = \frac{1}{\sqrt{2\pi}}\int_0^{\infty} z^{\frac{1}{2}}e^{-\frac{1}{2}z}dz$$

Let $y = \frac{1}{2}z$. Then

$$\begin{aligned}\mathrm{Var}(X) &= \frac{1}{\sqrt{2\pi}}\int_0^{\infty} 2^{\frac{3}{2}}y^{\frac{1}{2}}e^{-y}dy \\ &= \frac{2}{\sqrt{\pi}}\int_0^{\infty} y^{\frac{1}{2}}e^{-y}dy = \frac{2}{\sqrt{\pi}}(-1)\int_0^{\infty} y^{\frac{1}{2}}de^{-y} \\ &= \frac{2}{\sqrt{\pi}}(-1)\left\{y^{\frac{1}{2}}e^{-y}|_0^{\infty} - \frac{1}{2}\int_0^{\infty} y^{-\frac{1}{2}}e^{-y}dy\right\} \\ &= \frac{2}{\sqrt{\pi}}\frac{1}{2}\int_0^{\infty} y^{-\frac{1}{2}}e^{-y}dy \\ &= \frac{1}{\sqrt{\pi}}\Gamma(\frac{1}{2}) = \frac{1}{\sqrt{\pi}}\sqrt{\pi} = 1\end{aligned}$$