Probability and Statistics I

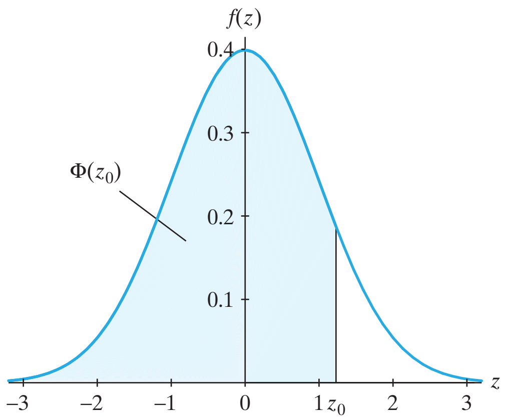
Mid-term Exam
SDS, CUHK(SZ)

March 18, 2023

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



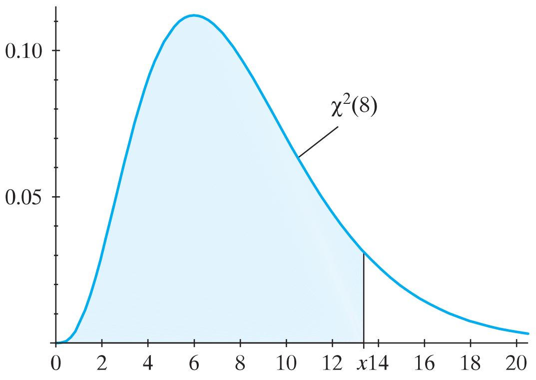
$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

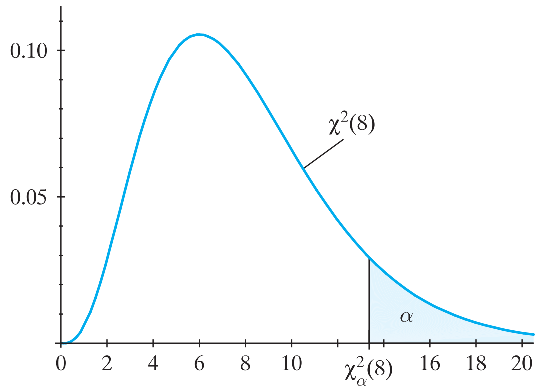
$$\Phi(-z) = 1 - \Phi(z)$$

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	$\boxed{8.672}$	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)

- **For each question, choose one and only one out of four given choices (A,B,C and D).**
- **3 points for each correct answer; 0 point for each incorrect or no answer.**

1. Let A, B, and C be three events in the sample space S. Suppose we know that $A \cup B \cup C = S$, $\mathbb{P}(A) = \frac{1}{2}$, $\mathbb{P}(B) = \frac{2}{3}$, $\mathbb{P}(A \cup B) = \frac{5}{6}$.

   (i) $\mathbb{P}(A \cap B) = 1/3$.

   (ii) A and B are independent.

   (iii) $\mathbb{P}(C \cap (A \cup B)') = 1/6$.

   (iv) If $\mathbb{P}(C \cap (A \cup B)) = \frac{5}{12}$, then $\mathbb{P}(C) = 1/2$.

   Determine which of the following four answers is correct where T stands for true and F stands for false.

   A. (T,T,T,F) 75%
   B. (F,T,T,F) 2%
   C. (T,T,F,F) 10%
   D. (T,F,T,F) 13%

Solution: A
Using the inclusion-exclusion principle, we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Thus,

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) \\ &= \frac{1}{2} + \frac{2}{3} - \frac{5}{6} \\ &= \frac{1}{3}.\end{aligned}$$

Since $P(A \cap B) = P(A)P(B)$, they are independent. We can write

$$\begin{aligned}C - (A \cup B) &= (C \cup (A \cup B)) - (A \cup B) \\ &= S - (A \cup B) \quad \text{(since } A \cup B \cup C = S\text{)} \\ &= (A \cup B)^c.\end{aligned}$$

Thus,

$$\mathbb{P}\left(C-(A\cup B)\right)=\mathbb{P}\left((A\cup B)^c\right)=1-\mathbb{P}(A\cup B)=\frac{1}{6}.$$

And we have

$$\mathbb{P}(C)=\mathbb{P}\left(C\cap(A\cup B)\right)+\mathbb{P}\left(C-(A\cup B)\right)=\frac{5}{12}+\frac{1}{6}=\frac{7}{12}.$$

Teaching Objective: inclusion-exclusion principle, set operations. Difficulty Level: Moderate.

2. A box contains three marbles: one red, one green, and one blue. Consider an experiment that consists of taking one marble from the box then replace it in the box and drawing a second marble from the box.

   (i) The sample space has 9 outcomes in total.

   (ii) If, at all times, each marble in the box is equally likely to be selected, the probability for each outcome is equal to each other.

   (iii) If, the first draw is without replacement and at all times each marble in the box is equally likely to be selected, then the probability for each outcome in the sample space is equal to each other.

   Determine which of the following four answers is correct where T stands for true and F stands for false.

   A. (T,T,F) 13%
   B. (T,T,T) 78%
   C. (T,F,T) 4%
   D. (F,T,F) 5%

Solution: B

1. Using R, G and B to denote the color red, green and blue. The sample space is

$$S=\{(R,R),(R,G),(R,B),(G,R),(G,G),(G,B),(B,R),(B,G),(B,B)\}$$

So the number of outcomes is 9.

2. As each marble in the box is equally likely to be selected, the probability for each point in the sample space is $\frac{1}{9}$.

3. The same as the previous one.

Teaching Objective: Counting technique, sample space. Difficulty Level: Fundamental.

3. A pair of fair dice is rolled.

   (i) The probability that the second die lands on a higher value than the first does is 5/12.

   (ii) The probability that the second die lands on the same value than the first does is 1/2.

   (iii) The probability that the second die lands on a smaller value than the first does is 1/12.

   Determine which of the following four answers is correct where T stands for true and F stands for false.

   A. (T,T,F) 1%
   B. (T,T,T) 0%
   C. (T,F,F) 98%
   D. (F,T,F) 1%

Solution: C
Since the dice are fair and distinct, each of the $6 \cdot 6$ rolls are equally likely. There are $\binom{6}{2}$ ways for second dice to be greater than first dice.

$$P(event) = \frac{\binom{6}{2}}{6 \cdot 6} = \frac{15}{36} = \frac{5}{12}.$$

Similarly, the probability of the same value is 6/36= 1/6. The remaining probability is 1-5/12- 1/6= 5/12.

Teaching Objective: enumeration method, counting technique. Difficulty Level: Moderate.

4. Suppose that we toss 2 fair dice. Define the following three events:

   a The sum of the dice is 6;

   b The first die equals 4;

   c The sum of the dice is 7;

   Determine which of the following four answers is correct.

   A. a and b are independent 2%
   B. b and c are independent 79%
   C. a, b and c are mutually independent 7%
   D. a, b and c are pairwise independent. 12%

   Solution: B
   $P(a) = 5/36, P(b) = 1/6, P(c) = 1/6$. And we can see that $P(a \cap b) = 1/36, P(b \cap c) = 1/36$. So b and c are independent, but a,b,c are not independent.

   Teaching Objective: Independent events. Difficulty Level: Fundamental.

5. A coin that, when flipped, comes up heads with probability p is flipped until either heads or tails has occurred twice. Let X be the required number of flips.

   (i) X follows a geometric distribution.

   (ii) The range of X is $\{2,3\}$.

   (iii) $E(X) = 3 - p^2 - (1-p)^2$.

   Determine which of the following four answers is correct where T stands for true and F stands for false.

   A. (T,F,T) 5%
   B. (T,F,F) 7%

C. (F,T,T) 74%
D. (F,F,F) 14%

Solution: C
Let X denote the number of flips. We stop when either Heads or Tails occur twice. There will be a minimum of 2 flips and a maximum of 3.

$$P(X=2) = p^2 + (1-p)^2$$

$$P(X=3) = 1 - P(X=2) = 2p(1-p)$$

$$E(X) = 2P(X=2) + 3P(X=3) = 3 - p^2 - (1-p)^2$$

Teaching Objective: Multiplication principle. Difficulty Level: Moderate.

6. For some constant c, the random variable X has the probability density function

$$f(x) = \begin{cases} cx^4 & 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(i) $c = 5/32$.
(ii) $E(X) = 5/3$.
(iii) Var (X)= 5/63.

Determine which of the following four answers is correct where T stands for true and F stands for false.

A. (T,F,T) 7%
B. (T,F,F) 9%
C. (F,T,T) 2%
D. (T,T,T) 82%

Solution: D
First we can determine the constant C by $\int_0^2 f(x)dx = 1$, the result is $C = 5/32$. Next, we can compute the mean of X by $E(X) = \int_0^2 xf(x)dx = 5/3$. Finally, we can compute the variance by $Var(X) = EX^2 - (EX)^2 = 5/63$.

Teaching Objective: Probability density function, expectation, variance. Difficulty Level: Moderate.

7. Roll two fair dice. The sample space is $\{(x, y)\}$ where $x, y \in (1, 2, 3, 4, 5, 6)$.

   (i) Event "the outcome of the first die is 1" is $\{1\}$

   (ii) Event "at least one of the two outcomes is 3" is $\{3\}$

   (iii) The probability of "the outcome of the first die is strictly larger than that of the second" equals $\frac{5}{12}$

   (iv) The probability of "the sum of the two outcomes is an odd number" equals $\frac{1}{2}$

   Determine which of the following four answers is correct where T stands for true and F stands for false.

   A. (F, T, T, F) 4%
   B. (F, F, T, T) 94%
   C. (F, T, F, T) 1%
   D. (T, F, F, T) 1%

   Solution: B
   (a) is not a valid event
   (b) is not a valid event
   (c) We count the number of samples (x, y) such that $x > y$. There are 15 in total. Hence the probability can be computed by $\frac{15}{36} = \frac{5}{12}$.
   (d) We count the number of samples (x, y) such that $x + y$ is an odd number. There are 18 in total. Hence the probability can be computed by $\frac{18}{36} = \frac{1}{2}$.

   Teaching Objective: sample space, event and probability. Difficulty Level: Fundamental.

8. Flip two fair coins. The sample space is $\{HH, HT, TH, TT\}$ where H and T denote Head and Tail respectively.

   (i) Events $\{HH\}$ and $\{TT\}$ are independent

   (ii) Events $\{HH, HT\}$ and $\{HT, TT\}$ are independent

(iii) Events "the first outcome is H" and "the second outcome is T" are mutually exclusive

(iv) Events "the second outcome is T" and "two outcomes are different" are independent

Determine which of the following four answers is correct where T stands for true and F stands for false.

A. (T,F,T,F) 7%
B. (T,F,F,T) 21%
C. (F,T,T,F) 4%
D. (F,T,F,T) 68%

Solution: D
(a) $P(\{HH\} \cap \{TT\}) = 0 \neq P(\{HH\}) \times P(\{TT\})$, hence not independent
(b) $P(\{HH, HT\} \cap \{HT, TT\}) = P(\{HT\}) = \frac{1}{4} = P(\{HH, HT\}) \times P(\{HT, TT\})$, hence independent
(c) The two events are $\{HH, HT\}$ and $\{HT, TT\}$. They are not mutually exclusive.
(d) The first event is $\{TT, HT\}$ and the second event is $\{HT, TH\}$. Similar to (b), they are independent.

Teaching Objective: independence, mutually exclusiveness. Difficulty Level: Fundamental.

9. You have three coins. The coins are identical except that two of them are fair and the other one is biased, with a probability of Head equal to 4/5. Suppose you pick two coins from the three uniformly at random, and flip, and the outcomes are two Heads. Conditional on that, what is the probability that you picked the biased coin? Redo the calculation if the outcomes are two Tails.

A. (a) $\frac{16}{21}$ for two Heads, (b) $\frac{4}{9}$ for two Tails 82%
B. (a) $\frac{6}{7}$ for two Heads, (b) $\frac{4}{9}$ for two Tails 7%
C. (a) $\frac{16}{21}$ for two Heads, (b) $\frac{5}{9}$ for two Tails 9%
D. (a) $\frac{6}{7}$ for two Heads, (b) $\frac{5}{9}$ for two Tails 1%

Solution: A

Let A denote the event that both coins you picked are fair. We want to find $P(A'|HH)$. We have $P(A) = \frac{1}{3}$ and $P(A') = \frac{2}{3}$. $P(HH|A) = \frac{1}{4}$, $P(HH|A') = \frac{1}{2}\times\frac{4}{5} = \frac{2}{5}$. By Bayes theorem, $P(A'|HH) = \frac{P(HH|A')\times P(A')}{P(HH)} = \frac{\frac{2}{5}\times\frac{2}{3}}{\frac{2}{5}\times\frac{2}{3}+\frac{1}{4}\times\frac{1}{3}} = \frac{16}{21}$.

For the case of two tails, we have $P(TT|A) = \frac{1}{4}$, $P(TT|A') = \frac{1}{2}\times\frac{1}{5} = \frac{1}{10}$. By Bayes theorem, $P(A'|TT) = \frac{P(TT|A')\times P(A')}{P(TT)} = \frac{\frac{1}{10}\times\frac{2}{3}}{\frac{1}{10}\times\frac{2}{3}+\frac{1}{4}\times\frac{1}{3}} = \frac{4}{9}$.

Teaching Objective: conditional probability, Bayes theorem. Difficulty Level: Fundamental.

10. Place 15 identical red balls and 12 identical green balls into 10 bins. Each ball is placed into a bin uniformly at random. Compute the probability of the following events: (a) each bin contains at least one red ball AND one green ball, and (b) each bin contains at least one red ball OR one green ball.

A. (a) $\frac{\binom{15}{9}\times\binom{12}{9}}{\binom{25}{9}\times\binom{22}{9}}$, (b) $\frac{\binom{26}{9}}{\binom{36}{9}}$ 8%

B. (a) $\frac{\binom{15}{9}\times\binom{12}{9}}{\binom{24}{9}\times\binom{21}{9}}$, (b) $\frac{\binom{26}{9}}{\binom{37}{10}}$ 11%

C. (a) $\frac{\binom{14}{9}\times\binom{11}{9}}{\binom{24}{9}\times\binom{21}{9}}$, (b) $\frac{\binom{26}{9}}{\binom{36}{9}}$ 70%

D. (a) $\frac{\binom{14}{9}\times\binom{11}{9}}{\binom{24}{9}\times\binom{22}{9}}$, (b) $\frac{\binom{26}{9}}{\binom{37}{10}}$ 10%

Solution: C

The total number of ways of placing red balls is $\binom{24}{9}$. The total number of ways of placing green balls is $\binom{21}{9}$. The total number of ways of placing red balls such that all bins must contain at least one red ball is $\binom{14}{9}$, and the total number of ways of placing green balls such that all bins must contain at least one green ball is $\binom{11}{9}$. Finally, the total number of ways of placing balls of both colors into bins such that all bins must contain at least one ball is $\binom{26}{9}$ and the total number of ways of placing the balls into bins without recording their colors is $\binom{36}{9}$ Thus the first probability is $\frac{\binom{14}{9}\times\binom{11}{9}}{\binom{24}{9}\times\binom{21}{9}}$ and the second probability is $\frac{\binom{26}{9}}{\binom{36}{9}}$

Teaching Objective: counting. Difficulty Level: Moderate.

11. Suppose Y is an exponentially distributed random variable with parameter 1. Compute (a) $P(Y > 5|Y > 3)$ and (b) $P(1 < \sqrt{Y} < 2)$

A. (a) e^{-1}, (b) $e^{-1} - e^{-\sqrt{2}}$ 0%

B. (a) e^{-2}, (b) $e^{-1} - e^{-\sqrt{2}}$ 7%

C. (a) e^{-2}, (b) $e^{-1} - e^{-4}$ 90%

D. (a) e^{-1}, (b) $e^{-1} - e^{-4}$ 3%

Solution: C

First we have $P(Y > 5|Y > 3) = \frac{P(\{Y>5\}\cap\{Y>3\})}{P(Y>3)} = \frac{P(Y>5)}{P(Y>3)} = \frac{e^{-5}}{e^{-3}} = e^{-2}$. Second, $P(1 < \sqrt{Y} < 2) = P(1^2 \leq Y \leq 2^2) = (1 - e^{-4}) - (1 - e^{-1}) = e^{-1} - e^{-4}$.

Teaching Objective: calculating probabilities for continuous RV. Difficulty Level: Fundamental.

12. Suppose Z is a standard normal random variable. Find $E[|Z|]$ and $Var(Z^2)$

A. (a) $E[|Z|] = 1$, (b) $Var(Z^2) = 3$ 18%

B. (a) $E[|Z|] = \sqrt{\frac{2}{\pi}}$, (b) $Var(Z^2) = 2$ 74%

C. (a) $E[|Z|] = 1$, (b) $Var(Z^2) = 3$ 4%

D. (a) $E[|Z|] = \sqrt{\frac{2}{\pi}}$, (b) $Var(Z^2) = 2$ 4%

Solution: B or D

First we compute the mean of $|Z|$. We have $E[|Z|] = \int_{-\infty}^{\infty} |z| f(z) dz = 2 \times \frac{1}{\sqrt{2\pi}} \int_0^{\infty} z e^{-\frac{z^2}{2}} dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-w} dw = \sqrt{\frac{2}{\pi}}$.

Next let $V = Z^2$. We compute its mgf: $M(t) = \int_{-\infty}^{\infty} e^{tz^2} f(z) dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz^2} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{1}{2}-t)z^2} dz = \frac{1}{\sqrt{1-2t}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2} dw = \frac{1}{\sqrt{1-2t}}$ for $t \leq 0.5$. Taking derivatives we have $E[V] = M'(0) = 1$ and $E[V^2] = M''(0) = 3$. Hence $Var(V) = 2$.

Teaching Objective: mean and variance derivation from mgf. Difficulty Level: Moderate.

13. A man first claimed that in tossing two unfair coins at once, we have only three possible outcomes: "two heads," "one head," and "no heads." Secondly, this man also claimed that each outcome has the same probability 1/3.

(i) {two heads, one head, no heads} are legitimate sample spaces.

(ii) If two unfair coins have different probabilities of heads, then it is possible to make the second claim true.

Determine which of the following four answers is correct where T stands for true and F stands for false.

A. (T,T) 26%
B. (T,F) 52%
C. (F,T) 10%
D. (F,F) 12%

Solution: (B)

(i) This is a legitimate sample space.

(ii) Suppose that the probability of heads for two coins is p and q, respectively. If the second claim is true, then we have

$$p * q = (1-p) * (1-q) = p * (1-q) + q * (1-p) = \frac{1}{3}$$

where the first equality leads to $p = 1 - q$, and thus the second equality leads to

$$(1-p) * p = p^2 + (1-p)^2$$

Thus, we obtain $3p^2 - 3p + 1 = 0$, which does not have a real solution.

Teaching Objective: sample space, random variables. Difficulty Level: Fundamental.

14. Let X be an exponential distribution with mean $\theta = 1/(3^{\frac{1}{3}})$. Find $Var(X^3)$.
    A. 2 10%
    B. 76 51%
    C. 80 19%
    D. 84 19%

Solution: (B)

We first find the moment generating function of X as follows:

$$\begin{aligned} M(t) = E[e^{tX}] &= \int_0^\infty e^{tX} \lambda e^{-\lambda X} dX \\ &= \lambda \int_0^\infty e^{(t-\lambda)X} dX = \frac{\lambda}{\lambda - t} \end{aligned}$$

Then we have

$E[X^3] = M^{(3)}(0) = \lambda * ((\lambda - t)^{-1})^{(3)}|_{t=0} = \lambda * 3 * 2 * 1 * (\lambda - t)^{-4}|_{t=0} = 3!\lambda^{-3}$

$E[X^6] = M^{(6)}(0) = \lambda * ((\lambda - t)^{-1})^{(6)}|_{t=0} = \lambda * 6! * (\lambda - t)^{-7}|_{t=0} = 6!\lambda^{-6}$

This leads to

$$Var[X^3] = E[X^6] - (E[X^3])^2 = 6!\lambda^{-6} - (3!\lambda^{-3})^2 = 684\lambda^{-6} = 76$$

Alternative method: On the one hand, we have

$$E[e^{tX}] = \frac{1}{1 - t/\lambda} = \sum_{k=0}^{\infty} (\frac{t}{\lambda})^k$$

On the other hand, we exploit the expression

$$E[e^{tX}] = \sum_{k=0}^{\infty} \frac{E[X^k]}{k!} t^k$$

This equivalence leads to $E[X^k] = k!\lambda^{-k}$.

Teaching Objective: moment generation function, exponential distribution. Difficulty Level: Hard.

15. 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 on average 2 requests are received per hour. What is the probability that during a given TWO-hour period, exactly 4 of the 5 operators receive no requests?

    A. $5(e^{-8} - e^{-10})$ 6% .

    B. $4(e^{-16} - e^{-20})$ 8% .

    C. $5(e^{-10} - e^{-16})$ 4% .

    D. $5(e^{-16} - e^{-20})$ 83% .

Solution: (D)
Let X be the number of requests per 2 hours. Then $X \sim \text{Poisson}(4)$. The probability that an operator gets no requests within an hour is $P(X = 0) = \frac{4^0 e^{-4}}{0!} = e^{-4}$. Let Y be the number of operators that receive no requests in a two-hour period. Then $Y \sim b(5, e^{-4})$ and hence the probability that 4 out of 5 operators receive no request is $P(Y = 4) = \binom{5}{4}(e^{-4})^4(1 - e^{-4})^{5-4} = 5(e^{-16} - e^{-20})$.

Teaching Objective: Poisson distribution, exponential distribution. Difficulty Level: Moderate

16. Let X be a discrete random variable with probability mass function

$$p(x) = \frac{x+1}{\lambda+1}\frac{\lambda^x}{x!}e^{-\lambda}, x = 0, 1, 2\cdots$$

where $\lambda > 0$. This distribution is a "tilted Poisson distribution" that has its mass pushed to the right and thus has heavier tails than the standard Poisson distribution. Find the mean of X.

A. $\frac{2\lambda}{\lambda+1}$ 9%
B. $\frac{\lambda^2+2\lambda}{\lambda+1}$ 48%
C. $\frac{2\lambda^2+\lambda}{\lambda+1}$ 20%
D. $\frac{\lambda^2+\lambda+1}{\lambda+1}$ 22%

Solution: (B)
This pmf can be regarded as the sum of two pmf

$$p(x) = \frac{1}{\lambda+1}\frac{\lambda^x}{x!}e^{-\lambda} + \frac{1}{\lambda+1}\frac{x\lambda^x}{x!}e^{-\lambda}$$

Since the mean and variance of Poisson distribution are both λ , we have

$$\sum_{x=0}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} = \lambda,$$
$$\sum_{x=0}^{\infty} x^2\frac{\lambda^x}{x!}e^{-\lambda} = \lambda + \lambda^2,$$

This leads to

$$E[X] = \frac{1}{\lambda+1}\sum_{x=0}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} + \frac{1}{\lambda+1}\sum_{x=0}^{\infty} x^2\frac{\lambda^x}{x!}e^{-\lambda} = \frac{\lambda^2+2\lambda}{\lambda+1}$$

Teaching Objective: Poisson distribution. Difficulty Level: Hard

17. Customers arrive at a travel agency according to a Poisson process at a rate of 10 per hour. What is the probability that less than 4 customers arrive in 2 hours?

A. $\int_2^{\infty} \frac{10^3}{2}x^2e^{-10x}dx$ 11%
B. $\int_2^{\infty} \frac{20^3}{2}x^2e^{-20x}dx$ 20%
C. $\int_2^{\infty} \frac{10^4}{6}x^3e^{-10x}dx$ 35%
D. $\int_2^{\infty} \frac{20^4}{6}x^3e^{-20x}dx$ 34%

Solution: (C)
The event A {less than 4 (or at most 3) customer arrive in 2 hours} is equivalent to the event B {the 4th customer arrives after 2 hours}. Assume $\bar{B}$ which means the 4th customer arrives in 2 hours, then we must have $\bar{A}$ which means at least 4 costumers arrives in 2 hours and hence B is necessary for A. Assume B, then obviously we must have A and hence B is also sufficient for A. We can safely conclude that A and B are equivalent. Note that when computing with the Poisson distribution, one should have $\lambda = 10 \times 2 = 20$ as the expected number of occurrence in given 2 hours. When computing with the Gamma distribution, $\lambda = 10$ since it is the expected number of occurrence in a unit of time.

Compute with the Gamma distribution: $\int_2^\infty \frac{10^4}{6} x^3 e^{-10x} dx$.

Teaching Objective: Gamma distribution, Poisson. Difficulty Level: Moderate

18. Consider the following 4 statements:

(i) If X follows a Gamma distribution with parameter $\alpha = 2$ and $\theta = 1$, then $E[X^4] = 120$
(ii) Consider the Gamma function with a positive integer α, then we could write it as $\Gamma(\alpha) = (\alpha - 1)!$
(iii) If $X \sim N(3, 1)$, then $E[X^3]$=27
(iv) If X follows a Chi-square distribution with 1 degree of freedom, then $P(X \geq 2.706) = 0.9$

A. All statements are true. 11%
B. Only (i), (ii) and (iii) are true. 18%
C. Only (iii) and (iv) are true. 6%
D. Only (i) and (ii) are true. 65%

Solution: (D)
(i). Using the moment generating function of gamma distribution, we have $E[X^4]$=$\alpha(\alpha+1)(\alpha+2)(\alpha+3)\theta^4 = 120$.
(ii). This a property of Gamma function and it can be proved by

using integration by part.
(iii). Let $X = Z + 3$, where $Z \sim N(0, 1)$. Notice that $E[Z] = 0$, $E[Z^2] = 1$ and $E[Z^3] = 0$, we have

$$\mathbb{E}[X^3] = \mathbb{E}[Z^3 + 9Z^2 + 27Z + 27] = 36$$

(iv). This can be checked through the table. The correct one is $P(X \leq 2.706) = 0.9$
Therefore, only (i) and (ii) are true.

Teaching Objective: Chi-square, Gamma distribution Difficulty Level: Moderate

19. Let $S = \{1, 2, 3, 4\}$ and consider the events $A = \{1, 4\}$, $B = \{2, 4\}$ and $C = \{3, 4\}$.

    (i) Suppose $P(\{i\}) = 1/4$, $i = 1, 2, 3, 4$. Are A and B independent?
    (ii) Are A, B, and C are pair-wise independent?
    (iii) Compute $P(A \cap B \cap C)$ and $P(A)P(B)P(C)$. Are these two quantities equal?
    (iv) Suppose $P(\{i\}) = i/10$, $i = 1, 2, 3, 4$. Are A and B independent?

    Determine which of the following four answers is correct where T stands for true and F stands for false.

    (A) (T, T, T, F) 10%
    (B) (T, T, F, F) 66%
    (C) (T, F, F, F) 6%
    (D) (F, F, F, F) 18%

Solution: B
For part (i), we have $P(A) = P(B) = P(C) = 0.5$, $P(A \cap B) = P(\{4\}) = 0.25 = P(A)P(B)$ so true. For part (ii), we can similarly verify that $P(A \cap C) = P(\{4\}) = 0.25 = P(A)P(C)$ and $P(B \cap C) = P(\{4\}) = 0.25 = P(B)P(C)$ so true. For part (iii) we have

$P(A \cap B \cap C) = P(\{1\}) = 0.25$ but $P(A)P(B)P(C) = 0.125$ so false, For part (iv), $P(A) = 4/100$, $P(B) = 8/100$ while $P(A \cap B) = P(\{4\}) = 4/10 \neq P(A)P(B)$ so false.

Teaching Objective: Independent Events. Difficulty Level: Fundamental.

20. Suppose that A and B are two events such that both $P(A)$ and $P(B)$ are in $(0, 1)$, and consider $P(A|B)$.

    (i) Is $P(A|B) = 0$ when A and B are mutually exclusive?

    (ii) Is $P(A|B) = P(A)$ when A and B are independent?

    (iii) Is $P(A|B) = 1$ if $B \subset A$?

    (iv) Is $P(A|B) \leq P(A)$ when $A \subset B$?

    Determine which of the following four answers is correct where T stands for true and F stands for false.

    (A) (T, T, T, T) 18%

    (B) (T, F, F, T) 4%

    (C) (T, T, T, F) 76%

    (D) (F, F, F, F) 2%

Solution: C
For part (i) since $A \cap B = \emptyset$ it follows that $P(A|B) = P(\emptyset)/P(B) = P(\emptyset) = 0$ so true. For part (ii), by independence, $P(A \cap B) = P(A)P(B)$ so $P(A|B) = P(A)P(B)/P(B) = P(A)$ so true. For part (iii), since $A \cap B = B$, we have $P(A|B) = P(B)/P(B) = 1$ so true. For part (iv), since $A \cap B = A$ we have $P(A|B) = P(A)/P(B) > P(A)$ so false.

Teaching Objective: Conditional Probability. Difficulty Level: Fundamental.

21. Let X_1 and X_2 be independent Bernoulli random variables with success probability p.

 (i) Is the probability of $P(X_1 + X_2 = 2|X_1 = 1) = p$?

 (ii) Is the probability of $P(X_1 + X_2 = 2|X_1 + X_2 \geq 1) = p/(2 - p)$?

 (iii) Alice believes that the two formulas above are correct and claims that families with two children are more likely to have two girls if the first born is a girl compared to families with two children where we only know that at least one of them is a girl. Is Alice correct in her claim under the further assumption that all the births are independent events with the same p, where $p \in (0, 1)$ is the probability of a girl?

 Determine which of the following four answers is correct where T stands for true and F stands for false.

 (A) (F, F, T) 2%

 (B) (T, F, T) 15%

 (C) (T, T, F) 22%

 (D) (T, T, T) 61%

 Solution: D
 $P(X_1+X_2 = 2|X_1 = 1) = P(X_1 = 1, X_2 = 1)/P(X_1 = 1) = p^2/p = p$ so (i) is true. Also, $P(X_1 + X_2 = 2|X_1 + X_2 \geq 1) = P(X_1 = 1, X_2 = 1)/(1 - P(X_1 + X_1 = 0)) = p^2/(1 - (1 - p)^2) = p/(2 - p)$, so part (ii) is true. Since $p > p/(2 - p)$ for all p such that $2 - p > 1$ implies that $p < 1$ we see that the claim is correct.

 Teaching Objective: Bernoulli Random Variables, Conditional Probability. Difficulty Level: Moderate.

22. Consider a Bernoulli process with success probability p and define $q = 1 - p$. Let X be the number of Bernoulli trials until the rth success and let N_k be a Binomial random variable with parameters $n = k$ and p.

 (i) Is $P(X > k) = P(N_k \leq r - 1)$?

(ii) For $r = 1$ would this yield $P(X > k) = q^k$ for all $k \in \{1, 2, \ldots\}$?

(iii) Is it true that for $r = 1$, $P(\{X > k + l\}|\{X > k\}) = P(X > l)$ for all positive integers l?

(iv) For arbitrary r is it $P(X = k) = P(N_k \le r-1) - P(N_{k-1} \le r-1)$ correct for all $k \in \{r, r+1, \ldots, \}$?

Determine which of the following four answers is correct where T stands for true and F stands for false.

(A) (T, T, T, T) 32%

(B) (T, F, F, T) 19%

(C) (T, T, T, F) 45%

(D) (F, F, F, F) 3%

Solution: C
(i) holds because the events $X > k$ and $N_k \le r - 1$ are equivalent. Part (ii) follows from (i) since $P(N_k \le 0) = q^k$. Part (iii) follows since the intersection of the events $\{X > k + l\}$ and $\{X > k\}$ is $\{X > k + l\}$, so $P(\{X > k + l\}|\{X > k\}) = P(\{X > k + l\})/P(\{X > k\}) = q^{k+l}/q^k = q^l = P(X > l)$. Part (iv) is false since the correct answer is $P(X = k) = P(N_{k-1} \le r-1) - P(N_k \le r-1)$.

Teaching objective: Geometric and Negative Binomial distributions and conditional distribution. Difficulty Level: Moderate.

23. Suppose that $X_1, X_2, \ldots, X_n$ are uniform $[0, 1]$ random variables, events associated with X_i, $i = 1, \cdots, n$ are mutually independent, and let $Y_n = \min(X_1, \ldots, X_n)$.

(i) Is $P(Y_n > y) = (1 - y)^n$?

(ii) Is $P(Y_n > z/n) = (1 - z/n)^n$?

(iii) Is $\lim_{n\to\infty} P(nY_n > z) = \exp(-z)$?

(iv) Is $\lim_{n\to\infty} E[nY_n] < 1$?

Determine which of the following four answers is correct where T stands for true and F stands for false.

(A) (T, T, T, T) 31%

(B) (T, T, T, F) 47%

(C) (T, F, F, T) 17%

(D) (F, F, F, F) 4%

Solution: B
(i) holds because the events $Y_n > y$ is equivalent to the event $X_i > y$ for all $i \in \{1, \ldots, n\}$, $P(X_i > y) = 1 - y$ and from independence. (ii) holds by applying (i) to $y = z/n$, (iii) standard argument for e^{-z}, (iv) is false since nY_n is in the limit an exponential random variable with parameter 1 and has expectation equal to 1.

Teaching objective: Exponential, Uniform, Independence. Difficulty Level: Moderate.

24. Let X be a random variable and $\hat{X} = a + bX$ for some constants a and b. Let π_p be the $100p$ percentile of X and $\hat{\pi}_p$ be the $100p$ percentile of $\hat{X}$.

(i) Suppose X is exponential with mean 1. Is $\pi_p = \ln(1 - p)$?

(ii) Suppose now that X is normal with mean zero and variance one. Is $\pi_p = \Phi^{-1}(p)$?

(iii) Is $\hat{\pi}_p = a + b\pi_b$ for all reals a and b when $b > 0$?

(iv) Suppose $a = 0$ and $b = \theta$, and X is exponential with mean one, is $\hat{X}$ exponential with mean θ?

(v) Suppose that X is normal, $a = \mu$ and $b = \sigma$, is $\hat{X}$ normal with mean μ and variance σ?

Determine which of the following four answers is correct where T stands for true and F stands for false.

(A) (T, T, T, T, T)

(B) (T, T, T, F, T)

(C) (T, T, T, T, F)

(D) (F, T, F, T, T)

Solution: A,B,C,D
(i) and (ii) follow from the definition of π_p and simple algebra. (iii) holds because $P(a + bX \leq a + b\pi_p) = P(X \leq \pi_p) = \pi_b$ for all a and all $b > 0$, (iv) follows because $P(\theta X > x) = P(X > x/\theta) = \exp(-x/\theta)$ which coincides with the tail of an exponential with mean θ. (v) is false because $\hat{X}$ is normal with mean μ and variance σ^2.

Teaching objective: Exponential, Normal, Percentiles Difficulty Level: Moderate.

II Regular Questions (28 points)

25. (14 points) The life of a certain type of automobile tire is normally distributed with mean 34,000 miles and standard deviation 4000 miles.

    (a) (3 points) What is the probability that such a tire lasts over 40,000 miles?

    (b) (3 points) What is the probability that it lasts between 30,000 and 35,000 miles?

    (c) (4 points) Given that it has survived 30,000 miles, what is the conditional probability that the tire survives another 10,000 miles?

    (d) (4 points) Suppose the lifetime follows exponential distribution with mean 34,000 miles. Answer the previous question (c) again.

Solution:

Define random variable X that marks the life time of a randomly selected automobile tyre. We are given that $X \sim \mathcal{N}(34000, 4000^2)$.

(a)

We are required to find

$$P(X > 40000) = 1 - P(X \le 40000) = 1 - P\left(\frac{X - 34000}{4000} \le \frac{40000 - 34000}{4000}\right)$$
$$= 1 - \Phi(1.5) = 1 - 0.93319 = 0.06681$$

(b)

We are required to find

$$P(30000 \le X \le 35000) = P\left(\frac{30000 - 34000}{4000} \le \frac{X - 34000}{4000} \le \frac{35000 - 34000}{4000}\right)$$
$$= P\left(-1 \le \frac{X - 34000}{4000} \le 0.25\right) = \Phi(0.25) - \Phi(-1)$$
$$= 0.59871 - 0.15866 = 0.44$$

(c)

We are required to find

$$P(X \ge 40000 | X \ge 30000) = \frac{P(X \ge 40000, X \ge 30000)}{P(X \ge 30000)} = \frac{P(X \ge 40000)}{P(X \ge 30000)}$$

The expression above is equal to

$$\frac{P(X \ge 40000)}{P(X \ge 30000)} = \frac{1 - P\left(\frac{X-34000}{4000} \le \frac{40000-34000}{4000}\right)}{1 - P\left(\frac{X-34000}{4000} \le \frac{30000-34000}{4000}\right)} = \frac{1 - \Phi(1.5)}{1 - \Phi(-1)}$$
$$= \frac{1 - 0.93319}{1 - 0.15866} = 0.0794$$

(d) We follow the same logic of (c) to get $P(X \ge 40000 | X \ge 30000) = P(X \ge 40000)/P(X \ge 30000) = P(X \ge 10000) = e^{-10/34}$.

26. (14 points)

(a) (8 points) Let X be a continuous random variable with its probability density function described by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-0.5(x-\mu)^2/\sigma^2}, x \in (-\infty, \infty). \quad (1)$$

Prove that

1. (4 points) its moment generating function $M_X(t)$ is equal to

$$M_X(t) = e^{\mu t + 0.5\sigma^2 t^2}, t \in (-\infty, \infty). \tag{2}$$

2. (2 points) $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.
3. (2 points) for any real constants a, b, $a + bX \sim N(a + b\mu, b^2\sigma^2)$.

(b) (6 points) Let Y be a continuous random variable whose moment generating function $M_Y(t)$ described by

$$M_Y(t) = e^{0.5t^2}, t \in (-\infty, \infty). \tag{3}$$

Prove that

1. (1 points) let $F(y)$ denote the cumulative distribution function of Y, prove that $F(-z_\alpha) = 1 - F(z_\alpha)$, where $\alpha \in (0, 1)$ is a constant, and z_α is the upper 100α percent point.
2. (5 points) Y^2 has a χ^2 distribution with degrees of freedom 1.

Solution: For part a) and part b), please refer to the lectures 11 and 12, except the third question in part a), whose proof is given below.

Let $Z = a + bX$. Then one way is to show that the mgf of Z is the same as the mgf of $N(a + b\mu, b^2\sigma^2)$.

The mgf of Z takes the form of

$$M_Z(t) = \mathrm{E}(e^{tZ}) = \mathrm{E}(e^{ta+tbX}) = e^{ta}\mathrm{E}(e^{tbX}) \tag{4}$$

$$= e^{ta}M_X(tb) = e^{ta}e^{\mu tb + 0.5\sigma^2 b^2 t^2} = e^{(a+\mu b)t + 0.5\sigma^2 b^2 t^2} \tag{5}$$

which indicates that $a + bX \sim N(a + b\mu, b^2\sigma^2)$.