### **Probability**

**1.1.** You have a bag with two coins. One will come up heads 40% of the time and the other will come up heads 60%. You pick a coin randomly, flip it and get a head. What is the probability it will be heads on the next flip?

Tag: Bayes Formula

Solution: Notations of events:

H – Get a head;

 $E_i$  – Picked coin i;

H' – Get a head next time.

From Bayes formula, we have:

$$P(E_1|H) = \frac{P(H|E_1)P(E_1)}{P(H|E_2)P(E_2) + P(H|E_1)P(E_1)} = \frac{0.4 * 0.5}{0.4 * 0.5 + 0.6 * 0.5} = 0.4$$

$$P(H'|H) = P(H'|E_1, H)P(E_1|H) + P(H'|E_2, H)P(E_2|H)$$
  
= 0.4 \* 0.4 + 0.6 \* 0.6 = 0.52

**1.2.** What is the Law of Large Numbers and The Central Limit Theorem? What are the differences between them?

Solution: The Law of Large Numbers:

For a infinite sequence of i.i.d Lebesgue integrable random variables  $\{X_1, X_2, \cdots\}$  with  $E(X_i) = \mu$ . Then the sample mean  $\bar{X}_n := \frac{X_1 + \cdots + X_n}{n}$  converge in probability to  $\mu$ . i.e. for any small  $\epsilon > 0$ ,

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| > \epsilon) = 0$$

The Central Limit Theorem:

For a infinite sequence of i.i.d random variables  $\{X_1, X_2, \cdots\}$  with  $E(X_i) = \mu$ 

and  $Var(X_i) = \sigma^2$ . Denote the sample mean by  $\bar{X_n} := \frac{X_1 + \dots + X_n}{n}$ . Then the random variable  $\frac{\bar{X_n} - \mu}{\sigma/\sqrt{n}}$  converge in distribution to N(0,1), i.e. the cumulative distribution function  $F_n$  converge to the normal cumulative distribution  $\Phi$ . The law of large number tells us that the sample mean is an unbiased estimator of expected value while the central limit theorem tells us about the asymptotic distribution of the same mean, concluding that it is also consistent.

**1.3.** Suppose you have a fair coin, and you flip it a million times. Estimate the probability that you get fewer than 499,000 heads.

Tag: Central Limit Theorem, Normal Distribution

Solution: From the CLT, the outcome distribution is closed to N(0.5M, 0.25M). The standard deviation is 0.5k = 500. So  $P(N < 499,000) = \Phi(\frac{499,000-500,000}{500}) = \Phi(-2) = 0.02275$ . (The number is checked from the table, should be acceptable to use  $\Phi(-2)$  as the final answer.)

1.4 (Unfair coin). In front of you is a jar of 1000 coins. One of the coins has two heads, and the rest are fair coins. You choose a coin at random, and flip it ten times, getting all heads. What is the probability it is one of the fair coins?

Tag: Bayes Formula

Solution: Notation of events:

F – Pick a fair coin;

S – Pick a two-head coin;

H – Flip 10 times, all heads.

$$\begin{split} P(F|H) &= \frac{P(H|F)P(F)}{P(H|F)P(F) + P(H|S)P(S)} \\ &= 0.5^{10}*0.999/(0.5^{10}*0.999 + 1*0.001) \approx 0.5 \end{split}$$

**1.5.** Starting at one vertex of a cube, and moving randomly from vertex to adjacent vertices, what is the expected number of moves until you reach the vertex opposite from your starting point?

П

Tag: Markov Process

Solution: Notation of states: (By the distance to the origin)

 $S_0$  – the origin;

 $S_1$  – the vertices next to origin;

 $S_2$  – the vertices next to opposite;

 $S_3$  – the opposite.

Denote  $E(S_i)$  to be the expected moves to get to the opposite. Suppose we are in  $S_0$ , then after one move, we are in  $S_1$ . Now suppose we are in  $S_1$ , then we have 1/3 chance we will move back to  $S_0$  and 2/3 chances to move to  $S_2$ .

2

Similar for analysis on  $S_2$ . Finally if we are in  $S_3$ , then  $E(S_3) = 0$ . Therefor, we have the equations:

$$E(S_0) = 1 + E(S_1)$$

$$E(S_1) = 1 + \frac{1}{3}E(S_0) + \frac{2}{3}E(S_2)$$

$$E(S_2) = 1 + \frac{1}{3}E(S_3) + \frac{2}{3}E(S_1)$$

$$E(S_3) = 0$$

Solve and get  $E(S_0) = 10$ ,  $E(S_1) = 9$ ,  $E(S_2) = 7$ .

**1.6.** Give an example of random variables that are normal, uncorrelated, and dependent.

Tag: Normal Distribution

Solution: Denote  $X \sim N(0,1)$  and Z has equal chance to be 1 or -1. Then Y = ZX. We can show that Y is normal, Cov(X,Y) = 0 and Y is dependent on X (simple from  $f(0,0) = p(0) \neq p(0)^2$ , where f is the joint density for (X,Y) and p if density function for normal distribution.)

**1.7.**  $X \sim N(\mu_X, \sigma_X^2)$  and  $Y \sim N(\mu_Y, \sigma_Y^2)$  are independent, and you know X+Y=s. What is the expected value of X?

Tag: Normal Distribution

Solution: We need to use the following lemma:

If (X, Y) are bivariate normal distributed, then  $(a_1X + b_1Y, a_2X + b_2Y)$  are also bivariate normal for any real coefficients  $a_1$ ,  $a_2$ ,  $b_1$ ,  $b_2$ . Moreover, if X, Y are uncorrelated, i.e. Cov(X, Y) = 0, then they are independent.

(It is unlikely you need to prove that. It may be given as hint in some cases) From the lemma, we can deduce the orthogonal decomposition of X as:

$$X = (X - \frac{Cov(X,Y)}{Var(X)}Y) + \frac{Cov(X,Y)}{Var(X)}Y$$

For this problem, let's consider (X,Z:=X+Y), which is bivariate normal from the lemma. We can see that  $E(Z)=\mu_X+\mu_Y$  and  $Var(Z)=\sigma_X^2+\sigma_Y^2$ ,

 $Cov(X,Z) = Cov(X,X) + Cov(X,Y) = \sigma_X^2$ . We need to calculate E(X|Z=s).

$$\begin{split} E(X|Z=s) &= E(X - \frac{Cov(X,Z)}{Var(X)}Z) + \frac{Cov(X,Z)}{Var(X)}Z|Z=s) \\ &= E(X - \frac{Cov(X,Z)}{Var(X)}Z|Z=s) + E(\frac{Cov(X,Z)}{Var(X)}Z|Z=s) \\ &= E(X - \frac{Cov(X,Z)}{Var(X)}Z) + \frac{Cov(X,Z)}{Var(X)}s \\ &= E(X - \frac{\sigma_X^2}{\sigma_X^2}(\mu_X + \mu_Y) + \frac{\sigma_X^2}{\sigma_X^2}s \\ &= s - \mu_Y \end{split}$$

**1.8.** You have a spinner that generates random numbers that are uniform between 0 and 1. You sum the spins until the sum is greater than one. What is the expected number of spins?

Solution: We first need to reorganize the formula for expectation:

$$E(n) = P(n = 1) + 2P(n = 2) + 3P(n = 3) + \cdots$$
  
=  $P(n > 1) + P(n > 2) + P(n > 3)$ 

Then we need to find the formula for  $P(n \geq i)$  for all  $i \in N$ . The event  $n \geq k+1$  is equivalent to  $X_1 + X_2 + \cdots + X_k \leq 1$ , where  $X_i$  are i.i.d. uniform on [0,1].  $(X_1, X_2, \cdots, X_k)$  are uniform in the area  $[0,1]^k$  and we need to calculate

the volume bounded by  $X_i \ge 0$  and  $\sum_{i=1}^k X_i \le 1$ .

$$V = \int_{0}^{1} \int_{0}^{1-x_{1}} \cdots \int_{0}^{1-(x_{1}+x_{2}+\cdots x_{n})} 1 dx_{n+1} dx_{n} \cdots dx_{1}$$

$$= \int_{0}^{1} \int_{0}^{1-x_{1}} \cdots \int_{0}^{1-(x_{1}+\cdots +x_{n-1})} 1 - (x_{1}+\cdots x_{n}) dx_{n} \cdots dx_{1}$$

$$= \int_{0}^{1} \int_{0}^{1-x_{1}} \cdots \int_{0}^{1-(x_{1}+\cdots +x_{n-1})} y dy dx_{1}$$

$$(y = 1 - (x_{1}+\cdots +x_{n}))$$

$$= \int_{0}^{1} \int_{0}^{1-x_{1}} \cdots \int_{0}^{1-(x_{1}+\cdots +x_{n-2})} \frac{1}{2} (1 - (x_{1}+\cdots +x_{n-1}))^{2} dx_{n-1} dx_{1}$$

$$(similar\ trick\ apply)$$

$$= \int_{0}^{1} \int_{0}^{1-x_{1}} \cdots \int_{0}^{1-(x_{1}+\cdots +x_{n-3})} \frac{1}{3!} (1 - (x_{1}+\cdots +x_{n-2}))^{3} dx_{n-2} dx_{1}$$

$$= \int_{0}^{1} \frac{1}{n!} y^{n} dy = \frac{1}{(n+1)!}$$

So our final result is  $E(n) = \sum_{k=0}^{\infty} \frac{1}{k!} = e$ .

(Remember the Taylor expansion of 
$$e^x = \sum_{n=1}^{\infty} \frac{x^n}{n!}$$
.)

**1.9.** A stick is broken randomly into 3 pieces. Two broken points are uniformly distributed on [0,1]. What is the probability of three pieces being able to form a triangle?

Solution: The whole probability space is  $[0,1] \times [0,1]$ , which has area 1. Denote the event that three pieces are able to form a triangle by E. Then the probability of E is just the area of E.

Let's determine the tuples (x,y) that belong to E. If x < y, then forming a triangle need to satisfy three conditions: x < (y-x)+(1-y), (y-x) < x+(1-y), (1-y) < x+(y-x), which is equivalent to  $x < \frac{1}{2}, y-x < \frac{1}{2}, y>\frac{1}{2}$ . The space by these three constraints is a right triangle above the diagonal. (Image in your head or draw it down on paper.) The area is  $\frac{1}{8}$ . Similar analysis apply to the situation x > y and the space is just the reflection from the line y = x, (as you just exchange x with y!) so it also has area  $\frac{1}{8}$ . So the total probability is  $\frac{1}{4}$ .

**1.10.** A stick is broken randomly into two pieces, i.e. the broken point is uniform random variable on [0,1]. The larger piece is then broken randomly into two pieces, i.e. the second broken points are uniform random variable on [0,X], where X is the length of the longer piece. What is the probability of the pieces being able to form a triangle?

Solution: It is obviously a variation of the previous problem, but the problem here is what is changed and what is not. A little thought may lead to the conclusion that the space of E is the same but the total probability space has change. Given  $x=\frac{1}{3}$ , y are just random number from  $\frac{1}{3}$  to 1. Now the total probability space is  $[0,1]\times[0,1]$  removing [0,x] for  $x\in[0,\frac{1}{2})$  and removing [x,1] for  $x\in[\frac{1}{2},1]$ . The area of the total probability space is  $\frac{3}{4}$ , so  $P(E)=\frac{1/4}{3/4}=1/3$ .

**1.11.** You play a game where you toss two fair coins in the air. You always win 1 dollar. However, if you have tossed 2 heads at least once, and 2 tails at least once, you surrender all winnings, and cannot play again. You may stop playing at anytime. What's your strategy?

Solution: If you have never tossed 2 heads or 2 tails, then next toss will guarantee you 1 dollar. After your first toss of 2 heads or 2 tails, then the next toss will have  $\frac{3}{4}$  chance of 1 dollar and  $\frac{1}{4}$  chance of -n dollar, where n is the dollars already wined. Just continue if expected value is greater than 0, i.e. n < 3, otherwise just stop.

1.12 (St. Petersburg Paradox). Consider the following game played by flipping a fair coin. The pot begins at a 1 dollar, and the pot doubles until a tail is flipped, at which point you receive the pot. Assume you can play as many times as you want. What would you pay to play this game?

Tag: Default Risk

Solution: The price should be equal to the expected payoff as it can be played many times. It can be deduced that the probability of getting  $2^{i-1}$  dollars is  $\frac{1}{2^i}$ . So  $E(X) = \sum_{i=1}^{\infty} 2^{i-1} \times \frac{1}{2^i} = \sum_{i=1}^{\infty} 1/2 = \infty$ . So that is the theoretical answer, which seems a bit unrealistic.

Assuming in the real world, a bank is try to pricing a product like this, it is obviously no way charge the clients infinitely dollars. So what can be done to make it more realistic? A possible idea here is to consider the default risk of the game host. Assume the total available money of the host is  $2^30$  dollar, then if you toss 31 head, the host will default and cannot pay you anything. Given such upper bound, the game may be priced with finite amount.

1.13 (Monte Hall Problem). You are on a game show, and there are 3 doors. Two of the doors conceal nothing, and one door conceals a prize. The game show host, Monte Hall, knows where the prize is. He lets you pick a door. You picked, say, door 1. If you pick the door of prize, he will randomly open one remaining door. If you pick the door of nothing, then he opens the other door of nothing. Suppose he open the door 2. Finally he then offers you the chance to switch doors. What should you do? Please justify your action. (If you have completely no idea, you may just say "Toss a coin and let the god guide me!". LOL)

Tag: Bayes formula

Solution: Notations of events:  $D_i$  – the price are inside door i;  $O_2$  – the host open the door 2.

$$P(D_1|O_2) = \frac{P(O_2|D_1)P(D_1)}{P(O_2|D_1)P(D_1) + P(O_2|D_2)P(D_2) + P(O_2|D_3)P(D_3)}$$

$$= \frac{\frac{1}{2} * \frac{1}{3}}{\frac{1}{2} * \frac{1}{3} + 0 * \frac{1}{2} + 1 * \frac{1}{3}} = \frac{1}{3}$$

Similarly, we can have  $P(D_2|O_2) = 0$  and  $P(D_3|O_2) = \frac{2}{3}$ . So you should switch to door 3.

**1.14.** What is the expected number of rolls of a fair die needed to get all six numbers?

Solution: The problem may seems complicated at the first glance but once formulated in the correct way, it is relatively easy to solve. The right way to decompose the process is by getting a new number one by one until we get 6. The first roll will give a new number whichever the number it is. Given we have 1 number, getting the next number number is just a geometric process of success rate  $\frac{5}{6}$ . Similarly, given the condition that we have i numbers, getting a new number is a geometric process of success rate  $\frac{6-i}{6}$ , which has expected trials of  $\frac{6}{6-i}$ . So the total expected rolls is just  $1+\frac{6}{5}+\cdots+\frac{6}{1}=14.7$ .

1.15. You have a bucket of unfair coins. Each coin has a probability of getting heads, p, which is uniformly distributed between zero and one. You pick a coin, and flip it 64 times, getting 48 heads. What is the expected value of p for your coin?

Tag: Bayesian Inference.

Solution: It helps if you know the beta distribution. (If not, google it!)  $\Box$ 

**1.16.** A room of 100 people put their business cards in a hat, then each person randomly draws a business card. What's the expected number of people who draw their own business card?

Solution: At the first glance, it seems complicated as the event that A draw her own business card will affect the event that B draw B's card. (Because if A draw B's card, then B can not draw his own, but if not, then B has a chance to draw his own.) It will be hard to calculate directly by summing n times the probability of n people draw their own card.

Denote X to be the number of people who draw their own card and we need to calculate E(X). Denote  $X_i$  to be a random variable that equals to 1 if guest i get her own card else 0, (Note here that  $X_i$ ,  $X_j$  are not independent.) and we have  $X = \sum_{i=1}^{100} X_i$ . Even  $X_i$ 's are dependent on each other, we still have  $E(X) = \sum_{i=0}^{100} E(X_i) = 100 \times \frac{1}{100} = 1$ .

1.17. A red ant and a black ant are at opposite vertices of a cube. Each randomly picks an edge to traverse and moves to the next vertex. They continue this until they meet. What is the expected number of edges each ant traverses?	
Solution: $\Box$	
<b>1.18.</b> If you roll a die repeatedly, what is the expected number of rolls until you see consecutive sixes?	
Tag: Markov Process	
Solution: $\Box$	
<b>1.19.</b> Alex and Beth take turns flipping a pair of coins. The first person to flip a pair of heads wins the game. Alex flips first. Beth eventually wins. What is the probability she flipped a pair of heads on her second turn?	
Tag: Bayes Formula	
Solution: $\Box$	
<b>1.20</b> (Coin toss game). A flips a fair coin 11 times, and B flips the coin 10 times. How likely is it that A flipped more heads than B?	
Source: Practical Guide	
Solution: $\Box$	
1.21 (Card game). A casino offers a simple card game. There are 52 cards in a deck with 4 cards for each value 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A. Each time the cards are thoroughly shuffled (so each card has equal probability of being selected). You pick up a card from the deck and the dealer picks another one without replacement. If you have a larger number, you win; if the numbers are equal or yours is smaller, the house wins. What is your probability of winning?	
Source: Practical Guide	
Solution: $\Box$	
1.22 (Drunk passenger). A line of 100 airline passengers are waiting to board a plane. They each hold a ticket of one of the 100 seats on the flight. For convenience, the n-th passenger holds the ticket for seat n. Being drunk, the first person in line picks a random seat (equal probability for all seats). All other passengers are sober. They will go to their seats unless it is occupied. In such case, they will pick a free seat randomly (also equal probability for all free seats). Suppose you hold ticket for seat 100, what is the probability that you end up in your seat (i.e. seat 100)?	
Source: Practical Guide	
Solution: $\Box$	

<b>1.23</b> (N points on a circle). Given N points drawn randomly on the circumference of a circle, what is the probability that they are all within a semicircle?
Source: Practical Guide
Solution: $\Box$
1.1 Counting and Combinations
$\textbf{1.24} \ (\text{Poker hands}). \ \textit{What is the probability of getting four-of-a-kind/Full House/two pair?}$
Source: Practical Guide
Solution: $\Box$
1.25 (Hopping rabbit). A rabbit sits at the bottom of a staircase with 10 stairs. It can only hop up only one or two stairs at a time. How many ways for it to ascend to the top of stairs.
Source: Practical Guide
Solution: $\Box$
<b>1.26</b> (Grid Jumping). You are on the grid $(0,0)$ . You may move right or up by 1 grid each step. How many ways are there for you to move to $(10,10)$ ?
Solution: $\Box$
1.27 (Balls to boxes). You have 5 balls label 1 to 5 and 5 boxes label 1 to 5. The balls are randomly put into the boxes. What is the probability of all balls are put to a box with a different number?
Source: Practical Guide
Solution: $\Box$
<b>1.28</b> (Same birthday). What is the minimum number $n$ that, among a group of $n$ people, the probability of two people have the same birthday is greater than $1/2$ ? (Assume a year always has $365$ days)
Source: Practical Guide
Solution: $\Box$

#### 1.2 Conditional Probability

**1.29** (Order Dice). Roll 3 dice one by one, what is the probability that the outcome is strictly increasing?

Source: Practical Guide	
Solution: $\Box$	
1.30 (Population die out). In a planet with 1 creature, this creature will die, stay the same, split into two or split into three with equal chance (0.25 for each). Each creature behave the same and independently. What is the probability that the creature will die out?	
Source: Practical Guide	
Solution: $\Box$	
1.31 (Candies in a jar). Given a jar with 10 red candies, 20 blue candies, 30 green candies, what is the probability that the jar still have blue and green candies after you have out all 10 red candies? (You are taking without replacement.)	
Solution:	
1.32 (Russian roulette).	
Source: Practical Guide	
Solution: $\Box$	
<b>1.33</b> (Aces). A poker of 52 cards is distributed to 4 players with each getting 13 card randomly. What is the probability that each play has an ace?	
Source: Practical Guide	
Solution: $\Box$	
<b>1.34</b> (GRE scores). A student is taking GRE exams with 100 questions, each for 1 point. She is getting first question right and the second question wrong. In the following questions, the difficulty will adjust to her previous performance. For the i-th question, she has probability $\frac{k}{i-1}$ to get the correct answer if she has got k points in the previous $i-1$ questions. What is the probability that she ends up with 50 points after she completes all 100 questions?	
Source: Practical Guide	
Solution: $\Box$	
1.35 (Cars on road). Assume that it has probability p to observe at least one car on the road during a period of 20 minutes, what is the probability to observe at least one car on the road during a period of 5 minutes? You can assume that on a fix length period, the probability is fixed and on two non-overlap periods,	

the events are independent.

Source: Practical Guide
Solution: $\Box$
<b>1.36</b> (Meeting Probability). A and B will arrive at the park at a random time from 17:00 to 18:00 and stay for 5 minutes. What is the probability that A meets $B$ ?
Source: Practical Guide
Solution: $\Box$
<b>1.37</b> (Moments of normal distribution). If $X \sim N(0,1)$ , then what is $E[X^n]$ ? Calculate the values for $n = 1, 2, 3, 4$ .
Solution: $\Box$
1.3 Expectation and Variance
1.38 (Connecting noodles). You have 100 noodles and 200 ends. Each time, you randomly choose two ends and connect them. The process last until no end left. What is the expected number of circles at the end of the process?
Source: Practical Guide
Solution: $\Box$
1.39 (Dice game). Consider a game that you roll a dice at get the face value. If you roll a 6, you can roll again. What is the fair price to enter the game?
Source: Practical Guide
Solution: $\Box$
1.40 (First ace). Each time, you draw a card from a regular 52 poker without replacement. What is the expected number of draws to get the first ace?
Source: Practical Guide
Solution: $\Box$
<b>1.41</b> (N balls in a jar). Assuming that there are N distinct balls in a jar and each draw will be replaced with the same ball. What is the expected number of draws to get a complete set of N balls? After $k$ draws, what is the expected numbers of distinct balls?
Source: Practical Guide
Solution: $\Box$

1.42 (Range of default). Assume A has default probability 50% and B has default probability 30% but the correlation is unknown. What is the range of probability that A or B default? What is the range of correlation of the event A defaults and the events B defaults?

Source: Practical Guide		
Solution:		
1.43 (Random ants). 500 ants are randomly placed on 1-foot string and move to one random end of the string at speed 1 foot/min. If two ants collide heads on, they just immediately turn around and keep on moving. The process will last until all ants fall off the string. What is the expected time length of this process?		
Source: Practical Guide		
Solution:		
1.44.		
Solution:		
1.45.		
Solution:		
1.46.		
Solution:		
1.47.		
Solution:		
1.48.		
Solution:		
1.49.		
Solution:		
1.50.		
Solution:		
1.51.		
Solution:		

**1.52** (Gambler's Ruin). A gambler start with i dollars and has probability p to win 1 dollar each time. He will stop if he has  $N(\dot{z}i)$  dollars or loses all his money. What is the probability that he end up losing all money?

Source: Practical Guide

Source. I factical Guide	
Calutian	
Solution:	

## Calculus

# Linear Algebra

### **Brain Teasers**

**4.1** (Chess tournament). A chess tournament has  $2^n$  players with skills  $1 > 2 > \cdots > 2^n$ . It is organized as a knockout tournament, so that after each round only the winner proceeds to the next round. Except for the final, opponents in each round are drawn at random. Let's also assume that when two players meet in a game, the player with better skills always wins. What's the probability that players 1 and 2 will meet in the final?

Also we have  $|1-\sqrt{2}| < \frac{1}{2}$ , hence  $(1-\sqrt{2})^{3000} < \frac{1}{2^{3000}} << 10^{-100}$ . So the 100th digit for  $(1-\sqrt{2})^{3000}$  is 0, hence the 100th digit of  $(1+\sqrt{2})^{3000}$  is 9.

**4.3.** How many numbers x are there less that  $10^6$  that the cubic of x ends with 11, i.e.  $x^3 \mod 100 = 11$ ?

Source: Practical Guide

Solution: