Probability

1. Terminology

- Sample Space: The set of all outcomes. It is denoted by S.
- Event: A subset of S.
- Pr(A): The probability of a subset A of the sample space.

  Note: Pr(x) Must satisfy these axioms:

  1. $0 \leq Pr(x) \leq 1$
  2. $Pr(\emptyset) = 0$
  3. $Pr(S) = 1$
  4. If $A \cap B = \emptyset$, then $Pr(A \cup B) = Pr(A) + Pr(B)$

- $\overline{A}$ is the complement of A.
  I.e. $\overline{A} = S - A$

  Note: $Pr(\overline{A}) + Pr(A) = 1$

- $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
- $Pr(A \cup B) \leq Pr(A) + Pr(B)$

- E.g. Suppose we toss a fair coin twice.

  The sample space is
  $S = \{HH, HT, TH, TT\}$

  $\{HH\}$ is an event.
  $\{HH, TT\}$ is another event.

  $Pr(\{HH\}) = \frac{1}{4}$

  $Pr(\{HH, HT\}) = \frac{1}{4} + \frac{1}{4}$
  $= \frac{1}{2}$

2. Independence:

- Two events, A and B, are independent iff $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

- Events A, B, C are independent iff:

  1. $Pr(A \cap B) = Pr(A) \cdot Pr(B)$
  2. $Pr(B \cap C) = Pr(B) \cdot Pr(C)$
  3. $Pr(A \cap C) = Pr(A) \cdot Pr(C)$
  4. $Pr(A \cap B \cap C) = Pr(A) \cdot Pr(B) \cdot Pr(C)$

- Generally, independence has to hold for all combinations.

- Independence is usually an assumption.
  E.g. We assume that a bunch of coin tosses are mutually independent.

- E.g. Suppose we toss 3 fair coins.

  Let A be the event that we get heads on the first toss.

  Let B be the event that we get heads on the 2nd toss.

  $A = \{HHH, HHT, HTH, HTT\}$
  $B = \{HHH, HHT, THH, THT\}$

  $Pr(A) = \frac{1}{2}$
  $Pr(B) = \frac{1}{2}$

  $Pr(1^{st}$ Toss is H and 2^{nd} Toss is H$)$
  $= Pr(A \cap B)$
  $= Pr(A) \cdot Pr(B)$
  $= (\frac{1}{2})(\frac{1}{2})$
  $= \frac{1}{4}$

  Since the first and second toss are unrelated, they are independent.

3. Conditional Probability:

- The probability of A given B is defined as:

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}, \quad Pr(B) \neq 0$$

Hilroy

- Thms:

1. $Pr(A|B) \cdot Pr(B) = Pr(A \cap B)$
   $= Pr(B|A) \cdot Pr(A)$

2. $Pr(A|B) \cdot Pr(B) + Pr(A|\bar{B}) \cdot Pr(\bar{B})$
   $= Pr(A)$

3. $Pr(\bar{A}|B) + Pr(A|B) = 1$

4. A and B are indep iff
   $Pr(A|B) = Pr(A)$

Proof for 2:

$P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$
$= P(A \cap B) + P(A \cap \bar{B})$
$= P(A)$

Proof for 3:

$P(\bar{A}|B) + P(A|B)$
$= \frac{P(\bar{A} \cap B)}{P(B)} + \frac{P(A \cap B)}{P(B)}$
$= \frac{P(\bar{A} \cap B) + P(A \cap B)}{P(B)}$
$= \frac{P(B)}{P(B)}$
$= 1$

Proof:

- If A and B are indep:

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(A) \cdot \cancel{Pr(B)}}{\cancel{Pr(B)}} = Pr(A)$$

- If $Pr(A|B) = Pr(A)$:

$$Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(A) \cdot Pr(B)$$

- E.g. Suppose we toss 2 fair coins. Given that at least one of the tosses result in heads, what is the probability that the 2nd toss will result in tails?

Soln:

Let A be the event that the 2nd toss will result in tails.

Let B be the event that at least one of the tosses will result in heads.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{1/4}{3/4}$$

$$= \frac{1}{3}$$

- E.g. Suppose that your code fails a test case. What is the probability that your code has a bug if we know that:

1. There is a 99% chance the buggy code fails the test.
2. There is a 1% chance the correct code fails the test.
3. There is a 0.1% chance that you write buggy code.

Hilroy

Soln:
Let $Pr(bug|F)$ be the probability that your code has a bug.

Let $Pr(F|bug) = 99\%$

Let $Pr(F|\overline{bug}) = 1\%$

Let $Pr(bug) = 0.1\%$

$$
\begin{aligned}
Pr(bug|F) &= \frac{Pr(F \cap Bug)}{Pr(F)} \\
&= \frac{Pr(F|bug) \cdot Pr(bug)}{Pr(F|bug) \cdot Pr(bug) + Pr(F|\overline{bug}) \cdot Pr(\overline{bug})} \\
&= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.99} \\
&\approx 0.09
\end{aligned}
$$

4. Baye's Thm:

- Suppose we partition the sample space into $B_1, B_2, \dots B_n$. Baye's Thm states that:

$$
\begin{aligned}
Pr(B_k|A) &= \frac{Pr(A \cap B_k)}{Pr(A)} \\
&= \frac{Pr(A \cap B_k)}{\sum_{i=1}^{n} Pr(A \cap B_i)} \\
&= \frac{Pr(A \cap B_k)}{\sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)} \\
&= \frac{Pr(A|B_k) \cdot Pr(B_k)}{\sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)}
\end{aligned}
$$

- E.g. Suppose there are 3 tech-support people who are equally likely to take a phone call.
If person #1 picks up, there is a 0.01 probability that the problem is unsolved.
If person #2 picks up, that probability is 0.02.
If person #3 picks up, that probability is 0.03.
Suppose I call them but my problem is unresolved. What is the probability that person #3 picks up?

Soln:

Let B_i be person #i who took my call.

Let A be the event that my problem is unresolved.

For each i, $Pr(B_i) = 1/3$

$Pr(A|B_1) = 0.01$
$Pr(A|B_2) = 0.02$
$Pr(A|B_3) = 0.03$

We are trying to solve for $Pr(B_3|A)$.

$$Pr(B_3|A) = \frac{Pr(A|B_3) \cdot Pr(B_3)}{Pr(A|B_1) \cdot Pr(B_1) + Pr(A|B_2) \cdot Pr(B_2) + Pr(A|B_3) \cdot B_3}$$

$$= \frac{0.03/3}{0.01/3 + 0.02/3 + 0.03/3}$$

$$0.5$$

5. Random Variables:

- A random variable is a function from the sample space to $\mathbb{R}$.

- The notion $Pr(X=k)$ means $Pr(\{o : X(o) = k\})$.

- E.g. Let the random variable X be the number of H's.
  Suppose we toss 3 fair coins.
  $Pr(X=2) = 3/8$
  This means the probability of getting 2 heads is 3/8.

- We can also have expressions like $2X$ and $X+Y$.

- E.g. Suppose we toss a fair coin until heads shows up.
  Let the random var X be the number of tosses until getting H.
  $Pr(X=1) = \frac{1}{2}$
  $Pr(X=2) = Pr(T) \cdot Pr(H)$
  $= (\frac{1}{2})(\frac{1}{2})$
  $= \frac{1}{4}$

$$Pr(X=3) = Pr(T) \cdot Pr(T) \cdot Pr(H)$$
$$= (\tfrac{1}{2})(\tfrac{1}{2})(\tfrac{1}{2})$$
$$= \frac{1}{8}$$

$$Pr(X=k) = (\tfrac{1}{2})^k$$

6. Expected Value:

- Average means expected value.
- The expected value of a random variable X is:
  $$E(X) = \sum_k k \cdot Pr(X=k),$$ k ranges over all possible values of X.

- Thm:
  Let c be a constant.

  1. $E(c \cdot X) = c \cdot E(X)$

  2. $E(c+X) = c + E(X)$ $\quad E(c) = c$

  3. $E(X+Y) = E(X) + E(Y)$

- E.g. A roulette wheel has 37 pockets. A ball randomly falls into one of them. You can bet on which pocket the ball falls into. If you guess wrong, the casino earns \$1. If you guess right, the casino loses \$35. What is the average (expected) casino earnings?

Hilroy

Soln:

$$E(X) = 1 \cdot Pr(X=1) + (-35) \cdot Pr(X=-35)$$
$$= \frac{36}{37} - 35\left(\frac{1}{37}\right)$$
$$= \frac{1}{37}$$

Expected Values Continued

- E.g. Suppose we toss a fair coin until we get 2 consecutive heads. What is the expected number of tosses?

Soln:

Let the expected number of coin flips be x.

The case analysis goes as follows:

a) If the first flip is tails, then we have wasted a flip. The probability of this is $\frac{1}{2}$ and the total number of flips required is $x+1$.

b) If the first flip is heads, but the second flip is tails, we have to start over. The probability of this is $\frac{1}{4}$ and the total number of flips required is $x+2$.

c) If both the first and second flips are heads, then we are done. The probability of this is $\frac{1}{4}$ and the total number of flips required is 2.

The equation we get is

$$x = (\tfrac{1}{2})(x+1) + (\tfrac{1}{4})(x+2) + (\tfrac{1}{4})(2)$$
$$= \frac{x+1}{2} + \frac{x+2}{4} + \frac{1}{2}$$
$$= \frac{2(x+1)+x+2+2}{4}$$
$$4x = 2x+2+x+2+2$$
$$= 3x+6$$
$$x = 6$$

$\therefore$ The expected number of tosses is 6.

- E.g. If we toss a fair coin until we get H followed by T, what is the expected number of tosses?

Soln:

Let x be the expected number of tosses.

Consider these cases:

a) If we get T, we wasted the flip, and have to start over. The probability of this is $\frac{1}{2}$ and the total number of flips required is $x+1$.

9 10 11

b) If we get H, we have a probability of $\frac{1}{2}$ to get T. Let Y be the expected additional number of tosses if you have just thrown H. The probability of getting H is $\frac{1}{2}$ and the number of flips required is Y+1.

With these 2 cases, we can set up a system of eqns to solve for X.

$$X = \frac{X+1}{2} + \frac{Y+1}{2}$$

$$Y = \frac{1}{2} + \frac{1}{2}(Y+1)$$

$$= \frac{1}{2}(Y+2)$$

$$2Y = Y+2$$

$$Y = 2$$

Plugging Y=2 into the first eqn, we get

$$X = \frac{X+1}{2} + \frac{3}{2}$$

$$2X = X+1+3$$

$$X = 4$$

∴ The expected number of flips is 4.

- E.g. What is the expected number of coin flips for getting H?

Soln:

Let x be the expected number of flips.

Consider the cases below:

a) We flip T. In this case, we have to flip another time. Therefore, the required number of flips is $x+1$. There is a $\frac{1}{2}$ chance of flipping T.

b) We flip H. In this case, the num of flips is 1. Furthermore the prob of getting H is $\frac{1}{2}$.

We can use the eqn $x = (\frac{1}{2})(1) + (\frac{1}{2})(x+1)$ to solve for x.

$$2x = x+1+1$$
$$x = 2$$

$\therefore$ The expected number of flips is 2.

7. Distributions:

- Let x be a random variable.
- We define the function $f(k)$ to be $Pr(X=k)$.
  I.e. $f(k) = Pr(x=k)$
- The distribution of x refers to that function.
- We say that X follows that distribution.
- If the dist's name is D, we also say that X is a D variable.
- If you know the dist, you know how to calculate various things, since it tells you about $Pr(x=k)$.
- Several common distributions are:

  a) Uniform Dist
  b) Bernouli Dist
  c) Binomial Dist
  d) Geometric Dist

Uniform
8. Dist:

- A uniform dist models a random draw from a range of integers.
  E.g. Throw a die, flip a fair coin

- X follows a uniform distribution over $R = \{m, \dots, n\}$ iff

$$Pr(X=k) = \frac{1}{|R|}$$
$$= \frac{1}{(n-m+1)}$$

I.e. $f(k) = \frac{1}{n-m+1}$

where k ranges from m to n, inclusive.

- $E(X) = \frac{m+n}{2}$

- E.g. If we throw a fair, six-sided die, what is the probability that 3 is facing up?

  Soln:
  $m=1, n=6$
  $$P(X=3) = \frac{1}{6-1+1}$$
  $$= \frac{1}{6}$$

9. Bernouli Dist:

- Models a single success or failure, with a prob of success p.

- Called an indicator random variable.

- X follows a Bernouli dist iff:
  $Pr(X=0) = 1-P$ (Failure)
  $Pr(X=1) = P$ (Success)

- $E(X) = 0(Pr(X=0)) + 1(Pr(X=1))$
  $= Pr(X=1)$
  $= P$

10. Binomial Dist:

- Sum of Bernouli Dist.

- This models tossing a coin n times where the tosses are mutually independent and each toss has a prob p of success.

- X follows a binominal dist iff:

  $Pr(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$

  where k ranges from 0 to n, inclusive.

  n is the number of independent trials.

  k is the number of successes.
  P is the probability of success.

- $E(x) = \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$
  $= n \cdot p$

- E.g. A fair coin is tossed 10 times. What is the probability of getting exactly 6 heads?

  Soln:
  $n = 10$
  $k = 6$
  $p = 0.5$
  $1 - p = 0.5$

  $\binom{10}{6} \cdot \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^{10-6}$
  $= \frac{10!}{(10-6)!6!} \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4$
  $= 210 (0.5)^{10}$
  $= 0.205$
  $= 20.5\%$

- E.g. You sell sandwiches. 70% of people buying a sandwich from you choose chicken while 30% choose something else. What is the probability that exactly 2 of the 3 next customers buyachicken sandwich?

  Soln:
  $n = 3,\ k = 2,\ p = 0.7,\ 1 - p = 0.3$
  $\frac{3!}{(3-2)!2!} \cdot (0.7)^2 \cdot (0.3)$
  $= 0.441$
  $= 44.1\%$

- E.g. You flip a coin.
  $Pr(H) = P$
  $Pr(T) = 1-P$
  What is the expected number of H in n tosses?

  Soln:
  Let Y be the number of H in n tosses. (Binomial)
  Let X_i be the number of H in the i^{th} toss. (Bernoulli)

  $$E[Y] = E[X_1 + X_2 + \dots + X_n]$$
  $$= E[X_1] + \dots + E[X_n]$$
  $$= P + P + \dots + P$$
  $$= nP$$

- E.g. If we have n customers who leave their hats checked at a coat check and will get a random hat when they exit, what is the expected number of ppl who get their own hat back?

  Soln:
  Let $X_i = 1$ if the i^{th} person gets back his/her hat. Otherwise, $X_i = 0$.

  Let Y be the number of ppl who get their hat back.

  $$E[Y] = E[X_1 + \dots + X_n]$$
  $$= E[X_1] + \dots + E[X_n]$$

$= P(X_1 = 1) + \dots + P(X_n = 1)$

$= \frac{(n-1)!}{n!} + \dots + \frac{(n-1)!}{n!}$

$= \frac{1}{n} + \dots + \frac{1}{n}$

$= 1$

Note: Another way you can think of this problem is say you have an array with n distinct elements, like this $[1, \dots n]$. If we scramble all the elements in the array, what is the expected number of elements that still remain in its index?

Each time you check if an element is in the right place, you don't care about the ordering of the other elements. There are $(n-1)!$ possible orderings for the other elements and $n!$ possible orderings for all elements. Therefore, to get the probability that the element you are checking is in the right place, divide $(n-1)!$ by $n!$ You get $\frac{1}{n}$.

- E.g. Consider the code below.

```
x=0
for i in range(n):
    d= random number in [0,1]
    if d ≤ 3/4:
        x = x+5
    else:
        x = x-1
```

What is the expected value of x after the loop?

Soln:
Let X_i be the value x changed by in the i^{th} iteration.

Let Y be the value of x after the loop.

$$
\begin{aligned}
E[Y] &= E[0 + X_1 + \dots + X_n] \\
&= E[0] + E[X_1] + \dots + E[X_n] \\
&= E[X_1] + \dots + E[X_n]
\end{aligned}
$$

$$
\begin{aligned}
E[X_i] &= (3/4)(5) + (1/4)(-1) \\
&= \frac{14}{4} \\
&= \frac{7}{2}
\end{aligned}
$$

$$
\begin{aligned}
E[Y] &= 7/2 + \dots + 7/2 \\
&= \frac{7n}{2}
\end{aligned}
$$

9 10 11

11. Geometric Dist:

- This models the number of failures before the first success.
- X follows a geo dist iff:

  $Pr(X=k) = (1-p)^{k-1} \cdot p, \quad k \geq 1$

  k is the number of independent trials.
  p is the probability of success.

- $E(X) = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p$

  $= \frac{1}{p}$

- E.g. What is the probability of flipping a fair coin 4 times before getting H?

  Soln:
  Another way to word this question is "What is the probability that I get the first H on my 4^{th} flip?"

  $k=4, \ p=0.5$

  $(1-0.5)^{4-1} \cdot 0.5$
  $= (0.5)^3 \cdot 0.5$
  $= 0.5^4$
  $= \frac{1}{16}$

9 10 11

- E.g. Consider the code below.

```
c=0
while random(1,4) ≠ 4:
    c=c+1
```

What is the expected value of c after the loop?

Soln:
Let C be the random variable for the final c.

Let R be the random var for the number of times random(1,4) is used.

Note: $C=R-1$

Since R has a geo dist, $E(R) = \frac{1}{1/4} = 4$.

$E(C) = E(R-1)$
$= 3$

12. Randomized Quicksort

- Quicksort can have a worst-case time of $\Theta(n^2)$ if we pick the max element as the pivot. The left side will be very big.
- If we do a randomized quicksort, the pivot is chosen randomly.
- At most, randomly picking a partition can happen n times.
- Let X be the number of comparisons. Note: X is a random variable. Then, it takes $O(n+X)$ time for randomized quicksort.
- If we can prove that $E(X) \in O(n \lg n)$, then we can say that the expected value of the run time is $O(n \lg n)$.
- Proof:
  - Note: Two elements are compared at most once.
  - Note: Two elements, z_i and z_j, are compared if either is a pivot and there is no z_k, $z_i < z_k < z_j$, that has been a pivot or is a pivot.

E.g. Consider the array.

[3, 5, 4, 9]

If we want to compare 3 with 5, then either 3 or 5 must be the pivot and 4 cannot be the pivot.

All other elements less than z_i or greater than z_j don't affect this.

E.g. From the array above, if 9 is the chosen pivot, then 3 and 5 would still be on the same side and could still be compared.

In summary:

- Let $z_i < z_k < z_j$
- If z_k is chosen as the pivot first, then z_i cannot be compared to z_j
- If z_i is the pivot, then z_i will be compared to z_j. The probability of choosing z_i as the pivot is $\frac{1}{j-i+1}$.

- If z_j is chosen as the pivot, then it will be compared to z_i. The probability of choosing z_j as the pivot is $\frac{1}{j-i+1}$.

- The probability of z_j being compared to z_i is $\frac{2}{j-i+1}$.

- Let z_i be the i^{th} smallest element in A, where A is an array of unsorted numbers.

- Let X_{ij} be a random variable s.t.

$$X_{ij} = \begin{cases} 1, & \text{If } z_i \text{ is compared to } z_j. \\ 0, & \text{otherwise} \end{cases}$$

- X = The sum of X_{ij} over all $i<j$.

This means:

$$E(X) = E\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right)$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E(X_{ij})$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Pr(z_i \text{ compared to } z_j)$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\frac{2}{j-i+1}\right)$$

$$= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \left(\frac{2}{k+1}\right), \quad k = j-i$$

$$< \sum_{i=1}^{n-1} 2\ln(n)$$

$$= 2(n-1)\ln(n)$$

$$\therefore E(x) \in O(n \lg n)$$

- There are 2 ways QS can incorporate randomization:

  1. The alg selects the pivots randomly. In this case, the input does not have to be random.
     E(running time) = Expected R.T.

  2. The input is random. Here, the alg does not need to be random.
     E(running time) = Average-Case Running Time

- To compute the average-case time of unrandomized QS, we can use the same math because each element is equally likely to be the pivot. The answer is still $O(n \lg n)$.