Probability Formula, Notations, Mutually Exclusive Events, Axioms, Permutation, Combination

Outline

1. Probability Notations
2. Mutually Exclusive Events
3. Axioms of Probability
4. Example 1
5. Basic Principle of Counting
6. Example 2
7. Permutation
8. Example 3
9. Example 4
10. Combination

Probability

Measure of likelihood that a particular event will occur. It is quantified as a number between 0 and 1, 0 indicates that an event will not happen and 1 indicates that the event will happen.

Probability Formula

Probability of an event A is

$$P(A) = \frac{Number\ of\ favorable\ outcomes}{Total\ number\ of\ possible\ outcomes}$$

Probability Notations (5)

1. $P(A) = 60\%, A\ is\ an\ event$
   It means that if experiment is repeated independently under the same conditions millions of times, A would happen 60% of the times. Note that P(60%) is an incorrect way to write.
2. P(A or B) means P(A or B or both). It can be written as $P(A \cup B)$.
3. P(A and B) is written as P(AB).
4. P(A given B) is written as P(A|B).
5. $P(A^c)$ means P(not A).

Mutually Exclusive Events

Two events A and B are mutually exclusive if they cannot occur at the same time.

$$P(AB) = 0$$

Axioms of Probability (5)

Axioms are initial assumptions or rules.

1. $P(A) \geq 0$
2. $P(A) + P(A^c) = 1$
3. If A_1, A_2, A_3, . . . are mutually exclusive, then
   $P\big(A_1\ or\ A_2\ or\ A_3 \ldots\big) = P\big(A_1\big) + P\big(A_2\big) + P\big(A_3\big) + \ldots$
   (Addition Rule for Mutually Exclusive Events)
4. $P(A\ or\ B) = P(A) + P(B) - P(AB)$
   $P(A\ or\ B\ or\ C)$
   $= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$
   (General Addition Rule)
   Note that probability is equivalent to area of polygons in Venn diagrams.
5. If A_1, A_2, . . . A_n are equally likely and mutually exclusive, and if
   $P\big(A_1\ or\ A_2\ or \ldots\ A_n\big) = 1$
   Then
   $P\big(A_k\big) = \frac{1}{n}$
   $P\big(A_1\ or\ A_2\ or \ldots\ A_k\big) = \frac{k}{n}$

Example 1 - You have 76, and the board is KQ54. What is P(straight)?

We have

$$52 - 2 - 4 = 46$$

Then

$$P\big(straight\big) = P(3\ on\ the\ river\ or\ 8\ on\ the\ river)$$
$$= P(3\ on\ the\ river) + P(8\ on\ the\ river) = \frac{4}{46} + \frac{4}{46} = \frac{4}{23} = 0.1739$$

Here 3 on the river and 8 on the river are mutually exclusive events, only one can occur at a time on the river.

Basic Principle of Counting

If there are a_1 distinct possible outcomes on trial 1, and for each of them, there are a_2 distinct possible outcomes on trial 2, then there are

$$p = a_1 \times a_2 \; distinct\ possible\ ordered\ outcomes\ on\ both$$

Example 2 - You get 1 card, your opponent gets 1 card. What are the no. of distinct possibilities?

Ordered, A♣ K♥ is not same as K♥ A♣

$$52 \times 51 = 2652$$

Permutation

Each outcome where order matters, is called a permutation. In general, with j experiments, each with a_i possibilities, the number of distinct outcomes where order matters is

$$p = a_1 \times a_2 \times \ldots a_j$$

Example 3 - How many permutations there are of a deck of cards?

$$52 \times 51 \times \ldots 1 = 52! = 8.06 \times 10^{67}$$

Example 4 - In hold 'em, how many distinct 2 card hands are possible?

Now, A♣K♥ is same as K♥A♣ since both cards are in our hand so order does not matter

Number of distinct hands where order doesn't matter is

$$\frac{52 \times 51}{2} = 1326$$

Combination

It is collection of outcomes, where order doesn't matter. In general, with n distinct objects, the number of ways to choose k different ones, where order doesn’t matter is

$$n\ choose\ k = C(n, k) = \frac{n!}{k!\,(n\ - k)!}$$

Conditional Probability, Independent Events, Multiplication Rule, Odds Ratios

Outline

1. Conditional Probability
2. Independent Events
3. General Multiplication Rule
4. Independent Events Examples
5. Example 1
6. Example 2
7. Odds Ratios
8. Example 3
9. Example 4
10. Example 5

Conditional Probability

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

Independent Events

Two events are independent if the occurrence of one event does not affect the occurrence the occurrence of the other.

$$P(AB) = P(A)P(B)$$

(Multiplication Rule for Independent Events)

Using conditional probability definition, A and B are independent if

$$P(A|B) = P(A)$$

General Multiplication Rule

$$P(AB) = P(A)P(B|A)$$
$$P(ABC \dots) = P(A)P(B|A)P(C|AB)$$

Independent Events Examples

1. Outcomes on different rolls of a die.
2. Outcomes on different flips of a coin.
3. Outcomes on different poker hands.
4. Outcomes when sampling from a large population.

Example 1 - What is P(you get AA on 1st hand and I get AA on 2nd hand)?

These events are independent of each other as hands are different. We have

$$= P(you\ get\ AA\ on\ 1st)P(I\ get\ AA\ on\ 2nd) = \frac{C(4,2)}{C(52,2)} \times \frac{C(4,2)}{C(52,2)}$$
$$= \frac{1}{221} \times \frac{1}{221} = \frac{1}{48841}$$

Example 2 - What is P(you get AA on 1st hand and I get AA on 1st hand)?

$$= P(you\ get\ AA\ on\ 1st)P(I\ get\ AA|you\ have\ AA) = \frac{1}{221} \times \frac{1}{50\ choose\ 2}$$
$$= \frac{1}{221} \times \frac{1}{1125} = \frac{1}{270725}$$

Odds Ratios

$$Odds\ Ratio\ of\ A = \frac{P(A)}{P(A^c)}$$
$$Odds\ against\ A = Odds\ Ratio\ of\ A^c = \frac{P(A^c)}{P(A)}$$

Example 3 - What is P(you get dealt AA and flop a house)?

Using general multiplication rule (events are not independent)

$$= P(you\ get\ dealt\ AA)P(you\ flop\ a\ full\ house|AA)$$
$$= \frac{C(4,2)}{C(52,2)} \times P(triplet\ or\ Axx|AA)$$
$$= \frac{6}{1326} \times \frac{12 \times C(4,3) + 2 \times 12 \times C(4,2)}{C(50,3)} = 0.00443\%$$

Getting a triplet or an Axx are mutually exclusive events, only one can occur, so we use addition rule above.

Example 4 - What is P(you are dealt A♦ K♦ and flop a royal flush)?

Recall that here order does not matter since getting dealt A♦K♦ is same as getting dealt K♦A♦. Using general multiplication rule

$$= P(you\ get\ dealt\ A♦K♦)P(you\ flop\ a\ royal\ flush|you\ have\ A♦K♦)$$
$$= P(you\ get\ dealt\ A♦K♦)P(flop\ contains\ 10♦J♦Q♦|you\ have\ A♦K♦)$$
$$= \frac{1}{C(52,2)} \times \frac{1}{C(50,3)} = \frac{1}{25989600}$$

Example 5 - Deal till the first ace appears. Let X = the next card after the ace. P(X = A♠)? P(X = 2♣)?

Divide the question into sub-parts, we have

a. How many permutations of the 52 cards are there?
   $a = 52!$
b. How many of these permutations have A♠ right after the 1st ace?
   i. How many perms of the other 51 cards are there?
      $b_1 = 51!$
   ii. For each of these, imagine putting the A♠ right after 1st ace. How many perms of A♠ right after 1st ace?
      $b_2 = 51!$
      Note that there is 1:1 correspondence between b_1 and b_2, i.e. between permutations of the other 51 cards and permutations of 52 cards s.t. A♠ is right after 1st ace.
   $b = 51!$

Recall that using probability definition, we have

$$P(A) = \frac{Number\ of\ favorable\ outcomes}{Total\ number\ of\ possible\ outcomes}$$

Using this we have answer to overall question is

$$\frac{b}{a} = \frac{51!}{52!} = \frac{1}{52} = 0.0192$$

Same goes for 2♣

Bayes' Rule

Outline

1. Example 1
2. Bayes' Rule Assumptions
3. Bayes' Rule
4. Example 2
5. Example 3

Example 1 - Which is more likely, give no information about your cards, flopping 3 of a kind or eventually making 4 of a kind?

First we find P(flop 3 of a kind), including the case where all 3 are on the board and not including full houses. Key idea is to forget the order. Consider all combinations of your 2 cards and the flop. We have a set of 5 cards, from which any such combo is equally likely.

$$P(flop\ 3\ of\ a\ kind) = \frac{No.of\ different\ 3\ of\ a\ kinds}{C(52,5)}$$

How many different 3 of a kind combinations are possible?

$$a = 13 \times C(4,3)\ different\ choices\ for\ the\ triple$$

For each such choice, there are C(12, 2) choices left for the numbers on the other 2 cards which aren't the 3 of a kind, and for each of these numbers, there are 4 possibilities for its suit. So we have,

$$P(flop\ 3\ of\ a\ kind) = \frac{13 \times C(4,3) \times C(12,2) \times C(12,2) \times 4 \times 4}{C(52,5)}$$
$$= 2.11\%$$

If we consider full house too, we have

$$P(flop\ 3\ of\ a\ kind\ or\ a\ full\ house) = \frac{13 \times C(4,3) \times C(48,2)}{C(52,5)} = 2.26\%$$

Now we find P(eventually make 4 of a kind), including the case where all 4 are on the board. Again, forget card order and consider all collections of 7 cards. Out of C(52, 7) different combinations, each equally likely, how many of involve 4 of a kind? There are 13 choices for the 4 of a kind. For each such choice, there are C(48, 3) possibilities for the other 3 cards. So we have,

$$P(4\ of\ a\ kind) = \frac{13 \times C(48,3)}{C(52,7)} = 0.168\%$$

Hence it is more likely that we flop a 3 of a kind than eventually making 4 of a kind.

Bayes' Rule Assumptions

Suppose that $B_1, B_2, \ldots B_n$ are disjoint events (different from mutually exclusive) and that exactly one of them must occur. Suppose you want P(B1 |A), but you only know P(A|B1), P(A|B2), etc., and also know P(B1), P(B2), . . . P(Bn)

Bayes' Rule

If $B_1, B_2, \ldots B_n$ are disjoint events with

$$P(B_1\ or \ldots B_n) = 1$$

Then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_j)P(B_j)}$$

Example 2 - If a test is 95% accurate and 1% of the population has a condition, then given a random person from the population, Find P(she has the condition|she tests positive).

Problem can be written as P(cond|+), we have

$$= \frac{P(+|cond)P(cond)}{(P(+|cond)P(cond)) + (P(+|no\ cond)P(no\ cond))}$$
$$= \frac{95\% \times 1\%}{(95\% \times 1\%) + (5\% \times 99\%)} = 16.1\%$$

Example 3 - Your opponent makes a huge bet of 10 times the big blind. Suppose she'd only do that with AA, or 54, or AK. Suppose P(huge bet|AA) is 5%, P(huge bet|54) is 70%, P(huge bet|AK) is 25%. What is P(AA|huge bet)?

$P(AA|huge\ bet) = 1.94\%$

Random Variable, Cumulative Distributive Function, Expected Value

Outline

1. Variable
2. Random Variable
3. Discrete and Continuous Random Variable
4. Example 1
5. Distribution of Random Variable
6. Cumulative Distribution Function (cdf)
7. Probability Mass Function (pmf) and Probability Density Function (pdf or density)
8. Expected Value
9. Expected Value of Discrete Random Variable/Probability Mass Function
10. Example 2
11. Example 3
12. Example 4
13. Difference between Sample Mean and Expected Value
14. Reasons why Expected Value applies to Poker
15. Expected Value Properties

Variable

A variable is something that can take different numeric values.

Random Variable

A random variable (X) can take different numeric values with different probabilities.

Discrete and Continuous Random Variable

Discrete	Continuous
X is discrete if all its possible values can be listed.	If X can take any value in an interval say [0, 1], then X is continuous

Example 1 - Two cards are dealt to you. Let X be 1 if you get a pair, and X is 0 otherwise. Find P(X is 1) and P(X is 0).

$$P(X\ is\ 1) = \frac{3}{51} = 5.9\%$$

$$P(X\ is\ 0) = \frac{48}{51} = 94.1\%$$

Distribution of Random Variable

Distribution of X means all the information about all the possible values X can take, along with their probabilities.

Cumulative Distribution Function (cdf)

Cumulative means total amount of something when all added together. Any random variable has a cumulative distribution function (cdf)

$$F(b) = P(X \leq b)$$

Probability Mass Function (pmf) and Probability Density Function (pdf or density)

Probability Mass Function (pmf)	Probability Density Function (pdf or density)
If X is discrete, then it has probability mass function. $f(b) = P(X = b)$	Continuous random variables are often characterized by their probability density function. $P(X\ is\ in\ B) = \int_B f(x)dx$

Expected Value

Mean of a random variable X in a probability distribution. Expected value of X represents a best guess of X. It is denoted E(X) or μ.

Expected Value of Discrete Random Variable/Probability Mass Function

For a discrete random variable X with pmf f(b), the expected value of X

$$E(X)\ or\ \mu = \sum b_i f(b_i)$$

The sum is over all possible values of b

Example 2 - 2 cards are dealt to you. X is 1 if pair, 0 otherwise. Find E(X).

From previous example, we have

$P(X\ is\ 1) = 5.9\%$

$\mathrm{P}(X\ is\ 0) = 94.1\%$

$E(X) = (1 \times 5.9\%) + (0 \times 94.1\%) = 5.9\%$

Example 3 - A coin is flipped, X equals 20 if heads, X equal 10 if tails. Find E(X).

$E(X) = (20 \times 50\%) + (10 \times 50\%) = 15$

Example 4 - Find expected value of a lotto ticket where f($10 million) is 1/C(52, 6) and f($0) is 1 - 1/20 million.

$$\frac{1}{C(52,6)} = \frac{1}{20\ million}$$

$$E(X) = \left(\$10\ million\ \times \frac{1}{20\ million}\right) + 0 = \$0.5$$

Difference between Sample Mean and Expected Value

Sample Mean	Expected Value
1. Average of a set of observations or data points collected from a sample. 2. Calculated by summing all the values in the sample and dividing by number of observation. $\bar{X} = \frac{\sum X_i}{n}$ 3. Represents the average value of the data you have collected.	1. Mean of a random variable in a probability distribution. 2. Calculated by summing the products of each possible value of the random variable and its probability. $E(X)\ or\ \mu = \sum b_i f(b_i)$ 3. Represents the long-term average value of random variable if the experiment were repeated many times.

Reasons why Expected Value applies to Poker (3)

1. In symmetric, winner-take-tall games, the optimal strategy is the one which use the myopic rule, i.e., given any choice of options, always choose the one that maximizes your expected value.
2. If you repeat an experiment over and over repeatedly, your long-term average will ultimately converge to the expected value.
3. Great way to check whether you are a long-term winning or losing player, or to verify if a certain strategy works or not, is to check whether the sample mean is positive and to see if it has converged to the expected value.

Expected Value Properties (2)

1. For any random variables X and any constants a and b
   $E(aX + b) = aE(X) + b$
2. $E(X + Y) = E(X) + E(Y)$
   unless E(X) is ∞ and E(X) is -∞, then E(X) + E(Y) is undefined.

Variance and Standard Deviation, Markov Inequality, Chebyshev Inequality, Moment Generating Functions

Outline

1. Example 1
2. Example 2
3. Variance and Standard Deviation
4. Example 3
5. Example 4
6. Markov Inequality
7. Proof of Markov Inequality if X is Discrete and if X is continuous
8. Chebyshev Inequality
9. Proof of Chebyshev Inequality
10. Example 5
11. Moment Generating Functions

Example 1 - What is P(flop 4 of kind)?

$$= \frac{13 \times 48}{C(52,5)} = \frac{1}{4165} = 0.024\%$$

Example 2 - Find P(flop 4 of a kind|pocket pair).

No matter which pocket pair you have, there are C(50, 3) possible flops, each equally likely. How many of them give you 4 of a kind? 48. For e.g. if you have 7♠7♥ then need to flop 7♦7♣ x, and there are 48 choices for x. So we have

$$P(flop\ 4\ of\ a\ kind|pocket\ pair) = \frac{48}{C(50,3)} = \frac{1}{408} = 0.245\%$$

Variance and Standard Deviation

Variance (denoted by V(X) or σ^2) measures the average degree to which each point differs from the mean.

$$V(X) = \sigma^2 = E\left((X-\mu)^2\right) = E\left(X^2\right) - \mu^2$$

(Can be derived using expected value properties)

Standard deviation (σ) indicates how far an observation would typically deviate from mean.

$$\sigma = \sqrt{V(X)}$$

Example 3 - Say X equals $4 if red card, X is $-5 if black. Calculate expected value, variance and standard deviation.

$$E(X) \; or \; \mu = (\$4 \times 0.5) + (\$-5 \times 0.5) = \$-0.5$$

$$E\left(X^2\right) = \left(\$4^2 \times 0.5\right) + \left(\$-5^2 \times 0.5\right) = (\$16 \times 0.5) + (\$25 \times 0.5)$$
$$= \$20.5$$

$$V(X) \; or \; \sigma^2 = E\left(X^2\right) - \mu^2 = \$20.5 - \${-0.5}^2 = \$20.25$$

$$SD \; \sigma = \$4.5$$

Example 4 - Say X equals $1 if red card, X is $-2 if black. Calculate expected value, variance and standard deviation.

$$E(X) \; or \; \mu = (\$1 \times 0.5) + (\$-2 \times 0.5) = \$-0.5$$

$$E\left(X^2\right) = \left(\$1^2 \times 0.5\right) + (\$-2 \times 0.5) = (\$1 \times 0.5) + (\$4 \times 0.5) = \$2.5$$

$$V(X) \; or \; \sigma^2 = E\left(X^2\right) - \mu^2 = \$2.5 - \${-0.5}^2 = \$2.25$$

$$SD \; \sigma = \$1.5$$

Markov Inequality

It states that if X takes only non-negative values, and c is any number > 0, then

$$P(X \geq c) \leq \frac{E(X)}{c}$$

Proof of Markov Inequality if X is Discrete and if X is Continuous

If X is discrete and non-negative, then

$$
\begin{aligned}
E(X) &= \sum_b bP(X = b) \\
&= \sum_{b<c} b(P(X = b) + \sum_{b\geq c} bP(X = b) \\
&\geq \sum_{b\geq c} bP(X = b) \\
&\geq \sum_{b\geq c} cP(X = b) \\
&= c\sum_{b\geq c} P(X = b) \\
&= cP(X \geq c)
\end{aligned}
$$

If X is continuous with pdf f(y)

$$
\begin{aligned}
E(X) &= \int yf(y)dy \\
&= \int_0^c yf(y)dy + \int_c^\infty yf(y)dy \\
&\geq \int_c^\infty yf(y)dy \\
&\geq \int_c^\infty cf(y)dy \\
&= c\int_c^\infty f(y)dy \\
&= cP(X \geq c)
\end{aligned}
$$

Thus

$$P(X \geq c) \leq \frac{E(X)}{c}$$

Chebyshev Inequality

It states that for any random variable Y with expected value μ and variance σ^2, and any real number a > 0

$$P(|Y - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

Proof of Chebyshev Inequality

It follows directly from the Markov inequality by letting

$$c = a^2$$

And

$$X = (Y - \mu)^2$$

Example 5 - Suppose the average time until a player is eliminated in a given tournament is 7 hours. Take a player at random, what does the Markov inequality say about the probability that the player lasts longer than 21 hours?

Let X = time till the player is eliminated. X is non-negative here and E(X) is 7 hours. Using Markov inequality we have

$$P(X \geq 21) \leq \frac{7}{21} = \frac{1}{3} = 33.3\%$$

Moments

Suppose X is a random variable. Then E(X), E(X^2), E(X^3), etc. are the moments of X.

Moment Generating Functions

$$\emptyset_X(t) = E\left(e^{tX}\right) \, is \, moment \, generating \, function \, of \, X$$

Take derivatives w.r.t. t of $\emptyset_X$(t) and evaluate at

$$t = 0$$

to get moments of X.