8. Probability theory: basics

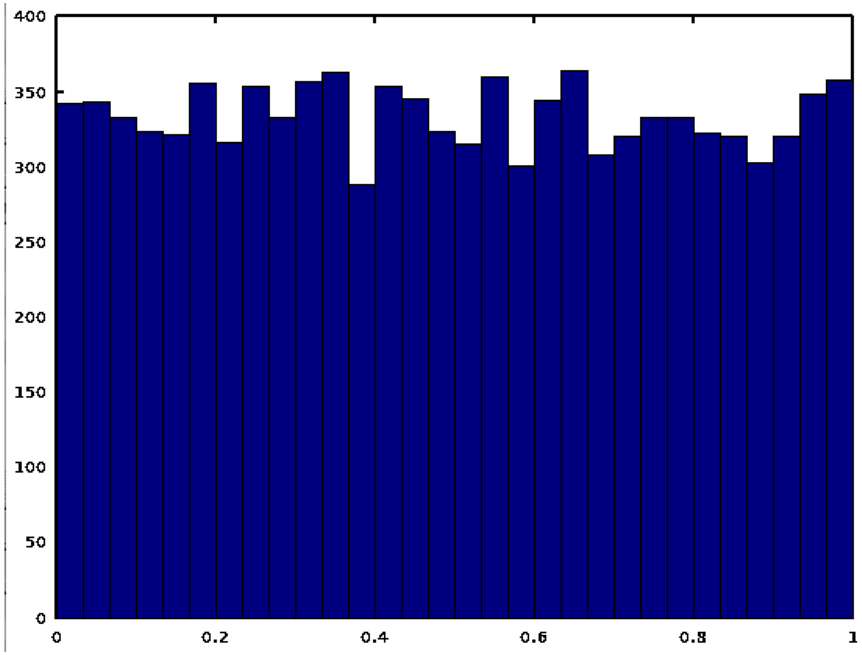
- Random numbers
- Conditional probability
- Joint probability
- Bayes’ theorem
- Marginal probability
- Posterior probability and prior probability
- Logical indexing of matrices

Random numbers

- `rand(m,n)` generates a $m \times n$ matrix of random numbers that are uniformly distributed on the interval (0,1)

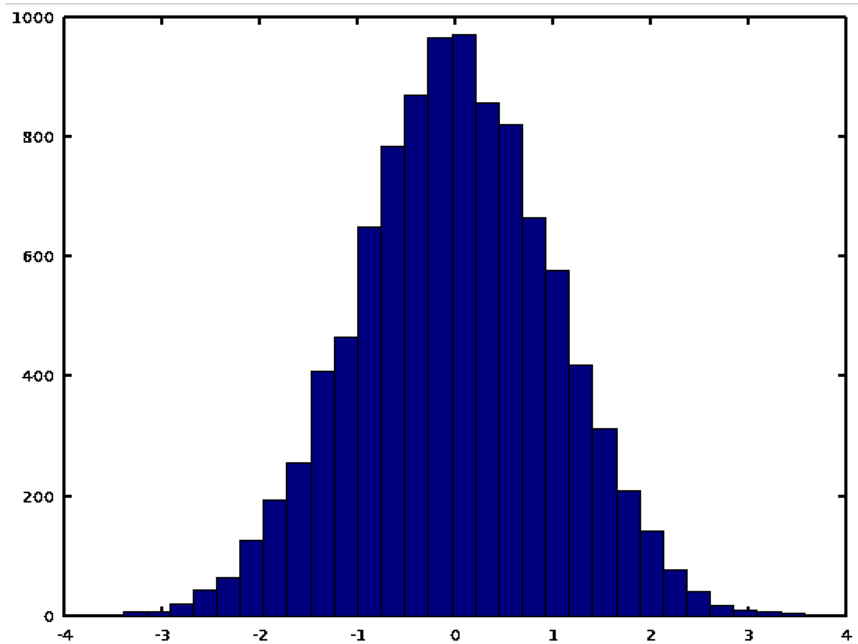
※ `hist` creates a histogram for a given set of numbers and plot it

```
>> hist(rand(10000,1),30)
```



- `randn(m,n)` generates a $m \times n$ matrix of random numbers distributed according to a normal distribution with zero mean and variance 1.

```
>> hist(randn(10000,1),30)
```



Probability theory(1/4)

- Suppose a red box and a blue box containing apples and oranges as shown below

2 apples, 6 oranges

3 apples, 1 orange

- Consider a trial of first choosing a box and then picking a fruit randomly from it
- Assume that the red box is chosen with probability 40% and the blue box is chosen with probability 60%
- Then, how can we answer questions like
  - What is the probability that an apple is picked?
  - When we know an apple is picked, what is the probability that the apple came from the blue box?

From 1st chapter in C. M. Bishop, Pattern Recognition and Machine Learning

Probability theory(2/4)

- Probabilistic variables
  - *B* = which box is selected; *B*=*r*: the red box and *B*=*b*: the blue box
  - *F* = which fruit is selected; *F*=*a*: an apple and *F*=*o*: an orange
- Probability of selecting each box:

$$p(B = r) = 4/10 \qquad p(B = b) = 6/10$$

- Conditional probability: probability of selecting an apple when the red box has been chosen
  - It is just the ratio of apples (2) to the number of fruits (8) in the box

$$p(F = a|B = r) = 1/4$$

  - Similarly, we have

$$p(F = o|B = r) = 3/4$$

$$p(F = a|B = b) = 3/4$$

$$p(F = o|B = b) = 1/4$$

Probability theory(3/4)

- Joint probability
  - What is the probability of selecting the red box AND an orange
  - From Bayes' theorem, it can be represented as follows:

$$p(F = o, B = r) = p(F = o|B = r)p(B = r)$$

	$B=r$	$B=b$
$F=a$	1/10	9/20
$F=o$	3/10	3/20

  - By the way, if the following holds true, we say that the two probabilistic variables are independent of each other

$$p(F = o, B = r) = p(F = o)p(B = r)$$

Monte Carlo methods

- Let's estimate the joint probabilities by simulating the above trial using random numbers; *Monte Carlo estimate of probabilities*
  - The following script first picks a box and then a piece of fruits randomly as is described above for, say, 10,000 trials; and counts the numbers of the cases (B, F) = (r, a), (r, o), (b, a), and (b, o), respectively

```
% box_fruit.m
num_bf = zeros(2,2);
for i=1:10000
  if rand(1,1) < 0.4, % red box (40%)
    if rand(1,1) < 2.0/8, % apple
      num_bf(1,1) += 1;
    else % orange
      num_bf(2,1) += 1;
    end
  else % blue box
    if rand(1,1) < 3.0/4, % apple
      num_bf(1,2) += 1;
    else % orange
      num_bf(2,2) += 1;
    end
  end
end
```

```
>> box_fruit
>> num_bf/sum(sum(num_bf))
ans =
   0.10090   0.44660
   0.29740   0.15510
```

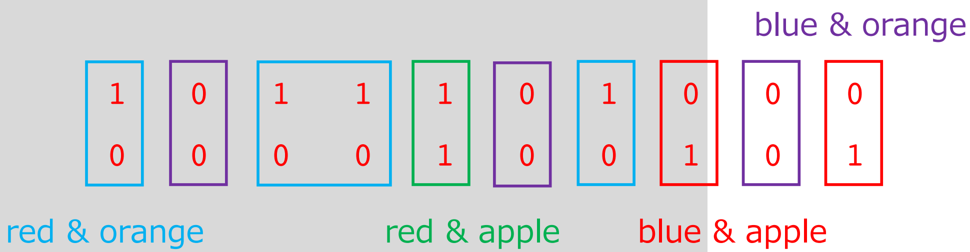
	B=r	B=b
F=a	1/10	9/20
F=o	3/10	3/20

Logical indexing of matrices

- The same results can be obtained much more efficiently

```
>> B = rand(1,10000) < 0.4; % 1 for red box; 0 for blue box
>>                           Each element of B is 1 if its corresponding element of the vector
>> B(1:10)                   on the right hand side satisfies the inequality and 0 otherwise
ans =      The boxes selected for the first 10 trials (out of 10,000)
   1    0    1    1    1    0    1    0    0    0

>> Frnd = rand(1,10000);
>>         The set of all indices of 1 elements in B
>> F(B==1) = Frnd(B==1) < 2/8; % 1 for apple; 0 for orange
>>         The set of all indices of 0 elements in B
>> F(B==0) = Frnd(B==0) < 3/4; % 1 for apple; 0 for orange
>>
>> F(1:10)
ans =      The fruits selected for the first 10 trials (out of 10,000)
   0    0    0    0    1    0    0    1    0    1
>> sum(F==1&B==1)/10000
ans = 0.09570
>> sum(F==1&B==0)/10000
ans = 0.45430
>> sum(F==0&B==1)/10000
ans = 0.29890
>> sum(F==0&B==0)/10000
ans = 0.15110
```



Probability theory(4/4)

- What is the probability of selecting an apple in a trial?
  - This kind of probabilities is called marginal probability
  - Answer is 11/20

$$p(F=a) = p(F=a, B=r) + p(F=a, B=b)$$

- We are told that the selected fruit is an orange; what is the probability that the selected box, from which the orange came, was the red box
  - Answer is 2/3

$$p(B=r|F=o) = \frac{p(B=r, F=o)}{p(F=o)}$$

  - Probabilities like this are called posterior probabilities; because it is the probabilities obtained after we have observed *F*
  - *Probabilities like* $p(B=r)$ are called prior probabilities; they are given in advance

Exercise 8.1

- Calculate a Monte Carlo estimate of $p(F=a)$ using logical indexing of matrices explained in a previous slide
- Calculate a Monte Carlo estimate of the posterior probability $p(B=r|F=o)$