20.110/2.772/5.601 Fall 2005

Recitation # 9
10/13/2005

1.

EXAMPLE 1.15 Counting sequences of coin flips and die rolls. You flip a coin $N = 4$ times. How many different sequences have three heads? According to Equation (1.19),

$$W(n_H, N) = \frac{N!}{n_H! n_T!} = \frac{4!}{3!1!} = 4.$$

They are $THHH$, $HTHH$, $HHTH$, and $HHHT$. How many different sequences have two heads?

$$W(2,4) = \frac{4!}{2!2!} = 6.$$

They are $TTHH$, $HHTT$, $THTH$, $HTHT$, $THHT$, and $HTTH$.

You flip a coin one hundred and seventeen times. How many different sequences have thirty-six heads?

$$W(36,117) = \frac{117!}{36!81!} \approx 1.84 \times 10^{30}.$$

We won't write the sequences out.

You roll a die fifteen times. How many different sequences have three 1's, one 2, one 3, five 4's, two 5's, and three 6's? According to Equation (1.18),

$$W = \frac{15!}{3!1!1!5!2!3!} = 151,351,200.$$

2.

EXAMPLE 1.17 Bose Einstein statistics. How many ways can n indistinguishable particles be put into M boxes, with any number of particles per box? This type of counting is needed to predict the properties of particles called *bosons*, such as photons and He^4 atoms. Bose Einstein statistics counts the ways that n particles can be distributed in M different energy levels, when several particles can occupy the same quantum mechanical energy levels. For now, our interest is not in the physics, but just in the counting problem. Figure 1.3 shows that one way to count the number of arrangements is to think of the system as a linear array of n particles interspersed with $M - 1$ movable walls that partition the system into M boxes (spaces between walls). There are $M + n - 1$ objects, counting walls plus particles. The n particles are indistinguishable from each other. The $M - 1$ walls are indistinguishable from the other walls. Because the walls are distinguishable from the particles, the number of arrangements is

$$W(n, M) = \frac{(M + n - 1)!}{(M - 1)!n!}. \tag{1.20}$$

3 Probabilities of sequences.

Assume that the four bases A, C, T, and G occur with equal likelihood in a DNA sequence of nine monomers.

(a) What is the probability of finding the sequence AAATCGAGT through random chance?

(b) What is the probability of finding the sequence AAAAAAAAA through random chance?

(c) What is the probability of finding any sequence that has four A's, two T's, two G's, and one C, such as that in (a)?

(a) Each base occurs with probability $1/4$. The probability of an A in position 1 is $1/4$, of A in position 2 is $1/4$, of A in position 3 is $1/4$, of T in position 4 is $1/4$, and so on. There are 9 bases. The probability of this specific sequence is $(1/4)^9 = 3.8 \times 10^{-6}$.

(b) Same answer as (a) above.

(c) Each specific sequence has the probability given above, but in this case there are many possible sequences which satisfy the requirement that we have 4 A's, 2 T's, 2 G's, and 1 C. How many are there? We start as we have done before, by assuming all nine objects are distinguishable. There are 9! arrangements of nine distinguishable objects in a linear sequence. (The first one can be in any of nine places, the second in any of the remaining eight places, and so on.) But we can't distinguish the four A's, so we have overcounted by a factor of 4!, and must divide this out. We can't distinguish the two T's, so we have overcounted by 2!, and must also divide this out. And so on. So the probability of having this composition is

$$\left[\frac{9!}{4!2!2!1!}\right]\left(\frac{1}{4}\right)^9 = 0.014.$$