2.5 *Calculating the probability of an event: the sample-point method*

- To find the probability of an event by the sample-point method, use the following steps:
  - Define the experiment and clearly determine how to describe one simple event
  - List the simple events associated with the experiment and test each to make certain that it cannot be decomposed. (this defines the sample space S)
  - Assign reasonable probabilities to the sample points in S, making certain that $P(E_i)\geq 0$ and $\Sigma P(E_i)=1$.
  - Define the event of interest, A, as a specific collection of sample points. (A sample point is in A if A occurs when the sample point occurs. Test all sample points in S to identify those in A.)
  - Find P(A) by summing the probabilities of the sample points in A.
- Example 2.2. An employer "randomly" selects two applicants for a job from five employer, with competence 1 (best) to 5 (worst). Evaluate:
  P(the employer selects the best and one of the two poorest applicants)=
  P(the employer selects at least one of the two bests)=
- Example 2.3. In a balanced coin tossing, compute P(exactly 2 of 3 tosses are heads)
- Example 2.4. When A and B plays tennis, the odds are two to one that A wins. If A and B plays two match, evaluate P(A wins at least one match).
- The method is direct and powerful (a bulldozer approach)
- Difficulties:
  - Prone to human errors: forget some sample points in S
  - Many sample spaces contain a very large number of sample points
- Detours:
  - Many sample spaces generated by experimental data contain subset of sample points that are equiprobable.
  - Use a computer to list the large number of sample points and sum their probabilities
  - The mathematical theory of counting, called "combinatorial analysis"
- HW: some questions from exercises 2.18~ 2.26
- Keywords: the sample-point method

2.6 Tools for counting sample points

- Aim: count the total number of sample points in the sample space S and in an event of interest
  - If a sample space contains N equiprobable sample points and an event A contains exactly n_a sample points, $P(A)=n_a/N$.
- *Theorem 2.1. With m elements $a_1, a_2, \ldots, a_m$ and n elements $b_1, b_2, \ldots, b_n$ it is possible to form mn(=m×n) pairs containing one element from each group.*
  - The "mn rule" can be extended to any number of sets. Given three sets of elements $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_n$ and $c_1, c_2, \ldots, c_p$, the number of distinct triplets (a_i, b_j, c_k) containing one element from each set is equal to mnp. (Why?)
- Example 2.5. (Tossing two dice) In tossing a pair of dice and observing the numbers on the upper fces, find the # of sample points in the sample space S.
- Example 2.6. Redo 2.3 (coin tossing) using the mn rule.
- Example 2.7. (Birthday problem) Consider an experiment that consists of recording the birthday for each of 20 randomly selected persons. Ignoring leap years and assuming that there are only 365 possible distinct birthdays, find the # of points in the sample space S for this experiment. If we assume that each of the possible sets of birthdays is equiprobable, what's the probability that each person in the 20 has a different birthday?
  - A sample point is an ordered sequence of 20 numbers, each corresponding to the birthday of a person from 365 days.
  - The mn rule tells us that there are $(365)^{20}$ such twenty-tuples. The sample space S contains $N=(365)^{20}$ sample points.
  - It's not feasible to list all the sample points
  - But if we assume them equiprobable, $P(E_i)=1/(365)^{20}$ for each sample event.
  - Let A be the event that each person has a different birthday.
  - $P(A)=n_a/N$ where n_a=(# of sample points in A) = |A|
  - A={20-tuples such that no two positions contain the same number}
  - $n_a=365 \times 364 \times \ldots \times 356$
  - P(A) = .5886
- The sample points associated with an experience often can be represented symbolically as a sequence of numbers or symbols.
  - In some instances, the total # of sample points equals the # of distinct ways that the respective symbols can be arranged in sequence

- Definition 2.7. An ordered arrangement of r distinct objects is called a "permutation". The number of ways of ordering n distinct objects taken r at a time will be designated by the symbol ${}_nP_r$. (P_r^n in the text)
- *Theorem 2.2. ${}_nP_r = n(n-1)(n-2)\ldots(n-r+1) = n!/(n-r)!$*
  - Proof:
  - Note n!=n(n-1)…(2)(1) and 0!=1.
- Example 2.8. (Office ruffle) The names of 3 employees are randomly drawn "without replacement" from a bowl containing the nams of 30 employees. The 1st, 2nd, and 3rd person each receives $100, $50 and $25 respectively. How many sample points are associated with the experiment?
- Example 2.9. (Manufacturing sequence) An assembly operation involves four steps that can be performed in any sequence. To study the assembly time for each of the sequences, how many different sequences need to be considered?
- *Theorem 2.3. The number of ways of partitioning n distinct objects into k distinct groups containing n_1, n_2, …, n_k objects is* $N = \binom{n}{n_1 n_2 \ldots n_k} = \frac{n!}{n_1! n_2! \ldots n_k!}$
  - Proof: application of "mn rule" and previous theorem.
  - The terms are called "multinomial coefficients" since they occur in the expansion of the multinomial term: $(y_1 + y_2 + \ldots + y_k)^n = \sum_{\substack{n_i = 0,1,\ldots,n \\ n_1 + n_2 + \ldots + n_k = n}} \binom{n}{n_1 n_2 \ldots n_k} y_1^{n_1} y_2^{n_2} \ldots y_k^{n_k}$
- Ex. 2.10. racial discrimination dispute in job assignment
- In many situations, the sample points are arrays of symbols in which the order of symbols is unimportant. (see Ex. 2.2)
- Definition 2.8. The number of "combinations" of n objects taken r at a time is the number of subsets, each of size r, that can be formed from the n objects. This number will be denoted nCr, C_r^n or $\binom{n}{r}$.
- *Theorem 2.4. The number of unordered subsets of size r chosen (without replacement) from n available objects is* $\binom{n}{r} = C_r^n = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$
  - Proof: Use theorem 2.3.

  - The terms are called "binomial coefficients" because they occur in the "binomial expansion" $(x+y)^n = \sum_{i=0}^{n} \binom{n}{r} x^{n-i} y^i$
- Example 2.11. # of ways of selecting two applicants out of five (cf. example 2.2)
- Example 2.12. (continuation of 2.11) Compute # of sample points in A={exactly one of the two best applicants is selected} and P(A)
- Example 2.13. When placing n orders from M distributors, for a particular distributor say I, compute P(I gets exactly k orders).
- Warning: Combinatorial analysis is a huge area!
- HW: some questions from exercises 2.27~2.55
- Keywords: mn rule; permutation; partitioning; combinations; multinomial/binomial expansions