10. Combinatorics

Throughout this section we sample k elements from a set of n elements at random such that

(1) the order in which the k elements are sampled is or is not relevant, (**ordered or unordered sampling**)
(2) each sample is or is not placed back in the set, (**sampling with or without replacement**)

Ordered Sampling with Replacement

Each single draw is the same experiment of sampling one element of n. Hence there are a total of n^k different outcomes.

Example 10.1. *(1) How many different functions* $f : \{1, \dots, 50\} \to \{1, \dots, 100\}$ *are there? For each* $k \in \{1, \dots, 50\}$ *we have* $f(k) \in \{1, \dots, 100\}$*. Hence we draw* 50 *samples from a set of* 100 *elements, where the order matters as the first sample is* $f(1)$*, the second sample is* $f(2)$ *and so on, and we allow replacement as* $f(1)$ *may equal* $f(2)$*, for instance. Overall, there are* 100^{50} *different functions.*

(2) How many different subsets does $A = \{x_1, \dots, x_n\}$ *have? Given any subset* B *of* A *we ask for each* x_i *whether* $x_i \in B$*. Hence* B *corresponds to a unique ordered list of* n *'yes/no' entries. Overall there are* 2^n *different such lists and so* A *has* 2^n *subsets.*

Ordered Sampling without Replacement

We consider each individual draw as a separate experiment. Then the first experiment has n possible outcomes, which reduces by one, for each subsequent experiment. Hence, overall there are

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

different outcomes.

Example 10.2. *In a school all pupils are born in the years 2005, 2006 and 2007 (no leap years). If* k *pupils attend a school event, what is the probability of at least two sharing the same birthday, however, not necessarily the same day of birth?*

We label the kids from 1 to k and let Ω be the set of all ordered lists of k birthdays. Then $|\Omega| = 365^k$, where each outcome is equally likely. Next let A be the event that at least two pupils share a birthday. Then A^c is the event of no two pupils share a birthday. Note that $P(A) = 1 - P(A^c) = 1 - \frac{|A^c|}{|\Omega|}$.

Next observe that A^c is precisely those ordered lists in Ω that have no repeated birthdays. Hence $|A^c| = 365 \cdot \ldots \cdot (366 - k) = \frac{365!}{(365-k)!}$. Thus

$$P(A) = 1 - \frac{|A^c|}{|\Omega|} = 1 - \frac{365!}{365^k \cdot (365 - k)!}$$

How many pupils need to attend the school event for there to be at least a 50 percent chance of a shared birthday? One can calculate that $P(A) \approx 0.4757$, for $k = 22$ and $P(A) \approx 0.5073$, for $k = 23$.

Fun Fact: *At the 2014 Football World Cup there were 32 teams of 23 players each and in exactly 16 teams at least two players with a shared birthday.*

Unordered Sampling without Replacement

Theorem 10.3. *The number of total outcomes in this scenario is*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Proof. Let x be the number of total outcomes. For each unordered sampling of k elements without replacement there are $k!$ different ways to order the sampling. Hence $x \cdot k!$ equals the number of ordered samplings of k elements without replacement, which we know equals $\frac{n!}{(n-k)!}$. This gives the result. □

Example 10.4. *(1) We draw three cards at random from a stack of 52 cards. Let Ω be the set of all possible combinations ignoring their order. Then*

$$|\Omega| = \binom{52}{3} = \frac{52 \cdot 51 \cdot 50}{6} = 22100.$$

Note that all outcomes are equally likely. Next let A be the event that exactly one ace has been drawn. This is like drawing one card from the four aces and two cards from the remaining 48. Hence

$$|A| = \binom{48}{2} \cdot \binom{4}{1} = \frac{48 \cdot 47}{2} \cdot 4 = 4512.$$

Now $P(A) = \frac{4512}{22100} = 0.2042$.

(2) A radio DJ wants to play 12 songs over the next hour, 4 each from the 60s, 70s and 80s. How many ways are there to arrange the song list by decade? In other words, the DJ wants to choose 4 of 12 slots to play a 60s song, 4 of the remaining 8 slots to play a 70s song and then fill the final 4 slots with 80s songs. Hence there are

$$\binom{12}{4} \cdot \binom{8}{4} \cdot \binom{4}{4} = 495 \cdot 70 \cdot 1 = 34650$$

different arrangements by decade.

Lemma 10.5. *Let $a, b \in \mathbb{R}$ and $n \geq 1$ an integer. Then*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Proof. Clearly, every summand of $(a+b)^n$ is a multiple of $a^k b^{n-k}$, for some $k \in \{0, 1, \ldots, n\}$. This multiple is the number of ways that one can pick k times the scalar a from the n factors of $(a+b)^n$. □

Lemma 10.6. Vandermonde's identity*:*

$$\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{i} \cdot \binom{n}{k-i}$$

Proof. Let $S = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$. Note that the number of subsets of S with k elements is $\binom{m+n}{k}$. On the other hand any such subset will contain i elements in $\{a_1, \ldots, a_m\}$ and $k-i$ elements in $\{b_1, \ldots, b_n\}$, for some $i \in \{0, 1, \ldots, k\}$. But there are $\binom{m}{i} \cdot \binom{n}{k-i}$ for each i. This proves the claim. □

Unordered Sampling with Replacement

Theorem 10.7. *The number of total outcomes in this scenario is*

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}.$$

Proof. Assume we sample from the set $\{x_1, \ldots, x_n\}$. For each $i \in \{1, \ldots, n\}$, let k_i denote the number of times we have drawn x_i. Then $k_1 + \ldots + k_n = k$ and each such n-tuple $(k_1, \ldots, k_n)$ represents a unique outcome of the sampling experiment. Next we replace k_i by k_i vertical bars. Now the sum $k_1 + \ldots + k_n$ corresponds to a unique sequence of k bars and $n-1$ plus signs. For instance, $3+0+1$ corresponds to $|||++|$. Such a sequence has $n+k-1$ spots, from which we sample (unordered and without replacement) k spots to place the bars. $\square$

Example 10.8. *Twenty people take the bus from Galway to Dublin. The bus stops five times and the bus driver records the number of people getting off at each stop. How many different possibilities exist? Let* $S1$ *to* $S5$ *denote the five stops. For each passenger we draw a sample from the set* $\{S1, \ldots, S5\}$*, where we allow replacement but the order does not matter, (that is, the order in which each group gets off the bus at a particular stop). Hence there are*

$$\binom{5+20-1}{20} = \binom{24}{4} = 10,626$$

possible outcomes.