Frequency and Probability

In many contexts the above two *distinct* concepts are used inter-changeably. What is the mathematics behind this?

Let us examine the photograph below which has various gray shades arranged in a 960×600 grid. If a small stone is thrown at the picture with aiming anywhere in particular, one might believe that each of the $N = 576000$ "pixels" is equally likely to be hit. (Though you would do well to avoid the tiger!)

Figure 1: Image from wallpapersafari.com

Formulating this as a probabilistic event, we let $H(i, j)$ be the event that the chalk hits the pixel at position (i, j). These events are mutually exclusive, exhaustive and equally likely; so $P(H(i, j)) = 1/N$ whatever i and j may be. For each gray shade we have an associated number c between 0 and 255 which represents it. Let $G(c)$ denote the event that the chalk hits a pixel with shade c. It is clear that $G(c)$ is just the "union" ($\vee$) of all $H(i, j)$ where the (i, j)-th pixel has shade c. It now follows that $P(G(c))$ is just the frequency with which the shade c is seen among all the N pixels.

In summary, when we have a population of distinct individuals (in the above case, pixels) each of which has some measurable feature (in the above case, gray shade), then the frequency with which the feature is seen in the population is the same as the probability of finding this feature in an individual from the population *providing* that it is *equally likely* for *any* individual from the population to be examined.

The last part above is very important and is a point where many market surveys and poll analyses go wrong.

Mathematical version

Mathematically, we can put this as follows. We have a certain population S from which we are picking individuals in a way that it is *equally likely* for any individual to be picked. If $E(s)$ denotes the event that the individual we picked was s, then:

- the probability $P(E(s)) = 1/N$ where N is the total size of S
- the events $E(s)$ are mutually exclusive and exhaustive

Now, we are looking for a feature/property of the individuals. This is characterised by giving a subset T of the population consisting of those individuals for whom this property holds. The *event* that we pick an individual who has this property is $E(T) = \vee_{s\in T} E(s)$; i.e. the $\vee$ of the events $E(s)$ for s in T. We then get the probability $P(E(T))$ to be m/N where m is the *count* of the number of individuals who have the property. This follows easily from the two statements above.

Let us examine two examples.

In the first case, we are "randomly" picking a student from the MTH202 class and looking for the probability that the student is from Hostel 5.

Note that when we use the term "random choice" it usually means that we wish to employ some way that makes each student *equally likely* to be picked. Thus, if S is the set of students, the event $E(s)$ that the student s is picked has probability $P(E(s)) = 1/200$, assuming that the class has 200 students. We can then *count* the number of students in S who belong to Hostel 5; let us say that number turns out to be 40. The event F that we are looking for is that a randomly chosen student is from hostel 5; its probability $P(F)$ is $40/200 = 1/5$.

In the second example, we flip a fair coin k times (each time independent of the previous times) looking for the probability of exactly r heads. The total population S of possible outcomes is all sequences of length k out of the set $\{0, 1\}$ where 0 denotes "tail" and 1 denotes "head". The set S has 2^k elements since each position in the sequence (of length k) has 2 possibilities. The subset T_r we are interested in has all sequences where 1 occurs r times and 0 occurs $k - r$ times. The size of this set is given by the classical formula

$$|T_r| = \binom{k}{r} = \frac{k(k-1)\cdots(k-r+1)}{r(r-1)\cdots 1} = \frac{k!}{r!(k-r)!}$$

This can be seen as follows. Every element of T_r is a permutation of the sequence which has r 1's followed by $k - r$ 0's. There $k!$ permutations possible of a sequence of length k. Of these there $r!(k-r)!$ which fix this particular sequence. By what you have learned in MTH101 you can see that the count is correct!

The probability of getting r heads in k independent flips of a fair coin is therefore

$$\frac{\binom{k}{r}}{2^k}$$