1. Reading: K. Rosen *Discrete Mathematics and Its Applications*, 4.5

2. The main message of this lecture:

**Probability with not necessarily equally likely outcomes, conditional probability, independent events:** all have natural and extremely useful mathematical definitions.

The classical definition of probability (Laplace) assumes that the sample space is finite $S = \{s_1, s_2, \ldots, s_n\}$, that all the outcomes $s_i$ are equally likely and introduces the formula for the probability of an event $E$:

$$p(E) = \frac{|E|}{|S|}.$$

We can present the same formula in a somewhat more natural way, via the probabilities of individual outcomes. Note that the probability of each single outcome $p_i = p(s_i) = \frac{1}{n}$, and that $p_1 + p_2 + \ldots + p_n = 1$. Then $p(E)$ equals to the sum of those $p(s)$ for which $s \in E$:

$$p(E) = \sum_{s \in E} p(s) = \frac{|E|}{|S|}.$$

**Definition 20.1.** Imagine that $S$ is still finite $S = \{s_1, s_2, \ldots, s_n\}$, but the outcomes $s_i$ are not necessarily equally likely. We assume that the probability of individual outcomes $p_i = p(s_i)$ are given and that

1. $0 \leq p(s) \leq 1$ for each $s \in S$
2. $p(s_1) + p(s_2) + \ldots + p(s_n) = 1$.

Then the formula number two for the probability of an event $E$ still applies:

$$p(E) = \sum_{s \in E} p(s).$$

**Example 20.2.** Biased coin: heads come up twice as often as tails. $S = \{H, T\}$, $p(H) + p(T) = 1$, $p(H) = 2p(T)$, $2p(T) + p(T) = 1$, therefore, $3p(T) = 1$, $p(T) = 1/3$, $p(H) = 2/3$.

**Theorem 20.3.** $p(E_1 \cap E_2) = p(E_1) + p(E_2) - p(E_1 \cup E_2)$

**Proof.** Similar to the Inclusion-Exclusion Principle, in

$$p(E_1) + p(E_2) = \sum_{s \in E_1} p(s) + \sum_{s \in E_2} p(s)$$

each element of the intersection $E_1 \cap E_2$ is counted twice. Subtracting $p(E_1 \cap E_2)$ to compensate this overcount we get the desired formula for $p(E_1 \cup E_2)$.

**Corollary 20.4.** $p(\overline{E}) = 1 - p(E)$. Indeed, since $p(S) = 1$, by 20.3, we have $1 = p(S) = p(E \cup \overline{E}) = p(E) + p(\overline{E})$. Thus $p(\overline{E}) + p(E) = 1$ and $p(\overline{E}) = 1 - p(E)$.

**Example 20.5.** Flipping a fair coin the probability of having at least one $T$ (event $E$) is $7/8$. Suppose we know that the first flip came up heads (event $F = \{HTT, HTH, HHT, HHH\}$) and we want to evaluate the ”new” probability of $E$ given $F$. A new provisional sample set is
The probability of success $p$, is an experiment with two outcomes: Success (probability $p$) and Failure (probability $q = 1 - p$). Examples: fair coin $p = 1/2$, biased coin $p = 2/3$, $q = 1/3$, etc.

**Theorem 20.12.** The probability of $k$ successes in $n$ independent Bernoulli trials with probability of success $p$ is $C(n,k) \cdot p^k \cdot q^{n-k}$.

**Proof.** Each $n$-trail with $k$ successes can be labelled by a string of $k$ $S$’s and $(n - k)$ $F$’s with the probability of such a string $p^k q^{n-k}$. There are $C(n,k)$ such trials, the event $E$ consists of all of them, therefore, $p(E) = C(n,k)p^k q^{n-k}$.

**Homework assignments.** (due Friday 03/16).

20A:Rosen4.5-6; 20B:Rosen4.5-10; 20C:Rosen4.5-16; 20D:Rosen4.5-26ac.