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 $p_i = p(s_i) = 1/n$, and that $p_1 + p_2 + \ldots + p_n = 1$. Then $p(E)$ equals to the sum of those $p(s)$ for which $s \in E$:

$$p(E) = \sum_{s \in E} p(s) = |E| \cdot \frac{1}{n} = \frac{|E|}{|S|}.$$

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. $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$
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 now down to F, we have four equally likely outcomes. Some of the outcomes from E (namely THH, TTH, THT, TTT) are no longer possible. To evaluate the "conditional" probability of E given F we have to take the ratio of what is left of E to what is left of S:

$$\frac{|E \cap F|}{|F|} = \frac{|E \cap F|/|S|}{|F|/|S|} = \frac{p(E \cap F)}{p(F)} = \frac{3/8}{4/8} = 3/4.$$

Definition 20.6. Conditional probability $p(E|F)$ ("the probability of E given F") is

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

The reason for this definition is similar to the one used in example 20.5: given a condition F we may regard it as a new sample space and adjust the formula accordingly.

Example 20.7. What is the conditional probability that a family of three children has more than one boy given they have at least one boy. $E = \{BBB,\ BBG,\ BGB,\ GBB\}$, $F = \{BBB,\ BBG,\ BGB,\ GBB,\ BGG,\ GBG, GGB\}$. Then $E \cap F = E$, since $E \subset F$.

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{p(E)}{p(F)} = \frac{4/8}{7/8} = 4/7.$$

What if the condition F is "the first child is a girl"? Then $F = \{GBB,\ GBG,\ GGB,\ GGG\}$, $E \cap F = \{GBB\}$, $p(E|F) = p(E \cap F)/p(F) = (1/8)/(4/8) = 1/4$.

Definition 20.8. If a condition F does not change the probability of an event E, we say that E and F are **independent**: $p(E) = p(E|F)$. Note that in a full accord to intuition, independence is symmetric: E is independent from F if and only if F is independent from E:

$$p(E) = \frac{p(E \cap F)}{p(F)} \quad \Leftrightarrow \quad p(E) \cdot p(F) = p(E \cap F) \quad \Leftrightarrow \quad p(F) = \frac{p(E \cap F)}{p(E)},$$

therefore, $p(E) = p(E|F) \quad \Leftrightarrow \quad p(F) = p(F|E)$.

Example 20.9. A fair coin is flipped twice. E="the first flip come up tails"=$\{TT,\ TH\}$, F="the second flip comes up tails"=$\{TT,\ HT\}$. Intuitively, E and F are independent. Let us check the formula. On one hand, $p(E) = 1/2$. On the other hand, $p(E \cap F) = p(\{TT\}) = 1/4$, $p(E|F) = p(E \cap F)/p(F) = (1/4)/(1/2) = 1/2$. Alternatively, one could check the condition $p(E \cap F) = p(E) \cdot p(F)$: $1/4 = (1/2) \cdot (1/2)$.

Example 20.10. The events E="more then one boy out of three kids" and F="the first baby is a girl" are not independent. Indeed, $p(E \cap F) = p(\{GBB\}) = 1/8$, whereas $p(E) \cdot p(F) = (1/2) \cdot (1/2) = 1/4$. So, $p(E|F) = (1/8)/(1/2) = 1/4$, i.e. E given F is twice less likely then E.

Definition 20.11. Bernoulli trial is an experiment with two outcomes: Success (probability p) and Failure (probability $q = 1 - p$). Examples: fair coin $p = q = 1/2$, 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.