CS206 **Review Sheet 2** **Sept. 24, 2018**

COUNTING AND COMBINATORICS: One basic, intuitive idea underlies most of what we will do with counting. We phrase it in the context of random experiments.

Suppose experiment $\mathcal{E}_1$ is "choose an element from $S_1 = \{a_1, \ldots, a_m\}$" and experiment $\mathcal{E}_2$ is "choose an element from $S_2 = \{b_1, \ldots, b_n\}$". Clearly the sample spaces are S_1 (of size $= m$) and S_2 (of size $= n$), respectively. We take as an axiom the following cartesian product principle: the composite experiment "do $\mathcal{E}_1$ and *then* do $\mathcal{E}_2$" has sample space $S = S_1 \times S_2 = \{(x_1, x_2) :$ where $x_1 \in S_1$ is the outcome of the first experiment $\mathcal{E}_1$, and $x_2 \in S_2$ is the outcome of the second, $\mathcal{E}_2\}$. S is the cartesian product of S_1 and S_2 and its size is $|S| = mn = |S_1||S_2|$. We call S the "composite sample space".

By induction this implies that if

- (1) experiment $\mathcal{E}_1$ has sample space S_1 ($|S_1| = n_1$),
- (2) experiment $\mathcal{E}_2$ has sample space S_2 ($|S_2| = n_2$),
- etc.
- (k) experiment $\mathcal{E}_k$ has sample space S_k ($|S_k| = n_k$)

then the composite experiment "do $\mathcal{E}_1$, then do $\mathcal{E}_2$, ..., and finally, do $\mathcal{E}_k$" has sample space $S = S_1 \times S_2 \times \cdots \times S_k$, the $k-$fold cartesian product whose size is $n_1 n_2 \cdots n_k$. We apply this idea throughout, and in particular in the following broad models for random experiments.

1. **Ordered Sampling:** We study two distinct ways to perform ordered sampling from a given set.
   (a) with replacement. Here, we are given a set $T = \{t_1, \ldots, t_n\}$ of size n. The experiment $\mathcal{E}$ is to sample r times from T, each time writing down *which* item of T we chose, and then replacing that item before making the next pick.
   The sample space is $S = \{(x_1, \ldots, x_r)$ where $x_i \in T$ denotes what you sampled on the i^{th} pick, $i = 1, \ldots, r\}$. $|S| = n^r$. (In class we observed that this experiment can also be viewed as a model for placing r balls in n boxes - for ball i we choose a box, $x_i \in \{1, \ldots, n\}$).
   (b) without replacement (permutations). We have a set $T = \{t_1, \ldots, t_n\}$. We sample r times from T, but we do *not* replace the item chosen on any of the steps. This forces $r \leq n$. The sample space is $S = \{(s_1, \ldots, s_r) : s_i \in T$ is what you chose on the i^{th} sample, $i = 1, \ldots, r\}$. Clearly $s_i \neq s_j$ when $i \neq j$ because the r samples must be *distinct* items from T. It follows that $|S| = n(n-1)\cdots(n-r+1)$, a product with r terms. We write $(n)_r$ as a shorthand for $|S|$ and note that
   $$(n)_r = n(n-1)\cdots(n-r+1). \tag{1}$$
   When $r = n$ or $n-1$, $(n)_r$ is written $n!$, a notation that is called "n factorial". Clearly $(n)_r = n!/(n-r)!$, and it is easy to verify the relations
   $$e\left(\frac{n}{e}\right)^n \leq n! \leq n^n.$$

2. **Unordered Sampling (combinations):** The experiment is "choose a group of $r \leq n$ elements from a set $T = \{t_1, \ldots, t_n\}$". The sample space S consists of all the subsets of T of size r. We write its size using the notation $\binom{n}{r}$ and say "n choose r", or "binomial coefficient of n things taken r at a time".
   - Fact 1:
   $$|S| = \binom{n}{r} = \frac{(n)_r}{r!} = \frac{n!}{r!(n-r)!}).$$

- Fact 2:
$$\binom{n}{r} = \binom{n}{n-r}.$$

- Fact 3 (Pascal): If $n > r \geq 1$,
$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

- Fact 5 (Binomial Theorem): Let a and b be reals and n a positive integer. Then
$$\sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j} = (a+b)^n.$$
(Proof by induction; base case $n = 1$ is easy.)
If we take $a = b = 1$ we get

- Fact 4: $\sum_{j=0}^{n} \binom{n}{j} = 2^n$ (number of subsets of an n element set). Other interesting cases are when $1 = a = -b$ and when $a = 1, b = x$.

A variety of examples were given in order to illustrate the calculation of event probabilities in sample spaces with equally likely probability measure. Ordered and unordered sampling featured in important ways.

3. **Stirlings Formula for $n!$:** The Stirling approximation to the value of $n!$ is the function
$$s_n = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$
It is a good approximation to $n!$ in that the ratio $n!/s_n \to 1$ as $n \to \infty$. In fact
$$e^{\frac{1}{12n+1}} \leq \frac{n!}{s_n} \leq e^{\frac{1}{12n}}.$$

4. **Generalized Binomial Coefficients:** Let x be a given real number and $r > 0$ an integer. Analogous to (1) we define
$$(x)_r = x(x-1)\cdots(x-r+1),$$
a product with r terms; we will agree that $(x)_0 = 1$. Then, analogous to the equation in Fact 1, we define the binomial coefficient.
$$\binom{x}{r} = \frac{(x)_r}{r!}.$$

5. **Inclusion/Exclusion Principle:** We have n events, $A_1, \ldots, A_n$ in a probability space (S, P). Their union has probability
$$P(\bigcup_{i=1}^{n} A_i) = S_1 - S_2 + \cdots + (-1)^{k+1} S_k + \cdots + (-1)^{n+1} S_n,$$
where $S_1 = \sum_{i=1}^{n} P(A_i)$, the sum over the n distinct single events, A_i, $S_2 = \sum P(A_i \cap A_j)$, the sum over the $\binom{n}{2}$ distinct pairs, i, j of events, and in general,
$$S_k = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}),$$
the sum over the $\binom{n}{k}$ distinct k-tuples, $i_1, \ldots, i_k$ of events. Events which are good candidates to attack by inclusion/exclusion involve unions, and are described using the terms

- "none" $= \bigcap(A_i^c) = (\bigcup A_i)^c$
- "at least 1" = "not none"$= \bigcup A_i = (\bigcap(A_i^c))^c$
- "all"$= \bigcap(A_i) = (\bigcup(A_i^c))^c$
- "not all" $= (\bigcap A_i)^c = \bigcup(A_i^c)$

Example 1: A computer has 4 output devices. There are 6 jobs currently in the system. The experiment $\mathcal{E}$ is "each job requests an output device". The sample space for $\mathcal{E}$ (6 balls in 4 boxes)

$$S = \{(d_1, \dots, d_6) | d_i \in \{1, 2, 3, 4\} \text{ denotes the device requested by job } i\},$$

and $|S| = 4^6$. We compute the probability of $A = \{\text{ALL deviced are requested}\}$. Let $A_i = \{\text{device } i \text{ is NOT requested}\}$ and note that $A^c = A_1 \cup A_2 \cup A_3 \cup A_4$, so by (1),

$$P(A) = 1 - P(A_1 \cup A_2 \cup A_3 \cup A_4) = 1 - [S_1 - S_2 + S_3 - S_4].$$

Now note that $P(A_i) = 3^6/4^6$ for each $i = 1, \dots, 4$, that $P(A_i \cap A_j) = 2^6/4^6$ for each of the 6 i, j pairs, that $P(A_i \cap A_j \cap A_k) = 1/4^6$ for each of the 4 distinct triples and that $S_4 = 0$. Therefore $P(A) = 1 - [4(3^6/4^6) - 6(2^6/4^6) + 4(1/4^6)] = 1560/4096$. This can be computed directly (we did it in class - it took some insight) but inclusion/exclusion makes it *automatic*.

Example 2 (derangements): Consider the sample space for the hat-check experiment with n hats (an ordered sample of size n from a set of size n - WITHOUT replacement, so $S = \{(h_1, \dots, h_n) | h_i \in \{1, \dots, n\}$ the hat given to person $i; h_i \neq h_j, i \neq j\}$. S has $|S| = n!$ outcomes, one for each permutation of $1, \dots, n$. The derangement event, A, is the set of outcomes where NOBODY gets their own hat. Let A_i be the event that person i gets their own hat and note that $A^c = A_1 \cup \dots \cup A_n$ so,

$$P(A) = 1 - P(A^c) = 1 - P(A_1 \cup \dots \cup A_n)$$
$$= 1 - [S_1 - S_2 + \dots + (-1)^{n+1} S_n].$$

We compute S_k. For each $i_1 < \dots < i_k$, $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = (n-k)!/n!$ because after person i_1 gets *his* hat, i_2 gets his, ... i_k gets his, the other $n-k$ hats are distributed to the other $n-k$ people in $(n-k)!$ different ways. Since there are $\binom{n}{k}$ such k-tuples, each with probability $(n-k)!/n!$,

$$S_k = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

This gives

$$P(A) = 1 - 1 + \frac{1}{2!} + \dots + (-1)^n \frac{1}{n!} \equiv q_n. \tag{2}$$

This is the n^{th} partial sum of the Taylor series for $1/e = .3678794412...$ to which q_n converges rapidly (e.g., $q_5 = .3666666...$, $q_6 = .368055555...$, $q_7 = .3678571429...$). The surprising fact is that once $n > 7$, the derangement probability is essentially constant $(= 1/e)$.

Let B_k be the event that *exactly* k people get their own hats. B_0 is the derangement and it is obvious that $P(B_n) = (n!)^{-1}$, Since B_k is derangement for the $n-k$ people who do not get their own hats, it is easy to show that

$$P(B_k) = \frac{q_{n-k}}{k!};$$

note that $P(B_{n-1}) = 0$.

6. **Partitions:** The partitioning experiment $\mathcal{E}$ takes a set $T = \{t_1, \ldots, t_n\}$ and partitions it into k subsets $T_1, \ldots, T_k$, $|T_i| = n_i > 0$, $n_1 + \cdots + n_k = n$. The subsets come in a fixed order (i.e., the first, second, etc.) but their elements are unordered. The sample space S is the collection of distinct partitions and its size, $N = |S|$ satisfies

$$N = \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-(n_1+\cdots n_{k-1})}{n_k} = \frac{n!}{(n_1!)(n_2!)\ldots(n_k!)},$$

the multinomial coefficient. We can also interpret N as the number of distinguishable permutations of n items where n_1 are of one kind and are indistinguishable from each other, n_2 are of a second kind and are indistinguishable from each other, etc. This is because we will form the permutation by first placing the n_1 indistinguishable items from T_1 in any of the n free locations (so $\binom{n}{n_1}$ possiblities), then placing the n_2 indistinguishable items from T_2 in any of the $n - n_1$ remaining, free locations (so $\binom{n-n_1}{n_2}$ possibilities), etc.