1.4 Lecture 4 (10/Oct): upper and lower bounds

1.4.1 Bonferroni inequalities

The union bound is a special case of the Bonferroni inequalities:

Let \( A_1, \ldots, A_n \) be events in some probability space, and \( A_i^c \) their complements. For \( S \subseteq [n] \) let \( A_S = \cap_{i \in S} A_i \).

For \( 0 \leq j \leq n \) let \( \binom{n}{j} \) denote the subsets of \([n]\) of cardinality \( j \).

**Lemma 6** For \( j \geq 1 \) let (see Fig. 1.1):

\[
\begin{align*}
m_j &= \sum_{S \in \binom{[n]}{j}} \Pr(A_S) \\
M_k &= \sum_{j=1}^k (-1)^{j+1} m_j = \sum_{j=1}^k (-1)^{j+1} \sum_{J \subseteq [n], |J| = j} \Pr(A_J)
\end{align*}
\]

Then:

\[
M_2, M_4, \ldots \leq \Pr(\bigcup A_i) \leq M_1, M_3, \ldots
\]

Moreover, \( \Pr(\bigcup A_i) = M_n \); this is known as the inclusion-exclusion principle.

Comment: Often, but not always, larger values of \( k \) give improved bounds. See the problem set.

**Proof:** The sample space is partitioned into \( 2^n \) measurable sets

\[
B_S = (\bigcap_{i \in S} A_i) \cap (\bigcap_{i \in S} A_i^c)
\]

Note that \( A_S = \bigcup_{T \subseteq S} B_T \), which, since the \( B_T \)'s are disjoint, gives \( \Pr(A_S) = \sum_{T \subseteq S} \Pr(B_T) \).

\[
m_j = \sum_{S \in \binom{[n]}{j}} \Pr(A_S) = \sum_{S \in \binom{[n]}{j}} \sum_{T \subseteq S} \Pr(B_T) = \sum_T \Pr(B_T) \binom{|T|}{j}
\]

Figure 1.1: \( m_2 \) (left), \( M_2 \) (right)
\[ M_k = \sum_{j=1}^{k} (-1)^{j+1} m_j \]

\[ = \sum_T \Pr(B_T) \sum_{j=1}^{k} (-1)^{j+1} \binom{|T|}{j} \]

\[ = \sum_{T \neq \emptyset} \Pr(B_T) \sum_{j=1}^{k} (-1)^{j+1} \binom{|T|}{j} \quad \text{because } \binom{0}{j} = 0 \text{ for } j \geq 1. \]

Observe \( \Pr(\cup A_i) = 1 - \Pr(B_\emptyset) = \sum_{T \neq \emptyset} \Pr(B_T) \). So

\[ M_k - \Pr(\cup A_i) = \sum_{T \neq \emptyset} \Pr(B_T) \sum_{j=0}^{k} (-1)^{j+1} \binom{|T|}{j} \]

where we have inserted the needed \( - \Pr(B_T) \) for \( T \neq \emptyset \) by starting the internal summation from \( j = 0 \).

The inequalities now follow from the claim that for \( t \geq 1, \)

\[ \sum_{j=0}^{k} (-1)^{j+1} \binom{t}{j} = \begin{cases} 0 & k \geq t \\ \geq 0 & k \text{ odd} \\ \leq 0 & k \text{ even} \end{cases} \]  \tag{1.13} \]

(For the inclusion-exclusion principle, note that once \( k \geq n \), all \( t \) fall into the first category.)

The first line follows by expanding \( (1 - 1)^t \) (and noting that all terms \( t < j \leq k \) have \( \binom{t}{j} = 0 \)).

For the remaining two lines we use the identity

\[ \binom{t}{j} - \binom{t}{j-1} = \binom{t-1}{j} - \binom{t-1}{j-2} \]  \tag{1.14} \]

(which holds for \( t, j \geq 1 \) with the interpretation \( \binom{a}{b} = 0 \) for \( a \geq 0, b < 0 \)).

Therefore when we group adjacent pairs \( j \) in the summation on the LHS of (1.13) (that is, \( \{k, k-1\}, \{k-2, k-3\}, \text{etc.} \), with 0 unpaired for \( k \) even), we obtain a telescoping sum, and so we have

For \( k \) odd:

\[ \sum_{j=0}^{k} (-1)^{j+1} \binom{t}{j} = \binom{t-1}{k} - \binom{t-1}{k-1} \geq 0 \]

For \( k \) even:

\[ \sum_{j=0}^{k} (-1)^{j+1} \binom{t}{j} = -\binom{t-1}{k} + \binom{t-1}{0} - \binom{t}{0} = -\binom{t-1}{k} \leq 0 \]

Comment: inclusion-exclusion is a special case of what is known in order theory as Möbius inversion.

### 1.4.2 Tail events: Borel-Cantelli

Here is a very fundamental application of the union bound.

**Definition 7** Let \( B = \{B_1, \ldots\} \) be a countable collection of events. \( \lim \sup B \) is the event that infinitely many of the events \( B_i \) occur.
Lemma 8 (Borel Cantelli I) Let $\sum_{i \geq 1} \Pr(B_i) < \infty$. Then $\Pr(\lim \sup B) = 0$.

$\lim \sup B$ is what is called a *tail event*: a function of infinitely many other events (in this case the $B_1, \ldots$) that is unaffected by the outcomes of any finite subset of them.

**Proof:** It is helpful to write $\lim \sup B$ as

$$\lim \sup B = \bigcap_{i \geq 0} \bigcup_{j \geq i} B_j.$$  

For every $i$, $\lim \sup B \subseteq \bigcup_{j \geq i} B_j$, so $\Pr(\lim \sup B) \leq \inf_i \Pr(\bigcup_{j \geq i} B_j)$. By the union bound, the latter is $\leq \inf_i \sum_{j \geq i} \Pr(B_j) = 0$. $\square$

### 1.4.3 B-C II: a partial converse to B-C I

Lemma 8 does not have a “full” converse.

To show a counterexample, we need to come up with events $B_i$ for which $\sum_{i \geq 1} \Pr(B_i) = \infty$ but $\Pr(\lim \sup B) = 0$. Here is an example. Pick a point $x$ uniformly from the unit interval. Let $B_i$ be the event $x < 1/i$.

You will notice that in this example the events are not independent. That is crucial, for B-C I does have the partial converse:

**Lemma 9 (Borel Cantelli II)** Suppose that $B_1, \ldots$ are independent events and that $\sum_{i \geq 1} \Pr(B_i) = \infty$. Then $\Pr(\lim \sup B) = 1$.

**Proof:** We’ll show that $(\lim \sup B)^c$, the event that only finitely many $B_i$ occur, occurs with probability 0. Write $(\lim \sup B)^c = \bigcup_{i \geq 0} \bigcap_{j \geq i} B_j^c$.

By the union bound (Cor. 3), it is enough to show that $\Pr(\bigcap_{j \geq i} B_j^c) = 0$ for all $i$. Of course, for any $I \geq i$, $\Pr(\bigcap_{j \geq i} B_j^c) \leq \Pr(\bigcap_{j \geq I} B_j^c)$.

By independence, $\Pr(\bigcap_{j \geq i} B_j^c) = \prod_{j \geq i} \Pr(B_j^c)$, so what remains to show is that

$$\lim_{I \to \infty} \prod_{i} \Pr(B_j^c) = 0. \quad (1.15)$$

(Note the LHS is decreasing in $I$.)

There’s a classic inequality we often use:

$$1 + x \leq e^x \quad (1.16)$$

which follows because the RHS is concave and the two sides agree in value and first derivative at a point (namely at $x = 0$).

Consequently if a finite sequence $x_i$ satisfies $\sum x_i \geq 1$ then $\prod (1 - x_i) \leq 1/e$.

Supposing (1.15) is false, fix $i$ for which it fails, let $q_i$ be the limit of the LHS, and let $I$ be sufficient that $\prod_{j=1}^I \Pr(B_j^c) \leq 2q_i$. Let $I'$ be sufficient that $\sum_{j=1}^{I'} \Pr(B_j) \geq 1$. Then $\prod_{j=1}^{I'} \Pr(B_j^c) \leq 2q_i/e$. Contradiction.