## 4. Double counting

The *double counting* principle states the following "obvious" fact: if the elements of a set are counted in two different ways, the answers are the same.

In terms of matrices the principle is as follows. Let $M$ be an $n \times m$ matrix with entries 0 and 1. Let $r_i$ be the number of 1s in the $i$-th row, and $c_j$ be the number of 1s in the $j$-th column. Then

$$\sum_{i=1}^{n} r_i = \sum_{j=1}^{m} c_j = \text{the total number of 1s in } M\,.$$

The next example is a standard demonstration of double counting. Suppose a finite number of people meet at a party and some shake hands. Assume that no person shakes his or her own hand and furthermore no two people shake hands more than once.

HANDSHAKING LEMMA. *At a party, the number of guests who shake hands an odd number of times is even.*

PROOF. Let $P_1, \ldots, P_n$ be the persons. We apply double counting to the set of ordered pairs $(P_i, P_j)$ for which $P_i$ and $P_j$ shake hands with each other at the party. Let $x_i$ be the number of times that $P_i$ shakes hands, and $y$ the total number of handshakes that occur. On one hand, the number of pairs is $\sum_{i=1}^{n} x_i$, since for each $P_i$ the number of choices of $P_j$ is equal to $x_i$. On the other hand, each handshake gives rise to two pairs $(P_i, P_j)$ and $(P_j, P_i)$; so the total is $2y$. Thus $\sum_{i=1}^{n} x_i = 2y$. But, if the sum of $n$ numbers is even, then evenly many of the numbers are odd. (Because if we add an odd number of odd numbers and any number of even numbers, the sum will be always odd). $\square$

This lemma is also a direct consequence of the following general identity, whose special version for graphs was already proved by Euler. For a point $x$, its *degree* or *replication number* $d(x)$ in a family $\mathcal{F}$ is the number of members of $\mathcal{F}$ containing $x$.

PROPOSITION 1.7. *Let $\mathcal{F}$ be a family of subsets of some set $X$. Then*

$$\sum_{x \in X} d(x) = \sum_{A \in \mathcal{F}} |A|\,. \tag{10}$$

PROOF. Consider the *incidence matrix* $M = (m_{x,A})$ of $\mathcal{F}$. That is, $M$ is a 0-1 matrix with $|X|$ rows labeled by points $x \in X$ and with $|\mathcal{F}|$ columns labeled by sets $A \in \mathcal{F}$ such that $m_{x,A} = 1$ if and only if $x \in A$. Observe that $d(x)$ is exactly the number of 1s in the $x$-th row, and $|A|$ is the number of 1s in the $A$-th column. $\square$

Graphs are families of 2-element sets, and the degree of a vertex $x$ is the number of edges incident to $x$, i.e., the number of vertices in its neighborhood. Proposition **??** immediately implies

THEOREM 1.8 (Euler 1736). *In every graph the sum of degrees of its vertices is two times the number of its edges, and hence, is even.*

The following identities can be proved in a similar manner (we leave their proofs as exercises):

$$\sum_{x \in Y} d(x) = \sum_{A \in \mathcal{F}} |Y \cap A| \quad \text{for any } Y \subseteq X. \tag{11}$$

$$\sum_{x \in X} d(x)^2 = \sum_{A \in \mathcal{F}} \sum_{x \in A} d(x) = \sum_{A \in \mathcal{F}} \sum_{B \in \mathcal{F}} |A \cap B|\,. \tag{12}$$

*Turán's number* $T(n,k,l)$ $(l \le k \le n)$ is the smallest number of $l$-element subsets of an $n$-element set $X$ such that every $k$-element subset of $X$ contains at least one of these sets.

PROPOSITION 1.9. *For all positive integers $l \le k \le n$,*

$$T(n,k,l) \ge \binom{n}{l} \Big/ \binom{k}{l}.$$

PROOF. Let $\mathcal{F}$ be a smallest $l$-uniform family over $X$ such that every $k$-subset of $X$ contains at least one member of $\mathcal{F}$. Take a 0-1 matrix $M = (m_{A,B})$ whose rows are labeled by sets $A$ in $\mathcal{F}$, columns by $k$-element subsets $B$ of $X$, and $m_{A,B} = 1$ if and only if $A \subseteq B$.

Let $r_A$ be the number of 1s in the $A$-th row and $c_B$ be the number of 1s in the $B$-th column. Then, $c_B \ge 1$ for every $B$, since $B$ must contain at least one member of $\mathcal{F}$. On the other hand, $r_A$ is precisely the number of $k$-element subsets $B$ containing a fixed $l$-element set $A$; so $r_A = \binom{n-l}{k-l}$ for every $A \in \mathcal{F}$. By the double counting principle,

$$|\mathcal{F}| \cdot \binom{n-l}{k-l} = \sum_{A \in \mathcal{F}} r_A = \sum_{B} c_B \ge \binom{n}{k},$$

which yields

$$T(n,k,l) = |\mathcal{F}| \ge \binom{n}{k} \Big/ \binom{n-l}{k-l} = \binom{n}{l} \Big/ \binom{k}{l},$$

where the last equality is another property of binomial coefficients . □

Our next application of double counting is from number theory: How many numbers divide at least one of the first $n$ numbers $1, 2, \dots, n$? If $t(n)$ is the number of divisors of $n$, then the behavior of this function is rather non-uniform: $t(p) = 2$ for every prime number, whereas $t(2^m) = m+1$. It is therefore interesting that the *average* number

$$\tau(n) = \frac{t(1) + t(2) + \cdots + t(n)}{n}$$

of divisors is quite stable: It is about $\ln n$.

PROPOSITION 1.10. $|\tau(n) - \ln n| \le 1$.

PROOF. To apply the double counting principle, consider the 0-1 $n \times n$ matrix $M = (m_{ij})$ with $m_{ij} = 1$ iff $j$ is divisible by $i$:

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 |
| 3 | | | 1 | | | 1 | | | 1 | | | 1 |
| 4 | | | | 1 | | | | 1 | | | | 1 |
| 5 | | | | | 1 | | | | | 1 | | |
| 6 | | | | | | 1 | | | | | | 1 |
| 7 | | | | | | | 1 | | | | | |
| 8 | | | | | | | | 1 | | | | |

The number of 1s in the $j$-th column is exactly the number $t(j)$ of divisors of $j$. So, summing over columns we see that the total number of 1s in the matrix is $T_n = t(1) + \cdots + t(n)$.

On the other hand, the number of 1s in the $i$-th row is the number of multipliers $i, 2i, 3i, \dots, ri$ of $i$ such that $ri \le n$. Hence, we have exactly $\lfloor n/i \rfloor$ ones in the $i$-th row. Summing over rows, we obtain that $T_n = \sum_{i=1}^{n} \lfloor n/i \rfloor$. Since $x - 1 < \lfloor x \rfloor \le x$ for every real number $x$, we obtain that

$$H_n - 1 \le \tau(n) = \frac{1}{n} T_n \le H_n,$$

where

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \ln n + \gamma_n, \qquad 0 \le \gamma_n \le 1 \tag{13}$$

is the $n$-th harmonic number. □

## 6. The inclusion-exclusion principle

The *principle of inclusion and exclusion* (sieve of Eratosthenes) is a powerful tool in the theory of enumeration as well as in number theory. This principle relates the cardinality of the union of certain sets to the cardinalities of intersections of some of them, these latter cardinalities often being easier to handle.

For any two sets $A$ and $B$ we have

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

In general, given $n$ subsets $A_1, \ldots, A_n$ of a set $X$, we want to calculate the number $|A_1 \cup \cdots \cup A_n|$ of points in their union. As the first approximation of this number we can take the sum

$$|A_1| + \cdots + |A_n|. \tag{17}$$

However, in general, this number is too large since if, say, $A_i \cap A_j \neq \emptyset$ then each point of $A_i \cap A_j$ is counted two times in (**??**): once in $|A_i|$ and once in $|A_j|$. We can try to correct the situation by subtracting from (**??**) the sum

$$\sum_{1 \le i < j \le n} |A_i \cap A_j|. \tag{18}$$

But then we get a number which is too small since each of the points in $A_i \cap A_j \cap A_k \neq \emptyset$ is counted three times in (**??**): once in $|A_i \cap A_j|$, once in $|A_j \cap A_k|$, and once in $|A_i \cap A_k|$. We can therefore try to correct the situation by adding the sum

$$\sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k|,$$

but again we will get a too large number, etc. Nevertheless, it turns out that after $n$ steps we will get the correct result. This result is known as the *inclusion-exclusion principle.* The following notation will be handy: if $I$ is a subset of the index set $\{1, \ldots, n\}$, we set

$$A_I := \bigcap_{i \in I} A_i,$$

with the convention that $A_\emptyset = X$.

Proposition 1.13 (Inclusion-Exclusion Principle). *Let $A_1, \ldots, A_n$ be subsets of $X$. Then the number of elements of $X$ which lie in none of the subsets $A_i$ is*

$$\sum_{I \subseteq \{1, \ldots, n\}} (-1)^{|I|} |A_I|. \tag{19}$$

Proof. The sum is a linear combination of cardinalities of sets $A_I$ with coefficients $+1$ and $-1$. We can re-write this sum as

$$\sum_I (-1)^{|I|} |A_I| = \sum_I \sum_{x \in A_I} (-1)^{|I|} = \sum_x \sum_{I : x \in A_I} (-1)^{|I|}.$$

We calculate, for each point of $X$, its contribution to the sum, that is, the sum of the coefficients of the sets $A_I$ which contain it.

First suppose that $x \in X$ lies in none of the sets $A_i$. Then the only term in the sum to which $x$ contributes is that with $I = \emptyset$; and this contribution is 1.

Otherwise, the set $J := \{i : x \in A_i\}$ is non-empty; and $x \in A_I$ precisely when $I \subseteq J$. Thus, the contribution of $x$ is

$$\sum_{I \subseteq J} (-1)^{|I|} = \sum_{i=0}^{|J|} \binom{|J|}{i} (-1)^i = (1-1)^{|J|} = 0$$

by the binomial theorem.

Thus, points lying in no set $A_i$ contribute 1 to the sum, while points in some $A_i$ contribute 0; so the overall sum is the number of points lying in none of the sets, as claimed. $\square$

For some applications the following form of the inclusion-exclusion principle is more convenient.

PROPOSITION 1.14. *Let $A_1, \ldots, A_n$ be a sequence of (not necessarily distinct) sets. Then*

$$|A_1 \cup \cdots \cup A_n| = \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} |A_I| \,. \tag{20}$$

PROOF. The left-hand of (**??**) is $|A_\emptyset|$ minus the number of elements of $X = A_\emptyset$ which lie in none of the subsets $A_i$. By Proposition **??** this number is

$$|A_\emptyset| - \sum_{I \subseteq \{1,\ldots,n\}} (-1)^{|I|} |A_I| = \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} |A_I| \,,$$

as desired. □

## 5. Density of 0-1 matrices

Let $H$ be an $m \times n$ 0-1 matrix. We say that $H$ is *$\alpha$-dense* if at least an $\alpha$-fraction of all its $mn$ entries are 1s. Similarly, a row (or column) is *$\alpha$-dense* if at least an $\alpha$-fraction of all its entries are 1s.

The next result says that any dense 0-1 matrix must either have one "very dense" row or there must be many rows which are still "dense enough."

LEMMA 2.13 (Grigni and Sipser 1995). *If $H$ is $2\alpha$-dense then either*
(a) *there exists a row which is $\sqrt{\alpha}$-dense, or*
(b) *at least $\sqrt{\alpha} \cdot m$ of the rows are $\alpha$-dense.*

Note that $\sqrt{\alpha}$ is larger than $\alpha$ when $\alpha < 1$.

PROOF. Suppose that the two cases do not hold. We calculate the density of the entire matrix. Since (b) does not hold, less than $\sqrt{\alpha} \cdot m$ of the rows are $\alpha$-dense. Since (a) does not hold, each of these rows has less than $\sqrt{\alpha} \cdot n$ 1s; hence, the fraction of 1s in $\alpha$-dense rows is strictly less than $(\sqrt{\alpha})(\sqrt{\alpha}) = \alpha$. We have at most $m$ rows which are not $\alpha$-dense, and each of them has less than $\alpha n$ ones. Hence, the fraction of 1s in these rows is also less than $\alpha$. Thus, the total fraction of 1s in the matrix is less than $2\alpha$, contradicting the $2\alpha$-density of $H$. □

Now consider a slightly different question: if $H$ is $\alpha$-dense, how many of its rows *or* columns are "dense enough"? The answer is given by the following general estimate due to Johan Håstad. This result appeared in the paper of Karchmer and Wigderson (1990) and was used to prove that the graph connectivity problem cannot be solved by monotone circuits of logarithmic depth.

Suppose that our universe is a Cartesian product $A = A_1 \times \cdots \times A_k$ of some finite sets $A_1, \ldots, A_k$. Hence, elements of $A$ are strings $\boldsymbol{a} = (a_1, \ldots, a_k)$ with $a_i \in A_i$. Fix now a subset of strings $H \subseteq A$ and a point $b \in A_i$. The *degree* of $b$ in $H$ is the number $d_H(b) = |\{\boldsymbol{a} \in H : a_i = b\}|$ of strings in $H$ whose $i$-th coordinate is $b$.

Say that a point $b \in A_i$ from the $i$-th set is *popular* in $H$ if its degree $d_H(b)$ is at least a $1/2k$ fraction of the average degree of an element in $A_i$, that is, if

$$d_H(b) \geq \frac{1}{2k} \frac{|H|}{|A_i|} \,.$$

Let $P_i \subseteq A_i$ be the set of all popular points in the $i$-th set $A_i$, and consider the Cartesian product of these sets:

$$P := P_1 \times P_2 \times \cdots \times P_k \,.$$

LEMMA 2.14 (Håstad). $|P| > \frac{1}{2}|H|$.

PROOF. It is enough to show that $|H \setminus P| < \frac{1}{2}|H|$. For every non-popular point $b \in A_i$, we have that

$$|\{\boldsymbol{a} \in H \,:\, a_i = b\}| < \frac{1}{2k}\frac{|H|}{|A_i|}\,.$$

Since the number of non-popular points in each set $A_i$ does not exceed the total number of points $|A_i|$, we obtain

$$\begin{aligned}|H \setminus P| &\leq \sum_{i=1}^{k}\sum_{b \notin P_i} |\{\boldsymbol{a} \in H \,:\, a_i = b\}| < \sum_{i=1}^{k}\sum_{b \notin P_i} \frac{1}{2k}\frac{|H|}{|A_i|} \\ &\leq \sum_{i=1}^{k} \frac{1}{2k}|H| = \frac{1}{2}|H|\,. \quad \square\end{aligned}$$

COROLLARY 2.15. *In any $2\alpha$-dense 0-1 matrix $H$ either a $\sqrt{\alpha}$-fraction of its rows or a $\sqrt{\alpha}$-fraction of its columns (or both) are $(\alpha/2)$-dense.*

PROOF. Let $H$ be an $m \times n$ matrix. We can view $H$ as a subset of the Cartesian product $[m] \times [n]$, where $(i, j) \in H$ iff the entry in the $i$-th row and $j$-th column is 1. We are going to apply Lemma **??** with $k = 2$. We know that $|H| \geq 2\alpha mn$. So, if $P_1$ is the set of all rows with at least $\frac{1}{4}|H|/|A_1| = \alpha n/2$ ones, and $P_2$ is the set of all columns with at least $\frac{1}{4}|H|/|A_2| = \alpha m/2$ ones, then Lemma **??** implies that

$$\frac{|P_1|}{m} \cdot \frac{|P_2|}{n} \geq \frac{1}{2}\frac{|H|}{mn} \geq \frac{1}{2} \cdot \frac{2\alpha mn}{mn} = \alpha\,.$$

Hence, either $|P_1|/m$ or $|P_2|/n$ must be at least $\sqrt{\alpha}$, as claimed. $\square$