Here are some notes about Kolmogorov complexity, to expand on what was said in class.

Lemma 0.1. *For every $n \geq 0$, there exists a string x of length n such that x is incompressible (meaning $K(x) \geq |x|$).*

Proof. There are 2^n strings of length n, but only $2^n - 1$ strings of length less than n. So there is no way that the strings of length less than n could unambiguously encode all of the strings of length n. $\square$

Theorem 0.2. *The Kolmogorov complexity K is not computable.*

Proof. Suppose (for the sake of contradiction) that K is computable. Let M be a Turing machine that on any input x halts with $K(x)$ on its tape. Use M to construct a Turing machine N that, on any input n (regarded as a base-2 integer), outputs some string x satisfying $K(x) \geq n$. (For example, N could try all strings of length n in lexicographic order, using M to compute K for each, until it found x with $K(x) \geq n$. The preceding lemma guarantees that such an x will be found.) Let m be any integer such that

$$m - \lceil \log_2 m \rceil - 1 > |N| + |\#|,$$

and let $x = N(m)$. Notice that m, when written in base 2, requires no more than $\lceil \log_2 m \rceil + 1$ bits. Thus $N\#m$ is a description of x, of length

$$\begin{aligned} |N\#m| &\leq |N| + |\#| + \lceil \log_2 m \rceil + 1 \\ &< m, \end{aligned}$$

by the definition of m. Therefore $K(x) < m$. But the definition of N guarantees that $K(x) \geq m$. From this contradiction we conclude that our initial assumption, that K is computable, was false. $\square$

Definition 0.3. *A **property** of strings over Σ is a function $f : \Sigma^* \to \{\mathrm{T}, \mathrm{F}\}$. A property f **holds for almost all strings** if*

$$\lim_{n \to \infty} \frac{\#\{x : |x| = n, f(x) = \mathrm{F}\}}{\#\{x : |x| = n\}} = 0.$$

The following mathematical lemma shows that we can replace "=" with "≤" in the above definition. Sipser uses this fact without proof. You may want to skip the proof on a first reading.

Lemma 0.4. *Let f be a property that holds for almost all strings. Then*

$$\lim_{n \to \infty} \frac{\#\{x : |x| \leq n, f(x) = \mathrm{F}\}}{\#\{x : |x| \leq n\}} = 0.$$

Proof. Let $\epsilon > 0$. We wish to show that there exists N such that for all $n \geq N$

$$\frac{\#\{x : |x| \leq n, f(x) = \mathrm{F}\}}{\#\{x : |x| \leq n\}} < \epsilon.$$

For the sake of brevity, let $L_n = \#\{x : |x| = n, f(x) = \mathrm{F}\}$. Because f holds for almost all strings, there exists an M such that for all $n > M$,

$$\frac{\#\{x : |x| = n, f(x) = \mathrm{F}\}}{\#\{x : |x| = n\}} < \frac{\epsilon}{2}.$$

That is, $L_n < \frac{\epsilon}{2} 2^n$ for all $n > M$. Pick N large enough so that

$$\sum_{i=0}^{M} L_i < \frac{\epsilon}{2} \left(2^{N+1} - 1\right).$$

Then for all $n \geq N$

$$\begin{aligned}
\#\{x : |x| \leq n, f(x) = \mathrm{F}\} &= \sum_{i=0}^{M} L_i + \sum_{i=M+1}^{n} L_i \\
&< \sum_{i=0}^{M} L_i + \sum_{i=M+1}^{n} \frac{\epsilon}{2} 2^i \\
&< \frac{\epsilon}{2}\left(2^{N+1} - 1\right) + \frac{\epsilon}{2}(2^{n+1} - 1) \\
&\leq \epsilon\left(2^{n+1} - 1\right) \\
&= \epsilon\, \#\{|x| \leq n\}.
\end{aligned}$$

This proves the lemma. □

Intuitively, a string generated at random should have no pattern and should not be compressible. The following theorem makes this intuition precise.

Theorem 0.5. *Let f be a computable property that holds for almost all strings. Let $b > 0$. Then $f(x) = \mathrm{F}$ for only finitely many strings that are incompressible by b.*

Proof. If f is false on only finitely many strings, then the theorem is obviously true. Henceforth assume that f is false on infinitely many strings. Denote these strings $s_0, s_1, s_2, \ldots$ in lexicographic order.

For any string x in the sequence $s_0, s_1, s_2, \ldots$, let i_x be its index in the list. That is, i_x is the unique number such that $s_{i_x} = x$. Let M be a Turing machine that on input i, regarded as a base-2 integer, outputs s_i. Then $M\#i_x$ is a description of x.

Fix $b > 0$. By the lemma, there exists a large N so that for all $n \geq N$

$$\frac{\#\{x : |x| \leq n, f(x) = \mathrm{F}\}}{\#\{x : |x| \leq n\}} < \frac{1}{2^{b+|M|+|\#|+1}}.$$

Using the fact that $\#\{x : |x| \leq n\} = 2^{n+1} - 1$, we have

$$\#\{x : |x| \leq n, f(x) = \mathrm{F}\} < \frac{2^{n+1}}{2^{b+|M|+|\#|+1}} = 2^{n-b-|M|-|\#|}.$$

If x is any string of length $n \geq N$ such that $f(x) = \mathrm{F}$, then $i_x < 2^{n-b-|M|-|\#|}$ and $|i_x| \leq n-b-|M|-|\#|$. This implies that

$$K(x) \leq |M\#i_x| \leq |M| + |\#| + n - b - |M| - |\#| = n - b.$$

So x is compressible by b.

We have shown that any string x of length at least N that fails f is compressible by b. There are only finitely many strings of length less than N. Therefore only finitely many x that fail f can be incompressible by b. □