4. The First and Second Moment Methods

- **Theorem 4.1** (Markov's Inequality). Let $X \geq 0$ be a nonnegative random variable with $E[X] < \infty$. Then for any $a > 0$,

$$\Pr(X \geq a) \leq E[X]/a.$$

Proof. In general terms:

$$E[X] = \int_0^\infty x\,dP = \int_0^a x\,dP + \int_a^\infty x\,dP \geq 0 + a\int_a^\infty dP$$
$$= a\Pr(X \geq a).$$

$$\Rightarrow \Pr(X \geq a) \leq E[X]/a.$$

[For clarity, let us review the proof in case X is integer-valued and $a \in \mathbb{N}_+$:

$$E[X] = \sum_{k=0}^{\infty} k\Pr(X=k) = \sum_{k=0}^{a-1} k\Pr(X=k) + \sum_{k=a}^{\infty} k\Pr(X=k)$$
$$\geq 0 + a\sum_{k=a}^{\infty}\Pr(X=k) = a\Pr(X \geq a).$$]

$\square$

- Corollary 1. If $X \geq 0$ is an integer-valued random variable with $E[X] = \mu$, then

$$\Pr(X \geq 1) \leq \mu, \quad \text{i.e.} \quad \Pr(X=0) \geq 1-\mu. \quad \square$$

- Let then X_n, $n \geq 0$, be a sequence of random variables counting the number of occurrences of some feature in a random structure of size n. If $\mu_n = E[X_n] \to 0$ as $n \to \infty$, then the feature almost surely doesn't occur at all in sufficiently large structures.
- This the "first moment method".

- Let's say that we to the contrary want to show that some interesting feature does occur almost surely in sufficiently large random structures, i.e. that

$$\Pr(X_n > 0) \to 1 \quad \text{as} \quad n \to \infty.$$

Now even $E[X_n] \to \infty$ isn't sufficient. (Consider e.g. a case where there are n equally probable structures of size n, one of which has 2^n occurrences of the feature and others 0.)

- However even this works out if the variance $\sigma_n^2 = \mathrm{Var}[X_n]$ stays small (more specifically if $\sigma_n^2/\mu_n^2 \to 0$).

- This is the "second moment method".

- Theorem 4.2 (Chebyshev's Inequality) Let X be a random variable with $E[X] = \mu$, $\mathrm{Var}[X] = \sigma^2 > 0$. Then for any $t > 0$,

$$\Pr(|X-\mu| \geq t\sigma) \leq \frac{1}{t^2}.$$

Proof. By definition, $\sigma^2 = E[(X-\mu)^2]$. Now the random variable $(X-\mu)^2$ is nonnegative, so by Markov's inequality:

$$\begin{aligned}\Pr(|X-\mu| \geq t\sigma) &= \Pr((X-\mu)^2 \geq t^2\sigma^2) \\ &\leq E[(X-\mu)^2] / t^2\sigma^2 \\ &= \frac{1}{t^2}. \quad \square\end{aligned}$$

- Corollary 2. If $\mu_n = E[X_n] > 0$ for n large, and $\frac{\sigma_n^2}{\mu_n^2} \to 0$ as $n \to \infty$, then $\Pr(X_n > 0) \to 1$ as $n \to \infty$.

Proof. If $X_n = 0$, then $|X_n - \mu_n| = \mu_n$. Choose $t_n = \frac{\mu_n}{\sigma_n}$ above; then:
$\Pr(X_n = 0) \leq \Pr(|X_n - \mu_n| \geq \mu_n) \leq \sigma_n^2/t_n^2 \to 0$ as $n \to \infty$. $\square$

- <u>Example 1.</u> Random graphs and threshold functions.

- Consider the family $\mathcal{G}_n$ of all (labelled, undirected) graphs on the vertex set $[n] = \{1, \dots, n\}$. Denote $N = \binom{n}{2}$. Then $|\mathcal{G}_n| = 2^N$.

  For a given $p \in [0,1]$, make $\mathcal{G}_n$ into a probability space $\mathcal{G}(n,p)$, by picking each edge uniformly and independently at random with prob. p.

  Thus, for any specific graph $H \in \mathcal{G}_n$ with $M \le N$ edges,
  $$\Pr(G_p = H) = p^M \underbrace{(1-p)}_{q}{}^{N-M}.$$

- This is called the <u>Erdös - Rényi</u> ensemble of random graphs. (Unfairly, because the ensemble was considered already before Erdös & Renyi's 1960 paper by E. Gilbert in 1959.)

- One of Erdös & Rényi's discoveries was that many structural properties of $\mathcal{G}(n,p)$ random graphs "emerge" at sharply defined <u>threshold</u> <u>densities</u> p.

- <u>Definition</u> Function $t = t(n)$ is a <u>threshold</u> for graph property Q if

  (i) $p(n) \ll t(n) \Rightarrow \Pr(G \in \mathcal{G}(n,p) \text{ has } Q) \longrightarrow 0$ as $n \to \infty$,

  (ii) $p(n) \gg t(n) \Rightarrow \Pr(G \in \mathcal{G}(n,p) \text{ has } Q) \longrightarrow 1$ as $n \to \infty$.

- Here "$f(n) \ll g(n)$" denotes $f(n) = o(g(n))$, i.e. $\frac{f(n)}{g(n)} \to 0$ as $n \to \infty$.

- Terminology: "random $G \in \mathcal{G}(n,p)$ has/doesn't have property Q <u>asymptotically</u> <u>almost</u> <u>surely</u> (a.a.s.)"