21 March 2021 22:44

Applications of Martingale tail inequalities

1. Pattern Matching

* Let $X = (x_1, x_2, \dots, x_n)$ be a sequence of characters chosen independently and uniformly at random from an alphabet Σ, where $|\Sigma| = s$.

* Let $B = (b_1, b_2, \dots, b_k)$ be a fixed string of k characters from Σ.

* Let F be no. of occurances of the fixed string B in the random string X.

$$X = (x_1, x_2, \dots x \dots, x_n)$$

k $\quad$ Σ

k

$= B \quad \left(\frac{1}{s}\right)^k \qquad \left(\frac{1}{s}\right)^k \qquad \left(\frac{1}{s}\right)^k$

$$\left(\frac{1}{s}\right) \qquad \to \left(\frac{1}{s}\right)^k$$

$$E[F] = \underline{(n-k+1)}\left(\frac{1}{s}\right)^k$$

Let $Z_0 = E[F]$

For $1 \le i \le n$,

$$Z_i = E[F \mid X_1, X_2, \dots, X_i]$$

$|Z_i - Z_{i-1}| \le k,\ \forall i \in [n]$

Azuma–Hoeffding.

The sequence $Z_0, Z_1, \dots, Z_n$ is a Doob martingale.

independent R.V.

$$F = f(\overset{x_1}{X_1}, \overset{x_2}{X_2}, \overset{x_3}{X_2}, \overset{2}{\overset{x_i}{X_i}}, \overset{x_n}{X_n})$$

satisfies the Lipschitz condition with bound k.

From McDiarmid Inequality, we know that

$$\Pr\left(\left|\overbrace{f(X_1, \dots, X_n)}^{F} - \overbrace{E[f(X_1, \dots, X_n)]}^{E[F]}\right| \ge \lambda\right) \le \frac{2}{\dots}$$

$$\leq \frac{2}{e^{\frac{2\lambda^2}{nk^2}}}$$

or

$$\Pr\left(|F - E[F]| \geq \varepsilon\sqrt{n}\right) \leq \frac{2}{e^{\frac{2\varepsilon^2}{k^2}}}$$

Balls and Bins

m balls $\qquad$ n bins

For Each ball, indep and uniformly at random, chosen a bin and put it inside that bin.

R.V. X_i : represents the bin into which i^{th} ball falls.

Let F be the no. of empty bins after all the m balls are thrown.

$$Z_0 = E[F]$$

$$Z_i = E[F \mid X_1, \ldots, X_i], \quad 1 \leq i \leq m$$

That means, $Z_m = F$.

That means, $Z_m = F$.

$$|Z_i - Z_{i-1}| \leq 1$$

$$\Pr\left(|\overset{F}{Z_m} - \overset{E[F]}{Z_0}| \geq \lambda\right) \leq \frac{2}{e^{\left(\frac{\lambda^2}{2m}\right)}}$$

$E[F] =$

$$\left(1-\frac{1}{n}\right)^m \quad \left(1-\frac{1}{n}\right)^m \quad \left(1-\frac{1}{n}\right)^m$$

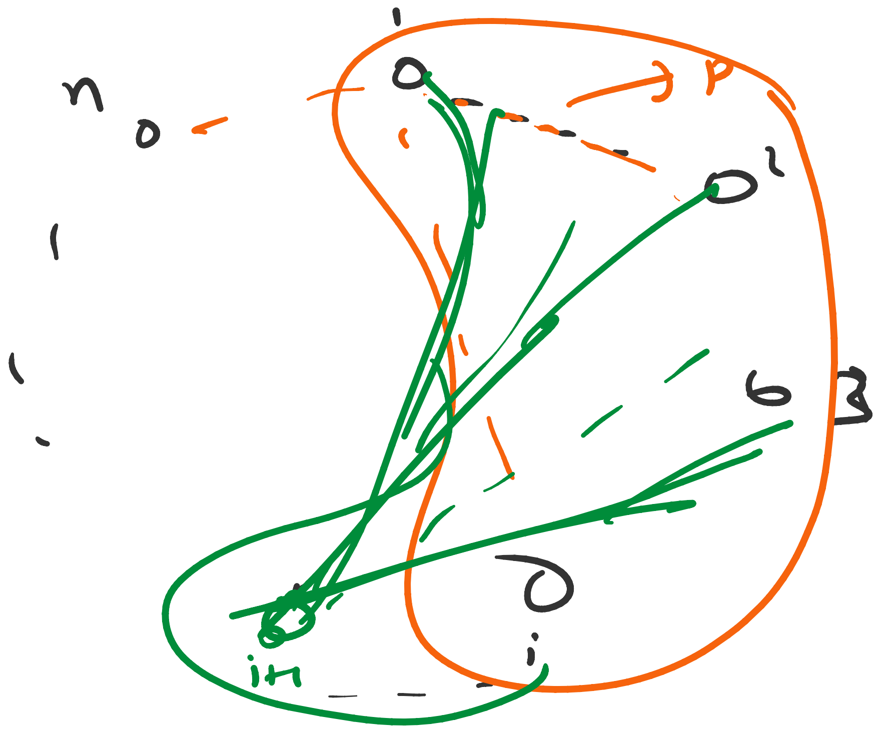
$$\underbrace{B_1 \quad B_2 \quad B_3 \quad B_4 \quad \cdots \quad B_n}_{n \text{ bins}}$$

$$E[F] = n\left(1-\frac{1}{n}\right)^m$$

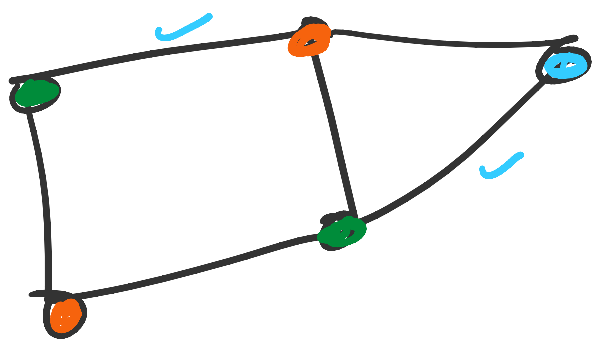
Chromatic number of random graphs

$G \sim \mathcal{G}(n,p)$, $0 \leq p \leq 1$

$G \sim \mathcal{G}(n,p)$, $0 \le p \le 1$



$\chi(G)$: min no of colors needed to color the vertices of G such that neighboring vertices get different colors.



Vertex exposure martingale.

Let G_i be the random subgraph of G

Let G_i be the random subgraph of G induced by the set of vertices $1, 2, \ldots, i$.

Let
$$Z_0 = E[X(G)]$$

$1 \le i \le n$,
$$Z_i = E[X(G) \mid G_1, G_2, \ldots, G_i]$$

$1 \le i \le n$,
$$|Z_i - Z_{i-1}| \le 1$$

Using Azuma-Hoeffding, we can say that

$$\Pr\left(|\overset{X(G)}{Z_n} - \overset{E[X(G)]}{Z_0}| \ge \lambda\sqrt{n}\right) \le \frac{2}{e^{\frac{\lambda^2}{2}}}$$