

به نام خدا

تمرین دوم
درس شبکه‌های اجتماعی

رضا منصوری خواه

810103246

استاد درس:

دکتر اسدیپور

بخش الف

از آنجا که گراف ما یک نوع گراف تصادفی است، برای محاسبه میانگین درجه تنها لازم است امید ریاضی درجه را برای یک گره محاسبه کنیم. این محاسبه شامل در نظر گرفتن دو نوع ارتباط ممکن است: ارتباطات درون انجمنی و ارتباطات برون انجمنی. در ابتدا به ارتباطات درون انجمنی توجه می‌کنیم و به احتمال p_{in} به گره‌های داخل انجمن متصل ایم که تعداد گره‌ها متصل برابر با $p * \left(\frac{N}{C} - 1\right)$ است. همچنین با احتمال p_{out} به گره‌های خارج از انجمن متصل ایم که تعداد این نوع گره‌های متصل برابر با $p_{out} * \left(N - \frac{N}{C}\right)$ است. پس با جمع تو مقدار میتوان به میانگین درجه برسیم.

$$\bar{k} = p * \left(\frac{N}{C} - 1\right) + p_{out} * \left(N - \frac{N}{C}\right)$$

در صورتی که N به سمت بینهایت حرکت کنیم. از آنجا که C مقدار ثابتی است پس مقدار $\left(N - \frac{N}{C}\right)$ برابر صفر خواهد بود و $\left(\frac{N}{C} - 1\right)$ برابر N خواهد بود. پس به صورت حدی برابر زیر خواهد بود.

$$\bar{k} = p * N$$

بخش ب

برای یک گره دلخواه α با فرض اینکه i تعداد گره‌های متصل به α در خارج از انجمن α باشد فرمول توزیع به صورت زیر خواهد بود.

$$P(k) = \sum_{i=0}^k \left(\left(\frac{N}{C} - 1\right)^{k-i} p^{k-i} (1-p)^{\left(\frac{N}{C}-1\right)-k+i} + \left(N - \frac{N}{C}\right)^i p_{out}^i (1-p_{out})^{\left(N-\frac{N}{C}\right)-i} \right)$$

در حالت حدی که تعداد کل نودها (N) به سمت بی‌نهایت میل می‌کند، توزیع دوجمله‌ای (Binomial) به توزیع پواسون (Poisson) تقریب می‌شود. این تقریب زمانی معتبر است که تعداد آزمایش‌ها بسیار زیاد و احتمال موفقیت در هر آزمایش بسیار کوچک باشد. به عبارت دیگر، اگر N بسیار بزرگ و p بسیار کوچک باشد به طوری که $Np = \lambda$ ثابت باقی بماند، توزیع دوجمله‌ای $\text{Bin}(N, p)$ به توزیع پواسون $\text{Poisson}(\lambda)$ میل می‌کند.

یکی از ویژگی‌های مهم توزیع پواسون این است که اگر دو متغیر تصادفی مستقل X_1, X_2 به ترتیب دارای توزیع‌های پواسون با پارامترهای λ_1 و λ_2 باشند، مجموع این دو متغیر تصادفی $X = X_1 + X_2$ نیز دارای توزیع پواسون با پارامتر $\lambda = \lambda_1 + \lambda_2$ خواهد بود.

کاربرد در گراف

در تحلیل گراف‌ها، این ویژگی برای مدل‌سازی تعداد اتصالات داخلی و خارجی نودها به کار می‌رود. به طور خاص، اگر تعداد اتصالات داخلی و تعداد اتصالات خارجی به ترتیب دارای توزیع‌های پواسون با پارامترهای λ_{in} و λ_{out} باشند، توزیع مجموع اتصالات نیز پواسون با پارامتر $\lambda = \lambda_{in} + \lambda_{out}$ خواهد بود.

بخش ج

دو همسایه گرهی α میتوانند چهار حالت زیر را داشته باشند که برای هر کدام احتمال وجود یال را میان دو همسایه حساب کرده و در انتها با هم جمع میکنیم.

- هر دو در انجمن α باشند . احتمال رخداد این حالت برابر $\left(\frac{kp}{p+p_{out}}\right)^2$ است . در این حال احتمال وجود یال میان این دو گره از آنجا که هر دو در یک انجمن هستند برابر p_{in} است پس احتمال رخداد این حالت برابر $\left(\frac{kp}{p+p_{out}}\right)^2 p$ است.
 - یکی در انجمن α و دیگری در خارج از آن باشند . احتمال رخداد این حالت برابر $\frac{\frac{kp}{p+p_{out}} * kp_{out}}{p+p_{out}}$ است . در این حال احتمال وجود یال میان این دو گره از آنجا که این دو گره در انجمن های متفاوت هستند برابر p_{out} است پس احتمال رخداد این حالت برابر $p_{out} * \frac{\frac{kp}{p+p_{out}} * kp_{out}}{p+p_{out}}$ است.
 - هر دو خارج از انجمن α ولی در یک انجمن یکسان باشند . احتمال رخداد این حالت برابر $\left(\frac{kp_{out}}{p+p_{out}}\right)^2$ است . در این حال احتمال وجود یال میان این دو گره از آنجا که هر دو در یک انجمن هستند برابر p_{in} است پس احتمال رخداد این حالت برابر $p \left(\frac{kp_{out}}{p+p_{out}}\right)^2$ است.
 - هر دو خارج از انجمن α و در دو انجمن متفاوت باشند . احتمال رخداد این حالت برابر $\left(\frac{kp_{out}}{p+p_{out}}\right)^2$ است . در این حال احتمال وجود یال میان این دو گره از آنجا که دو انجمن متفاوت هستند برابر p_{out} است پس احتمال رخداد این حالت برابر $p_{out} \left(\frac{kp_{out}}{p+p_{out}}\right)^2$ است.
- $$C = \left(\frac{kp}{p+p_{out}}\right)^2 p + \frac{\frac{kp}{p+p_{out}} * kp_{out}}{p+p_{out}} * p_{out} + \left(\frac{kp_{out}}{p+p_{out}}\right)^2 p + \left(\frac{kp_{out}}{p+p_{out}}\right)^2 p_{out}$$
- که از آنجا که به N وابسته نیست پس در N های بینهایت نیز مقدار ثابتی خواهد داشت.

بخش د

فاصله بین دو گره در یک انجمن برابر با $\overline{d} = \frac{\log \frac{n}{c}}{\log k}$ است (مقدار \overline{k} با بخش الف محاسبه میشود) حال فرض میکنیم دو گره به صورت تصادفی انتخاب میکنیم و فاصله این دو گره را محاسبه میکنیم. این دو گره به احتمال $\frac{1}{c} = \frac{N}{N}$ در یک انجمن قرار دارند در این حالت فاصله برابر با \overline{d} است همچنین به احتمال $1 - \frac{1}{c}$ در دو انجمن متفاوت هستند در این صورت با فرض اینکه در N های بزرگ با احتمال تقریباً صد در صد بین هر دو انجمن یک یال قرار دارد پس فاصله دو گره برابر با $\frac{\overline{d}}{2} + 1 + \frac{\overline{d}}{2}$ است که فرمول نهایی برابر با مقدار زیر خواهد بود.

$$\bar{d} = \frac{\bar{d} * 1}{c} + \left(1 - \frac{1}{c}\right) * (\bar{d} + 1) = 1 + \bar{d} - \frac{1}{c}$$

که مقدار بالا در N به سمت بینهایت برابر با $\frac{\log N}{\log k}$ خواهد بود . در فرمول گفته شده \bar{k} برابر با مقدار حدی محاسبه شده از بخش الف است.