باسمه تعالی

پروژه‌ی درس آمار و احتمال مهندسی

متاگراف

استاد درس

دکتر محمّدعلی مدّاح‌علی

آخرین مهلت تحویل:

۱۵ بهمن ۱۴۰۰

۱ بلاتشبیه مقدّمه!

احتمالا یادتان هست که در چهارم اکتبر سال ۲۰۲۱ شرکت متا (فیس‌بوک وقت!) و زیر مجموعه‌هایش یعنی واتس‌اپ، اینستاگرام و ... برای شش ساعت از دسترس خارج شد. در آن زمان شایعه شد که یک نابغه‌ی ۱۳ ساله‌ی چینی، این شرکت را هک کرده است! در این پروژه می‌خواهیم فرض کنیم که چنین اتفاقی افتاده است و این هکر ۱۳ ساله با هک‌کردن این شرکت، ایمیل تمام کاربرانش را بدست آورده است!

شکل ۱: شبکه‌ی دوستی افراد در فیس‌بوک

۲ درخت دوستی بنشان![1]

ابتدا یک تعریف برای رابطه‌ی دوستی بین دو شخص ارائه می‌کنیم:

تعریف ۱ (رابطه‌ی دوستی). بین شخص A و شخص B رابطه‌ی دوستی وجود دارد، اگر و تنها اگر شخص A شخص B را دنبال کرده باشد و برعکس.

هکر قصّه‌ی ما از بین ایمیل‌هایی که از هک کردن فیس‌بوک به دست آورده است، n ایمیل را به صورت تصادفی انتخاب می‌کند. فرض کنید بین این n نفر در دیتابیس شرکت m رابطه دوستی وجود داشته باشد (یعنی تعداد دوستی‌های واقعی برابر m است!). از آنجا که این هکر فقط ایمیل‌ها را به دست آورده است و اطّلاعی در مورد روابط دوستی بین افراد ندارد، تصمیم می‌گیرد روابط دوستی را بین تمام کاربران (تمام ایمیل‌ها) به صورت تصادفی و مستقل از هم در نظر بگیرد. او برای این کار، بین هر دو ایمیل یک سکّه می‌اندازد و اگر سکّه شیر بیاید بین آن‌ها رابطه‌ی دوستی در نظر می‌گیرد و اگر خط بیاید، بین آن‌ها رابطه‌ی دوستی در نظر نمی‌گیرد. فرض می‌کنیم احتمال شیر آمدن سکه p است. این شبکه‌ی تصادفی ایجاد شده را **متاگراف**[2] می‌نامیم.

[1] درخت دوستی بنشان که کام دل به بار آرد/ نهال دشمنی برکن که رنج بی‌شمار آرد [حافظ]

[2] MetaGraph

ابتدا می‌خواهیم درستیِ روابط دوستی که این هکر جوان به دست آورده است را بررسی کنیم.

پرسش تئوری ۱. احتمال اینکه این هکر موفّق شده باشد تمام روابط دوستی را به درستی تعیین کند بر حسب n و p و m بیابید.

پرسش تئوری ۲. تنها در این پرسش، فرض کنید این هکر مقدار دقیق m را می‌داند و در نتیجه m رابطه‌ی دوستی بین این n نفر به صورت تصادفی برقرار می‌کند. احتمال اینکه او تمام روابط دوستی را به درستی تعیین کند بر حسب n و m بیابید.

پرسش تئوری ۳. احتمال اینکه این هکر ۲۰ درصد از روابط دوستی بین این n نفر را به درستی تعیین کند بیابید.

هکر جوان ما از متاگراف خوشش آمده و می‌خواهد بیشتر با آن آشنا شود. به عنوان اوّلین مسئله، او می‌خواهد ببیند که در متاگرافش به طور میانگین چند رابطه‌ی دوستی وجود دارد.

پرسش شبیه‌سازی ۱. برای $n = ۱۰۰۰$، $p = ۰/۰۰۳۴$ و $m = ۳۰۰۰$ برنامه‌ای بنویسید که اختصاص روابط دوستی را به تعداد ۱۰ بار تکرار کند و هر بار تعداد روابط دوستی را ذخیره کرده و در پایان میانگین تمام مقادیر به دست آمده را محاسبه کند. آیا این مقدار میانگین، تقریباً (با حدّاکثر خطای ۵ درصد) با مقدار m برابر است؟

پرسش تئوری ۴. به ازای n و p بیان‌شده در پرسش شبیه‌سازی ۱، این مقدار میانگین را به صورت تئوری بدست آورید. در حالت کلی چه رابطه‌ای بین n، p و m باید برقرار باشد تا این مقدار میانگین تقریباً با مقدار m برابر شود؟

۳ خلوت‌گزیده را به تماشا چه حاجت است؟[۳]

هکر داستان ما، در زندگی به شدّت منزوی است و می‌خواهد بداند در متاگرافش چند نفر مانند او زندگی می‌کنند؟ برای این کار، ابتدا منزوی‌بودن را تعریف می‌کنیم:

تعریف ۲ (فرد گوشه‌گیر و فرد اجتماعی). در یک متاگراف، اگر هر فرد به طور میانگین L دوست داشته باشد آنگاه یک فرد را گوشه‌گیر می‌نامیم اگر کم‌تر از L دوست داشته باشد و او را اجتماعی می‌نامیم اگر بیش‌تر از L دوست داشته باشد.

پرسش شبیه‌سازی ۲. به ازای $n = ۱۰۰۰$ و $p = ۰/۰۰۰۱۶$ برنامه ای بنویسید که با ۱۰ بار تکرار، متوسّط تعداد افراد اجتماعی را بیابد.
علاوه بر این، برای اینکه دید بهتری از توزیع تعداد دوستان یک فرد داشته باشید، نموداری رسم کنید که محور افقی آن تعداد دوستان و محور عمودی آن متوسّط تعداد افرادی است که آن تعداد دوست دارند.

پرسش تئوری ۵. به ازای n و p بیان‌شده در پرسش شبیه‌سازی ۲، هر فرد به طور میانگین چند دوست دارد؟

پرسش تئوری ۶. برای n و p بیان‌شده در پرسش شبیه‌سازی ۲، اگر یک نفر را به صورت تصادفی انتخاب کنیم، احتمال اینکه اجتماعی باشد چقدر است؟ همچنین امید ریاضی تعداد افراد اجتماعی را بیابید.

[۳] خلوت گزیده را به تماشا چه حاجت است/ چون کوی دوست هست به صحرا چه حاجت است [حافظ]

۴ هواداران کویش را چو جان خویشتن دارم؟[۴]

گفتیم که هکر محترم و بزرگوار داستان ما بسیار منزوی است، در نتیجه از دارِ دنیا تنها یک دوست برای خودش دارد. نکته‌ی جالب توجّه آنست که دوستِ هکر عزیز ما بر خلاف خود او منزوی نیست و تعدادی دوست دیگر هم دارد که به تعبیر شعرا، رقیبان[۵] هکر قصّه‌ی ما محسوب می‌شَوند و همه‌ی مشکلات از همین نقطه شروع می‌شوند...

تعریف ۳. (خاصیت تراگذری) در رابطه‌ی دوستی بین سه شخص A، B و C خاصیت تراگذری برقرار است به شرطی که اگر A با B دوست باشد و B با C دوست باشد آنگاه میان A و C نیز رابطه‌ی دوستی برقرار باشد.

تعریف ۴. (خاصیت رقابت) در رابطه‌ی دوستی بین سه شخص A، B و C خاصیت رقابت برقرار است به شرطی که A با B دوست باشد و B با C دوست باشد، امّا میان A و C رابطه‌ی دوستی‌ای برقرار نباشد (و چه بسا مقداری هم رابطه‌ی این دو شکرآب باشد!)

هکر جوان ما می‌خواهد بررسی کند که در چند درصد روابط دوستی متاگرافش خاصیت تراگذری و در چند درصد آن‌ها خاصیت رقابت برقرار است؟

پرسش شبیه‌سازی ۳. برنامه ای بنویسید که بعد از ۵ بار اختصاص روابط دوستی به صورت تصادفی بین $n = ۳۰۰۰$ نفر با احتمال $p = ۰/۰۱$ میانگین تعداد روابط دوستی دارای خاصیت تراگذری و میانگین روابط دوستی دارای خاصیت رقابت را حساب کند.

پرسش شبیه‌سازی ۴. پرسش شبیه‌سازی ۳ را با $p = ۰/۴$ تکرار کنید.

پرسش تئوری ۷. امید ریاضی تعداد روابط دوستی دارای خاصیت تراگذری و امید ریاضی تعداد روابط دوستی دارای خاصیت رقابت را محاسبه کنید.

پرسش تئوری ۸. در یک متاگراف، چه کسری از کلّ روابط دوستی بین سه نفر، با فرض آن‌که هر کدام از این سه نفر حداقل با یکی از دو نفر دیگر دوست باشد، دارای خاصیت تراگذری هستند؟ آیا شبیه‌سازی‌های شما با عددی که از محاسبه‌ی تئوری به دست می‌آورید هم‌خوانی دارند؟ نتیجه را تحلیل کنید.

تا این‌جا، وجود و عدم وجود رابطه‌ی دوستی میان دو دوست از یک شخص پرداختیم. حال می‌خواهیم ببینیم که اگر یک نفر را به صورت تصادفی انتخاب کنیم، به طور میانگین چند رابطه‌ی دوستی بین همه‌ی دوست‌های او وجود دارد؟ به تعبیر دیگر، اگر یک «جانان» داشته باشیم، چند رابطه‌ی دوستی میان «هواداران کوی»ش برقرار است؟

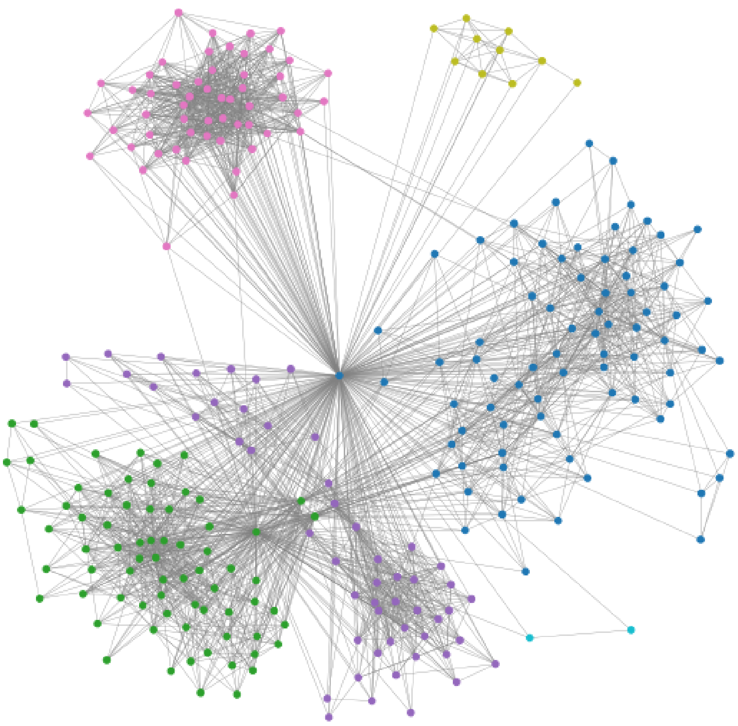
پرسش شبیه‌سازی ۵. برای متاگرافی با $n = ۱۰۰۰$ و $p = ۰/۰۰۳$ میانگین تعداد روابط دوستی میان دوستان یک شخص را با کمک شبیه‌سازی محاسبه کنید.

پرسش تئوری ۹. به نظر شما اگر متوسّط تعداد روابط دوستی میان دوستان یک شخص را با استفاده از روابط دوستی واقعی که در دیتابیس شرکت متا وجود دارد محاسبه کنیم، مقدار به دست آمده از میانگین محاسبه‌شده در پرسش شبیه‌سازی ۵ بیشتر می‌شود یا کمتر؟ چرا؟ (راهنمایی: به شکل‌های موجود در پروژه توجّه کنید!)

پرسش تئوری ۱۰. امید ریاضی تعداد روابط دوستی میان دوستان یک شخص در متاگرافی با n رأس و احتمال p را بیابید. (جواب بسته لازم است!)

---

[۴] مرا عهدیست با جانان که تا جان در بدن دارم/ هواداران کویش را چو جان خویشتن دارم [حافظ]
[۵] ز رقیب دیو سیرت به خدای خود پناهم/ مگر آن شهاب ثاقب مددی دهد خدا را [حافظ]



شکل ۲: مثالی از روابط دوستی یک فرد در یک شبکه‌ی اجتماعی

۵ من از دیار حبیبم نه از بلاد غریب![۶]

ابتدا برای اینکه به درک بهتری برسیم، مدل ریاضی مسئله را قدری دقیق‌تر می‌کنیم.
می‌دانیم یک گراف، یک زوج مرتّب مانند $G = (V, E)$ است که در آن V مجموعه‌ای ناتهی از رئوس و E مجموعه‌ی یال‌های آن است. حال گرافی را در نظر بگیرید که رأس های آن کاربران فیس‌بوک و یال‌های آن نشان‌دهنده‌ی وجود رابطه‌ی دوستی بین کاربران مختلف هستند. به تعبیر دیگر، دو کاربر i و j با یک‌دیگر دوست هستند اگر و تنها اگر میان رئوس متناظر آن‌ها در گراف یک یال وجود داشته باشد. با مدلی که ذکر کردیم، کاری که هکر جوان در بخش‌های قبلی انجام می‌داد، تنها ایجاد تعدادی یال تصادفی روی مجموعه‌ی این رأس‌ها بود. چنین گراف‌هایی را گراف‌های تصادفی می‌نامیم که خانواده‌های گوناگون و خواصّ جالبی دارند. در ابتدای پروژه نام متاگراف را برای گراف تصادفی مورد استفاده در این پروژه انتخاب کردیم. برای سادگی، متاگرافی که n رأس دارد و هر دو رأس آن با احتمال p به هم متّصل هستند را با $\mathcal{G}(n,p)$ نشان می‌دهیم.
یکی از ویژگی‌های بسیار جالب متاگراف‌ها این است که قانون «جهان کوچک» در آن‌ها برقرار است. طبق قانون «جهان کوچک»، هر دو نفر در دنیا با احتمال نزدیک به یک با حدّاکثر ۶ واسطه هم‌دیگر را می‌شناسند. یعنی به طور مثال اگر دو فرد A و B را به صورت تصادفی انتخاب کنیم و مجموعه‌ی دوستانِ A، دوستانِ دوستانِ A، دوستانِ دوستانِ دوستانِ A و ... را بررسی کنیم، حدّاکثر بعد از ۶ مرحله به فرد B می‌رسیم.
در این بخش، می‌خواهیم وجود این ویژگی را در متاگراف (که روابط دوستی در آن، برخلاف جهان واقعی، به صورت تصادفی چیده شده‌اند) تحقیق کنیم.

[۶] نماز شام غریبان چو گریه آغازم/ به مویه‌های غریبانه قصه پردازم
به یاد یار و دیار آن چنان بگریم زار/ که از جهان ره و رسم سفر براندازم
من از دیار حبیبم نه از بلاد غریب/ مهیمنا به رفیقان خود رسان بازم [حافظ]

ابتدا به بررسی میانگین فاصله‌ی دو شخص در یک متاگراف می‌پردازیم. لازم به ذکر است که فاصله‌ی دو شخص، حدّاقلّ تعداد روابط دوستی‌ای است که آن‌ها را به یک‌دیگر متّصل می‌کند.

پرسش شبیه‌سازی ۶. برنامه‌ای بنویسید که میانگین فاصله‌ی دو شخص در $\mathcal{G}(n,p)$ را به ازای $n = ۱۰۰۰$ و $p = ۰/۰۰۳۳$ محاسبه کند.

حال به بررسی حدّاکثر فاصله‌ی دو شخص در یک متاگراف می‌پردازیم.

پرسش شبیه‌سازی ۷. متاگرافی با $n = ۵۰$ و $p = ۰/۳۴$ را ۱۰۰ بار به طور تصادفی تولید کنید. هر بار جفت کاربری را پیدا کنید که بیشترین فاصله را از هم در گراف دارند. میانگین بیشترین فاصله بین دو کاربر روی این ۱۰۰ متاگراف را به دست آورید.

پرسش شبیه‌سازی ۸. با ثابت (و برابر با مقدار بیان‌شده در پرسش شبیه‌سازی ۷) نگه داشتن p تعداد رأس‌ها (n) را در بازه‌ی $[۱۰, ۲۰۰]$ با گام ۱۰ واحد تغییر دهید و برای هر n پرسش شبیه‌سازی ۷ را تکرار کنید. در نهایت نمودار میانگین حدّاکثر فاصله بین جفت افراد (که میانگین روی ۱۰۰ نمونه‌ی مختلف متاگراف با مشخّصات مشابه گرفته می‌شود) را به صورت تابعی از n رسم کنید. این نمودار چه فرمی دارد؟ با افزایش n رفتار این نمودار به چه صورتی است؟

متغیر توصیف شده در بالا (حدّاکثر فاصله‌ی بین دو رأس در متاگراف) را قطر متاگراف می‌نامیم.

اکنون می‌خواهیم پدیده‌ای که در قسمت بالا به کمک شبیه‌سازی مشاهده کردید را اثبات کنیم. ابتدا توجّه داشته باشید که اغلب خاصیت‌های جالب گراف‌های تصادفی (که متاگراف هم یکی از آن‌هاست) با میل‌کردن n به سمت بی‌نهایت ظاهر می‌شوند. از طرف دیگر، پدیده‌هایی از جهان واقعی که با گراف‌های تصادفی مدل می‌شوند، معمولاً دارای تعداد زیادی رأس هستند و بررسی حالتی که n به سمت بی‌نهایت میل می‌کند بسیار مهم و کلیدی است.

پرسش تئوری ۱۱. برای دو رأس u و v متغیّر تصادفی برنولی $I_{u,v}$ را به این صورت تعریف می‌کنیم:

$$I_{u,v} = \begin{cases} ۰ & \text{دو رأس } u,v \text{ همسایه‌ی مشترک داشته باشند} \\ ۱ & \text{دو رأس } u,v \text{ همسایه‌ی مشترک نداشته باشند} \end{cases}.$$

$\mathbb{P}[I_{u,v} = ۱]$ را محاسبه کنید.

پرسش تئوری ۱۲. برای متاگراف $\mathcal{G}(n,p)$ متغیّر تصادفی X_n را به صورت «تعداد جفت راس هایی از متاگراف که همسایه مشترکی ندارند» تعریف می‌کنیم. $\mathbb{E}[X_n]$ را بیابید.

پرسش تئوری ۱۳. با استفاده از نامساوی مارکف کران بالایی برای $\mathbb{P}[X_n \geq ۱]$ بیابید. سپس با میل دادن n به سمت بی‌نهایت رفتار این کران را بررسی کنید.

پرسش تئوری ۱۴. نتیجه بگیرید که وقتی n عدد خیلی بزرگی باشد، قطر متاگراف با احتمال بالا کران بالا دارد و همچنین مقدار این کران بالا را نیز مشخص کنید. آیا قطر متاگراف برای nهای بزرگ به p وابسته است؟ آیا نتیجه‌ای که از اثبات تئوری گرفتید با نتیجه‌ی شبیه‌سازی تطابق دارد؟

۶ وَ إن یَکاد بخوانید![۷]

هکر جوان ما تا به حال عضو هیچ گروه دوستی‌ای نبوده است و در نتیجه به بررسی گروه‌های دوستی در متاگراف علاقه‌مند می‌شود. او ابتدا با گروه‌های دوستی سه نفره آغاز می‌کند...

پرسش شبیه‌سازی ۹. متاگرافی با $n = ۱۰۰$ و $p = ۰/۳۴$ را ۱۰۰ بار به طور تصادفی تولید کنید و هر بار تعداد حلقه‌های دوستی ۳ نفره را در آن بشمارید. میانگین تعداد حلقه‌های دوستی در این ۱۰۰ گراف را به دست آورید.

پرسش شبیه‌سازی ۱۰. تعداد رأس ها (n) را در بازه‌ی $[۱۰, ۱۰۰]$ و با گام ۱۰ تغییر دهید و برای هر n، p را به صورت تابعی از n و برابر

$$p(n) = \frac{۶۰}{n^۲}$$

قرار دهید. برای هر n میانگین تعداد حلقه‌های دوستی ۳ نفره را به روش پرسش شبیه‌سازی ۹ حساب کنید و این میانگین را در یک نمودار بر حسب n رسم کنید.
آیا با افزایش n میانگین به عدد خاصی میل می‌کند؟ این رفتار را چگونه توجیه می‌کنید؟

پرسش شبیه‌سازی ۱۱. پرسش شبیه‌سازی ۱۰ را با $p = ۰/۳۴$ تکرار کنید. آیا میانگین تعداد حلقه‌های دوستی با افزایش n به عدد خاصی میل می‌کند؟

پرسش شبیه‌سازی ۱۲. این بار از

$$p = \frac{۱}{n}$$

استفاده کنید. n را در بازه‌ی $[۵۰, ۱۲۰۰]$ با گام ۵۰ تغییر دهید. مجدّداً نمودار میانگین تعداد حلقه‌های دوستی را بر حسب n رسم کنید. همچنین میانگین تجمّعی[۸] این نمودار را رسم کنید. آیا به عدد خاصی میل می‌کند؟

حال می‌خواهیم نتیجه‌ی به دست آمده در بخش شبیه‌سازی را اثبات کنیم. برای این کار، متغیّر تصادفی $T_{۳,n}$ را برابر با تعداد حلقه‌های ۳ نفره‌ی دوستی موجود در $\mathcal{G}(n,p)$ تعریف می‌کنیم.

پرسش تئوری ۱۵. متغیّر تصادفی برنولی $I_{u,v,w}$ را به این صورت تعریف می‌کنیم:

$$I_{u,v,w} = \begin{cases} ۱ & \text{سه رأس } u,v,w \text{ تشکیل یک مثلّث بدهند} \\ ۰ & \text{سه رأس } u,v,w \text{ تشکیل یک مثلّث ندهند} \end{cases}.$$

$\mathbb{P}[I_{u,v,w} = ۱]$ را محاسبه کنید.

پرسش تئوری ۱۶. $\mathbb{E}[T_{۳,n}]$ را بر حسب n و p محاسبه کنید.

پرسش تئوری ۱۷. با استفاده از نامساوی مارکف کران بالایی برای $\mathbb{P}[T_{۳,n} \geq ۱]$ بیابید. سپس قرار دهید

$$p(n) = \frac{۱}{n^۲}$$

و با میل دادن n به سمت بی‌نهایت رفتار این کران را بررسی کنید.

---

[۷] حضور خلوت انس است و دوستان جمعند/ و ان یکاد بخوانید و در فراز کنید [حافظ]

[8] cumulative mean

رفتار حدّی تعداد حلقه‌های دوستی ۳ نفره را برای حالتی که p رابطه‌ی معکوس با مجذور n دارد را بررسی کردیم. اکنون می‌خواهیم این رفتار را برای $p = c$ (c ثابت است) و $p(n) = \frac{1}{n}$ هم بررسی کنیم و در نهایت مقایسه‌ای بین این سه حالت انجام دهیم. مطالعه‌ی رفتار حدّی تعداد مثلّث‌ها برای این دو حالت اندکی مفصّل‌تر از $p(n) = \frac{1}{n^2}$ است و گام به گام با هم پیش می‌رویم...

قضیه ۰.۱ اگر برای دنباله‌ی متغیّرهای تصادفی $\{Y_n\}_{n=1}^{\infty}$ که مقادیر آن‌ها همواره صحیح و نامنفی است، دو شرط زیر برقرار باشد:

$$\forall n \in \mathbb{N} : \mathbb{E}[Y_n] > 0$$
$$\lim_{n\to\infty} \frac{\mathrm{Var}[Y_n]}{\mathbb{E}^2[Y_n]} = 0$$

آن‌گاه:

$$\lim_{n\to\infty} \mathbb{P}[Y_n \geq 1] = 1.$$

پرسش تئوری ۰.۱۸ قضیه‌ی ۱ را اثبات کنید.

پرسش تئوری ۰.۱۹ با توجّه به تعریف $I_{u,v,w}$ در پرسش تئوری ۱۵، $\mathbb{E}[I_{u,v,w}I_{u',v',w'}]$ را محاسبه کنید.
(راهنمایی: سه حالت مختلف ممکن است رخ دهد. برای هر حالت جداگانه محاسبه کنید.)

پرسش تئوری ۰.۲۰ $\mathbb{E}[T_{3,n}^2]$ را بر حسب n و p محاسبه کنید. می‌توانید برای ساده‌شدن محاسبات از تقریب‌های زیر استفاده کنید:

$$\binom{n}{3} \approx \binom{n-3}{3} \approx \frac{n^3}{6}$$
$$\binom{n-3}{2} \approx \frac{n^2}{2}$$
$$n-3 \approx n$$

پرسش تئوری ۰.۲۱ نشان دهید اگر $p = c$ باشد که c یک ثابت است، آن‌گاه:

$$\lim_{n\to\infty} \frac{\mathrm{Var}[T_{3,n}]}{\mathbb{E}^2[T3,n]} = 0.$$

پرسش تئوری ۰.۲۲ نتیجه بگیرید $\lim_{n\to\infty} \mathbb{P}[T_{3,n} \geq 1] = 1$. این نتیجه را تفسیر کرده و با شبیه‌سازی مقایسه کنید.

حال حالتی را بررسی می‌کنیم که $p = \frac{c}{n}$ باشد.

تعریف ۵ (گشتاور فاکتوریل). گشتاور فاکتوریل مرتبه‌ی r متغیّر تصادفی X را به صورت زیر تعریف می‌کنیم:

$$\mathbb{E}\big[(X)_r\big] := \mathbb{E}\big[X(X-1)\dots(X-r+1)\big].$$

تعریف ۶ (تابع مولّد گشتاور فاکتوریل). تابع مولّد گشتاور فاکتوریل متغیّر تصادفی X را به صورت زیر تعریف می‌کنیم:

$$M_X(t) = \mathbb{E}\big[t^X\big].$$

به سادگی می‌توان دید که:

$$\mathbb{E}\big[(X)_r\big] = \frac{\partial^r}{\partial t^r} M_X(t)\Big|_{t=1}.$$

قضیه ۲. برای متغیّر تصادفی صحیح و نامنفی X داریم:

$$\mathbb{E}\big[(X)_r\big] = \lambda^r,$$

اگر و تنها اگر تابع جرم احتمال آن به صورت زیر باشد:

$$\mathbb{P}[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}$$

به تعبیر دیگر، شرط لازم و کافی برای آن‌که متغیّر تصادفی X، یک متغیّر تصادفی پواسون با پارامتر λ باشد، آنست که داشته باشیم $\mathbb{E}\big[(X)_r\big] = \lambda^r$.

پرسش تئوری ۲۳. قضیه‌ی ۲ را اثبات کنید.

پرسش تئوری ۲۴. با استفاده از نتایج پرسش تئوری ۲۰ و پرسش تئوری ۱۶ نشان دهید که اگر قرار دهیم:

$$p = \frac{c}{n}$$

و تعریف کنیم:

$$\lambda = \frac{c^{۳}}{۶},$$

آن‌گاه برای nهای بزرگ داریم:

$$\mathbb{E}\big[(T_{۳,n})_{۱}\big] \sim \lambda,$$
$$E\big[(T_{۳,n})_{۲}\big] \sim \lambda^{۲}.$$

پرسش تئوری ۲۵. (امتیازی) برای هر $r \geq ۲$ نشان دهید که با میل کردن n به سمت بی‌نهایت داریم:

$$E\big[(T_{۳,n})_r\big] \sim \lambda^r.$$

می‌توانید از این فرض ساده کننده برای nهای بزرگ استفاده کنید: احتمال وجود r مثلّث در متاگراف به صورتی که حدّاقل ۲ مثلّث از میان آن‌ها یک رأس مشترک داشته باشند، بسیار کمتر از احتمال وجود r مثلّث است که هیچ‌کدام هیچ رأس مشترکی ندارند.
همچنین می توانید از تقریب $\left(\binom{n}{۳}p^{۳}\right)^r \sim \lambda^r$ استفاده کنید.

با توجّه به دو پرسش آخر و قضیه‌ی ۲ می‌توان نتیجه گرفت که اگر $p = \frac{c}{n}$ باشد، آن‌گاه تعداد مثلّث‌ها در متاگراف $\mathcal{G}(n,p)$ یک متغیّر تصادفی پواسون با پارامتر $\lambda = \frac{c^{۳}}{۶}$ است.

۷ قومی به جد و جهد نهادند وصل دوست، قومی دگر حواله به تقدیر می‌کنند![۹]

هکر داستان ما می‌خواهد بررسی کند که آیا در متاگرافش همه‌ی افراد از طریق روابط دوستی با هم مرتبط می‌شوند یا نه؟

پرسش شبیه‌سازی ۱۳. متاگرافی با $n = ۱۵۰$ و $p = ۰/۲$ را ۱۰۰ بار تولید کنید و با توجّه به نتایج شبیه‌سازی، احتمال اینکه فردی بدون دوست وجود داشته باشد را بیابید. همچنین احتمال هم‌بند بودن گراف را بیابید.

پرسش شبیه‌سازی ۱۴. تعداد رأس‌ها را در بازه‌ی $[۱۰, ۱۵۰]$ و با گام ۱۰ تغییر دهید و برای هر n، قرار دهید:

$$p = \frac{۴\ln(n)}{n}.$$

حال برای هر n احتمال عدم وجود فردی بدون دوست و احتمال هم‌بند بودن متاگراف را به روش پرسش شبیه‌سازی ۱۳ حساب کنید. و نمودار این دو احتمال را برحسب n در یک نمودار رسم کنید.

پرسش شبیه‌سازی ۱۵. پرسش شبیه‌سازی ۱۴ را با

$$p = \frac{۴}{n}$$

تکرار کنید.

حال می‌خواهیم مشاهدات بخش شبیه‌سازی را به صورت تئوری هم بررسی کنیم.

قضیه ۳. اگر $p = \frac{a\ln(n)}{n}$ باشد و بدانیم $a \geq ۳$، آن‌گاه داریم:

$$\lim_{n\to\infty} \mathbb{P}[\mathcal{G}(n,p) \text{ هم‌بند بودن}] = ۱.$$

در چند مساله بعدی می‌خواهیم قدم به قدم قضیه‌ی ۳ را ثابت کنیم.
برای این کار، متغیر تصادفی X را این‌گونه تعریف می‌کنیم که اگر گراف همبند باشد، $X = ۰$ و اگر گراف شامل k مؤلّفه‌ی هم‌بندی باشد، $X = k$. توجّه داریم که طبق تعریف برای هیچ گرافی $X = ۱$ نمی‌شود.
X را می‌توانیم به صورت زیر بنویسیم:

$$X = X_۱ + X_۲ + \ldots + X_{n-۱}.$$

که X_j تعداد مؤلّفه‌های هم‌بندی‌ای از گراف است که دقیقاً j رأس دارند.
اگر گراف n رأس داشته باشد، می‌توان به $\binom{n}{j}$ حالت j رأس از آن انتخاب کرد. هر کدام از این حالت‌ها را با اندیس $i \in [۱, \binom{n}{j}]$ نشان می‌دهیم. متغیّر تصادفی برنولی $K_i^{(j)}$ را به این صورت تعریف می‌کنیم:

$$K_i^{(j)} = \begin{cases} ۱ & \text{اگر } i\text{اُمین حالت انتخاب } j \text{ رأس از } n \text{ رأس، تشکیل یک مؤلّفه‌ی هم‌بندی بدهد.} \\ ۰ & \text{در غیر این صورت} \end{cases}.$$

حال می‌توانیم اثبات قضیه‌ی ۳ را آغاز کنیم.

[۹] حافظ

پرسش تئوری ۲۶. ثابت کنید:

$$\mathbb{P}[K_i^{(j)} = ۱] \leq (۱ - p)^{j(n-j)}.$$

پرسش تئوری ۲۷. با توجه به تعریف X_j، کران بالایی برای $\mathbb{E}[X_j]$ بر حسب p و n و j به دست آورید.

پرسش تئوری ۲۸. با توجه به تعریف X کران بالایی برای $\mathbb{E}[X]$ بیابید و رفتار این کران را وقتی n به سمت بی‌نهایت میل می‌کند بررسی کنید.

می‌توانید از این فرض ساده‌کننده استفاده کنید که با p فرض شده در قضیه‌ی ۳ و nهای به اندازه‌ی کافی بزرگ، به ازای هر $j \geq ۱$ داریم:

$$\binom{n}{j}(۱ - p)^{j(n-j)} \leq \frac{۱}{n^{۱/۷}}.$$

پرسش تئوری ۲۹. اثبات قضیه‌ی ۳ را کامل کنید.

۸ نکات مهم!

لطفاً به نکات زیر دقّت کنید:

۱. این پروژه بخشی از نمره‌ی شما در این درس را تشکیل خواهد داد.

۲. پروژه به صورت گروهی است. به این معنا که شما برای انجام پروژه به گروه‌های دونفری تقسیم می‌شوید (گروه‌بندی با خود شما و به صورت اختیاری است) و پروژه را انجام خواهید داد. در روز تحویل پروژه، ممکن است از هرجای گزارش یا کد بارگذاری شده از هریک از دونفر سؤال پرسیده شود و میزان آمادگی او در پاسخ به پرسش‌ها در نمره‌ی نهایی بسیار مؤثّر است. در نهایت نمره‌ی هردو عضو یک گروه یکسان خواهد بود.

۳. تمامی شبیه‌سازی‌ها باید با کمک زبان Python انجام شود. شما تنها مجاز به استفاده از کتاب‌خانه‌های `networkx`، `numpy` و `matplotlib` هستید. اگر روی عنوان هر کتاب‌خانه کلیک کنید، به راهنمای آن کتاب‌خانه هدایت می‌شوید.

۴. تحویل پروژه به صورت گزارش و کدهای نوشته‌شده است. گزارش باید شامل پاسخ پرسش‌ها، تصاویر و نمودارها و نتیجه‌گیری‌های لازم باشد. توجّه کنید که قسمت عمده‌ی نمره‌ی بخش شبیه‌سازی را گزارش شما و نتیجه‌ای که از خروجی کد می‌گیرید تشکیل می‌دهد. همچنین تمیزی گزارش بسیار مهم است. کدها و گزارش را در یک فایل فشرده‌شده در سامانه‌ی درس‌افزار آپلود کنید.

۵. اگر برای پاسخ به پرسش‌ها، از منبعی (کتاب، مقاله، سایت و...) کمک گرفته‌اید، حتماً به آن ارجاع دهید.

۶. نوشتن گزارش کار با LaTeX نمره‌ی امتیازی دارد.

۷. در متن پروژه، پرسش‌های شبیه‌سازی با رنگ سبز و پرسش‌های تئوری با رنگ آبی مشخص شده‌اند.

۸. بخش‌های تئوری گزارش که در قالب پرسش‌های تئوری طرح شده‌اند را می‌توانید روی کاغذ بنویسید و تصویر آن‌ها را در گزارش خود بیاورید، ولی توصیه‌ی برادرانه می‌کنم که این کار را نکنید! همچنین در صورت انجام این کار، نمره‌ی امتیازی حروف‌چینی گزارش با LaTeX را از دست خواهید داد.

۹. اگر در فهم صورت پروژه ابهام یا پرسشی دارید، در تالار گفتگوی مخصوص پروژه که در سامانه‌ی درس‌افزار ایجاد شده است، آن را مطرح کنید. طراحان پروژه در اسرع وقت پاسخ خواهند داد.

۱۰. درصورت مشاهده‌ی تقلّب، نمره‌ی هردو گروه صفر منظور خواهد شد.

موفّق باشید