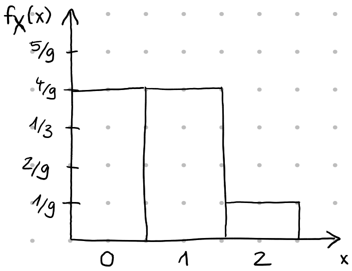
RECAP: Woche 6

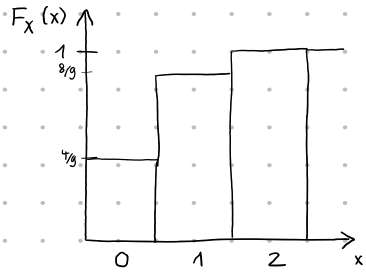
Eine Zufallsvariable (ZV) ist eine Abbildung $X: \Omega \to \mathbb{R}$, wobei ihr Wertebereich
$W_X := X(\Omega) = \{x \in \mathbb{R} \mid \exists \omega \in \Omega \text{ mit } X(\omega) = x\}$

Dichtefunktion $f_X: \mathbb{R} \to [0,1], \quad x \mapsto Pr[X=x]$

Verteilungsfunktion: $F_X: \mathbb{R} \to [0,1], \quad x \mapsto Pr[X \leq x]$

Aufgabe: Ein normaler Würfel wird zweimal geworfen.
Die ZV X bezeichnet die Gesamtzahl der Würfe mit Ergebnis "1 oder 6".
Zeichne Dichte- und Verteilungsfunktion.





$Pr[X=0] = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$

$Pr[X=1] = \frac{4}{9}$

$Pr[X=2] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$

Indikatorvariable $I_A(\omega) := \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{sonst} \end{cases}$ für ein Ereignis A

$E[X_A] = Pr[A]$

$$E[X_A] = \sum_{x \in W_X} Pr[X=x] \cdot x = Pr[\bar{A}] \cdot 0 + Pr[A] \cdot 1$$
$$= Pr[A]$$

Varianz

Definition 2.39. Für eine Zufallsvariable X mit $\mu = \mathbb{E}[X]$ definieren wir die *Varianz* $\mathrm{Var}[X]$ durch

$$\mathrm{Var}[X] := \mathbb{E}[(X - \overset{E[X]}{\mu})^2] = \sum_{x \in W_X} (x - \mu)^2 \cdot \Pr[X = x].$$

Die Grösse $\sigma := \sqrt{\mathrm{Var}[X]}$ heisst *Standardabweichung* von X.

Satz 2.40. Für eine beliebige Zufallsvariable X gilt

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Beweis: $\mathrm{Var}[X] = E[(X-\mu)^2] = E[X^2 - 2X\mu + \mu^2]$

$$= E[X^2] - E[2\mu X] + \mu^2$$

$$= E[X^2] - 2\mu E[X] + \mu^2$$

$$= E[X^2] \underbrace{- 2E[X]^2 + E[X]^2}_{-E[X]^2}$$

$$= E[X^2] - E[X]^2$$

Satz 2.41. Für eine beliebige Zufallsvariable X und $a, b \in \mathbb{R}$ gilt

$$\mathrm{Var}[a \cdot X + b] = a^2 \cdot \mathrm{Var}[X].$$

Beweis:

$$E[aX] \cdot E[aX]$$

$$a^2 E[X]^2$$

Definition 2.42. Für eine Zufallsvariable X nennen wir $\mathbb{E}[X^k]$ das k-*te Moment* und $\mathbb{E}[(X - \mathbb{E}[X])^k]$ das k-*te zentrale Moment*.

Verteilungen

1. Bernoulli-Verteilung

Für eine ZV X mit $W_X = \{0,1\}$ und $f_X(x) = \begin{cases} p & , x=1 \\ 1-p & , x=0 \\ 0 & , \text{sonst} \end{cases}$ gilt $X \sim \text{Bernoulli}(p)$

$E[X] = p$
$Var[X] = p(1-p)$, da $Var[X] = p(1-p)$

$\binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^4$

2. Binomialverteilung

Eine ZV X mit $W_X = \{0,1,\dots,n\}$ und Dichte $f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x \in \{0,1,\dots,n\} \\ 0 & \text{sonst} \end{cases}$
heisst binomialverteilt mit Parameter n und p, $X \sim Bin(n,p)$

$E[X] = np$
$Var[X] = np(1-p)$

3. Geometrische Verteilung

Wenn ein einzelner Versuch mit W'keit p gelingt,
so ist die Anzahl der Versuche bis zum Erfolg geometrisch verteilt, $X \sim Geo(p)$

$f_X(i) = \begin{cases} p(1-p)^{i-1} & \text{für } i \in \mathbb{N} \\ 0 & \text{sonst} \end{cases}$ $\quad F_X(n) = Pr[X \leq n] = \sum_{i=1}^{n} p(1-p)^{i-1} = 1-(1-p)^n$

$E[X] = \frac{1}{p}$
$Var[X] = \frac{1-p}{p^2}$

„gedächtnislos":

Satz 2.45. Ist $X \sim Geo(p)$, so gilt für alle $s, t \in \mathbb{N}$:

$$Pr[X \geq s+t \mid X > s] = Pr[X \geq t].$$

Beweis: $Pr[X \geq s+t \mid X>s] = \frac{Pr[X \geq s+t]}{Pr[X>s]} = \frac{(1-p)^{s+t-1}}{(1-p)^s} = (1-p)^{t-1} = Pr[X \geq t]$

4. Negative Binomialverteilung

"verallgemeinerte geom. Verteilung", Warten auf den n-ten Erfolg
falls $n=1$: $Z \sim Geo(p)$
#Versuche bis zum n-ten erfolgreichen Experiment

$f_Z(z) = \binom{z-1}{n-1} \cdot p^n \cdot (1-p)^{z-n}$

übrigen n-1 Erfolge können irgendwo auf z-1 Versuche aufgeteilt werden

$Z = \sum_{i=1}^{n} X_i$, wobei jedes $X_i \sim Geo(p)$
$E[Z] = \frac{n}{p}$

5. Poisson-Verteilung

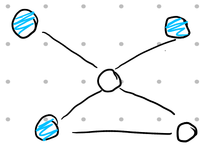
Motivation: jedes Ereignis sehr unwahrscheinlich, aber viele Ereignisse möglich

$f_X(i) = \begin{cases} \frac{e^{-\lambda}\lambda^i}{i!} & \text{für } i \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$

$E[X] = Var[X] = \lambda$

Stabile Mengen

gegeben: $G=(V,E)$ mit $|V|=n$ und $|E|=m$
gesucht: möglichst grosse Teilmenge $S\subseteq V$, so dass $G[S]$ keine Kanten erhält



Algorithmus besteht aus 2 Runden:
1) Löschen jedes Knotens mit W'keit $1-p$
2) Für jede übrig gebliebene Kante einen der beiden Knoten löschen

$X :=$ Anzahl der Knoten, die die erste Runde überleben $\quad X=\sum_{i=1}^{n} X_i$
X_i: Indikatorvariable, dass der i-te Knoten überlebt $\quad X_i \sim \text{Bernoulli}(p)$

$$E[X]=E[\sum_{i=1}^{n} X_i]=\sum_{i=1}^{n} E[X_i]=\sum_{i=1}^{n} Pr[X_i=1]=np$$

$Y :=$ Anzahl der Kanten, die die erste Runde überleben $\quad Y=\sum_{i=1}^{m} Y_i$
Y_i: Indikatorvariable, dass Kante i überlebt $\quad Y_i \sim \text{Bernoulli}(p^2)$

$$E[Y]=E[\sum_{i=1}^{m} Y_i]=\sum_{i=1}^{m} \underbrace{E[Y_i]}_{p^2}=mp^2$$

$S :=$ Anzahl der Knoten in der stabilen Menge

$S \geq X-Y$

$$E[S]\geq E[X-Y]=E[X]-E[Y]=np-mp^2$$

maximal für $p=\frac{n}{2m}$

Mehrere Zufallsvariablen

$\Pr[X=x, Y=y] = \Pr[\{w \in \Omega \mid X(w)=x, Y(w)=y\}]$

gemeinsame Dichte
der ZV X und Y $f_{X,Y}(x,y)$

$$\Pr[X=x] = \sum_{y \in W_Y} \Pr[X=x, Y=y]$$

Randdichten: $f_X(x) = \sum_{y \in W_Y} f_{X,Y}(x,y) \qquad f_Y(y) = \sum_{x \in W_X} f_{X,Y}(x,y)$

gemeinsame Verteilung $F_{X,Y}(x,y) := \Pr[X \leq x, Y \leq y] = \sum_{x' \leq x} \sum_{y' \leq y} f_{X,Y}(x', y')$

Unabhängigkeit von Zufallsvariablen

Definition 2.52. Zufallsvariablen $X_1, \ldots, X_n$ heissen unabhängig genau dann, wenn für alle $(x_1, \ldots, x_n) \in W_{X_1} \times \ldots \times W_{X_n}$ gilt

$$\Pr[X_1 = x_1, \overset{\wedge}{\ldots}, X_n = x_n] = \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n].$$

Lemma 2.53. Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen und sind $S_1, \ldots, S_n \subseteq \mathbb{R}$ beliebige Mengen, dann gilt

$$\Pr[X_1 \in S_1, \ldots, X_n \in S_n] = \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n].$$

Beweis:

Korollar 2.54. Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen und ist $I = \{i_1, \ldots, i_k\} \subseteq [n]$, dann sind $X_{i_1}, \ldots, X_{i_k}$ ebenfalls unabhängig.

Beweis mit Lemma 2.53 mit $S_i = \begin{cases} W_{X_i} & \text{falls } i \notin I \\ \{x_i\} & \text{falls } i \in I \end{cases}$

Satz 2.55. Seien $f_1, \ldots, f_n$ reellwertige Funktionen ($f_i \colon \mathbb{R} \to \mathbb{R}$ für $i = 1, \ldots, n$). Wenn die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind, dann gilt dies auch für $f_1(X_1), \ldots, f_n(X_n)$.

Beweis mit Lemma 2.53 mit $S_i = \{x \mid f(x) = z_i\}$ und $z_i \in W_{f(X_i)}$

Zusammengesetzte Zufallsvariablen

Satz 2.60. *(Linearität des Erwartungswerts)* Für Zufallsvariablen $X_1, \ldots, X_n$ und $X := a_1X_1 + \cdots + a_nX_n$ mit $a_1, \ldots, a_n \in \mathbb{R}$ gilt

$$\mathbb{E}[X] = a_1\mathbb{E}[X_1] + \cdots + a_n\mathbb{E}[X_n].$$

Satz 2.61. *(Multiplikativität des Erwartungswerts)* Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[X_1 \cdot \ldots \cdot X_n] = \mathbb{E}[X_1] \cdot \ldots \cdot \mathbb{E}[X_n].$$

Beweis: $n=2$ $\quad E[X \cdot Y] = \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot \Pr[X=x, Y=y] \overset{\text{Unabh.}}{=} \sum_{x \in W_X} \sum_{y \in W_Y} xy \Pr[X=x] \cdot \Pr[Y=y]$

$$= \left(\sum_{x \in W_X} x \cdot \Pr[X=x]\right)\left(\sum_{y \in W_Y} y \cdot \Pr[Y=y]\right)$$

$$= E[X] \cdot E[Y]$$

$\blacksquare$

Satz 2.62. Für unabhängige Zufallsvariablen $X_1, \dots, X_n$ und $X := X_1 + \dots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \dots + \mathrm{Var}[X_n].$$

Beweis: $\mathrm{Var}[Z]$ $\quad Z = X+Y$

$$= E[(X+Y)^2] - E[X+Y]^2 = E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

$$= E[X^2] + 2E[X]E[Y] + E[Y^2]$$

$$= (E[X] + E[Y])^2$$
$$= E[X]^2 + 2E[X]E[Y] + E[Y]^2$$

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$	$\forall X, Y$
$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$	$\forall X, Y$ unabhängig
$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$	$\forall X, Y$ unabhängig
$\mathrm{Var}[X \cdot Y] \neq \mathrm{Var}[X] \cdot \mathrm{Var}[Y]$	i.A. (auch für unabhängige ZV)

Waldsche Identität

Satz 2.65 (Waldsche Identität). N und X seien zwei unabhängige Zufallsvariable, wobei für den Wertebereich von N gilt: $W_N \subseteq \mathbb{N}$. Weiter sei

$$Z := \sum_{i=1}^{N} X_i,$$

wobei $X_1, X_2, \dots$ unabhängige Kopien von X seien. Dann gilt:

$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X]. \Rightarrow$$

MINIQUIZFRAGEN

Für jede Zufallsvariable X gilt: $\mathbb{E}[X^2] = (\mathbb{E}[X])^2$.

Falsch, weil sonst $Var[X]=0 \quad \forall X$

Sind X und Y zwei positive Zufallsvariablen für einen Wahrscheinlichkeitsraum Ω, so dass für alle $\omega \in \Omega$ gilt: $X(\omega) \leq Y(\omega)$, dann gilt:

$\mathbb{E}[2X] \geq \mathbb{E}[Y]$ Falsch, $\Omega = \{1,2\}$, Laplace, $X(\omega)=\omega$, $Y(\omega)=\omega^3$
$E[2X]=3 \quad E[Y]=0{,}5+4=4{,}5$ $\quad 2E[X]=\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 4$ $\quad E[Y]=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 8$

$\mathbb{E}[2X] \leq \mathbb{E}[Y]$ Falsch, $\Omega=\{1,2\}$, Laplace, $X(\omega)=\omega$, $Y(\omega)=\omega^2$

$\mathbb{E}[X-Y] \leq 0$ Richtig, Linearität des EW

$\mathbb{E}[XY] \leq \mathbb{E}[Y^2]$ Richtig

$= \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot Pr[X=x, Y=y]$

Wir werfen einen fairen, sechsseitigen Würfel 4 mal. Sei X die Summe der Augen der vier Würfe, dann ist X binomial verteilt.

Falsch

Ben wirft eine (faire) Münze so lange bis zum ersten Mal Kopf kommt. Anschliessend wirft er einen Würfel so lange bis eine sechs kommt. Berechnen Sie den Erwartungswert für die Gesamtanzahl Würfe (Münze *und* Würfel).

8

Wir werfen einen Würfel so lange bis eine 5 kommt. Mit X bezeichnen wir die Anzahl der benötigten Würfe. Danach werfen wir den Würfel nochmals X mal und bezeichnen mit Y die Anzahl gerader Augenzahlen, die wir dabei sehen.

Bestimme $\mathbb{E}[Y]$.

3

Bestimmen Sie diejenigen Aussagen die für *jede* Zufallsvariable X gelten:

$\mathrm{Var}[X] \geq \mathbb{E}[X]$ Falsch, $X(\omega)=5 \quad \forall \omega \in \Omega$, $Var[X]=0$, $E[X]=5$

$\mathrm{Var}[X] \geq 0$ Richtig

$\mathbb{E}[X - \mathbb{E}[X]] = 0$ Richtig

$\mathrm{Var}[X - \mathrm{Var}[X]] = 0$ Falsch, $Var[X+b]=Var[X]$

Seien X und Y zwei Zufallsvariabeln mit $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. Dann sind X und Y unabhängig.

Falsch

Seien X, Y unabhängige Zufallsvariablen mit $Var(X) = 1$ und $Var(Y) = 4$.
Dann ist $Var(X - Y) = -3$.

$Var[-X] = Var[X]$ Falsch
$Var[(-1)\cdot X] = (-1)^2 \cdot Var[X]$

Chinese Whispers

In this exercise, you are supposed to compute probabilities for the children's game *Chinese Whispers* (also known as *Broken Telephone*).

The game consists of n people lined up in a straight line $1, \ldots, n$, and a start player—in this case, you. You devise a message, which may be either a 0 or a 1, and whisper it to the first person in the line (person 1). The message then propagates in this fashion throughout the line, meaning the i-th person whispers it to the $(i+1)$-st and so on until the message reaches the n-th person. That person then shouts it out loud in order to compare it with your original message.

However, this can go wrong, as each person in the line may misinterpret a message with some probability. Namely, the i-th person interprets and whispers forward the opposite of what they hear (i.e., 0 instead of 1, or 1 instead of 0) with probability p_i.

Knowing the probabilities p_i for all $i \in \{1, \ldots, n\}$, can you compute the probability of the final message matching your original message? Note that we only require that the start message and final message are the same; we do **not** require that all intermediate steps were correct.

Input

The first line of the input file contains a number $t \leq 30$ of test cases. Each of the t test cases is described as follows:

- It starts with a line consisting of two integers n and m, separated by a space, denoting the number of people participating in the game other than yourself $(1 \leq n \leq 10^4)$ and your starting message $(m \in \{0, 1\})$.
- The following line consists of n real numbers $p_1, p_2, \ldots, p_n$, separated by spaces, denoting the probability that the i-th person forwards the opposite of what they receive $(0 \leq p_i \leq 1$, for all $i \in \{1, \ldots, n\})$.

Output

For each test case, output a single line with the probability that your message matches what the n-th person shouts out loud. Each output value should be a real number between 0 and 1. Your solution will be accepted if it has an absolute or relative error of at most 10^{-3}.

Points

There are three groups of test sets, worth 100 points in total:

1. For the first group of test sets, worth 20 points, you may assume that $n \leq 3$.
2. For the second group of test sets, worth 30 points, you may assume that $n \leq 20$.
3. For the third group of test sets, worth 50 points, there are no additional assumptions.

DP[i,j] = W'keit, dass die letzte Person das richtige Wort versteht, wenn Person i Wort j verstanden hat

$DP[i,j] = p[i] \cdot DP[i+1, (j+1) \bmod 2] + (1-p[i]) \cdot DP[i+1, j]$

Base Case: $DP[n,m] = 1$, $DP[n, (m+1) \bmod 2] = 0$

Result: $DP[0,m]$

```java
import algorithms.*;

class Main {
    public static void main(String[] args) {
        // Uncomment this line if you want to read from a file
        In.open("public/custom.in");
        Out.compareTo("public/custom.out");

        int t = In.readInt();
        for (int i = 0; i < t; i++) {
            testCase();
        }

        // Uncomment this line if you want to read from a file
        // In.close();
    }

    static int n, m;
    static double[] p;
    static double[][] dp;

    public static void testCase() {
      n = In.readInt();
      m = In.readInt();
      p = new double[n];
      dp = new double[n+1][2];
      for(int i=0; i<n; i++){
        p[i] = In.readDouble();
      }
      for(int i=0; i<=n; i++){
        for(int j=0; j<2; j++) dp[i][j] = -1;
      }
      // dp[i][j] prob that the last person understands the right word, when person i understood word j
      Out.println(solve(0, m));
    }
```

```java
    public static double solve(int i, int j){
      if(i==n){
        if(j==m) return 1;
        return 0;
      }
      if(dp[i][j] != -1) return dp[i][j];

      dp[i][j] = p[i]*solve(i+1, (j+1)%2) + (1-p[i])*solve(i+1, j);
      return dp[i][j];
    }
}
```

Schranken

Satz 2.67. *(Ungleichung von Markov)* Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Oder äquivalent dazu $\Pr[X \geq t \cdot \mathbb{E}[X]] \leq 1/t$.

$$E[X] = \sum_{x \in W_X} x \cdot Pr[X=x] \geq \sum_{\substack{x \in W_X \\ x \geq t}} x \cdot Pr[X=x] \geq t \cdot \sum_{\substack{x \in W_X \\ x \geq t}} Pr[X=x]$$
$$= t \cdot Pr[X \geq t]$$

Satz 2.68. *(Ungleichung von Chebyshev)* Sei X eine Zufallsvariable und $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

$$\Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

oder äquivalent dazu $\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\mathrm{Var}[X]}] \leq 1/t^2$.

$$Pr[|X-E[X]| \geq t] = Pr[\underbrace{(X-E[X])^2}_{Y} \geq t^2] \overset{\text{markov}}{\leq} \frac{E[Y]}{t^2} = \frac{Var[X]}{t^2}$$

$E(\downarrow)$

Anwendung:

$$Pr[X \geq t + E[X]] = Pr[X - E[X] \geq t] \leq Pr[|X-E[X]| \geq t]$$

$X - E[X] \leq -t$

Sie findet in ihrem Schrank eine grosse Schachtel voller Perlen und ohne zu beachten welche Farben die Perlen haben zieht sie eine Perle nach der andern und reiht sie aneinander um eine schöne Kette zu bilden. Sie konstruiert eine Kette mit $n \geq 3$ farbigen Perlen. Jede Perle ist rot mit Wahrscheinlichkeit $\frac{1}{2}$ und sonst blau, unabhängig von den anderen Perlen. Die Perlen formen natürlich einen Kreis, der Verschluss der Kette wird ignoriert. Wir sind interessiert an der Zufallsvariable X, der Anzahl Farbenwechsel entlang der Kette.

Wir definieren $X_1, \ldots, X_n$ wie folgt. Für $1 \leq i \leq n-1$, sei X_i die Indikatorvariable für das Ereignis dass die Perlen i und $i+1$ unterschiedliche Farben haben und sei X_n die Indikatorvariable dass Perlen n und 1 unterschiedliche Farben haben. Offensichtlich gilt $X = X_1 + X_2 + \cdots + X_n$.

Sei $1 \leq i \leq n$. Berechnen Sie $\mathbb{E}[X_i]$ und $\mathbb{E}[X]$.

$E[X_i] = \frac{1}{2}$

$E[X] = \frac{n}{2}$

Sei $t > 0$. Benutzen Sie die Markov Ungleichung um eine obere Schranke für die Wahrscheinlichkeit $\Pr[X \geq \mathbb{E}[X] + t]$ zu finden.

$Pr[X \geq E[X]+t] \leq \frac{E[X]}{E[X]+t} = \frac{\frac{n}{2}}{\frac{n}{2}+t} = \frac{n}{n+2t}$

Sei $1 \leq i, j \leq n$. Was ist der Erwartungswert $\mathbb{E}[X_i \cdot X_j]$? (Ihre Antwort sollte von i, j abhängen.) Benutze dies zusammen mit der Formel $X^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} X_i \cdot X_j$ um $\mathbb{E}[X^2]$ und die Varianz von X zu berechnen.

$E[X^2] = \sum_{i=1}^{n} \sum_{j=1}^{n} E[X_i \cdot X_j]$

Fall ①: $i=j$ n-mal

$X_i \cdot X_j = X_i^2 \qquad E[X_i \cdot X_j] = \frac{1}{2}$

Fall ②: $|i-j| \equiv_n 1$ 2n-mal

$X_i \cdot X_j = 1 \Leftrightarrow X_i = 1, X_j = 1$

$E[X_i \cdot X_j] = \frac{1}{4}$

Fall ③: $|i-j| \geq 2$ n^2-3n-mal

$E[X_i \cdot X_j] = E[X_i] \cdot E[X_j] = \frac{1}{4}$

$E[X^2] = n \cdot \frac{1}{2} + 2n \cdot \frac{1}{4} + \frac{n^2-3n}{4} = n + \frac{n^2-3n}{4} = \frac{n^2+n}{4}$

$Var[X] = E[X^2] - E[X]^2 = \frac{n^2+n}{4} - \frac{n^2}{4} = \frac{n}{4}$

Sei $t > 0$. Benutzen Sie die Chebyshev Ungleichung um eine obere Schranke für $\Pr[X \geq \mathbb{E}[X] + t]$ zu berechnen.

Zusätzliche Aufgaben:

Aufgabe 1 – *Dominante Menge*

Sei $G = (V, E)$ ein Graph. Eine 'dominante Menge' $W \subseteq V$ ist eine Knotenmenge, sodass für jeden Knoten $v \in V$ gilt, dass entweder v selbst oder ein Nachbar von v in W enthalten ist. In dieser Aufgabe betrachten wir einen randomisierten Algorithmus, der eine dominante Menge findet und dem Algorithmus aus der Vorlesung zum Finden einer stabilen Menge ähnelt.

Wir nehmen an, dass G Minimalgrad mindestens $d > 1$ hat, d.h. jeder Knoten $v \in V$ hat Grad $deg(v) \geq d$. Der Algorithms besteht aus zwei Runden. In der ersten Runde markieren wir jeden Knoten unabhängig von den anderen Knoten mit Wahrscheinlichkeit p (wobei $0 \leq p \leq 1$ gegeben ist). In der zweiten Runde betrachten wir jeden Knoten $v \in V$, wenn weder v noch einer seiner Nachbarn in der ersten Runde markiert wurden, so markieren wir v. Man sieht leicht, dass die Menge der markierten Knoten nach der zweiten Runde eine dominante Menge ist. Wir analysieren die Grösse dieser dominanten Menge im Folgenden.

(a) Sei X die Anzahl der Knoten, die in der ersten Runde markiert werden. Berechnen Sie $\mathbb{E}[X]$.

(b) Sei $v \in V$ ein beliebiger (aber fixer) Knoten. Berechnen Sie die Wahrscheinlichkeit, dass weder v noch einer der Nachbarn von v markiert wurde genau (die Antwort darf von v Abhängen). Finden Sie eine obere Schranke für diese Wahrscheinlichkeit, die nur von d und p abhängt (und nicht von v).

(c) Sei Y die Anzahl der Knoten, die in der zweiten Runde markiert werden. Verwenden Sie (b) um eine obere Schranke für $\mathbb{E}[Y]$ zu finden.

(d) Zeigen Sie, dass die erwartete Grösse der dominanten Menge höchstens $n(p + e^{-p(d+1)})$ ist. *Hinweis:* Sie dürfen die Ungleichung $1 - x \leq e^{-x}$ ohne Beweis verwenden.

(e) Zeigen Sie, dass G eine dominante Menge der Grösse höchstens $n \cdot \frac{1+\ln(d+1)}{d+1}$ enthält.

Aufgabe 1 – *Establish Dominance*

Sei $G = (V, E)$ ein Graph mit $V = [n]$. Wir definieren den Wahrscheinlichkeitsraum $\Omega = \{0, 1\}^n$. Wobei für $\omega = \omega_1\omega_2 \dots \omega_n$ gilt das $\Pr[\omega_v = 1] = p$ für alle $v \in V$ unanbängig von allen anderen $\omega_{v'}$. Es ist zu beachten, dass, obwohl unser Algorithmus zweistufig ist, wir den Wahrscheinlichkeitsraum so definieren können, da die zweite Stufe des Algorithmus keinen "neuen" Zufall braucht. In der ersten Runde fügen wir alle Knoten v für die gilt $\omega_v = 1$ in unsere dominante Menge ein. In der zweiten Runde dann alle Knoten w für die gilt $\omega_w = 0$ und $\omega_u = 0$ für alle $u \in N(w)$.

(a) Wir definieren Indikatorzufallsvariablen für jeden Knoten $v \in V$

$$X_v = \begin{cases} 1 & \text{falls } v \text{ in der ersten Runde makiert wurde,} \\ 0 & \text{falls } v \text{ in der ersten Runde nicht markiert wurde.} \end{cases}$$

Dann gilt $X = \sum_{v \in V} X_v$. Des Weiteren gilt

$$\mathbb{E}[X_v] = \Pr[X_v = 1] = \Pr[\omega_v] = p.$$

Durch die Linearität des Erwartungswerts folgt

$$\mathbb{E}[X] = \sum_{v \in V} \mathbb{E}[X_v] = np.$$

(b) Wir definieren das Ereignis

$$F_v = \text{'weder } v \text{ noch einer seiner Nachbarn wurde in der ersten Runde markiert'.}$$

Für einen beliebigen Knoten w definieren wir das Ereignis

$$E_w = \text{'}w \text{ wurde in der ersten Runde nicht markiert'.}$$

Es ist leicht zu sehen das gilt

$$\Pr[E_w] = \Pr[\omega_w = 0] = 1 - p.$$

Man beachte das gilt

$$F_v = \bigcap_{w \in \{v\} \cup N(v)} E_w.$$

Wir wissen das die E_w's unanhängig sind, da wir die Knoten in der ersten Stufe unabhängig voneinander zur dominanten Menge hinzufügen. Daraus folgt

$$\Pr[F_v] = \prod_{w \in \{v\} \cup N(v)} \Pr[E_w] = (1-p)^{\deg(v)+1} \leq (1-p)^{d+1},$$

wobei der letzte Schritt folgt weil $1 - p \leq 1$ und somit $(1-p)^{\deg(v)+1}$ monoton fallend in $\deg(v)$ ist.

(c) Wir definieren Indikatorzufallsvariablen für jeden Knoten $v \in V$

$$Y_v = \begin{cases} 1 & \text{falls } v \text{ in der zweiten Runde makiert wurde,} \\ 0 & \text{falls } v \text{ in der zweiten Runde nicht markiert wurde.} \end{cases}$$

Es ist leicht zu sehen, dass nach der Beschreibung des Algorithmus Y_v genau dann 1 ist, wenn F_v eintritt. Aus der Berechnung aus (b) folgt

$$\mathbb{E}[Y_v] = \Pr[Y_v = 1] = \Pr[F_v] \leq (1-p)^{d+1}.$$

Aus der Linearität des Erwartungswerts folgt für $Y = \sum_{v \in V} Y_v$, dass

$$\mathbb{E}[Y] = \sum_{v \in V} \mathbb{E}[Y_v] \leq n(1-p)^{d+1}.$$

(d) Sei Z die Anzahl der Knoten in der dominierenden Menge, die der Algorithmus liefert. Es gilt

$$Z = X + Y.$$

Aus der Linearität des Erwartungswerts folgt:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y] \leq np + n(1-p)^{d+1}.$$

Durch die in der Aufgabenstellung gegebene Ungleichung erhalten wir $(1-p)^{d+1} \leq e^{-p(d+1)}$. Daraus folgt

$$\mathbb{E}[Z] \leq n\left(p + e^{-p(d+1)}\right).$$

(e) Wir wollen den in (d) erhaltenen Term in Abhängigkeit von p optimieren. Hierfür berechnen wir die erste Ableitung.

$$\frac{d}{dp} n\left(p + e^{-p(d+1)}\right) = n\left(1 - (d+1)e^{-p(d+1)}\right)$$

Wir setzen

$$n\left(1 - (d+1)e^{-p(d+1)}\right) \stackrel{!}{=} 0.$$

Auflösen nach p gibt

$$p = \frac{\ln(d+1)}{d+1}.$$

Es ist leicht zu überprüfen das $p \in [0,1]$.

Man sieht (zum Beispiel über die zweite Ableitung oder anders) leicht, dass es sich bei diesem Wert von p um ein Minimum handelt (und keinen Sattelpunkt/Maximum)

Einsetzten in unsere Formel für $\mathbb{E}[Z]$ gibt

$$\mathbb{E}[Z] \leq n\left(\frac{\ln(d+1)}{d+1} + e^{-\ln(d+1)}\right) = n\left(\frac{1+\ln(d+1)}{d+1}\right).$$

Dies bedeutet, dass für $p = \frac{\ln(d+1)}{d+1}$ die erwartete Grösse der dominanten Menge höchstens $n\left(\frac{1+\ln(d+1)}{d+1}\right)$ ist. Insbesondere bedeutet dies, dass eine solche dominierende Menge existieren muss (generell gibt es immer Elementarereignisse $\omega_1, \omega_2 \in \Omega$, so dass $X(\omega_1) \geq \mathbb{E}[X]$ und $X(\omega_2) \leq \mathbb{E}[X]$).

Aufgabe 1 – *Der kleine Bruder des Coupon Collectors*

Die neue Serie von Sammelkarten mit Bildern berühmter Fussballspieler ist gerade erschienen. Die Serie besteht aus n verschiedenen Sammelkarten. Sie kaufen Pakete mit je einer Sammelkarte, die unabhängig, uniform zufällig aus den n Sammelkarten ausgewählt wurde. Da Sie nicht an doppelten Bildern interessiert sind, geben Sie alle Bilder, die Sie bereits besitzen, an Ihren kleinen Bruder weiter.

(a) Nehmen Sie an, dass $\sqrt{n}$ eine gerade Zahl ist. Sie haben bereits $2\sqrt{n}$ Bilder gekauft. Zeigen Sie, dass Ihre Bruder mit Wahrscheinlichekeit mindestens 0.5 mindestens ein Bild erhalten hat.

(b) Sie wollen sowohl Ihre eigene als auch die Sammlung Ihres Bruders vervollständigen. Berechnen Sie die asymptotisch beste obere Schranke (in $O(\cdot)$ Notation) auf die erwartete Anzahl von Bildern, die Sie hierfür kaufen müssen.

(c) Messaldo ist der Lieblingsspieler Ihres Bruders. Berechnen Sie die erwartete Anzahl von Bildern, die Sie kaufen müssen, bis Ihr Bruder das Bild von Messaldo erhält.

(d) Sei T die Zahl der Bilder, die Sie kaufen müssen, bis Sie selbst $n/2$ verschiedene Bilder gesammelt hat. Berechnen Sie $\mathbb{E}[T]$. Geben Sie das Ergebnis zusätzlich (möglichst einfach) in O-Notation an.

Aufgabe 1 – *Der kleine Bruder des Coupon Collectors*

(a) Wir berechnen die Wahrscheinlichkeit, dass die ersten $2\sqrt{n}$ Karten, die wir ziehen, paarweise verschieden sind. Formal definieren wir als E_i das Eregeinis, dass unsere ersten i Karten paarweise verschiedenen sind. Es genügt dann zu zeigen, dass $\Pr[E_{2\sqrt{n}}] \leq 1/2$ ist, da unser kleiner Bruder genau dann eine Karte bekommt, wenn $\overline{E}_{2\sqrt{n}}$ eintritt.

Wenn wir bereits i verschiedene Karten gezogen haben, ist die Warhscheinlichkeit, dass wir beim $(i+1)$-ten Mal wiederum eine neue Karten ziehen $(1-\frac{i}{n})$. Als Formel ausgedrückt bedeutet das

$$\Pr[E_{i+1} \mid E_i] = 1 - \frac{i}{n}$$

Mithilfe der Ungleichung $1 - x \leq e^{-x}$, die für alle reelen Zahlen x gilt, können wir E_k für $k = 2\sqrt{n}$ mit unserer obigen Beobachtung wie folgt abschätzen:

$$\Pr[E_k] = \prod_{i=1}^{k} \Pr[E_i \mid E_{i-1}] = \prod_{i=1}^{k-1}\left(1-\frac{i}{n}\right) \leq \prod_{i=1}^{k-1} e^{-\frac{i}{n}} \leq e^{-\frac{1}{n}\sum_{i=1}^{k-1} i} = e^{\frac{(k-1)k}{2n}} \leq e^{-1} \leq 1/2.$$

Für die vorletzte Ungeleichung haben wir verwendet, dass $k = 2\sqrt{n}$ und $k-1 \geq \sqrt{n}$, woraus $k(k-1) \geq 2n$ folgt.

(b) Sei T die erwartete Anzahl Karten, die wir brauchen, bis wir alle Karten gezogen haben (ganz unabhängig von unserem kleinen Bruder). Aus der Vorlesung wissen wir, dass $\mathbb{E}[T] = \Theta(n \log n)$.

Sei B die Anzahl Karten, die wir ziehen müssen, bis unser Bruder alle Karten hat. Offensichtlich gilt $B \geq T$ und daher auch $\mathbb{E}[B] \geq \Omega(n \log n)$.

Wir wollen nun zeigen, dass diese untere Schranke bestmöglich ist, also dass $\mathbb{E}[B] \leq O(n \log n)$. Dazu machen wir ein kleines Gedankenexperiment und stellen uns vor, wir hätten eine grosse Schwester, mit der wir ebenfalls eine Variante des Coupon-Collector-Spiels spielen[1]. Für diese Variante warten wir, bis wir allen Karten selbst gezogen haben. Danach geben wir jede Karte, die wir ziehen, an unsere Schwester weiter. Sei S die Anzahl Karten, die wir ziehen müssen, bis unsere Schwester alle Karten hat. Man sieht leicht, dass $\mathbb{E}[S] = 2\,\mathbb{E}[T] = O(n \log n)$, da wir zweimal hintereinander das Coupon-Collector Spiel spielen.

Ausserdem kann man sehen, dass $B \leq S$: Dafür stellen wir uns vor, dass wir mit unseren Geschwistern gleichzeitig spielen (und wenn wir eine Karte ziehen, die wir laut Spielregeln beiden gleichzeitig geben müssen, dann fertigen wir eine Kopie an und geben die Karte tatsächlich beiden[2]). In dieser Variante des Spiels hat unser kleiner Bruder immer mindestens so schnell alle Karten, wie unsere grosse Schwester und daher gilt $B \leq S$. Daraus folgt $\mathbb{E}[B] \leq \mathbb{E}[S] = O(n \log n)$.

Somit haben wir gezeigt, dass $\mathbb{E}[B] = \Theta(n \log n)$.

(c) Sei X die Anzahl Karten, die wir ziehen müssen bis wir einmal Messaldo ziehen. Sei Y die Anzahl Karten, die wir ziehen müssen nachdem wir zum ersten Mal Messaldo hatten bis wir ihn zum zweiten Mal ziehen. Dann erhält unser Bruder nach genau $X+Y$ Karten sein erstes Mal Messaldo.

Bei jeder Karte ist die Wahrscheinlichkeit Messaldo zu ziehen $1/n$. Daher sind sowohl X als auch Y geometrisch verteilt mit Parameter $p = 1/n$ und Erwartungswert n.

Laut Linearität des Erwartungswerts gilt $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = n + n = 2n$.

(d) Sei, genau wie in der Volresung, X_i die Anzahl Karten, die wir ziehen müssen, um von der i-ten bis zur $(i+1)$-ten Karten zu kommen. Es gilt $\mathbb{E}[X_i] = \frac{n}{n-i}$ wie wir in der Vorlesung gesehen haben.

Die Zeit X bis zur $n/2$-ten Karte ist $X = X_0 + \ldots + X_{n/2-1}$. Mit der Linearität des Erwartungswert erhält man

$$\mathbb{E}[X] = \sum_{i=0}^{n/2-1} \frac{n}{n-i}$$

Man sieht $\mathbb{E}[X] \geq \sum_{i=0}^{n/2-1} 1 = n/2$ (da jeder Summand mindestens 1 ist) und $\mathbb{E}[X] \leq \sum_{i=0}^{n/2-1} 2 = n$ (da jeder Summand höchstens 2 ist). Daraus folgt $\mathbb{E}[X] = \Theta(n)$.

Aufgabe 2 – *Varianz*

Wir betrachten dieselbe Situation wie in Aufgabe 2 von Extra-Blatt 6. Gegeben sei also ein Graph $G = (V, E)$ mit n Knoten und m Kanten, und wir betrachten den Laplace-Raum $\Omega = \{S \mid S \subseteq V\}$ und die Zufallsvariable $X :=$"Anzahl Kanten über den Schnitt $(S, V \setminus S)$". Ausserdem definieren wir für jede Kante $e \in E$ die Indikatorvariable X_e für das Event $E_e :=$"e läuft über den Schnitt $(S, V \setminus S)$". Offenbar ist $X = \sum_e X_e$.

(a) Seien $e, f \in E$. Berechnen Sie $\mathbb{E}[X_e \cdot X_f]$.
Hinweis: Unterscheiden Sie die Fälle, dass e und f sich in 0, 1, und 2 Knoten überschneiden.

(b) Berechnen Sie $\mathbb{E}[X^2]$.
Hinweis: Benutzen Sie $X^2 = \sum_{e,f \in E} X_e \cdot X_f$ und Teil (a).

(c) Berechnen Sie aus $\mathbb{E}[X^2]$ die Varianz und die Standardabweichung von X.

(d) Betrachten Sie folgenden Algorithmus zur Erzeugung eines Schnittes, der fast Grösse $m/2$ (oder mehr) hat:

1. Wähle $S \subseteq V$ uniform zufällig.
2. IF $|E(S, V \setminus S)| \geq \frac{m}{2} - \sqrt{m}$ THEN return S ELSE repeat from 1.

Geben Sie eine (möglichst gute) obere Schranke für die erwartete Anzahl Schleifendurchläufe des Algorithmus an.

Aufgabe 2 – *Varianz*

Den beschriebenen Laplace-Raum können wir erzeugen, indem wir für jeden Knoten v unabhängig voneinander entscheiden, ob er in S oder in $V \setminus S$ landet (jeweils mit Wahrscheinlichkeit $1/2$). Das sieht man sofort, da dann für jedes $S \subseteq V$ gilt $\Pr[S] = 2^{-|S|} \cdot 2^{-|V \setminus S|} = 2^{-n}$. Wir werden daher von nun an von dieser Beschreibung ausgehen.

(a) Ist $e \cap f = \emptyset$, so sind die Indikatorvariablen X_e und X_f unabängig, womit $\mathbb{E}[X_e \cdot X_f] = \mathbb{E}[X_e] \cdot \mathbb{E}[X_f] = 1/4$ gilt.
Überschneiden sich e und f hingegen in einem Knoten, so können wir o.b.d.A. annehmen, dass $e = \{u, v\}$ und $f = \{v, w\}$. Folglich ist $X_e \cdot X_f = 1$, falls entweder $u, w \in S$ und $v \notin S$ oder falls $u, w \notin S$ und $v \in S$. Somit gilt

$$\mathbb{E}[X_e \cdot X_f] = \Pr[u \in S, w \in S, v \notin S] + \Pr[u \notin S, w \notin S, v \in S] = (1/2)^3 + (1/2)^3 = 1/4.$$

Alternativ kann man auch argumentieren, dass X_e und X_f auch in diesem Fall unabhängig sind, da die Events "$u \in S$" und "$w \in S$" unabhängig voneinander sind.

Im dritten Fall überschneiden sich e und f in 2 Konten. Dann ist $e = f$ und somit $X_e \cdot X_f = X_e^2 = X_e$. (Die letzte Ungleichung gilt, da X_e nur die Werte 0 und 1 annehmen kann, und $0^2 = 0$ und $1^2 = 1$ gelten.) Also ist $\mathbb{E}[X_e \cdot X_f] = \mathbb{E}[X_e] = 1/2$.

(b) Da es insgesamt m^2 Paare $e, f \in E$ gibt und mit Hilfe der Linearität des Erwartungswertes erhalten wir

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}[\sum_{e,f \in E} X_e \cdot X_f] \\
&= \sum_{e,f \in E} \mathbb{E}[X_e \cdot X_f] \\
&= \sum_{e \in E} \mathbb{E}[X_e] + \sum_{e \neq f \in E} \mathbb{E}[X_e \cdot X_f] \\
&= m/2 + (m^2 - m) \cdot 1/4 = \frac{m^2 + m}{4}
\end{aligned}$$

(c)

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{m^2 + m}{4} - \left(\frac{m}{2}\right)^2 = m/4.$$

(d) Als erstes überlegen wir uns, wie wir Wahrscheinlichkeit $\Pr[|E(S, V \setminus S)| < \frac{m}{2} - \sqrt{m}]$ auf die Form von Chebyshevs Ungleichung überführen können:

$$\begin{aligned}
\Pr[|E(S, V \setminus S)| < \tfrac{m}{2} - \sqrt{m}] &= \Pr[X < \mathbb{E}[X] - \sqrt{m}] \\
&= \Pr[\sqrt{m} < \mathbb{E}[X] - X] \\
&\leq \Pr[\sqrt{m} < |\mathbb{E}[X] - X|] \\
&= \Pr[\sqrt{m} < |X - \mathbb{E}[X]|] \\
&\leq \Pr[|X - \mathbb{E}[X]| \geq \sqrt{m}]
\end{aligned}$$

Somit gilt dank Chebyshevs Ungleichung, dass

$$\Pr[|E(S, V \setminus S)| < \tfrac{m}{2} - \sqrt{m}] \leq \mathrm{Var}[X]/m = 1/4.$$

Nun können wir die Erfolgswahrscheinlichkeit des 2. Schrittes des Algorithmus abschätzen durch

$$\Pr[|E(S, V \setminus S)| \geq \tfrac{m}{2} - \sqrt{m}] = 1 - \Pr[|E(S, V \setminus S)| < \tfrac{m}{2} - \sqrt{m}] \geq 3/4,$$

womit wir eine obere Schranke von $4/3$ für die erwartete Anzahl Schleifendurchläufe erhalten.

Exercise 1 – *Probabilistic method*

This is one of the most beautiful applications of so called "Probabilistic Method", due to Paul Erdős — a method of using probability theory to show existence of specific combinatorial objects, that on its face have nothing to do with probability.

Specifically, as mentioned in Lecture 7, we will show that every k, there is a graph G that has no triangle (so the largest complete subgraph is just K_2), but has chromatic number at least k.

Let us consider a random graph on n vertices, such that every edge is included independently with probability $p = \alpha/n$, where $\alpha = \log n$, and let $\beta = m/n$ for some $m < n$ be a small constant (it will be chosen s.t. $\beta \approx 1/2k$).

(a) Show that if T is a random variable denoting the number of triangles in G, then $\mathbb{E}[T] \leq \alpha^3$.

(b) Let M be a random variable denoting the number of independent sets of size $n\beta$ in G. Show that

$$\mathbb{E}[M] \leq 2^n \exp(-\alpha\beta^2 n/3).$$

Conclude that when n is large enough, probability that G has an independent set of size $n\beta$ is at most $\exp(-100)$.

(c) Show that this implies that for every n larger than some n_0, there is a graph G with at most $10\alpha^3 = 10\log^3 n$ triangles, and no independent set of size $n\beta$.

(d) Show if n is large enough, this implies that there is a graph G' with no triangles, and chromatic number at least $2/\beta$.

Solution 1

(a) For any three distinct vertices $x, y, z \in [n]$, Let $T_{x,y,z}$ be an indicator random variable denoting whether triangle $\{x, y, z\}$ is present in G (i.e. all three edges $\{x, y\}, \{y, z\}$ and $\{z, x\}$ are present in G). On one hand, since the edges are independent, we have $\mathcal{E}T_{x,y,z} = p^3$, and therefore

$$\mathcal{E}[T] = \sum_{(x,y,z)\in\binom{[n]}{3}} \mathcal{E}[T_{x,y,z}] = \binom{n}{3} p^3 \leq n^3 p^3 \leq \alpha^3.$$

(b) For a given subset $U \subset [n]$ of size βn, let M_U be an indicator random variable denoting that U is independent set in G. For n large enough we have $\binom{\beta n}{2} = \frac{\beta n(\beta n - 1)}{2} \leq \beta^2 n^3/3$. Therefore

$$\mathcal{E}[M_U] = (1-p)^{\binom{\beta n}{2}} \leq (1-p)^{\beta^2 n^2/3} \leq \exp(-pn^2\beta^2/3) \leq \exp(-n\alpha\beta^2/3).$$

Since $M = \sum_{U\in\binom{n}{\beta n}} M_U$ and $\binom{n}{\beta n} \leq 2^n$ (the number of subsets of size βn is at most the number of all subsets, we have

$$\mathcal{E}M \leq 2^n \exp(-n\alpha\beta^2/3) = \exp(-n\alpha\beta^2/3 + n\ln 2).$$

If $\alpha = \log n$ and β is a small constant, then for n large enough, we have $-n\log n\beta^2/3 + n\ln(2) < -100$, and hence $\mathcal{E}[M] < \exp(-100)$. By Markov inequality,since $M \geq 0$, we have $\Pr[M > 1] \leq \mathcal{E}[M] \leq \exp(-100)$.

(c) Let A be an event that G has more than $3\alpha^3$ triangles, and B be an event that G has independent set of size n. By union bound, we have $\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$, and by Markov inequality $\Pr[A] \leq 1/3$. By previous exercise, $\Pr[B] \leq \exp(-100)$. Hence $\Pr[A \cup B] \leq 1/3 + \exp(-100) \leq 1/2$. So there exist a graph for which neither A nor B holds.

(d) Consider a graph G' where we take graph G as in the previous sub-exercise and remove all vertices in all triangles in G. We remove at most $3\alpha^3 = 3\log^3 n$ vertices, so if n is large enough, $3\log^3 n < n/2$, and the graph G' has $n' > n/2$ vertices. Clearly, by construction G' is triangle free, and moreover has no independent set of size βn. As such, the chromatic number of G' is at least $n'/(\beta n) \geq n/(2\beta n) \approx (1/2\beta)$.

Exercise 2 – *Extra exercise*

Adjust those ideas to show that for every t and k there is a graph with no cycles shorter than t, and chromatic number at least k.