Primary simulations with different schemes of "taboo"-ing

Name: Somak Laha (MB 2016) Date: December 7, 2021

Multiple simulations of "Depreferential attachment random graph through preferential attachment" (denoted by $(G_t^{(m,p)})_{t\geq 1}$) has been performed to obtain ideas about distribution of various statistics of the model through simulation. In this document I shall mainly focus on the statistic **degree sequence**. For $(G_t^{(m,p)})_{t\geq 1}$, the notation $P_k(t)$ stands for the number of k-degree vertices at time t.

$$P_k(t) = \frac{1}{t}\sum_{i=1}^{t} \mathbb{1}_{\{D_i(t)=k\}}.$$

For simplification, we have taken $m = 1$, i.e., when a new vertex comes, it comes with only $m = 1$ many half-edges. Also, three different methods have been taken for choosing the *taboo* vertices.

- **Hard-code:** In this method, the model has a parameter $p\,(0 < p < 1)$. At time point $t \geq 1$, we choose the $\lfloor (t+1)p \rfloor$ vertices in such a way that all the other vertices in $G_{t-1}^{(1,p)}$ had degree less than the chosen ones. The chosen vertices are made taboo.
- **Soft-code:** In this method too, the model has a parameter $p\,(0 < p < 1)$. At time point $t \geq 1$, we choose the $\lfloor (t+1)p \rfloor$ vertices in a preferential way, i.e., choosing a vertex v_j is proportional to its degree in $G_{t-1}^{(1,p)}$.
- **Soft-code (alternative):** This method does not have the parameter p. Instead, in this method, at time $t \geq 1$, one will choose only one vertex preferentially. That chosen vertex will be made taboo.

Based on these different taboo-ing schemes, multiple simulations have been performed by using three different values of the parameter p. For each of the cases, 500 simulations have been performed to compute mean, variance and standard deviation of $P_k(2000)$ as function of k.

1 Hard-code

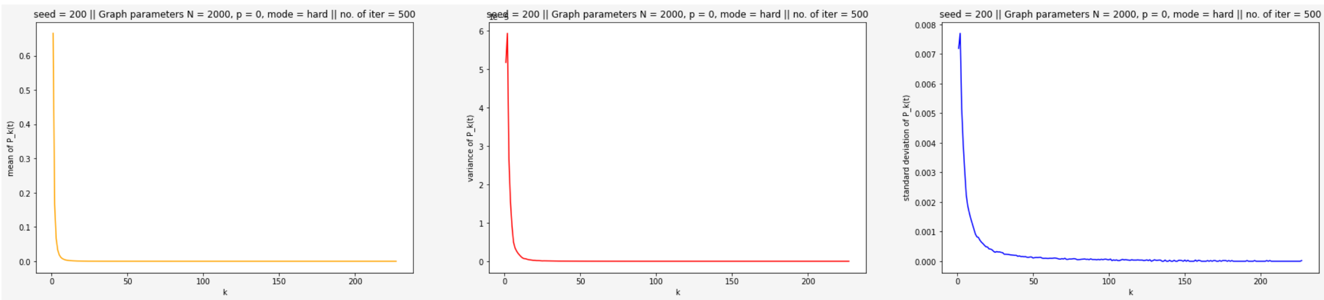
Consider the 'hard-code' taboo-ing scheme. In the following plots, we show the mean, variance and standard deviation respectively of $P_k(2000)$ with $p \in \{0.01, 0.025, 0.05, 0.1, 0.25, 0.5\}$ as function of k.

Please note that, in the following plots, we used the following colour code

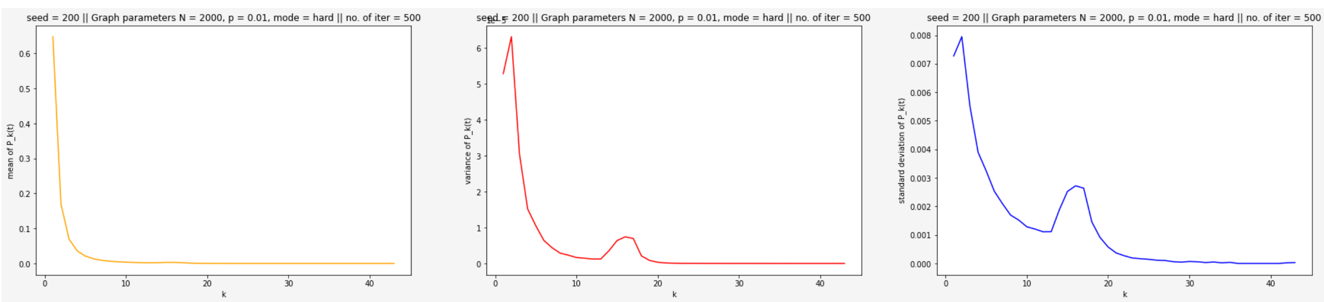
- Yellow: Mean of $P_k(2000)$
- Red: Variance of $P_k(2000)$
- Blue: Standard deviation of $P_k(2000)$

First, for reference, let us have a look at the degree sequence $P_k(2000)$ of a *preferential attachment model.*

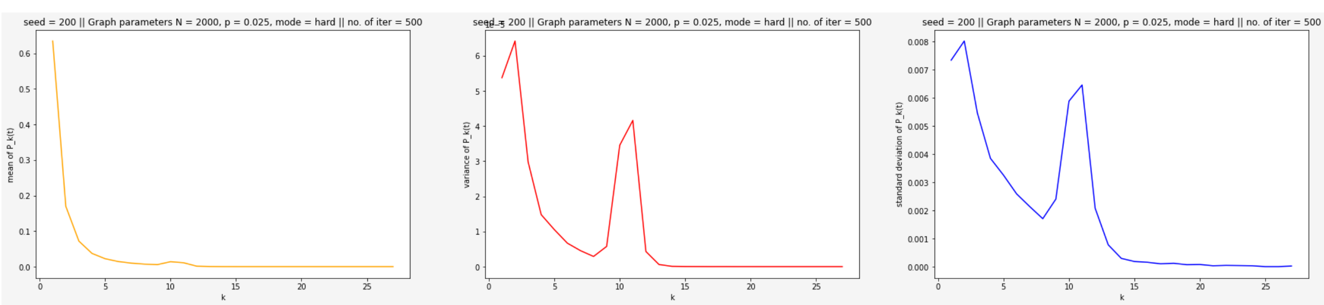
Figure 1: Mean, variance and standard deviation of $P_k(2000)$ as function of k in a *preferential attachment model*



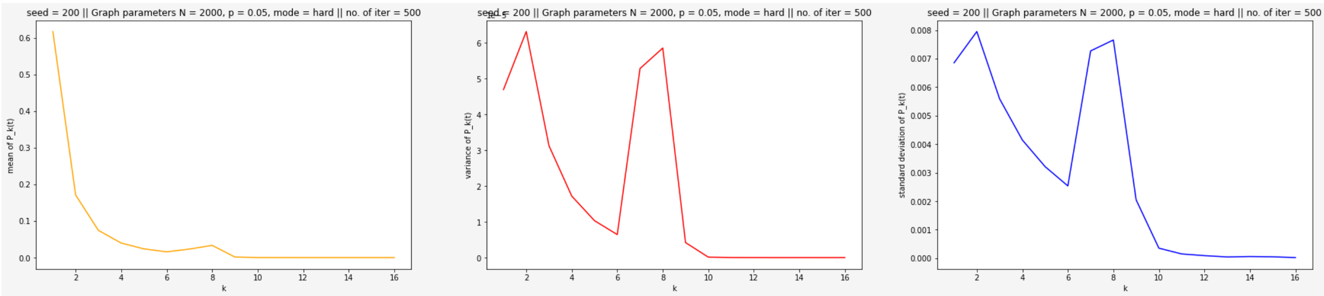
1.1 p = 0.01



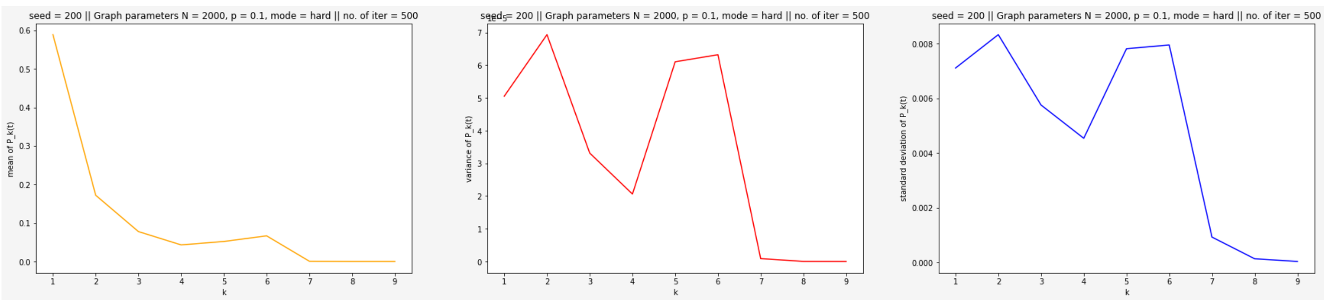
1.2 p = 0.025



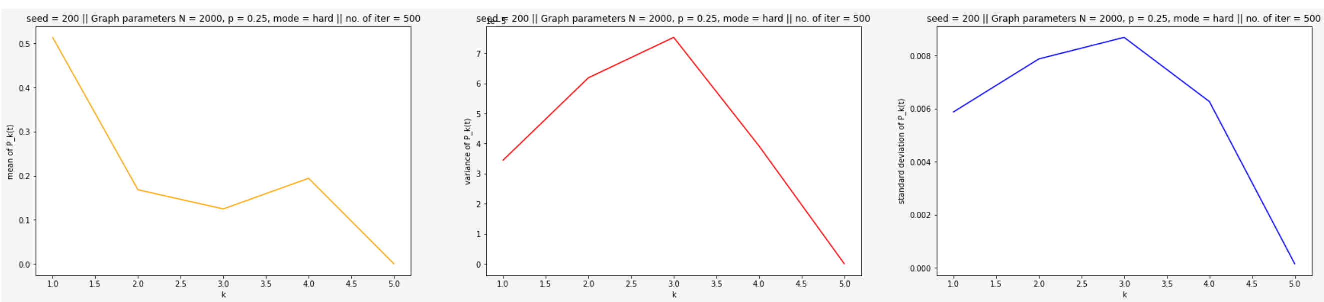
1.3 p = 0.05



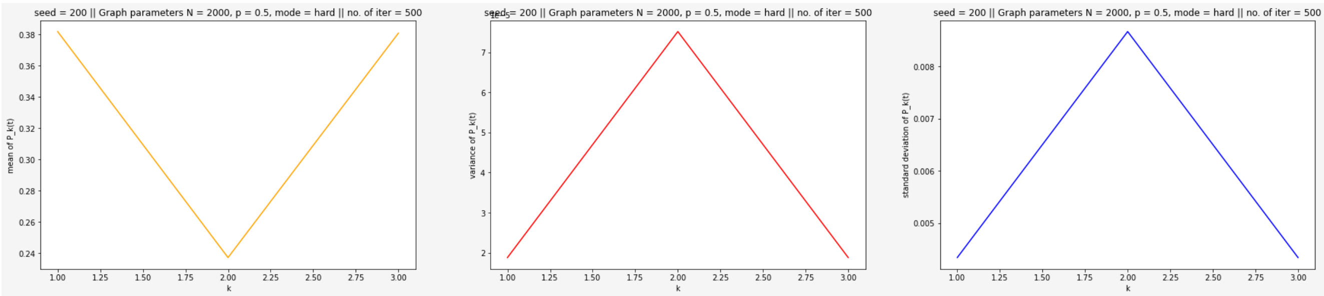
1.4 p = 0.1



1.5 p = 0.25



1.6 p = 0.5

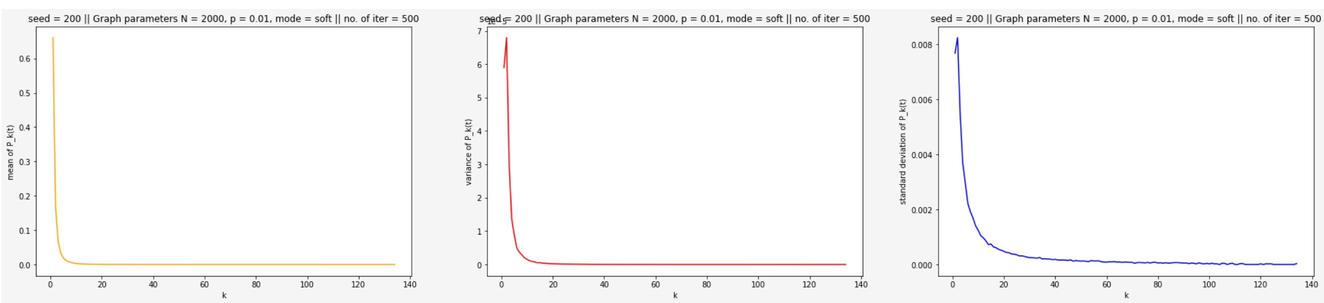


We observe that, for hard-coding scheme, we know that the maximal degree is bounded above by $\left\lceil \frac{1}{p} \right\rceil$. Also, unlike the preferential attachment model, $E(P_k(t))$ and $V(P_k(t))$ are not strictly decreasing functions of k. Instead, there seems to be a bump in the plot of expectation and variation.

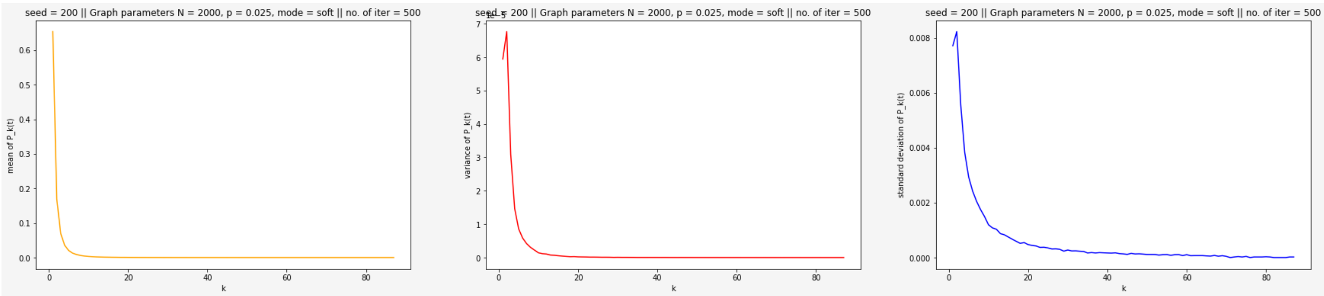
2 Soft-code

Now we adopt the 'soft-code' taboo-ing scheme. The parameter p is varied as before.

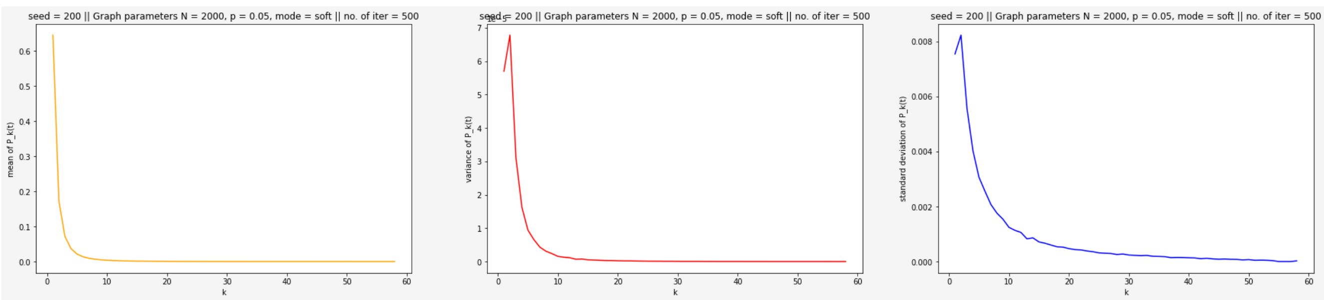
2.1 p=0.01



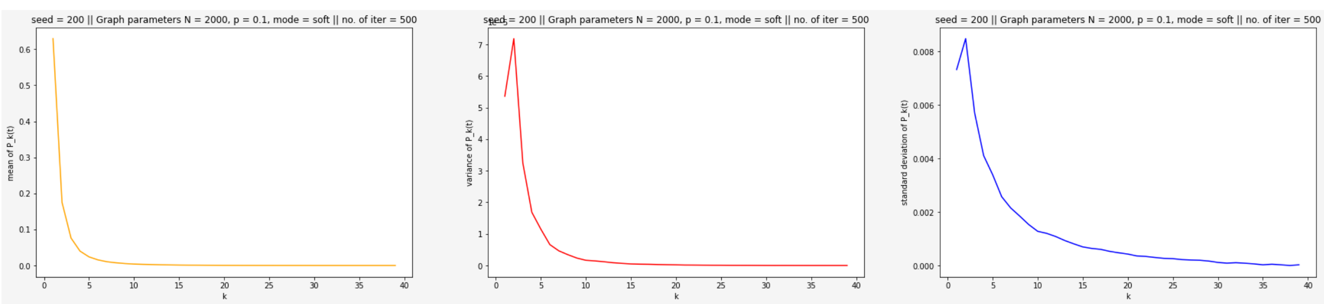
2.2 p=0.025



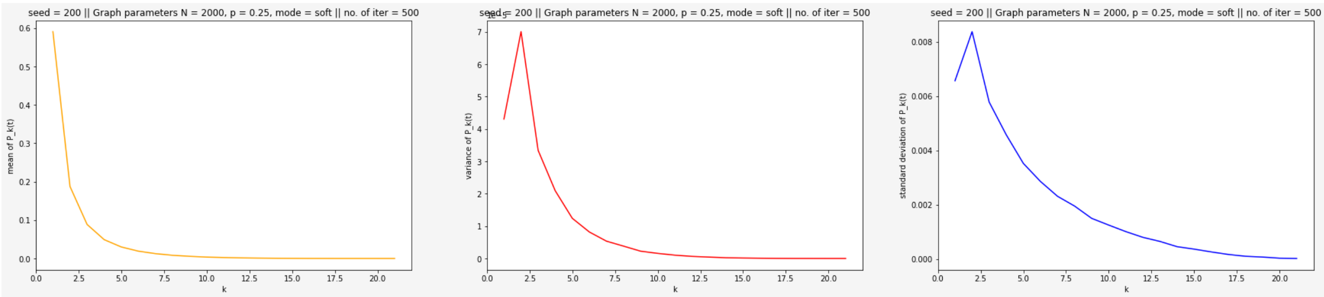
2.3 p=0.05



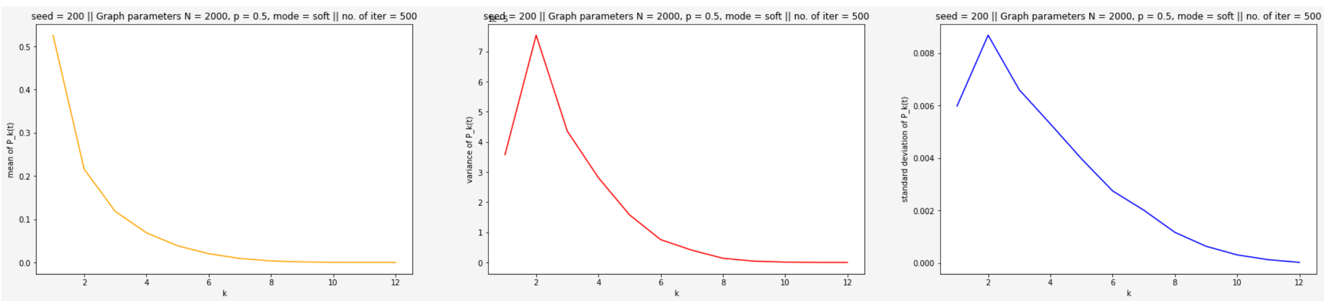
2.4 p=0.1



2.5 p=0.25



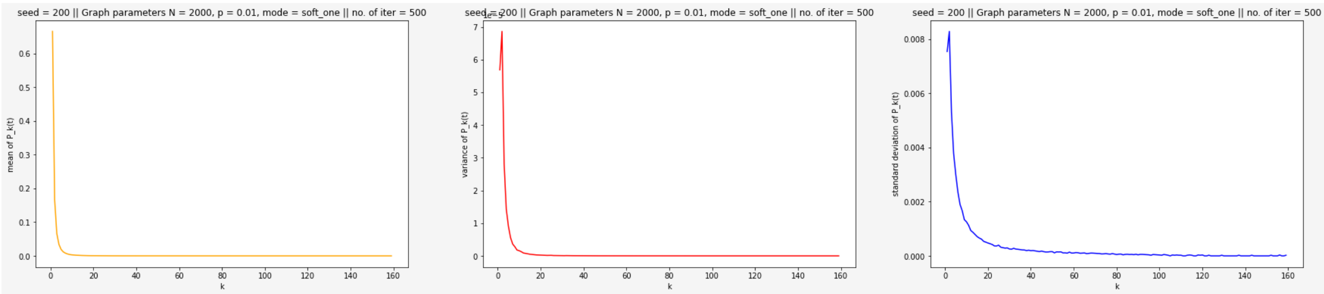
2.6 p=0.5



We observe that for this scheme of taboo-ing, the graphs looks very similar to the ones of *preferential attachment model*, mostly for small values of p. It is also important notice that the plots are truncated at some earlier value, i.e., the *preferential attachment model* has vertices of maximum degree around 200 while, the graphs obtained from soft-coding scheme has vertices of maximum degree strictly less than 200. We might imagine that the plots for *preferential attachment model* are truncated to obtain the plots for soft-coding scheme.

3 Soft-code (alternative)

For this last part, we adopt the 'soft-code (alternative)' taboo-ing scheme.This scheme has no parameter p.



As imagined, these plots also look very similar to those for *preferential attachment model*. But, these ones are also truncated at 160 for $t = 2000$.