Distribution of the Number of Triangles in an Erdös-Rényi RG

Muriel Pérez

November 3, 2016

output: md_document variant: markdown_github

Distribution of the Number of Triangles in an Erdös-Rényi RG

Here I investigate numericaly the limiting distribution of the number of triangles in a $ER(n, p = \lambda/n)$ graph. I want to confirm that as the number of nodes $n \to \infty$, the number of triangles $\Delta_n \to Pois(\lambda^3/6)$ in distribution.

First, I show the R package that I used, igraph.

igraph

There is a nice R package to work with graphs called igraph.

```
library(igraph)
```

igraph is a collection of network analysis tools. Among them, there are functions to easily sample $ER(n, p = \lambda/n)$ graphs, plot them, and count the number of triangles, which is what we need. Let's generate a graph with $n = 100$, $\lambda = 3$, and count the triangles with the function count_triangles, which returns how many triangles each vertex is part of.

```
set.seed(2016)
rg <- sample_gnp(n = 100, p = 3/100)
count_triangles(rg) %>%
    sum/3
```

```
## [1] 6
```

we can also plot rg (See Figure 1).

```
plot(rg, vertex.label = NA, vertex.size = 2)
```

Random sample from $ER(n, p = \lambda/n)$

Now, we want to confirm that as the number of nodes $n \to \infty$, the number of triangles $\Delta_n \to Poiss(\lambda^3/6)$ in distribution. For that, I generate a random sample of $N = 1000$ random graphs for $\lambda = 2$, and $n = 10, 20, 100, 200, 500, 1000$, and then analyze the empirical distribution of Δ_n in each case. Then I explote what occurs when lambda changes. First, I will try to use two graphical methods: plotting the probability mass functions and the qqplot. Then, I will use the χ^2 statistic using as a reference distribution $Poiss(\lambda^3/6)$ as evidence of convergence.

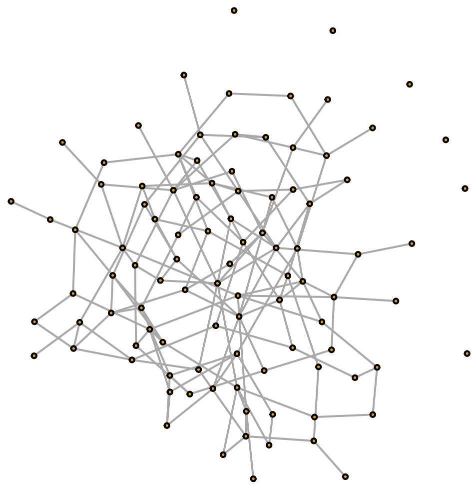

Figure 1: A realization of $ER(n = 100, p = 3/100)$

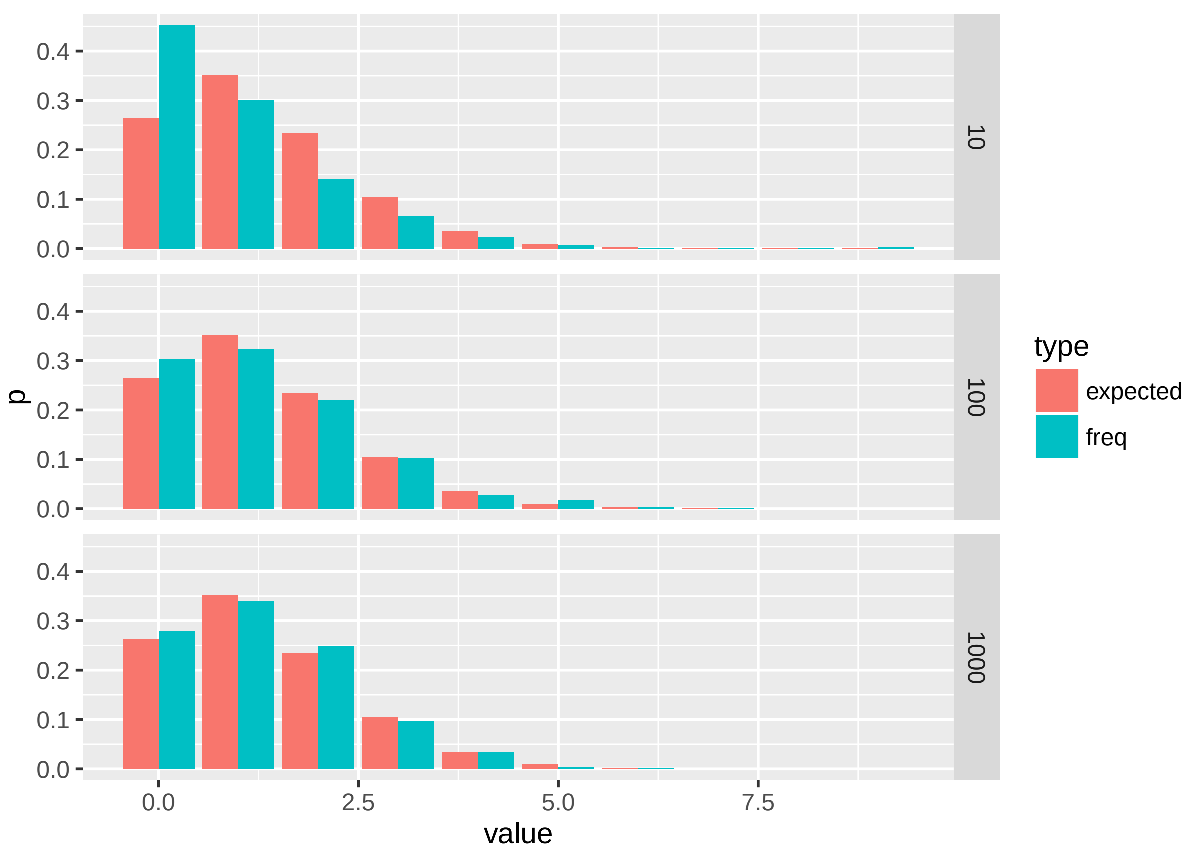

Figure 2: Observed frequency distribution and expected Poisson distribution

We can compare the empirical distribution with the theoretical probability mass function. For low numbers of n, it is obvious that the distributions do not coincide, while for larger numbers, they get closer.

In the same fashion, the qqplots show that for lower numbers of n, the distribution of Δ_n deviates from the Poisson distribution, but it gets closer as n grows larger.

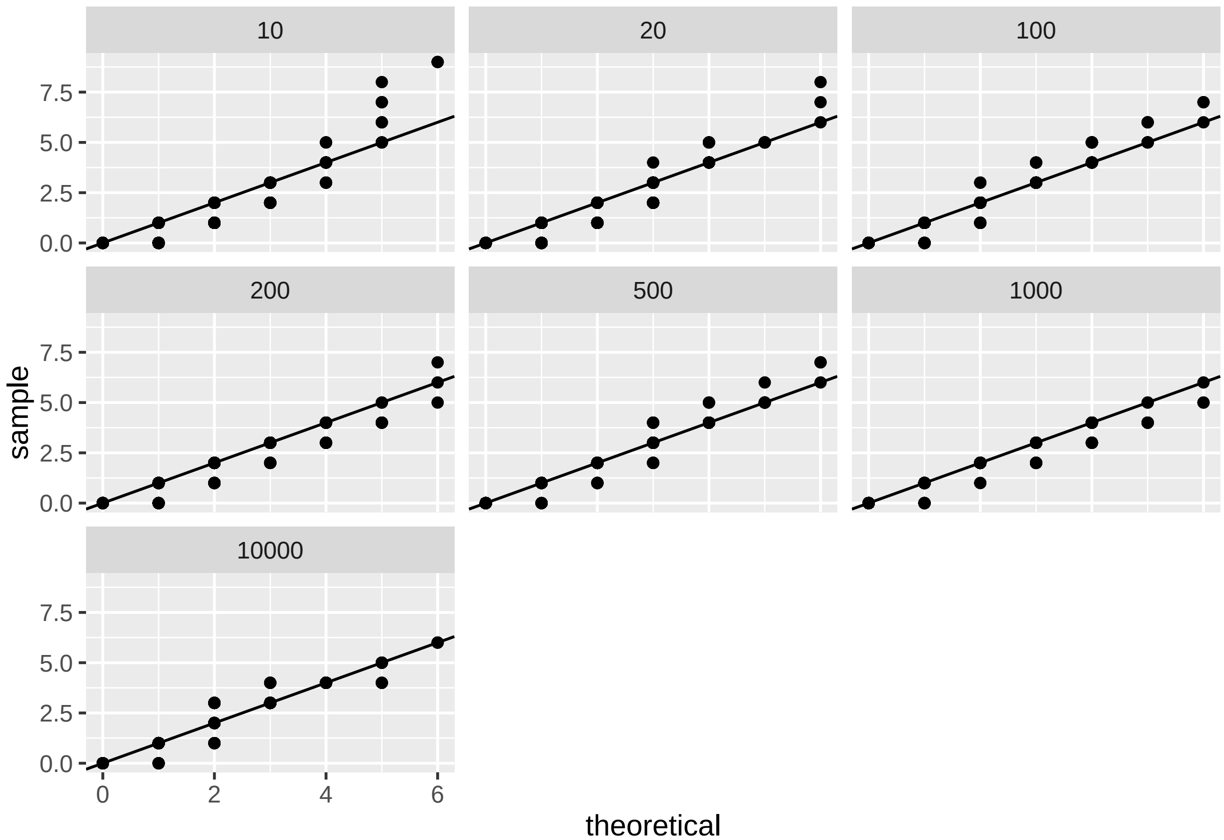

Figure 3: QQ-plots of the sample versus the Poisson distribution quantiles

Finally, we compute the χ^2 for each n, and plot it

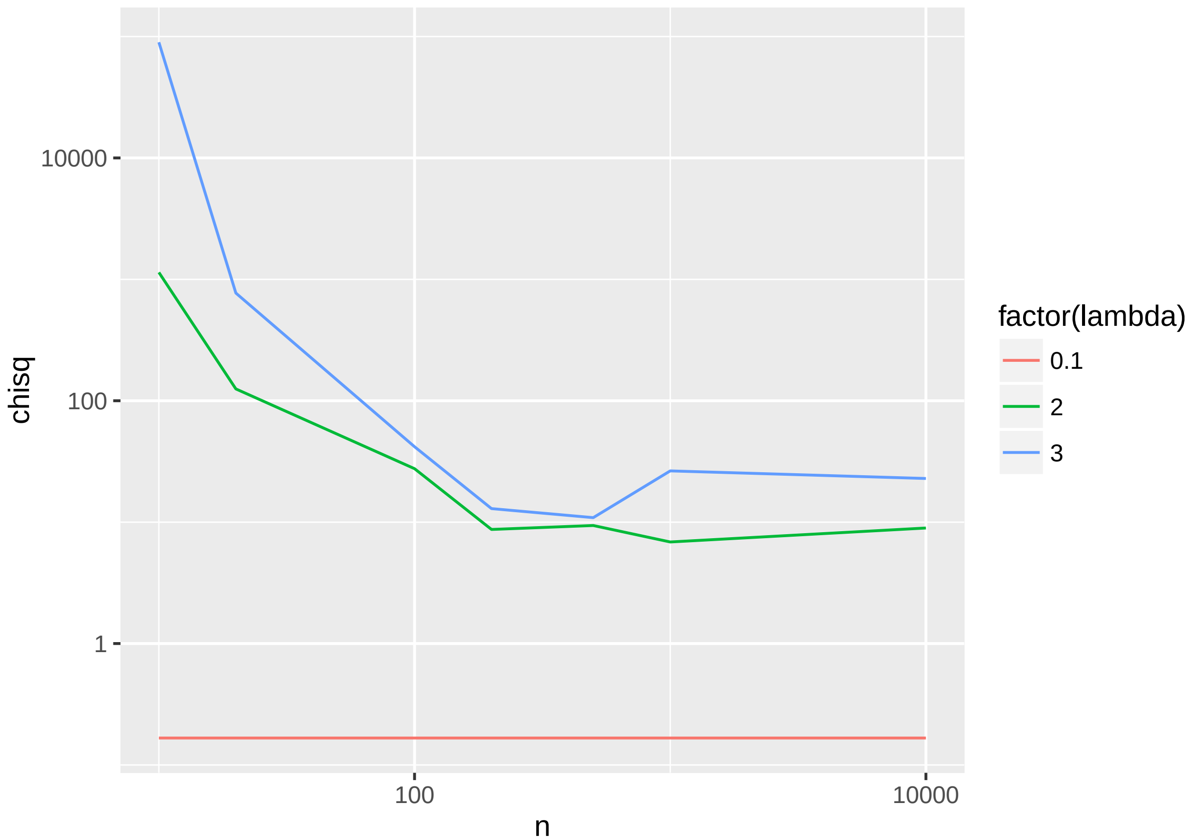

Figure 4: chi^2 statistic as a function of n.