Ramsey Objects and Delaporte

Aaron Robertson
Colgate University

with Maria Dascalu (undergraduate at Colgate)
and
Will Cipolli (statistician at Colgate)

July 2018

Notation

Notation

The *complete graph* on k vertices will be denoted by K_k.

Notation

The *complete graph* on k vertices will be denoted by K_k.

Ramsey's Theorem: The minimal integer n such that any 2-coloring of the edges of K_n admits a monochromatic K_k subgraph will be denoted $R(k)$.

Notation

The *complete graph* on k vertices will be denoted by K_k.

Ramsey's Theorem: The minimal integer n such that any 2-coloring of the edges of K_n admits a monochromatic K_k subgraph will be denoted $R(k)$.

So, $R(3) = 6$; $R(4) = 18$; $43 \leq R(5) \leq 48$, etc.

Also,

$$R(k) > \frac{k\sqrt{2}}{e} \cdot 2^{k/2}$$

is the best-known general lower bound.

A Different Approach

Since these Ramsey numbers are notoriously difficult to compute, perhaps instead of requiring all 2-colored complete graphs to contain a monochromatic subgraphs, we just require almost all.

A Different Approach

Since these Ramsey numbers are notoriously difficult to compute, perhaps instead of requiring all 2-colored complete graphs to contain a monochromatic subgraphs, we just require almost all.

Define the random variable over the equally likely sample space of all 2-color complete graphs on n vertices

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs.}$$

A Different Approach

Since these Ramsey numbers are notoriously difficult to compute, perhaps instead of requiring all 2-colored complete graphs to contain a monochromatic subgraphs, we just require almost all.

Define the random variable over the equally likely sample space of all 2-color complete graphs on n vertices

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs.}$$

Then we can define the "almost-all" number, for $0 < \alpha < 1$, as

$$\widehat{R}_k(\alpha) = \min_{n \in \mathbb{Z}^+} \{n : \mathbb{P}(X_k(n) > 0) > \alpha\}.$$

A Different Approach

Since these Ramsey numbers are notoriously difficult to compute, perhaps instead of requiring all 2-colored complete graphs to contain a monochromatic subgraphs, we just require almost all.

Define the random variable over the equally likely sample space of all 2-color complete graphs on n vertices

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs.}$$

Then we can define the "almost-all" number, for $0 < \alpha < 1$, as

$$\widehat{R}_k(\alpha) = \min_{n \in \mathbb{Z}^+} \{n : \mathbb{P}(X_k(n) > 0) > \alpha\}.$$

In order to determine these numbers, we need to know (if possible) the distribution of $X_k(n)$.

A Different Approach

Godbole, Skipper, and Sunley investigated the distribution of $X_k(n)$.

In particular, they proved – **for large k** (with conditions on n) – that

$$\mathbb{P}(X_k(n) = j) \approx \frac{\lambda^j e^{-\lambda}}{j!} \qquad \text{where} \qquad \lambda = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}.$$

A Different Approach

Godbole, Skipper, and Sunley investigated the distribution of $X_k(n)$.

In particular, they proved – **for large k** (with conditions on n) – that

$$\mathbb{P}(X_k(n) = j) \approx \frac{\lambda^j e^{-\lambda}}{j!} \qquad \text{where} \qquad \lambda = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}.$$

In other words, they showed that

$$X_k(n) \sim \text{Poisson}\left(\frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}\right).$$

A Different Approach

Godbole, Skipper, and Sunley investigated the distribution of $X_k(n)$.

In particular, they proved – **for large k** (with conditions on n) – that

$$\mathbb{P}(X_k(n) = j) \approx \frac{\lambda^j e^{-\lambda}}{j!} \qquad \text{where} \qquad \lambda = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}.$$

In other words, they showed that

$$X_k(n) \sim \text{Poisson}\left(\frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}\right).$$

This is an **asymptotic** result.

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

What is the motivation behind this?

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

What is the motivation behind this?

Let $\{Y_i\}_{i=1}^{n}$ be a sequence of indicator random variables each with the same small probability of being 1.

Let $Y = \displaystyle\sum_{i=1}^{n} Y_i$ with n large.

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

What is the motivation behind this?

Let $\{Y_i\}_{i=1}^{n}$ be a sequence of indicator random variables each with the same small probability of being 1.

Let $Y = \displaystyle\sum_{i=1}^{n} Y_i$ with n large.

If the Y_i are independent, then Y follows a Binomial distribution.

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

What is the motivation behind this?

Let $\{Y_i\}_{i=1}^{n}$ be a sequence of indicator random variables each with the same small probability of being 1.

Let $Y = \displaystyle\sum_{i=1}^{n} Y_i$ with n large.

If the Y_i are independent, then Y follows a Binomial distribution.

Since $P(Y_i = 1) = p$ for all i with p small, we know that the Poisson distribution well-approximates the Binomial (even for relatively small n).

The Poisson Paradigm

The Poisson Paradigm is that Ramsey object counts tend to be (asymptotically) well-approximated by a Poisson distribution.

What is the motivation behind this?

Let $\{Y_i\}_{i=1}^{n}$ be a sequence of indicator random variables each with the same small probability of being 1.

Let $Y = \displaystyle\sum_{i=1}^{n} Y_i$ with n large.

If the Y_i are independent, then Y follows a Binomial distribution.

Since $P(Y_i = 1) = p$ for all i with p small, we know that the Poisson distribution well-approximates the Binomial (even for relatively small n).

Why does the Poisson distribution make sense for $X_k(n)$?

The Poisson Paradigm for $X_k(n)$

Consider a randomly edge-colored K_n and define the indicator random variables Y_i with $Y_i = 1$ precisely when the i^{th} K_k is monochromatic, for $i = 1, 2, \ldots, \binom{n}{k}$.

Then

$$X_k(n) = \sum_{i=1}^{\binom{n}{k}} Y_i.$$

The Poisson Paradigm for $X_k(n)$

Consider a randomly edge-colored K_n and define the indicator random variables Y_i with $Y_i = 1$ precisely when the i^{th} K_k is monochromatic, for $i = 1, 2, \ldots, \binom{n}{k}$.

Then

$$X_k(n) = \sum_{i=1}^{\binom{n}{k}} Y_i.$$

We know that $Y_i \not\perp Y_j$ if the i^{th} and j^{th} K_ks share at least one edge (or, equivalently, at least two vertices) so the assumption that the Y_i are independent is **not** met.

The Poisson Paradigm for $X_k(n)$

Consider a randomly edge-colored K_n and define the indicator random variables Y_i with $Y_i = 1$ precisely when the i^{th} K_k is monochromatic, for $i = 1, 2, \ldots, \binom{n}{k}$.

Then

$$X_k(n) = \sum_{i=1}^{\binom{n}{k}} Y_i.$$

We know that $Y_i \not\perp Y_j$ if the i^{th} and j^{th} K_ks share at least one edge (or, equivalently, at least two vertices) so the assumption that the Y_i are independent is **<u>not</u>** met.

However, for $n \gg k$ this is unlikely:

$$P(Y_i \not\perp Y_j) \leq \frac{\binom{k}{2}\binom{n-k+2}{k-2}}{\binom{n}{k}} \underset{\substack{\uparrow \\ \text{Stirling}}}{\sim} \left(\frac{ek^2}{n}\right)^2 \to 0.$$

The Poisson Paradigm for $X_k(n)$

Hence, “almost all” of the Y_i are independent.

The Poisson Paradigm for $X_k(n)$

Hence, “almost all” of the Y_i are independent.

Similar analyses work for other Ramsey object, e.g., arithmetic progressions, many systems of equations, general graphs (vs. complete graphs), etc. Hence, the paradigm.

The Poisson Paradigm for $X_k(n)$

Hence, “almost all” of the Y_i are independent.

Similar analyses work for other Ramsey object, e.g., arithmetic progressions, many systems of equations, general graphs (vs. complete graphs), etc. Hence, the paradigm.

Our focus here is on the classical numbers $R(k)$ and the distribution of the respective random variable

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs.}$$

The Poisson Paradigm for $X_k(n)$

Hence, “almost all” of the Y_i are independent.

Similar analyses work for other Ramsey object, e.g., arithmetic progressions, many systems of equations, general graphs (vs. complete graphs), etc. Hence, the paradigm.

Our focus here is on the classical numbers $R(k)$ and the distribution of the respective random variable

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs.}$$

Because of this paradigm, we should investigate the Poisson distribution first.

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n, p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n, p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

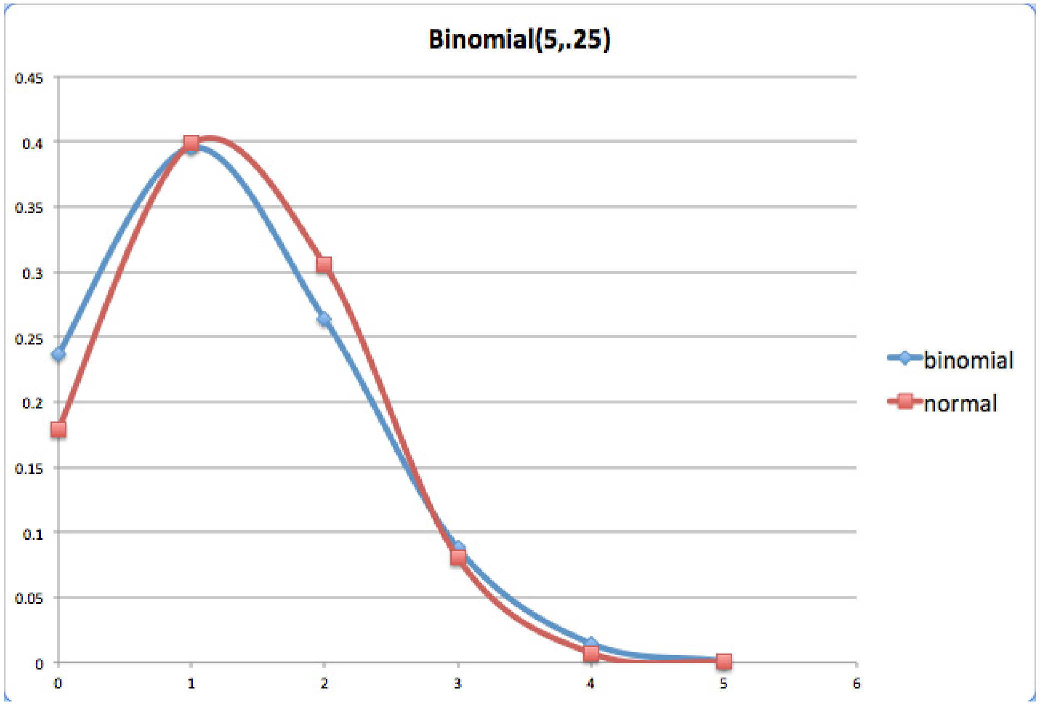
Let's take $p = .25$, so that, based on our rule of thumb, we shouldn't use the Normal approximation unless $n \geq 120$.

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n,p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

Let's take $p = .25$, so that, based on our rule of thumb, we shouldn't use the Normal approximation unless $n \geq 120$.



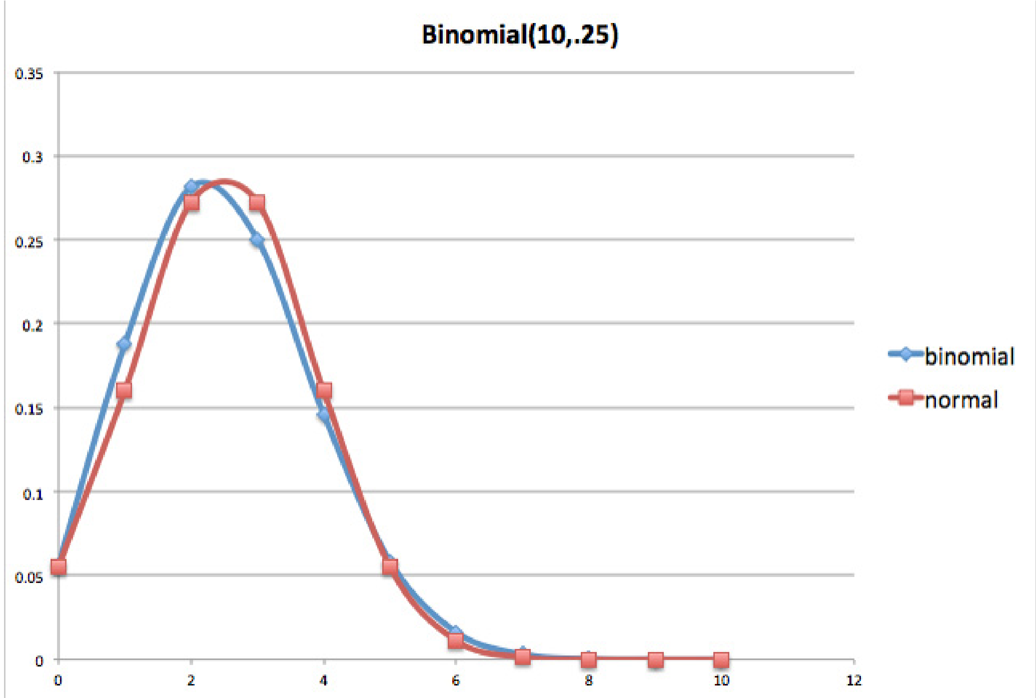
$n = 5$

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n, p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

Let's take $p = .25$, so that, based on our rule of thumb, we shouldn't use the Normal approximation unless $n \geq 120$.



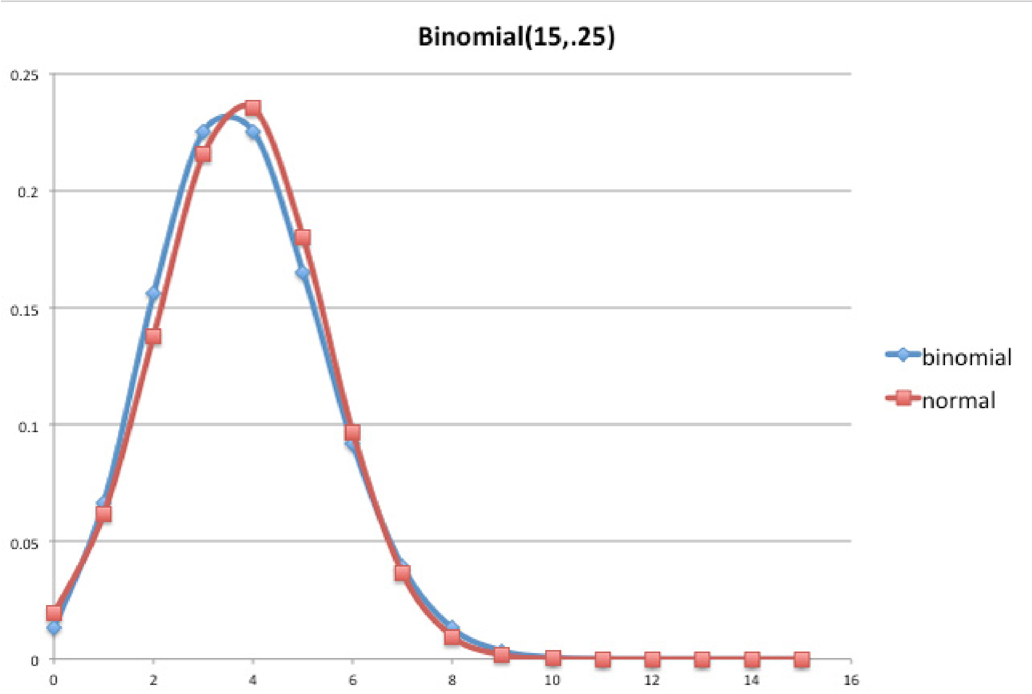
$n = 10$

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n,p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

Let's take $p = .25$, so that, based on our rule of thumb, we shouldn't use the Normal approximation unless $n \geq 120$.



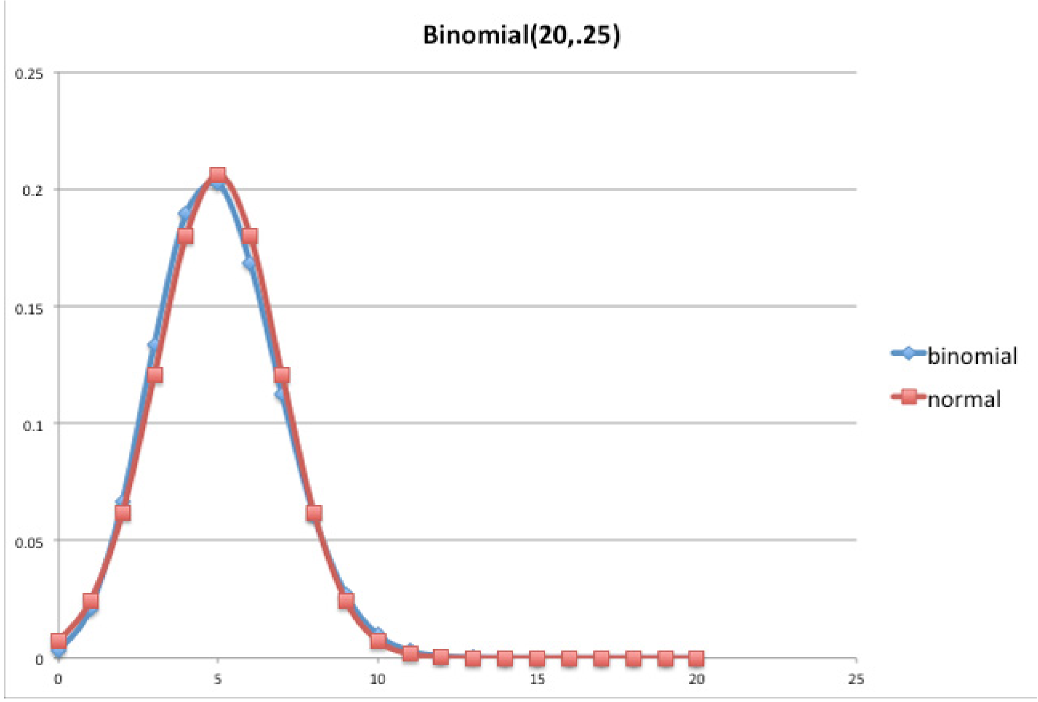
$n = 15$

$X_k(n) \sim$ Poisson?

We know that some asymptotic results are actually very good for small values.

For example, we know that Binomial$(n, p) \sim$ Normal$(np, np(1-p))$ for large n (typical requirement: $\min(np, n(1-p)) \geq 30$).

Let's take $p = .25$, so that, based on our rule of thumb, we shouldn't use the Normal approximation unless $n \geq 120$.



$n = 20$

$X_k(n) \sim$ Poisson?

Back to our question: Is

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs}$$

approximately Poisson?

$X_k(n) \sim$ Poisson?

Back to our question: Is

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs}$$

approximately Poisson?

First, we need something to compare the Poisson distribution against.

$X_k(n) \sim$ Poisson?

Back to our question: Is

$$X_k(n) = \text{number of monochromatic } K_k \text{ subgraphs}$$

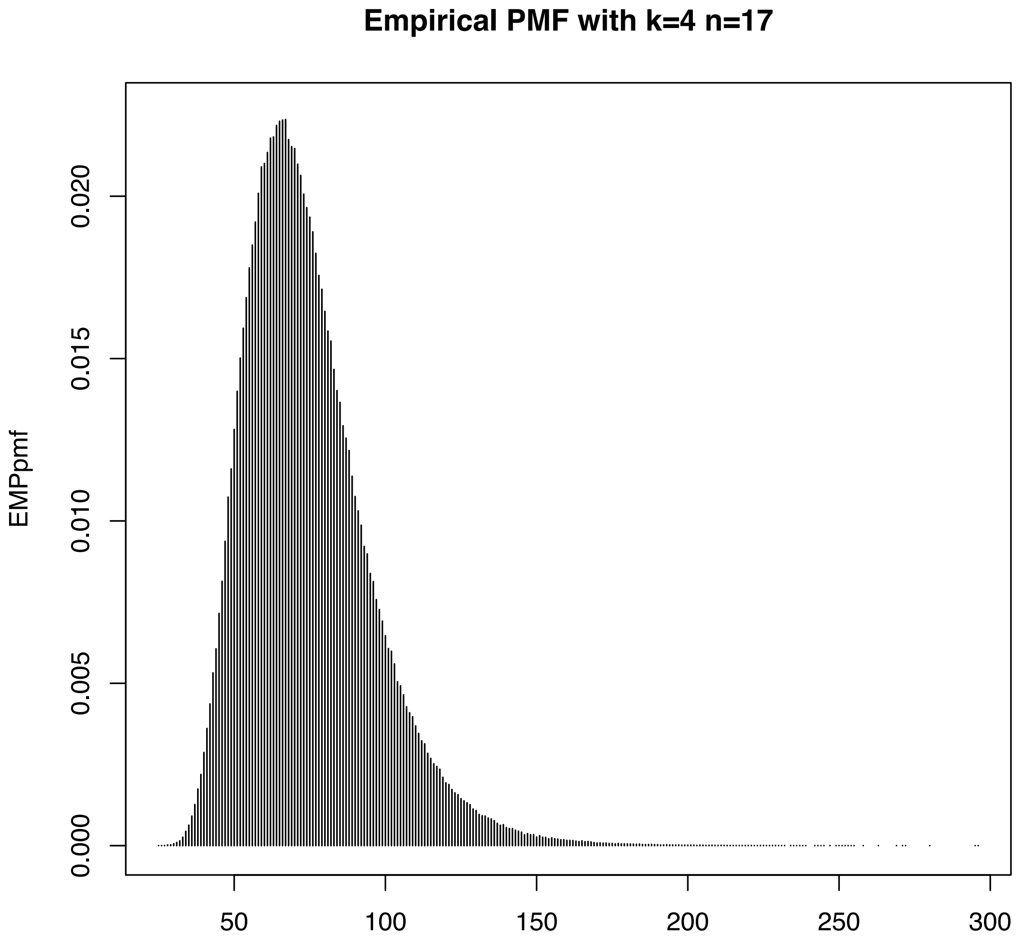
approximately Poisson?

First, we need something to compare the Poisson distribution against.

Maria (the undergrad) created an efficient code to generate random 2-colorings on a given number of vertices and count the resulting number of monochromatic complete graphs on k vertices for given k.

Empirical Histograms

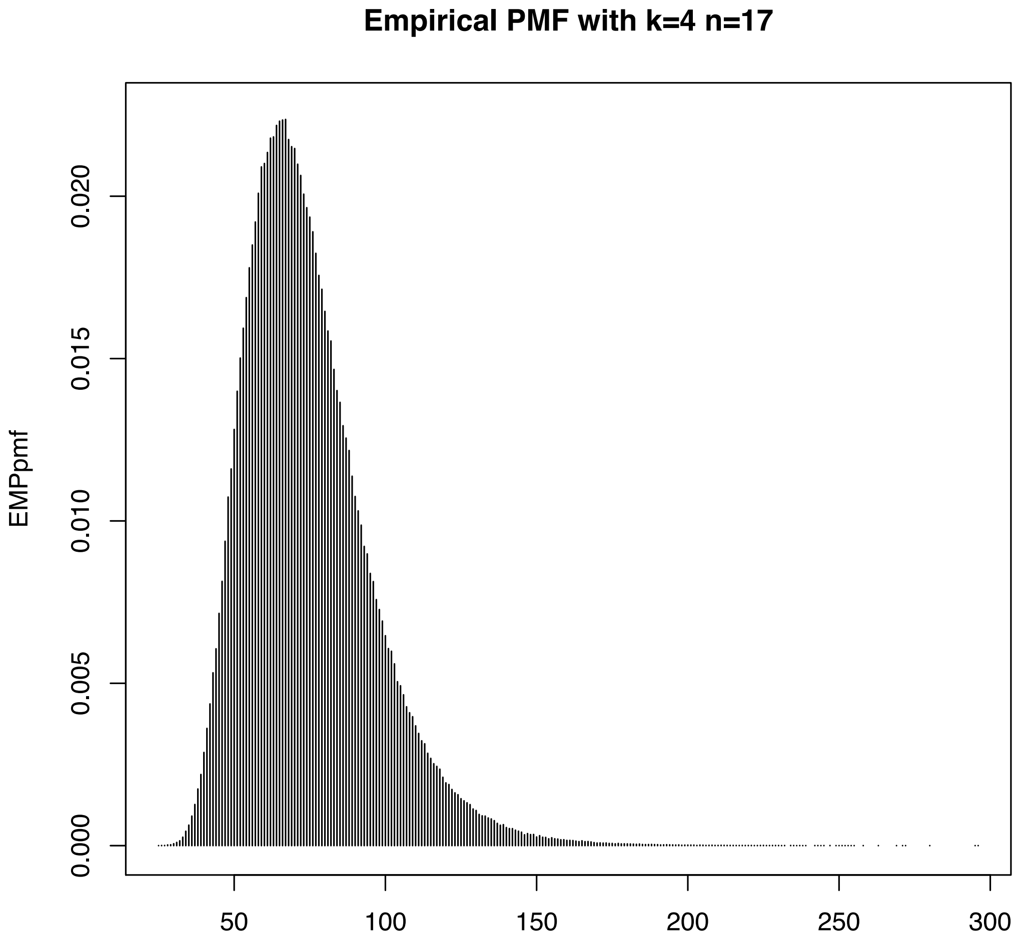
Here is an empirical histogram for $k = 4$ with $n = 17$ based on a million random graphs (i.e., an empirical pmf of $X_4(17)$):



Empirical Histograms

Here is an empirical histogram for $k = 4$ with $n = 17$ based on a million random graphs (i.e., an empirical pmf of $X_4(17)$):

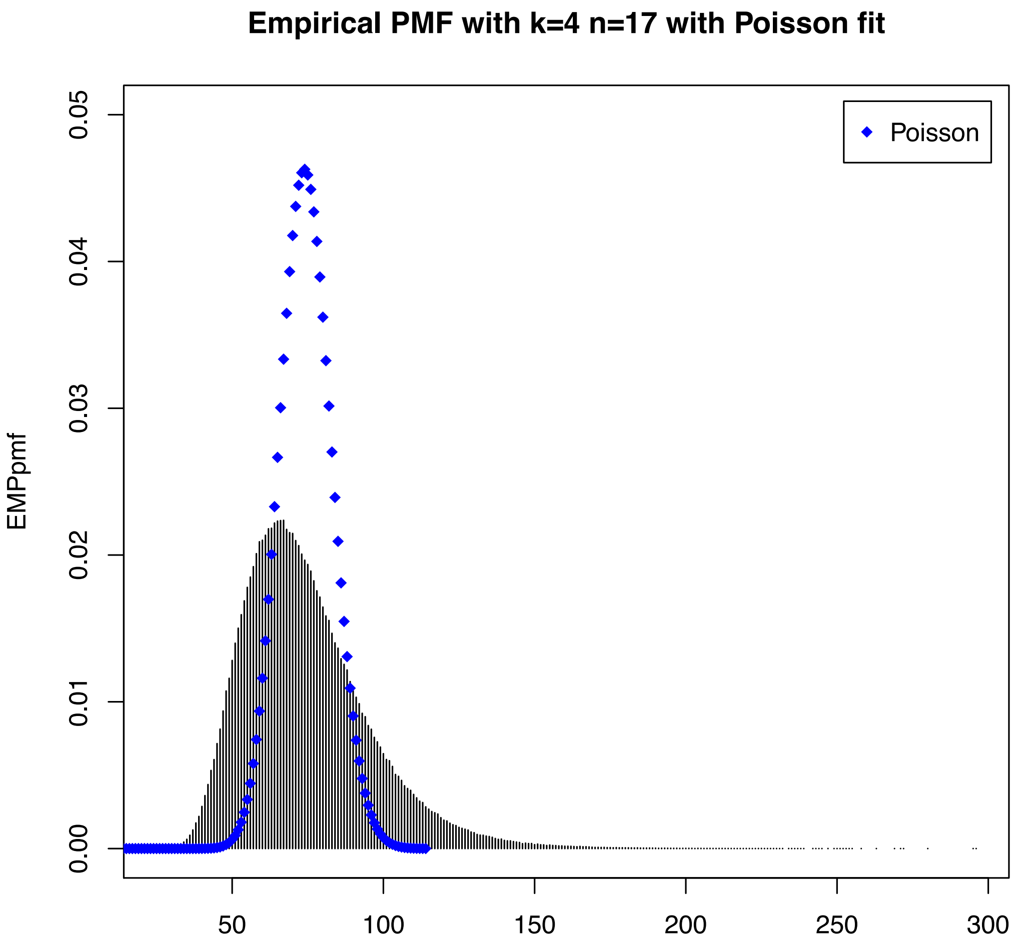
Is the Poisson close?



Empirical Histograms ∼ Poisson?

NO

(and this is the BEST-fitting Poisson based on the MLE for the parameter)

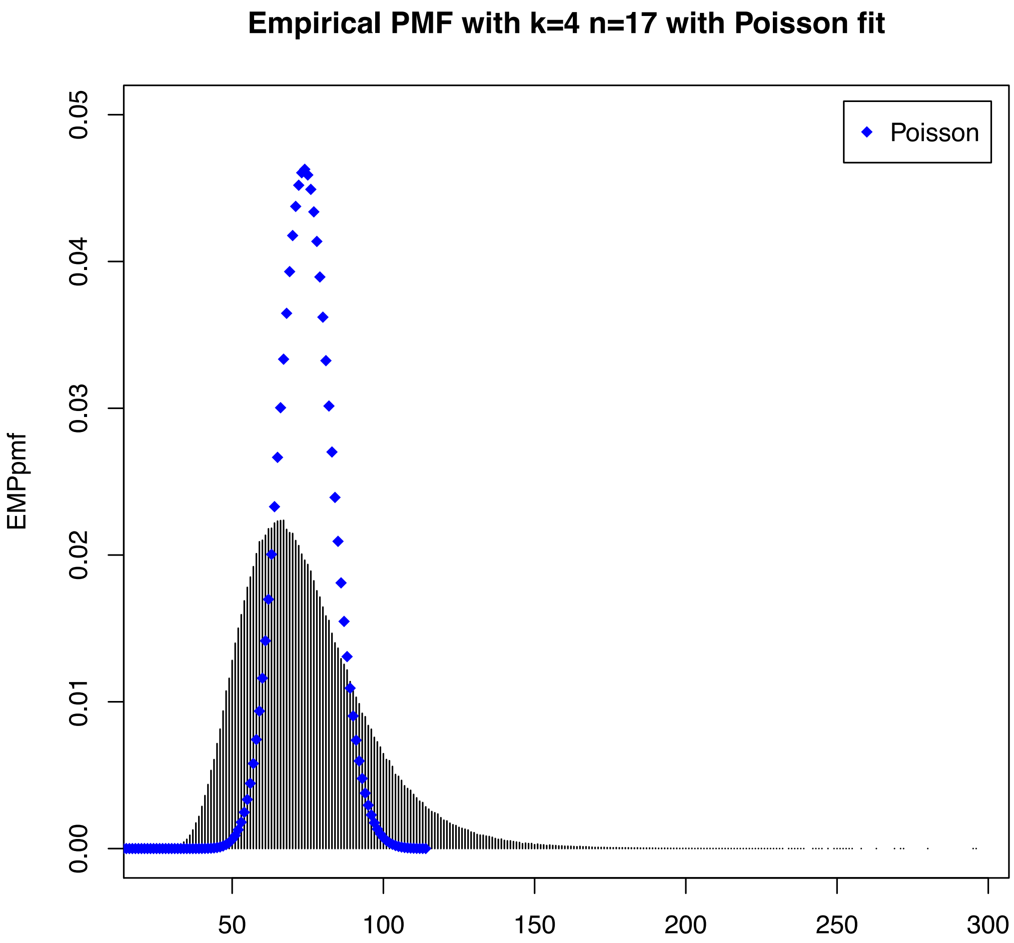


Empirical Histograms $\sim$ Poisson?

NO

(and this is the BEST-fitting Poisson based on the MLE for the parameter)

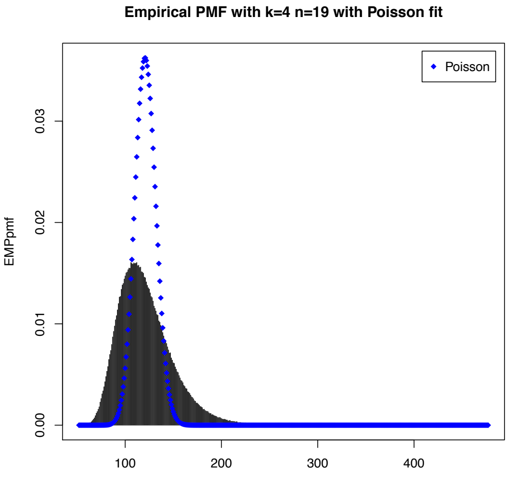
Could this be a fluke?

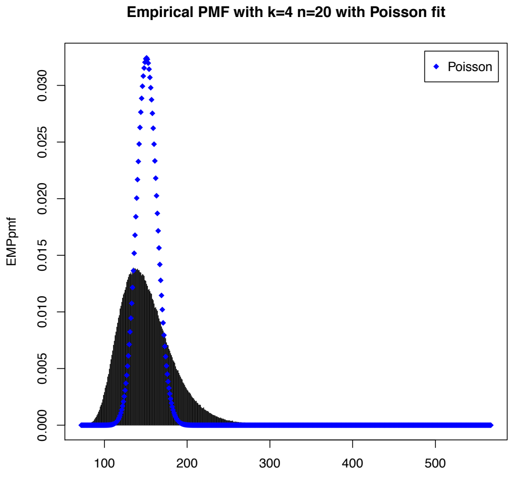


Empirical Histograms $\sim$ Poisson?

NOT a fluke:

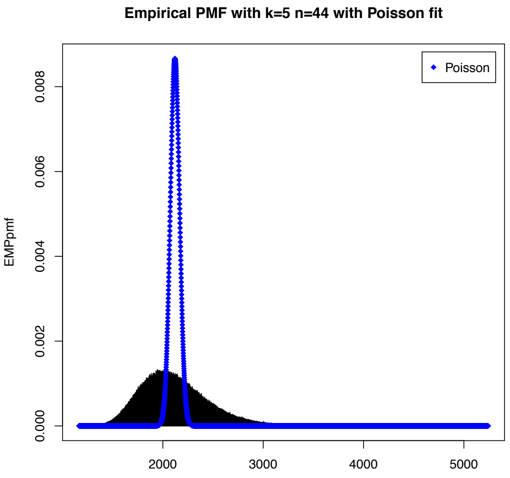
$k = 4, n = 19$

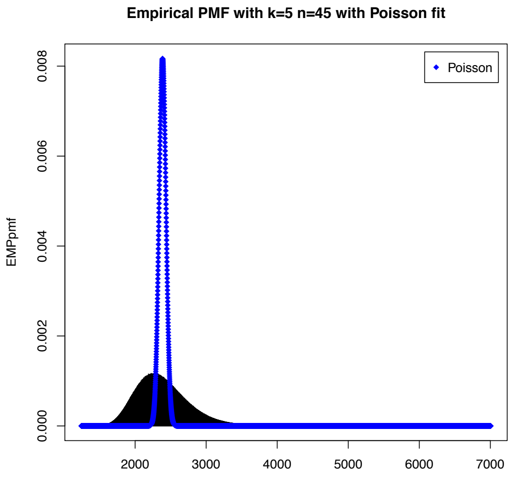

Empirical PMF with k=4 n=19 with Poisson fit


Empirical PMF with k=4 n=20 with Poisson fit

$k = 4, n = 20$

$k = 5, n = 44$


Empirical PMF with k=5 n=44 with Poisson fit


Empirical PMF with k=5 n=45 with Poisson fit

$k = 5, n = 45$

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

So, we turn to compound distributions by letting the Poisson parameter be a random variable itself.

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

So, we turn to compound distributions by letting the Poisson parameter be a random variable itself.

Consider $T \sim \mathrm{Poisson}(\lambda)$ where λ is a random variable.

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

So, we turn to compound distributions by letting the Poisson parameter be a random variable itself.

Consider $T \sim \mathrm{Poisson}(\lambda)$ where λ is a random variable.

A common choice is for $\lambda \sim \mathrm{Gamma}$ to correct for the overdispersion of the empirical pmf (Poisson has mean equal to variance and this is too restrictive).

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

So, we turn to compound distributions by letting the Poisson parameter be a random variable itself.

Consider $T \sim \text{Poisson}(\lambda)$ where λ is a random variable.

A common choice is for $\lambda \sim \text{Gamma}$ to correct for the overdispersion of the empirical pmf (Poisson has mean equal to variance and this is too restrictive).

In this situation, the result is that $T \sim \text{Negative Binomial}$.

If Not Poisson, Then What?

We know that its distribution must be asymptotically Poisson.

So, we turn to compound distributions by letting the Poisson parameter be a random variable itself.

Consider $T \sim \text{Poisson}(\lambda)$ where λ is a random variable.

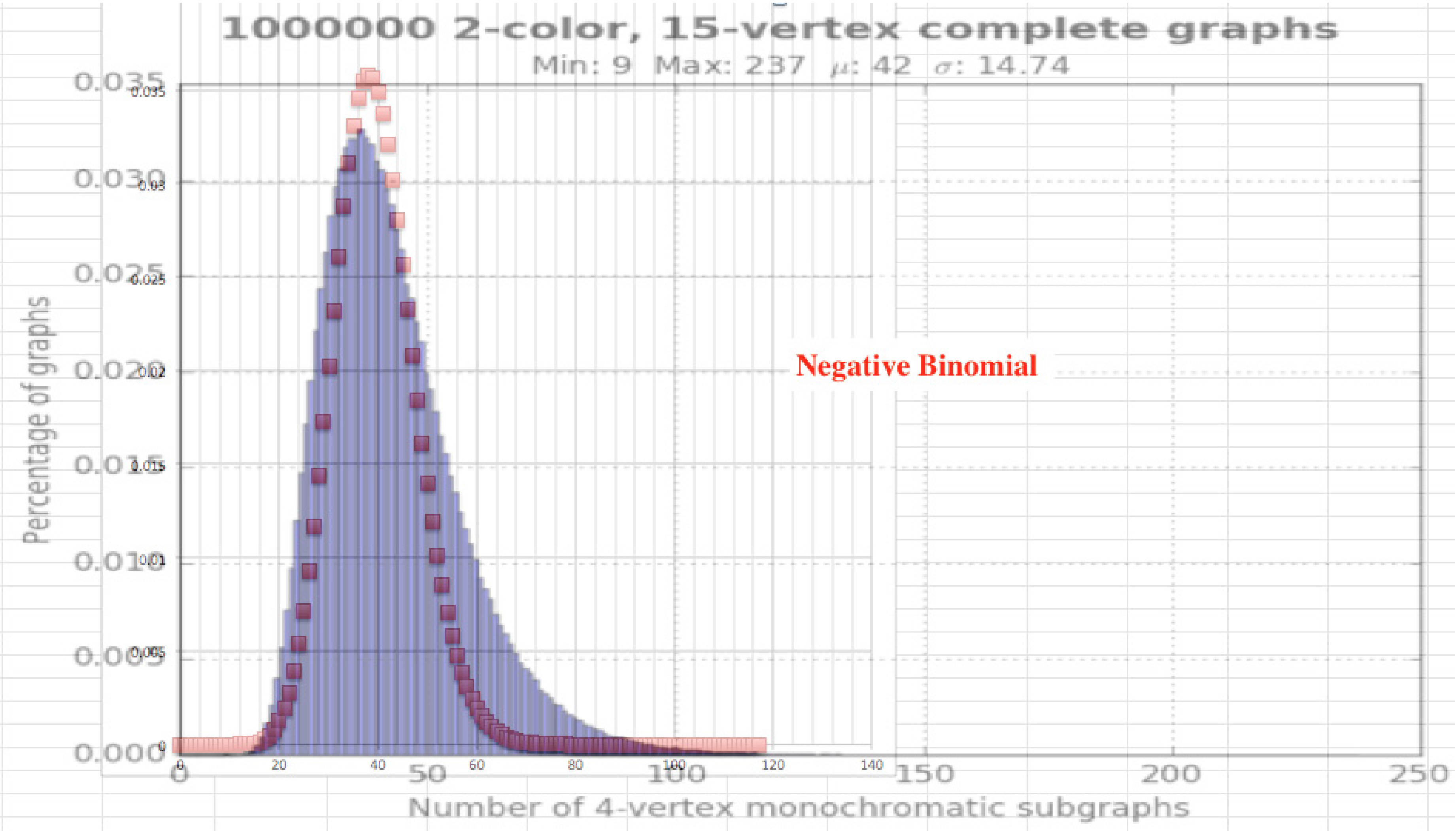
A common choice is for $\lambda \sim \text{Gamma}$ to correct for the overdispersion of the empirical pmf (Poisson has mean equal to variance and this is too restrictive).

In this situation, the result is that $T \sim \text{Negative Binomial}.$

How does this do?

If Not Poisson, Then What?

It’s closer, but not great:



If Not Negative Binomial, Then What?

Next, we try $\lambda \sim$ Poisson.

If Not Negative Binomial, Then What?

Next, we try $\lambda \sim$ Poisson. This leads to the Generalized Poisson distribution.

If Not Negative Binomial, Then What?

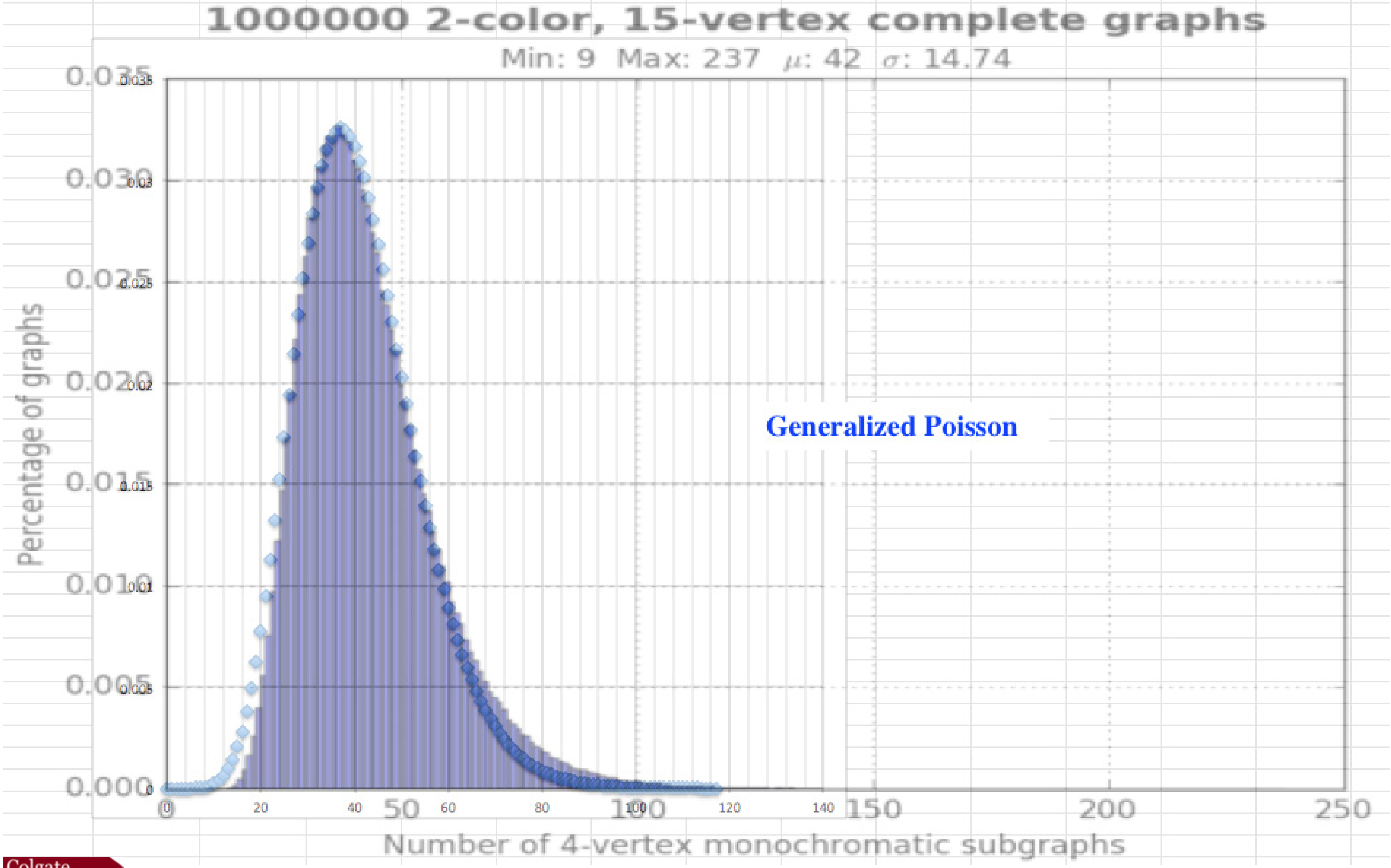
Next, we try $\lambda \sim$ Poisson. This leads to the Generalized Poisson distribution.

Do we have a winner?

If Not Negative Binomial, Then What?

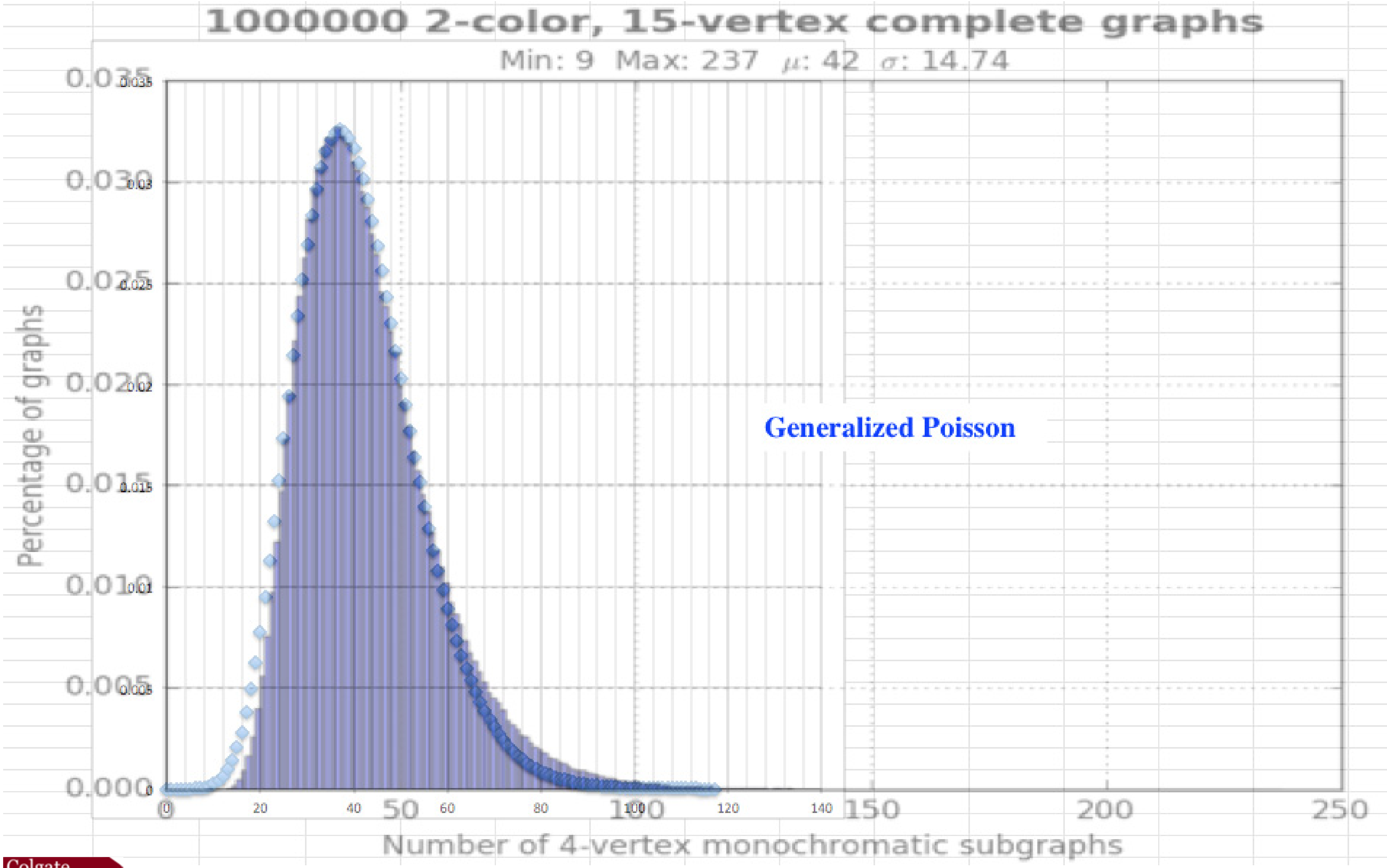
Next, we try $\lambda \sim$ Poisson. This leads to the Generalized Poisson distribution.

Do we have a winner?



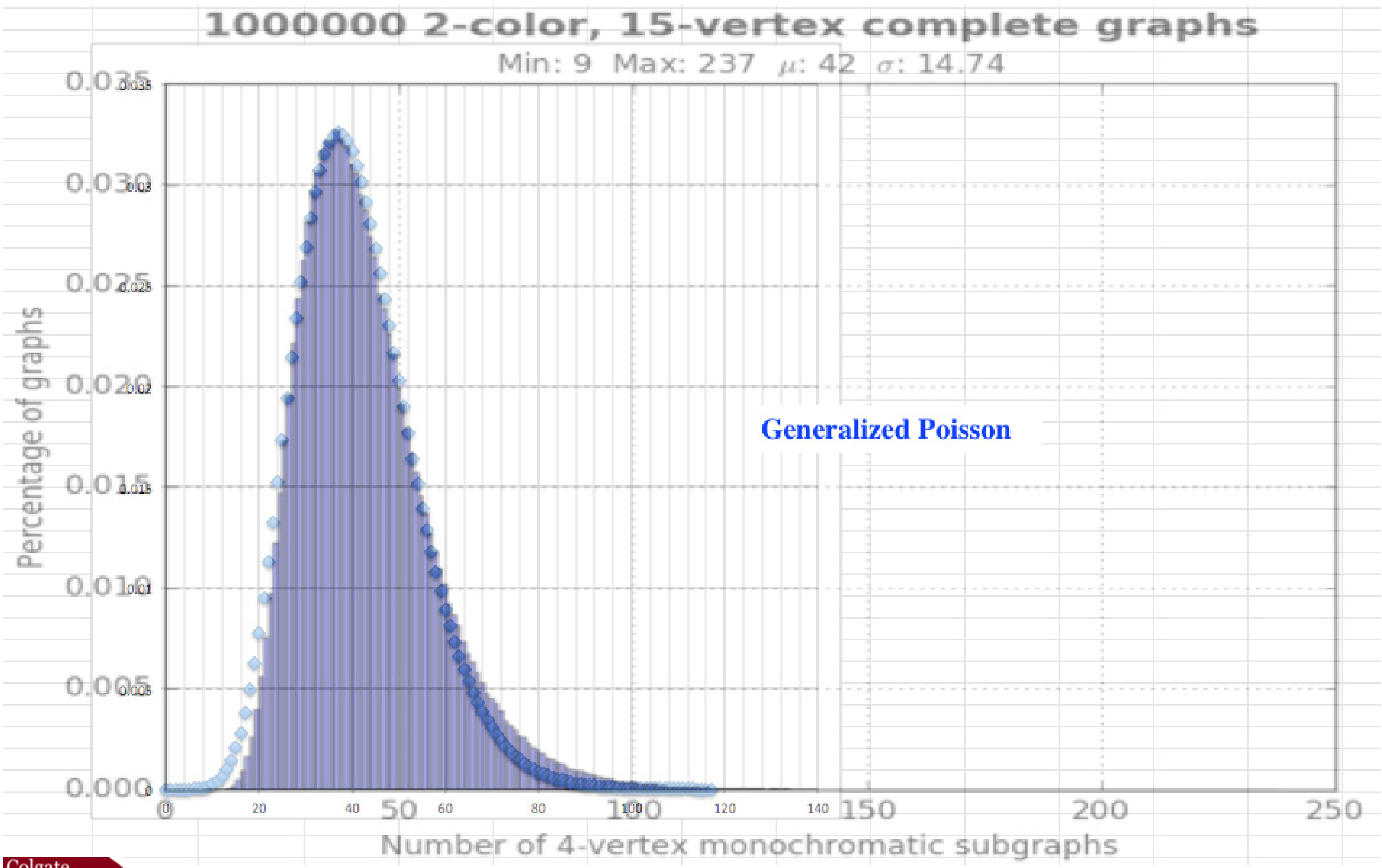
If Not Negative Binomial, Then What?

Looks pretty good, but



If Not Negative Binomial, Then What?

Looks pretty good, but **we're mathematicians not statisticians!**



What's left?

Let's think about the Poisson paradigm.

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

How do we account for the dependent parts?

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

How do we account for the dependent parts?

Let's ask the obvious question:

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

How do we account for the dependent parts?

Let's ask the obvious question: Can we take a cue from car insurance?

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

How do we account for the dependent parts?

Let's ask the obvious question: Can we take a cue from car insurance?

Consider a small company who wants the distribution of accidents of cars they insure. Most accidents will be with cars insured by a different company, but sometimes the accident will occur between cars insured by the same small company. This won't happen often, but has a non-zero probability. So, most, but not all, cars will have accidents independent of other cars from the same small company.

What's left?

Let's think about the Poisson paradigm.

Most – but not all – of the random variables indicating whether or not the i^{th} K_k is monochromatic are independent.

So "most" of the random variable should be Poisson.

How do we account for the dependent parts?

Let's ask the obvious question: Can we take a cue from car insurance?

Consider a small company who wants the distribution of accidents of cars they insure. Most accidents will be with cars insured by a different company, but sometimes the accident will occur between cars insured by the same small company. This won't happen often, but has a non-zero probability. So, most, but not all, cars will have accidents independent of other cars from the same small company.

How does this translate into a compound distribution?

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

We can think of Y as a contagion driver (if one K_k is monochromatic, then other K_ks that share vertices are more likely to be monochromatic; monochromaticity is contagious).

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

We can think of Y as a contagion driver (if one K_k is monochromatic, then other K_ks that share vertices are more likely to be monochromatic; monochromaticity is contagious).

Polyá's Urn model was devised to model disease contagion.

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

We can think of Y as a contagion driver (if one K_k is monochromatic, then other K_ks that share vertices are more likely to be monochromatic; monochromaticity is contagious).

Polyá's Urn model was devised to model disease contagion.

In *Limit Distributions for Large Polyá Urns* by Chauvin, Pouyanne, and Sahnoun (2011) we see that the associated martingale (a stochastic process where the expected value of the next random variable given the current state is equal to the value of the current random variable – requires large urns) follows a Gamma distribution.

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

We can think of Y as a contagion driver (if one K_k is monochromatic, then other K_ks that share vertices are more likely to be monochromatic; monochromaticity is contagious).

Polyá's Urn model was devised to model disease contagion.

In *Limit Distributions for Large Polyá Urns* by Chauvin, Pouyanne, and Sahnoun (2011) we see that the associated martingale (a stochastic process where the expected value of the next random variable given the current state is equal to the value of the current random variable – requires large urns) follows a Gamma distribution.

So, let's investigate $\lambda = c + G$, where c is constant and G is a Gamma random variable.

What's left?

We let $\lambda = c + Y$ where c is constant (accounting for the independence) and Y is a random variable (accounting for the dependence).

We can think of Y as a contagion driver (if one K_k is monochromatic, then other K_ks that share vertices are more likely to be monochromatic; monochromaticity is contagious).

Polyá's Urn model was devised to model disease contagion.

In *Limit Distributions for Large Polyá Urns* by Chauvin, Pouyanne, and Sahnoun (2011) we see that the associated martingale (a stochastic process where the expected value of the next random variable given the current state is equal to the value of the current random variable – requires large urns) follows a Gamma distribution.

So, let's investigate $\lambda = c + G$, where c is constant and G is a Gamma random variable.

So, we are considering a Poisson(λ) random variable where $\lambda = c + G$ where c is a parameter and G is a Gamma random variable.

The resulting distribution is . . .

The resulting distribution is . . .

$$\mathbb{P}(X_k(n) = j) = \sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!} \left(\frac{\beta}{1+\beta}\right)^i \left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!},$$

a convolution of a Negative Binomial random variable and a Poisson random variable.

The Delaporte Distribution

The resulting distribution is . . .

$$\mathbb{P}(X_k(n) = j) = \sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!} \left(\frac{\beta}{1+\beta}\right)^i \left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!},$$

a convolution of a Negative Binomial random variable and a Poisson random variable.

This relatively obscure distribution is called the Delaporte distribution, whose name comes from Delaporte (1959) who used it to model car accidents.

The Delaporte Distribution

It does have (under certain conditions) the asymptotic Poisson property that is needed:

The Delaporte Distribution

It does have (under certain conditions) the asymptotic Poisson property that is needed:

Theorem. Let $n, k \in \mathbb{Z}^+$ with $k \geq 3$. Define $D \sim$ Delaporte(λ, α, β), and $P \sim$ Poisson$(\lambda + \alpha\beta)$. Then the $\mathrm{mgf}(D) \to \mathrm{mgf}(P)$ as $k \to \infty$ under the following assumptions:

- $\alpha \sim \dfrac{\binom{n}{k}}{n^{k-1}}$
- $\beta \sim \dfrac{n^{k-2}}{2^{\binom{k}{2}}}$
- $n = o\left(k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}\right)$ (so that $\alpha\beta \to 0$)

The Delaporte Distribution

It does have (under certain conditions) the asymptotic Poisson property that is needed:

Theorem. Let $n, k \in \mathbb{Z}^+$ with $k \geq 3$. Define $D \sim$ Delaporte(λ, α, β), and $P \sim$ Poisson$(\lambda + \alpha\beta)$. Then the $\mathrm{mgf}(D) \to \mathrm{mgf}(P)$ as $k \to \infty$ under the following assumptions:

- $\alpha \sim \dfrac{\binom{n}{k}}{n^{k-1}}$
- $\beta \sim \dfrac{n^{k-2}}{2^{\binom{k}{2}}}$

We'll attempt to justify before presenting the proof.

- $n = o\left(k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}\right)$ (so that $\alpha\beta \to 0$)

We know that $\mu = \mathbb{E}(X_k(n)) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$. Using Zeilberger's Maple package `SMCramsey` that accompanies his article *Symbolic moment calculus II: why is Ramsey theory sooooo eeeenormously hard?* we find the leading terms for the second and third moments about the mean for X_k for small k:

We know that $\mu = \mathbb{E}(X_k(n)) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$. Using Zeilberger's Maple package `SMCramsey` that accompanies his article *Symbolic moment calculus II: why is Ramsey theory sooooo eeeenormously hard?* we find the leading terms for the second and third moments about the mean for X_k for small k:

k	$\mathbb{E}((X_k-\mu)^2)$	$\mathbb{E}((X_k-\mu)^3)$
3	$\binom{n}{3}\cdot\frac{3}{2^4}$	$\binom{n}{3}\cdot\frac{6n}{2^6}$
4	$\binom{n}{4}\cdot\frac{12n}{2^{10}}$	$\binom{n}{4}\cdot\frac{24n^3}{2^{15}}$
5	$\binom{n}{5}\cdot\frac{15n^2}{2^{18}}$	$\binom{n}{5}\cdot\frac{15n^5}{2^{27}}$
6	$\binom{n}{6}\cdot\frac{10n^3}{2^{28}}$	$\binom{n}{6}\cdot\frac{10n^7}{3\cdot 2^{42}}$
7	$\binom{n}{7}\cdot\frac{35n^4}{2\cdot 2^{42}}$	$\binom{n}{7}\cdot\frac{35n^9}{6\cdot 2^{64}}$
8	$\binom{n}{8}\cdot\frac{84n^5}{15\cdot 2^{56}}$	$\binom{n}{8}\cdot\frac{42n^{11}}{255\cdot 2^{84}}$

$$\sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha) i!} \left(\frac{\beta}{1+\beta}\right)^{i} \left(\frac{1}{1+\beta}\right)^{\alpha} \frac{\lambda^{j-i} e^{-\lambda}}{(j-i)!}$$

$$\mathbb{E}(X_k) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$$

k	$\mathbb{E}((X-\mu)^2)$	$\mathbb{E}((X-\mu)^3)$
3	$\binom{n}{3} \cdot \frac{3}{2^4}$	$\binom{n}{3} \cdot \frac{6n}{2^6}$
4	$\binom{n}{4} \cdot \frac{12n}{2^{10}}$	$\binom{n}{4} \cdot \frac{24n^3}{2^{15}}$
5	$\binom{n}{5} \cdot \frac{15n^2}{2^{18}}$	$\binom{n}{5} \cdot \frac{15n^5}{2^{27}}$
6	$\binom{n}{6} \cdot \frac{10n^3}{2^{28}}$	$\binom{n}{6} \cdot \frac{10n^7}{3 \cdot 2^{42}}$
7	$\binom{n}{7} \cdot \frac{35n^4}{2 \cdot 2^{42}}$	$\binom{n}{7} \cdot \frac{35n^9}{6 \cdot 2^{64}}$
8	$\binom{n}{8} \cdot \frac{84n^5}{15 \cdot 2^{56}}$	$\binom{n}{8} \cdot \frac{42n^{11}}{255 \cdot 2^{84}}$

$$\mathbb{E}(X_k) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$$

k	$\mathbb{E}((X-\mu)^2)$	$\mathbb{E}((X-\mu)^3)$
3	$\binom{n}{3} \cdot \frac{3}{2^4}$	$\binom{n}{3} \cdot \frac{6n}{2^6}$
4	$\binom{n}{4} \cdot \frac{12n}{2^{10}}$	$\binom{n}{4} \cdot \frac{24n^3}{2^{15}}$
5	$\binom{n}{5} \cdot \frac{15n^2}{2^{18}}$	$\binom{n}{5} \cdot \frac{15n^5}{2^{27}}$
6	$\binom{n}{6} \cdot \frac{10n^3}{2^{28}}$	$\binom{n}{6} \cdot \frac{10n^7}{3 \cdot 2^{42}}$
7	$\binom{n}{7} \cdot \frac{35n^4}{2 \cdot 2^{42}}$	$\binom{n}{7} \cdot \frac{35n^9}{6 \cdot 2^{64}}$
8	$\binom{n}{8} \cdot \frac{84n^5}{15 \cdot 2^{56}}$	$\binom{n}{8} \cdot \frac{42n^{11}}{255 \cdot 2^{84}}$

$$\sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!} \left(\frac{\beta}{1+\beta}\right)^i \left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!}$$

$$\mu = \mathbb{E}(D) = \lambda + \alpha\beta$$

$$\mathbb{E}((D-\mu)^2) = \lambda + \alpha\beta(1+\beta)$$

$$\mathbb{E}((D-\mu)^3) = \lambda + \alpha\beta(1+3\beta+2\beta^2).$$

$$\mathbb{E}(X_k) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$$

k	$\mathbb{E}((X-\mu)^2)$	$\mathbb{E}((X-\mu)^3)$
3	$\binom{n}{3} \cdot \frac{3}{2^4}$	$\binom{n}{3} \cdot \frac{6n}{2^6}$
4	$\binom{n}{4} \cdot \frac{12n}{2^{10}}$	$\binom{n}{4} \cdot \frac{24n^3}{2^{15}}$
5	$\binom{n}{5} \cdot \frac{15n^2}{2^{18}}$	$\binom{n}{5} \cdot \frac{15n^5}{2^{27}}$
6	$\binom{n}{6} \cdot \frac{10n^3}{2^{28}}$	$\binom{n}{6} \cdot \frac{10n^7}{3 \cdot 2^{42}}$
7	$\binom{n}{7} \cdot \frac{35n^4}{2 \cdot 2^{42}}$	$\binom{n}{7} \cdot \frac{35n^9}{6 \cdot 2^{64}}$
8	$\binom{n}{8} \cdot \frac{84n^5}{15 \cdot 2^{56}}$	$\binom{n}{8} \cdot \frac{42n^{11}}{255 \cdot 2^{84}}$

$$\sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!} \left(\frac{\beta}{1+\beta}\right)^i \left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!}$$

$$\mu = \mathbb{E}(D) = \lambda + \alpha\beta$$
$$\mathbb{E}((D-\mu)^2) = \lambda + \alpha\beta(1+\beta)$$
$$\mathbb{E}((D-\mu)^3) = \lambda + \alpha\beta(1+3\beta+2\beta^2).$$

We have $\mathbb{E}((D-\mu)^2) = \mathbb{E}(D) + \alpha\beta^2$, so

$$\alpha\beta^2 \sim \binom{n}{k} \frac{n^{k-3}}{2^{2\binom{k}{2}-2}}.$$

$$\mathbb{E}(X_k) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$$

k	$\mathbb{E}((X-\mu)^2)$	$\mathbb{E}((X-\mu)^3)$
3	$\binom{n}{3} \cdot \frac{3}{2^4}$	$\binom{n}{3} \cdot \frac{6n}{2^6}$
4	$\binom{n}{4} \cdot \frac{12n}{2^{10}}$	$\binom{n}{4} \cdot \frac{24n^3}{2^{15}}$
5	$\binom{n}{5} \cdot \frac{15n^2}{2^{18}}$	$\binom{n}{5} \cdot \frac{15n^5}{2^{27}}$
6	$\binom{n}{6} \cdot \frac{10n^3}{2^{28}}$	$\binom{n}{6} \cdot \frac{10n^7}{3 \cdot 2^{42}}$
7	$\binom{n}{7} \cdot \frac{35n^4}{2 \cdot 2^{42}}$	$\binom{n}{7} \cdot \frac{35n^9}{6 \cdot 2^{64}}$
8	$\binom{n}{8} \cdot \frac{84n^5}{15 \cdot 2^{56}}$	$\binom{n}{8} \cdot \frac{42n^{11}}{255 \cdot 2^{84}}$

$$\sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!} \left(\frac{\beta}{1+\beta}\right)^i \left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!}$$

$$\begin{aligned}
\mu &= \mathbb{E}(D) = \lambda + \alpha\beta \\
\mathbb{E}((D-\mu)^2) &= \lambda + \alpha\beta(1+\beta) \\
\mathbb{E}((D-\mu)^3) &= \lambda + \alpha\beta(1+3\beta+2\beta^2).
\end{aligned}$$

We have $\mathbb{E}((D-\mu)^2) = \mathbb{E}(D) + \alpha\beta^2$, so

$$\alpha\beta^2 \sim \binom{n}{k} \frac{n^{k-3}}{2^{2\binom{k}{2}-2}}.$$

Looking at the third moments, we have evidence to suggest that

$$2\alpha\beta^3 \sim \binom{n}{k} \frac{n^{2k-5}}{2^{3\binom{k}{2}-3}}.$$

$$\mathbb{E}(X_k) = \frac{\binom{n}{k}}{2^{\binom{k}{2}-1}}$$

k	$\mathbb{E}((X-\mu)^2)$	$\mathbb{E}((X-\mu)^3)$
3	$\binom{n}{3} \cdot \frac{3}{2^4}$	$\binom{n}{3} \cdot \frac{6n}{2^6}$
4	$\binom{n}{4} \cdot \frac{12n}{2^{10}}$	$\binom{n}{4} \cdot \frac{24n^3}{2^{15}}$
5	$\binom{n}{5} \cdot \frac{15n^2}{2^{18}}$	$\binom{n}{5} \cdot \frac{15n^5}{2^{27}}$
6	$\binom{n}{6} \cdot \frac{10n^3}{2^{28}}$	$\binom{n}{6} \cdot \frac{10n^7}{3 \cdot 2^{42}}$
7	$\binom{n}{7} \cdot \frac{35n^4}{2 \cdot 2^{42}}$	$\binom{n}{7} \cdot \frac{35n^9}{6 \cdot 2^{64}}$
8	$\binom{n}{8} \cdot \frac{84n^5}{15 \cdot 2^{56}}$	$\binom{n}{8} \cdot \frac{42n^{11}}{255 \cdot 2^{84}}$

$$\sum_{i=0}^{j} \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)i!}\left(\frac{\beta}{1+\beta}\right)^i\left(\frac{1}{1+\beta}\right)^\alpha \frac{\lambda^{j-i}e^{-\lambda}}{(j-i)!}$$

$$\begin{aligned}
\mu = \mathbb{E}(D) &= \lambda + \alpha\beta \\
\mathbb{E}((D-\mu)^2) &= \lambda + \alpha\beta(1+\beta) \\
\mathbb{E}((D-\mu)^3) &= \lambda + \alpha\beta(1+3\beta+2\beta^2).
\end{aligned}$$

We have $\mathbb{E}((D-\mu)^2) = \mathbb{E}(D) + \alpha\beta^2$, so

$$\alpha\beta^2 \sim \binom{n}{k}\frac{n^{k-3}}{2^{2\binom{k}{2}-2}}.$$

Looking at the third moments, we have evidence to suggest that

$$2\alpha\beta^3 \sim \binom{n}{k}\frac{n^{2k-5}}{2^{3\binom{k}{2}-3}}.$$

Taking the ratio of these last two expressions yields

$$\alpha \sim \frac{\binom{n}{k}}{n^{k-1}} \quad \text{and} \quad \beta \sim \frac{n^{k-2}}{2^{\binom{k}{2}}}.$$

Proof. Since D is a convolution of a Negative Binomial random variable with success probability $\frac{\beta}{1+\beta}$ and mean $\alpha\beta$ and a Poisson random variable with mean λ, using the moment generating functions of these, we easily have

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^{\alpha}}.$$

Proof. Since D is a convolution of a Negative Binomial random variable with success probability $\frac{\beta}{1+\beta}$ and mean $\alpha\beta$ and a Poisson random variable with mean λ, using the moment generating functions of these, we easily have

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha}.$$

Isolating the denominator, we have

$$(1-\beta(e^t-1))^\alpha = \left(1 - \frac{\alpha\beta(e^t-1)}{\alpha}\right)^\alpha.$$

Proof. Since D is a convolution of a Negative Binomial random variable with success probability $\frac{\beta}{1+\beta}$ and mean $\alpha\beta$ and a Poisson random variable with mean λ, using the moment generating functions of these, we easily have

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha}.$$

Isolating the denominator, we have

$$(1-\beta(e^t-1))^\alpha = \left(1-\frac{\alpha\beta(e^t-1)}{\alpha}\right)^\alpha.$$

Under the given assumptions, for large k (and n) we have

$$\left(1-\frac{\alpha\beta(e^t-1)}{\alpha}\right)^\alpha \approx e^{-\alpha\beta(e^t-1)}, \quad |t| < \frac{1}{n^{k-2}}$$

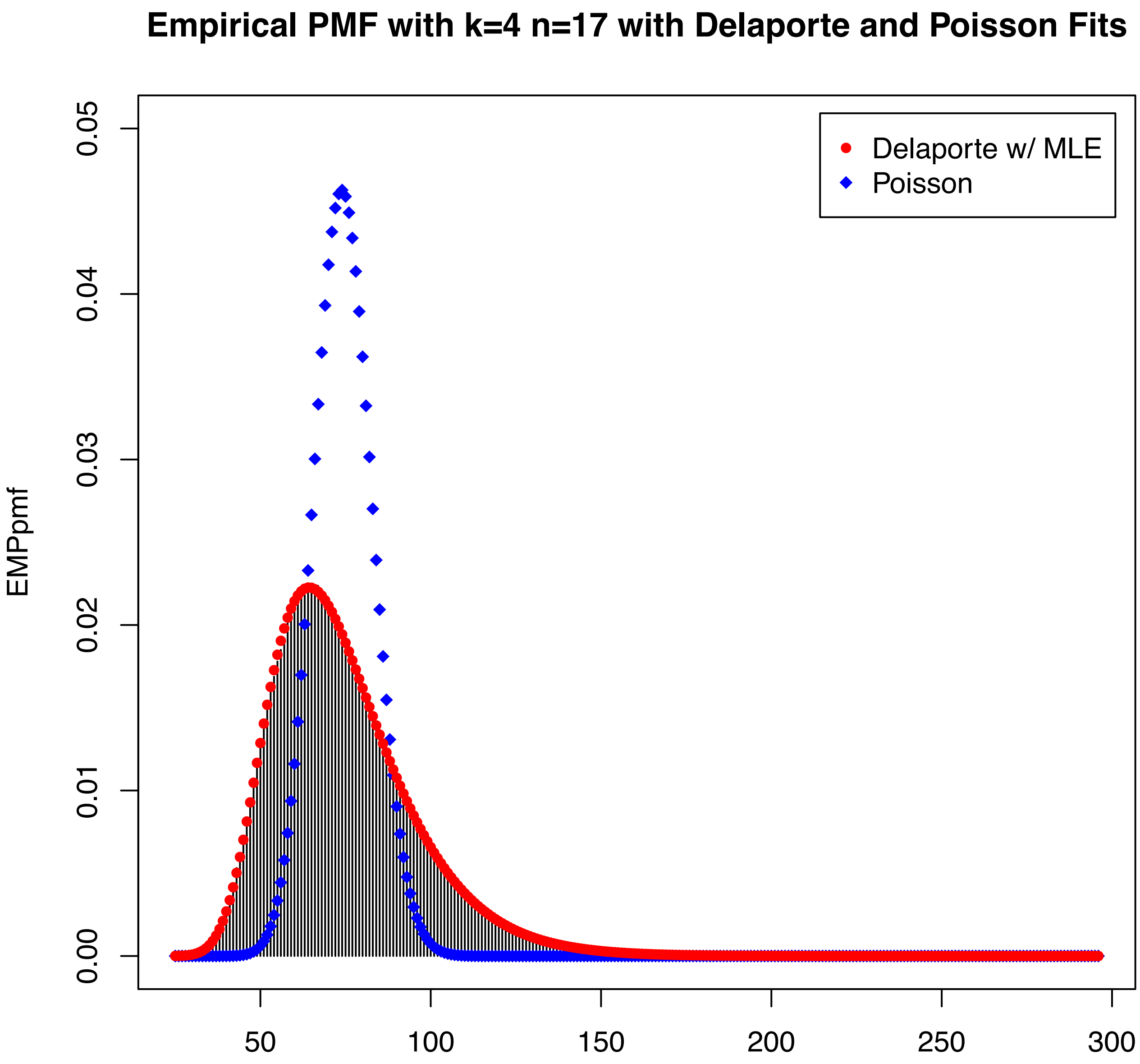
so that mgf$(D) \approx e^{(\lambda+\alpha\beta)(e^t-1)}$ =mgf(P) on an interval including $t = 0$. ☐

The Delaporte Distribution

Does it do better than Poisson?

The Delaporte Distribution

Does it do better than Poisson?

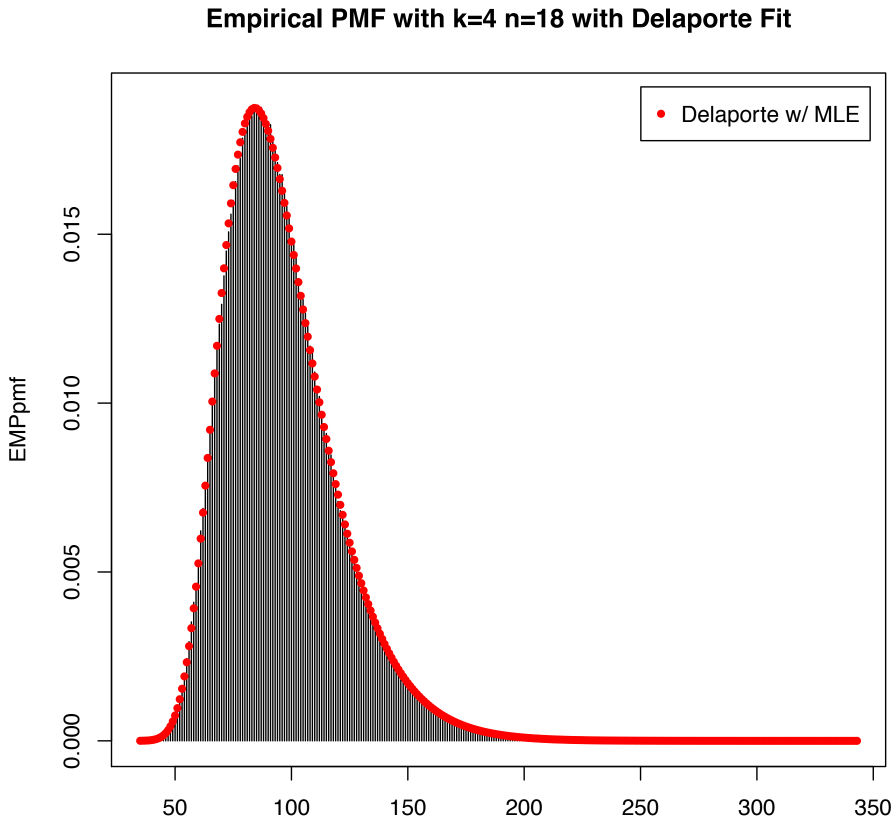


The Delaporte Distribution

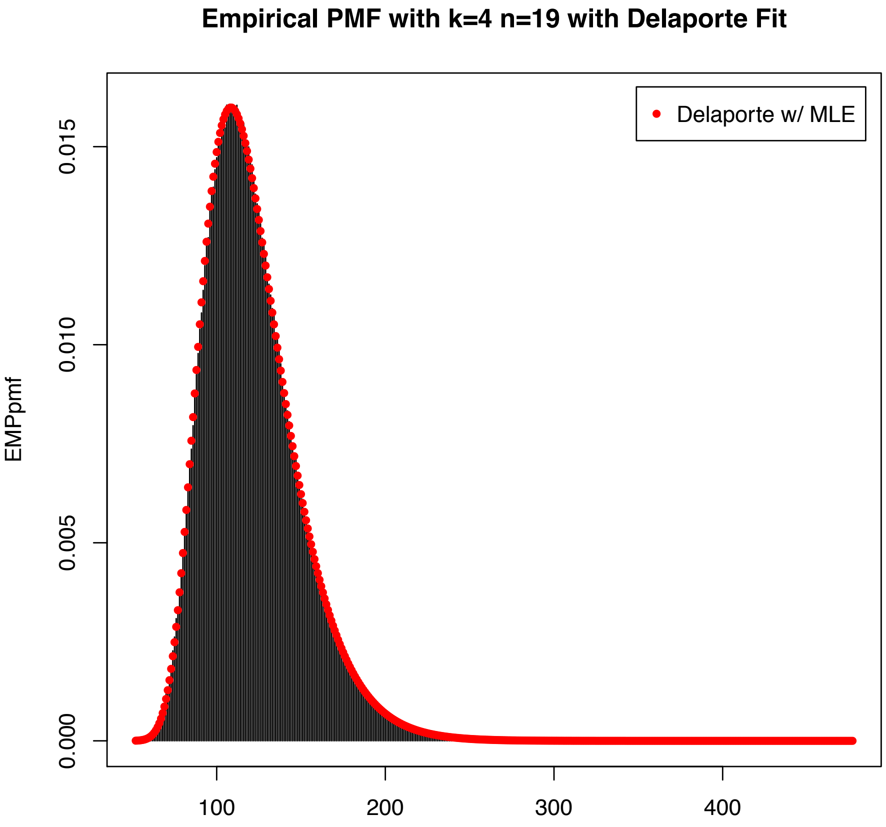
Lest you think that was a fluke:

The Delaporte Distribution

Lest you think that was a fluke:


Empirical PMF with k=4 n=18 with Delaporte Fit

$$k = 4, n = 18$$


Empirical PMF with k=4 n=19 with Delaporte Fit

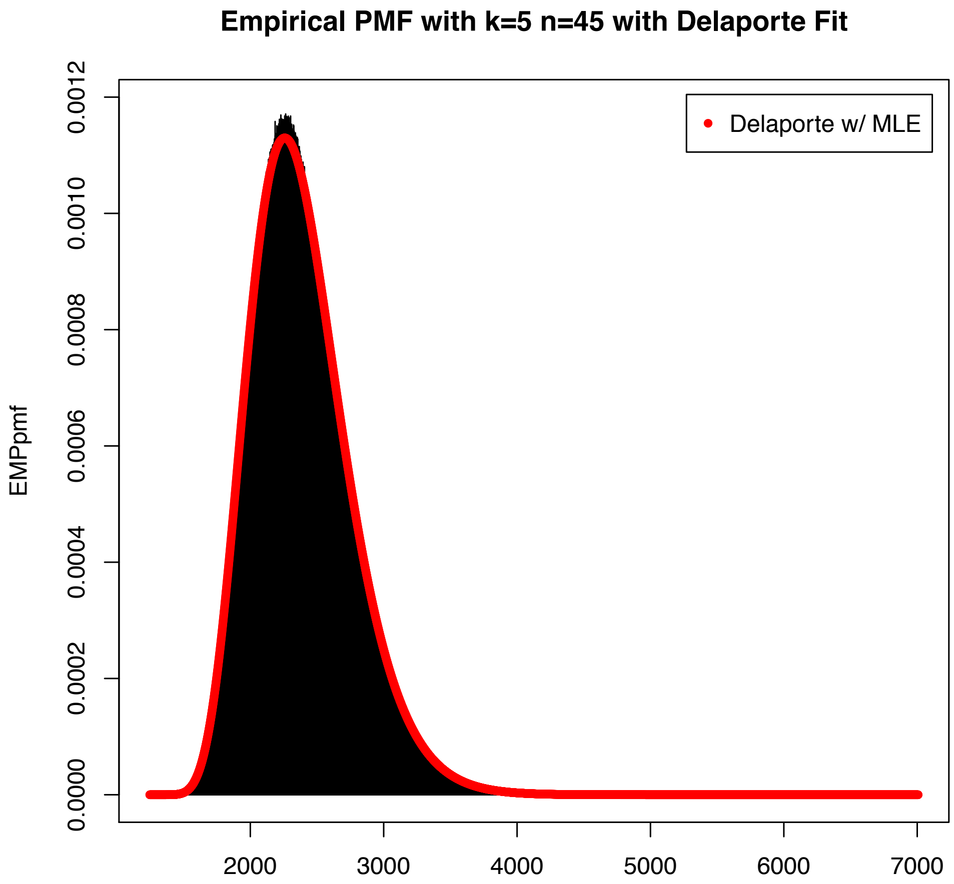
$$k = 4, n = 19$$

The Delaporte Distribution

Three flukes?

The Delaporte Distribution

Three flukes?



$$k = 5, n = 45$$

Consequences

Assume that the distribution of $X_k(n)$ is Delaporte (BIG Assumption) so that

$$\mathbb{P}(X_k(n) = 0) = \frac{1}{e^{\lambda}(1+\beta)^{\alpha}}.$$

Consequences

Assume that the distribution of $X_k(n)$ is Delaporte (BIG Assumption) so that

$$\mathbb{P}(X_k(n) = 0) = \frac{1}{e^\lambda(1+\beta)^\alpha}.$$

We had defined: $\widehat{R}_k(\tau) = \min_{n \in \mathbb{Z}^+} \{n : \mathbb{P}(X_k(n) > 0) > \tau\}$.

Consequences

Assume that the distribution of $X_k(n)$ is Delaporte (BIG Assumption) so that

$$\mathbb{P}(X_k(n) = 0) = \frac{1}{e^{\lambda}(1+\beta)^{\alpha}}.$$

We had defined: $\widehat{R}_k(\tau) = \min\limits_{n \in \mathbb{Z}^+} \{n : \mathbb{P}(X_k(n) > 0) > \tau\}$.

It would be nice to let $\tau = 1 - \frac{1}{2^{\binom{n}{k}}}$ so that we want $\mathbb{P}(X_k(n) = 0) < \frac{1}{2^{\binom{n}{k}}}$.

But this doesn't occur for any calculated parameters and values of n that are computationally feasible (and can't happen asymptotically).

Consequences

Assume that the distribution of $X_k(n)$ is Delaporte (BIG Assumption) so that

$$\mathbb{P}(X_k(n) = 0) = \frac{1}{e^{\lambda}(1+\beta)^{\alpha}}.$$

We had defined: $\widehat{R}_k(\tau) = \min_{n \in \mathbb{Z}^+} \{n : \mathbb{P}(X_k(n) > 0) > \tau\}$.

It would be nice to let $\tau = 1 - \frac{1}{2^{\binom{n}{k}}}$ so that we want $\mathbb{P}(X_k(n) = 0) < \frac{1}{2^{\binom{n}{k}}}$.

But this doesn't occur for any calculated parameters and values of n that are computationally feasible (and can't happen asymptotically).

Using Zeilberger's package and setting Delaporte moments equal to Ramsey graph moments and letting $\tau \approx 2^{-kn^2}$ (and using $\alpha \approx \frac{n}{k}$) gets you in roughly the area of $R(k)$ values and bounds for small k. Unfortunately, unless making assumptions about one of λ, α, β, the calculation time is too much for $k \geq 6$.

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

We know that

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha} \qquad \text{and} \qquad \mathrm{mgf}(P) = e^{(\lambda+\alpha\beta)(e^t-1)}$$

so that we need $(1-\beta(e^t-1))^\alpha \rightarrow e^{-\alpha\beta(e^t-1)}$.

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

We know that

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha} \qquad \text{and} \qquad \mathrm{mgf}(P) = e^{(\lambda+\alpha\beta)(e^t-1)}$$

so that we need $(1-\beta(e^t-1))^\alpha \rightarrow e^{-\alpha\beta(e^t-1)}$.

Hence, we need $\alpha\beta \rightarrow 0$.

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

We know that

$$\text{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha} \qquad \text{and} \qquad \text{mgf}(P) = e^{(\lambda+\alpha\beta)(e^t-1)}$$

so that we need $(1-\beta(e^t-1))^\alpha \rightarrow e^{-\alpha\beta(e^t-1)}$.

Hence, we need $\alpha\beta \rightarrow 0$.

Based on our evidence-based assumptions for α and β, we require

$$n = o\left(k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}\right)$$

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

We know that

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^\alpha} \qquad \text{and} \qquad \mathrm{mgf}(P) = e^{(\lambda+\alpha\beta)(e^t-1)}$$

so that we need $(1-\beta(e^t-1))^\alpha \rightarrow e^{-\alpha\beta(e^t-1)}$.

Hence, we need $\alpha\beta \rightarrow 0$.

Based on our evidence-based assumptions for α and β, we require

$$n = o\left(k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}\right)$$

so that if $n > k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}$ then mgf(Delaporte) $\not\rightarrow$ mgf(Poisson).

Consequences

We do have a very interesting by-product of our proof that mgf(D) $\rightarrow$ mgf(P):

We know that

$$\mathrm{mgf}(D) = \frac{e^{\lambda(e^t-1)}}{(1-\beta(e^t-1))^{\alpha}} \qquad \text{and} \qquad \mathrm{mgf}(P) = e^{(\lambda+\alpha\beta)(e^t-1)}$$

so that we need $(1-\beta(e^t-1))^{\alpha} \rightarrow e^{-\alpha\beta(e^t-1)}$.

Hence, we need $\alpha\beta \rightarrow 0$.

Based on our evidence-based assumptions for α and β, we require

$$n = o\left(k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}\right)$$

so that if $n > k^{1+\frac{1}{k-1}} \cdot 2^{\frac{k}{2}}$ then mgf(Delaporte) $\not\rightarrow$ mgf(Poisson).

But we know that, asymptotically, we have a Poisson distribution and the Delaporte sure appears to be a fantastic fit.

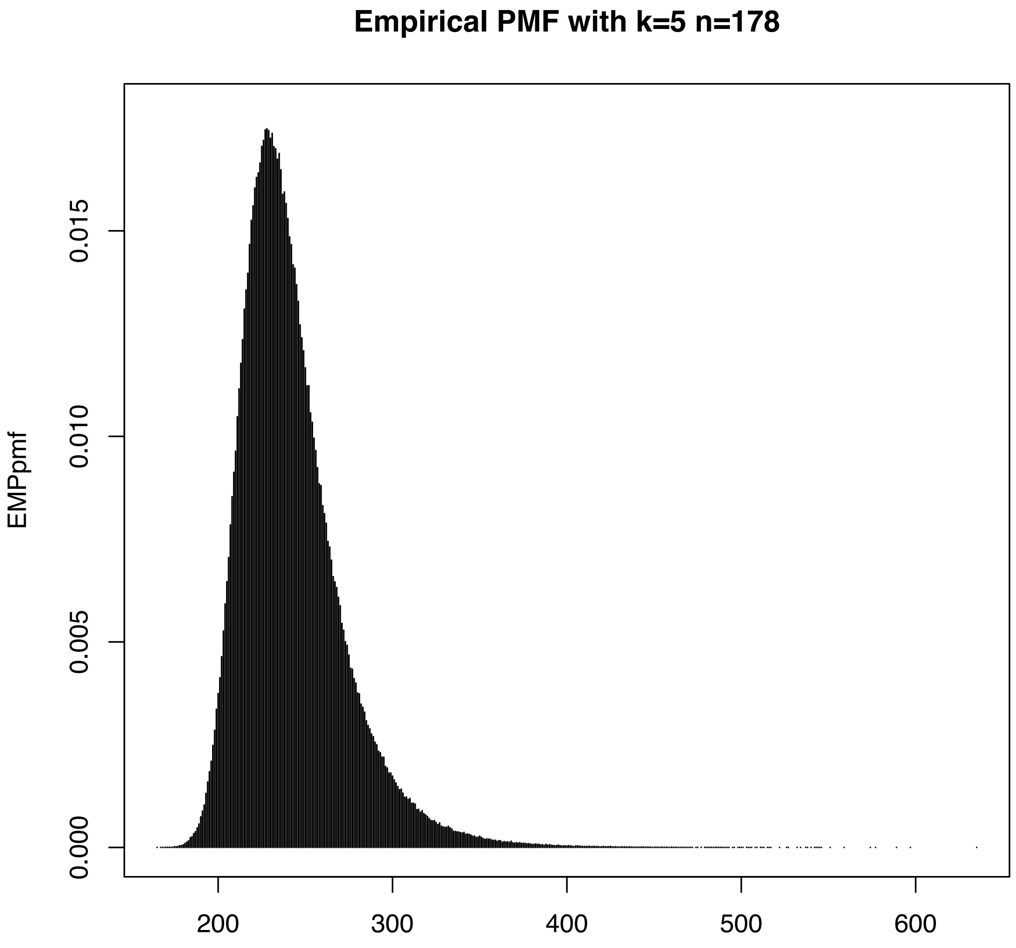
Arithmetic Progressions

Maria was very quick at programming, so after she finished with the graphs, she coded the similar problem for arithmetic progressions.

She counted the number of monochromatic k-term arithmetic progressions in 2-colorings of $\{1, 2, \ldots, n\}$ and produced the resulting sample histograms.

Arithmetic Progressions

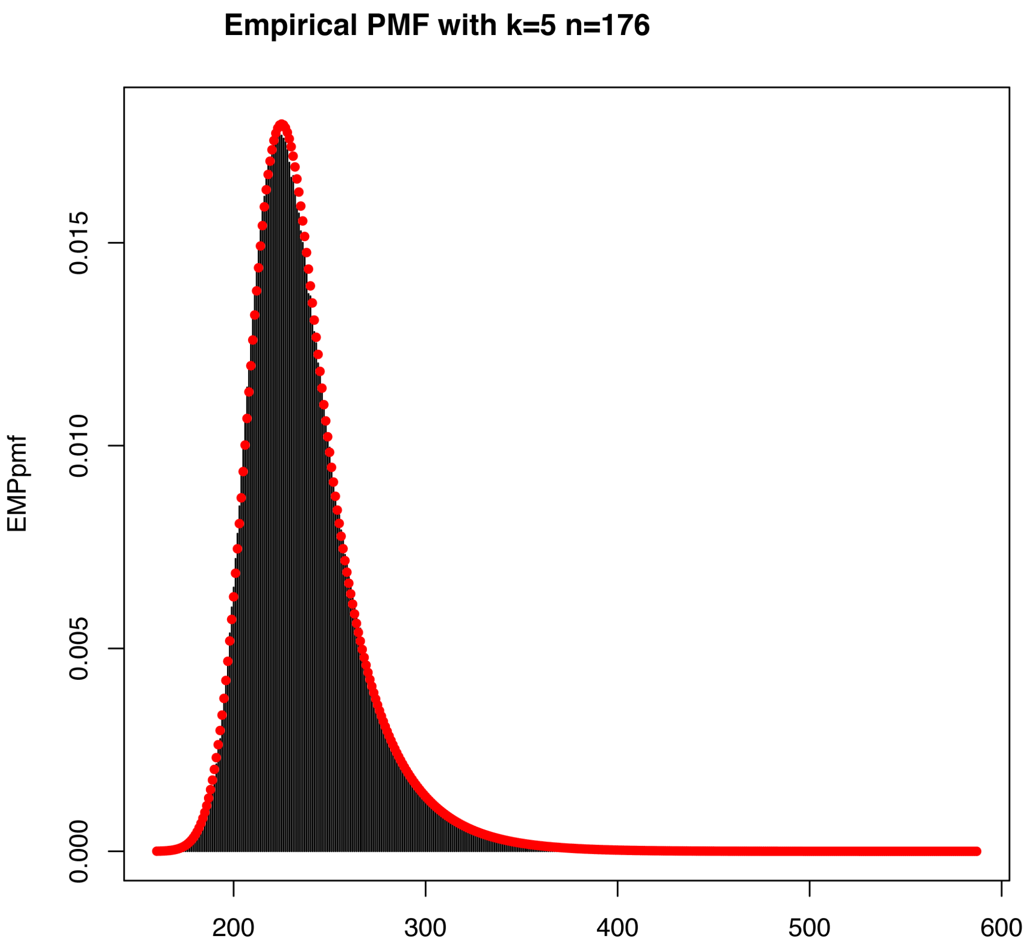
Looks like they also follow a formula.



Arithmetic Progressions

After much less searching we again found a very good fit:

$k = 5, n = 176$



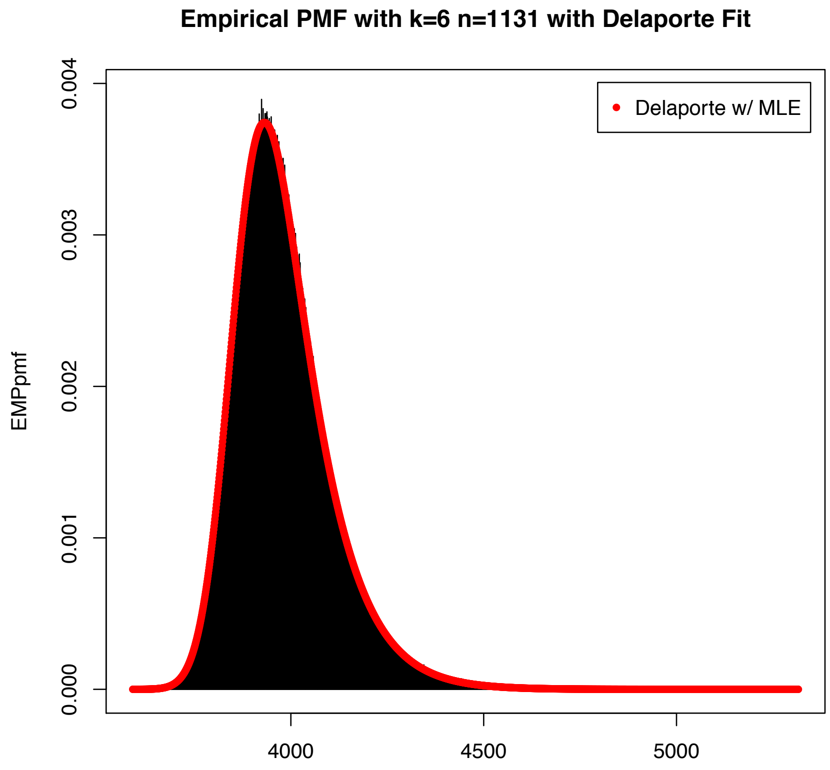
where the overlay uses MLEs.

Arithmetic Progressions

We again found a very good fit:

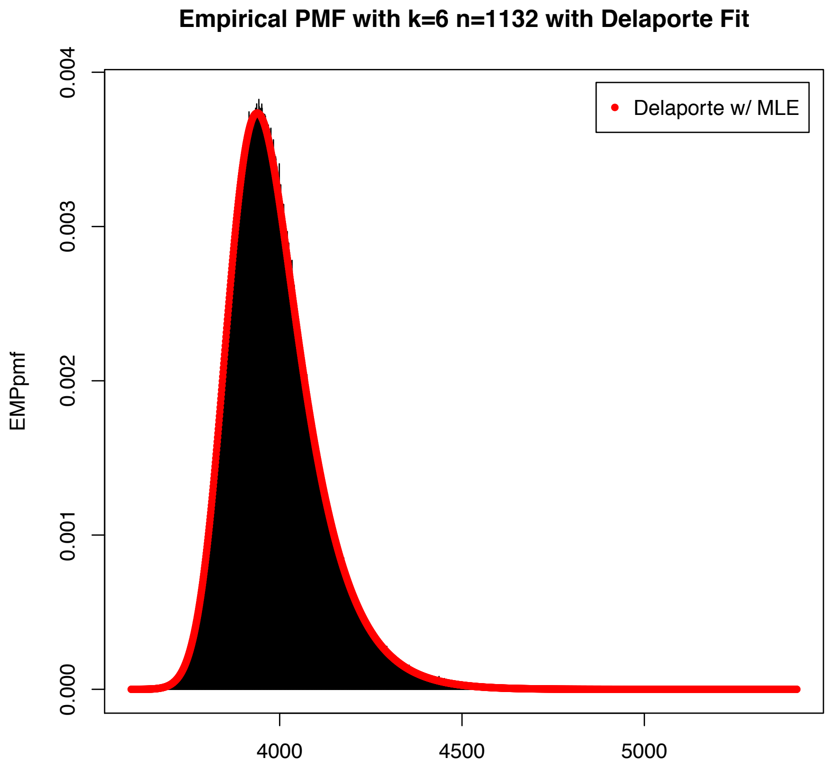
Arithmetic Progressions

We again found a very good fit: the Delaporte distribution.



Empirical PMF with k=6 n=1131 with Delaporte Fit

$$k = 6, n = 1131$$



Empirical PMF with k=6 n=1132 with Delaporte Fit

$$k = 6, n = 1132$$

Is there a Delaporte paradigm for Ramsey objects?

Thank You!

`www.aaronrobertson.org`