Stochastic Dominance in the Poisson Paradigm

Notes by Albert Zhang

April 8, 2020

This set of notes is an attempt to organize my thoughts on stochastic dominance of certain binomial-type random variables using Poisson random variables. The broad idea is that binomials are harder to work with due to their lack of independence. Much of what follows borrows heavily from my lecture notes for CS 271; the last section on random geometric graphs was a homework problem for the class.

1 Preliminaries

Recall the Poisson Point Process (PPP), which describes an arrival process in some region of space. We will be considering PPP's in both a discrete setting (Balls & Bins) and a continuous setting (Random Geometric Graphs).

In the discrete case, we have a finite collection of n "points", which later we'll call bins, in space. To each point we'll associate an independent arrival process modeled by a Poi(1) random variable. Note that the sum of the arrivals across all points is thus given by a Poi(n) random variable.

In the continuous case, we will be observing an arrival process on the unit square in $\mathbb{R}^2$. Given the parameter λ, the number of arrivals in a subset S of the unit square is distributed as $\text{Poi}(\lambda \times \text{area}(S))$. The key property here is that arrivals in disjoint regions of space are independent.

With the definitions out of the way, let's cover some bounds that we will need later.

Theorem 1.1 (Vanilla Chernoff)

Let $X = \sum_{i=1}^n X_i$ be a sum of independent $[0,1]$-valued random variables with $\mathbb{E}[X] = \mu$. For $\beta > 0$ we have the upper tail bounds:

$$\mathbb{P}[X \geq (1+\beta)\mu] \leq \begin{cases} \exp\left(-\frac{\beta^2\mu}{2+\beta}\right) & \beta > 1 \\ \exp\left(-\frac{\beta^2\mu}{3}\right) & 0 < \beta \leq 1. \end{cases}$$

Proof. Can be found in my lecture notes for CS 271. □

Theorem 1.2 (Chernoff for Poisson I)
Suppose $X \sim \text{Poi}(\mu)$. Then for $\lambda > 0$, we have the upper tail bound

$$\mathbb{P}[X \geq \mu + \lambda] \leq \exp\left\{-(\mu+\lambda)\ln\frac{\mu+\lambda}{\mu} + \lambda\right\}.$$

For $\mu > \lambda > 0$, we have the lower tail bound,

$$\mathbb{P}[X \leq \mu - \lambda] \leq \exp\left\{-(\mu-\lambda)\ln\frac{\mu-\lambda}{\mu} - \lambda\right\}.$$

Proof. By Markov's inequality, we have

$$\begin{aligned}\mathbb{P}[X \geq m] &= \mathbb{P}[e^{Xt} \geq e^{mt}] \quad \text{for } t > 0 \\ &\leq e^{-mt}\mathbb{E}[e^{Xt}].\end{aligned}$$

We may compute

$$\begin{aligned}\mathbb{E}[e^{Xt}] &= \sum_{k=0}^{\infty} e^{kt} \cdot \frac{\mu^k e^{-\mu}}{k!} \\ &= \frac{e^{-\mu}}{e^{-\mu e^t}} \cdot \sum_{k=0}^{\infty} \frac{(\mu e^t)^k e^{-\mu e^t}}{k!} \\ &= e^{\mu e^t - \mu},\end{aligned}$$

where in the second line the series sums to 1 since it is the distribution of a $\text{Poi}(\mu e^t)$ random variable. Therefore, our bound becomes

$$e^{-mt+\mu e^t-\mu}. \tag{1}$$

The first derivative is

$$e^{-mt+\mu e^t-\mu}(-m+\mu e^t). \tag{2}$$

The second derivative is

$$e^{-mt+\mu e^t-\mu}(\mu e^t + (-m+\mu e^t)^2) \geq 0.$$

Hence the function is convex in t, so we may set (2) to 0, which implies

$$\mu e^t = m,$$

so that

$$t = \ln\frac{m}{\mu}.$$

Plugging this back into (1) gives us

$$e^{-m\ln\frac{m}{\mu}+m-\mu}.$$

So, if we take $m = \mu + \lambda$, this gives us

$$\mathbb{P}[X \geq \mu + \lambda] \leq \exp\left\{-(\mu+\lambda)\ln\frac{\mu+\lambda}{\mu} + \lambda\right\}. \tag{3}$$

The proof for the lower tail bound proceeds similarly. We have

$$\begin{aligned}
\mathbb{P}[X \leq m] &= \mathbb{P}[e^{Xt} \leq e^{mt}] \quad \text{for } t > 0 \\
&= \mathbb{P}[e^{-Xt} \geq e^{-mt}] \\
&\leq e^{mt}\mathbb{E}[e^{-Xt}] \\
&= e^{mt}\sum_{k=0}^{\infty} e^{-kt}\frac{\mu^k e^{-\mu}}{k!} \\
&= e^{mt}e^{-\mu+\mu e^{-t}}\sum_{k=0}^{\infty}\frac{(\mu e^{-t})^k e^{-\mu e^{-t}}}{k!} \\
&= e^{mt+\mu e^{-t}-\mu}.
\end{aligned}$$

The first derivative is

$$e^{mt+\mu e^{-t}-\mu}(m - \mu e^{-t}). \tag{4}$$

The second derivative is

$$e^{mt+\mu e^{-t}-\mu}(\mu e^{-t} + (m - \mu e^{-t})^2) \geq 0,$$

so our bound is convex in t. Optimizing over t using (4), we take $t = -\ln\frac{m}{\mu}$. Plugging this back in, along with $m = \mu - \lambda$, gives us our lower tail bound

$$\mathbb{P}[X \leq \mu - \lambda] \leq \exp\left\{-(\mu-\lambda)\ln\frac{\mu-\lambda}{\mu} - \lambda\right\}. \tag{5}$$

□

Corollary 1.3 (Chernoff for Poisson II)
If $X \sim \text{Poi}(\mu)$ as before, we have the variants, for any $\beta > 0$,

$$\begin{aligned}
\mathbb{P}[X \geq (1+\beta)\mu] &\leq \exp\{-\mu(-\beta + (1+\beta)\ln(1+\beta))\} \\
\mathbb{P}[X \leq (1-\beta)\mu] &\leq \exp\{-\mu(\beta + (1-\beta)\ln(1-\beta))\}.
\end{aligned}$$

Proof. Following immediately by taking $\lambda = \mu\beta$ in Theorem 1.2. □

2 Balls & Bins

In the standard balls & bins model, we throw m balls into n bins independently and u.a.r. We are concerned with the issue of load-balancing. In particular, what is the maximum load of any bin? With the Poisson paradigm, we will be able to prove an asymptotic result by translating probabilities under the standard balls and bins model to an easier to work with Poisson model.

The idea behind the following lemma essentially that of stochastic dominance. Since we only care about bounds and asymptotic behavior, it suffices to work in a Poisson model which "dominates" the standard balls and bins model. This will work to our advantage since multinomial coefficients tend to be much harder to work with than terms appearing in a Poisson distribution.

To get some intuition for why we want to use Poisson to dominate our balls and bins model, recall that in the limit, if we keep $m/n = \lambda$ constant, $\text{Bin}(m, 1/n) \xrightarrow{d} \text{Poi}(\lambda)$.

Lemma 2.1

Let $\mathcal{E}$ be any event that depends only on the bin loads such that $\mathbb{P}[\mathcal{E}]$ is monotonically increasing with m. Then

$$\mathbb{P}_X[\mathcal{E}] \leq 4\mathbb{P}_Y[\mathcal{E}],$$

where $\mathbb{P}_X$ is the standard balls and bins model, and $\mathbb{P}_Y$ is the Poisson model, where for each bin we associate an independent $\text{Poi}(m/n)$ variable.

Proof. First, we note that the joint distribution of the balls and bins model is equal to that of n independent $Y_i \sim \text{Poi}(\lambda)$ variables conditioned on $\sum Y_i = m$. Explicitly,

$$\mathbb{P}[X_{1:n} = k_{1:n}] = \frac{1}{n^m} \cdot \frac{m!}{k_1! \dots k_n!} = \frac{\prod_{i=1}^n \frac{e^{-\lambda}\lambda^{k_i}}{k_i!}}{\frac{e^{-n\lambda}(n\lambda)^m}{m!}} = \frac{\prod_{i=1}^n \mathbb{P}[Y_i = k_i]}{\mathbb{P}[\sum_{i=1}^n Y_i = m]}.$$

So, for an event $\mathcal{E}$, this tells us

$$\mathbb{P}_X[\mathcal{E}] = \mathbb{P}_Y\left[\mathcal{E} | \sum_{i=1}^n Y_i = m\right]$$

Now, we have

$$\begin{aligned}
\mathbb{P}_Y[\mathcal{E}] &= \sum_{k=0}^{\infty} \mathbb{P}_Y\left[\mathcal{E} | \sum_{i=1}^n Y_i = k\right] \mathbb{P}_Y\left[\sum_{i=1}^n Y_i = k\right] \\
&\geq \mathbb{P}_Y\left[\mathcal{E} | \sum_{i=1}^n Y_i = m\right] \mathbb{P}_Y\left[\sum_{i=1}^n Y_i \geq m\right] \\
&\geq \mathbb{P}_Y\left[\mathcal{E} | \sum_{i=1}^n Y_i = m\right] \cdot \frac{1}{4} \\
&= \mathbb{P}_X[\mathcal{E}] \cdot \frac{1}{4}.
\end{aligned}$$

In the second line we used monotonicity, and in the third we used the fact that for any $Y \sim \text{Poi}(\lambda)$, we have

$$\mathbb{P}[Y \geq \mathbb{E}[Y]] \geq 1/4, \tag{6}$$

and applied to the sum of independent Poissons $Y = \sum Y_i$, which is itself Poisson. $\square$

Armed with this lemma, we can now show:

Theorem 2.2

For the balls and bins model with $m = n$ (so $\lambda = 1$), the maximum load of any bin is

$$\Theta\left(\frac{\ln n}{\ln\ln n}\right)$$

asymptotically almost surely (a.a.s.).

Proof. It suffices to show that the maximum load lies in

$$\left((1-\epsilon)\frac{\ln n}{\ln\ln n}, (1+\epsilon)\frac{\ln n}{\ln\ln n}\right)$$

for any $\epsilon > 0$ a.a.s. For notation, define $c_1 = 1+\epsilon$ and $c_2 = 1-\epsilon$, and the events

$$\mathcal{E}_1 := \text{some bin contains more than } c_1\tfrac{\ln n}{\ln\ln n} \text{ balls,}$$
$$\mathcal{E}_2 := \text{no bin contains more than } c_2\tfrac{\ln n}{\ln\ln n} \text{ balls.}$$

We will show that $\mathbb{P}[\mathcal{E}_i] = o(1)$ for $i = 1, 2$.

Note that since $\mathcal{E}_i$ are monotonic (for decreasing, a similar bound holds). Therefore, by Lemma 2.1, we may work with independent Poi(1) variables. We have the following useful bounds:

$$\frac{1}{ek!} \leq \mathbb{P}[Y_i \geq k] \leq \frac{1}{ek!}\left(1 + \frac{1}{k+1} + \frac{1}{(k+1)(k+2)} + \cdots\right) \leq \frac{1}{k!}.$$

Setting $k = c_1\frac{\ln n}{\ln\ln n}$, we get

$$\ln\mathbb{P}[Y_i \geq k] \leq -\ln k! \sim -k\ln k = -c_1 \cdot \frac{\ln n}{\ln\ln n}(\ln\ln n + \ln c_1 - \ln\ln\ln n) \sim -c_1\ln n,$$

which implies $\mathbb{P}[Y_i \geq k] = o(n^{-1})$. Taking a union bound, gives us $\mathbb{P}[\mathcal{E}_1] = o(1)$.

Next, to show $\mathbb{P}[\mathcal{E}_2] = o(1)$, we note that

$$\mathbb{P}_Y[\mathcal{E}_2] = (1 - \mathbb{P}[Y_i \geq k])^n \leq \left(1 - \frac{1}{ek!}\right)^n \leq e^{-\frac{n}{ek!}}.$$

Taking $k = c_2\frac{\ln n}{\ln\ln n}$, we can show that the exponent $\frac{n}{ek!} \to \infty$ as $n \to \infty$. It follows that $\mathbb{P}[\mathcal{E}_2] = o(1)$ as well. $\square$

3 Random Geometric Graphs

In the n-point random geometric graph model, n points are placed independently and u.a.r. in a unit square, and each point is connected by an edge to its k nearest neighbors. Call the resulting undirected graph G. Note that this model can be viewed as a continuous version of the balls and bins model in $\mathbb{R}^2$. We are concerned with thresholds of k for when G will be connected. In a similar spirit to the previous section, we will find a threshold by translating the original n-point model into a PPP model on the unit square.

Theorem 3.1

In the n-point model on the unit square, $k = \Theta(\log n)$ is a threshold for G to be connected in the sense that

1. We can pick a sufficiently large constant c_1 so that if $k = c_1 \log n$, then w.h.p. G is connected as $n \to \infty$.
2. We can pick a sufficiently small constant c_2 so that if $k = c_2 \log n$, then w.h.p. G is disconnected as $n \to \infty$.

Proof of 1. Partition the unit square up into smaller squares of area $\frac{\log n}{n}$ (boundary/rounding issues are neglible here). There are $\frac{n}{\log n}$ squares of area $\frac{\log n}{n}$. We will show that w.h.p. every square contains at least one point. This can be done by a union bound:

$$\begin{aligned}
\mathbb{P}[\text{at least one square has no pt}] &\leq \frac{n}{\log n}\mathbb{P}[\text{particular square has no pt}] \\
&= \frac{n}{\log n} \cdot \left(1 - \frac{\log n}{n}\right)^n \\
&\leq \frac{n}{\log n} \cdot e^{-\log n} \\
&= \frac{1}{\log n}.
\end{aligned}$$

Thus,

$$\mathbb{P}[\text{every square has} \geq 1 \text{ pt}] \geq 1 - \frac{1}{\log n},$$

so that every square contains at least one point w.h.p.

Now, call D_p the disc of radius $\sqrt{10\frac{\log n}{n}}$ centered at a point p, and X_p the number of points in D_p. Then

$$\mathbb{E}[X_p] = n \cdot \pi \cdot \left(10\frac{\log n}{n}\right) = 10\pi \log n.$$

Note that if D_p is not fully contained in the unit square, then the above expectation would be smaller, and the following Chernoff bound would still hold, since the value of

β would only increase. In particular, we have by the Chernoff bound in Theorem 1.1,

$$\mathbb{P}[X_p \geq 20\pi \log n] \leq e^{-\frac{10\pi \log n}{3}} = n^{-\frac{10}{3}\pi}.$$

But since $\frac{10}{3}\pi > 1$, we may apply a union bound to get

$$\mathbb{P}[\exists p : X_p \geq 20\pi \log n] \leq n \cdot n^{-\frac{10}{3}\pi} \to 0.$$

Thus, for $c = 20\pi$, the number of points within each D_p is at most $c \log n$ w.h.p.

Therefore, w.h.p. there is at least one point in each smaller square, and since the farther a point can be from another point in a neighboring square is

$$\sqrt{5} \cdot (\text{side length of square}) = \sqrt{5} \cdot \sqrt{\frac{\log n}{n}} < \sqrt{10 \frac{\log n}{n}},$$

we see that every point must be connected to all of the points in its neighboring squares w.h.p. But then w.h.p. the entire graph is connected. □

The proof of the second claim is quite a bit more involved. The rough idea is to stochastically dominate the original n point model with a PPP of parameter n. Then we will show that a particular toy system of concentric discs occurs w.h.p. in the PPP model, so that by translating it back to the original model we may deduce the claim.

Definition 3.2 (Toy System)**.** Let r be such that $\pi r^2 = \frac{k+1}{n}$. The reason for this choice will be clearer through the computations. Consider a system of three concentric open discs (D_1, D_3, D_5) of radii $(r, 3r, 5r)$, respectively, such that the largest disc D_5 is contained fully in the unit square. We call the system of discs "bad" if the following hold:

(i) D_1 contains at least $k + 1$ points.

(ii) $D_3 \setminus D_1$ contains no points.

(iii) The intersection of $D_5 \setminus D_3$ with each disc of radius $1.5r$ centered at a set of points spaced evenly at distance $0.01r$ around the boundary of D_3 contains at least $k+1$ points.

Lemma 3.3

If there exists a bad system of discs among our point set, then the resulting graph G is disconnected.

Proof. Suppose our set of discs is bad. We want to show that there exist no connections between points in D_1 and points in $D_5 \setminus D_3$. The farthest any two points can be in D_1 is less than $2r$, e.g. they are diametrically opposite. On the other hand, the closest any point in D_1 can be to a point in $D_5 \setminus D_3$ is more than $3r - r = 2r$, where they must lie on the same ray from the center (we can show this using triangle inequality: let O be the center, P be a point close to ∂D_1, Q the point on the ∂D_3 on the same radius as P, and Q' any other point along ∂D_3. Then by triangle inequality, $OP + PQ' \geq OQ = 3r = OP + PQ$, so then $PQ' \geq PQ$). Therefore the closest k points to any point in D_1 all belong to D_1.

Similarly, the closest k points to any point in $D_5 \setminus D_3$ have to be at least as close as the k other points in one of the radius $1.5r$ "semicircles" its contained in, so in particular the closest k points to any point in $D_5 \setminus D_3$ are in $D_5 \setminus D_3$ itself. Thus the graph will be disconnected, since there are no edges between points in D_1 and $D_5 \setminus D_3$. $\square$

Now, this is where we introduce the PPP model. Note that the number of points in two disjoint regions are not independent in the original n point model (this is for similar reason to why the bins in the binomial model are not independent, either). However, if we consider a PPP with parameter n on the unit square, we will have the convenient property of independence for disjoint subsets. Call an event $\mathcal{E}$ an *occupancy event* if it depends only on the number of points in some fixed region of the unit square. The next result says that for occupancy events that become unlikely fast (as n grows) in the PPP model, then they will also become unlikely in the original model.

Lemma 3.4

Suppose an occupancy event $\mathcal{E}$ holds in the PPP model with probability $\delta(n) = o(1/\sqrt{n})$, then $\mathcal{E}$ holds in the original n point model with probability tending to zero as $n \to \infty$.

Proof. Let $\mathbb{P}_X$ be defined over the original n point model, and $\mathbb{P}_Y$ be defined over the Poisson Point Process model. Note that an occupancy event may be of the form

$$\mathcal{E} = \{k \text{ points lie in region } A\colon\ k \in S\},$$

where S is some subset of $\{0, 1, 2, \ldots, n\}$. For convenience, we will also use A to denote the area of the region A. In the original model, we have

$$\mathbb{P}_X[\mathcal{E}] = \sum_{k \in S} \mathbb{P}_X[k \text{ points in } A] = \sum_{k \in S} \binom{n}{k} A^k (1-A)^{n-k}.$$

For each term in the summation, we can write

$$\begin{aligned}
\binom{n}{k} A^k (1-A)^{n-k} &= \frac{n^n}{n^n} \cdot \frac{n!}{(n-k)!k!} \cdot A^k (1-A)^{n-k} \\
&= \frac{\frac{(nA)^k e^{-nA}}{k!} \cdot \frac{(n(1-A))^{n-k} e^{-n(1-A)}}{(n-k)!}}{\frac{n^n e^{-n}}{n!}} \\
&= \frac{\mathbb{P}[\mathrm{Poi}(nA) = k] \cdot \mathbb{P}[\mathrm{Poi}(n(1-A)) = n-k]}{\mathbb{P}[\mathrm{Poi}(n) = n]}.
\end{aligned}$$

Using the fact that $\mathbb{P}[\mathrm{Poi}(n) = n] \geq \frac{1}{e\sqrt{n}}$, we have

$$\begin{aligned}
\mathbb{P}_Y[\mathcal{E}] &= \sum_{k \in S} \mathbb{P}_Y[k \text{ points in } A] \\
&= \sum_{k \in S} \mathbb{P}[\mathrm{Poi}(nA) = k] \\
&\geq \mathbb{P}[\mathrm{Poi}(n) = n] \sum_{k \in S} \frac{\mathbb{P}[\mathrm{Poi}(nA) = k] \cdot \mathbb{P}[\mathrm{Poi}(n(1-A)) = n-k]}{\mathbb{P}[\mathrm{Poi}(n) = n]} \\
&= \mathbb{P}[\mathrm{Poi}(n) = n] \mathbb{P}_X[\mathcal{E}] \\
&\geq \frac{1}{e\sqrt{n}} \mathbb{P}_X[\mathcal{E}].
\end{aligned}$$

Thus, if $\mathbb{P}_Y[\mathcal{E}] \leq \delta(n) \ll \frac{1}{\sqrt{n}}$, then

$$\mathbb{P}_X[\mathcal{E}] \leq e\sqrt{n}\mathbb{P}_Y[\mathcal{E}] = O(\sqrt{n}) \cdot o(1/\sqrt{n}) \to 0$$

as $n \to \infty$. □

We are now ready to prove claim 2 of Theorem 3.1.

Proof of 2. We will first show that the probability in the PPP model that any particular toy system of discs is bad is $\Omega(n^{-1+\epsilon})$, for some $\epsilon > 0$, then use this to compute the probability there exists a bad system throughout the entire unit square.

For a particular system of discs to be bad, it must satisfy each of the three properties. Due to PPP, the probabilities of each of the three properties are multiplicative since events of disjoint regions are independent. So we may compute the probabilities of the three events separately.

1. The area of D_1 is $\pi r^2 = \frac{k+1}{n}$, so we have

   $$\mathbb{P}[\mathrm{Poi}(k+1) \geq k+1],$$

   which from (6) we know to be at least $1/4$.

2. The area of $D_3 \setminus D_1$ is $9\pi r^2 - \pi r^2 = 8\pi r^2 = 8 \cdot \frac{k+1}{n}$. So we have

   $$\mathbb{P}[\mathrm{Poi}(8(k+1)) = 0] = \frac{(8(k+1))^0 e^{-8(k+1)}}{0!} = e^{-8(c_2 \log n + 1)} = n^{-8c_2} e^{-8}.$$

   Now, fix a small $\epsilon > 0$. We can take $c_2 > 0$ small enough so that

   $$n^{-1+\epsilon} \leq n^{-8c_2} e^{-8}.$$

   Then the probability above is at least $n^{-1+\epsilon}$.

3. There are $\frac{6\pi r}{0.01r} = 600\pi \le 1885$ circles spaced around the boundary of D_3. We compute the probability that one of the "semicircles" (the intersection of one of the discs of radius $1.5r$ with $D_5 \setminus D_3$) contains less than $k+1$, and then take a union bound. In particular, note that the mean number of points lying in the "semicircle" is at least the mean number of points lying in the semicircle contained in the "semicircle", whose diameter is tangent to ∂D_3. The latter mean is distributed as $\mathrm{Poi}(n \cdot \pi(1.5r)^2/2)$, so the mean number of points lying in the "semicircle" is at least $n \cdot \frac{\pi(1.5r)^2}{2} = 1.125(k+1)$. So if X is the number of points in our "semicircle" region, then by a Chernoff bound from Corollary 1.3, we get

$$\begin{aligned}\mathbb{P}[X < k+1] &= \mathbb{P}[X < 1.125(k+1) - 0.125(k+1)] \\ &\le \exp\{(k+1)\ln(1.125) - 0.125(k+1)\} \\ &\le \exp\{-0.007(k+1)\}.\end{aligned}$$

So the probability that some "semicircle" has less than $k+1$ points can be computed by a union bound, and is less than

$$1 - 1885e^{-0.007(k+1)}.$$

And, as $n \to \infty$, we know that $k = c_2 \log n \to \infty$, so this quantity approaches 1. Therefore, we may write the probability that condition (iii) is satisfied as $1 - o(1)$.

Putting this all together, the probability all three properties of a bad set are satisfied is given by

$$\mathbb{P}[(D_1, D_3, D_5) \text{ are bad}] \ge \left(\frac{1}{4}\right)\left(n^{-1+\epsilon}\right)(1 - o(1)) = \Omega(n^{-1+\epsilon}).$$

Now, note that we can fit $O\left(\frac{n}{\log n}\right)$ disjoint discs of radius $5r$ into the unit square. For each of these discs, we have a set of 3 concentric discs which is bad independently with probability at least $\Omega(n^{-1+\epsilon})$. So, the probability that there is no bad set is at most the probability that none of these sets are bad, which is (for constants c, c', c'')

$$\left(1 - \Omega(n^{-1+\epsilon})\right)^{O\left(\frac{n}{\log n}\right)} \le \left(1 - \frac{c' n^\epsilon}{n}\right)^{\frac{c'' n}{\log n}} \le \exp\{-c' n^\epsilon c'' / \log n\} = \exp\{-c n^\epsilon / \log n\}.$$

So, the probability that there are no bad sets is $o(1/\sqrt{n})$ in the PPP model. By Lemma 3.4 this implies that the probability there are no bad sets in the original model goes to zero as $n \to \infty$. Thus G is connected w.h.p. as $n \to \infty$. □