Convergence in distribution for subset counts between random sets

Dudley Stark
School of Mathematical Sciences
Queen Mary, University of London
London E1 4NS, United Kingdom
D.S.Stark@maths.qmul.ac.uk

Submitted: Jan 8, 2003; Accepted: Aug 27, 2004; Published: Sep 9, 2004
Mathematics Subject Classifications: 60C05, 60F05

Abstract

Erdős posed the problem of how many random subsets need to be chosen from a set of n elements, each element appearing in each subset with probability $p = 1/2$, in order that at least one subset is contained in another. Rényi answered this question, but could not determine the limiting probability distribution for the number of subset counts because the higher moments diverge to infinity.

The model considered by Rényi with p arbitrary is denoted by $\mathcal{P}(m, n, p)$, where m is the number of random subsets chosen. We give a necessary and sufficient condition on $p(n)$ and $m(n)$ for subset counts to be asymptotically Poisson and find rates of convergence using Stein's method. We discuss how Poisson limits can be shown for other statistics of $\mathcal{P}(m, n, p)$.

1 Introduction

Erdős posed the following problem which Rényi [5] solved. Subsets $S_1, S_2, \ldots, S_m$ are chosen randomly from the set $[1, n] := \{1, 2, 3, \ldots, n\}$, where $m, n \geq 1$. For each $r \in [1, n]$ and $i \in [1, m]$, the event $A_{r,i} := \{r \in S_i\}$ has probability $\mathbb{P}(A_{r,i}) = 1/2$ and the $A_{r,i}$ are mutually independent. How large does $m = m(n)$ need to be so that the probability approaches 1 that $S_i \subseteq S_j$ for some pair $i, j \in [1, m]$, $i \neq j$?

The model studied by Rényi with $\mathbb{P}(A_{r,i}) = p$ will be denoted by $\mathcal{P}(m, n, p)$ and may be considered to be a model of a random Boolean lattice when sets S_i containing identical elements are identified. A different random lattice model has been studied recently (in [3, 4], for example) in which each of the possible 2^n subsets are present independently and with probability p.

Let X be the number of pairs (S_i, S_j), $1 \leq i < j \leq n$, for which either $S_i \subseteq S_j$ or $S_j \subseteq S_i$. If we define $I_{(i,j)}$, $i, j \in [1, n]$, $i \neq j$, by $I_{(i,j)} := I[S_i \subseteq S_j]$, where $I[\cdot]$ equals 1 if and only if the expression in brackets is true and otherwise equals 0, then

$$\begin{aligned} X &= \sum_{1\le i\ne j\le n} I_{(i,j)} - \sum_{1\le i<j\le n} I\left[S_i = S_j\right] && (1)\\ &= W - \sum_{1\le i<j\le n} I\left[S_i = S_j\right], && (2)\end{aligned}$$

where

$$W = \sum_{1\le i\ne j\le n} I_{(i,j)}. \tag{3}$$

Let ε_r^i be the indicator variable of the event $A_{r,i}$. For each $i, j \in [1, m]$ and $r \in [1, n]$, in the model studied by Rényi we have

$$\mathbb{P}(\varepsilon_r^i \le \varepsilon_r^j) = \mathbb{P}(\varepsilon_r^i = 0) + \mathbb{P}(\varepsilon_r^i = 1)\mathbb{P}(\varepsilon_r^j = 1) = 3/4.$$

The expectation $\mathbb{E}X$ may be calculated by noting that

$$\mathbb{E}I_{(i,j)} = \mathbb{P}\left(\bigcap_{r\in[1,n]} \{\varepsilon_r^i \le \varepsilon_r^j\}\right) = (3/4)^n$$

and

$$\mathbb{E}I[X_i = X_j] = (1/2)^n$$

give

$$\mathbb{E}X = m(m-1)(3/4)^n - \binom{m}{2}(1/2)^n \sim m^2(3/4)^n. \tag{4}$$

Rényi showed that for any fixed $c > 0$, if $m = m(n)$ satisfies

$$m \sim c\left(\frac{2}{\sqrt{3}}\right)^n, \tag{5}$$

which by (4) is equivalent to $\lim_{n\to\infty} \mathbb{E}X = c^2$, then

$$\mathbb{P}(X \ge 1) = 1 - e^{-c^2}.$$

In the model $\mathcal{P}(m, n, p)$, (1) becomes, with $q = 1 - p$,

$$\mathbb{E}X = m(m-1)(q+p^2)^n - \binom{m}{2}(p^2+q^2)^n \tag{6}$$

and when p is fixed $\mathbb{E}X$ converges to c^2 iff m satisfies

$$m \sim c\,(q+p^2)^{-n/2}. \tag{7}$$

It is natural to suppose that the distribution of X would be approximately Poisson. Rényi used sieve methods for his results and was not able to prove a Poisson limit for

X because higher moments than the fourth diverge to ∞. Poisson limits were shown, however, for the probabilities $\mathbb{P}(X = k)$ when $k \leq 3$.

If p is fixed, then the argument in [5] extends easily to show that all moments of X of a high enough order diverge to ∞ when the first moment converges. Suppose that m satisfies (7). The τth moment of $\mathbb{E}X^\tau$ is bounded below by

$$\begin{aligned}
\mathbb{E}X^\tau &\geq \sum_{i_1 < i_2 \cdots < i_\tau < j} \mathbb{P}\left(A_{i_1,j} \cap A_{i_2,j} \cap \cdots \cap A_{i_\tau,j}\right) \\
&\geq \binom{m}{\tau+1} \mathbb{P}\left(\bigcap_{r=1}^{n} \{\varepsilon_r^j = 1\}\right) \\
&\sim \frac{m^{\tau+1}}{(\tau+1)!} p^n \\
&\sim \frac{c^{\tau+1}}{(\tau+1)!} \left(p(q+p^2)^{-(\tau+1)/2}\right)^n .
\end{aligned}$$

Thus, the τth moment diverges whenever $\tau > 2\frac{\log p}{\log(q+p^2)} - 1$.

The nonconvergence of moments indicates that it is not possible to get good approximation results for subset counts between random sets by using sieve methods. We use Stein's method to show the convergence to the Poisson distribution. Erdős' problem is thus a natural example where sieve methods fail to show convergence in distribution and Stein's method succeeds. Stein's method has the advantage that it gives a rate of convergence and gives Poisson approximation bounds even when moments do not converge. For a comprehensive account of Stein's method see [1]. In applying Stein's method to subset counts in $\mathcal{P}(m, n, p)$ we were able to use the "coupling" version of Stein's method and consequently were able to obtain rates of convergence in a straightforward way by calculating certain covariances. This was not possible with other statistics of $\mathcal{P}(m, n, p)$ analysed by Rényi, for which it seems necessary to apply the "local" version of Stein's method.

The total variation distance between the distributions of two random variables X_1, X_2 defined on a finite or countable state space S is defined to be

$$d_{\mathrm{TV}}(\mathcal{L}(X_1), \mathcal{L}(X_2)) = \frac{1}{2} \sum_{s \in S} \left|\mathbb{P}(X_1 = s) - \mathbb{P}(X_2 = s)\right| . \tag{8}$$

It is well known that

$$d_{\mathrm{TV}}(\mathcal{L}(X_1), \mathcal{L}(X_2)) = \min_{\text{couplings}} \mathbb{P}(X_1 \neq X_2), \tag{9}$$

where the minimum is taken over all couplings of X_1 and X_2 on the same probability space.

Theorem 1 *Suppose that X and W are defined as (1) and (3). Let*

$$\lambda := \mathbb{E}W = \binom{m}{2}(1 - qp)^n.$$

Under these assumptions,

$$\begin{aligned}
&d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \\
&\leq \frac{1-e^{-\lambda}}{\lambda}\left(\left[\binom{m}{2}+1\right](1-qp)^{2n} + (m-2)\left[(p^3+q)^n + (p+q^3)^n\right]\right. \\
&\quad \left. -(2m-3)(q^2+p^2)^n\right).
\end{aligned}$$

and

$$\begin{aligned}
&d_{\mathrm{TV}}(\mathcal{L}(X), \mathrm{Po}(\lambda)) \\
&\leq \frac{1-e^{-\lambda}}{\lambda}\left(\left[\binom{m}{2}+1\right](1-qp)^{2n} + (m-2)\left[(p^3+q)^n + (p+q^3)^n\right]\right. \\
&\quad \left. +\left[\binom{m}{2}-(2m-3)\right](q^2+p^2)^n\right).
\end{aligned}$$

It follows that if $p = p(n)$ and $m = m(n)$ are chosen in such a way that $\mathbb{E}X$ converges, then X is asymptotically Poisson if and only if simultaneously $np \to \infty$ and $n(1-p) \to \infty$ as $n \to \infty$.

In Section 2 we use Stein's method to prove Theorem 1. In Section 3 we discuss the Poisson approximation of other statistics of $\mathcal{P}(m, n, p)$ considered by Rényi.

2 A coupling for subset counts between random sets

In this section we prove Theorem 1.

It is convenient initially to work with the random variable W. Note that (1) implies

$$\begin{aligned}
\mathbb{P}(W \neq X) &= \mathbb{P}\left(\bigcup_{i<j}\{S_i = S_j\}\right) \\
&\leq \binom{m}{2}\mathbb{P}(S_1 = S_2) = \binom{m}{2}\left(p^2+q^2\right)^n. \qquad (10)
\end{aligned}$$

Suppose that Γ is a finite or countable index set and that $\{I_\alpha : \alpha \in \Gamma\}$ are indicator variables, possibly dependent. Let W denote $W = \sum_{\alpha\in\Gamma} I_\alpha$. We set $\Gamma_\alpha = \Gamma \setminus \{\alpha\}$ and suppose that Γ_α can be decomposed into disjoint sets Γ_α^+, Γ_α^-, Γ_α^0 such that $\Gamma_\alpha = \Gamma_\alpha^+ \cup \Gamma_\alpha^+ \cup \Gamma_\alpha^0$ which have certain properties. We suppose that for each $\alpha \in \Gamma$ that there exist random variables $(J_{\beta,\alpha},\, \beta \in \Gamma)$ defined on the same probability space as $(I_\beta,\, \beta \in \Gamma)$ with

$$\mathcal{L}(J_{\beta,\alpha},\, \beta \in \Gamma) = \mathcal{L}(I_\beta,\, \beta \in \Gamma | I_\alpha = 1).$$

Moreover, we assume that

$$J_{\beta,\alpha}\begin{cases} \leq I_\beta & \text{if } \beta \in \Gamma_\alpha^- \\ \geq I_\beta & \text{if } \beta \in \Gamma_\alpha^+. \end{cases}$$

The random variables $(I_\beta,\, \beta \in \Gamma_\alpha^-)$ are said to be *negatively related* to I_α The random variables $(I_\beta,\, \beta \in \Gamma_\alpha^+)$ are said to be *positively related* to I_α. Let $\pi_\alpha = \mathbb{E}I_\alpha$ and let $\lambda = \mathbb{E}W$. Then Theorem 2.C of [1] gives the total variation distance bound

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \;&\le\; \frac{1-e^{-\lambda}}{\lambda}\left(\sum_{\alpha\in\Gamma}\pi_\alpha^2 + \sum_{\alpha\in\Gamma}\sum_{\beta\in\Gamma_\alpha^-}|\mathrm{Cov}(I_\alpha, I_\beta)|\right. \\ &\quad\left. +\; \sum_{\alpha\in\Gamma}\sum_{\beta\in\Gamma_\alpha^+}\mathrm{Cov}(I_\alpha, I_\beta) + \sum_{\alpha\in\Gamma}\sum_{\beta\in\Gamma_\alpha^0}(\mathbb{E}I_\alpha I_\beta + \pi_\alpha\pi_\beta)\right) \qquad (11)\end{aligned}$$

We will next construct the couplings needed to apply (11) to the problem of approximating the number of subset counts in $\mathcal{P}(m,n,p)$. We can express W as $W = \sum_{(i,j)\in\Gamma} I_{(i,j)}$ where

$$\Gamma = \{(i,j) : i \in [1,m], j \in [1,m], i \neq j\}.$$

The equivalent to Γ_α in this setting is

$$\Gamma_{(i,j)} = \{(k,l) \in \Gamma : (k,l) \neq (i,j)\}.$$

Now, define

$$\Gamma_{(i,j)}^- = \{(k,l) \in \Gamma_{(i,j)} : \{k,l\}\cap\{i,j\} = \emptyset\}\cup\{(k,l) \in \Gamma_{(i,j)} : l = i\}\cup\{(k,l) \in \Gamma_{(i,j)} : k = j\},$$

$$\Gamma_{(i,j)}^+ = \{(k,l) \in \Gamma_{(i,j)} : k = i\} \cup \{(k,l) \in \Gamma_{(i,j)} : l = j\},$$

$$\Gamma_{(i,j)}^0 = \emptyset.$$

Clearly, $\Gamma_{(i,j)} = \Gamma_{(i,j)}^- \cup \Gamma_{(i,j)}^+ \cup \Gamma_{(i,j)}^0$. It will be shown that the indicator variables indexed by $\Gamma_{(i,j)}^-$ are negatively related to $I_{(i,j)}$ and that the indicators indexed by $\Gamma_{(i,j)}^+$ are positively related to $I_{(i,j)}$.

We will now define the coupling defining $J_{(k,l),(i,j)}$ for $(k,l) \in \Gamma_{(i,j)}$. Observe that, conditional on $I_{(i,j)} = 1$, each $(\varepsilon_r^i, \varepsilon_r^j)$, $r \in [1,n]$, equals one of $(0,0), (0,1)$ or $(1,1)$ with the following probabilities

$$\mathbb{P}((\varepsilon_r^i, \varepsilon_r^j) = (0,0)) \;=\; \frac{q^2}{1-qp}, \qquad (12)$$

$$\mathbb{P}((\varepsilon_r^i, \varepsilon_r^j) = (0,1)) \;=\; \frac{qp}{1-qp}, \qquad (13)$$

$$\mathbb{P}((\varepsilon_r^i, \varepsilon_r^j) = (1,1)) \;=\; \frac{p^2}{1-qp}. \qquad (14)$$

Given a realization of the ε_r^i, we construct $J_{(k,l),(i,j)}$ by choosing new values of the ε_r^l, which we denote by $\tilde\varepsilon_r^l$, for each $l \in [1,m]$ and $r \in [1,n]$. If $l \notin \{i,j\}$, then we set $\tilde\varepsilon_r^l = \varepsilon_r^l$. If $(\varepsilon_r^i, \varepsilon_r^j) \in \{(0,0),(0,1),(1,1)\}$, then we set $(\tilde\varepsilon_r^i, \tilde\varepsilon_r^j) = (\varepsilon_r^i, \varepsilon_r^j)$. If $(\varepsilon_r^i, \varepsilon_r^j) = (1,0)$, then choose $(\tilde\varepsilon_r^i, \tilde\varepsilon_r^j)$ to be one of $(0,0)$, $(0,1)$, or $(1,1)$ randomly and with probabilities given by

(12), (13), and (14). We let $J_{(k,l),(i,j)} = 1$ if $\tilde{\varepsilon}_r^k \le \tilde{\varepsilon}_r^l$ for all $r \in [1,n]$ and let $J_{(k,l),(i,j)} = 0$ otherwise.

We will show that $J_{(k,l),(i,j)} \le I_{(k,l)}$ for each $(k,l) \in \Gamma^-_{(i,j)}$. Clearly $J_{(k,l),(i,j)} = I_{(k,l)}$ for all (k,l) such that $\{k,l\} \cap \{i,j\} = \emptyset$. The way the coupling is defined implies that $\tilde{\varepsilon}_r^i \le \varepsilon_r^i$ and $\tilde{\varepsilon}_r^j \ge \varepsilon_r^j$ for all r, i, j. Therefore, we have $J_{(k,i),(i,j)} \le I_{(k,i)}$. In the same way it follows that $J_{(j,l),(i,j)} \le I_{(j,l)}$.

An analogous argument shows that $J_{(k,l),(i,j)} \ge I_{(k,l)}$ for all $(k,l) \in \Gamma^+_{(i,j)}$, hence the requirements on the indicator sets Γ^-_α, Γ^+_α, and Γ^0_α in (11) have been shown to be satisfied and we now proceed with calculating the covariances appearing therein.

If $\{k,l\} \cap \{i,j\} = \emptyset$, then

$$\operatorname{Cov}(I_{(k,l)}, I_{(i,j)}) = 0. \tag{15}$$

Suppose now that $k \in [1,m] \setminus \{i\}$. We have $\operatorname{Cov}(I_{(k,i)}, I_{(i,j)}) = \mathbb{E}(I_{(k,i)} I_{(i,j)}) - (1-qp)^{2n}$. It happens that $I_{(k,i)} I_{(i,j)} = 1$ if and only if $(\varepsilon_r^k, \varepsilon_r^i, \varepsilon_r^j)$ takes on one of the values $(0,0,0)$, $(0,0,1)$, $(0,1,1)$, $(1,1,1)$ for each r. Hence,

$$\mathbb{E}(I_{(k,i)} I_{(i,j)}) = (q^3 + q^2 p + q p^2 + p^3)^n = (q^2 + p^2)^n$$

and

$$\operatorname{Cov}(I_{(k,i)}, I_{(i,j)}) = (p^2 + q^2)^n - (1-qp)^{2n}. \tag{16}$$

Similarly,

$$\operatorname{Cov}(I_{(j,l)}, I_{(i,j)}) = (p^2 + q^2)^n - (1-qp)^{2n}, \tag{17}$$

for all $l \in [1,m] \setminus \{j\}$. Note that both (16) and (17) include the case $k = j$, $l = i$, so there are $(m-1)+(m-1)-1 = 2m-3$ terms that contribute covariances of the form (16) and (17). It is easily checked that $(p^2+q^2)^n \le (1-qp)^{2n}$. The covariances for $(k,l) \in \Gamma^+_{(i,j)}$ equal, for $k,l \in [1,m] \setminus \{i,j\}$,

$$\operatorname{Cov}(I_{(i,l)}, I_{(i,j)}) = (p^3 + q)^n - (1-qp)^{2n}. \tag{18}$$

$$\operatorname{Cov}(I_{(k,j)}, I_{(i,j)}) = (p + q^3)^n - (1-qp)^{2n}. \tag{19}$$

There are $m-2$ terms which contribute covariances of the form (18) and $m-2$ which contribute covariances of the form (19). The covariances for $(k,l) \in \Gamma^+_{(i,j)}$ are all nonnegative.

Substituting the covariances (15) through (19) in (11) gives

$$\begin{aligned}
&d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \\
&\le \frac{1-e^{-\lambda}}{\lambda} \left(\binom{m}{2} (1-qp)^{2n} + (2m-3)\left[(1-qp)^{2n} - (p^2+q^2)^n\right] \right. \\
&\qquad \left. +(m-2)\left[(p^3+q)^n - (1-qp)^{2n}\right] + (m-2)\left[(p+q^3)^n - (1-qp)^{2n}\right] \right) \\
&= \frac{1-e^{-\lambda}}{\lambda} \left(\left[\binom{m}{2} + 1\right] (1-qp)^{2n} + (m-2)\left[(p^3+q)^n + (p+q^3)^n\right] \right. \\
&\qquad \left. -(2m-3)(q^2+p^2)^n \right).
\end{aligned}$$

By (10), (8) and (9), we have

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(X), \mathrm{Po}(\lambda)) &\leq d_{\mathrm{TV}}(\mathcal{L}(X), \mathcal{L}(W)) + d_{\mathrm{TV}}(\mathcal{L}(W)), \mathrm{Po}(\lambda)) \\ &\leq \binom{m}{2}(p^2+q^2)^n + d_{\mathrm{TV}}(\mathcal{L}(W)), \mathrm{Po}(\lambda)). \end{aligned}$$

Since $(p^2+q^2)^n \leq (1-qp)^{2n}$, the bound on $d_{\mathrm{TV}}(\mathcal{L}(X), \mathrm{Po}(\lambda))$ is of the same order as the bound on $d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda))$. If $\binom{m}{2}(1-qp)^{2n} = \lambda(1-p+p^2)^n = o(1)$, then $(m-2)(p^3+q)^n \leq m(1-p+p^3)^n = o(1)$ Similarly $(m-2)(p+q^3)^n = o(1)$ and $(2m-3)(p^2+q^2)^n = o(1)$. The factor $(1-e^{-\lambda})/\lambda$ is bounded above by $\min(\lambda^{-1}, 1) \leq 1$. Thus, the total variation distances in Theorem 1 converge to 0 as long as $\lambda(1-p+p^2)^n = o(1)$. This condition holds if, for example, λ is bounded and p can be written as $p = \omega_1(n)/n$ and $p = 1 - \omega_2(n)/n$ where $\omega_1(n) \to \infty$ and $\omega_2(n) \to \infty$ as $n \to \infty$.

The fact that the range of p for which X has a Poisson limit when λ converges cannot be extended beyond intervals of the form $[\omega_1(n)/n, 1-\omega_2(n)/n]$ is shown considering $p = c/n$ with c constant and fixing m. The distribution of X cannot converge weakly to a Poisson distribution because $X \leq m^2$ and a Poisson distributed variable is unbounded. The actual limiting distribution of X can be found by the following argument. Let $Y = |\{i \in [1,m] : S_i = \emptyset\}|$. Then Y is asymptotically Binomial(m, e^{-c}) distributed. The expected number of elements of $[1,n]$ occurring in more than one S_i is $n(1-q^m - mpq^{m-1}) = O(n^{-1}) = o(1)$, hence by the first moment method the random variable X asymptotically almost surely equals $\binom{Y}{2} + Y(m-Y)$ which does not have a Poisson limit by the observation above. A similar remark may be shown for $p = 1-c/n$ by redefining Y as $Y = |\{i \in [1,m] : S_i = [1,n]\}|$ and considering the number of elements of $[1,n]$ occurring in at most $m-2$ of the S_i.

This completes the proof of Theorem 1.

3 Other statistics of randomly chosen sets

Rényi [5] considered other statistics of $\mathcal{P}(m,n,p)$. For example, he considered the number of triples (i,j,k) for which $S_k = S_i \cup S_j$; the number of triples (i,j,k) for which $S_k = S_i \cap S_j$; and the number of r-tuples $(i_1, i_2, \ldots, i_r)$ for which $S_{i_1} \subseteq S_{i_2} \subseteq \cdots \subseteq S_{i_r}$. These results were extended by Bognár [2] to a general theory of relations on $\mathcal{P}(m,n,p)$.

We could not find a direct coupling for general relations of $\mathcal{P}(m,n,p)$ as was done in Section 2 for subset counts. It is possible, however, to apply the "local" version of Stein's method, which follows from the "coupling" version (see Corollary 2.C.5 of [1]). We indicate how this is done by a sketch of an application of the local version of Stein's method to the number of triples (i,j,k) for which $S_k = S_i \cup S_j$.

In the local version of Stein's method, for each $\alpha \in \Gamma$ sets $\Gamma_\alpha^{\mathrm{s}}$ and $\Gamma_\alpha^{\mathrm{w}}$ are defined such that $\Gamma = \Gamma_\alpha^{\mathrm{s}} \cup \Gamma_\alpha^{\mathrm{w}}$ in such a way that I_α is not very dependent on the indicators

$\{I_\beta; \beta \in \Gamma_\alpha^{\mathrm{w}}\}$. Theorem 1.A of [1] then gives the bound

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda) &\leq \min(1, \lambda^{-1}) \sum_{\alpha \in \Gamma} \left(p_\alpha^2 + p_\alpha \mathbb{E} Z_\alpha + \mathbb{E}(I_\alpha Z_\alpha)\right) \\ &\qquad + \min(1, \lambda^{-1/2}) \sum_\alpha \eta_\alpha, \end{aligned}$$

where

$$p_\alpha = \mathbb{E} I_\alpha,$$
$$Z_\alpha = \sum_{\beta \in \Gamma_\alpha^{\mathrm{s}}} I_\beta,$$

and

$$\eta_\alpha = \mathbb{E}\left|\mathbb{E}\{I_\alpha | (I_\beta, \beta \in \Gamma_\alpha^{\mathrm{w}}\} - p_\alpha\right|.$$

If I_α is independent of the indicators $(I_\beta, \beta \in \Gamma_\alpha^{\mathrm{w}})$, then $\eta_\alpha = 0$ and

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda) \leq \min(1, \lambda^{-1}) \sum_{\alpha \in \Gamma} \left(p_\alpha^2 + p_\alpha \mathbb{E} Z_\alpha + \mathbb{E}(I_\alpha Z_\alpha)\right). \tag{20}$$

Consider the indicators $I_{i,j;k}$ where i, j are unordered. There are $\binom{n}{2}(n-2)$ such indicators in total. We have $I_{i,j;k} = 1$ if and only if for each $\tau \in [1, n]$, exactly one of the following four options occurs:

1) $\tau \notin S_i$, $\tau \notin S_j$, $\tau \notin S_k$
2) $\tau \in S_i$, $\tau \notin S_j$, $\tau \in S_k$
3) $\tau \notin S_i$, $\tau \in S_j$, $\tau \in S_k$
4) $\tau \in S_i$, $\tau \in S_j$, $\tau \in S_k$

(This is the normal disjunctive form decomposition of the relationship $S_i \cup S_j = S_k$ used in [2].) Thus, $\mathbb{E} I_{1,2;3} = (q^3 + 2qp^2 + p^3)^n$, which is the p_α in (20). We define

$$\Gamma_{i,j;k}^{\mathrm{w}} = \{(l_1, l_2; l_3) : l_1, l_2, l_3 \notin \{i, j, k\}\} \quad \text{and} \quad \Gamma_{i,j;k}^{\mathrm{s}} = \Gamma \setminus \Gamma_{i,j;k}^{\mathrm{w}}.$$

For the analysis in this paragraph we will assume that λ converges and that $p = q = 1/2$. The indicators indexed by $\Gamma_\alpha^{\mathrm{w}}$ are clearly mutually independent of $I_{i,j;k}$. The first two terms in (20) vanish asymptotically. The last term is $\min(1, \lambda^{-1}) \sum_{\alpha \in \Gamma} \mathbb{E}(I_\alpha Z_\alpha) \leq \mathbb{E}(Z_{1,2;3} | I_{1,2;3} = 1)$. The main contribution to the last expression comes from the elements of $\Gamma_\alpha^{\mathrm{s}}$ of the form $(l_1, l_2; k)$, $l_1, l_2 \notin \{i, j, k\}$. It is easy to check that $\mathbb{P}(I_{l_1,l_2;k} = 1 | I_{i,j;k} = 1) = (5/8)^n$. Under our assumption that λ converges and $p = q = 1/2$, $m \asymp 2^{n/3}$. Hence $d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda) = O\left((2^{2/3} 5/8)^n\right) = o(1)$.

Clearly, similar calculations could be done for other Boolean relations on $\mathcal{P}(m, n, p)$.

References

[1] Barbour, A. D., Holst, L., Janson, S. *Poisson Approximation*, Oxford University Press, Oxford, 1992.

[2] Bognár, K. On random sets. *Magyar Tud. Akad. Mat. Kutató Int. Köł.* **7** (1961) 425–440.

[3] Kohayakawa, Y., Kreuter, B., Osthus, D. The length of random subsets of Boolean lattices. *Random Structures Algorithms* **16** (2000), 177–194.

[4] Osthus, D., Maximum antichains in random subsets of a finite set. *J. Combin. Theory Ser. A* **90** (2000), 336–346.

[5] Rényi, A., Egy általános módszer valószínűségszámítási tételek bizonyítására és annak néhány alkalmazása. *MTA III. Oszt. Közl.* **11** (1961) 79–105. (Translated into English in *Selected Papers of Alfréd Rényi*, vol. 2, pp. 581–602, Budapest: Akadémiai Kiadoó.)