Stein Chen Method for Poisson Approximation

Suppose $A_1, \dots, A_k$ are (rare) events and $X := \sum_{i=1}^{k} \mathbb{1}_{A_i}$ counts the number of these events that occur. Suppose also that they are 'mostly' independent, and let D be a dependency digraph for the family $\{A_1, \dots, A_k\}$. We would like to show that, under mild conditions, the distribution of X is close to that of a Poisson distribution.

For ease of notation, write

$$p_i := \mathbb{P}(A_i), \qquad p_{ij} := \mathbb{P}(A_i \cap A_j), \qquad \mu := \mathbb{E}[X] = \sum_{i=1}^{k} p_i,$$

$$\Delta := \sum_{i,j\,:\,i\to j} p_{ij} \qquad \text{and} \qquad S := \sum_{i,j\,:\,i\to j \text{ or } i=j} p_i p_j,$$

where $i \to j$ indicates that the directed edge from i to j lies in the dependency digraph D, so each A_i is independent of the family $\{A_j : j \neq i,\, i \not\to j\}$.

Theorem 1 (Stein–Chen). *With the above notation and for any $A \subseteq \mathbb{N}$,*

$$\big|\mathbb{P}(X \in A) - \mathbb{P}(Z \in A)\big| \leqslant \tfrac{1-e^{-\mu}}{\mu}\big(S + \Delta\big),$$

where $Z \sim \mathrm{Po}(\mu)$ is a Poisson random variable with mean μ.

Remark. We note that $\frac{1-e^{-\mu}}{\mu} \leqslant 1$ for any $\mu > 0$, so a simpler but slightly weaker form would be to omit this factor.

Before we start the proof, we note that for $Z \sim \mathrm{Po}(\mu)$ and $n \geqslant 1$,

$$n\,\mathbb{P}(Z = n) = \mu\,\mathbb{P}(Z = n-1). \tag{1}$$

Indeed,

$$n\,\mathbb{P}(Z = n) = ne^{-\mu}\frac{\mu^n}{n!} = \mu e^{-\mu}\frac{\mu^{n-1}}{(n-1)!} = \mu\,\mathbb{P}(Z = n-1).$$

Proof. Fix $A \subseteq \mathbb{N}$ and introduce a function $g = g_A$ by defining $g_A(0) := 0$ and

$$g_A(n) := \frac{\mathbb{P}(Z \in A \text{ and } Z < n) - \mathbb{P}(Z \in A)\mathbb{P}(Z < n)}{n\mathbb{P}(Z = n)}$$

for $n > 0$. Now, by (1),

$$\begin{aligned}\mu g(n+1)\mathbb{P}(Z = n) &= g(n+1)\cdot(n+1)\mathbb{P}(Z = n+1)\\ &= \mathbb{P}(Z \in A, Z \leqslant n) - \mathbb{P}(Z \in A)\mathbb{P}(Z \leqslant n),\end{aligned}$$

so

$$\begin{aligned}(\mu g(n+1) - ng(n))\mathbb{P}(Z = n) &= \big(\mathbb{P}(Z \in A, Z \leqslant n) - \mathbb{P}(Z \in A)\mathbb{P}(Z \leqslant n)\big)\\ &\quad - \big(\mathbb{P}(Z \in A, Z < n) - \mathbb{P}(Z \in A)\mathbb{P}(Z < n)\big)\\ &= \mathbb{P}(Z \in A, Z = n) - \mathbb{P}(Z \in A)\mathbb{P}(Z = n)\\ &= (\mathbb{1}_{\{n\in A\}} - \mathbb{P}(Z \in A))\mathbb{P}(Z = n).\end{aligned}$$

Hence $\mu g(n+1) - ng(n) = \mathbb{1}_{\{n\in A\}} - \mathbb{P}(Z \in A)$. Now substituting in X and taking expectations we get

$$\mathbb{E}[\mu g(X+1) - Xg(X)] = \mathbb{P}(X \in A) - \mathbb{P}(Z \in A).$$

Write $X_S = \sum_{i\in S} \mathbb{1}_{A_i}$ (and $X_i = X_{\{i\}} = \mathbb{1}_{A_i}$). Then

$$\mathbb{P}(X \in A) - \mathbb{P}(Z \in A) = \mathbb{E}[\mu g(X+1) - Xg(X)] = \sum_i \mathbb{E}[p_i g(X+1) - X_i g(X)].$$

Now fix i and write $D = \{j : i \to j\} = \{j_1, j_2, \ldots, j_d\}$ and $I = \{j \neq i : i \not\to j\}$ so that $[k]$ is the disjoint union of I, D and $\{i\}$ and $X = X_{I\cup D\cup\{i\}}$. By considering the cases $X_j = 0$ or $X_j = 1$ we have $g(Y + X_j) - g(Y) = X_j(g(Y+1) - g(Y))$ for any Y and j. Hence we can write

$$\begin{aligned}
p_i g(X+1) - X_i g(X) &= p_i(g(X_{I\cup D\cup\{i\}} + 1) - g(X_{I\cup D} + 1)) + (p_i - X_i)g(X_{I\cup D} + 1) \\
&= p_i(g(X_{I\cup D\cup\{i\}} + 1) - g(X_{I\cup D} + 1)) \\
&\quad + \sum_{r=1}^{d} (p_i - X_i)(g(X_{I\cup\{j_1,\ldots,j_r\}} + 1) - g(X_{I\cup\{j_1,\ldots,j_{r-1}\}} + 1) \\
&\quad + (p_i - X_i)g(X_I + 1) \\
&= p_i X_i(g(X_{I\cup D} + 2) - g(X_{I\cup D} + 1)) \\
&\quad + \sum_{r=1}^{d} (p_i - X_i)X_{j_r}(g(X_{I\cup\{j_1,\ldots,j_{r-1}\}} + 2) - g(X_{I\cup\{j_1,\ldots,j_{r-1}\}} + 1)) \\
&\quad + (p_i - X_i)g(X_I + 1).
\end{aligned}$$

Now assume there is a constant c such that for all $n \geqslant 1$, $|g(n+1) - g(n)| \leqslant c$. Then, taking expectations and noting that $g(X_I + 1)$ is independent of X_i, we have

$$\begin{aligned}
\big|\mathbb{E}[p_i g(X+1) - X_i g(X)]\big| &\leqslant cp_i\mathbb{E}[X_i] + \sum_{r=1}^{k} c\mathbb{E}[|(p_i - X_i)X_{d_r}|] \\
&\leqslant cp_i^2 + \sum_{j\colon i\to j} c(p_i p_j + p_{ij}).
\end{aligned}$$

Summing over i then gives

$$|\mathbb{P}(X \in A) - \mathbb{P}(Z \in A)| \leqslant c(S + \Delta).$$

The result now follows from the following lemma. □

Lemma 2. *For any $A \subseteq \mathbb{N}$ and any $n \geqslant 1$, $|g_A(n+1) - g_A(n)| \leqslant (1-e^{-\mu})/\mu$.*

Proof. We first consider the case when $A = \{i\}$. If $i < n$ then

$$\begin{aligned} g_{\{i\}}(n) &= \frac{\mathbb{P}(Z=i) - \mathbb{P}(Z=i)\mathbb{P}(Z<n)}{n\mathbb{P}(Z=n)} = \frac{\mathbb{P}(Z=i)\mathbb{P}(Z \geqslant n)}{n\mathbb{P}(Z=n)} \\ &= \mathbb{P}(Z=i)\sum_{m=0}^{\infty} \frac{\mathbb{P}(Z=n+m)}{n\mathbb{P}(Z=n)} = \mathbb{P}(Z=i)\sum_{m=0}^{\infty} \frac{\mu^m}{n(n+1)\cdots(n+m)}, \end{aligned}$$

which is a positive, decreasing function of n. On the other hand, if $i \geqslant n$ then

$$\begin{aligned} g_{\{i\}}(n) &= \frac{0 - \mathbb{P}(Z=i)\mathbb{P}(Z<n)}{n\mathbb{P}(Z=n)} = -\frac{\mathbb{P}(Z=i)\mathbb{P}(Z<n)}{n\mathbb{P}(Z=n)} \\ &= -\mathbb{P}(Z=i)\sum_{m=0}^{n-1} \frac{\mathbb{P}(Z=n-m-1)}{n\mathbb{P}(Z=n)} = -\mathbb{P}(Z=i)\sum_{m=0}^{n-1} \frac{(n-1)\cdots(n-m)}{\mu^{m+1}}, \end{aligned}$$

which is a negative, decreasing function of n.

Thus $g_{\{i\}}(n+1) - g_{\{i\}}(n)$ is negative unless $i = n$. But for a general A, $g_A(n) = \sum_{i \in A} g_{\{i\}}(n)$, so

$$g_A(n+1) - g_A(n) \leqslant g_{\{n\}}(n+1) - g_{\{n\}}(n).$$

As $g_{\mathbb{N}}(n) = 0$, $g_A(n) = -g_{A^c}(n)$, so applying the above to A^c we have

$$|g_A(n+1) - g_A(n)| \leqslant g_{\{n\}}(n+1) - g_{\{n\}}(n).$$

Finally,

$$\begin{aligned} g_{\{n\}}(n+1) - g_{\{n\}}(n) &= \frac{\mathbb{P}(Z=n)\mathbb{P}(Z \geqslant n+1)}{(n+1)\mathbb{P}(Z=n+1)} - \frac{-\mathbb{P}(Z=n)\mathbb{P}(Z<n)}{n\mathbb{P}(Z=n)} \\ &= \frac{1}{\mu}\mathbb{P}(Z>n) + \frac{1}{n}\mathbb{P}(Z<n) \\ &= \frac{1}{\mu}\mathbb{P}(Z>n) + \sum_{m=1}^{n} \frac{1}{n}\mathbb{P}(Z=m-1) \\ &= \frac{1}{\mu}\mathbb{P}(Z>n) + \sum_{m=1}^{n} \frac{m}{n\mu}\mathbb{P}(Z=m) \\ &\leqslant \frac{1}{\mu}\mathbb{P}(Z>n) + \frac{1}{\mu}\mathbb{P}(0 < Z \leqslant n) \\ &= \frac{1}{\mu}\mathbb{P}(Z>0) = \frac{1-e^{-\mu}}{\mu}. \end{aligned}$$

□