

# STATS310A - Lecture 14

Persi Diaconis  
Scribed by Michael Howes

11/09/21

## Contents

<b>1</b>	<b>Recap</b>	<b>1</b>
<b>2</b>	<b>Stein's equation</b>	<b>2</b>
<b>3</b>	<b>Proof of the Poisson approximation</b>	<b>5</b>
<b>4</b>	<b>References</b>	<b>8</b>

## 1 Recap

Today we will continue with Stien's method and Poisson approximation. As before  $\mathcal{P}_\lambda$  will be used to denote the Poisson distribution with parameter  $\lambda > 0$ . That is, for  $j \in \mathbb{N} = \{0, 1, 2, \dots\}$ , we have  $\mathcal{P}_\lambda(\{j\}) = \frac{e^{-\lambda} \lambda^j}{j!}$ . As before we will say that a random variable  $Z$  is  $\text{Poisson}(\lambda)$  to mean that  $Z$  has distribution  $\mathcal{P}_\lambda$  and so

$$\mathbb{P}(Z \in A) = \mathcal{P}_\lambda(A).$$

Suppose we have a finite index set  $I$  and random variables  $\{X_i\}_{i \in I}$  such that  $X_i$  takes values 0, 1. Suppose  $\mathbb{P}(X_i = 1) = \mathbb{E}[X_i] = p_i$  and that  $\mathbb{P}(X_i = 1, X_j = 1) = \mathbb{E}[X_i X_j] = p_{ij}$ . Let

$$W = \sum_{i \in I} X_i,$$

and

$$\lambda = \sum_{i \in I} p_i = \mathbb{E}[W].$$

Also define  $\mathbb{P}_W$  to be the probability measure  $\mathbb{P}_W(A) = \mathbb{P}(W \in A)$  for  $A \subseteq \mathbb{N}$ . Suppose that we have a dependency graph  $\Gamma$  for  $\{X_i\}_{i \in I}$ . That is, for all subsets  $A, B \subseteq I$ , if  $A$  and  $B$  are disjoint and there are no edges between  $A$  and  $B$  in  $\Gamma$ , then

$$\{X_i\}_{i \in A} \text{ and } \{X_j\}_{j \in B},$$

are independent. For  $i \in I$  we define  $N_i$  to be the neighbourhood of  $i$  in  $\Gamma$ . That is,

$$N_i = \{j \in I : \text{there is an edge from } i \text{ to } j \text{ in } \Gamma\} \cup \{i\}.$$

We wish to prove

**Theorem 1.** *With notation as above*

$$\|\mathcal{P}_\lambda - \mathbb{P}_W\|_{TV} \leq \min(3, \lambda^{-1}) \left( \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + \sum_{i \in I} \sum_{j \in N_i} p_{ij} \right),$$

where  $\|\cdot\|_{TV}$  is the total variation distance.

## 2 Stein's equation

The key idea is the following proposition:

**Proposition 1.** *A random variable is  $\text{Poisson}(\lambda)$  if and only if for all bounded  $f : \mathbb{N} \rightarrow \mathbb{R}$ ,*

$$\mathbb{E}[Zf(Z)] - \lambda \mathbb{E}[f(Z+1)] = 0.$$

We'll need the following analytic lemma.

**Lemma 1 (\*\*).** *For all  $A \subseteq \mathbb{N}$  and  $\lambda > 0$ , there exists a unique function  $f : \mathbb{N} \rightarrow \mathbb{R}$  such that*

- i.  $f(0) = 0$ ,*
- ii. For all  $j \in \mathbb{N}$ ,  $\lambda f(j+1) - jf(j) = \delta_A(j) - \mathcal{P}_\lambda(A)$*
- iii. For all  $j \in \mathbb{N}$ ,  $|f(j)| \leq 1.25$ .*
- iv. For all  $j \in \mathbb{N}$ ,  $|f(j+1) - f(j)| \leq \min(3, \lambda^{-1})$ .*

*Proof.* Starting at  $j = 0$  we can set

$$f(j+1) = \frac{1}{\lambda} (jf(j) + \delta_A(j) - \mathcal{P}_\lambda(A)).$$

The function  $f$  is well-defined by recursion and unique by induction. Thus there exists a unique function  $f$  satisfying items *i.* and *ii.* We wish to show that  $f$  satisfies items *iii.* and *iv.* If we multiply the equation

$$\lambda f(j+1) - jf(j) = \delta_A(j) - \mathcal{P}_\lambda(A),$$

by  $\frac{\lambda^j}{j!}$ , we get the equation

$$\frac{\lambda^{j+1}}{j!} f(j+1) - \frac{\lambda^j}{(j-1)!} f(j) = \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)),$$

for  $j \geq 1$  and for  $j = 0$  we have

$$\lambda f(1) = \lambda (\delta_A(0) - \mathcal{P}_\lambda(A)).$$

Thus

$$\begin{aligned} \frac{\lambda^k}{(k-1)!} f(k) &= \lambda f(1) + \sum_{j=1}^{k-1} \left( \frac{\lambda^{j+1}}{j!} f(j+1) - \frac{\lambda^j}{(j-1)!} f(j) \right) \\ &= \sum_{j=0}^{k-1} \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)) \\ &= - \sum_{j=k}^{\infty} \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)). \end{aligned}$$

The last equality hold because

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)) &= e^\lambda \sum_{j=0}^{\infty} \delta_A(j) \mathcal{P}_\lambda(\{j\}) - \mathcal{P}_\lambda(A) \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\ &= e^\lambda \mathcal{P}_\lambda(A) - \mathcal{P}_\lambda(A) e^\lambda \\ &= 0. \end{aligned}$$

Taking absolute values we get

$$\begin{aligned} |f(k)| &= \frac{(k-1)!}{\lambda^k} \left| \sum_{j=0}^{k-1} \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)) \right| \\ &\leq \frac{(k-1)!}{\lambda^k} \sum_{j=0}^{k-1} \frac{\lambda^j}{j!} |\delta_A(j) - \mathcal{P}_\lambda(A)| \\ &= \frac{1}{\lambda} \sum_{j=0}^{k-1} \frac{(k-1)!}{\lambda^{k-j-1} j!} |\delta_A(j) - \mathcal{P}_\lambda(A)| \\ &\leq \frac{1}{\lambda} \sum_{j=0}^{k-1} \frac{(k-1)!}{\lambda^{k-j-1} j!}. \end{aligned}$$

The last equality holds since  $\delta_A(j), \mathcal{P}_\lambda(A) \in [0, 1]$ . We will now perform a change of variables and sum over  $j' = k - j - 1$  so that  $j = k - j' - 1$ . We thus have

$$\begin{aligned} |f(k)| &\leq \frac{1}{\lambda} \sum_{j=0}^{k-1} \frac{(k-1)!}{\lambda^{k-j-1} j!} \\ &= \frac{1}{\lambda} \sum_{j'=0}^{k-1} \frac{(k-1)!}{\lambda^{j'} (k-1-j')!} \\ &= \frac{1}{\lambda} \sum_{j'=0}^{k-1} \frac{(k-1)(k-2) \dots (k-j')}{\lambda^{j'}} \\ &\leq \frac{1}{\lambda} \sum_{j'=0}^{k-1} \left( \frac{k-1}{\lambda} \right)^{j'} \\ &\leq \frac{1}{\lambda} \sum_{j'=0}^{\infty} \left( \frac{k-1}{\lambda} \right)^{j'}. \end{aligned}$$

If  $k < \lambda + 1$ , then above series is convergent and we have

$$\begin{aligned} |f(k)| &\leq \frac{1}{\lambda} \left( \frac{1}{1 - \frac{k-1}{\lambda}} \right) \\ &= \frac{1}{\lambda - k + 1}. \end{aligned}$$

In particular when  $k \leq \lambda + \frac{1}{5}$ ,

$$|f(k)| \leq \frac{1}{4/5} = 1.25.$$

Using

$$\frac{\lambda^k}{(k-1)!} f(k) = - \sum_{j=k}^{\infty} \frac{\lambda^j}{j!} (\delta_A(j) - \mathcal{P}_\lambda(A)),$$

we also have

$$\begin{aligned} |f(k)| &\leq \frac{(k-1)!}{\lambda^k} \sum_{j=k}^{\infty} \frac{\lambda^j}{j!} |\delta_A(j) - \mathcal{P}_\lambda(A)| \\ &\leq \frac{(k-1)!}{\lambda^k} \sum_{j=k}^{\infty} \frac{\lambda^j}{j!} \\ &= \frac{1}{k} \sum_{j=k}^{\infty} \frac{\lambda^{j-k} k!}{j!} \\ &= \frac{1}{k} \sum_{m=0}^{\infty} \frac{\lambda^m k!}{(m+k)!} \\ &= \frac{1}{k} \sum_{m=0}^{\infty} \frac{\lambda^m}{(m+k)(m+k-1)\dots(k+1)} \\ &\leq \frac{1}{k} \sum_{m=0}^{\infty} \left( \frac{\lambda}{k+1} \right)^m \end{aligned}$$

If  $k > \lambda - 1$ , then the above series is convergent and so

$$|f(k)| \leq \frac{1}{k} \left( \frac{1}{1 - \frac{\lambda}{k+1}} \right) = \frac{k+1}{k(k+1-\lambda)}.$$

In particular if  $k > \lambda + 1/5$  and  $k \geq 2$ , then  $\frac{k+1}{k} \leq \frac{3}{2}$  and so

$$|f(k)| \leq \frac{k+1}{k(1+1/5)} = \frac{5(k+1)}{6k} \leq 1.25.$$

For  $k < 2$ , we have  $f(0) = 0$  and

$$|f(1)| = \frac{1}{\lambda} |\delta_A(1) - \mathcal{P}_\lambda(A)|,$$

which is maximized when  $A = \{0\}$  or  $A = \mathbb{N} \setminus \{0\}$ . In both these cases,

$$|\delta_A(1) - \mathcal{P}_\lambda(A)| = 1 - e^{-\lambda},$$

and thus

$$|f(1)| \leq \frac{1}{\lambda} (1 - e^{-\lambda}) \leq 1.$$

Thus we have shown *iii.* To show *iv.* we need to bound  $|f(j+1) - f(j)|$ . By the triangle inequality

$$|f(j+1) - f(j)| \leq |f(j+1)| + |f(j)| \leq 2 \times 1.25 \leq 3.$$

For homework show that

$$|f(j+1) - f(j)| \leq \lambda^{-1},$$

for  $\lambda \geq \frac{1}{3}$ . □

We are now ready to prove proposition 1.

*Proof.* First suppose that  $Z$  is  $\text{Poisson}(\lambda)$  and  $f : \mathbb{N} \rightarrow \mathbb{R}$  is bounded, then

$$\begin{aligned}
 \mathbb{E}[Zf(Z)] &= \sum_{j=0}^{\infty} jf(j)\mathcal{P}_{\lambda}(\{j\}) \\
 &= \sum_{j=0}^{\infty} jf(j)\frac{e^{-\lambda}\lambda^j}{j!} \\
 &= \sum_{j=1}^{\infty} jf(j)\frac{e^{-\lambda}\lambda^j}{j!} \\
 &= \sum_{j=1}^{\infty} f(j)\frac{e^{-\lambda}\lambda^j}{(j-1)!} \\
 &= \sum_{k=0}^{\infty} f(k+1)\frac{e^{-\lambda}\lambda^{k+1}}{k!} \\
 &= \lambda \sum_{k=0}^{\infty} f(k+1)\mathcal{P}_{\lambda}(\{k\}) \\
 &= \lambda \mathbb{E}[f(Z+1)].
 \end{aligned}$$

Alternative one can note that the equation

$$\mathbb{E}[Zf(Z)] = \lambda \mathbb{E}[f(Z+1)],$$

is linear in  $f$  and thus reduce the result to the case when  $f$  is an indicator function (that is, use a (1),(2),(3) argument).

Now conversely suppose that for all bound  $f : \mathbb{N} \rightarrow \mathbb{R}$ ,

$$\mathbb{E}[Zf(Z)] = \lambda \mathbb{E}[f(Z+1)].$$

Let  $A$  be a subset of  $\mathbb{N}$  and let  $f$  be as in Lemma (\*\*). We then have

$$\begin{aligned}
 \mathbb{P}(Z \in A) - \mathcal{P}_{\lambda}(A) &= \sum_{j=0}^{\infty} (\delta_A(j) - \mathcal{P}_{\lambda}(A)) \mathbb{P}(Z = j) \\
 &= \sum_{j=0}^{\infty} (jf(j) - \lambda f(j+1)) \mathbb{P}(Z = j) \\
 &= \mathbb{E}[Zf(Z) - \lambda f(Z+1)] \\
 &= 0.
 \end{aligned}$$

So  $\mathbb{P}(Z \in A) = \mathcal{P}_{\lambda}(A)$  and  $Z$  is  $\text{Poisson}(\lambda)$ . □

### 3 Proof of the Poisson approximation

We are now ready to prove

$$\|\mathbb{P}_W - \mathcal{P}_{\lambda}\|_{TV} \leq \min(3, \lambda^{-1}) \left( \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + \sum_{i \in I} \sum_{j \in N_i} p_{ij} \right),$$

where the notation is as at the start of this lecture.

*Proof.* Since

$$\|\mathbb{P}_W - \mathcal{P}_\lambda\|_{TV} = \sup_{A \subseteq \mathbb{N}} |\mathbb{P}_W(A) - \mathcal{P}_\lambda(A)|,$$

it suffices to show that

$$|\mathbb{P}_W(A) - \mathcal{P}_\lambda(A)| \leq \min(3, \lambda^{-1}) \left( \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + \sum_{i \in I} \sum_{j \in N_i} p_{ij} \right),$$

for all  $A \subseteq \mathbb{N}$ . Fix such an  $A$  and define  $\Delta = \mathbb{P}_W(A) - \mathcal{P}_\lambda(A)$ . Let  $f$  be as in Lemma (\*\*). Then, as seen in the previous proof,

$$\begin{aligned} \Delta &= \mathbb{P}(W \in A) - \mathcal{P}_\lambda(A) \\ &= \mathbb{E}[Wf(W) - \lambda f(W+1)] \\ &= \mathbb{E} \left[ \sum_{i \in I} X_i f(W) - p_i f(W+1) \right] \\ &= \sum_{i \in I} \mathbb{E}[X_i f(W) - p_i f(W+1)] \end{aligned}$$

For every  $i$ , let  $W_i = W - X_i$  and  $V_i = \sum_{j \in N_i^c} X_j$ . Note that by the definition of a dependency graph,  $V_i$  is independent of  $X_i$ . Note also that

$$\begin{aligned} X_i f(W) &= \begin{cases} 0 & \text{if } X_i = 0, \\ f(W_i + 1) & \text{if } X_i = 1. \end{cases} \\ &= X_i f(W_i + 1). \end{aligned}$$

Thus

$$\begin{aligned} \Delta &= \sum_{i \in I} \mathbb{E}[(X_i - p_i)f(W_i + 1) + p_i(f(W_i + 1) - f(W + 1))] \\ &= \sum_{i \in I} \mathbb{E}[(X_i - p_i)(f(W_i + 1) - f(V_i + 1))] + \mathbb{E}[p_i(f(W_i + 1) - f(W + 1))] \quad (1) \\ &= \sum_{i \in I} \mathbb{E}[(X_i - p_i)(f(W_i + 1) - f(V_i + 1))] + \sum_{i \in I} \mathbb{E}[p_i(f(W_i + 1) - f(W + 1))] \\ &= (I) + (II). \end{aligned}$$

The equality in (1) holds because  $V_i$  and  $X_i$  are independent and thus for each  $i$ ,

$$\begin{aligned} \mathbb{E}[(X_i - p_i)f(W_i + 1)] &= \mathbb{E}[(X_i - p_i)(f(W_i + 1) - f(V_i + 1))] + \mathbb{E}[(X_i - p_i)f(V_i + 1)] \\ &= \mathbb{E}[(X_i - p_i)(f(W_i + 1) - f(V_i + 1))] + \mathbb{E}[(X_i - p_i)]\mathbb{E}[f(V_i + 1)] \\ &= \mathbb{E}[(X_i - p_i)(f(W_i + 1) - f(V_i + 1))]. \end{aligned}$$

We will now bound the absolute value of the sums (I) and (II). We'll start with (II) which is simpler. For each  $i$ ,  $f(W_i + 1) = f(W + 1)$  if  $X_i = 0$  and otherwise  $W_i + 1$  and  $W + 1$  differ by 1. Thus,

$$|f(W_i + 1) - f(W + 1)| \leq X_i \min(3, \lambda^{-1}).$$

And so we have

$$\begin{aligned}
 \left| \sum_{i \in I} \mathbb{E} [p_i (f(W_i + 1) - f(W + 1))] \right| &\leq \sum_{i \in I} p_i \mathbb{E} [|f(W_i + 1) - f(W + 1)|] \\
 &\leq \min(3, \lambda^{-1}) \sum_{i \in I} p_i \mathbb{E}[X_i] \\
 &= \min(3, \lambda^{-1}) \sum_{i \in I} p_i p_i.
 \end{aligned}$$

The sum  $(I)$  is trickier but similar ideas can be used to bound it. For a fixed  $i$ , let  $X'_1, \dots, X'_m$  be an enumeration of the variables in  $N_i \setminus \{i\}$ . We then have

$$\begin{aligned}
 |f(W_i + 1) - f(V_i + 1)| &= \left| f\left(1 + V_i + \sum_{k=1}^m X'_k\right) - f(1 + V_i) \right| \\
 &= \left| \sum_{j=1}^m X'_j \left( f\left(1 + V_i + \sum_{k=1}^j X'_k\right) - f\left(1 + V_i + \sum_{k=1}^{j-1} X'_k\right) \right) \right| \\
 &\leq \sum_{j=1}^m X'_j \left| f\left(1 + V_i + \sum_{k=1}^j X'_k\right) - f\left(1 + V_i + \sum_{k=1}^{j-1} X'_k\right) \right| \\
 &\leq \min(3, \lambda^{-1}) \sum_{j=1}^m X'_j \\
 &= \min(3, \lambda^{-1}) \sum_{j \in N_i \setminus \{i\}} X_j.
 \end{aligned}$$

Thus we have that

$$\begin{aligned}
 \left| \sum_{i \in I} \mathbb{E} [(X_i - p_i)(f(W_i + 1) - f(V_i + 1))] \right| &\leq \min(3, \lambda^{-1}) \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} \mathbb{E}[|X_i - p_i| X_j] \\
 &= \min(3, \lambda^{-1}) \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} \mathbb{E}[X_i X_j] + p_i \mathbb{E}[X_j] \\
 &= \min(3, \lambda^{-1}) \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + p_i p_j.
 \end{aligned}$$

Thus combining our bounds on  $(I)$  and  $(II)$  we have

$$\begin{aligned}
 |\Delta| &\leq \min(3, \lambda^{-1}) \left( \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + p_i p_j + p_i p_i \right) \\
 &= \min(3, \lambda^{-1}) \left( \sum_{i \in I} \sum_{j \in N_i \setminus \{i\}} p_{ij} + \sum_{i \in I} \sum_{j \in N_i} p_{ij} \right)
 \end{aligned}$$

□

## 4 References

Three references are

- “Poisson Approximation and the Chen-Stein Method” by Arratia, Goldstein and Gord.
- “Exchangeable pairs and Poisson approximation” by Chatterjee, Diaconis and Meckes.
- “An Introduction to Stein’s Method” by Barbour and Chen. This is a textbook which is available online through Stanford Libraries.