Método de Stein:

Haces approx de leyes de proba complicadas mediante leyes "Menos complicadas" ← { • Gaussianas. • <u>Poisson</u>

Ejemplo de aprox. Poisson.

$$X_n = \sum_{k=1}^{n} \xi_k^n \qquad \xi_k^n \leftarrow \begin{array}{l}\text{independientes}\\ \text{Bernoulli}(p_k).\end{array}$$

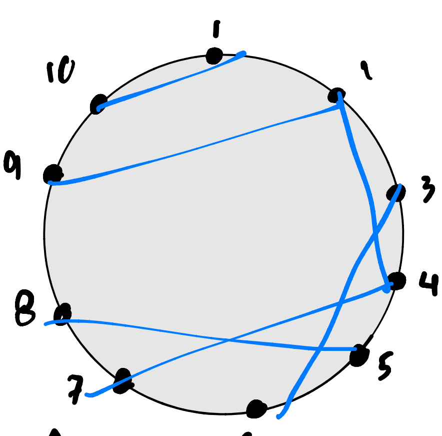
En algunas circunstancias:

$$\mathcal{L}(X_n) \approx \text{Ley Poisson}(\mathbb{E}[X_n]).$$

Ejemplos:

1). X_n = conteo en contextos de

- De fondo hay una gráfica:



↑ realización de una gráfica.

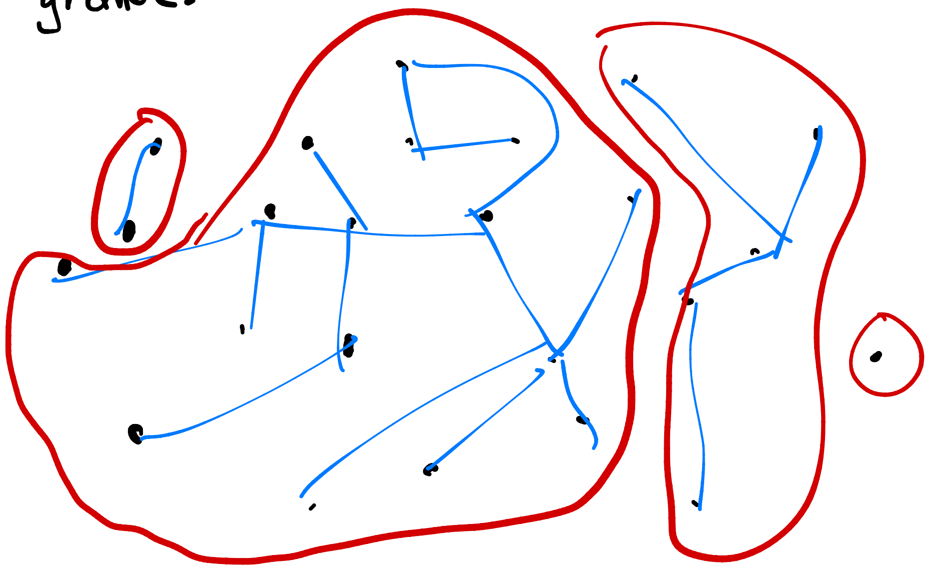
Conteo de "cruces"

$$= \sum_{\substack{(a,b)\\(c,d)}} \mathbb{1}\{\text{hay un cruce entre } (a,b) \text{ y } (c,d)\} \ \circledast$$

Pregunta central:

$\circledast \approx$ Algo "sencillo"?

2.-) Ejemplo: una gráfica grande:



Se estudia la conectividad de una gráfica.

Conteo de # componentes conexas ¿$\approx$ Algo simple?

Motivación:

Problema:

$\{\xi_k\}_{k\geq 1}$ ← variables aleatorias independientes, con $\mathbb{E}[\xi_k]=0$, y definimos

$$\sigma_n^2 = \sum_{k=1}^{n} \mathbb{E}[\xi_k^2].$$

$$S_n = \frac{1}{\sigma_n}\sum_{k=1}^{n} \xi_k$$

Pregunta: $d_K(S_n, N) \approx$? $N \sim \text{Normal}(0,1)$?

Heurística de Stein:

es "razonable" pensar que es equivalente

a) $$\sup_{z\in\mathbb{R}} |\mathbb{P}[S_n \le z] - \mathbb{P}[N \le z]| \approx 0$$

$$\sup_{z\in\mathbb{R}} |\mathbb{E}[\mathbb{1}_{\{S_n \le z\}}] - \mathbb{E}[\mathbb{1}_{\{N \le z\}}]|$$

b) $$\sup_{f\in\mathcal{C}} |\mathbb{E}[S_n f(S_n) - f'(S_n)]| \approx 0.$$

$\uparrow$

familia de funciones de prueba

Equivalencia entre a) y b):

considerar la ecuación

$$x f(x) - f'(x) = \mathbb{1}_{\{x \le z\}} - \mathbb{E}[\mathbb{1}_{\{N \le z\}}] \quad \circledast$$

llamamos f_z a la solución a $\circledast$.

Notar que

$$|\mathbb{E}[S_n f_z(S_n) - f_z'(S_n)]|$$
$$= |\mathbb{E}[\mathbb{1}_{\{S_n \le z\}} - \mathbb{E}[\mathbb{1}_{\{N \le z\}}]]|$$
$$= |\mathbb{P}[S_n \le z] - \mathbb{P}[N \le z]|.$$

$\Rightarrow$ tomando $\sup_{z \in \mathbb{R}}$

$\Rightarrow$

$$d_K(S_n, N) = \sup_z |\mathbb{P}[S_n \leq z] - \mathbb{P}[N \leq z]|.$$

$$= \sup_z |\mathbb{E}[S_n f_z(S_n) - f_z'(S_n)]|$$

Claim:

$f_z \leftarrow$ sale de de $\mathbb{1}_{[0,z]}(\cdot)$.

pero tiene mejores propiedades.

$\|f_z\|_\infty, \|f_z'\|_\infty \leq 10$

Por lo tanto,

$$d_K(S_n, N) \leq \sup_{\substack{f:\mathbb{R}\to\mathbb{R} \\ \|f\|_\infty, \|f'\|_\infty \leq 10}} |\mathbb{E}[\underbrace{S_n f(S_n) - f'(S_n)}_{\mathcal{L}[f](S_n)}]|,$$

con $\mathcal{L}[f](x) = x f(x) - f'(x)$