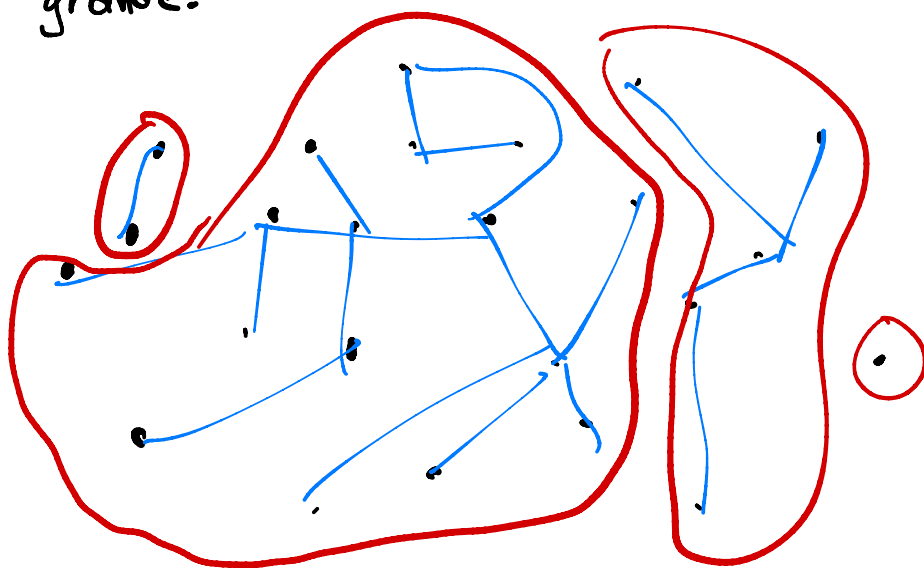



2.-) Ejemplo: una gráfica
grande:



Se estudia la conectividad de una gráfica.
Cuento de # componentes conexas \approx Algo simple?

Motivación:

Problema:

$$\{z_k\}_{k \geq 1}$$

\leftarrow variables aleatorias independientes,
con $E[z_k] = 0$, y definimos

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

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

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

Heurística de Stein:
es "razonable" pensar que es equivalente

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

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

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

↑
familia de
funciones de prueba

Equivalencia entre a) y b):

considerar la ecuación

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

llamamos f_z a la solución a $\textcircled{*}$.

Notar que

$$|E[S_n f_z(S_n) - f_z'(S_n)]|$$

$$= |E[\mathbb{1}_{\{S_n \leq z\}} - E[\mathbb{1}_{\{N \leq z\}}]]|$$

$$= |P\{S_n \leq z\} - P\{N \leq z\}|.$$

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

\Rightarrow

$$\begin{aligned}d_K(S_n, N) &= \sup_z |P\{S_n \leq z\} - P\{N \leq z\}| \\ &= \sup_z |E[S_n f_z(S_n) - f'_z(S_n)]|\end{aligned}$$

Claim:

$f_z \leftarrow$ sale de de $[0, z]$ (o).

pero tiene mejores propiedades.

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

Por lo tanto,

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

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

con $L[f](x) = xf(x) - f'(x)$

Pregunta:

Heurística de Stein:

$\mathcal{C} \leftarrow$ cjtto de funciones de prueba.

$$d_{\mathcal{C}}(X, N) := \sup_{h \in \mathcal{C}} |\mathbb{E}[h(X) - h(N)]|$$

$X \leftarrow$ v.a.

$N \leftarrow \sim N(0,1)$

↓ intercambio mediante
heurística de Stein:

$$\sup_{h \in \mathcal{C}} |\mathbb{E}[X f_h(X) - f_h'(X)]|$$

resolver:

$$\underbrace{X f(X) - f'(X)}_{\mathcal{L}[f](X)} = \mathbb{E}[h(X) - h(N)]$$

Nota:

f_h tiene "mejores" propiedades que h .

Ejemplo:

$$\mathcal{C} = \left\{ \mathbb{1}_{(-\infty, z]}(x); z \in \mathbb{R} \right\}.$$

Claim: si $h \in \mathcal{C} \Rightarrow f_h$ es absolutamente
continua; \downarrow

Lema 2.3.

libro Normal approximation

$$\|f_h'\|_{\infty} \leq 1.$$

by Stein Method.

$$\Rightarrow d_{\mathcal{C}}(X, N) \leq \sup_{f \in \tilde{\mathcal{C}}} |\mathbb{E}[X f(X) - f'(X)]|.$$

con $\hat{C} = \{ f: \mathbb{R} \rightarrow \mathbb{R}; f' \text{ existe c.s.} \}$
 y $\|f'\|_\infty \leq 1$

Nota:

Si $h: \mathbb{R}^d \rightarrow \mathbb{R}^d$ y consideramos.

$$\bar{x} \nabla f(\bar{x}) - \Delta f(\bar{x}) = h(\bar{x}) - E[h(\bar{N})]$$

$L(f)(\bar{x}) \qquad \bar{N} \sim \mathcal{N}(0, I_d).$

$$f(\bar{x}) = \int_{\mathbb{R}^d} (h(\bar{x} \bar{e}^0 + \sqrt{1 - \bar{e}^{20}} \bar{z}) - E[h(\bar{N})]) F_N(dz)$$

Si $C = \{ h: \mathbb{R}^d \rightarrow \mathbb{R}^d : h \text{ es } L\text{-Lipchitz} \}$.

$d_C(x, N) \leftarrow$ solución explícita. y está en términos de una ec. de transporte (libro de C. Villani).

Hay situaciones en las que

$d_{TV}(x, \text{Otra cosa})$ es aproximable.

Otra situación:

- Expansiones de Edgeworth.

$\mu_{X_n} \leftarrow$ medida de proba de X_n

$$\tilde{\mu}_n = \mu_{\frac{X_n - \mu_n}{\sigma_n}} \approx \gamma$$

medida Gaussiana estandar d.

$$\int f(x) \mu_n(dx) \approx \int f(x) \gamma(dx) + \int f(x) \underbrace{\hat{\gamma}(dx)}_{\substack{\uparrow \\ \text{medida compleja.}}}$$

medida real

Hasta el momento, hemos visto que

$$d_K(S_n, N) \leq \sup_{f \in \mathcal{C}} |\mathbb{E}[X f(X) - f'(X)]|$$

$$\mathcal{C} = \left\{ f: \mathbb{R} \rightarrow \mathbb{R} \text{ t.g.} \quad X f_n(X) - f'_n(X) = h(X) \cdot \mathbb{E}[h(N)] \right. \\ \left. \text{con } h(x) = \mathbb{1}_{(-\infty, 2)}(x) \right\}.$$

¿Cómo manejamos a

$$\mathbb{E}[X f(X) - f'(X)]?$$

Caso que consideraremos:

$$X = S_n = \frac{1}{\sigma_n} \sum_{k=1}^n Z_k, \quad \text{con } Z_k \text{ son centradas } \Leftarrow \\ \text{e indep. (NO necesariamente i.i.d.)}$$

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

Vamos a probar que $S_n \xrightarrow{\text{ley}} N(0, 1)$

$$\text{si } \sum_{k=1}^n \mathbb{E} \left[\left| \frac{Z_k}{\sigma_n} \right|^3 \right] \rightarrow 0$$

De hecho, $\sum_{k=1}^n \mathbb{E} \left[\left| \frac{Z_k}{\sigma_n} \right|^3 \right]$ es la velocidad de convergencia.

Idea:

Acutar el funcional

$$|\mathbb{E}[S_n f(S_n) - f'(S_n)]|$$

$$S_n = \sum_{k=1}^n \zeta_k^n, \quad \text{con} \quad \zeta_k^n = \frac{\zeta_k}{\sigma_n}$$

$$\mathbb{E}[S_n f(S_n)] = \sum_{k=1}^n \mathbb{E}[\zeta_k^n f(\sum_{j=1}^n \zeta_j^n)]$$

$$= \sum_{k=1}^n \mathbb{E} \left[\zeta_k^n \left(f \left(\sum_{\substack{1 \leq j \leq n \\ k \neq j}} \zeta_j^n + \zeta_k^n \right) - f \left(\sum_{\substack{1 \leq j \leq n \\ k \neq j}} \zeta_j^n \right) \right) \right]$$

heurística

\approx

$$\sum_{k=1}^n \mathbb{E} [|\zeta_k^n|^2] \mathbb{E} \left[f' \left(\sum_{\substack{1 \leq j \leq n \\ k \neq j}} \zeta_j^n \right) \right]$$

$$\approx \left(\sum_{k=1}^n \mathbb{E} [|\zeta_k^n|^2] \right) \mathbb{E} [f'(S_n)] \quad \sum_{k=1}^n \frac{\mathbb{E} [\zeta_k^2]}{\sigma_n^2} = 1$$

Moral e ja: $\mathbb{E}[S_n f(S_n)] \approx \mathbb{E}[f'(S_n)]$.

Formalidades:

$$\mathbb{E}[S_n f(S_n)] = \sum_{k=1}^n \mathbb{E} \left[z_k^n \left(f(S_n^{(k)} + z_k^n) - f(S_n^{(k)}) \right) \right]$$

$$S_n^{(k)} := \sum_{\substack{1 \leq j \leq n \\ k \neq j}} z_j^n$$

Ahora,

$$\underbrace{f(S_n^{(k)} + z_k^n) - f(S_n^{(k)})}_{= f(S_n^{(k)} + \lambda z_k^n) \Big|_{\lambda=0}^{\lambda=1}} = z_k^n \int_0^1 f'(S_n^{(k)} + \lambda z_k^n) d\lambda = *$$

$$+ f'(S_n^{(k)})$$

$$* = z_k^n f'(S_n^{(k)}) + z_k^n \int_0^1 \int_0^1 f''(S_n^{(k)} + \gamma \lambda z_k^n) \lambda z_k^n d\gamma d\lambda$$

\Rightarrow

$$\left| \mathbb{E}[S_n f(S_n)] - \sum_{k=1}^n \mathbb{E}[z_k^n f'(S_n^{(k)})] \right|$$

$$\leq \|f''\|_{\infty} \sum_{k=1}^n \mathbb{E}[|z_k^n|^3]$$

⇒

$$|\mathbb{E}[S_n f(S_n)] - \sum_{k=1}^n \mathbb{E}[R_k^2 f'(S_{n,k})]|$$

$$\leq \|f''\|_\infty \sum_{k=1}^n \mathbb{E}[|Z_k^n|^3] + \sum_{k=1}^n |\mathbb{E}[R_k^2 (f'(S_{n,k}^{(k)}) - f'(S_{n,k}))]|$$

$| \sim | \leq \|f''\|_\infty |Z_k^n|$

$$\leq 2 \|f''\|_\infty \sum_{k=1}^n \mathbb{E}[|Z_k^n|^3]. \quad \square$$

Conclusión:

Teorema:

$$\sup_{h \in \mathcal{C}} |\mathbb{E}[h(S_n) - h(N)]| \leq 2 \left(\sup_{h \in \mathcal{C}} \|f_h''\|_\infty \right) \sum_{k=1}^n \mathbb{E}[|Z_k^n|^3].$$

$$\frac{\sum Z_k}{\sigma_n}$$



versión del teorema de Berry-Esseen.

Caso particular: $\{Z_k\}_{k=1}^n$ son i.i.d

$$\sup_{h \in \mathcal{C}} |\mathbb{E}[h(S_n) - h(N)]| \leq 2 \left(\sup_{h \in \mathcal{C}} \|f_h''\|_\infty \right) \frac{\mathbb{E}[|Z_1|^3]}{\mathbb{E}[|Z_1|^2]^{3/2}} \cdot \frac{1}{\sqrt{n}}$$