

Near-independence; Janson's inequality.

Set-up. $X = \{X_1, \dots, X_n\}$ $P = (p_1, \dots, p_n)$

X_P random set: select X_i w. prob. p_i indep.

$E_1, \dots, E_k \subseteq X$, $A_i :=$ event that $E_i \subseteq X_P$

$Z := \sum_{i=1}^k 1_{A_i} = \sum_{i=1}^k 1[E_i \subseteq X_P]$ "# of events that hold".

Define

$$\mu := E[Z] = \sum_{i=1}^k P(A_i)$$

$i \sim j$ if $i \neq j$, $E_i \cap E_j \neq \emptyset$

$$\Delta := \sum_i \sum_{j \sim i} P(A_i \cap A_j)$$

NB: $A_i = \{A : A \supseteq E_i\}$ "principle up-set".

Note the graphs with $\triangle$ as subgraph is a principle up-set.

However, set of graphs with:  vertices 1 and 2 connected is NOT. (Since multiple paths to connect)

Thm Janson's Ineq. $P[Z=0] \leq e^{-\mu + \frac{\Delta}{2}}$

Proof Write $r_i = P(A_i | A_1^c \cap \dots \cap A_{i-1}^c)$.

$$\text{Then } P(Z=0) = P(A_1^c \cap \dots \cap A_k^c)$$

$$= P(A_1^c) \cdot P(A_2^c | A_1^c) \dots P(A_k^c | A_1^c \cap \dots \cap A_{k-1}^c)$$

$$= \prod_{i=1}^k (1 - r_i) \leq \prod_{i=1}^k e^{-r_i} \quad \text{since } 1-x \leq e^{-x}$$

$$= e^{-\sum_{i=1}^k r_i}$$

Thus it suffices to show $\sum_{i=1}^k r_i \geq \mu - \frac{\Delta}{2}$.

(NB if $r_i = P(A_i)$, then get $\sum r_i = \mu$, so want $r_i \geq P(A_i) - \delta$ for small δ).

Pf (of Janson's ineq. ct'd).

Fix i , consider $r_i = P(A_i | A_i^c \cap \dots \cap A_{i-1}^c)$

$$D_i := \bigcap_{\substack{j < i \\ j \neq i}} A_j^c$$

dep. part

$$D_0 := \bigcap_{\substack{j < i \\ j \neq i}} A_j^c$$

indep. part

NB:

$$D_i^c = \left(\bigcap_{\substack{j < i \\ j \neq i}} A_j^c \right)^c = \bigcup_{\substack{j < i \\ j \neq i}} A_j$$

Note $P(A_i | D_0) = P(E_i \leq X_F | \bigcap_{\substack{j < i \\ j \neq i}} E_j \leq X_F)$

$$= P(E_i \leq X_F)$$

$$= P(A_i)$$

since $E_j \wedge E_i = \emptyset$

for all j above

used that A_i principle up-set.

$$r_i := P(A_i | D_0 \cap D_i) = \frac{P(A_i \cap D_0 \cap D_i)}{P(D_0 \cap D_i)} \geq \frac{P(A_i \cap D_0 \cap D_i)}{P(D_0)}$$

$$\Rightarrow r_i \geq P(A_i \cap D_i | D_0) = \underbrace{P(A_i | D_0)}_{= P(A_i)} - \underbrace{P(A_i \cap D_i^c | D_0)}_{\text{wts small.}}$$

$$D_0 = \bigcap_{\substack{j < i \\ j \neq i}} A_j^c \quad \text{down-set (since } A_j \text{ up-sets, } A_j^c \text{ down-set)}$$

similarly $A_i \cap D_i^c$ up-set, since A_i, D_i^c up-sets ($\because D_i$ down-set)

$$\Rightarrow P(A_i \cap D_i^c | D_0) \leq P(A_i \cap D_i^c) = P(A_i \cap \left(\bigcup_{\substack{j < i \\ j \neq i}} A_j \right))$$

$$\hookrightarrow = P\left(\bigcup_{\substack{j < i \\ j \neq i}} (A_i \cap A_j)\right) \leq \sum_{\substack{j < i \\ j \neq i}} P(A_i \cap A_j)$$

Pf (Janson's neg ct'd II)

$$\therefore r_i \geq P(A_i) - \sum_{j < i} P(A_i \cap A_j)$$

$$\text{r.e.} - \sum_i r_i \geq \sum_{i=1}^k P(A_i) - \sum_i \sum_{\substack{j < i \\ j \sim i}} P(A_i \cap A_j) = \mu - \frac{\Delta}{2} \quad \square$$

— Aside: general up-sets.

Remark: Janson's neg. also true for $A_1, \dots, A_k$ general up-sets.

where $i \sim j$ if A_i, A_j dependent.

Claim graph defined by " $\sim$ " is a dependency digraph
when $A_1, \dots, A_k$ are upsets. (Not true for general events).

↑ part of proof for general case.

Prop. For any $0 \leq t \leq \mu$:

$$P(Z \leq \mu - t) \leq \exp \left(- \frac{\Phi(-t/\mu) \mu^2}{\mu + \Delta} \right) \leq \exp \left(- \frac{t^2}{2(\mu + \Delta)} \right)$$

where $\Phi(x) := (1+x) \log(1+x) - x$ with $\Phi(-1) := 1$.

(no proof)

Same setting as Janson's neg.

Thm $\hookrightarrow$ if $\Delta \geq \mu$, then $P(Z=0) \leq e^{-\frac{\mu^2}{2\Delta}}$.

Pf By Janson's neg, for any $S \subseteq K$. Janson's neg.

$$P(Z=0) := P\left(\bigcap_{i=1}^k A_i^c\right) \leq P\left(\bigcap_{i \in S} A_i^c\right) \leq e^{-\mu_S + \frac{\Delta_S}{2}}$$

where $\mu_S = \sum_{i \in S} P(A_i) = \sum_{i \in S} 1_{[i \in S]} P(A_i)$

$$\Delta_S = \sum_{i \in S} \sum_{\substack{j \neq i \\ j \in S}} P(A_i \cap A_j)$$

fix $r \in [0, 1]$,
Now select S randomly : for $i=1, \dots, k$ indep include $i \in S$ w. prob. r .

$\Rightarrow \mu_S, \Delta_S$ now random variables.

$$E[\mu_S] = \sum_{i=1}^k E[\overbrace{1_{[i \in S]}}^{=r}] P(A_i) = r \cdot \mu$$

$$E[\Delta_S] = \sum_i \sum_{\substack{j \neq i \\ j \in S}} E[\overbrace{1_{[i \in S, j \in S]}}^{=r^2}] \cdot P(A_i \cap A_j) = r^2 \Delta$$

$$\therefore E\left[\mu_S - \frac{\Delta_S}{2}\right] = r\mu - r^2\Delta \Rightarrow \exists S' \text{ s.t. } \mu_{S'} - \frac{\Delta_{S'}}{2} \geq r\mu - r^2\Delta$$

$$\therefore P(Z=0) \leq e^{-r\mu + \frac{r^2\Delta}{2}}$$

take $r = \frac{\mu}{\Delta} \leq 1$ (since assumed $\mu \leq \Delta$) $r\mu = \frac{\mu^2}{\Delta}$

$$\Rightarrow P(Z=0) \leq e^{-\frac{\mu^2}{\Delta} + \frac{\mu^2}{2\Delta}} = e^{-\frac{\mu^2}{2\Delta}}$$

□

Cor $P(Z=0) \leq \exp(-\min(\mu/2, \mu^2/2\Delta))$.

Pf $\Delta < \mu$: apply orig. Janson's meg.

$\Delta \geq \mu$: " last thm. $\square$.

Compare: Second-Moment-Method (SMM)

vs Janson's meg.

Chebyshev.

SMM $P(Z=0) \leq \frac{\text{Var}(Z)}{\mu^2} \leq \frac{\Delta + \mu}{\mu^2}$

where
$$\begin{aligned} \text{Var}(Z) &= \sum_{i,j} P(A_i \cap A_j) - P(A_i) \cdot P(A_j) \\ &= \sum_{i,j} P(A_i \cap A_j) - P(A_i)P(A_j) \\ &\quad + \underbrace{\sum_i P(A_i) - P(A_i)^2}_{\leq \mu} \quad \Delta \\ &\leq \Delta + \mu. \end{aligned}$$

if $\mu \geq L, \mu^2/\Delta \geq L$ SMM: $P(Z=0) \leq 2/L$.

Janson meg $P(Z=0) \leq \exp(-\min\{\mu/2, \mu^2/2\Delta\})$.

if $\mu \leq L, \mu^2/\Delta \geq L$ $\leq \exp(-L/2)$ $\leftarrow$ exponentially decreasing error prob!