Markov’s Inequality

Let $X$ be a non-negative r.v. Then for any $t > 0$,

$$\Pr[X \geq t] \leq \frac{EX}{t}$$

**Proof:**

$$EX = \sum_x x \cdot \Pr[X = x]$$

$$= \sum_{x < t} x \cdot \Pr[X = x] + \sum_{x \geq t} x \cdot \Pr[X = x]$$

$$\geq 0$$

$$\geq t \sum_{x \geq t} \Pr[X = x]$$

$$= t \cdot \Pr[X \geq t]$$

---

Chebyshev’s Inequality

Let $X$ be any r.v. Then for any $t > 0$,

$$\Pr[|X - EX| \geq t] \leq \frac{\text{Var} X}{t^2}$$

The most "general" inequality
with no assumptions on $X$.

**Proof:**

$$\Pr[|X - EX| \geq t] = \Pr[(X - EX)^2 \geq t^2]$$

$$\leq \frac{1}{t^2} \mathbb{E}(X - EX)^2$$

$$= \frac{\text{Var} X}{t^2}$$
\[
\frac{\text{Exp}(X - \text{Exp}(X))}{t^2} \leq \frac{\text{Var}(X)}{t^2}
\]

Often times, we will be interested in r.v.'s \( Z \) which are sum of independent rvs.

\[ Z = X_1 + X_2 + \ldots + X_n \]

and each \( X_i \) is an indicator rv with

\[ \text{Pr}[X_i = 1] = p_i \]

\[
\therefore \text{Exp} Z = \sum_{i=1}^{n} p_i =: \mu
\]

\[
\text{Var} Z = \sum_{i=1}^{n} \text{Var} X_i = \sum_{i=1}^{n} p_i (1 - p_i) \leq \mu
\]

is a sum of independent random rvs.

\[
\text{Pr}[|Z - \mu| \geq \delta \mu] \leq ??
\]

Chebyshov gives us

\[
\leq \frac{\text{Var} Z}{\delta^2 \mu^2} \leq \frac{1}{\mu \delta^2}
\]

The “Chernoff Bound” gives us an exponentially better bound.
Chernoff Bound

\[ Z = X_1 + \ldots + X_n \text{ where each } X_i \text{ is ind. rv} \]

\[ \mathbb{E} X_i = p_i \quad \text{and} \quad \mu = \sum p_i \]

\[ P[Z \geq (1+\delta)\mu] \leq e^{-\mu \left((1+\delta)\ln(1+\delta) - \delta\right)} \]

...even for \(-1 < \delta \leq 0\)

"Interesting regions"

- when \(\delta\) is "small" (think \(\delta < 1\))
  \[ \ln(1+\delta) \approx \delta \]

\[ P[Z \geq (1+\delta)\mu] \leq e^{-\mu \delta^2/4} \quad \text{if} \quad -1 \leq \delta \leq 1 \]

- when \(\delta\) is "large" (\(\delta \gg 1\))
  and so only one "tail" is interesting

\[ P[Z \geq (1+\delta)\mu] \leq e^{-\mu \delta^2 / 2} \quad \text{if} \quad \delta \gg 10 \]

A similar stunt is done for the "other tail"

\[ P[Z \leq (1-\delta)\mu] \leq e^{-\mu \left[(1-\delta)\ln(1-\delta) + \delta\right]} \]

\[ \leq e^{-\mu \delta^2/2} \quad \text{if} \quad |\delta| \leq 1 \]

(which'll be the case)

\[ \text{(TRICK): Look at } e^{tZ} \]

\[ P[Z \geq (1+\delta)\mu] = P[e^{tZ} \geq e^{(1+\delta)\mu t}] \]

for any \(t > 0\).

Markov's Inequality then gives...

\[ P[Z \geq (1+\delta)\mu] \leq \mathbb{E}[e^{tZ}] \cdot e^{-(1+\delta)\mu t} \]
\[ P[Z > (1+\delta)\mu] \leq \mathbb{E}[e^{tZ}] \cdot e^{-(1+\delta)t\mu} \]

\[ \mathbb{E}[e^{tZ}] = \mathbb{E}\left[ e^{t\sum x_i} \right] = \prod_{i=1}^{n} \mathbb{E}[e^{tx_i}] \]

Only place where independence is used.

\[ \mathbb{E}[e^{tx_i}] = p_i e^t + (1-p_i) \]
\[ = 1 + p_i (e^t - 1) \]
\[ \leq e^{p_i(e^t-1)} \quad \text{(only other trick)} \]

Putting things together, we get ...

\[ P[Z > (1+\delta)\mu] \leq e^{-(1+\delta)t\mu} \cdot e^{\sum p_i(e^t-1)} \]
\[ = e^{-(1+\delta)t\mu + \mu(e^t-1)} \]
\[ = e^{-\mu\left[t(1+\delta) - (e^t-1)\right]} \]

For all \( t > 0 \)

In particular (after calculus) for \( t = \ln(1+\delta) \)

\[ P[Z > (1+\delta)\mu] \leq e^{-\mu\left[(1+\delta)\ln(1+\delta) - \delta\right]} \]
Generalization (Exercise)

If \( u \geq 1 \), \( P[Z \geq (1+\varepsilon)u] \leq e^{-\varepsilon [1+(1+\varepsilon)2\ln(1+\varepsilon) - 8]} \)

Applications

1. Toss \( n \) coins. \( X_i = \begin{cases} 1 & \text{if tails} \\ 0 & \text{if heads} \end{cases} \)
   
   \[ Z = \sum X_i = \# \text{ of tails}. \]
   
   \[ \mathbb{E}Z = n/2 \]
   
   \[ P \left[ |Z - \frac{n}{2}| \geq 8 \cdot \frac{n}{2} \right] \leq e^{-\frac{n\delta^2}{4}} \quad \text{for } \delta \leq 1 \]

In particular if \( \delta = \frac{2}{\sqrt{n}} \), we get

\[ P \left[ |Z - \frac{n}{2}| \geq \sqrt{n} \right] \leq e^{-1} \]

& more generally

\[ P \left[ |Z - \frac{n}{2}| \geq c \sqrt{n} \right] \leq e^{-c^2} \]

\[ \delta = \frac{2c}{\sqrt{n}} \quad \& \quad c \ll \sqrt{n} \]

2. Discrepancy of a set system
Discrepancy of a set system

- m sets \( S_1, \ldots, S_m \subseteq \{1, \ldots, n^3\} \)

- Want to find a coloring
  \[ \sigma: \{1, \ldots, n^3\} \rightarrow \{ \pm 1, -1 \}^3 \]
  s.t. \( \max_{i=1}^m |\sigma(S_i)| \) is minimized

- Equivalently,
  \[ \sigma: \{1, \ldots, n^3\} \rightarrow \{0, 1\}^3 \]
  s.t. \( \max_{i=1}^m |\sigma(S_i) - \frac{|S_i|}{2}| \) is min.

- By the previous application, even a random coloring, i.e., \( \sigma(i) = \{1, 0, -1\} \) give for every \( S_i \) ...

\[ \Pr \left[ \left| \sigma(S_i) - \frac{|S_i|}{2} \right| \geq t \sqrt{|S_i|} \right] \leq e^{-t^2/2} \]

\[ \forall i = 1, \ldots, m, \]

\[ \Pr \left[ \left| \sigma(S_i) - \frac{|S_i|}{2} \right| \geq \sqrt{4 \ln m} \right] \leq e^{-\frac{2 \ln m}{m^2}} = \frac{1}{m^2} \]

\[ \therefore \sum_{i=1}^m \left| \sigma(S_i) - \frac{|S_i|}{2} \right| < 1 \]
\[
\Pr \left[ \exists i = 1 \ldots m : \left| \sigma(s_i) - \frac{1}{2} \right| > \sqrt{\frac{\ln m}{m}} \right] < \frac{1}{m}
\]

\[
\Pr \left[ \exists i \in \mathcal{E}_i \right] \leq \sum_i p[x_i]
\]

Every set system has a discrepancy of \(O(\sqrt{\ln m})\)

---

**Minimizing Congestion in Directed Networks**

**Input:**
- Directed Graph \(G = (V, A)\)
- Source-sink pairs \(\xi(s_i, t_i), (s_2, t_2) \ldots (s_k, t_k)\)

**Output:** Paths \(p_1, p_2, \ldots, p_k\) where

\[p_i \text{ goes from } s_i \rightarrow t_i\]

**Obj:** Minimize \(\max_e \text{ congestion } (e)\)

\[\text{\# of } p_i \text{'s containing } e\]

**Example:**

\[
\begin{array}{c}
s_1 \quad \rho_1 \quad s_1 \quad t_1
\end{array}
\]
LP relaxation
- \( P_i \) := set of paths from \( s_i \) to \( t_i \)
- Variable: \( x_p \)

\[
\min \sum_{i=1}^{k} \sum_{p \in P_i} x_p \leq C
\]

\[
\forall i : \sum_{p \in P_i} x_p = 1
\]
- \( x_p \in [0,1] \)
- \( C \geq 1 \)

As described, this LP has exponentially many variables.

For now, let's assume we can solve the LP \& that \( x_p > 0 \) for only polynomially many paths.

Algorithm:
For each \( i = 1 \ldots k \),

Sample one \( p \in P_i \) w.p. \( x_p \)

Analysis
- Fix edge \( e \)
- Fix edge \( e \)
- Define \( X_{e,i} = \begin{cases} 1 & \text{if the } i^{th} \text{ path contains } e \\ 0 & \text{otherwise} \end{cases} \)
- \( \text{Cong}(e) = \sum_{i=1}^{k} X_{e,i} \)
- \( \text{ALG} = \max_{e} \text{Cong}(e) \)
- \( \mathbb{E}[\text{Cong}(e)] = \sum_{i=1}^{k} \mathbb{E}[X_{e,i}] \)
- \( = \sum_{i=1}^{k} \sum_{p \in P_i : e \in p} X_p \leq L_P \)
- Also note: \( X_{e,1}, X_{e,2}, \ldots, X_{e,k} \) are independent.
- Just to contrast, \( X_{e,i} \neq X_{e',i} \) for the same \( i \) are NOT independent.
- \( \Pr[\text{Cong}(e) \geq (1 + \delta) L_P] \leq e^{-\frac{L_P \delta \ln \delta}{2}} \)
- (for \( \delta > 10 \)) \( \ldots \) using the generalization
- \( \text{if } \frac{L_P \delta \ln \delta}{2} \geq 2 \ln m \), then
if \( \frac{LP \cdot \ln s}{2} \geq 2 \ln m \), then

\[
P\left[ \exists e : \text{Cong}(e) \geq (1+\delta)LP \right] \leq \frac{1}{m}
\]

\[
\Rightarrow \quad \forall e, \text{Cong}(e) \leq (1+\delta)LP \quad \forall e.
\]

\[
\therefore \text{if } S \geq \frac{8 \ln m}{\ln \ln m}, \text{ this is indeed true.}
\]

\[
\therefore \ln S \geq \frac{8 \ln m}{\ln \ln m}
\]

\[
\Rightarrow \quad LP \geq 1
\]

\[
\therefore \text{The Randomized Algorithm is an}
\]

\[
O\left(\frac{\ln m}{\ln \ln m}\right) - \text{factor approx algo.}
\]

\textbf{Note:}

If LP turns out to be "large", i.e.,

\[
LP \geq \ln m , \text{ then } S \text{ can be chosen to be } \Theta(1) \text{, and there is a constant factor approximation.}
\]

\textbf{Solving the LP}

One way to solve the LP is to write an LP which has variables \( f_{i,e} \) where \( i = 1 \ldots k, e \in A \). s.t.
for all i, the \( f_i \)’s constitute an \( \text{unit flow} \) from \( s_i \rightarrow t_i \).

Given \( f_i \)’s, we can perform FLOW DECOMPOSITION to get \( x^* \)’s satisfying the LP-constraint above. This would be the “easier” way to do things.

In class, we looked at a different way (and I missed a crucial step). Here is the full proof.

\[
\begin{align*}
\min & \quad C \\
\text{s.t.} & \quad \sum_{p \in P_i} x_p = 1 \\
& \quad \sum_{e \in E} x_p \geq 0
\end{align*}
\]

\[
\begin{align*}
\max & \quad \sum_{e \in E} \beta_e \\
\text{s.t.} & \quad \sum_{e \in E} \beta_e \leq 1 \\
& \quad \sum_{p \in P_i} \beta_e \geq 0 \\
& \quad \beta \geq 0, \quad \alpha - \text{free}
\end{align*}
\]

* The dual has a separation oracle. Given any \( (\alpha^*, \beta^*) \), since \( \beta \geq 0 \), we can use Dijkstra’s shortest path algorithm to separate.

\[
\therefore \quad \text{We can obtain } (\alpha^*, \beta^*) \text{ a soln to the DUAL LP.}
\]

* By complementary slackness, we know that

\[
\forall p \in P_i : \quad x^*_p > 0 \implies \sum_{e \in E} \beta_e^* = \alpha^*_i \quad \text{must be true.}
\]
\[ \forall p \in P_i : x^* > 0 \implies \sum P_i^* = \alpha_i \] must be true.

.: In the primal LP, we may only keep the variables \{ x_p : p \in P_i \} only for the paths \( p \) at \( \beta^*(p) = \alpha_i \).

But even this could be exponentially many. We don't straightaway know how to enumerate all shortest paths.

Resolution of (5) is that we only care about “linearly independent” collection of paths (and # of cols is \# edges); resolution of (6) is less obvious — how to find a linearly independent “all” of shortest paths. Sounds like an interesting problem.

* Instead, let me give a more general resolution.

\[
\begin{array}{cc}
\text{PRIMAL} & \text{DUAL} \\
\text{exp many vars} & \text{poly many vars} \\
\text{poly n constr.} & \text{exp. many constr.} \\
\end{array}
\]

But Sep. oracle.

What does ellipsoid do? It calls the Sep. oracle at most \( T = \text{poly}(n) \) times to give us the solution to the dual.

Let the Sep. hyperplanes of the dual returned by the s.o. (in our case the shortest path alg.) be \( \{ a_1^T x \geq b_1, a_2^T x \geq b_2, \ldots, a_T^T x \geq b_T \} \).

Look at the new LP

\[
\begin{array}{c}
\text{DUAL} \\
\text{DUAL} \\
\end{array}
\]

Lecture 14 Page 12
DUAL
poly many vars
\[ a_i^T x \geq b_i \]
\[ a_2^T x \geq b_2 \]
\[ a_i^T x \geq b_i \]
only the exp. hyperplanes that we encountered in ellipsoid.

For our problem, DUAL' would look like

\[
\max \sum_{e \in E} \beta_e
- \sum_{e \in E} \beta_e \leq 1
\]

\[
\sum_{e \in P} \beta_e \geq \alpha_i \quad \text{for some}
\]

\[
\text{total } \# \text{ of } P_i \text{'s } \leq T
= \text{poly}(n)
\]

MAIN OBS: If we run ellipsoid on DUAL', then

since ellipsoid is deterministic we will have the SAME BEHAVIOUR on DUAL' as in DUAL.

Whenever we asked for a sep. oracle in DUAL
and returned \( a_i^T x \geq b_i \), say, the same sep. hyp. is present in DUAL' as well!
$\Rightarrow DUAL = DUAL'$

$PRIMAL = PRIMAL'$

The only variables that survive are the $T = \text{poly}(m)$ variables $x_{ij}$. 