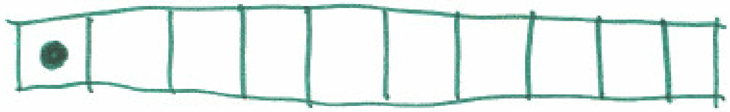
Markov's Inequality

- Many times we have a random variable X and we would like to know how likely it is that X is close to $E[X]$.

- For example, last time we saw that the expected number of moves required for our token to fall off a line of n spaces, starting from the leftmost space, is n:



$$\Pr(\text{left}) = \Pr(\text{right}) = \frac{1}{2}$$

- How likely is it that our token has not fallen off after $2n$ moves? after n^2 moves?

- <u>Markov's Inequality</u> gives us a bound.

<u>Ex</u>: Let $\Omega = \{H, T\}$ $\quad Pr(\{H\}) = 1/2$

$Pr(\{T\}) = 1/2$,

- $\alpha \geq 0$
- $X(H) = \alpha$
- $X(T) = -\alpha$

$$E[X] = \alpha \cdot \frac{1}{2} + (-\alpha) \cdot \frac{1}{2} = 0$$

- With probability 1, $|X(\omega) - \underbrace{E[X]}_{=0}| = \alpha$

Thm (Markov's Inequality)

Let X be a non-negative r.v. (i.e. $\forall \omega \in \Omega$ $X(\omega) \geq 0$). If $t > 0$, then

$$\Pr(X \geq t) \leq \frac{E[X]}{t}.$$

Proof:

$$E[X] = \sum_a a \cdot \Pr(X=a)$$

(b/c X is non-negative)

$$= \sum_{0 \leq a < t} a \cdot \Pr(X=a) + \sum_{a \geq t} a \cdot \Pr(X=a)$$

$$\geq \sum_{a \geq t} a \cdot \Pr(X=a)$$

$$\geq t \sum_{a \geq t} \Pr(X=a)$$

$$= t \cdot \Pr(X \geq t)$$

and dividing by t gives the result. ∎

Application: let X be the number of moves before the token falls off.

By Markov's Inequality,

$$Pr(X \geq t) \leq \frac{E[X]}{t}$$

$$= \frac{n}{t}$$

So, the probability that the token is still alive after $t = 2n$ $(t = n^2)$ is

$$Pr(X \geq 2n) \leq \frac{n}{2n} = \frac{1}{2}$$

$$Pr(X \geq n^2) \leq \frac{n}{n^2} = \frac{1}{n}$$

Therefore with probability at least $1 - \frac{1}{n}$ (i.e. with high probability) the token has fallen after n^2 moves.

The Probabilistic Method: Introduction.

- Method is based on two simple ideas:

(1) If $Pr(B) > 0$, then there is at least 1 outcome $\omega \in \Omega$ in which B occurs (i.e. $\omega \in B$)

(2) If $E[X] = \alpha$, then there is at least 1 outcome $\omega_0 \in \Omega$ for which $X(\omega_0) \leq \alpha$ and at least 1 outcome $\omega_1 \in \Omega$ for which $X(\omega_1) \geq \alpha$.

The strategy to use (1) is as follows. Suppose we want to prove something exists. We construct a probability space and hope that a random object $\omega \in \Omega$ has the desired property. We often consider a collection $A_1, \ldots, A_n$ of "bad" events which prevent the desired property and bound $Pr(\bigcup A_j) < 1$.

The Union Bound.

Prop Let $A_1, A_2, \dots, A_n$ be events and let $A = A_1 \cup A_2 \cup \dots \cup A_n$. We have

$$\Pr(A) \le \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_n).$$

Pf: By induction on n. When $n=1$ the statement is clear. Suppose $n \ge 2$ and let $B = A_1 \cup A_2 \cup \dots \cup A_{n-1}$ and $C = A - B$. Note that if $\omega \in A$ and $\omega \notin B$, it must be that $\omega \in A_n$, so that $C \subseteq A_n$. Also, note $B \cap C = \emptyset$, so that

$$\Pr(A) = \Pr(B \cup C) = \Pr(B) + \Pr(C)$$

(by I.H.) $\le \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_{n-1}) + \Pr(C)$

(because $C \subseteq A_n$, $\Pr(C) \le \Pr(A_n)$.) $\le \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_{n-1}) + \Pr(A_n)$

Example (1): A company organizes its employees into r task forces; each task force has size n and there is no limit on the number of task force groups to which an employee may belong.

The company repairs heaters and air conditioners. ~~Each~~ The company can train each employee in heater repair or airconditioner repairs (but not both.)

Show that after the task forces are organized, the company can train its employees so that each task force has at least one person trained in heater repair and at least one person trained in air conditioner repair, provided that $r \leq 2^{n-1} - 1$.

Solution: Model the problem with set theory.

Each task force is a set of n people; call the task forces $S_1, S_2, \ldots, S_r$. Let $\mathcal{U}$ be the set of employees; so $S_j \subseteq \mathcal{U}$, $|S_j| = n$.

We want to show that we can color the elts of $\mathcal{U}$ with two colors red/blue so that each S_j contains a red employee (heater expert) and a blue employee (air conditioner expert).

Choose a random coloring. That is, for each $x \in \mathcal{U}$ assign x to be red with prob $1/2$ and blue with prob $1/2$; make the assignments independently.

Let B be the event that each S_j contains representatives of both colors.

Want to show $\Pr(B) > 0$. We show instead $\Pr(\overline{B}) < 1$. Let $A = \overline{B}$ be the event that at least one set S_j fails to contain representatives from both colors.

We can break A down into r bad events: let A_j be the event that S_j fails to contain a red elt and a blue elt.

Note: $A = A_1 \cup A_2 \cup \cdots \cup A_r$, but this is not ~~necessarily~~ a disjoint union of events.

Still, by the union bound,

$$\Pr(A) = \Pr(A_1 \cup \cdots \cup A_r)$$

$$\leq \Pr(A_1) + \cdots + \Pr(A_r)$$

$$= r \cdot \frac{2}{2^n}$$

because each A_j has $\Pr(A_j) = \frac{2}{2^n}$.

Therefore if $r \leq 2^{n-1} - 1$, then

$$\Pr(A) \leq r \cdot \frac{2}{2^n} = \frac{r}{2^{n-1}} < 1$$

so there must be at least one way of coloring the elts of $\mathcal{U}$ in which A does not occur and thus in which B does occur. ■

Remark: This is an example of a non-constructive proof technique. Although we know a good coloring exists, we don't know how to find it deterministically any faster than using a brute-force search.

An area of active research, called Derandomization, tries to find ways of converting these Randomized Algorithms into Deterministic Algs.

Example (2):

Recall that if G is a graph, a <u>dominating set</u> is a set $S \subseteq V(G)$ such that each vertex in G is ~~adjacent either~~ in S or has a neighbor in S.

<u>Thm</u> If G is an n-vertex graph in which each vertex has degree at least k and $0 \leq p \leq 1$, then G has a dominating set of size at most $n(p + (1-p)^{k+1})$.

<u>Pf</u>: for each vertex $v \in V(G)$, flip a coin with bias p and put $v \in S_0$ if the coin comes up heads; that is, we define a set $S_0 \subseteq V(G)$ such that we independently choose $v \in S_0$ with prob p and $v \notin S_0$ with prob. $1-p$.

Now, S_0 need not be a dominating set; let

$$S_1 = \{ v \in V(G) \mid v \notin S_0 \text{ and } v \text{ has no neighbors in } S_0 \}.$$

Note that $S = S_0 \cup S_1$ is a dominating set in G. Let $X = |S|$ -- note that X is a random variable.

Moreover, $S_0 \cap S_1 = \emptyset$, so $|S| = |S_0| + |S_1|$; let $Y = |S_0|$ and $Z = |S_1|$. We have that $X = Y + Z$.

Hence, by linearity of expectation,

$$E[X] = E[Y] + E[Z]$$

Using linearity of expectation, $E[Y] = np$; the details are a (straight-forward) exercise.

What about $E[Z]$? Again, use linearity of expectation. For each $v \in V(G)$ define

$$Z_v = \begin{cases} 1 & v \in S_1 \\ 0 & v \notin S_1 \end{cases}$$

so that $Z = \sum_{v \in V(G)} Z_v$ and hence

$$E[Z] = \sum_{v \in V(G)} E[Z_v]$$

$$= \sum_{v \in V(G)} Pr(v \in S_1)$$

(b/c v and all nbrs of v must flip tails for $v \in S_1$ to occur)

$$= \sum_{v \in V(G)} (1-p)^{d(v)+1}$$

$$\leq \sum_{v \in V(G)} (1-p)^{k+1}$$

$$= n(1-p)^{k+1}$$

Therefore $E[X] = E[Y] + E[Z]$

$$\leq np + n(1-p)^{k+1}$$

$$= n(p + (1-p)^{k+1})$$

so there must be some way to flip the coins that leads to a dominating set of size at most $n(p+(1-p)^{k+1})$. ∎

<u>Cor</u> If G is an n-vertex graph with minimum degree k, then G has a dominating set of size at most $\frac{n}{k+1}(1+\ln(k+1))$.

<u>Pf</u>: By the theorem and our inequality $1-x \le e^{-x}$, for each $0 \le p \le 1$, G has a dominating set of size at most

$$n(p+(1-p)^{k+1}) \le n(p+e^{-p(k+1)}).$$

Calculus tells us that $n(p+e^{-p(k+1)})$ is minimized by choosing $p=\frac{1}{k+1}(\ln(k+1))$ and the bound follows, after substituting our formula for p and simplifying.