Name: Chih-Wei Chang	**Andrew ID: chihweic**
Dimension Reduction and Nearest Neighbors	**Date: November 11, 2008**

1. Johnson-Lindenstrauss lemma
   Projecting n points in a d-dimensional space into $K \geq 2\varepsilon^{-2} \ln n$ dimenstions. There is a mapping $f : R^d \rightarrow R^k$ such that $\forall u, v; (1-\varepsilon)||u-v||^2 \leq ||f(u)-f(v)||^2 \leq (1+\varepsilon)||u-v||^2$

   Idea:

   (a) project onto random heperplane, measure distance, and multiply by $\frac{d}{K}$
   (b) Show that squared length of random vector is sharply concentrated around the mean, not distorted by more than $(1 \pm \varepsilon)$ with probability $\frac{1}{n^2}$

   $pr[Y \geq t] \leq \frac{E[Y]}{t}$– Markov Inequality
   $pr[X - \mu_x \geq t\sigma_x] \leq \frac{1}{t^2}, \sigma_x = \sqrt{E[(x-\mu_x)^2]}$, ChebyChev Inequality
   $E[L_k^2] = \frac{k}{d}$, $x_1^2 + x_2^2 + ... + x_d^2 = 1$, choosing k from d (view this process as the flipping coins).

2. Chernoff Bounds
   $x_1, x_2, ..., x_n$ independent posisson trials (0 or 1 outcome, each with probability p of being 1).
   $X = \sum_{i=1} nx_i, \mu_x = np$
   $pr[X > (1+\sigma)\mu] < (\frac{e^\sigma}{(1+\sigma)^{1+\sigma})^\mu}$ $pr[X < (1-\sigma)\mu] < e^{-\mu\sigma/2}$ i.e, $n = 1000, p = 0.5, \sigma = 0.1$

3. proof of Chernoff bounds
   Consider $E[e^{tx}]$,the moment generating function$= 1 + tE[x] + \frac{t^2}{2}e[x^2] + \frac{t^3}{6}e[x^3]$
   $e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + ... + \frac{t^ix^i}{i!}$
   $E[e^{tX}] = E[e^{t\sum_{i=1}^n x_i}] = E[\prod_{i=1}^n e^{tx_i}] = \prod_{i=1}^n E[e^{tx_i}] = \prod_{i=1}^n (pe^t + (1-p)) = \prod_{i=1}^n (1 + p(e^t - 1)) = (1 + p(e^t-1))^n < e^{p(e^t-1)n} = e^{\mu(e^t-1)}$

   $pr[x > (1+\sigma)\mu] = pr[e^{tx} > e^{t(1+\sigma)\mu}] < \frac{E[e^{tx}]}{e^{t(1+\sigma)\mu}} = (\frac{e^{(e^t-1)}}{e^{t(1+\sigma)}})^\mu$