15-859 Algorithms for Big Data Assignment 1

Fan Pu Zeng
fzeng@andrew.cmu.edu

1: Scratcy Scratch

Proof. Hello!

$$
\begin{aligned}
&(a)\\
&\{b\}\\
&[c]
\end{aligned}
$$

□

$$|\mathbf{Pr}\left[S \leq u\right] - \mathbf{Pr}\left[Z \leq u\right]| \leq \text{const} \cdot \beta,$$

where the exact constant depends on the proof, with the best known constant being .5600, and $\beta = \sum_{i=1}^{n} \mathbf{E}\left[|X_i|^3\right]$.

$$Z_n = \frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{\sigma}$$

$$\mathbf{E}\left[e^{tX}\right]$$

$$
\begin{aligned}
M(0) &= \mathbf{E}\left[e^{tX}\right]\Big|_{t=0}\\
&= \mathbf{E}\left[1\right]\\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
M(t) &= \mathbf{E}\left[e^{tX}\right] & &= \mathbf{E}\left[1\right]\\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
M^{(k)}(0) &= \frac{d^k}{dt^k}\mathbf{E}\left[e^{tX}\right]\Big|_{t=0}\\
&= \frac{d^k}{dt^k}\mathbf{E}\left[e^{tX}\right]\Big|_{t=0}
\end{aligned}
$$

$$
\begin{aligned}
M^{(k)}(t) &= \frac{d^k}{dt^k} \mathbf{E}\left[e^{tX}\right] \\
&= \frac{d}{dt} \mathbf{E}\left[X^{k-1} e^{tX}\right] && \text{(by IH)} \\
&= \frac{d}{dt} \int f(x) x^{k-1} e^{tx} \, dx \\
&= \int \frac{d}{dt} f(x) x^{k-1} e^{tx} \, dx \\
&= \int f(x) x^k e^{tx} \, dx \\
&= \mathbf{E}\left[X^k e^{tX}\right].
\end{aligned}
$$

$$
Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A_i
$$

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} A_i &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} \\
&= \sqrt{n} \sum_{i=1}^{n} \frac{X_i - \mu}{n\sigma} \\
&= \sqrt{n} \frac{\overline{X}_n - \mu}{\sigma} \\
&= Z_n.
\end{aligned}
$$

$$
\begin{aligned}
M_{Z_n}(t) &= \mathbf{E}\left[e^{tZ_n}\right] \\
&= \mathbf{E}\left[\exp\left(t \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A_i\right)\right] && \text{(by equivalent definition of } Z_n\text{)} \\
&= \prod_{i=1}^{n} \mathbf{E}\left[\exp\left(\frac{t}{\sqrt{n}} A_i\right)\right] && \text{(by independence of } A_i\text{'s)} \\
&= \prod_{i=1}^{n} M_{A_i}(t/\sqrt{n}) && \text{(definition of } M_{A_i}\text{)} \\
&= M_{A_i}(t/\sqrt{n})^n.
\end{aligned}
$$

$$
f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)(a)}}{n!} (x-a)^n
$$

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)(a)}}{n!}(x-a)^n$$

Our first three moments are

$$\begin{aligned}
M_{A_i}(0) &= \mathbf{E}\left[e^{tA_i}\right]\Big|_{t=0} \\
&= \mathbf{E}[1] \\
&= 1, \\
M'_{A_i}(0) &= \mathbf{E}[A_i] && \text{(by the kth moment property proved previously)} \\
&= 0, \\
M''_{A_i}(0) &= \mathbf{E}\left[A_i^2\right] && \text{(by the kth moment property proved previously)} \\
&= \mathbf{E}\left[A_i^2\right] - \mathbf{E}[A_i]^2 + \mathbf{E}[A_i]^2 \\
&= \mathbf{Var}(A_i) + \mathbf{E}[A_i]^2 && (\mathbf{Var}(A_i) = \mathbf{E}\left[A_i^2\right] - \mathbf{E}[A_i]^2) \\
&= 1 + 0 \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
M_{A_i}(t/\sqrt{n}) &\approx M_{A_i}(0) + M'_{A_i}(0) + M''_{A_i}(0)\frac{t^2}{2n} \\
&= 1 + 0 + \frac{t^2}{2n} \\
&= 1 + \frac{t^2}{2n}
\end{aligned}$$

$$\begin{aligned}
M_{Z_n}(t) &= M_{A_i}(t/\sqrt{n})^n \\
&\approx \left(1 + \frac{t^2}{2n}\right)^n \\
&\to e^{t^2/2}. && \text{(by identity } \lim_{n\to\infty}(1+x/n)^n \to e^x\text{)}
\end{aligned}$$

$$\begin{aligned}
M_Z &= \mathbf{E}\left[e^{tZ}\right] \\
&= \int f_Z(x)e^{tx}\,dx \\
&= \int \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}e^{tx}\,dx && \text{(subst. pdf of standard Gaussian)} \\
&= \int \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2+tx}\,dx \\
&= \int \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2+\frac{1}{2}t^2}\,dx && \text{(completing the square)}
\end{aligned}$$

$$
\begin{aligned}
&= e^{\frac{1}{2}t^2} \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} \, dx && (e^{\frac{1}{2}t^2} \text{ does not depend on } x) \\
&= e^{\frac{1}{2}t^2} \cdot 1 \\
&= e^{\frac{1}{2}t^2},
\end{aligned}
$$

where the second last step comes from the fact that

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}$$

is a probability distribution of a Gaussian with mean

$$t$$

and variance 1, and therefore the integral integrates to 1.