Chapter 2 Concentration of sums of independent random variables

Desh Raj

May 29, 2018

Exercise 2.1.4.

Solution Given, $g \sim \mathcal{N}(0,1)$. On differentiating, we get $g'(t) = -tg(t)$.

$$
\begin{aligned}
\mathbb{E}\, X^2 \mathbb{1}_{X>t} &= \int_t^\infty x^2 g(x) dx \\
&= \int_t^\infty x(xg(x)) dx \\
&= x \int_t^\infty xg(x) dx + \int_t^\infty g(x) dx && \text{(Integration by parts)} \\
&= -xg(x) \Big]_t^\infty + \mathbb{P}(X > t) \\
&= tg(t) + \mathbb{P}(X > t) && \text{(Taking limit of the first term)}
\end{aligned}
$$

This is the desired result.

Exercise 2.2.3. Bounding the hyperbolic cosine

Solution We have,

$$
\begin{aligned}
\cosh(x) &= \frac{e^x + e^{-x}}{2} \\
&= \frac{1}{2}\left(\sum_{i=0}^\infty \frac{x^i}{i!} + \sum_{i=0}^\infty \frac{(-x)^i}{i!} \right) && \text{(Using Taylor series expansion)} \\
&= \sum_{i=0}^\infty \frac{x^{2i}}{(2i)!} \geq \sum_{i=0}^\infty \frac{x^{2i}}{2^i i!} \\
& && \text{(Consider the denominator of the } i^{th} \text{ term. Each product term in } (2i)! \text{ is strictly greater than that in } 2^i i!.) \\
&= \exp\left(\frac{x^2}{2}\right). && \text{(Again using Taylor series expansion)}
\end{aligned}
$$

Exercise 2.2.7.

Solution Given, $X_1, \ldots, X_N$ are independent random variables such that $X_i \in [m_i, M_i]$. Let $S_N = X_1 + \ldots + X_N$. Then, for $t > 0$, we have

$$
\begin{aligned}
\mathbb{P}\{S_N - \mathbb{E}\, S_N \geq t\} &= \mathbb{P}\{e^{s(S_N - \mathbb{E}\, S_N)} \geq e^{st}\} && \text{(For some } s > 0\text{)} \\
&\leq \frac{\mathbb{E}\, e^{s(S_N - \mathbb{E}\, S_N)}}{e^{st}} && \text{(Using Markov's inequality)} \\
&= e^{-st} \prod_{i=1}^{N} \mathbb{E}\left[e^{s(X_i - \mathbb{E}\, X_i)}\right] \\
&\leq e^{-st} \prod_{i=1}^{N} \exp\left(\frac{s^2(M_i - m_i)^2}{8}\right) && \text{(Using Hoeffding's lemma, since } \mathbb{E}(X_i - \mathbb{E}\, X_i) = 0\text{)} \\
&= \exp\left(-st + \frac{s^2}{8} \sum_{i=1}^{N} (M_i - m_i)^2\right).
\end{aligned}
$$

Let $g(s) = \exp\left(-st + \frac{s^2}{8}\sum_{i=1}^{N}(M_i - m_i)^2\right)$. $g(s)$ achieves its minimum at $s = \frac{4t}{\sum_{i=1}^{N}(M_i - m_i)^2}$. Substituting this value in the inequality, we get

$$
\mathbb{P}\{S_N - \mathbb{E}\, S_N \geq t\} \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{N}(M_i - m_i)^2}\right).
$$

Exercise 2.2.8.

Solution Let X_i denote the Bernoulli random variable indicating that the answer is wrong in the i^{th} turn. Then $\mathbb{E}\, X_i = \frac{1}{2} - \delta$, and $X_i \in [0, 1]$. We have to bound the probability that the final answer is wrong with probability ϵ, i.e., $\mathcal{P}$(majority of decisions are wrong) $\leq \epsilon$. From Hoeffding's inequality, we have

$$
\begin{aligned}
&\mathcal{P}\left(\sum_{i=1}^{N} (X_i - \mathbb{E}\, X_i) \geq t\right) \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^{N}(M_i - m_i)^2}\right) \\
\Rightarrow &\mathcal{P}\left(\sum_{i=1}^{N} \left(X_i - \frac{1}{2} + \delta\right) \geq t\right) \leq \exp\left(\frac{-2t^2}{N}\right) && \text{(Since } M_i = 1 \text{ and } m_i = 0\text{)}
\end{aligned}
$$

To get a wrong answer finally, we require that at least half of all answers must be wrong, i.e.

$$
\begin{aligned}
&\sum_{i=1}^{N} X_i \geq \frac{N}{2} \\
\Rightarrow &\sum_{i=1}^{N} \left(X_i - \frac{1}{2} + \delta\right) \geq \delta N
\end{aligned}
$$

Plugging $t = \delta N$ in earlier inequality, we get

$$
\mathcal{P}\left(\sum_{i=1}^{N} \left(X_i - \frac{1}{2} + \delta\right) \geq \delta N\right) \leq \exp\left(-2N\delta^2\right)
$$

For this probbility to be bounded by ϵ, we require that

$$\begin{aligned}\exp(-2N\delta^2) &\leq \epsilon \\ \Rightarrow -2N\delta^2 &\leq \log(\epsilon) && \text{(Taking log both sides)} \\ \Rightarrow N &\geq \frac{\log(\epsilon^{-1})}{2\delta^2}.\end{aligned}$$

Exercise 2.2.9.

Solution Part 1

We want to bound the probability that the sample mean deviates from the true mean by a quantity greater than ϵ. Using Chebyshev's inequality, we have

$$\mathcal{P}\left(\|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\| \geq \epsilon\right) \leq \frac{\operatorname{var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)}{\epsilon^2} \tag{1}$$

Since the samples are drawn i.i.d, we have

$$\begin{aligned}\operatorname{var}(X_1 + \ldots + X_N) &= \operatorname{var}(X_1) + \ldots + \operatorname{var}(X_N) \\ \Rightarrow \operatorname{var}(X_1 + \ldots + X_N) &= N\sigma^2 && \text{(Since var}(X_i)\text{=}\sigma^2\text{)} \\ \Rightarrow \frac{1}{N^2}\operatorname{var}(X_1 + \ldots + X_N) &= \frac{\sigma^2}{N} && \text{(Dividing both sides by } \sigma^2\text{)}\end{aligned}$$

$$\Rightarrow \operatorname{var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{\sigma^2}{N}$$

Plugging this value in 1, we get

$$\mathcal{P}\left(\|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\| \geq \epsilon\right) \leq \frac{\sigma^2}{N\epsilon^2}$$

We want that an ϵ-accurate estimate should be computed with probability at least 3/4, i.e., the error probability should be at most 1/4. Using this bound, we get $\frac{\sigma^2}{N\epsilon^2} \leq \frac{1}{4}$, which implies $N = \mathcal{O}(\frac{\sigma^2}{\epsilon^2})$.

Part 2

Using Theorem 2.2.6 with $X_i \in [-\sigma, \sigma]$, we get

$$\begin{aligned}\mathcal{P}\left(\frac{1}{N}\sum_{i=1}^{N} X_i - \mu \geq \epsilon\right) &= \mathcal{P}\left(\sum_{i=1}^{N}(X_i - \mu) \geq N\epsilon\right) && \text{(Multiplying both sides by } N\text{)} \\ &\leq \exp\left(-\frac{2N^2\epsilon^2}{\sum_{i=1}^{N}(2\sigma)^2}\right) && \text{(From Theorem 2.2.6)} \\ &= \exp\left(-\frac{2N^2\epsilon^2}{4N\sigma^2}\right) \\ &= \exp\left(-\frac{N\epsilon^2}{2\sigma^2}\right)\end{aligned}$$

We want that an ϵ-accurate estimate should be computed with probability at least $1-\delta$, i.e., the error probability should be at most δ. This means that

$$\begin{aligned}
&\exp\left(-\frac{N\epsilon^2}{2\sigma^2}\right) \leq \delta \\
\Rightarrow &-\frac{N\epsilon^2}{2\sigma^2} \leq \log(\delta) && \text{(Taking log both sides)} \\
\Rightarrow &N = \mathcal{O}(\log(\delta^{-1})\frac{\sigma^2}{\epsilon^2}),
\end{aligned}$$

which is the desired result.

Exercise 2.2.10.

Solution Part 1

Let f_{X_i} denote the probability density function. Since X_i is non-negative, it's MGF is given as

$$M_{X_i}(-t) = \int_0^\infty e^{-tx} f_{X_i}(x)dx$$

Since $e^{-tx} \geq 0$, $e^{-tx} f_{X_i} \leq e^{-tx}\max\|f_{X_i}\| = e^{-tx}$ (since $\max\|f_{X_i}\| = 1$). Using this in above equation, we get

$$M_{X_i}(-t) \leq \int_0^\infty e^{-tx}dx = \frac{1}{t}$$

Part 2

We have

$$\begin{aligned}
\mathcal{P}\left(\sum_{i=1}^N X_i \leq \epsilon N\right) &= \mathcal{P}\left(\sum_{i=1}^N -X_i \geq -\epsilon N\right) \\
&= \mathcal{P}\left(\sum_{i=1}^N -\frac{X_i}{\epsilon} \geq -N\right) \\
&= \mathcal{P}\left(\exp(\sum_{i=1}^N -\frac{X_i}{\epsilon} \geq \exp(-N)\right) && \text{(Since exp is monotonously increasing)} \\
&\leq e^N\,\mathbb{E}\exp\left(\sum_{i=1}^N -\frac{X_i}{\epsilon}\right) && \text{(Using Markov inequality)} \\
&= e^n \prod_{i=1}^N \mathbb{E}\exp\left(-\frac{X_i}{\epsilon}\right) \\
&\leq e^N \prod_{i=1}^N \epsilon && \text{(Using result in part 1)} \\
&= (e\epsilon)^N.
\end{aligned}$$