Concentration Inequalities Practice

Kelvin Lee

December 3, 2020

1 Problems

Problem 1 (EECS126 Fa20 MT1). Let's say that for some real $b \geq 1$ that a random variable X is b-reasonable if it satisfies:

$$\mathbb{E}[X^4] \leq b\left(\mathbb{E}[X^2]\right)^2.$$

Suppose X is b-reasonable and that $\mathbb{E}[X] = 0$ and $\mathrm{Var}(X) = \sigma^2$. Prove that

$$\mathbb{P}[|X| \geq t\sigma] \leq \frac{b}{t^4}.$$

Solution 1. We know that $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] = \sigma^2 \implies \sigma = \sqrt{\mathbb{E}[X^2]}$. Then we have $\mathbb{P}[|X| \geq t\sigma] = \mathbb{P}\left[|X| \geq t\sqrt{\mathbb{E}[X^2]}\right]$. Taking the fourth power on both sides removes the absolute value sign because X^4 is non-negative and we have $\mathbb{P}\left[X^4 \geq t^4\mathbb{E}[X^2]^2\right]$. Then we can apply the Markov Inequality to X^4:

$$\mathbb{P}\left[X^4 \geq t^4(\mathbb{E}[X^2])^2\right] \leq \frac{\mathbb{E}[X^4]}{t^4(\mathbb{E}[X^2])^2} \leq \frac{b(\mathbb{E}[X^2])^2}{t^4(\mathbb{E}[X^2])^2} = \frac{b}{t^4}.$$

Note that $t^4(\mathbb{E}[X^2])^2$ is a constant.

Problem 2 (EECS126 Fa18 Final). Let $X_1, \ldots, X_n \overset{i.i.d}{\sim}$ Lognormal (μ, σ^2). Let $Y_k := \left(\Pi_{i=1}^k X_i\right)^{1/k}$. (If X is *log-normally* distributed, then $\ln(X)$ is $\mathcal{N}(\mu, \sigma^2)$)

(a) Find $\mathbb{E}\left[\ln\left(Y_k\right)\right]$.

(b) Find a lower bound on n such that $\mathbb{P}\left(\left|\ln Y_n - \mathbb{E}\left[\ln Y_n\right]\right| > 0.01\right) < 0.05$. You may leave your answer in terms of $\Phi(x)$, the normal CDF. Use the fact that $\mathbb{P}(-1.96 < X < 1.96) = 0.95$.

Solution 2.

(a)

$$\mathbb{E}\left[\ln\left(Y_k\right)\right] = \mathbb{E}\left[\frac{1}{k}\sum_{i=1}^{k} \ln X_i\right] = \frac{1}{k}\cdot k\mathbb{E}[\ln X_1] = \mu.$$

(b) Notice that $\ln Y_n$ is a normal random variable. Let's first compute its variance:

$$\mathrm{Var}\left(\ln Y_n\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} \ln X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\left(\ln X_i\right) = \frac{1}{n^2}\cdot n\,\mathrm{Var}(\ln X_1) = \frac{\sigma^2}{n},$$

where the second inequality follows from the independence of $\{\ln X_i\}_{i=1}^n$. Then we know that $\ln Y_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Also note that $\frac{\ln Y_i - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$. Then

$$\begin{aligned}
\mathbb{P}\left(\left|\ln Y_n - \mathbb{E}\left[\ln Y_n\right]\right| > 0.01\right) &= \mathbb{P}\left(\left|\frac{\ln Y_n - \mathbb{E}[\ln Y_n]}{\sigma/\sqrt{n}}\right| > \frac{0.01}{\sigma/\sqrt{n}}\right) \\
&= \mathbb{P}\left(|\mathcal{N}(0,1)| > \frac{0.01\sqrt{n}}{\sigma}\right) < 0.05 \\
&\implies 2\left(\Phi\left(-\frac{0.01\sqrt{n}}{\sigma}\right)\right) < 0.05 \\
&\implies -\frac{0.01\sqrt{n}}{\sigma} < -1.96 \\
&\implies n > 196^2\sigma^2.
\end{aligned}$$

We multiply by two on the third line because $\Phi\left(-\frac{0.01\sqrt{n}}{\sigma}\right)$ represents the one-tailed probability of X that is less than $-\frac{0.01\sqrt{n}}{\sigma}$ but we need a two-tailed probability due to the absolute value sign (normal distribution is symmetrical). The second to last line uses the fact that $\mathbb{P}(-1.96 < X < 1.96) = 0.95$, which implies that the probability of X being outside of the interval $[-1.96, 1.96]$ is 0.05. We need the probability to be less than that and so we must restrict $-\frac{0.01\sqrt{n}}{\sigma}$ to be less than -1.96.