11 Mar 2022

Chernoff Bound

"The probability that $X_1 + \cdots + X_n$ differs from $\mathbb{E}[X_1 + \cdots + X_n]$ by more than $t\sqrt{n}$ tends to zero exponentially in t, when $X_1, \ldots, X_n$ are independent."

The cumulant generating function of a distribution

If X is a $\mathbb{R}$-valued random var, its cumulant generating function is

$$K_X(t) = \ln \mathbb{E}[e^{tX}].$$

If X, Y are independent

$$\begin{aligned} K_{X+Y}(t) &= \ln\left(\mathbb{E}[e^{t(X+Y)}]\right) \\ &= \ln\left(\mathbb{E}[e^{tX} e^{tY}]\right) \\ &= \ln\left(\mathbb{E}[e^{tX}]\,\mathbb{E}[e^{tY}]\right) \\ &= \ln \mathbb{E}(e^{tX}) \quad \ln \mathbb{E}(e^{tY}) \\ &= K_X(t) + K_Y(t) \end{aligned}$$

If $X \sim N(0,1)$, for any t,

$$\begin{aligned} \mathbb{E}[e^{tX}] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{2}x^2}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2 - 2tx)}\, dx \end{aligned}$$

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

$$= e^{\frac{1}{2}t^2}$$

$$K_x(t) = \ln\left(e^{\frac{1}{2}t^2}\right) = \frac{1}{2}t^2.$$

If $X_1, \dots, X_n$ are identically distributed when $\mathbb{E}X_i = 0$ $\forall i$ and $K_{X_i}(t) = K(t) \dots$

$$K(t) = \ln\left(\mathbb{E}\left(e^{tX_i}\right)\right)$$

$$= \ln\left(\mathbb{E}\left(1 + tX_i + \frac{1}{2}t^2 X_i^2 + \dots\right)\right)$$

$$= \ln\left(1 + t\,\mathbb{E}(X_i) + \frac{1}{2}t^2\,\mathbb{E}(X_i^2) + \dots\right)$$

$$= t\,\mathbb{E}(X_i) + \frac{1}{2}t^2\,\mathrm{Var}(X_i) + O(t^3).$$

So, $\mathrm{Var}(X_i) = \sigma^2$. $K_{X_i}(t) = \frac{1}{2}\sigma^2 t^2 + O(t^3).$

$$Y = \frac{1}{\sqrt{n}}(X_1 + \dots + X_n)$$

$$K_Y(t) = \ln\left(\mathbb{E}\left(e^{\frac{1}{\sqrt{n}}(X_1 + \dots X_n)t}\right)\right)$$

$$= K_{X_1 + \dots + X_n}\left(\frac{t}{\sqrt{n}}\right)$$

$$= n \cdot K\left(\frac{t}{\sqrt{n}}\right)$$

$$= n \cdot \frac{1}{2}\sigma^2\left(\frac{t}{\sqrt{n}}\right)^2 + O\left(n \cdot \left(\frac{t}{\sqrt{n}}\right)^3\right)$$

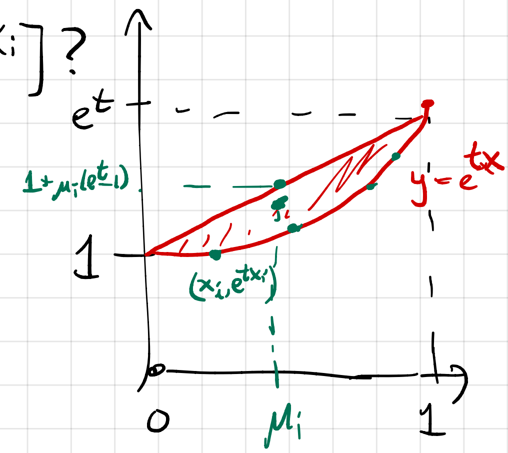
$$= \frac{1}{2}\sigma^2 t^2 + O\left(\frac{t^3}{\sqrt{n}}\right)$$

Say $X_1, X_2, \dots, X_n$ all take values in $[0,1]$.

Then what can we say about $K_{X_i}(t)$?

Let $\mu_i = \mathbb{E}X_i$.

What can we say about $\mathbb{E}[e^{tX_i}]$?



$$e^{t\mu_i} \le \mathbb{E}[e^{tX_i}] \le 1 + \mu_i(e^t - 1)$$

Using $\ln(1+x) \le x$,

$$K_{X_i}(t) \le \mu_i (e^t - 1).$$

Suppose $X = X_1 + \dots + X_n$ and $X_1, \dots, X_n$ indep.

$$K_X(t) = \sum_i K_{X_i}(t) \le \left(\sum_i \mu_i\right)(e^t - 1)$$

$$= \mathbb{E}[X](e^t - 1).$$

What can we say about $\Pr(X > (1+\varepsilon)\mathbb{E}[X])$?

If $t > 0$, $X > (1+\varepsilon)\mathbb{E}(X)$

$$\Updownarrow$$

$$e^{tX} > e^{t(1+\varepsilon)\mathbb{E}(X)}.$$

$$\Pr\left(e^{tX} > e^{t(1+\varepsilon)\mathbb{E}(X)}\right) \le \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\varepsilon)\mathbb{E}[X]}} \quad \text{(Markov's)}$$

$$= \frac{e^{K_X(t)}}{e^{t(1+\varepsilon)\mathbb{E}[X]}}$$

$$= e^{K_X(t) - (1+\varepsilon)\mathbb{E}(X)\, t}$$

$$\leq e^{(e^t - 1 - (1+\varepsilon)t)\mathbb{E}(X)}$$

Set t to minimize $e^t - 1 - (1+\varepsilon)t$:

$$e^t = 1+\varepsilon$$

$$t = \ln(1+\varepsilon)$$

$$e^t - 1 - (1+\varepsilon)t = \varepsilon - (1+\varepsilon)\ln(1+\varepsilon)$$

$$= \varepsilon - (1+\varepsilon)\left(\varepsilon - \tfrac{1}{2}\varepsilon^2 + \tfrac{1}{3}\varepsilon^3 - \tfrac{1}{4}\varepsilon^4 + \cdots\right)$$

$$= \varepsilon - \left(\varepsilon + \tfrac{1}{2}\varepsilon^2 - \tfrac{1}{6}\varepsilon^3 + \tfrac{1}{12}\varepsilon^4 - \tfrac{1}{20}\varepsilon^5 + \cdots\right)$$

$$< -\tfrac{1}{3}\varepsilon^2 \quad \text{for } 0 \leq \varepsilon < 1.$$

Chernoff: If $X_1, \ldots, X_n$ are indep, $[0,1]$-valued

$$\Pr(X_1 + \cdots + X_n > (1+\varepsilon)\mathbb{E}(X_1 + \cdots + X_n))$$
$$< e^{-\frac{1}{3}\varepsilon^2 \mathbb{E}[X_1 + \cdots + X_n]}$$

$$\Pr(X_1 + \cdots + X_n < (1-\varepsilon)\mathbb{E}(X_1 + \cdots + X_n))$$
$$< e^{-\frac{1}{2}\varepsilon^2 \mathbb{E}(X_1 + \cdots + X_n)}$$