Berry-Esseen Theorem

- $X_1, X_2, \cdots$ i.i.d. $E|X_1|^3 < \infty$
- $\left| P\left(\frac{S_n - nEX_1}{\sqrt{n}\,\sigma} \le x \right) - \Phi(x) \right| \le C \frac{E|X_1 - EX_1|^3}{\sigma^3 \sqrt{n}}$

  where $c \le 2$

$X_1, X_2, \cdots$ i.i.d. Bernolli(p)

$S_n \overset{D}{=} \text{Binomial}(n, p)$

$$P(S_n > x) = P\left(\frac{S_n - np}{\sqrt{np(1-p)}} > \frac{x - np}{\sqrt{np(1-p)}} \right)$$

$$\approx P\left(N(0,1) > \frac{x - np}{\sqrt{np(1-p)}} \right)$$

e.g. choose $x = n$

$P(S_n > n) = 0$

relative error of Normal approx. is terrible for $x = n$

How to fix this ?

$$P(S_n > x) = 1 - \Phi\left(\frac{x - nEX_1}{\sqrt{n}\,\sigma} \right) + O\left(\frac{1}{\sqrt{n}}\right)$$

$F(x)$

$\bar{F}(x) = 1 - F(x)$

if

$$\bar{\Phi}\left(\frac{x - nEX_1}{\sqrt{n}\,\sigma} \right) = O\left(\frac{1}{\sqrt{n}}\right)$$

the error term dominates

$\downarrow$

Normal approximation no longer works

$$P(S_n > nEX_1 + x\sqrt{n}\,\sigma)$$

when x is too big

normal approx will be poor

What do we use to replace Normal Approx when x is large ?

$P(S_n > n EX_1 + x_n \sqrt{n}\, \sigma)$

make x a func of n

Large Deviations

$$X_n = O(\sqrt{n})$$

- $X_1, X_2, \ldots$ i.i.d
- $P(S_n > na) \qquad (a > EX_1)$
- $E \exp(\theta X_i) < \infty \quad . \quad \theta > 0 \qquad$ "light-tailed"

$$P(S_n > na) \approx e^{-nI(a)}$$ rate function

Heavy-tailed Setting

If the X_i's are heavy-tailed, we get different behavior:

$$P(X_1 > x) = (1+x)^{-\alpha}, \quad x \geq 0$$

$$\Rightarrow P(X_1 > na) \sim n^{-\alpha} x^{-\alpha} \quad \text{as} \quad n \to \infty$$

asymptotic to

$a_n \sim b_n$ iff $a_n / b_n \to 1$ as $n \to \infty$

$$P(S_n > x) \sim n\, P(X_1 > x) \quad \text{as} \quad x \to \infty$$

any x very big

The sum gets large when a single one of the X_i's is big (heavy-tailed).

"single big jump"

Light-tailed Setting

$$P(S_n > na) \approx e^{-nI(a)}$$

$$\frac{1}{n} \log P(S_n > na) \to -I(a) \quad \text{as} \quad n \to \infty$$

Markov Inequality:

$$P(W > w) \le \exp(-\theta w)\, E\exp(\theta W), \quad \theta > 0$$

$$\Rightarrow \quad P(S_n > na) \le \exp(-\theta n a)\, E\exp(\theta S_n)$$

$$E\exp(\theta S_n) = E\exp\left(\theta \sum_i^n X_i\right)$$

$$= \prod_{i=1}^n E\exp(\theta X_i)$$

$$= (E\exp(\theta X_i))^n$$

$$= e^{n\varphi(\theta)}$$

$$\varphi(\theta) = \log E\exp(\theta X_i)$$ "cumulant generating fcn"

$$P(S_n > na) \le \exp(-n(\theta a - \varphi(\theta)))$$

Choose θ to maximize $\theta a - \phi(\theta)$

$$\frac{d}{d\theta} = a - \varphi'(\theta(a)) = 0$$

$$\boxed{\varphi'(\theta(a)) = a}$$

$$P(S_n > na) \le \exp(-n I(a))$$

$$I(a) \triangleq \theta a - \phi(\theta(a))$$

$$\varphi(\theta) = \log E e^{\theta X_1}$$

$$\varphi'(\theta) = \frac{1}{Ee^{\theta X_1}} \frac{d}{d\theta} E e^{\theta X_1}$$

$$= \frac{1}{Ee^{\theta X_1}} E X_1 e^{\theta X_1}$$

Dominant Convergence Thm
$\rightarrow$ bring d into E

$$\varphi'(\theta) = \frac{\int_{\mathbb{R}} x e^{\theta x} F(dx)}{\int_{\mathbb{R}} e^{\theta x} F(dx)}$$

$$= \int_{\mathbb{R}} x F_\theta(dx)$$

$$F_\theta(dx) = \frac{e^{\theta x} F(dx)}{\int_{\mathbb{R}} e^{\theta y} F(dy)}$$

"exponentially tilted version of $F(\cdot)$"

($f(x) \Rightarrow e^{\theta x} f(x)$, exponentially more likely)

$\theta > 0$, $\int_{\mathbb{R}} x F_\theta(dx) > \int_{\mathbb{R}} x F(dx)$

$$\varphi'(\theta) = \int_{\mathbb{R}} x F_\theta(dx)$$

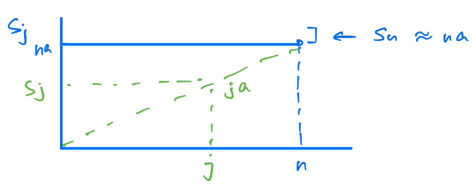
$$\varphi'(\theta(a)) = a$$

means

$$\int x F_{\theta(a)}(dx) = a$$

In other words,

$$E_{\theta(a)} X_1 = a$$



$$P(S_n > na) \in \exp(-n(\underbrace{a\theta(a) - \varphi(\theta(a))}_{I(a) > 0}))$$

$\{S_n > na\}$ very rare

$$E(X_1 \mid S > na)$$
$$\approx E(X_1 \mid S \approx na)$$
$$\approx E(X_1 \mid S = na)$$

$$E(\underbrace{X_1, \dots + X_n}_{\sum_{i=1}^{n} E[X_1 \mid S_n = na)} \mid S_n = na) = na$$

$$n\, E[X_1 \mid S_n = na) = na$$

$$E[X_1 \mid S_n = na) = a$$

We've seen the exponential tilts arise in the upper bound for $P(S_n > na)$

Maybe it appears intrinsically in deciding the lower bound on $P(S_n > na)$

$$P(S_n \in B) = \int_{\{x_1+\cdots x_n \in B\}} f(x_1)f(x_2)\cdots f(x_n)\, dx_1 \cdots dx_n$$

$$= \int_{\{x_1+\cdots+x_n \in B\}} \frac{\prod_1^n f(x_i)}{\prod_1^n f_\theta(x_i)} \prod_1^n f_\theta(x_i)\, dx_i$$

$$= E_\theta\, I(S_n \in B) \prod_{i=1}^n \frac{f(x_i)}{f_\theta(x_i)}$$

$$\frac{f(x)}{f_\theta(x)} = \frac{f(x)}{e^{\theta x} f(x) e^{-\varphi(\theta)}} = e^{-\theta x + \varphi(\theta)}$$

$$\Rightarrow \quad P(S_n \in B) = E_\theta\, I(S_n \in B)\, e^{-\theta \sum_{i=1}^n x_i + n\varphi(\theta)}$$

$$= E_\theta\, I(S_n \in B)\, e^{-\theta S_n + n\varphi(\theta)}$$

$$P(S_n > na) = E_\theta\, I(S_n > na)\, e^{-\theta S_n + n\varphi(\theta)}$$

Choose $\theta = \theta(a)$

$$P(S_n > na) = E_{\theta(a)}\, I(S_n > na)\, e^{-\theta(a) S_n + n\varphi(\theta(a))}$$

Under $P_{\theta(a)}$, $X_1, \cdots, X_n$ i.i.d.

with common tilted $F_{\theta(a)}(\cdot)$

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} E_{\theta(a)} X_1 = a$$

$$S_n \approx na$$

"Change of Measure"

We modified the measure from the "nominal" measure

to a specially chosen prob. that makes the calculation easier