Feynman Liang
STAT 205A
Lecture #12
10/4/2016

1 Large Deviation Theorem

If $a_n \sim ce^{\beta n}$ as $n \to \infty$, then $\frac{1}{n} \log a_n \to \beta =$ "asymptotic growth rate."

Today $\beta < 0$. ~~growth~~ decay

IID (X_i). $S_n = \sum_i^n X_i$. $\mathbb{E}X = \mu$. Fix $a > \mu$, $P(X \geq a) > 0$.

Consider $P\left(\frac{S_n}{n} \geq a\right)$. This $\to 0$ as $n \to \infty$ by WLLN. How fast? We already have

Proposition 1.1 (General large-deviation inequality)**.**

$$P(Y \geq y) \leq \inf_{\theta \geq 0} \frac{\mathbb{E}e^{\theta Y}}{e^{\theta y}} \tag{1.1}$$

Definition 1.2. The *transform* of X

$$\phi(\theta) = \mathbb{E}\exp(\theta X) \tag{1.2}$$

Assume $\theta^* = \sup\{\theta : \phi(\theta) < \infty\} > 0$. Going back to "how fast?"

By general LD inequality and independence of X_i

$$\begin{aligned}
P\left(\frac{S_n}{n} \geq a\right) &= P(S_n \geq an) && (1.3)\\
&\leq \inf_\theta \frac{\mathbb{E}\exp(\theta S_n)}{\exp(\theta an)} = \inf_\theta \frac{\mathbb{E}\prod_{i=1}^n \exp(\theta X_i)}{\exp(\theta an)} && (1.4)\\
&= \inf_\theta \frac{\prod_{i=1}^n \mathbb{E}\exp(\theta X_i)}{\exp(\theta an)} && (1.5)\\
&= \left(\inf_\theta \frac{\phi(\theta)}{e^{\theta a}}\right)^n && (1.6)\\
\frac{1}{n}\log P\left(\frac{S_n}{n} \geq a\right) &\leq \inf_\theta \left[\log\phi(\theta) - a\theta\right] && (1.7)
\end{aligned}$$

So we have

$$\forall n : \frac{1}{n}\log P\left(\frac{S_n}{n} \geq a\right) \leq \inf_\theta \left[\log\phi(\theta) - a\theta\right] \tag{1.8}$$

Theorem 1.3. *As $n \to \infty$*

$$\lim \frac{1}{n}\log P\left(\frac{S_n}{n} \geq a\right) = \inf_\theta \left[\log\phi(\theta) - a\theta\right] \tag{1.9}$$

1.1 Proof outline

Three steps:

(a) Analysis of $\phi(\theta)$

(b) Tilting lemma

(c) Put together

Lemma 1.4. $\phi'(0+) = \mu$

$$\frac{d}{d\theta}\phi(\theta) = \frac{d}{d\theta}\mathbb{E}e^{\theta X} \stackrel{?}{=} \mathbb{E}\frac{d}{d\theta}e^{\theta X} = \mathbb{E}[Xe^{\theta X}] \qquad \forall\theta \tag{1.10}$$

$$\theta = 0 : \phi'(0+) = \mathbb{E}X \tag{1.11}$$

How to justify "?" in detail?

Proof. Know $\frac{e^{\theta X}-1}{\theta} \to X$ as $\theta \downarrow 0$. **Want:** $\mathbb{E}\left[\frac{e^{\theta X}-1}{\theta}\right] \to \mathbb{E}X$.

Tool for going from sequence convergence to convergence of an expectation: Dominated Convergence Theorem.

$$x > 0 : e^{\theta X} - 1 = \int_0^{\theta x} e^y dy \leq \theta x e^{\theta x}] \tag{1.12}$$

$$x < 0 : |e^{\theta x} - 1| = \int_{\theta x}^0 e^y dy \leq |\theta x| \tag{1.13}$$

$$\implies |e^{\theta x} - 1| \leq \theta |x| \max(1, e^{\theta x}) \tag{1.14}$$

For $0 < \theta \leq \theta_0$

$$\mathbb{E}\left[\frac{e^{\theta X}-1}{\theta}\right] \leq |X| \max(1, e^{\theta_0 X}) \tag{1.15}$$

Hypothesis: $\exists\theta_1$ such that $\mathbb{E}E^{\theta_1 X} < \infty$.
Choose $\theta_0 < \theta_1$
$\implies \mathbb{E}\left[|X|\max(1, e^{\theta X})\right] < \infty$
Now $\mathbb{E}\left[\frac{e^{\theta X}-1}{\theta}\right] \ll |X|\max(1, e^{\theta X})$ and $|X|\max(1, e^{\theta X})$ is integrable.
Applyl DCT. □

Lemma 1.5. $\phi'(0+) = \mu$ *and for* $0 < \theta < \theta^*$

$$\phi'(\theta) = \mathbb{E}[Xe^{\theta X}] \tag{1.16}$$

$$\phi''(\theta) = \mathbb{E}[X^2 e^{\theta X}] \tag{1.17}$$

Proof. Write out as integrals, apply Fubini-Tonelli theorem. □

Suppose X discrete (so $\phi(\theta) = \sum_x e^{\theta x} P(X = x)$). Fix θ. Define a dist for $\hat{X}$ by

$$P(\hat{X} = x) = \frac{e^{\theta x} P(X = x)}{\phi(\theta)} \tag{1.18}$$

$$\mathbb{E}\hat{X} = \sum_x x P(\hat{X} = x) = \frac{\sum_x x e^{\theta x} P(X = x)}{\phi(\theta)} \tag{1.19}$$

$$= \frac{\mathbb{E} X e^{\theta X}}{\phi(\theta)} = \frac{\phi'(\theta)}{\phi(\theta)} = \frac{d}{d\theta} \log \phi(\theta) \tag{1.20}$$

$$\mathbb{E}(\hat{X}^2) = \frac{\mathbb{E}[X^2 e^{\theta X}]}{\phi(\theta)} = \frac{\phi''(\theta)}{\phi(\theta)} \tag{1.21}$$

$$\mathrm{Var}(\hat{X}) = \mathbb{E}(\hat{X}^2) - (\mathbb{E}\hat{X})^2 \tag{1.22}$$

$$= \frac{\phi''(\theta)}{\phi(\theta)} - \left(\frac{\phi'(\theta)}{\phi(\theta)}\right)^2 \tag{1.23}$$

$$= \frac{d}{d\theta}\left(\frac{\phi'(\theta)}{\phi(\theta)}\right) = \frac{d}{d\theta} \log \phi(\theta) \tag{1.24}$$

The transform of X, $\phi(\theta)$, encodes information abouut the moments.
For general X, define distribution of $\hat{X}$ by Radon-Nikodyn density ($\frac{d\nu}{d\mu}$)

$$\frac{dP(\hat{X} \in \cdot)}{dP(X \in \cdot)}(x) = \frac{e^{\theta x}}{\phi(\theta)} \tag{1.25}$$

Lemma 1.6. $\mathbb{E}\hat{X} = \frac{d}{d\theta} \log \phi(\theta)$, $Var\hat{X} = \frac{d^2}{d\theta^2} \log \theta$

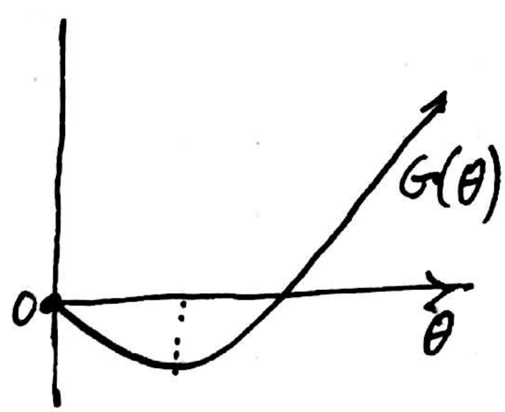
1.1.1 Study $G(\theta) = \log \phi(\theta) - a\theta$

$$G'(0+) = (\log \phi(\theta))' - a = \mu - a \tag{1.26}$$

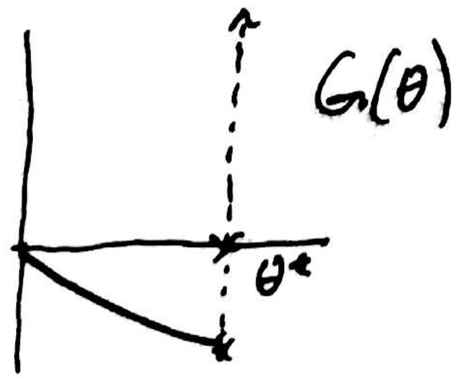
$$G''(\theta) = \mathrm{Var} X_\theta > 0 \text{ on } 0 < \theta < \theta^* \tag{1.27}$$

$$G(0) = 0 \tag{1.28}$$

So G is strictly convex on $(0, \theta^*)$. Easy to show $G(\theta) \to \infty$ as $\theta \to \infty$.



Find $\inf_\theta$ of $G(\theta)$ by solving $G'(\theta) = 0 \implies \frac{\phi'(\theta)}{\phi(\theta)} = a$.



Case 1: $\exists$ solution $\theta_a \in (0, \theta^*)$ of equation $\frac{\phi'(\theta)}{\phi(\theta)} = a$
Assume case 1. Choose $\theta \in (\theta_a, \theta^*)$. Consider tilded distribution $\hat{X} = \hat{X}_\theta$.

$$\mathbb{E}\hat{X} = \frac{d}{d\theta} \log \phi(\theta) > \frac{d}{d\theta} \log \phi(\theta) \mid_{\theta=\theta_a} \tag{1.29}$$

TODO: ??? $\mathbb{E}\hat{X} > a$ and $\mathbb{E}\hat{X}_\theta \downarrow a$ as $\theta \downarrow \theta_a$ [Check!].
Fix $b > \mathbb{E}\hat{X}_\theta$.
Trick: Apply WLLN to tilted $(\hat{X}_i)$

$$\frac{P(\hat{X}_1 = x_1, \cdots, \hat{X}_n = x_n)}{P(X_1 = x_1, \cdots, X_n = x_n)} = \frac{e^{\theta \sum_i^n x_i}}{\phi^n(\theta)} \implies \underbrace{\frac{P(\hat{S}_n = s)}{P(S_n = s)}}_{\text{interp as Radon-Nikodyn density}} = \frac{e^{\theta s}}{\phi^n(\theta)} \tag{1.30}$$

$$\frac{P(y_1 \le \hat{S}_n \le y_2)}{P(y_1 \le S_n \le y_2)} \le \frac{e^{\theta y_2}}{\phi^n(\theta)} \tag{1.31}$$

$$P(a \le \frac{S_n}{n} \le b) \ge e^{-\theta b_n} \phi^n(\theta) \tag{1.32}$$

$$P(a \le \frac{S_n}{n} \le b) \to 1 \qquad \text{as } n \to \infty \tag{1.33}$$

where we use $y_1 = an, y_2 = bn$.

$$\liminf_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} \ge a) \ge -b\theta + \log \phi(\theta) \tag{1.34}$$

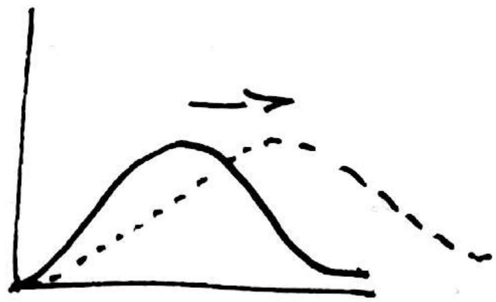
Let $\theta \downarrow \theta_a$.

$$\liminf_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} \ge a) \ge -b\theta_a + \log \phi(\theta_a) \tag{1.35}$$

True $\forall b > a$, let $b \downarrow a$

$$\liminf_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} \geq a) \geq -a\theta_a + \log \phi(\theta_a) = G(\theta_a) \tag{1.36}$$

Why is it called tilting?



2 Intro to next lecture

Suppose X, Y continuous.

TODO: Format better

$$P(x,y) = P(X=x, Y=y) \leftrightarrow \qquad f(x \tag{2.1}$$

$$\text{marginal dist} p_X(x) = P(X=x) \leftrightarrow \qquad f_X \tag{2.2}$$

$$\text{conditional dist of } Y \text{ given } X = x p_{Y|X}(y|x) = P(Y=y|X=x) \leftrightarrow \quad y \mapsto f_{Y|X}(y|x) \text{ conditional density c} \tag{2.3}$$

$$\text{reltiaon} p(x,y) = p_X(x) p_{Y|X}(y,x) \leftrightarrow \tag{2.4}$$