Feynman Liang
STAT 205A
Lecture #5
9/8/2016

1 Expectation and Inequalities

1.1 Expectation (Undergrad version)

(1) $\mathbb{E}X$ is the limit of $\frac{X_1+X_2+\cdots+X_n}{n}$ for iid (will prove as SLLN)

(2) $\mathbb{E}X$ is fair stake for random payoff X (conceptual basis of martingale theory)

(3) $\mathbb{E}X = \sum_i iP(X=i)$ or $\int xf(x)dx$ (change of variable in MT, last lecture)

(4) $\mathbb{E}h(X) = \sum_i h(i)P(X=i)$ or $\int h(x)f(x)dx$ (change of variable in MT, last lecture)

(5) abstract rules: $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$ even if dependent

1.2 Measure-theoretic version

Let $X : (\Omega, \mathcal{F}, P \to \mathbb{R}$ be a random variable on a probability space.

Definition 1.1. The *expectation* $\mathbb{E}X := \int_\Omega X(\omega)P(d\omega)$

Expectation is well-defined if:

(a) $\mathbb{E}X < \infty$ or $0 \leq X \leq \infty$, where $0 \leq \mathbb{E}X \leq \infty$

(a) $\implies -\infty < \mathbb{E}X < \infty$

From definition 1.1, can use proporties of abstract $\int$

- $\mathbb{E}1_A = P(A)$
- $\mathbb{E}(c_1X_1 + c_2X_2) = c_1\mathbb{E}X_1 + c_2\mathbb{E}X_2$ (Linearity)
- (Monotone Convergence): If $0 \leq X_1 \leq X_2 \leq \cdots \leq \infty$, $X_n \uparrow X$a.s., then $\mathbb{E}X_n \uparrow \mathbb{E}X \leq \infty$
  (a) a.s. means for all ω outside some A where $P(A) = 0$
  (b) To prove this for a.s., consider $0 \leq X_1 1_{A^c} \leq X_2 1_{A^c} \leq \cdots$, then $X_n 1_{A^c} \uparrow X 1_{A^c} \forall \omega$ and $\mathbb{E}X_n 1_{A^c} \uparrow \mathbb{E}X 1_{A^c}$

Example 1.2. $X \geq 0$. $\mathbb{E}X < \infty \implies P(X < \infty) = 1$. However, $P(X \leq \infty) = 1 \not\Longrightarrow \mathbb{E}X < \infty$.

Consider $P(X = i) \sim ci^{-3/2}$.

1.3 Inequalities

Lemma 1.3 (Markov's Inequality)**.** *If $X \geq 0$, $\mathbb{E}X < \infty$, then $P(X \geq x) \leq \frac{\mathbb{E}X}{x}$, $0 < x < \infty$.*

Definition 1.4. If $\mathbb{E}X^2 < \infty$, the *variance* $\mathrm{Var}(X) := \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}(X - \mathbb{E}X)^2$ and $0 \leq \mathrm{Var}(X) < \infty$.

Lemma 1.5 (General form of Markov's inequality)**.** *Let $\phi : \mathbb{R} \to [0, \infty)$ be increasing. Then $P(X \geq x) \leq \frac{\mathbb{E}\phi(X)}{\phi(X)}$ provided not indeterminate (e.g. $\frac{0}{0}$).*

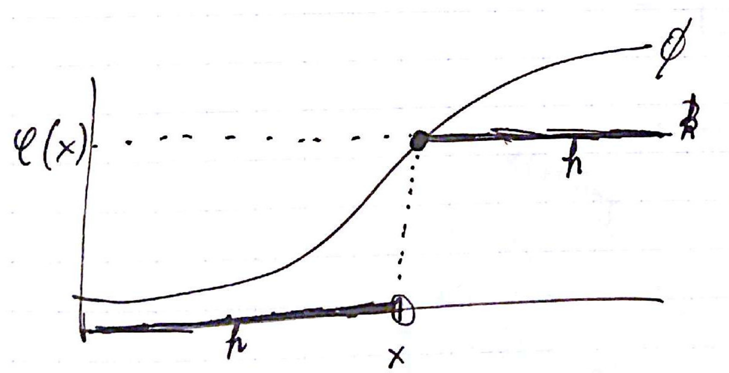

Figure 1: Illustration of $h(x) \leq \phi(x) \forall x$

Proof. Define $h(y) = \begin{cases} 0, & \text{if } y < x \\ \phi(x), & \text{if } y \geq x \end{cases} = \phi(x) 1_{y \geq x}$.

Clear $h(y) \leq \phi(y) \forall y$.

$\mathbb{E}\phi(X) \geq \mathbb{E}h(X) = \phi(X)\mathbb{E}1_{X \geq x} = \phi(x)P(X \geq x)$ □

Lemma 1.3 is lemma 1.5 with $\phi(x) = x^+ = \max(0, x)$.

Lemma 1.6 (Chebychev's Inequality)**.** *If $Var(X) < \infty$, then $P(|X - \mathbb{E}X| \geq x) \leq \frac{Var(X)}{x^2}$ for $0 < x < \infty$.*

Proof. Take $Y = |X - \mathbb{E}X|$ and $\phi(x) = (x^+)^2$ in lemma 1.5. For $x > 0$

$$P(Y \geq x) \leq \frac{\mathbb{E}Y^2}{x^2} = \frac{\mathrm{Var}(X)}{x^2} \tag{1.1}$$

□

Another case is to take $\phi(x) = e^{\theta x}$ for parameter $\theta > 0$ and $0 < x < \infty$

$$P(X \geq x) \leq \frac{\mathbb{E}e^{\theta X}}{e^{\theta x}} \tag{1.2}$$

In particular

Lemma 1.7 (Basic Large Deviation Inequality). *For* $0 < x < \infty$

$$P(X \geq x) \leq \inf_{\theta > 0} \frac{\mathbb{E}e^{\theta X}}{e^{\theta x}} \tag{1.3}$$

(Only useful if $P(X \geq x) \to 0$ *exponentially fast)*

Example 1.8. $X \sim \text{Poisson}(\lambda)$, $\mathbb{E}X = \lambda$, $\text{Var}X = \lambda$.
By lemma 1.5: $P(X \geq x) \leq \lambda/x$.
By lemma 1.6: $P(X \geq x) \leq \frac{\lambda}{(x-\lambda)^2}$

$$\mathbb{E}e^{\theta X} = \sum_i e^{\theta i} e^{-\lambda} \lambda^i = e^{-\lambda} e^{\lambda e^\theta} \tag{1.4}$$

$$P(X \geq x) \leq \inf_\theta \exp(\underbrace{-\theta x - \lambda + \lambda e^\theta}_{(*)}) \tag{1.5}$$

$$= \exp(-x \log \frac{x}{\lambda} - \lambda + x) \tag{1.6}$$

$$\frac{d}{d\theta}(*) = -x + \lambda e^\theta \tag{1.7}$$

Take θ such that $\lambda e^\theta = x$.

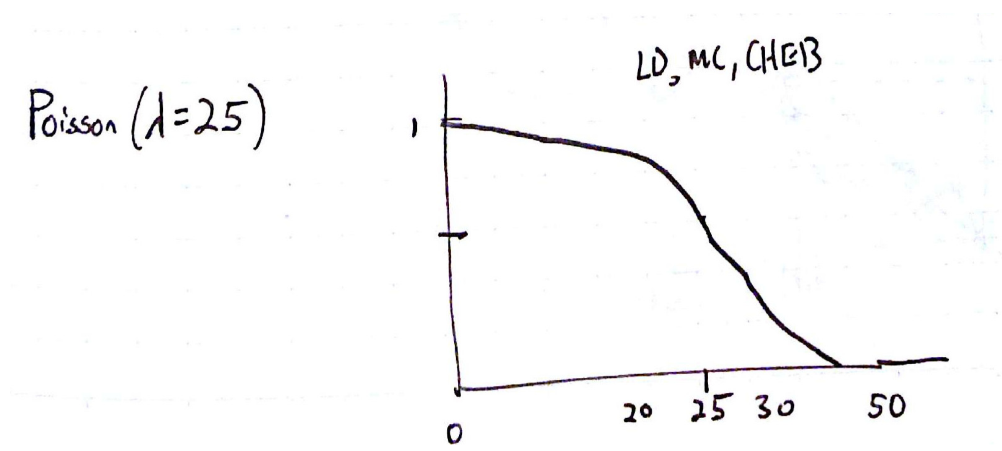

Figure 2: TODO: Draw LD, MC, and Cheb bounds

Lemma 1.9 (Cauchy-Schwarz Inequality).

$$|\mathbb{E}(XY)| \leq \sqrt{(\mathbb{E}X^2)(\mathbb{E}Y^2)} \tag{1.8}$$

Proof (Trick!) Recall quadratic equation: for $a > 0$

$$ax^2 + 2bx + c \geq 0 \forall x \iff b^2 \leq ac \tag{1.9}$$

Applying

$$\mathbb{E}\underbrace{(X + xY)^2}_{\geq 0\ \forall x} = \underbrace{\mathbb{E}(Y^2)}_{a>0} \cdot x^2 + 2\underbrace{\mathbb{E}(XY)}_{b} \cdot x + \underbrace{\mathbb{E}X^2}_{c} \tag{1.10}$$

$$\implies b^2 \leq ac \tag{1.11}$$

□

Example 1.10. Given $(x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, take $P(X = x_i, Y = y_i) = \frac{1}{n}$ for $1 \leq i \leq n$. Then C-S yields

$$\left|\frac{1}{n}\sum_i x_i y_i\right| \leq \sqrt{\left(\frac{1}{n}\sum_i x_i^2\right)\left(\frac{1}{n}\sum_i y_i^2\right)} \tag{1.12}$$

Definition 1.11. ϕ is *convex* if $\forall x < y, \lambda \in [0,1], \phi(x + \lambda(y - x)) + \lambda(\phi(y) - \phi(x))$.

In practice, $\phi''(x) \geq 0 \implies \phi$ is convex.

Lemma 1.12 (Jensen's inequality). *Interval $I \subset \mathbb{R}$, let $\phi : I \to \mathbb{R}$ be convex. Then $\phi\mathbb{E}Xx \leq \mathbb{E}\phi X$ provided both expectations are well-defined.*

Proof. Intuition:

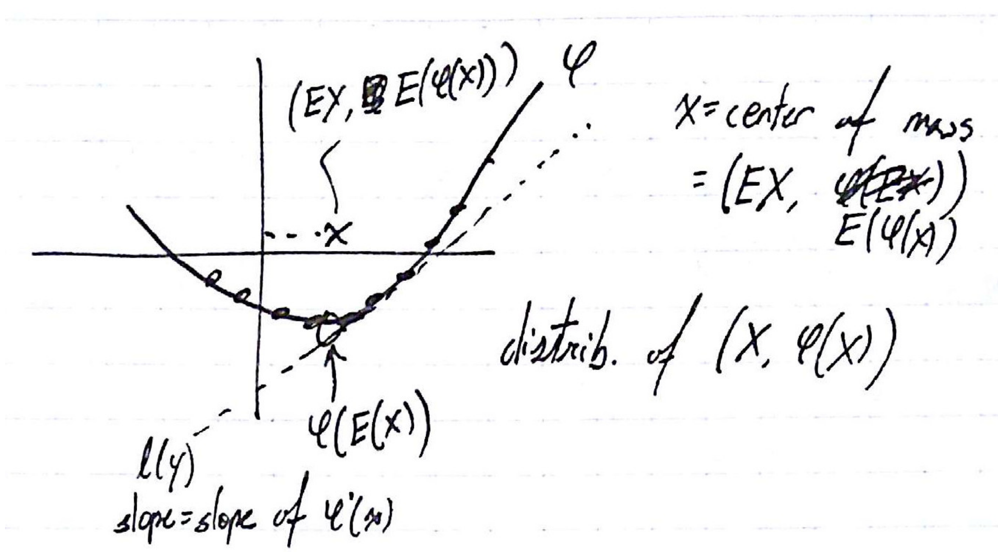

Figure 3: Illustration of Jensen's inequality and tangent line

Given x and convex ϕ, $\exists$ tangent line $l(y)$ such that $l(y) \leq \phi(y)\ \forall y$ and $l(x) = \phi(x)$.

Set $x = \mathbb{E}X$, take tangent $l(\cdot)$ at x.

$$\begin{aligned}
\phi(X) &\geq l(x) && (1.13)\\
\mathbb{E}\phi(X) &\geq \mathbb{E}l(x) && (1.14)\\
&= l(\mathbb{E}X) \qquad\qquad l \text{ linear} && (1.15)\\
&= l(x) = \phi(x) = \phi(\mathbb{E}X) && (1.16)
\end{aligned}$$

□

Example 1.13. $\phi(x) = |x|^p$, $1 \leq p$. Then Jensen's inequality says

$$|\mathbb{E}Y|^p \leq \mathbb{E}|Y|^p \tag{1.17}$$

Applying this with $0 < a < b < \infty$, $y = |X|^a$, $p = \frac{b}{a}$, shows

$$(\mathbb{E}|X|^a)^{b/a} \leq \mathbb{E}|X|^b \tag{1.18}$$

$$(\mathbb{E}|X|^a)^{1/a} \leq \left(\mathbb{E}|X|^b\right)^{1/b} \tag{1.19}$$

The L^p *norm* is $\|X\|_p := (\mathbb{E}|X|^p)^{1/p}$, $p \in [1, \infty)$ so this result says $p \mapsto \|X\|_p$ is increasing on $p \in [1, \infty)$.

Example 1.14. For $x \in (0, \infty)$, consider

(1) $\phi(x) = 1/x$

(2) $\phi(x) = -\log x$,

If $x > 0$, then $\mathbb{E}\phi(X) \geq \phi(\mathbb{E}X)$. Applying Jensen's

(1) $\mathbb{E}\frac{1}{x} \geq \frac{1}{\mathbb{E}X} \iff \mathbb{E}X \geq \frac{1}{\mathbb{E}\frac{1}{X}}$

(2) $-\mathbb{E}\log X \geq -\log \mathbb{E}X \iff \mathbb{E}X \geq e^{\mathbb{E}\log X}$

Consider $(x_i)_{i=1}^n > 0$, $P(X = x_i) = \frac{1}{n}$ $1 \leq i \leq n$.

$$\underbrace{\frac{1}{n}\sum_i}_{\text{Arithmetic mean}} \geq \underbrace{\frac{1}{\frac{1}{n}\sum_i \frac{1}{x_i}}}_{\text{Harmonic mean}} \tag{1.20}$$

$$\frac{1}{n}\sum_i \geq e^{\frac{1}{n}\sum_i \log x_i} = \underbrace{\left(\prod_i x_i\right)^{1/n}}_{\text{Geometric mean}} \tag{1.21}$$