# LECTURE 19

- **Readings:** Sections 5.1–5.3

## Lecture outline

- $M_n = (X_1 + X_2 + \cdots + X_n)/n$ and its limits
- Markov inequality
- Chebyshev inequality
- Convergence in probability
- Weak law of large numbers (convergence in probability of $M_n$)



## Fourth quarter of the course

- Combining probabilistic modeling and data
- Ch. 5: Limit theorems (omit Section 5.5)
  - Emphasis on sample mean of sample $X_1, X_2, \ldots, X_n$
$$M_n \;=\; \frac{X_1 + X_2 + \cdots + X_n}{n}$$
    but also look at other sequences
  - A rationale for the importance of the normal distribution
- Make **inferences** about parameters of a model from data
- Ch. 8: Bayesian statistical inference
  - data come from a model with random parameters
- Ch. 9: Classical statistical inference
  - data come from a model with non-random parameters



## Sample mean $M_n$

- Let $X_1, X_2, \ldots$ be independent and identically distributed with $\mathbf{E}[X_i] = \mu$ and $\mathrm{var}(X_i) = \sigma^2$
- Can we use $n$ samples to estimate $\mu$?
- Form sample mean
$$M_n \;=\; \frac{X_1 + X_2 + \cdots + X_n}{n}$$
  (a random variable)
  - $\mathbf{E}[M_n] \;=$
  - $\mathrm{var}(M_n) \;=$
  - What happens as $n \to \infty$?



## Markov inequality

- Any nonnegative-valued random variable with a mean has a limited probability of being much larger than its mean
- Let $X$ be any random variable that takes only nonnegative values. Let $a$ be any positive number. Then
$$\mathbf{P}(X \geq a) \;\leq\; \mathbf{E}[X]/a.$$
  - Proof: Define a convenient function of $X$:
$$Y_a \;=\; \begin{cases} 0, & \text{for } X \in [0, a); \\ a, & \text{for } X \in [a, \infty). \end{cases}$$
$$\mathbf{E}[X] \;\geq\; \mathbf{E}[Y_a] \;=\; a\,\mathbf{P}(X \geq a)$$



## Chebyshev inequality

- Any random variable with a mean and a variance is unlikely to differ greatly from its mean
- Let $X$ be a random variable with $\mathbf{E}[X] = \mu$ and $\mathrm{var}(X) = \sigma^2$. Let $c$ be any positive number. Then
$$\mathbf{P}(|X - \mu| \geq c) \;\leq\; \frac{\sigma^2}{c^2}.$$
  - Proof: Apply Markov inequality to $(X - \mu)^2$.



## Chebyshev inequality (2)

- Alternate proof for continuous $X$: Let
$$g(x) \;=\; \begin{cases} 0, & \text{for } |x - \mu| < c; \\ c^2, & \text{for } |x - \mu| \geq c, \end{cases}$$
  so $(x - \mu)^2 \geq g(x)$ for all $x$.
$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)\,dx \\ &\geq \int_{-\infty}^{\infty} g(x)\,f_X(x)\,dx \\ &= c^2\,\mathbf{P}(|X - \mu| \geq c) \end{aligned}$$
- Alternate expression (set $c = k\sigma$):
$$\mathbf{P}(|X - \mu| \geq k\sigma) \;\leq\; \frac{1}{k^2}$$

## Deterministic convergence

- Sequence of numbers $a_n$ **converges** to number $a$ means "$a_n$ eventually gets and stays (arbitrarily) close to $a$"
- Formally: Sequence $a_n$ **converges** to $a$ when, for every $\epsilon > 0$, there exists $n_0$ such that $|a_n - a| \leq \epsilon$ for every $n \geq n_0$.

## Convergence in probability

- Sequence of random variables $Y_n$ **converges in probability** to a number $a$ means "almost all of the PMF/PDF of $Y_n$ eventually gets concentrated (arbitrarily) close to $a$"
- Formally: For every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}(|Y_n - a| \geq \epsilon) \;=\; 0$$



## Convergence in probability: Examples

- Let $Y_1, Y_2, \ldots$ be a sequence of Bernoulli random variables with

$$\mathbf{P}(Y_n = 1) \;=\; 1/n$$

- Let $Y_1, Y_2, \ldots$ be a sequence of random variables with

$$Y_n \text{ continuous uniform over } [n, n + 1/n]$$



## Convergence of the sample mean (Weak law of large numbers)

- $X_1, X_2, \ldots$ i.i.d. with finite mean $\mu$ and variance $\sigma^2$

$$M_n \;=\; \frac{X_1 + \cdots + X_n}{n}$$

- $\mu \;=\; \mathbf{E}[M_n] \;=$
- $\text{var}(M_n) \;=$
- Apply Chebyshev inequality to $M_n$:

$$\mathbf{P}(|M_n - \mu| \geq \epsilon) \;\leq\; \frac{\text{var}(M_n)}{\epsilon^2} \;=\; \frac{\sigma^2}{n\epsilon^2}.$$

Since $\epsilon > 0$ is arbitrary and

$$\lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0,$$

$M_n$ converges in probability to $\mu$.



## The pollster's problem

- $f$: fraction of population that "..."
- $i$th (randomly selected) person polled: $X_i = \begin{cases} 1, & \text{if yes;} \\ 0, & \text{if no.} \end{cases}$
- $M_n = (X_1 + \cdots + X_n)/n$ is fraction of "yes" in our sample
- Goal: "95% confidence in being within 1% error"

$$\mathbf{P}(|M_n - f| \geq 0.01) \leq 0.05$$

- Use Chebyshev's inequality:

$$\mathbf{P}(|M_n - f| \geq 0.01) \;\leq\; \frac{\sigma^2_{M_n}}{(0.01)^2} \;=\; \frac{\sigma^2_x}{n(0.01)^2} \;\leq\; \frac{1}{4n(0.01)^2}$$

- If $n = 50,000$, then $\mathbf{P}(|M_n - f| \geq 0.01) \leq 0.05$



## Different scalings of $M_n$

- $X_1, X_2, \ldots$ i.i.d. with finite mean $\mu$ and variance $\sigma^2$
- Let $S_n \;=\; X_1 + X_2 + \cdots + X_n$

– $M_n \;=\; S_n/n$

– $S_n$

– $S_n/\sqrt{n}$