Aula 5

Aula passada

- Valor esperado condicional
- Espaço amostral contínuo, função densidade
- Limitantes para probabilidade
- Desigualdades de Markov, Chebyshev, Chernoff
- *with high probability*

Aula de hoje

- Limitante da união
- Método do primeiro momento
- Lei dos grandes números (fraca e forte)
- Erro e confiança

Pergunta da Aula Passada

- Calcular valor exato para cauda é bem mais difícil que calcular limitante superior
  - N~Bin(10000,0.002) → μ=200

$$P[N \geq 500] = \sum_{i=500}^{10000} \binom{10000}{i} \left(\frac{2}{1000}\right)^{i} \left(1 - \frac{2}{1000}\right)^{10000-i}$$

Número muito grande!

Número muito pequeno!

- Problema: multiplicar número grande por pequeno!
- Por Chernoff, temos

$$P[X \geq (1+1.5)\mu] \leq \left(\frac{e^{1.5}}{(1+1.5)^{1+1.5}}\right)^{200}$$

Número de bom tamanho, que já é a probabilidade!

Limitante da União

- O muito famoso *Union Bound*
  - muito usado na computação e matemática
  - muito útil para lidar com muitos eventos que não necessariamente são mutuamente exclusivos ou independentes
- Sejam A e B dois eventos em um espaço amostral
- Temos que

$$P[A \vee B] = P[A+B] = P[A] + P[B] - P[A \wedge B]$$

$P[A \cup B] \leq P[A] + P[B]$ <= $P[A] + P[B]$

- Se A e B são mutuamente exclusivos, temos igualdade (pois interseção é vazia)

Limitante da União

- Seja A_i uma sequência de eventos de um espaço amostral, com $i = 1,\ldots, n$
- Temos que

$$P[.\cup_i A_i] = P[\sum_i^n A_i] \leq \sum_i^n P[A_i]$$

$P[A_i] = P[A_j] \ \forall i, j$

- Se A_i forem identicamente distribuídos (mesma probab)

$$P[.\cup_i A_i] = P[\sum_i^n A_i] \leq \sum_i^n P[A_i] = nP[A_1]$$

- Caso contrário, ainda temos

$$P[.\cup_i A_i] = P[\sum_i^n A_i] \leq \sum_i^n P[A_i] \leq n \max_i P[A_i]$$

Exemplo

- Jogar um dado honesto três vezes. Qual a probabilidade de sair 6 ao menos uma vez?
- Seja X_i o resultado da *i*-ésima rodada

$$P[X_1=6 \vee X_2=6 \vee X_3=6] \leq P[X_1=6]+P[X_2=6]+P[X_3=6]$$

$$.=3\left(\frac{1}{6}\right)=\frac{1}{2}$$

- Qual é a probabilidade exata?
- Complemento de não sair 6 em nenhuma rodada!

$$1-P[X_1 \neq 6 \wedge X_2 \neq 6 \wedge X_3 \neq 6]$$

$$= 1-P[X_1 \neq 6]P[X_2 \neq 6]P[X_3 \neq 6]$$ (por independência)

$$= 1-\left(\frac{5}{6}\right)^3 = 0.42$$ ← Limitante deu um bom resultado!

Exemplo 2

- Jogar um dado honesto 10 vezes. Qual a probabilidade de sair 6 ao menos uma vez?
- Seja X_i o resultado da *i*-ésima rodada

$$P[.\cup_i \{X_i = 6\}] = P[\sum_{i=1}^{10} \{X_i = 6\}] \leq 10\left(\frac{1}{6}\right) = \frac{5}{3}$$

Nada útil!!!

- Limitante da união demanda parcelas com probabilidade pequena e/ou pequeno número de parcelas!
  - caso contrário, resultado não é útil

Bolas e Urnas

- Considere *kn* bolas jogadas aleatoriamente sobre *n* urnas, para *k* >= 1
- Qual a prob de termos ao menos uma urna vazia (p_0)?
- X_i : v.a. indicadora que urna *i* está vazia

$$P[X_i] = \underbrace{\left(1-\frac{1}{n}\right)}_{\text{prob. da bola não cair na urna}}^{kn} \quad \longleftarrow \text{prob. urna } i \text{ vazia}$$

kn → em kn tentativas

- Limitante da união

$$p_0 = P[.\cup_i^n \{X_i\}] = P\left[\sum_{i=1}^{n} \{X_i\}\right] \le n\left(1-\frac{1}{n}\right)^{kn}$$

- Exemplos
- n=10, k=3 → p_0 = 0.42
- n=100, k=2 → p_0 = 13.4 (nada útil)

Método do Primeiro Momento

- Seja A_n uma sequência de eventos sobre o respectivo espaço amostral (n é algum parâmetro do modelo)
- Muitas vezes queremos entender se e quando a probabilidade de um evento vai a zero

$$\lim_{n \to \infty} P[A_n] = 0$$

É uma forma de computar a convergência do valor esperado

Para onde vai o limite da probabilidade do evento?

- Em particular, seja X_n uma v.a. que assume valores inteiros e não negativos, parametrizada por um parâmetro n
  - X_n conta ocorrências de alguma coisa
- Considere o evento $X_n > 0$ → Para ser mais geral, trabalhamos com a v.A. ao invés do evento
- Queremos entender $\lim_{n \to \infty} P[X_n > 0]$

Método do Primeiro Momento

- Neste caso, temos o seguinte resultado
- Se $\lim_{n \to \infty} E[X_n] = 0$ então $\lim_{n \to \infty} P[X_n > 0] = 0$

Se $E[X_n] \to 0$, então $X_n \to 0$. Logo, $P[X_n > 0] \to 0$

- Ou seja, se E[X_n] = 0, X_n assume valor 0 com probabilidade que vai a 1 quando n → oo → X_n só pode assumir um valor: 0.
  - não há ocorrências do evento que X_n conta
- Prova: desigualdade de Markov!
  - P[X >= k] <= E[X] / k
  - para k=1, temos que P[X >= 1] <= E[X], e como X é inteiro, temos que P[X >= 1] = P[X > 0] <= E[X]
  - logo, se E[X] vai a zero, P[X > 0] também vai a zero
- Abordagem conhecida por método do primeiro momento

Bolas e Urnas

- Considere *kn* bolas jogadas aleatoriamente sobre *n* urnas, para *k* >= 1
- Qual valor de k (em função de n) para que a prob de termos ao menos uma urna vazia (p_0) seja zero?
- X_i : v.a. indicadora que urna *i* está vazia

$$Y = \sum_{i=1}^{n} X_i$$

← Y = número de urnas vazias (em sistemas com *n* urnas)

- Determinar quando E[Y] vai a zero
    - condição suficiente pelo método do primeiro momento

$$E[Y] = E[\sum_{i}^{n} X_i] = \sum_{i}^{n} E[X_i] = n\underbrace{(1-1/n)^{kn}}_{e^{k} \text{ quando } n \to \infty}$$

- Se k = ω(log n) então E[*Y*] → 0
- Logo, se o número de bolas > n log n, não teremos nenhuma urna vazia (com certeza, conforme n cresce)!

Lei dos Grandes Números

- Todos devem conhecer, ao menos intuitivamente!
- Motivação: Jogar um dado honesto com seis faces n vezes

- X_i : resultado da i-ésima jogada
- $N_1(n)$: número de vezes que o resultado é 1
- $F_1(n)$: fração de vezes que o resultado é 1

$$N_1(n) = \sum_{i=1}^{n} I(X_i = 1) \qquad F_1(n) = \frac{N_1(n)}{n}$$

Binomial

→ Normalizada (tende a p se $n \to \infty$)

- Quanto vale $F_1(10) = ?$
- $F_1(100) = ?$
- $F_1(1000) = ?$

N_1 terá uma distribuição binomial com $p = 1/\#faces$, pois é a soma de indicadoras.

Lei dos Grandes Números

- $F_1(n)$ converge para $P[X_i = 1]$ quando $n \rightarrow \infty$
- Frequência relativa do resultado de experimento aleatório converge para sua probabilidade!

Conexão da teoria com a prática!

- Resultado fundamental em probabilidade e estatística
- Atribui significado físico a um conceito abstrato
  - probabilidade existe!
  - números aleatórios quando muitos, convergem (lei dos "muitos" números)

Lei dos Grandes Números

- Seja X_i uma sequência de v.a. iid, tal que

Condição da lei dos grandes números

  - $\mu = E[X_i]$, $\sigma^2 = Var[X_i]$

$$X = M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

chamada de média amostral

estimador sem viés

- M_n é uma v.a. Qual seu valor esperado, variância?

$$E[M_n] = \frac{1}{n} E[\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

$$Var[M_n] = \frac{1}{n^2} Var[\sum_{i=1}^{n} X_i] = \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

A variância da média amostral diminui com n

Lei Fraca dos Grandes Núm.

- M_n possui mesmo valor esperado que X_i e variância que vai a zero com n

$$E[M_n] = \mu \qquad Var[M_n] = \frac{\sigma^2}{n}$$

- Lei fraca dos grandes números
  - se μ finito, para qualquer $\varepsilon > 0$, temos

$$\lim_{n \to \infty} P[|M_n - \mu| < \epsilon] = 1$$

- Chamado de “convergência em probabilidade”
- Probabilidade de M_n estar ε de distância da média vai a 1, para qualquer ε positivo (ex. $\varepsilon = 10^{-10}$)

Lei Fraca dos Grandes Núm.

- Quem poderá nos ajudar a provar este resultado (assumindo σ^2 é finito)?
- Para qualquer ε > 0, $\lim_{n\to\infty} P[|M_n-\mu|<\epsilon]=1$

$$E[M_n]=\mu \qquad Var[M_n]=\frac{\sigma^2}{n}$$

lembrando que chebyshev é uma distribuição de cauda, então temos valores pequenos

- Lembrando desigualdade de Chebyshev

$$P[|M_n-\mu_{M_n}|\geq k\sigma_{M_n}]\leq\frac{1}{k^2}$$

$$\sigma_{M_n}^2=\frac{\sigma^2}{n} \rightarrow \sigma_{M_n}=\frac{\sigma}{\sqrt{n}}$$

- Aplicando, temos

$$k\sigma_{M_n}=\epsilon \rightarrow k=\frac{\epsilon\sqrt{n}}{\sigma} \quad \sigma_{M_n}^{-1}$$

$$VAR[X_i]=\sigma^2$$

$$VAR[M_n]=\sigma_{M_n}^2$$

Lei Fraca dos Grandes Núm.

- Usando a complementar

$$P[|M_n - \mu| < \epsilon] = 1 - P[|M_n - \mu| > \epsilon]$$

- Usando Chebyshev

$$P[|M_n - \mu| \geq k\,\sigma_{M_n}] \leq \frac{1}{k^2} = \frac{\sigma^2}{\epsilon^2 n}$$

- Substituindo acima, temos

$$P[|M_n - \mu| < \epsilon] = 1 - P[|M_n - \mu| > \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2 n}$$

- Cujo limite vai a 1 com n → infinito

$$\boxed{P[|M_n - \mu| < \varepsilon] \geq 1 - \frac{\sigma^2}{\varepsilon^2 n}}$$

Lei Forte dos Grandes Núm.

- M_n possui mesmo valor esperado que X_i e variância que vai a zero com n

$$E[M_n]=\mu \qquad Var[M_n]=\frac{\sigma^2}{n}$$

- Para μ finito, temos

↳ Lembrando que M_n é uma V.A. que possui uma probabilidade dessa V.A. ser igual ao seu valor esperado.

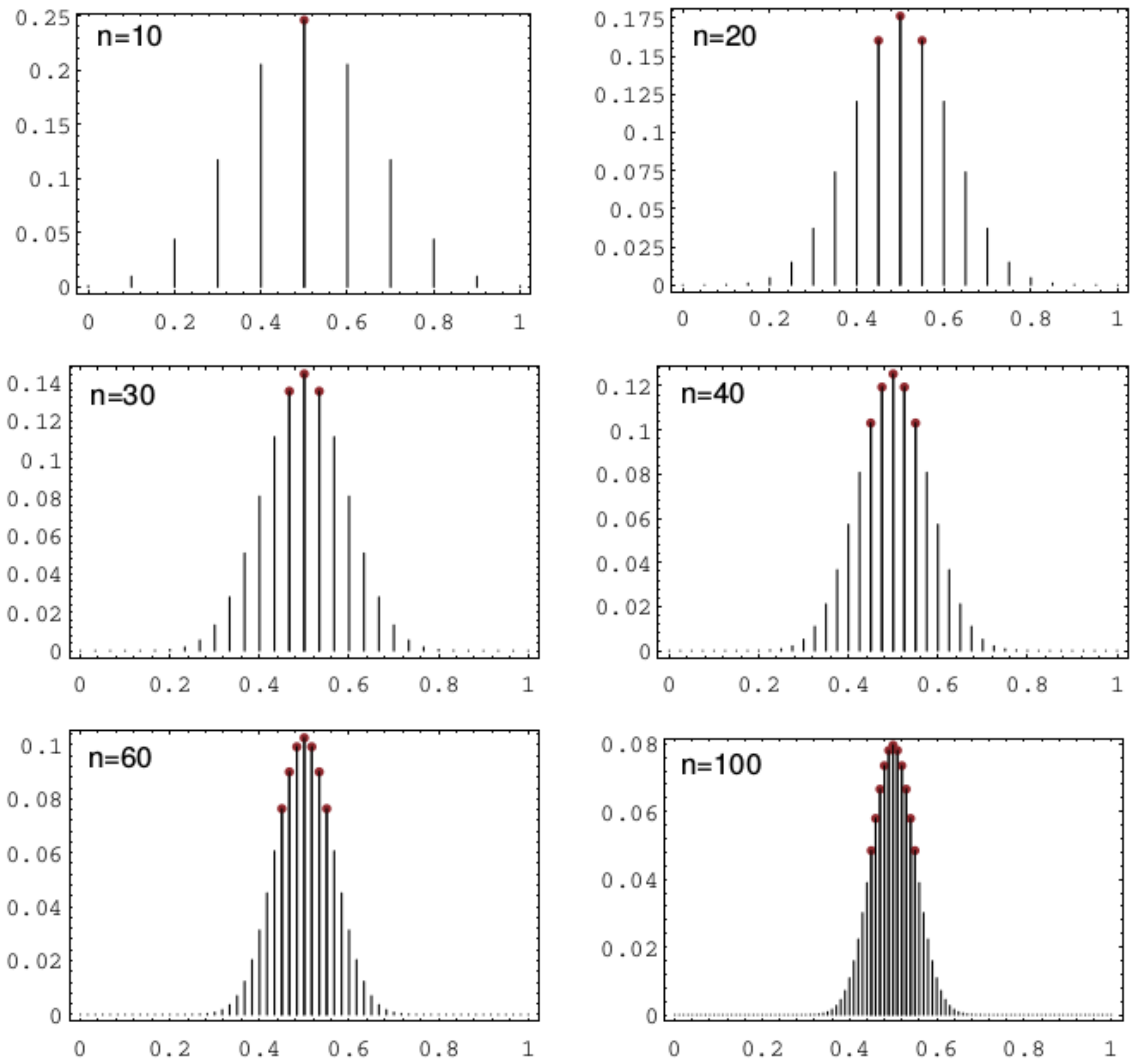
$$P[\lim_{n\to\infty} M_n=\mu]=1$$

- Chamado de “convergência quase certamente” (*almost surely*)

Quando $n\to\infty$, essa prob. vai a 1.

- Resultado bem mais forte (não temos ε)
  - M_n de fato converge para sua média!

Exemplo



- Moeda honesta, fração de caras

$$E[M_n] = \frac{1}{2}$$

$$Var[M_n] = \frac{1}{4n}$$

- Conforme *n* aumenta, M_n fica mais centrada!

Figueiredo 2018

Calculando Erro e Confiança

- Podemos usar Chebyshev para calcular precisão e confiança na lei dos grande números
- Seja precisão ε, confiança β

→ Parâmetros

$$P[M_n \in [\mu - \epsilon, \mu + \epsilon]] > \beta$$

Nosso problema é escolher n de forma que essa desigualdade seja verdade

- Dado precisão ε, confiança β (além de μ e σ^2), podemos calcular valor de *n* para atingir esta meta

→ Também posso estimar σ^2 por média amostral

- Lembrando

$$P[|M_n - \mu| < \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2 n} \longrightarrow 1 - \frac{\sigma^2}{\epsilon^2 n} = \beta$$

  - Confiança: relação linear com n
  - Precisão: relação quadrática com n
  - Implicações importantes!

$$n = \frac{\sigma^2}{\varepsilon^2 (1 - \beta)}$$

Exemplo

- Suponha uma moeda enviesada, com probabilidade de cara sendo 45%
- Você quer testar se moeda é enviesada. Quantas vezes lançar a moeda?
- Supor ε = 0.01 e β = 0.95
- Temos $\mu = 0.45$, $\sigma^2 = 0.45*0.55$ p(1-p)

$$P[M_n \in [0.44, 0.46]] > 1 - \frac{(0.45*0.55)}{(0.01)^2 n} = 0.95$$

- Logo, n = 49500

Limitante da união: $P[A \cup B] \leq P[A] + P[B]$

↳ só é útil quando $P[A_i]$ for pequeno

Método do 1º momento: Se $\lim_{n \to \infty} \mathbb{E}[X_n] = 0$ então $\lim_{n \to \infty} P[X_n \geq 0] = 0$

Se $\mathbb{E}[X_n] \to 0$, então $X_n \to 0$. Logo, $P[X_n \geq 0] \to 0$

Lei dos Grandes Números: $\mathbb{E}[M_n] = \mu$ $\quad Var[M_n] = \frac{\sigma^2}{n}$

Lei FRACA dos grandes números: $\underbrace{\lim_{n \to \infty} P[|M_n - \mu| < \varepsilon] = 1}_{\text{Convergência em probabilidade}}$

↳ Grau de exigência ε

Lei FORTE dos grandes números: $\underbrace{P[\lim_{n \to \infty} M_n = \mu] = 1}_{\text{convergência almost surely}}$

Confiança:

$$P[|M_n - \mu| < \underset{\text{precisão}}{\varepsilon}] \geq 1 - \frac{\sigma^2}{\varepsilon^2 n} = \beta \to \text{confiança}$$