Aula 4

Aula passada

- Binomial, Geométrica, Zeta
- Valor esperado
- Variância
- Distribuição conjunta
- Independência de v.a.
- Valor esperado condicional
- Espaço amostral contínuo, função densidade

Aula de hoje

- Limitantes para probabilidade
- Desigualdades de Markov, Chebyshev, Chernoff
- *with high probability*
- Limitante da União

Limitantes para Probabilidade

- Calcular probabilidade de um evento pode ser difícil
  - analiticamente intratável
  - computacionalmente intratável
- Calcular um limitante inferior ou superior para probabilidade pode ser mais fácil

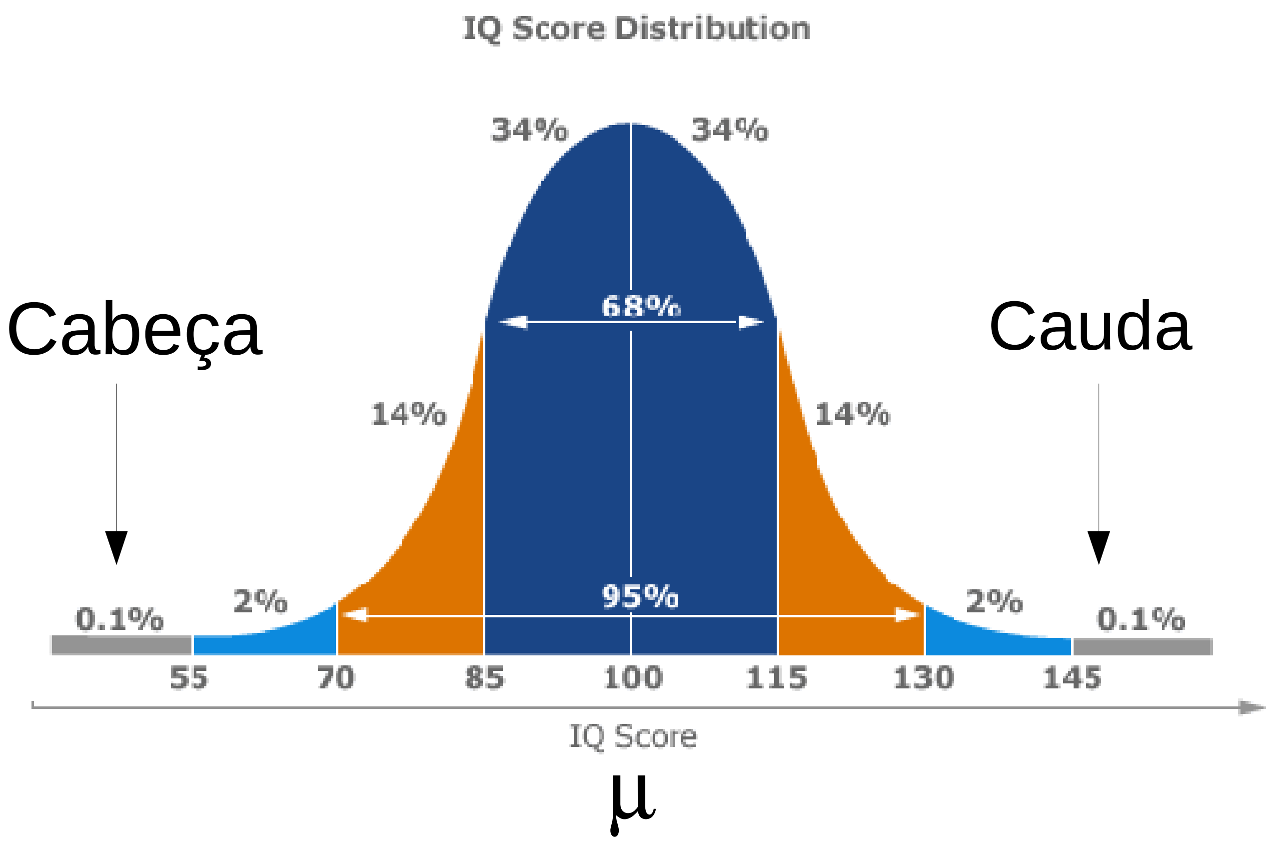
$$P[A] \leq U_A \longleftarrow \text{U_A é um limitante superior}$$

$$P[A] \geq L_A \longleftarrow \text{L_A é um limitante inferior}$$

- Em geral, estamos interessados na probabilidade da cauda ou cabeça da distribuição

Cauda e Cabeça

- Seja X uma v.a com $\mu = E[X]$
- Cauda: valores de X bem maiores que μ
- Cabeça: valores de X bem menores que μ



- Exemplos
- $P[X > k\mu]$: prob. da cauda, $k > 1$
- $P[X < k\mu]$: prob. da cabeça, $k < 1$
- Probabilidade de eventos extremos, mais raros

Exemplo

- Jogar um dado honesto de 10 faces 50 vezes
- *N* = número de vezes que o resultado foi um primo
- X_i = resultado do dado na *i*-ésima rodada

$$N = \sum_i I(X_i)$$ ← v.a. indicadora de número primo (vale 1 quando argumento é primo)

$$P[N \geq 40] = ?$$

- Qual é a distribuição de *N*?
- $P[X_i = 1] = 2/5$
- *N*~Bin(50, 2/5)

$$P[N \geq 40] = \sum_{i=40}^{50} \binom{50}{i} \left(\frac{2}{5}\right)^i \left(\frac{3}{5}\right)^{50-i}$$

- difícil calcular coeficientes de Newton
- somatório poderia ser muito longo

Desigualdade de Markov

- Importante limitante superior para probabilidade de um evento
  - relação entre valor esperado e probabilidade
- Para qualquer v.a. *X* não negativa e constante a>0, temos:

$$P[X \geq a] \leq \frac{E[X]}{a}$$

Só faz sentido para *a* > E[*X*], *a* está na cauda da distrib.

- Prova
  - I(X >= a) : v.a. indicadora do evento X >= a
  - Então aI(X >=a) <= X
  - Aplicando esperança dos dois lados, temos
    E[aI(X >=a)] <= E[X]

    P[X >= a] <= E[X] / a

Voltando ao Exemplo

- N~Bin(50, 2/5) $$P[N \geq 40] = \sum_{i=40}^{50} \binom{50}{i} \left(\frac{2}{5}\right)^{i} \left(\frac{3}{5}\right)^{50-i}$$

- Podemos aplicar desigualdade de Markov

$$P[X \geq a] \leq \frac{E[X]}{a}$$

- E[N] = 50*2/5 = 20

$$P[N \geq 40] \leq \frac{20}{40} = \frac{1}{2}$$ ← Chance de ver 40 ou mais primos é menor do que 1/2

Desigualdade de Chebyshev

- Outro importante limitante superior para probabilidade de um evento
  - relação entre valor esperado, variância e probabilidade
- Mais precisa que desigualdade de Markov
  - Markov foi aluno de Chebyshev (russos)
- Para qualquer v.a. *X* com valor esperado μ e variância σ^2, e qualquer *k*>0, temos

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

← prob de *X* estar *k* desvios padrão longe da média

- Prova
  - $Y = (X - \mu)^2$ $a = (k\sigma)^2$
  - Aplicar desigualdade de Markov usando *Y* e *a*

Caso Interessante

- Se $k = \sqrt{2}$, então temos:

$$P[|X - \mu| \geq 1.41\,\sigma] \leq \frac{1}{2}$$

- Probabilidade de X estar fora do intervalo [μ – 1.41σ , μ + 1.41σ] é menor do que ½
  - vale para qualquer distribuição da v.a.

Voltando ao Exemplo

- N~Bin(50, 2/5) $$P[N \geq 40] = \sum_{i=40}^{50} \binom{50}{i} \left(\frac{2}{5}\right)^{i} \left(\frac{3}{5}\right)^{50-i}$$
- Podemos aplicar desigualdade de Chebyshev

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

- μ = 50*(2/5) = 20 , σ^2 = 50*(2/5)*(3/5) = 12
- {N >= 40} = {N – μ >= 20}
- kσ = 20 → k = 10/sqrt(3)

$$P[N \geq 40] \leq P[|N - \mu| \geq 20] \leq \frac{1}{(10/\sqrt{3})^2} = \frac{3}{100} = 0.03$$

- Resultado melhor que por Markov!

Desigualdade de Chernoff

- Limitante superior para para soma de v.a. independentes

$$X = Y_1 + \ldots + Y_n$$

- Resultado muito importante e muito usado
  - perfeito para Binomial (soma de Bernoulli)
  - muitas variações das desigualdades (mais fáceis de usar)
- Seja Y_i ~Bern(p), μ = E[X] = np e qualquer $\delta > 0$, temos

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

← prob da cauda (depois da média)

$$P[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^{\mu}$$

← prob da cabeça (antes da média)

Voltando ao Exemplo

- N~Bin(50, 2/5) $$P[N \geq 40] = \sum_{i=40}^{50} \binom{50}{i} \left(\frac{2}{5}\right)^{i} \left(\frac{3}{5}\right)^{50-i}$$
- Podemos aplicar desigualdade de Chernoff

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

- μ = 50*(2/5) = 20
- (1+δ)μ = 40 → δ = 1

$$P[N \geq (1+1)20] \leq \left(\frac{e^{1}}{(1+1)^{1+1}}\right)^{20} = \frac{e^{20}}{2^{40}} = 0.00044$$

- Resultado bem melhor que por Chebyshev!

Desigualdade de Chernoff

- Considere Y_i resultado de moeda honesta é cara, iid

$$X = Y_1 + \ldots + Y_n$$

- X = número de caras em *n* jogadas, μ = n/2
- Onde está quase toda a “massa” da distribuição?
  - prob. da cauda vai a zero com *n*
- Ou seja, qual o valor de λ tal que P[X > μ + λ] < 1/n

$$P[X \geq (1+\delta)\mu] \leq e^{-\delta^2 \mu/3}$$

← variação da desigualdade de Chernoff

- (1+δ)μ = μ + λ → δ = λ/μ

$$P[X \geq \mu + \lambda] \leq e^{-\left(\frac{\lambda}{\mu}\right)^2 \mu/3} = e^{-\frac{2\lambda^2}{3n}} = 1/n$$

$$\lambda = \sqrt{3/2\, n \ln n}$$

With High Probability (whp)

- Seja *n* um parâmetro de um modelo probabilístico
  - ex. número de rodadas de um dado, vértices no grafo G(n,p)
- Seja *A*(*n*) um evento no respectivo espaço amostral
- *A*(*n*) ocorre *with high probability* (whp) quando

$$P[A(n)] \geq 1 - \frac{1}{n^{\alpha}}$$ ← para algum α > 1 constante

- Repare que $\lim_{n \to \infty} P[A(n)] = 1$ ← convergência em probabilidade
- Do exemplo anterior temos:

$$X \leq \frac{n}{2} + \sqrt{3/2\, n \ln n} = \frac{n}{2} + \theta\left(\sqrt{n \ln n}\right)$$, w.h.p.

  - ou seja, número de caras será praticamente sempre menor do que a média + sqrt(*n* log *n*), ao jogar *n* vezes

Figueiredo 2018

Voltando ao Exemplo

- Se $\lambda = \sqrt{3/2\, n \ln n}$
- Então $P[X \geq \mu + \lambda] \leq 1/n$
- Exemplo: *n*=1000 (lançar moeda 1000 vezes)
  - μ = 500
  - $\lambda = \sqrt{3/2\, n \ln n} = \sqrt{1500 \ln 1000} = 101.8$
- Então

$$P[X \geq 500 + 102] = P[X \geq 602] \leq 0.001$$

- Ou seja, observar 602 caras ou mais é bastante raro, ao jogar uma moeda honesta 1000 vezes
- Podemos apostar com bastante segurança!

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

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

$$<= P[A] + P[B]$$

- 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

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

- 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

$$P[X_i] = \left(1 - \frac{1}{n}\right)^{kn} \longleftarrow \text{prob. urna } i \text{ vazia}$$

- 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)

Figueiredo 2018