

Aula 5

Roteiro

- Lei dos grandes números (fraca e forte)
- Erro e confiança
- Margem de erro
- Falácia do apostador

Fração Relativa de Resultados

- Modelo de jogar um dado honesto n vezes
- X_i : resultado da i -ésima jogada
- $Y_i = I(X_i = 1)$: indicadora que resultado é número 1
- $N_1(n)$: número de vezes que resultado é 1
- $F_1(n)$: fração de vezes que resultado é 1



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

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

$F_1(n)$ é aleatório!

Lei dos Grandes Números

- $F_1(n)$ converge quando $n \rightarrow$ infinito
- Frequência relativa converge para a probabilidade do evento

Conexão da teoria com a prática!



- Resultado fundamental em probabilidade e estatística
- Atribui significado físico a um conceito abstrato
 - a fração de resultados aleatórios quando muitos, convergem (lei dos “muitos” números)
 - probabilidade possui implicação prática

Média Amostral



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

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

- Seja $M_n = \frac{1}{n} \sum_{i=1}^n X_i$ ← média amostral (média aritmética dos n valores)

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

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

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

Lei Fraca dos Grandes Núm.

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

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

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

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

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



Provando a Lei

- Prova assumindo que σ^2 é finito
- Para qualquer $\varepsilon > 0$, $\lim_{n \rightarrow \infty} P[|M_n - \mu| < \varepsilon] = 1$

$$E[M_n] = \mu_{M_n} = \mu \quad \text{Var}[M_n] = \sigma_{M_n}^2 = \frac{\sigma^2}{n}$$

- Lembrando desigualdade de Chebyshev

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

- Aplicando, temos

$$k \sigma_{M_n} = \varepsilon \rightarrow k = \frac{\varepsilon \sqrt{n}}{\sigma}$$

Provando a Lei

- Usando a probabilidade 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 $n \rightarrow \infty$

Lei Forte dos Grandes Núm.

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

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

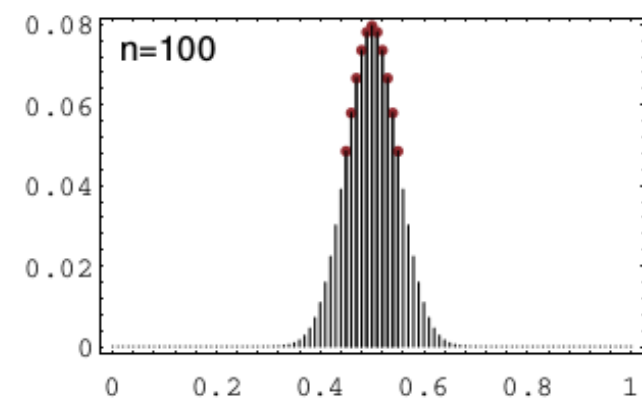
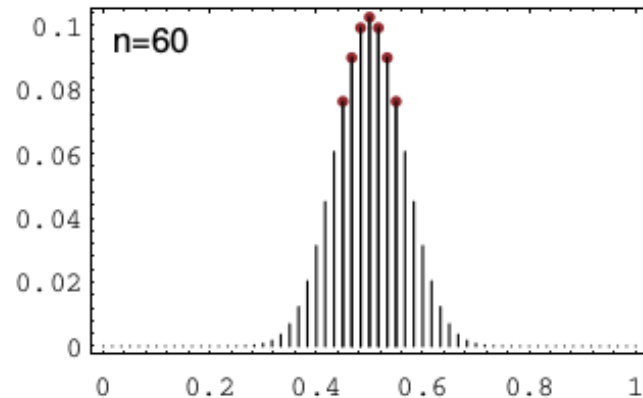
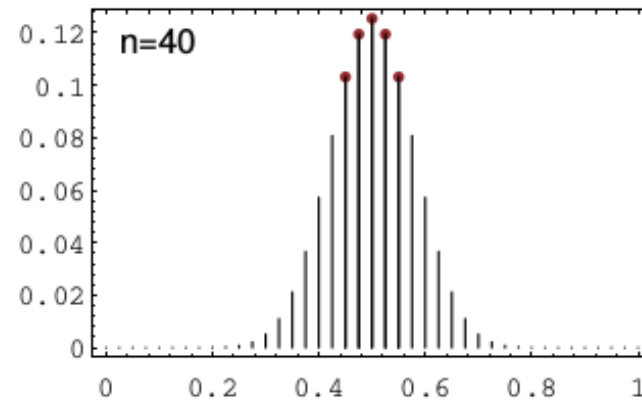
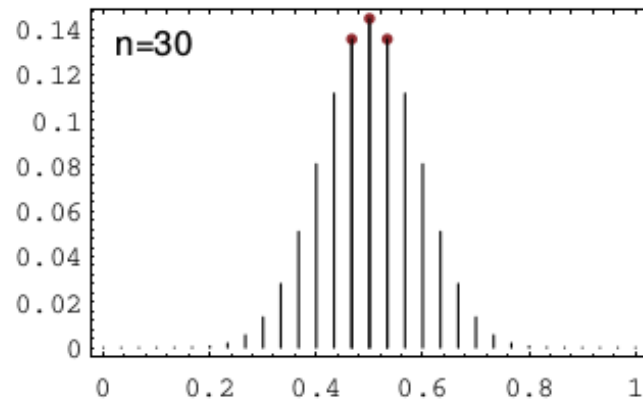
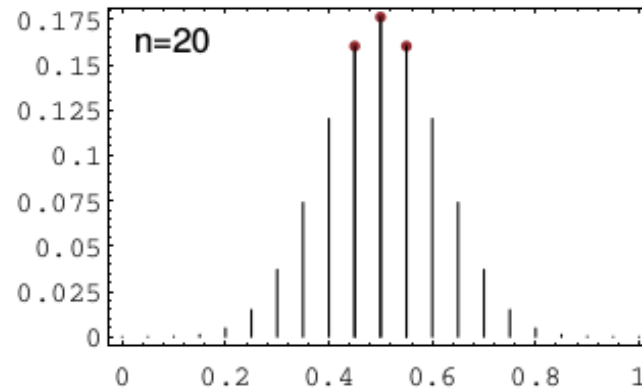
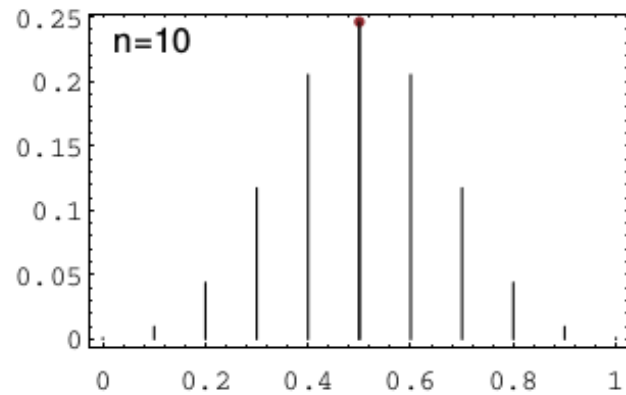
- Para μ, σ^2 finitos, temos

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

- Chamado de “convergência quase certamente” (*almost surely*)
- Resultado bem mais forte (não temos ε)
 - M_n de fato converge para sua média!



Exemplo



- Moeda honesta, fração de caras

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

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

- Conforme n aumenta, M_n fica mais centrada!

Erro e Confiança



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

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

- Dado precisão ϵ , confiança β (além de μ e σ^2), determinar valor de n para garantir esta relação

Erro e Confiança



- Lembrando

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

- Resolvendo para n , temos

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

- Confiança: influencia n linearmente
- Precisão: influencia n de forma quadrática
- Implicações importantes
 - aumentar a precisão (reduzir ϵ) demanda mais do que aumentar a confiança (aumentar β)

Exemplo



- Suponha uma moeda enviesada, com probabilidade de cara sendo 45%
- Você quer testar se moeda é mesmo enviesada
Quantas vezes lançar a moeda?
- Supor $\varepsilon = 0.01$ e $\beta = 0.95$
- Temos $\mu = 0.45$, $\sigma^2 = 0.45*0.55$

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

- Logo, $n = 49500$

Estimando Margem de Erro

- Pesquisa com 1000 cariocas revelou que entre praia e cachoeira, 70% preferem praia
- Qual a margem de erro da pesquisa com 90% de confiança?
- Modelar cada pessoa com uma Bernoulli, $X_i = 1$ se a pessoa prefere praia
- Supor $p = \mu = 0.7$ (resultado da pesquisa), e então $\sigma^2 = 0.7 \cdot 0.3 = 0.21$
- Calcular ε dado $n = 1000$ e $\beta = 0.9$

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

- Logo, $\varepsilon = 0.046 = 4.6\%$

Falácia do Apostador

- Se um evento ocorre mais frequentemente que o esperado no passado então ele vai ter menos chances de ocorrer no futuro (ou vice-versa)
 - falácia comum em cassinos e jogos
- Ex. jogo de roleta sai vermelho 10 vezes seguidas, então a chance de sair vermelho na próxima vez é menor
- Seja X_i indicadora da cor vermelha na i -ésima rodada
 - Assumindo sequência iid, temos

$$P[X_{11}=1 | X_1=1 \wedge \dots \wedge X_{10}=1] = P[X_{11}=1]$$

- Passado não influencia aleatoriedade do futuro, pois já foi observado e é independente
 - nem na teoria (assumindo iid), nem na prática (assumindo condições para iid)

Falácia e Lei

- Lei dos Grandes Números não torna a falácia verdadeira?

Não! São afirmações diferentes

- Lei diz que fração relativa das observações converge para sua respectiva probabilidade
- Falácia diz que observações passadas influenciam a probabilidade de observações futuras
- Fonte de muita confusão e discussão, ao menos desde *Laplace* no final do século XVIII
 - falácia é indicativo da dificuldade de compreender aleatoriedade e independência em modelos aleatórios