

Aula 3

Aula passada

- Espaço amostral
- Probabilidade
- Eventos
- Independência
- Exclusão mútua
- Probabilidade total
- Regra de Bayes
- Variável aleatória
- Função de distribuição

Aula de hoje

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

Distribuição Binomial

- Considere uma sequência iid de n v.a. de Bernoulli

$$X_1, X_2, \dots, X_n$$

- onde $X_i \sim \text{Bernoulli}(p)$ para $i = 1, 2, \dots, n$

- Seja Z a soma destas v.a.

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

- Z possui distribuição Binomial, com parâmetros n e p
 - $Z \sim \text{Bin}(n, p)$
- Que valores que Z pode assumir ?

Distribuição Binomial

- Qual a expressão para a distribuição Binomial?
 - probabilidade da soma ser igual a i

$$f_Z(i) = P[Z = i]$$

- Quanto vale?

$$f_Z(i) = P[Z = i] = \binom{n}{i} p^i (1-p)^{n-i}$$

- Podemos obter expressão a partir de princípios elementares
 - ver exercício da lista
- Exemplos
 - número de caras ao jogar uma moeda 20 vezes
 - grau do vértice em um grafo com n vértices onde cada aresta incidente ocorre com prob p (modelo $G(n,p)$)

Distribuição Geométrica

- Considere uma sequência iid de Bernoulli

X_1, X_2, \dots

0 0 0 1 $\rightarrow Z=4$
0 0 1 $\rightarrow Z=3$

- onde $X_i \sim \text{Bernoulli}(p)$ para $i = 1, 2, \dots, n$

- Seja Z o menor valor tal que $X_Z = 1$

- posição do primeiro valor 1 da sequência

$$Z = \min \{ i \mid X_i = 1 \}$$

- Z possui distribuição Geométrica, com parâmetro p

- $Z \sim \text{Geo}(p)$ \rightarrow parâmetro da Bernoulli que gera cada amostra

- Que valores que Z pode assumir ? \rightarrow da geométrica

$\{0, 1, \dots, \infty\}$

Distribuição Geométrica

- Qual a expressão para a distribuição Geométrica?

$$f_Z(i) = P[Z = i] = (1 - p)^{i-1} p, i = 1, 2, \dots$$

Handwritten notes:
↗ $i-1$ saídas 0
↘ última saída 1

- Para ser função de probabilidade, precisamos

$$\sum_{i>0} f_Z(i) = 1 \quad \longleftarrow \text{Verificar!}$$

- Podemos obter expressão a partir de princípios elementares
 - ver exercício da lista
- Exemplos
 - número de vezes que moeda é jogada até primeira cara
 - número de elementos inseridos em tabela *hash* até colisão com elemento em uma posição fixa

Distribuição Zeta

- Seja Z uma v.a. com distribuição Zeta com parâmetro $s > 1$

$$f_Z(i) = P[Z = i] = \frac{C(s)}{i^s}, i = 1, 2, \dots$$

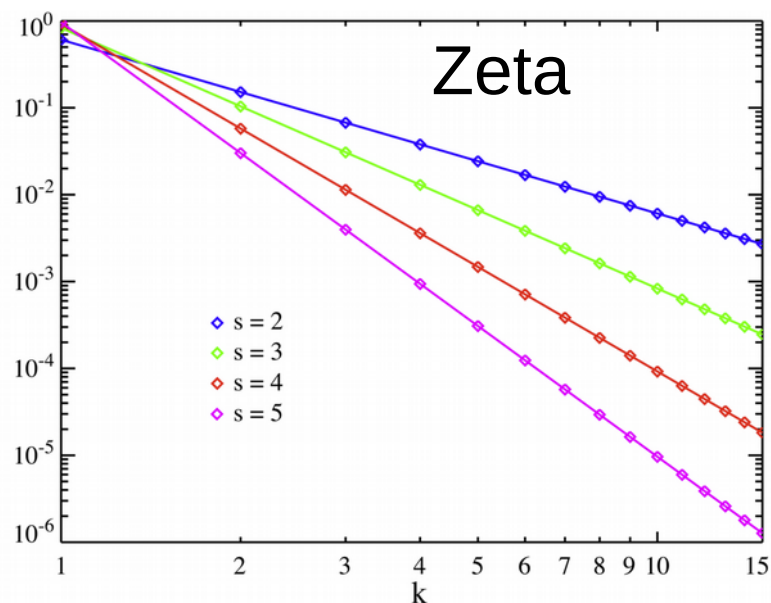
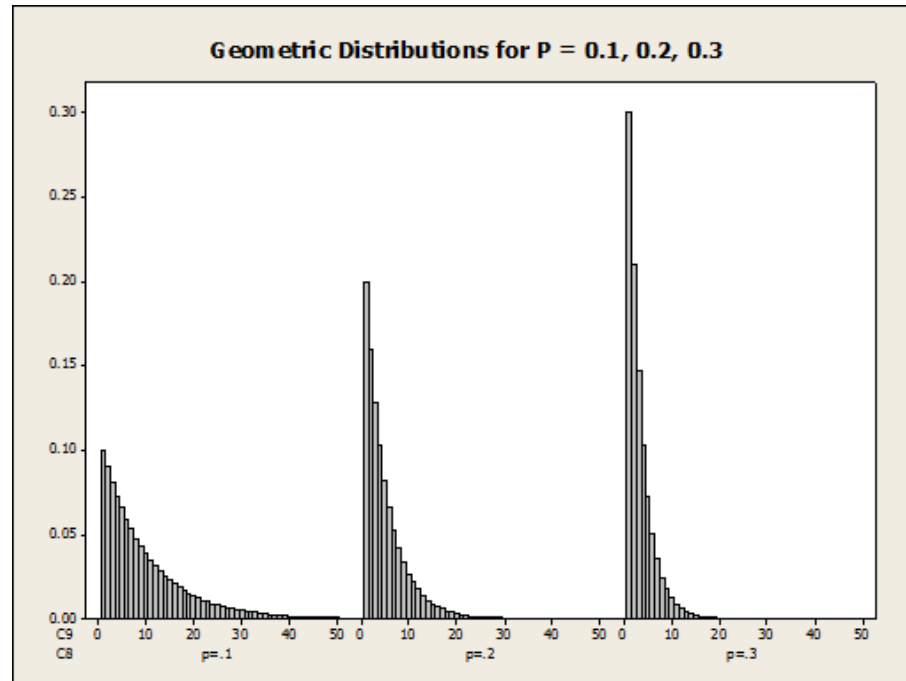
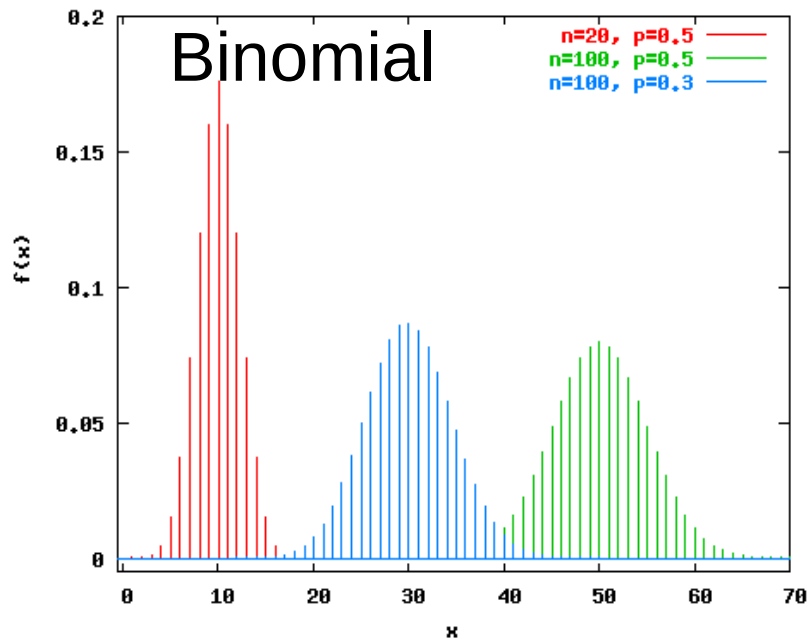
- onde $C(s)$ é a constante de normalização
 - conhecida por função zeta de Riemann

$$\sum_{i>0} \frac{1}{i^s} = \frac{1}{C(s)} \quad \longleftarrow \text{ Definida para todo } s > 1$$

- Zeta é uma distribuição em lei de potência
 - *cauda pesada*: probabilidade decai bem mais devagar do que exponencial
- Exemplos
 - número de vezes que uma palavra ocorre na wikipedia
 - número de seguidores de um perfil no Twitter

Gráficos da Distribuições

- “A picture is worth a thousand words”



- Diferença fundamental entre funções
- Binomial é “simétrica”, Geométrica tem um “lado”, Zeta tem “cauda pesada”

Valor Esperado

- Função de probabilidade caracteriza por completo comportamento de uma v.a.
 - muitas vezes é desconhecida
- Resumo do comportamento de uma v.a.

Valor Esperado

$$\mu_X = E[X] = \sum_{i \in O_X} i f_X(i) \leftarrow \text{média ponderada dos valores que } X \text{ pode assumir}$$

- Não é aleatório (é simplesmente um número)
- Representa o comportamento médio da v.a.
- Cada distribuição tem o seu valor esperado

Exemplos de Valor Esperado

- $X \sim \text{Bernoulli}(p) \rightarrow E[X] = ?$ $= p \cdot 1 + 0 \cdot (1-p) = p$

- aplicar definição e fazer as contas!

- $X \sim \text{Bin}(n, p) \rightarrow E[X] = ?$

$$E[X] = np$$

- $X \sim \text{Geo}(p) \rightarrow E[X] = ?$ \longrightarrow

$$E[X] = 1/p$$

- $X \sim \text{Zeta}(s) \rightarrow E[X] = ?$

$$E[X] = C(s-1)/C(s)$$

, para $s > 2$

$$E[\text{Bin}(n, p)] = E\left[\sum_{i=1}^n \text{Bernoulli}(p)\right] = \sum_{i=1}^n E[\text{Bernoulli}(p)] = \sum_{i=1}^n p = np$$

Função de v.a.

- Seja g uma função qualquer, tal que

$$g: \mathbb{Z} \rightarrow \mathbb{R}$$

- Seja X uma v.a., podemos aplicar g a X

$g(X)$ ← Define uma nova v.a. com imagem nos reais, mas ainda é discreta (contável)!

- Valor esperado da função da v.a. (mesma definição)

$$E[g(X)] = \sum_{i \in O_X} g(i) f_X(i)$$

- Em geral, $E[g(X)] \neq g(E[X])$

- valor esperado da função e função do valor esperado

Variância

- Outro resumo do comportamento de uma v.a.
- Medida de “dispersão” ao redor da média

$$g(X) = (X - \mu)^2 \longleftarrow \begin{array}{l} \text{quadrado da diferença com o} \\ \text{valor esperado} \end{array}$$

- Variância: valor esperado da função acima

$$\sigma_X^2 = \text{Var}[X] = E[(X - \mu_X)^2]$$

- Desvio padrão: raiz quadrada da variância

$$\sigma_X = \sqrt{\text{Var}[X]} = \sqrt{\sigma_X^2}$$

- Não é aleatório (é simplesmente um número)
- Representa a dispersão média da v.a.
- Cada distribuição tem a sua variância

Exemplos de Variância

- $X \sim \text{Bernoulli}(p) \rightarrow E[X] = ?$

- aplicar definição e fazer as contas!

- $X \sim \text{Bin}(n, p) \rightarrow \text{Var}[X] = ?$

- $X \sim \text{Geo}(p) \rightarrow \text{Var}[X] = ?$

- $X \sim \text{Zeta}(s) \rightarrow \text{Var}[X] = ?$

$$\text{Var}[X] = np(1-p)$$

$$\text{Var}[X] = (1-p)/p^2$$

$$\text{Var}[X] = C(s-2)/C(s-1)$$

, para $s > 3$

Propriedades

- Linearidade da esperança (outro nome para valor esperado)

$$E[X + Y] = E[X] + E[Y] \leftarrow \text{muito usada!}$$

- Seja X e Y duas v.a. independentes

$$E[XY] = E[X]E[Y]$$

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

- Seja $Y = aX + b$, para constantes a, b

$$E[Y] = aE[X] + b$$

$$\text{Var}[Y] = a^2 \text{Var}[X]$$

Distribuição Conjunta

- Até agora, eventos sobre uma única v.a.
 - muitas vezes interesse em eventos que envolvem mais de uma v.a.
- Seja X e Y duas v.a. definidas sobre um mesmo espaço amostral

$$f_{XY}(i, j) = P[X = i \wedge Y = j] \quad \leftarrow \text{Distribuição conjunta de } X \text{ e } Y$$

- Distribuição de X ou Y podem ser obtidas da conjunta

$$f_X(i) = \sum_{j \in O_Y} f_{XY}(i, j)$$

$$f_Y(j) = \sum_{i \in O_X} f_{XY}(i, j)$$

\leftarrow Também chamada de distribuição marginal

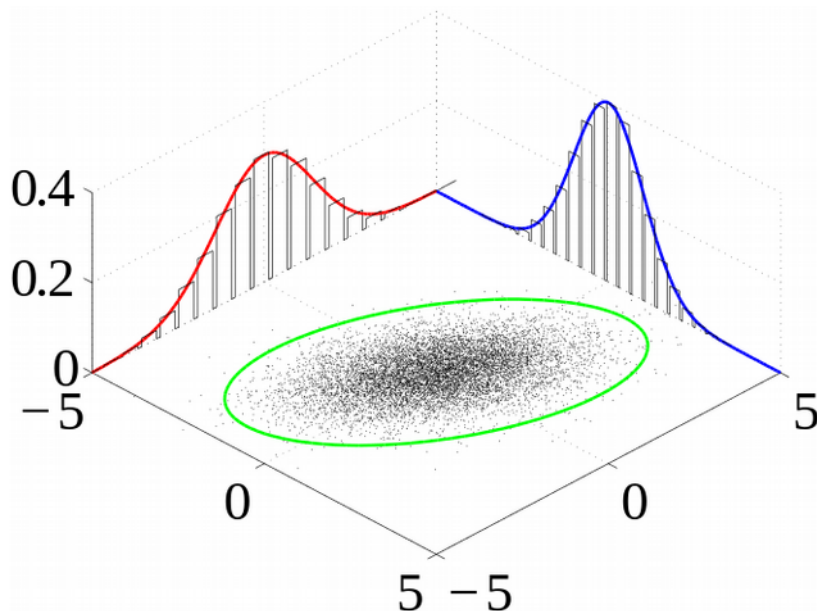
Exemplo



- $S = \text{alfabeto}$, $|S| = 26$, $p_x = 1/26$
- X é uma v.a. tal que $X(a)=1$, $X(b)=2$, $X(c)=3$, ..., $X(z)=26$
- Y é uma v.a. tal que $Y(\text{vogal})=1$, $Y(\text{consoante})=2$

$$f_{XY}(1,1)=? \quad f_{XY}(1,2)=?$$

$$f_{XY}(2,1)=? \quad f_{XY}(2,2)=?$$



- Figura com distribuições marginais de X e Y e distribuição conjunta (nuvem de pontos)
- Apenas ilustrativa

Independência entre v.a.

- Seja X e Y duas v.a. são independentes sse

$$\begin{aligned}f_{XY}(i, j) &= P[X = i \wedge Y = j] \\ &= P[X = i]P[Y = j] = f_X(i)f_Y(j)\end{aligned}$$

- Exemplo
- Dois dados honestos com k faces
- Joga moeda enviesada com prob p para escolher o dado
- X = face observada, Y = dado escolhido

$$f_{XY}(i, j) = \frac{p}{k}, \text{ se } j = 1$$

$$f_{XY}(i, j) = \frac{1-p}{k}, \text{ se } j = 2$$

Esperança Condicional

- Seja X e Y duas v.a. definidas sobre um mesmo espaço amostral

$$f_{X|Y}(i, j) = P[X = i | Y = j] \quad \leftarrow \text{Distribuição condicional de } X \text{ dado } Y$$

- Valor esperado condicional
 - valor esperado restrito a um subconjunto do espaço amostral (definido pelo valor de outra v.a.)

$$E[X | Y = j] = \sum_{i \in O_X} i f_{X|Y}(i, j)$$

- segue definição de valor esperado (nada de novo)!

Esperança Condicional

- Muito útil para calcular valores esperados
 - uso similar a prob conditional para calcular marginal
- Seja $E[X | Y=j]$ uma função da v.a. Y
 - repare que $E[X|Y]$ é uma v.a. ao considerar j aleatório

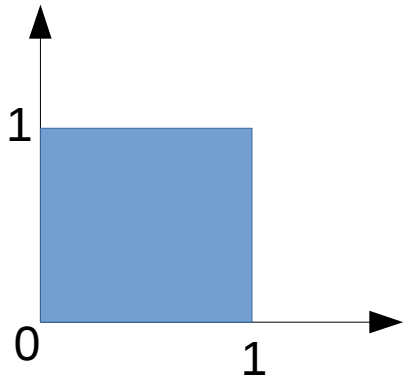
- Temos que

$$E[X] = E[E[X|Y]] \quad \leftarrow \text{propriedade da torre da esperança}$$

- Exemplo: $S = N_1 + N_2 + \dots + N_K$
 - parcela $N_i \sim \text{Bin}(n, p)$, iid, e $K \sim \text{Geo}(q)$
 - número de parcelas é aleatório
- $E[S] = ?$

Espaço Amostral Contínuo

- Até agora, espaço amostral era discreto (contável)
- Espaço amostral contínuo (não contável)
 - ex. números reais, pontos do quadrado $[0,1] \times [0,1]$, etc



Problema: Como dar probabilidade a cada ponto do espaço amostral?

Solução: Dar probabilidade a “subconjuntos” do espaço amostral

- pedaços contíguos do espaço amostral tem uma “densidade” de probabilidade
- eventos representam tais pedaços

Função de Densidade

- Seja A um evento qualquer de S
- Dizemos que $f(x)$ é uma função densidade sse

$$P[A] = \int_A f(x) dx \quad \leftarrow \text{Área da função densidade dentro do espaço definido por } A$$

- Mesma restrições que antes

$$0 \leq \int_A f(x) dx \leq 1 \quad \int_S f(x) dx = 1$$

- Exemplo: $S = [a, b]$, para a, b constantes

$$f(x) = \frac{1}{b-a}, x \in [a, b] \quad P[A] = \int_{a'}^{b'} f(x) dx = \frac{b' - a'}{b - a}$$

- $A = [a', b']$ com $a' \geq a, b' \leq b$

Segue Tudo

- Todos os conceitos são equivalentes
 - independência, exclusão mútua, prob. totais, etc
 - trocar somatório por integral

- Definir v.a. para S contínuo

$X : S \rightarrow \mathbb{R}$ \longleftarrow v.a. contínua, com imagem não contável

- Todos os conceitos sobre v.a. são equivalentes
 - função de distribuição (chamada de densidade)
 - valor esperado, variância, propriedades
 - distribuição conjunta, etc
 - trocar somatório por integral nas definições, usando CDF quando necessário