Universidade Federal do Rio Grande do Norte
Instituto Metrópole Digital
Bacharelado em Tecnologia da Informação
IMD0033 - Probabilidade

Modelos Probabilísticos

Natal/RN
5 de junho de 2016

Sumário

1 Modelos Probabilísticos Discretos

1.1 Distribuição Uniforme Discreta

Uma variável aleatória X segue o modelo uniforme discreto, denotada por $X\ Ud(M)$, quando M é o conjunto de valores equiprováveis $x_1, x_2, \cdots, x_n$ que X assume.

Tem uma função de probabilidade dada por:

$$P(X == x_i) = \frac{1}{n}, \forall i = 1, 2, \cdots, n \tag{1}$$

1.1 Uma rifa tem 100 bilhetes numerados de 1 a 100. Eu tenho 5 bilhetes consecutivos numerados de 21 a 25 e meu colega tem outros 5 bilhetes com os números 1, 11, 29, 68 e 93. Quem tem maior possibilidade de ser sorteado?

1.2 Distribuição Binomial

Em cada ensaio considera-se somente a ocorrência ou não-ocorrência de um certo evento que será denominado sucesso (S) e cuja não-ocorrência será denominada falha (F). Os ensaios são independentes e a probabilidade de sucesso, que denotaremos por p, é a mesma para cada ensaio. A probabilidade de falha será denotada por $1 - p$.

Seja X o número de sucessos obtidos na realização de n ensaios de Bernoulli independentes. Diremos que X tem distribuição binomial com parâmetros n e p, em que p é a probabilidade de sucesso em cada ensaio, se sua função de probabilidade for dada por

$$f(x) = P(X == x) = \binom{n}{k} p^k (1-p)^{1-k}, \text{ para } k = 0, 1, \cdots, n \tag{2}$$ [1]

2.1x Suponha que numa linha de produção a probabilidade de se obter uma peça defeituosa (sucesso) é $p = 0,1$. Toma-se uma amostra de 10 peças para serem inspecionadas. Qual a probabilidade de se obter:

(a) Uma peça defeituosa?

(b) Nenhuma peça defeituosa?

(c) Duas peças defeituosas?

(d) No mínimo duas peças defeituosas?

(e) No máximo duas peças defeituosas?

[1] $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

2.2x O professor da disciplina de Estrutura de Dados elaborou um prova de múltipla escolha, consistente em 10 questões cada uma com 5 alternativas cada questão. Suponha que nenhum dos estudantes que vão a fazer a prova não vai as aulas e não estudou para a prova. O professor estabeleceu que para aprovar deve conter corretamente ao menos 6 questões. Se 200 alunos se apresentaram, quantos alunos aprovaram a disciplina?

2.3 Uma família tem 6 crianças. Supondo que a probabilidade de uma criança em particular ser menina seja $\frac{1}{2}$, encontre a probabilidade de existir 3 meninos e 3 meninas. Depois, encontre a probabilidade de existir menos meninos que meninas.

1.2.1 Valor esperado e Variância

$E(x) = np$
$V(x) = np(1-p)$

2.4 Uma urna contém 20 bolas brancas e 30 bolas vermelhas. Uma bola é retirada da urna e a variável aleatória X denota o número de bolas vermelhas obtidas. Calcule a média, a variância e o desvio-padrão de X

1.3 Distribuição Geométrica

Consideremos uma sequência ilimitada de ensaios de Bernoulli, com probabilidade de sucesso p em cada ensaio. Designemos sucesso por S e falha por F. Realizamos os ensaios até que ocorra o primeiro sucesso.

Seja X a variável aleatória que fornece o número de falhas até o primeiro sucesso. A variável X tem distribuição Geométrica com parâmetro p, $0 < p < 1$, se sua função de probabilidade é dada por

$$P(X = x) = (1-p)^x p, j = 0, 1, \cdots \tag{3}$$

1.3.1 Valor esperado e Variância

$E(x) = \frac{1-p}{p}$
$V(x) = \frac{1-p}{p^2}$

3.1x Um dado honesto é lançado sucessivas vezes até que apareça pela primeira vez a face 1. Seja X a variável aleatória que conta o número de ensaios até que corra o primeiro 1. Qual a probabilidade de obtermos 1 no terceiro lançamento?

1.4 Distribuição Binomial Negativa

Imagine as mesmas condições em que foi definida a distribuição geométrica, porém agora queremos repetir o experimento até que se obtenha o r-ésimo sucesso

Assim, a função de probabilidade de X é dada por:

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, x = r, r+1, r+2, \cdots \tag{4}$$

1.4.1 Valor esperado e Variância

$$E(x) = \frac{r}{p}$$
$$V(x) = \frac{r(1-p)}{p^2}$$

4.1x João é um jogador de basquete cuja probabilidade de acertar um arremesso livre é 0,70. Qual a probabilidade de João acertar seu terceiro arremesso na sua quinta tentativa.

4.2 Suponha que a probabilidade de nascimento de um menino seja 0.5. Imagine que um casal quer ter exatamente duas meninas e terá filhos até essa condição ser satisfeita.

(a) Qual é a probabilidade da família ter x filhos homens?

(b) Qual é a probabilidade da família ter quatro filhos?

(c) Qual é a probabilidade de a família ter no máximo quatro filhos?

(d) Quantos filhos homens espera-se que essa família tenha?

(e) Quantos filhos espera-se que essa família tenha?

1.5 Distribuição de Poisson

Na prática muitos experimentos consistem em observar a ocorrência de eventos discretos em um intervalo contínuo (unidade de medida)

A distribuição de Poisson é empregada em experimentos, nos quais não se está interessado no número de sucessos obtidos em n tentativas, como ocorre no caso da distribuição Binomial, mas sim no número de sucessos ocorridos durante um intervalo contínuo, que pode ser um intervalo de tempo, espaço, etc.

$$f(x) = \frac{e^{-\mu}}{x!} \tag{5}$$

1.5.1 Esperança e Variância

$E(X) = \mu = \lambda$

$V(X) = \sigma^2 = \lambda$ $\quad \lambda$: Média dos eventos discretos em uma unidade de medida.

5.1x As consultas num banco de dados ocorrem de forma independente e aleatório seguindo a distribuição de Poisson. Suponha que a média de consultas é 3 a cada 4 minutos. Qual é a probabilidade que banco de dados seja consultado no máximo 2 em um intervalo de 2 minutos?

1.5.2 Propriedade do Modelo de Poisson

Se $X_1, \cdots, X_n$ são variáveis aleatórias independentes, com distribuição de Poisson com parâmetros $\mu_1, \cdots, \mu_n$ respectivamente, neste caso, a variável que soma essas variáveis tem distribuição de Poisson com parâmetro $\mu = \mu_1, \cdots, \mu_n$

$$Y = X_1 + \cdots + X_n$$

1.6 Distribuição Hipergeométrica

Uma amostra de tamanho n é selecionada aleatoriamente sem reposição de uma população de N itens. Na população, k itens podem ser classificados como sucessos e $N \breve{} k$ itens podem ser classificados como fracassos.

Onde:

N : O número de itens na população

k : O número de itens na população que são classificados como sucessos.

n : O número de itens na amostra.

x : O número de itens na amostra selecionada que são classificados como sucessos.

$$P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \tag{6}$$

6.1x Em problemas de controle de qualidade, lotes dom N itens são examinados. O número de itens com defeito (atributo A), K, é conhecido. Colhemos uma amostra de n itens e determinamos x. Exemplo, um lote com N=100 peças, K=10 são defeituosas. Escolhendo n=5 peças sem reposição, a probabilidade de não se obter peças defeituosas é:

1.6.1 Esperança e Variância

$$E[X] = \frac{nK}{N} = np \qquad \text{se } p = \frac{K}{N}$$
$$V[X] = np\frac{N-n}{N-1}$$

1.7 Distribuição Multinomial

Um experimento multinomial consiste de n tentativas repetidas. Cada tentativa tem um número discreto resultados possíveis. Em qualquer tentativa dada, a probabilidade de que um particular resultado ocorrerá é constante. As tentativas são independentes; isto é, o resultado de uma tentativa não afeta o resultado das outras tentativas.

$$P = [\frac{n!}{n_1! * n_2! * \cdots * n_k!}](p_1^{n_1} * p_2^{n_2} * \cdots * p_k^{n_k}) \tag{7}$$

7.1x Suponha uma carta de baralho sendo extraída aleatoriamente de um maço de jogo de baralho, e depois então devolvida ao maço. Este exercício é repetido 5 vezes. Qual é a probabilidade de se extraírem 1 espada, 1 copa, 1 ouros e 2 paus?

2 Modelos Probabilísticos Contínuos

2.1 Distribuição Uniforme

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{Caso contrário} \end{cases}$$

2.1.1 Esperança e Variância

$$E[X] = \frac{b+a}{2} V[X] \qquad = \frac{(b-a)^2}{12}$$

2.1.2 Função Distribuição de Probabilidade

$$f(x) = \begin{cases} 0 & \text{se } x \leq a < a \\ \frac{x-a}{b-a} & \text{se } a \leq x \leq b \\ 1 & \text{se } x \geq b \end{cases}$$

2.2 Modelo Exponencial

O modelo exponencial é muito aplicado quando o interesse é descrever em termos probabilísticos o tempo (espaço) até a ocorrência de um evento de interesse.

$f(x) = \lambda e^{-\lambda x}$

$P(a \leq x \leq b) = e^{-\lambda a} - e^{-\lambda b}$

$P(X \leq x) = 1 - e^{-\lambda x}$

$P(X > x) = e^{-\lambda x}$

2.2.1 Esperança e Variância

$E[X] = \dfrac{1}{\lambda}$

$V[X] = \dfrac{1}{\lambda^2}$

2.2.2 Falta de memória

Permite a translação da origem no calcula de probabilidade. A distribuição **exponencial** possui essa propriedade.

Suponha que X representa o tempo de vida de um equipamento, a probabilidade do equipamento durar pelo menos t+s unidades de tempo, sabendo-se que ele já durou s, é igual a probabilidade de um equipamento novo durar pelo menos t unidades de tempo.

A informação "idade" do equipamento pode ser esquecida e o que importa é quanto tempo a mais queremos que ele dure.

2.3 Distribuição Normal

A distribuição normal conhecida também como distribuição gaussiana é sem dúvida a mais importante distribuição contínua. Sua importância se deve a vários fatores, entre eles podemos citar o teorema central do limite, o qual é um resultado fundamental em aplicações práticas e teóricas, pois ele garante que mesmo que os dados não sejam distribuídos segundo uma normal a média dos dados converge para uma distribuição normal conforme o número de dados aumenta. Além disso diversos estudos práticos tem como resultado uma distribuição normal. Podemos citar como exemplo a altura de uma determinada população em geral segue uma distribuição normal. Entre outras características físicas e sociais tem um comportamento gaussiano, ou seja, segue uma distribuição normal.

Como a fórumula e tanto a integral dela são impráticas de calcular, utilizamos a tabela 1.

Tabela da Distribuição Normal Reduzida $P(0 \leq Z \leq z)$

Z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Figura 1: Tabela de Distribuição Normal Reduzida.

2.3.1 Normal Reduzida

$$P(a \leq X \leq b) = P(0 \leq Z \leq z) = \frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}$$

3 Teorema Central do Limite

Se, ao invés de tirarmos uma única amostra (digamos, 100 coletas), tirarmos várias amostras de tamanho "n" (digamos, 20 amostras compostas por cinco coletas: 20x5=100 coletas) e analisarmos a distribuição das médias (ou a soma das amostras) de cada amostra de tamanho n, observaremos que:

- À medida que o tamanho "n" da amostra aumenta, a distribuição das médias amostrais tende a uma distribuição normal
- A média das médias amostrais será a média populacional $\mu = \dfrac{\Sigma x}{n}$
- O desvio padrão das médias amostrais será $\dfrac{\sigma}{\sqrt{n}}$
- Observações importantes:
  - Quando maior o tamanho das amostras, a distribuição das médias será mais próxima de uma distribuição normal.
  - Para $n > 30$, a distribuição das médias amostrais pode ser aproximada satisfatoriamente por uma distribuição normal.

- Se a distribuição da variável x for originalmente uma distribuição normal, então a distribuição das médias amostrais terá distribuição normal para qualquer tamanho amostral n

Então, o que o TCL tem de mais? Ele nos diz que qualquer que seja a forma da sua distribuição original, suas médias resultam numa distribuição.

Para encontrarmos a distribuição da média, basta conhecermos a média da população e o desvio.

4 Gabarito

2.1 - (a) 0,3875 (b) 0,3486 (c) 0,1937 (d) 0,2639 (e) 0,9298

2.2 - Aproximadamente 2 alunos.

3.1 - 0,09645

4.1 - 0,18522

4.1 - 0,808847

6.1 - 0,584

7.1 - 0,05859