Modelo Uniforme

- Exemplo: Uma rifa tem 100 bilhetes numerados de 1 a 100. 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 probabilidade de ser sorteado?
  - espalhar os números é a melhor forma de ganhar o sorteio?
  - assumindo honestidade da rifa, todos os números têm a mesma probabilidade de ocorrência, com $\frac{1}{100}$ para cada um.
  - como eu e meu colega temos 5 bilhetes, temos a mesma probabilidade de ganhar a rifa: $\frac{5}{100} = \frac{1}{20}$
  - assim, a probabilidade de ganhar depende somente da quantidade de bilhetes que se tem na mão, independente da numeração.

Profa. V.A. González-López, IMECC/UNICAMP

Modelo Uniforme

- Modelo
  - X é uma variável aleatória cujos possíveis valores são representados por $\Omega = \{x_1, x_2, ..., x_k\}$
  - X segue o modelo uniforme discreto se atribui a mesma probabilidade $\frac{1}{k}$ a cada um desses possíveis valores
  - $P(X = x_j) = \frac{1}{k} \quad \forall \ 1 \leq j \leq k$
  - $E(X) = \sum_{i=1}^{k} p_i x_i = \frac{1}{k} \sum_{i=1}^{k} x_i$
  - $Var(X) = \sum_{i=1}^{k} p_i (x_i - E(X))^2 = \frac{1}{k} \sum_{i=1}^{k} (x_i - E(X))^2$
- Compare este modelo com a estatística descritiva do início de semestre...

Profa. V.A. González-López, IMECC/UNICAMP

Propriedade Geral da Variância

- Definição: $Var(X) = E([X - E(X)]^2)$

$$
\begin{aligned}
&= E([X - \mu_X]^2) = E(X^2 - 2X\mu_x + \mu_X^2) \\
&= E(X^2) - 2\mu_X E(X) + \mu_X^2 = \mu_{X^2} - 2\mu_X\mu_X + \mu_X^2 \\
&= \mu_{X^2} - 2\mu_X^2 + \mu_X^2 = \mu_{X^2} - \mu_X^2 \\
&= E(X^2) - [E(X)]^2
\end{aligned}
$$

- Então: $Var(X) = E(X^2) - [E(X)]^2$

Modelo Uniforme

- Retomando o modelo da uniforme
  - já sabemos que $E(X) = \frac{1}{k}\sum_{i=1}^{k} x_i$
  - portanto $[E(X)]^2 = \frac{1}{k^2}(\sum_{i=1}^{k} x_i)^2$
  - $E(X^2) = \sum_{i=1}^{k} p_i x_i^2 = \frac{1}{k}\sum_{i=1}^{k} x_i^2$
  - finalmente $Var(X) = \frac{1}{k}[\sum_{i=1}^{k} x_i^2 - \frac{1}{k}(\sum_{i=1}^{k} x_i)^2]$

Modelo Uniforme

Profa. V.A. González-López, IMECC/UNICAMP

- Exemplo: $X =$ resultado obtido no lançamento de um dado honesto

x	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

  - $E(X) = \frac{1}{6} \times (1+2+3+4+5+6) = \frac{21}{6} = 3.5$
  - $Var(X) = \frac{1}{6}[(1+4+9+16+25+36) - \frac{1}{6} \times (21)^2] = \frac{35}{2} = 17.5$

Modelo Uniforme

- Cálculo da f.d.a. de uma variável uniforme discreta:
  - $F(x) = P(X \leq x) = \frac{n(x)}{k}$
  - $n(x) =$ número de x_i's tais que $x_i \leq x$
- Retomando o exemplo do dado:
  - $F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{2}{6}$
  - $F(2.5) = P(X \leq 2.5) = P(X = 1) + P(X = 2) = \frac{2}{6}$

Profa. V.A. González-López, IMECC/UNICAMP

Modelo Bernoulli

Profa. V.A. González-López, IMECC/UNICAMP

- Ensaios tipo Bernoulli: estamos interessados na ocorrência de um sucesso ou fracasso.
- Exemplo: uma pessoa é escolhida ao acaso entre os moradores de uma cidade, e é perguntado a ela se concorda com um projeto. As possíveis respostas são apenas "Sim"ou "Não".
- $X = \begin{cases} 1, & \text{evento de interesse ocorre} \\ 0, & \text{caso contário} \end{cases}$
- $\Omega = \{0, 1\}$
- $P(X = 1) = P(sucesso) = p \Rightarrow P(X = 0) = P(fracasso) = 1 - p$

Modelo Bernoulli

- Forma geral de escrever a probabilidade de uma variável Bernoulli:

  $P(X = x) = p^x(1-p)^{1-x}$ em que $x = 0, 1$

- $E(X) = 0 \times (1-p) + 1 \times p = p$
- $E(X^2) = 0^2 \times (1-p) + 1^2 \times p = p$
- $Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p)$
- f.d.a.: $\mathrm{F(x)} = \begin{cases} 0, & x < 0 \\ 1-p, & x \in [0,1) \\ 1, & x \geq 1 \end{cases}$

Modelo Bernoulli

Profa. V.A. González-López, IMECC/UNICAMP

- Exemplo: lançamos um dado e consideramos como sucesso, a obtenção da face 5. Supondo que o dado é honesto:

x	0	1
$p(x)$	$\frac{5}{6}$	$\frac{1}{6}$

  - $P(X = x) = p^x(1-p)^{1-x} = (\frac{1}{6})^x(\frac{5}{6})^{1-x}$ em que $x = 0, 1$
  - $E(X) = \frac{1}{6}$
  - $E(X^2) = \frac{1}{6}$
  - $Var(X) = \frac{1}{6} - \frac{1}{36} = \frac{6-1}{36} = \frac{5}{36}$

Modelo Binomial

- A repetição de ensaios Bernoulli independentes dá origem a uma variável aleatória Binomial.
- Exemplo: Sabe-se que a eficiência de uma vacina é de 80%. Um grupo de 3 indivíduos é sorteado, dentre a população vacinada, e cada um submetido a testes para averiguar se está imunizado. Nesse caso, consideramos como sucesso, a imunização.
  - $X_i = \begin{cases} 1, & \text{indivíduo i está imunizado} \\ 0, & \text{caso contrário} \end{cases}$
  - pelo enunciado, sabe-se que $P(X_i = 1) = p = 0.8$

Modelo Binomial

Profa. V.A. González-López, IMECC/UNICAMP

- As respostas associadas aos indivíduos X_1, X_2 e X_3 são independentes e cada uma é uma variável Bernoulli
- Se o interesse está em estudar $X =$ número de indivíduos imunizados no grupo, X poderá assumir valores em $\Omega = \{0, 1, 2, 3\}$
- Note que $X = X_1 + X_2 + X_3$

Modelo Binomial

evento	P(evento)	X
$X_1 = 0,\ X_2 = 0,\ X_3 = 0$	$(0.2)^3$	0
$X_1 = 1,\ X_2 = 0,\ X_3 = 0$	$0.8 \times (0.2)^2$	1
$X_1 = 0,\ X_2 = 1,\ X_3 = 0$	$0.8 \times (0.2)^2$	1
$X_1 = 0,\ X_2 = 0,\ X_3 = 1$	$0.8 \times (0.2)^2$	1
$X_1 = 1,\ X_2 = 1,\ X_3 = 0$	$(0.8)^2 \times 0.2$	2
$X_1 = 1,\ X_2 = 0,\ X_3 = 1$	$(0.8)^2 \times 0.2$	2
$X_1 = 0,\ X_2 = 1,\ X_3 = 1$	$(0.8)^2 \times 0.2$	2
$X_1 = 1,\ X_2 = 1,\ X_3 = 1$	$(0.8)^3$	3

Modelo Binomial

- Assim, as probabilidades de cada valor possível de X são:

x	0	1	2	3
$p(x)$	$(0.2)^3$	$3 \times 0.8 \times (0.2)^2$	$3 \times (0.8)^2 \times 0.2$	$(0.8)^3$

- E o comportamento de X é completamente determinado pela função

$$P(X = x) = \binom{3}{x} (0.8)^x (0.2)^{3-x}$$

onde $x = 0, 1, 2, 3$.

Modelo Binomial

- Modelo Geral: Considere a repetição de n ensaios X_i Bernoulli independentes e todos com a mesma probabilidade de sucesso p. A variável aleatória $X = X_1 + \dots + X_n$ que representa o total de sucessos pertence ao modelo Binomial com parâmetros n e p.
  - A probabilidade de se observar x é dada pela expressão geral

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \forall \ x = 0, 1, \dots, n;$$

  - $E(X) = np$; $Var(X) = np(1-p)$;
  - notação: $X \sim Bin(n, p)$

Modelo Binomial

- No exemplo da vacina, temos então $n = 3$ e $p = 0.8$
  - $X \sim Bin(3, 0.8)$
  - $E(X) = 3 \times 0.8$
  - $Var(X) = 3 \times 0.8 \times 0.2$

Propriedades da Esperança e Variância

Profa. V.A. González-López, IMECC/UNICAMP

- A esperança de uma soma de variáveis aleatórias é a soma das esperanças dessas variáveis
  X e Y variáveis aleatórias $\Rightarrow E(X+Y) = E(X) + E(Y)$
- A variância de uma soma de duas variáveis aleatórias independentes é a soma da variância dessas variáveis
  $Var(X+Y) = Var(X) + Var(Y)$
- Sendo que duas variáveis aleatórias são independentes $\Leftrightarrow$
  $P(X=x, Y=y) = P(X=x)P(Y=y)$

Distribuição Hipergeométrica

- População dividida em duas características
- Extrações casuais sem reposição
- Detalhes:
  - N objetos
  - r têm a característica A
  - $N - r$ têm a característica B
  - um grupo de n elementos é escolhido ao acaso, dentre os N possíveis, sem reposição
- Objetivo: calcular a probabilidade de que este grupo de n elementos contenha x elementos com a característica A

Profa. V.A. González-López, IMECC/UNICAMP

Distribuição Hipergeométrica

- Modelo Geral:

$$P(X=x)=\frac{\begin{pmatrix} r \\ x \end{pmatrix}\begin{pmatrix} N-r \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

  onde $0 \leq x \leq min\{r,n\}$

- X registra o número de elementos dentre os n sorteados, que possuem a característica A
- X tem distribuição Hipergeométrica

Distribuição Hipergeométrica

- Notação: $X \sim Hip(N, n, r)$
- Se X tem distribuição Hipergeométrica com parâmetros N, n, r, então:
  - $E(X) = \frac{nr}{N}$
  - $Var(X) = \frac{nr}{N}\left(1 - \frac{r}{N}\right)\frac{(N-n)}{(N-1)}$

Distribuição Hipergeométrica

Profa. V.A. González-López, IMECC/UNICAMP

- Aplicação: Controle de Qualidade

  Suponha um lote com $N = 100$ elementos a ser analisado. São escolhidas $n = 5$ peças sem reposição. Sabendo que neste lote de 100 elementos, $r = 10$ são defeituosos, a probabilidade de não se obter nenhuma peça defeituosa na amostra retirada é:

$$P(X=0) = \frac{\binom{10}{0}\binom{100-10}{5-0}}{\binom{100}{5}} = \frac{\binom{90}{5}}{\binom{100}{5}} \approx 0.584$$

Distribuição Hipergeométrica

Profa. V.A. González-López, IMECC/UNICAMP

- A probabilidade de se obter pelo menos uma peça defeituosa é:

  $\sum_{i=1}^{5} P(X = i) = 1 - P(X = 0) \approx 0.416$
- $E(X) = \frac{nr}{N} = \frac{5 \times 10}{100} = 0.5$
- $Var(X) = \frac{nr}{N}\left(1 - \frac{r}{N}\right)\frac{(N-n)}{(N-1)} = \frac{5 \times 10}{100}\left(1 - \frac{10}{100}\right)\frac{(100-5)}{(100-1)} \approx 0.432$

Distribuição de Poisson

Profa. V.A. González-López, IMECC/UNICAMP

- Considerando a distribuição $Bin(n,p)$, quando temos grandes valores para n e p pequeno, podemos usar a seguinte aproximação para a probabilidade:

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \approx \frac{e^{-np}(np)^x}{x!}$$

$\forall \;\; x = 0, 1, 2, \ldots$

- Geralmente considera-se o critério $np \leq 7$ para usar essa aproximação

Distribuição de Poisson

Profa. V.A. González-López, IMECC/UNICAMP

- Uma variável aleatória X tem distribuição de Poisson com parâmetro $\lambda > 0$, se sua função de probabilidade é dada por:

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad \forall \; x = 0, 1, 2, \ldots$$

- $E(X) = Var(X) = \lambda$
- λ é chamada de taxa de ocorrência
- Notação: $X \sim P(\lambda)$

Distribuição de Poisson

- Exemplo: $X \sim Bin(100, 0.065)$, deseja-se obter $P(X = 10)$
  - $\lambda = np = 100 \times 0.065 = 6.5 \leq 7$
  - no modelo Binomial:
    $$P(X = 10) = \binom{100}{10} (0.065)^{10} (0.935)^{100-10} = 0.055$$
  - no modelo Poisson: $P(X = 10) = \frac{e^{-6.5}(6.5)^{10}}{10!} \approx 0.056$

Distribuição de Poisson

Profa. V.A. González-López, IMECC/UNICAMP

- Exemplo: A emissão de partículas radioativas tem sido modelada através de uma distribuição de Poisson, com o valor do parâmetro dependendo da fonte utilizada. Supondo que o número de partículas alfa emitidas por minuto é uma variável aleatória seguindo o modelo Poisson com parâmetro $\lambda = 5$, ou seja, a taxa de ocorrência é de 5 emissões a cada minuto. Estamos interessados em calcular a probabilidade de haver mais de duas emissões por minuto.

Distribuição de Poisson

Profa. V.A. González-López, IMECC/UNICAMP

- $X \sim P(5)$
- $P(X > 2) = \sum_{x=3}^{\infty} \frac{e^{-5}5^x}{x!} = 1 - \sum_{x=0}^{2} \frac{e^{-5}5^x}{x!} \approx 0.875$
- $E(X) = \lambda = 5$
- $Var(X) = \lambda = 5$