AULA 16 - Distribuição de Poisson e Geométrica

Susan Schommer

Introdução à Estatística Econômica - IE/UFRJ

Distribuição de Poisson

- Em muitas situações nos deparamos com a situação em que o número de ensaios n é grande $(n \to \infty)$ e p é pequeno $(p \to 0)$.
- Isso para o cálculo da função binomial, nos leva a algumas dificuldades, pois, como podemos analisar, para n muito grande e p pequeno, fica relativamente difícil calcularmos a probabilidade de k sucessos a partir do modelo binomial, isto é, utilizando a função de probabilidade.
- Para modelos desse tipo, iremos usar a distribuição de Poisson.
- Tal expressão é devida a Poisson (matemático e físico francês) no qual publicou o trabalho em 1938.
- Essa distribuição é muito utilizada para calcular probabilidades de ocorrências de defeitos "raros" em sistemas e componentes.

Distribuição de Poisson

- Considere o modelo binomial

$$p(k) = \mathbb{P}(X = k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}.$$

- Podemos reescrever a expressão acima da seguinte forma

$$\begin{aligned}
P(X = k) &= \frac{n!}{k!(n-k)!} p^k \frac{n^k}{n^k} \left(1 - \frac{np}{n}\right)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \frac{(np)^k}{n^k} \left(1 - \frac{np}{n}\right)^{n-k} \quad \text{tomando} \quad \lambda = np \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} 1 \underbrace{\left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)}_{\lim_{n\to\infty} = 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{n-k}}_{\lim_{n\to\infty} = e^{-\lambda}}
\end{aligned}$$

para $k = 0, 1, \ldots$ e $e \approx 2,718$.

$$\lim_{n\to\infty} \mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Distribuição de Poisson

Definition

Uma variável aleatória discreta X segue a distribuição de Poisson com parâmetro λ, $\lambda > 0$, se sua função de probabilidade for dada por

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Utilizamos a notação $X \sim$ Poisson(λ) ou $X \sim$ Po(λ).
O parâmetro λ indica a taxa de ocorrência por unidade medida.

Distribuição de Poisson

Example

Considere um processo que têm uma taxa de $0,2$ defeitos por unidade. Qual a probabilidade de uma unidade qualquer apresentar:

a) dois defeitos? $\mathbb{P}(X = 2) = \dfrac{e^{-0,2}(0,2)^2}{2!} = 0,0164$

b) um defeito? $\mathbb{P}(X = 1) = \dfrac{e^{-0,2}(0,2)^1}{1!} = 0,1637$

c) zero defeito? $\mathbb{P}(X = 0) = \dfrac{e^{-0,2}(0,2)^0}{0!} = 0,8187$

Distribuição de Poisson

Example

Suponha que 10% das crianças de um determinado bairro do Rio de Janeiro prefiram sorvete de baunilha ao de chocolate. Qual a probabilidade de que, se entrevistarmos 10 crianças deste bairro, exatamente 2 duas prefiram soverte de baunilha?
Podemos resolver pela binomial com $n = 10$ e $p = 0,1$ ou pela Poisson $\lambda = np = 1$ (é menos exata, mas mais simples).
Binomial: $\mathbb{P}(X = 2) = \left(\begin{array}{c} 10 \\ 2 \end{array} \right) (0,1)^2 (0,9)^8 \approx 0,194$
Poisson: $\mathbb{P}(X = 2) = \frac{e^{-1}}{2!} \approx 0,184$
Embora pela Poisson tenhamos um erro associado a aproximação da distribuição binomial (que é a distribuição exata), este erro, em muitos casos, não chega a ser significativo.

Distribuição de Poisson

- Seja X uma variável aleatória discreta com distribuição de Poisson, com parâmetro λ, ou seja, $X \sim$ Poisson(λ)
- O valor esperado de X, que frequentemente é chamado de taxa de defeitos, é dado por:

$$E(X) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} \lambda \left[\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\right]$$
$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \left[\frac{\lambda^{k-1}}{(k-1)!}\right] = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \left[\frac{\lambda^k}{k!}\right] = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

- Variância:

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \lambda(\lambda+1) - \lambda^2 = \lambda.$$

Distribuição Geométrica

- Suponha-se que realizemos um experimento ϵ e que estejamos interessados apenas na ocorrência ou não-ocorrência de algum evento A.
- Como na explicação da distribuição binomial, que realizemos ϵ repetidamente, que as repetições sejam independentes e que em cada repetição $P(A) = p$ e $P(\underline{A}) = 1 - p$.
- Repetimos o experimento até que A ocorra pela primeira vez. Na binomial, o número de repetições era predeterminado, enquanto aqui é uma variável aleatória.
- Defina-se a variável aleatória X como o número de repetições necessárias para obter a primeira ocorrência de A e sendo as primeiras $(j-1)$ repetições de ϵ que derem o resultado $\underline{A}$.

$$\mathbb{P}(X = j) = (1-p)^{j-1}p, \quad j = 1, 2, \ldots$$

Geométrica (conta o número ensaios para se obter um sucesso)

Definition

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

$$\mathbb{P}(X = j) = (1 - p)^{j-1} p, \quad j = 1, 2, \ldots$$

Usaremos a notação X $\sim$ Geo(p).

Geométrica

Example (Conta o número de ensaios para se obter um sucesso)

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 ocorra o primeiro 1. Qual a probabilidade de obtermos 1 no terceiro lançamento.
Como o dado é honesto, a probabilidade de, em um lançamento, obtermos qualquer face é igual a $1/6$. Neste caso, a probabilidade de se obter a face 1 (sucesso) é $1/6$ e a probabilidade de se obter qualquer outra face (fracasso) é $5/6$. Podemos definir a v.a.

$$Y = \begin{cases} 1, & \text{se obtemos 1 no lançamento do dado} \\ 0, & \text{caso contrário} \end{cases}$$

Distribuição Geométrica

Example (Continuação: Conta o número de ensaios para se obter um sucesso)

Neste caso, $Y \sim$ Bernoulli$(1/6)$ e, se definirmos X como sendo a variável que representa o número de lançamentos até a obtenção do primeiro sucesso (aparecimento da face 1), temos que $X \sim$ Geo$(1/6)$. Portanto, se estamos interessados no cálculo da probabilidade de obter 1 no terceiro lançamento, precisamos calcular $\mathbb{P}(X = 3)$, ou seja,

$$\mathbb{P}(X = 3) = (1 - p)^2 p = \left(1 - \frac{1}{6}\right)^2 \frac{1}{6} = \left(\frac{5}{6}\right)^2 \frac{1}{6} = \frac{5^2}{6^3} \approx 0,115741$$

Distribuição Geométrica

- O valor esperado e variância no caso em que conta o número ensaios para se obter um sucesso

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} j(1-p)^{j-1}p = p\sum_{j=1}^{\infty} j(1-p)^{j-1} = \frac{p}{p^2} = \frac{1}{p}.$$

$$\mathsf{Var}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$