Principais Distribuições

Vicente Garibay Cancho

11 de outubro de 2023

Principais modelos Discretos e Contínuos

Modelos discretos

- Uniforme
- Bernoulli
- Binomial
- Geométrico
- Binomial Negativa
- Poisson
- Hipergeometrica

Modelos contínuos

- Uniforme
- Exponencial
- Gama
- Weibull
- Beta
- Normal

Modelo Uniforme

Definição

Uma variável aleatória X tem uma distribuição uniforme discreta se e somente se sua distribuição de probabilidade é dada por

$$f_X(x;N) = \begin{cases} \frac{1}{N}, & x = x_1, \ldots, x_N \\ 0, & \text{caso contrario} \end{cases}$$

onde $x_i \neq x_j$ e $i \neq j$.

A média e variância desta distribuição são

$$E(X) = \mu_X = \sum_{i=1}^{N} x_i f(x_i;N) = \sum_{i=1}^{N} x_i \frac{1}{N} = \frac{x_1 + \cdots + x_k}{N}$$

e

$$Var(X) = \sigma_X^2 = E[(X-\mu)^2] = \sum_{i=1}^{N} (x_i - \mu)^2 \frac{1}{N} = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}.$$

Modelo Uniforme

No caso especial em que $x_i = i$, a distribuição uniforme discreta torna-se

$$f_X(x; N) = \begin{cases} \frac{1}{N}, & x = 1, 2, \ldots, N \\ 0, & \text{caso contrario} \end{cases}$$

aplica-se, por exemplo, ao número de pontos ao lançar um dado equilibrado.

Exercício-1

Se a variável aleatória X tem distribuição uniforme discreta, $f(x; k) = 1/k$, para $x = 1, \ldots, k$, mostre que

(a) A média: $\mu_X = \frac{k+1}{2}$,

(b) A variância: $\sigma_X^2 = \frac{k^2-1}{12}$,

(c) A função geradora de momentos:

$$m_X(t) = \frac{e^t(1 - e^{kt})}{k(1 - e^t)}$$

Modelo Bernoulli

Na prática muitos experimentos admitem apenas dois resultados. Exemplos:

- Uma peça é classificada como boa ou defeituosa;
- O resultado de um exame médico para detecção de uma doença é positivo ou negativa.
- Um entrevistado concorda ou não com a afirmação feita;
- No lançamento de um dado ocorre ou não face 6;
- No lançamento de uma moeda ocorre cara ou coroa.
- Situações com alternativos dicotômicas, podem ser representadas genericamente por resposta do tipo sucesso (0)-fracasso (1).

Esses experimentos recebem o nome de Ensaio de Bernoulli e originam uma v.a. com distribuição de Bernoulli (ou modelo Bernoulli).

Modelo Bernoulli

Uma variável aleatória (v.a.) X tem distribuição de Bernoulli com parâmetro θ $(0 < \theta < 1)$, se somente se sua função de probabilidade (f.p.) é dada por:

$$f_X(x;\theta) = \begin{cases} \theta^x(1-\theta)^{1-x}, & \text{se } x = 0, 1 \\ 0, & \text{caso contrario} \end{cases}$$

Equivalentemente,

$$f_X(x;\theta) = \theta^x(1-\theta)^{1-x} I_{\{0,1\}}(x),$$

onde

$$I_A(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases}.$$

Notação: $X \sim \text{Bernoulli}(\theta)$.

Modelo Bernoulli

Propriedades

Se $X \sim \text{Bernoulli}(\theta)$, então

(1) a média e variância são

$$E[X] = \theta, \text{ e } \; Var(X) = \theta(1-\theta),$$

respectivamente.

(2) a função geradora de momentos (f.g.m.) é

$$m_X(t) = E(e^{tX}) = (1-\theta) + \theta e^t, \quad t \leq 1.$$

(3) o r-ésimo momento,

$$\mu'_r = E(X^r) = \theta, \; r = 1, 2, \ldots$$

Modelo Binomial

Repetições independentes de ensaios de Bernoulli dão origem ao modelo Binomial.

Exemplo-1

Suponha que uma moeda é lançada 3 vezes e probabilidade de cara seja θ em cada lançamento. Determinar a distribuição de probabilidade da variável X, número de caras nos 3 lançamentos.

- Denotemos, S: sucesso, ocorrer cara e F:fracasso, ocorrer coroa.
- O espaço amostral para o experimento de lançar um moeda 3 vezes é:
$$S = \{FFF.FFS, FSF, SFF, FSS, SFS, SSF, SSS\}.$$

Modelo Binomial

S	Probabilidade	$X(w) = x$
FFF	$(1-\theta)^3$	0
SFF	$\theta(1-\theta)^2$	1
FSF	$\theta(1-\theta)^2$	1
FFS	$\theta(1-\theta)^2$	1
SSF	$\theta^2(1-\theta)$	2
SFS	$\theta^2(1-\theta)$	2
FSS	$\theta^2(1-\theta)$	2
SSS	θ^3	3

Dai tem-se

$$P(X=0) = P(\{FFF\}) = (1-\theta)^3$$
$$P(X=1) = P(\{SFF, FSF, FFS\}) = 3\theta(1-\theta)^2$$
$$P(X=2) = P(\{SSF, SFS, FSS\}) = 3\theta^2(1-\theta)$$
$$P(X=1) = P(\{SSS\}) = \theta^2$$

Modelo Binomial

A distribuição de probabilidade da v.a X é dada por

x	0	1	2	3
$f_X(x) = P(X = x)$	$(1-\theta)^3$	$3\theta(1-\theta)^2$	$3\theta^2(1-\theta)$	θ^3

O comportamento de X , pode ser representado pela seguinte função

$$f_X(x) = \begin{cases} \binom{3}{x}\theta^x(1-\theta)^{n-x}, & x = 0, 1, 2, 3. \\ 0 \quad , & \text{c.c.} \end{cases}$$

onde $\binom{3}{x} = \frac{3!}{x!(3-x)!}$ é o coeficiente binomial.

Modelo Binomial

Uma v.a. X tem distribuição binomial com parâmetros n e θ se sua f.p. é dada por:

$$f_X(x; n, \theta) = \begin{cases} \binom{n}{x}\theta^x(1-\theta)^{n-x}, & x = 0, 1, \ldots, n. \\ 0 \quad , & \text{c.c.} \end{cases}$$

Notação: $X \sim \text{Binomial}(n, \theta)$.

A função de distribuição acumulada da v.a $X \sim \text{Binomial}(n, p)$ é dado por

$$F_X(x; n, \theta) = P[X \leq x] = \begin{cases} \sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i}\theta^i(1-\theta)^{n-i}, & 0 \leq x \leq n \\ 0 \quad , & \text{c.c.} \end{cases}$$

onde $\lfloor x \rfloor$ é o maior número inteiro menor ou igual a x.

Modelo binomial

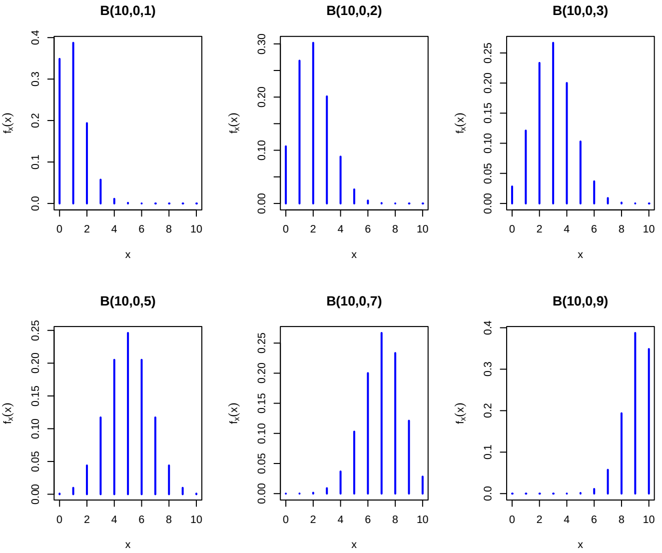


Figura: Função de probabilidade da distribuição binomial

Modelo binomial

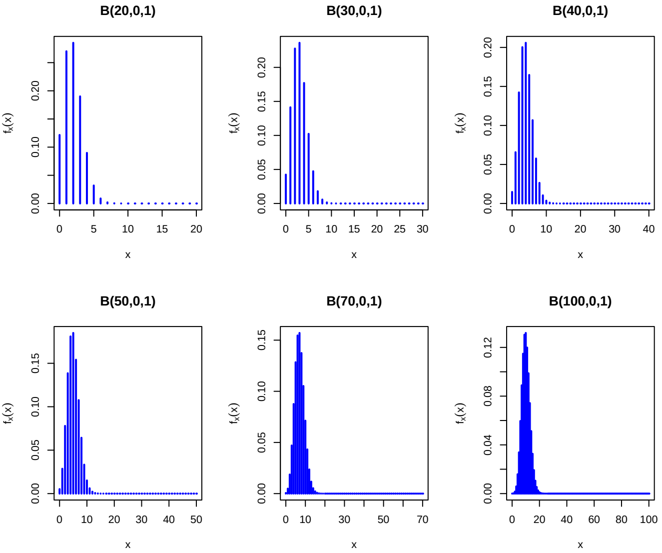


Figura: Função de probabilidade da distribuição binomial

Modelo binomial

Propriedades

Se $X \sim Binomial(n, \theta)$, então

(1) a média e variância são

$$E[X] = n\theta, \text{ e, } Var(X) = n\theta(1-\theta),$$

respectivamente;

(2) a f.g.m. é dada por

$$m_X(t) = [(1-\theta) + \theta e^t]^n, \ t \leq 1.$$

Modelo binomial

Exemplo-1

O professor da disciplina de Estatística 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 vão as aulas e não estudaram para a prova. O professor estabeleceu que para aprovar deve contestar corretamente ao menos 6 questões. Se 200 alunos se apresentaram, quantos alunos aprovaram a disciplina?

Solução:

- Seja a v.a. X : número de questões respondidas corretamente nas 10 questões. Então o evento de interesse é:
- S:questão respondida corretamente, $\Longrightarrow P(S) = 1/5$
- F:questão respondida incorretamente, $\Longrightarrow P(F) = 4/5$
- Assim, $X \sim \text{Binomial}(10, 1/5)$

Modelo Binomial

- Assim,

$$f_X(x;10,1/4) = \begin{cases} \binom{10}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{10-x}, & x = 0,1,\ldots,10. \\ 0 \quad , & \text{c.c.} \end{cases}$$

- A probabilidade de aprovar a prova um aluno é

$$P(X \geq 6) = 1 - P(X < 6) = 1 - F(5) = 0,00637$$

- Portanto, dos 200 alunos que fizeram a prova aprovariam: $200(0,00637) \approx 2$, alunos
- A função de probabilidade, função de distribuição acumulada, função quantil e a geração de v.a´s binomial no R

```
dbinom(x, size, prob, log = FALSE)
pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
rbinom(n, size, prob)
```

Aplicação da distribuição binomial numa amostra

Uma amostra aleatória de n elementos de uma população pode considerar-se como um experimento que consiste em n ensaios independentes nos seguintes casos:.

- Quando a amostra é extraída com ou sem reposição de uma população infinita;
- Quando a amostra é sorteada com reposição de uma população finita.

Exemplo

Numa população grande de *Drosophila melanogaster*, o 25% da população tem mutação nas aças. Foram selecionados uma amostra aleatória de 300 moscas para um exame de mutação das aças. Defina X como o número de moscas que tem mutação nas aças na amostra. Determine $E(X)$ e $Var(X)$

Solução:

1. $X(w)$: número de moscas que tem mutação de aças na amostra de 300 moscas.
2. A população é grande (infinita), não interessa como foi extraída a amostra (com ou sem reposição), aplica-se a distribuição binomial com $n = 300$ e $\theta = 0,25$.
3. A função de probabilidade de X é dada por

$$f_X(x) = \begin{cases} \binom{300}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{300-x}, & x = 0, 1, \ldots, 300 \\ 0, & c.c \end{cases}$$

4. A média e variância da variável aleatória X

$$E[X] = n\theta = 300/4 = 75, \ Var(X) = n\theta(1-\theta) = 75 \times \frac{3}{4} = \frac{225}{4}$$

Exemplo

Um lote contém 100 peças dos quais 10 estão com defeito. Sorteia-se uma amostra de 5 peças do lote com reposição. Defina X a v.a que denota número de peças com defeito na amostra. Determine: (a) a f.p de X; (b) $E[X]$ e $Var[X]$.

Solução:

1. $X(w)$: número de peças com defeito na amostra de 5 peça.
2. A população finita, amostra com reposição, aplica-se a distribuição binomial com $n = 5$ e $\theta = 10/100 = 0,1$.
3. A função de probabilidade de X é dada por

$$f_X(x) = \begin{cases} \binom{300}{x}(0,1)^x(0,9)^{10-x}, & x = 0, 1, \ldots, 5 \\ 0, & c.c \end{cases}$$

4. A média e variância da variável aleatória X

$$E[X] = n\theta = 5(0,1) = 0,5 \;\; Var(X) = n\theta(1-\theta) = 0,5 \times 0,9 = 0,45$$

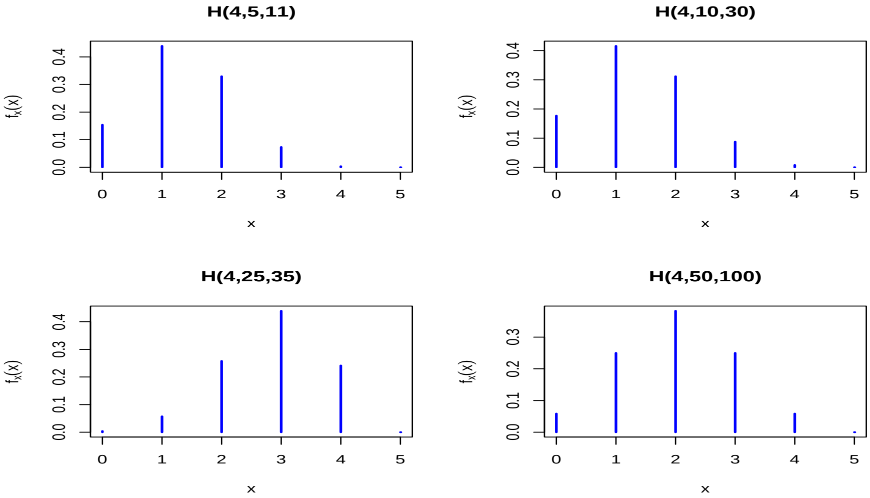
Modelo Hipergeométrica

- Suponha uma população finita de N elementos. Suponha que M ($M \leq N$) elementos da população está na classe de interesse.
- A classe de interesse dependerá, da situação em consideração. Podem ser defeituosos (versus não defeituosos) no caso de um lote de produção, ou pacientes curados (versus não curados) no tratamento de pacientes com certa doença.
- Seleciona-se, sem reposição, uma amostra aleatória de tamanho n e a variável de interesse, X, é o número de itens na amostra que pertence à classe de interesse.
- Pode-se mostrar que a f.p. da v.a. X é dado por

$$f_X(x) = \begin{cases} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, & x = 0, 1, \ldots \min(n, M) \\ 0, & c.c \end{cases}$$

Notação $X \sim H(n, M, N)$.

Modelo Hipergeométrica



Modelo Hipergeométrica

A função de distribuição acumulada da v.a $X \sim Po(\mu)$ é dado por

$$F(x;\mu) = \begin{cases} \sum_{i=0}^{\lfloor x \rfloor} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}, & x \geq 0 \\ 0, & c.c \end{cases}$$

Propriedades

Se $X \sim H(n, M, N)$, então

- $E[X] = n\left[\frac{M}{N}\right]$,
- $Var(X) = n\left[\frac{M}{N}\right]\left[1 - \frac{M}{N}\right]\left[\frac{N-n}{N-1}\right]$

A f.p, FDA, função quantil e a geração de v.a´s com distribuição Hipergeométrica no R

```
dhyper(x, m, n, k, log = FALSE)
phyper(q, m, n, k, lower.tail = TRUE, log.p = FALSE)
qhyper(p, m, n, k, lower.tail = TRUE, log.p = FALSE)
rhyper(nn, m, n, k)
```

Exemplo

Um lote contém 100 peças dos quais 10 estão com defeito. Sorteia-se uma amostra de 5 peças do lote sem reposição. Defina X a v.a que denota número de peças com defeito na amostra. Determine: (a) a f.p de X; (b) $E[X]$ e $Var[X]$.

Solução:

1. $X(w)$: número de peças com defeito na amostra de 5 peça.
2. A população finita, amostra sem reposição, aplica-se a distribuição hipergeometrica com $n = 5$ $N = 100$ e $M = 10$
3. A função de probabilidade de X é dada por

$$f_X(x) = \begin{cases} \frac{\binom{10}{x}\binom{90}{5-x}}{\binom{100}{5}}, & x = 0, 1, \ldots, 5 \\ 0, & c.c \end{cases}$$

4. A média e variância da variável aleatória X

$$E[X] = n\left[\frac{M}{N}\right] = 5[\frac{10}{100}] = 0,5$$

$$Var(X) = 5[\frac{10}{100}][\frac{90}{100}]\left[\frac{100-5}{100-1}\right] = 0,4275$$

Modelo Geométrico

Em uma série de ensaios independentes de Bernoulli e todos com a mesma probabilidade de sucesso θ. A variável aleatória X, que conta o número total de fracassos até primeiro sucesso é denominada de variável aleatória geométrica com parâmetro θ e sua função de probabilidade é dado por:

$$f_X(x;\theta) = \begin{cases} \theta(1-\theta)^x, & x = 0,1,2,\ldots \\ 0 \quad , & \text{c.c.} \end{cases}$$

Notação: $X \sim \mathrm{G}(\theta)$.
A função de distribuição acumulada da v.a $X \sim \mathrm{G}(\theta)$

$$F_X(x;\theta) = \begin{cases} 0 & x < 0 \\ 1-(1-\theta)^{\lfloor x \rfloor + 1}, & x \geq 0 \end{cases}$$

Modelo Geometrico

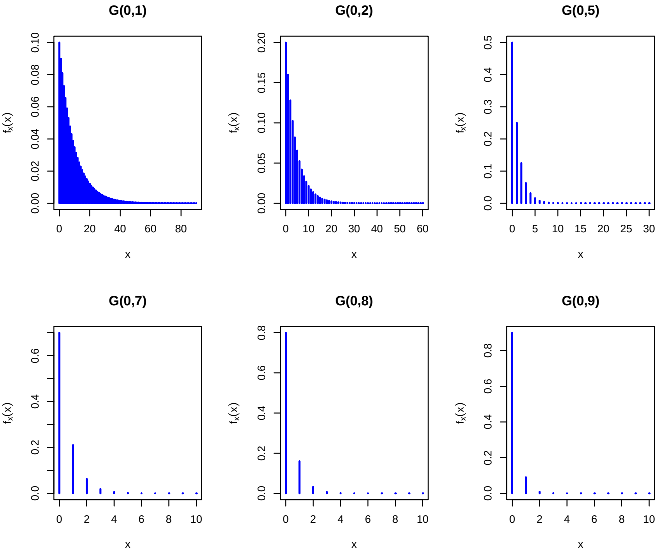


Figura: Função de probabilidade do Modelo Geométrico

Modelo Geométrico

Propriedades

Se $X \sim G(\theta)$, então

1) $E[X] = (1-\theta)/\theta$ e $Var(X) = (1-\theta)/\theta^2$.
2) A f.g.m. é dado por $m_X(t) = (1+\theta e^t)^{-1}$.
3) A falta de memoria

$$P(X > j + k | X > j) = P(X > k), \forall k,\ j = 1, 2, \ldots$$

Modelo Geométrico

Exemplo-2

Certo experimento deve ser realizado até se obter um resultado bem sucedido. Os experimentos são independentes, e o custo de realização do experimento é de \$ 25.000; não entanto, se resulta um fracasso o custo é de \$5000 para o preparo para a próxima prova. O experimentador gostaria de determinar o custo esperado do projeto. Se a probabilidade de sucesso de uma única prova é de 0.25.

Solução:

- Se X : é o número de experimentos desnecessários para se obter um resultado bem sucedido.
- então, $\sim G(1/4)$.
- A f.p de X é dada por

$$f_X(x; 1/4) = \begin{cases} (\frac{1}{4})(\frac{3}{4})^x, & x = 0, 1, \dots \\ 0, & c.c \end{cases}$$

Modelo Geométrico

Continuação do Exemplo-2

- Seja a v.a C que representa o custo
- Assim,

$$C = 25.000(X + 1) + 5000X = 30.000X + 25.500$$

- O custo esperado resulta:

$$E(C) = 30.000(X){+}25.500 = 30.000(\frac{0.75}{0.25}){+}25.000 = 115.000.$$

(b) Suponha que o experimentador tenha \$500.000 de orçamento. Qual é a probabilidade que o trabalho experimental custar no máximo esse valor?

$$\begin{aligned} P(C \leq 500.000) &= P(30.000X + 25.500 \leq 500.000) = P(X \leq 15,816) \\ &= F_X(15.816) = 0,9899. \end{aligned}$$

Modelo Geométrico

A função de probabilidade, função de distribuição acumulada, função quantil e a geração de v.a´s com distribuição Geométrico no R

```
dgeom(x, prob, log = FALSE)
pgeom(q, prob, lower.tail = TRUE, log.p = FALSE)
qgeom(p, prob, lower.tail = TRUE, log.p = FALSE)
rgeom(n, prob)
```

No Exemplo,

```
> pgeom(15.817, 0.25)
[1] 0.9899774
```

Modelo Binomial negativa

Em uma série de ensaios independentes de Bernoulli e todos com a mesma probabilidade de sucesso θ. A variável aleatória X, que conta o número total de fracassos até r-ésimo sucesso. Neste caso a v.a X tem distribuição binomial negativa com parâmetros r e θ e sua função de probabilidade é dado por

$$f_X(x; r, \theta) = \begin{cases} \binom{x+r-1}{r-1}\theta^r(1-\theta)^x, & x = 0, 1, \ldots. \\ 0 \quad , & \text{c.c.} \end{cases}$$

Notação: $X \sim \text{BN}(r, \theta)$.
A função de distribuição acumulada da v.a $X \sim \text{BN}(r, \theta)$

$$F_X(x; r, \theta) = \begin{cases} 0 & x < 0 \\ \sum_{j=1}^{\lfloor x \rfloor} \binom{j+r-1}{r-1}\theta^r(1-\theta)^j & x \geq 0 \end{cases}$$

Modelo Binomial negativa

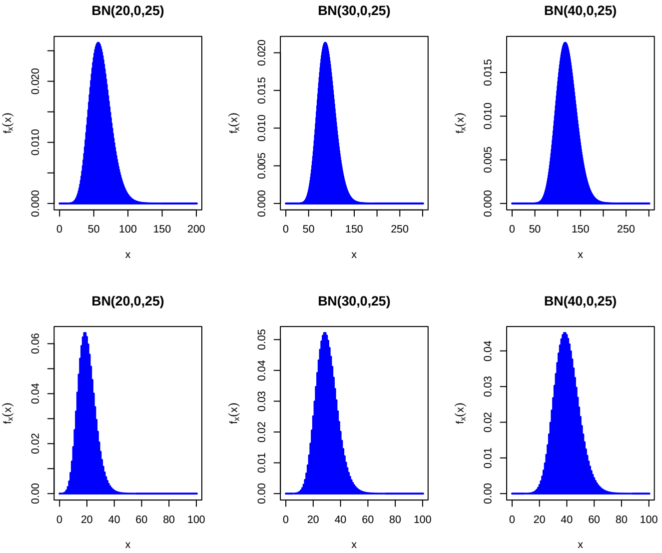


Figura: Função de probabilidade do Modelo Binomial negativo

Modelo Binomial negativa

Propriedades

Se $X \sim BN(r, \theta)$, então

1) $E[X] = r(1-\theta)/\theta$ e $Var(X) = r(1-\theta)/\theta^2$.
2) A f.g.m. é dado por $m_X(t) = (1+\theta e^t)^{-r}$.

- A função de probabilidade, função de distribuição acumulada, função quantil e a geração de v.a´s com distribuição binomial negativa no R

```
dnbinom(x, size, prob, mu, log = FALSE)
pnbinom(q, size, prob, mu, lower.tail = TRUE, log.p = FALSE)
qnbinom(p, size, prob, mu, lower.tail = TRUE, log.p = FALSE)
rnbinom(n, size, prob, mu)
```

Modelo Binomial negativa

Exemplo-4

A probabilidade com que um bit transmitido através de um canal digital de transmissão seja recebido com erro é de 0,1. Considere que as transmissões sejam eventos independentes. Qual é a probabilidade de que seja necessário no máximo 10 transmissões até que ocorra o quarto erro?

Solução

- Seja X a v.a que denota o número de bits transmitidos sem erros até o quarto erro.
- $X \sim BN(4, 0, 1)$;
- Assim

$$f(x; r, \theta) = \begin{cases} \binom{x+4-1}{4-1} \left(\frac{1}{10}\right)^4 \left(\frac{9}{10}\right)^x, & x = 0, 1, \ldots . \\ 0 \quad , & \text{c.c.} \end{cases}$$

- $P(X \leq 6) = F(6) = pnbinom(6, 4, 1/10) = 0, 0127952.$

Modelo Poisson

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

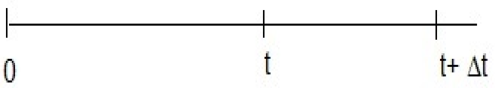
Exemplo:

- Número de consultas a uma base de dados em um minuto.
- Número de acidentes de trabalho por semana em uma empresa industrial.
- Número de machas (falhas) por metro quadrado no esmaltado de uma geladeira.
- Número de chamadas que chegam a uma central telefônica de uma empresa num intervalo de tempo (digamos de 8,0 a 12,0).
- Número de autos que chegam ao Campus entre 7,0 a.m. a 10,0 a.m.
- Número de microorganismos por cm cúbico de água contaminada.

Modelo Poisson

- Processo de Poisson;
- Aproximação do modelo Binomial.

Desenvolvimento a partir de um processo de Poisson



- A variável de interesse, X_t número de chegadas que ocorrem no intervalo $[0, t]$.
- O suporte da v.a X_t é $R_{X_t} = \{0, 1, 2, \ldots\}$.

Modelo Poisson

Hipóteses

- O número de chegadas durante intervalos de tempo disjuntos sejam variáveis aleatórias independentes,
- Existe uma quantidade positiva tal que, para qualquer intervalo de tempo pequeno, Δt, os seguintes postulados sejam válidos.

  1. A probabilidade de que exatamente uma chegada ocorrerá num intervalo de comprimento Δt é aproximadamente $\lambda \Delta t$
  2. A probabilidade de que ocorrerão zero chegadas no intervalo é aproximadamente $1 - \lambda \Delta t$
  3. A probabilidade de que duas ou mais chegadas ocorrerão no intervalo é igual a zero.

onde λ é denominado de taxa média de chegada ou taxa média de ocorrência

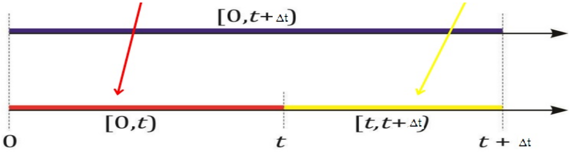
Modelo Poisson

Seja

$$p(x) = P(X_t = x) = p_x(t), \; x = 0, 1, 2, \ldots$$

Fixamos o tempo t e obtemos

$$\begin{aligned} p_0(t+\Delta t) &= P(0 \text{ ocorrências no intervalo } [0, t+\Delta t)) \\ &= P(0 \text{ ocorrências em } [0,t) \text{ e } 0 \text{ ocorrências em } [\mathrm{t}, t+\Delta t)\ ) \\ &= p_0(t) p_0(\Delta t) \end{aligned}$$



$$p_0(t+\Delta t) \approx [1 - \lambda \Delta t] p_0(t)$$

Modelo Poisson

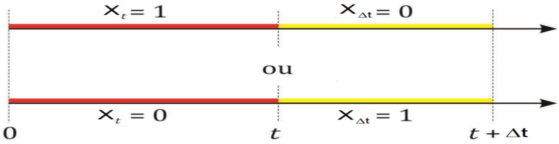
de modo que

$$\left[\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t}\right]=-\lambda p_0(t)$$

e

$$\lim_{\Delta t\longrightarrow 0}\left[\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t}\right]=p_0'(t)=-\lambda p_0(t)$$

$$p_1(t+\Delta t) = p_1(t)p_0(\Delta t) + p_0(t)p_1(\Delta t)$$



$$p_1(t+\Delta t) \approx \lambda \Delta t p_0(t) + [1 - \lambda \Delta t] p_1(t)$$

Daí

$$\frac{p_1(t+\Delta t) - p_1(t)}{\Delta t} = \lambda p_0(t) - p_1(t)$$

e

$$\lim_{\Delta t \longrightarrow 0} \left[\frac{p_1(t+\Delta t) - p_1(t)}{\Delta t} \right] = p_1'(t) = \lambda p_0(t) - p_1(t)$$

Modelo Poisson

Para $x > 0$,

$$p_x(t + \Delta t) = \lambda \Delta t p_{x-1}(t) - [1 - \lambda \Delta t] p_x$$

de modo que

$$\left[\frac{p_x(t + \Delta t) - p_x(t)}{\Delta t} \right] = -\lambda p_{x-1}(t) - \lambda p_x(t)$$

e

$$\lim_{\Delta t \longrightarrow 0} \left[\frac{p_x(t + \Delta t) - p_x(t)}{\Delta t} \right] = p'_x(t) = -\lambda p_{x-1}(t) - \lambda p_x(t)$$

Modelo Poisson

Assim temos um sistema de equações diferenciais:

$$p_0'(t) = -\lambda p_0(t) \tag{1}$$

e

$$p_x'(t) = -\lambda p_{x-1}(t) - \lambda p_x(t),\ x = 1, 2, \ldots \tag{2}$$

A solução dessas equações é

$$p_x(t) = (\lambda t)^x e^{-\lambda t}/x!,\ \ x = 0, 1, 2, \ldots \tag{3}$$

Assim, para t fixo fazemos $\mu = \lambda t$ e obtemos a distribuição de Poisson como

$$p(x) = \begin{cases} \frac{e^{-\mu}\mu^x}{x!}, & x = 0, 1, \ldots \\ 0, & c.c \end{cases}$$

Modelo Poisson

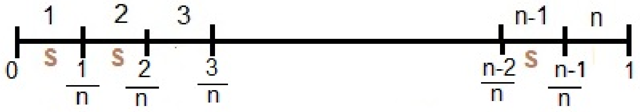
Suposições-modelo binomial

Considere que o intervalo pode ser dividido em subintervalos com comprimento suficientemente pequeno tal

- a probabilidade de mais uma contagem em um subintervalo seja zero,
- a probabilidade de uma contagem em um subintervalo seja a mesma para todos os subintervalos e proporcional ao comprimento de subintervalo e
- a contagem em cada subintervalo seja independente de outros subintervalos.

Modelo Poisson

- Assim,



- Seja X a v.a que conta o sucessos do intervalo de comprimento unitário.
- Pela suposições: $P(S) = \frac{\mu}{n}$.
- A probabilidade de x sucessos nos n sub-intervalos

$$P(X = x) = \binom{n}{x} \left(\frac{\mu}{n}\right)^x \left(1 - \frac{\mu}{n}\right)^{n-x}, \; x = 0, 1, \ldots, n.$$

onde $E(X) = \mu$ é número médio de sucessos no intervalo de comprimento unitário.

Modelo Poisson

Note que

$$
\begin{aligned}
P(X=x) &= \binom{n}{x}\left(\frac{\mu}{n}\right)^{x}\left(1-\frac{\mu}{n}\right)^{n-x} = \frac{n!}{x!(n-x)!}\times\frac{\mu^x}{n^x}\left(1-\frac{\mu}{n}\right)^{n-x} \\
&= \frac{(n-x+1)!}{x!}\times\frac{\mu^x}{n^x}\left(1-\frac{\mu}{n}\right)^{n-x} \\
&= \left(1-\frac{1}{n}\right)\left(1-\frac{1}{n}\right)\times\cdots\times\left(1-\frac{x-1}{n}\right)\frac{\mu^x}{x!}\frac{\left(1-\frac{\mu}{n}\right)^{n}}{\left(1-\frac{\mu}{n}\right)^{x}}
\end{aligned}
$$

Fazendo $n \longrightarrow \infty$, tem-se

$$P[X=x] = \frac{\mu^x e^{-\mu}}{x!}.$$

Modelo Poisson

Uma variável aleatória discreta X tem distribuição Poisson com parâmetro μ se sua função de probabilidade é dada por

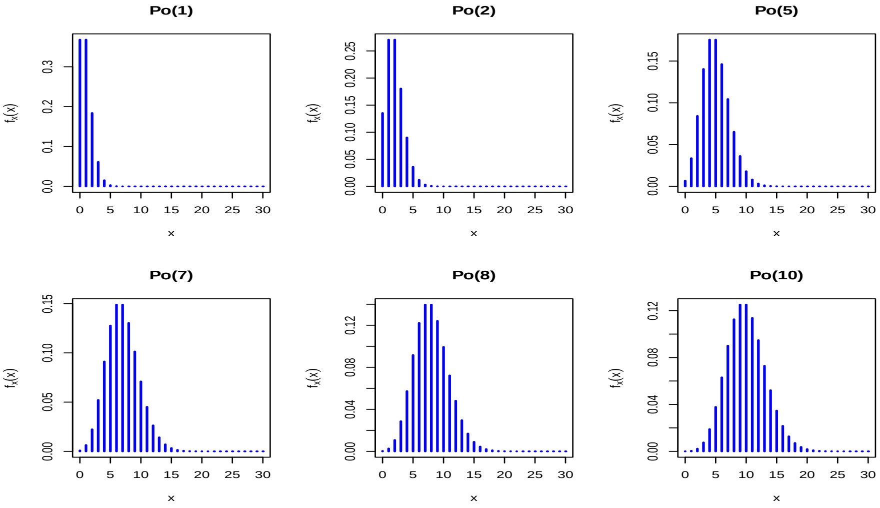
$$f_X(x;\mu) = \begin{cases} \frac{e^{-\mu}\mu^x}{x!}, & x = 0, 1, \ldots \\ 0, & c.c \end{cases}$$

onde X é o número de eventos discretos em t unidades de medida.

- λ : é a media de eventos discretos em uma unidade de medida,
- t : unidade de medida
- $\mu = \lambda t$: é a media de eventos discretos em t unidades de medida

Notação: $X \sim Po(\mu)$.

Modelo Poisson



Modelo Poisson

A função de distribuição acumulada da v.a $X \sim Po(\mu)$ é dado por

$$F_X(x;\mu) = \begin{cases} \sum_{i=0}^{\lfloor x \rfloor} \frac{e^{-\mu}\mu^i}{i!}, & x \geq 0 \\ 0, & c.c \end{cases}$$

Propriedades

Se $X \sim Po(\mu)$, então

(1) $E[X] = Var(X) = \mu$.

(2) a f.g.m. é: $M_X(t) = \exp(\mu(e^t - 1))$.

A função de probabilidade, função de distribuição acumulada, função quantil e a geração de v.a´s com distribuição Poisson no R

```
dpois(x, lambda, log = FALSE)
ppois(q, lambda, lower.tail = TRUE, log.p = FALSE)
qpois(p, lambda, lower.tail = TRUE, log.p = FALSE)
rpois(n, lambda)
```

Modelo Poisson

Exemplo-5

Contaminação é um problema de fabricação de discos ópticos de armazenagem. O número de partículas de contaminação que ocorrem em um disco óptico tem uma distribuição de Poisson e o número médio de partículas por centímetro quadrado da superfície é 0,1. A área do disco em estudo é 100 centímetros quadrados. Encontre a probabilidade de que 12 partículas ocorram na área de um disco sob estudo.

Solução:

- Se X : número de partículas na área de um disco sob estudo, então, $X \sim Po(\mu)$, onde $\mu = \lambda t = 0,1 \times 100 = 10$.
- A f.p é dado por

$$f_X(x; 10) = \frac{e^{-10}10^x}{x!}, \; x = 0, 1, \ldots$$

- $P[X = 12] = \frac{e^{-10}10^x}{x!} = dpois(12, 10) = 0,0947$

Exercicio

1. Para os modelos apresentados, mostre que são realmente uma função probabilidade.
2. Para os modelos introduzidos, mostre as propriedades do texto;
3. Para os modelos Binomial, Binomial negativa e Poisson determine os coeficientes de assimétria e curtose.