Variáveis e vetores aleatórios discretos

Prof. Caio Azevedo

Distribuição binomial

- Dizemos que uma vad (variável aleatória discreta) $Y \sim \text{binomial}(n,p)$, $n \in \{0,1,2,...\}$, $p \in (0,1)$ se sua fdp (função de probabilidade) é dada por:

$$f_Y(y) \equiv f_Y(y,p) \equiv p(y) \equiv p(y,p) \quad = \quad \binom{n}{y} p^y (1-p)^{n-y} \mathbb{1}_{\{0,1,2,\ldots,n\}}(y)$$

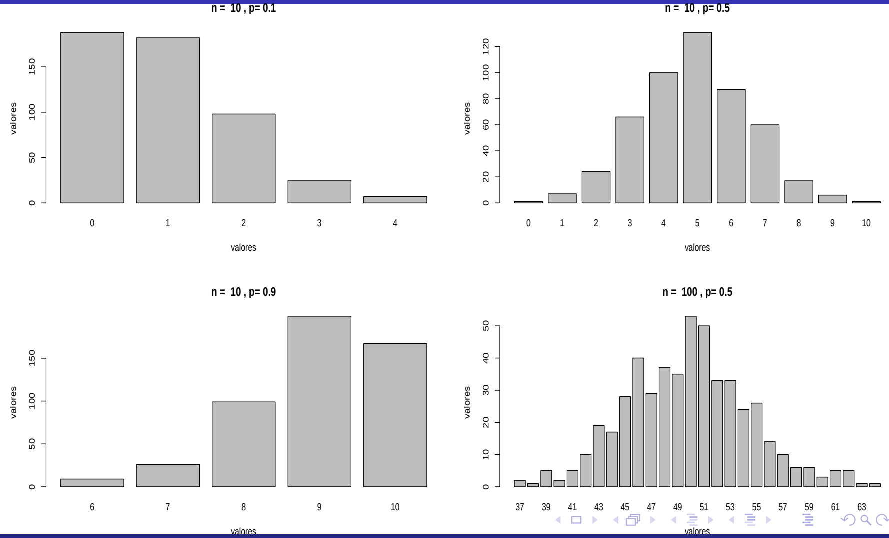
- Temos que sua fgm é dada por:

$$\begin{aligned} M_Y(t) &= \mathcal{E}(e^{tY}) = \sum_{y=0}^{n} e^{ty} \binom{n}{y} p^y (1-p)^{n-y} \\ &= \sum_{y=0}^{n} \binom{n}{y} (e^t p)^y (1-p)^{n-y} = (1 + e^t p - p)^n \end{aligned}$$

Distribuição binomial

- $\frac{\partial M_Y(t)}{\partial t} = np(1 + e^t p - p)^{n-1} e^t$ e $\frac{\partial^2 M_Y(t)}{\partial t^2} =$
$np\left((n-1)(1 + e^t p - p)^{n-2} e^{2t} p + (1 + e^t p - p)^{n-1} e^t\right)$.
- Logo, $\mathcal{E}(Y) = \left.\frac{\partial M_Y(t)}{\partial t}\right|_{t=0} = np$ e
$\mathcal{E}(Y^2) = \left.\frac{\partial^2 M_Y(t)}{\partial t^2}\right|_{t=0} = np(np + 1 - p) = n^2 p^2 + np(1-p)$. Logo
$\mathcal{V}(Y) = np(1-p)$
- Se $n = 1$, então $Y \sim \text{Bernoulli}(p)$.

Histogramas da distribuição binomial



Distribuição de Poisson

- Dizemos que uma vad $Y \sim \text{Poisson}(\lambda)$, $\lambda \in (0, \infty)$ se sua fdp é dada por:

$$f_Y(y) = \frac{e^{-\lambda}\lambda^y}{y!}\mathbb{1}_{\{0,1,2,\ldots,\}}(y)$$

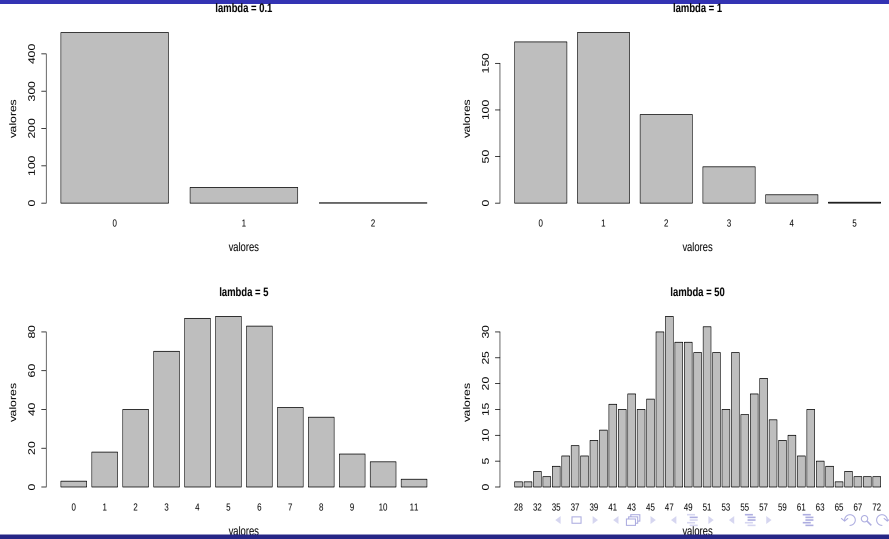
- Temos que sua fgm é dada por:

$$\begin{aligned} M_Y(t) &= \mathcal{E}(e^{tY}) = \sum_{y=0}^{\infty} e^{ty}\frac{e^{-\lambda}\lambda^y}{y!} \\ &= e^{-\lambda}\sum_{y=0}^{\infty}\frac{(e^t\lambda)^y}{y!} = e^{-\lambda}e^{e^t\lambda} = e^{\lambda(e^t-1)} \end{aligned}$$

Distribuição de Poisson

- $\dfrac{\partial M_Y(t)}{\partial t} = \lambda e^{\lambda(e^t - 1) + t}$ e $\dfrac{\partial^2 M_Y(t)}{\partial t^2} = \lambda(\lambda e^t + 1)e^{\lambda(e^t-1)+t}$.
- Logo, $\mathcal{E}(Y) = \left.\dfrac{\partial M_Y(t)}{\partial t}\right|_{t=0} = \lambda$ e $\mathcal{E}(Y^2) = \left.\dfrac{\partial^2 M_Y(t)}{\partial t^2}\right|_{t=0} = \lambda + \lambda^2$. Logo $\mathcal{V}(Y) = \lambda$

Histogramas da distribuição de Poisson



Distribuição geométrica

- Dizemos que uma vad $Y \sim$ geométrica(p), $p \in (0,1)$ se sua fdp é dada por:

$$f_Y(y) = p(1-p)^y \mathbb{1}_{\{0,1,2,\ldots,\}}(y)$$

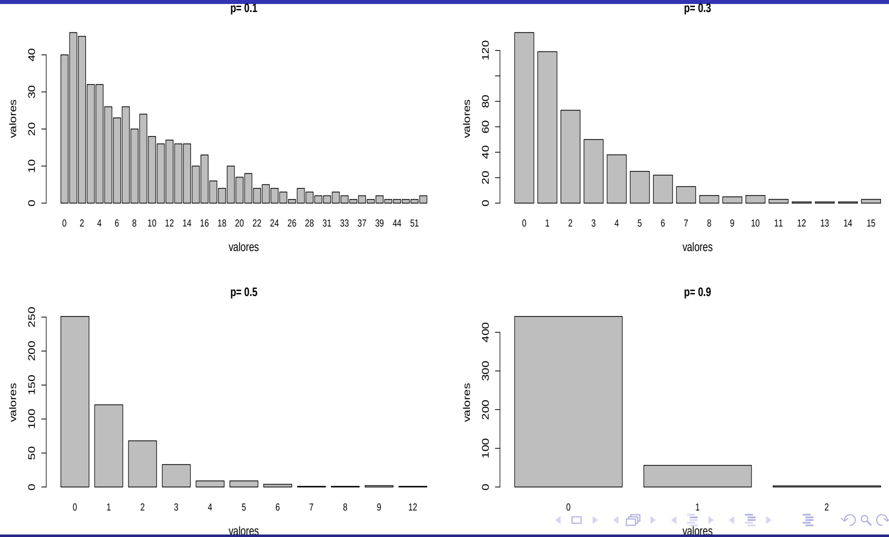
- Temos que sua fgm é dada por:

$$\begin{aligned} M_Y(t) &= \mathcal{E}(e^{tY}) = \sum_{y=0}^{\infty} e^{ty} p(1-p)^y = p \sum_{y=0}^{\infty} (e^t(1-p))^y \\ &= \frac{p}{1-e^t(1-p)} \end{aligned}$$

Distribuição geométrica

- $\dfrac{\partial M_Y(t)}{\partial t} = -\dfrac{(p-1)pe^t}{((p-1)e^t+1)^2}$ e
  $\dfrac{\partial^2 M_Y(t)}{\partial t^2} = \dfrac{(p-1)pe^t((p-1)e^t-1)}{((p-1)e^t+1)^3}$.
- Logo, $\mathcal{E}(Y) = \left.\dfrac{\partial M_Y(t)}{\partial t}\right|_{t=0} = \dfrac{1-p}{p}$ e
  $\mathcal{E}(Y^2) = \left.\dfrac{\partial^2 M_Y(t)}{\partial t^2}\right|_{t=0} = \dfrac{1-p^2}{p^2}$. Logo $\mathcal{V}(Y) = \dfrac{1-p}{p^2}$

Histogramas da distribuição geométrica



Distribuição hipergeométrica

- Dizemos que uma vad (variável aleatória discreta) $Y \sim$ hipergeométrica(N, A, n), $N \in \{1, 2, ...\}$, $A \in \{0, 1, .., N\}$ e $n \in \{0, 1, ..., N\}$ se sua fdp (função de probabilidade) é dada por:

$$f_Y(y) = \frac{\binom{A}{y}\binom{N-A}{n-y}}{\binom{N}{y}} \mathbb{1}_{\{a,\ldots,b\}}(y)$$

em que $a = max\{0, n - N + A\}$ e $b = min\{n, A\}$, embora, na prática, trabalhemos com $a = 0$ e $b = n$.

Distribuição multinomial

- Dizemos que um vetor aleatório

  $\boldsymbol{Y} = (Y_1, ..., Y_k)' \sim \text{multinomial}(n, \boldsymbol{p}^*)$, $\boldsymbol{p}^* = (p_1, ..., p_k)'$,

  $\sum_{i=1}^{k} p_i = 1, p_i \in (0,1), \forall i$, $\sum_{i=1}^{k} y_i = n, y_i \in \{0, 1, ..., n\}$, ou

  $\boldsymbol{Y} = (Y_1, ..., Y_{k-1})' \sim \text{multinomial}(n, \boldsymbol{p})$, $\boldsymbol{p} = (p_1, ..., p_{k-1})'$,

  $\sum_{i=1}^{k} p_i < 1, p_i \in (0,1), \forall i$, $\sum_{i=1}^{k-1} y_i < n$, $y_i \in \{0, 1, ..., n\}, \forall i$, se

  sua fdp é dada por (em que A é o suporte da distribuição,

  $p_k = 1 - \sum_{i=1}^{k-1} p_i$, $y_k = n - \sum_{i=1}^{k-1} y_i$):

Distribuição multinomial

$$
\begin{aligned}
f_{\boldsymbol{Y}}(\boldsymbol{y}) &= \frac{n!}{\prod_{i=1}^{k} y_i!}\left(\prod_{i=1}^{k} p_i\right)\mathbb{1}_A(\boldsymbol{y}) \\
&= \frac{n!}{\prod_{i=1}^{k} y_i!}\left(\prod_{i=1}^{k} p_i\right)\mathbb{1}_{\{0,1,\ldots,n\}}(y_1)\mathbb{1}_{\{0,1,\ldots,n-y_1\}}(y_2)\times\ldots \\
&\times \mathbb{1}_{\{0,1,\ldots,n-\sum_{i=1}^{k-2} y_i\}}(y_{k-1})
\end{aligned}
$$

- $Y_i \sim \text{binomial}(n, p_i), i = 1, \ldots, n.$

Distribuição multinomial

- Função geradora de momentos ($\boldsymbol{t} = (t_1, ..., t_{k-1})'$)

$$
\begin{aligned}
M_{\boldsymbol{Y}}(\boldsymbol{t}) &= \sum_{y_1=0}^{n} \sum_{y_2=0}^{n-y_1} \cdots \sum_{y_{k-1}=0}^{n-\sum_{i=1}^{k-2} y_i} \frac{n!}{\prod_{i=1}^{k-1} y_i!(n - \sum_{i=1}^{k-1} y_i)!} \prod_{i=1}^{k-1} (p_i e^{t_i})^{y_i} p_k^{y_k} \\
&= \left(\sum_{i=1}^{k-1} p_i e^{t_i} + p_k^{x_k} \right)^n
\end{aligned}
$$

- $\dfrac{\partial^2}{\partial t_i \partial t_j} M_{\boldsymbol{Y}}(\boldsymbol{t}) = n(n-1) p_i e^{t_i} \left[p_i e^{t_i} + p_j e^{t_j} + 1 - p_i - p_j \right] p_j e^{t_j}.$

Distribuição multinomial

- Assim $\mathcal{E}(Y_i Y_j) = \frac{\partial^2}{\partial t_i \partial t_j} M_{\boldsymbol{Y}}(\boldsymbol{t})\Big|_{\boldsymbol{t}=\boldsymbol{0}} = n(n-1)p_i p_j$.
- $Cov(Y_i, Y_j) = \mathcal{E}(Y_i Y_j) - \mathcal{E}(Y_i)\mathcal{E}(Y_j) = n(n-1)p_i p_j - n^2 p_i p_j = -np_i p_j$.
- $Corre(Y_i, Y_j) = \frac{Cov(Y_i, Y_j)}{DP(Y_i)DP(Y_j)} = -\frac{p_i p_j}{\sqrt{p_i(1-p_i)p_j(1-p_j)}}$
- Exercício: obtenha os estimadores de MV dos parâmetros da multinomial. Dica: utilize multiplicadores de Lagrange.