

DUAS CARACTERIZAÇÕES DA DISTRIBUIÇÃO GEOMÉTRICA

ATRAVÉS DE ESTATÍSTICAS DE ORDEM

SUELI APARECIDA MINGOTTI

ORIENTADOR

PROF. DR. PUSHPA NARAYAN RATHIE

Tese apresentada no Instituto de
Matemática, Estatística e Ciência da
Computação da Universidade Estadual
de Campinas, como requisito parcial
para a obtenção do Título de Mestre
em Ciências.

CAMPINAS
Estado de São Paulo - Brasil
1982

UNICAMP
BIBLIOTECA CENTRAL

CONTEÚDO

CAPÍTULO 1

INTRODUÇÃO	1
------------------	---

CAPÍTULO 2

NOTAÇÕES, DEFINIÇÕES E PROPRIEDADES BÁSICAS

2.1. DEFINIÇÕES E NOTAÇÕES	7
2.2. PROPRIEDADES BÁSICAS	14

CAPÍTULO 3

CARACTERIZAÇÕES DA DISTRIBUIÇÃO GEOMÉTRICA

3.1. INTRODUÇÃO	25
3.2. CARACTERIZAÇÕES DA DISTRIBUIÇÃO GEOMÉTRICA	26
3.3. CASOS ESPECIAIS	60

BIBLIOGRAFIA	69
--------------------	----

AGRADECIMENTOS

Desejo apresentar meus sinceros agradecimentos a várias pessoas que contribuíram na elaboração desta tese : À Professora Lilian Tong Sheng, orientadora inicial deste trabalho, pelo estímulo, pela orientação segura e objetiva e pela grande dedicação e amizade ; Ao Professor Pushpa Narayan Rathie, que em face da viagem da Professora Lilian ao exterior, gentilmente, me orientou , na fase final deste trabalho , possibilitando sua conclusão ; Ao amigo e Professor Roswílcio José Moreira Góis pelo incentivo e apoio que me prestou e aos meus familiares que comigo viveram todos os momentos, difíceis e alegres, e sempre com grande paciência e carinho.

SUMMARY

In this thesis, we present two new characterizations of geometric distribution by using the order statistics and constant regression property.

For this purpose, we consider a random sample of size n , $n \geq 2$, drawn from the distribution of a discrete random variable with the set of possible values of which as $\{0, 1, 2, \dots\}$ and their corresponding order / statistics as

$$Y_1 \leq Y_2 \leq \dots \leq Y_n$$

We define two statistics

$$Z_k = \frac{1}{n - k} \sum_{i=k+1}^n (Y_i - Y_k) \quad , \quad 1 \leq k \leq n-1$$

and

$$W_k = \frac{1}{k - 1} \sum_{i=1}^{k-1} (Y_k - Y_i) \quad , \quad 2 \leq k \leq n$$

and we characterize the geometric distribution by the following properties :

$$E [Z_k / Y_k = y, Y_m = y, Y_{m+1} > y] = \beta \quad , \quad \text{q.c.}$$

for some arbitrarily fixed k, m , $1 \leq k \leq n-1$, $k \leq m < n$

or

$$E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \beta \quad , q.c.$$

for some arbitrarily fixed $k, m, 2 \leq k \leq n, 2 \leq m \leq k$

where $\beta > 1$ is a constant and y is a non-negative integer.

The main ideas used to demonstrate our results are the conditional expectation and permutations, which are parallel to that used in the paper of Wang and Srivastava (1980) referring a characterization of the exponential and related distributions by linear regression.

The presentation of this work is divided in three chapters.

Chapter 1 contains the presentation of earlier work on characterizing the geometric distribution using order statistics and the principle ideas and technique which motivates this study.

Chapter 2 presents the basic definitions, notations and lemmas which are necessary for development of our characterizations.

Finally, in the chapter 3, we conclude our work, presenting two / principles theorems regarding the characterization of the geometric distribution and some special cases as well as their proofs.

SUMÁRIO

A proposta desta tese é a apresentação de duas novas caracterizações da distribuição geométrica, através de estatísticas de ordem e da propriedade de regressão constante.

Para isso, consideraremos, inicialmente, uma amostra aleatória de tamanho n , $n \geq 2$, de uma variável aleatória discreta, não degenerada, que / assume apenas os valores inteiros, não negativos, bem como suas correspondentes estatísticas de ordem, isto é,

$$Y_1 \leq Y_2 \leq \dots \leq Y_n$$

Definiremos, então, as seguintes estatísticas :

$$Z_k = \frac{1}{n - k} \sum_{i=k+1}^n (Y_i - Y_k) \quad , \quad 1 \leq k \leq n-1$$

$$W_k = \frac{1}{k - 1} \sum_{i=1}^{k-1} (Y_k - Y_i) \quad , \quad 2 \leq k \leq n$$

e baseados nas seguintes hipóteses

$$E \left[Z_k / Y_k = y, Y_m = y, Y_{m+1} > y \right] = \beta \quad , \quad q.c.$$

para algum k e m , fixos e arbitrários, tais que $1 \leq k \leq n-1$, $k \leq m < n$

ou

$$E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \beta, \text{ q.c.}$$

para algum k e m , fixos e arbitrários, tais que $2 \leq k \leq n$, $2 \leq m \leq k$

onde $\beta > 1$ é uma constante e y é um inteiro não negativo,

procuraremos caracterizar a distribuição geométrica.

A idéia que será utilizada para a demonstração desses resultados está fundamentada, basicamente, no emprego da esperança condicional e permutações, sendo paralela àquela desenvolvida por Wang e Srivastava (1980), no trabalho referente à caracterização da distribuição exponencial e algumas / distribuições relacionadas, através de regressão linear.

A apresentação deste nosso trabalho se encontra dividida em 3 capítulos.

O capítulo 1, nos dá uma idéia geral sobre alguns trabalhos executados até o presente momento, referentes à caracterização da distribuição / geométrica pelo uso de estatísticas de ordem. Contém ainda, as idéias principais que motivaram a realização deste nosso estudo, bem como a descrição das técnicas básicas que serão por nós utilizadas.

No capítulo 2, encontraremos algumas definições, notações e resul

tados essenciais ao desenvolvimento de nossas caracterizações.

Finalmente, no capítulo 3, concluiremos o nosso trabalho, apresentando os dois teoremas principais da caracterização da distribuição geométrica e alguns casos especiais decorrentes destes teoremas, bem como suas respectivas demonstrações.

CAPITULO 1

INTRODUÇÃO

Muitos estudos já foram desenvolvidos com referência à caracterização da distribuição geométrica utilizando-se certas propriedades de estatísticas de ordem. Fergunson (1965) provou o seguinte teorema :

TEOREMA 1.1. Sejam X e Y duas variáveis aleatórias independentes, discretas e não degeneradas. Então $U = \text{Min}(X,Y)$ e $V = X-Y$ são independentes se, e somente se, X e Y tem distribuição geométrica com os mesmos parâmetros de escala e locação.

Gupta (1970) estabeleceu o seguinte resultado :

TEOREMA 1.2. Sejam $X_1, X_2, \dots, X_n$ variáveis aleatórias discretas, independentes e identicamente distribuídas e sejam $Y_1 \leq Y_2 \leq \dots \leq Y_n$ as correspondentes estatísticas de ordem. Então a independência de $Y_k - Y_j$ e Y_i para alguma tripla i, j, k , com $1 \leq i \leq j < k \leq n$, caracteriza a distribuição geométrica.

Srivastava (1974) mostrou que se tomarmos uma amostra aleatória de tamanho n , $n \geq 2$, de uma população cuja distribuição é a mesma que a de uma variável aleatória discreta X , não limitada, cujo suporte é o conjunto $\{ \alpha + \beta i, i=1,2,\dots, \beta \neq 0 \}$, os eventos $\{ Y_1=Y_2=\dots=Y_n \}$ e $\{ Y_1 \}$ serão

independentes, se e somente se, a distribuição de X for geométrica.

Motivado por esse resultado de Srivastava, Galambos (1975) conseguiu estendê-lo considerando o caso em que o conjunto dos possíveis valores de X (finito ou infinito) pode ser arranjado numa sequência crescente, isto é, ele não precisa ser necessariamente da forma $\{\alpha + \beta i, i=1,2,\dots, \beta \neq 0\}$.

Recentemente, El-Newehi e Govindarajulu (1979) caracterizaram a distribuição geométrica usando :

i) a independência dos eventos $\{Y_1\}$ e $\{Y_k = Y_1\}$, para algum k , fixo e arbitrário, $2 \leq k \leq n$.

ii) a independência dos eventos $\{Y_1\}$ e $\{Y_2 - Y_1 \in B\}$, onde $B = \{m\}$ ou $B = \{m, m+1, \dots\}$, $m \geq 1$, inteiro positivo, fixo e arbitrário.

O próprio Govindarajulu (1980) generalizou os resultados dados acima e alguns outros obtidos por Ferguson [(1965) e (1967)] .

Wang e Srivastava (1980) também demonstraram dois teoremas muito interessantes, referentes à caracterização da distribuição exponencial, e que nos serviram de base para o desenvolvimento inicial de nosso trabalho. Eles provaram o seguinte :

Sejam $X_1, X_2, \dots, X_n$, $n \geq 2$, uma amostra aleatória de uma população cuja função distribuição é F_X , contínua, e sejam $Y_1 < Y_2 < \dots < Y_n$ as correspondentes estatísticas de ordem. Sejam, também, as seguintes estatísticas :

$$(1.1.) \quad Z_k = \frac{1}{n-k} \sum_{i=k+1}^n (Y_i - Y_k) \quad , \quad 1 \leq k \leq n-1$$

$$(1.2.) \quad W_k = \frac{1}{k-1} \sum_{i=1}^{k-1} (Y_k - Y_i) \quad , \quad 2 \leq k \leq n$$

Então

TEOREMA 1.3. Suponha que F_X seja contínua com primeiro momento finito. Então para algum k , $1 \leq k \leq n-1$

$$(1.3.) \quad E \left[Z_k / Y_k = y \right] = \beta \quad , \quad \text{q.c. } (dF)^*$$

onde $\beta > 0$ é uma constante, se e somente se, F_X tem a forma

$$(1.4.) \quad F_X(x) = 1 - e^{-\frac{1}{\beta}(x-\mu)} \quad , \quad x > \mu > -\infty$$

TEOREMA 1.4. Suponha que F_X seja contínua com primeiro momento finito. Então, para algum k , $2 \leq k \leq n$

$$(1.5.) \quad E \left[W_k / Y_k = y \right] = \beta \quad , \quad \text{q.c. } (dF)$$

onde $\beta > 0$ é uma constante, se e somente se, a distribuição de $-X$ tem a forma

$$(1.4.).$$

Sabendo-se que a distribuição geométrica guarda uma íntima relação com a distribuição exponencial, com propriedades análogas para o caso / discreto, poderia ser questionado se a distribuição geométrica não poderia ser caracterizada através de regressão constante e de suas estatísticas de ordem, assim como ocorre com a distribuição exponencial.

Neste trabalho nós nos propomos a responder, parcialmente, esta pergunta.

A idéia utilizada para a demonstração de nossos resultados é paralela à empregada por Wang e Srivastava (1980).

Inicialmente, nós iremos considerar $X_1, X_2, \dots, X_n$, $n \geq 2$, variáveis aleatórias discretas, não degeneradas, independentes e identicamente distribuídas, e com função distribuição comum F_X . Consideraremos, então, as suas correspondentes estatísticas de ordem $Y_1 \leq Y_2 \leq \dots \leq Y_n$ e as estatísticas dadas como em (1.1.) e (1.2.), e provaremos que

$$(1.6.) \quad E \left[Z_k / Y_k=y, Y_m=y, Y_{m+1} > y \right] = \beta, \text{ q.c.}$$

para algum k e m , fixos e arbitrários, tais que $1 \leq k \leq n-1$, $k \leq m < n$, se e somente se, a distribuição de X é geométrica, como também

$$(1.7.) \quad E \left[W_k / Y_{m-1} < y, Y_m=y, Y_k=y \right] = \beta, \text{ q.c.}$$

para algum k e m , fixos e arbitrários, tais que $2 \leq k \leq n$, $2 \leq m \leq k$, se e somente se, a distribuição de $-X$ é geométrica, sendo que, em ambos os casos

* q.c. (dF) = quase certamente em distribuição.

$\beta > 1$ é uma constante e y é um inteiro não negativo, fixo.

Para a demonstração desses resultados empregamos uma técnica que se fundamenta, principalmente, na decomposição dos conjuntos $\{Y_k=y, Y_m=y, Y_{m+1} > y\}$ e $\{Y_{m-1} < y, Y_m=y, Y_k=y\}$ numa sequência finita de conjuntos disjuntos $A_{\sigma(r)\delta(m-r)}$ e $A_{\sigma(m-1)\delta(s)}$, respectivamente, onde $0 \leq r \leq k-1$, $k-m+1 \leq s \leq n-m+1$, e que definem alguma permutação específica das variáveis aleatórias $X_1 \dots X_n$. Desta forma, em vez de considerarmos $E [Z_k / Y_k=y, Y_m=y, Y_{m+1} > y]$ nós consideraremos $E [Z_k / A_{\sigma(r)\delta(m-r)}]$ para cada $A_{\sigma(r)\delta(m-r)}$, o que acarretará maior facilidade na resolução de nosso problema, uma vez que obteremos a seguinte relação:

$$(1.8.) \quad E [Z_k / Y_k=y, Y_m=y, Y_{m+1} > y] = \frac{n-m}{n-k} E [(X_n-y) / X_n > y] = \beta, \text{ q.c.}$$

Calculando-se, então, o lado direito de (1.8.) nós chegaremos a seguinte equação de diferenças finitas de primeira ordem:

$$(1.9.) \quad F_X(y+1) - \left(\frac{\alpha-1}{\alpha} \right) F_X(y) = \frac{1}{\alpha}$$

$$\text{onde} \quad \alpha = \left(\frac{n-k}{n-m} \right) \beta, \quad \alpha > 1$$

que resolvida, de acordo com algumas condições iniciais e algumas propriedades da função densidade de probabilidade e da função distribuição, nos dará a seguinte solução para $F_X(y)$:

$$(1.10.) \quad F_X(y) = 1 - \left(\frac{\alpha - 1}{\alpha} \right)^{y+1}, \quad y \in \{0, 1, 2, \dots\}$$

ou seja, chegaremos à conclusão que X tem distribuição geométrica com parâmetro $\frac{\alpha - 1}{\alpha}$, $\alpha > 1$.

O resultado referente a W_k , será provado pelo mesmo processo, ou seja, pela decomposição do conjunto $\{Y_{m-1} < y, Y_m = y, Y_k = y\}$ nos conjuntos $A_{\sigma(m-1)} \delta(s)$, nós teremos a relação:

$$(1.11.) \quad E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \frac{m-1}{k-1} E \left[(y - X_1) / X_1 < y \right] = \beta, \text{ q.c.}$$

que também nos fornecerá uma equação de diferenças finitas de primeira ordem, que resolvida determinará a função distribuição F_{-X} .

Em ambos os casos, a prova da condição necessária, será feita por cálculo direto.

Além destes resultados, nós discutiremos no final de nosso trabalho, alguns casos especiais interessantes, referentes às nossas caracterizações da distribuição geométrica.

CAPITULO 2

NOTAÇÕES, DEFINIÇÕES E PROPRIEDADES BÁSICAS

Neste capítulo, apresentaremos algumas definições e notações que serão utilizadas no decorrer deste trabalho e as propriedades básicas que / serão necessárias, sendo que algumas delas serão demonstradas.

2.1. DEFINIÇÕES E NOTAÇÕES

Inicialmente vamos apresentar algumas definições.

DEFINIÇÃO 2.1.1. Seja X uma variável aleatória discreta cujo suporte é o conjunto $\{ 0, 1, 2, \dots \}$. Dizemos que X tem distribuição geométrica com parâmetro $(1 - q)$, $0 < q < 1$, se a função distribuição $F_X(x)$ é da seguinte / forma :

$$(2.1.1.) \quad F_X(x) = P[X \leq x] = 1 - q^{x+1} \quad , \quad x = 0, 1, 2, \dots$$

$$= 0 \quad , \quad \text{caso contrário}$$

DEFINIÇÃO 2.1.2. Sejam $X_1, X_2, \dots, X_n$ variáveis aleatórias independentes e identicamente distribuídas, com função distribuição comum F_{X_1} . As X_i , $i=1, \dots, n$, ordenadas em ordem crescente são chamadas estatísticas de ordem da amostra selecionada e são representadas por

$$(2.1.2.) \quad Y_1 \leq Y_2 \leq \dots \leq Y_n$$

onde Y_k denota a k -ésima estatística de ordem da amostra aleatória de tamanho n , ($n \geq 2$, inteiro) selecionada.

As próximas notações que seguem são relativas à técnica que será por nós utilizada na demonstração das caracterizações da distribuição geométrica.

NOTAÇÃO 2.1.1. Sejam r e s dois números inteiros tais que $0 \leq r \leq n$, $0 \leq s \leq n$.

Consideraremos então :

$\sigma(r)$ um subconjunto de tamanho r do conjunto $\{ 1, 2, \dots, n \}$, isto é

$$\sigma(r) \subset \{ 1, 2, \dots, n \}$$

e

$\delta(s)$ um subconjunto de tamanho s do conjunto $\{ 1, 2, \dots, n \}$, isto é

$$\delta(s) \subset \{ 1, 2, \dots, n \}$$

Denotaremos por $\sigma_0(r)$ o seguinte conjunto

$$\sigma_0(r) = \{ 1, 2, \dots, r \}$$

e por $\delta_r(s)$ o conjunto

$$\delta_r(s) = \{ r+1, r+2, \dots, r+s \}$$

NOTAÇÃO 2.1.2. Para $\sigma(r) \cap \delta(s) = \emptyset$, onde $0 \leq r \leq k-1$, $k-r \leq s \leq n-r$, e $1 \leq k \leq n$ é fixo e arbitrário, e para y inteiro não negativo fixo, denotaremos por $A_{\sigma(r)\delta(s)}$ o conjunto :

$$A_{\sigma(r)\delta(s)} = \left\{ (X_1, \dots, X_n) : \begin{array}{ll} X_j < y & , j \in \sigma(r) \\ X_j = y & , j \in \delta(s) \\ X_j > y & , j \notin \sigma(r) \cup \delta(s) \end{array} \right\}$$

Para que possamos ter uma melhor visualização da formação desses conjuntos, apresentaremos o seguinte exemplo :

EXEMPLO 2.1.1. Sejam X_1, X_2, X_3, X_4 , variáveis aleatórias independentes e identicamente distribuídas ; $n=4$, $k=3$.

Então temos que $0 \leq r \leq 2$ e para cada r tomado teremos a seguinte variação de s :

para $r=0$, $3 \leq s \leq 4$

para $r=1$, $2 \leq s \leq 3$

para $r=2$, $1 \leq s \leq 2$

Desta forma teremos que tomar todos os conjuntos $\sigma(r)$ de tamanhos 0,1 e 2 , e todos os conjuntos $\delta(s)$ de tamanhos 1,2,3 e 4.

Os conjuntos tomados são os seguintes :

$$\sigma(0) = \emptyset$$

$$\sigma_{(1)}(1) = \{1\} \quad \sigma_{(2)}(1) = \{2\} \quad \sigma_{(3)}(1) = \{3\} \quad \sigma_{(4)}(1) = \{4\}$$

$$\sigma_{(1)}(2) = \{1,2\} \quad \sigma_{(2)}(2) = \{1,3\} \quad \sigma_{(3)}(2) = \{1,4\} \quad \sigma_{(4)}(2) = \{2,3\}$$

$$\sigma_{(5)}(2) = \{2,4\} \quad \sigma_{(6)}(2) = \{3,4\}$$

$$\delta_{(1)}(1) = \{1\} \quad \delta_{(2)}(1) = \{2\} \quad \delta_{(3)}(1) = \{3\} \quad \delta_{(4)}(1) = \{4\}$$

$$\delta_{(1)}(2) = \{1,2\} \quad \delta_{(2)}(2) = \{1,3\} \quad \delta_{(3)}(2) = \{1,4\} \quad \delta_{(4)}(2) = \{2,3\}$$

$$\delta_{(5)}(2) = \{2,4\} \quad \delta_{(6)}(2) = \{3,4\}$$

$$\delta_{(1)}(3) = \{1,2,3\} \quad \delta_{(2)}(3) = \{1,2,4\} \quad \delta_{(3)}(3) = \{1,3,4\} \quad \delta_{(4)}(3) = \{2,3,4\}$$

$$\delta(4) = \{1,2,3,4\}$$

Então construiremos os seguintes conjuntos disjuntos

$$A_{\sigma(0)\delta_{(1)}(3)} = \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 = y, x_3 = y, x_4 > y \}$$

$$A_{\sigma(0)\delta_{(2)}(3)} = \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 = y, x_3 > y, x_4 = y \}$$

$$A_{\sigma(0)\delta_{(3)}(3)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 > y, X_3 = y, X_4 = y \}$$

$$A_{\sigma(0)\delta_{(4)}(3)} = \{ (X_1, X_2, X_3, X_4) : X_1 > y, X_2 = y, X_3 = y, X_4 = y \}$$

$$A_{\sigma(0)\delta_{(4)}} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 = y, X_3 = y, X_4 = y \}$$

$$A_{\sigma(1)(1)\delta_{(4)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 < y, X_2 = y, X_3 = y, X_4 > y \}$$

$$A_{\sigma(1)(1)\delta_{(5)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 < y, X_2 = y, X_3 > y, X_4 = y \}$$

$$A_{\sigma(1)(1)\delta_{(6)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 < y, X_2 > y, X_3 = y, X_4 = y \}$$

$$A_{\sigma(2)(1)\delta_{(2)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 < y, X_3 = y, X_4 > y \}$$

$$A_{\sigma(2)(1)\delta_{(3)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 < y, X_3 > y, X_4 = y \}$$

$$A_{\sigma(2)(1)\delta_{(6)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 > y, X_2 < y, X_3 = y, X_4 = y \}$$

$$A_{\sigma(3)(1)\delta_{(1)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 = y, X_3 < y, X_4 > y \}$$

$$A_{\sigma(3)(1)\delta_{(3)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 > y, X_3 < y, X_4 = y \}$$

$$A_{\sigma(3)(1)\delta_{(5)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 > y, X_2 = y, X_3 < y, X_4 = y \}$$

$$A_{\sigma(4)(1)\delta_{(1)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 = y, X_3 > y, X_4 < y \}$$

$$A_{\sigma(4)(1)\delta_{(2)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 = y, X_2 > y, X_3 = y, X_4 < y \}$$

$$A_{\sigma(4)(1)\delta_{(4)}(2)} = \{ (X_1, X_2, X_3, X_4) : X_1 > y, X_2 = y, X_3 = y, X_4 < y \}$$

$$\begin{aligned}
A_{\sigma(1)(1)\delta(4)(3)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 < y, x_3 = y, x_4 = y \} \\
A_{\sigma(2)(1)\delta(3)(3)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 < y, x_3 = y, x_4 = y \} \\
A_{\sigma(3)(1)\delta(2)(3)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 = y, x_3 < y, x_4 = y \} \\
A_{\sigma(4)(1)\delta(1)(3)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 = y, x_3 = y, x_4 < y \} \\
A_{\sigma(1)(2)\delta(3)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 < y, x_3 = y, x_4 > y \} \\
A_{\sigma(1)(2)\delta(4)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 < y, x_3 > y, x_4 = y \} \\
A_{\sigma(2)(2)\delta(2)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 = y, x_3 < y, x_4 > y \} \\
A_{\sigma(2)(2)\delta(4)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 > y, x_3 < y, x_4 = y \} \\
A_{\sigma(3)(2)\delta(2)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 = y, x_3 > y, x_4 < y \} \\
A_{\sigma(3)(2)\delta(3)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 > y, x_3 = y, x_4 < y \} \\
A_{\sigma(4)(2)\delta(1)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 < y, x_3 < y, x_4 > y \} \\
A_{\sigma(4)(2)\delta(4)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 > y, x_2 < y, x_3 < y, x_4 = y \} \\
A_{\sigma(5)(2)\delta(1)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 < y, x_3 > y, x_4 < y \} \\
A_{\sigma(5)(2)\delta(3)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 > y, x_2 < y, x_3 = y, x_4 < y \} \\
A_{\sigma(6)(2)\delta(1)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 > y, x_3 < y, x_4 < y \} \\
A_{\sigma(6)(2)\delta(2)(1)} &= \{ (x_1, x_2, x_3, x_4) : x_1 > y, x_2 = y, x_3 < y, x_4 < y \} \\
A_{\sigma(1)(2)\delta(6)(2)} &= \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 < y, x_3 = y, x_4 = y \}
\end{aligned}$$

$$A_{\sigma(2)}^{(2)} \delta_{(5)}^{(2)} = \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 = y, x_3 < y, x_4 = y \}$$

$$A_{\sigma(3)}^{(2)} \delta_{(4)}^{(2)} = \{ (x_1, x_2, x_3, x_4) : x_1 < y, x_2 = y, x_3 = y, x_4 < y \}$$

$$A_{\sigma(4)}^{(2)} \delta_{(3)}^{(2)} = \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 < y, x_3 < y, x_4 = y \}$$

$$A_{\sigma(5)}^{(2)} \delta_{(2)}^{(2)} = \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 < y, x_3 = y, x_4 < y \}$$

$$A_{\sigma(6)}^{(1)} \delta_{(1)}^{(2)} = \{ (x_1, x_2, x_3, x_4) : x_1 = y, x_2 = y, x_3 < y, x_4 < y \}$$

Neste exemplo, com $n=4$ e $k=3$, é interessante notar que para cada r e s fixos temos

$$\binom{n}{r} \binom{n-r}{s} = \binom{4}{r} \binom{4-r}{s}$$

conjuntos disjuntos $A_{\sigma(r)} \delta_{(s)}$, e que no total temos

$$\sum_{r=0}^1 \sum_{s=3-r}^{4-r} \binom{4}{r} \binom{4-r}{s} = 39$$

conjuntos disjuntos $A_{\sigma(r)} \delta_{(s)}$

2.2. PROPRIEDADES BÁSICAS

Apresentaremos, agora, algumas propriedades da distribuição geométrica e alguns Lemas que utilizaremos no desenvolvimento das demonstrações dos teoremas do capítulo 3.

LEMA 2.2.1. Sejam X_1 , X_2 duas variáveis aleatórias independentes e identicamente distribuídas, com função densidade de probabilidade dada por :

$$(2.2.1.) \quad P[X_1 = k] = p^k (1-p) \quad , \quad k=0,1,2,\dots \quad , \quad 0 < p < 1$$

isto é, X_1 e X_2 tem distribuição geométrica com parâmetro $(1-p)$.

Então, Y_1 e $Y_2 - Y_1$ são independentes.

PROVA

Temos que para n, t inteiros tais que $n > 0$, $t \geq 0$

$$\begin{aligned} (2.2.2.) \quad & P[Y_2 - Y_1 = n, Y_1 = t] \\ &= P[X_1 = t, X_2 = n+t] + P[X_1 = n+t, X_2 = t] \\ &= 2 P[X_1 = t, X_2 = n+t] \\ &= 2 p^t (1-p) p^{n+t} (1-p) \\ &= 2 p^{n+2t} (1-p)^2 \\ &= P[Y_2 - Y_1 = n] P[Y_1 = t] \end{aligned}$$

pois

$$\begin{aligned}
 (2.2.3.) \quad & P [Y_2 - Y_1 = n] \\
 &= \sum_{i=0}^{\infty} P [X_1 = i, X_2 = n+i] + P [X_1 = n+i, X_2 = i] \\
 &= 2 \sum_{i=0}^{\infty} P [X_1 = i, X_2 = n+i] \\
 &= 2 \sum_{i=0}^{\infty} P [X_1 = i] P [X_2 = n+i] \\
 &= 2 (1-p)^2 p^n \sum_{i=0}^{\infty} p^{2i} \\
 &= 2 (1-p)^2 p^n \frac{1}{1-p^2}
 \end{aligned}$$

e

$$\begin{aligned}
 (2.2.4.) \quad & P [Y_1 = t] \\
 &= P [Y_1 \geq t] - P [Y_1 \geq t+1] \\
 &= P [X_1 \geq t, X_2 \geq t] - P [X_1 \geq t+1, X_2 \geq t+1] \\
 &= p^{2t} - p^{2t+2} \\
 &= p^{2t} (1 - p^2)
 \end{aligned}$$

Considerando-se , agora, o caso em que $n=0$ e $t \geq 0$, teremos

$$\begin{aligned}
 (2.2.5.) \quad & P [Y_2 - Y_1 = 0, Y_1 = t] \\
 &= P [Y_2 = Y_1 = t] \\
 &= P [X_1 = X_2 = t] \\
 &= p^{2t} (1-p)^2 \\
 &= P [Y_2 - Y_1 = 0] \cdot P [Y_1 = t]
 \end{aligned}$$

pois

$$\begin{aligned}
 (2.2.6.) \quad & P [Y_2 - Y_1 = 0] \\
 &= \sum_{i=0}^{\infty} P [X_1 = X_2 = i] \\
 &= (1-p)^2 \sum_{i=0}^{\infty} p^{2i} \\
 &= (1-p)^2 \frac{1}{1-p^2}
 \end{aligned}$$

e

$$(2.2.7.) \quad P [Y_1 = t] = p^{2t} (1-p^2)$$

Portanto, de (2.2.2.) e (2.2.5.) podemos concluir , imediatamente, que $Y_2 - Y_1$ e Y_1 são independentes .

LEMA 2.2.2. Se $G(i) = P[X \geq i] = q^{i-1}$, $i=1,2,\dots$, $0 < q < 1$, isto é X tem distribuição geométrica com parâmetro $(1-q)$, então

- i) Os eventos $\{Y_1\}$ e $\{Y_{k+1} - Y_k \geq m\}$ são mutuamente independentes para todo k ($1 \leq k \leq n-1$) e todo $m \geq 1$.
- ii) Os eventos $\{Y_1\}$ e $\{Y_{k+1} - Y_k = m\}$ são mutuamente independentes para todo k , ($1 \leq k \leq n-1$) e todo $m \geq 0$.

Nós não demonstraremos aqui, este resultado , mas a sua prova poderá ser encontrada no artigo de Govindarajulu (1980).

Motivados pelos resultados apresentados nos dois lemas anteriores e pelo trabalho desenvolvido por Wang e Srivastava (1980) com relação à caracterização da distribuição exponencial através da propriedade de regressão constante de algumas estatísticas, que são funções de diferenças de estatísticas de ordem, nós definimos

$$(2.2.8.) \quad d_i = Y_{i+1} - Y_i \quad , \quad 0 \leq i \leq n-1$$

com $Y_0 \equiv 0$

e procuramos estudar se a distribuição geométrica poderia, também, ser caracte

rizada através da propriedade de regressão constante de estatísticas que fossem funções das diferenças d_i .

O nosso trabalho responderá, parcialmente, esta questão.

LEMA 2.2.3. Para k, m fixos e tais que $1 \leq k \leq n-1$, $k \leq m < n$, os conjuntos $A_{\sigma(r)\delta(m-r)}$, $0 \leq r \leq k-1$, onde $\sigma(r) \cap \delta(m-r) = \emptyset$, são disjuntos e para y inteiro não negativo, fixo, tem-se que :

$$(2.2.9.) \quad \{ Y_k = y, Y_m = y, Y_{m+1} > y \} = \bigcup_{r \in \sigma, \delta} A_{\sigma(r)\delta(m-r)}$$

Para uma melhor visualização deste resultado, basta olharmos o exemplo 2.1.1., apresentado anteriormente.

Podemos verificar também que para cada r fixo, existem $\binom{n}{r} \binom{n-r}{m-r}$ conjuntos disjuntos $A_{\sigma(r)\delta(m-r)}$ que são equiprováveis, pois as variáveis X_j , $j=1, \dots, n$ são independentes e identicamente distribuídas, e podemos ver que

$$(2.2.10.) \quad P [A_{\sigma(r)\delta(m-r)}] = P [A_{\sigma_0(r)\delta_r(m-r)}]$$

Vamos denotar

$$(2.2.11.) \quad P_{r,m-r} = P [A_{\sigma_0(r)\delta_r(m-r)}]$$

Desta forma , poderemos estabelecer o seguinte Lema :

LEMA 2.2.4.

$$P \left[A_{\sigma(r)\delta(m-r)} / Y_k = y, Y_m = y, Y_{m+1} > y \right] = \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}$$

onde $1 \leq k \leq n-1$, $k \leq m < n$ e $p_{i,m-i} = P[A_{\sigma_0(i)\delta_i(m-i)}]$

PROVA

Pelo Lema 2.2.3. temos que

$$\{ Y_k = y, Y_m = y, Y_{m+1} > y \} = \bigcup_r \bigcup_{\sigma, \delta} A_{\sigma(r)\delta(m-r)}$$

Então

$$\begin{aligned} (2.2.12.) \quad & P \{ Y_k = y, Y_m = y, Y_{m+1} > y \} \\ &= P \left\{ \bigcup_r \bigcup_{\sigma, \delta} A_{\sigma(r)\delta(m-r)} \right\} \end{aligned}$$

$$\begin{aligned}
&= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} P[A_{\sigma(r) \delta(m-r)}] \\
&= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} P[A_{\sigma_0(r) \delta_r(m-r)}] \\
&= \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} p_{r, m-r} \\
&= \sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i, m-i}
\end{aligned}$$

Desta forma teremos que :

$$\begin{aligned}
&P [A_{\sigma(r) \delta(m-r)} / Y_k = Y, Y_m = Y, Y_{m+1} > Y] \\
&= \frac{P [A_{\sigma(r) \delta(m-r)} \cap \{ Y_k = Y, Y_m = Y, Y_{m+1} > Y \}]}{P [Y_k = Y, Y_m = Y, Y_{m+1} > Y]} \\
&= \frac{P [A_{\sigma(r) \delta(m-r)}]}{P \{ Y_k = Y, Y_m = Y, Y_{m+1} > Y \}}
\end{aligned}$$

$$\begin{aligned}
&= \frac{P \left[A_{\sigma_0(r) \delta_r(m-r)} \right]}{P \{ Y_k = y, Y_m = y, Y_{m+1} > y \}} \\
&= \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}
\end{aligned}$$

LEMA 2.2.5. Para k, m fixos e tais que $2 \leq k \leq n, 2 \leq m \leq k$, os conjuntos $A_{\sigma(m-1) \delta(s)}$, $k-m+1 \leq s \leq n-m+1$, onde $\sigma(m-1) \cap \delta(s) = \emptyset$, são disjuntos, e para y inteiro não negativo, fixo, tem-se que :

$$(2.2.13.) \quad \{ Y_{m-1} < y, Y_m = y, Y_k = y \} = \bigcup_s \bigcup_{\sigma, \delta} A_{\sigma(m-1) \delta(s)}$$

Este resultado pode ser visualizado no exemplo 2.1.1., apresentado anteriormente.

Podemos verificar também que para cada s fixo, existem $\binom{n}{m-1} \binom{n-m+1}{s}$ conjuntos disjuntos $A_{\sigma(m-1) \delta(s)}$ que são equiprováveis, pois as variáveis aleatórias $X_j, j=1, \dots, n$, são independentes e identicamente distribuídas e podemos ver também que

$$(2.2.14.) \quad P \left[A_{\sigma(m-1)\delta(s)} \right] = P \left[A_{\sigma_0(m-1)\delta_{m-1}(s)} \right]$$

Denotaremos

$$(2.2.15.) \quad p_{m-1,s} = P \left[A_{\sigma_0(m-1)\delta_{m-1}(s)} \right]$$

Desta forma, poderemos estabelecer o seguinte Lema :

LEMA 2.2.6.

$$P \left[A_{\sigma(m-1)\delta(s)} / Y_{m-1} < Y, Y_m = Y, Y_k = Y \right] = \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$\text{onde } 2 \leq k \leq n, 2 \leq m \leq k \quad \text{e} \quad p_{m-1,i} = P \left[A_{\sigma_0(m-1)\delta_{m-1}(i)} \right]$$

PROVA

Pelo Lema 2.2.5. temos que

$$\{ Y_{m-1} < Y, Y_m = Y, Y_k = Y \} = \bigcup_s \bigcup_{\sigma, \delta} A_{\sigma(m-1)\delta(s)}$$

Então

$$\begin{aligned}
 (2.2.16.) \quad & P \{ Y_{m-1} < y, Y_m = y, Y_k = y \} \\
 &= P \left\{ \bigcup_s \bigcup_{\sigma, \delta} A_{\sigma(m-1)\delta}(s) \right\} \\
 &= \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} P [A_{\sigma(m-1)\delta}(s)] \\
 &= \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} P [A_{\sigma_0(m-1)\delta_{m-1}}(s)] \\
 &= \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} p_{m-1,s} \\
 &= \sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i}
 \end{aligned}$$

Desta forma teremos que :

$$\begin{aligned}
 & P [A_{\sigma(m-1)\delta}(s) / Y_{m-1} < y, Y_m = y, Y_k = y] \\
 &= \frac{P [A_{\sigma(m-1)\delta}(s) \cap \{ Y_{m-1} < y, Y_m = y, Y_k = y \}]}{P \{ Y_{m-1} < y, Y_m = y, Y_k = y \}}
 \end{aligned}$$

$$= \frac{P \left[A_{\sigma(m-1)} \delta(s) \right]}{P \left\{ Y_{m-1} < Y, Y_m = Y, Y_k = Y \right\}}$$

$$= \frac{P \left[A_{\sigma_0(m-1)} \delta_{m-1}(s) \right]}{P \left\{ Y_{m-1} < Y, Y_m = Y, Y_k = Y \right\}}$$

$$= \frac{P_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} P_{m-1,i} \right)}$$

CAPÍTULO 3

CARACTERIZAÇÕES DA DISTRIBUIÇÃO GEOMÉTRICA

A proposta deste capítulo é a apresentação dos dois teoremas fundamentais que permitem as caracterizações da distribuição geométrica e de alguns casos especiais decorrentes destes teoremas. A idéia utilizada para a demonstração destes teoremas é paralela à utilizada por Wang e Srivastava (1980), sobre uma caracterização da distribuição exponencial e algumas distribuições relacionadas através de regressão linear, e baseia-se, principalmente, no uso de esperança condicional e permutações.

3.1. INTRODUÇÃO

Sejam $X_1, X_2, \dots, X_n$ uma amostra aleatória de tamanho n , $n \geq 2$, de uma variável aleatória cuja função distribuição é $F_X(x)$, onde X é uma variável aleatória discreta, não degenerada, cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e sejam

$$Y_1 \leq Y_2 \leq \dots \leq Y_n$$

as correspondentes estatísticas de ordem.

Consideremos a função distribuição :

$$\begin{aligned} (3.1.1.) \quad F_X(x) &= 1 - q^{x+1}, \quad x = 0, 1, 2, \dots \\ &= 0, \quad \text{caso contrário} \end{aligned}$$

onde $0 < q < 1$ e $(1 - q)$ é o parâmetro da distribuição de X .

Definamos , então, as seguintes estatísticas :

$$(3.1.2.) \quad Z_k = \frac{1}{n-k} \sum_{i=k+1}^n (Y_i - Y_k) \quad , \quad 1 \leq k \leq n-1$$

$$(3.1.3.) \quad W_k = \frac{1}{k-1} \sum_{i=1}^{k-1} (Y_k - Y_i) \quad , \quad 2 \leq k \leq n$$

Um dos resultados apresentados por Wang e Srivastava (1980) refere-se à caracterização da distribuição exponencial que é feita baseada nas estatísticas (3.1.2.) e (3.1.3.) para o caso em que F_X é contínua e no fato dessas estatísticas terem regressão constante sobre um conjunto apropriado.

Tendo-se em vista, a grande semelhança existente entre as distribuições exponencial e geométrica, motivados pelo trabalho executado por Wang e Srivastava (1980) resolvemos investigar com mais profundidade se a distribuição geométrica não poderia, também, ser caracterizada através de estatísticas convenientes que possuíssem a propriedade de regressão constante sobre um conjunto adequadamente escolhido.

3.2. CARACTERIZAÇÕES DA DISTRIBUIÇÃO GEOMÉTRICA

Nesta seção apresentaremos os dois teoremas fundamentais das caracterizações da distribuição geométrica e as suas respectivas demonstrações.

TEOREMA 3.2.1. Seja $F_X(x)$ a função distribuição de uma variável aleatória X , discreta, não degenerada, cujo suporte é o conjunto $\{ 0,1,2,\dots \}$. Suponha que $E(X)$ existe. Então, para algum k e m , onde k e m são inteiros, fixos

e arbitrários tais que , $1 \leq k \leq n-1$, $k \leq m \leq n$, $n \geq 2$

$$(3.2.1.) \quad E \left[Z_k / Y_k = y, Y_m = y, Y_{m+1} > y \right] = \beta \quad , \text{ q.c.}$$

onde y é um inteiro não negativo, fixo e $\beta > 1$ é uma constante, se e somente se, $F_X(x)$ é dada por (3.1.1.) com o parâmetro q dado por

$$q = \frac{\alpha - 1}{\alpha} \quad , \text{ onde } \alpha = \left(\frac{n - k}{n - m} \right) \beta \quad , \quad 0 < q < 1$$

TEOREMA 3.2.2. Seja $F_X(x)$ a função distribuição de uma variável aleatória X , discreta, não degenerada, que assume os valores x , tais que $-x \in \{0, 1, 2, \dots\}$. Suponha que $E(X)$ existe. Então, para algum k e m , onde k e m são inteiros fixos e arbitrários tais que , $2 \leq k \leq n$, $2 \leq m \leq k$,

$$(3.2.2.) \quad E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \beta \quad , \text{ q.c.}$$

onde y é um inteiro não negativo, fixo e $\beta > 1$ é uma constante, se e somente se, $F_{-X}(x)$ é dada por (3.1.1.) , isto é,

$$F_{-X}(x) = 1 - q^{x+1} \quad , \quad x \in \{0, 1, 2, \dots\}$$

$$\text{com } q = \frac{\alpha - 1}{\alpha} \quad , \text{ onde } \alpha = \left(\frac{k - 1}{m - 1} \right) \beta \quad , \quad 0 < q < 1$$

Antes de procedermos as demonstrações dos teoremas enunciados vamos estabelecer e provar alguns lemas básicos necessários ao desenvolvimento / destas demonstrações.

LEMA 3.2.1. Para Y_k, Y_m, Y_{m+1} e $A_{\sigma(r)\delta(m-r)}$ como definidos anteriormente,

$$\begin{aligned}
 (3.2.3.) \quad & E \left[\sum_{i=k+1}^n (Y_i - Y_k) / Y_k = y, Y_m = y, Y_{m+1} > y \right] \\
 &= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} E \left[\sum_{j \notin \sigma(r) \cup \delta(m-r)} (X_j - y) / A_{\sigma(r)\delta(m-r)} \right] \times \\
 &\quad \times P \left[A_{\sigma(r)\delta(m-r)} / Y_k = y, Y_m = y, Y_{m+1} > y \right]
 \end{aligned}$$

PROVA

Desde que, pelo Lema 2.2.3.

$$\{ Y_k = y, Y_m = y, Y_{m+1} > y \} = \bigcup_r \bigcup_{\sigma, \delta} A_{\sigma(r)\delta(m-r)}$$

onde $A_{\sigma(r)\delta(m-r)}$ são disjuntos, para todo r , $0 \leq r \leq k-1$,

teremos que sobre $A_{\sigma(r)\delta(m-r)}$ vale a seguinte relação

$$(3.2.4.) \quad \sum_{i=k+1}^n (Y_i - Y_k) = \sum_{j \notin \sigma(r) \cup \delta(m-r)} (X_j - y)$$

Desta forma podemos calcular a

$$E \left[\sum_{i=k+1}^n (Y_i - Y_k) / Y_k = Y, Y_m = Y, Y_{m+1} > Y \right]$$

pelo seguinte caminho

$$\begin{aligned} & E \left[\sum_{i=k+1}^n (Y_i - Y_k) / Y_k = Y, Y_m = Y, Y_{m+1} > Y \right] \\ &= \sum_{\{Y_k = Y, Y_m = Y, Y_{m+1} > Y\}} \sum_{i=k+1}^n (Y_i - Y) \frac{P[Y_i = Y_i, k+1 \leq i \leq n]}{P[Y_k = Y, Y_m = Y, Y_{m+1} > Y]} \\ &= \sum_{\substack{r \in \sigma, \delta \\ \cup \sigma(r) \delta(m-r)}} \sum_{i=k+1}^n (Y_i - Y) \frac{P[Y_i = Y_i, k+1 \leq i \leq n]}{P[Y_k = Y, Y_m = Y, Y_{m+1} > Y]} \\ &= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} \sum_{i=k+1}^n (Y_i - Y) \frac{P\{ [Y_i = Y_i, k+1 \leq i \leq n] \cap A_{\sigma(r) \delta(m-r)} \}}{P[Y_k = Y, Y_m = Y, Y_{m+1} > Y]} \\ &= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} \sum_{j \notin \sigma(r) \cup \delta(m-r)} (x_j - Y) \frac{P[X_j = x_j, j \notin \sigma(r) \cup \delta(m-r)]}{P[Y_k = Y, Y_m = Y, Y_{m+1} > Y]} \times \\ &\quad \times \frac{P[A_{\sigma(r) \delta(m-r)}]}{P[A_{\sigma(r) \delta(m-r)}]} \end{aligned}$$

Como $A_{\sigma(r) \delta(m-r)}$ é um subconjunto do conjunto $\{Y_k = Y, Y_m = Y, Y_{m+1} > Y\}$ temos que

$$\begin{aligned}
& E \left[\sum_{i=k+1}^n (Y_i - Y_k) / Y_k = y, Y_m = y, Y_{m+1} > y \right] \\
&= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} \left[\sum_{j \notin \sigma(r) \cup \delta(m-r)} (X_j - y) \frac{P\{[X_j = x_j, j \notin \sigma(r) \cup \delta(m-r)] \cap A_{\sigma(r) \delta(m-r)}\}}{P[A_{\sigma(r) \delta(m-r)}]} \right] \\
&\times \frac{P\{A_{\sigma(r) \delta(m-r)} \cap [Y_k = y, Y_m = y, Y_{m+1} > y]\}}{P[Y_k = y, Y_m = y, Y_{m+1} > y]} \\
&= \sum_{r=0}^{k-1} \sum_{\sigma, \delta} E \left[\sum_{j \notin \sigma(r) \cup \delta(m-r)} (X_j - y) / A_{\sigma(r) \delta(m-r)} \right] \times \\
&\times P[A_{\sigma(r) \delta(m-r)} / Y_k = y, Y_m = y, Y_{m+1} > y]
\end{aligned}$$

Portanto, o Lema 3.2.1. está provado.

Como as variáveis aleatórias X_j , $j=1, \dots, n$ são independentes e identicamente distribuídas, podemos observar que as seguintes relações são válidas :

$$\begin{aligned}
\text{i)} \quad & E \left[\sum_{j \notin \sigma_0(r) \cup \delta_r(m-r)} (X_j - y) / A_{\sigma_0(r) \delta_r(m-r)} \right] = E \left[\sum_{j \notin \sigma_0(r)} (X_j - y) / A_{\sigma_0(r) \delta_r(m-r)} \right] \\
\text{ii)} \quad & E \left[(X_j - y) / A_{\sigma_0(r) \delta_r(m-r)} \right] = E \left[(X_n - y) / A_{\sigma_0(r) \delta_r(m-r)} \right]
\end{aligned}$$

para $j > m$, $k \leq m < n$, onde

$$\sigma_0(r) = \{1, 2, \dots, r\} \quad \text{e} \quad \delta_r(m-r) = \{r+1, r+2, \dots, k, \dots, m\}$$

O resultado em (i) decorre, imediatamente, do fato de serem as variáveis aleatórias X_j , $j=1, \dots, n$, identicamente distribuídas.

O resultado em (ii) pode ser provado da seguinte forma :

$$\begin{aligned} & E \left[(X_j - y) / A_{\sigma_0(r)\delta_r(m-r)} \right] \\ &= E \left[(X_j - y) / X_1 < y, \dots, X_r < y, X_{r+1} = \dots = X_k = \dots = X_m = y, X_{m+1} > y, \dots, X_j > y, \dots, X_n > y \right] \\ &= E \left[(X_n - y) / X_1 < y, \dots, X_r < y, X_{r+1} = \dots = X_k = \dots = X_m = y, X_{m+1} > y, \dots, X_j > y, \dots, X_n > y \right] \\ &= E \left[(X_n - y) / A_{\sigma_0(r)\delta_r(m-r)} \right] \end{aligned}$$

LEMA 3.2.2. Para Z_k , Y_k , Y_m e Y_{m+1} , como anteriormente definidos, temos :

$$(3.2.5.) \quad E \left[Z_k / Y_k = y, Y_m = y, Y_{m+1} > y \right] = \frac{n-m}{n-k} E \left[(X_n - y) / X_n > y \right]$$

onde $1 \leq k \leq n-1$ e $k \leq m < n$

PROVA

Para provar este Lema, nós usaremos o Lema 3.2.1., bem como o resultado mencionado no Lema 2.2.4., isto é,

$$(3.2.6.) \quad P[A_{\sigma(r)\delta(m-r)} / Y_k=Y, Y_m=Y, Y_{m+1}>Y] = \frac{P_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} P_{i,m-i} \right)}$$

onde

$$P_{r,m-r} = P[A_{\sigma_0(r)\delta_r(m-r)}] \quad \text{e} \quad P_{i,m-i} = P[A_{\sigma_0(i)\delta_i(m-i)}]$$

Temos que :

$$\begin{aligned} & E [Z_k / Y_k=Y, Y_m=Y, Y_{m+1}>Y] \\ &= \frac{1}{(n-k)} E \left[\sum_{i=k+1}^n (Y_i - Y_k) / Y_k=Y, Y_m=Y, Y_{m+1}>Y \right] \\ &= \frac{1}{(n-k)} \sum_{r=0}^{k-1} \sum_{\sigma, \delta} E \left[\sum_{j \notin \sigma(r) \cup \delta(m-r)} (X_j - Y) / A_{\sigma(r)\delta(m-r)} \right] \times \\ & \quad \times P [A_{\sigma(r)\delta(m-r)} / Y_k=Y, Y_m=Y, Y_{m+1}>Y] \\ &= \frac{1}{(n-k)} \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} E \left[\sum_{j \notin \sigma_0(r)} (X_j - Y) / A_{\sigma_0(r)\delta_r(m-r)} \right] \frac{P_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} P_{i,m-i} \right)} \end{aligned}$$

$$= \frac{1}{(n-k)} \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} E \left[\sum_{j=m+1}^n (X_j - Y) / A_{\sigma_0(r)} \delta_r(m-r) \right] \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}$$

$$= \frac{1}{(n-k)} \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} (n-m) E \left[(X_n - Y) / A_{\sigma_0(r)} \delta_r(m-r) \right] \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}$$

$$= \frac{n-m}{n-k} \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} E \left[(X_n - Y) / X_1 < Y, \dots, X_r < Y, X_{r+1} = \dots = X_k = \dots = X_m = Y, X_{m+1} > Y, \dots \right]$$

$$\dots, X_n > Y \Big] \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}$$

$$= \frac{n-m}{n-k} \sum_{r=0}^{k-1} \binom{n}{r} \binom{n-r}{m-r} E \left[(X_n - Y) / X_n > Y \right] \frac{p_{r,m-r}}{\left(\sum_{i=0}^{k-1} \binom{n}{i} \binom{n-i}{m-i} p_{i,m-i} \right)}$$

$$= \frac{n-m}{n-k} E \left[(X_n - Y) / X_n > Y \right]$$

Deve ser observado que existem $\binom{n}{r} \binom{n-r}{m-r}$ termos sob o somatório $\sum_{\sigma, \delta}$, $(n-m)$ termos sob o somatório $\sum_{j \notin \sigma_0(r)}$, que $E[(X_n - y) / A_{\sigma_0(r)} \delta_r(m-r)]$ é constante para todo r , $0 \leq r \leq k-1$, e ainda que a última igualdade vem do fato da variável aleatória X_n ser independente de todas as demais variáveis $X_1, \dots, X_{n-1}$.

Desta forma, a prova do lema 3.2.2. está completa.

LEMA 3.2.3. Se F_X é a função distribuição de uma variável aleatória X , não degenerada, cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e $\alpha > 1$ é uma constante e

$$E[(X_n - y) / X_n > y] = \alpha, \text{ q.c.}$$

Então

$$(3.2.7.) \quad F_X(y+1) - \left(\frac{\alpha - 1}{\alpha} \right) F_X(y) = \frac{1}{\alpha}$$

PROVA

Sem perda de significância, vamos supor que $P(X_n > y) > 0$, isto é,

$F_X(y) \neq 1$, pois caso contrário nada existe para ser discutido.

Então

$$\begin{aligned} E[(X_n - y) / X_n > y] &= E[(X - y) / X > y] \\ &= \sum_{x=y+1}^{\infty} (x-y) \frac{P[X=x]}{P[X > y]} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{P[X > y]} \left\{ \sum_{x=y+1}^{\infty} x P[X=x] - y \sum_{x=y+1}^{\infty} P[X=x] \right\} \\
&= \frac{1}{P[X > y]} \left\{ \sum_{x=0}^{\infty} x P[X=x] - \sum_{x=0}^y x P[X=x] - y P[X > y] \right\} \\
&= \frac{1}{P[X > y]} \left\{ E(X) - \sum_{x=0}^y x P[X=x] \right\} - y
\end{aligned}$$

pois temos por hipótese que $E(X)$ existe .

Como $P[X > y] = 1 - F_X(y)$, temos que

$$\begin{aligned}
(3.2.8.) \quad E[(X_n - y) / X_n > y] \\
= \frac{1}{[1 - F_X(y)]} \left\{ E(X) - \sum_{x=0}^y x P[X=x] \right\} - y
\end{aligned}$$

e utilizando-se a hipótese dada, isto é,

$$E[(X_n - y) / X_n > y] = \alpha$$

teremos

$$\alpha = \frac{1}{[1 - F_X(y)]} \left\{ E(X) - \sum_{x=0}^y x P[X=x] \right\} - y$$

ou, equivalentemente

$$(3.2.9.) \quad (\alpha+y) [1 - F_X(y)] = E(X) - \sum_{x=0}^y x P[X=x]$$

Como ambos os lados de (3.2.9.) são finitos, podemos fazer uso do cálculo de diferenças finitas e calcular a diferença de primeira ordem (Δ) de ambos os lados de (3.2.9.), sendo que a diferença de primeira ordem de uma função $G(x)$, qualquer, onde x assume apenas valores inteiros consecutivos, é definida por :

$$(3.2.10.) \quad \Delta G(x) = G(x+1) - G(x)$$

Desta forma obteremos :

$$(3.2.11.) \quad \Delta \{ (\alpha+y) [1 - F_X(y)] \} = \Delta \left\{ E(X) - \sum_{x=0}^y x P[X=x] \right\}$$

e isto implica que,

$$\Delta \{ (\alpha+y) [1 - F_X(y)] \} = \Delta E(X) - \Delta \sum_{x=0}^y x P[X=x]$$

ou, equivalentemente

$$\Delta \{ (\alpha+y) [1 - F_X(y)] \} = \Delta E(X) - \Delta \sum_{x=0}^y x [F_X(x) - F_X(x-1)]$$

e por definição da diferença finita Δ , obtemos

$$(\alpha+y) [1-F_X(y+1) - 1+F_X(y)] + [1-F_X(y+1)] = -(y+1) [F_X(y+1) - F_X(y)]$$

que é equivalente a

$$[(\alpha+y) - (y+1)] [F_X(y) - F_X(y+1)] + [1 - F_X(y+1)] = 0$$

Portanto, temos que

$$(\alpha-1) F_X(y) - \alpha F_X(y+1) + 1 = 0$$

ou ainda,

$$F_X(y+1) - \left(\frac{\alpha-1}{\alpha} \right) F_X(y) = \frac{1}{\alpha}$$

o que completa a prova do Lema 3.2.3.

LEMA 3.2.4. A equação de diferenças finitas de primeira ordem dada por

$$(3.2.12.) \quad F_X(y+1) - \left(\frac{\alpha-1}{\alpha} \right) F_X(y) = \frac{1}{\alpha}$$

onde F_X é uma função de distribuição, tem a seguinte solução para $F_X(y)$:

$$(3.2.13.) \quad F_X(y) = 1 - \left(\frac{\alpha - 1}{\alpha} \right)^{y+1}, \quad y \in \{0, 1, 2, \dots\}$$

onde $\alpha > 1$ é uma constante.

PROVA

Temos que a equação dada em (3.2.12.) é uma equação de diferenças / não homogênea, e portanto para resolvê-la temos, primeiramente, que encontrar a solução da equação homogênea associada a (3.2.12.), ou seja

$$(3.2.14.) \quad \bar{F}_X(y+1) - \left(\frac{\alpha - 1}{\alpha} \right) \bar{F}_X(y) = 0$$

Como esta equação de diferenças é linear com coeficientes constantes, temos que sua solução é da forma :

$$(3.2.15.) \quad \bar{F}_X(y) = \lambda^y$$

Então, de (3.2.14.) e (3.2.15.), nós obteremos

$$(3.2.16.) \quad \lambda^{y+1} - \left(\frac{\alpha - 1}{\alpha} \right) \lambda^y = 0$$

ou ainda

$$(3.2.17.) \quad \lambda^y \left[\lambda - \left(\frac{\alpha - 1}{\alpha} \right) \right] = 0$$

ou, finalmente

$$(3.2.18.) \quad \lambda - \left(\frac{\alpha - 1}{\alpha} \right) = 0$$

pois $\lambda^y \neq 0$, para todo y

Desta forma, temos que

$$(3.2.19.) \quad \lambda = \left(\frac{\alpha - 1}{\alpha} \right)$$

e portanto a solução geral da equação (3.2.14.) será dada por

$$(3.2.20.) \quad \bar{F}_X(y) = C \left(\frac{\alpha - 1}{\alpha} \right)^y$$

onde C é uma constante .

Agora, vamos buscar uma solução particular da equação (3.2.12.) .

Para isso vamos considerar uma solução geral do tipo

$$(3.2.21.) \quad F_X^*(y) = A, \text{ onde } A \text{ é uma constante.}$$

Substituindo-se $F_X^*(y)$ em (3.2.12.) obteremos :

$$(3.2.22.) \quad A - \left(\frac{\alpha - 1}{\alpha} \right) A = \frac{1}{\alpha}$$

ou seja

$$(3.2.23.) \quad A = 1$$

Portanto, temos que a solução geral da equação (3.2.12.) será

$$F_X(y) = \bar{F}_X(y) + F_X^*(y)$$

isto é,

$$(3.2.24.) \quad F_X(y) = C \left(\frac{\alpha - 1}{\alpha} \right)^y + 1$$

Para determinarmos a constante C , vamos considerar o seguinte fato :

Tem-se por hipótese que o suporte da variável aleatória X é o conjunto $\{ 0, 1, 2, \dots \}$ e logo tem-se que para todo y inteiro negativo, $F_X(y) = 0$.

Consideremos, v um inteiro negativo qualquer. Então

$$(3.2.25.) \quad F_X(v) = C \left(\frac{\alpha - 1}{\alpha} \right)^v + 1 = 0$$

e isto implica que

$$(3.2.26.) \quad C = - \left(\frac{\alpha - 1}{\alpha} \right)^{-v}$$

e portanto

$$(3.2.27.) \quad F_X(y) = 1 - \left(\frac{\alpha - 1}{\alpha} \right)^{y-v}$$

Precisamos, então, determinar o valor de v .

Lembrando-se que

$$(3.2.28.) \quad \sum_{y=0}^{\infty} P[X=y] = 1$$

pois X é uma variável aleatória cujo suporte é $\{0, 1, 2, \dots\}$, e que

$$\begin{aligned} (3.2.29.) \quad P[X=y] &= F_X(y) - F_X(y-1) \\ &= \left(\frac{\alpha - 1}{\alpha} \right)^{y-v-1} - \left(\frac{\alpha - 1}{\alpha} \right)^{y-v} \\ &= \left(\frac{\alpha - 1}{\alpha} \right)^{y-v} \left(\frac{1}{\alpha - 1} \right) \end{aligned}$$

Então, temos

$$\begin{aligned}
\sum_{y=0}^{\infty} P[X=y] &= \frac{1}{\alpha - 1} \sum_{y=0}^{\infty} \left(\frac{\alpha - 1}{\alpha} \right)^{y-\nu} \\
&= \frac{1}{\alpha - 1} \left(\frac{\alpha - 1}{\alpha} \right)^{-\nu} \sum_{y=0}^{\infty} \left(\frac{\alpha - 1}{\alpha} \right)^y \\
&= \frac{1}{\alpha - 1} \left(\frac{\alpha - 1}{\alpha} \right)^{-\nu} \alpha
\end{aligned}$$

pois $\sum_{y=0}^{\infty} \left(\frac{\alpha - 1}{\alpha} \right)^y$ é uma série geométrica infinita de razão $\frac{\alpha - 1}{\alpha} < 1$,

$\alpha > 1$.

Logo

$$\sum_{y=0}^{\infty} P[X=y] = \left(\frac{\alpha - 1}{\alpha} \right)^{-\nu-1}$$

e por (3.2.28.) segue que

$$(3.2.30.) \quad \left(\frac{\alpha - 1}{\alpha} \right)^{-\nu-1} = 1$$

o que implica que

$$(3.2.31.) \quad v = -1$$

e portanto, temos finalmente que

$$(3.2.32.) \quad F_X(y) = 1 - \left(\frac{\alpha - 1}{\alpha} \right)^{y+1}$$

onde $y \in \{0, 1, 2, \dots\}$ e $\alpha > 1$

e isto completa a prova do Lema 3.2.4.

Passaremos, agora, mediante a utilização dos Lemas apresentados e provados nesta secção, a demonstrar a condição suficiente do Teorema 3.2.1. já enunciado.

DEMONSTRAÇÃO DO TEOREMA 3.2.1.

(I) CONDIÇÃO SUFICIENTE

Considerando-se que, por hipótese, a função de distribuição $F_X(x)$ é a de uma variável aleatória cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e que

$$(3.2.33.) \quad E [Z_k / Y_k = y, Y_m = y, Y_{m+1} > y] = \beta, \text{ q.c.}$$

onde $\beta > 1$, $1 \leq k \leq n-1$, $k \leq m < n$

e tendo se em vista o Lema 3.2.2. apresentado, isto é,

$$E \left[Z_k / Y_k=Y, Y_m=Y, Y_{m+1} > Y \right] = \frac{n-m}{n-k} E \left[(X_n - Y) / X_n > Y \right]$$

obtemos

$$(3.2.34.) \quad E \left[(X_n - Y) / X_n > Y \right] = \alpha$$

$$\text{onde } \alpha = \left(\frac{n-k}{n-m} \right) \beta$$

sendo que $\alpha > 1$, pois $\frac{n-k}{n-m} \geq 1$ e $\beta > 1$

Desta forma, pelo Lema 3.2.3. obtemos a seguinte equação de diferenças finitas de primeira ordem

$$(3.2.35.) \quad F_X(y+1) - \left(\frac{\alpha-1}{\alpha} \right) F_X(y) = \frac{1}{\alpha}$$

cuja solução, pelo Lema 3.2.4., tem a forma

$$(3.2.36.) \quad F_X(y) = 1 - \left(\frac{\alpha-1}{\alpha} \right)^{y+1}$$

onde $y \in \{0, 1, 2, \dots\}$ e $\alpha > 1$

e isto completa a prova da condição suficiente do teorema dado .

(II) CONDIÇÃO NECESSÁRIA

Para mostrarmos a condição necessária, basta verificarmos por cálculo direto que com a função distribuição F_X dada em (3.2.36.) a equação (3.2.9.) é verdadeira, o que é equivalente a mostrar que a equação

$$(3.2.37.) \quad (\alpha+y) [1 - F_X(y)] = \sum_{x=y+1}^{\infty} x [F_X(x) - F_X(x-1)]$$

é verdadeira.

Com efeito, substituindo-se (3.2.36.) no lado direito de (3.2.37.), e tendo em vista que $0 < \frac{\alpha-1}{\alpha} < 1$, teremos

$$\begin{aligned} (3.2.38.) \quad & (\alpha+y) [1 - F_X(y)] \\ &= \sum_{x=y+1}^{\infty} x \left[\left(\frac{\alpha-1}{\alpha} \right)^x - \left(\frac{\alpha-1}{\alpha} \right)^{x+1} \right] \\ &= \frac{1}{\alpha} \sum_{x=y+1}^{\infty} x \left(\frac{\alpha-1}{\alpha} \right)^x \\ &= \frac{1}{\alpha} \{ (y+1) \left(\frac{\alpha-1}{\alpha} \right)^{y+1} + (y+2) \left(\frac{\alpha-1}{\alpha} \right)^{y+2} + (y+3) \left(\frac{\alpha-1}{\alpha} \right)^{y+3} + \dots \} \\ &= \frac{1}{\alpha} \left(\frac{\alpha-1}{\alpha} \right)^{y+1} \{ (y+1) + (y+2) \left(\frac{\alpha-1}{\alpha} \right) + (y+3) \left(\frac{\alpha-1}{\alpha} \right)^2 + \dots \} \end{aligned}$$

$$= \frac{1}{\alpha} \left(\frac{\alpha - 1}{\alpha} \right)^{y+1} [\alpha(y+\alpha)] = (y+\alpha) \left(\frac{\alpha - 1}{\alpha} \right)^{y+1}$$

Agora, tomando-se , $F_X(y)$ dada por (3.2.36.) no lado esquerdo de (3.2.37.), vamos obter uma identidade, o que completa a prova da condição / necessária, e portanto temos que o teorema 3.2.1. está completamente demonstrado.

Vamos provar, agora, o Teorema 3.2.2.

Para isso, vamos, inicialmente, estabelecer alguns lemas básicos à sua demonstração, sendo que o processo a ser utilizado nesta demonstração é semelhante ao usado no Teorema 3.2.1.

LEMA 3.2.5. Para Y_k, Y_m, Y_{m+1} e $A_{\sigma(m-1)\delta(s)}$, $k-m+1 \leq s \leq n-m+1$

como definidos anteriormente

$$\begin{aligned} (3.2.39.) \quad & E \left[\sum_{i=1}^{k-1} (Y_k - Y_i) / Y_{m-1} < y, Y_m = y, Y_k = y \right] \\ &= \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} E \left[\sum_{j \in \sigma(m-1)} (y - X_j) / A_{\sigma(m-1)\delta(s)} \right] \times \\ &\times P \left[A_{\sigma(m-1)\delta(s)} / Y_{m-1} < y, Y_m = y, Y_k = y \right] \end{aligned}$$

PROVA

Como pelo Lema 2.2.5.

$$\{ Y_{m-1} < y, Y_m = y, Y_k = y \} = \bigcup_{s \in \sigma, \delta} \bigcup_{\sigma(m-1) \delta(s)} A$$

onde os conjuntos $A_{\sigma(m-1) \delta(s)}$ são disjuntos, para todo s , $k-m+1 \leq s \leq n-m+1$

teremos que sobre $A_{\sigma(m-1) \delta(s)}$ vale a seguinte relação

$$(3.2.40.) \quad \sum_{i=1}^{k-1} (Y_k - Y_i) = \sum_{j \in \sigma(m-1)} (y - x_j)$$

Desta forma podemos calcular a

$$E \left[\sum_{i=1}^{k-1} (Y_k - Y_i) / Y_{m-1} < y, Y_m = y, Y_k = y \right]$$

através do seguinte caminho

$$E \left[\sum_{i=1}^{k-1} (Y_k - Y_i) / Y_{m-1} < y, Y_m = y, Y_k = y \right]$$

$$= \frac{\sum_{\{Y_{m-1} < Y, Y_m = Y, Y_k = Y\}} \sum_{i=1}^{k-1} (Y - Y_i) P[Y_i = Y_i, 1 \leq i \leq k-1]}{P[Y_{m-1} < Y, Y_m = Y, Y_k = Y]}$$

$$= \frac{\sum_{\substack{U \\ s}} \sum_{\substack{U \\ \sigma, \delta}} \sum_{\substack{A \\ \sigma(m-1)\delta(s)}} \sum_{i=1}^{k-1} (Y - Y_i) P[Y_i = Y_i, 1 \leq i \leq k-1]}{P\{[Y_{m-1} < Y, Y_m = Y, Y_k = Y]\}}$$

$$= \frac{\sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} \sum_{i=1}^{k-1} (Y - Y_i) P\{[Y_i = Y_i, 1 \leq i \leq k-1] \cap A_{\sigma(m-1)\delta(s)}\}}{P[Y_{m-1} < Y, Y_m = Y, Y_k = Y]}$$

$$= \frac{\sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} \sum_{j \in \sigma(m-1)} (Y - x_j) P[X_j = x_j, j \in \sigma(m-1)]}{P[Y_{m-1} < Y, Y_m = Y, Y_k = Y]}$$

$$= \frac{\sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} \sum_{j \in \sigma(m-1)} (Y - x_j) P[X_j = x_j, j \in \sigma(m-1)]}{P[A_{\sigma(m-1)\delta(s)}]} \times$$

$$\times \frac{P[A_{\sigma(m-1)\delta(s)}]}{P[Y_{m-1} < Y, Y_m = Y, Y_k = Y]}$$

Como $A_{\sigma(m-1)\delta(s)}$ é um subconjunto do conjunto $\{Y_{m-1} < Y, Y_m = Y, Y_k = Y\}$

temos que

$$E \left[\sum_{i=1}^{k-1} (Y_k - Y_i) / Y_{m-1} < Y, Y_m = Y, Y_k = Y \right]$$

$$= \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} \left[\sum_{j \in \sigma(m-1)} (Y - X_j) \frac{P \{ [X_j = x_j, j \in \sigma(m-1)] \cap A_{\sigma(m-1)\delta(s)} \}}{P [A_{\sigma(m-1)\delta(s)}]} \right]$$

$$\times \frac{P \{ A_{\sigma(m-1)\delta(s)} \cap [Y_{m-1} < Y, Y_m = Y, Y_k = Y] \}}{P [Y_{m-1} < Y, Y_m = Y, Y_k = Y]}$$

$$= \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} E \left[\sum_{j \in \sigma(m-1)} (Y - X_j) / A_{\sigma(m-1)\delta(s)} \right] \times$$

$$\times P [A_{\sigma(m-1)\delta(s)} / Y_{m-1} < Y, Y_m = Y, Y_k = Y]$$

Portanto o Lema 3.2.5. está provado .

Pelas mesmas razões dadas para o desenvolvimento do Lema 3.2.2., teremos que

$$i) \quad E \left[\sum_{j \in \sigma(m-1)} (y - X_j) / A_{\sigma(m-1)\delta(s)} \right] = E \left[\sum_{j \in \sigma_0(m-1)} (y - X_j) / A_{\sigma_0(m-1)\delta_{m-1}(s)} \right]$$

$$ii) \quad E \left[(y - X_j) / A_{\sigma_0(m-1)\delta_{m-1}(s)} \right] = E \left[(y - X_1) / A_{\sigma_0(m-1)\delta_{m-1}(s)} \right]$$

para $j < m$, $2 \leq m \leq k$, onde

$$\sigma_0(m-1) = \{1, 2, \dots, m-1\} \quad \text{e} \quad \delta_{m-1}(s) = \{m, m+1, \dots, m+s-1\}$$

Vamos utilizar também o resultado provado no Lema 2.2.6., isto é

$$P \left[A_{\sigma(m-1)\delta(s)} / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

onde

$$p_{m-1,s} = P \left[A_{\sigma_0(m-1)\delta_{m-1}(s)} \right] \quad \text{e} \quad p_{m-1,i} = P \left[A_{\sigma_0(m-1)\delta_{m-1}(1)} \right]$$

Estes fatos, conjuntamente, nos permite o estabelecimento do seguinte lema

LEMA 3.2.6. Para W_k , Y_{m-1} , Y_m e Y_k já definidos, temos

$$(3.2.41) \quad E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \frac{m-1}{k-1} E \left[(y - X_1) / X_1 < y \right]$$

PROVA

Utilizando-se o Lema 3.2.5., teremos

$$\begin{aligned} & E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] \\ &= \frac{1}{k-1} E \left[\sum_{i=1}^{k-1} (Y_k - Y_i) / Y_{m-1} < y, Y_m = y, Y_k = y \right] \\ &= \frac{1}{k-1} \sum_{s=k-m+1}^{n-m+1} \sum_{\sigma, \delta} E \left[\sum_{j \in \sigma(m-1)} (y - X_j) / A_{\sigma(m-1)} \delta(s) \right] \times \\ &\times P \left[A_{\sigma(m-1)} \delta(s) / Y_{m-1} < y, Y_m = y, Y_k = y \right] \end{aligned}$$

$$= \frac{1}{k-1} \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} E \left[\sum_{j \in \sigma_0(m-1)} (Y-X_j) / A_{\sigma_0(m-1)} \delta_{(m-1)}(s) \right] \times$$

$$\times \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$= \frac{1}{k-1} \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} E \left[\sum_{j=1}^{m-1} (Y-X_j) / A_{\sigma_0(m-1)} \delta_{m-1}(s) \right] \times$$

$$\times \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$= \frac{1}{k-1} \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} (m-1) E \left[(Y-X_1) / A_{\sigma_0(m-1)} \delta_{m-1}(s) \right] \times$$

$$\times \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$= \frac{m-1}{k-1} \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} E[(Y-X_1)/X_1 < Y, \dots, X_{m-1} < Y, X_m = Y, \dots, X_k = Y, \dots,$$

$$\dots X_{m+s-1} = Y, X_{m+s} > Y, \dots, X_n > Y] \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$= \frac{m-1}{k-1} \sum_{s=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{s} E[(Y-X_1)/X_1 < Y] \frac{p_{m-1,s}}{\left(\sum_{i=k-m+1}^{n-m+1} \binom{n}{m-1} \binom{n-m+1}{i} p_{m-1,i} \right)}$$

$$= \frac{m-1}{k-1} E[(Y-X_1)/X_1 < Y]$$

Deve ser observado que existem $\binom{n}{m-1} \binom{n-m+1}{s}$ termos sob o somatório $\sum_{\sigma, \delta}$,
 $(m-1)$ termos sob o somatório $\sum_{j \in \sigma_0(m-1)}$, que $E[(Y-X_1)/A_{\sigma_0(m-1)\delta_{m-1}(s)}]$ é

constante para todo s , $k-m+1 \leq s \leq n-m+1$ e ainda que a última igualdade decorre do fato de ser a variável aleatória X_1 independente de todas as demais variáveis $X_2, \dots, X_n$.

E, portanto, a prova do Lema 3.2.6. está completa.

Observando-se, agora, que

$$E \left[(y - X_1) / X_1 < y \right] = E \left[-X_1 - (-y) / -X_1 > (-y) \right]$$

onde $-X_1$ assume os valores $-x \in \{0, 1, 2, \dots\}$,

teremos que tudo o que foi aplicado na demonstração do teorema 3.2.1. para a variável aleatória X_n poderá ser aplicado à variável aleatória $-X_1$. Desta forma, teremos que o Lema 3.2.3. poderá ser expresso como :

LEMA 3.2.7. Se F_{-X} é a função distribuição da variável aleatória $-X$, discreta, não degenerada, cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e $\alpha > 1$ é uma constante e

$$E \left[(y - X_1) / X_1 < y \right] = \alpha, \text{ q.c.}$$

Então

$$(3.2.42.) \quad F_{-X}(y+1) - \left(\frac{\alpha - 1}{\alpha} \right) F_{-X}(y) = \frac{1}{\alpha}$$

Consequentemente, o Lema 3.2.4. poderá ser expresso como :

LEMA 3.2.8. A equação de diferenças finitas de primeira ordem dada por (3.2.42.) onde F_{-X} é uma função de distribuição, tem a seguinte solução para $F_{-X}(y)$

$$(3.2.43.) \quad F_{-X}(y) = 1 - \left(\frac{\alpha - 1}{\alpha} \right)^{y+1}$$

onde $y \in \{ 0, 1, 2, \dots \}$ e $\alpha > 1$

Podemos, agora, demonstrar o teorema 3.2.2.

DEMONSTRAÇÃO DO TEOREMA 3.2.2.

(I) CONDIÇÃO SUFICIENTE

Considerando-se que, por hipótese, a função $F_{-X}(x)$ é a de uma variável aleatória cujo suporte é o conjunto $\{ 0, 1, 2, \dots \}$ e que

$$(3.2.44.) \quad E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \beta, \text{ q.c.}$$

onde $\beta > 1$ é uma constante e $2 \leq k \leq n$, $2 \leq m \leq k$

e, tendo se em vista o resultado apresentado no Lema 3.2.6., isto é,

$$E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \frac{m-1}{k-1} E \left[(y - X_1) / X_1 < y \right]$$

obtemos

$$(3.2.45.) \quad E[(y-X_1)/X_1 < y] = E[(-X_1 - (-y)) / -X_1 > (-y)] = \alpha$$

onde $\alpha = \left(\frac{k-1}{m-1} \right) \beta$

sendo que $\alpha > 1$, pois $\frac{k-1}{m-1} > 1$ e $\beta > 1$

Desta forma, pelo Lema 3.2.7. obtemos a seguinte equação de diferenças de primeira ordem

$$(3.2.46.) \quad F_{-X}(y+1) - \left(\frac{\alpha-1}{\alpha} \right) F_{-X}(y) = \frac{1}{\alpha}$$

cuja solução, pelo Lema 3.2.8., tem a seguinte forma

$$(3.2.47.) \quad F_{-X}(y) = 1 - \left(\frac{\alpha-1}{\alpha} \right)^{y+1}$$

onde $y \in \{0, 1, 2, \dots\}$ e $\alpha > 1$

e isto completa a prova da condição suficiente do teorema 3.2.2.

(II) CONDIÇÃO NECESSÁRIA

Para mostrarmos a condição necessária, basta verificarmos que

$$(3.2.48.) \quad E \left[(y - X_1) / X_1 < y \right] = \alpha, \text{ q.c.}$$

$$\text{onde} \quad \alpha = \left(\frac{k-1}{m-1} \right) \beta, \quad \beta > 1$$

é verdadeira quando consideramos a função F_{-X} dada por (3.2.47.), pois sabemos pelo Lema 3.2.6. que

$$(3.2.49.) \quad E \left[W_k / Y_{m-1} < y, Y_m = y, Y_k = y \right] = \frac{m-1}{k-1} E \left[(y - X_1) / X_1 < y \right]$$

Mas, nós temos que

$$(3.2.50.) \quad E \left[(y - X_1) / X_1 < y \right] = E \left[(-X_1 - (-y)) / -X_1 > (-y) \right]$$

$$= \sum_{-x = -(y+1)}^{\infty} (-x - (-y)) \frac{P[-X = -x]}{P[-X > (-y)]}$$

onde $-y \in \{0, 1, 2, \dots\}$ e $-x \in \{0, 1, 2, \dots\}$

ou, ainda

$$\begin{aligned}
 (3.2.51.) \quad E \left[(Y-X_1) / X_1 < Y \right] \\
 &= \sum_{x=y+1}^{\infty} (x-y) \frac{P[-X=x]}{P[-X > y]} \\
 &= \frac{1}{[1 - F_{-X}(y)]} \sum_{x=y+1}^{\infty} (x-y) P[-X=x] \\
 &= \frac{1}{[1 - F_{-X}(y)]} \sum_{x=y+1}^{\infty} (x-y) [F_{-X}(x) - F_{-X}(x-1)]
 \end{aligned}$$

onde $y \in \{0, 1, \dots\}$ e $x \in \{0, 1, 2, \dots\}$

Substituindo-se, agora, em (3.2.51.) a função $F_{-X}(y)$ dada por (3.2.47.) e sabendo-se que $0 < \frac{\alpha - 1}{\alpha} < 1$, teremos

$$\begin{aligned}
 (3.2.52.) \quad E \left[(Y-X_1) / X_1 < Y \right] \\
 &= \left(\frac{\alpha - 1}{\alpha} \right)^{-y-1} \sum_{x=y+1}^{\infty} (x-y) \left[\left(\frac{\alpha - 1}{\alpha} \right)^x - \left(\frac{\alpha - 1}{\alpha} \right)^{x+1} \right]
 \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\alpha - 1}{\alpha} \right)^{-y-1} \frac{1}{\alpha} \sum_{x=y+1}^{\infty} (x-y) \left(\frac{\alpha - 1}{\alpha} \right)^x \\
&= \left(\frac{\alpha - 1}{\alpha} \right)^{-y-1} \frac{1}{\alpha} \left\{ \left(\frac{\alpha - 1}{\alpha} \right)^{y+1} + 2 \left(\frac{\alpha - 1}{\alpha} \right)^{y+2} + \dots \right\} \\
&= \left(\frac{\alpha - 1}{\alpha} \right)^{-y-1} \frac{1}{\alpha} \left(\frac{\alpha - 1}{\alpha} \right)^{y+1} \left\{ 1 + 2 \left(\frac{\alpha - 1}{\alpha} \right) + 3 \left(\frac{\alpha - 1}{\alpha} \right)^2 + \dots \right\} \\
&= \frac{1}{\alpha} \left\{ \alpha + \left(\frac{\alpha - 1}{\alpha} \right) \alpha^2 \right\} \\
&= \alpha
\end{aligned}$$

onde $\alpha = \left(\frac{k-1}{m-1} \right) \beta$, $\beta > 1$

e isto completa a prova da condição necessária do teorema 3.2.2., e logo temos que o teorema dado está completamente provado.

3.3. CASOS ESPECIAIS

Nesta seção apresentaremos alguns casos especiais dos teoremas mencionados na seção 3.2.. Alguns serão dados em forma de teoremas e outros em forma de corolários.

Como já dissemos, no capítulo introdutório de nosso trabalho, Ferguson (1965) provou o seguinte teorema de caracterização da distribuição geométrica:

TEOREMA 3.3.1. Sejam X_1, X_2 variáveis aleatórias independentes e identicamente distribuídas, não degeneradas e com função distribuição comum F_X , discreta. Então, $Y_2 - Y_1$ e Y_1 são independentes, se e somente se, X tem distribuição geométrica.

Agora, se observarmos melhor o nosso teorema 3.2.1., nós veremos que, se o aplicarmos, diretamente, para o caso em que $n = 2$ e $m = k = 1$, nós obteremos um teorema de caracterização da distribuição geométrica através das seguintes condições :

i) $E(X)$ existe.

ii) $E [Y_2 - Y_1 / Y_1 = y, Y_2 > y] = \beta$, q.c.

com $\beta > 1$.

Portanto, poderemos estabelecer o seguinte teorema :

TEOREMA 3.3.2. Sejam X_1, X_2 , variáveis aleatórias discretas, não degeneradas, independentes e identicamente distribuídas, cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e cuja função distribuição é F_X . Suponha que $E(X)$ existe. Então :

$$(3.3.1.) \quad E \left[Y_2 - Y_1 / Y_1 = y, Y_2 > y \right] = \beta, \text{ q.c.}$$

onde y é um inteiro não negativo, fixo, e $\beta > 1$ é uma constante, se e somente se, F_X tem a forma

$$(3.3.2.) \quad F_X(x) = 1 - q^{x+1}, \quad x \in \{0, 1, 2, \dots\}$$

com
$$q = \frac{\beta - 1}{\beta}, \quad 0 < q < 1$$

É interessante notar que quando $E(X)$ existe, a condição (3.3.1.) do Teorema 3.3.2. é mais fraca que a condição exigida por Ferguson no Teorema 3.3.1.. A fim de mostrarmos este fato, vamos estabelecer o seguinte Lema:

LEMA 3.3.1. Sejam X_1, X_2 variáveis aleatórias dadas como no teorema (3.3.1.) e suponha que $E(X)$ existe. Então, se $Y_2 - Y_1$ e Y_1 são independentes

$$(3.3.3.) \quad E \left[Y_2 - Y_1 / Y_1 = y, Y_2 > y \right] = \beta, \text{ q.c.}$$

onde β é uma constante.

PROVA

É evidente que, se $Y_2 - Y_1$ e Y_1 são independentes

$$(3.3.4.) \quad E [Y_2 - Y_1 / Y_1 = y] = \alpha$$

onde α é uma constante.

Agora, temos que

$$\begin{aligned} (3.3.5.) \quad E [Y_2 - Y_1 / Y_1 = y] &= E [Y_2 - Y_1 / Y_1 = y, Y_2 \leq y] \frac{P [Y_1 = y, Y_2 \leq y]}{P [Y_1 = y]} + \\ &+ E [Y_2 - Y_1 / Y_1 = y, Y_2 > y] \frac{P [Y_1 = y, Y_2 > y]}{P [Y_1 = y]} \\ &= E [Y_2 - Y_1 / Y_1 = y, Y_2 > y] \frac{P [Y_1 = y, Y_2 > y]}{P [Y_1 = y]} \end{aligned}$$

$$\begin{aligned}
&= E \left[Y_2 - Y_1 / Y_1=y, Y_2 > y \right] \frac{P[Y_1=y, Y_2 - Y_1 > 0]}{P[Y_1=y]} \\
&= E \left[Y_2 - Y_1 / Y_1=y, Y_2 > y \right] \frac{P[Y_1=y] P[Y_2 - Y_1 > 0]}{P[Y_1=y]} \\
&= E[Y_2 - Y_1 / Y_1=y, Y_2 > y] P[Y_2 - Y_1 > 0]
\end{aligned}$$

Utilizando-se, então, o resultado dado em (3.3.4.) e observando-se que

$$(3.3.6.) \quad P[Y_2 - Y_1 > 0]$$

é independente de y , nós podemos concluir que

$$(3.3.7.) \quad E[Y_2 - Y_1 / Y_1=y, Y_2 > y] = \beta$$

onde β é uma constante, dada por

$$(3.3.8.) \quad \beta = \frac{\alpha}{P[Y_2 - Y_1 > 0]}$$

e isto completa a prova do Lema 3.3.1..

Agora, se observarmos a prova do Lema 3.3.1., nós veremos que se $Y_2 - Y_1$ e Y_1 forem independentes, então automaticamente, teremos que a $E[Y_2 - Y_1 / Y_1 = y, Y_2 > y]$ é constante, o que prova a nossa afirmação de que a condição (3.3.1.) do Teorema (3.3.2.) é mais fraca que a condição de Ferguson para o caso em que $E(X)$ existe.

O teorema que apresentaremos agora, também, é muito importante, no tocante à caracterização da distribuição geométrica, e sua demonstração é conseguida mediante a aplicação do teorema 3.2.1.

TEOREMA 3.3.3. Seja $F_X(x)$ a função distribuição de uma variável aleatória X , discreta, não degenerada, cujo suporte é o conjunto $\{0, 1, 2, \dots\}$ e suponha que $E(X)$ existe. Seja $\bar{X}$ a média aritmética da amostra aleatória de tamanho n , selecionada de X , isto é,

$$(3.3.9.) \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Então para algum m , inteiro fixo e arbitrário, $1 \leq m < n$, $n \geq 2$

$$(3.3.10.) \quad E[(\bar{X} - y) / Y_1 = y, Y_m = y, Y_{m+1} > y] = C, \text{ q.c.}$$

onde y é um inteiro não negativo, fixo, e $C > 1$ é uma constante, se e somente se, $F_X(x)$ tem a forma dada por (3.11.) com o parâmetro q , dado por

$$q = \frac{\alpha - 1}{\alpha}, \text{ onde } \alpha = \left(\frac{n}{n-m} \right) C, \quad \alpha > 1, \quad 0 < q < 1$$

DEMONSTRAÇÃO

Temos, inicialmente, que $\bar{X}$ pode ser expressa como

$$(3.3.11.) \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n Y_i$$

Desta forma, teremos que a expressão em (3.3.10) poderá ser escrita /

como

$$(3.3.12.) \quad E \left[(\bar{X} - y) / Y_1=y, Y_m=y, Y_{m+1}>y \right] = \frac{n-1}{n} E[Z_1 / Y_1=y, Y_m=y, Y_{m+1}>y]$$

onde

$$Z_1 = \frac{1}{n-1} E \left[\sum_{i=2}^n (Y_i - Y_1) / Y_1=y, Y_m=y, Y_{m+1}>y \right]$$

pois

$$\begin{aligned}
 & E \left[(\bar{X} - Y) / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right] \\
 &= \frac{1}{n} E \left[\left(\sum_{i=1}^n Y_i - nY \right) / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right] \\
 &= \frac{n-1}{n} \left\{ E \left[\frac{1}{n-1} \sum_{i=1}^n (Y_i - Y) / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right] \right\} \\
 &= \frac{n-1}{n} E \left[\frac{1}{n-1} \sum_{i=2}^n (Y_i - Y) / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right] \\
 &= \frac{n-1}{n} E \left[Z_1 / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right]
 \end{aligned}$$

Utilizando-se, agora, a hipótese dada em (3.3.10.) teremos

$$(3.3.13.) \quad E \left[Z_1 / Y_1=Y, Y_m=Y, Y_{m+1} > Y \right] = \beta, \text{ q.c.}$$

$$\text{onde} \quad \beta = \left(\frac{n}{n-1} \right) C, \quad \beta > 1$$

Mas, pelo teorema 3.2.1., já apresentado, a expressão em (3.3.13.) ocorre, se e somente se, $F_X(x)$ tem a forma

$$(3.3.14.) \quad F_X(x) = 1 - q^{x+1}$$

$$\text{com } x \in \{0, 1, 2, \dots\} \text{ e } q = \frac{\alpha - 1}{\alpha}$$

$$\text{onde } \alpha = \left(\frac{n}{n-m} \right) C, \quad \alpha > 1, \quad 0 < q < 1$$

Desta forma, temos que o teorema 3.3.3. está completamente provado.

Os dois corolários a seguir, encerram a apresentação deste nosso trabalho.

COROLÁRIO 3.3.1. Seja $F_X(x)$ definida como no teorema 3.2.1. e suponha que $E(X)$ existe. Então

$$(3.3.15.) \quad E \left[(Y_n - Y_{n-1}) / Y_{n-1} = y, Y_n > y \right] = \beta, \text{ q.c.}$$

onde y é um inteiro não negativo, fixo e $\beta > 1$ é uma constante, se e somente se, a variável aleatória X tem a distribuição geométrica com o parâmetro q , dado por

$$q = \frac{\beta - 1}{\beta}, \quad 0 < q < 1$$

COROLÁRIO 3.3.2. Seja $F_X(x)$ definida como no teorema 3.2.2. e suponha que $E(X)$ existe. Então

$$(3.3.16.) \quad E \left[(Y_2 - Y_1) / Y_1 < y, Y_2 = y \right] = \beta, \text{ q.c.}$$

onde y é um inteiro não negativo, fixo e $\beta > 1$ é uma constante, se e somente se, a variável aleatória X tem distribuição geométrica com o parâmetro q , dado por

$$q = \frac{\beta - 1}{\beta}, \quad 0 < q < 1$$

As demonstrações destes dois corolários são imediatas, pois o corolário 3.3.1. é uma aplicação direta do teorema 3.2.1. para o caso em que $m = k = n-1$ e o corolário 3.3.2. é uma aplicação direta do teorema 3.2.2. para o caso em que $m = k = 2$.

BIBLIOGRAFIA

- [1] El-Newehi, E. and Govindarajulu, Z. (1979) Characterizations of geometric distribution and discrete IFR (DFR) distributions using order statistics. J. Statist. Planning and Inference, 3 , 85-90.
- [2] Ferguson, T.S. (1965) A characterization of the geometric / distribution . Amer. Math. Monthly, 72 , 256-260.
- [3] Ferguson, T.S. (1967) On characterizing distributions by properties of order statistics . Sankhyā, Ser. A, 29 , 265-278.
- [4] Galambos, J. (1975) Characterizations of probability distributions by properties of order statistics. Statistical Distributions in / Scientific Work, Vol. 2 : Characterizations and Applications. Editors G. P. Patil, S. Kotz and J.K. Ord, D. Reidel, Boston, II , 89-101.
- [5] Galambos, J. and Kotz, S. (1978) Characterizations of probability. Lecture notes in mathematics, n° 675 , Springer Verlag, Berlin.
- [6] Govindarajulu, Z. (1980) . Characterization of the geometric / distribution using properties of order statistics. J. Statist. / Planning and Inference, 4 , 237-247.

- [7] Guelfond,A.O. (1963) Calcul des différences finies. Dunod ,
Paris.

- [8] Gupta,R.C. (1970) Some characterizations of probability /
distributions . Ph. D. thesis, Wayne State University.

- [9] Kotz,S. (1974) Characterizations of statistical distributions :
a supplement to recent surveys. Int. Stat. Rev. , Vol. 42,nº 1 ,
39-65.

- [10] Rohatgi,V.K. (1976) An introduction to probability theory and /
mathematical statistics. John Wiley and Sons, New York.

- [11] Srivastava,R.C. (1974) Two characterizations of the geometric /
distribution. J. Amer. Statist. Assoc. , 69 , 267-269.

- [12] Wang,Y.H. and Srivastava,R.C. (1980) A characterization of the /
exponential and related distributions by linear regression. Ann. of
Statistics, 8 , nº 1 , 217-220.