Um estudo teórico sobre a função de densidade Hipergeométrica confluente G e seus casos particulares

Mariana Gusmão & Natália S. M. Fonseca

Universidade Federal de Alfenas

marye_mb2011@hotmail.com
natalia.martins@unifal-mg

Introdução

Uma distribuição de probabilidade é utilizada para descrever o comportamento de fenômenos aleatórios, em outras palavras, experiências que não são possíveis de se antever o resultado.A distribuição relaciona uma variável aleatória com a sua probabilidade de ocorrência, sendo essa discreta ou contínua. Se contínua, definimos sua função distribuição(f.d) em um espaço contínuo da seguinte forma $F_x(x) = \int_{-\infty}^{x} f_x(x)dx$. A função densidade de probabilidade(f.d.p) definida por $f_x(x)$, apresenta as seguintes propriedades:

$$f_x(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$P[a \leq X \leq b] = \int_{a}^{b} f_x(x)dx$$

A f.d.p é de suma importância na área de estatística pois é essa função que descreve a probabilidade relativa de uma variável aleatória assumir um valor dado.

No presente trabalho foi realizado um estudo da função densidade de probabilidade hipergeométrica confluente G, quando G assume o modelo Weibull. Esta distribuição foi apresentada por Rodrigues, Chaves e Castellares (2012) e é utilizada para ánalise de sobrevivência.

A análise de sobrevivência consiste em uma importante área da ciência estatstica, visto que sua aplicabilidade apresenta grande diversidade, tendo como principal elemento o tempo até a ocorrência do evento de interesse. Dado sua importância, a busca por novos modelos de distribuição é recorrente no meio estatístico.

Objetivos

- Estudar as propriedades da função densidade de probabilidade da família de distribuições Hipergeométrica confluênte G.
- Estudar o caso particular da Hipergeométrica confluênte: hipergeométrica confluente Weibull.
- Apresentar as função de distribuição, densidade e de sobrevivência do caso particular estudado.

Distribuição Hipergeométrica confluente G

A distribuição Hipergeométrica confluente G de Rodrigues, Chaves e Castellares (2012) é uma generalização da família de distribuições Beta G, utilizada para ánalise de sobrevivência. Os autores mostram sua potencialidade por meio da modelagem de um conjunto de dados reais de trinta e cinco crianças com deficiência do hormônio de crescimento.

A fd hipergeométrica confluente G é dada por

$$F(x) = \frac{1}{B(a,b)_1F_1(a,a+b,-c)} \int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt, \tag{1}$$

a fdp associada a esta distribuição é

$$f(x) = \frac{g(x)G(x)^{a-1}[1-G(x)]^{b-1} exp[-cG(x)]}{B(a,b)_1F_1(a,a+b,-c)}, \qquad g(x) = \frac{\partial G(x)}{\partial x}, \tag{2}$$

e a função de sobrevivência é

$$S(x) = 1 - \frac{1}{B(a,b)_1F_1(a,a+b,-c)} \int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt. \tag{3}$$

São apresentados alguns casos particulares desse modelo de distribuio, onde G assume as distribuições normal, log-normal, weibull, gama, pareto, gumbel e weibull inversa.

Nesse estudo foram feitas as derivadas e integrais para chegarmos nos resultados apresentados pelos autores, e uma função no software R para desenvolver os gráficos da fdp hipergeométrica confluente Weibull.

Inicialmente demonstramos que $\frac{\partial F(x)}{\partial x} = f(x)$.

Demonstração.

$$\begin{aligned} \frac{\partial F(x)}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{1}{B(a,b)_1F_1(a,a+b,-c)} \int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt \right) \\ &= \frac{1}{B(a,b)_1F_1(a,a+b,-c)} \frac{\partial}{\partial x}\left(\int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt \right) \end{aligned} \tag{4}$$

De acordo com o Teorema fundamental do calculo(TFC) se f for continua no intervalo $[a, b]$, então a função F definida por: $F(x) = \int_a^x f(t)dt$ é continua e diferenciável neste intervalo. E $F'(x) = f(x)$, ou seja, $\frac{\partial}{\partial x} \int_a^x f(t)dt = f(x)$.

Vamos denominar a função $\int_0^u t^{a-1}(1-t)^{b-1} exp(-ct)dt$ de $H(u)$. Segundo o TFC $\frac{\partial H(u)}{\partial u} = h(u)$, onde $h(t) = t^{a-1}(1-t)^{b-1} exp(-ct)$. Entretanto na integral de interesse presente em (4) $u = G(x)$ assim será necessário realizar a derivação pela Regra da cadeia. Logo teremos $H[G(x)]' = H'[G(x)].G'(x) = h[G(x)].g(x)$, sendo assim

$$\frac{\partial}{\partial x}\left(\int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt \right) = G(x)^{a-1}[1-G(x)]^{b-1} exp[-cG(x)]g(x) \tag{5}$$

Substituindo (5) em (4) temos

$$\frac{\partial F(x)}{\partial x} = \frac{1}{B(a,b)_1F_1(a,a+b,-c)} g(x)G(x)^{a-1}[1-G(x)]^{b-1} exp[-cG(x)]$$

□

Para encontrarmos a fd e a fdp da distribuição hipergeométrica confluente Weibull basta substituirmos $G(x)$ e $g(x)$ pela fd e a fdp, respectivamente, do modelo weibull.

A fd da distribuição Weibull é dada por

$$\mathbf{G(x) = 1 - exp[-(sx)^k]}, \tag{6}$$

e a fdp dessa distribuição é

$$\mathbf{g(x) = ks^k x^{k-1} exp[-(sx)^k]}. \tag{7}$$

Substituindo a fd e a fdp desse modelo em (5), encontramos a fdp hipergeométrica confluente Weibull, que pode ser denotada por HCW(a,b,c,k,s),

$$f(x) = \frac{ks^k x^{k-1} \left\{1 - exp\left[-(sx)^k\right]\right\}^{a-1} exp\left\{cexp\left[-(sx)^k\right] - b(sx)^k\right\}}{exp(c)B(a,b)_1F_1(a,a+b,-c)}. \tag{8}$$

De forma análoga, substituindo $G(x)$ em (3) obtemos a função de sobrevivência hipergeométrica confluente Weibull,

$$S(x) = 1 - \frac{1}{B(a,b)_1F_1(a,a+b,-c)} \int_0^{1-exp[-(sx)^k]} t^{a-1}(1-t)^{b-1} exp(-ct)dt. \tag{9}$$

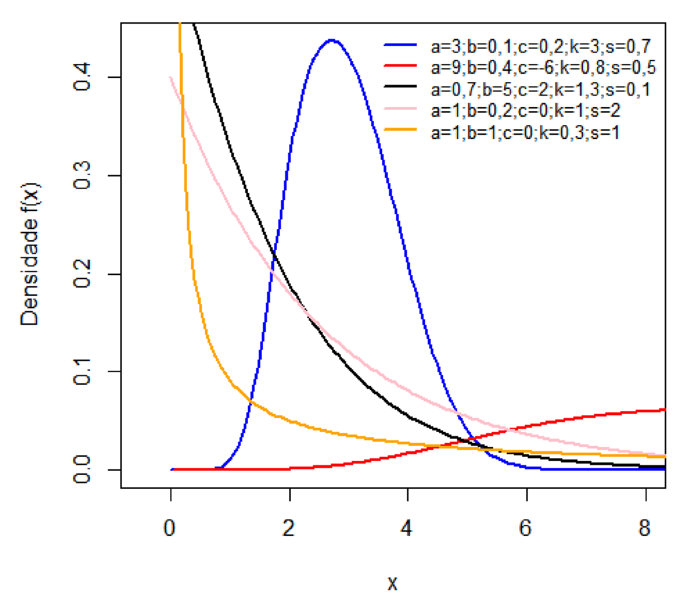

Figura 1: Funções densidade de probabilidade HCW(a,b,c,k,s).

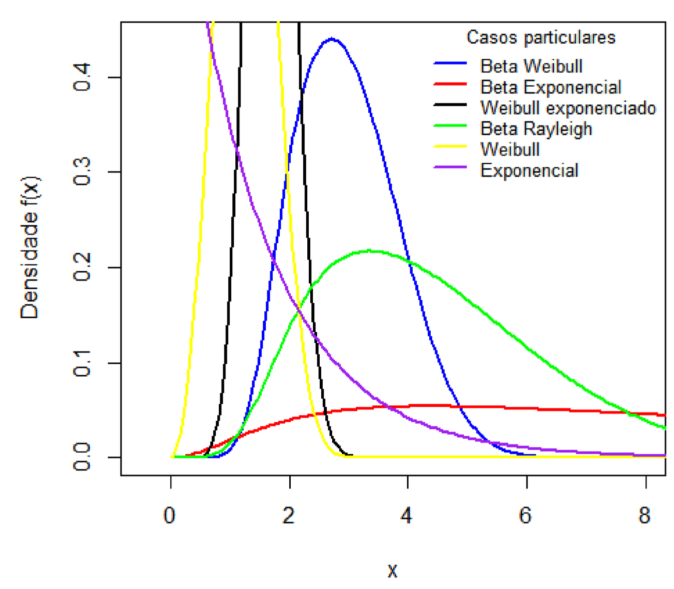

Figura 2: Casos particulares HCW(a,b,c,k,s).

Considerações finais

- Cálculo matemático

  Como algumas funções presentes na estatística não são vistas no decorrer do curso de matemática, se tornou necessário um período de tempo estudando funções de distribuição(fd), mais especificamente, a fd hipergeométrica cofluente G. A ajuda de professores de matemática também se mostrou necessária na busca da primitiva de $\int_0^{G(x)} t^{a-1}(1-t)^{b-1} exp(-ct)dt$, presente na f.d em estudo, além dos conhecimentos prévios adqueridos no curso.

  Entretanto, após um certo período sem êxito, optamos por realizar a derivada da fd por meio do Teorema fundamental do calculo, visto que nosso objetivo não estava centrado no calculo da integral e sim em encontrar a fdp do modelo em estudo, e pelo teorema não era necessário encontrar a primitiva da integral.

- Gráficos

  Com a criação da função HCW(a,b,c,k,s) no software R vimos que a função densidade de probabilidade se assemelha a outras já existentes para a ánalise de sobrevivência. Alem disso, está função só é eficaz para perídos curtos de tempo.

Referências

[1] Luiz Gonzaga. MORETTIN. *Estatística Básica: Probabilidade*. Makron books, 7ed edition, 1999.

[2] RODRIGUES, J A; CHAVES, L. M.; CASTELLARES, F. Uma nova classe de distribuies generalizadas. Ten. Mat. Apl. Comput., 13(2):167-178, Julho 2012.

[3] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.