SHORT COMMUNICATIONS

The density curve of *F* distribution*

LIU Xiaopeng and LIU Kunhui**

(College of Sciences, Beijing Jiaotong University, Beijing 100044, China)

Received June 25, 2003; revised August 29, 2003

Abstract Employing the properties of special function, we discuss the positional relation between two density curves with different parameters for F distribution in this paper. Some varying regularities about the position of density curve of F distribution have been obtained.

Keywords: ***F* distribution, density curve, parameter, Γ-function.**

Denote by $F(m,n)$ the F distribution with parameters m and n, also by $f_{m,n}(x)$ the probability density function of $F(m,n)$. $y=f_{m,n}(x)$ is called the density curve of $F(m,n)$.

It can be proved that

Theorem 1. Suppose that m, n_1 and n_2 are all positive constants and $n_1<n_2<\infty$. If $0<m\leqslant 2$, the equation $f_{m,n_1}(x)=f_{m,n_2}(x)$, $0<x<\infty$ has only one root $\tilde{x}\in(1,\infty)$, and $f_{m,n_1}(x)<f_{m,n_2}(x)$ when $x\in(0,\tilde{x})$, $f_{m,n_1}(x)>f_{m,n_2}(x)$ when $x\in(\tilde{x},\infty)$. If $m>2$, the equation $f_{m,n_1}(x)=f_{m,n_2}(x)$, $0<x<\infty$ has two roots $\underline{x}\in(0,1)$ and $\bar{x}\in(1,\infty)$, $f_{m,n_1}(x)<f_{m,n_2}(x)$ when $x\in(\underline{x},\bar{x})$ and $f_{m,n_1}(x)>f_{m,n_2}(x)$ when $x\in(0,\underline{x})\cup(\bar{x},\infty)$.

Theorem 2. Suppose that m and n_0 are all positive constants, n is a positive variable parameter. When $0<m\leqslant 2$ and $n\neq n_0$, denote by $\tilde{x}(n)\in(1,\infty)$ the only root of equation $f_{m,n}(x)=f_{m,n_0}(x)$, $0<x<\infty$. When $m>2$ and $n\neq n_0$, denote by $\underline{x}(n)\in(0,1)$ and $\bar{x}(n)\in(1,\infty)$ the two roots of equation $f_{m,n}(x)=f_{m,n_0}(x)$, $0<x<\infty$. Then $x=\tilde{x}(n)$, $x=\underline{x}(n)$ and $x=\bar{x}(n)$ are all continuously derivable on $(0,n_0)$ and (n_0,∞), and satisfy the equation

$$\left[q(m,n)-\frac{1}{2}\ln\left(1+\frac{m}{n}x\right)+\frac{1}{2}\,\frac{\frac{m}{n}\left(1+\frac{m}{n}\right)x}{1+\frac{m}{n}x}\right]\mathrm{d}n=\frac{m^2(n-n_0)(x-1)}{2(n+mx)(n_0+mx)}\mathrm{d}x,$$

where

$$q(m,n)\triangleq\lim_{k\to\infty}\left\{\left(\frac{1}{n}+\frac{1}{n+2}+\cdots+\frac{1}{n+2k}\right)-\left(\frac{1}{m+n}+\frac{1}{m+n+2}+\cdots+\frac{1}{m+n+2k}\right)\right\}-\frac{m}{2}\frac{1}{n}.$$

Besides, $\lim\limits_{n\to n_0}\underline{x}(n)$, $\lim\limits_{n\to n_0}\bar{x}(n)$ and $\lim\limits_{n\to n_0}\tilde{x}(n)$ exist and are all roots of the equation

$$q(m,n_0)+\int_0^x\frac{m^2(1-t)}{2(n_0+mt)^2}\mathrm{d}t=0,\quad 0<x<\infty.$$

Theorem 3. Suppose that m_1, m_2 and n are all positive constants and $m_1<m_2<\infty$. If $0<n\leqslant 2$, the equation $f_{m_1,n}(x)=f_{m_2,n}(x)$, $0<x<\infty$ has only one root $\tilde{x}\in(0,1)$, $f_{m_1,n}(x)<f_{m_2,n}(x)$ when $x\in(\tilde{x},\infty)$ and $f_{m_1,n}(x)>f_{m_2,n}(x)$ when $x\in(0,\tilde{x})$. If $n>2$, the equation $f_{m_1,n}(x)=f_{m_2,n}(x)$, $0<x<\infty$ has two roots $\underline{x}\in(0,1)$ and $\bar{x}\in(1,\infty)$, $f_{m_1,n}(x)<f_{m_2,n}(x)$ when $x\in(\underline{x},\bar{x})$ and $f_{m_1,n}(x)>f_{m_2,n}(x)$ when $x\in(0,\underline{x})\cup(\bar{x},\infty)$.

* Supported by the National Natural Science Foundation of China (Grant No. 19671004)

** To whom correspondence should be addressed. E-mail: lkhww@263.net

Theorem 4. Suppose m_0 and n are all positive constants, m is a positive variable parameter. When $0<n\leqslant 2$ and $m\neq m_0$, denote by $\tilde{x}(m)\in(0,1)$ the only root of equation $f_{m,n}(x)=f_{m_0,n}(x)$, $0<x<\infty$. When $n>2$ and $m\neq m_0$, denote by $\underline{x}(m)\in(0,1)$ and $\bar{x}(m)\in(1,\infty)$ the two roots of equation $f_{m,n}(x)=f_{m_0,n}(x)$, $0<x<\infty$. Then $x=\tilde{x}(m)$, $x=\underline{x}(m)$ and $x=\bar{x}(m)$ are all continuously derivable on $(0,m_0)$ and (m_0,∞), and satisfy the equation

$$\Big\{\lim_{k\to\infty}\Big[\Big(\frac{1}{m}+\frac{1}{m+2}+\cdots+\frac{1}{m+2k}\Big)-\Big(\frac{1}{m+n}+\frac{1}{m+n+2}+\cdots+\frac{1}{m+n+2k}\Big)\Big]+\frac{1}{2}+\frac{1}{2}\ln m+\frac{1}{2}\ln x-\frac{(n+m)x}{2(n+mx)}-\frac{1}{2}\ln(n+mx)\Big\}\mathrm{d}m+\Big[\frac{n^2(m-m_0)(1-x)}{2(n+m_0x)(n+mx)x}\Big]\mathrm{d}x=0.$$

Besides, $\lim_{m\to m_0}\tilde{x}(m)$, $\lim_{m\to m_0}\underline{x}(m)$ and $\lim_{m\to m_0}\bar{x}(m)$ exist and are all roots of the equation

$$\lim_{k\to\infty}\Big[\Big(\frac{1}{m_0}+\frac{1}{m_0+2}+\cdots+\frac{1}{m_0+2k}\Big)-\Big(\frac{1}{m_0+n}+\frac{1}{m_0+n+2}+\cdots+\frac{1}{m_0+n+2k}\Big)\Big]+\frac{1}{2}\ln m_0-\frac{1}{2}\ln(m_0+n)-\int_x^1\frac{n^2(1-t)}{2(n+m_0t)^2t}\mathrm{d}t=0,\quad x\in(0,\infty).$$

Theorem 5. Suppose m_0 and n are all positive constants, m is a positive variable parameter. Then $\forall n>0$ and $\forall x>0$, $\lim_{m\to\infty}f_{m,n}(x)$ exists and

$$f_{\infty,n}(x)\triangleq\lim_{m\to\infty}f_{m,n}(x)=\frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}x^{-\frac{n+2}{2}}\mathrm{e}^{-\frac{n}{2x}},$$

$$\forall n>0,\ \forall x>0;$$

$$\sup_{x>0}f_{\infty,n}(x)=f_{\infty,n}(x)\big|_{x=\frac{n}{n+2}}=\frac{(n+2)^{\frac{n+2}{2}}}{2^{\frac{n}{2}}n\Gamma\left(\frac{n}{2}\right)}\mathrm{e}^{-\frac{n+2}{2}},\quad\forall n>0.$$

If $n>2$, $\forall m_0>0$, the equation $f_{\infty,n}(x)=f_{m_0,n}(x)$, $0<x<\infty$ has two roots $\underline{x}(\infty)\in(0,1)$ and $\bar{x}(\infty)\in(1,\infty)$, $f_{\infty,n}(x)>f_{m_0,n}(x)$ when $x\in(\underline{x}(\infty),\bar{x}(\infty))$ and $f_{\infty,n}(x)<f_{m_0,n}(x)$ when $x\in(0,\underline{x}(\infty))\cup(\bar{x}(\infty),\infty)$. If $0<n\leqslant 2$, $\forall m_0>0$, the equation $f_{\infty,n}(x)=f_{m_0,n}(x)$, $0<x<\infty$ has only one root $\tilde{x}(\infty)\in(0,1)$, $f_{\infty,n}(x)<f_{m_0,n}(x)$ when $x\in(0,\tilde{x}(\infty))$ and $f_{\infty,n}(x)>f_{m_0,n}(x)$ when $x\in(\tilde{x}(\infty),\infty)$.

Besides, $\forall n>0$, the density curve $y=f_{m,n}(x)$ has some stability when $m\to\infty$.

Theorem 6. Suppose m_1, m_2, n_1, n_2 are all positive constants and $m_1<m_2$. Let

$$\rho=\frac{\left(\frac{1}{m_1}-\frac{1}{m_2}\right)}{\left(\frac{1}{n_2}-\frac{1}{n_1}\right)};\quad g(x)=\ln\frac{f_{m_2,n_2}(x)}{f_{m_1,n_1}(x)},\quad x\in(0,\infty).$$

Then

If $n_2>n_1$, the equation $f_{m_1,n_1}(x)=f_{m_2,n_2}(x)$, $x\in(0,\infty)$ has two roots $x_1\in(0,1)$ and $x_2\in(1,\infty)$. $f_{m_2,n_2}(x)>f_{m_1,n_1}(x)$ when $x\in(x_1,x_2)$ and $f_{m_2,n_2}(x)<f_{m_1,n_1}(x)$ when $x\in(0,x_1)\cup(x_2,\infty)$.

If $n_2<n_1$, then $\rho\in(0,\infty)$ and

(1) $\rho=1$. The equation $f_{m_1,n_1}(x)=f_{m_2,n_2}(x)$, $x\in(0,\infty)$ has only one root x_1, $f_{m_2,n_2}(x)<f_{m_1,n_1}(x)$ when $x\in(0,x_1)$ and $f_{m_2,n_2}(x)>f_{m_1,n_1}(x)$ when $x\in(x_1,\infty)$.

(2) $0<\rho<1$, then $g(1)<g(\rho)$ and

(i) $g(1)>0$, the conclusion is the same as (1).

(ii) $g(1)=0$, the equation $f_{m_1,n_1}(x)=f_{m_2,n_2}(x)$, $x\in(0,\infty)$ has two roots $x_1\in(0,\rho)$ and $x_2=1$. When $x_1<x\neq 1$, there is $f_{m_2,n_2}(x)>f_{m_1,n_1}(x)$, and when $x<x_1$, there is $f_{m_2,n_2}(x)<f_{m_1,n_1}(x)$.

(iii) $g(1)<0<g(\rho)$, the above equation has three roots $x_1\in(0,\rho)$, $x_2\in(\rho,1)$ and $x_3\in(1,\infty)$. When $x\in(0,x_1)\cup(x_2,x_3)$ there is $f_{m_2,n_2}(x)<f_{m_1,n_1}(x)$ and when $x\in(x_1,x_2)\cup(x_3,\infty)$ there is $f_{m_2,n_2}(x)>f_{m_1,n_1}(x)$.

(iv) $g(\rho)=0$, the above equation has two roots $x_1=\rho$ and $x_2 \in (1, \infty)$. When $x_2 > x \neq \rho$ there is $f_{m_2, n_2}(x) < f_{m_1, n_1}(x)$ and when $x > x_2$ there is $f_{m_2, n_2}(x) > f_{m_1, n_1}(x)$.

(v) $g(\rho) < 0$, the above equation has only one root $x_1 \in (1, \infty)$, $f_{m_2, n_2}(x) < f_{m_1, n_1}(x)$ when $x < x_1$ and $f_{m_2, n_2}(x) > f_{m_1, n_1}(x)$ when $x > x_1$.

(3) $\rho > 1$, then $g(1) > g(\rho)$ and the problem can be discussed by analysing the following cases: i) $g(\rho) > 0$. ii) $g(\rho) = 0$. iii) $g(\rho) < 0 < g(1)$. iv) $g(1) = 0$. v) $g(1) < 0$. The corresponding conclusions are also similar to (2).

Theorem 7. Suppose m_0, n_0, m_1, n_1 and x_1 are all positive constants, $(m_1, n_1) \neq (m_0, n_0)$ and $f_{m_1, n_1}(x_1) = f_{m_0, n_0}(x_1)$. If $U(m_1, n_1, x_1) \neq 0$, then there exists an only continuously derivable function $x = x(m, n)$ on some neighborhood of (m_1, n_1), which satisfies $x(m_1, n_1) = x_1$ and

$$f_{m, n}(x) = f_{m_0, n_0}(x),$$

$$V(m, n, x) + U(m, n, x) \frac{\partial x}{\partial m} = 0,$$

$$T(m, n, x) + U(m, n, x) \frac{\partial x}{\partial n} = 0,$$

where

$$U(m, n, x) = \frac{(1-x)[n n_0(m - m_0) + m m_0 x(n - n_0)]}{2x(n_0 + m_0 x)(n + mx)},$$

$$\begin{aligned} V(m, n, x) = & \lim_{k \to \infty}\left[\left(\frac{1}{m} + \frac{1}{m+2} + \cdots + \frac{1}{m+2k}\right) - \left(\frac{1}{m+n} + \frac{1}{m+n+2} + \cdots + \frac{1}{m+n+2k}\right)\right] \\ & + \frac{1}{2} + \frac{1}{2}\ln m + \frac{1}{2}\ln x - \frac{1}{2}\ln(n + mx) - \frac{(m+n)x}{2(n+mx)}, \end{aligned}$$

$$\begin{aligned} T(m, n, x) = & \lim_{k \to \infty}\left[\left(\frac{1}{n} + \frac{1}{n+2} + \cdots + \frac{1}{n+2k}\right) - \left(\frac{1}{n+m} + \frac{1}{n+m+2} + \cdots + \frac{1}{n+m+2k}\right)\right] \\ & - \frac{m}{2} \cdot \frac{1}{n} - \frac{1}{2}\ln(n + mx) + \frac{1}{2}\ln n + \frac{m+n}{2n} \cdot \frac{mx}{n+mx}; \end{aligned}$$

$$m > 0, \quad n > 0, \quad x > 0.$$

Specially, there are two continuously derivable functions $x = x_1(m, n)$ and $x = x_2(m, n)$ on region $\{(m, n): m > m_0, n > n_0\}$, which satisfy the above three equations and $0 < x_1(m, n) < 1 < x_2(m, n)$.

Note. All the cases of Theorem 6 are indeed existent but the limit of $x(m, n)$ is not certainly existent when $m \to m_0$ and $n \to n_0$, where $x(m, n)$ is some solution function of equation $f_{m, n}(x) = f_{m_0, n_0}(x)$ in Theorem 7. Several examples can be given to explain the above facts.

The basic introduction to F distribution is given in [1~2].

References

1 DeGroot, M. H. Probability and Statistics, 2nd ed., California: Addison-Wesley Publishing Company, 1984, 499~506.

2 Hogg, R. V. et al. Probability and Statistical Inference, 3rd ed., New York: Macmillan Publishing Company, 1988, 266~276.