

Phân phối F

Phân phối F chỉ là một tỷ lệ của hai phân phối chi bình phương riêng biệt và nó được sử dụng để so sánh phương sai của hai mẫu. Kết quả là nó có hai bậc tự do khác nhau, một bậc cho mỗi mẫu.

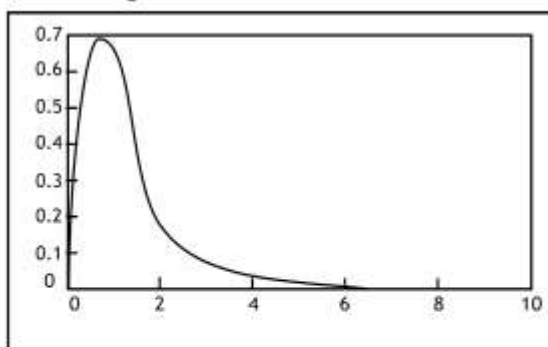
Đây là hàm mật độ xác suất của phân phối F:

$$f(x) = \begin{cases} \frac{\left(\int_0^\infty x^{\frac{v_1+v_2-1}{2}} e^{-x} dx \right) \times (v_1)^{\frac{v_1}{2}} \times (v_2)^{\frac{v_2}{2}}}{\left(\int_0^\infty x^{\frac{v_1-1}{2}} e^{-x} dx \right) \times \left(\int_0^\infty x^{\frac{v_2-1}{2}} e^{-x} dx \right)} \times \frac{x^{\frac{v_1-1}{2}}}{(v_1 \times x + v_2)^{\frac{v_1+v_2}{2}}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

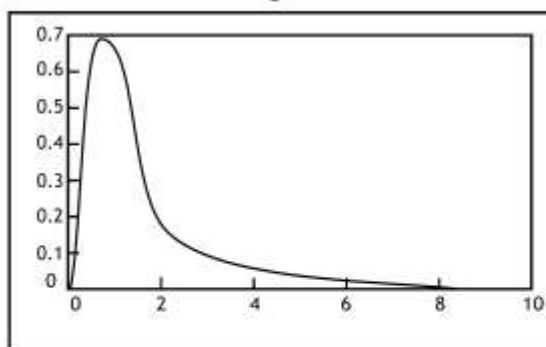
Nếu hàm mật độ xác suất của X là hàm được trình bày ở trên, trong thống kê, chúng ta nói, “ X tuân theo phân phối F với bậc tự do v_1 và v_2 ”.

Khi $v_1 = 5$ và $v_2 = 10$ và khi $v_1 = 10$ và $v_2 = 5$, chúng ta nhận được các đường cong hơi khác nhau, như trong Hình 1-5.

Khi $v_1 = 5$ và $v_2 = 10$

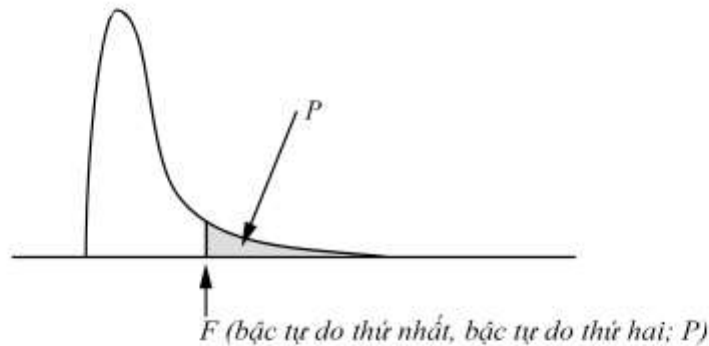


Khi $v_1 = 10$ và $v_2 = 5$



Hình 1-5: Đường cong mật độ phân bố F cho 5 và 10 bậc tự do tương ứng (trái) và 10 và 5 bậc tự do tương ứng (phải)

Hình 1-6 cho thấy đồ thị của phân bố F với bậc tự do v và v . Điều này cho thấy giá trị F dưới dạng một điểm trên trục hoành và tổng diện tích của phần tô bóng ở bên phải là xác suất P để một biến có phân phối F có giá trị lớn hơn giá trị F đã chọn.



Hình 1-6: xác suất P mà giá trị x vượt quá giá trị F tới hạn

Bảng 1-7 cho thấy bảng phân phối F khi $p = 0,05$.

Bảng 1-7: Bảng phân phối xác suất F với $p = 0,05$

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10
1	161.4	199.5	215.7	224.6	230.2	264.0	236.8	238.9	240.5	241.9
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.6	9.3	9.1	9.0	8.9	8.9	8.8	8.8	8.8
4	7.7	6.9	6.6	6.4	6.3	6.2	6.1	6.0	6.0	6.0
5	6.6	5.8	5.4	5.2	5.1	5.0	4.9	4.8	4.8	4.7
6	6.0	5.1	4.8	4.5	4.4	4.3	4.2	4.1	4.1	4.1
7	5.6	4.7	4.3	4.1	4.0	3.9	3.8	3.7	3.7	3.6
8	5.3	4.5	4.1	3.8	3.7	3.6	3.5	3.4	3.4	3.3
9	5.1	4.3	3.9	3.6	3.5	3.4	3.3	3.2	3.2	3.1
10	5.0	4.1	3.7	3.5	3.3	3.2	3.1	3.1	3.0	3.0
11	4.8	4.0	3.6	3.4	3.2	3.1	3.1	2.9	2.9	2.9
12	4.7	3.9	3.5	3.3	3.1	3.0	2.9	2.8	2.8	2.8

Sử dụng bảng phân phối F tương tự như sử dụng bảng phân phối chi bình phương, chỉ khác là lần này các tiêu đề cột trên cùng cung cấp bậc tự do cho một mẫu và nhãn hàng cung cấp bậc tự do cho mẫu kia. Một bảng riêng biệt được sử dụng cho mỗi giá trị p chung.

Trong Bảng 1-7, khi $v = 1$ và $v = 12$, giá trị tới hạn là 4,7. Điều này có nghĩa là khi chúng ta thực hiện một kiểm định thống kê, chúng ta tính toán thống kê kiểm định và so sánh nó với giá trị tới hạn là 4,7 từ bảng này; nếu thống kê kiểm định được tính toán lớn hơn 4,7, kết quả của chúng ta được coi là có ý nghĩa thống kê. Trong bảng này, đối với bất kỳ thống kê kiểm định nào lớn hơn số trong bảng, giá trị p nhỏ hơn 0,05. Điều này có nghĩa là khi $v_1 = 1$ và $v_2 = 12$, xác suất của một thống kê F từ 4,7 trở lên xảy ra khi giả thuyết H_0 của bạn là đúng 5%, vì vậy chỉ có 5% cơ hội bác bỏ giả thuyết H_0 khi nó thực sự đúng.

Hãy xem một ví dụ khác. Bảng 1-8 cho thấy bảng phân phối F khi $p = 0,01$.

Bảng 1-8: Bảng phân phối F với $p = 0,01$

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10
1	4052.2	4999.3	5403.5	5624.3	5764.0	5859.0	5928.3	5981.0	6022.4	6055.9
2	98.5	99.0	99.2	99.3	99.3	99.3	99.4	99.4	99.4	99.4
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2
4	21.2	18.8	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1
6	13.7	10.9	9.8	9.1	8.7	8.5	8.3	8.1	8.0	7.9
7	12.2	9.5	8.5	7.8	7.5	7.2	7.0	6.8	6.7	6.6
8	11.3	8.6	7.6	7.0	6.6	6.4	6.2	6.0	5.9	5.8
9	10.6	8.0	7.0	6.4	6.1	5.8	5.6	5.5	5.4	5.6
10	10.0	7.6	6.6	6.0	5.6	5.4	5.2	5.1	4.9	4.8
11	9.6	7.2	6.2	5.7	5.3	5.1	4.9	4.7	4.6	4.5
12	9.3	6.9	6.0	5.4	5.1	4.8	4.6	4.5	4.4	4.3

Bây giờ khi $v_1 = 1$ và $v_2 = 12$, giá trị tới hạn là 9,3. Xác suất để một thống kê mẫu lớn hoặc lớn hơn 9,3 sẽ xảy ra nếu giả thuyết H_0 là đúng chỉ 0,01. Do đó, có một xác suất rất nhỏ là bạn sẽ bác bỏ sai giả thuyết H_0 . Chú ý rằng khi $p = 0,01$, giá trị tới hạn lớn hơn khi $p = 0,05$. Đối với v_1 và v_2 không đổi, khi giá trị p giảm xuống, giá trị tới hạn sẽ tăng lên.

**Biên dịch: Anh Tuấn
(còn tiếp)**

(Tiếp theo trang 57)

[11] Hsu TH, Shiao LH, Hsieh C, Chang DM. 2002. A comparison of the chemical composition and bioactive ingredients of the Chinese medicinal mushroom DongChongXiaCao, its counterfeit and mimic, and fermented mycelium of *Cordyceps sinensis*. *Food Chem* 78: 463-469.

[12] Imtiaj A, Lee TS. 2007. Screening of antibacterial and antifungal activities from Korean wild mushrooms. *World J Agric Sci* 3: 316-321.

[13] Kim HG, Shrestha B, Lim SY, Yoon DH, Chang WC, Shin DJ, Han SK, Park SM, Park JH, Park HI, Sung JM, Jang Y, Chung N, Hwang KC, Kim TW. 2006. Cordycepin inhibits lipopolysaccharide-induced inflammation by the suppression of NF-kappaB through Akt and p38 inhibition in RAW 264.7 macrophage cells. *Eur J Pharmacol* 545: 192-199.