

ANOVA

Tjipto Juwono, Ph.D.

Desember 2017



SURYA
UNIVERSITY

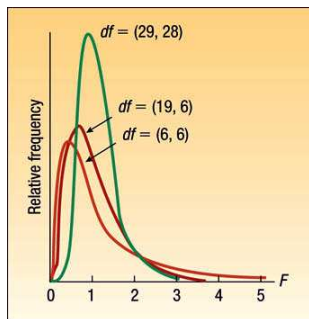
Penggunaan Distribusi- F

- 1 Untuk menguji apakah dua sampel berasal dari populasi dengan variance yang sama.
- 2 Untuk membandingkan beberapa mean populasi secara simultan. Proses membandingkan beberapa mean secara simultan disebut *Analysis of Variance* (ANOVA).

Dalam kedua kasus di atas, populasi-populasi nya harus mempunyai distribusi normal, dan data-datanya paling sedikit harus berupa interval-scale.

Karakteristik Distribusi- F

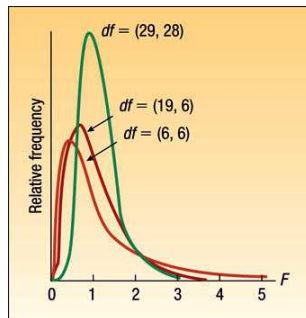
- 1 Merupakan sekelompok beberapa Distribusi- F ("Family").
- 2 Anggota kelompok Distribusi- F ditentukan oleh dua parameter: (1) Derajat kebebasan (df) dari pembilang, dan (2) df dari penyebut.
- 3 Distribusi- F adalah distribusi yang kontinu.



Gambar 1: Distribusi- F

Karakteristik Distribusi- F

- 1 Nilai F tidak dapat negatif.
- 2 Distribusi- F positively-skewed.
- 3 Bersifat asimptotik. Ketika $F \rightarrow \infty$, kurva mendekati sumbu- X tetapi tidak pernah menyentuhnya.



Uji Kesamaan Variance

Pengujian Kesamaan Variance

Hipotesa:

$$\begin{aligned}H_0 : \sigma_1^2 &= \sigma_2^2 \\H_1 : \sigma_1^2 &\neq \sigma_2^2\end{aligned}\quad (1)$$

Statistik:

$$F = \frac{s_1^2}{s_2^2}\quad (2)$$

Penting: Karena nilai $F > 1$, maka selalu letakkan variance yang lebih besar pada *nominator* (pembilang)

Contoh

Sebuah jasa penyewaan mobil ingin membandingkan waktu yang dibutuhkan untuk ke bandara melalui dua rute, yaitu rute-25 dan I75. Diperoleh data yang ditunjukkan pada tabel di kanan ini. Dengan significance level 0.1, apakah ada perbedaan variasi dalam waktu tempuh melalui dua rute tersebut? (Perhatikan: yang dibandingkan bukan rata-rata waktu tempuh, tetapi variasi-nya, Rata-rata: $\bar{X}_1 = 58.29$, $\bar{X}_2 = 59.00$. Tidak ada perbedaan yang signifikan.).

U.S. Route 25	Interstate 75
52	59
67	60
56	61
45	51
70	56
54	63
64	57
	65

Gambar 2: Data waktu tempuh dalam menit

Contoh

① Hipotesa

$$\begin{aligned}H_0 : \sigma_1^2 &= \sigma_2^2 \\H_1 : \sigma_1^2 &\neq \sigma_2^2\end{aligned}\tag{3}$$

② Significance level $\alpha = 0.1$.

③ Statistik: Distribusi-F

Contoh

- ① Aturan pengambilan keputusan. H_0 ditolak jika

$$F > F_{\alpha/2, df1, df2}$$

$$F > F_{0.1/2, 7-1, 8-1}$$

$$F > F_{0.05, 6, 7}$$

$$F > 3.87 \tag{4}$$

Contoh

Degrees of Freedom for Denominator	Degrees of Freedom for Numerator			
	5	6	7	8
1	230	234	237	239
2	19.3	19.3	19.4	19.4
3	9.01	8.94	8.89	8.85
4	6.26	6.16	6.09	6.04
5	5.05	4.95	4.88	4.82
6	4.39	4.28	4.21	4.15
7	3.97	3.87	3.79	3.73
8	3.69	3.58	3.50	3.44
9	3.48	3.37	3.29	3.23
10	3.33	3.22	3.14	3.07

Gambar 3: Nilai Kritis untuk Dist.- F , $\alpha = 0.05$

Contoh

- ⑤ Hitung F , dan ambil keputusan, serta menjelaskan artinya.

$$\begin{aligned} F &= \frac{s_1^2}{s_2^2} \\ &= \frac{(8.9947)^2}{(4.3753)^2} \\ &= 4.23 \end{aligned} \tag{5}$$

$F_{\text{hitung}} > F_{\text{kritis}}$, sehingga H_0 ditolak. Kesimpulannya ada perbedaan variasi dalam waktu tempuh melalui kedua rute.

Membandingkan Mean dari Dua Atau Lebih Populasi

Uji hipotesa membandingkan mean dari dua atau lebih populasi

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1 : \text{Mean - mean tidak semuanya sama} \quad (6)$$

H_0 ditolak jika $F > F_{\alpha, k-1, n-k}$

ANOVA-Statistik- F

- 1 Jika ada k populasi yang diambil sampelnya, maka degree of freedom untuk numerator adalah $df = k - 1$.
- 2 Jika ada n total data, maka degree of freedom untuk denominator adalah $df = n - k$.
- 3 Statistik test dihitung dengan rumus:

$$F = \frac{SST/(k - 1)}{SSE/(n - k)} \quad (7)$$

Contoh

Sebuah perusahaan survey melakukan riset tentang kepuasan konsumen kepada empat perusahaan jasa. Setiap responden diberi 25 pertanyaan, yang jawabannya adalah sangat baik (4), baik (3), cukup (2), dan kurang (1). Hasilnya dijumlahkan dan ditunjukkan pada tabel berikut.

Northern	WTA	Pocono	Branson
94	75	70	68
90	68	73	70
85	77	76	72
80	83	78	65
	88	80	74
		68	65
		65	

Pertanyaan: Adakah perbedaan antara mean dari kepuasan konsumen dari keempat sampel itu? Gunakan significance level 0.01.

Contoh

- 1 Nyatakan Hipotesa, H_0 dan H_1

$$H_0 : \mu_N = \mu_W = \mu_P = \mu_B$$

$$H_1 : \text{Tidak semua mean sama} \quad (8)$$

- 2 Nyatakan significance level, $\alpha = 0.01$
- 3 Statistik: F

Contoh

- 1 Nyatakan aturan pengambilan keputusan. H_0 ditolak jika:

$$F > F_{\alpha, k-1, n-k}$$

$$F > F_{0.01, 4-1, 22-4}$$

$$F > F_{0.01, 3, 18}$$

$$F > 5.09 \tag{9}$$

Contoh

- 5 Hitung F dan buat keputusan dan jelaskan

Sumber Variasi	SS	df	MS	F
Treatments	SST	k-1	$SST/(k-1)=MST$	MST/MSE
Error	SSE	n-k	$SSE/(n-k)=MSE$	
Total	SSTotal	n-1		

$$\begin{aligned}SSTotal &= \sum (X - \bar{X}_G)^2 \\SSE &= \sum (X - \bar{X}_c)^2 \\SST &= SSTotal - SSE\end{aligned}\quad (10)$$

Contoh

	Northern	WTA	Pocono	Branson	Total
	94	75	70	68	
	90	68	73	70	
	85	77	76	72	
	80	83	78	65	
		88	80	74	
			68	65	
			65		
Column total	349	391	510	414	1,664
n	4	5	7	6	22
Mean	87.25	78.20	72.86	69.00	75.64

$$\bar{X}_G = \frac{1664}{22} = 75.64 \quad (11)$$

$$\bar{X}_C = 87.25, 78.20, 72.86, 69.00 \quad (12)$$

Contoh

Sumber Variasi	SS	df	MS	F
Treatments	890.68	3	296.89	8.99
Error	594.41	18	33.02	
Total	1485.09	21		

Contoh

Hasil $F_{\text{hitung}} = 8.99$ lebih besar daripada $F_{\text{critical}} = 5.09$. Sehingga H_0 ditolak. Mean dari keempat kelompok itu tidak semuanya sama. Ada perbedaan dalam tingkat kepuasan konsumen dari keempat group.