

ONE-WAY ANOVA

ONE-WAY ANALYSIS OF VARIANCE (ONE-WAY ANOVA)

- Fungsi: memeriksa apakah ada perbedaan di antara beberapa buah populasi dalam hal nilai rata-rata
- Merupakan perluasan dari *pooled variance t-test for the difference between two means*)
- Tokoh: R. A. Fisher (membangun prinsip-prinsip mendasar yang pertama dalam ANOVA)
- Pertama kali diterapkan di bidang biologi dan pertanian

CONTOH PENERAPAN

- Apakah beberapa buah kantor cabang penjualan suatu produk memiliki rata-rata hasil penjualan yang sama?
- Apakah beberapa pabrik yang memproduksi suatu barang sama dalam hal rata-rata persentase barang cacat yang dihasilkan?
- Apakah beberapa metode pelatihan yang berbeda yang diterapkan kepada pegawai baru memberikan hasil yang berbeda?
- Apakah ada perbedaan rata-rata penghasilan penduduk di antara Kab. Bandung, Kab. Bandung Barat, dan Kota Cimahi?

ASUMSI-ASUMSI DALAM ANOVA

- Semua populasi yang diamati memiliki variansi yang sama (syarat *homoscedasticity*)
- Setiap populasi yang diamati berdistribusi normal
- Setiap sampel merupakan sampel acak yang saling bebas

LANGKAH-LANGKAH PENYELESAIAN

- $H_0: \mu_1 = \mu_2 = \dots = \mu_k$
- $H_1: \text{Di antara } \mu_1, \mu_2, \dots, \mu_k \text{ ada yang berbeda nilai}$
- Taraf nyata: $\alpha = \dots$
- Statistik uji: $F = \frac{MSA}{MSE}$
- Daerah kritis: $F > F_{\alpha; (v_1, v_2)}$; $v_1 = k - 1$, $v_2 = n - k$
- Tentukan nilai statistik uji (hasil *sampling*)
- Keputusan: (tolak H_0 atau jangan tolak H_0)
- Kesimpulan:

CONTOH KASUS

- Untuk mengetahui apakah di antara Kabupaten Bandung, Bandung Barat, dan Cimahi terdapat perbedaan rata-rata penghasilan penduduknya, diambil sampel acak penduduk masing-masing daerah tersebut. Dari populasi penduduk Kab. Bandung diambil 10 sampel, dari Bandung Barat 5 sampel, dan dari Cimahi juga 5 sampel. Hasil *sampling* tersebut (dalam satuan juta Rupiah) diringkaskan dalam tabel di samping. Apakah terdapat perbedaan rata-rata penghasilan penduduk di ketiga daerah tersebut? Gunakan taraf nyata 0,05. Anggaplah masing-masing populasi berdistribusi normal dan memiliki variansi yang sama.

Kab. Bandung	Kab. Bandung Barat	Kota Cimahi
4,2	1,5	1,8
3,7	4,2	4,2
2,8	2,0	5,4
3,2	4,1	3,8
4,3	3,2	4,8
3,9		
6,1		
5,2		
3,7		
2,9		

Apakah ketiga daerah tsb. berbeda dalam hal rata-rata penghasilan penduduknya?

- Garis besar penyelesaian:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1 : Di antara μ_1, μ_2 , dan μ_3 ada yang berbeda nilai

$\alpha = \dots$

$$\text{Statistik uji: } F = \frac{MSA}{MSE}$$

Daerah kritis: $F > F_{\alpha; (v_1, v_2)}$; $v_1 = k - 1$, $v_2 = n - k$

Tentukan nilai statistik uji (hasil *sampling*)

Keputusan: (tolak H_0 atau jangan tolak H_0)

Kesimpulan:

URAIAN PENGHITUNGAN NILAI STATISTIK UJI F

$$F = \frac{MSA}{MSE}$$

$$MSA = \frac{SSA}{k - 1} \quad ; \quad SSA = \sum_{j=1}^k n_j (\bar{X}_j - \bar{X}_G)^2$$

$$MSE = \frac{SSE}{n - k} \quad ; \quad SSE = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$$\bar{X}_G = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \dots + n_k \bar{X}_k}{n_1 + n_2 + \dots + n_k}$$

URAIAN PERHITUNGAN STATISTIK UJI F CONTOH KASUS

i	j		
	1	2	3
1	4,2	1,5	1,8
2	3,7	4,2	4,2
3	2,8	2,0	5,4
4	3,2	4,1	3,8
5	4,3	3,2	4,8
6	3,9		
7	6,1		
8	5,2		
9	3,7		
10	2,9		
JUMLAH	40,0	15,0	20,0
n_j	10	5	5
\bar{X}_j	4,0	3,0	4,0

$$F = \frac{MSA}{MSE}$$

$$MSA = \frac{SSA}{k - 1} ; SSA = \sum_{j=1}^k n_j (\bar{X}_j - \bar{X}_G)^2$$

$$MSE = \frac{SSE}{n - k} ; SSE = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$$\bar{X}_G = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \dots + n_k \bar{X}_k}{n_1 + n_2 + \dots + n_k} \quad \bar{X}_G = \frac{40 + 15 + 20}{10 + 5 + 5} = \frac{75}{20} = 3,75$$

$$SSA = 10(4,0 - 3,75)^2 + 5(3,0 - 3,75)^2 + 5(4,0 - 3,75)^2 = 3,75$$

$$MSA = \frac{3,75}{3 - 1} = \frac{3,75}{2} = 1,875$$

$$SSE = [(4,2 - 4,0)^2 + (3,7 - 4,0)^2 + \dots + (2,9 - 4,0)^2] \\ + [(1,5 - 3,0)^2 + (4,2 - 3,0)^2 + \dots + (3,2 - 3,0)^2] \\ + [(1,8 - 4,0)^2 + (4,2 - 4,0)^2 + \dots + (4,8 - 4,0)^2] \\ = 9,46 + 5,94 + 7,52 = 22,92$$

$$MSE = \frac{22,92}{20 - 3} \approx 1,35$$

$$F = \frac{1,875}{1,35} \approx 1,39$$

PENENTUAN DAERAH KRITIS

- $\alpha = 0,05$
- $df_1 = k - 1 = 3 - 1 = 2$
- $df_2 = n - k = 20 - 3 = 17$
- $F_{0,05;(2,17)} = 3,59$
- Daerah penolakan: $F > 3,59$

Degrees of f										
		13	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
14	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
15	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
16	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
17	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68
	.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75

PENYELESAIAN CONTOH KASUS

- $H_0: \mu_1 = \mu_2 = \mu_3$ [Rata-rata penghasilan penduduk sama di antara ketiga daerah]
- H_1 : Ada daerah yang rata-rata penghasilan penduduknya berbeda
- $\alpha = 0,05$
- Statistik uji: $F = \frac{MSA}{MSE}$
- Daerah kritis/daerah penolakan: $F > 3,59$
- Nilai statistik uji (hasil *sampling*): $F = 1,39$
- Keputusan: Jangan tolak H_0
- Kesimpulan:
Tidak ada perbedaan signifikan dalam hal rata-rata penghasilan penduduk di ketiga daerah tersebut.

TABEL RINGKASAN ANOVA

Sumber Variasi	Jumlah Kuadrat	df	Rata-rata Kuadrat	F
Treatment	SSA	$k - 1$	MSA	MSA/MSE
Error	SSE	$n - k$	MSE	
Total	SST	$n - 1$		

Sumber Variasi	Jumlah Kuadrat	df	Rata-rata Kuadrat	F
Treatment	3,75	2	1,875	1,39
Error	22,92	17	1,35	
Total	26,67	19		