

Rancob 5.1: Asumsi pada tes parametrik

nuzul

Analisis data yang dipelajari dalam MK rancob ini hanya yang termasuk dalam kategori statistik parametrik

Statistik parametrik adalah suatu cabang statistic dimana data yang akan diolah harus memenuhi asumsi-asumsi berikut:

- Data yang akan dianalisis merupakan sample dari populasi yang memiliki distribusi normal, dengan rerata sebesar μ dan standar deviasi sebesar σ
- Terdapat hubungan yang linier antara respons variable (variable yang diamati = y) dengan factor (variable yang memberikan pengaruh = x)
- Independen: antar titik data tidak memiliki hubungan yang sistematis
- Homoscedaticity: dalam satu factor, varians error antar level adalah setara atau homogen

PERTEMUAN 4



When To Use ANOVA

		Independent Variable (X)	
		Continuous	Categorical
Dependent Variable (Y)	Continuous	Regression	ANOVA
	Categorical	Logistic Regression	Chi-Square (χ^2) Test

The tool depends on the data type. ANOVA is used with an attribute (categorical) input and a continuous response.

Variabel bebas berupa kategori (X): variable yang jumlahnya terbatas, biasanya fixed (sudah ditentukan oleh peneliti), berupa grouping (misal: perlakuan) dalam suatu penelitian.

Contoh:

- Perlakuan 1, Perlakuan 2, Perlakuan 3
- Dosis A, Dosis B, Dosis C
- Bangsa X, Bangsa Y, Bangsa Z
- dst

Variabel tidak bebas bersifat kontinyu (Y): adalah variable yang diamati; untuk tiap dua nilai Y yang berurutan, selalu bisa disisipi nilai-nilai lagi diantaranya.

Contoh:

- Bobot Badan
- Ukuran tubuh
- Produksi telur
- Produksi susu

Mengapa asumsi itu penting?

- Misal mbak A punya gebetan temen seangkatan, belum resmi sih
- Udah deket nih, kemarin habis nongkrong di wedangan jugak
- Tapi, kenapa sampe sekarang si Co sama sekali ga kontak ya?!
- Maka mbak A akan mulai menyimpulkan:
 - Apa si Co marah ya? Tapi kenapa?



Mungkin mbak A berasumsi bahwa si Co pegang HaPe dan buka WA setiap saat, makanya dia membuat kesimpulan tsb; OKI...Asumsi itu harus berdasar

Pelajaran apa yang dapat diambil?

- **Jika Asumsi mbak A salah, maka dia tidak akan dapat melihat peristiwa yang sebenarnya**
- **Jika Asumsi yang dibuat tidak benar, maka kesimpulan yang diambil tidak akan ada benarnya pula**
- **Asumsi harus dibuktikan supaya kesimpulan yang diambil tidak salah**

Statistik Parametrik

- Statistik parametrik, dibuat berdasarkan satu set asumsi tentang data
- Agar kesimpulan yang diambil tepat, data perlu diuji apakah memenuhi asumsi-asumsi yang disyaratkan
- Jika asumsi-asumsi tersebut tidak terpenuhi, maka kesimpulan yang diambil bisa jadi memiliki nilai kebenaran yang rendah



Asumsi dasar tes parametrik

- Data yang akan diuji dengan tes parametrik harus memenuhi asumsi2 disamping
- Semakin banyak asumsi yang terpenuhi, hasil analisis parametrik akan semakin dapat dipercaya
- Tes parametrik memiliki power yang lebih besar dari non-parametrik dikarenakan asumsi2 tersebut
 - Akan lebih akurat mendeteksi jika ada perbedaan antar perlakuan
 - Peluang untuk mengambil kesimpulan yang salah (berdasarkan hasil analisis data) akan semakin kecil

Data berupa interval atau rasio

Data berdistribusi normal

Sampel acak dan independen

Homogenitas varians

Asumsi dasar tes parametrik

Data berupa interval atau rasio

Data berdistribusi normal

Sampel acak dan independen

Homogenitas varians

Data berupa interval atau rasio

- Response variable harus berupa interval, atau rasio atau sering juga disebut variable kontinu (continuous variable)
- Parametrik statistic tidak dapat dilakukan pada data yang bukan berupa interval atau rasio.
- Mudah diverifikasi: cukup dengan melihat datanya
 - Jika diantara dua nilai data yg berturut-turut masih bisa disisipi nilai lagi ditengahnya berarti dia merupakan variable yg kontinu

Asumsi dasar tes parametrik

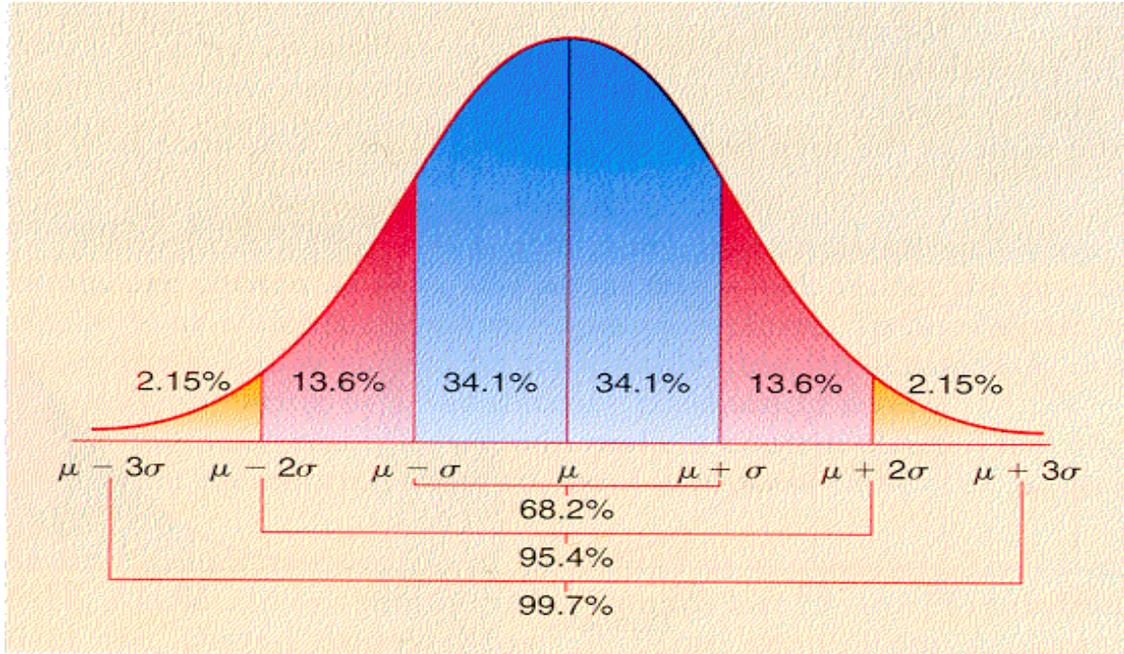
Data berupa interval atau rasio

Data berdistribusi normal

Sampel acak dan independen

Homogenitas varians

Distribusi normal



- Sebaran datanya akan menyerupai lonceng seperti gambar di samping
- Mean, median dan modus saling berhimpitan pada satu nilai
- Nilai-nilai pada dan dekat dengan rata-rata akan lebih sering muncul dibandingkan nilai yang lebih menjauhi rata-rata
- Semakin jauh suatu nilai dari rata-rata maka frekuensi keberadannya pada data semakin kecil.
- Standar deviasi (SD) menunjukkan sebaran data di sekitar mean.
- $\pm 68\%$ data akan berisi nilai-nilai di sekitar $\text{mean} \pm 1$ kali SD
- $\pm 95\%$ data akan berisi nilai-nilai di sekitar $\text{mean} \pm 2$ kali SD; dan
- $\pm 99\%$ data akan berisi nilai-nilai di sekitar $\text{mean} \pm 3$ kali SD

- Contoh: Mahasiswa Peternakan satu Angkatan terdiri dari 100 orang.
- Kita abaikan factor jenis kelamin.
- Kemudian diambil data tinggi badan. Rerata tinggi badan 160 cm dengan sd 10cm. Lalu Mhs diperintahkan untuk berbaris sesuai dengan aturan di slide sebelumnya.
- Maka mhs dengan TB 160 ± 10 cm ada di tengah (150-170cm).
- Kemudian yg tingginya 140-180cm di $\text{mean} \pm 2\text{sd}$, sisanya ada di $\text{mean} \pm 3\text{sd}$.
- Maka akan terbentuk barisan seperti gambar disamping.
- Fenomena ini berlaku pada banyak peristiwa yg terjadi di alam.



Lalu apa yang harus berdistribusi normal?

- Ada beberapa pendapat



Data yang akan diolah

2. Populasi dari data yang akan diolah – diasumsikan bahwa data yang tersedia merupakan sample dari populasi dengan distribusi normal
3. Errornya
4. Central limit theorem → distribusi mean dari sample-sample yang diambil dari populasi

- Cara cek normalitas:

- Histogram, QQ plot
- Kolmogorov-Smirnov test

Asumsi dasar tes parametrik

Data berupa interval atau rasio

Data berdistribusi normal

Sampel acak dan independen

Homogenitas varians

Sampel acak dan independen

- Tiap-tiap individu data di dalam sample tidak saling berhubungan
- Hal ini bisa disetting sejak peneliti mendesain penelitian
- Salah satu pentingnya pengacakan

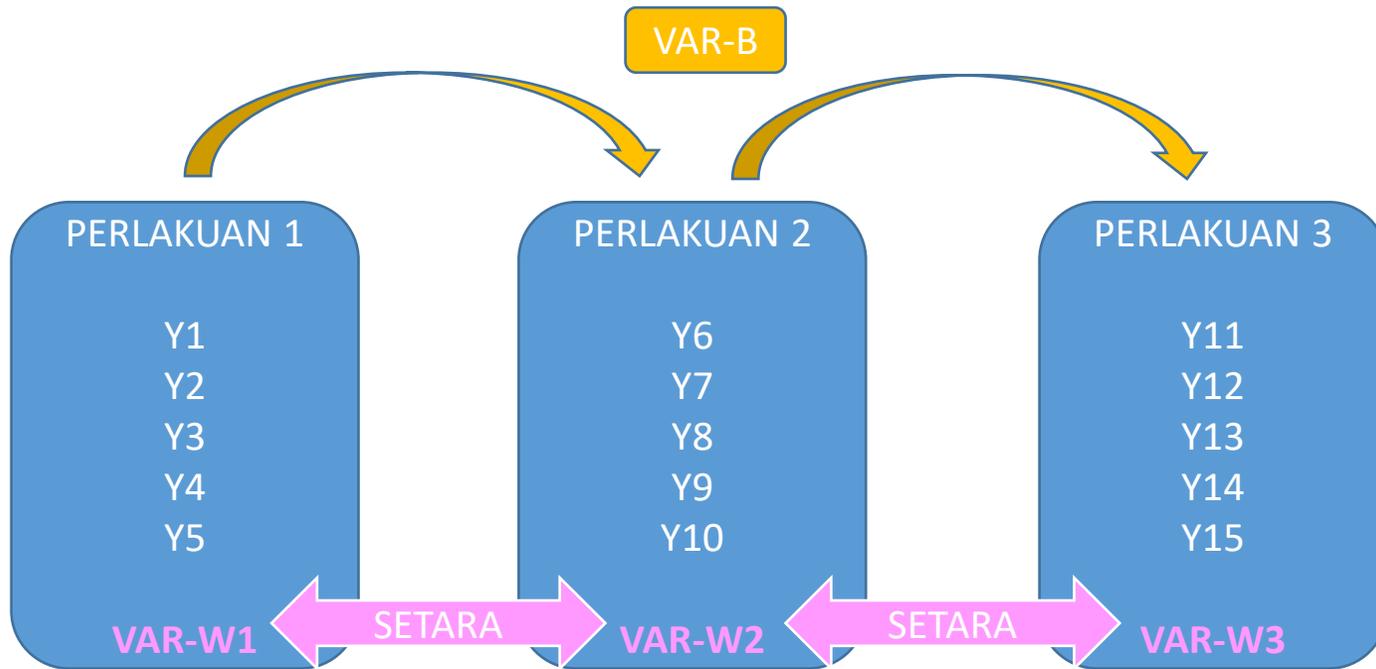
Asumsi dasar tes parametrik

Data berupa interval atau rasio

Data berdistribusi normal

Sampel acak dan independen

Homogenitas varians



Var-B = Variance between (antar) perlakuan
 Var-W = variance within (dalam) perlakuan

Cara cek kesetaraan varians adalah dengan Levene's test

- Pada saat kita mendesain eksperimen, diusahakan agar factor selain yang diberikan dikontrol seseragam mungkin
- Di dalam tiap perlakuan akan ada beberapa ulangan. Kondisi di dalam tiap perlakuan seragam. Kondisi antar perlakuan berbeda sesuai dengan perlakuannya
- Namun, setelah eksperimen dilakukan dan diperoleh data. Dalam perlakuan yang sama pasti ada variasi antar ulangan, disebut dengan Var-W atau varians error.
- Karena kondisi dalam tiap perlakuan dibuat seragam, maka variasi yang muncul dianggap berasal dari error.
- Dan besaran Var-W setara antar perlakuan

Thank you

