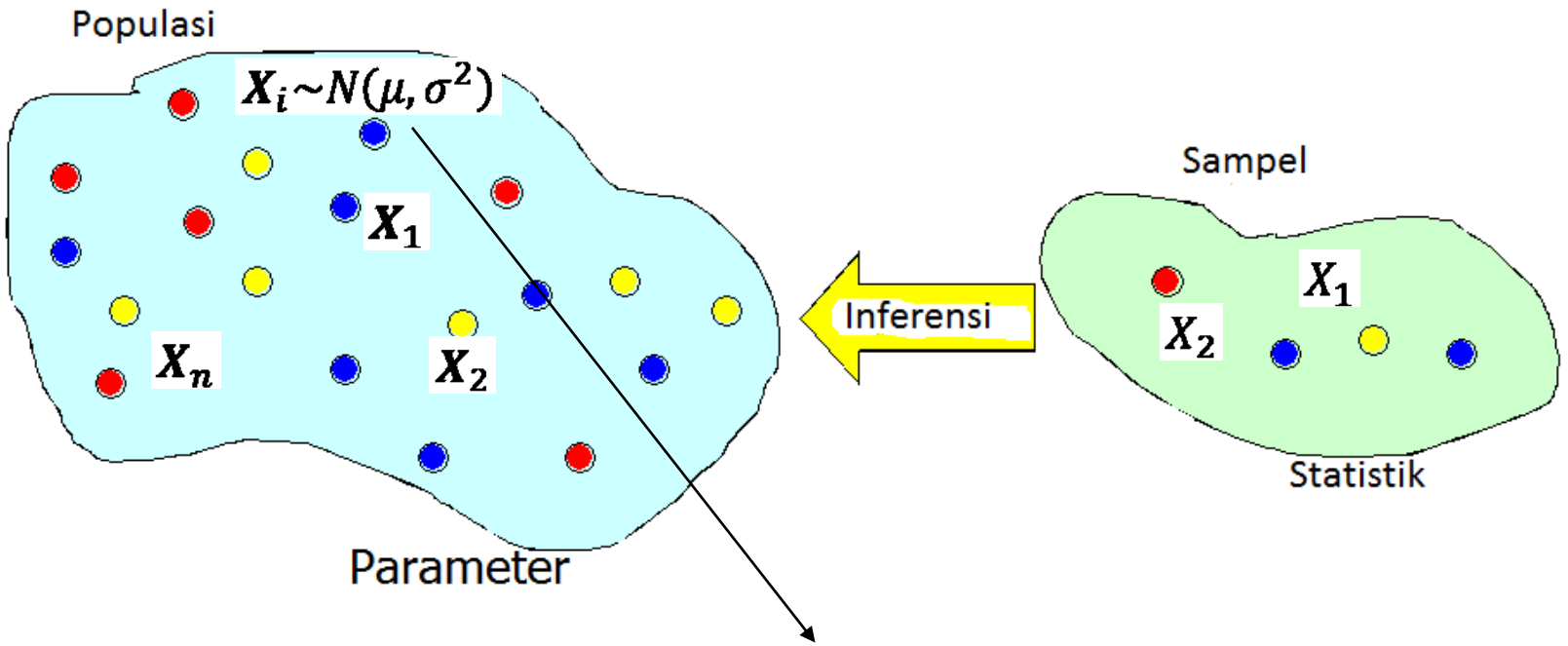
KD 1 PENGANTAR

STATISTIKA NONPARAMETRIK

Statistika Inferensi

- **Ingat kembali konsep tentang statistika parametrik! Misal populasi mengikuti distribusi Normal**



Sampel yang mengikuti populasi berdistribusi Normal $N(\mu, \sigma^2)$ maka harus :
→ Normal
→ Variansi Homogen (homoskedastisitas)

→ **Statistika parametrik**

Beda Statistika Parametrik vs Nonpar?

- Statistik Non-Parametrik,
- Statistik bebas sebaran (tidak mensyaratkan bentuk sebaran parameter populasi, baik normal atau tidak)
- Skala pengukuran sosial, yakni nominal dan ordinal
- Jumlah sampel biasanya kecil
- *Distribution free methods can be used for small samples when the central limit theorem may not apply*

Keunggulan statnonpar :

- Tidak membutuhkan asumsi normalitas.
- lebih mudah dikerjakan dan lebih mudah dimengerti
- dapat digantikan data numerik (nominal) dengan jenjang (ordinal).
- Pengujian hipotesis dilakukan secara langsung pada pengamatan
- Walaupun tidak terikat pada distribusi normal populasi, tetapi dapat digunakan pada populasi berdistribusi normal.

Kelemahan :

- mengabaikan beberapa informasi tertentu.
- Hasil pengujian hipotesis tidak setajam statistik parametrik.
- Hasil tidak dapat diekstrapolasikan ke populasi studi seperti pada statistik parametrik→ dikarenakan statistik non-parametrik mendekati eksperimen dengan sampel kecil dan umumnya membandingkan dua kelompok tertentu

Suatu Metode Statistika dikatakan nonparametrik jika memenuhi sedikitnya salah satu dari kriteria di bawah ini

- **Menggunakan skala pengukuran nominal**
- **Menggunakan skala pengukuran ordinal**
- **Metode menggunakan skala pengukuran interval atau rasio, dimana fungsi distribusi variabel acak yang menghasilkan data tidak diketahui (distribution-free) atau diketahui kecuali parameternya yang tidak diketahui (nonparametric)**

Contoh penggunaan Uji

Decision Table for Inferential Statistical Tests Employed with Interval/ Ratio Data

Number of samples One independent variable	Hypothesis evaluated	Test
Single sample	Hypothesis about a population mean	The single-sample *z* test (Test 1) (σ known) The single-sample *t* test (Test 2) (σ unknown)
	Hypothesis about a population parameter/characteristic other than the mean	The single-sample chi-square test for a population variance (Test 3) The single-sample test for evaluating population skewness (Test 4) The single-sample test for evaluating population kurtosis (Test 5) The mean square successive difference test (for serial randomness) (Test 10g) The D'Agostino–Pearson test of normality (Test 5a) Procedures for identifying outliers (Test 11f)

Cont..

Two samples	Two independent samples	Hypothesis about difference between two independent population means	The *t* test for two independent samples (Test 11) The *z* test for two independent samples (Test 11e) The single-factor between-subjects analysis of variance (Test 21) The single-factor between-subjects analysis of covariance (Test 21j)
		Hypothesis about two independent population variances	Hartley's F_{max} test for homogeneity of variance/ *F* test for two population variances (Test 11a)
	Two dependent samples	Hypothesis about difference between two dependent population means	The *t* test for two dependent samples (Test 17) Sandler's *A* test (Test 17d) The *z* test for two dependent samples (Test 17e) The single-factor within-subjects analysis of variance (Test 24)
		Hypothesis about two dependent population variances	The *t* test for homogeneity of variance for two dependent samples (Test 17a)

Cont…

Two or more samples	Two or more independent samples	Hypothesis about difference between two or more independent population means	The single-factor between-subjects analysis of variance (Test 21) The single-factor between-subjects analysis of covariance (Test 21j)
		Hypothesis about two or more independent population variances	Hartley's F_{max} test for homogeneity of variance/ F test for two population variances (Test 11a)
	Two or more dependent samples	Hypothesis about difference between two or more dependent population means	The single-factor within-subjects analysis of variance (Test 24)
		Hypothesis about two or more dependent population variances	See discussion of sphericity assumption under the single-factor within-subjects analysis of variance (Test 24)
Two independent variables		Hypothesis about difference between two or more population means	The between-subjects factorial analysis of variance (Test 27) The factorial analysis of variance for a mixed design (Test 27i) The within-subjects factorial analysis of variance (Test 27j)

Decision Table for Inferential Statistical Tests Employed with Ordinal/Rank-Order Data

Number of samples		Hypothesis evaluated	Test
Single sample		Hypothesis about a population median or the distribution of data in a single population	The Wilcoxon signed-ranks test (Test 6) The Kolmogorov–Smirnov goodness-of-fit test for a single sample (Test 7) The Lilliefors test for normality (Test 7a) The single-sample test for the median (Test 9b)
Two samples	Two independent samples	Hypothesis about two independent population medians, or some other characteristic (other than variability) of two independent populations	The Mann–Whitney *U* test (Test 12) The randomization test for two independent samples (Test12a) The bootstrap (Test 12b) (The bootstrap can also be employed for evaluating hypotheses concerning variability, as well as for evaluating various hypotheses for single sample designs and designs involving two or more independent or dependent samples.)* The jackknife (Test 12b) (The jackknife can also be employed for evaluating hypotheses concerning variability, as well as for evaluating various hypotheses for single sample designs and designs involving two or more independent or dependent samples.)* The Kolmogorov–Smirnov test for two independent samples (Test 13)

Cont..

	Two dependent samples	Hypothesis about the ordering of data in two dependent populations	The Wilcoxon matched-pairs signed-ranks test (Test 18) The binomial sign test for two dependent samples (Test 19)
Two or more samples	Two or more independent samples	Hypothesis about two or more independent population medians, or some other characteristic of two or more independent populations	The Kruskal–Wallis one-way analysis of variance by ranks (Test 22) The Jonckheere–Terpstra test for ordered alternatives (Test 22a) The van der Waerden normal-scores test for *k* independent samples (Test 23) The median test for independent samples (Test 16e)
	Two or more dependent samples	Hypothesis about two or more dependent population medians	The Friedman two-way analysis of variance by ranks (Test 25) The Page test for ordered alternatives (Test 25a)

Decision Table for Measures of Correlation/Association

Interval/ratio data	Bivariate	The Pearson product-moment correlation coefficient (Test 28)
	More than two sets of scores	The intraclass correlation coefficient (Test 24i)
	Multivariate	The multiple correlation coefficient (Test 28k) The partial correlation coefficient 28l) The semipartial correlation coefficient (Test 28m)
Ordinal/rank order data	Bivariate/two sets of ranks	Spearman's rank-order correlation coefficient (Test 29) Kendall's tau (Test 30) Goodman and Kruskal's gamma (for ordered contingency tables) (Test 32)
	More than two samples/sets of ranks	Kendall's coefficient of concordance (Test 31)

Categorical/ nominal data	Two dichotomous variables*	The contingency coefficient (Test 16f) The phi coefficient (Test 16g) Yule's *Q* (Test 16i) The odds ratio (Test 16j) Cohen's kappa (Test 16k) Binomial effect size display (Test 28r)
	Two nondichotomous variables	The contingency coefficient (Test 16f) Cramér's phi coefficient (Test 16h) The odds ratio (Test 16j) Cohen's kappa (Test 16k)
Other bivariate correlational measures for which interval/ratio data are employed or implied for at least one of the variables		Omega squared (One variable, interval/ratio data; second variable, two or more nominal levels) (Tests 11c/17c/21g/24g/27g) Eta squared (One variable, interval/ratio data; second variable, two or more nominal levels) (Test 11d (two nominal levels); Test 21h) Cohen's *d* index (and *g* index sample analogue) (Test 11b/17b) (One variable, interval/ratio data; second variable, two nominal levels) (with Test 2a for one variable) Cohen's *f* index (One variable, interval/ratio data; second variable, two or more nominal levels) (Test21i/24h/27h) The point-biserial correlation coefficient (One variable, interval/ratio data; second variable represented by dichotomous categories) (Test 28h) The biserial correlation coefficient (One variable, interval/ratio data; second variable, an interval/ratio variable expressed in form of dichotomous categories) (Test 28i) The tetrachoric correlation coefficient (Two interval/ratio variables, both of which are expressed in the form of dichotomous categories) (Test 28j)