ANALISIS PEUBAH GANDA

SEBARAN NORMAL GANDA

Dr. Ir. I Made Sumertajaya, M.Si

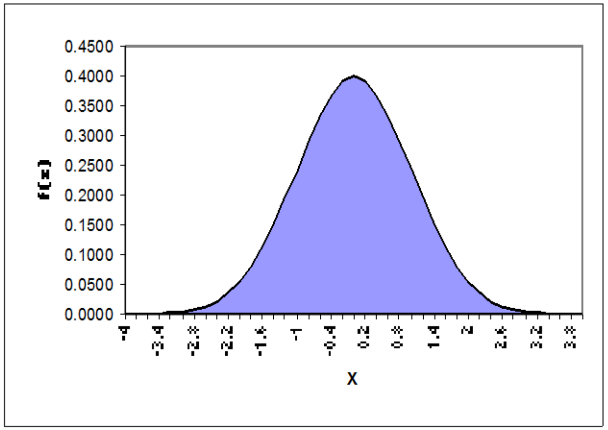
SEBARAN NORMAL GANDA

- Fungsi kepekatan normal ganda (*multivariate normal*) adalah generalisasi dari fungsi kepekatan *univariate* normal dengan p $\geq$ 2 dimensi.
- Fungsi kepekatan dari peubah acak x yang menyebar normal dengan nilai tengah μ dan ragam σ^2 adalah :

$$f(x) = \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-[(x-\mu)/\sigma]^2/2}$$

- **dimana -∞ < x < ∞**

- Plot dari fungsi di atas akan menghasilkan kurva berbentuk genta yang memiliki ciri-ciri sebagai berikut:
  - Simetrik terhadap nilai tengah (μ)
  - Mean, median, modus berada pada titik yang sama
  - P (μ - θ < x < μ + σ) = 0.683
  - P (μ - 2θ < x < μ + 2σ) = 0.954



Fungsi Kepekatan Peluang Normal Ganda

- Fungsi kepekatan bersama dari p peubah acak yang menyebar normal dan saling bebas adalah :

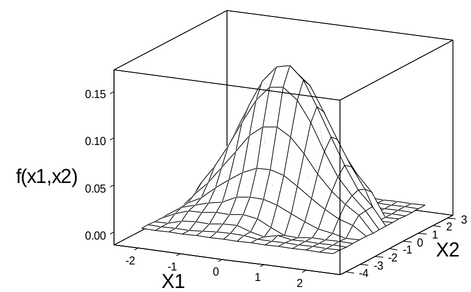
$$f(x_1,\dots,x_p) = \frac{1}{(2\pi)^{p/2}\sigma_1\dots\sigma_p} Exp\left[-1/2\sum_{i=1}^{p}\left(\frac{x_i-\mu_i}{\sigma_i}\right)\right]^2$$

- Bentuk $[\ (xi - \mu)/\sigma]^2$ dari eksponen fungsi sebaran normal mengukur jarak kuadrat dari x_i ke μ_i dalam unit simpangan baku.
- Bentuk ini dapat digeneralisasikan untuk vektor $\underline{\mathbf{x}}$ dari pengamatan beberapa peubah sebagai :

  $(\underline{\mathbf{x}} - \underline{\mu})' \Sigma^{-1} (\underline{\mathbf{x}} - \underline{\mu})$

- Secara umum fungsi kepekatan peluang normal bersama untuk p peubah dapat ditulis sebagai berikut :

$$f(\underline{x}) = \frac{1}{(2\pi)^{p/2}\left|\Sigma\right|^{1/2}} Exp\left[-1/2(\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})\right]$$



- dimana $-\infty < x_i < \infty$, i = 1,2,...p.
- Fungsi kepekatan normal berdimensi p ini dapat ditulis sebagai $N_p(\mu,\Sigma)$ yang analog dengan kasus *univariate*.

Countur
Sebaran Normal Ganda

- Eksponen $(x - \mu)' \Sigma^{-1} (x - \mu)$ dari kepekatan normal ganda memperlihatkan persamaan ellipsoid dalam ruang peubah berdimensi p jika bentuk ini ditulis dalam sebuah persamaan terhadap sebuah nilai konstanta positif c.

  $(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu) = c^2$
- Dengan pusat ellips adalah μ dan absis $\pm c \sqrt{\lambda_i}\, \mathbf{e}_i$ dimana,

  $\Sigma \mathbf{e}_i = \lambda_i \mathbf{e}_i$, $i = 1,2,...p$.

Sifat-sifat Sebaran Normal Ganda

- Kombinasi linier dari semua komponen peubah x juga menyebar normal. Jika $\mathbf{X}_p \sim N_p\,(\, \underline{\mu}\, , \Sigma\,)$, maka kombinasi linear :

  $\underline{\mathbf{a}}'\underline{\mathbf{X}} = a_1 X_1 + a_2 X_2 + \ldots + a_p X_p$ menyebar

  $N\,(\, \underline{\mathbf{a}}'\, \underline{\mu}\ , \underline{\mathbf{a}}'\, \Sigma\, \underline{\mathbf{a}}\,)$
- Jika $\underline{X}_p \sim N_p\ (\, \underline{\mu}\, , \Sigma\,)$ maka semua anak gugus dari X juga menyebar normal
- Jika $\underline{X}_1$ dan $\underline{X}_2$ saling bebas, dan menyebar $N_{q1}\,(\underline{\mu}_1, \Sigma_{11})$ dan $N_{q2}\,(\mu_2, \Sigma_{22})$ maka sebaran bersyarat $[\underline{X}_1 / \underline{X}_2]$ adalah normal ganda :

$$N_{q1+q2}\left(\left[\frac{\mu_1}{\mu_2}\right],\left[\begin{matrix}\Sigma_{11} & 0 \\ \hline 0 & \Sigma_{22}\end{matrix}\right]\right)$$

- Jika $X \sim N_p\ (\mu, \Sigma)$ dengan $|\Sigma| > 0$ maka:
  - $(x - \mu)'\, \Sigma^{-1}\, (x - \mu) \sim \chi_{(p)}^2$ dimana χ_p^2 menyatakan sebaran khi kuadrat dengan derajat bebas p.
  - Selang kepercayaan $(1-\alpha)100\% \rightarrow$

    $$(\underline{\mathbf{x}}-\underline{\boldsymbol{\mu}})'\Sigma^{-1}(\underline{\mathbf{x}}-\underline{\boldsymbol{\mu}}) \leq \chi_{(\alpha,\, p)}^2$$

Eksplorasi Sebaran Normal Ganda

- **Untuk mengevaluasi apakah data yang dimiliki menyebar normal ganda dapat ditelusuri secara ekplorasi**
- **Seperti halnya untuk kasus *univariate* penelusuran sebaran normal ganda dapat juga memanfaatkan plot quantil-quantil → quantil khi-kuadrat**

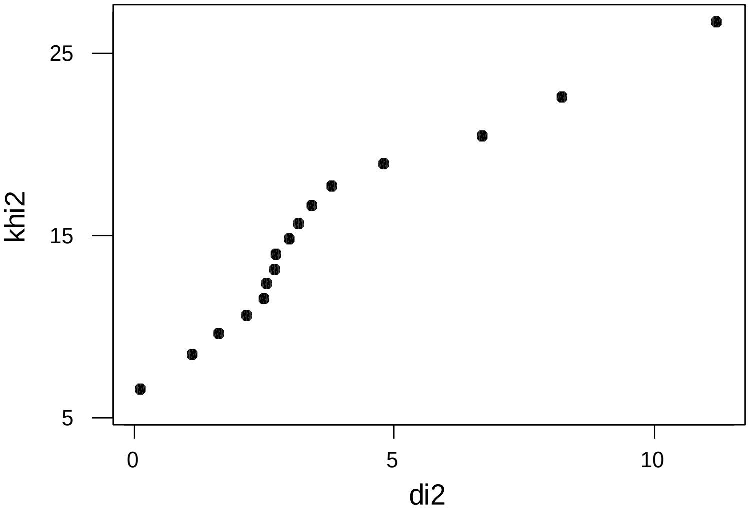
Beberapa tahapan yang harus dilakukan dalam menyusun Plot Kuantil χ^2 adalah sebagai berikut:

1. Hitung: $d_i^2 = (\underline{x}_i - \underline{\mu})' \Sigma^{-1} (\underline{x}_i - \underline{\mu})$
2. Beri peringkat k untuk nilai d_i^2
3. Carilah nilai khi-kuadrat dari nilai (k–1/2)/n dengan derajat bebas p, misal $\chi_p^2\left(\frac{k - \frac{1}{2}}{n}\right)$
4. Buat plot d_i^2 dengan $\chi_p^2\left(\frac{k - \frac{1}{2}}{n}\right)$
5. Jika plot tersebut membentuk garis lurus maka data tersebut menyebar normal ganda p.

Contoh kasus

Obs	x1	x2	x3	x4	d_i^2	k	p(k)	$\chi_{(4)}^2$
1	1.29	0.98	0.99	1.05	8.25	15	0.91	22.57
2	1.28	1.08	1.06	1.08	6.71	14	0.84	20.42
3	0.79	1.06	1.01	1.05	11.22	16	0.97	26.70
4	0.80	1.01	1.01	1.05	2.55	6	0.34	12.30
5	1.39	1.03	1.03	1.04	2.19	4	0.22	10.59
6	0.89	0.97	0.99	1.02	2.72	7	0.41	13.11
7	1.40	1.06	1.05	1.06	2.74	8	0.47	13.92
8	0.72	1.00	1.02	1.03	3.81	12	0.72	17.65
9	1.03	1.00	1.00	1.01	1.13	2	0.09	8.41
10	0.69	0.97	0.99	1.01	3.00	9	0.53	14.76
11	1.29	1.05	1.03	1.04	1.64	3	0.16	9.61
12	1.17	1.00	1.00	1.01	3.43	11	0.66	16.59
13	0.82	1.00	1.01	1.01	2.50	5	0.28	11.47
14	1.23	1.05	1.03	1.02	3.17	10	0.59	15.64
15	1.09	1.02	1.02	1.03	0.13	1	0.03	6.56
16	1.00	1.04	1.05	1.03	4.82	13	0.78	18.89

- Korelasi antara d_i^2 dengan kuantil khi-kuadrat adalah 0.941
- Bandingkan dengan tabel, jika lebih besar dari tabel berarti normal ganda



Penanganan Masalah

- Jika data tidak normal maka dapat didekati dengan transformasi, misal transformasi Box-Cox
- Lakukan eksplorasi pengamatan-pengamatan ekstrem, buat box-plot untuk data d_i^2 (aturan main sama seperti kasus univariate