

MODEL DISTRIBUSI WEIBULL

DALAM APLIKASI ANALISIS DATA UJI HIDUP

Data durasi seperti data interval kelahiran dan lamanya waktu hingga terjadinya kematian biasanya berkaitan dengan kelangsungan hidup (survival) suatu objek yang sedang diamati. Data survival dapat digunakan dalam 2 prosedur statistik, yaitu : prosedur non parametrik dan prosedur parametrik. Prosedur non-parametrik biasanya hanya membahas tentang individu dalam kelompok dan waktu survivalnya saja, tapi tidak membahas peubah-peubah yang mempengaruhinya. Ada 2 prosedur yang lazim digunakan adalah Log-Rank Test dan Wilcoxon Test. Karena prosedur non parametrik hanya membahas tentang individu dalam kelompok dan waktu survivalnya saja, maka kita memerlukan suatu pendekatan berdasarkan model statistic yang mampu menjelaskan variable-variable yang berhubungan atau yang mempengaruhi individu dalam waktu survival. Biasanya, model yang digunakan adalah model proportional hazard. Sering kali data survival tidak sesuai dengan prosedur statistic standar yang digunakan dalam analisis data, karena data survival secara umum tidak berdistribusi secara simetris, sehingga sebagai konsekuensinya tidak beralasan untuk mengasumsikan bahwa data survival berdistribusi normal. Kesulitan ini dapat dipecahan dengan mentransformasi data untuk memberikan distribusi yang lebih simetris. Jika asumsi dari sebuah distribusi peluang yang khusus untuk datanya adalah valid, kesimpulan berdasarkan asumsi akan lebih tepat. Prosedur yang sebaiknya digunakan adalah parametric, biasanya model eksponensial dan model weibull.

Dalam prakteknya asumsi dari sebuah fungsi hazard yang konstan atau waktu survivalnya berdistribusi eksponensial jarang bisa dipertahankan karena fungsi hazard tidak selalu konstan, sehingga dalam makalah ini kita akan membahas model weibull, dimana distribusi weibull dapat digunakan jika dalam suatu data survival fungsi hazard meningkat atau menurun secara monoton dengan peningkatan waktu survival. Karena distribusi weibull miring atau tidak simetris maka hasil yang lebih tepat dan dapat dipertahankan untuk lokasi dari distribusinya adalah median masa hidup.

1. Model fungsi hazard

Misalkan T adalah variable acak waktu ketahanan hidup yang mempunyai fungsi densitas peluang $f(t)$, dan fungsi distribusinya sebagai berikut :

$$F(t) = P(T < t) = \int_0^t f(u)du$$

Dimana peluang $P(T < t)$ = peluang hidup sampai dengan t. Fungsi ketahanan hidup dirumuskan dengan :

$$\begin{aligned} S(t) &= P(T \geq t) = 1 - P(T < t) \\ &= 1 - F(t) \end{aligned}$$

dimana $P(T \geq t)$ = peluang hidup melebihi t.

Hal terpenting yang berkaitan dengan analisis data ketahanan hidup adalah laju kematian, dimana laju kematian tersebut diakomodir dalam fungsi hazard, yaitu peluang bahwa suatu individu mati pada saat t, dengan syarat bahwa ia dapat bertahan sampai dengan waktu t tersebut. Fungsi hazard ini menyatakan peluang mati sesaat (laju kerusakan sesaat) suatu individu yang bertahan sampai dengan waktu t tersebut.

Pada saat sebuah model distribusi untuk masa hidup telah ditentukan dalam bentuk fungsi densitas peluang, fungsi survivor dan fungsi hazard yang sesuai dapat dihasilkan dari hubungan berikut :

$$S(t) = 1 - \int_0^t f(u)du$$

dan

$$h(t) = \frac{f(t)}{S(t)} = \frac{d}{dt} [\ln[S(t)]]$$

dimana $f(t)$ adalah fungsi densitas peluang dari masa hidupnya. Sebuah pendekatan alternative untuk menentukan sebuah bentuk fungsional bagi fungsi hazard, yang darinya fungsi survivor dan fungsi densitas peluangnya dapat ditentukan dari persamaan berikut :

$$S(t) = \exp\{-H(t)\}$$

dan

$$f(t) = h(t)S(t) = - \frac{dS(t)}{dt}$$

dimana

$$H(t) = \int_0^t h(u)du$$

Bentuk $H(t) = \int_0^t h(u)du$ disebut Kumulatif Hazard Rate.

Variable lain yang mempengaruhi analisis data ketahanan hidup disebut dengan kovariat.

$H_s(t)$: hazard bagi individu kelompok standar pada waktu t

$H_b(t)$: hazard bagi individu kelompok metode baru waktu t

$$H_b(t) = \Psi h_s(t) \quad ; \Psi = \text{konstanta}$$

dimana jika $\Psi < 1$ maka $h_b(t) < h_s(t)$ dan jika $\Psi > 1$ maka $h_b(t) > h_s(t)$.