

ABSTRAK

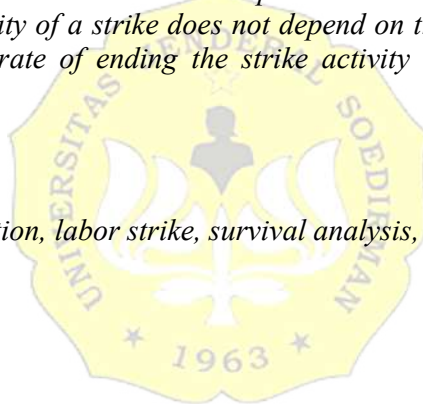
Analisis survival adalah metode analisis yang menghasilkan fungsi survival dan fungsi *hazard* dengan menggunakan data lamanya waktu. Satuan waktu tersebut dimulai dari suatu titik awal sampai dengan terjadinya suatu 'kejadian' atau titik akhir. Kejadian yang diperhatikan tidak selalu sesuatu yang '*terminate*' yaitu kejadian yang sekali saja terjadi dan berhenti, misalnya kematian. Titik akhir atau kejadian bisa juga berupa status yang lebih umum, misalnya pemogokan. Pemogokan tidak termasuk dalam kasus yang bersifat *terminate*, karena setelah pemogokan berlangsung bisa saja individu tersebut melakukan pemogokan yang berulang. Pemogokan adalah tindakan sekelompok karyawan yang direncanakan dan dilaksanakan secara bersama untuk menyampaikan tuntutan atau aspirasi. Pemogokan berakibat pada penghentian atau perlambatan pekerjaan. Analisis survival yang diterapkan pada data pemogokan bertujuan untuk melihat peluang berakhirnya mogok kerja bergantung atau tidak bergantung pada lama waktu (durasi) pemogokan. Hasil dari fungsi *hazard* yang diperoleh dengan metode non-parametrik menunjukkan adanya ketergantungan durasi pemogokan. Fungsi *hazard* yang diperoleh dengan metode parametrik dengan mengasumsikan data distribusi eksponensial menghasilkan fungsi *hazard* yang konstan. Oleh karena itu, peluang mogok kerja yang terjadi tidak ada ketergantungan terhadap durasi pemogokan, yang berarti laju berakhirnya aktivitas mogok kerja tidak bergantung kepada durasi pemogokan.

Kata kunci: analisis survival, fungsi *hazard*, fungsi survival, pemogokan buruh

ABSTRACT

Survival analysis is an analytical method that produces a survival function and a hazard function using data for the length of time. The unit of time starts from a clear starting point until the occurrence of an 'event' or end point. The event or end point that is considered is not always something that is 'terminated', namely an event that only happens once and stops, for example, death. The end point or event can also be a more general state, such as a strike. Strikes are not included in cases that are terminated, because after the strike, the individual does not end (die) and the individual may carry out repeated strikes. A strike is an action by a group of employees that is planned and carried out jointly to convey demands or aspirations. Strikes result in the termination or deceleration of work. Survival analysis applied to strike data aims to see the probability of ending the strike depending or not depending on the length of time (duration) of the strike. The results of the hazard function obtained by the non-parametric method show that there is a dependence of the duration value. The hazard function obtained from the parametric method by assuming the exponential distribution data produces a constant hazard function. Therefore, the probability of a strike does not depend on the duration of the strike, which means that the rate of ending the strike activity does not depend on the duration of the strike.

Keywords: *hazard function, labor strike, survival analysis, survival function*



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Menurut Miranti (2017: 7), analisis survival atau analisis ketahanan adalah suatu metode yang berkaitan dengan waktu dari suatu objek yang dimulai dari waktu asal (*time origin*) atau titik awal (*start point*) sampai terjadinya kejadian tertentu yang telah ditetapkan sebagai peristiwa kegagalan (*failure event*) atau titik akhir (*end point*). Pada analisis survival diasumsikan hanya ada satu peristiwa kegagalan (*failure event*) yang menjadi amatan, meskipun sebenarnya bisa saja terjadi lebih dari satu *failure event* dalam penelitian yang sama. *Failure event* yang populer adalah kesembuhan dari suatu penyakit dan kematian akibat suatu penyakit.

Salah satu tujuan dasar analisis survival menurut Kleinbaum dan Klein (2012: 16), adalah untuk mengestimasi fungsi survival dan/atau fungsi *hazard* dari data survival. Fungsi survival yaitu peluang bertahan setelah suatu waktu tertentu dan *hazard rate* merupakan laju kegagalan *event*. Menurut Kiefer (1988: 648), penerapan analisis survival yang aktual dan potensial dalam bidang ekonomi misalnya, waktu antara perdagangan di pasar keuangan, daya tahan produk, waktu antar perpindahan pekerjaan, masa hidup perusahaan, waktu untuk penemuan dari investasi penelitian, periode pengembalian pinjaman luar negeri, jarak pembelian barang tahan lama (atau modal pengganti) dan durasi lamanya pemogokan kerja.

Pemogokan adalah tindakan karyawan yang direncanakan dan dilaksanakan secara bersama-sama oleh sekelompok karyawan untuk menghentikan atau memperlambat pekerjaan. Mengenai pemogokan karyawan, Kennan (1985: 5) memberikan pendapatnya bahwa studi empiris tentang fluktuasi siklus dalam aktifitas pemogokan memiliki sejarah yang panjang dan pemogokan bersifat prosiklis yang dapat diterima secara luas di kalangan ekonom tenaga kerja. Bukti dasar atas kedua kesimpulan empiris tersebut diberikan oleh Rees (Kennan, 1985: 5) yang menerapkan alat siklus bisnis *Mitchell-Bums-NBER* pada data bulanan durasi pemogokan buruh yang terjadi selama periode pemogokan. Menurut Rees (1952: 371), tujuan menerapkan alat siklus bisnis *Mitchell-Bums-NBER* untuk

mengidentifikasi pola keteraturan pemogokan buruh dengan siklus bisnis. Hal ini penting diidentifikasi karena pada masa itu di Amerika Serikat pemogokan buruh sering menjadi subjek undang-undang atau tindakan pemerintah lainnya. Kesalahan dalam diagnosis pola keteraturan pemogokan buruh dapat menyebabkan kesalahan dalam kebijakan publik.

Hasil yang diperoleh Rees (Kennan, 1985: 5) menunjukkan bahwa hubungan antara pemogokan individu dan kegiatan ekonomi tampak tidak teratur untuk siklus bisnis, artinya pemogokan memang tidak mengikuti siklus bisnis yang ada. Temuan empiris ini memiliki pengaruh besar pada perkembangan teori ekonomi pemogokan. Jika bukan karena temuan ini, para ahli ekonomi pemogokan cenderung akan terus setuju dengan pendapat Hicks (Kennan, 1985: 5) bahwa pemogokan disebabkan oleh perilaku irasional atau keyakinan yang salah dan dibiarkan saja.

Kiefer (1988) menganalisis fungsi *hazard* pada data pemogokan bulan Juni di industri manufaktur Amerika Serikat tahun 1968 sampai dengan tahun 1976 menggunakan metode non-parametrik, metode parametrik dan dengan menambahkan variabel penjelas pada model *proportional hazard*. Pada metode non-parametrik diasumsikan bahwa distribusi waktu survival tidak mengikuti suatu distribusi tertentu. Sementara itu, pada metode parametrik diasumsikan bahwa distribusi waktu survival mengikuti suatu distribusi tertentu. Tujuan yang ingin dicapai dengan meneliti data pemogokan buruh adalah menganalisis fungsi *hazard* untuk melihat ketergantungan durasi pemogokan. Pemogokan disebabkan adanya hubungan antara pemogokan individu dengan kegiatan ekonomi bukan disebabkan oleh keyakinan yang salah, maka ketergantungan durasi pemogokan dapat diketahui.

Metode non-parametrik pada data bulanan durasi pemogokan buruh dengan adanya data tersensor (data yang tidak rampung sampai penelitian berakhir) sedikit berbeda dengan analisis pada umumnya sehingga perlu adanya modifikasi. London (1988) mengkaji ulang bahasan mengenai durasi pemogokan pada bukunya dalam Bab 11 yang sebagian besar didasarkan pada Kiefer (1988).

Analisis survival untuk durasi pemogokan pada data tersensor dapat dikerjakan dengan menggunakan metode non-parametrik dan metode parametrik. Pada metode non-parametrik untuk data durasi pemogokan sebagian besar didasarkan pada artikel yang ditulis oleh Kiefer (1988) sementara metode parametrik sebagian besar didasarkan pada London (1988). Penulis tertarik untuk mengkaji ulang bahasan ini karena merupakan suatu pembahasan yang jarang dilakukan oleh banyak orang.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, rumusan masalah dalam tugas akhir ini adalah:

1. Bagaimana deskripsi data pada pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976?
2. Bagaimana penerapan fungsi *hazard* dan fungsi survival dengan metode non-parametrik pada data pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976?
3. Bagaimana penerapan fungsi *hazard* dan fungsi survival dengan metode parametrik pada data pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. data yang digunakan merupakan data sekunder, yaitu data pemogokan di industri manufaktur Amerika Serikat setiap bulan Juni pada tahun 1968 sampai 1976;
2. metode parametrik hanya berdasarkan distribusi eksponensial.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

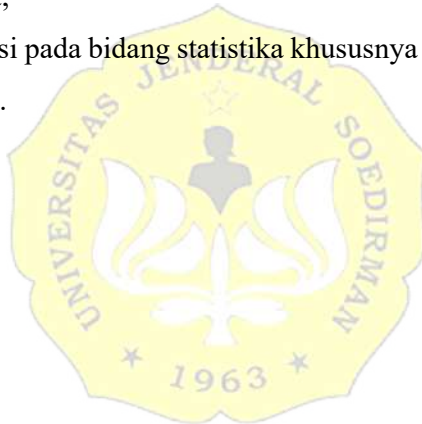
1. mendeskripsikan data pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976;

2. menerapkan fungsi *hazard* dan fungsi survival dengan metode non-parametrik dan pada data pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976; dan
3. menerapkan fungsi *hazard* dan fungsi survival dengan metode parametrik pada data pemogokan bulan Juni di industri manufaktur Amerika Serikat pada tahun 1968 sampai 1976.

1.5 Manfaat Penelitian

Berdasarkan tujuan penelitian, manfaat dari penelitian ini adalah:

1. menambah pengetahuan mengenai analisis survival dalam bidang ekonomi sehingga dapat digunakan sebagai acuan untuk membuat penelitian bagi mahasiswa lainnya;
2. menambah referensi pada bidang statistika khususnya tentang analisis survival di bidang ekonomi.



BAB 2

TINJAUAN PUSTAKA

Pada bab ini akan dibahas tentang pengertian durasi pemogokan, analisis survival, fungsi survival dan fungsi *hazard*, tipe penyensoran data, dan analisis survival pada data pemogokan dengan metode non-parametrik dan parametrik.

2.1 Durasi Pemogokan

Menurut Bureau of Labor Statistics (2021), pemogokan adalah penghentian sementara pekerjaan oleh sekelompok pekerja (tidak harus anggota serikat pekerja) untuk mengungkapkan keluhan atau menegakkan tuntutan. Pemogokan diprakarsai oleh para pekerja suatu perusahaan. Informasi tentang pemogokan yang diperoleh dari Bureau of Labor Statistics (2021), merupakan laporan dari Layanan Mediasi dan Konsiliasi Federal (*Federal Mediation and Conciliation Service*), merupakan data yang akan dianalisis pada penelitian ini. Pihak-pihak yang terlibat dalam penghentian kerja (manager, asosiasi, serikat pekerja) dihubungi untuk memverifikasi durasi dan jumlah pekerja yang menganggur karena penghentian.

Durasi pemogokan menyatakan lamanya waktu yang dihabiskan oleh para pekerja dengan tidak masuk kerja pada hari kerja, karena memiliki masalah atau tuntutan kepada pihak perusahaan tempat bekerja. Menurut Bureau of Labor Statistics (2019: 2), perhitungan pemogokan dimulai sejak tanggal yang sudah direncanakan secara resmi oleh para pekerja sampai dengan tanggal berakhirnya pemogokan berdasarkan kesepakatan awal (sementara) oleh kedua belah pihak. Tanggal berakhirnya mungkin bukan tanggal sebenarnya yang telah disepakati oleh para pekerja dengan perusahaan untuk melakukan mogok kerja, atau tanggal pekerja kembali bekerja. Tanggal berakhirnya mungkin lebih lambat dari tanggal kesepakatan awal. Selain itu, jika pemogokan terjadi pada akhir pekan, hari libur federal atau hari kerja dengan pengaturan *shift* yang tidak penuh, maka tidak dihitung sebagai hari pemogokan.

2.2 Analisis Survival

Analisis survival sudah banyak diterapkan dalam berbagai bidang. Sebagai contoh, menurut Gayatri (2005: 36), ahli demografi menggunakan analisis survival untuk mengukur dan menganalisis angka mortalitas, ahli asuransi menggunakannya untuk menghitung premi yang harus dibayar peserta asuransi, ahli kedokteran menggunakannya untuk menghitung efektivitas pengobatan atau memperkirakan lama hidup seorang pasien ketika diagnosis diberikan, ahli keperawatan menggunakan analisis ini untuk mengukur kemungkinan seseorang berisiko terkena *phelebitis* sejak mendapatkan terapi intra vena atau mengukur lamanya waktu yang dibutuhkan untuk perawatan luka dalam kondisi tertentu hingga sembuh.

Menurut Kleinbaum dan Klein (2012: 4), analisis survival adalah kumpulan prosedur statistika dengan variabel *outcome* yang diteliti adalah waktu (*time*) sampai suatu peristiwa terjadi. Menurut Harlan (2017: 1), analisis survival (analisis kesintasan) adalah prosedur statistika untuk menganalisis data dengan waktu sampai terjadinya suatu peristiwa tertentu (*time until an event occurs*) sebagai variabel respon. Untuk istilah waktu, satuan yang digunakan dapat berupa tahun, bulan, minggu, atau hari sejak awal pengamatan terhadap individu sampai suatu peristiwa terjadi; atau, waktu dapat merujuk pada usia individu ketika suatu peristiwa terjadi.

Tujuan dasar analisis survival menurut Kleinbaum dan Klein (2012: 16), adalah sebagai berikut:

1. untuk mengestimasi dan menginterpretasikan fungsi survival dan/ atau fungsi *hazard* dari data survival;
2. untuk membandingkan fungsi survival dan fungsi *hazard*; dan
3. untuk menilai hubungan variabel *explanatory* dengan waktu survival.

Secara umum metode untuk mengestimasi fungsi survival dan fungsi *hazard* dibagi menjadi tiga metode yaitu metode non-parametrik, semi parametrik dan parametrik. Untuk metode non-parametrik, diasumsikan bahwa distribusi yang mendasari waktu survival tidak mengikuti suatu distribusi tertentu. Pada metode non-parametrik terdapat dua cara untuk menghasilkan fungsi survival, yaitu metode Tabel Kehidupan (*Life Table*) dan metode *Product Limit Kaplan-Meier*. Metode

semi parametrik dapat digunakan untuk tujuan dasar analisis survival ketiga, yaitu untuk menilai hubungan variabel *explanatory* dengan waktu survival. Pada tujuan ini biasanya membutuhkan penggunaan beberapa bentuk pemodelan matematika seperti *Cox Proportional Hazard*. Sedangkan pada metode parametrik diasumsikan bahwa distribusi yang mendasari waktu survival mengikuti suatu distribusi tertentu.

2.3 Fungsi Survival dan Fungsi Hazard

Misalkan variabel acak non negatif T menyatakan waktu ketahanan hidup atau waktu hingga terjadinya suatu peristiwa. Variabel acak T mempunyai fungsi distribusi peluang dengan fungsi densitas $f(t)$, sehingga fungsi distribusi kumulatif bagi T , adalah peluang yang menyatakan waktu ketahanan hidup bernilai lebih kecil dari t , didefinisikan sebagai berikut (Hosmer dkk., 2008: 16):

$$F(t) = P(T \leq t). \quad (2.1)$$

Jika T merupakan variabel acak kontinu, maka persamaan (2.1) didefinisikan sebagai

$$F(t) = \int_0^t f(u) du.$$

Fungsi survival, dilambangkan dengan $S(t)$, menyatakan peluang bahwa individu dapat bertahan melebihi waktu yang ditentukan. Menurut Hosmer dkk. (2008: 16), fungsi survival didefinisikan sebagai:

$$S(t) = P(T > t) = 1 - P(T \leq t).$$

Karakteristik fungsi survival atau $S(t)$ menurut Harlan (2017: 3), antara lain:

1. monoton tak naik (*non increasing*), artinya fungsi survival $S(t)$ akan mengecil atau konstan dengan bertambahnya nilai t ;
2. pada waktu $t = 0$, $S(t) = S(0) = 1$, artinya pada awal studi belum ada subjek yang mengalami kegagalan atau $P(T > 0) = 1$; dan
3. pada waktu $t \rightarrow \infty$, $\lim_{t \rightarrow \infty} S(t) = 0$, yaitu jika secara teoritis periode studi diperpanjang tanpa batas, suatu saat tidak ada lagi subjek yang *survive/* bertahan.

Menurut Kleinbaum dan Klein (2012: 11), fungsi *hazard*, dilambangkan dengan $h(t)$ menggambarkan potensi (kegagalan) sesaat per satuan waktu untuk terjadinya peristiwa pada individu, dengan syarat bahwa individu bertahan hingga waktu t . Menurut Harlan (2017: 2), fungsi *hazard* bukan peluang, melainkan berupa *rate* (kelajuan) sehingga fungsi *hazard* disebut juga sebagai *hazard rate*. Dengan demikian, *hazard rate* dimungkinkan memiliki nilai yang lebih besar dari 1. Batasan yang digunakan pada fungsi *hazard* hanyalah $h(t) \geq 0$. Fungsi *hazard* didefinisikan sebagai:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T \geq t)}{\Delta t}$$

Penjabaran antara hubungan fungsi survival dan fungsi hazard adalah sebagai berikut:

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T \geq t)}{\Delta t} \\ h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{Pr(T \geq t) \Delta t} \\ h(t) &= \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{S(t) \Delta t} \\ h(t) &= \lim_{\Delta t \rightarrow 0} \frac{[1 - F(t)] - [1 - F(t + \Delta t)]}{S(t) \Delta t} \\ h(t) &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \cdot \frac{1}{S(t)} \\ h(t) &= \frac{f(t)}{S(t)}. \end{aligned} \tag{2.2}$$

Fungsi densitas $f(t)$ adalah turunan pertama $F(t)$ terhadap t , yaitu:

$$f(t) = \frac{dF(t)}{dt} = \frac{d(1 - S(t))}{dt}$$

Fungsi $h(t)$ pada persamaan (2.2) dapat dinyatakan dalam persamaan sebagai berikut (Miranti, 2017):

$$h(t) = \frac{\left(\frac{d(1 - S(t))}{dt} \right)}{S(t)}$$

$$\begin{aligned}
 h(t) &= \frac{-\frac{d(S(t))}{dt}}{S(t)} \\
 h(t) &= -\frac{d(S(t))}{dt} \cdot \frac{d \ln(S(t))}{dS(t)} \\
 h(t) &= -\frac{d \ln(S(t))}{dt}. \tag{2.3}
 \end{aligned}$$

Apabila kedua ruas pada persamaan (2.3) diintegrasikan akan diperoleh hubungan fungsi $h(t)$ dan fungsi $S(t)$ sebagai berikut (London, 1988: 16):

$$\int_0^t h(u) du = -\ln S(t),$$

sehingga fungsi *hazard* kumulatif didefinisikan sebagai:

$$H(t) = -\ln S(t). \tag{2.4}$$

Interpretasi fungsi *hazard* kumulatif menurut Harlan (2017: 2), yaitu:

1. Ukuran risiko total yang terakumulasi sampai dengan waktu t , dan
2. Risiko kumulatif dan peluang survival berbanding terbalik.

Sementara itu, fungsi survival kumulatif adalah:

$$S(t) = \exp(-H(t)).$$

2.4 Tipe Penyensoran Data

Menurut penjelasan Harlan (2017: 4), data tersensor terbagi menjadi 3, yaitu:

1. Tersensor kanan (*right-censored*) yang mungkin terjadi pada salah satu keadaan berikut:
 - a. Pada saat studi berakhir subjek belum mengalami kegagalan.
 - b. Subjek mengundurkan diri dari penelitian dan tidak dapat diamati lebih lanjut (*lost to follow-up, withdrawal*).
 - c. Subjek mengalami kegagalan lain yang menyebabkan pengamatan tidak dapat diteruskan (*competing risk*).
2. Tersensor interval (*interval-censored*) jika pengamatan tidak dilakukan secara kontinu, melainkan setiap akhir interval tertentu. Kegagalan dapat terjadi di tengah suatu interval tanpa diketahui secara tepat waktu survivalnya.

3. Tersensor kiri (*left-censored*) jika kegagalan telah terjadi sebelum subjek memasuki studi, tetapi tidak diketahui secara tepat saat terjadinya.

2.5 Analisis Survival untuk Data Pemogokan Metode Non-parametrik

Sifat khas dari analisis survival adalah data yang dianalisis menunjukkan lamanya waktu. Satuan waktu tersebut dimulai dari suatu titik awal yang jelas sampai dengan terjadinya suatu ‘kejadian’ atau titik akhir. Kejadian atau titik akhir yang diperhatikan tidak selalu sesuatu yang ‘*terminate*’ yaitu kejadian yang sekali saja terjadi dan berhenti, misalnya kematian. Titik akhir atau kejadian bisa juga berupa status yang lebih umum, misalnya pemogokan. Pemogokan tidak termasuk dalam kasus yang bersifat *terminate*, karena setelah pemogokan berlangsung individu tersebut tidak berakhir (mati) dan bisa saja individu tersebut melakukan pemogokan yang berulang. Tabel 2.1 merupakan gambaran dari struktur data untuk data pemogokan:

Tabel 2.1 Struktur data pemogokan

Waktu	n	Durasi				
Bulan/Tahun	n_1	t_{11}	t_{12}	t_{13}	...	t_{1j}
Bulan/Tahun	n_2	t_{21}	t_{22}	t_{23}	...	t_{2j}
Bulan/Tahun	n_3	t_{31}	t_{32}	t_{33}	...	t_{3j}
:	:	:	:	:	...	:
Bulan/Tahun	n_i	t_{i1}	t_{i2}	t_{i3}	...	t_{ij}

Pada Tabel 2.1 waktu menunjukkan bulan dan tahun terjadinya durasi pemogokan, n menunjukkan jumlah individu yang melakukan pemogokan pada waktu tersebut, sehingga banyaknya n akan sama dengan baris waktunya. Kemudian durasi menunjukkan lamanya setiap individu melakukan pemogokan, dengan durasi pemogokan yang satu dengan yang lainnya bisa memiliki nilai yang sama ataupun tidak. Tidak ada batasan maksimal atau minimal untuk nilai n .

Menurut Kiefer (1988: 657), taksiran fungsi survival untuk sampel n pengamatan tanpa sensor adalah sebagai berikut :

$$S(t) = n^{(-1)}.$$

Modifikasi diperlukan untuk memungkinkan terjadinya pengamatan yang tersensor. Misalkan durasi dalam sampel ukuran n diurutkan dari yang terkecil sampai terbesar, $t_1 < t_2 < \dots < t_k$, maka jumlah durasi k akan kurang dari n karena adanya durasi pemogokan yang sama atau apabila adanya beberapa pengamatan yang tersensor. Jumlah *event* (frekuensi *event*) yang terjadi pada durasi t_j dilambangkan dengan d_j , untuk $j = 1, 2, \dots, k$. Sedangkan jumlah pengamatan yang disensor antara t_j dan t_{j+1} dilambangkan dengan m_j , untuk $j = 1, 2, \dots, k$, sehingga apabila tidak ada durasi yang tersensor maka nilai m_j sama dengan nol.

Risiko pemogokan sebelum waktu t_j dilambangkan dengan r_j , yang didefinisikan dengan:

$$r_j = \sum_{j \geq i}^k (m_j + d_j). \quad (2.5)$$

Peluang bersyarat mengakhiri pemogokan pada durasi t_j , diestimasi oleh (London, 1988: 273):

$$\hat{q}_j = \frac{d_j}{r_j}.$$

Fungsi *hazard* untuk (t_j) diestimasi dengan:

$$h(t_j) = \frac{d_j}{r_j}, \quad (2.6)$$

yaitu jumlah *event* (frekuensi pemogokan) yang diselesaikan pada durasi t_j dibagi dengan jumlah risiko pada durasi t_j . Penaksir untuk fungsi survival adalah

$$S(t_j) = \prod_{j=1}^k \frac{(r_j - d_j)}{r_j} = \prod_{j=1}^k (1 - h(t_j)) \quad (2.7)$$

yang merupakan penaksir Kaplan-Meier atau *product limit estimator*.

Menurut London (1988: 269), apabila fungsi *hazard* meningkat maka mengindikasikan bahwa adanya ketergantungan durasi positif, yaitu indikasi bahwa semakin besar nilai durasi maka peluang berakhirnya *event* akan semakin tinggi. Contoh dalam durasi pemogokan, apabila terindikasi ketergantungan durasi positif maka semakin lama mogok kerja berlangsung, semakin besar pula peluang

pemogokan itu akan berakhir. Mungkin karena adanya peningkatan intensitas pencarian kerja atau permasalahan upah minimum yang tidak dipermasalahkan lagi oleh para pekerja.

Sebaliknya, apabila fungsi *hazard* menurun maka mengindikasikan bahwa adanya ketergantungan durasi negatif, yaitu indikasi bahwa semakin besar nilai durasi maka peluang berakhirnya *event* akan semakin kecil. Contoh dalam durasi pemogokan, apabila terindikasi ketergantungan durasi negatif maka semakin lama mogok kerja berlangsung, semakin kecil pula peluang pemogokan itu akan berakhir. Mungkin karena kesempatan kerja yang hampir habis.

Apabila fungsi *hazard* konstan, maka tidak memiliki ketergantungan terhadap nilai durasi, yang berarti peluang berakhirnya *event* tidak bergantung kepada nilai durasinya. Ini mengarah kepada distribusi eksponensial yang sudah dikenal dengan kaidah “*memoryless*”.

2.6 Metode Parametrik

Pada metode parametrik, secara khusus hanya akan membahas distribusi eksponensial saja dan *maximum likelihood estimation* untuk data tersensor yang berdistribusi eksponensial.

2.6.1 Uji Kecocokan Distribusi

Uji Kecocokan Kolmogorov-Smirnov Uji kecocokan merupakan suatu metode untuk pengujian hipotesis statistik yang digunakan untuk mengetahui apakah suatu data mengikuti suatu distribusi teoritis tertentu atau tidak. Salah satu uji kecocokan distribusi yang ada adalah uji *Kolmogorov-Smirnov*. Menurut Siegel (Desriawati, 2019: 16), uji *Kolmogorov-Smirnov* adalah suatu uji *goodness-of-fit*. Artinya, yang diperhatikan adalah tingkat kesesuaian antara distribusi serangkaian harga sampel (skor yang diobservasi) dengan suatu distribusi teoritis tertentu.

Tes ini menetapkan apakah skor-skor dalam sampel dapat secara masuk akal dianggap berasal dari suatu populasi dengan distribusi teoritis tertentu. Singkatnya, tes ini mencakup perhitungan distribusi frekuensi kumulatif yang

akan terjadi dibawah distribusi teoritisnya, serta membandingkan distribusi frekuensi itu dengan distribusi frekuensi kumulatif hasil observasi. Tes ini menetapkan suatu titik dimana kedua distribusi itu, yakni yang teoritis dan yang terobservasi memiliki perbedaan terbesar.

Hipotesis yang digunakan adalah sebagai berikut:

H_0 : Data berdistribusi teoritis tertentu.

H_1 : Data tidak berdistribusi teoritis tertentu.

Nilai statistik Uji *Kolmogorov-Smirnov* adalah:

$$KS = \text{maksimum } |F_0(X) - S_N(X)|$$

dengan $F_0(X)$ adalah suatu fungsi distribusi frekuensi kumulatif teoritis dibawah H_0 dan $S_N(X)$ adalah distribusi kumulatif yang diobservasi dari suatu sampel random dengan N observasi. Kriteria pengujian adalah tolak H_0 apabila nilai $KS >$ Nilai Tabel *Kolmogorov-Smirnov*. Sementara itu, jika menggunakan nilai *p-value* maka kriteria pengujiannya adalah tolak H_0 apabila nilai *p-value* $< \alpha$.

2.6.2 Distribusi Eksponensial

Distribusi eksponensial memiliki fungsi densitas sebagai berikut:

$$f(t) = \theta e^{-\theta t}$$

untuk $\theta > 0$ dan $t \geq 0$. Menurut Fitria dkk. (2016), fungsi distribusi kumulatif, fungsi survival dan fungsi *hazard* dari distribusi eksponensial adalah

$$F(t; \theta) = 1 - e^{-\theta t} \quad (2.8)$$

$$S(t; \theta) = e^{-\theta t}$$

$$h(t; \theta) = \theta$$

untuk menaksir nilai parameter atau nilai θ akan menggunakan metode kemungkinan maksimum (*Maximum Likelihood*).

Telah disinggung bahwa distribusi eksponensial dikenal dengan properti *memoryless*. Menurut Stephanie (2017), *memoryless* berarti bahwa peluang dari sebuah distribusi tidak tergantung pada nilai sebelumnya, setiap waktu dapat ditandai sebagai waktu yang pertama. Menurut Ross (1996: 35), variabel acak X dikatakan *without memory*, atau *memoryless*, jika

$$P\{T > t_1 + t | T > t\} = P\{T > t_1\} \quad (2.9)$$

untuk $t_1 \geq 0, t \geq 0$. Jika T adalah variabel acak yang menyatakan lama masa hidup (*lifetime*) untuk sebuah instrument selama minimal $t_1 + t$, maka persamaan (2.9) menyatakan bahwa peluang instrumen hidup, dengan syarat dapat bertahan hidup pada waktu t adalah sama dengan peluang awal, peluang bertahan hidup pada waktu t_1 .

Pembuktian dengan menggunakan definisi peluang bersyarat, maka persamaan (2.9) menjadi (Myers, 2012)

$$P(T > t_1 + t | T > t) = \frac{P(T > t_1 + t \text{ dan } T > t)}{P(T > t)}.$$

Jika $T > t_1 + t$, maka $T > t$ *redundant*, sehingga

$$P(T > t_1 + t | T > t) = \frac{P(T > t_1 + t)}{P(T > t)} \quad (2.10)$$

dengan menggunakan fungsi densitas peluang distribusi eksponensial,

$$P(T > t_1 + t | T > t) = \frac{P(T > t_1 + t)}{P(T > t)} = \frac{e^{-\theta(t_1+t)}}{e^{-\theta t}}. \quad (2.11)$$

Dengan demikian, didapat hasil sebagai berikut

$$P(T > t_1 + t | T > t) = e^{-\theta t_1} \quad (2.12)$$

dari persamaan (2.12), terbukti peluang bersyarat tidak bergantung pada t . Artinya distribusi eksponensial memiliki sifat *memoryless* karena nilai sebelumnya tidak berpengaruh pada nilai setelahnya.

2.6.3 Maximum Likelihood Estimation

Syarat utama untuk menggunakan *maximum likelihood estimation* adalah harus mengetahui distribusi dan/atau fungsi densitas dari sebuah data. Setelah mengetahui distribusi dan/atau fungsi densitas langkah-langkah untuk menaksir parameter adalah sebagai berikut:

1. mencari fungsi *likelihood*;
2. mencari logaritma natural dari fungsi *likelihood*; dan