

## DAFTAR ISI

Bagian 1 .....	4
Uji .....	5
Import Library.....	5
Import Data.....	5
Mengedit Data .....	5
Load 1960 Iowa Standard Mortality.....	6
Computation of One-Sample, Log-Rank Test .....	7
Uji Hipotesis .....	9
Tujuan.....	9
Hipotesis.....	9
Tingkat Signifikansi .....	9
Statistik Uji .....	9
Titik Kritis.....	9
Aturan Keputusan.....	10
Kesimpulan .....	10
Bagian 2.....	11
JAWABAN BAGIAN 2 NOMOR 1.....	12
Uji & Tujuan .....	12
Data dan Statistik Deskriptif.....	12
Hipotesis.....	12
Taraf Signifikansi .....	13
Estimasi Kaplan-Meier.....	13
Statistik Uji .....	13
Aturan Keputusan & Hasil.....	14
Kesimpulan .....	14
JAWABAN BAGIAN 2 NOMOR 2.....	15
Uji.....	15
Tujuan.....	15
Hipotesis.....	15
Taraf Signifikansi .....	15
Statistik Uji .....	15
• <b>Estimasi Kaplan-Meier.....</b>	16

Aturan Keputusan.....	17
Kesimpulan .....	17
JAWABAN BAGIAN 2 NOMOR 3.....	18
Uji.....	18
Tujuan.....	18
Hipotesis.....	18
Taraf Signifikansi .....	18
Uji statistik .....	18
Aturan Keputusan : .....	20
Kesimpulan :.....	21
JAWABAN BAGIAN 2 NOMOR 4.....	22
Untuk Penyakit NHL.....	22
Untuk Penyakit HOD .....	22
Untuk Jenis Penanganan Allo .....	23
Untuk Jenis Penanganan Auto.....	24
Bagian 3. Lampiran .....	25
Bagian 1.....	25
Bagian 2 Nomor 1 .....	25
Bagian 2 Nomor 2 .....	26
Bagian 2 Nomor 3 .....	27
Bagian 2 Nomor 4 .....	28

## Bagian 1

[25 poin]. Uji 1 sampel

### EXAMPLE 7.1

In section 6.3, we examined models for excess and relative mortality in a sample of 26 Iowa psychiatric patients described in section 1.15. We shall now use the one-sample log-rank statistic to test the hypothesis that the hazard rate of this group of patients is the same as the hazard rate in the general Iowa population, given by the standard 1960 Iowa

### ter 7 Hypothesis Testing

mortality table. To perform this test, we will use the sex specific survival rates. Time  $T_j$  is taken as the  $j$ th individual's age at death or the end of the study, and the left-truncation time  $L_j$ , is this individual's age at entry into the study. We obtain  $H(t)$  as  $-\ln[S(t)]$  from the appropriate column of Table 6.2. Table 7.1 shows the calculations to compute  $O(71)$  and  $E(71)$ .

The test statistic is  $\chi^2 = (15 - 4.4740)^2 / 4.4740 = 24.7645$  which has a chi-squared distribution with one degree of freedom. Here the  $p$ -value of this test is close to zero, and we can conclude that the mortality rates of the psychiatric patients differ from those of the general public.

Mengacu pada prosedur contoh di atas (sumber: Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (Vol. 1230). New York: Springer., link ebook), lakukan hal serupa untuk menguji apakah ada perbedaan yang signifikan pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum.

No	Gender	Age at diagnosis	Follow up
1	M	34	2
2	F	40	7
3	M	37	8
4	M	59	5+
5	M	54	5
6	M	45	6
7	F	38	1
8	M	44	8+
9	F	64	2
10	F	42	10+

Cuplikan data adalah sebagai berikut (data lengkap lihat lampiran file Excel):

Follow up menyatakan waktu pengamatan pasien sejak mulai terdiagnosis hingga meninggal (atau keluar dari studi, untuk yang tersensor, dalam tahun).

Tuliskan secara lengkap mulai dari hipotesis hingga kesimpulan, lampirkan tampilan proses perhitungan (file excel atau lainnya yang relevan).

## Uji

Akan digunakan uji 1 sampel, *log rank test*, sesuai dengan instruksi pada soal. Akan diuji apakah terdapat perbedaan yang signifikan pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum.

### Import Library

```
# Import Library ----  
library(survival)  
library(readxl)
```

### Import Data

```
# Import Data ----  
raw <- read_excel("C:/Users/marye/Downloads/data project 2.xlsx",  
                  col_names = c("No", "gender", "age_entry", "follow_up"),  
                  skip = 1)
```

### Mengedit Data

```
# Preprocessing Data ----  
  
# Ubah tipe data tibble menjadi dataframe  
data<- as.data.frame(raw)  
# Hapus kolom "No"  
data$No <- NULL  
# Menambah kolom "Status"  
data <- data %>%  
  mutate(status = ifelse(grepl("\\\\+$", follow_up), 0, 1))  
# Jika tersensor, status = 0. Jika data Lengkap, status = 1  
data$follow_up <- gsub("\\\\+$", "", data$follow_up)  
  
# Menambah kolom age_exit  
data$follow_up <- as.numeric(data$follow_up)  
data$age_exit <- data$age_entry + data$follow_up
```

Didapat data yang sudah dilakukan pengeditan/manipulasi sebagai berikut:

Subject $j$	Sex	Age at Entry $L_i$	Follow Up	Status $d_i$	Age at Exit $T_j$
1	M	34	2	1	36
2	F	40	7	1	47

3	M	37	8	1	45
4	M	59	5	0	64
5	M	54	5	1	59
6	M	45	6	1	51
7	F	38	1	1	39
8	M	44	8	0	52
9	F	64	2	1	66
10	F	42	10	0	52
11	F	62	4	1	66
12	M	56	5	1	61
13	M	43	4	1	47
14	M	53	7	0	60
15	F	47	8	1	55
16	M	30	1	1	31
17	F	62	18	0	80
18	F	42	11	1	53
19	M	70	19	1	89
20	F	44	20	1	64
21	F	28	8	0	36
22	F	27	7	1	34
23	M	62	8	1	70
24	M	44	6	1	50
25	F	46	5	1	51

## Load 1960 Iowa Standard Mortality

Data tabel ini didapat dari Table 6.2 1960 Iowa Standard Mortality pada buku referensi (Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (Vol. 1230). New York: Springer). Tabel mortalitas terdiri dari gender perempuan dan laki-laki. Pada tabel, hanya terdapat kolom Survival Function, untuk mempermudah perhitungan selanjutnya dibuat kolom baru yakni Kumulatif Hazard dengan rumus berikut:

$$H(t) = -\ln[S(t)]$$

Sehingga dihasilkan cuplikan datanya sebagai berikut:

```
# 1960 Iowa Standard Mortality ----

## Male (From Table 6.2) ----
raw_male <- read_excel("G:/My Drive/Kuliah - Maryesta/Semester 4/Model Survival/Tugas or Project/Project 2/male_iowa.xlsx", )
raw_male$survival <- as.numeric(raw_male$survival)
male <- as.data.frame(raw_male)
# Membuat kolom untuk kumulatif hazard
male$hazard_cum <- -log(male$survival)

## Female (From Table 6.2) ----
raw_female <- read_excel("G:/My Drive/Kuliah - Maryesta/Semester 4/Model Survival/Tug as or Project/Project 2/female_iowa.xlsx", )
```

```

raw_female$survival <- as.numeric(raw_female$survival)
female <- as.data.frame(raw_female)
# Membuat kolom untuk kumulatif hazard
female$hazard_cum <- -log(female$survival)

```

Age	Survival Function	Cumulative Hazard
18	0.96394	0.03672623
19	0.96246	0.03826277
20	0.96088	0.03990575
21	0.95919	0.04166610
22	0.95739	0.04354445
...	...	...

Male:

Age	Survival Function	Cumulative Hazard
18	0.97372	0.02663149
19	0.97317	0.02719649
20	0.97263	0.02775154
21	0.97210	0.02829660
22	0.97158	0.02883167
...	...	...

Female:

## Computation of One-Sample, Log-Rank Test

Dibuat tabel perhitungan. Di dapat pada data 17 dan 19 dengan Age at Exit ( $T_j$ ) secara berurutan yakni 80 dan 89. Namun, kedua usia ini tidak terdapat di Tabel 1960 Iowa Standard Mortality. Untuk menangani hal ini, akan diambil nilai  $H_0(T_j)$  dari range usia terbesar, yakni 77.

```

# Tabel ----
## Hazard Entry ----
# Membuat kolom baru "hazard_entry" dalam dataframe "data"
data$hazard_entry <- ifelse(data$gender == "M",
                           male$hazard_cum[match(data$age_entry, male$age)],
                           female$hazard_cum[match(data$age_entry, female$age)])
## Hazard Exit ----
# Membuat kolom baru "hazard_exit" dalam dataframe "data"
data$hazard_exit <- ifelse(data$gender == "M",
                           male$hazard_cum[match(data$age_exit, male$age)],
                           female$hazard_cum[match(data$age_exit, female$age)])

## Mengatasi NA ----
# Data NA akan diganti dengan kumulatif hazard umur 77 (umur max pada tabel)
# Mencari posisi NA dalam kolom hazard_exit
na_positions <- which(is.na(data$hazard_exit))

# Mengganti nilai NA dengan nilai terakhir dari tabel male$hazard_cum atau female$hazard_cum
for (i in na_positions) {
  if (data$gender[i] == "M") {
    data$hazard_exit[i] <- tail(male$hazard_cum, 1)
  } else {
    data$hazard_exit[i] <- tail(female$hazard_cum, 1)
  }
}

## Hazard Exit - Hazard Entry ----
data$hazard_en_ex<- data$hazard_exit - data$hazard_entry

```

Di dapat hasil komputasi sebagai berikut:

Subject <i>j</i>	Sex	Age at Entry <i>L<sub>i</sub></i>	Follow Up	Status <i>d<sub>i</sub></i>	Age at Exit <i>T<sub>j</sub></i>	$H_0(L_j)$	$H_0(T_j)$	$H_0(T_j) - H_0(L_j)$
1	M	34	2	1	36	0.06213076	0.06564847	0.003517715
2	F	40	7	1	47	0.04496599	0.06047213	0.015506148
3	M	37	8	1	45	0.06762595	0.09263077	0.025004824
4	M	59	5	0	64	0.23183118	0.34208194	0.110250762
5	M	54	5	1	59	0.16126090	0.23183118	0.070570282
6	M	45	6	1	51	0.09263077	0.13241202	0.039781247
7	F	38	1	1	39	0.04205192	0.04345045	0.001398529
8	M	44	8	0	52	0.08813200	0.14118345	0.053051448
9	F	64	2	1	66	0.17161903	0.20052626	0.028907228
10	F	42	10	0	52	0.04835026	0.07968194	0.031331678
11	F	62	4	1	66	0.14779260	0.20052626	0.052733659
12	M	56	5	1	61	0.18563103	0.27040551	0.084774482
13	M	43	4	1	47	0.08416452	0.10349559	0.019331065
14	M	53	7	0	60	0.15074150	0.25027069	0.099529190
15	F	47	8	1	55	0.06047213	0.09494825	0.034476112
16	M	30	1	1	31	0.05613658	0.05754437	0.001407788
17	F	62	18	0	80	0.14779260	0.52797178	0.380179181
18	F	42	11	1	53	0.04835026	0.08449092	0.036140659
19	M	70	19	1	89	0.54709204	0.95163298	0.404540938
20	F	44	20	1	64	0.05244132	0.17161903	0.119177712
21	F	28	8	0	36	0.03247154	0.03954156	0.007070024
22	F	27	7	1	34	0.03177967	0.03741115	0.005631485
23	M	62	8	1	70	0.29235966	0.54709204	0.254732376
24	M	44	6	1	50	0.08813200	0.12431683	0.036184830
25	F	46	5	1	51	0.05751259	0.07519780	0.017685202

Lalu dilakukan juga perhitungan untuk  $O(\tau)$  dan  $E(\tau)$  sebagai berikut:

```
## Menghitung O(tau) dan E(tau)
(o_tau <- sum(data$status))

## [1] 19

(e_tau <- sum(data$hazard_en_ex))

## [1] 1.932915
```

Sehingga didapat:

$O(\tau)$  = observed number of events at or prior to time  $\tau = 19$

$$E(\tau) = V[Z(\tau)] = \sum_{j=1}^n [H_0(T_j) - H_0(L_j)] = 1.932915$$

## Uji Hipotesis

### Tujuan

Untuk menguji apakah ada perbedaan yang signifikan pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum.

### Hipotesis

$H_0$ : Tidak ada perbedaan yang signifikan pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum

$H_1$ : Terdapat perbedaan yang signifikan pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum

atau

$$H_0: h(t) = h(t); t \leq \tau$$

$$H_1: h(t) \neq h(t); \text{for some } t \leq \tau$$

### Tingkat Signifikansi

$$\alpha = 0.05$$

### Statistik Uji

Pada bagian sebelumnya telah didapat:

$O(\tau)$  = observed number of events at or prior to time  $\tau = 19$

$$E(\tau) = V[Z(\tau)] = \sum_{j=1}^n [H_0(T_j) - H_0(L_j)] = 1.932915$$

Maka

```
## Statistik Uji ----  
chi_sq = ((o_tau-e_tau)^2) / e_tau  
chi_sq  
## [1] 150.6975
```

$$\chi^2 = \frac{Z(\tau)}{V[Z(\tau)]} = \frac{[O(\tau) - E(\tau)]^2}{E(\tau)} = \frac{[19 - 1.932915]^2}{1.932915} = 150.6975$$

### Titik Kritis

```
## Titik kritis ----  
titik_kritis=qchisq(p = 1-0.025, df = 1, lower.tail = TRUE) #Chi sq(1-0.025, df=1)  
titik_kritis  
## [1] 5.023886
```

$$\chi^2_{db=1,1-\frac{\alpha}{2}} = 5.023886$$

## Aturan Keputusan

$H_0$  ditolak jika  $\chi^2 > \chi^2_{db=1,1-\frac{\alpha}{2}}$ . Didapat bahwa pada pengujian ini:

$$\begin{aligned}\chi^2 &> \chi^2_{db=1,1-\frac{\alpha}{2}} \\ 150.6975 &> 5.023886\end{aligned}$$

Sehingga aturan keputusannya  $H_0$  ditolak

```
## p-value ----  
pchisq(q=chi_sq, df=1, lower.tail=FALSE)  
## [1] 1.220404e-34
```

Selain itu, didapat pula  $p - value = 1.220404e - 34 < 0.05$  sehingga  $H_0$  ditolak

## Kesimpulan

Dengan signifikansi  $\alpha = 0.05$ , didapat bahwa  $\chi^2 > \chi^2_{db=1,1-\frac{\alpha}{2}}$  dan  $p - value < 0.05$  maka  $H_0$  ditolak.

Dapat disimpulkan bahwa **ada perbedaan yang signifikan** pada tingkat mortalitas pasien kanker lidah dengan tingkat mortalitas populasi Iowa secara umum ( $h(t) \neq h(t)$ ; for some  $t \leq \tau$ )

## Bagian 2.

Lakukan prosedur pengujian hipotesis yang sesuai, dengan langkah-langkah pengujian yang jelas, mulai dari penyusunan hipotesis hingga pengambilan kesimpulan, untuk menjawab pertanyaan berikut ini. Lampirkan codes (screenshot/link file) dari pengolahan data yang dilakukan.

**TABLE 1.5**

*Times to death or relapse (in days) for patients with bone marrow transplants for Hodgkin's and non-Hodgkin's lymphoma*

Allo NHL				Auto NHL				Allo HOD				Auto HOD			
$T_i$	$\delta_i$	$Z_1$	$Z_2$												
28	1	90	24	42	1	80	19	2	1	20	34	30	1	90	73
32	1	30	7	53	1	90	17	4	1	50	28	36	1	80	61
49	1	40	8	57	1	30	9	72	1	80	59	41	1	70	34
84	1	60	10	63	1	60	13	77	1	60	102	52	1	60	18
357	1	70	42	81	1	50	12	79	1	70	71	62	1	90	40
933	0	90	9	140	1	100	11					108	1	70	65
1078	0	100	16	176	1	80	38					132	1	60	17
1183	0	90	16	210	0	90	16					180	0	100	61
1560	0	80	20	252	1	90	21					307	0	100	24
2114	0	80	27	476	0	90	24					406	0	100	48
2144	0	90	5	524	1	90	39					446	0	100	52
				1037	0	90	84					484	0	90	84
												748	0	90	171
												1290	0	90	20
												1345	0	80	98

Sumber: Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (Vol. 1230). New York: Springer., hal 11-12; link ebook)

Data di atas adalah mengenai 43 pasien transplantasi tulang belakang di The Ohio State University Bone Marrow Transplant Unit. Semua pasien menderita salah satu tipe ini: Hodgkin's disease (HOD), atau non-Hodgkin's limpoma (NHL). Masing-masing pasien diberi salah satu penanganan ini: allogeneic (Allo) transplant, atau autogenic (Auto) transplant.  $T_i$  menyatakan waktu pasien berada dalam studi (sejak terdiagnosis),  $\delta_i=1$  jika pasien meninggal atau mengalami relaps, dan 0 jika mengalami penyensoran.  $Z_1$  adalah pretransplant Karnofsky score, dan  $Z_2$  adalah waktu tunggu untuk mendapatkan transplant (dalam bulan), terhitung dari pasien terdiagnosis.

1. [20 poin] Apakah ada perbedaan kondisi ketahanan hidup dari pasien HOD dengan pasien NHL? Uji apa yang kalian gunakan? Jelaskan.

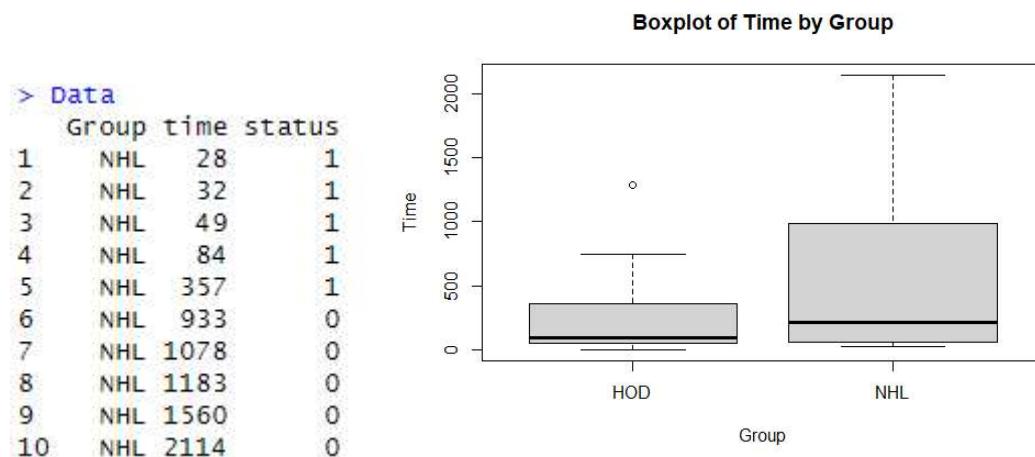
## JAWABAN BAGIAN 2 NOMOR 1

### Uji & Tujuan

Untuk mengetahui apakah ada perbedaan kondisi ketahanan hidup dari pasien HOD dengan pasien NHL akan digunakan uji k-sampel. Perhatikan bahwa HOD merupakan pasien yang menderita penyakit Hodgkin dan NHL adalah pasien yang tidak menderita Hodgkin dengan setiap pasien (baik penderita maupun tidak) mendapatkan masing-masing penanganan allogenic dan autogenic. Walaupun demikian, karena hanya akan dicari apakah ada perbedaan kondisi ketahanan hidup dari pasien HOD dan NHL maka dianggap hanya ada 2 sampel (HOD dan NHL) dengan mengabaikan penanganan yang diberikan. Berdasarkan hal tersebut, maka akan digunakan uji 2-sampel pada soal ini.

### Data dan Statistik Deskriptif

Pada uji ini akan dibuat data table dengan nilai waktu pasien dalam studi  $T_i$  dan  $\delta_i$  yang menyatakan indikator keadaan pasien. Berikut merupakan cuplikan data yang lebih lengkap dapat dilihat pada lampiran



Untuk melengkapi statistik deskriptif, maka dibuatlah boxplot seperti di atas. Dengan boxplot tersebut dapat dilihat bahwa pasien NHL memiliki waktu studi kasus yang cenderung lebih lama dibandingkan pasien HOD. Untuk melakukan uji 2 sample untuk mengetahui apakah ada perbedaan kondisi ketahanan hidup antara pasien HOD dan NHL berikut adalah hipotesis null dan hipotesis alternatif dari kasus ini

### Hipotesis

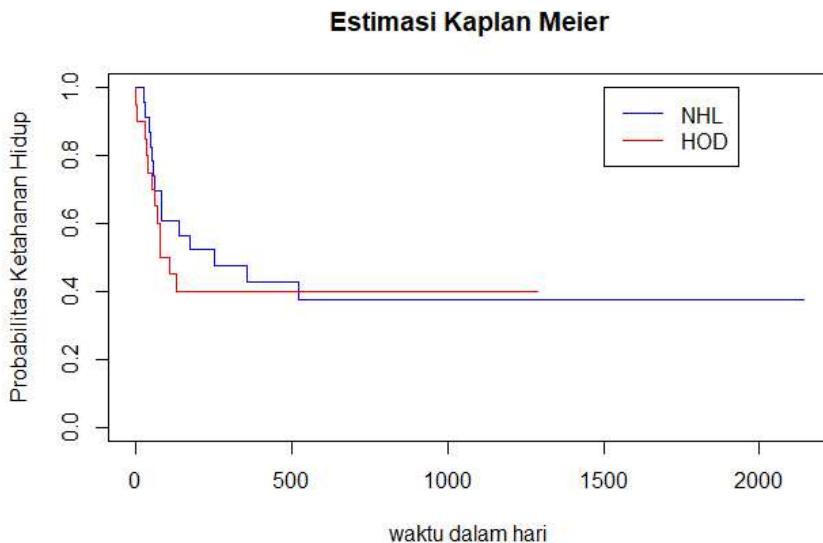
$$H_0: h_1(t) = h_2(t), \forall t \leq \tau \text{ (Tidak ada perbedaan ketahanan hidup antara pasien HOD dan NHL)}$$

$H_t: h_1(t) \neq h_2(t)$  pada beberapa  $t \leq \tau$  (ada perbedaan ketahanan hidup antara pasien HOD dan NHL)

## Taraf Signifikansi

Akan digunakan taraf signifikansi  $\alpha = 0.05$

## Estimasi Kaplan-Meier



Dengan bantuan R didapatkan juga plot estimasi kaplan meier yang telihat tidak terlalu jauh pada rentang waktu tertentu antara NHL dan HOD. Untuk memastikannya akan dilakukan uji hipotesis menggunakan statistik uji.

## Statistik Uji

Menggunakan bantuan R akan dilakukan pengujian yang akan menghasilkan  $p\text{-value}$  dengan code berikut

```
> result<-survdiff(Surv(time, status) ~ Group, data=Data);result
Call:
survdiff(formula = Surv(time, status) ~ Group, data = Data)

      N Observed Expected (O-E)^2/E  (O-E)^2/V
Group=HOD 20       12     10.6    0.184    0.315
Group=NHL 23       14     15.4    0.127    0.315

chisq= 0.3  on 1 degrees of freedom, p= 0.6
```

## Aturan Keputusan & Hasil

Hipotesis null dikatakan ditolak jika nilai *p-value* kurang dari nilai taraf signifikansi dan sebaliknya. Berdasarkan hasil *p-value* yang didapatkan dengan taraf signifikansi yang telah ditentukan maka didapatkan bahwa  $p-value = 0.6 > 0.05$  dimana tidak cukup bukti untuk menolak hipotesis null sehingga hipotesis null tidak ditolak.

Karena hipotesis null diterima maka tidak ada perbedaan kondisi ketahanan hidup antara pasien HOD dan NHL pada setiap waktu baik itu menerima penanganan allo maupun auto.

## Kesimpulan

Berdasarkan uji 2 sample didapatkan kesimpulan bahwa tidak ada perbedaan kondisi ketahanan hidup antara pasien HOD dengan NHL.

2. [20 poin] Apakah ada perbedaan kondisi ketahanan hidup antara pasien yang mendapat penanganan Allo dengan pasien yang mendapat penanganan Auto? Uji apa yang kalian gunakan? Jelaskan.

## JAWABAN BAGIAN 2 NOMOR 2

### Uji

Uji yang digunakan adalah uji 2 sampel (uji k sampel) karena ada dua kelompok pasien yang dibedakan berdasarkan jenis penanganannya dan oleh sebab itu k nya bernilai 2. Alasan lainnya karena di dalam uji ini membandingkan antara 2 perlakuan yaitu mendapat penanganan Allo dan mendapat penanganan Auto yang mana melihat apakah ada perbedaan kondisi ketahanan hidup dari dua kelompok yaitu antara pasien yang mendapat penanganan Allo dan pasien yang mendapat penanganan Auto. Oleh sebab itu, uji yang tepat untuk digunakan untuk kasus ini adalah uji 2 sampel.

### Tujuan

Untuk melihat apakah ada perbedaan kondisi ketahanan hidup dari dua kelompok yaitu antara pasien yang mendapat penanganan Allo dan pasien yang mendapat penanganan Auto.

### Hipotesis

$H_0 : h_1(t) = h_2(t)$  untuk semua  $t \leq \tau$  (Tidak ada perbedaan ketahanan hidup antara pasien yang mendapat penanganan Allo dan pasien yang mendapat penanganan Auto)

$H_1 : \text{tidak demikian}$  (Terdapat perbedaan ketahanan hidup antara pasien yang mendapat penanganan Allo dan pasien yang mendapat penanganan Auto)

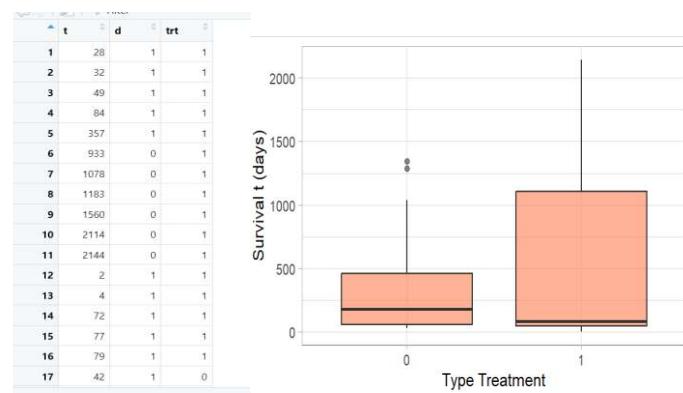
### Taraf Signifikansi

Alpha = 0,05

### Statistik Uji

- **Data dan Statistik Deskriptif**

Salah satu eksplorasi deskriptif yang dapat kita lakukan adalah memvisualisasikan waktu bertahan hidup di setiap perlakuan adalah kita dapat menggunakan boxplot.



Keterangan Gambar : trt atau *type treatment* = pasien yang mendapat penanganan Allo ditandai dengan *type treatment* 0 dan pasien yang mendapat penanganan Auto ditandai dengan *type treatment* 1, t = menyatakan waktu pasien berada dalam studi (sejak terdiagnosis), dan d = 1 jika pasien meninggal atau mengalami relaps, dan 0 jika mengalami penyensoran.

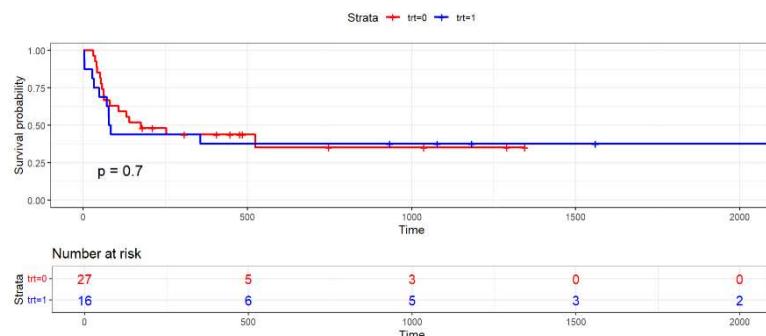
Dari grafik menunjukkan bahwa terdapat perbedaan pada pasien dengan penanganan Allo dan penanganan Auto, waktu hidup pasien semakin singkat bagi pasien yang mendapat penanganan Auto dibandingkan dengan pasien dengan penanganan Allo.

Fungsi survival dari kedua kelompok pasien tersebut tersaji pada tabel berikut

trt=0							trt=1						
	time	n.risk	n.event	survival	std.err		time	n.risk	n.event	survival	std.err		
				0.765	0.995					0.938	0.0605		
trt=0				0.735	0.981		trt=1			1	0.875	0.0827	
30	27	1	0.963	0.0363		36	14	1	0.812	0.0976			
36	26	1	0.926	0.0504		41	13	1	0.750	0.1083			
41	25	1	0.889	0.0605		42	12	1	0.688	0.1159			
42	24	1	0.852	0.0684		52	11	1	0.625	0.1210			
52	23	1	0.815	0.0748		53	10	1	0.562	0.1240			
53	22	1	0.778	0.0800		57	9	1	0.500	0.1250			
57	21	1	0.741	0.0843		62	8	1	0.438	0.1240			
62	20	1	0.704	0.0879		63	7	1	0.375	0.1210			
63	19	1	0.667	0.0907		81	6	1	0.312	0.1159			
81	18	1	0.630	0.0929		108	5	1	0.250	0.1159			
108	17	1	0.593	0.0946		132	4	1	0.188	0.1159			
132	16	1	0.556	0.0956		140	3	1	0.125	0.1159			
140	15	1	0.519	0.0962		176	2	1	0.063	0.1159			
176	14	1	0.481	0.0962		252	1	1	0.000	0.1159			
252	11	1	0.438	0.0969		524	0	1	0.000	0.1159			
				0.350	0.1102					0.058	0.1159		

### • Estimasi Kaplan-Meier

Akan disajikan secara grafik buat fungsi survivalnya dan hasil perhitungan fungsi survivalnya :



Terlihat bahwa pada grafik fungsi survival yang berwarna biru (pasien yang mendapat penanganan Auto) terlihat lebih curam dibandingkan yang berwarna merah (pasien yang mendapat penanganan Allo ) artinya kemampuan pasien yang mendapat penanganan Auto untuk survive lebih rendah. Dari grafik terlihat bahwa p-value yang didapat adalah 0.7.

```

Call:
survdiff(formula = surv(t, d) ~ trt, data = mydata)

          N Observed Expected (O-E)^2/E  (O-E)^2/V
trt=0    27      16     16.92   0.0496   0.143
trt=1    16      10      9.08   0.0923   0.143

chisq= 0.1 on 1 degrees of freedom, p= 0.7

```

Terlihat bahwa pada allo (trt = 0) dan auto (trt = 1) nilai *expected* nya lebih tinggi dari *observed* artinya kondisi pasien lebih buruk karena jumlah kematian lebih tinggi daripada seharusnya.

## Aturan Keputusan

Jika p-value < 0.05 maka  $H_0$  ditolak

Jika p-value > 0.05 maka  $H_0$  tidak ditolak

Didapat nilai  $p\text{-value} = 0.7 > 0.05 = \alpha$ , maka  $H_0$  tidak ditolak

## Kesimpulan

∴ Dengan alpha = 0.05, didapat nilai  $p\text{-value} = 0.7 > 0.05 = \alpha$ , maka  $H_0$  tidak ditolak. Disimpulkan bahwa tidak terdapat perbedaan kondisi ketahanan hidup yang signifikan antar pasien yang mendapat penanganan Allo dan pasien yang mendapat penanganan Auto.

3. [20 poin] Apakah ada perbedaan kondisi ketahanan hidup antara keempat kelompok pasien tersebut (Allo NHL, Auto NHL, Allo HOD, Auto HOD)? Uji apa yang kalian gunakan? Jelaskan.

## JAWABAN BAGIAN 2 NOMOR 3

### Uji

Pada analisis ini, kami akan menggunakan uji log-rank untuk membandingkan kondisi hidup antara 4 kelompok pasien Allo NHL, Auto NHL, Allo HOD, Auto HOD, yang masing-masing memiliki sampel berukuran 11, 12, 5, dan 15.

### Tujuan

Metode uji log-rank bertujuan untuk memeriksa apakah terdapat perbedaan yang signifikan antar kelompok pada masing-masing kurva survivalnya

### Hipotesis

$$H_0 : h_1(t) = h_2(t) = h_3(t) = h_4(t), \quad t \leq \tau$$

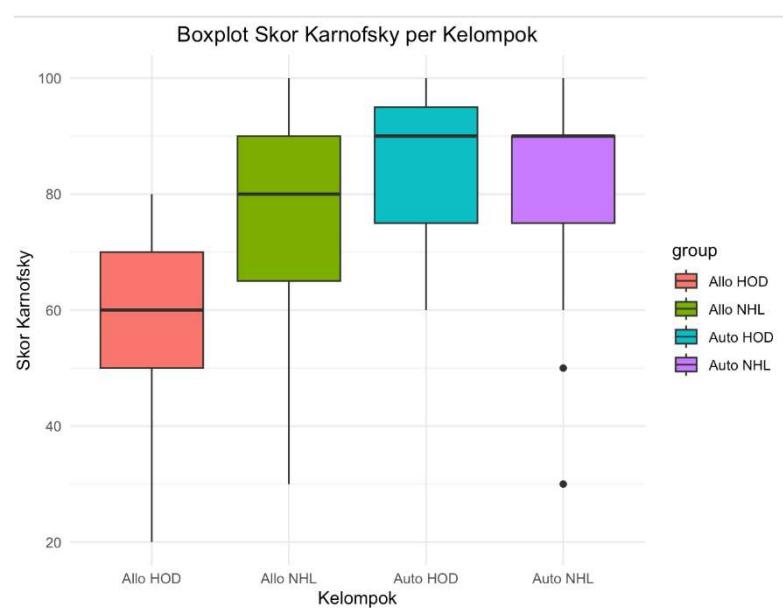
$$H_1: h_j(t) \neq h_k(t), \text{ untuk beberapa } t \leq \tau, j \neq k; j,k = 1,2,3,4$$

### Taraf Signifikansi

Alpha = 0,05

### Uji statistik

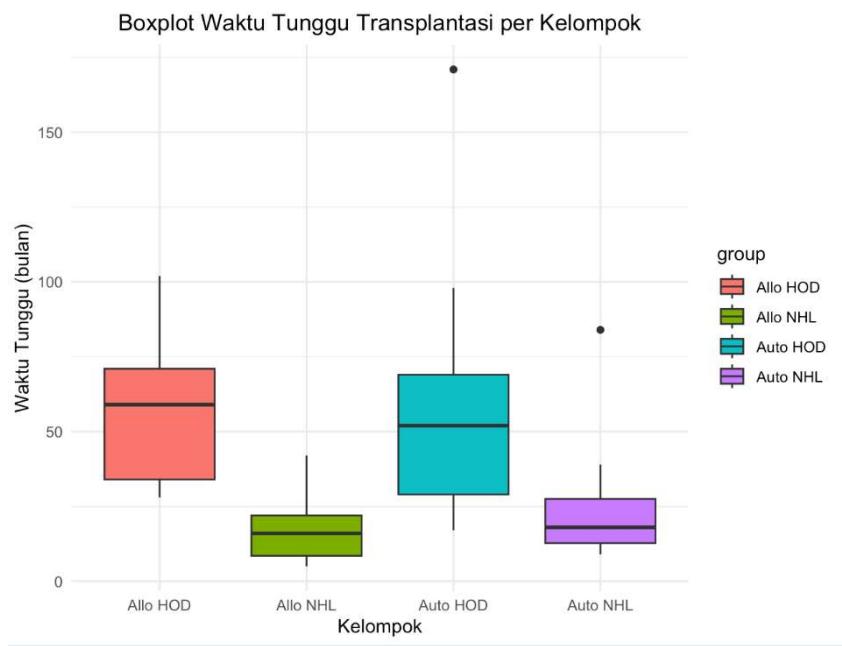
Berikut merupakan Skor Karnofsky untuk masing – masing kelompok



Berdasarkan boxplot diatas, dapat kita amati bahwa :

- Kelompok Allo HOD memiliki median 60 dengan range nilai antara 20 sampai 80. Hal ini menunjukkan bahwa kondisi fisik pasien dalam kelompok tersebut seragam tetapi memiliki nilai yang cukup rendah
- Kelompok Auto HOD memiliki median sebesar 90 dengan range nilai antara 60 sampai 100. Hal ini menunjukkan bahwa hampir seluruh pasien dalam kelompok ini memiliki kondisi fisik yang baik
- Kelompok Allo NHL memiliki nilai median 80 dengan range nilai antara 30 sampai 100. Hal ini mengindikasikan bahwa rata-rata kondisi fisik dari pasien pada kelompok ini adalah baik
- Kelompok Auto NHL, median tidak dapat ditentukan karena terdapat beberapa outlier pada data. Sehingga, pada boxplot diatas tidak terlihat median untuk kelompok ini. Hal ini menunjukkan bahwa kelompok Auto NHL memiliki variasi yang signifikan antar setiap pasiennya

Berikut merupakan Boxplot Waktu Tunggu Transplantasi Masing-Masing Kelompok



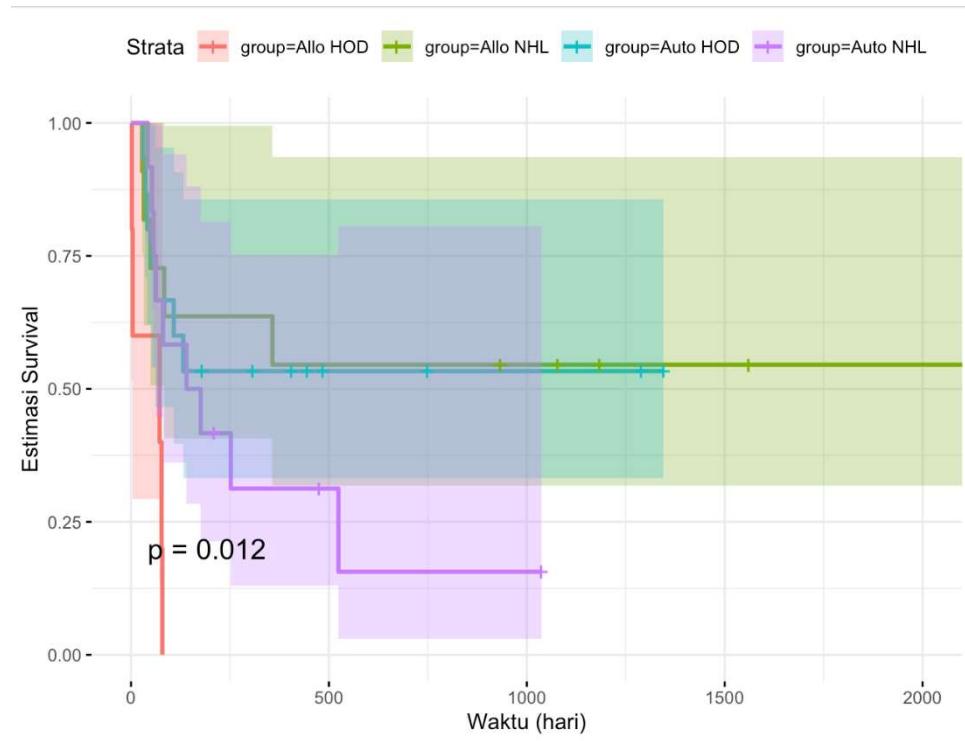
Berdasarkan boxplot diatas, bisa disimpulkan bahwa :

- Kelompok Allo HOD memiliki waktu tunggu transplantasi yang relatif singkat dengan median kisaran 60 hari dengan range waktu kisaran 30 hingga 100 hari. Oleh karena itu, proses transplantasi pada pasien dalam kelompok ini cenderung cepat
- Kelompok Auto HOD memiliki median waktu tunggu transplantasi kisaran 50 hari dengan range waktu tunggu kisaran 20 hingga 90 hari. Hal ini menunjukkan bahwa waktu tunggu dari kelompok Auto HOD mirip dengan kelompok Allo HOD. Akan tetapi, pada kelompok

ini variasi dalam waktu tunggu lebih sedikit dikarenakan range antara nilai minimum dan maksimumnya

- Kelompok Allo NHL memiliki median waktu tunggu transplantasi sekitar 20 hari dan memiliki rentang waktu tunggu kisaran 10 hingga 40 hari
- Kelompok Auto NHL memiliki median waktu tunggu sekitar 20 hari dengan range waktu dari 15 hingga 35 hari.

### Plot Kaplan – Meier



Berdasarkan plot diatas, hasil analisis menggunakan bantuan R menunjukkan bahwa nilai p-value yang didapat hanya sebesar 0.012. Nilai p yang lebih kecil dari 0.05 ini mengindikasikan adanya perbedaan yang signifikan pada fungsi survival antar kelompok.

Dengan bantuan perhitungan menggunakan R yang terdapat pada lampiran, didapat

$$\chi^2 = 10.96153$$

Aturan Keputusan :

Tolak  $H_0$  jika  $p - value < 0.05$

Jika  $p$ -value  $> 0.05$  maka  $H_0$  tidak ditolak

Didapat nilai  $p$ -value =  $0.012 < 0.05 = \alpha$ , maka  $H_0$  ditolak

### Kesimpulan :

$H_0$  ditolak karena nilai  $p$ -value  $< 0.05$  dan dapat disimpulkan bahwa terdapat perbedaan yang signifikan secara statistik pada fungsi survival diantara keempat kelompok pasien Allo HOD, Auto HOD, Allo NHL, dan Auto NHL. Hal ini dapat dibuktikan dengan nilai  $p$ -value =  $0.012$  dan  $\chi^2 = 10.96153$ . Oleh karena itu, dengan menggunakan metode analisis uji log-rank, dapat disimpulkan bahwa jenis penyakit, jenis transplantasi, waktu tunggu transplantasi, dan skor Karnofsky memiliki pengaruh terhadap estimasi survival pada pasien

4. [15 poin] Apakah ada beda efek dari jenis penanganan (Allo dan Auto) pada kedua tipe penyakit? Jelaskan alasannya.

## JAWABAN BAGIAN 2 NOMOR 4

Untuk melihat apakah ada perbedaan efek dari jenis penanganan (Allo dan Auto) pada kedua tipe penyakit, kita dapat melakukan uji stratifikasi dan membagi data berdasarkan jenis penyakit yaitu NHL dan HOD dengan mempertimbangkan faktor jenis penanganan.

### Untuk Penyakit NHL

- Tujuan : Untuk melihat apakah terdapat perbedaan efek jenis penanganan (Allo dan Auto) pada pasien yang menderita penyakit NHL

- Menentukan Hipotesis

$$H_0 : h_{allo(NHL)}(t) = h_{auto(NHL)}(t), t \leq \tau$$

$$H_1 : \text{tidak demikian}$$

- Menentukan taraf signifikansi

Akan digunakan  $\alpha = 0.05$

- Statistik uji

Perhatikan kode beserta output R berikut

```
> # Read data from URL
> url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo_Auto_NHL.DAT"
> dfNHL <- read.table(url, sep = ";", header = TRUE)
> # Membuat objek survival dari data
> surv_obj_NHL <- Surv(time = dfNHL$ftime, event = dfNHL$fustat)
> # Melakukan uji strata
> strata_test_NHL <- survdiff(surv_obj_NHL ~ dfNHL>Type)
> # Melihat hasil uji strata
> strata_test_NHL
Call:
survdiff(formula = surv_obj_NHL ~ dfNHL>Type)

      N Observed Expected (O-E)^2/E (O-E)^2/V
dfNHL>Type=Allo 11      5    7.31    0.728     1.58
dfNHL>Type=Auto 12      9    6.69    0.794     1.58

Chisq= 1.6  on 1 degrees of freedom, p= 0.2
```

- Aturan keputusan dan kesimpulan

$H_0$  ditolak jika p-value <  $\alpha$ . Terlihat bahwa nilai p-value = 0.2 >  $\alpha = 0.05$ , sehingga  $H_0$  gagal ditolak. Akibatnya tidak ada perbedaan efek jenis penanganan (Allo dan Auto) pada pasien yang menderita penyakit NHL

### Untuk Penyakit HOD

- Tujuan : Untuk melihat apakah terdapat perbedaan efek jenis penanganan (Allo dan Auto) pada pasien yang menderita penyakit NHL.

- Menentukan Hipotesis

$$H_0 : h_{allo(HOD)}(t) = h_{auto(HOD)}(t), t \leq \tau$$

$$H_1 : \text{tidak demikian}$$

- Menentukan taraf signifikansi

Akan digunakan  $\alpha = 0.05$

- Statistik uji

Perhatikan kode beserta output R berikut

```
> # Read data from URL
> url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo_Auto_HOD.DAT"
> dfHOD <- read.table(url, sep = ";", header = TRUE)
> # Membuat objek survival dari data
> surv_obj_HOD <- Surv(time = dfHOD$futime, event = dfHOD$fustat)
> # Melakukan uji strata
> strata_test_HOD <- survdiff(surv_obj_HOD ~ dfHOD$Type)
> # Melihat hasil uji strata
> strata_test_HOD
Call:
survdiff(formula = surv_obj_HOD ~ dfHOD$Type)

      N Observed Expected (O-E)^2/E (O-E)^2/V
dfHOD$Type=Allo  5      5     1.89    5.095    6.36
dfHOD$Type=Auto 15      7    10.11    0.955    6.36

Chisq= 6.4 on 1 degrees of freedom, p= 0.01
```

- Aturan keputusan dan kesimpulan

$H_0$  ditolak jika  $p\text{-value} < \alpha$ . Terlihat bahwa nilai  $p\text{-value} = 0.01 < \alpha = 0.05$ , sehingga  $H_0$  ditolak. Akibatnya memang terdapat perbedaan efek jenis penanganan (Allo dan Auto) pada pasien yang menderita penyakit NHL.

Beikutnya kita akan melakukan uji stratifikasi dengan membagi menjadi kelompok Allo dan Auto

### Untuk Jenis Penanganan Allo

- Tujuan : Untuk melihat apakah terdapat perbedaan efek jenis penanganan Allo antara pasien yang menderita penyakit NHL dengan antara pasien yang menderita penyakit HOD.
- Menentukan Hipotesis  
 $H_0 : h_{allo(NHL)}(t) = h_{allo(HOD)}(t), t \leq \tau$   
 $H_1 : \text{tidak demikian}$
- Menentukan taraf signifikansi  
Akan digunakan  $\alpha = 0.05$
- Statistik uji

Perhatikan kode beserta output R berikut

```
> # Read data from URL
> url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo.DAT"
> dfAllo <- read.table(url, sep = ";", header = TRUE)
> # Membuat objek survival dari data
> surv_obj_Allo <- Surv(time = dfAllo$futime, event = dfAllo$fustat)
> # Melakukan uji strata
> strata_test_Allo <- survdiff(surv_obj_Allo ~ dfAllo$Type)
> # Melihat hasil uji strata
> strata_test_Allo
Call:
survdiff(formula = surv_obj_Allo ~ dfAllo$Type)

      N Observed Expected (O-E)^2/E (O-E)^2/V
dfAllo$Type=HOD  5      5     1.86    5.31     7.05
dfAllo$Type=NHL 11      5    8.14    1.21     7.05

Chisq= 7 on 1 degrees of freedom, p= 0.008
```

- Aturan keputusan dan kesimpulan

$H_0$  ditolak jika p-value <  $\alpha$ . Terlihat bahwa nilai p-value = 0.008 <  $\alpha$  = 0.05, sehingga  $H_0$  ditolak. Akibatnya memang terdapat perbedaan efek jenis penanganan Auto antara pasien yang menderita penyakit NHL dengan antara pasien yang menderita penyakit HOD.

## Untuk Jenis Penanganan Auto

- Tujuan : Untuk melihat apakah terdapat perbedaan efek jenis penanganan Auto antara pasien yang menderita penyakit NHL dengan antara pasien yang menderita penyakit HOD.

- Menentukan Hipotesis

$$H_0 : h_{auto(NHL)}(t) = h_{auto(HOD)}(t), t \leq \tau$$

$$H_1 : \text{tidak demikian}$$

- Menentukan taraf signifikansi

Akan digunakan  $\alpha = 0.05$

- Statistik uji

Perhatikan kode beserta output R berikut

```
> # Read data from URL
> url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Auto.DAT"
> dfAuto <- read.table(url, sep = ";", header = TRUE)
> # Membuat objek survival dari data
> surv_obj_Auto <- Surv(time = dfAuto$futime, event = dfAuto$fustat)
> # Melakukan uji strata
> strata_test_Auto <- survdiff(surv_obj_Auto ~ dfAuto$Type)
> # Melihat hasil uji strata
> strata_test_Auto
Call:
survdiff(formula = surv_obj_Auto ~ dfAuto$Type)

      N Observed Expected (O-E)^2/E (O-E)^2/V
dfAuto$Type=HOD 15      7     8.78    0.362    0.807
dfAuto$Type=NHL 12      9     7.22    0.441    0.807

  chisq= 0.8 on 1 degrees of freedom, p= 0.4
```

- Aturan keputusan dan kesimpulan

$H_0$  ditolak jika p-value <  $\alpha$ . Terlihat bahwa nilai p-value = 0.4 >  $\alpha$  = 0.05, sehingga  $H_0$  gagal ditolak. Akibatnya tidak terdapat perbedaan efek jenis penanganan Auto antara pasien yang menderita penyakit NHL dengan antara pasien yang menderita penyakit HOD.

## Bagian 3. Lampiran

### Bagian 1

[https://drive.google.com/drive/folders/1G\\_f9VIH-zDLfEZEg5l9h-G1Rt2rc7ORp](https://drive.google.com/drive/folders/1G_f9VIH-zDLfEZEg5l9h-G1Rt2rc7ORp)

### Bagian 2 Nomor 1

	Group	time	status
1	NHL	28	1
2	NHL	32	1
3	NHL	49	1
4	NHL	84	1
5	NHL	357	1
6	NHL	933	0
7	NHL	1078	0
8	NHL	1183	0
9	NHL	1560	0
10	NHL	2114	0
11	NHL	2144	0
12	NHL	42	1
13	NHL	53	1
14	NHL	57	1
15	NHL	63	1
16	NHL	81	1
17	NHL	140	1
18	NHL	176	1
19	NHL	210	0
20	NHL	252	1
21	NHL	476	0

	Group	time	status
22	NHL	524	1
23	NHL	1037	0
24	HOD	2	1
25	HOD	4	1
26	HOD	72	1
27	HOD	77	1
28	HOD	79	1
29	HOD	30	1
30	HOD	36	1
31	HOD	41	1
32	HOD	52	1
33	HOD	62	1
34	HOD	108	1
35	HOD	132	1
36	HOD	180	0
37	HOD	307	0
38	HOD	406	0
39	HOD	446	0
40	HOD	484	0
41	HOD	748	0
42	HOD	1290	0

Showing 22 to 43 of 43 entries, 3 total columns

```

1 #kelompok G no 2.1
2 library(survival)
3 Data <- data.frame( Group = c(rep("NHL", length=23 ), rep("HOD", length=20)),
4                      time = c(c(28, 32, 49, 84, 357, 933, 1078, 1183, 1560, 2114,
5                               2144, 42, 53, 57, 63, 81, 140, 176, 210, 252, 476,
6                               524, 1037),c(2, 4, 72, 77, 79, 30, 36, 41, 52, 62,
7                               108, 132, 180, 307, 406, 446, 484,
8                               748, 1290, 134)),
9                      status = c(c(1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
10                           1, 1, 0, 1, 0, 1, 0), c(1, 1, 1, 1, 1, 1, 1, 1, 1,
11                           1, 1, 1, 0, 0, 0, 0, 0, 0, 0)))
12 view(Data)
13 #boxplot
14 boxplot(time ~ Group, data = Data,
15           xlab = "Group", ylab = "Time",
16           main = "Boxplot of Time by Group")
17
18 #Kaplan Meiler estimasi
19 NHL <- survfit(surv(time, status) ~ 1,
20                  data = subset(Data, Group == "NHL"))
21 HOD <- survfit(Surv(time, status) ~ 1,
22                  data = subset(Data, Group == "HOD"))
23 plot(NHL, conf.int="none", col='blue',
24       xlab = "waktu dalam hari",
25       ylab = "Probabilitas Ketahanan Hidup")
26 lines(HOD, conf.int = "none", col="red")
27 legend(1500, 1, c("NHL", "HOD"),
28         col = c('blue', 'red'), lty=1)
29 title(main = "Estimasi Kaplan Meier")
30
31 #Uji 2 sample
32 result<-survdiff(surv(time, status) ~ Group, data=Data);result

```

## Bagian 2 Nomor 2

```

1 library(readxl) #menggunakan library readxl buat membaca data dr excel
2 #mengupload data ke rstudio
3 mydata <- read_excel("C:/Users/RACHMANIA AZZAHRA S/Videos/Adobe Premiere Pro Captured Video/Semester 4/Modsur(ok.xlsx")
4 view(mydata) #melihat data
5
6 #melihat detail data
7 str(mydata)
8 head(mydata)
9 mydata
10
11 #menggunakan library ggplot buat membuat plot membandingkan waktu survival dari 2 perlakuan yang berbeda
12 library(ggplot2)
13
14 #waktu bertahan hidup seiring perubahan perlakuan dari penanganan Allo ke penanganan Auto
15 ggplot(data = mydata) +
16   geom_boxplot(aes(x = factor(trt),
17                     y = t),
18               fill = "coral",
19               alpha = .6) +
20   labs(x = "Type Treatment",
21        y = "Survival t (months)") +
22   theme(axis.title = element_text(size = 18),
23         axis.text = element_text(size = 18)) +
24   theme_light()
25
26 #Tabel fungsi survival
27 summary(mydata_surv)

```

```

29 #menggunakan library survminer buat membuat plot fungsi survival
30 library(survminer)
31 ggsurvplot(
32   mydata_surv,
33   data = mydata,
34   size = 1, # change line size
35   palette = c("red", "blue", "purple", "green"),# custom color palettes
36   conf.int = F, # Add confidence interval
37   pval = T, # Add p-value
38   risk.table = T, # Add risk table
39   risk.table.col = "strata",# Risk table color by groups
40   risk.table.height = 0.3, # useful to change when you have multiple groups
41   ggtheme = theme_bw() # change ggplot2 theme
42 )
43
44
45 #menghitung statistik uji hipotesis
46 mydata_hypo <- survdiff(formula = Surv(t, d) ~ trt,
47                         data = mydata)
48 mydata_hypo

```

## Bagian 2 Nomor 3

```

library(ggplot2)
library(survival)
library(survminer)

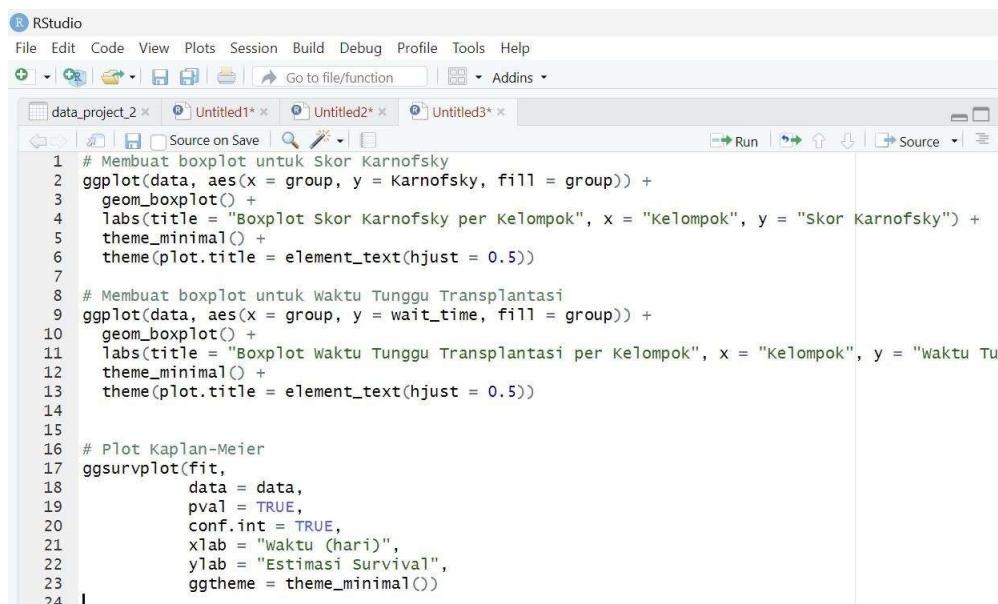
# Input data
data <- data.frame(
  time = c(
    # Auto NHL
    28, 32, 49, 84, 357, 933, 1078, 1183, 1560, 2114, 2144,
    # Auto NHL
    42, 57, 63, 81, 140, 176, 210, 252, 476, 524, 1037,
    # Auto NHL
    2, 4, 72, 77, 79,
    # Auto HOD
    30, 36, 41, 52, 62, 108, 132, 180, 307, 406, 446, 484, 748, 1290, 1345),
  status = c(
    # Auto NHL
    1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0,
    # Auto NHL
    1, 1, 1, 1, 1, 1, 0, 1, 0,
    # Auto HOD
    1, 1, 1, 1,
    # Auto HOD
    1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0),
  group = c(
    # Auto NHL
    rep('Auto NHL', 11),
    # Auto NHL
    rep('Auto NHL', 12),
    # Auto HOD
    rep('Auto HOD', 5),
    # Auto HOD
    rep('Auto HOD', 15)),
  Karnofsky = c(
    # Auto NHL
    90, 90, 60, 70, 90, 100, 90, 80, 80, 90,
    # Auto NHL
    80, 90, 30, 60, 50, 100, 80, 90, 90, 90, 90,
    # Auto HOD
    20, 30, 80, 60, 70,
    # Auto HOD
    90, 80, 60, 60, 70, 60, 100, 100, 100, 100, 90, 90, 80),
  wait_time = c(
    # Auto NHL
    24, 7, 8, 10, 42, 9, 16, 16, 20, 27, 5,
    # Auto NHL
    18, 17, 9, 13, 12, 11, 38, 16, 21, 24, 39, 84,
    # Auto HOD
    34, 28, 59, 102, 71,
    # Auto HOD
    73, 61, 34, 18, 40, 65, 17, 61, 24, 48, 52, 84, 171, 20, 98))
)

# Membuat objek survival
surv_object = survfit(Surv(time, status) ~ group)

# Melakukan uji log-rank
log_rank_test <- survdiff(surv_object ~ group, data = data)

# Membuat data frame untuk ringkasan uji log-rank
log_rank_summary <- data.frame(
  groups = names(log_rank_test$n),      # Menggunakan $n untuk nama kelompok
  observed = log_rank_test$obs,         # Jumlah observasi yang terjadi
  expected = log_rank_test$exp,         # Jumlah observasi yang diharapkan
  chisq = rep(log_rank_test$chisq, length(log_rank_test$n)), # Mengulang nilai chisq
  pvalue = rep(log_rank_test$pvalue, length(log_rank_test$n)) # Mengulang nilai pvalue)
print(log_rank_summary)

```



```
1 # Membuat boxplot untuk Skor Karnofsky
2 ggplot(data, aes(x = group, y = Karnofsky, fill = group)) +
3   geom_boxplot() +
4   labs(title = "Boxplot Skor Karnofsky per Kelompok", x = "Kelompok", y = "Skor Karnofsky") +
5   theme_minimal() +
6   theme(plot.title = element_text(hjust = 0.5))
7
8 # Membuat boxplot untuk Waktu Tunggu Transplantasi
9 ggplot(data, aes(x = group, y = wait_time, fill = group)) +
10  geom_boxplot() +
11  labs(title = "Boxplot Waktu Tunggu Transplantasi per Kelompok", x = "Kelompok", y = "Waktu Tu")
12  theme_minimal() +
13  theme(plot.title = element_text(hjust = 0.5))
14
15
16 # Plot Kaplan-Meier
17 ggsurvplot(fit,
18             data = data,
19             pval = TRUE,
20             conf.int = TRUE,
21             xlab = "Waktu (hari)",
22             ylab = "Estimasi survival",
23             ggtheme = theme_minimal())
24
```

## Bagian 2 Nomor 4

```
library(survival)
# Read data from URL
url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo_Auto_NHL.DAT"
dfNHL <- read.table(url, sep = ";", header = TRUE)

# Membuat objek survival dari data
surv_obj_NHL <- Surv(time = dfNHL$futime, event = dfNHL$fustat)
# Melakukan uji strata
strata_test_NHL <- survdiff(surv_obj_NHL ~ dfNHL$Type)
# Melihat hasil uji strata
strata_test_NHL

# Read data from URL
url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo_Auto_HOD.DAT"
dfHOD <- read.table(url, sep = ";", header = TRUE)

# Membuat objek survival dari data
surv_obj_HOD <- Surv(time = dfHOD$futime, event = dfHOD$fustat)
# Melakukan uji strata
strata_test_HOD <- survdiff(surv_obj_HOD ~ dfHOD$Type)
# Melihat hasil uji strata
strata_test_HOD
```

```
# Read data from URL
url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Allo.DAT"
dfAllo <- read.table(url, sep = ";", header = TRUE)

# Membuat objek survival dari data
surv_obj_Allo <- Surv(time = dfAllo$futime, event = dfAllo$fustat)
# Melakukan uji strata
strata_test_Allo <- survdiff(surv_obj_Allo ~ dfAllo>Type)
# Melihat hasil uji strata
strata_test_Allo

# Read data from URL
url <- "https://raw.githubusercontent.com/rahmataffandi/Modsur2/main/Auto.DAT"
dfAuto <- read.table(url, sep = ";", header = TRUE)

# Membuat objek survival dari data
surv_obj_Auto <- Surv(time = dfAuto$futime, event = dfAuto$fustat)
# Melakukan uji strata
strata_test_Auto <- survdiff(surv_obj_Auto ~ dfAuto>Type)
# Melihat hasil uji strata
strata_test_Auto
```