

# **Tarea 1**

Estudiantes

**Jhonatan Smith Garcia Muñoz**

**Diego Cano**

**Juan Manuel Sanchez**

Docente

**Juan Carlos Salazar Uribe**

Asignatura

**Introducción al análisis de supervivencia**



Sede Medellín

17 de marzo de 2023

# Índice

<b>1. Ejercicio 1</b>	<b>4</b>
<b>2. Ejercicio 2</b>	<b>4</b>
2.1. Estimacion de KM para grupo 1 . . . . .	7
2.2. Estimacion con NA - Grupo 1 . . . . .	8
2.3. Estimacion con NA - Grupo 2 . . . . .	8
2.4. ejercicio 3 . . . . .	9
<b>3. Ejercicio 4</b>	<b>10</b>

## 1. Ejercicio 1

En las siguientes situaciones identifique: el evento de interes y la variable respuesta.

### Solucion

a) Se realiza un seguimiento durante 13 años a un grupo de adultos de al menos 60 años de edad para ver que tanto tiempo permanecen vivos.

- Evento de interes: Fallecimiento
- Variable respuesta: Tiempo de supervivencia en años para los adultos de al menos 60 años de edad

b) Un grupo de pacientes con un trasplante de corazon son monitoreados para ver cuanto tiempo sobreviven

- Evento de interes: La muerte
- Tiempo de supervivencia para pacientes que han sido sometidos a trasplante de corazon.

## 2. Ejercicio 2

Resuelva a MANO, usando los datos sobre Hepatitis Crónica Activa (CAH). Estos datos son de un ensayo clinico descrito en Kirk (1980). 44 pacientes con CAH se aleatorizaron a una droga llamada prednisolona o a un grupo de control sin tratamiento. El tiempo de supervivencia en meses, despues de ser admitido al ensayo es la variables respuesta. Con estos datos obtenga una estimación de KM y el de NA. Note que debe evidenciar tablas de vida. No olvide el analisis descriptivo. Interprete

### Solucion

A continuacion se muestra ña estructura general de la base de datos. Tenga presente que se tienen ensayos clinicos para 44 pacientes, donde existen dos grupos: Control con tratamiento y control sin tratamiento. 1 Representa fallas y 0 representa las censuras. Se presentan los primeros 6 datos de la tabla de manera ilustrativa.

```
##
## -- Column specification -----
## cols(
##   id = col_double(),
##   group = col_double(),
##   time = col_double(),
##   status = col_double()
## )

## Warning: 1 parsing failure.
## row col expected actual
## 23 -- 4 columns 5 columns 'C:/Users/jhsga/OneDrive/Escritorio/GitHub/Asignaturas-UN
```

id	group	time	status
1	1	2	1
2	1	2	1
3	1	12	1
4	1	54	1
5	1	56	0
6	1	68	1

Esta informacion puede ser resumida de la siguiente manera:

Tabla 1: Tabla de vida grupo 1

$t_j$	$d_j$	$q_j$	$n_j$
$t_0=0$	0	0	22(22 unidades sobreviven $\geq 0$ meses )
$t_1=2$	2	0	22(22 unidades sobreviven $\geq 2$ meses )
$t_2=12$	1	0	20(20 unidades sobreviven $\geq 12$ meses )
$t_3=54$	1	1	19(19 unidades sobreviven $\geq 54$ meses )
$t_4=68$	1	0	17(17 unidades sobreviven $\geq 68$ meses )
$t_5=89$	1	0	16(16 unidades sobreviven $\geq 89$ meses )
$t_6=96$	2	5	15(15 unidades sobreviven $\geq 96$ meses )
$t_7=143$	1	1	8(8 unidades sobreviven $\geq 143$ meses )
$t_8=146$	1	2	6(6 unidades sobreviven $\geq 146$ meses )
$t_9=168$	1	2	3(3 unidades sobreviven $\geq 168$ meses )
TOTAL	11	11	11+11=22

Con esto, es necesario analizar el tiempo promedio  $\hat{T}$  y el Hazar  $\hat{h}$  para el grupo 1 que se define respectivamente como:

$$\bar{T} = \frac{1}{n} \sum_{i=1}^n t_i$$

$$\bar{T} = \frac{1}{22} \sum_{i=1}^{22} t_i = 109.3636$$

$$\bar{h} = \frac{11}{22(109.3636)} = 0.004571905$$

Tabla 2: Tabla de vida Grupo 2

$t_j$	$d_j$	$q_j$	$n_j$
$t_0=0$	0	0	22(22 unidades sobreviven $\geq 0$ meses)
$t_1=2$	1	0	22(22 unidades sobreviven $\geq 2$ meses )
$t_2=3$	1	0	21 (21 unidades sobreviven $\geq 3$ meses)
$t_3=4$	1	0	20(20 unidades sobreviven $\geq 4$ meses )

$t_j$	$d_j$	$q_j$	$n_j$
$t_4=7$	1	0	19(19 unidades sobreviven $\geq 7$ meses )
$t_5=10$	1	0	18(18 unidades sobreviven $\geq 10$ meses)
$t_6=22$	1	0	17(17 unidades sobreviven $\geq 22$ meses )
$t_7=28$	1	0	16(16 unidades sobreviven $\geq 28$ meses )
$t_8=29$	1	0	15(15 unidades sobreviven $\geq 29$ meses )
$t_9=32$	1	0	14(14 unidades sobreviven $\geq 32$ meses)
$t_{10}=37$	1	0	13(13 unidades sobreviven $\geq 37$ meses )
$t_{11}=40$	1	0	12(12 unidades sobreviven $\geq 40$ meses )
$t_{12}=41$	1	0	11(11 unidades sobreviven $\geq 41$ meses)
$t_{13}=54$	1	0	10(10 unidades sobreviven $\geq 54$ meses )
$t_{14}=61$	1	0	9 (9 unidades sobreviven $\geq 61$ meses )
$t_{15}=63$	1	0	8(8 unidades sobreviven $\geq 63$ meses)
$t_{16}=71$	1	6	7 unidades sobreviven $\geq 71$ meses )
TOTAL	16	6	16+6=22

Ahora para el grupo 2, el tiempo promedio  $\hat{T}$  y  $\hat{h}$  es:

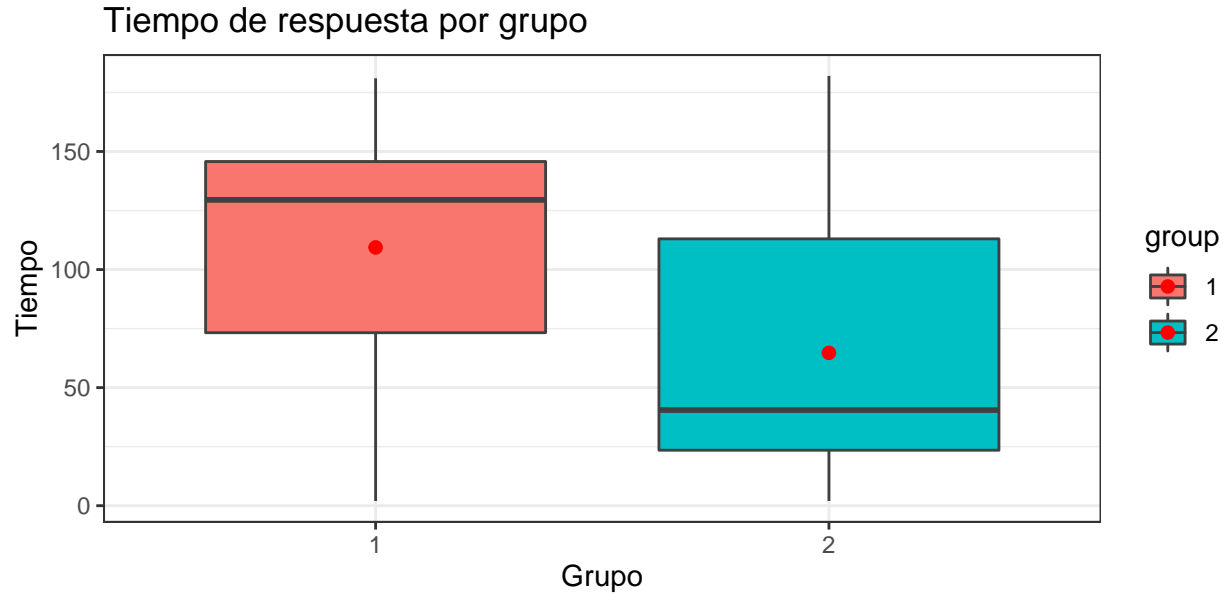
$$\bar{T} = \frac{1}{n} \sum_{i=1}^n t_i$$

$$\bar{T} = \frac{1}{22} \sum_{i=1}^{22} t_i = 64.72727$$

$$\bar{h} = \frac{11}{22(109.3636)} = 0.007724719$$

A mayor Hazard es menor la probabilidad de supervivencia (Evidentemente, por eso el nombre), en este caso, se ve claramente que  $\hat{h}_1 < \hat{h}_2$  y a su vez;  $\hat{T}_1 > \hat{T}_2$ . Basado en esto, se podría concluir que la supervivencia del grupo 2 es menor que la del grupo 1. Esto podría evidenciarse en el siguiente análisis gráfico.

```
df$group = df$group %>% as.factor()
ggplot(df, aes(x = group, y = time, group = group, fill = group)) +
  geom_boxplot( color = "#404040")+
  stat_summary(fun = mean, geom = "point", shape = 20,
               size = 3, color = "red", position = position_dodge(width = 0.75))+
  labs(title = "Tiempo de respuesta por grupo",
        x = "Grupo",
        y = "Tiempo") +
  theme_bw()
```



Note que claramente las medias de ambos procesos (puntos rojos) se encuentran uno por encima de la otra. Esto da un indicio que los individuos del grupo 1 en comparación a los individuos del grupo 2 presentan un tiempo de supervivencia promedio mayor.

En el grupo 1 el 50 % de los datos esta aproximadamente por encima de 120 meses y el 50 % por debajo, mientras que en el grupo 2 el 50 % esta aproximadamente por encima de 40 meses y el 50 % por debajo, es decir, que se presenta una diferencia significativa en el promedio de los tiempos hasta que el individuo falla entre el grupo 1 y el grupo 2

Dado esto, se procede a realizar entonces las estimaciones con KM (Kaplan-Meier):

## 2.1. Estimacion de KM para grupo 1

$t_{(j)}$	$d_j$	$q_j$	$n_j$	$\hat{S}_{KM}(t_j)$
0	0	0	22	$1 - \frac{0}{22} = 1$
2	2	0	22	$1 \times \left(1 - \frac{2}{22}\right) = \frac{10}{11} \approx 0.9091$
12	1	0	20	$\frac{10}{11} \times \left(1 - \frac{1}{20}\right) = \frac{19}{22} \approx 0.8636$
54	1	1	19	$\frac{19}{22} \times \left(1 - \frac{1}{19}\right) = \frac{9}{11} \approx 0.8182$
68	1	0	17	$\frac{9}{11} \times \left(1 - \frac{1}{17}\right) = \frac{144}{187} \approx 0.7701$
89	1	0	16	$\frac{144}{187} \times \left(1 - \frac{1}{16}\right) = \frac{135}{187} \approx 0.7221$
96	2	5	15	$\frac{135}{187} \times \left(1 - \frac{2}{15}\right) = \frac{117}{187} \approx 0.6257$
143	1	1	8	$\frac{117}{187} \times \left(1 - \frac{1}{8}\right) = \frac{819}{146} \approx 0.5475$
146	1	2	6	$\frac{819}{1496} \times \left(1 - \frac{1}{6}\right) = \frac{1365}{2992} \approx 0.4562$
168	1	2	3	$\frac{1365}{2992} \times \left(1 - \frac{1}{3}\right) = \frac{455}{1496} \approx 0.3041$

## Estimacion de KM para el grupo 2

$t_{(j)}$	$d_j$	$q_j$	$n_j$	$\hat{S}_{KM}(t_j)$
0	0	0	22	$1 - \frac{0}{22} = 1$
2	1	0	22	$1 \times \left(1 - \frac{1}{22}\right) = \frac{21}{22} \approx 0.9545$
3	1	0	21	$\frac{21}{22} \times \left(1 - \frac{1}{21}\right) = \frac{10}{11} \approx 0.9091$
4	1	0	20	$\frac{10}{11} \times \left(1 - \frac{1}{20}\right) = \frac{19}{22} \approx 0.8636$
7	1	0	19	$\frac{19}{22} \times \left(1 - \frac{1}{19}\right) = \frac{9}{11} \approx 0.8182$
10	1	0	18	$\frac{9}{11} \times \left(1 - \frac{1}{18}\right) = \frac{17}{22} \approx 0.7727$
22	1	0	17	$\frac{17}{22} \times \left(1 - \frac{1}{17}\right) = \frac{8}{11} \approx 0.7273$
28	1	0	16	$\frac{8}{11} \times \left(1 - \frac{1}{16}\right) = \frac{15}{22} \approx 0.6818$
29	1	0	15	$\frac{15}{22} \times \left(1 - \frac{1}{15}\right) = \frac{7}{11} \approx 0.6364$
32	1	0	14	$\frac{7}{11} \times \left(1 - \frac{1}{14}\right) = \frac{13}{22} \approx 0.5909$
37	1	0	13	$\frac{13}{22} \times \left(1 - \frac{1}{13}\right) = \frac{6}{11} \approx 0.5455$
40	1	0	12	$\frac{6}{11} \times \left(1 - \frac{1}{12}\right) = \frac{1}{2} \approx 0.5$
41	1	0	11	$\frac{1}{2} \times \left(1 - \frac{1}{11}\right) = \frac{5}{11} \approx 0.4545$
54	1	0	10	$\frac{5}{11} \times \left(1 - \frac{1}{10}\right) = \frac{9}{22} \approx 0.4091$
61	1	0	9	$\frac{9}{22} \times \left(1 - \frac{1}{9}\right) = \frac{4}{11} \approx 0.3636$
63	1	0	8	$\frac{4}{11} \times \left(1 - \frac{1}{8}\right) = \frac{7}{22} \approx 0.3182$
71	1	6	7	$\frac{7}{22} \times \left(1 - \frac{1}{7}\right) = \frac{3}{11} \approx 0.2727$

## 2.2. Estimacion con NA - Grupo 1

Falla	Número	meses	$\hat{h}(t)$	$\hat{H}(t)$	$S(t)$
1-2	2		(2)/(22)	(2)/(22)~0.0909	0.9131
3	12		(1)/(20)	(1)/(22)+(1)/(20)=0.3586	0.6414
4	54		(1)/(19)	(31)/(220)+(1)/(19)=0.8209	0.1791
5	68		(1)/(17)	(809)/(4180)+(1)/(17)=0.7601	0.2399
6	89		(1)/(16)	(17933)/(77960)+(1)/(16)=0.7209	0.2791
7-8	96		(2)/(15)	(89497)/(284240)+(2)/(15)=0.6385	0.3615
9	143		(1)/(8)	(382187)/(852720)+(1)/(8)=0.5638	0.4362
10	146		(1)/(6)	0.5732+(1)/(6)=0.7397	0.2603
11	168		(1)/(3)	0.7399+(1)/(3)=1.0742	0.9258

## 2.3. Estimacion con NA - Grupo 2

Falla	Número	meses	$\hat{h}(t)$	$\hat{H}(t)$	$S(t)$
1	2		(1)/(22)	(1)/(22)~0.0455	0.9556
2	3		(1)/(21)	(1)/(22)+(1)/(21)=0.9111	0.0889
3	4		(1)/(20)	(43)/(462)+(1)/(20)=0.8667	0.1333

Falla	Número	meses	hat(h)_(t)	hat(H)_(t)	S(t)
	4	7	(1)/(19)	(661)/(4620)+(1)/(18223)	(17797)/(87780)~~0.195
	5	10	(1)/(18)	(17)/(87799)+(1)/(18778)	(66167)/(263340)~~0.25
	6	22	(1)/(17)	~~0.3101	0.7334
	7	28	(1)/(16)	~~0.3716	0.6896
	8	29	(1)/(15)	~~0.4393	0.6445
	9	32	(1)/(14)	~~0.5107	0.6001
	10	37	(1)/(13)	~~0.5876	0.5557
	11	40	(1)/(12)	~~0.6709	0.5112
	12	41	(1)/(11)	~~0.7618	0.4668
	13	54	(1)/(10)	~~0.8618	0.4224
	14	61	(1)/(9)	~~0.9730	0.3779
	15	63	(1)/(8)	~~1.0980	0.3335
	16	71	(1)/(7)	~~1.2408	0.2891

## 2.4. ejercicio 3

Usando SAS o R y los datos del numeral anterior, obtenga estimaciones para la curva de supervivencia usando KM y NA

```
df$group<-as.factor(df$group)
df=data.frame(df)
KM_fit <- survfit(Surv(time, status) ~ group, data = df)
summary(KM_fit)
```

```
## Call: survfit(formula = Surv(time, status) ~ group, data = df)
##
##               group=1
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    2      22       2   0.909  0.0613    0.797    1.000
##   12      20       1   0.864  0.0732    0.732    1.000
##   54      19       1   0.818  0.0822    0.672    0.996
##   68      17       1   0.770  0.0904    0.612    0.969
##   89      16       1   0.722  0.0967    0.555    0.939
##   96      15       2   0.626  0.1051    0.450    0.870
##  143       8       1   0.547  0.1175    0.359    0.834
##  146       6       1   0.456  0.1285    0.263    0.793
##  168       3       1   0.304  0.1509    0.115    0.804
##
##               group=2
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    2      22       1   0.955  0.0444    0.871    1.000
##    3      21       1   0.909  0.0613    0.797    1.000
##    4      20       1   0.864  0.0732    0.732    1.000
```



##	7	19	1	0.818	0.0822	0.672	0.996
##	10	18	1	0.773	0.0893	0.616	0.969
##	22	17	1	0.727	0.0950	0.563	0.939
##	28	16	1	0.682	0.0993	0.513	0.907
##	29	15	1	0.636	0.1026	0.464	0.873
##	32	14	1	0.591	0.1048	0.417	0.837
##	37	13	1	0.545	0.1062	0.372	0.799
##	40	12	1	0.500	0.1066	0.329	0.759
##	41	11	1	0.455	0.1062	0.288	0.718
##	54	10	1	0.409	0.1048	0.248	0.676
##	61	9	1	0.364	0.1026	0.209	0.632
##	63	8	1	0.318	0.0993	0.173	0.587
##	71	7	1	0.273	0.0950	0.138	0.540

### 3. Ejercicio 4

Usando de nuevo los datos de CAH y A MANO obtenga estimaciones para  $H(t)$  usando  $-\text{Log}(S(t))$  y NA. Grafique las estimaciones a mano. Note que se deben evidenciar las tablas usadas en el proceso. Interprete.

*Solución*

\*Estimación con  $-\text{Log}(S(t))$  - Grupo 1

$t_{(j)}$	$d_j$	$q_j$	$n_j$	$\hat{S}_{KM}(t_j)$	$\hat{H}_{NA}(t_j)$
0	0	0	22	$1 - \frac{0}{22} = 1$	0
2	2	0	22	$1x \left(1 - \frac{2}{22}\right) = \frac{10}{11} \approx 0.9091$	0.09530
12	1	0	20	$\frac{10}{11}x \left(1 - \frac{1}{20}\right) = \frac{19}{22} \approx 0.8636$	0.14664
54	1	1	19	$\frac{19}{22}x \left(1 - \frac{1}{19}\right) = \frac{9}{11} \approx 0.8182$	0.20065
68	1	0	17	$\frac{9}{11}x \left(1 - \frac{1}{17}\right) = \frac{144}{187} \approx 0.7701$	0.26123
89	1	0	16	$\frac{144}{187}x \left(1 - \frac{1}{16}\right) = \frac{155}{187} \approx 0.7221$	0.32559
96	2	5	15	$\frac{135}{187}x \left(1 - \frac{2}{15}\right) = \frac{117}{187} \approx 0.6257$	0.46888
143	1	1	8	$\frac{117}{187}x \left(1 - \frac{1}{8}\right) = \frac{819}{1496} \approx 0.5475$	0.6024
146	1	2	6	$\frac{819}{1496}x \left(1 - \frac{1}{6}\right) = \frac{165}{2952} \approx 0.4562$	0.7848
168	1	2	3	$\frac{1365}{2992}x \left(1 - \frac{1}{3}\right) = \frac{455}{1496} \approx 0.3041$	1.1904

Estimación con  $-\text{Log}(S(t))$  - Grupo 2

$t_{(j)}$	$d_j$	$q_j$	$n_j$	$\hat{S}_{KM}(t_j)$	$\hat{H}_{NA}(t_j)$
0	0	0	22	$1 - \frac{0}{22} = 1$	0
2	1	0	22	$1 \times \left(1 - \frac{1}{22}\right) = \frac{21}{22} \approx 0.9545$	0.04657
3	1	0	21	$\frac{2}{22} \times \left(1 - \frac{1}{21}\right) = \frac{10}{11} \approx 0.9091$	0.0953
4	1	0	20	$\frac{10}{11} \times \left(1 - \frac{1}{20}\right) = \frac{19}{22} \approx 0.8636$	0.14665
7	1	0	19	$\frac{19}{22} \times \left(1 - \frac{1}{19}\right) = \frac{9}{11} \approx 0.8182$	0.20065
10	1	0	18	$\frac{9}{11} \times \left(1 - \frac{1}{18}\right) = \frac{17}{22} \approx 0.7727$	0.25786
22	1	0	17	$\frac{17}{22} \times \left(1 - \frac{1}{17}\right) = \frac{8}{11} \approx 0.7273$	0.31842
28	1	0	16	$\frac{8}{11} \times \left(1 - \frac{1}{16}\right) = \frac{15}{22} \approx 0.6818$	0.38302
29	1	0	15	$\frac{15}{22} \times \left(1 - \frac{1}{15}\right) = \frac{7}{11} \approx 0.6364$	0.45193
32	1	0	14	$\frac{7}{11} \times \left(1 - \frac{1}{14}\right) = \frac{13}{22} \approx 0.5909$	0.52611
37	1	0	13	$\frac{13}{22} \times \left(1 - \frac{1}{13}\right) = \frac{6}{11} \approx 0.5455$	0.60605
40	1	0	12	$\frac{6}{11} \times \left(1 - \frac{1}{12}\right) = \frac{1}{2} \approx 0.5$	0.69315
41	1	0	11	$\frac{1}{2} \times \left(1 - \frac{1}{11}\right) = \frac{5}{11} \approx 0.4545$	0.78856
54	1	0	10	$\frac{5}{11} \times \left(1 - \frac{1}{10}\right) = \frac{9}{22} \approx 0.4091$	0.8938
61	1	0	9	$\frac{9}{22} \times \left(1 - \frac{1}{9}\right) = \frac{4}{14} \approx 0.3636$	1.0117
63	1	0	8	$\frac{4}{14} \times \left(1 - \frac{1}{8}\right) = \frac{7}{22} \approx 0.3182$	1.14508
71	1	6	7	$\frac{7}{22} \times \left(1 - \frac{1}{7}\right) = \frac{3}{11} \approx 0.2727$	1.29938

Despues de esto, se procede a realizar la estimacion de  $H(t)$  por el medio de Nelson-Alen. Se adjunta el resultado en la siguiente tabla.

- Estimación con NA - Grupo 1:

Falla Número	meses	$\hat{h}_t$	$\hat{H}_t$
1 - 2	2	$\frac{2}{22}$	$\frac{2}{22} \approx 0.0909$
3	12	$\frac{1}{20}$	$\frac{2}{22} + \frac{1}{20} = \frac{31}{220} \approx 0.1409$
4	54	$\frac{1}{19}$	$\frac{31}{220} + \frac{1}{19} = \frac{809}{8997} \approx 0.1935$
5	68	$\frac{1}{17}$	$\frac{809}{8997} + \frac{1}{17} = \frac{193033}{771060} \approx 0.2524$
6	89	$\frac{1}{16}$	$\frac{17933}{8997} + \frac{1}{16} = \frac{24240}{382187} \approx 0.3149$
7 - 8	96	$\frac{2}{15}$	$\frac{77160}{89497} + \frac{2}{15} = \frac{24240}{152720} \approx 0.4482$
9	143	$\frac{1}{8}$	$\frac{382187}{852720} + \frac{1}{8} \approx 0.5732$
10	146	$\frac{1}{6}$	$0.5732 + \frac{1}{6} \approx 0.7399$
11	168	$\frac{1}{3}$	$0.7399 + \frac{1}{3} \approx 1.0732$

- Estimación con NA - Grupo 2:

Falla Número	meses	$\hat{h}_t$	$\hat{H}_t$
1	2	$\frac{1}{22}$	$\approx 0.0455$
2	3	$\frac{1}{21}$	$\frac{1}{22} + \frac{1}{21} = \frac{43}{462} \approx 0.0931$
3	4	$\frac{1}{20}$	$\frac{43}{462} + \frac{1}{20} = \frac{661}{4620} \approx 0.1431$
4	7	$\frac{1}{19}$	$\frac{661}{4620} + \frac{1}{19} = \frac{1779}{87880} \approx 0.1957$
5	10	$\frac{1}{18}$	$\frac{1779}{87880} + \frac{1}{18} = \frac{66167}{661670} \approx 0.2513$
6	22	$\frac{1}{17}$	$\frac{66167}{661670} + \frac{1}{17} = \frac{263340}{661670} \approx 0.3101$
7	28	$\frac{1}{16}$	$0.3101 + \frac{1}{16} \approx 0.3716$
8	29	$\frac{1}{15}$	$0.3716 + \frac{1}{15} \approx 0.4393$
9	32	$\frac{1}{14}$	$0.4393 + \frac{1}{14} \approx 0.5107$
10	37	$\frac{1}{13}$	$0.5107 + \frac{1}{13} \approx 0.5876$
11	40	$\frac{1}{12}$	$0.5876 + \frac{1}{12} \approx 0.6709$
12	41	$\frac{1}{11}$	$0.6709 + \frac{1}{11} \approx 0.7618$
13	54	$\frac{1}{10}$	$0.7618 + \frac{1}{10} \approx 0.8618$
14	61	$\frac{1}{9}$	$0.8618 + \frac{1}{9} \approx 0.9730$
15	63	$\frac{1}{8}$	$0.9730 + \frac{1}{8} \approx 1.0980$
16	71	$\frac{1}{7}$	$1.0980 + \frac{1}{7} \approx 1.2408$