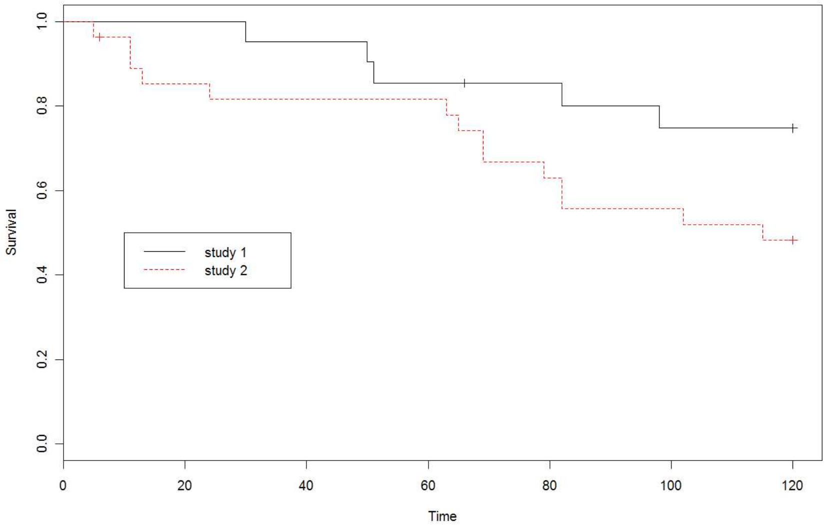
Example: Sea sickness 2

- The "rocking" study on sea sickness was repeated with 28 subjects.
- 2 times higher frequency (0.333 Hz) and acceleration (0.222 G).
- From the table, it seems that the times are shorter in the second study!

Time (min.)	Vomit (yes = 1)	Study
30	1	1
50	1	1
50	0	1
...	...	...
120	0	1
5	1	2
6	0	2
11	1	2
...	...	...
120	0	2

```
> study<-c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,
+ 2,2,2,2,2,2,2,2)
> Time<-c(30,50,50,51,66,82,98,120,120,120,120,120,120,120,120,120,120,120,120,120,120,5,6,11,11,
+ 13,24,63,65,69,69,79,82,82,102,115,120,120,120,120,120,120,120,120,120,120,120,120,120)
> Status<-c(1,1,0,1,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,
+ 0,0,0,0,0,0,0)

> fit<-survfit(Surv(Time,Status)~study)
> summary(fit)
Call: survfit(formula = Surv(Time, Status) ~ study)

                study=1
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   30     21       1    0.952  0.0465        0.866        1.000
   50     20       1    0.905  0.0641        0.788        1.000
   51     18       1    0.854  0.0778        0.715        1.000
   82     16       1    0.801  0.0894        0.644        0.997
   98     15       1    0.748  0.0981        0.578        0.967

                study=2
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    5     28       1    0.964  0.0351        0.898        1.000
   11     26       2    0.890  0.0599        0.780        1.000
   13     24       1    0.853  0.0679        0.730        0.997
   24     23       1    0.816  0.0744        0.682        0.976
   63     22       1    0.779  0.0797        0.637        0.952
   65     21       1    0.742  0.0841        0.594        0.926
   69     20       2    0.668  0.0906        0.512        0.871
   79     18       1    0.630  0.0928        0.472        0.841
   82     17       2    0.556  0.0956        0.397        0.779
  102     15       1    0.519  0.0961        0.361        0.746
  115     14       1    0.482  0.0962        0.326        0.713
```

```
> plot(fit[1],xlab="Time",ylab="Survival")
> lines(fit[2],col="red",lty=2)
> legend(10,0.5,legend=c("study 1","study 2"),lty=c(1,2),col=c("black","red"))
```



```
data vomit2;
input time vomit study;
cards;
30  1 1
50  1 1
50  0 1
51  1 1
66  0 1
82  1 1
98  1 1
...
120 0 2
120 0 2
120 0 2
120 0 2
120 0 2
120 0 2
;
run;

proc lifetest data=vomit2;
by study;
time time*vomit(0);
run;
```

study=1

The LIFETEST Procedure

Product-Limit Survival Estimates

time	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000	1.0000	0	0	0	21
30.000	0.9524	0.0476	0.0465	1	20
50.000	0.9048	0.0952	0.0641	2	19
50.000*	.	.	.	2	18
51.000	0.8545	0.1455	0.0778	3	17
66.000*	.	.	.	3	16
82.000	0.8011	0.1989	0.0894	4	15
98.000	0.7477	0.2523	0.0981	5	14
120.000*	.	.	.	5	13
120.000*	.	.	.	5	12
120.000*	.	.	.	5	11
120.000*	.	.	.	5	10
120.000*	.	.	.	5	9
120.000*	.	.	.	5	8
120.000*	.	.	.	5	7
120.000*	.	.	.	5	6
120.000*	.	.	.	5	5
120.000*	.	.	.	5	4
120.000*	.	.	.	5	3
120.000*	.	.	.	5	2
120.000*	.	.	.	5	1
120.000*	.	.	.	5	0

NOTE: The marked survival times are censored observations.

Summary Statistics for Time Variable time

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower	Upper)
75	.	.	.
50	.	.	.
25	98.000	51.000	.

Mean	Standard Error
89.259	4.789

NOTE: The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.

Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
21	5	16	76.19

study=2

Product-Limit Survival Estimates

time	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000	1.0000	0	0	0	28
5.000	0.9643	0.0357	0.0351	1	27
6.000*	.	.	.	1	26
11.000	.	.	.	2	25
11.000	0.8901	0.1099	0.0599	3	24
13.000	0.8530	0.1470	0.0679	4	23
24.000	0.8159	0.1841	0.0744	5	22
63.000	0.7788	0.2212	0.0797	6	21
65.000	0.7418	0.2582	0.0841	7	20
69.000	.	.	.	8	19
69.000	0.6676	0.3324	0.0906	9	18
79.000	0.6305	0.3695	0.0928	10	17
82.000	.	.	.	11	16
82.000	0.5563	0.4437	0.0956	12	15
102.000	0.5192	0.4808	0.0961	13	14
115.000	0.4821	0.5179	0.0962	14	13
120.000*	.	.	.	14	12
120.000*	.	.	.	14	11
120.000*	.	.	.	14	10
120.000*	.	.	.	14	9
120.000*	.	.	.	14	8
120.000*	.	.	.	14	7
120.000*	.	.	.	14	6
120.000*	.	.	.	14	5
120.000*	.	.	.	14	4
120.000*	.	.	.	14	3
120.000*	.	.	.	14	2
120.000*	.	.	.	14	1
120.000*	.	.	.	14	0

Summary Statistics for Time Variable time

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower	Upper)
75	.	.	.
50	115.000	69.000	.
25	65.000	13.000	82.000

Mean	Standard Error
84.739	7.709

NOTE: The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.

Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
28	14	14	50.00

Difference in survival curves at fixed points

At a fixed point t_0, we look at the different survival probabilities,

$$H_0: \; S_1(t_0) = S_2(t_0) = \ldots = S_K(t_0) \quad \text{vs} \quad H_a: \text{at least one} \neq .$$

Like in regression, we rewrite this hypothesis as

$$H_0: CS(t_0) = 0$$

with $K-1 \times K$ contrast matrix C and vector $S(t_0)$ given by

$$C = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & -1 \\ 0 & 1 & 0 & \ldots & 0 & -1 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -1 \end{bmatrix} \quad \text{and} \quad S(t_0) = \begin{bmatrix} S_1(t_0) \\ S_2(t_0) \\ \vdots \\ S_K(t_0) \end{bmatrix}.$$

Since we have that

$$\hat{S}_j(t_0) \approx N(S_j(t_0), \mathrm{Var}(S_j(t_0))), \quad j = 1, \ldots, K$$

Hence, under H_0,

$$T = (C\hat{S}(t_0))^t [C\hat{V}C^t]^{-1} (C\hat{S}(t_0)) \approx \chi^2_{K-1}$$

with $\hat{V}$ diagonal matrix containing $\widehat{\mathrm{Var}}(\hat{S}_j(t_0))), j = 1, \ldots, K$.

Remarks:

- Simple method to compare survival curves, but fixed point t_0 has to be chosen in advance!
- More time points $t_{01}, \ldots, t_{0m} \Rightarrow$ correction multiple testing (ex: Bonferroni).

Returning to the example, we compare the survival probability in both groups at time-points $t_0 = 30, 60, 90$.

We have the hypotheses,

$$H_0 : S_1(t_0) = S_2(t_0) \quad \text{vs} \quad H_a : S_1(t_0) \neq S_2(t_0).$$

and the test is, under H_0,

$$T = \frac{(\hat{S}_1(t_0) - \hat{S}_2(t_0))^2}{\widehat{\mathrm{Var}}(\hat{S}_1(t_0)) + \widehat{\mathrm{Var}}(\hat{S}_2(t_0))} \approx \chi^2_1.$$

t_0	$\hat{S}_1(t_0)$	$\hat{S}_2(t_0)$	T	p-value
30	0.9524 (0.0465)	0.8159 (0.0744)	2.4205	0.1198
60	0.8545 (0.0778)	0.8159 (0.0744)	0.1286	0.7199
90	0.7477 (0.0981)	0.5563 (0.0956)	1.9525	0.1623

Log-rank test

Instead of looking at fixed time points, we want to compare the whole survival function of different groups,

$$H_0 : S_1(t) = S_2(t) = \ldots = S_K(t),\ 0 < t < \tau.$$

Since the true survival functions are unknown in each group, we go for a nonparametric test.

Before we derive the test statistic for this nulhypothesis. We look at a simpler setting with $K = 2$,

$$H_0 : S_1(t) = S_2(t),\ 0 < t < \tau.$$

In this setting it is more clear which idea's we will use here.

Suppose we observe, from populations $j = 0, 1$,

$$(T_{j1}, \delta_{j1}), (T_{j2}, \delta_{j2}), \ldots, (T_{jn_j}, \delta_{jn_j}).$$

Under H_0, both populations are equal. Hence we can order the uncensored lifetimes. Denote by $\tau_1, \ldots, \tau_k$ be the k ordered, distinct death times.

At the l-th death time, we construct a 2×2 contingency table,

Popul / Died	yes	no	Total
0	d_{0l}	$n_{0l} - d_{0l}$	n_{0l}
1	d_{1l}	$n_{1l} - d_{1l}$	n_{1l}
Total	d_l	$n_l - d_l$	n_l

where d_{jl} is the number of deaths and n_{jl} is the number at risk in population j at this time.

Now, under H_0 and conditional on the marginals, d_{jl} has a hypergeometric distribution.

$$d_{jl} \sim \text{Hypergeometric}(n_l, d_l, n_{jl}).$$

Therefore, we find that the mean and variance of d_{jl} are given by

$$\begin{aligned} \mathrm{E}[d_{jl}] &= \frac{n_{1l} d_l}{n_l} \\ \mathrm{Var}(d_{jl}) &= \frac{n_l - n_{1l}}{n_l - 1} n_{1l} \frac{d_l}{n_l} \left(1 - \frac{d_l}{n_l}\right) \\ &= \frac{n_{1l} n_{0l} d_l (n_l - d_l)}{n_l^2 (n_l - 1)}. \end{aligned}$$

If the sample size at each death time is sufficiently large, we approximate the hypergeometric distribution by a normal distribution.

$$d_{1l} - \mathrm{E}[d_{1l}] = d_{1l} - \frac{n_{1l}d_l}{n_l} \approx N\left(0, \frac{n_{1l}n_{0l}d_l(n_l - d_l)}{n_l^2(n_l - 1)}\right).$$

Assuming that the contingency tables at different death times are independent, we find the log-rank test,

$$T = \frac{\left[\sum_{l=1}^{k}\left(d_{1l} - \frac{n_{1l}d_l}{n_l}\right)\right]^2}{\sum_{l=1}^{k}\frac{n_{1l}n_{0l}d_l(n_l-d_l)}{n_l^2(n_l-1)}}$$

which is, under H_0, approximately χ^2 distributed with df 1.

Some remarks:

- By the hypergeometric distribution, we get in the denominator a finite sample correction factor.
- The numerator can be interpreted as $\sum_{l=1}^{k}(O_l - E_l)$ where O_l is the observed number of deaths in group 1, and E_l is the expected number, given the risk set. furthermore E_l is the proportion of deaths in group 1 among those at risk.
- It does not matter which group we choose to sum over because $\sum_{\text{groups}}(O_l - E_l) = 0$.

```
proc lifetest data=vomit2;
time time*vomit(0);
strata study;
run;

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for time over Strata

        Rank Statistics

study       Log-Rank    Wilcoxon

1            -3.8607     -149.00
2             3.8607      149.00

Covariance Matrix for the Log-Rank Statistics

study             1             2

1           4.64782      -4.64782
2          -4.64782       4.64782

...

       Test of Equality over Strata

                                      Pr >
Test      Chi-Square      DF    Chi-Square

Log-Rank      3.2069       1      0.0733
Wilcoxon      3.1816       1      0.0745
-2Log(LR)     3.4928       1      0.0616
```

The log-rank test is a special case of the Tarone-Ware class of tests.

$$T = \frac{\left[\sum_{l=1}^{k} w_l \left(d_{1l} - \frac{n_{1l} d_l}{n_l}\right)\right]^2}{\sum_{l=1}^{k} w_l^2 \frac{n_{1l} n_{0l} d_l (n_l - d_l)}{n_l^2 (n_l - 1)}}$$

where $w_l \geq 0$ are weights.

Test	w_l
Log-rank	1
Wilcoxon or Gehan	n_l
Peto-peto	$\tilde{S}(t_i)$
Harrington-Fleming (p,q)	$\hat{S}(t_i)^p (1 - \hat{S}(t_i))^q, \ p, q \geq 0$

with $\tilde{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i + 1}\right)$.

In a practical data analysis, the choice of the weights is important.

For example,

- Log-rank test has optimal power to detect alternatives in which the hazard are proportional.
- Wilcoxon-Gehan test is more "sensitive" to "early" differences in survival curves (at later times)
- Harrington-Fleming test with $p = 0, q > 0$ is "sensitive" to "late" differences.

Note: the decision about the choice of the test should be made before seeing the data!

```
proc lifetest data=vomit2;
time time*vomit(0);
strata study/test=(logrank wilcoxon peto);
run;

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for time over Strata

              Rank Statistics

study       Log-Rank    Wilcoxon        Peto

1            -3.8607     -149.00     -3.0632
2             3.8607      149.00      3.0632

...

Covariance Matrix for the Peto Statistics

study            1             2

1          2.94876      -2.94876
2         -2.94876       2.94876

       Test of Equality over Strata

                                   Pr >
Test      Chi-Square      DF    Chi-Square

Log-Rank      3.2069       1      0.0733
Wilcoxon      3.1816       1      0.0745
Peto          3.1822       1      0.0744
```

```
> survdiff(Surv(Time,Status)~study)
Call: survdiff(formula = Surv(Time, Status) ~ study)

         N Observed Expected (O-E)^2/E (O-E)^2/V
study=1 21        5     8.86      1.68      3.21
study=2 28       14    10.14      1.47      3.21

Chisq= 3.2  on 1 degrees of freedom, p= 0.0733

> survdiff(Surv(Time,Status)~study,rho=1)
Call: survdiff(formula = Surv(Time, Status) ~ study, rho = 1)

         N Observed Expected (O-E)^2/E (O-E)^2/V
study=1 21     4.01      7.2      1.41      3.22
study=2 28    11.49      8.3      1.22      3.22

Chisq= 3.2  on 1 degrees of freedom, p= 0.0728
```

When we have K populations, we generalize the previous results.

At the l-th ordered death time, we now construct a $K \times 2$ contingency table,

Popul / Died	yes	no	Total
1	d_{1l}	$n_{1l} - d_{1l}$	n_{1l}
2	d_{2l}	$n_{2l} - d_{2l}$	n_{2l}
...			...
K	d_{Kl}	$n_{Kl} - d_{Kl}$	n_{Kl}
Total	d_l	$n_l - d_l$	n_l

where d_{jl} is the number of deaths and n_{jl} is the number at risk in population j at this time.

Under H_0, the vector $\mathbf{O}_l$ of the observed deaths in groups 1 to $K-1$ at time l,

$$\mathbf{O}_l = (d_{1l}, \ldots, d_{K-1})$$

has a multivariate hypergeometric distribution with mean $\mathbf{E}_l$ and covariance matrix $\mathbf{V}_l$, given by

$$\begin{aligned}
\mathbf{E}_l &= \left(\frac{d_l n_{1l}}{n_l}, \ldots, \frac{d_l n_{(K-1)l}}{n_l}\right) \\
\mathbf{V}_{jjl} &= \frac{n_{jl}(n_l - n_{jl}) d_l (n_l - d_l)}{n_l^2 (n_l - 1)} \\
\mathbf{V}_{jml} &= \frac{-n_{jl} n_{ml} d_l (n_l - d_l)}{n_l^2 (n_l - 1)}.
\end{aligned}$$

Approximating this distribution by a multivariate normal distribution and using an analogous construction as before, we get a test statistic T,

$$T = (\mathbf{O} - \mathbf{E})\mathbf{V}^{-1}(\mathbf{O} - \mathbf{E})^t$$

which is approximately χ^2 distributed with df $K - 1$.

Hereby $\mathbf{O} = \sum_{l=1}^{k} \mathbf{O}_l$, $\mathbf{E} = \sum_{l=1}^{k} \mathbf{E}_l$, $\mathbf{V} = \sum_{l=1}^{k} \mathbf{V}_l$ are the sums over the k distinct death times.

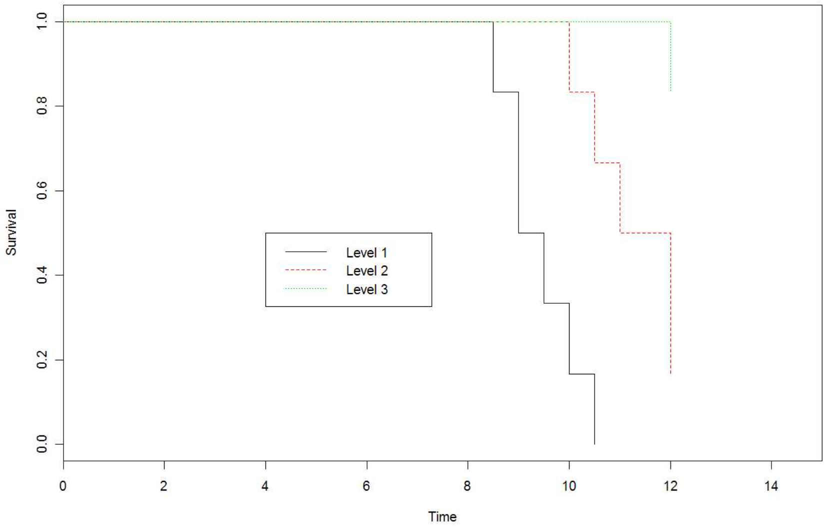
Example: Performance testing

- Test-subject were asked to perform a certain test and the time needed was recorded.
- 3 different noise distractions are applied.
- Did the different noise distractions influence the time to finish the test?

Noise level		
9.0	10.0	12.0
9.5	12.0	12.0^+
9.0	12.0^+	12.0^+
8.5	11.0	12.0^+
10.0	12.0	12.0^+
10.5	10.5	12.0^+

```
> Level<-c(1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3)
> Time<-c(9.0,9.5,9.0,8.5,10.0,10.5,10.0,12.0,12.0,11.0,12.0,10.5,12.0,12.0,12.0,12.0,12.0,12.0)
> Censor<-c(1,1,1,1,1,1,1,1,0,1,1,1,1,0,0,0,0,0)

> survfit(Surv(Time,Censor)~Level)
Call: survfit(formula = Surv(Time, Censor) ~ Level)

        n events median 0.95LCL 0.95UCL
Level=1 6      6   9.25     9.0     Inf
Level=2 6      5  11.50    10.5     Inf
Level=3 6      1    Inf     Inf     Inf

> fit<-survfit(Surv(Time,Censor)~Level)
> summary(fit)
Call: survfit(formula = Surv(Time, Censor) ~ Level)

                Level=1
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
  8.5      6       1    0.833   0.152       0.5827        1.000
  9.0      5       2    0.500   0.204       0.2246        1.000
  9.5      3       1    0.333   0.192       0.1075        1.000
 10.0      2       1    0.167   0.152       0.0278        0.997
 10.5      1       1    0.000      NA           NA           NA

                Level=2
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
 10.0      6       1    0.833   0.152       0.5827        1.000
 10.5      5       1    0.667   0.192       0.3786        1.000
 11.0      4       1    0.500   0.204       0.2246        1.000
 12.0      3       2    0.167   0.152       0.0278        0.997

                Level=3
 time       n.risk      n.event     survival      std.err lower 95% CI upper 95% CI
 12.000       6.000       1.000        0.833        0.152        0.583        1.000
```

```
> plot(fit[1],xlab="Time",ylab="Survival",xlim=c(0.0,15.0))
> lines(fit[2],col="red",lty=2)
> lines(fit[3],col="green",lty=3)
> legend(4,0.5,legend=c("Level 1","Level 2","Level 3"),lty=c(1,2,3),col=c("black","red","green"))

> survdiff(Surv(Time,Censor)~Level)
Call: survdiff(formula = Surv(Time, Censor) ~ Level)

        N Observed Expected (O-E)^2/E (O-E)^2/V
Level=1 6        6     1.57   12.4463   17.2379
Level=2 6        5     4.53    0.0488    0.0876
Level=3 6        1     5.90    4.0660    9.4495

 Chisq= 20.4  on 2 degrees of freedom, p= 3.75e-05

> survdiff(Surv(Time,Censor)~Level,rho=1)
Call: survdiff(formula = Surv(Time, Censor) ~ Level, rho = 1)

        N Observed Expected (O-E)^2/E (O-E)^2/V
Level=1 6     5.17     1.39   10.2756   16.2355
Level=2 6     3.00     3.28    0.0235    0.0536
Level=3 6     0.50     4.00    3.0625    8.4750

Chisq= 18.3  on 2 degrees of freedom, p= 0.000105
```



Survival
1.0
0.8
0.6
0.4
0.2
0.0
Time
0
2
4
6
8
10
12
14
Level 1
Level 2
Level 3

```
data noise;
input Time censor level;
cards;
9.0 1 1
9.5 1 1
9.0 1 1
8.5 1 1
10.0 1 1
10.5 1 1
10.0 1 2
12.0 1 2
12.0 0 2
11.0 1 2
12.0 1 2
10.5 1 2
12.0 1 3
12.0 0 3
12.0 0 3
12.0 0 3
12.0 0 3
12.0 0 3
;
run;

proc lifetest data=noise;
time Time*censor(0);
strata level/test=(logrank wilcoxon peto);
run;
```

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for Time over Strata

Rank Statistics

level	Log-Rank	Wilcoxon	Peto
1	4.4261	68.000	3.4232
2	0.4703	-5.000	-0.3476
3	-4.8964	-63.000	-3.0756

...

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	20.3844	2	<.0001
Wilcoxon	18.3265	2	0.0001
Peto	18.0014	2	0.0001

Sometimes we want to compare survival curves for three or more ordered groups. For example: different tumor stages (T0,T1,T2), age-groups,...

$$H_0: \quad S_1(t) = S_2(t) = \ldots = S_K(t), t \leq \tau$$

$$H_a: \quad \begin{cases} S_1(t) \geq S_2(t) \geq \ldots \geq S_K(t), t \leq \tau, \text{ at least one } > \\ \text{or} \\ S_1(t) \leq S_2(t) \leq \ldots \leq S_K(t), t \leq \tau, \text{ at least one } < \end{cases}$$

In such case one may use a logrank test for trend:

$$T = \frac{\sum_{j=1}^{K} a_j(O_j - E_j)}{\sqrt{\sum_{j=1}^{K}\sum_{k=1}^{K} a_j a_k V_{jk}}} \approx N(0,1), \text{ under } H_0,$$

where $a_1 < a_2 < \ldots < a_K$ are ordered of scores (mostly $a_j = j$) and

$$
\begin{aligned}
O_j &= \sum_{l=1}^{k} d_{jl} \\
E_j &= \sum_{l=1}^{k} \frac{n_{jl} d_l}{n_l}.
\end{aligned}
$$

For the ordered alternative hypothesis, the test for trend has got a higher statistical power than the "usual" logrank.

```
proc lifetest data=noise;
time Time*censor(0);
strata level/trend;
run;
```

The LIFETEST Procedure

Scores for Trend Test

level	Score
1	1
2	2
3	3

Trend Tests

Test	Test Statistic	Standard Error	z-Score	Pr > |z|
Log-Rank	-9.3224	2.1960	-4.2451	<.0001
Wilcoxon	-131.0000	32.2452	-4.0626	<.0001

```
> survdiff(Surv(Time,Censor)~Level)
Call: survdiff(formula = Surv(Time, Censor) ~ Level)

        N Observed Expected (O-E)^2/E (O-E)^2/V
Level=1 6        6     1.57   12.4463   17.2379
Level=2 6        5     4.53    0.0488    0.0876
Level=3 6        1     5.90    4.0660    9.4495

Chisq= 20.4  on 2 degrees of freedom, p= 3.75e-05

> OB<-survdiff(Surv(Time,Censor)~Level)$obs
> OB
[1] 6 5 1

> EX<-survdiff(Surv(Time,Censor)~Level)$exp
> EX
[1] 1.573950 4.529692 5.896359

> V<-survdiff(Surv(Time,Censor)~Level)$var
> V
[,1]       [,2]       [,3]
[1,]  1.1364441 -0.5619089 -0.5745352
[2,] -0.5619089  2.5244614 -1.9625525
[3,] -0.5745352 -1.9625525  2.5370877

> a<-c(1,2,3)
> test<-a%*%(OB-EX)
> stderror<-sqrt(t(a)%*%V%*%a)
> zscore<-test/stderror
> Pvalue<-2*pnorm(abs(zscore),lower.tail=FALSE)
> data.frame(test,stderror,zscore,Pvalue)
       test stderror    zscore       Pvalue
1 -9.322409 2.196042 -4.245095 2.185007e-05
```

With the log-rank test, we compare two or more survival curves.

However, sometimes there are confounding variables which also affect the outcome and for which we need to adjust for.

If the confounding variable has M levels, we get

$$H_0 : S_{1m}(t) = S_{2m}(t) = \ldots = S_{Km}(t),\ t \leq \tau,\ m = 1, \ldots, M.$$

Hence, we let the shape of the survival function differ for each level of the confounding variable.

To set up a test statistic, we divide the data according the M levels of the confounding variable and construct a 2×2 contingency table for each ordered death time at each level,

Popul / Died	yes	no	Total
1	d_{m0l}	$n_{m0l} - d_{m0l}$	n_{m0l}
2	d_{m1l}	$n_{m1l} - d_{m1l}$	n_{m1l}
Total	d_{ml}	$n_{ml} - d_{ml}$	n_{ml}

Let $\mathbf{O}_m$ be the sum of the observed O's by applying the log-rank calculations in level m. Similar for $\mathbf{E}_m$, the sum of the expected E's and $\mathbf{V}_m$, the sum of the v's.

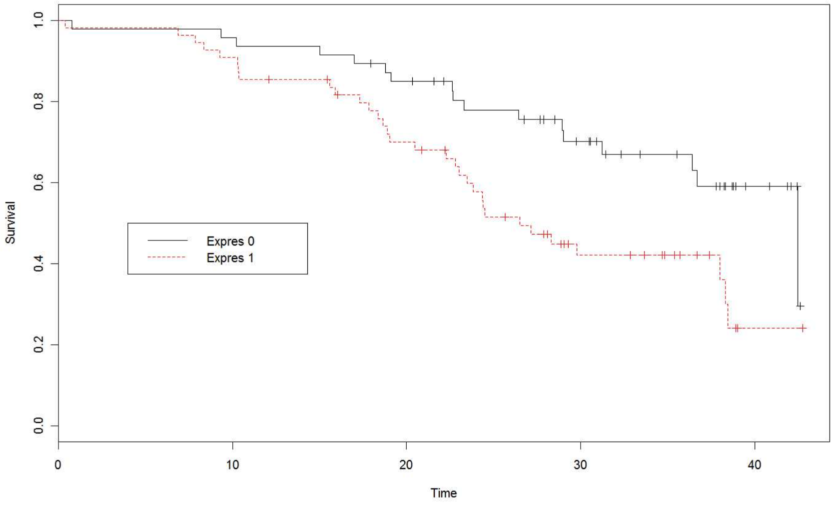
The stratified log-rank is

$$Z = \frac{\sum_{m=1}^{M} (\mathbf{O}_m - \mathbf{E}_m)}{\sqrt{\sum_{m=1}^{M} \mathbf{V}_m}}.$$

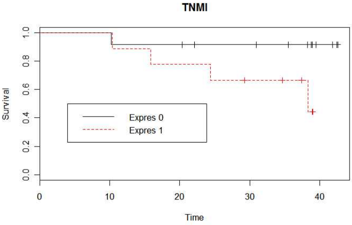
NSCLC

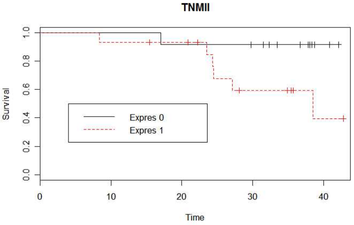
- Laudanski et al., Eur Respir J (2001).
- In this study, we had 102 patients who were operated from long cancer.
- The severity of the cancer was expressed in three TNM (Tumor, Nodes, Metastasis) categories: I, II, IIIa.
- The expression of the P53 protein was found from tumor biopsies.
- We are interested on the effect of this protein on the survival time of a patient.

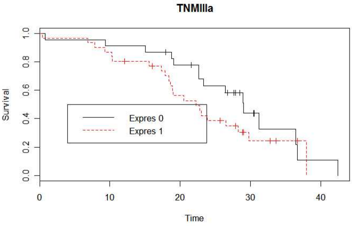
```
> NSCLC<-read.table("C:/Werk/Roel/Onderwijs/Theorie/SurvivalLeuven/NSCLC.txt",header=T,sep="\t")
>
> fit<-survfit(Surv(survtime,survind)~expres,data=NSCLC)
> summary(fit)

> plot(fit[1],xlab="Time",ylab="Survival")
> lines(fit[2],col="red",lty=2)
> legend(4,0.5,legend=c("Expres 0","Expres 1"),lty=c(1,2),col=c("black","red"))
>
> survdiff(Surv(survtime,survind)~expres,data=NSCLC)
Call: survdiff(formula = Surv(survtime, survind) ~ expres, data = NSCLC)

          N Observed Expected (O-E)^2/E (O-E)^2/V
expres=0 47       17     26.2      3.25       7.1
expres=1 55       32     22.8      3.75       7.1

Chisq= 7.1  on 1 degrees of freedom, p= 0.00769

> survdiff(Surv(survtime,survind)~expres+strata(tnm),data=NSCLC)
Call: survdiff(formula = Surv(survtime, survind) ~ expres + strata(tnm), data = NSCLC)

          N Observed Expected (O-E)^2/E (O-E)^2/V
expres=0 47       17     25.5      2.82      6.05
expres=1 55       32     23.5      3.06      6.05

Chisq= 6.1  on 1 degrees of freedom, p= 0.0139
```



Survival
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Time
Expres 0
Expres 1


TNMI
Survival
Time
Expres 0
Expres 1


TNMII
Survival
Time
Expres 0
Expres 1


TNMIIIa
Survival
Time
Expres 0
Expres 1

```
> survdiff(Surv(survtime,survind)~tnm,data=NSCLC)
Call: survdiff(formula = Surv(survtime, survind) ~ tnm, data = NSCLC)

       N Observed Expected (O-E)^2/E (O-E)^2/V
tnm=1 21        5     13.2      5.13      7.30
tnm=2 27        7     15.8      4.91      7.32
tnm=3 54       37     19.9     14.60     26.23

Chisq= 26.3  on 2 degrees of freedom, p= 1.94e-06

> OB<-survdiff(Surv(survtime,survind)~tnm,data=NSCLC)$obs
> OB
[1]  5  7 37

> EX<-survdiff(Surv(survtime,survind)~tnm,data=NSCLC)$exp
> EX
[1] 13.24541 15.81566 19.93893

> V<-survdiff(Surv(survtime,survind)~tnm,data=NSCLC)$var
> V
[,1]       [,2]       [,3]
[1,]  9.309469 -4.417150 -4.892319
[2,] -4.417150 10.623178 -6.206028
[3,] -4.892319 -6.206028 11.098346

> a<-c(1,2,3)
> test<-a%*%(OB-EX)
> stderror<-sqrt(t(a)%*%V%*%a)
> zscore<-test/stderror
> Pvalue<-2*pnorm(abs(zscore),lower.tail=FALSE)
> data.frame(test,stderror,zscore,Pvalue)
     test   stderror  zscore       Pvalue
1 25.30649 5.494766 4.605562 4.113525e-06
```

```
proc lifetest data=nsclc;
time Survtime*Survind(0);
strata expres;
run;
```

```
Testing Homogeneity of Survival Curves for Survtime over Strata


        Rank Statistics

expres      Log-Rank    Wilcoxon

0            -9.2411     -648.00
1             9.2411      648.00

...

       Test of Equality over Strata

                                   Pr >
Test      Chi-Square      DF    Chi-Square

Log-Rank      7.1051       1      0.0077
Wilcoxon      6.5750       1      0.0103
-2Log(LR)     5.2939       1      0.0214
```

```
proc lifetest data=nsclc;
time Survtime*Survind(0);
test expres;
strata tnm;
run;
```

Univariate Chi-Squares for the Log-Rank Test

Variable	Test Statistic	Standard Deviation	Chi-Square	Pr > Chi-Square
expres	-8.4782	3.4457	6.0541	0.0139

Covariance Matrix for the Log-Rank Statistics

Variable	expres
expres	11.8730

Forward Stepwise Sequence of Chi-Squares for the Log-Rank Test

Variable	DF	Chi-Square	Pr > Chi-Square	Chi-Square Increment	Pr > Increment
expres	1	6.0541	0.0139	6.0541	0.0139

```
proc lifetest data=nsclc;
time Survtime*Survind(0);
strata tnm/trend;
run;
```

Summary of the Number of Censored and Uncensored Values

Stratum	tnm	Total	Failed	Censored	Percent Censored
1	1	21	5	16	76.19
2	2	27	7	20	74.07
3	3	54	37	17	31.48
Total		102	49	53	51.96

Trend Tests

Test	Test Statistic	Standard Error	z-Score	Pr > |z|
Log-Rank	25.3065	5.4948	4.6056	<.0001
Wilcoxon	1502.0000	403.6583	3.7210	0.0002