Wertrationalität: Kosten × Zeit (t_1)

		B 1	B 2
A	1	m_{11}	m_{12}
A	2	m_{21}	m_{22}

		B ja 1	B 2 nein	
A	ja 1	74	53	127
A	nein 2	67	127	194

$$\ln(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

$$m_{ij} = e^{\lambda} \cdot e^{\lambda_i^A} \cdot e^{\lambda_j^B} \cdot e^{\lambda_{ij}^{AB}}$$

```
*** FREQUENCIES ***
                   nij        mij
  A B         observed   estimated
  1 1           74.000      74.000
  1 2           53.000      53.000
  2 1           67.000      67.000
  2 2          127.000     127.000
```

Saturiertes Modell: beobachtete = geschätzte Häufigkeiten

A \ B	1	2	Zeilenmittelwerte
1	ln74 = 4,3041	ln53 = 3,9703	4,1372
2	ln67 = 4,2047	ln127 = 4,8442	4,5244
Spaltenmittelwerte	4,2544	4,4072	$\boxed{4,3308}$

$$\lambda_1^A = -0{,}1936$$

$$\lambda_2^A = +0{,}1936$$

$$\lambda_1^B = -0{,}0764 \qquad \lambda_2^B = +0{,}0764$$

$$\boxed{\lambda = \sum \ln m_{ij} / k = 4{,}3308}$$

k Anzahl der Zellen

$$\ln(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

Loglineare Haupteffekte

+/- Abweichungen der Zeilen- bzw. Spaltenmittelwerte vom Gesamtdurchschnitt der logarithmierten Zellhäufigkeiten

Haupteffekt **gleich Null** …

… bei gleichen Zeilenmittelwerten bzw. gleichen Spaltenmittelwerten

… bzw. bei Gleichheit der Produkte im Zähler und Nenner von …

$$\boxed{\lambda_1^A = \frac{1}{4}\ln\frac{m_{11}\cdot m_{12}}{m_{21}\cdot m_{22}}}$$

$$\boxed{\lambda_1^B = \frac{1}{4}\ln\frac{m_{11}\cdot m_{21}}{m_{12}\cdot m_{22}}}$$

Nulleffekt resultiert ….

… bei **gleich**wahrscheinlichen Kategorien (*Rand*häufigkeiten) *beider* Variablen;

[bei [a] gleich oder [b] symmetrisch um die Haupt- oder Nebendiagonale verteilten *Zell*häufigkeiten]

… wenn zugehörige Variable gleichwahrscheinliche Kategorien (gleiche Rand-Häufigkeiten) aufweist und die Zellhäufigkeiten symmetrisch um die Längs- bzw. Querachse der Tabelle verteilt sind [c].

$$\lambda_1^A = \frac{1}{4}\left(\left(\ln m_{11} + \ln m_{12}\right) - \left(\ln m_{21} + \ln m_{22}\right)\right) = \frac{1}{4}\left(\ln m_{11} + \ln m_{12} - \ln m_{21} - \ln m_{22}\right)$$

35	15	50
25	25	50
60	40	100

Gleichverteilte Zellhäufigkeiten

[a]

25	25	50
25	25	50
50	50	100

[b]

10	40	50
40	10	50
50	50	100

Symmetrisch um die Haupt- oder Nebendiagonale verteilte *Zell*häufigkeiten

[c]

20	30	50
20	30	50
40	60	100

Zugehörige Variable hat gleichwahrscheinliche Kategorien (gleiche Randhäufigkeiten) und die Zellhäufigkeiten sind symmetrisch um die Querachse der Tabelle verteilt.

A \ B	1	2
1	0,2305	0,1651
2	0,2087	0,3956

P(AB) 1,0

$$\ln\left(m_{ij}\right) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

$$\lambda_{ij}^{AB} = \ln m_{ij} - \left(\lambda + \lambda_i^A + \lambda_j^B\right)$$

Im Beispiel: $\lambda_{ij}^{AB} = 0,2433$

A \ B	1	2
1	0,1738	0,2219
2	0,2655	0,3389

P(AB) 1,0

$$\ln\left(m_{ij}\right) = \lambda + \lambda_i^A + \lambda_j^B$$

```
*** FREQUENCIES ***

  A B       observed   estimated
  1 1        74.000      55.785
  1 2        53.000      71.215
  2 1        67.000      85.215
  2 2       127.000     108.785
```

$\lambda = 4,3554$ $\quad \lambda_1^A = -0,2118$ $\quad \lambda_1^B = -0,1221$

		B: 1	B: 2	
A	1	0,2305	0,1651	0,3956
A	2	0,2087	0,3956	0,6044
				1,0

P(AB)

A	1	0,3956
A	2	0,6044
		1,0

P(A)

	B: 1	B: 2	
A: 1	0,5827	0,4173	1,0
A: 2	0,3454	0,6546	1,0

P(B|A)

$$P(AB) = P(A) \times P(B|A)$$

$$\frac{P(AB)}{P(A)} = P(B|A)$$

$$\lambda_1^A = -0{,}2118 \qquad \lambda_1^B = -0{,}0764 \qquad \lambda_{ij}^{AB} = 0{,}2433$$

$$\lambda_1^A = \ln m_1 - ((\ln m_1 + \ln m_2)/2) = \ln 127 - ((\ln 127 + \ln 194)/2)$$

LEM

```
*** INPUT ***

  * Wertrationalitaet Kosten x
Zeit t1 dichotom
  res 0
  lat 0
  man 2
  dim 2 2
  lab A B
  sub AB
  mod A B|A {AB}
  fre [74 53 67 127]
  cas 321
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.3956  (0.0273)
  2               0.6044  (0.0273)

* P(B|A) *

  1 | 1           0.5827  (0.0438)
  2 | 1           0.4173  (0.0438)
  1 | 2           0.3454  (0.0341)
  2 | 2           0.6546  (0.0341)
```

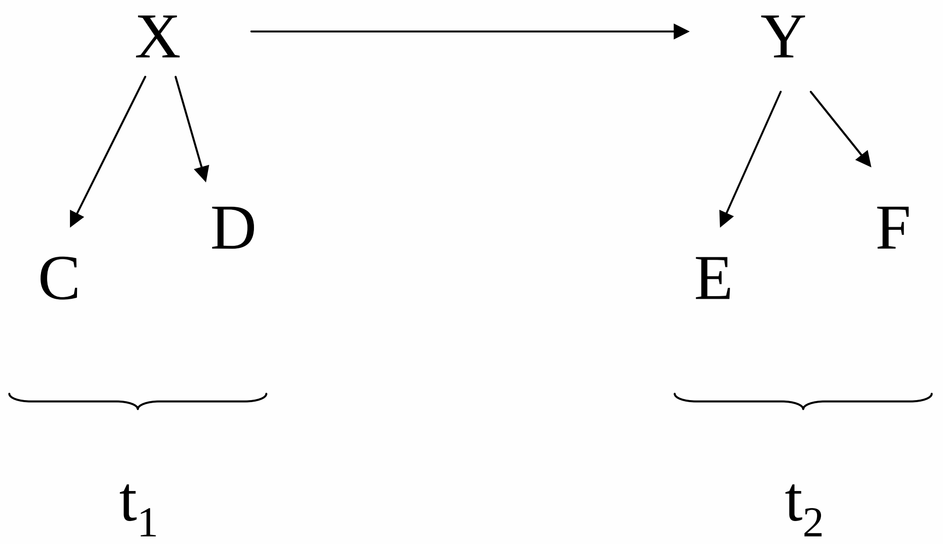
```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect            beta  std err  z-value   exp(beta)     Wald  df  prob
  A
   1             -0.2118   0.0571   -3.712      0.8091
   2              0.2118                         1.2359    13.78   1 0.000

* TABLE AB [or P(B|A)] *

  effect            beta  std err  z-value   exp(beta)     Wald  df  prob
  B
   1             -0.0764   0.0587   -1.301      0.9264
   2              0.0764                         1.0794     1.69   1 0.193
  AB
   1 1            0.2433   0.0587    4.143      1.2755
   1 2           -0.2433                         0.7840
   2 1           -0.2433                         0.7840
   2 2            0.2433                         1.2755    17.16   1 0.000
```



X, Y	Latent categorcial variables
C, D, E, F	Manifest kategoriale indicator variables

```
*** INPUT ***

  res 2
  lat 2
  man 4
  dim 2 2 2 2 2 2 2 2
  lab R S X Y C D E F
  sub CDEF CD EF -
  mod X Y|X {XY}
      C|X D|X
      E|Y F|Y
      R S|R
  dat wr.dat
  cas 336
```

*** LATENT CLASS OUTPUT ***

			X 1	X 1	X 2	X 2	
		Nein	X 1	X 1	X 2	X 2	Ja
		Nein	Y 1	Y 2	Y 1	Y 2	Ja
			0.4909	0.1491	0.0056	0.3544	
Ja	C	1	0.0851	0.0851	0.9563	0.9563	
Nein	C	2	0.9149	0.9149	0.0437	0.0437	
Ja	D	1	0.1988	0.1988	0.8669	0.8669	
Nein	D	2	0.8012	0.8012	0.1331	0.1331	
Ja	E	1	0.1520	0.8904	0.1520	0.8904	
Nein	E	2	0.8480	0.1096	0.8480	0.1096	
Ja	F	1	0.2924	0.9248	0.2924	0.9248	
Nein	F	2	0.7076	0.0752	0.7076	0.0752	

		Y: 1	Y: 2	
X	1	0,4909	0,1491	0,64
	2	0,0056	0,3544	0,36

P(XY) *P(X)*

		Y: 1	Y: 2	
X	1	0,767	0,233	1,0
	2	0,016	0,984	1,0

P(Y |X)

X = 1 [2], wenn C, D =2 [1]

Y = 1 [2], wenn E, F = 2 [1]

$$P(XY) = P(X) \cdot P(Y|X)$$

$$P(Y|X) = \frac{P(XY)}{P(X)}$$

```
*** STATISTICS ***

  X-squared             = 13.2637 (0.2093)

  Degrees of freedom    = 10

  Number of parameters = 14 (+1)
```

```
*** PSEUDO R-SQUARED MEASURES ***

* P(Y|X) *
                        baseline   fitted   R-squared
  entropy                0.6931    0.3763    0.4570

* P(C|X) *
                        baseline   fitted   R-squared
  entropy                0.6723    0.2508    0.6269

* P(D|X) *
                        baseline   fitted   R-squared
  entropy                0.6857    0.4604    0.3285
```

Aus dem pdf-File: „Datenmatrix, Mittlere Werte und Streuungsmaße“ (Statistik I)

Entropie $H = -\sum_{i=1}^{I} \frac{n_i}{n} \ln \frac{n_i}{n}$

Zahlen-beispiele ..

	p_i	p_i	p_i	p_i
1	0,5	0,4	0,3	0,2
2	0,5	0,6	0,7	0,8
	0,69	**0,67**	**0,61**	**0,50**

.. unter Ver-wendung des natürlichen Logarithmus *ln*

z.B.: $-1\times[(0{,}5 \times \log_e(0{,}5))+(0{,}5 \times \log_e(0{,}5))]=0{,}69$
z.B.: $-1\times[(0{,}4 \times \log_e(0{,}4))+(0{,}6 \times \log_e(0{,}6))]=0{,}67$

	Gruppe 1	Gruppe 2	p_i
1	0,75	0,17	0,4
2	0,25	0,83	0,6
Entropie	**0,56**	**0,46**	**0,67**
Entropie (gewichtet)	0,56*0,40 = 0,22	0,46*0,60 = 0,28	
n	**40**	**60**	**100**

$$Pseudo-R^2 = \frac{S_y^2 - S_e^2}{S_y^2} = \frac{0{,}67 - (0{,}22 + 0{,}28)}{0{,}67} = \frac{0{,}17}{0{,}67} = \underline{\underline{0{,}25}}$$

```
*** INPUT ***

    res 2
    lat 2
    man 4
    dim 2 2 2 2 2 2 2 2
    lab R S X Y C D E F
    sub CDEF CD EF -
    mod X Y|X {XY}
        C|X D|X
        E|Y F|Y
        R S|R
    dat wr.dat
    cas 336
```

C	D	E	F	freq
1	1	1	1	49
1	1	1	2	5
1	1	2	1	7
1	1	2	2	2
1	2	1	1	9
1	2	1	2	1
1	2	2	1	3
1	2	2	2	4
2	1	1	1	7
2	1	1	2	1
2	1	2	1	10
2	1	2	2	8
2	2	1	1	22
2	2	1	2	12
2	2	2	1	18
2	2	2	2	45

CD [R=1 S=2]

C	D	E	F	freq
1	1	?	?	13
1	2	?	?	5
2	1	?	?	6
2	2	?	?	19

EF [R=2 S=1]

C	D	E	F	freq
?	?	1	1	8
?	?	1	2	0
?	?	2	1	0
?	?	2	2	7

- [R=2 S=2]

C	D	E	F	freq
?	?	?	?	75

CDEF [R=1, S=1]

		Y_l	
		$l=1$	$l=2$
X_q	$q=1$	m_{q1}	m_{q2}
	$q=2$	m_{q1}	m_{q2}

Zur Äquivalenz von loglinearen und Logit Modellen:

$$\ln\left(\frac{m_{q1}}{m_{q2}}\right) = \beta_0 + \beta_1 x_q = 2\lambda_1^Y + 2\lambda_{q1}^{XY} \qquad \text{wenn:} \quad \underline{\lambda_1^Y = -\lambda_2^Y} \quad \text{und} \quad \underline{\lambda_{q1}^{XY} = -\lambda_{q2}^{XY}}$$

$$\ln\left(\frac{m_{q1}}{m_{q2}}\right) = \ln(m_{q1}) - \ln(m_{q2})$$

$$\ln(m_{q1}) = \lambda + \lambda_q^X + \lambda_1^Y + \lambda_{q1}^{XY} \qquad \ln(m_{q2}) = \lambda + \lambda_q^X + \lambda_2^Y + \lambda_{q2}^{XY}$$

$$\ln(m_{q1}) - \ln(m_{q2}) = \lambda - \lambda + \lambda_q^X - \lambda_q^X + \lambda_1^Y - \underline{\lambda_2^Y} + \lambda_{q1}^{XY} - \underline{\lambda_{q2}^{XY}}$$
$$= \lambda_1^Y - \underline{\left(-\lambda_1^Y\right)} + \lambda_{q1}^{XY} - \underline{\left(-\lambda_{q1}^{XY}\right)} = 2\lambda_1^Y + 2\lambda_{q1}^{XY}$$