

# Mesure de la force de la liaison entre deux variables binaires : risque relatif et odds-ratio

- Une maladie et un facteur de risque :

	Maladie	
	oui	non
FR = oui	a	b
FR = non	c	d

$$RR = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}}$$

- Une maladie et un facteur de risque :

	Maladie	
	oui	non
FR = oui	a	b
FR = non	c	d

$$OR = \frac{\frac{a}{b}}{\frac{c}{d}}$$

- Odds-ratio *versus* risque relatif : avantages et inconvénients

- Inconvénient : le sens

- Inconvénient : le sens
- Avantages :
  - La régression logistique, utile pour démêler des relations complexes entre variables binaires, estime des OR et pas des RR

- Inconvénient : le sens
- Avantages :
  - La régression logistique, utile pour démêler des relations complexes entre variables binaires, estime des OR et pas des RR
  - Dans une enquête cas-témoin (par exemple une étude où l'on compare 300 patients avec un infarctus à 300 sujets sains), le RR n'a aucun sens alors que l'OR si

- Inconvénient : le sens
- Avantages :
  - La régression logistique, utile pour démêler des relations complexes entre variables binaires, estime des OR et pas des RR
  - Dans une enquête cas-témoin (par exemple une étude où l'on compare 300 patients avec un infarctus à 300 sujets sains), le RR n'a aucun sens alors que l'OR si
  - Si la maladie est peu fréquente :  $OR = RR$



- Une enquête en population générale :

	enrhumés	
	oui	non
Tabac: oui	30	300
Tabac: non	30	600

$$RR = \frac{\frac{30}{(300+30)}}{\frac{30}{(600+30)}} = 1,91$$

$$OR = \frac{\frac{30}{300}}{\frac{30}{600}} = 2$$

- Une enquête cas-témoins :

	enrhumés	
	oui	non
Tabac: oui	30	30
Tabac: non	30	60

$$RR = \frac{\frac{30}{(30+30)}}{\frac{30}{(60+30)}} = 1,5$$

$$OR = \frac{\frac{30}{30}}{\frac{30}{60}} = 2$$

```
> smp.c$ed.b <- ifelse(smp.c$ed>2,1,0)
> str(smp.c)
'data.frame': 799 obs. of 10 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7 6 8 6 3 2 6 6 ...
 $ dep.cons : int   0 0 0 0 1 0 1 0 1 0 ...
 $ scz.cons : int   0 0 0 0 0 0 0 0 0 0 ...
 $ grav.cons: int   1 2 2 1 2 1 5 1 5 5 ...
 $ n.enfant : int   2 7 2 0 1 3 5 2 1 2 ...
 $ rs       : int   2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int   1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int   1 1 2 2 2 1 2 2 1 2 ...
 $ ed.b     : num   0 0 1 0 0 0 1 0 1 0 ...
> table(smp.c$ed.b,smp.c$ed,deparse.level=2,useNA="always")
      smp.c$ed
smp.c$ed.b  1    2    3 <NA>
0          315 155    0    0
1           0   0 222    0
<NA>       0   0   0 107
```

```
> smp.c$ed.b <- ifelse(smp.c$ed>2,1,0)
> str(smp.c)
'data.frame': 799 obs. of 10 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7 6 8 6 3 2 6 6 ...
 $ dep.cons : int   0 0 0 0 1 0 1 0 1 0 ...
 $ scz.cons : int   0 0 0 0 0 0 0 0 0 0 ...
 $ grav.cons: int   1 2 2 1 2 1 5 1 5 5 ...
 $ n.enfant : int   2 7 2 0 1 3 5 2 1 2 ...
 $ rs       : int   2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int   1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int   1 1 2 2 2 1 2 2 1 2 ...
 $ ed.b     : num   0 0 1 0 0 0 1 0 1 0 ...
```

```
> table(smp.c$ed.b, smp.c$ed, deparse.level=2, useNA="always")
```

	smp.c\$ed			
smp.c\$ed.b	1	2	3	<NA>
0	315	155	0	0
1	0	0	222	0
<NA>	0	0	0	107

```
> smp.c$ed.b <- ifelse(smp.c$ed>2,1,0)
> str(smp.c)
'data.frame': 799 obs. of 10 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7 6 8 6 3 2 6 6 ...
 $ dep.cons : int   0 0 0 0 1 0 1 0 1 0 ...
 $ scz.cons : int   0 0 0 0 0 0 0 0 0 0 ...
 $ grav.cons: int   1 2 2 1 2 1 5 1 5 5 ...
 $ n.enfant : int   2 7 2 0 1 3 5 2 1 2 ...
 $ rs       : int   2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int   1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int   1 1 2 2 2 1 2 2 1 2 ...
 $ ed.b     : num   0 0 1 0 0 0 1 0 1 0 ...
```

```
> table(smp.c$ed.b, smp.c$ed, deparse.level=2, useNA="always")
```

	smp.c\$ed			
smp.c\$ed.b	1	2	3	<NA>
0	315	155	0	0
1	0	0	222	0
<NA>	0	0	0	107

```
> smp.c$ed.b <- ifelse(smp.c$ed>2,1,0)
> str(smp.c)
'data.frame': 799 obs. of 10 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7 6 8 6 3 2 6 6 ...
 $ dep.cons : int   0 0 0 0 1 0 1 0 1 0 ...
 $ scz.cons : int   0 0 0 0 0 0 0 0 0 0 ...
 $ grav.cons: int   1 2 2 1 2 1 5 1 5 5 ...
 $ n.enfant : int   2 7 2 0 1 3 5 2 1 2 ...
 $ rs       : int   2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int   1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int   1 1 2 2 2 1 2 2 1 2 ...
 $ ed.b     : num   0 0 1 0 0 0 1 0 1 0 ...
```

```
> table(smp.c$ed.b, smp.c$ed, deparse.level=2, useNA="always")
```

	smp.c\$ed			
smp.c\$ed.b	1	2	3	<NA>
0	315	155	0	0
1	0	0	222	0
<NA>	0	0	0	107

```
> library(Epi)
> twoby2(1-smp.c$ed.b,1-smp.c$dep.cons)
2 by 2 table analysis:
-----
Outcome      : 0
Comparing    : 0 vs. 1

      0      1      P(0) 95% conf. interval
0 126  96  0.5676    0.5016    0.6312
1 135 335  0.2872    0.2481    0.3298

                                95% conf. interval
      Relative Risk: 1.9760    1.6456    2.3726
      Sample Odds Ratio: 3.2569    2.3361    4.5408
Conditional MLE Odds Ratio: 3.2508    2.3037    4.6035
      Probability difference: 0.2803    0.2020    0.3549

      Exact P-value: 0
      Asymptotic P-value: 0
-----
```

```
> library(Epi)
> twoby2(1-smp.c$ed.b,1-smp.c$dep.cons)
2 by 2 table analysis:
-----
Outcome      : 0
Comparing    : 0 vs. 1

      0      1      P(0) 95% conf. interval
0 126  96  0.5676    0.5016    0.6312
1 135 335  0.2872    0.2481    0.3298

                                95% conf. interval
      Relative Risk: 1.9760    1.6456    2.3726
      Sample Odds Ratio: 3.2569    2.3361    4.5408
Conditional MLE Odds Ratio: 3.2508    2.3037    4.6035
      Probability difference: 0.2803    0.2020    0.3549

      Exact P-value: 0
      Asymptotic P-value: 0
-----
```



```
> library(Epi)
> twoby2(1-smp.c$ed.b,1-smp.c$dep.cons)
2 by 2 table analysis:
-----
Outcome      : 0
Comparing    : 0 vs. 1

      0      1      P(0) 95% conf. interval
0 126  96  0.5676   0.5016   0.6312
1 135 335  0.2872   0.2481   0.3298

      Relative Risk: 1.9760  1.6456  2.3726
      Sample Odds Ratio: 3.2569  2.3361  4.5408
Conditional MLE Odds Ratio: 3.2508  2.3037  4.6035
      Probability difference: 0.2803  0.2020  0.3549

      Exact P-value: 0
      Asymptotic P-value: 0
-----
```

# Conclusion

Introduction à la statistique avec R > Risque relatif et odds-ratio



```
smp.c$ed.b <- ifelse(smp.c$ed>2,1,0)
str(smp.c)
table(smp.c$ed.b, smp.c$ed, deparse.level=2, useNA="always")
library(Epi)
twoby2(1-smp.c$ed.b, 1-smp.c$dep.cons)
```

