













Calibration in the large: property of the full sample

Predicted prob. of survival (1 = survived, 0 = not) for 10 Titanic passengers of different sex

	Probability	Actual		Probability	Actual
	0.12	0		0.35	0
	0.72	1		0.26	0
	0.43	1		0.29	0
	0.27	0		0.66	1
	0.52	1		0.38	0

Mean probability of survival of model:

0.4

Observed frequency of survival in sample:

0.4













Calibration in the large:

0.4 = 0.4



Calibration in the small: property of subsets

Predicted prob. of survival (1 = survived, 0 = not) for 10 Titanic passengers of different sex

	Probability	Actual		Probability	Actual
	0.12	0		0.35	0
	0.72	1		0.26	0
	0.43	1		0.29	0
	0.27	0		0.66	1
	0.52	1		0.38	0

Mean probability of survival of model:

male: 0.25

female: 0.50

Observed frequency of survival in sample:

male: 0.0

female: 0.67

Calibration in the small:







male: $0.25 \neq 0.0$



female: $0.5 \neq 0.67$



Predicted prob. of survival (1 = survived, 0 = not) for 6 Titanic passengers of different gender

	Probability	Actual
	0.2	0
	0.8	0
	0.9	1
	0.1	0
	0.7	1
	0.3	1







Calibration in the large:
property of full sample

Mean predicted
probability of survival:
0.5

Observed probability
of survival in sample:
0.5

Calibration in the large:
 $0.5 = 0.5$ ✓

Predicted prob. of survival (1 = survived, 0 = not) for 6 Titanic passengers of different gender

	Probability	Actual
	0.2	0
	0.8	0
	0.9	1
	0.1	0
	0.7	1
	0.3	1

Calibration in the large:
property of full sample

Mean predicted
probability of survival:

0.5

Observed probability
of survival in sample:

0.5

Calibration in the large:

$0.5 = 0.5$ ✓

Calibration in the small:
property of subsets

Mean predicted
probability of survival:

male: 0.20

female: 0.80

Observed probability
of survival in sample:

male: 0.33

female: 0.67

Calibration in the small:

male: $0.20 \neq 0.33$ ✗

female: $0.80 \neq 0.67$ ✗