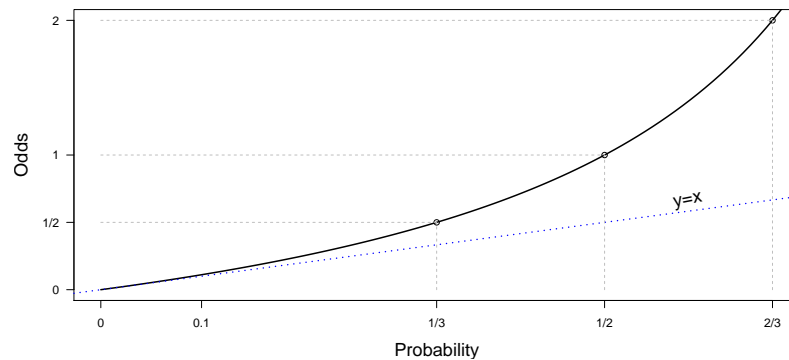


"Non-binary" categorical data has more than 2 values (A, B, C, \dots). We talk about this data in terms of the entire *distribution*, i.e. the probability function ($\mathbf{Prob}(A), \mathbf{Prob}(B), \mathbf{Prob}(C), \dots$). Of course we could omit one of these, since it would be implicit, but it is inconvenient.

MORE
OO



A few remarks on odds

- if an event has odds θ , then its probability p is $p = \theta/(1 + \theta)$
E.g.
 - $\theta = 2$ corresponds to $p = 2/3$.
 - $\theta = 1$ corresponds to $p = 1/2$.
 - $\theta = 1/100$ corresponds to $p = 1/101$.
- there are multiple systems of betting (sports) 'odds', that are not odds in the sense of this course!
- in a betting game where you stand to win 1 unit of money, your stake S (if this is kept when winning) should not exceed the odds
 - 'expected' profit $= 1p - S(1 - p) \geq 0$ is equivalent to $S \leq \frac{p}{1-p} = \theta$
 - E.g. if you are offered x units of money for a game you think has odds 2 (in your favor) then do not bet more than $2x$. (Betting $2x$ makes the game *fair*.)

Odds ratio (OR)

Probabilities cannot be retrieved from the OR alone.

If you know the 'denominator' probability (p_2) then the 'numerator' probability (p_1) can be calculated

$$p_1 = \frac{OR \cdot p_2}{1 + (OR - 1) \cdot p_2}.$$

For small values of p_2 and 'moderate' values of OR

$$p_1 \approx OR \cdot p_2,$$

i.e.

$$OR \approx \frac{p_1}{p_2} = RR.$$

Relational measures

	Bleeding		Sum
	Yes	No	
dabigatran	27	320	347
placebo	8	363	371
Sum	35	683	718

Measures for risk of dabigatran versus placebo

- (Risk ratio (RR))** $= \frac{27/347}{8/371} \approx 3.6$
- Odds ratio (OR)** $= \frac{27/320}{8/363} \approx 3.8$
- Risk difference** $= 27/347 - 8/371 \approx 0.056$

Fishers exact test

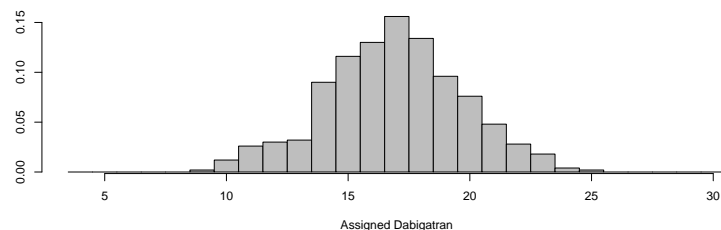
	Bleeding		Sum
	Yes	No	
dabigatran	X	(347-X)	347
placebo	(35-X)	(683-347+X)	371
Sum	35	(683)	718

Suppose that whether a person bleeds or not is complete independent of intervention. (H_0 : "odds ratio = 1".)

Then the 35 individuals who bled should be a random sample of the study population (of size 718) and we would expect that $X/35 = 347/718 \approx 48\%$.

We would expect X to be 17-18, but is $X = 27$ within some acceptable range of possibilities?

Simulate (500 times) the experiment of randomly selecting 35 people from the study population and record the number who got dabigatran (15, 15, 16, 18, 16...)

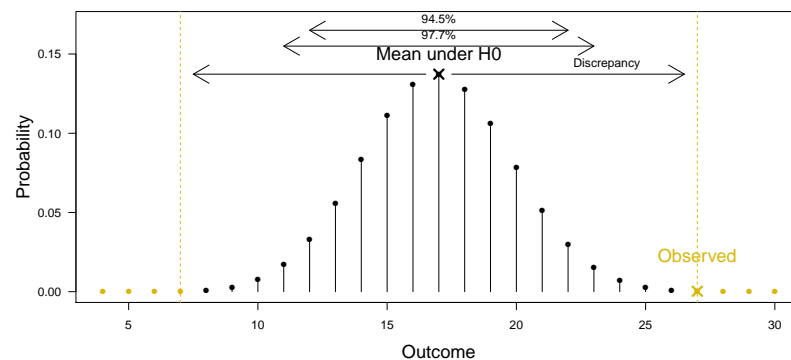


However, we can calculate *exactly* what the distribution of X is *given* H_0 (in this case).

The p -value is the probability of a discrepancy the size of that between the observed and the expected.

p -value in Fishers exact test

Sum the yellow values to get $p = 0.00045$.



More on Fishers exact test

My software produced the following output:

Fisher's Exact Test for Count Data

```
data: Dabigatran_example
p-value = 0.0004458
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.659358 9.877595
sample estimates:
odds ratio
 3.821942
```

So odds ratio is between 1.7 and 9.9. (Allows for test of model.)
Probabilities are small, so risk of dabigatran is (approx.) between 1.7 and 9.9 times larger than placebo risk.

Absolute risk

What can we say about the *absolut* risk of bleeding with dabigatran?

This was covered by Lars in Lecture 3! (Genotype example.)

The risk estimate $27/347 = 0.078$ has a standard error (SE) given by

$$\sqrt{\frac{0.078(1 - 0.078)}{347}} = 0.0144.$$

This yields a 95% confidence interval given by

$$(0.078 \pm 1.96 \cdot 0.0144) = (0.050, 0.11).$$

(This allows for test of model.)

What is the *difference* in risk between dabigatran and placebo?
 This has (almost) been covered by Lars. One needs to know that for two **independent** estimators (having SE_1 and SE_2) the SE for their difference is given by

$$\sqrt{SE_1^2 + SE_2^2}.$$

Risk	Estimate	Standard error
dabigatran	$p_1 = 27/347 = 0.078$	$\sqrt{p_1(1 - p_1)/347} = 0.0144$
placebo	$p_2 = 8/371 = 0.022$	$\sqrt{p_2(1 - p_2)/371} = 0.0075$
difference	$p_1 - p_2 = 0.056$	$\sqrt{0.0144^2 + 0.0075^2} = 0.0162$

We get a 95% confidence interval for the difference with

$$(0.056 \pm 1.96 \cdot 0.0162) = (0.024, 0.088).$$

(This allows for test of model.)

It certainly seems so (but it will actually return again later in the lecture!)

Summary of the dabigatran example:

Quantity	Estimate	Confidence interval
p_1	0.078	(0.050, 0.11)
$p_1 - p_2$	0.056	(0.024, 0.088)
OR (p_1 vs. p_2)	3.82	(1.7, 9.9)

To prove a new product is hypoallergenic it should provoke no more skin reactions than current market leader.
 To test a new product 40 individuals got both products applied to to patches of skin and observed for reaction (yes/no)

id	new	market
1	no	no
2	yes	no
3	no	no
\vdots	\vdots	\vdots
40	yes	yes

The following table is *not* appropriate to answer that question.

	no	yes
old	22	18
new	32	8

The rows are dependent.

Note:

One *can* estimate (and get confidence intervals) for p_1 and p_2 (risk of skin reaction with new and old, respectively).
 But it is harder to quantify the SE for risk difference and OR, due to the dependence.

		new		Σ
		no	yes	
market	no	17	5	22
	yes	15	3	18
Σ		32	8	40

McNemars test only considers the pairs where the results are different.

With new product there are 8 reactions but 3 of them would have happened anyway. The new product 'creates' 5 reactions, whereas the market leader 'creates' 15 reactions.

Switching from the market leader to the new product would benefit 15, make it worse for 5, and have no effect on 20. (In terms of skin reactions.)

A test statistic can be calculated. p -value for H_0 : 'no difference' is approx 4%.

Twins being randomized to intervention or placebo:

		Placebo	
		improvement	non
Intervention	improvement	a	b
	non	c	d

Before/after data:

		Before	
		event	non
After	event	a	b
	non	c	d

One of Mendel's pea-experiments was a (dihybrid) cross between the genes for round/wrinkled seeds and yellow/green seeds.

Type	RY	RG	WY	WG	Sum
Count (O)	315	108	101	32	558

According to his theory, these should appear in ratios of 9:3:3:1. So, we have a model for X = "the type":

Value v	RY	RG	WY	WG
$\text{Prob}(X = v)$	9/16	3/16	3/16	1/16

χ^2 tests are applied to tabulated data (i.e. the 'counts'), typically categorical data.

Like the t -test, we can use χ^2 to compare a sample against a model or, compare 2 or more samples against each other.

We can apply χ^2 -tests to binary data, e.g. the Dabigatran data.

χ^2 tests typically calculate a test statistic Q according to the formula

$$Q = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}.$$

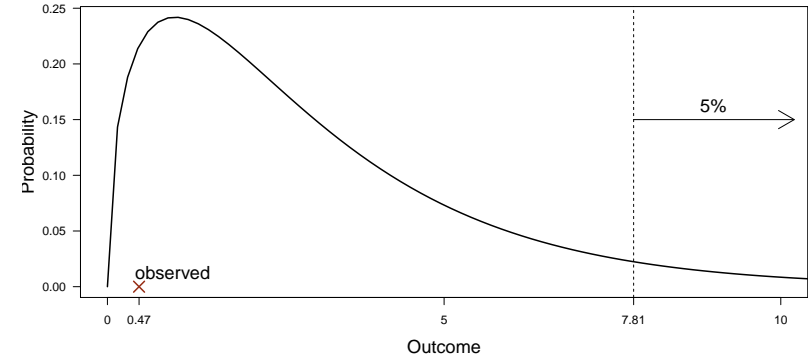
Q is compared to a χ^2 distribution with a parameter (degrees of freedom) that depends on the situation. (This is not an exact test.)

χ^2 -analysis:

Type	RY	RG	WY	WG	Sum
Data (O)	315	108	101	32	558
H_0 model (p)	9/16	3/16	3/16	1/16	1
Expected ($E = 558 \times p$)	313.9	104.6	104.6	34.9	558
Q , i.e. $(O - E)^2/E$	0.004	0.111	0.124	0.241	0.479
Residuals $(O - E)/\sqrt{E}$	0.127	0.367	-0.318	-0.467	

If H_0 is correct then Q should be (approximately) $\chi^2(3)$.
(3 = the number of categories − 1.)

The observed test statistic 0.47 is compatible with H_0 .

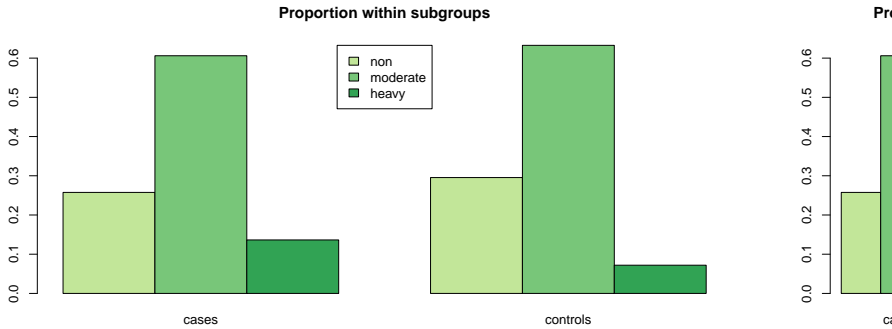


A case control study of coronary heart disease and drinking (none, moderate, heavy). Cases were matched on age, gender and smoking habits.

	non	moderate	heavy	sum
cases	34	80	18	132
controls	156	334	38	528
sum	190	414	56	660

Does drinking habits differ between cases and controls?
If they do not (H_0), their distributions should be close to

	non	moderate	heavy
	28.8% (190/660)	62.7% (414/660)	8.48% (56/660)



Comparison of open surgery (OS) and percutaneous nephrolithotomy (PN) for removal of kidney stones.
(Data illustrates Simpson's paradox.)

	Total		Adjusted for size			
	OS	PN	Small stones		Large stones	
	OS	PN	OS	PN	OS	PN
Success	273	289	81	234	192	55
Failure	77	61	6	36	71	25
Odds (for success)	3.5	4.7	13.5	6.5	2.7	2.2
Odds ratio (OS / PN)	0.75		2.1		1.2	

Here it seems like we should adjust for stone size.
(The **Mantel-Haenszel** test is a way to analyse several contingency tables.)

- Chapters 23-25: Petrie & Sabin. *Medical Statistics at a Glance*, Wiley-Blackwell (2009).
- Grant, R. L.: Converting an odds ratio to a range of plausible relative risks for better communication of research findings, *BMJ* **348** (2014) 7 pages.

In observational studies we typically gather more information. E.g.

Ind.	Bleeding	DE Dose	Age	Gender	Weight	...
1	Yes	50	75	M	83	...
2	No	75	64	F	77	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

When medicine is not randomized a simple cross tabulation analysis of 'Bleeding' versus 'DE Dose' is likely to be confounded.

One way to deal with this is logistic regression (more on that in Lecture 10).