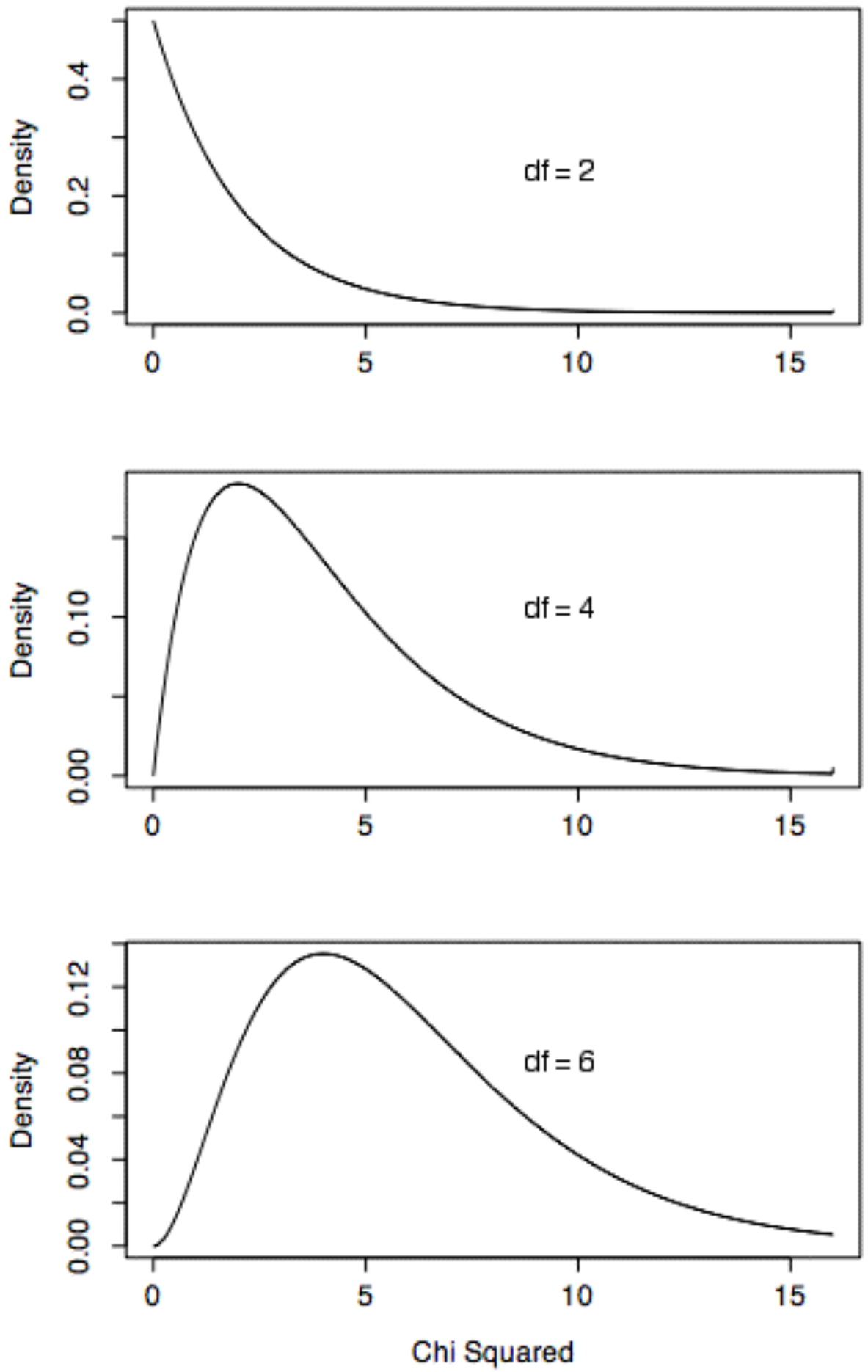
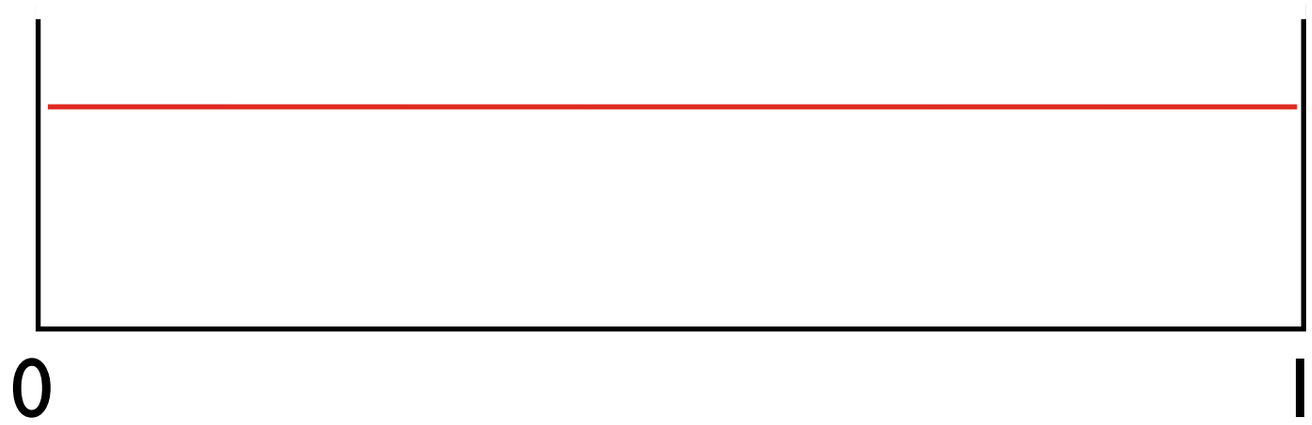
Distributions have tails.


Density
0.0
0.2
0.4
df = 2
0
5
10
15
Density
0.00
0.10
df = 4
0
5
10
15
Density
0.00
0.04
0.08
0.12
df = 6
0
5
10
15
Chi Squared

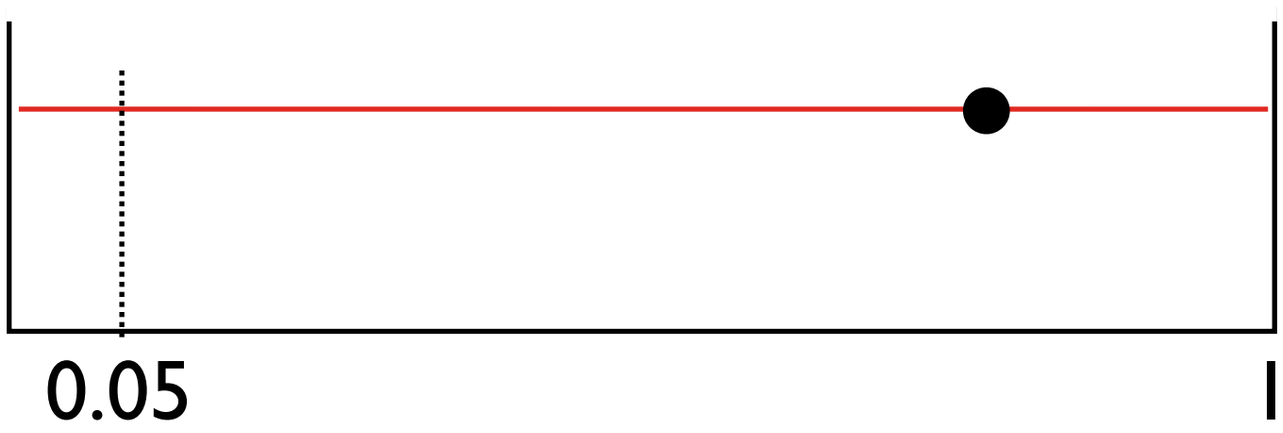
P-values



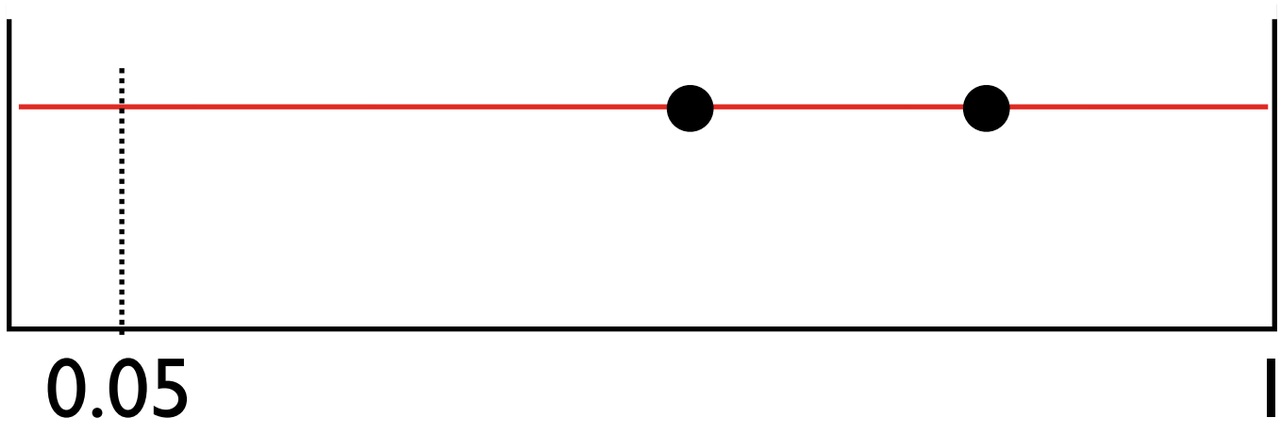
P-values are uniformly distributed under the null hypothesis

α = the probability of having a *single* false positive
(type I error)

If you are conducting a single test, then α is your p-value



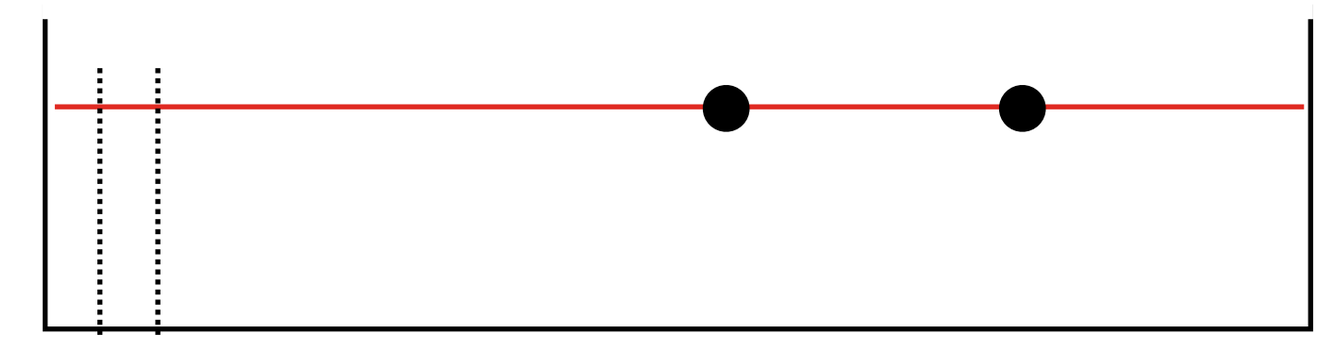
If you are conducting multiples tests, then α is *not* your p-value



Bonferroni correction

To maintain a probability, α, of a single false positive, then p-value cut-off must become:

$= \alpha/m$



0.025

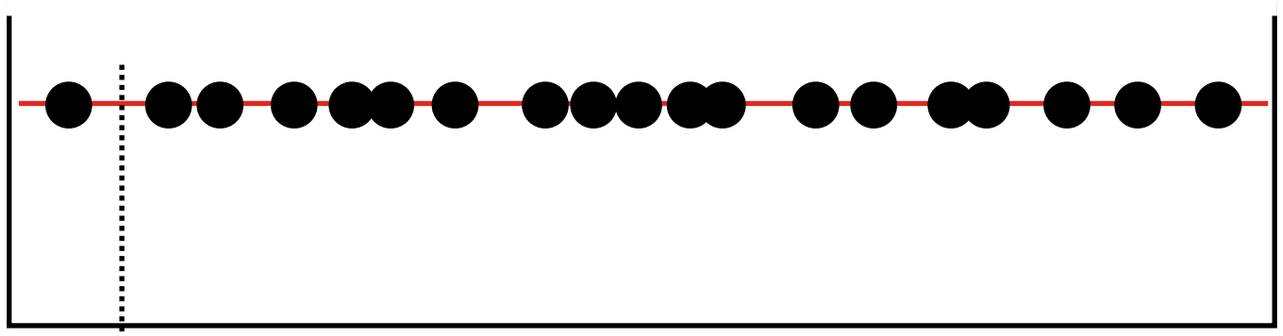
Dunn-Sidak correction

$$= 1-(1-\alpha)^{1/m}$$

False Discovery Rate (FDR)

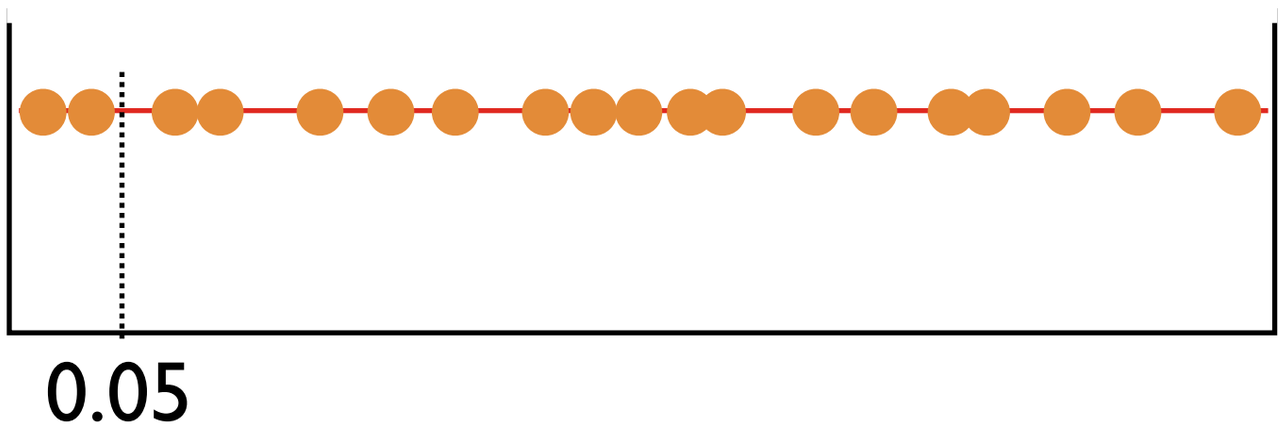
Instead of controlling probability of a single false positive, simply control fraction of false positives *among your significant tests*

P-values



If there are no true positives, p-values are uniformly distributed

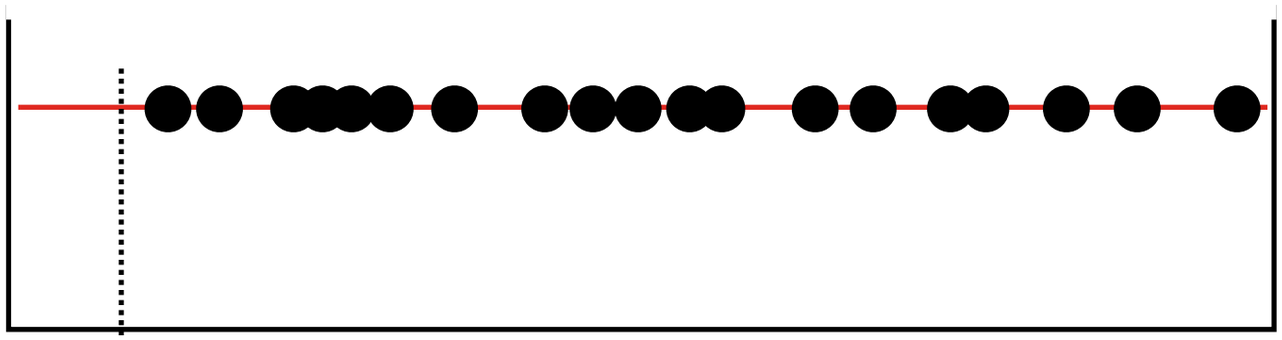
False discovery rate



FDR= the number significant expected/
the number significant observed

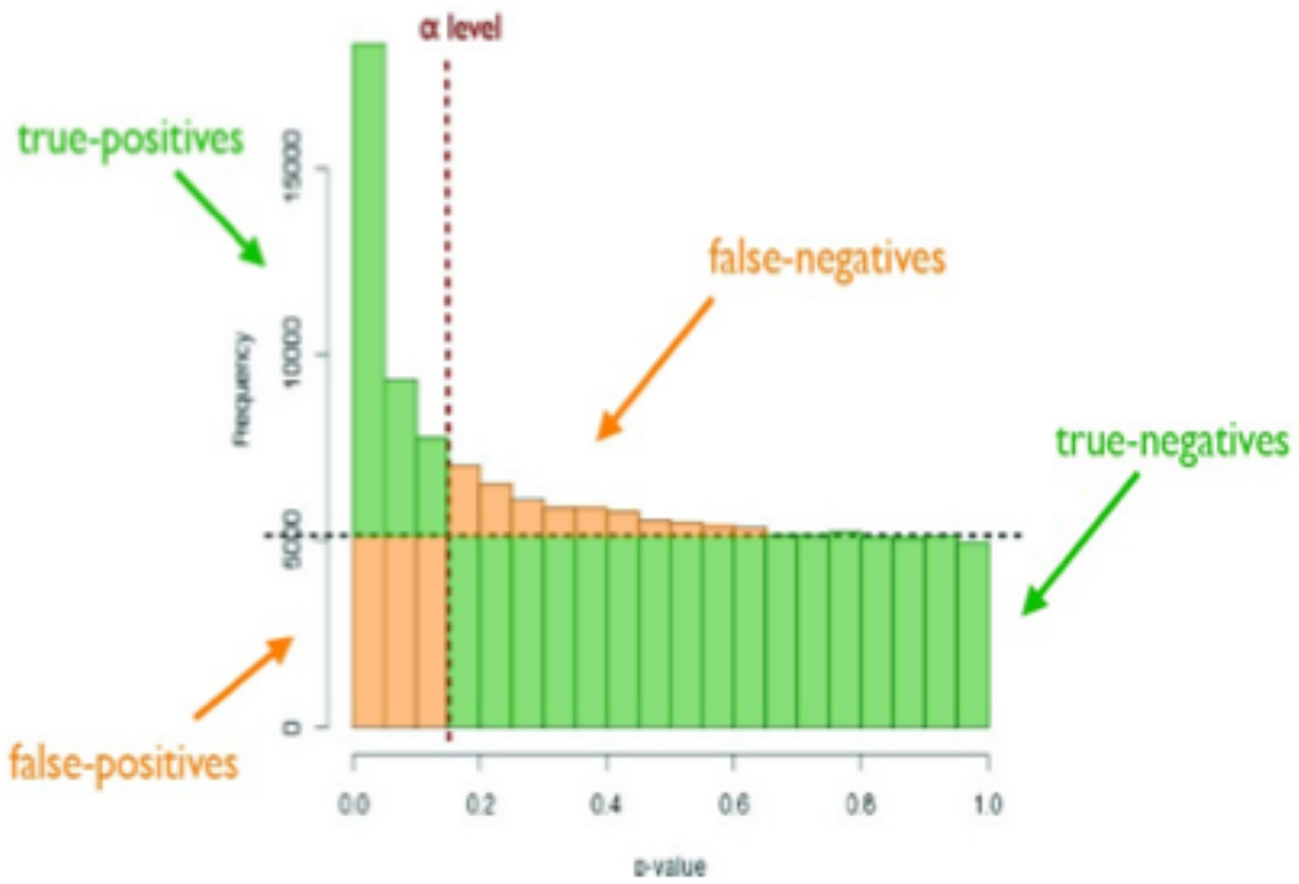
(the expected number is itself binomially distributed)

False discovery rate



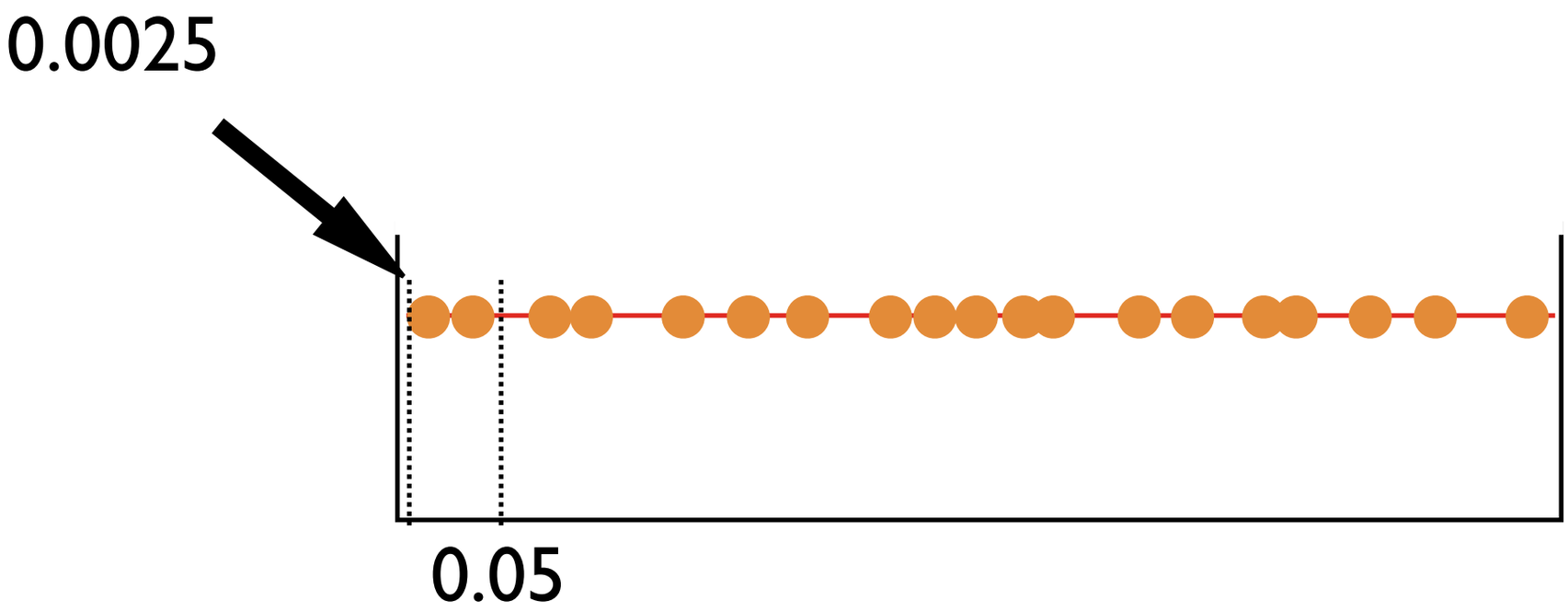
Even random data should have 5% of tests significant at 0.05

Type I and Type II error



Real data has true positives and true negatives,
as well as false positives (type I) and false negatives (type II)

False discovery rate



FDR helps to avoid false negatives
(type II error)

False Discovery Rate

$$FDR = \frac{p(i)*m}{i}$$

where p(*i*) is the p-value of the *i*th test, and
i is the rank of this test in the whole list

m=100

Bonf

B-H

P-value	rank	FDR
0.0001	1	0.01
0.0002	2	0.01
0.001	3	0.033
0.001	4	0.025
0.01	5	0.2
	6	
	7	
	8	
	9	
0.05	10	0.5

There are many ways of testing for significance, and many different cut-offs used.

Your choice really depends on what you want to do next.