HIGH DIMENSIONAL STATISTICS

LECTURE 2: MULTIPLE TESTING

Msc programme in Mathematics and Statistics

University of Helsinki

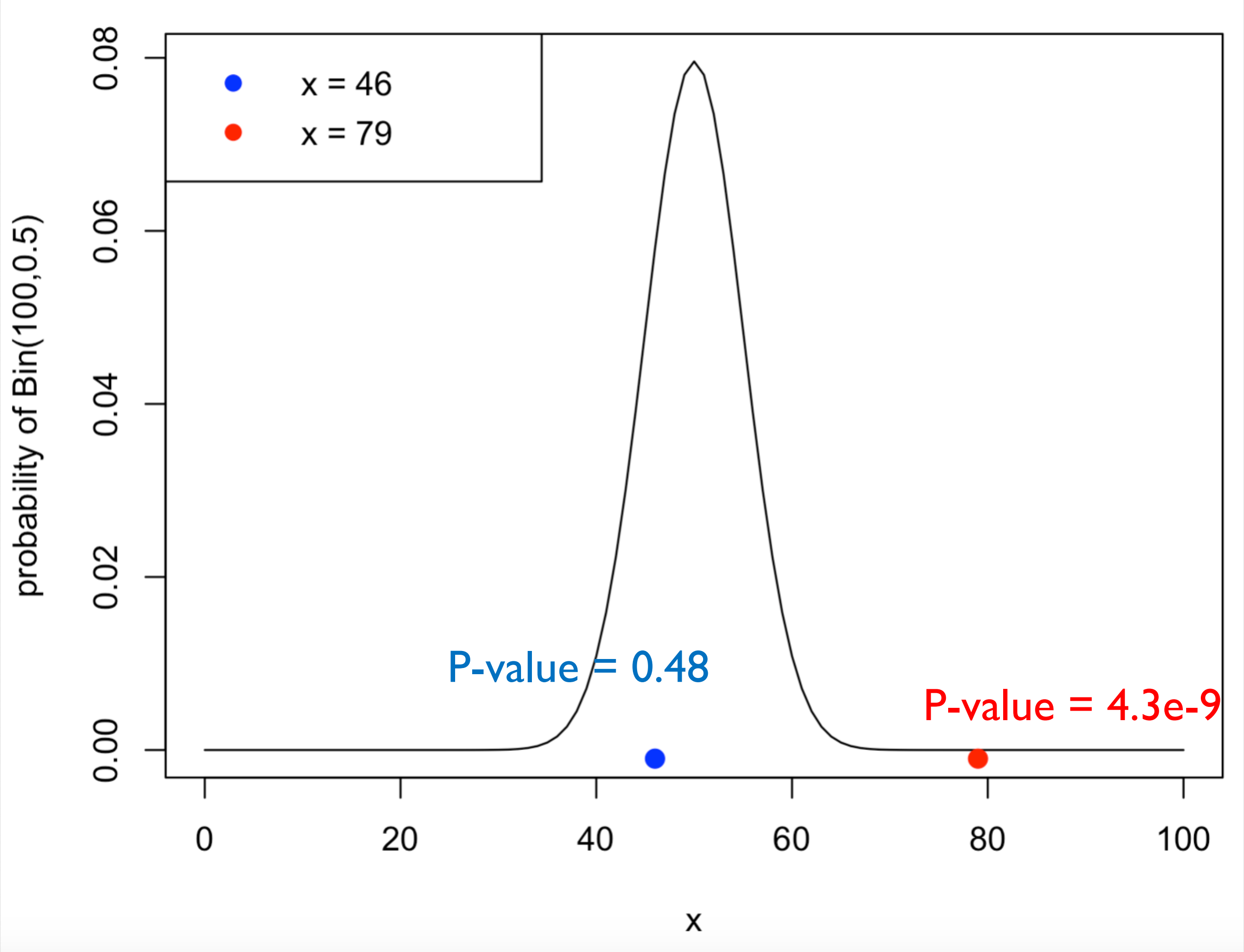
Matti Pirinen

THINK ABOUT COIN TOSSING

- We want to study whether a coin is fair, i.e., ends head-side up in 50% of tosses, on average
- We toss the coin n times and observe *x* heads
- The probability model is $x \sim \text{Bin}(n, \theta)$, where θ is the "success probability" of heads
- The **null hypothesis** is that the coin is fair, i.e., $\theta = 0.50$
- If we observe (i) $x = 46$ or (ii) $x = 79$ when $n = 100$, what can we infer?
  - (i) $\hat{\theta} = 0.46$, 95%CI = (0.360, 0.563), P-value = 0.48
  - (ii) $\hat{\theta} = 0.79$, 95%CI = (0.697, 0.865), P-value = 4.3e-9

```
> binom.test(x = 46, n = 100, p = 0.5)

        Exact binomial test

data:  46 and 100
number of successes = 46, number of trials = 100, p-value = 0.4841
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.3598434 0.5625884
sample estimates:
probability of success
                  0.46

> binom.test(x = 79, n = 100, p = 0.5)

        Exact binomial test

data:  79 and 100
number of successes = 79, number of trials = 100, p-value =
4.337e-09
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.6970846 0.8650563
sample estimates:
probability of success
                  0.79
```

NULL DISTRIBUTION BIN(N = 100, $\theta = 0.5$)



- Under the null, the expected value of x is 50 (=$n \cdot \theta$)
- The farther away from 50 the observed x is, the smaller is the probability of such an observation to happen under the null
- An observation that is very unlikely under the null hypothesis, makes us consider that the null hypothesis may not hold
- We can measure the consistency between the observation and the null hypothesis by **P-value**:
  - P-value is the probability that under the null hypothesis we would get at least as extreme observation as what we have actually observed
  - P-value is a tail probability of the null distribution
- Smaller P-value means that the observation is less consistent with the null hypothesis

P-VALUE IS NOT PROBABILITY OF ANY HYPOTHESIS

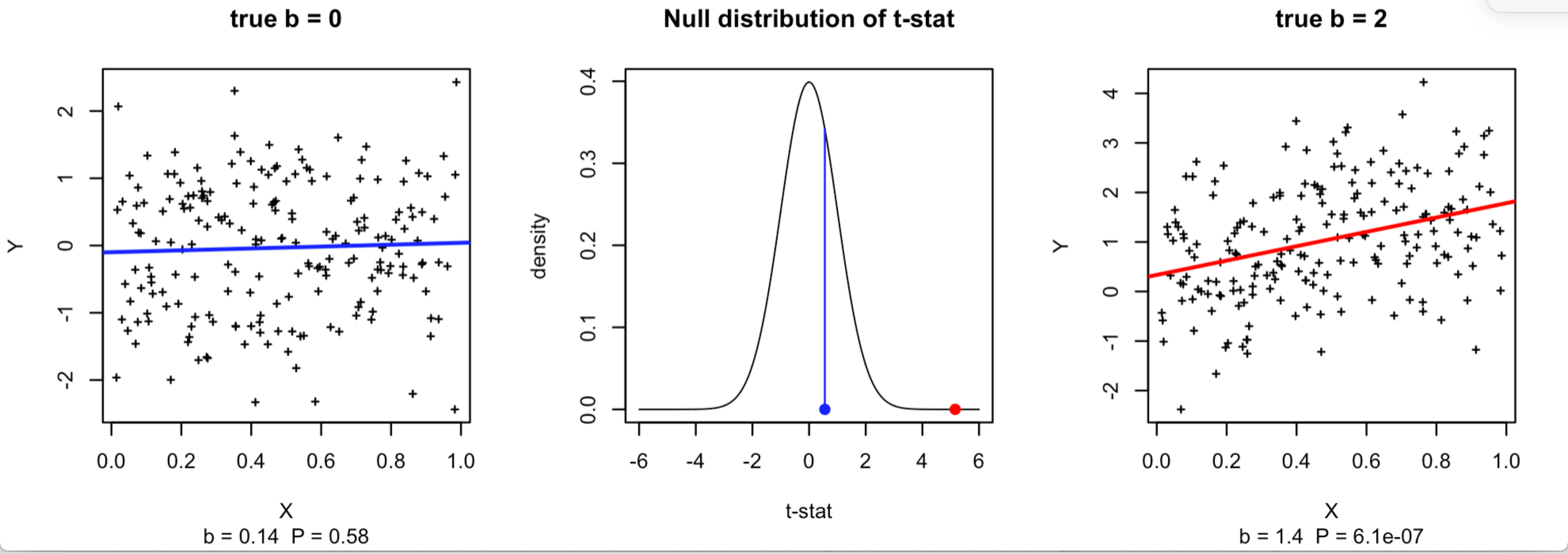
- P-value is probability of observing certain kind of data sets under the null hypothesis, i.e., it is of the form

  P-value = Pr(observing at least as extreme data as x | NULL holds)

- P-value is NOT of form ~~Pr(NULL holds | observation x)~~
  - This form would rather be a Bayesian posterior probability

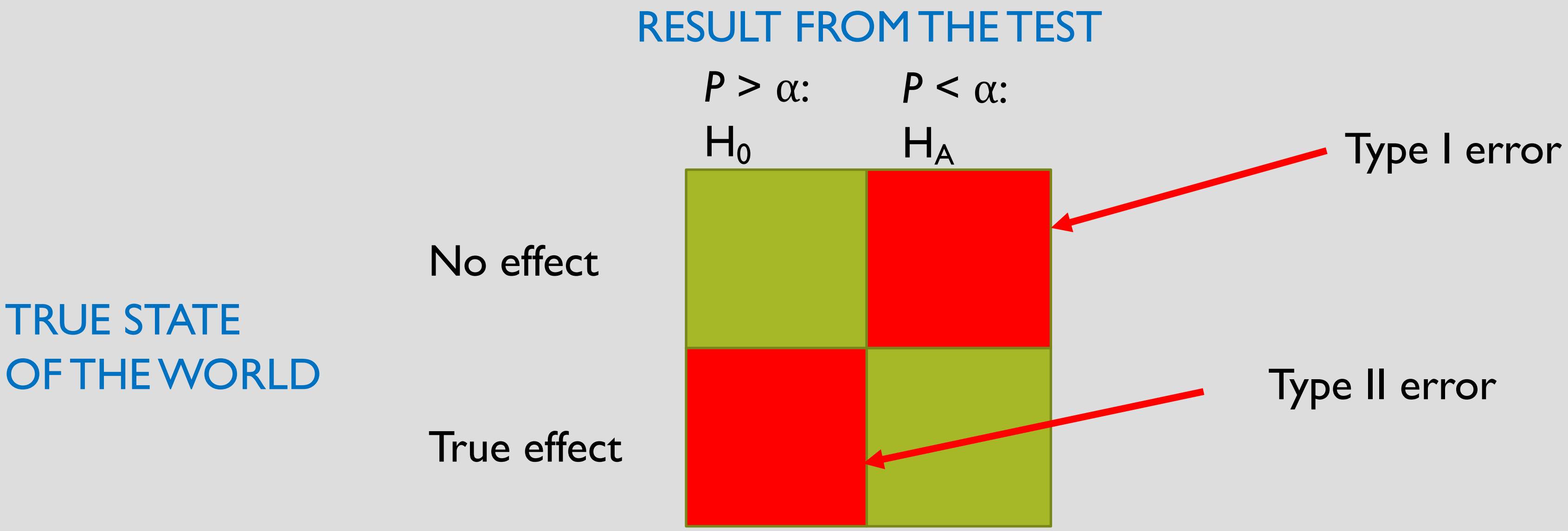
- P-value tells how probable certain kinds of data sets are to occur under the null hypothesis
- P-value cannot tell how probable the null hypothesis is given the observation
  - For this we will need the concept of Bayesian posterior probability

P-VALUES IN LINEAR REGRESSION



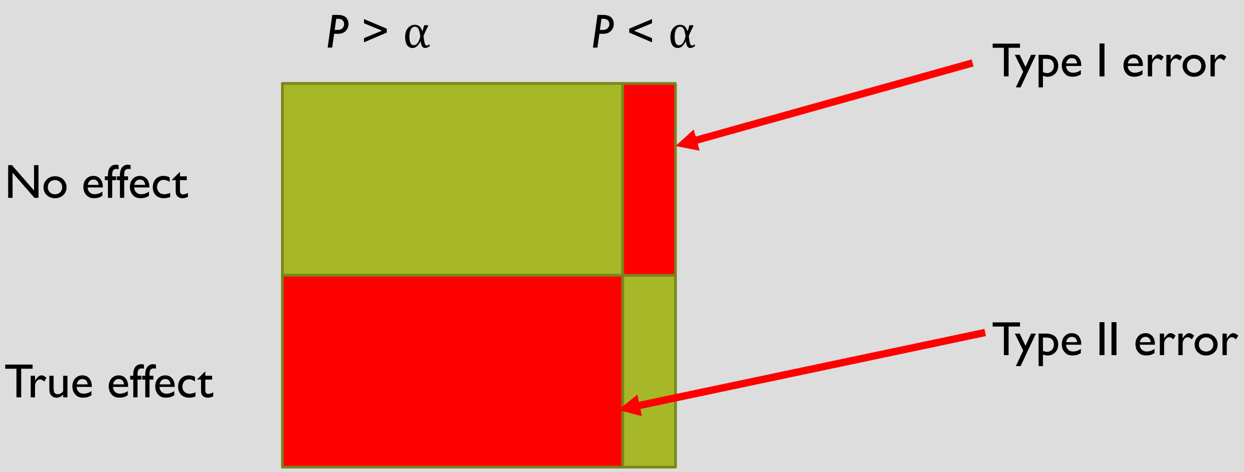
HYPOTHESIS TESTING FRAMEWORK

- Type I error: seeing an effect that is not real, a false positive
- Type II error: failing to see an effect that is real, a false negative
- Significance level α affects how likely different types of errors are to occur

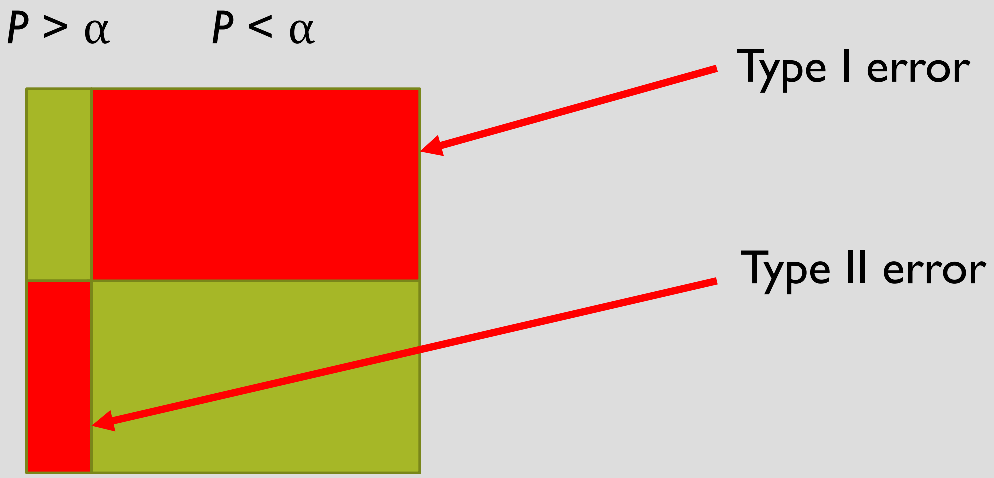


HYPOTHESIS TESTING FRAMEWORK

- Type I error: seeing an effect that is not real, a false positive
- Type II error: failing to see an effect that is real, a false negative
- With small significance level we rarely make Type I errors, but may often make Type II errors
- With large significance level we rarely make Type II errors, but may often make Type I errors



α small, e.g. 0.0001



α large, e.g. 0.5