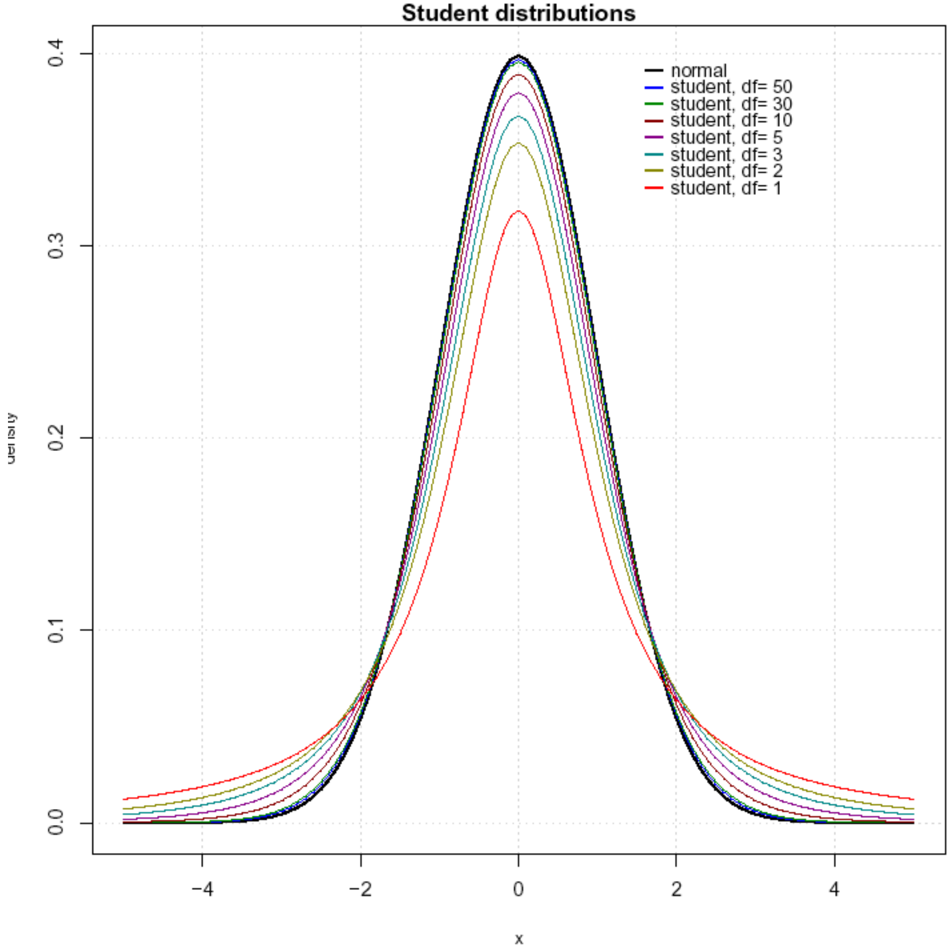
Tests of conformity

One-tailed Student test of conformity

- One-tailed tests
  - $H_0: m \geq m_0\,;\, H_A: m < m_0$
  - $H_0: m \leq m_0\,;\, H_A: m > m_0$
- Principle of the test
  - Estimate the difference between m and m_0
  - Compare this estimation with the theoretical distribution
- Usually, the variance is a priori not know, and has to be estimated
  - The theoretical distribution is thus the *Student (t)*
  - We estimate a single parameter (the standard error).We thus have $k=n-1$ degrees of freedom
  - α is placed on a single tail → use the value for $t_{1-\alpha}$ in Student theoretical distribution.

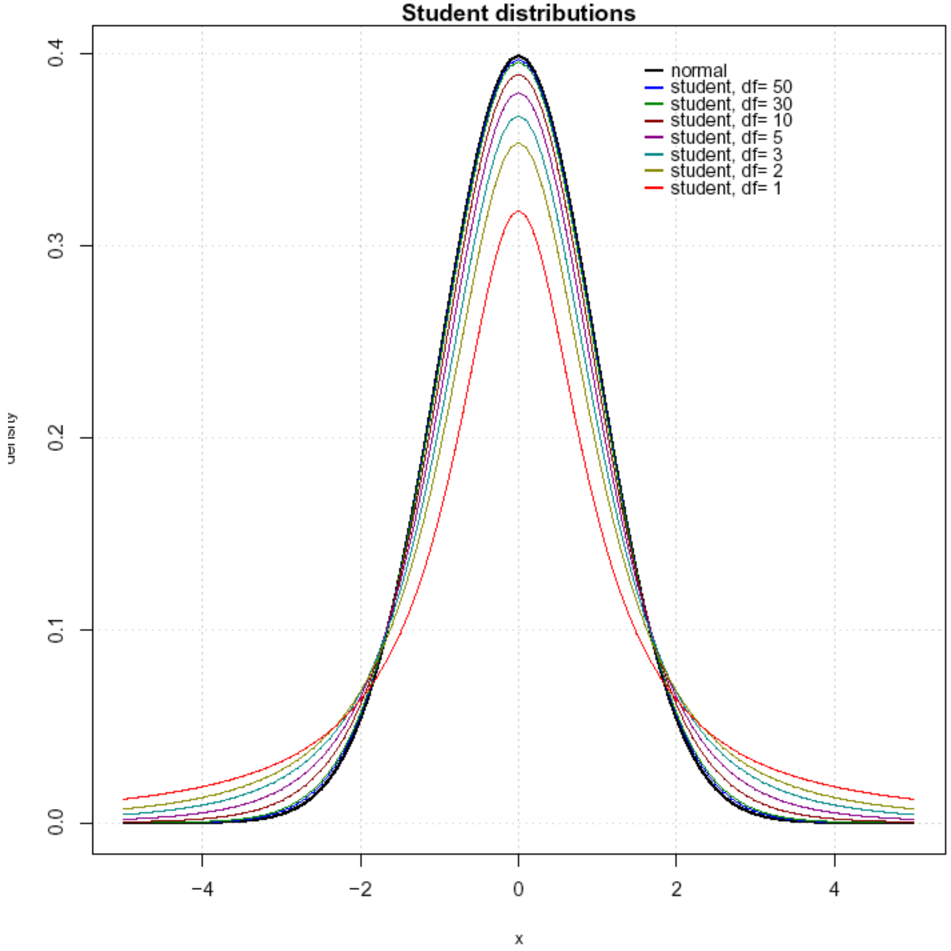


$$t_{obs} = \frac{\bar{x} - m_0}{\hat{\sigma}/\sqrt{n}} = \frac{\bar{x} - m_0}{s/\sqrt{n-1}}$$

$H_0: m \leq m_0$	Reject H_0 if	$t_{obs} \geq -t_{1-\alpha}$
$H_0: m \geq m_0$	Reject H_0 if	$t_{obs} \leq -t_{1-\alpha}$

Two-tailed Student test of conformity

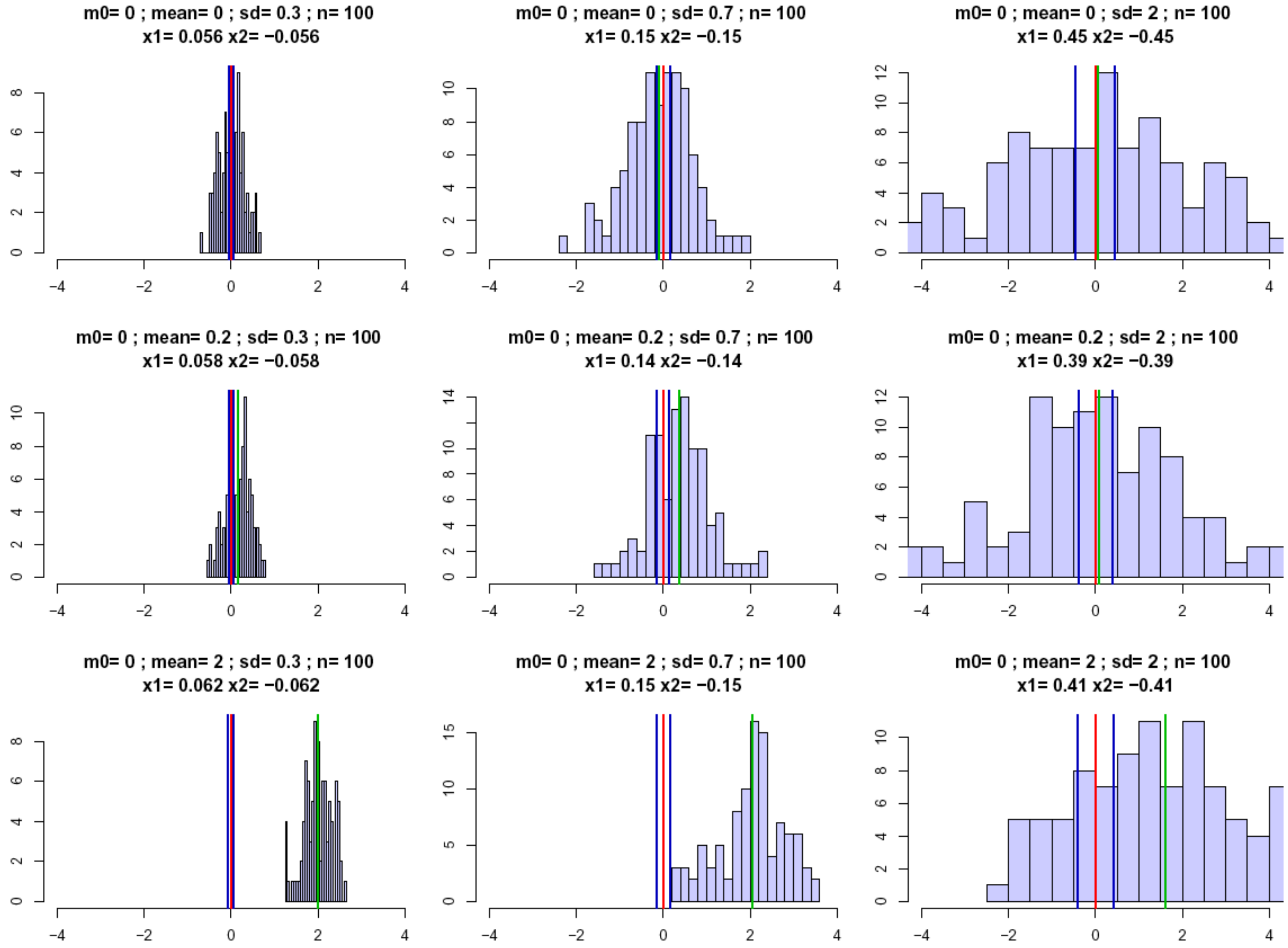
- Two-tailed tests
  - $H_0: m = m_0 ; H_A: m \neq m_0$
- Principle of the test
  - Estimate the absolute value of difference between m and m_0
  - Compare this estimation with the theoretical distribution
- Usually, the variance is a priori not know, and has to be estimated
  - The theoretical distribution is thus the *Student (t)*
  - We estimate a single parameter (the standard error).We thus have $k=n-1$ degrees of freedom
  - α is ***shared between the two tails*** → use the value for $t_{1-\alpha/2}$ in Student theoretical distribution.



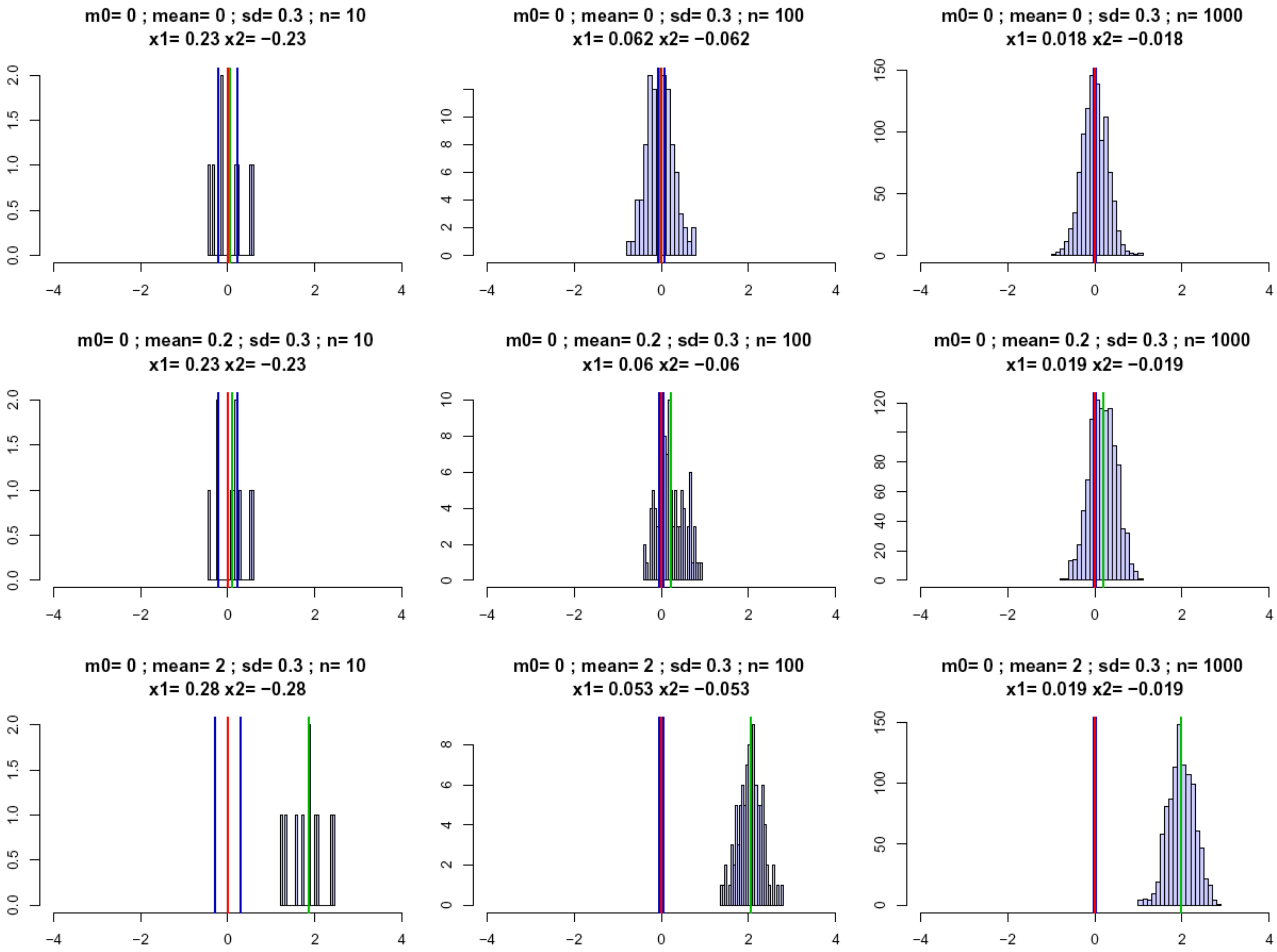
$$t_{obs} = \frac{|\bar{x} - m_0|}{\hat{\sigma}/\sqrt{n}} = \frac{|\bar{x} - m_0|}{s/\sqrt{n-1}}$$

$H_0: m = m_0$ Reject H_0 if $t_{obs} \geq t_{1-\alpha/2}$

Conformity of the mean - effect of mean and variance



Conformity of the mean - effect of sample size



Example: conformity of the mean with DNA chip data

- For a given DNA chip with 5921 spots, we calculated the following parameters
  - mean log-ratio = 0.0112
  - standard deviation = 0.4653009
- Test whether the mean of log-ratios is zero, with a risk of 5% first type error.
- What would be the probability to be wrong if you reject the null hypothesis ?

Student test of conformity with one microarray chip

- We perform a two-tailed test.
- The risk α is thus shared between the two tails → use the value for $t_{1-\alpha/2}$ in Student's table.
- To calculate t_{obs} we use the absolute value of the difference between the observed and the expected mean.
- In this case, the observed Student statistics is smaller than the theoretical one.
- The probability to be wrong if we reject the null hypothesis is 0.065, thus larger than our α.
- We can thus not reject the null hypothesis (we accept it).

$$H_0 : m = 0$$

$$\bar{x} = 0.0112$$

$$s = 0.4653009$$

$$n = 5921$$

$$m_0 = 0$$

$$\hat{\sigma} = s\sqrt{\frac{n}{n-1}} = 0.4653795$$

$$t_{obs} = \frac{|\bar{x} - m_0|}{\hat{\sigma}/\sqrt{n}} = \frac{|\bar{x} - m_0|}{s/\sqrt{n-1}} = 1.848477$$

$$df = n - 1 = 5920$$

$$\alpha = 0.05$$

$$t_{theor} = t_{1-\alpha/2,df} = 1.9604$$

$$t_{obs} < t_{theor} \rightarrow A(H_0)$$

$$P.value = 0.06458322$$

Exercise : single-gene test

- DNA chip data
  - measurement of gene expression in methionine+minimal medium (red channel) versus minimal medium (green channel)
  - 3 repetitions (DNA chip). The log-ratio is calculated for each chip.
- Question:
  - A given gene has the following log-ratios: 2.0, 3.1, 0.3.
  - Is this gene activated in presence of methionine ?