浙江大学爱丁堡大学联合学院

ZJU-UoE Institute

# The mathematics of the *t*-test
# Why and when does it work?

Nicola Romanò - nicola.romano@ed.ac.uk

(based on slides by Rob Young)

# This statistics lecture contains trade secrets!


- William Sealy Gosset, Head Experimental Brewer at Guinness
- Developed statistical techniques to assess quality of the finished product based on sampling during production.
- He published his 1908 paper under the pseudonym "Student" which is where the Student in Student's *t*-distribution comes from.

# Student's original paper

(actually, Gosset's!)

VOLUME VI MARCH, 1908 No. 1

BIOMETRIKA.

## THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

### *Introduction.*

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong. In a great number of cases the question finally turns on the value of a mean, either directly, or as the mean difference between the two quantities.

# Rationale for the *t*-test

- What do we mean when we say two values are "different"?
- Is the difference in values greater than what we might expect?

# Rationale for the *t*-test

- What do we mean when we say two values are ‘different’?

- Is the difference in values greater than what we might expect?

- Compare the sample mean to a known value or distribution.
  - Null hypothesis: the values are equal, they come from the same population

- Appropriate for small sample sizes (n < 100).

- Generally, we assume a normal distribution of the underlying population.

# Learning objectives

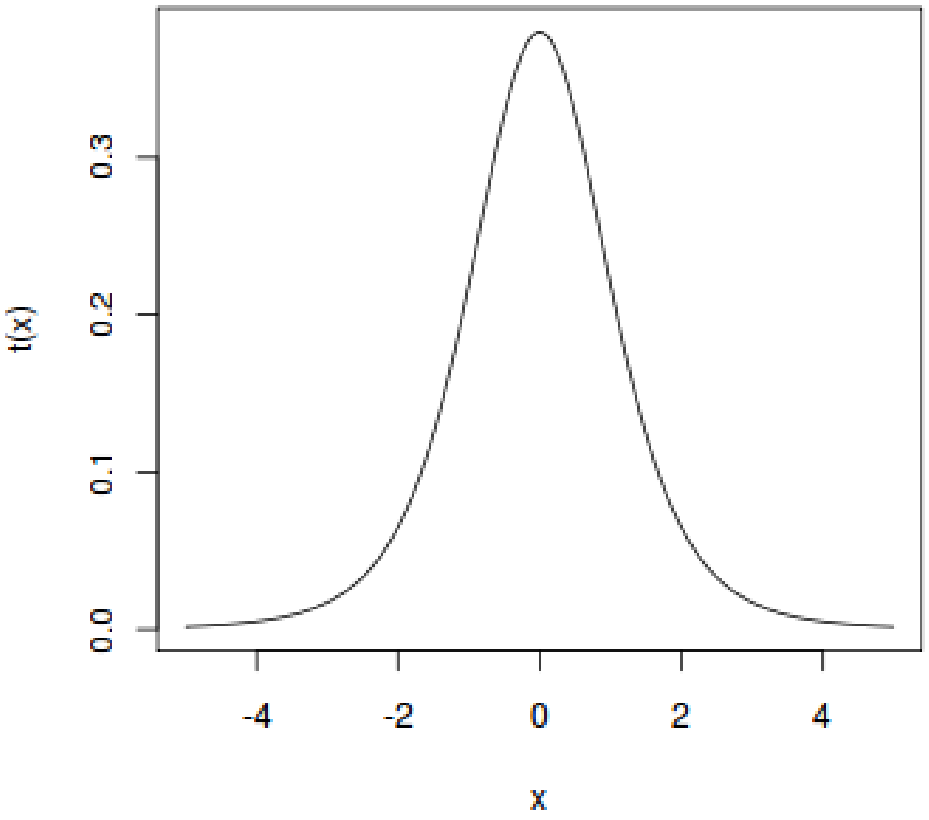
After this lecture, you should be able to:

- **Understand the mathematics behind the *t*-test.**

- Use the Student's *t*-distribution to determine the significance of a given sample.

- Describe the assumptions that need to be met to apply the *t*-test appropriately.

# Student's *t*-distribution




- Continuous probability distribution.
- Symmetric and bell-shaped.
- Derived from a small sample size where the population standard deviation is unknown.

# The only equation in this lecture: How to calculate the *t*-statistic

$$t = \frac{Z}{s} = \frac{\text{Difference in values}}{\text{Uncertainty in values}}$$

standard error of the mean

# The only equation in this lecture: How to calculate the *t*-statistic

sample mean

population mean

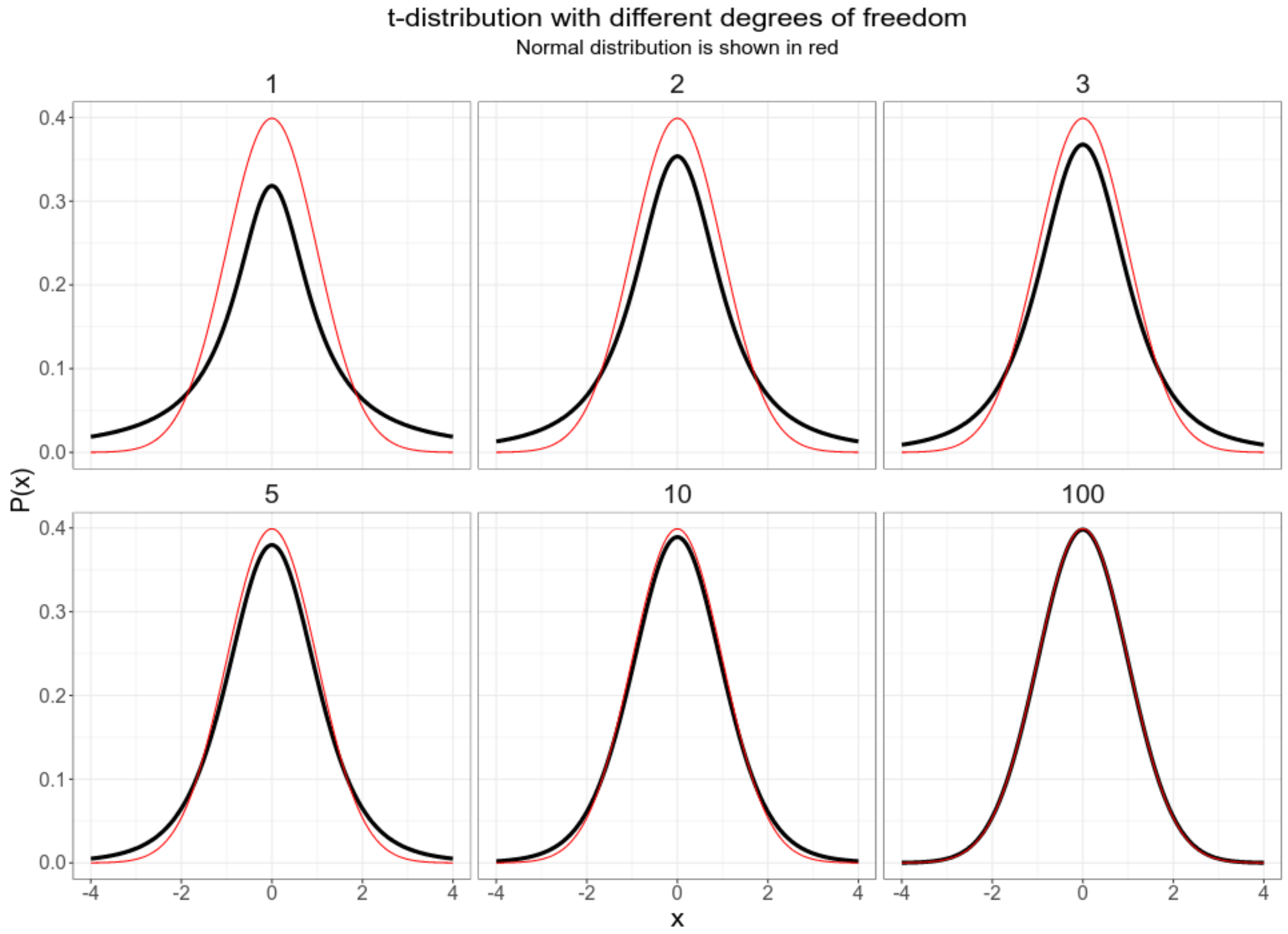
$$t = \frac{Z}{s} = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}}$$

standard error of the mean

estimated population standard deviation

sample size

# Student's *t*-distribution: approaching the normal distribution




- The *t*-distribution is a generalization of the normal distribution.
- **Symmetric and bell-shaped**.
- Its distribution depends on a single parameter, the **degrees of freedom** (d.o.f, see next slide)
- It is heavier tailed compared to the **normal distribution** at small d.o.f.
- As d.o.f. increase, it converges to the normal distribution

# So just what are these degrees of freedom?

The **degrees of freedom** (d.o.f.) in a statistical calculation represent the number of independent pieces of information that are used to calculate that statistic.

In other words, how many values involved in a calculation have freedom to vary, without changing the result?

In the case of t-test (and in several other cases...) this means

d.o.f. = n-1

(n=sample size)

# …or to put it another way

## The Freedom to Vary

First, forget about statistics. Imagine you're a fun-loving person who loves to wear hats. You couldn't care less what a degree of freedom is. You believe that variety is the spice of life.

Unfortunately, you have constraints. You have only 7 hats. Yet you want to wear a different hat every day of the week.

On the first day, you can wear any of the 7 hats. On the second day, you can choose from the 6 remaining hats, on day 3 you can choose from 5 hats, and so on.

When day 6 rolls around, you still have a choice between 2 hats that you haven't worn yet that week. But after you choose your hat for day 6, you have no choice for the hat that you wear on Day 7. You *must* wear the one remaining hat. You had 7-1 = 6 days of "hat" freedom—in which the hat you wore could vary!

That's kind of the idea behind degrees of freedom in statistics. Degrees of freedom are often broadly defined as the number of "observations" (pieces of information) in the data that are free to vary when estimating statistical parameters.

# Learning objectives

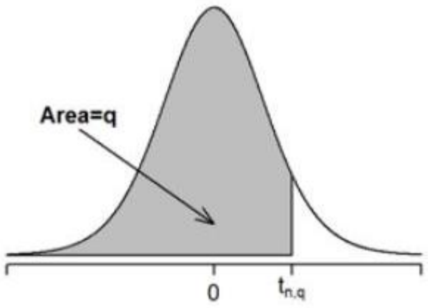
After this lecture, you should be able to:

- Understand the mathematics behind the *t*-test.

- **Use the Student's *t*-distribution to determine the significance of a given sample.**

- Describe the assumptions that need to be met to apply the *t*-test appropriately.

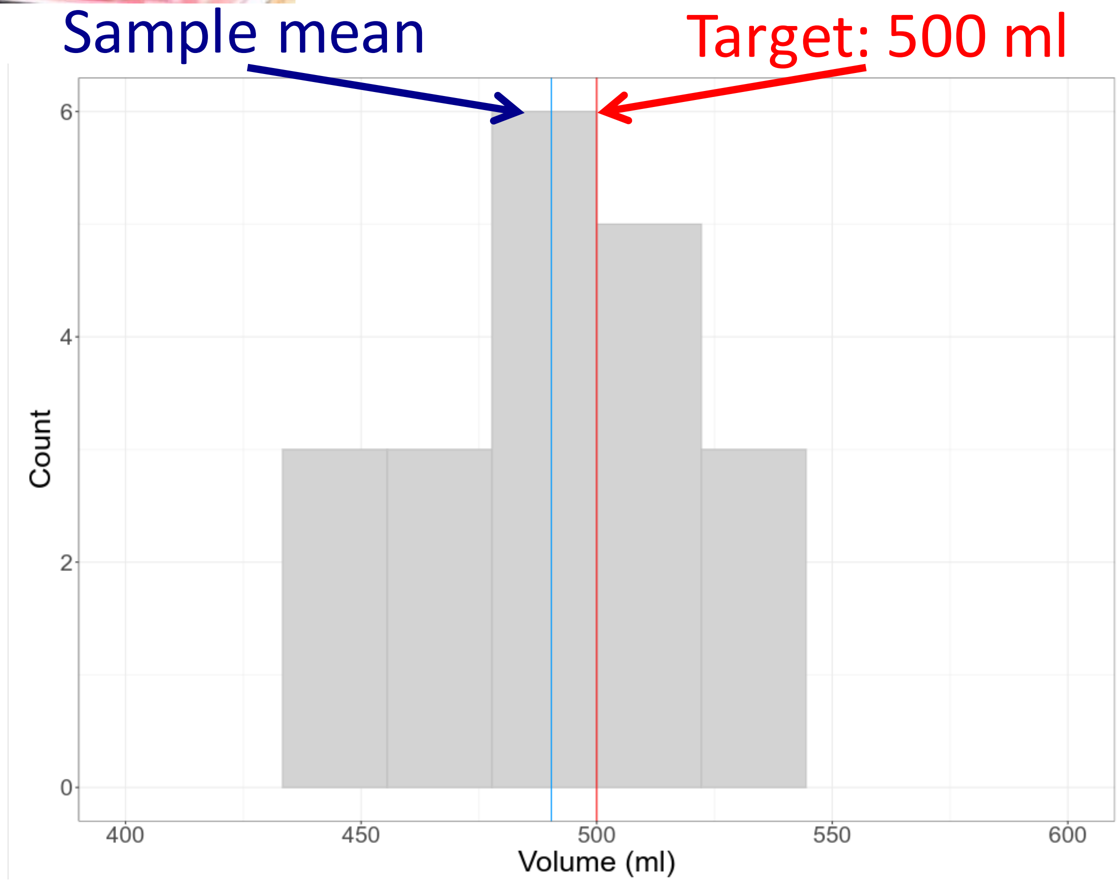
# Critical values of the Student's *t*-distribution




| | $q = 0.6$ | 0.75 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.9975 | 0.999 | 0.9995 |
|---|---|---|---|---|---|---|---|---|---|---|
| $n = 1$ | 0.3249 | 1.0000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 127.321 | 318.309 | 636.619 |
| 2 | 0.2887 | 0.8165 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 14.089 | 22.327 | 31.599 |
| 3 | 0.2767 | 0.7649 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 7.453 | 10.215 | 12.924 |
| 4 | 0.2707 | 0.7407 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 5.598 | 7.173 | 8.610 |
| 5 | 0.2672 | 0.7267 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 4.773 | 5.893 | 6.869 |
| 6 | 0.2648 | 0.7176 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 4.317 | 5.208 | 5.959 |
| 7 | 0.2632 | 0.7111 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.029 | 4.785 | 5.408 |
| 8 | 0.2619 | 0.7064 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 3.833 | 4.501 | 5.041 |
| 9 | 0.2610 | 0.7027 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 3.690 | 4.297 | 4.781 |
| 10 | 0.2602 | 0.6998 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 3.581 | 4.144 | 4.587 |
| 11 | 0.2596 | 0.6974 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 3.497 | 4.025 | 4.437 |
| 12 | 0.2590 | 0.6955 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.428 | 3.930 | 4.318 |
| 13 | 0.2586 | 0.6938 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.372 | 3.852 | 4.221 |
| 14 | 0.2582 | 0.6924 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.326 | 3.787 | 4.140 |

- Each *t*-value has an associated probability, determined by the degrees of freedom.
- Critical values can be looked up in Statistical Tables like this.

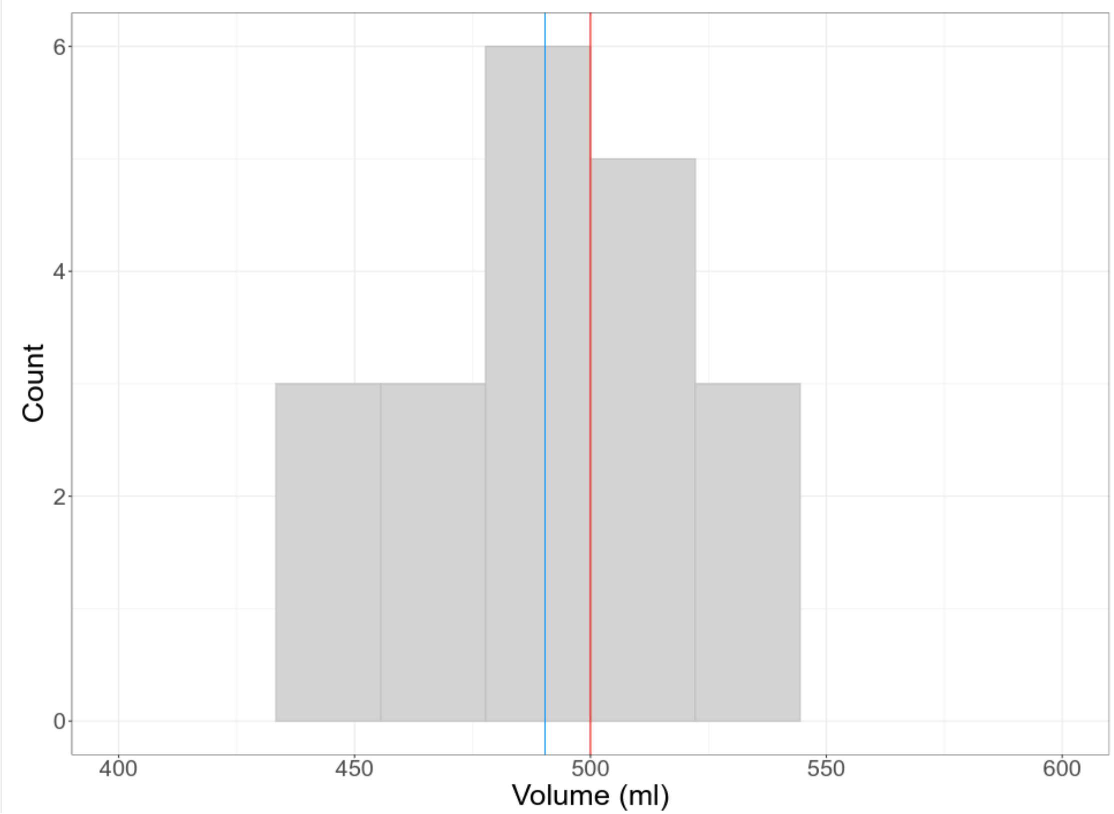

# Is the factory filling each bottle with enough Guinness?



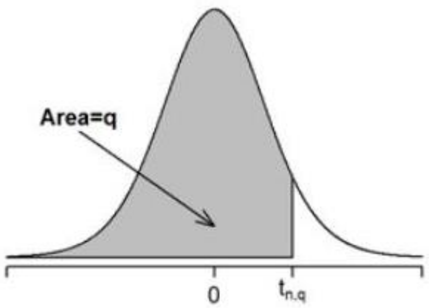

- 10 million glasses are drunk (and assumed to be brewed) every day, so we can't check them all!
- The volume of 20 bottles have been measured.
- The mean volume is 490.4 ml, but the required volume is 500 ml.

# Calculating the *t*-statistic

$$t = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}} = \frac{490.4 - 500}{25.7/\sqrt{20}} = -1.67$$



# Is this *t*-statistic significant?

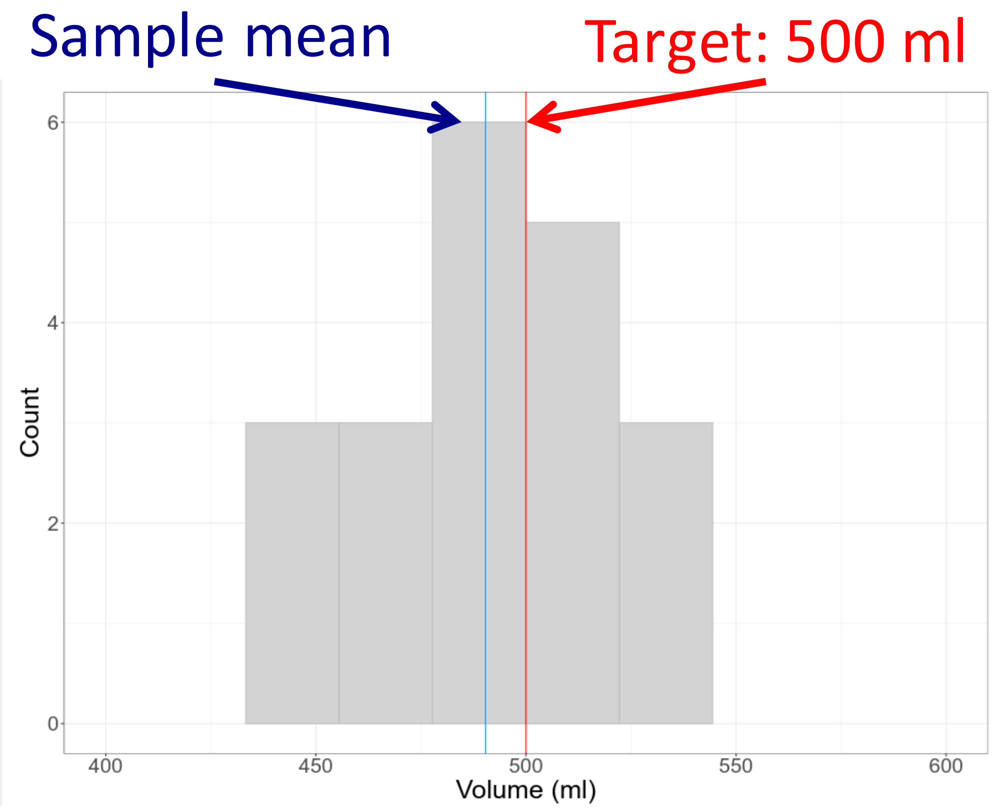



| d.f. | $t_{.100}$ | $t_{.050}$* | $t_{.025}$** | $t_{.010}$ | $t_{.005}$ | d.f. |
|---|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 1 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 2 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 3 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 4 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 6 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 7 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 8 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 9 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 10 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 11 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 12 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 13 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 14 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 15 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 16 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 17 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 18 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 19 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 20 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 21 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 22 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 23 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 24 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 25 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 26 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 27 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 28 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 29 |
| inf. | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | inf. |

There is only a 5% probability that a sample with 10 degrees of freedom will have a t value greater than 1.812.
* one tail 5% α risk ** two tail 5% α risk

- Observed $t$ = -1.67
- How many degrees of freedom (20 bottles)?

# In R, use t.test(SAMPLE, mu = VALUE)




```
> t.test(bottles$Volume, mu = 500)

One Sample t-test data:
bottles$Volume
t = -1.6687, df = 19, p-value = 0.1116

alternative hypothesis: true mean is not equal to 500
95 percent confidence interval: 478.3062 502.4473
sample estimates: mean of x 490.3767
```

- Much easier than looking up tables of critical values!
- Reports a *p*-value, here it is $p = 0.11$.

# Learning objectives

After this lecture, you should be able to:

- Understand the mathematics behind the *t*-test.

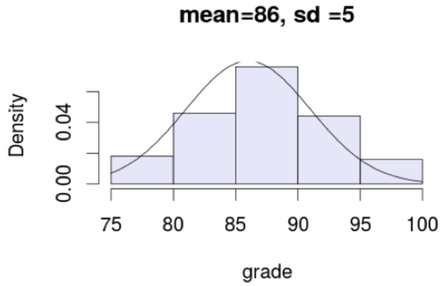
- Use the Student's *t*-distribution to determine the significance of a given sample.

- **Describe the assumptions that need to be met to apply the *t*-test appropriately.**

# Assumptions required for using the *t*-statistic

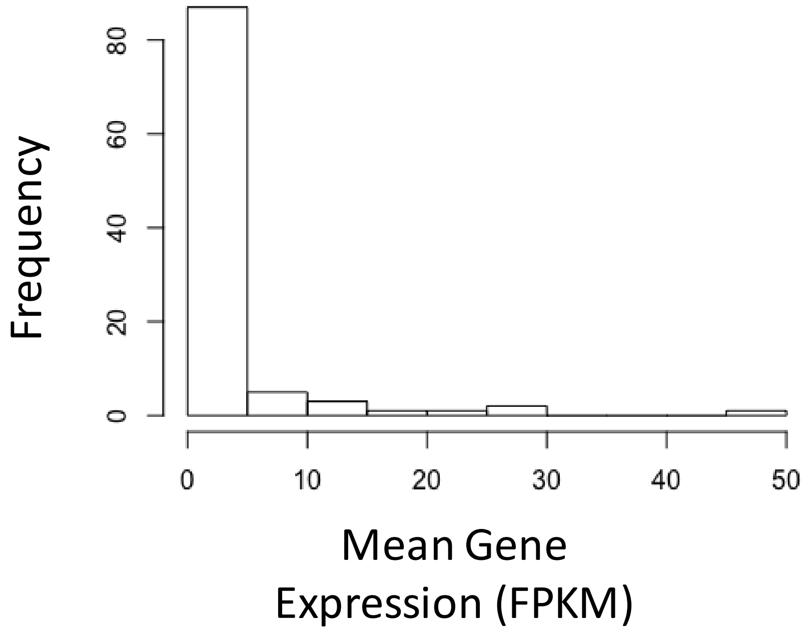
1. Data is continuous and randomly-selected.
   → See lecture on sampling

2. The sample is normally distributed.
   → See lecture on sampling distributions



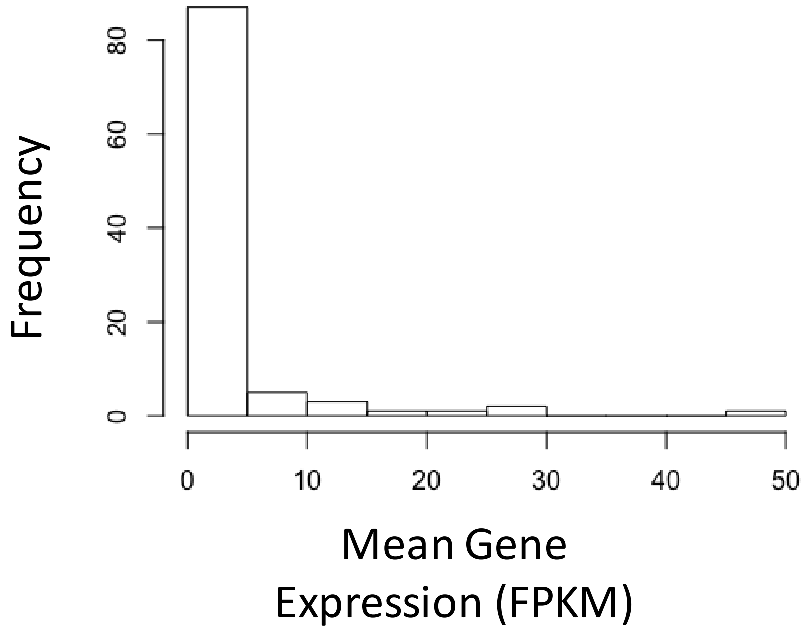

3. The mean and standard error are independent.
   → Nearly always true, but could test by simulation.

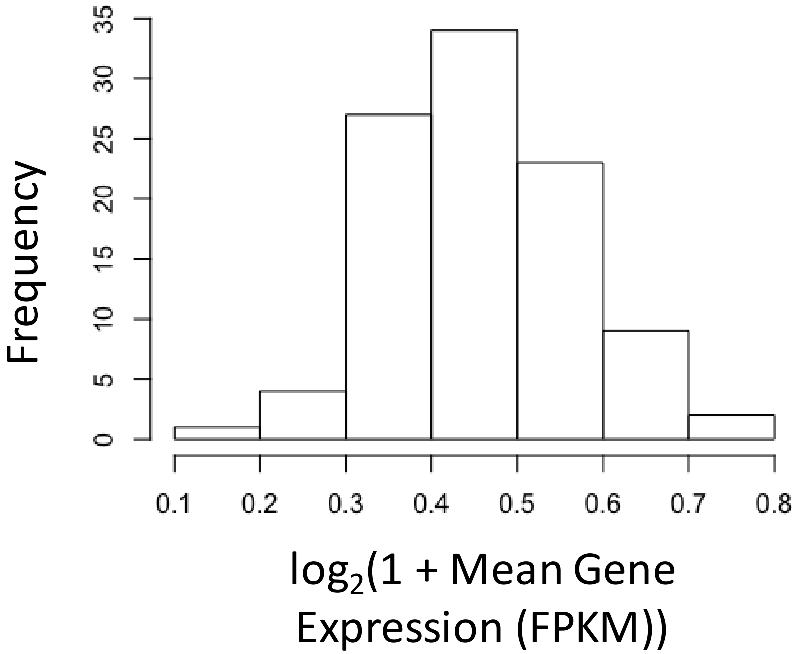
# Assumption 2/3: but what if my sample is not normally distributed?




- Gene expression measurements are frequently very far from the normal or *t*-distribution.

# Assumption 2/3: but what if my sample is not normally distributed?



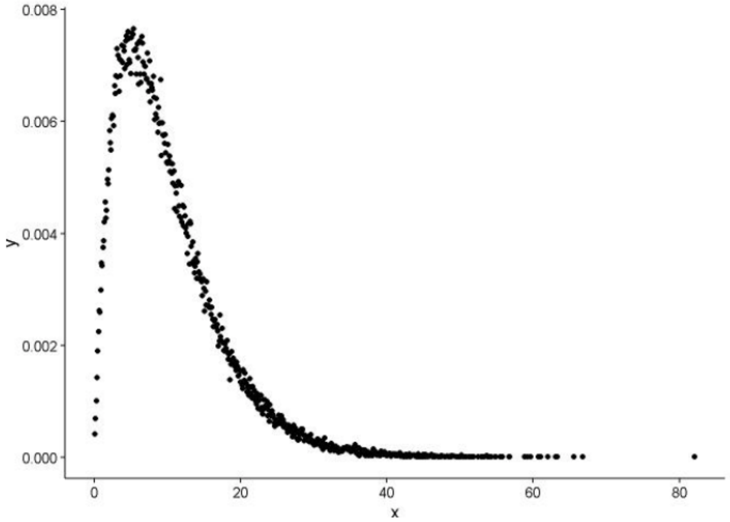

- Gene expression measurements are frequently very far from the normal or *t*-distribution.
- Transformations, such as the logarithmic, can make your data more ‘normal’.

# Assumption 2/3: but what if my transformed sample is still not normally distributed???

# Assumption 2/3: but what if my transformed sample is still not normally distributed???

Comparing area under curve to the right/left of a value, e.g. can I pass through the right side without getting significantly wet

e.g. getting less than 1% of water




- Could try simulations or sampling from a known distribution

  → see lecture on comparing two means using simulation

# Assumptions required for using the *t*-statistic

1. Data is continuous and randomly-selected.
   $\rightarrow$ see lecture on sampling

2. The sample is normally distributed.
   $\rightarrow$ see lecture on sampling distributions

3. The mean and standard error are independent.
   $\rightarrow$ nearly always true, but can be tested (problem sheet)

# Learning objectives

Now you should be able to:

- Understand the mathematics behind the *t*-test.
- Use the Student's *t*-distribution to determine the significance of a given sample.
- Describe the assumptions that need to be met to apply the *t*-test appropriately.

浙江大学爱丁堡大学联合学院
ZJU-UoE Institute

# The mathematics of the *t*-test
## Why and when does it work?

Nicola Romanò - nicola.romano@ed.ac.uk