

# Error analysis in biology

Marek Gierliński  
Division of Computational Biology

Hand-outs available at <http://is.gd/statlec>

Errors, like straws, upon the surface flow;  
He who would search for pearls must dive below  
*John Dryden (1631-1700)*

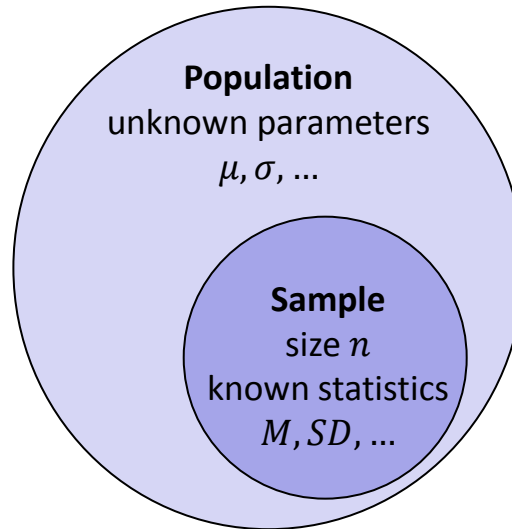
# Previously on Errors...

## Random errors

- measurement error
- reading error
- counting error
- **sampling error**

**Statistical estimator** is a sample attribute used to estimate a population parameter

- mean, median, mode
- variance, standard deviation
- correlation

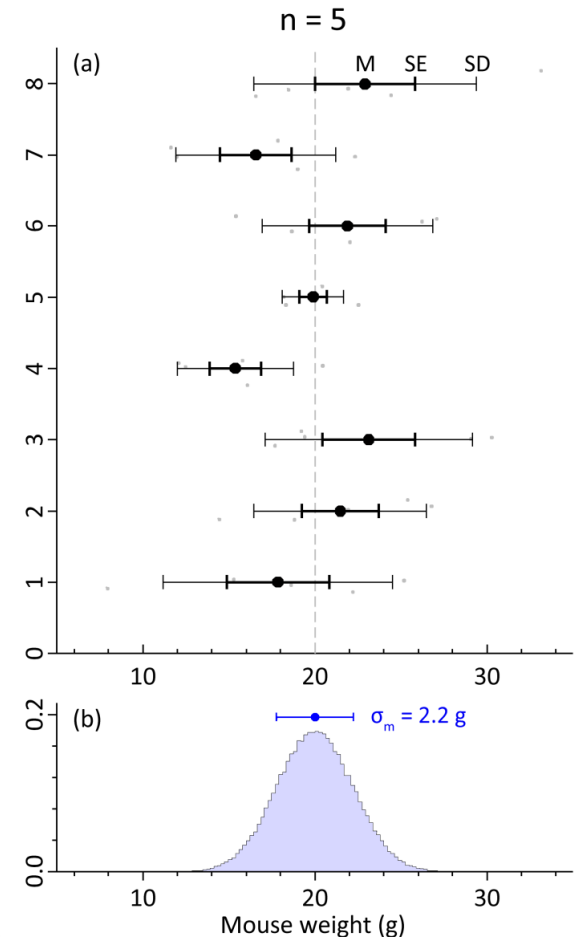


## Standard deviation

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - M)^2}$$

## Standard error

$$SE = \frac{SD}{\sqrt{n}}$$



**Sampling distribution** of the mean – distribution of sample means from repeated experiments

# Standard error of the mean

## Hypothetical experiment

- 10,000 samples of 5 mice
- Build a distribution of sample means
- Width of this distribution is the true uncertainty of the mean

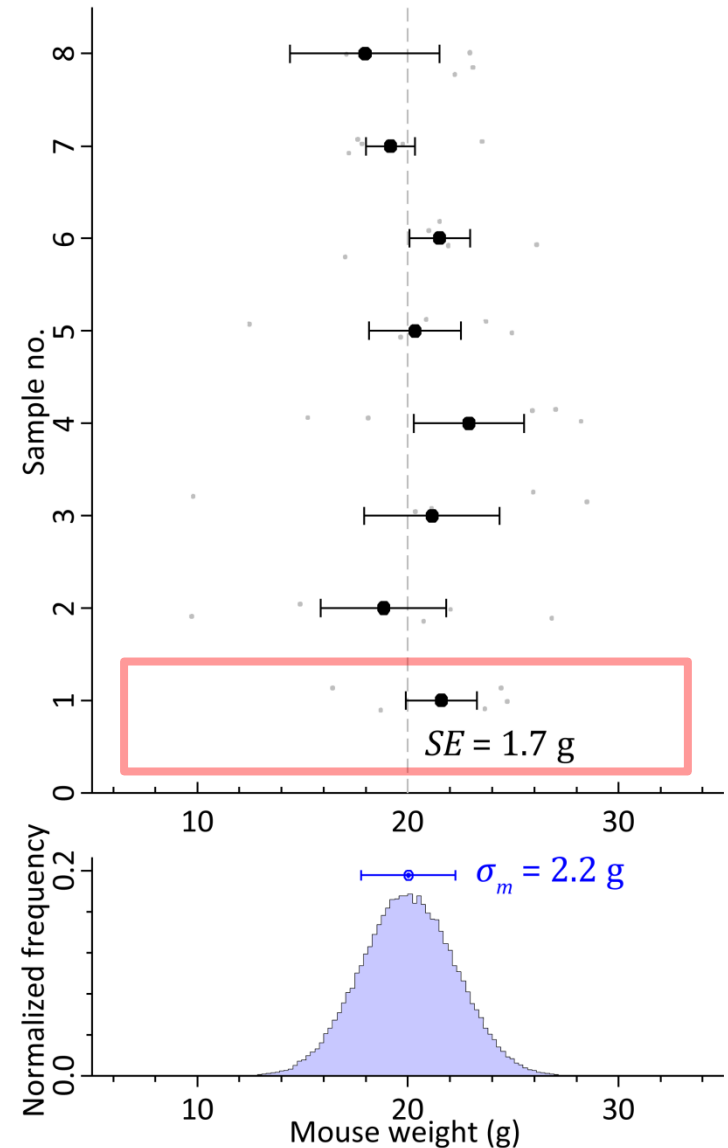
$$\sigma_m = \frac{\sigma}{\sqrt{n}} = 2.2 \text{ g}$$

## Real experiment

- 5 mice
- Measure body mass:  
16.4, 18.7, 23.7, 24.4, 24.7 g
- Find standard error

$$SE = \frac{SD}{\sqrt{n}} = 1.7 \text{ g}$$

***SE is an approximation of  $\sigma_m$***

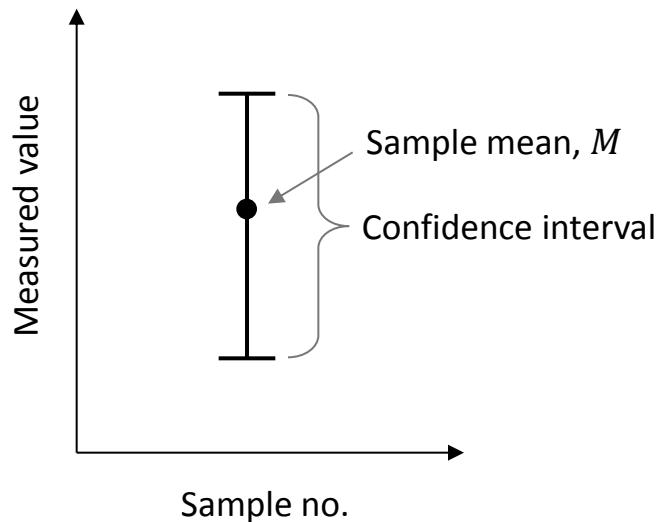
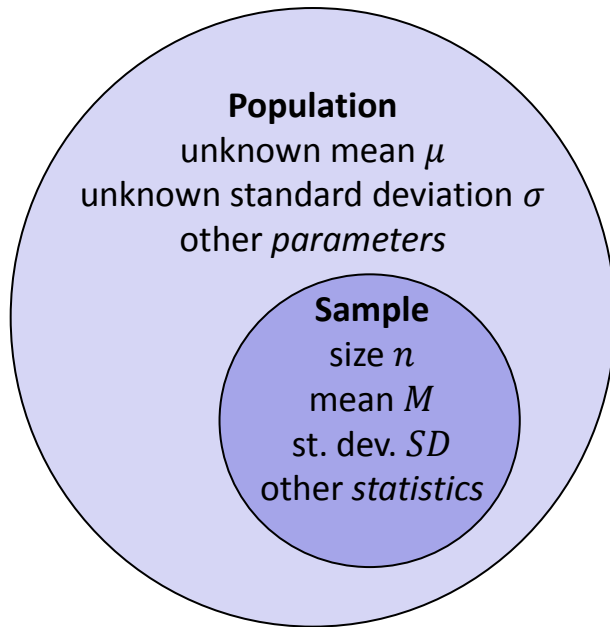


## 4. Confidence intervals I

“Confidence is what you have before you understand the problem”

*Woody Allen*

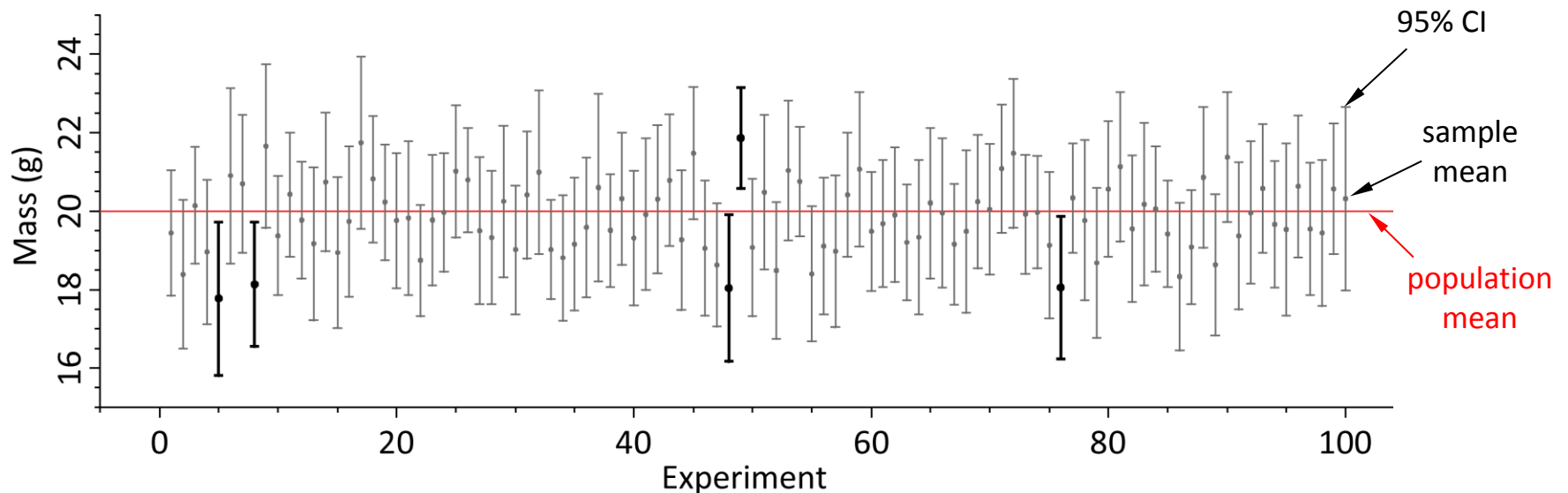
# Confidence intervals



- Sample mean,  $M$ , is a statistical estimator of the true mean,  $\mu$
- How good is  $M$ ?
  
- Confidence interval: a range  $[M_L, M_U]$  around  $M$ , where we think the true mean is, with a *certain confidence*
  
- This can be done for any population parameter
  - mean
  - median
  - standard deviation
  - correlation
  - proportion
  - etc.

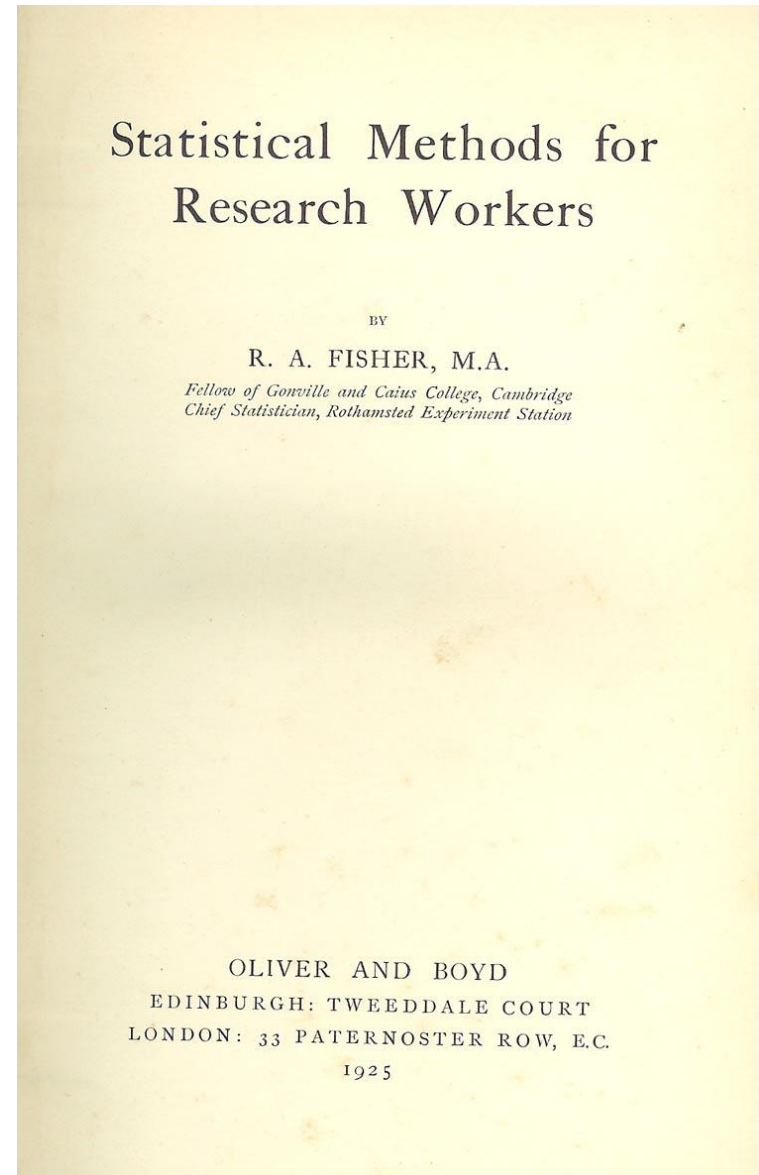
# What is confidence?

- Consider a 95% confidence interval for the mean  $[M_L, M_U]$
- This **does not mean** there is a 95% probability of finding the true mean in the calculated interval
- The true population mean is a constant number, not a random variable!
- Repeated experiments: in 95% of cases the true mean would be within the calculated interval, and in 5% of cases outside it



# Why 95%?

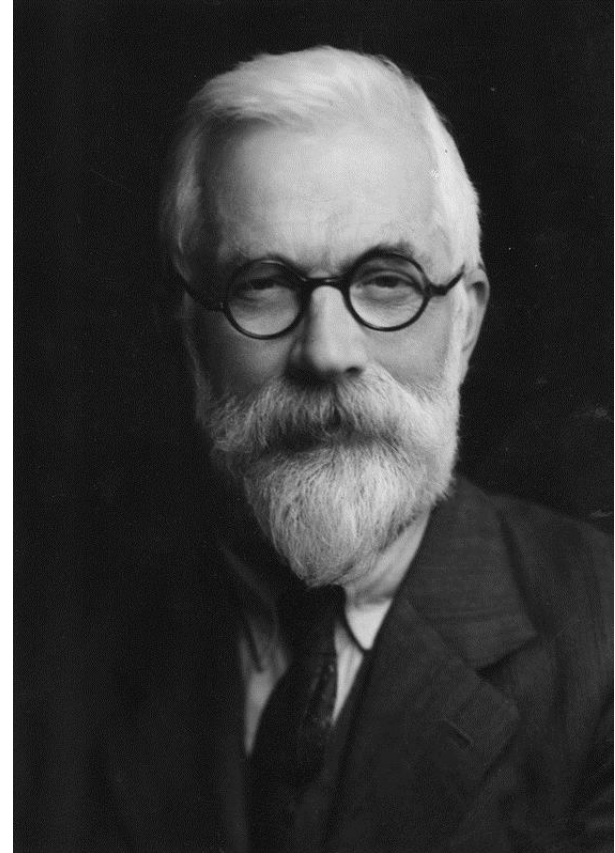
- Textbook by Ronald Fisher (1925)
- He thought 95% confidence interval was “convenient” as it resulted in 1 false indication in 20 trials
- He published tables for a few probabilities, including  $p = 5\%$
- The book had become one of the most influential textbooks in 20<sup>th</sup> century statistics
- However, there is nothing special about 95% confidence interval or  $p$ -value of 5%



# Ronald Fisher

---

- Probably the most influential statistician of the 20<sup>th</sup> century
- Also evolutionary biologists
- Went to Harrow School and then Cambridge
- Arthur Vassal, Harrow's schoolmaster:  
*I would divide all those I had taught into two groups: one containing a single outstanding boy, Ronald Fisher; the other all the rest*
- Didn't like administration and admin people: "an administrator, not the highest form of human life"

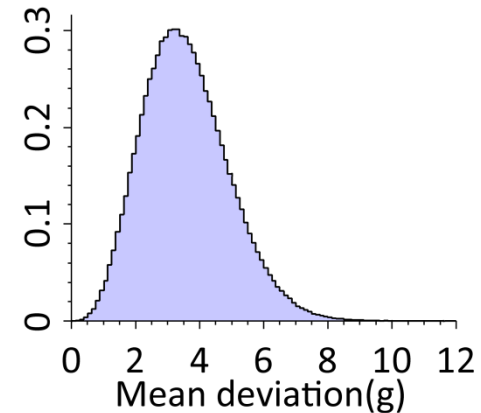
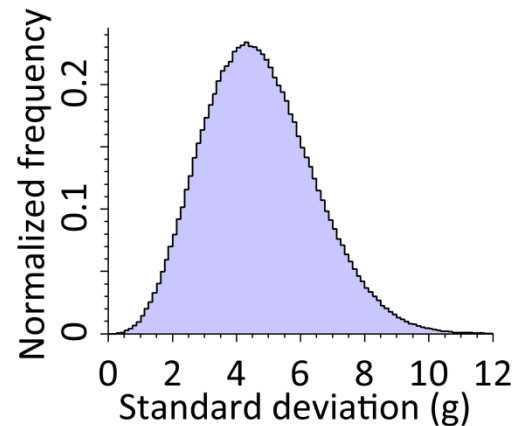
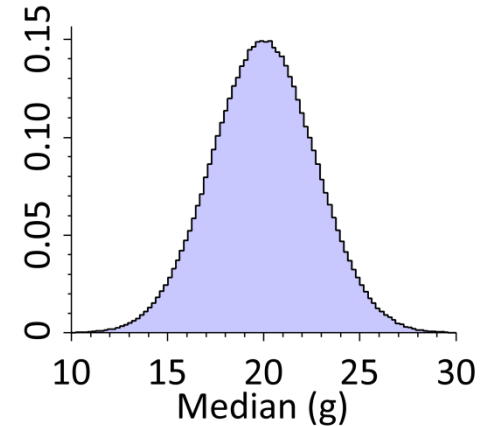
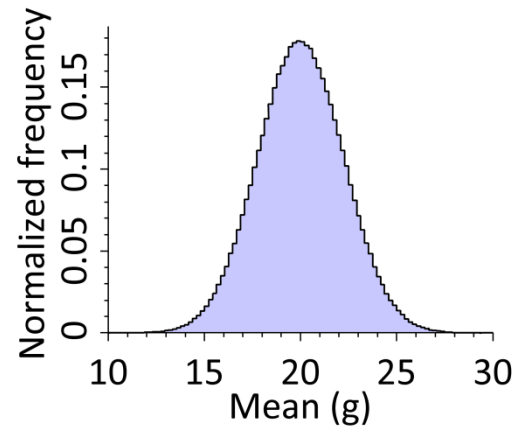


Ronald Fisher (1890-1962)



# Sampling distribution

- *Gedankenexperiment*
- Consider an unknown population you wish to characterize
- Draw lots of samples of size  $n$  from this population
- Calculate an estimator from each sample, for example
- Build a distribution of the estimator
- This is a *sampling distribution*
- Width of the sampling distribution is a standard error

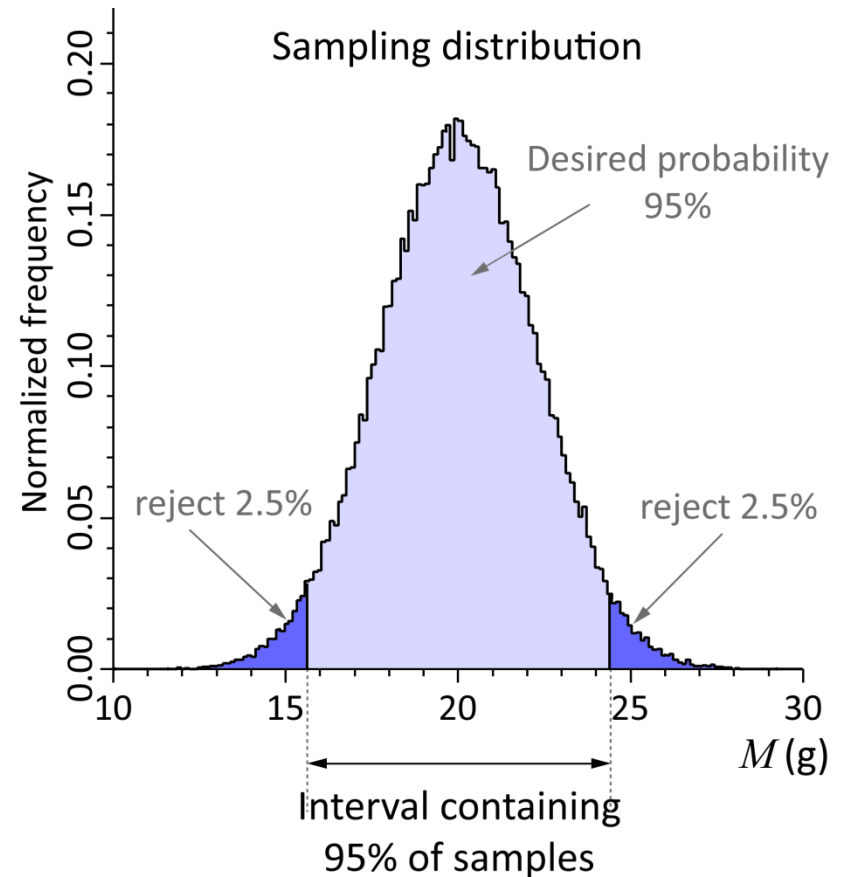


Examples of sampling distribution

$10^6$  samples of  $n = 5$  from  $\mathcal{N}(20, 5)$

# Sampling distribution of the mean

- The distribution curve represents all samples
- Keep the region corresponding to the required confidence, e.g. 95%
- Reject 2.5% on each side
- This gives confidence interval for our estimator

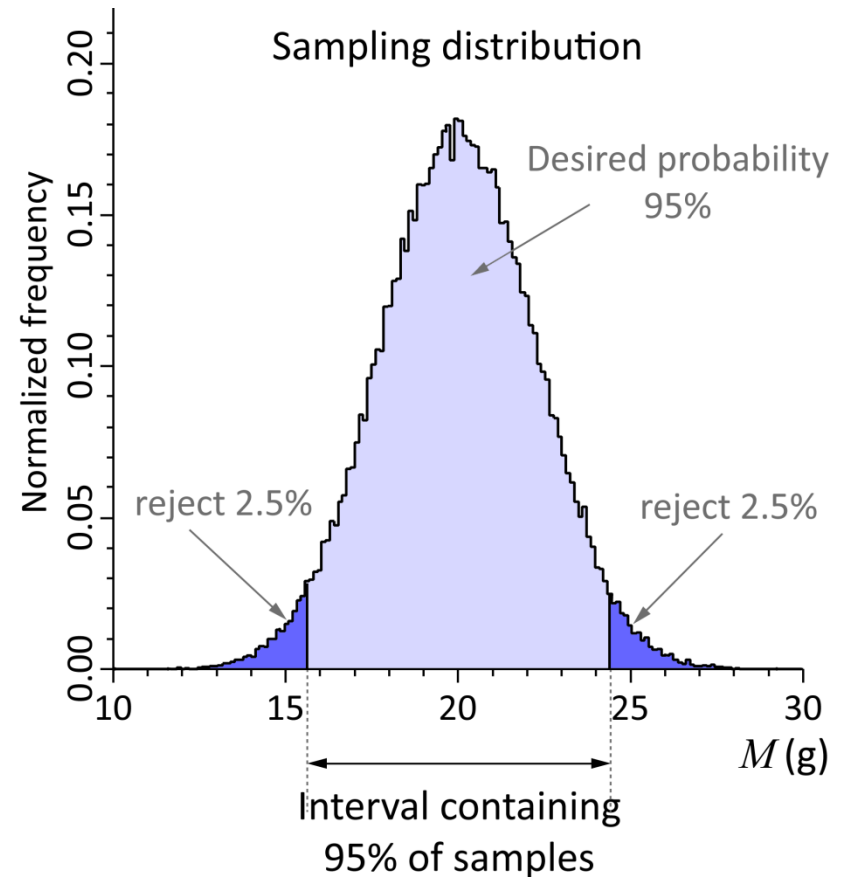


100,000 samples of 5 mice from normal population  
with  $\mu = 20$  g and  $\sigma = 5$  g

Mean body weight calculated for each sample

# Sampling distribution of the mean

- The distribution curve represents all samples
- Keep the region corresponding to the required confidence, e.g. 95%
- Reject 2.5% on each side
- This gives confidence interval for our estimator
- In real life you can't draw thousands of samples!
- Instead you can use a *known probability distribution* to calculate probabilities



100,000 samples of 5 mice from normal population  
with  $\mu = 20$  g and  $\sigma = 5$  g

Mean body weight calculated for each sample

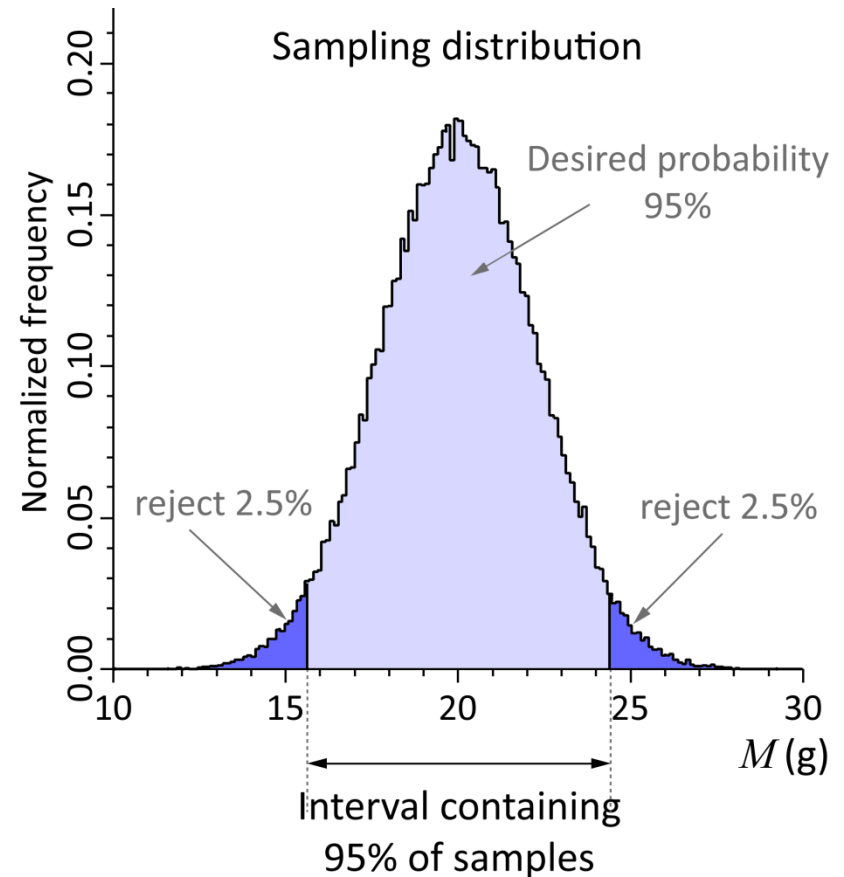
# Sampling distribution of the mean

- Sampling distribution of the mean is not known

- For the given sample find  $M$ ,  $SD$  and  $n$
- Let us define a statistic

$$t = \frac{M - \mu}{SE}$$

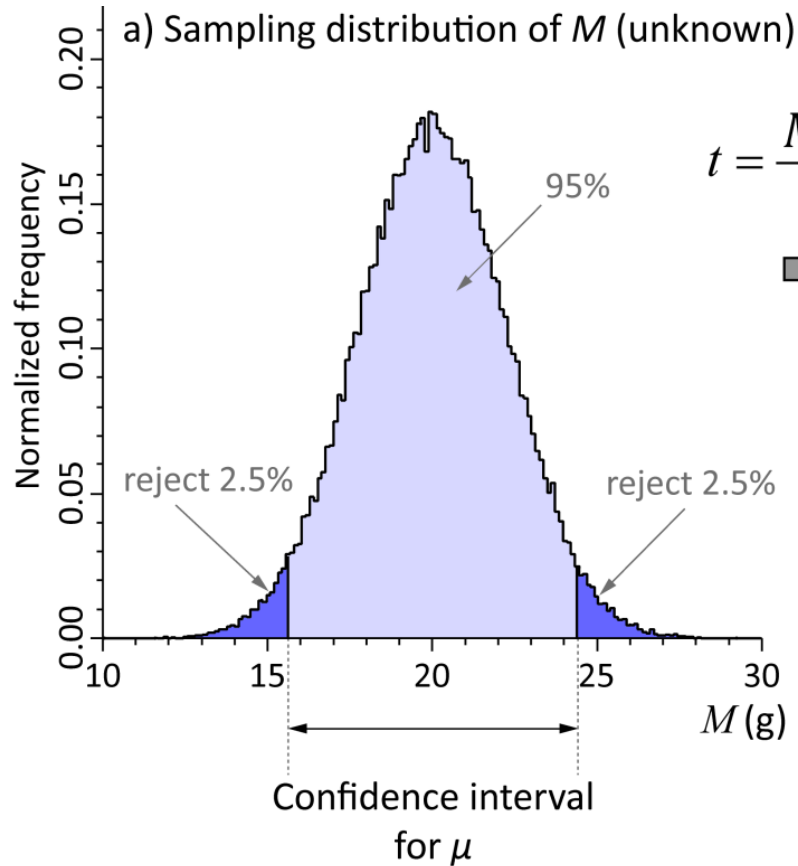
- Mathematical trick – we cannot calculate  $t$
- *Gedankenexperiment*: create a sampling distribution of  $t$




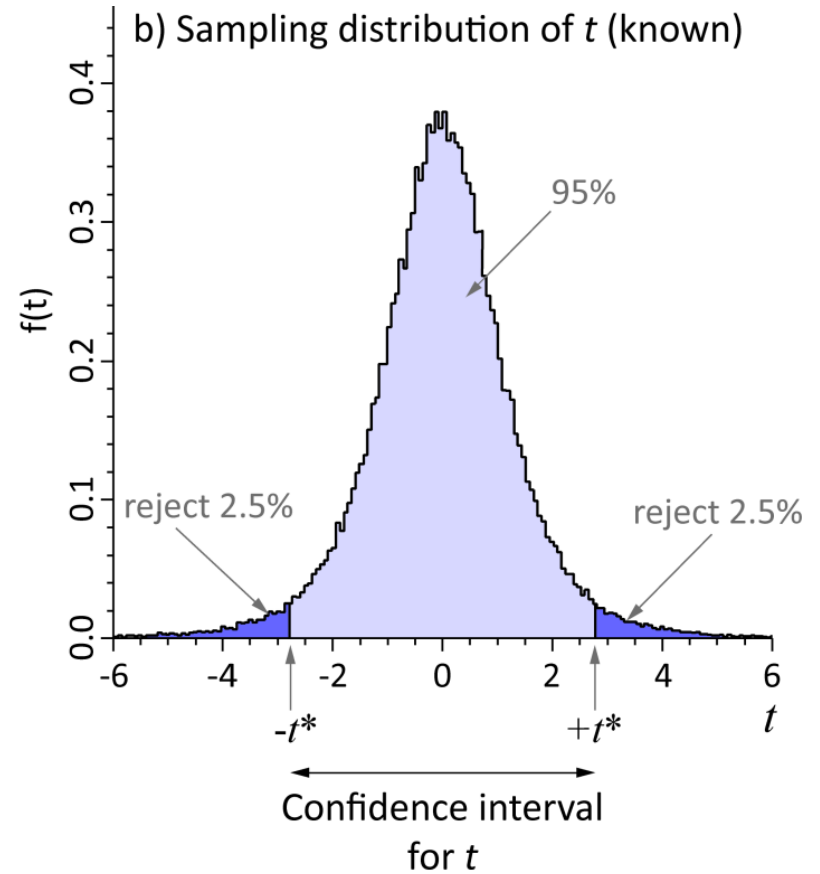
100,000 samples of 5 mice from normal population  
with  $\mu = 20$  g and  $\sigma = 5$  g

Mean body weight calculated for each sample

# Sampling distribution of t-statistic



$$t = \frac{M - \mu}{SE}$$




# Confidence interval of the mean

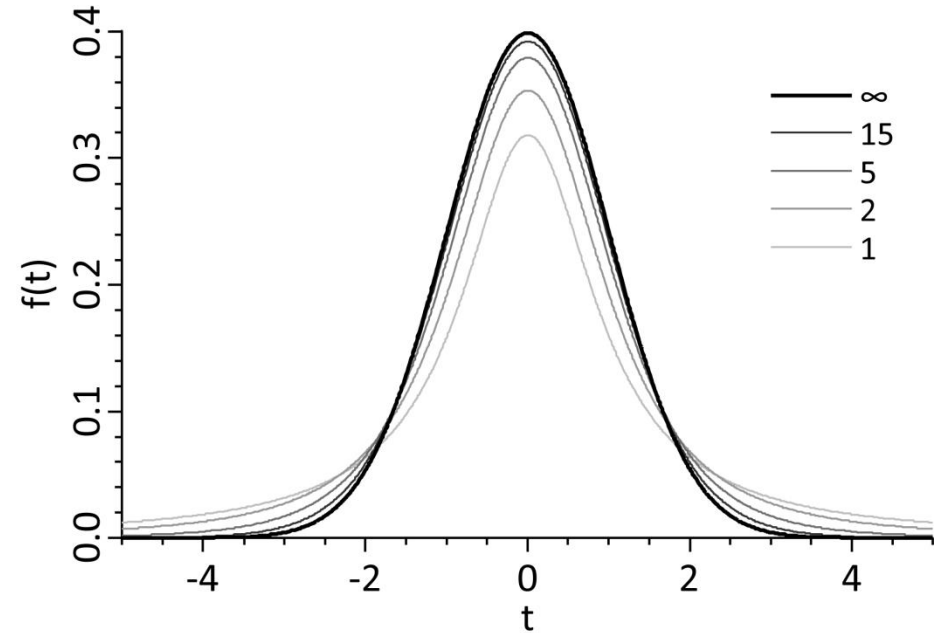
- Statistic

$$t = \frac{M - \mu}{SE}$$

has a *known* sampling distribution:  
Student's t-distribution with  $n - 1$   
degrees of freedom

- We can calculate probabilities!

Student's t-distribution



This is the  $t$ -distribution for 1, 2, 5, 15 and  $\infty$  degrees of freedom. For large number of d.o.f. this turns into a Gaussian distribution.

# William Gosset

---

- Brewer and statistician
- Developed Student's  $t$ -distribution
- Worked for Guinness, who prohibited employees from publishing any papers
- Published as "Student"
- Worked with Fisher and developed the  $t$ -statistic in its current form
- Always worked with experimental data
- Progenitor bioinformatician?



William Sealy Gosset (1876-1937)

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.

If the number of experiments be very large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty:—(1) owing to the “error of random sampling” the mean of our series



# William Gosset's calculator

---



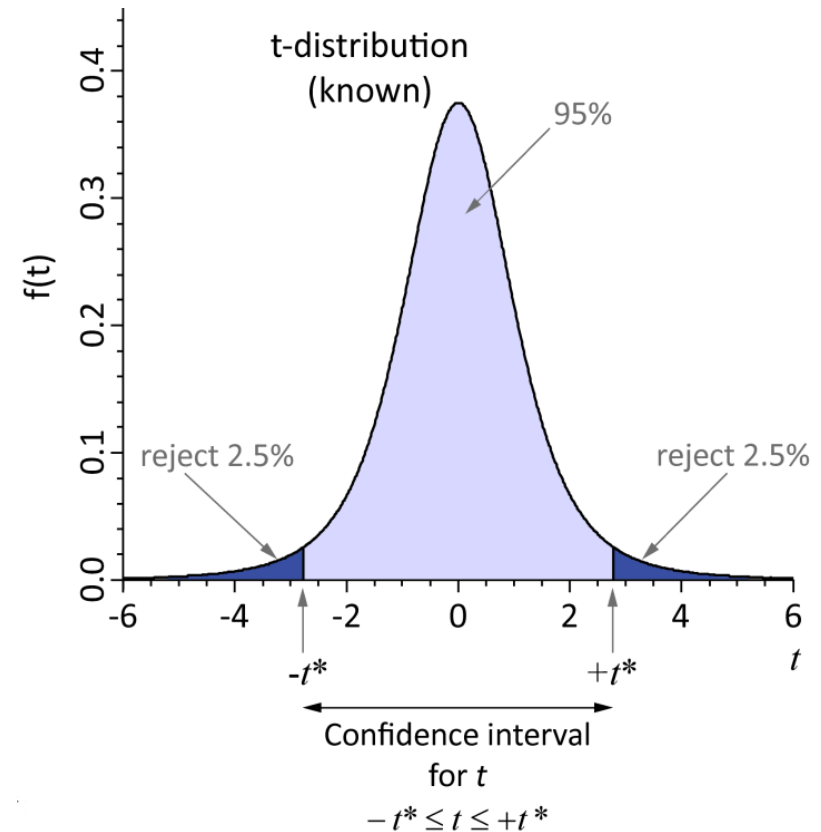
# Confidence interval of the mean

- Statistic

$$t = \frac{M - \mu}{SE}$$

has a *known* sampling distribution:  
Student's *t*-distribution with  $n - 1$   
degrees of freedom

- We can find a critical value of  $t^*$  to cut off required confidence interval
- Use tables of *t*-distribution or any statistical package
- Confidence interval on  $t$  is  $[-t^*, +t^*]$



Tables of *t*-distribution usually give critical values of  $t^*$  for the tail probability.

# Confidence interval of the mean

- We used transformation

$$t = \frac{M - \mu}{SE}$$

- Confidence interval on  $t$  is  $[-t^*, +t^*]$

- Find  $\mu$  from the equation above

$$\mu = M + tSE$$

- From limits on  $t$  we find limits on  $\mu$ :

$$M_L = M - t^*SE$$

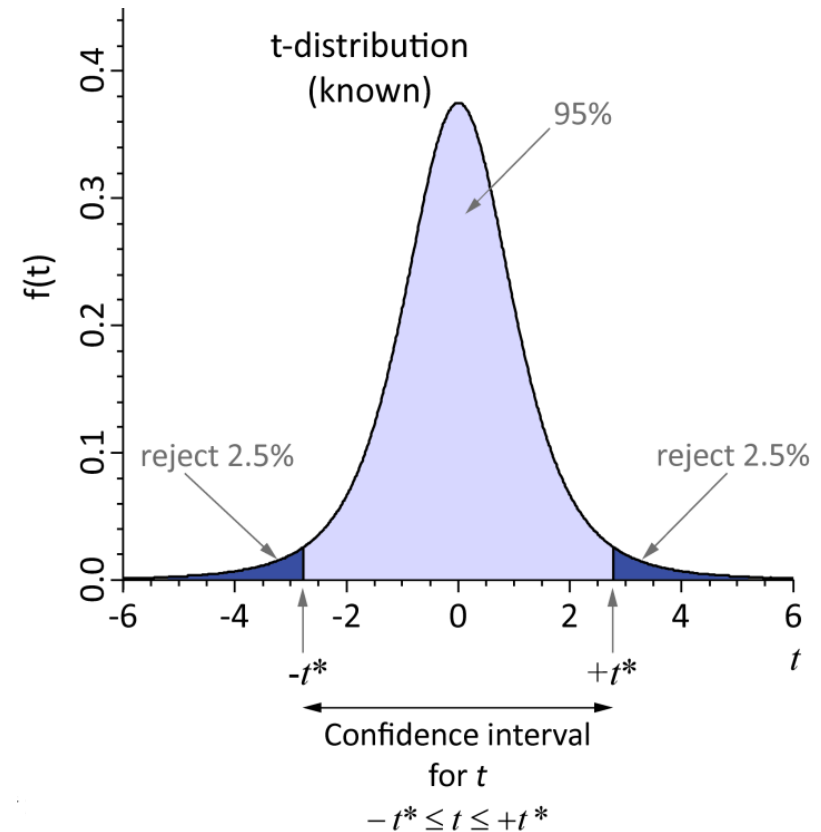
$$M_U = M + t^*SE$$

- Or

$$\mu = M \pm t^*SE$$

- Confidence interval is a scaled standard error

$$CI = t^*SE$$



Tables of t-distribution usually give critical values of  $t^*$  for the tail probability.

# Exercise: 95% confidence interval for the mean

- We have 7 mice with measured body weights 16.8, 21.8, 29.2, 23.3, 19.5, 18.2 and 26.3 g

- Estimators from the sample

$$M = 22.16 \text{ g}$$

$$SD = 4.46 \text{ g}$$

$$SE = 1.69 \text{ g}$$

- Find the 95% confidence interval for the mean

Tail Probabilities							
One Tail		0.10	0.05	0.025	0.01	0.005	
Two Tails		0.20	0.10	0.05	0.02	0.01	
<hr/>							
D	1		3.078	6.314	12.71	31.82	63.66
E	2		1.886	2.920	4.303	6.965	9.925
G	3		1.638	2.353	3.182	4.541	5.841
R	4		1.533	2.132	2.776	3.747	4.604
E	5		1.476	2.015	2.571	3.365	4.032
E	6		1.440	1.943	2.447	3.143	3.707
S	7		1.415	1.895	2.365	2.998	3.499
	8		1.397	1.860	2.306	2.896	3.355
O	9		1.383	1.833	2.262	2.821	3.250
F	10		1.372	1.812	2.228	2.764	3.169
	11		1.363	1.796	2.201	2.718	3.106
F	12		1.356	1.782	2.179	2.681	3.055
R	13		1.350	1.771	2.160	2.650	3.012
E	14		1.345	1.761	2.145	2.624	2.977
E	15		1.341	1.753	2.131	2.602	2.947
D	16		1.337	1.746	2.120	2.583	2.921
O	17		1.333	1.740	2.110	2.567	2.898
M	18		1.330	1.734	2.101	2.552	2.878
	19		1.328	1.729	2.093	2.539	2.861
	20		1.325	1.725	2.086	2.528	2.845
			1.323	1.721	2.080	2.519	2.830
			1.321	1.717	2.075	2.512	2.815
			1.319	1.713	2.070	2.506	2.801
			1.317	1.710	2.065	2.500	2.787
			1.315	1.706	2.060	2.494	2.773
			1.313	1.703	2.055	2.488	2.759
			1.311	1.700	2.050	2.482	2.745
			1.309	1.697	2.045	2.476	2.731
			1.307	1.694	2.040	2.470	2.717
			1.305	1.691	2.035	2.464	2.703
			1.303	1.688	2.030	2.458	2.689
			1.301	1.685	2.025	2.452	2.675
			1.299	1.682	2.020	2.446	2.661
			1.297	1.679	2.015	2.440	2.647
			1.295	1.676	2.010	2.434	2.633
			1.293	1.673	2.005	2.428	2.619
			1.291	1.670	2.000	2.422	2.605
			1.289	1.667	1.995	2.416	2.591
			1.287	1.664	1.990	2.410	2.577
			1.285	1.661	1.984	2.404	2.563
			1.283	1.658	1.978	2.398	2.549
			1.281	1.655	1.972	2.392	2.535
			1.279	1.652	1.966	2.386	2.521
			1.277	1.649	1.960	2.380	2.507
			1.275	1.646	1.954	2.374	2.493
			1.273	1.643	1.948	2.368	2.479
			1.271	1.640	1.942	2.362	2.465
			1.269	1.637	1.936	2.356	2.451
			1.267	1.634	1.930	2.350	2.437
			1.265	1.631	1.924	2.344	2.423
			1.263	1.628	1.918	2.338	2.409
			1.261	1.625	1.912	2.332	2.395
			1.259	1.622	1.906	2.326	2.381
			1.257	1.619	1.900	2.320	2.367
			1.255	1.616	1.894	2.314	2.353
			1.253	1.613	1.888	2.308	2.339
			1.251	1.610	1.882	2.302	2.325
			1.249	1.607	1.876	2.296	2.311
			1.247	1.604	1.870	2.290	2.297
			1.245	1.601	1.864	2.284	2.283
			1.243	1.598	1.858	2.278	2.269
			1.241	1.595	1.852	2.272	2.255
			1.239	1.592	1.846	2.266	2.241
			1.237	1.589	1.840	2.260	2.227
			1.235	1.586	1.834	2.254	2.213
			1.233	1.583	1.828	2.248	2.199
			1.231	1.580	1.822	2.242	2.185
			1.229	1.577	1.816	2.236	2.171
			1.227	1.574	1.810	2.230	2.157
			1.225	1.571	1.804	2.224	2.143
			1.223	1.568	1.798	2.218	2.129
			1.221	1.565	1.792	2.212	2.115
			1.219	1.562	1.786	2.206	2.101
			1.217	1.559	1.780	2.200	2.087
			1.215	1.556	1.774	2.194	2.073
			1.213	1.553	1.768	2.188	2.059
			1.211	1.550	1.762	2.182	2.045
			1.209	1.547	1.756	2.176	2.031
			1.207	1.544	1.750	2.170	2.017
			1.205	1.541	1.744	2.164	2.003
			1.203	1.538	1.738	2.158	1.989
			1.201	1.535	1.732	2.152	1.975
			1.199	1.532	1.726	2.146	1.961
			1.197	1.529	1.720	2.140	1.947
			1.195	1.526	1.714	2.134	1.933
			1.193	1.523	1.708	2.128	1.919
			1.191	1.520	1.702	2.122	1.905
			1.189	1.517	1.696	2.116	1.891
			1.187	1.514	1.690	2.110	1.877
			1.185	1.511	1.684	2.104	1.863
			1.183	1.508	1.678	2.098	1.849
			1.181	1.505	1.672	2.092	1.835
			1.179	1.502	1.666	2.086	1.821
			1.177	1.499	1.660	2.080	1.807
			1.175	1.496	1.654	2.074	1.793
			1.173	1.493	1.648	2.068	1.779
			1.171	1.490	1.642	2.062	1.765
			1.169	1.487	1.636	2.056	1.751
			1.167	1.484	1.630	2.050	1.737
			1.165	1.481	1.624	2.044	1.723
			1.163	1.478	1.618	2.038	1.709
			1.161	1.475	1.612	2.032	1.695
			1.159	1.472	1.606	2.026	1.681
			1.157	1.469	1.600	2.020	1.667
			1.155	1.466	1.594	2.014	1.653
			1.153	1.463	1.588	2.008	1.639
			1.151	1.460	1.582	2.002	1.625
			1.149	1.457	1.576	1.996	1.611
			1.147	1.454	1.570	1.990	1.597
			1.145	1.451	1.564	1.984	1.583
			1.143	1.448	1.558	1.978	1.569
			1.141	1.445	1.552	1.972	1.555
			1.139	1.442	1.546	1.966	1.541
			1.137	1.439	1.540	1.960	1.527
			1.135	1.436	1.534	1.954	1.513
			1.133	1.433	1.528	1.948	1.499
			1.131	1.430	1.522	1.942	1.485
			1.129	1.427	1.516	1.936	1.471
			1.127	1.424	1.510	1.930	1.457
			1.125	1.421	1.504	1.924	1.443
			1.123	1.418	1.498	1.918	1.429
			1.121	1.415	1.492	1.912	1.415
			1.119	1.412	1.486	1.906	1.401
			1.117	1.409	1.480	1.900	1.387
			1.115	1.406	1.474	1.894	1.373
			1.113	1.403	1.468	1.888	1.359
			1.111	1.400	1.462	1.882	1.345
			1.109	1.397	1.456	1.876	1.331
			1.107	1.394	1.450	1.870	1.317
			1.105	1.391	1.444	1.864	1.303
			1.103	1.388	1.438	1.858	1.289
			1.101	1.385	1.432	1.852	1.275
			1.099	1.382	1.426	1.846	1.261
			1.097	1.379	1.420	1.840	1.247
			1.095	1.376	1.414	1.834	1.233
			1.093	1.373	1.408	1.828	1.219
			1.091	1.370	1.402	1.822	1.205
			1.089	1.367	1.396	1.816	1.191
			1.087	1.364	1.390	1.810	1.177
			1.085	1.361	1.384	1.804	1.163
			1.083	1.358	1.378	1.798	1.149
			1.081	1.355	1.372	1.792	1.135
			1.079	1.352	1.366	1.786	1.121
			1.077	1.349	1.360	1.780	1.107
			1.075	1.346	1.354	1.774	1.093
			1.073	1.343	1.348	1.768	1.079
			1.071	1.340	1.342	1.762	1.065
			1.069	1.337	1.336	1.756	1.051
			1.067	1.334	1.330	1.750	1.037
			1.065	1.331	1.324	1.744	1.023
			1.063	1.328	1.318	1.738	1.009
			1.061	1.325	1.312	1.732	0.995
			1.059	1.322	1.306	1.726	0.981
			1.057	1.319	1.300	1.720	0.967
			1.055	1.316	1.294	1.714	0.953
			1.053	1.313	1.288	1.708	0.939
			1.051	1.310	1.282	1.702	0.925
			1.049	1.307	1.276	1.696	0.911
			1.047	1.304	1.270	1.690	0.897
			1.045	1.301	1.264	1.684	0.883
			1.043	1.298	1.258	1.678	0.869
			1.041	1.295	1.252	1.672	0.855
			1.039	1.292	1.246	1.666	0.841
			1.037	1.289	1.240	1.660	0.827
			1.035	1.286	1.234	1.654	0.813
			1.033	1.283	1.228	1.648	0.799
			1.031	1.280	1.222	1.642	0.785
			1.029	1.277	1.216	1.636	0.771
			1.027	1.274	1.210	1.630	0.757
			1.025	1.271	1.204	1.624	0.743
			1.023	1.268	1.198	1.618	0.729
			1.021	1.265	1.192	1.612	0.715
			1.019	1.262	1.186	1.606	0.701
			1.017	1.259	1.180	1.600	0.687
			1.015	1.256	1.174	1.	

# Exercise: 95% confidence interval for the mean

- We have 7 mice with measured body weights 16.8, 21.8, 29.2, 23.3, 19.5, 18.2 and 26.3 g

- Estimators from the sample

$$M = 22.16 \text{ g}$$

$$SD = 4.46 \text{ g}$$

$$SE = 1.69 \text{ g}$$

- Critical value from t-distribution for one-tail probability 0.025 and 6 degrees of freedom

$$t^* = 2.447$$

- Half of the confidence interval is

$$CI = t^* SE = 4.14 \text{ g}$$

- Estimate of the mean with 95% confidence is

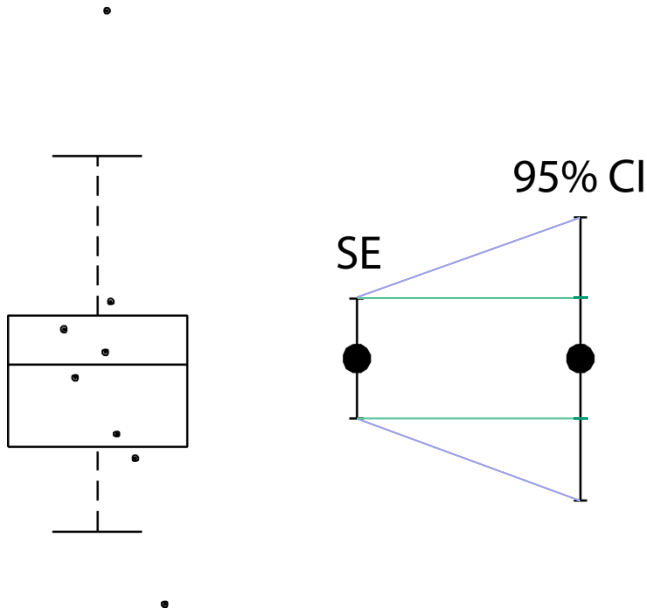
$$\mu = 22 \pm 4 \text{ g}$$

~~$$22.16 \pm 4.14 \text{ g}$$~~

Tail Probabilities						
One Tail	0.10	0.05	0.025	0.01	0.005	
Two Tails	0.20	0.10	0.05	0.02	0.01	
<hr/>						
D	1	3.078	6.314	12.71	31.82	63.66
E	2	1.886	2.920	4.303	6.965	9.925
G	3	1.638	2.353	3.182	4.541	5.841
R	4	1.533	2.132	2.776	3.747	4.604
E	5	1.476	2.015	2.571	3.365	4.032
E	6	1.440	1.943	2.447	3.143	3.707
S	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
O	9	1.383	1.833	2.262	2.821	3.250
F	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
F	12	1.356	1.782	2.179	2.681	3.055
R	13	1.350	1.771	2.160	2.650	3.012
E	14	1.345	1.761	2.145	2.624	2.977
E	15	1.341	1.753	2.131	2.602	2.947
D	16	1.337	1.746	2.120	2.583	2.921
O	17	1.333	1.740	2.110	2.567	2.898
M	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	23	1.323	1.721	2.080	2.520	2.837
	24	1.322	1.719	2.078	2.518	2.835
	30	1.319	1.714	2.071	2.510	2.827
	40	1.315	1.708	2.064	2.503	2.820
	60	1.312	1.704	2.060	2.390	2.815
	65	1.295	1.697	1.997	2.385	2.654
	70	1.294	1.667	1.994	2.381	2.648
	80	1.292	1.664	1.990	2.374	2.639
	100	1.290	1.660	1.984	2.364	2.626
	150	1.287	1.655	1.976	2.351	2.609
	200	1.286	1.653	1.972	2.345	2.601
<hr/>						
Two Tails	0.20	0.10	0.05	0.02	0.01	
One Tail	0.10	0.05	0.025	0.01	0.005	
Tail Probabilities						

# Confidence interval vs standard error

---



How many standard errors  
are in a confidence interval?

What is the confidence of  
the standard error?

# Confidence interval vs standard error

- Confidence interval is a scaled standard error

$$CI = t_{n-1}^*(\gamma) \times SE$$

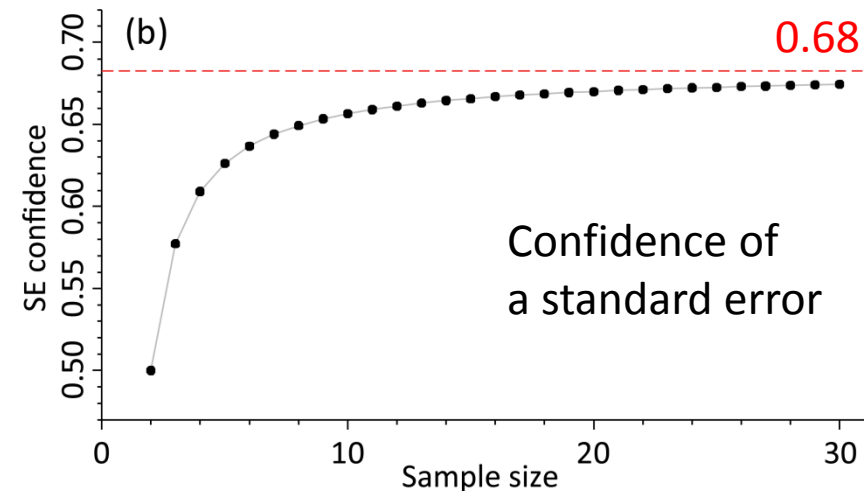
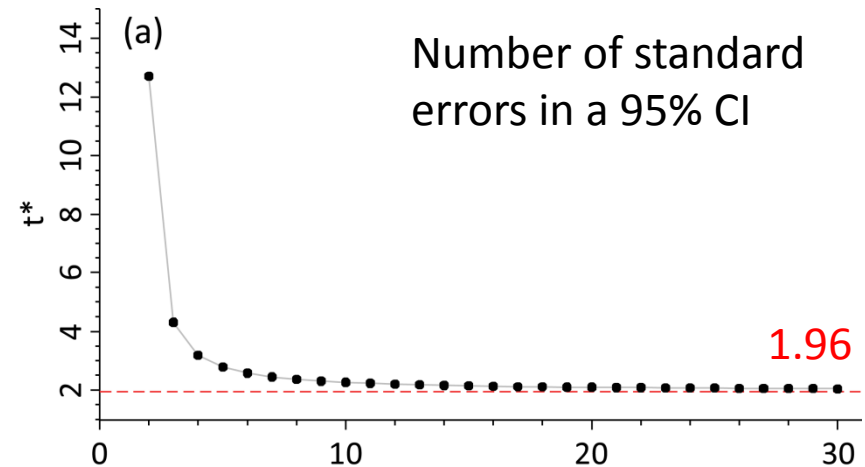
- Scaling factor,  $t^*$ , depends on

- confidence,  $\gamma$
- sample size,  $n$

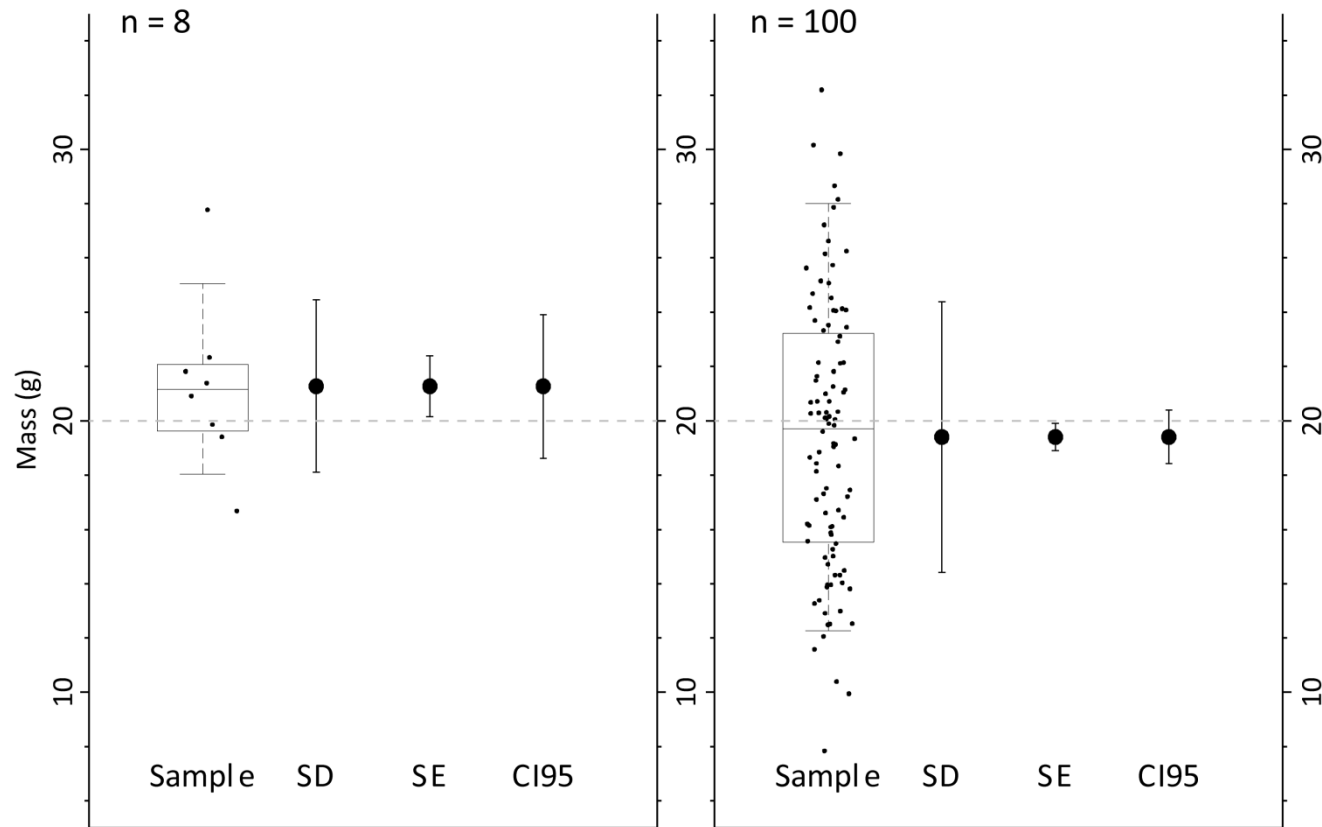
- We can find how  $t^*$  depends on the sample size
- We can find confidence corresponding to  $t^* = 1$

- 95% CI is huge for  $n = 2$
- SE has no simple probabilistic interpretation

- Use confidence intervals!



# SD, SE and 95% CI



- Normal population of  $\mu = 20$  g and  $\sigma = 5$  g
- Sample of  $n = 8$  and  $n = 100$
- Whiskers in the box plot encompass 90% of data – nothing to do with 90% confidence interval



# Exercise: confidence intervals

---

- Experiment where a reporter measures transcriptional activity of a gene
- Day 1: 3 biological replicates, day 2: 5 biological replicates (the same experiment)
- Normalized data:

Day 1	0.89	0.92	0.89		
Day 2	0.55	0.76	0.61	0.83	0.75

- 95% confidence intervals for the mean:
  - Day 1: [0.86, 0.94]
  - Day 2: [0.56, 0.84]
- What can you say about these results? What else can you do with these data?

# Exercise: confidence intervals

---

- Experiment where a reporter measures transcriptional activity of a gene
- Day 1: 3 biological replicates, day 2: 5 biological replicates (the same experiment)
- Normalized data:

Day 1	0.89	0.92	0.89		
Day 2	0.55	0.76	0.61	0.83	0.75

- 95% confidence intervals for the mean:
  - Day 1: [0.86, 0.94]
  - Day 2: [0.56, 0.84]
- What can you say about these results? What else can you do with these data?
- You could pool data together to get [0.66, 0.89]
- But  $t$ -test between two days gives  $p = 0.03$
- Perhaps something had changed between day 1 and day 2?

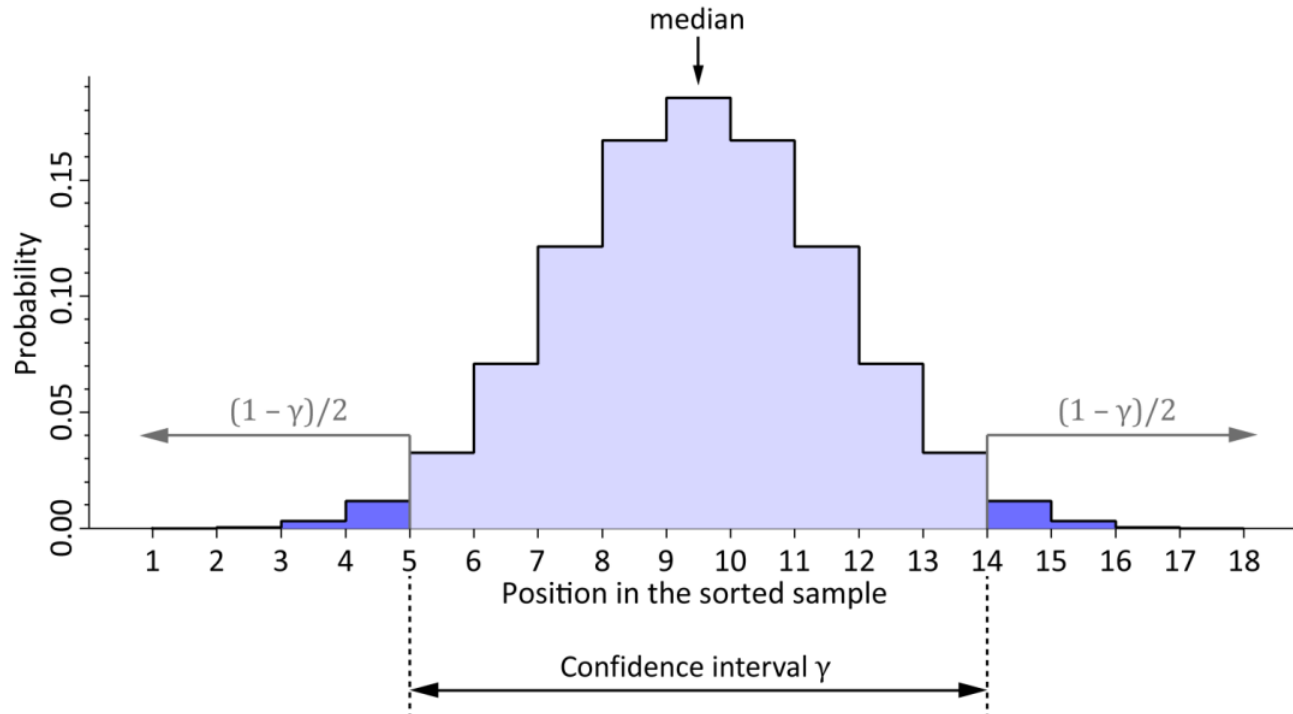
# Confidence interval of the median

- Median is quite often used in biology
- Do you know how to find its uncertainty?
- Consider sample  $x_1, x_2, \dots, x_n$
- Let (unknown) population median be  $\Theta$
- $P(x_i < \Theta) = \frac{1}{2}$  and  $P(x_i > \Theta) = \frac{1}{2}$
- It is like tossing a coin

- Sort the sample and find the probability of  $\Theta$  between  $k$  and  $n - k$  points:

$$P(x_{(1)} \leq \dots \leq x_{(k)} \leq \Theta \leq x_{(k+1)} \dots \leq x_{(n)})$$

- Binomial distribution
- You can find precise confidence limits by interpolating in this discrete distribution



# Confidence interval of the median: approximation

- Sample  $x_1, x_2, \dots, x_n$
- Sorted sample  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$
- Find two limiting indices:

$$L = \left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \sqrt{\frac{n}{4}} \right\rfloor$$

$\lfloor x \rfloor$  - floor,  $\lfloor 3.4 \rfloor = 3$

$\lceil x \rceil$  - ceiling,  $\lceil 3.4 \rceil = 4$

$$U = n - L$$

- Standard error of the median

$$\widetilde{SE} = \frac{x_{(U)} - x_{(L+1)}}{2}$$

- Confidence intervals

$$\widetilde{M}_L = \widetilde{M} - t^* \widetilde{SE}$$

$$\widetilde{M}_U = \widetilde{M} + t^* \widetilde{SE}$$

- Here,  $t^*$  is the critical value from t-distribution with  $U - L - 1$  degrees of freedom

## Example

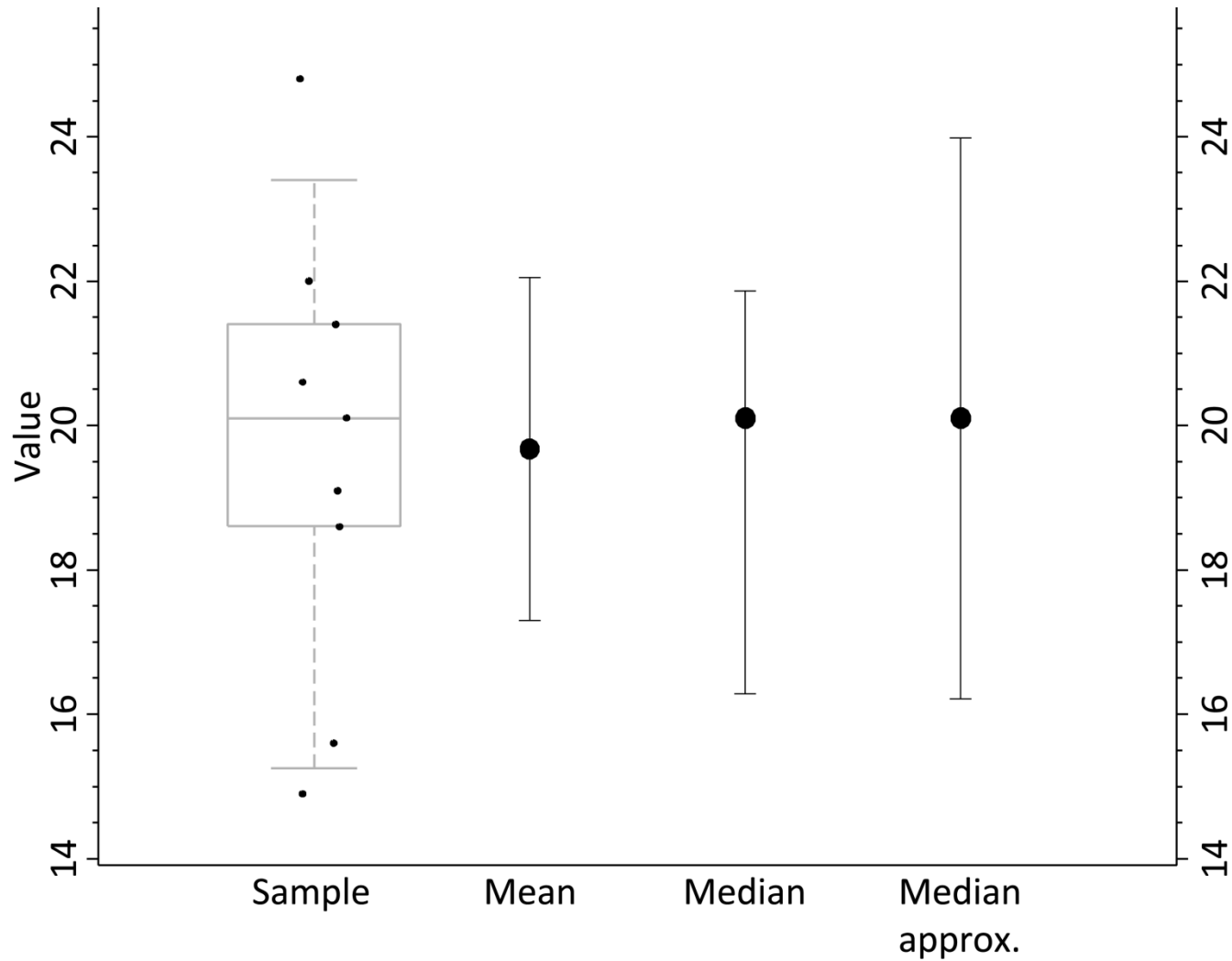
- Weighed 9 mice: 14.9, 22.0, 15.6, 19.1, 21.4, 20.6, 20.1, 24.8, 18.6 g

- Sorted sample:

$i$	1	2	3	4	5	6	7	8	9
$x_{(i)}$	14.9	15.6	18.6	19.1	20.1	20.6	21.4	22.0	24.8

- Median is  $\widetilde{M} = 20.1$  g
- $L = \lfloor 4.5 \rfloor - \lceil 1.5 \rceil = 4 - 2 = 2$
- $U = 9 - 2 = 7$
- $\widetilde{SE} = \frac{21.4 - 18.6}{2} = 1.4$  g
- $t^* = 2.776$  for 4 d.o.f.
- 95% CI is  $[16.2, 24.0]$  g

# Confidence interval of the median: example





Hand-outs available at <http://is.gd/statlec>

Please leave your feedback forms on the table by the door

