Statistics – Student´s t-Distribution

Student´s t vs. Normal Distribution? What to use when?

Student´s t-Distribution

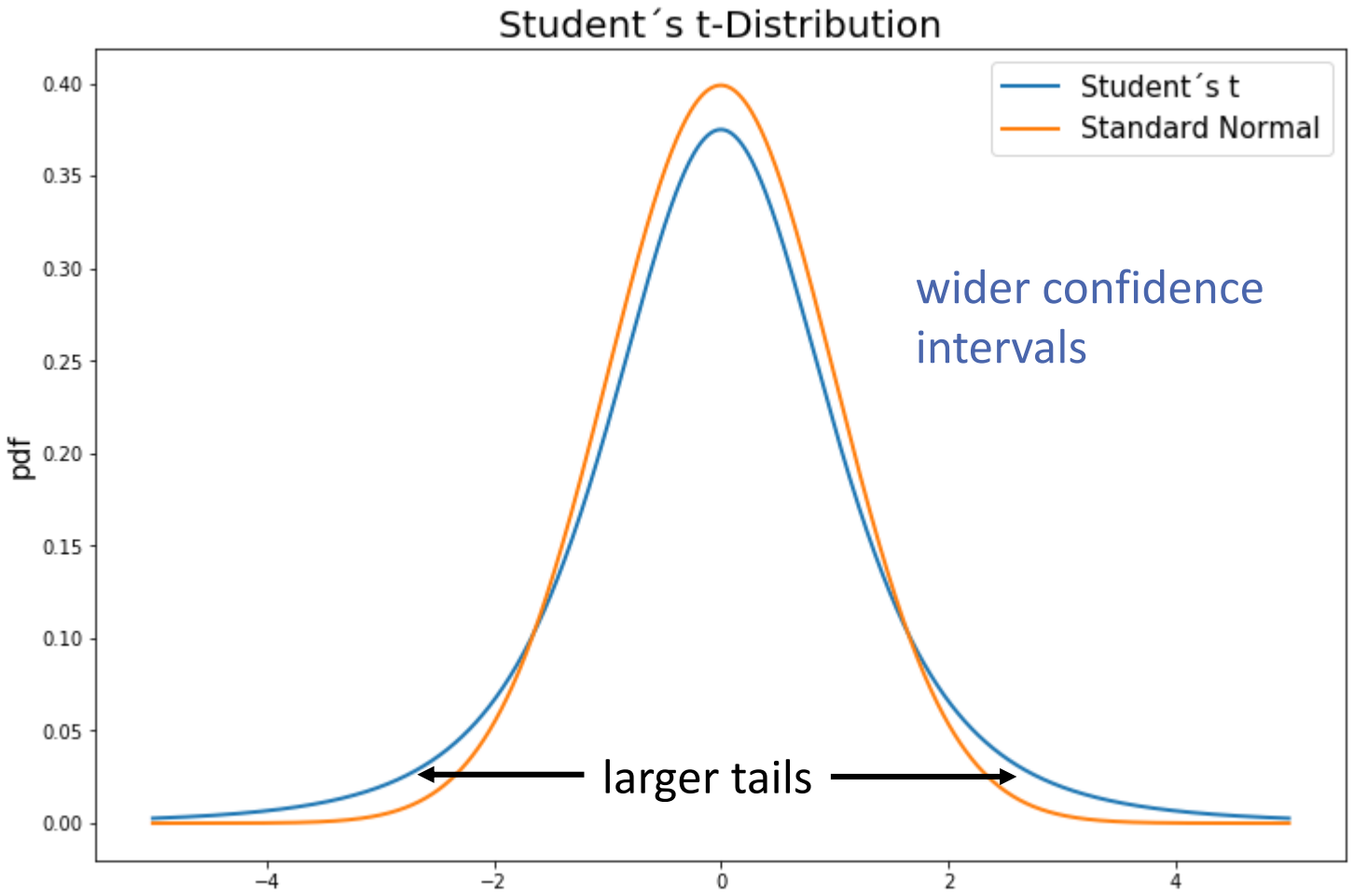
Confidence Intervals

To create a Confidence Interval Estimate for the population mean, we require:

- Point Estimate for the mean → sample mean
- Standard Error → ~~population std~~ / $\sqrt{n}$
- Point Estimate for the std → sample std

→ estimates are less reliable/precise (in particular for small sample sizes)
→ Normal Distribution not appropriate, we need a Distribution with more probabilities in the tails (more conservative)

Use Student´s t-Distribution



properties and t-tables

Properties

- symmetrical around the mean
- completely defined and described by the “degrees of freedom df” (sample size – 1).
- more probabilities in the tails (“fatter tails”)
- for large sample sizes, it approaches the standard normal distribution
- t-scores can be converted to probabilities and vice versa (depending on df) → t-table / scipy.stats

Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).

df/p	0.40	0.25						0.0005
1	0.324920							636.6192
2	0.288675							
3								
4								
22								
23								
24	0.256175							
25	0.256060							3.7251
26	0.255955	0.684043						3.7066
27	0.255858	0.683685	1.313703			2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%