The pdf of the central *t*-distribution on *df* degrees of freedom

$$f_t(t|df) = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df\pi}\Gamma\left(\frac{df}{2}\right)}\left(1+\frac{1}{df}t^2\right)^{-\frac{df+1}{2}} .$$

Typical (i.e., historical) use: we have $\hat{\mu}$ (often it is a simple mean) with estimated standard error, $\hat{\text{se}}$. For inference with

$$t = \frac{\hat{\mu}}{\hat{\text{se}}}$$

and compute a *P*-value (say 2-sided) as

$$P = 2\int_{|t|}^{\infty} f_t(x|df)dx.$$

A simpler, likelihood inferential proceedure exists based on the pivitol quantity

$$t_\mu = \frac{\hat{\mu}-\mu}{\hat{\text{se}}}$$

taken as having a central *t*-distribution. In the likelihood $\hat{\mu}$ and $\hat{\text{se}}$ are considered fixed. That likelihood in μ, standardized to a maximum value of 1 is

$$L(\mu) = \left(1+\frac{1}{df}\left(\frac{\hat{\mu}-\mu}{\hat{\text{se}}}\right)^2\right)^{-\frac{df+1}{2}}.$$

The best supported estimate is $\hat{\mu}$, and evidence about other values of μ is then relative to this MLE. Use $L(\mu)$ to compute evidence as e:1 odds. I prefer to have "e" ≥ 1 hence flip $L(\mu)$ over and use as odds ***against*** another μ in favor of $\hat{\mu}$

$$\left(1+\frac{1}{df}\left(\frac{\hat{\mu}-\mu}{\hat{\text{se}}}\right)^2\right)^{\frac{df+1}{2}} :1 .$$

For example, if df $= 7$ and we consider $\mu = 0$, for which case say we have $t = 2.36462 = \hat{\mu}/\hat{\text{se}}$. Then odds against $\mu = 0$, relative to $\hat{\mu}$ are computed as

$(1+\frac{1}{7}(2.36462)^2)^4$:1

10.47 :1

Note 1. Take this as one-sided because we pay attention to the sign of $\hat{\mu}$.
Note 2. For this case, 2-sided $P = 0.05$.

For the “normal distribution” case odds are $\exp(z^2/2)$:1 against $\mu = 0$.

Strength of inference is expressed as odds, not as probability; easy to compute, and likelihood-based.

What is wrong with this - if anything? Why don't we (statistics) use it for inference **given the data?** Probability theory as already developed is still to be used in planning stages, such as sample size assessment. However, this way our thinking and methods are different before we have the data vs. after we have the data. There exists more thoughts and literature about this, but I am keeping it short here.

Well, not that short.

One can do fudicial inference here (discredited though it is) - it also can be considered Bayesian arising from using an improper prior $d\mu$ (at least I think so):

$$f_t(\mu|df) = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df\pi}\Gamma\left(\frac{df}{2}\right)\hat{\text{se}}}\left(1+\frac{1}{df}\left(\frac{\hat{\mu}-\mu}{\hat{\text{se}}}\right)^2\right)^{-\frac{df+1}{2}} .$$

For example, let $\hat{\mu} = 4$ and $\hat{\text{se}} = 1.6916037,$ hence $t = 2.36462$. Numerical integration (using SAS) verifies this pdf integrates to 1 for this example (and others). From said numerical results with this pdf I find $\Pr\{0 \leq \mu \leq 8\} = 0.95$, an equal tail and shortest 95% interval.

From the central *t*-distribution, $t_{7,0.975} = 2.36462$, hence the usual frequentist 95% confidence interval is $\hat{\mu} \pm \hat{\text{se}} * t_{7,0.975}$ which is 0 to 8. Same interval, different construction and math-probability interpretations. In terms of evidence (i.e. relative likelihood) the interval $0 \leq \mu \leq 8$ **excludes** all values of μ for which the evidence against them, relative to $\hat{\mu} = 4$, is greater than 10.47 to 1. This is an inferential statement about a parameter, given the data, hence not at all like a *P*-value. Also, it does not seem to have multiple definitions-derivations like probability-based interval inference. Likelihood provides an alternative to *P*-values and associate intervals and this possible usage has been known about for going on 100 years.