$$\mathcal{L}(\mu,\sigma|x) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

So our null hypothesis is $\mu = 0$, our alternative hypothesis is $\mu \neq 0$. We can then write the likelihood ratio test as

$$\begin{aligned}
\log\lambda &= -2\log\left(\frac{\sup\mathcal{L}(\mu_0,\sigma|x)}{\sup\mathcal{L}(\mu,\sigma|x)}\right) \\
&= -2\log\left(\sup\mathcal{L}(\mu_0,\sigma|x)\right) + 2\log\left(\sup\mathcal{L}(\mu,\sigma|x)\right)
\end{aligned}$$

We can write

$$\log\left(\sup\mathcal{L}(\mu_0,\sigma|x)\right) = \sup\log(\mathcal{L}(\mu_0,\sigma|x)) = \sup\left(-\frac{n}{2}\log\left(2\pi\sigma^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu_0)^2}{2\sigma^2}\right)$$

Now maximizing over σ we have

$$\frac{d}{d\sigma}\left(-\frac{n}{2}\log\left(2\pi\sigma^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu_0)^2}{2\sigma^2}\right) = 0$$

we get

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2$$

and so

$$\log\left(\sup\mathcal{L}(\mu_0,\sigma|x)\right) = \sup\log(\mathcal{L}(\mu_0,\sigma|x)) = \left(-\frac{n}{2}\log\left(2\pi\hat{\sigma}_0^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu_0)^2}{2\hat{\sigma_0}^2}\right) = -\frac{n}{2}\left(\log\left(2\pi\hat{\sigma}_0^2\right)+1\right)$$

Now in the case where we can float μ we have

$$\log\left(\sup\mathcal{L}(\mu,\sigma|x)\right) = \left(-\frac{n}{2}\log\left(2\pi\hat{\sigma}^2\right) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\hat{\sigma}^2}\right) = -\frac{n}{2}\left(\log\left(2\pi\hat{\sigma}^2\right)+1\right)$$

with

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i\text{and}\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

So our likelihood ratio is given by

$$\begin{aligned}
\lambda &= -\frac{n}{2}\left(\log\left(2\pi\hat{\sigma}_0^2\right)+1\right) + \frac{n}{2}\left(\log\left(2\pi\hat{\sigma}_0^2\right)+1\right) \\
&= \frac{n}{2}\left(\log\left(2\pi\hat{\sigma}^2\right) - \log\left(2\pi\hat{\sigma}_0^2\right)\right) \\
&= \frac{n}{2}\log\left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right) \\
&= \frac{n}{2}\log\left(\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x}_i)^2}\right)
\end{aligned}$$

Further simplifying this noting $\sum_{i=1}^{n}\left(x_i-\mu_0\right)^2=\sum_{i=1}^{n}\left(x_i-\bar{x}+\bar{x}-\mu_0\right)^2=\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2+n\left(\bar{x}-\mu_0\right)^2$ we can write above as

$$\lambda=\frac{n}{2}\log\left(\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2+n\left(\bar{x}-\mu_0\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\right)=\frac{n}{2}\log\left(1+\frac{n\left(\bar{x}-\mu_0\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}_i\right)^2}\right)=\frac{n}{2}\log\left(1+\frac{\left(\bar{x}-\mu_0\right)^2}{\hat{\sigma}^2}\right)$$

The right formula is the same as the t-test and so will follow the t-distribution.

In the case were we had sampled from a different gaussian we see that on average we have

$$\lambda=\frac{n}{2}\log\left(1+\frac{\left(\mu-\mu_0\right)^2}{\hat{\sigma}}\right)$$

This deviation will grow linearly with the number of samples. Now finally, how do we quantify this into human terms. We need one more important thing: Wilk's theorem.

Wilk's theorem states that if a hypthoesis h_0 with p_0 unkonwn(floated) parameters and an alternative hypthoesis h with p unknown(floated) parameters. Then under some general condiations and when the null hypothesis is true the likelihood ratio $-2\log\lambda$ converges to a χ^2distribution with $\nu=p-p_0$ degrees of freedom in the limit of $n\to\infty$, or in other words

$$-2\log\lambda\sim\chi^2_\nu$$

Now in the scenario above we had one free parameter with the null hypothesis σ and two free parameters with the alternative hypothesis $\mu$and σ consequently $\nu=1$ or in this case

$$-2\log\lambda=\frac{n}{2}\log\left(1+\frac{\left(\mu-\mu_0\right)^2}{\hat{\sigma}}\right)\sim\chi^2_1$$

Now recall that a $\chi^2_1=\frac{(E-O)^2}{\sigma^2}$ consequently we have

$$-2\log\lambda=\frac{n}{2}\log\left(1+\frac{\left(\mu-\mu_0\right)^2}{\hat{\sigma}}\right)\sim\frac{E-O}{\sigma^2_\mu}$$

which means that σ_μ occurs when

$$\sigma_\mu\to 2\log\lambda(\mu+\sigma_\mu)-2\log\lambda(\mu)=1$$

This is how we get uncertainties on parameters, its all Wilks' theorem.

In this case, we can taylor expand $\log\lambda$ about $\mu=\mu_0$(this is known as a wald expansion) to get an analytic form. This gives us

$$2\log\lambda\approx n\left(\frac{\left(\mu-\mu_0\right)^2}{\hat{\sigma}^2}\right)\sim\frac{E-O}{\sigma^2_\mu}$$

or in other words we get the well known uncertainty on the mean scales as $\frac{1}{\sqrt{n}}$ ie $\sigma_\mu=\frac{\sigma}{\sqrt{n}}$