Tues afternoon, 23rd Sept 2025

Tutorial 3: Gaussian, king of the distributions 👑

---

1. Conguate priors: Gaussian

Problem curtesy of Jakob Kollmüller.

Let’s implement the temperature measurement as discussed in the lecture. The prior distribution over temperatures is the Gaussian $P(x) = \mathcal{N}(x \,|\, x_0 = 295, T = 3^2)$ and the measurement noise is given by $P(n) = \mathcal{N}(x \,|\, n_0 = 0, N = 1)$.

1a) Draw a true temperature from the prior distribution $x_{\text{true}} \sim P(x)$, a noise realization $n \sim P(n)$, and generate data by adding the true temperature and measurement noise: $y = x_{\text{true}} + n$.

1b) Plot the Prior distribution and indicate the data and ground truth as vertical lines.

1c) Calculate posterior mean and variance and add the posterior distribution to the previous plot.

1d) Repeat this procedure for various noise levels (but constant true temperature) and discuss your findings.

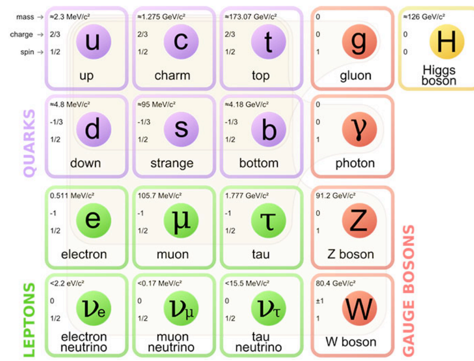
1e) 1e) For a sample of 100 (x,y) values, plot the posterior mean vs. the true value for the different noise levels you considered in (d).

1f) (Bonus) How does the mean squared error compare for the meas vs. the posterior mean? What's the relative improvement for the noise levels?

2. Combination of statistical results

The statistical combination of results… it's a complicated business, with high profile physics results often taking many PhD-student years to complete.

However, armed with your new stats toolkit for seeing Gaussians as conjugate priors, we can do these calculations with the formula we derived in lecture today!



Let's consider a concrete example: the **top quark** is the heavist known fundamental particle: it's as heavy as a gold atom, and yet compressed in a space so tiny it's "point-like" based on the highest resolution "microscope" we've been able to study it with (a.k.a, the Large Hadron Collider (LHC)).

Due to its eminent status as the heaviest particle (and correspondingly large interactions with the Higgs boson), the top mass has important implications for the future stability of the universe.

Both of the major LHC experiments, ATLAS and CMS have measured this quantity, with the results (and papers describing the analyses) given below:

ATLAS: $m_t = 172.71 \pm 0.48 \text{ GeV}$
https://www.arxiv.org/abs/1810.01772

CMS: $m_t = 172.52 \pm 0.42 \text{ GeV}$
https://www.arxiv.org/abs/1509.04044

Your task: Interpret m_t as a random variable and use Gaussian likelihood and prior to calculate the posterior probability of m_t given these two measurements.

Since CMS analysis came first, let it be the prior, and then the ATLAS measurement can be the likelihood of the new observation.

Oh wait! ATLAS and CMS actually *have* already published this combination! Check out their paper and see how your result compares: https://arxiv.org/abs/2402.08713

So these conjugate priors / posteriors formalism is super useful to see if a combination is "worth it" before embarking on such a problem 😉
Or … for sanity checking results your results once you have them 🌈