Sampling

Sampling is a stochastic approach to approximating the evidence.

The main thing we want to do is replace the integral with a summation:

$$\hat{f} = \frac{1}{L} \sum_{l=1}^{L} f(z^{(l)})$$

where the points $z^{(l)}$ where we evaluate the function are drawn from the true distribution $p(z)$.

This is useful since the integral is what will make the evidence intractable most of the time. The summation is useable because $\mathbb{E}[\hat{f}] = \mathbb{E}[f]$

> note that the variance of the estimator shrinks with the number of samples

This relies on us having a uniform distribution. The change-of-variable method helps us normalise a distribution (such as a Gaussian or Beta) to a uniform one.

To normalise, we want to formulate the cumulative distribution function:

$$z = f^{-1}(y) = \int_{-\infty}^{y} p(y) dy$$

This tells us the probability that our distribution will take a value less than or equal to y.

If we inverse this function, we can give it a probability and get back the value y which will give you said probability.

We can now do 'change of variables', where we draw samples from a (i.e uniform) distribution, by drawing samples from **another** (e.g. Gaussian, Beta) distribution and transforming these to be samples from the desired (i.e. uniform) distribution.

If we can formulate the cdf for the distribution, we are done!

If we can't, we need to sample through other methods. There are a bunch of different methods, but Karl has said in the summary that you don't need to know them.

To sum up:

- Sampling is not dependent on dimension
- It will often converge to the correct answer with infinite time
  - but we have no guarantees until we get there

- It is quite time consuming