Laplace Approximation

Essentially what we want to do is find the mode of the posterior (i.e. through a MAP estimate) and then fit a Gaussian to it (by identifying elements in the resulting expression as Gaussian.)

We can't compute our posterior $p(z)$, but we can compute $f(z)$ if we know our likelihood and prior.

$$p(z) = \frac{1}{Z} f(z) = \frac{f(z)}{\int f(z) dz}$$

Method:

1. Find the the mode of $f(z)$ by using MAP
   - we're finding the stationary points of our posterior
   - $f(z)$ has the same modes as $p(z)$
2. Make a taylor expansion around the mode
   - this will give us a funky equation
   - we can cancel out a lot of terms though because we know the derivative of our mode will be 0 (as we are at a maxima), and the second derivative will be negative
3. Take the exponential to get our Gaussian function
   - $f(z) \approx f(z_0) e^{-\frac{1}{2} A (z - z_0)^2}$
   - the A is our Gaussian's precision
4. We want to find an approximation to $p(z)$ - we need to normalize our function to a distribution
   - $p(z) = \frac{1}{Z} f(z) \approx q(z)$
5. Assume that $q(z)$ is a Gaussian, i.e. $f(z_0) = p(mean)$
   - $q(z) = (\frac{A}{2} \pi)^{\frac{1}{2}} e^{-\frac{A}{2}(z - z_0)^2}$
   - not really sure what this means

We can then use $q(z)$ in place of $p(w|t)$ when computing predictions (replace the true posterior with our 'fake' one).

This can be bad if the posterior is far from a Gaussian - we could over/under fit.