

## Mixtures of normal distributions

[Nematrian website page: [MixturesOfNormalDistributions](#), © Nematrian 2023]

1. Suppose we have two Normal distributions,  $N_1 \cong N(0, \sigma_1^2)$  and  $N_2 \cong N(0, \sigma_2^2)$ , with the same mean (which without loss of generality is set to zero) but with different standard deviations. Suppose that we form a *mixture*,  $M$ , of these two different distributions, i.e. a distribution where the probability of drawing from the first distribution is  $p_1$  and from the second distribution is  $p_2$  where  $p_1 + p_2 = 1$ . The probability density functions of the two distributions in isolation and of the mixture are:

$$N_1: \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-x^2/(2\sigma_1^2)}$$

$$N_2: \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-x^2/(2\sigma_2^2)}$$

$$M: \frac{p_1}{\sqrt{2\pi\sigma_1^2}} e^{-x^2/(2\sigma_1^2)} + \frac{p_2}{\sqrt{2\pi\sigma_2^2}} e^{-x^2/(2\sigma_2^2)}$$

2. The (excess) kurtosis of the mixture is as follows:

$$\frac{3(p_1\sigma_1^4 + p_2\sigma_2^4)}{(p_1\sigma_1^2 + p_2\sigma_2^2)^2} - 3$$

3. If  $p_1 = 0$  or  $p_2 = 0$  or  $\sigma_1^2 = \sigma_2^2$  then the mixture simplifies to a single normal distribution and thus has (excess) kurtosis of zero. However in other situations the excess kurtosis is greater than zero. This corresponds to the observation that the weighted arithmetic average of different positive numbers is greater than their geometric average, as long as more than one is given a positive weight.

4. More generally, if we have a mixture of several different normal distributions (e.g. we postulate that we are in the presence of time-varying volatility and we are measuring time averaged statistics), and if each of these normal distributions has the same underlying mean, then we can expect our time averaged probability distribution to appear to exhibit excess kurtosis.

5. If the distributions have different means then the magnitude of the excess kurtosis also depends on the means. In the above example, if  $N_1 \cong N(\mu_1, \sigma_1^2)$  and  $N_2 \cong N(\mu_2, \sigma_2^2)$  then the probability density of the mixture becomes:

$$M: \frac{p_1}{\sqrt{2\pi\sigma_1^2}} e^{-(x-\mu_1)^2/(2\sigma_1^2)} + \frac{(1-p_1)}{\sqrt{2\pi\sigma_2^2}} e^{-(x-\mu_2)^2/(2\sigma_2^2)}$$

and its (excess) kurtosis becomes:

$$\frac{6p_1(1-p_1)(A_0 + A_1p_1 + A_2p_1^2)}{B}$$

where

$$A_0 = \frac{(\mu_1 - \mu_2)^4}{6} + (\mu_1 - \mu_2)^2(\sigma_1^2 - \sigma_2^2) + \frac{(\sigma_1 - \sigma_2)^2(\sigma_1 + \sigma_2)^2}{2}$$

$$A_1 = -(\mu_1 - \mu_2)^2((\mu_1 - \mu_2)^2 + 2(\sigma_1 - \sigma_2)(\sigma_1 + \sigma_2))$$

$$A_2 = (\mu_1 - \mu_2)^4$$

$$B = ((\mu_1 - \mu_2)^2 p_1^2 - ((\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)(\sigma_1 + \sigma_2)) p_1 - \sigma_2^2)^2$$

6. If the means are different enough (relative to differences in the standard deviations) then the excess kurtosis can be negative, e.g. the excess kurtosis of an equal mixture (i.e.  $p_1 = p_2 = 0.5$ ) of  $N(2, 1^2)$  and  $N(-2, 1^2)$  has an excess kurtosis of  $-32/25$ . This is consistent with the observation that if the standard deviations of the every component element of the mixture are arbitrarily small relative to the spread of the means then any probability density function can be approximated arbitrarily accurately by giving suitable weights to different elements of the mixture, as long as there are enough suitably spaced contributors to the mixture. The weight to give to the element with mean  $\mu$  in this formulation is the probability density at  $x = \mu$  of the desired overall probability distribution,  $p(x)$ .