Some questions on the basic rules of probability and probability distributions.

1. *Basic rules of probability.* Draw two sets A and B using Venn diagrams. Determine the meaning of conditional and joint probability in terms of ratios of areas of specific regions in the Venn diagram. Prove the product rule of probability using the Venn diagram. What does the statement that A and B are independent imply in terms of the Venn diagram?

2. *Maximum likelihood estimation*. Given observations $x_1,\dots,x_N$ drawn from a Gaussian distribution $\mathcal{N}\left(\mu,\sigma^2\right)$, show that the maximum likelihood estimates of μ and σ^2 are given

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i-\hat{\mu}\right)^2.$$

3. *Bayes theorem*. A Gaussian distribution can be parametrized as

$$p\left(x|\mu,\beta^2\right) = \sqrt{\frac{\beta^2}{2\pi}}\exp\left[\beta^2\frac{(x-\mu)}{2}\right]$$

where β^2 is the *precision* or inverse variance. Suppose μ to be known and assume a Gamma prior for β^2 with parameters θ and k,

$$p\left(\beta^2|k,\theta\right) = \frac{1}{Z}\beta^{2(k-1)}\exp\left(-\beta^2/\theta\right).$$

Show that, given observations $x_1,\dots,x_N$, the posterior distribution over the precision β^2 is again a Gamma distribution with parameters

$$k' = k+\frac{N}{2}$$

$$\theta' = \frac{\theta}{1+1/2\theta\sum_{i=1}^{N}\left(x_i-\mu\right)^2}.$$

The mean of a Gamma distribution is given by $k\theta$. Show that the mean of the posterior distribution over β^2 tends, as N grows, to

$$k'\theta' \to \frac{N}{\sum_{i=1}^{N}\left(x_i-\mu\right)^2}$$

which is the inverse of the maximum likelihood estimate of the variance.