1. Recall that (joint) conjugate prior for $\theta = (\mu, \sigma^2)$ in the normal example involves defining the prior on μ conditional on σ^2. However, if we instead make the priors on μ and σ^2 independent then we can still make use of conjugacy within a Gibbs sampler. Let $\pi(\mu) \sim N(\mu_0, \sigma_0^2)$ and $\pi(\sigma^2) \propto 1/\sigma^2$. Assume that the data Y_i are iid $N(\mu, \sigma^2)$.

   (a) Prove that $\mu | Y_1, \ldots, Y_n, \sigma^2$ is normally distributed and find the appropriate parameter updates.

   (b) Prove that $\sigma^2 | Y_1, \ldots, Y_n, \mu$ is inverse gamma and find the appropriate parameter updates.

   (c) Consider the *ChickWeight* dataset within the *datasets* package in *R*. This dataset consists of 4 groups of chicks on different diets. Take each chick and create a new variable denoting their overall change in weight (include only those chicks that were completely observed). For each group, estimate the mean and variance of these differences using a Gibbs sampler with the discussed priors and likelihood. You may choose the hyper parameters as you like, but they should be relatively "weak" in the sense that they don't have much of an impact on the posterior. On a single graph, plot the kernel density estimates of the posteriors for the different group means. Comment on how the groups compare (you don't need a formal statistical test, this is all based on visual inspection).