Solution to Exercise 4.6

$\star$**4.6** Let the random variables x_1 and x_2 be distributed as bivariate normal, with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and covariance σ_{12}. Using the result of Exercise 4.5, write down the joint density of x_1 and x_2 in terms of the parameters just specified. Then find the marginal density of x_1.

What is the density of x_2 conditional on x_1? Show that the mean of x_2 conditional on x_1 can be written as $\mathrm{E}(x_2 \,|\, x_1) = \beta_1 + \beta_2 x_1$, and solve for the parameters β_1 and β_2 as functions of the parameters of the bivariate distribution. How are these parameters related to the least-squares estimates that would be obtained if we regressed realizations of x_2 on a constant and realizations of x_1?

If we modify the density (4.74) so as to take account of the nonzero means, the joint density of x_1 and x_2 can be seen to be

$$\frac{1}{2\pi}\frac{1}{(1-\rho^2)^{1/2}\sigma_1\sigma_2}\times \\ \exp\Big(\frac{-1}{2(1-\rho^2)}\Big(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\Big)\Big).$$

The correlation between x_1 and x_2 is $\rho = \sigma_{12}/(\sigma_1\sigma_2)$. If we replace ρ in the exponent above by this expression, the exponent becomes

$$\frac{-1}{2(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)}\big(\sigma_2^2(x_1-\mu_1)^2 - 2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2) + \sigma_1^2(x_2-\mu_2)^2\big).$$

Thus, since $(1-\rho^2)^{1/2}\sigma_1\sigma_2 = (\sigma_1^2\sigma_2^2 - \sigma_{12}^2)^{1/2}$, the joint density can be expressed as

$$\frac{1}{2\pi}\frac{1}{(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)^{1/2}}\exp\Big(\frac{-1}{2(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)}\times \\ \big(\sigma_2^2(x_1-\mu_1)^2 - 2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2) + \sigma_1^2(x_2-\mu_2)^2\big)\Big). \tag{S4.08}$$

Since the marginal densities that correspond to a bivariate normal density are normal, the marginal density of x_1 is $\mathrm{N}(\mu_1, \sigma_1^2)$. From (4.10), the PDF is

$$\frac{1}{\sigma_1\sqrt{2\pi}}\exp\Big(-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2\Big). \tag{S4.09}$$

The density of x_2 conditional on x_1 is the joint density divided by the marginal density of x_1. The ratio of the factors that multiply the exponentials in (S4.08) and (S4.09) is

$$\frac{1}{\sqrt{2\pi}}\Big(\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}\Big)^{-1/2}.$$

The ratio of the exponential factors in expressions (S4.08) and (S4.09) is the exponential of the difference of the exponents. After a little algebra, this can be expressed as

$$\exp\biggl(\frac{-1}{2(\sigma_2^2-\sigma_{12}^2/\sigma_1^2)}\times$$
$$\Bigl((x_2-\mu_1)^2-\frac{2\sigma_{12}}{\sigma_1^2}(x_1-\mu_1)(x_2-\mu_2)+\frac{\sigma_{12}^2}{\sigma_1^4}(x_1-\mu_1)^2\Bigr)\biggr)$$
$$=\exp\biggl(\frac{-1}{2(\sigma_2^2-\sigma_{12}^2/\sigma_1^2)}\Bigl(x_2-\mu_2-\frac{\sigma_{12}}{\sigma_1^2}(x_1-\mu_1)\Bigr)^2\biggr).$$

Putting the two factors together, we find that the conditional density is

$$\frac{1}{\sqrt{2\pi}\,(\sigma_2^2-\sigma_{12}^2/\sigma_1^2)^{1/2}}\exp\biggl(\frac{-1}{2(\sigma_2^2-\sigma_{12}^2/\sigma_1^2)}\Bigl(x_2-\mu_2-\frac{\sigma_{12}}{\sigma_1^2}(x_1-\mu_1)\Bigr)^2\biggr).$$

This is just the normal distribution with conditional mean

$$\mu_2+\frac{\sigma_{12}}{\sigma_1^2}(x_1-\mu_1) \tag{S4.10}$$

and conditional variance $\sigma_2^2-\sigma_{12}^2/\sigma_1^2$. The conditional mean (S4.10) can also be written as

$$\mathrm{E}(x_2\,|\,x_1)=\mu_2-\frac{\sigma_{12}}{\sigma_1^2}\mu_1+\frac{\sigma_{12}}{\sigma_1^2}x_1=\beta_1+\beta_2x_1,$$

where

$$\beta_1=\mu_2-\frac{\sigma_{12}}{\sigma_1^2}\mu_1 \quad\text{and}\quad \beta_2=\frac{\sigma_{12}}{\sigma_1^2}.$$

Thus $\mathrm{E}(x_2\,|\,x_1)$ has the form of a linear regression function, with coefficients that depend on the parameters of the bivariate normal distribution.

Suppose that $\boldsymbol{x}_2$ is an n–vector of realizations of x_2, $\boldsymbol{x}_1$ is an n–vector of realizations of x_1, and $\boldsymbol{\iota}$ is a vector of 1s. Then, if we ran the regression

$$\boldsymbol{x}_2=\beta_1\boldsymbol{\iota}+\beta_2\boldsymbol{x}_1+\boldsymbol{u},$$

we would obtain the estimates

$$\hat\beta_2=\frac{\boldsymbol{x}_1{}^{\top}\boldsymbol{M}_{\iota}\boldsymbol{x}_2}{\boldsymbol{x}_1{}^{\top}\boldsymbol{M}_{\iota}\boldsymbol{x}_1}\quad\text{and}\quad \hat\beta_1=\frac{1}{n}\boldsymbol{\iota}^{\top}\boldsymbol{x}_2-\hat\beta_2\frac{1}{n}\boldsymbol{\iota}^{\top}\boldsymbol{x}_1.$$

It is easy to see that the plims of $\hat\beta_1$ and $\hat\beta_2$ are β_1 and β_2, respectively.