Summary:

If we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known) and our prior beliefs about μ follow a $N(b, 1/d)$ distribution then, after incorporating the data, our (posterior) beliefs about μ follow a $N(B, 1/D)$ distribution.

Notice that the way prior information and observed data combine is through the parameters of the normal distribution:

$$b \to \frac{db + n\tau\bar{x}}{d + n\tau} \qquad \text{and} \qquad d^2 \to d + n\tau.$$

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$B = \alpha b + (1 - \alpha)\bar{x},$$

for some $\alpha \in (0, 1)$. This equation for the posterior mean, which can be rewritten as

$$E(\mu|\boldsymbol{x}) = \alpha E(\mu) + (1 - \alpha)\bar{x},$$

arises in other models and is known as the *Bayes linear rule*.

The changes in our beliefs about μ are summarised in Table 2.6. Notice that the posterior mean is greater than the prior mean if and only if the likelihood mode (sample mean) is greater than the prior mean, that is

$$E(\mu|\boldsymbol{x}) > E(\mu) \qquad \Longleftrightarrow \qquad Mode[L(\mu|\boldsymbol{x})] > E(\mu).$$

lso, the standard deviation of the posterior distribution is smaller than that of the prior distribution.

	Prior (2.19)	Likelihood (2.20)	Posterior (2.22)
$Mode(\mu)$	b	$\bar{x}$	$(db + n\tau\bar{x})/(d + n\tau)$
$E(\mu)$	b	–	$(db + n\tau\bar{x})/(d + n\tau)$
$Precision(\mu)$	d	–	$d + n\tau$

Table 2.6: Changes in beliefs about μ.