

These problems are more appropriate for homework than an exam as they are messy...  
 In a Bayesian model,  $X_1, \dots, X_n | \mu \sim N(\mu, 5^2)$  with prior distribution  $\mu \sim N(100, 10^2)$ .

1. What is  $E(X_1)$ ?

Solution:  $E(X_1) = E(E(X_1 | \mu)) = E(\mu) = 100$ .

2. Write an expression for  $P(X_1 > 110)$ .

Solution: The conditional probability  $P(X_1 > 110 | \mu)$  is  $1 - \Phi((110 - \mu)/5)$ , and the answer is to average this expression against the prior density.

$$\begin{aligned} P(X_1 > 110) &= \int_{-\infty}^{\infty} P(X_1 > 110 | \mu) \pi(\mu) d\mu \\ &= \int_{-\infty}^{\infty} \int_{110}^{\infty} \frac{1}{\sqrt{2\pi}5} e^{-(x-\mu)^2/(2 \cdot 25)} dx \frac{1}{\sqrt{2\pi}10} e^{-(\mu-100)^2/(2 \cdot 100)} d\mu \end{aligned}$$

It turns out that it is possible to simplify this expression. Here is the solution to a general problem of finding  $P(X < t)$  where  $X | \mu \sim N(\mu, \sigma^2)$  and  $\mu \sim N(\mu_0, \sigma_0^2)$ . The solution uses this algebraic identity found by completing the square,

$$a(\mu - c)^2 + b(\mu - d)^2 = (a + b) \left( \mu - \frac{ac + bd}{a + b} \right)^2 + \left( \frac{ab}{a + b} \right) (c - d)^2$$

and also this algebraic identity relating sums of variances and their sums of their reciprocals.

$$\sigma\sigma_0\sqrt{\sigma^{-2} + \sigma_0^{-2}} = \sqrt{\sigma^2 + \sigma_0^2}$$

This derivation leaves out a lot of the algebra, but proceeds by switching the order of integration and writing the inner integral in terms of a normal density.

$$\begin{aligned} P(X < t) &= \int_{-\infty}^{\infty} P(X < t | \mu) \pi(\mu) d\mu \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^t \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}} d\mu \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi}\sigma\sigma_0} \left( \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2} - \frac{(\mu-\mu_0)^2}{2\sigma_0^2}} d\mu \right) dx \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2 + \sigma_0^2}} e^{-\frac{(x-\mu_0)^2}{2(\sigma^2 + \sigma_0^2)}} \\ &\quad \times \left( \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}(\sigma^{-2} + \sigma_0^{-2})^{-1/2}} e^{-\frac{\left(\mu - \frac{\sigma^{-2}x + \sigma_0^{-2}\mu_0}{\sigma^{-2} + \sigma_0^{-2}}\right)^2}{2(\sigma^{-2} + \sigma_0^{-2})^{-1}}} d\mu \right) dx \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2 + \sigma_0^2}} e^{-\frac{(x-\mu_0)^2}{2(\sigma^2 + \sigma_0^2)}} dx \\ &= \int_{-\infty}^{(t-\mu_0)/\sqrt{\sigma^2 + \sigma_0^2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\ &= \Phi\left(\frac{t - \mu_0}{\sqrt{\sigma^2 + \sigma_0^2}}\right) \end{aligned}$$

Note that this derivation implies that the marginal distribution of  $X$  is  $X \sim N(\mu_0, \sigma^2 + \sigma_0^2)$ .

It follows that the numerical answer to the original problem is then  $P(X_1 > 110) = 1 - \Phi((110 - 100)/\sqrt{100 + 25}) \doteq 1 - \Phi(0.8944) = 0.1855$ .

3. Suppose that for sample  $s$ ,  $n = 50$  and  $\bar{x} = 92$ . Find the posterior density of  $\mu$ .

Solution:

$$\begin{aligned}\pi(\mu | s) &\propto \pi(\mu) f_\mu(s) \\ &\propto e^{-(\mu-100)^2/(2 \cdot 10^2)} e^{-(50/2 \cdot 5^2)(92-\mu)^2} \\ &\propto e^{-(\mu-\mu_1)^2/(2\eta)}\end{aligned}$$

where

$$\mu_1 = \left(\frac{0.01}{2.01}\right) 100 + \left(\frac{2}{0.01}\right) 92 = 92.0398$$

and

$$\eta = \left(\frac{1}{100} + \frac{50}{25}\right)^{-1} = 0.4975$$

which implies that  $\mu | s \sim N(\mu_1, \eta)$ .

4. Write an expression for  $P(X_{51} > 110 | s)$ .

Solution: Here we integrate the probability against the posterior density instead of the prior density.

$$P(X_{51} > 110 | s) = \int_{-\infty}^{\infty} P(X_{51} > 100 | \mu) \pi(\mu | s) d\mu$$

By the work from before,  $X_{51} | s \sim N(\mu_1, \sigma^2 + \eta)$  which for this data implies that  $X_{51} | s \sim N(92.0398, 25.4975)$  so the probability is  $1 - \Phi((110 - 92.0398)/\sqrt{25.4975}) = 1 - \Phi(3.557) \doteq 0.00019$ .