Prediction of a success probability using Beta distributions as prior and posterior
(Example 4.7.5, p. 225)

Let \( X \) be the number of patients out of the first 40 in a clinical trial who have success as their outcome. Let \( P \) be the probability that an individual patient is a success.

The conditional p.f. of \( X \) given \( P = p \) is the binomial p.f. with \( n = 40 \)

\[
g_1(x|p) = \binom{40}{x} p^x (1 - p)^{40-x}.
\]

As prior distribution for \( P \), we use a uniform p.d.f. on the interval \((0, 1)\). This is a special case of a Beta distribution for \( \alpha = 1, \beta = 1 \). The Beta p.d.f. is

\[
x^{\alpha - 1} (1 - x)^{\beta - 1} / B(\alpha, \beta)
\]

where \( B(\alpha, \beta) \) is the normalizing constant which makes the integral one (\( B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx \)). It has been shown before (in the example Castaneda v. Partida, p. 304) that with a Beta \((\alpha, \beta)\) prior on \( P \), and a conditional binomial distribution with parameters \( n, p \) for \( X \), the posterior of \( P \) is again Beta with parameters

\[
\hat{\alpha} = \alpha + x, \hat{\beta} = n - x + \beta.
\]

Now let us compute \( E(P|x) \) which is the best predictor of \( P \) when \( P \) has its posterior distribution. Since this distribution is Beta \((\hat{\alpha}, \hat{\beta})\) with expectation

\[
E(P|x) = \frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}}
\]

(cf. p. 306, the expectation of a Beta distribution) we obtain

\[
E(P|x) = \frac{\alpha + x}{\alpha + x + n - x + \beta} = \frac{1 + x}{1 + x + 40 - x + 1}
\]

\[
= \frac{x + 1}{42}.
\]

Note that this predictor is very close to the observed sample proportion of success from \( X \)

\[
\hat{p} = \frac{x}{40}.
\]

Compare this predictor with the best predictor of \( P \) before observing \( X \) which is just \( E(P) \) (expectation of its prior distribution). Since the prior distribution is uniform we have \( E(P) = 1/2 \), with M.S.E. (mean square error) of prediction

\[
E (P - E(P))^2 = Var(P) = \frac{1}{12}.
\]
A computation of the error of prediction when $X$ can be observed gives

$$E[Var(P|X)] = E \left[ E \left( (P - E(P|X))^2 | X \right) \right] = 0.003968$$

which is much smaller than 1/12. Here we use the formula for the variance of the Beta distribution with parameters $(\tilde{\alpha}, \tilde{\beta})$ on p. 306: if $P$ has this distribution (conditionally on $X$) then

$$Var(P|X) = \frac{\tilde{\alpha}\tilde{\beta}}{(\tilde{\alpha} + \tilde{\beta})^2 (\tilde{\alpha} + \tilde{\beta} + 1)} = \frac{(1 + x)(40 - x + 1)}{(42)^2 (43)}$$

hence

$$E(Var(P|X)) = \frac{E(X + 1)(41 - X)}{(42)^2 (43)}$$

To compute this, we note (p. 226) that the marginal p.f. of $X$ is $f_1(x) = 1/41$ for $x = 0, \ldots, 40$, that is the uniform distribution on $0, \ldots, 40$. 
