Normal Model

sampling model $p(y|\theta,\sigma^2) = \text{dnorm}(y,\theta,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\theta}{\sigma}\right)^2},\ y \in \mathbb{R}$

conditional prior $p(\theta|\sigma^2) = \text{dnorm}(y, \mu_0, \tau_0)$

posterior $p(\theta|\sigma^2, y_1, \ldots, y_n) = \text{dnorm}(\theta, \mu_n, \tau_n)$, where $\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}}$ and $\tau_n^2 = \frac{1}{\frac{1}{\tau_0^2}+\frac{n}{\sigma^2}}$

set $\tau_0^2 = \frac{\sigma^2}{\kappa_0}$, $\mu_n = \frac{\kappa_0}{\kappa_0+n}\mu_0 + \frac{n}{\kappa_0+n}\bar{y}$ and $\tau_n^2 = \frac{\sigma^2}{\kappa_0+n} = \frac{\sigma^2}{\kappa_n}$ $(\kappa_n = \kappa_0 + n)$

prior $p(\theta,\sigma^2) = p(\theta|\sigma^2)p(\sigma^2) = \text{dnorm}(\theta, \mu_0, \tau_0 = \frac{\sigma}{\sqrt{\kappa_0}})p(\sigma^2)$, $p(\frac{1}{\sigma^2}) = \text{dgamma}(\frac{1}{\sigma^2}, \frac{\nu_0^2}{2}, \frac{\nu_0^2}{2}\sigma_0^2)$

posterior $p(\theta|\sigma^2, y_1, \ldots, y_n) = \text{dnorm}(\theta, \mu_n, \tau_n = \frac{\sigma}{\sqrt{\kappa_n}})$, $p(\frac{1}{\sigma^2}|y_1,\ldots,y_n) = \text{dgamma}(\frac{1}{\sigma^2}, \frac{\nu_n^2}{2}, \frac{\nu_n^2}{2}\sigma_n^2)$,

$\nu_n = \nu_0 + n$ and $\sigma_n^2 = \frac{1}{\nu_n}[\nu_0^2\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_n}(\bar{y}-\mu_0)2]$, $s^2 = \frac{1}{n-1}\sum_{l=1}^{n}(y_l - \bar{y})^2$

```
> # Fig 5.4 Joint posterior distributions of (theta, sigma^-2) and  (theta, sigma^2)
> # http://www2.stat.duke.edu/~pdh10/FCBS/Replication/chapter5.R
> # prior
> # 1/sigma^2 ~ dgamma(nu_0/2, nu_0 sigma^2_0/2)
> s2_0 <- 0.01 ; nu_0 <- 1
> # theta|sigma^2 ~ dnorm(mu_0, sigma^2/k_0)
> mu_0 <- 1.9 ; k_0 <- 1
>
> ## data
> y <- c(1.64, 1.70, 1.72, 1.74, 1.82, 1.82, 1.82, 1.90, 2.08)
> n <- length(y) ; ybar<-mean(y) ; s2 <- var(y)
>
> ## posterior inference
> k_n <- k_0 + n ; nu_n <- nu_0 + n
> mu_n <- (k_0 * mu_0 + n * ybar) / k_n
> s2_n <- (nu_0 * s2_0 + (n-1) * s2 + k_0 * n * (ybar - mu_0)^2 / k_n) / nu_n
> mu_n
```

```
[1] 1.814
```

```
> s2_n
```

```
[1] 0.015324
```

```
> # density of inverse gamma beta^alpha / Gamma(alpha) * x^-(alpha+1) exp(-beta/x)
> dinvgamma<-function(x,a,b) {
+   ld <- a * log(b) - lgamma(a) - (a+1) * log(x) - b / x
+   exp(ld)
+ }
>
> # grid size
> gs <- 100
> theta_g <- seq(1.6, 2.0, length = gs)
> is2_g <- seq(15, 160 , length = gs)
> s2_g <- seq(0.001, 0.045, length = gs)
>
> # log density theta inverse sigma square / sigma square
> ld.th.is2 <- ld.th.s2 <- matrix(0, gs, gs)
```

```
> for(i in 1:gs) {
+   for(j in 1:gs) {
+     ld.th.is2[i,j]<- dnorm(theta_g[i], mu_n, 1/sqrt(is2_g[j]*k_n), log = TRUE) +
+                       dgamma(is2_g[j], shape = nu_n/2, rate = nu_n*s2_n/2, log = TRUE)
+     ld.th.s2[i,j]<- dnorm(theta_g[i], mu_n, sqrt(s2_g[j]/k_n), log = TRUE) +
+                       log(dinvgamma(s2_g[j], nu_n/2, nu_n * s2_n/2 ))
+   }
+ }
>
> par(mfrow = c(1,2), mar = c(3,3,1,1), mgp = c(1.75, 0.75, 0))
> grays <- gray((10:0)/10)
> image(theta_g, is2_g, exp(ld.th.is2), col = grays, xlab = expression(theta),
+        ylab = expression(sigma^{-2}) )
> image(theta_g, s2_g, exp(ld.th.s2), col = grays, xlab = expression(theta),
+        ylab = expression(sigma^2) )
```