Multiple Param model

e.g. Gaussian with unknown mean $\mu$ & unknown variance $\sigma$

Goal: based on i.i.d observations $\{y_1, \dots, y_n\}$, infer $\mu$

$\boxed{\text{Prior}}$ $p(\mu, \sigma^2) \propto p(\mu) p(\sigma^2) \propto 1 * \frac{1}{\sigma^2} \propto \frac{1}{\sigma^2}$

(Alternatives? V!)

$\boxed{\text{Likelihood}}$ $p(y_1, \dots, y_n | \mu, \sigma^2) = \prod_{i=1}^{n} p(y_i | \mu, \sigma^2) \propto$

$$\sigma^{-n} \exp\left(-\frac{\sum_{i=1}^{n} (y_i - \mu)^2}{2\sigma^2}\right)$$

$\boxed{\text{posterior}}$

$$p(\mu, \sigma^2 | y) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)S_n^2 + n(\bar{y}_n - \mu)^2\right]\right)$$

$$\bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i \quad , \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y}_n)^2$$

☆ Quantities to be derive / estimated:

① $E[\mu | y_1, \dots, y_n]$

② $Var[\mu | y_1, \dots, y_n]$

③ and find $[a, b]$ s.t. $\int_a^b p[\mu | y_1, \dots, y_n] d\mu = 0.95$

★ Strategies in general:

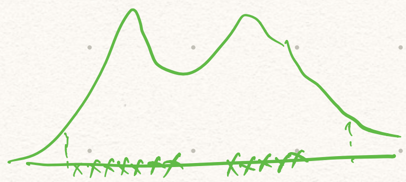
① analytical integration $\xrightarrow{\text{derive}}$ $p(\mu \mid y_1, \dots, y_n)$ $\xrightarrow[\text{analytically numerically}]{\text{calculate}}$

$$\mathbb{E}[\mu \mid y_1, \dots, y_n] = \int_{\mathbb{R}} \mu \, p[\mu \mid y_1, \dots, y_n] \, d\mu$$

② numerical approximation ( Monte Carlo methods ):

if I have $\mu^{(i)} \overset{i.i.d.}{\sim} p(\mu \mid y_1, y_2, \dots, y_n)$ then

$$\mathbb{E}(\mu \mid y_1, \dots, y_n) \approx \sum_{i=1}^{M} \mu^{(i)} / M$$



Method 1.

$$p(\mu \mid y_1, \dots, y_n) = \int_0^\infty p(\mu, \sigma^2 \mid y_1, \dots, y_n) \, d\sigma^2$$

$$\propto \left[ 1 + \frac{n(\mu - \bar{y}_n)^2}{(n-1) S_n^2} \right]^{-n/2}$$

check it using the inverse Gamma density (inverse Chi-squared)

$$\boxed{\int e^{-x^2/2} \, dx = \sqrt{2\pi}}$$

$$\left. \frac{\mu - \bar{y}_n}{S_n / \sqrt{n}} \right| y_1, \dots, y_n \sim t_{n-1} \text{ (student-T)}$$

$$\mu \mid y_1, \dots, y_n \sim \bar{y}_n + \frac{S_n}{\sqrt{n}} t_{n-1}$$

expectance = 0 for a t-dis'

① $\mathbb{E}[\mu \mid y_1, \dots, y_n] = \mathbb{E}\left[ \bar{y}_n + \frac{S_n}{\sqrt{n}} t_{n-1} \,\middle|\, y_1, \dots, y_n \right] = \bar{y}_n$ (analytical solution)

In practice:

② draw $\mu^{(i)} \sim \bar{y}_n + \frac{S_y}{\sqrt{n}} t_{n-1}$, $i = 1 \ldots, M$ ↳ Large number

(can use R)

$\mathbb{E}[\mu | y_1, y_2, \ldots, y_n] \approx \sum_{i=1}^{M} \mu^{(i)} / M$ (generic)

$Var[\mu | y_1, y_2, \ldots, y_n] \approx \sum_{i=1}^{M} (\mu^{(i)} - \bar{\mu})^2 / (M-1)$

③ What if $\int p(\mu, \sigma^2 | y_1, \ldots, y_n) d\sigma^2$ cannot be found?:

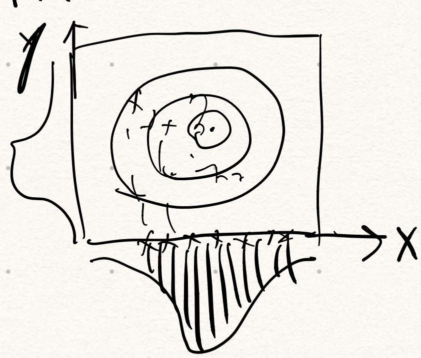
☆ draw $\{\mu^{(i)}, \sigma^{(i)2}\}$ from $p(\mu, \sigma^2 | y_1, \ldots, y_n)$, $i = 1 \cdots, M$

$\mu^{(i)} \sim p[\mu | y_1, \ldots, y_n]$

$(\sigma^{(i)})^2 \sim p[\sigma^2 | y_1, \ldots, y_n]$

$\mathbb{E}[\mu | y_1, \ldots, y_n] \approx \sum_{i=1}^{M} \mu^{(i)} / M$

Multivariate Gaussian distribution




Now, how to perform this step?

Recall $p(\mu, \sigma^2 | y_1, \dots, y_n) \propto \sigma^{-n-2} \exp\left(-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\right)$

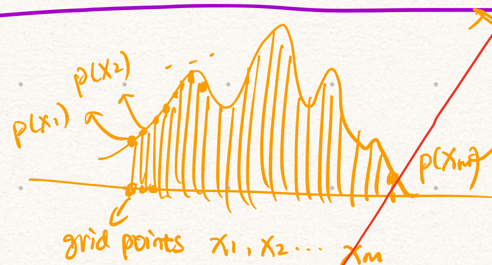
given
a p.m.f

| | | | | | |
|---|---|---|---|---|---|
| -1 | 0.5 | 0 | 0.5 | 1 | 1.5 |
| 0.6 | 0.1 | 0.15 | | | |

sample( ___ , prob= ___ ) in R

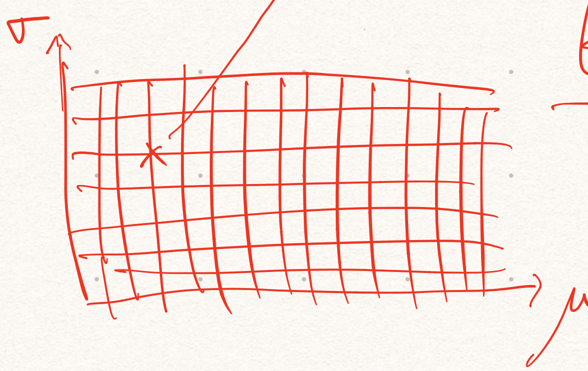
Now suppose

(discretize idea)




Similar concept

if 2 dimension:




Grid Approximation

typically even space

How to find $x_1$, $x_m$ efficiently?

④ we're dealing with. $p(\mu, \sigma^2 | y_1, \dots, y_n)$

What if we know $\sigma$? $\longrightarrow \mu$

What if we know $\mu$? $\longrightarrow \sigma^2$?

Gaussian conjugate model (Single param model)

$\mu | \sigma^2, y_1, y_2, \ldots, y_n \sim N(\bar{\mu}_n, \bar{\sigma}_n^2)$

$\sigma^2 | y_1, \cdots, y_n, \mu$ <u>another Single param model</u>

likelihood $p(y_1, \cdots, y_n | \mu, \sigma^2)$ the same

prior $p(\sigma^2) \propto \frac{1}{\sigma^2}$ (noninformative)

posterior $p(\sigma^2 | \mu, y_1, ~~y_n~~ \ldots y_n) \propto \sigma^{-n-2} \exp\left(-\frac{\sum_{i=1}^{n} (y_i - \mu)^2}{2\sigma^2}\right)$

$\propto$ inverse Gamma

$\sim$ inverse $-\chi^2 \left(n, \sum_{i=1}^{n} (y_i - \mu)^2\right)$

(check this result)

Now how to combine them together:

Sampling Strategy. (Gibbs Sampler)

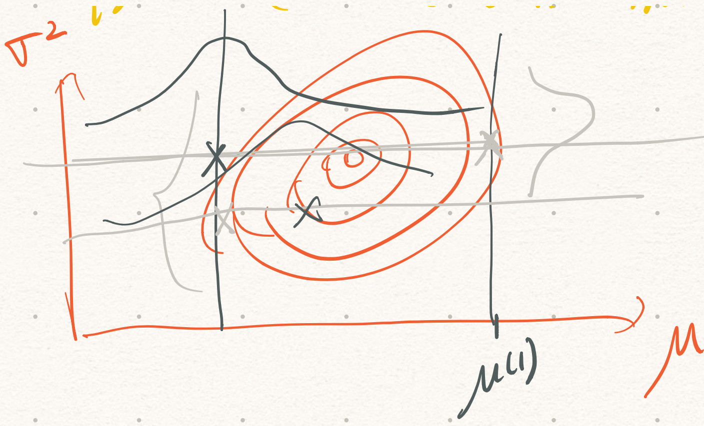
?

For $t = 1, 2, \ldots, M$, do

sample $\mu^{(t+1)}$ from $p(\mu | \sigma^{2(t)}, y_1, \ldots, y_n)$

sample $(\sigma^{(t+1)})^2$ from $p(\sigma^2 | \mu^{(t+1)}, y_1, \ldots, y_n)$

output $\{\mu^{(t)}, (\sigma^2)^{(t)}\}_{t=1,\ldots,M}$




iteratively.