ベイズ統計入門

条件付確率

- 事象Fが起こったことが既知であるという条件の下で、Eが起こる確率を条件付確率（conditional probability）という。

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

- 定義式を変形すると、確率の乗法公式となる。

$$P(E \cap F) = P(F)P(E \mid F) = P(E)P(F \mid E)$$

事象の独立

- ある事象の生起する確率が、他のある事象が生起するかどうかによって変化しないとき、2つの事象は独立であるという。

$$P(E) = P(E \mid F)$$

$$P(E \cap F) = P(E)P(F)$$

- いま、$F_1, F_2, \ldots, F_k$ が互いに素である集合で、かつ $F_1 \cup F_2 \cup \ldots \cup F_k = \Omega$ であるとすると、

$$P(E) = P(E \mid F_1)P(F_1) + P(E \mid F_2)P(F_2) + \ldots. + P(E \mid F_k)P(F_k)$$

ベイズの定理

- 以上から、ベイズの定理（Bayes' Theory）が成立する。

$$P(E_i | F) = \frac{P(F | E_i)P(E_i)}{P(F | E_1)P(E_1) + P(F | E_2)P(E_2) + ... + P(F | E_k)P(E_k)}$$

（覚え方のこつ）

データがFであるときにパラメータEiが得られる確率は、データがFでかつパラメータがEiである確率を、データがFである確率でわったもの

ベイズ定理の本質

- 前頁のベイズの定理の一般形において、右辺の分母は分子の和となっている。したがって分子も知れば分母も知る。このとき母数を定数とみなして以下のように略記する

$$P(E_i | F) \propto P(F | E_i) P(E_i)$$

- すなわち、パラメータの尤度×パラメータの事前分布＝パラメータの事後分布

確率変数(1)

- 確率が定義される基礎空間Ωの要素は事象や命題である。今、Ωの要素ωにある1つの実数xを対応させる。命題が真であるかどうかが観測するまで確定することのできない不確実性を持つとき、Ωの要素と対応するxも同じ不確実性を持つ。Ωについて定義された確率関数によってxについてその生起確率を定義でき、このようなxの全体をXとおき、確率変数(random variable)と呼ぶ。特定の値xが生起することを$X=x$と表す。

確率変数（2）

- 実験の結果生ずるXのすべての可能な値によって構成される集合を標本空間（sample space）と呼ぶ。標本空間上に定義される確率を確率分布と呼ぶ。
- Xの可能な値が不連続であるとき、xの分布を離散分布（discrete distribution）という。Xが連続の値をとるとき、xの分布を連続分布（continuous distribution）という。
- $P(X=x)$ が関数 $p(x)$ によって表されるとき、$p(x)$ を確率密度関数（probability density function）と呼ぶ。またXが連続であるとき、以下で表される $f(x)$も確率密度関数と呼ぶ。

$$P(a \le X \le b) = \int_a^b f(x)dx$$

確率変数の期待値

- 確率変数のデータの平均値は以下で表現する。

$$\bar{x} = \sum_{i=1}^{t} \frac{x(i) f_i}{n} = \sum_{i=1}^{t} x(i) \hat{p}_i$$

$x(i)$: 確率変数Xの取りうる値

f_i : 各々の観測値の数

$\hat{p}_i$: n個の観測値のうち、$x(i)$をとる観測値の割合

- 期待値は以下で表現する。

$$E(X) = \int_{\Psi} xp(x)dx$$

Ψ : 標本空間の取りうる値

$$E(g(X)) = \int_{\Psi} g(x)p(x)dx$$

確率変数の分散

- 確率変数の分散は以下で表現する。確率変数の分散とは、確率変数Xが平均して$E(X)$からどれくらい離れているかを示す。Xの分布のばらつきともいえる。

$$V(X) = E\left\{(X - E(X))^2\right\} = E\left(X^2\right) - (E(X))^2$$

- $\sqrt{V(X)}$ を標準偏差とよぶ。
- 以下で表現することもある。

$$E(X) \equiv \mu_x$$

$$V(X) \equiv \sigma_x^2$$

$$\sqrt{V(X)} \equiv \sigma_x$$

多変量分布

- 基礎空間の要素を1つの確率変数ではなく、複数の確率変数の組として表すことがある。この場合の標本空間は2次元ユークリッド空間の部分空間である。
- X、Yの分布が $p(x,y)$ によって表現されるとき、xとyが連続であればその生起確率は以下で表現し、$p(x,y)$ を同時確率密度関数と呼ぶ。

$$P\{(X,Y)\in R\}=\iint_R p(x,y)dxdy$$

周辺分布

- 以下の時、$p(x)$ を周辺確率密度関数といい、この関数が示す分布を周辺分布（marginal distribution）と呼ぶ。

yの値にかかわらず$a \le X \le b$となる確率は、

$$\int_a^b p(x)dx = \int_a^b \int_{-\infty}^{\infty} p(x, y)dydx$$

すなわち、

$$p(x) = \int_{-\infty}^{\infty} p(x, y)dy$$

*ある変数の周辺分布はもう一つの変数を積分消去することによって得られる。これを周辺化とよぶ。

条件付分布

- $Y_1 = y_1$ が与えられた時の Y_2 の条件付分布の確率密度関数は以下で与えられる

$$p(y_2|y_1) = \frac{p(y_1, y_2)}{p(y_1)}$$

共分散

- 2変数 Y_1, Y_2の同時確率密度（同時分布）が与えられている時、$\mu_1 = E(Y_1), \mu_2 = E(Y_2)$ とすると、以下をそれぞれ共分散、相関係数と呼ぶ

$$Cov(Y_1, Y_2) = E(E(Y_1 - E(Y_1))E(Y_2 - E(Y_2)))$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_1 - \mu_1)(y_2 - \mu_2) p(y_1, y_2) dy_1 dy_2$$

$$\rho = \frac{Cov(Y_1, Y_2)}{\sqrt{V(Y_1)}\sqrt{V(Y_2)}}$$

ベイズ定理のまとめ

$$P(\theta \mid y) = \frac{P(y \mid \theta)P(\theta)}{P(y)}$$

$$P(\theta \mid y) = \frac{L(y \mid \theta)P(\theta)}{M(y)} \propto L(y \mid \theta)P(\theta)$$

$$P(y, \theta) = P(\theta \mid y)P(y) = P(y \mid \theta)P(\theta)$$

$P(\theta \mid y)$: 事後密度関数 (posterior density function)

$L(y \mid \theta)$: 尤度 (likelihood)

$P(\theta)$: 事前分布 (prior distribution)

$M(y) = \int L(y \mid \theta)P(\theta)d\theta$: 周辺尤度 (marginal likelihood)

要は分子をパラメータに関して積分
＝データyの得られる尤度

1変量正規分布のベイズ推定

- 正規分布からN個のランダムサンプリング

$$y_1,\ldots,y_n \overset{i.i.d.}{\sim} N\left(\mu,\sigma^2\right)$$

- 分散σ^2を既知とし、μを推定する
- 尤度関数

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

- 事前分布

$$\mu \sim N\left(\mu_0,\sigma_0^2\right)$$

$$p(\mu)=\frac{1}{\sqrt{2\pi\sigma_0}}\exp\left(-\frac{\left(\mu-\mu_0\right)^2}{2\sigma_0^2}\right)$$

補足：iidとは

- Independently and identically distributed:独立に同一の分布に従うの意味。次のように用いる

$$Y_i \sim \text{i.i.d.}\, N\left(\mu, \sigma^2\right)$$

1変量正規分布のベイズ推定(続き)

- 事後分布(事前分布と尤度関数をかける)

$$p(\mu|y) \equiv p(\mu|y_1,\ldots,y_n) \propto p(\mu) \bullet p(y_1,\ldots,y_n|\mu)$$

$$\propto \exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2} - \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2} \right)$$

$$\sum(y_i-\mu)^2 = \sum(y_i-\bar{y}+\bar{y}-\mu)^2 = \sum(y_i-\bar{y})^2 + n(\bar{y}-\mu)^2$$

$$= (n-1)s^2 + n(\bar{y}-\mu)^2$$

$$\propto \exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2} - \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2} \right) \propto \exp\left(-\frac{(\mu-\mu')^2}{2\sigma'^2} \right)$$

ここでμ'はμの事後平均をさす

1変量正規分布のベイズ推定（続き）

- 事後分布（続き）

$$\mu|\mathbf{y} \sim N(\mu', \sigma'^2)$$

縮約推定（データの平均を事前平均へ縮約）

$$\begin{cases} \mu' = \dfrac{\dfrac{1}{\sigma_0^2}\mu_0 + \dfrac{n}{\sigma^2}\bar{y}}{\dfrac{1}{\sigma_0^2} + \dfrac{n}{\sigma^2}} = \omega\mu_0 + (1-\omega)\bar{y} \\ \sigma'^2 = \dfrac{1}{\dfrac{1}{\sigma_0^2} + \dfrac{n}{\sigma^2}}, \omega = \dfrac{\dfrac{1}{\sigma_0^2}}{\dfrac{1}{\sigma_0^2} + \dfrac{n}{\sigma^2}} \end{cases}$$

精度表示（精度＝分散の逆数）

$$\tau = \frac{1}{\sigma_0^2}, \gamma = \frac{1}{\sigma^2} \Rightarrow \gamma' = \tau + n\gamma$$ （事後精度＝事前精度＋データ精度）

$$\mu' = \frac{\tau\mu + n\gamma\bar{y}}{\tau + n\gamma}$$ 事後平均＝$\dfrac{\text{事前精度}\bullet\mu_0 + \text{データ精度}\bullet\bar{y}}{\text{事前精度}+\text{データ精度}}$

1変量正規分布のベイズ推定②

- 平均 μ を既知とし、分散 σ^2 を推定する
- 尤度関数

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

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

- 事前分布

$$\sigma^2 \sim Inv-\chi^2\left(v_0,\sigma_0\right) \propto \left(\frac{\sigma_0^2}{\sigma^2}\right)^{\frac{v_0}{2}+1} \exp\left(-\frac{v_0\sigma_0^2}{2\sigma^2}\right)$$

$$=-\chi^2\left(v_0,\sigma_0\right)=-\sigma_0\chi^2\left(v_0\right)$$

補足:カイ二乗分布と逆カイ二乗分布

- いずれまた機会があれば

1変量正規分布のベイズ推定②

- 事後分布

$$p\left(\sigma^2 \middle| y\right) \propto \left(\frac{\sigma_0^2}{\sigma^2}\right)^{\frac{v_0}{2}+1} \exp\left(-\frac{v_0\sigma_0^2}{2\sigma^2}\right) \bullet \left(\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\frac{v}{\sigma^2}\right)$$

$$\propto \left(\sigma^2\right)^{-\frac{n+v_0}{2}+1} \exp\left(-\frac{1}{2\sigma^2}\left(v_0\sigma_0^2 + nv\right)\right)$$

$$\sigma^2 \middle| y \sim Inv-\chi^2\left(v_0 + n, \frac{v_0\sigma_0^2 + nv}{\sigma_0 + n}\right)$$

1変量正規分布のベイズ推定③

- μ、σ 未知
- 共役事前分布

$$\mu \mid \sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{k_0}\right)$$

$$\sigma^2 \sim Inv-\chi^2\left(v_0, \sigma_0^2\right)$$

$$p\left(\mu, \sigma^2\right) \propto \sigma^{-1}\left(\sigma^2\right)^{-\left(\frac{v_0}{2}+1\right)} \exp\left(-\frac{1}{2\sigma^2}\left(v_0\sigma_0^2 + k_0\left(\mu_0-\mu\right)^2\right)\right)$$

$$N-Inv-\chi^2\left(\mu_0, \frac{\sigma_0^2}{k_0}; v_0, \sigma_0^2\right)$$

1変量正規分布のベイズ推定③

- 尤度関数と同時事後分布

$$p(\mu,\sigma^2|y) \propto \sigma^{-1}(\sigma^2)^{-\left(\frac{v_0}{2}+1\right)} \exp\left(-\frac{1}{2\sigma^2}\left(v_0\sigma_0^2 + k_0(\mu-\mu_0)^2\right)\right)$$

$$\bullet (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\left((n-1)s^2 + n(\bar{y}-\mu)^2\right)\right)$$

$$= N - Inv - \chi^2\left(\mu_n, \frac{\sigma_n^2}{k_n}; v_n, \sigma_n^2\right)$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$\mu_n = \frac{k_0}{k_0+n}\mu_0 + \frac{n}{k_0+n}\bar{y}$$

$$k_n = k_0 + n; v_n = v_0 + n$$

$$v_n\sigma_n^2 = v_0\sigma_0^2 + (n-1)s^2 + \frac{k_0 n}{k_0+n}(\bar{y}-\mu_0)^2$$

1変量正規分布のベイズ推定③

- 条件付事後分布

$$\mu \mid \sigma^2, y \sim N\left(\mu_n, \frac{\sigma^2}{k_n}\right)$$

$$= N\left(\frac{\frac{k_0}{\sigma^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{k_0}{\sigma^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{k_0}{\sigma^2} + \frac{n}{\sigma^2}}\right)$$

- 周辺事後分布

$$\sigma^2 \mid y \sim Inv-\chi^2\left(v_n, \sigma_n^2\right)$$

多変量正規分布のベイズ推定①

- 尤度関数

$$y|\mu,\Sigma \sim N_d(\mu,\Sigma)$$

Σ : 分散共分散行列

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

$$=|\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2}tr\left(\Sigma^{-1}S_0\right)\right)$$

$$S_0=\sum_{i=1}^{n}(y_1-\mu)(y_1-\mu)'$$

多変量正規分布のベイズ推定①

- Σ既知
- 事前分布 $\mu \sim N(\mu_0, \Lambda_0)$

$$p(\mu) \propto \exp\left(-\frac{1}{2}\left((\mu-\mu_0)'\Lambda_0^{-1}(\mu-\mu_0)\right)\right)$$

- 事後分布

$$\mu | y, \Sigma \sim N(\mu_n, \Sigma_n)$$

$$p(\mu | y, \Sigma) \propto \exp\left(-\frac{1}{2}\left((\mu-\mu_0)'\Lambda_0^{-1}(\mu-\mu_0)+\sum_{i=1}^{n}(y_1-\mu)'\Sigma^{-1}(y_1-\mu)\right)\right)$$

$$\propto \exp\left(-\frac{1}{2}\left((\mu-\mu_n)'\Lambda_n^{-1}(\mu-\mu_n)\right)\right)$$

$$\begin{cases} \mu_n = \left(\Lambda_0^{-1}+n\Sigma^{-1}\right)^{-1}\left(\Lambda_0^{-1}\mu_0+n\Sigma^{-1}\bar{y}\right) \\ \Lambda_n^{-1} = \Lambda_0^{-1}+n\Sigma^{-1} \end{cases}$$

多変量正規分布のベイズ推定②

- Σ未知
- 事前分布

$$Normal-Inv-Wishart\left(\mu_0, \Lambda_0/k_0; v_0, \Lambda_0\right)$$

$$p(\mu,\Sigma) \propto |\Sigma|^{-((v_0+d)/2+1)} \exp\left(-\frac{1}{2}tr\left(\Lambda_0\Sigma^{-1}\right)-\frac{k_0}{2}(\mu-\mu_0)'\Sigma^{-1}(\mu-\mu_0)\right)$$

$$\mu|\Sigma \sim N\left(\mu_0, \Sigma/k_0\right)$$

$$\Sigma \sim Inv-Wishart\left(v_0, \Lambda_0\right)$$

- 事後分布

$$\mu|y,\Sigma \sim N\left(\mu_n, \Sigma/k_n\right)$$

$$\Sigma|y \sim Inv-Wishart\left(v_n, \Lambda_n\right)$$

$$\mu_n = \frac{k_0}{k_0+n}\mu_0 + \frac{n}{k_0+n}\bar{y}$$

$$k_n = k_0 + n; v_n = v_0 + n$$

$$\Lambda_n = \Lambda_0 + S + \frac{k_0 n}{k_0+n}(\bar{y}-\mu_0)(\bar{y}-\mu_0)'$$

$$S = \sum_{i=1}^{n}(y_i-\bar{y})(y_i-\bar{y})'$$

ベイズを理解するためには

- ここまでの数式の導出を完全に理解するためには、確率統計を一から丁寧にさらっていかなくてはならない
- それはまるでコンピュータを使うのにマシン語を習うようなもの

線形回帰モデルのベイズ推測

- モデル

$$Y = X\beta + \varepsilon; \varepsilon \sim N(0, \Sigma); \Sigma \text{ known}$$

- 事前分布

$$\beta \sim N(\mu, \Sigma_\beta); \mu, \Sigma_\beta \text{ ハイパーパラメー}$$

- 古典的ベイズ：ハイパーパラメータを指定
- 経験ベイズ：ハイパーパラメータをデータから推定
- 階層的ベイズ：ハイパーパラメータに確率分布を設定（第2段階事前分布として階層構造を形成）

$$\begin{cases} \mu \sim f(\mu) \\ \Sigma_\beta \sim g(\Sigma_\beta) \end{cases}$$

線形回帰モデルのベイズ推測
古典的ベイズ(Classical Bayes)

- 共役事前分布(正規分布)がもっとも簡単
- 望ましくない性質を与える場合あり
  - Unbounded Risk:事前分布が悪ければベイズ推定も話しにならぬ
  - ハイパーパラメータ(事前分布パラメータ)の選択が困難
  - 事前分布とデータが矛盾してしまう

線形回帰モデルのベイズ推測
経験ベイズ（Emprical Bayes）

- 回帰のパラメータと共に、ハイパーパラメータも推定（最尤法＝ML、モーメント法＝MOMなど）
- 問題点
  - ハイパーパラメータの推定に伴う不確実性を考慮しない（小標本で問題）
  - Admissibleではない
  - 推定された事前分布の効果が漸近的にゼロへ収束しない

線形回帰モデルのベイズ推測
階層的ベイズ（Hierachical Bayes）

- 2段階のベイズ推定手続きを行う
- 経験ベイズ推定のハイパーパラメータをベイズで推定
- ハイパーパラメータに事前分布を設定（通常は無情報事前分布、non-informative）

線形回帰モデルのベイズ推測

- モデル $y_t = x_t^{'}\beta + \varepsilon_t ; t = 1,2,\ldots, T$

  $y_t : t$期の従属変数

  $x_t : t$期の $(k \times 1)$ 次元説明変数ベクトル

  $\beta : (k \times 1)$ 次元パラメータベクト ル

  $\varepsilon_t : t$期の誤差項（独立な正規分布 $N(0, \sigma^2)$ を仮定）

- 行列表記 $y = X\beta + \varepsilon$
- 尤度関数

$$L(y|\beta, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\left\{-\frac{1}{2}\sigma^{-2}(Y - X\beta)'(Y - X\beta)\right\}$$

線形回帰モデルのベイズ推測
事前分布の特定化

- 非報知的事前分布（Non-Informative Prior）
  - 一様（uniform）事前分布（省略）
  - Jeffreys事前分布（省略）
- 共役事前分布
  - 古典的ベイズ推定
  - 経験ベイズ推定
  - 階層的ベイズ推定

線形回帰モデルのベイズ推測
古典的ベイズ推定

- σ^2既知
- 尤度関数 $$L(y|\beta) \propto (\sigma^2)^{-T/2} \exp\left\{-\frac{1}{2}\sigma^{-2}(Y-X\beta)'(Y-X\beta)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\sigma^{-2}\left[s^2+(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\right]\right\}$$

$$\hat{\beta}=(X'X)^{-1}X'Y$$

$$s^2=(Y-X\hat{\beta})'(Y-X\hat{\beta})$$

- 事前分布 $\beta \sim N(\mu, \Sigma_\beta)$
- 事後分布 $\beta|y \sim N(\mu', \Sigma'_\beta)$

$$\begin{cases} \mu' = \left(\dfrac{1}{\sigma^2}(X'X)+\Sigma_\beta^{-1}\right)\left(\dfrac{1}{\sigma^2}(X'X)\hat{\beta}+\Sigma_\beta^{-1}\mu\right) \\ \Sigma'_\beta = \left(\dfrac{1}{\sigma^2}(X'X)+\Sigma_\beta^{-1}\right)^{-1} \end{cases}$$

線形回帰モデルのベイズ推測
リッジ回帰

- 前頁の式を以下の様に変形するとリッジ回帰となる

$$\mu = 0$$

$$\Sigma_\beta = v^2 I_k$$

$$r = \sigma^2 / v^2$$

とすると、

$$\mu' = E(\beta|y) = \left(X'X + rI_k\right)^{-1}\left(X'X\right)\hat{\beta} = \left(X'X + rI_k\right)^{-1} X'y$$

線形回帰モデルのベイズ推測
経験ベイズ推定

- データのmarginal densityを直接利用
- デンジョン・ルールを直接推定
- 事後分布のパラメータを推定（階層的ベイズの特殊ケース）

- すべて省略

線形回帰モデルのベイズ推測 階層的ベイズ推定①

- 事前分布

$$\beta \sim N\left(\mu, v^2 \Omega\right); \Omega \text{ known}$$

v^2 : ハイパーパラメータ（事前分布のパラメータ）

ある分布 $\mathrm{p}\left(\mathrm{v}^2\right)$ を仮定

- 条件付事後分布

$$\beta \mid y, v^2 \sim N\left(\mu', \Sigma'_\beta\right)$$

$$\mu' = \left(\frac{1}{\sigma^2}\left(X'X\right) + v^2\Omega^{-1}\right)\left(\frac{1}{\sigma^2}\left(X'X\right)\hat{\beta} + v^2\Omega^{-1}\mu\right)$$

$$\Sigma'_\beta = \left(\frac{1}{\sigma^2}\left(X'X\right) + v^2\Omega^{-1}\right)^{-1}$$

条件付ベイズ推定量 : $\hat{\beta}_{v^2} = \mu'$

線形回帰モデルのベイズ推測
階層的ベイズ推定②

- 階層的ベイズ推定量

$$\widehat{\beta}_{HR} = \int \widehat{\beta}_{v^2} P\left(v^2 \middle| y\right) dv^2$$

$$= \int \left(\frac{1}{\sigma^2}\left(X'X\right) + v^2\Omega^{-1} \right)\left(\frac{1}{\sigma^2}\left(X'X\right)\hat{\beta} + v^2\Omega^{-1}\mu \right) P\left(v^2 \middle| y\right) dv^2$$

$$P\left(v^2\right) \propto 1; (improper\text{の場合})$$

$$P\left(v^2 \middle| y\right) = \frac{P\left(y \middle| v^2\right) P\left(v^2\right)}{\int P\left(y \middle| v^2\right) P\left(v^2\right) dv^2}$$

$$= \frac{\left(\sigma^2 + v^2\right)^{-K} \exp\left(-Q / \left\{2\left(\sigma^2 + v^2\right)\right\}\right) dv^2}{\int \left(\sigma^2 + v^2\right)^{-K} \exp\left(-Q / \left\{2\left(\sigma^2 + v^2\right)\right\}\right) dv^2}$$

$$Q = \left(\beta - \widehat{\beta}\right)' X' X \left(\beta - \widehat{\beta}\right)$$

線形回帰モデルのベイズ推測
階層的ベイズ推定③

- 前頁の式の分母の積分は解析的には評価不可能であるため、以下のアルゴリズムを用いる
- 完全条件付分布とアルゴリズム

$\beta | y, v^2 \sim N\left(\mu', \Sigma'_{\beta}\right)$

$v^2 | y, \beta \sim v^{-K} \exp\left(-Q/2v^2\right)$;逆ガンマ分布 (K, Q)

(i) v_0^2を固定して、β_0を$N\left(\mu', \Sigma'_{\beta}\right)$からサンプリング

(ii) β_0を固定して、v_0^2を逆ガンマ分布からサンプリング

(i),(ii)を繰り返して、$\{v_0^2, v_1^2, v_2^2 \ldots,\}\{\beta_0, \beta_1, \beta_2 \ldots,\}$

これらの経験分布（ヒストグラム）が事後分布$p\left(\beta, v^2 | y\right)$に収束

周辺だけを見れば、$p\left(\beta | y\right)$、$p\left(v^2 | y\right)$

点推定 :$\frac{1}{M}\sum_{i=0}^{M} \beta_i \to E\left(\beta | y\right), \frac{1}{M}\sum_{i=0}^{M} v_i^2 \to E\left(v^2 | y\right)$

事後分布評価

- モンテカルロ法
  - 棄却・受容(Accept/Reject)法
  - MCMC（マルコフチェーンモンテカルロ）法
- Gibbsサンプリング
- Metropolis-Hastings(M-H)サンプリング
  - ランダム・ウォーク(RW)アルゴリズム
  - 独立M-Hアルゴリズム