Комоедица — белорусский народный праздник, посвящённый пробуждению медведя :)

1. Докажем свойства оценок максимального правдоподобия!

   Пусть $L(y|\theta)$ — функция правдоподобия, y — вектор-столбец из n наблюдений, θ — вектор-столбец из m параметров. Кроме того, введём дополнительные обозначения, $\ell(\theta) = \ln L(y|\theta)$, $s(\theta) = \partial\ell/\partial\theta$. Буква s сокращает слово «score».

   Для наших целей мы определим информацию Фишера как $I = \operatorname{Var}(s(theta))$. То есть информация Фишера — это ковариационная матрица первых производных лог-функции правдоподобия. По определению.

   а) Чтобы взбодриться, укажите размеры векторов и матриц $s(\theta)$, $\operatorname{E}(s(\theta))$, $\operatorname{Var}(s(\theta))$.

   б) Собрав всю силу воли в кулак, найдите $\operatorname{E}(1)$.

   в) Запишите $\operatorname{E}(1)$ с помощью интеграла по dy и функции правдоподобия $L()$.

   г) Продифференциировав обе части найденного тождества по θ_j, найдите $\int \frac{\partial L}{\partial \theta_j}\, dy$.

   д) Найдите $\operatorname{E}\left(\frac{\partial \ell}{\partial \theta_j}\right)$.

   е) Найдите $\operatorname{E}(s(\theta))$.

   ж) Докажите, что $I = \operatorname{E}(s(\theta)s(\theta)')$.

   з) Вспомнив магию дифференциирования ещё раз, найдите $\int \frac{\partial^2 L}{\partial\theta_j\partial\theta_i}\, dy$.

   и) Найдите $\operatorname{E}\left(\frac{\partial^2 L}{\partial\theta_j\partial\theta_i}\frac{1}{L}\right)$.

   к) Выразите $\frac{\partial^2 \ell}{\partial\theta_j\partial\theta_i}$ через $\frac{\partial \ell}{\partial\theta_j}$, $\frac{\partial \ell}{\partial\theta_i}$ и $\frac{\partial^2 L}{\partial\theta_j\partial\theta_i}$.

   л) Докажите, что $\operatorname{E}(\frac{\partial^2 \ell}{\partial\theta_j\partial\theta_i}) = -\operatorname{E}\left(\frac{\partial \ell}{\partial\theta_j}\frac{\partial \ell}{\partial\theta_i}\right)$.

   м) Докажите, что $I = -\operatorname{E}\left(\frac{\partial^2 \ell}{\partial\theta\partial\theta'}\right)$.

2. ML в линейных моделях:

   Можно смело считать первое упражнение сделанным, то есть использовать тот факт, что $I = -E(\frac{\partial^2 \ell}{\partial\theta\partial\theta'})$.

   Рассмотрим задачу линейной регрессии, $y = X\beta + u$, где $u \sim \mathcal{N}(0;\sigma^2)$. Для удобства определим x_i' — i-ую строку матрицы X и будем считать регрессоры неслучайными.

   В нашем случае $\theta = (\beta', \sigma^2)'$.

   Хинты для забывших матричное дифференцирование: $\frac{\partial Ar}{\partial r} = A$, $\frac{\partial^2 r'Ar}{\partial r\partial r'} = A + A'$, $\frac{\partial r'Ar}{\partial r'} = Ar + A'r$.

   а) Выпишите $\ell(\theta)$ в виде суммы.

   б) Выпишите вектор $s(\theta)$ в виде $s(\theta) = \begin{pmatrix} ? \\ ? \end{pmatrix}$, где первый элемент — это сразу вектор производных по всем β одним махом.

в) Найдите ML оценки $\hat{\theta}$.

г) Докажите, что $L(\hat{\theta}) = a \cdot RSS^b$, где a и b — некоторые константы. Забейте на a и найдите b.

д) Найдите $\frac{\partial^2 \ell}{\partial\theta\partial\theta'}$ в виде четырёх блоков:

$$\frac{\partial^2 \ell}{\partial\theta\partial\theta'} = \begin{pmatrix} \frac{\partial^2 \ell}{\partial\beta\partial\beta'} & \frac{\partial^2 \ell}{\partial\beta\partial\sigma^2} \\ \frac{\partial^2 \ell}{\partial\sigma^2\partial\beta'} & \frac{\partial^2 \ell}{\partial\sigma^2\partial\sigma^2} \end{pmatrix}.$$

е) В предыдущем пункте все блоки должны были получится ненулевые. Однако найдите I и пара блоков занулится :)

ж) Найдите I^{-1}.

3. Выведите формулу для R^2 в регрессии вектора $y = X\beta + u$.

4. LM-тест в линейных моделях.

Обозначим: $\hat{\theta}_R$ и $\hat{\theta}_{UR}$ — ограниченные и неограниченные экстремумы правдоподобия, а $\hat{u}_R$ и $\hat{u}_{UR}$ — соответствующие остатки.

Определим LM статистику как $LM = s(\hat{\theta}_R)'\widehat{\mathrm{Var}}^{-1}(s)s(\hat{\theta}_R)$.

Будем считать первое упражнение сделанным, поэтому

$$LM = s(\hat{\theta}_R)'I^{-1}(\hat{\theta}_R)s(\hat{\theta}_R).$$

Также можно считать сделанным второе упражнение, поэтому:

$$s = \begin{pmatrix} \frac{X'u}{\sigma^2} \\ \frac{u'u - n\sigma^2}{2\sigma^4} \end{pmatrix}, \quad I^{-1} = \begin{pmatrix} \sigma^2(X'X)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}.$$

а) Кстати, а чему равно $s(\hat{\theta}_{UR})$?

б) Найдите $s(\hat{\theta}_R)$ и $I^{-1}(\hat{\theta}_R)$.

в) Выведите формулу для LM статистики, содержащую только X, $\hat{u}_R$ и n.

г) Какую регрессию надо построить, чтобы R^2 в ней оказался таким, что $LM = nR^2$?

5. Исследовательница Елизавета оценила модель множественной регрессии, $y = X\beta + u$. Затем Елизавета проверяет гипотезу о незначимости отдельного коэффициента β_j двумя способами: через t-статистику и через F-статистику с ограниченной и неограниченной регрессией.

Докажите, что $t^2 = F$.

6. Василий обнаружил странную монетку и решил произвести над ней эксперименты. Выпадение орла он кодирует $y_i = 1$, решки — $y_i = 0$.

При известном параметре p, наблюдаемые $y_1, \ldots, y_n$ независимы и имеют распределение Бернулли, $y_i|p \sim Bernoulli(p)$. Априорно по мнению Василия параметр p имеет бета-распределение, $p \sim beta(\alpha, \beta)$, где α и β — некоторые неслучайные константы, описывающие мнения Василия.

Функция плотности бета-распределения имеет вид:

$$f(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$$

Василий подкинул неизвестную монетку 100 раз и оказалось, что орёл выпал 70 раз и решка — 30.

а) При каких α и β априорное распределение совпадает с равномерным?

б) Найдите апостериорное распределение параметра p.

в) Найдите апостериорный прогнозный закон распределения для y_{n+1}.

г) Проинтерпретируйте смысл чисел α и β.

7. В i-ый день Тимофей встречает y_i покемонов. Тимофей предполагает, что при известном параметре λ, наблюдаемые y_1, ..., y_n независимы и имеют пуассоновское распределение, $y_i|\lambda \sim Pois(\lambda)$. Априорно по мнению Тимофея параметр λ имеет гамма-распределение, $p \sim Gamma(shape = \alpha, rate = \beta)$, где α и β — некоторые константы, определяющие мнение Тимофея о встречаемости покемонов.

   Функция плотности гамма-распределения имеет вид:

$$f(\lambda) \propto \lambda^{\alpha-1} \exp(-\lambda\beta)$$

   За прошедшие 100 дней Тимофей встретил 70 покемонов.

   а) Найдите апостериорное распределение параметра λ.

   б) Найдите апостериорный прогнозный закон распределения y_{n+1}.

   в) Проинтерпретируйте смысл констант α и β.

8. Андрей генерирует случайные величины X_i и Y_i по следующим принципам. Начинает Андрей с $X_0 = 0$. При $i \geq 1$ Василий генерирует Y_i из нормального распределения $Y_i|X_{i-1} \sim \mathcal{N}(0.5X_{i-1}+2, 1)$. Затем Андрей генерирует X_i из нормального распределения $X_i|Y_i \sim \mathcal{N}(0.5Y_i + 4, 1)$.

   а) Как в пределе распределена величина X_i?

   б) Как в пределе распределена величина Y_i?

9. Василий генерирует случайные величины X_i и Y_i по следующим принципам. Начинает Василий с $X_0 = 0$. При $i \geq 1$ Василий генерирует Y_i из нормального распределения $Y_i|X_{i-1} \sim \mathcal{N}(X_{i-1}, 1)$. Затем Василий генерирует независимую от Y_i величину $Z_i \sim \mathcal{N}(0; 1)$. Если оказалось, что $Z_i > Y_i$, то Василий берёт $X_i = 1$, и $X_i = 0$ иначе.

   а) Как в пределе распределена величина X_i?

   б) Как в пределе распределена величина Y_i?

10. Рассмотрим линейную модель $y = X\beta + u$, причем $u \sim \mathcal{N}(0; \sigma^2 I)$.

    Пусть априорно считается, что $\beta \sim \mathcal{N}(0; \tau I)$.

    Константы X, τ и σ^2 известны. Для удобства все X и y центрированы.

    а) Найдите апостериорное распределение β с учётом наблюдаемых y.

    б) Найдите, при каком λ апостериорное среднее совпадёт с результатом гребневой регрессии

    $$\min_{\beta} \sum (y_i - \hat{y}_i)^2 + \lambda \sum \beta_j^2.$$

    в) Как надо изменить целевую функцию гребневой регрессии, чтобы результат совпал с апостериорным средним β при априорном распределении $\beta \sim \mathcal{N}(0; \Sigma)$?