Семинар 19.

Рассмотрим применение тестов W, LR и LM для тестирования гипотез о параметрах распределения.

Пусть требуется протестировать систему (нелинейных) ограничений относительно вектора неизвестных параметров

$$H_0: \begin{cases} g_1(\theta) = 0 \\ g_2(\theta) = 0 \\ \dots \\ g_r(\theta) = 0 \end{cases}$$

где $g_i(\theta)$ — функция, которая задаёт i-ое ограничение на вектор параметров θ, $i = 1, \dots, r$.

$$\frac{\partial g}{\partial \theta'} = \begin{pmatrix} \partial g_1/\partial\theta' \\ \partial g_2/\partial\theta' \\ \vdots \\ \partial g_r/\partial\theta' \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial\theta_1} & \frac{\partial g_1}{\partial\theta_2} & \dots & \frac{\partial g_1}{\partial\theta_k} \\ \frac{\partial g_2}{\partial\theta_1} & \frac{\partial g_2}{\partial\theta_2} & \dots & \frac{\partial g_2}{\partial\theta_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_r}{\partial\theta_1} & \frac{\partial g_r}{\partial\theta_2} & \dots & \frac{\partial g_r}{\partial\theta_k} \end{pmatrix}$$

$$\frac{\partial g'}{\partial \theta} = \begin{pmatrix} \frac{\partial g'_1}{\partial\theta} & \frac{\partial g'_2}{\partial\theta} & \dots & \frac{\partial g'_r}{\partial\theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial\theta_1} & \frac{\partial g_2}{\partial\theta_1} & \dots & \frac{\partial g_r}{\partial\theta_1} \\ \frac{\partial g_1}{\partial\theta_2} & \frac{\partial g_2}{\partial\theta_2} & \dots & \frac{\partial g_r}{\partial\theta_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_1}{\partial\theta_k} & \frac{\partial g_2}{\partial\theta_k} & \dots & \frac{\partial g_r}{\partial\theta_k} \end{pmatrix}, \ \frac{\partial \ell}{\partial\theta} = \begin{pmatrix} \frac{\partial\ell}{\partial\theta_1} \\ \frac{\partial\ell}{\partial\theta_2} \\ \vdots \\ \frac{\partial\ell}{\partial\theta_k} \end{pmatrix}$$

$$I(\theta) = -E\left(\frac{\partial^2\ell}{\partial\theta\partial\theta'}\right) = -E\begin{pmatrix} \frac{\partial^2\ell}{\partial\theta_1\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_1\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_1\partial\theta_k} \\ \frac{\partial^2\ell}{\partial\theta_2\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_2\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_2\partial\theta_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2\ell}{\partial\theta_k\partial\theta_1} & \frac{\partial^2\ell}{\partial\theta_k\partial\theta_2} & \dots & \frac{\partial^2\ell}{\partial\theta_k\partial\theta_k} \end{pmatrix}$$ — информационная матрица Фишера

$\Theta_{UR} := \Theta$ — множество допустимых значений вектора неизвестных параметров без учёта ограничений

$\Theta_R := \{\theta \in \Theta : g(\theta) = 0\}$ — множество допустимых значений вектора неизвестных параметров с учётом ограничений

$\hat\theta_{UR} \in \Theta_{UR}$ — точка максимума функции ℓ на множестве Θ_{UR}

$\hat\theta_R \in \Theta_R$ — точка максимума функции ℓ на множестве Θ_R

Тогда для тестирования гипотезы H_0 можно воспользоваться одной из следующих ниже статистик.

$LR := -2(\ell(\hat\theta_R) - \ell(\hat\theta_{UR})) \overset{a}{\sim} \chi^2_r$ — статистика отношения правдоподобия

$W := g'(\hat\theta_{UR}) \cdot \left[\frac{\partial g}{\partial\theta'}(\hat\theta_{UR}) \cdot I^{-1}(\hat\theta_{UR}) \cdot \frac{\partial g'}{\partial\theta}(\hat\theta_{UR})\right]^{-1} g(\hat\theta_{UR}) \overset{a}{\sim} \chi^2_r$ — статистика Вальда

$LM := \left[\frac{\partial\ell}{\partial\theta}(\hat\theta_R)\right]' \cdot I^{-1}(\hat\theta_R) \cdot \left[\frac{\partial\ell}{\partial\theta}(\hat\theta_R)\right] \overset{a}{\sim} \chi^2_r$ — статистика множителей Лагранжа

1. Величины $X_1, ..., X_n$ — независимы и нормально распределены, $N(\mu, \sigma^2)$. По 100 наблюдениям $\sum_{i=1}^n X_i = 100$ и $\sum_{i=1}^n X_i^2 = 900$.

a) Найдите ML оценки неизвестных параметров μ и σ^2.

b) Постройте 95%-ые доверительные интервалы для μ и σ^2.

c) С помощью LR, LM и W теста проверьте гипотезу о том, что $\sigma^2 = 1$ и одновременно $\mu = 2$.

(a)

$$\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \quad = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i-\mu)^2}{2\sigma^2}\right), \quad \ln = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$\text{FOC:} \quad \frac{\partial \ln}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu), \quad \frac{\partial \ln}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$\frac{\partial^2 \ln}{\partial \mu^2} = -\frac{n}{\sigma^2}, \quad \frac{\partial^2 \ln}{\partial \mu\, \partial(\sigma^2)} = -\frac{1}{(\sigma^2)^2}\sum_{i=1}^{n}(X_i-\mu), \quad \frac{\partial^2 \ln}{\partial(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^{n}(X_i-\mu)^2$$

Так как даны $\sum X_i$ и $\sum X_i^2$, можно вывести, что $\sum(X_i-\mu)^2 = \sum X_i^2 - \sum 2\mu X_i + \sum \mu^2 = \sum X_i^2 - 2\mu\sum X_i + n\mu^2$.

Из условий первого порядка следует, что ММП-оценка матожидания $\hat{\mu}_{\text{ML}}$ — это выборочное среднее, а дисперсии $\hat{\sigma}^2_{\text{ML}}$ — выборочная дисперсия (без коррекции на одну степень свободы):

$$\hat{\mu}_{\text{ML}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{100}{100} = 1, \quad \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{1}{100}(900 - 2\cdot 1\cdot 100 + 100\cdot 1^2) = \frac{800}{100} = 8$$

(b)

$$I(\theta) = -\left(\frac{\partial^2 \ln}{\partial \theta^2}\right), \quad I(\hat{\theta}) = \begin{pmatrix} \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & \frac{n}{2(\hat{\sigma}^2)^2} \end{pmatrix} = \begin{pmatrix} \frac{100}{8} & 0 \\ 0 & \frac{100}{128} \end{pmatrix}, \quad I^{-1}(\hat{\theta}) = \begin{pmatrix} \frac{2}{25} & 0 \\ 0 & \frac{32}{25} \end{pmatrix}$$

Так как ММП-оценки асимптотически нормальны, то 95%-й доверительный интервал для вектора неизвестных параметров выглядит как

$$\begin{pmatrix} \hat{\mu} \pm z_{\frac{\alpha}{2}}\sqrt{(\hat{\mu})} \\ \hat{\sigma}^2 \pm z_{\frac{\alpha}{2}}\sqrt{(\hat{\mu})} \end{pmatrix} \approx \begin{pmatrix} 1 \pm 1.96\sqrt{\frac{2}{25}} \\ 8 \pm 1.96\sqrt{\frac{32}{25}} \end{pmatrix} \approx \begin{pmatrix} [0.446; 1.554] \\ [5.783; 10.217] \end{pmatrix}$$

(c) Тест Вальда выглядит следующим образом:

$$W = \big(\mathbf{c}(\hat{\theta}) - q\big)'(CI^{-1}(\hat{\theta})C')^{-1}\big(\mathbf{c}(\hat{\theta}) - q\big)\chi^2_r$$

За C обозначено $\frac{\partial \mathbf{c}(\theta)}{\partial \theta}$, за I — информационная матрица Фишера $\left(I(\theta) = -\left(\frac{\partial^2 \ln}{\partial \theta^2}\right)\right)$.

В данном случае нулевая гипотеза $\mathbf{c}(\theta) = q$ записывается как $\mathbf{c}(\theta) = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, поэтому все статистики имеют две степени свободы хи-квадрата. $C = \frac{\partial \mathbf{c}}{\partial \theta} = \begin{pmatrix} \frac{\partial c_1}{\partial \mu} & \frac{\partial c_2}{\partial \mu} \\ \frac{\partial c_1}{\partial \sigma^2} & \frac{\partial c_2}{\partial \sigma^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{c}(\theta) - q = \begin{pmatrix} 1 \\ 8 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \end{pmatrix}$, поэтому расчётная статистика выглядит следующим образом:

$$W = \begin{pmatrix} -1 & 7 \end{pmatrix} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{100}{8} & 0 \\ 0 & \frac{100}{128} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} -1 \\ 7 \end{pmatrix} = \begin{pmatrix} -1 & 7 \end{pmatrix} \begin{pmatrix} \frac{25}{2} & 0 \\ 0 & \frac{25}{32} \end{pmatrix} \begin{pmatrix} -1 \\ 7 \end{pmatrix} = 50.78$$

Тест отношения правдоподобия:

$$LR = -2\big(\ln_{\mathrm{R}} - \ln_{\mathrm{UR}}\big)\chi^2_r$$

$$\begin{aligned} LR &= -2\big(\ln(q) - \ln(\hat{\theta})\big) = \\ &= -2\left(-\frac{n}{2}\ln\sigma_0^2 - \frac{1}{2\sigma_0^2}\left(\sum X_i^2 - 2\mu_0 \sum X_i + n\mu_0^2\right) + \frac{n}{2}\ln\hat{\sigma}^2 + \frac{1}{2\hat{\sigma}^2}\left(\sum X_i^2 - \right.\right. \\ &\quad - 2\hat{\mu}\sum X_i + n\hat{\mu}^2 == -2\left(-\frac{100}{2}\ln 1 - \frac{1}{2}(900 - 2\cdot 2\cdot 100 + 100\cdot 2^2) + \right. \\ &\quad + \frac{100}{2}\ln 8 + \frac{1}{16}(900 - 2\cdot 1\cdot 100 + 100\cdot 1) \approx 592 \end{aligned}$$

Тест множителей Лагранжа:

$$LM = S(\theta_0)' I^{-1}(\theta_0) S(\theta_0)\chi^2_r$$

$$I(\theta_0) = \begin{pmatrix} \frac{n}{\sigma_0^2} & 0 \\ 0 & \frac{n}{2(\sigma_0^2)^2} \end{pmatrix} = \begin{pmatrix} 100 & 0 \\ 0 & 50 \end{pmatrix}, \quad I^{-1}(\theta_0) = \begin{pmatrix} \frac{1}{100} & 0 \\ 0 & \frac{1}{50} \end{pmatrix}$$

$$S(\theta_0) = \begin{pmatrix} \frac{1}{\sigma_0^2}(100 - 100\mu_0) \\ -\frac{100}{2\sigma_0^2} + \frac{1}{2(\sigma_0^2)^2}(900 - 200\mu_0 + 100\mu_0^2) \end{pmatrix} = \begin{pmatrix} -100 \\ 400 \end{pmatrix}$$

$$LM = \begin{pmatrix} -100 & 400 \end{pmatrix} \begin{pmatrix} \frac{1}{100} & 0 \\ 0 & \frac{1}{50} \end{pmatrix} \begin{pmatrix} -100 \\ 400 \end{pmatrix} = 3300$$

Для уровня значимости 5 % критическое значение χ^2_2 равно ≈ 5.99, поэтому во всех трёх тестах гипотеза $H_0\colon \theta = \theta_0$ отвергается.