30 september 2025 - Lec 1

LIKELIHOOD-BASED INFERENCE

We observe a random sample $(y_1, \dots, y_n) = y^{obs}$

We formulate a statistical model $Y \sim p(y;\theta)$ $\theta \in \Theta$

LIKELIHOOD FUNCTION: the function of θ $p(y^{obs};\theta)$, $\theta \in \Theta$, denoted with $L(\theta) = L(\theta; y^{obs})$

LOG-LIKELIHOOD FUNCTION: the logarithmic transformation of $L(\theta)$, denoted with $\ell(\theta) = \log L(\theta)$

MAXIMUM LIKELIHOOD ESTIMATE: if it exists, is the value $\hat\theta$ such that $L(\hat\theta) \geq L(\theta)$ for all $\theta \in \Theta$

(equivalently, $\ell(\hat\theta) \geq \ell(\theta)$)

before observing the sample, we have the MAXIMUM LIKELIHOOD ESTIMATOR $\hat\Theta = \hat\Theta(Y)$ (it is a random variable)

SCORE FUNCTION $\ell_*(\theta) = \dfrac{\partial \ell(\theta)}{\partial \theta}$ first derivative

LIKELIHOOD EQUATION $\ell_*(\theta) = 0$

the MLE is a solution of the likelihood equation

HESSIAN $\ell_{**}(\theta) = \dfrac{\partial^2 \ell(\theta)}{\partial\theta\,\partial\theta^T}$ second derivative

example: $Y \sim Pois(\lambda)$ $p(y;\lambda) = e^{-\lambda}\dfrac{\lambda^y}{y!}$

I observe $(y_1, y_2) = (3, 4)$

Assuming $\underline{y}$ is a random sample, $p((y_1,y_2);\lambda) = p(y_1;\lambda)\,p(y_2;\lambda)$

$$= e^{-\lambda}\frac{\lambda^{y_1}}{y_1!} \cdot e^{-\lambda}\frac{\lambda^{y_2}}{y_2!} = e^{-2\lambda}\frac{\lambda^{y_1+y_2}}{y_1!\,y_2!}$$

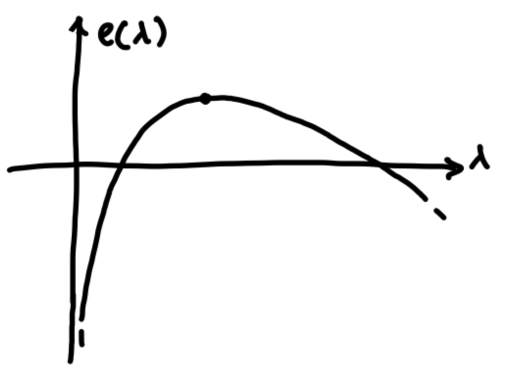
likelihood function $L(\lambda) = L(\lambda;(y_1,y_2)) = e^{-2\lambda}\dfrac{\lambda^7}{3!\,4!}$

it is equivalent to consider $k \cdot L(\lambda)$

$\Rightarrow$ $L(\lambda) = e^{-2\lambda}\lambda^7$

log-likelihood function $\ell(\lambda) = -2\lambda + 7\log\lambda$

maximum likelihood estimate is $\hat\lambda = \underset{\lambda \in (0,+\infty)}{\operatorname{argmax}}\ \ell(\lambda)$



$\ell'(\lambda) = \dfrac{d}{d\lambda}\ell(\lambda) = -2 + \dfrac{7}{\lambda}$

$\ell'(\lambda) = 0 \Rightarrow -2 + \dfrac{7}{\lambda} = 0 \Rightarrow \dfrac{7}{\lambda} = 2 \Rightarrow \hat\lambda = 3.5$

is it a max?

$\ell''(\lambda) = -\dfrac{7}{\lambda^2} < 0$ ok!

Before observing the data $\ell(\lambda;y) = -2\lambda + (y_1+y_2)\log\lambda$ and $\hat\lambda = \dfrac{\left(\sum_{i=1,2} y_i\right)}{2} = \bar{y}$

If I look at $\hat\lambda$ as a function of the r.v. $\hat\Lambda = \hat\Lambda(Y) = \bar{Y} = \frac{1}{2}(Y_1+Y_2)$

we can study the properties of $\hat\Lambda$

$$\mathbb{E}[\hat\Lambda] = \mathbb{E}[\bar Y] = \mathbb{E}\left[\tfrac{1}{2}(Y_1+Y_2)\right] = \tfrac{1}{2}(\mathbb{E}[Y_1] + \mathbb{E}[Y_2]) = \tfrac{1}{2}\cdot 2\lambda = \lambda$$

$$var(\hat\Lambda) = var(\bar Y) = var\left(\tfrac{1}{2}(Y_1+Y_2)\right) \overset{\perp\!\!\!\perp}{=} \tfrac{1}{4}(var(Y_1) + var(Y_2)) = \tfrac{1}{4}\cdot 2\lambda = \tfrac{\lambda}{2}$$

LIKELIHOOD RATIO TEST

It is a general procedure to perform tests on nested models (i.e. the simpler model can be obtained starting from the more complex model through constraints on the parameters)

consider a test where both the null and the alternative hypotheses are simple:

$$\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta = \theta_1 \end{cases}$$

one way to decide between H_0 and H_1 is to compare the likelihood evaluated at θ_0 and θ_1

$$\frac{L(\theta_0; y^{obs})}{L(\theta_1; y^{obs})}$$

If θ_0 is better than θ_1, $L(\theta_0) > L(\theta_1)$. Hence large values of the ratio suggest the acceptance of H_0.

In general, when the hypotheses are not simple

$$\begin{cases} H_0: \theta \in \Theta_0 \\ H_1: \theta \in \Theta \setminus \Theta_0 \end{cases}$$

the likelihood ratio test is based on the test statistic

$$\lambda_{LR} = -2\log\left[\frac{\sup_{\theta\in\Theta_0} L(\theta)}{\sup_{\theta\in\Theta} L(\theta)}\right]$$

$$= -2\left[\ell(\theta_0) - \ell(\hat\theta)\right]$$

The exact distribution has to be determined on a case-by-case basis, however

ASYMPTOTIC DISTRIBUTION (holds for large n)

under some regularity conditions, $\lambda_{LR} \overset{\cdot}{\sim} \chi^2_q$ under H_0

where $q = \dim(\Theta) - \dim(\Theta_0)$