MEM-264 Applied Statistics

Department of Mathematics and Applied Mathematics, University of Crete

Costas Smaragdakis (kesmarag@uoc.gr)

5th Lecture - 22-02-2021

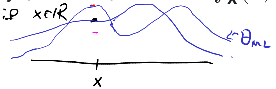
$X_1, \dots, X_n$ iid $\quad x_j \quad f(x_j;\theta)$

Definition : Likelihood function

$\underset{\sim}{X} = (X_1, \dots, X_n)^T$

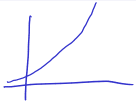
For a realization $\mathbf{x}$ of a random variable $\mathbf{X}$ with a probability density (or mass) function $f_{\mathbf{X}}(\mathbf{x};\theta)$, the likelihood function $\mathcal{L}$ is defined by

$$\mathcal{L}(\theta;\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x};\theta)$$



Definition : Log-likelihood function

$$\ell(\theta;\mathbf{x}) = \log \mathcal{L}(\theta;\mathbf{x})$$



maximum

Definition : Likelihood estimator

An estimator t is called the **maximum likelihood estimator** of θ if

$$t(\mathbf{x}) = \underset{\theta\in\Theta}{\arg\sup}\, \mathcal{L}(\theta;\mathbf{x}) = \underset{\theta\in\Theta}{\arg\sup}\, \ell(\theta;\mathbf{x})$$

- We will denote the maximum likelihood estimator as $\hat{\theta}_{ml}$ or simply $\hat{\theta}$.

(mle)
the maximum likelihood estimator of $g(\theta)$ is given by

$g(\hat{\theta}_{mle})$

where $\hat{\theta}_{mle}$ is the maximum likelihood estimator of θ

Example $\underset{\sim}{X} = (X_1, \dots, X_n)^T$ $\quad \underset{\sim}{x} = (x_1, \dots, x_n)^T$ $\quad \arg\max \log L(\theta; x)$

Let $X_1, \dots, X_n$ be an i.i.d sample and $X_1 \sim \mathcal{N}(\mu, \sigma^2)$. Find the maximum likelihood estimator of $\theta = (\mu, \sigma^2)^T$.

$$f(x_1; \overbrace{\mu, \sigma^2}^{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(x_1-\mu)^2}{2\sigma^2} \right\}$$

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}; \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^2} \right\} = L(\mu, \sigma^2; \underset{\sim}{x})$$

$$\ell(\mu, \sigma^2; x) = -n/2 \left\{ \log 2\pi + \log \sigma^2 \right\} - \sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^2}$$

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2; x) = \sum_{j=1}^{n} \frac{(x_j-\mu)}{\sigma^2} = 0 \Rightarrow \sum_{j=1}^{n} (x_j - \mu) = 0 \Rightarrow \hat{\mu}_{ml} = \frac{\sum x_j}{n} = \bar{x}_n$$

$\sigma^2 = \tau$

$$\frac{\partial}{\partial \tau} \ell(\mu, \tau; x)$$

$$\frac{\partial}{\partial \sigma^2} \ell(\mu, \sigma^2; x) = -n/2 \frac{1}{\sigma^2} + \sum_{j=1}^{n} \frac{(x_j-\mu)^2}{2\sigma^4} = 0 \Rightarrow \hat{\sigma}^2_{ml} = \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x}_n)^2$$

Example

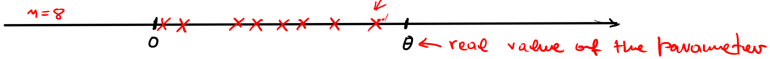
$\theta > 0$

Let $X_1, \dots, X_n$ be an i.i.d sample and $X_1 \sim \mathcal{U}(0, \theta)$. Find the maximum likelihood estimator of θ.

$$f_{X_1}(x_1;\theta) = \begin{cases} \frac{1}{\theta}, & x_1 \in (0,\theta) \\ 0, & \text{otherwise} \end{cases} = \frac{1}{\theta}\, 1\{x \in (0,\theta)\}$$

$$\lambda(\theta; \underset{\sim}{x}) = \frac{1}{\theta^n} \prod_{j=1}^{n} 1\{x_j \in (0,\theta)\} = \frac{1}{\theta^n}\, 1\{\min\{x_1,\dots,x_n\} > 0\}\, 1\{\max\{x_1,\dots,x_n\} < \theta\}$$

$\max\{x_1,\dots,x_n\} = \arg\sup \lambda(\theta; \underset{\sim}{x})$



Binomial distributions

Let $X_1, \dots, X_n$ be an i.i.d sample and $X_1 \sim \text{Bin}(N, \theta)$

$$\sum_{j=1}^{n} X_j \sim \text{Bin}(nN, \theta)$$

Poisson

~~Normal~~ distributions

Let $X_1, \dots, X_n$ be an i.i.d sample and $X_1 \sim \text{Po}(\theta)$

$$\sum_{j=1}^{n} X_j \sim \text{Po}(n\theta)$$

Normal distributions

Let $X_1, \dots, X_n$ be an i.i.d sample and $X_1 \sim \mathcal{N}(\mu, \sigma^2)$

$$\sum_{j=1}^{n} X_j \sim \mathcal{N}(n\mu, n\sigma^2)$$

Exercise

Let $X \sim \text{Bin}(4, \theta)$ and $x = 3$ a single realization of X. Find the maximum likelihood estimator of θ.
Repeat the estimation for the independent observations $\mathbf{x} = (3, 1)^T$. X_1, X_2 iid $X_1 \sim \text{Bin}(4, \theta)$

$$S_n^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - \bar{X}_n)^2$$

$$S_n = \sqrt{S_n^2}$$

Let $X_1, \ldots, X_n$ be an i.i.d sample with $X_1 \sim \mathcal{N}(\mu, \sigma^2)$.

$$\frac{S_n}{\sqrt{n}}$$

We define the random variable T as

$$T = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}$$

$$\bar{X}_n \sim \mathcal{N}\left(\mathbb{E}\{X_1\}, \frac{\mathrm{Var}\{X_1\}}{n}\right)$$

$$\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

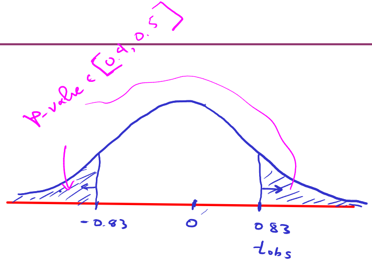
The distribution of T is called t-distribution with $n-1$ degrees of freedom.

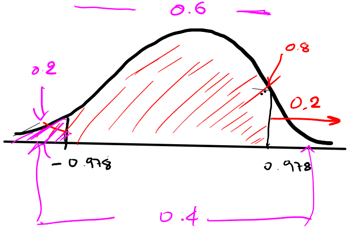
Student's t-distribution

p-value

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

t_{obs}





Let $X_1, \ldots, X_4$ be an i.i.d sample and $X_1 \sim \mathcal{N}(\mu, \sigma^2)$. Accept or reject the null hypothesis.

$$H_0 : \mu = 2 \quad \text{vs} \quad H_1 : \mu \neq 2 \quad \text{observations} : \mathbf{x} = (1.5, 2.4, 1.6, 4.5)^T$$

· Let's assume that H_0 is true.

$X \sim \mathcal{N}(2, \sigma^2)$

unknown

$T = t(\underset{\sim}{x}) = \dfrac{\bar{X}_n - 2}{S_n / \sqrt{n}}$

$t_{obs} = 0.829$

$\bar{x}_n = 2.5$

$T \sim t_{df=3} = t_3$

$s_n \simeq 1.21$

p-value $\in [0.4, 0.5]$

There is no evidence against the null hypothesis !!

Example