Introduction to statistics

author: Alejandro Cáceres date:
autosize: true

Barcelona East School of Engineering
Universitat Politècnica de Catalunya (UPC)

Problems on estimation

Objectives

- Problems on estimation

Tema 7: Summary (Method of moments)

Model	f(x)	E(X)	Parameter estimates from $E(X) = 1/n\sum_i x_i = \bar{x}$ $E(X^2) = 1/n\sum_i x_i^2$, ...
Bernoulli	$p^x(1-p)^{1-x}$	p	$\hat{p} = \bar{x}$
Binomial	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$\hat{p} = \frac{\bar{x}}{n}$
Shifted geometric	$p(1-p)^{x-1}$	$\frac{1}{p}$	$\hat{p} = \frac{1}{\bar{x}}$
Negative Binomial	$\binom{x+r-1}{x}p^r(1-p)^x$	$r\frac{1-p}{p}$	$\hat{p} = \frac{r}{\bar{x}-r}$
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}$	λ	$\hat{\lambda} = \bar{x}$
Exponential	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\hat{\lambda} = \frac{1}{\bar{x}}$
Normal	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	$\hat{\mu} = \bar{x}$,$\hat{\sigma}^2 = \frac{1}{n}\sum_i x_i^2 - \bar{x}^2$

Tema 7: Summary (Maximum likelihood)

- Imagine we make n observations and obtain the values $(x_1, \ldots x_n)$ is

The likelihood function, the probability of having observed $(x_1, \ldots x_n)$ is
$L(\theta) = \Pi_{i=1..n} f(x_i; \theta)$

We can take the log of L,
$\ln L(\theta) = \sum_i \ln(f(x_i;\theta))$

this is called the **log-likelihood** function

Computing the maximum of $\ln L(\theta)$ gives us an estimate for θ which we call $\hat{\theta}$

Problem 1

Consider:

- $P(X=0) = 1/2$
- $P(X=1) = a$
- $P(X=-1) = 1/2 - a$
- $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ for $a \in (0, 1/2)$

 a. for $T = \frac{\bar{X}}{2} + \frac{1}{4}$ compute $E(T)$, $V(T)$

$E(T) = \frac{E(\bar{X})}{2} + \frac{1}{4}$

and

$E(\bar{X}) = E(X) = \sum_{x=-1,0,1} xP(X=x)$
$= -1 * P(X=-1) + 0 * P(X=0) + 1 * P(X=1) = 2a - 1/2$

then

$E(T) = a - 1/4 + 1/4 = a$ and thus $E(T)$ is an **unbiased** estimator of a

$V(T) = V(\frac{\bar{X}}{2} + \frac{1}{4}) = \frac{1}{4}V(\bar{X}) = \frac{1}{4}\frac{V(X)}{n}$

so we need to find $V(X)$

Remember: $V(X) = E(X^2) - E(X)^2$ we miss $E(X^2)$

$E(X^2) = \sum_{x=-1,0,1} x^2 P(X=x) =$
$= (-1)^2 * P(X=-1) + 0^2 * P(X=0) + 1^2 * P(X=1) = a + \frac{1}{2} - a = \frac{1}{2}$

Then

$V(X) = E(X^2) - E(X)^2$
$= \frac{1}{2} - (2a - \frac{1}{2})^2 = \frac{1}{4} + 2a - 4a^2$

putting everything together

$V(T) = \frac{1}{4}\frac{V(X)}{n} = \frac{1/4 + 2a - 4a^2}{4n}$

since $V(T) = \sigma_T \to 0$, when $n \to 0$ then T is a **consistent** estimator

Problem 2

Consider:

- $E(\bar{X}) = E(X) = \mu$ then $\bar{X}$ is an **unbiased** estimator of $E(X) = \mu$

 a. is $E(\bar{X})$ an **unbiased** estimator of $E(X)^2 = \mu^2$

compute: $E(\bar{X}^2)$ (second moment about the origin)

Remember: $V(\bar{X}) = E(\bar{X}^2) - E(\bar{X})^2$

then $E(\bar{X}^2) = E(\bar{X})^2 + V(\bar{X}) = E(X)^2 + \frac{V(X)}{n} = \mu^2 + \frac{\sigma_X^2}{n}$

as $E(\bar{X}^2) \neq \mu^2$ then $E(\bar{X}^2)$ is a **biased** estimator of μ^2

Problem 3

Consider:

$$f(x) = \begin{cases} (1+\theta)x^\theta, & \text{if } x \in (0,1) \\ 0, & otherwise \end{cases}$$

a. compute $E(X)$

$E(X) = \int_0^1 x(1+\theta)x^\theta dx = \int_0^1 (1+\theta)x^{1+\theta} dx$

$= \frac{(1+\theta)x^{2+\theta}}{2+\theta}\Big|_0^1 = \frac{1+\theta}{2+\theta}$

b. compute $\hat{\theta}$ using the method of moments

$E(X) = \mu_1' = \frac{1}{n}\sum_i x_i = \bar{x}$

$\frac{1+\hat{\theta}}{2+\hat{\theta}} = \bar{x}$

solving for $\hat{\theta}$

$\hat{\theta} = \frac{1}{1-\bar{x}} - 2$

c. if the result of a random sample is $x_1 = 0.92; x_2 = 0.79; x_3 = 0.90; x_4 = 0.65; x_5 = 0.86.$ Compute $\hat{\theta}$

$\bar{x} = \frac{0.92+0.79+0.90+0.65+0.86}{5} = 0.824$

then

$\hat{\theta} = \frac{1}{1-\bar{x}} - 2 = \frac{1}{1-0.824} - 2 = 3.6818$

Problem 4

Consider:

$$f(x) = \begin{cases} \frac{x}{\theta}e^{\frac{-x^2}{2\theta}}, & \text{if } x > 0 \\ 0, & otherwise \end{cases}$$

a. Compute $\hat{\theta}$ by maximum likelihood

If we have a set of observations $(x_1, \dots x_n)$, the probability of having observed those numbers is given by the likelihood function:

$L(\theta) = \Pi_{i=1..n}\frac{x}{\theta}e^{\frac{-x_i^2}{2\theta}}$

$= x_1 * \dots * x_n \theta^{-n} e^{\sum_i \frac{-x_i^2}{2\theta}}$

the log likelihood is:

$\ln L(\theta) = \sum_i \ln(x_i) - n\ln(\theta) - \sum_i \frac{x_i^2}{2\theta}$

deriving with respect to θ then and equalling to zero in $\hat{\theta}$

$\frac{d \ln L(\theta)}{d\theta}\Big|_{\hat{\theta}} = -\frac{n}{\hat{\theta}} + \sum_i \frac{x_i^2}{2\hat{\theta}^2} = 0$

solving for $\hat{\theta}$

$\hat{\theta} = \frac{1}{2n}\sum_i x_i^2$

b. for a random experiment of 4 random sample with values:

$x_1 = 16.88; x_2 = 10.23; x_3 = 4.59; x_4 = 6.66; x_5 = 13.68$

compute the estimate of $\hat{\theta}$

since: $\hat{\theta} = \frac{1}{2n}\sum_i x_i^2$

then

$\hat{\theta} = \frac{1}{2*5}(16.88^2 + 10.23^2 + 4.59^2 + 6.66^2 + 13.68^2) = 64.21534$

Problem 5

Consider:

$$f(t) = \begin{cases} \lambda e^{-\lambda(t-\tau)}, & \text{if } t \geq \tau \\ 0, & otherwise \end{cases}$$

a. For random sample: $T_1 \ldots . T_n$ and τ is known then compute $\hat{\lambda}$ by maximum likelihood

If we have a set of observations $(t_1, \ldots t_n)$, the probability of having observed those numbers is given by the likelihood function:

$L(\lambda) = \Pi_{i=1..n}\lambda e^{-\lambda(t_i-\tau)}$

the log likelihood is:

$\ln L(\lambda) = \sum_i \ln(\lambda) + \sum_i \ln(e^{-\lambda(t_i-\tau)}) = n\ln(\lambda) - \lambda\sum_i t_i + n\lambda\tau$

deriving with respect to λ then and equalling to zero in $\hat{\lambda}$

$\frac{d \ln L(\lambda)}{d\lambda}\Big|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}} + n\tau - \sum_i t_i = 0$

solving for $\hat{\lambda}$

$\hat{\lambda} = \frac{n}{\sum_i t_i - n\tau} = \frac{1}{\bar{t}-\tau}$

a. For random sample: $T_1 \ldots . T_n$ and τ is known then compute $\hat{\lambda}$ by them method of moments

For the method of moments we must have

$E(T) = \bar{t}$

since we are given $E(T) = \tau + \frac{1}{\lambda}$; and because $E(Z) = E(T-\tau) = \frac{1}{\lambda}$ for an exponential probability function then

$\hat{\lambda} = \frac{1}{\bar{t} - \tau}$, the same estimate as maximum likelihood

c. For random sample: $T_1 \ldots T_n$ and λ is known then compute $\hat{\tau}$ by them method of moments

Again

$E(T) = \bar{t}$

but this time it gives us the equation for the estimate of τ

$\hat{\tau} + \frac{1}{\lambda} = \bar{t}$ solving for $\hat{\tau}$ then

$\hat{\tau} = \bar{t} - \frac{1}{\lambda}$

Problem 6

Consider:

$$f(x) = \begin{cases} 2\alpha x e^{-\alpha x^2}, & \text{if } x \geq 0 \\ 0, & otherwise \end{cases}$$

a. Compute $\hat{\alpha}$ by maximum likelihood

Remember problem 10

$$f(x) = \begin{cases} \frac{x}{\theta} e^{\frac{-x^2}{2\theta}}, & \text{if } x > 0 \\ 0, & otherwise \end{cases}$$

then $2\alpha = \frac{1}{\theta}$, we can re-parametrize and by maximum likelihood we had found

$\hat{\theta} = \frac{1}{2n} \sum_i x_i^2$ then $\hat{\alpha} = \frac{n}{\sum_i x_i^2}$

Problem 7

Consider the exponential density re-parametrized by

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}}, & \text{if } x \geq 0 \\ 0, & otherwise \end{cases}$$

then $E(X) = \beta$ and $V(X) = \beta^2$

a. compute $\hat{\beta}$ by maximum likelihood

then

$L(\beta) = \Pi_{i=1..n} \frac{1}{\beta} e^{-\frac{x_i}{\beta}}$

$\ln L(\frac{1}{\beta}) = \sum_i \ln(\frac{1}{\beta}) + \sum_i \ln(e^{-\frac{x_i}{\beta}}) = -n \ln(\beta) - \frac{1}{\beta} \sum_i x_i$

deriving with respect to β then and equalling to zero in $\hat{\beta}$

$\frac{d \ln L(\beta)}{d\beta}\Big|_{\hat{\beta}} = -\frac{n}{\hat{\beta}} + \frac{1}{\hat{\beta}^2} \sum_i x_i = 0$

solving for $\hat{\beta}$

$\hat{\beta} = \bar{x}$ which for the exponential function (see table) is $\hat{\beta} = \hat{\lambda}^{-1}$

c. compute the $E(B)$ and $V(B)$, where B is the estimator (random variable) whose values give us the estimate $\hat{\beta}$

Then

$B = \bar{X}$

and

$E(B) = E(\bar{X}) = E(X) = \beta$ therefore B is an **unbiased** estimator of β, or $\hat{\beta}$ is **unbiased** estimate of β

$V(B) = V(\bar{X}) = \frac{V(X)}{n}$ and then when $n \to \infty$, $V(B) \to 0$ and therefore we say that estimator is **consistent**

Problem 8

Consider

$$f(x) = \begin{cases} \frac{2(\theta - x)}{\theta^2}, & \text{if } x \in [0, \theta] \\ 0, & otherwise \end{cases}$$

a. For random sample: $X_1 \dots . X_n$ and compute $\hat{\theta}$ by the method of moments

The method of moments proposes to find the estimate $\hat{\theta}$ from the equation

$E(X) = \bar{x}$ therefore we need to find $E(X)$

$E(X) = \int_0^\theta x \frac{2(\theta - x)}{\theta^2} dx = \frac{2}{\theta^2} (\frac{\theta x^2}{2} - \frac{x^3}{3}) \Big|_0^\theta = \frac{\theta}{3}$

Therefore, we have

$E(X) = \frac{\hat{\theta}}{3} = \bar{x}$ and solving for $\hat{\theta}$

then

$\hat{\theta} = 3\bar{x}$