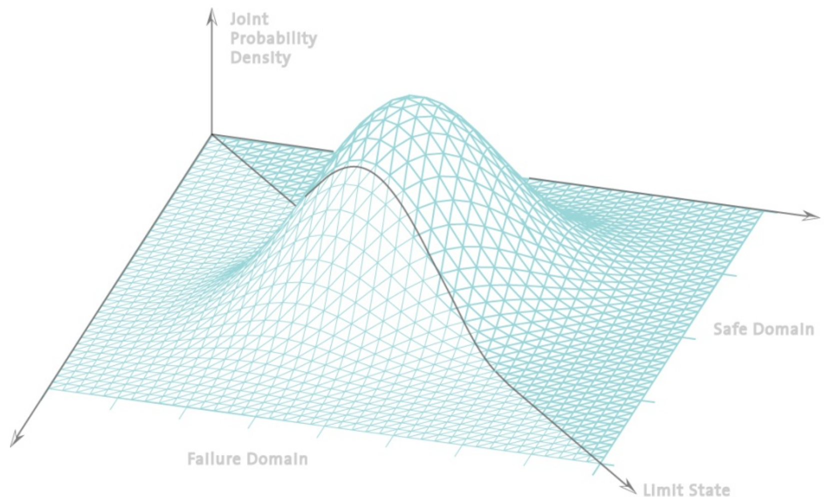


Exercise Tutorial 10

Statistics and Probability Theory

Prof. Dr. Michael Havbro Faber
Swiss Federal Institute of Technology Zurich
ETHZ

Exercise 8.7 (Group Exercise)

To rebuild a car park, the arrival times of cars were measured. The time interval between arriving cars are shown in the table.

a. Check graphically, if the time interval of car arrivals can be represented by an Exponential distribution.

b. Calculate the mean value of the time interval of car arrivals. Under the assumption that the time interval is Exponential distributed, determine the mean value of the time interval from the graph produced in part (a.).

i	Time interval (seconds)	*i*	Time interval (seconds)
1	1.52	7	30.4
2	6.84	8	30.4
3	9.12	9	34.2
4	10.64	10	60.8
5	15.2	11	78.28
6	21.28	12	95.76

a. Check graphically, if the time interval of car arrivals can be represented by an exponential distribution.

i	$i/(n+1)$	Time interval (seconds)
1	1/13	1.52
2	2/13	6.84
3	3/13	9.12
4	4/13	10.64
5	5/13	15.2
6	6/13	21.28
7	7/13	30.4
8	8/13	30.4
9	9/13	34.2
10	10/13	60.8
11	11/13	78.28
12	12/13	95.76

Probability paper

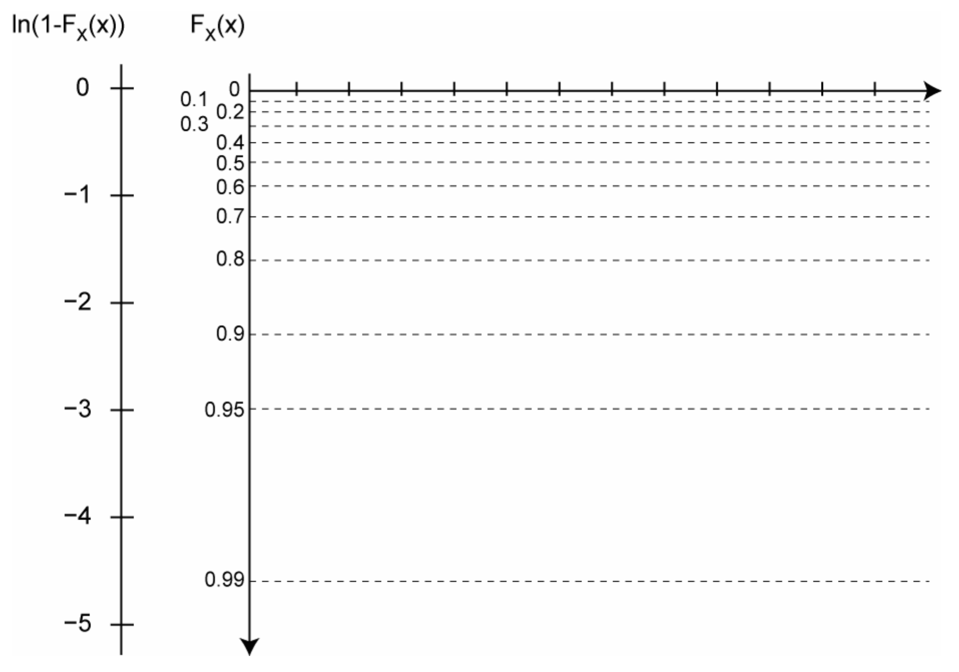
- Exponential distribution

$$F_X(x) = 1 - e^{-\lambda x} \longrightarrow \underbrace{\ln\left(1 - F_X(x)\right)}_{= y} = -\lambda x$$

Probability paper

- Exponential distribution

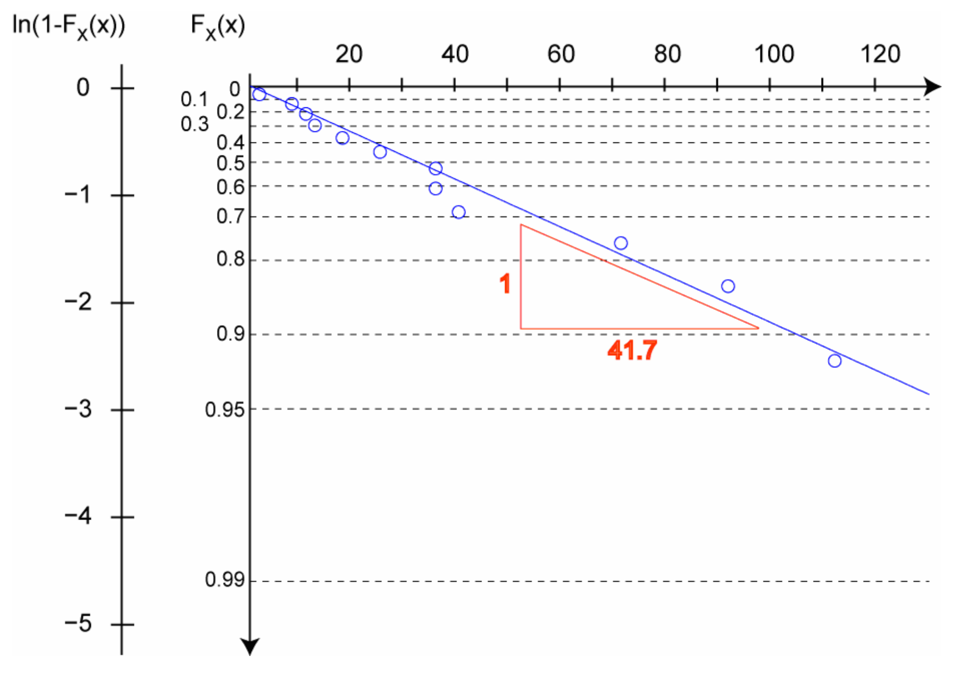
$$F_X(x) = 1 - e^{-\lambda x} \longrightarrow \underbrace{\ln\left(1 - F_X(x)\right)}_{= y} = -\lambda x$$



Probability paper

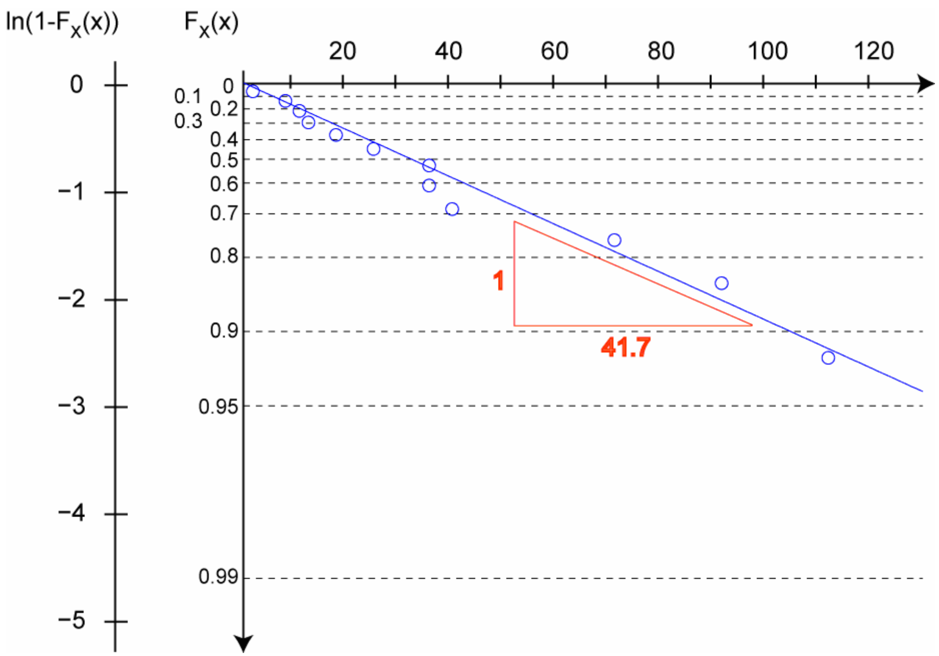
- Exponential distribution

$$F_X(x) = 1 - e^{-\lambda x} \longrightarrow \underbrace{\ln\left(1 - F_X(x)\right)}_{=y} = -\lambda x$$



b. Calculate the mean value of the time interval of car arrivals. Under the assumption that the time interval is Exponential distributed, determine the mean value of the time interval from the graph produced in part (a.).

$$\underbrace{\ln\left(1-F_X(x)\right)}_{=y} = -\lambda x$$



From the formula it turns out that the tangent of the straight line is equal to -λ .

$$-\lambda = \frac{-1}{41.7} \Leftrightarrow \lambda = \frac{1}{41.7}$$

The mean value is thus equal to 41.7.

Probability paper

- Exponential distribution

$$F_X(x) = 1 - e^{-\lambda x} \longrightarrow \underbrace{\ln\left(1 - F_X(x)\right)}_{=y} = -\lambda x$$

- Normal distribution

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \longrightarrow \underbrace{\Phi^{-1}\left(F_X(x)\right)}_{=y} = \frac{x-\mu}{\sigma} = \frac{1}{\sigma}x - \frac{\mu}{\sigma}$$

- Gumbel distribution

$$F_X(x) = \exp\left(-\exp(-a(x-b))\right) \longrightarrow \underbrace{-\ln\left(-\ln\left(F_X(x)\right)\right)}_{=y} = a(x-b) = ax - ab$$

Estimation of parameters

There are several ways to estimate the parameters of distributions from observed data. For instance,

1. Method of moments

2. Maximum likelihood method

Let’s start with the method of moments.

The method of moments

Moment in "probability" sense

Moments are defined as:

$$m_k = \int x^k f_X(x)dx \qquad \text{for continuous case}$$

$$m_k = \sum_j x_j{}^k p(x_j) \qquad \text{for discrete case}$$

The parameters of the distribution are estimated so that the moments match.

Calculate directly from assumed distribution

$$\boxed{m_k = \int x^k f_X(x)dx} = E[X^k] = \boxed{\frac{1}{n}\sum_{i=1}^{n} \hat{x}_i{}^k}$$

$$\boxed{m_k = \sum_j x_j{}^k p(x_j)} = E[X^k] = \boxed{\frac{1}{n}\sum_{i=1}^{n} \hat{x}_i{}^k}$$

Calculate from obtained sample

The method of moments

A simple example

We want to estimate the parameter of the Exponential distribution.

The probability density function is given as $f_X(x \mid \lambda) = \lambda \exp(-\lambda x)$.

The first moment m_1 is

$$\left.\begin{aligned} m_1 &= \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda} \\ \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i &= 0.48 \end{aligned}\right\} \quad \lambda = 1/0.48 = 2.08$$

Realization

1	0.02
2	0.08
3	0.14
4	0.21
5	0.30
6	0.40
7	0.52
8	0.69
9	0.94
10	1.50

The maximum likelihood method

Another method is the maximum likelihood method
Let’s start with an example.

There is a coin. We are interested in the probability that the tail comes out, p.
We tossed the coin 20 times and observed 6 tails and 14 heads. How can we estimate the parameter p?

Note that p is the parameter of Bernoulli trials, and is the probability of success.

The maximum likelihood method

We can never know the truth, but…

If p=0.1, what is the probability that 6 tails and 14 heads come out?

$$L(0.1) = \binom{20}{6} \times 0.1^{6} \times 0.9^{14} = 0.0089 \leftarrow P[6\,tails\,and\,14\,heads \mid p = 0.1]$$

If p=0.3, what is the probability that 6 tails and 14 heads come out?

$$L(0.3) = \binom{20}{6} \times 0.3^{6} \times 0.7^{14} = 0.192 \leftarrow P[6\,tails\,and\,14\,heads \mid p = 0.3]$$

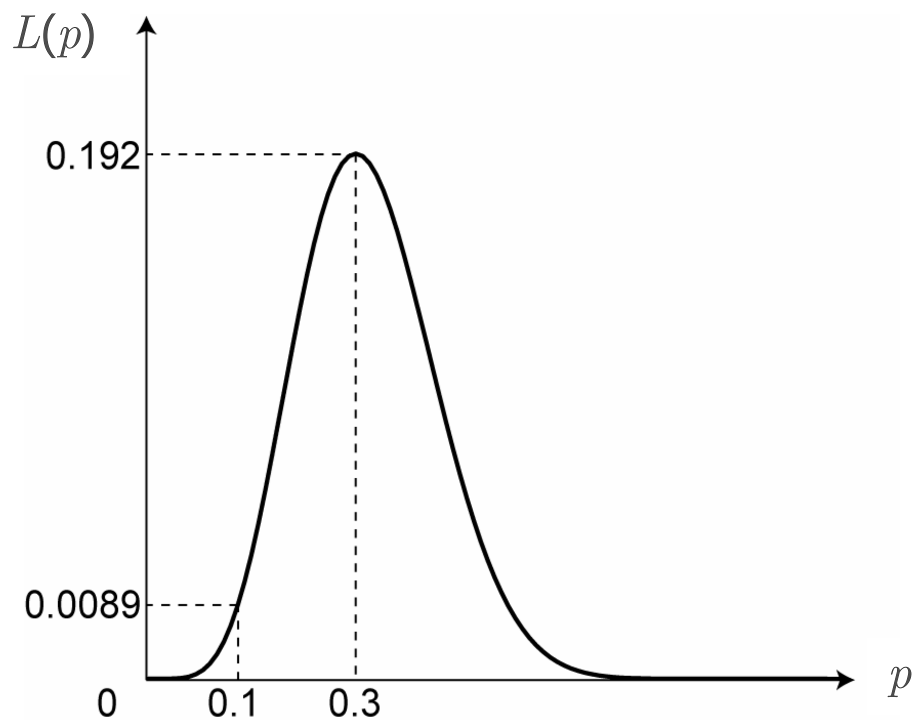
How should we interpret these numbers?

Since the probability that 6 tails and 14 heads come out given p=0.1 is much smaller than that given p=0.3, we may believe that the outcome comes from "p=0.3" rather than "p=0.1".

These are not a probability of "p=0.3" nor "p=0.1".

The maximum likelihood method

We can calculate $L(p)$ for all p ($0 \leq p \leq 1$) and obtain the following figure.



"$L(p)$ is called "likelihood function.""

The maximum likelihood method

The likelihood function

Let X be a random variable whose probability density function is $f_X(x \mid \theta)$.

When you make n trials, the joint probability density of outcomes is

$$f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f_X(x_i \mid \theta)$$ ← before observation

This is the probability density for any given θ.

However, we can regard this formulation as the function of θ
after the observation of outcomes $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$.

In this interpretation of the above formulation, we call it a “likelihood function”

$$L(\theta \mid \hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) = \prod_{i=1}^{n} f_X(\hat{x}_i \mid \theta)$$ ← after observation (θ is not known)

The maximum likelihood method

The maximum likelihood method (MLM)

The maximum likelihood estimator θ is obtained as the value of θ that maximizes the likelihood function $L(\theta)$.

This is equivalent to maximize the "log-likelihood function" defined as:

$$l(\theta \mid \hat{\mathbf{x}}) = \ln L(\theta \mid \hat{\mathbf{x}}) = \ln \prod_{i=1}^{n} f_X(\hat{x}_i \mid \theta) = \sum_{i=1}^{n} \ln f_X(\hat{x}_i \mid \theta)$$

Coin example again!

To be maximized is $$L(p) = \binom{20}{6} p^6 (1-p)^{14}$$

The log-likelihood function is $$l(p) = \ln L(p) = \ln \binom{20}{6} + 6 \ln p + 14 \ln(1-p)$$

The MLM estimate is obtained as: $$\frac{\partial l}{\partial p} = \frac{6}{p} - \frac{14}{1-p} = 0 \quad \Rightarrow \quad p = 0.3$$

Exercise 9.1

In order to model the concrete compressive strength of a certain concrete production, 20 samples were measured and the result of measurements is shown in the table. It is assumed that the population of the samples follows the Normal distribution $N(\mu,\sigma^2)$.

1) Describe the likelihood function.
2) Estimate the unknown parameters (μ,σ) with the maximum likelihood method.
3) Estimate the unknown parameters with the method of moments.

No. of sample	Compressive strength [MPa]	No. of sample	Compressive strength [MPa]
1	24.4	11	33.3
2	27.6	12	33.5
3	27.8	13	34.1
4	27.9	14	34.6
5	28.5	15	35.8
6	30.1	16	35.9
7	30.3	17	36.8
8	31.7	18	37.1
9	32.2	19	39.2
10	32.8	20	39.7

1) The likelihood function is

$$L(\mu,\sigma\,|\,\hat{\mathbf{x}}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]$$

to be maximized. Instead of the likelihood function we adopt the log-likelihood function.

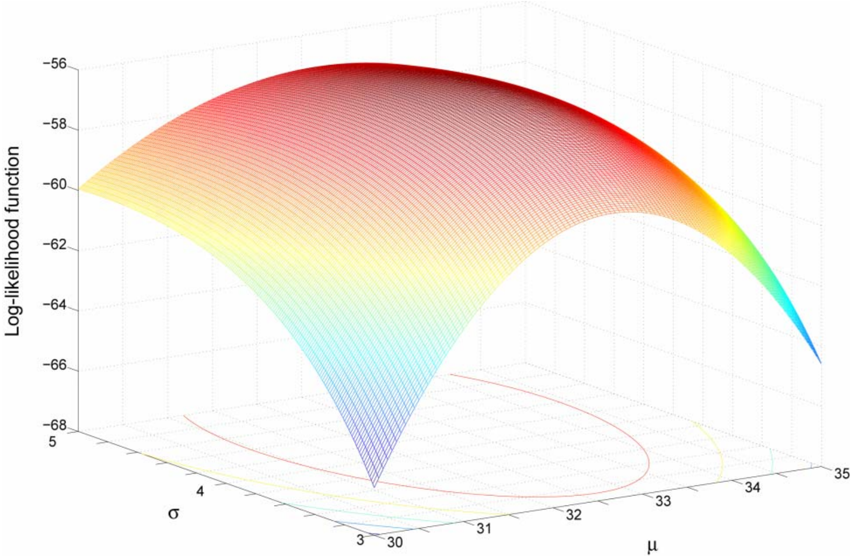
$$l = \ln L = \sum_{i=1}^{n} \ln\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]\right]$$

$$= -n\ln(\sqrt{2\pi}) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2$$

2) The maximum likelihood estimates are obtained as the parameters at which the likelihood function (accordingly the log-likelihood function) is maximized.

$$l = \ln L = \sum_{i=1}^{n} \ln\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(\hat{x}_i-\mu)^2}{2\sigma^2}\right]\right]$$

$$= -n\ln(\sqrt{2\pi}) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2$$



$$\frac{\partial l}{\partial \mu} = \frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(\hat{x}_i-\mu) = 0$$

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2 = 0$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i-\mu)^2$$

$$\longrightarrow \quad \mu = 32.67 \qquad \sigma = 4.04$$

3) With the method of moments, the parameters are estimated as:

$$m_1 = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] dx = \mu$$

$$m_2 = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] dx = \sigma^2 + \mu^2$$

$$\left. \begin{aligned} \mu &= m_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i \\ \sigma^2 &= m_2 - m_1^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} \hat{x}_i \right)^2 \end{aligned} \right\} \quad \begin{aligned} \mu &= 32.67 \\ \sigma &= 4.04 \end{aligned}$$

Exercise 9.2

What happens if the Exponential distribution is assumed instead of the Normal distribution in exercise 9.1?

a. Estimate the unknown parameters of the Exponential distribution with the maximum likelihood method.

b. Draw the cumulative distribution functions together with observed cumulative distribution.

The likelihood function is

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda \hat{x}_i}$$

The corresponding log-likelihood function is

$$l = \ln L = \sum_{i=1}^{n} \ln\left(\lambda e^{-\lambda \hat{x}_i}\right)$$

$$= n \ln \lambda - \lambda \sum_{i=1}^{n} \hat{x}_i$$

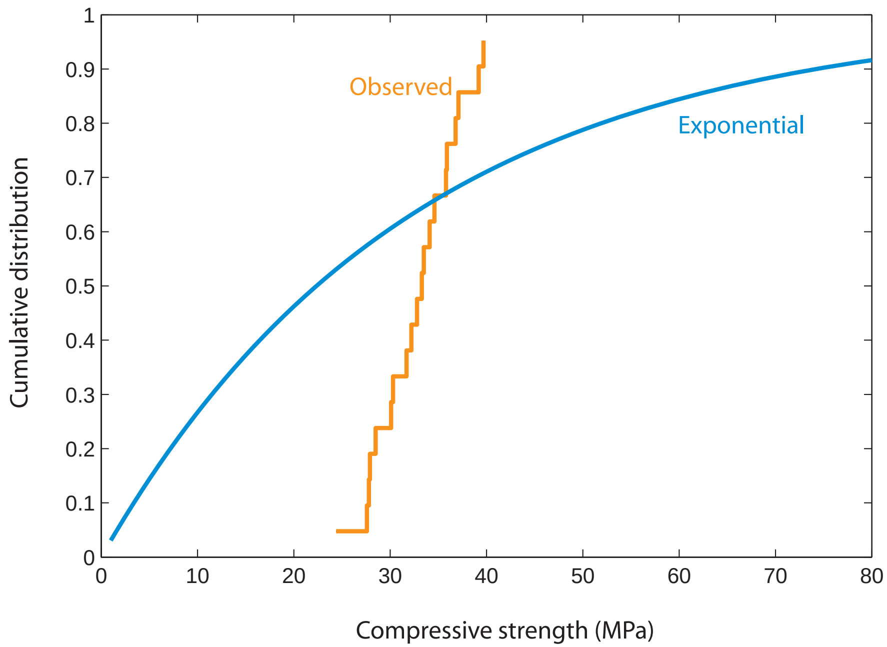
The maximum likelihood method estimate is obtained from

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \hat{x}_i = 0$$

It is

$$\lambda = \frac{n}{\sum_{i=1}^{n} \hat{x}_i} \longrightarrow \lambda = 0.031$$

Do you think that the Exponential distribution with the maximum likelihood estimate models the population of the sample nicely???



How to judge that? We will see this the next week!

Exercise 9.3 (Group exercise)

It is known that the data shown in the table are the realizations of a random variable X characterized by the cumulative distribution function,
$F_X(x) = x^{\alpha}, 0 \leq x \leq 1$ with unknown parameter α.
Estimate the parameter α in the following methods.

a. Estimate the unknown parameter α with the *method of moments*.

b. Estimate the unknown parameter α with the *maximum likelihood method*.

c. Draw the cumulative distribution function with the estimated parameter in (b.) and the observed cumulative distribution.

Exercise 9.3 (Group exercise)

a) Method of moments

"sample first moment = analytical first moment"

The sample first moment is calculated by

$$\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i$$

The analytical first moment is calculated by

$$\int x f(x\,|\,\alpha)\,dx \longleftarrow f(x\,|\,\alpha) = \frac{dF_X(x\,|\,\alpha)}{dx} = \frac{dx^{\alpha}}{dx}$$

Equate and solve

$$\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i = \int x f(x\,|\,\alpha)\,dx$$

Exercise 9.3 (Group exercise)

b) Maximum likelihood method

The likelihood function is described as:

$$L(\alpha) = \prod_{i=1}^{n} f(\hat{x}_i \mid \alpha) \longleftarrow f(x \mid \alpha) = \frac{dF_X(x \mid \alpha)}{dx} = \frac{dx^{\alpha}}{dx}$$

Maximize the log likelihood function with respect to α:

$$l(\alpha) = \sum_{i=1}^{n} \ln(f(\hat{x}_i \mid \alpha))$$