PAGE No.
DATE / /

MLE for Bernoulli.

pmf : $f(x_i) = p^{x_i}(1-p)^{1-x_i}$ — we have to find MLE of

$$L = \prod_{i=1}^{n} f(x_i|\theta)$$

$$L = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$L = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \dots p^{x_n}(1-p)^{1-x_n}$$

$$\log L = \log\left[p^{x_1}(1-p)^{1-x_1}\right] + \log\left[p^{x_2}(1-p)^{1-x_2}\right] \dots + \log\left[p^{x_n}(1-p)^{1-x_n}\right]$$

$$\log L = \log p^{x_1} + \log(1-p)^{1-x_1} + \log p^{x_2} + \log(1-p)^{1-x_2} + \dots \log p^{x_n} + \log(1-p)^{1-x_n}$$

$$\log L = x_1 \log p + (1-x_1)\log(1-p) + x_2 \log p + (1-x_2)\log(1-p) + \dots x_n \log p + (1-x_n)\log(1-p)$$

$$\log L = \log p\,(x_1 + x_2 + \dots x_n) + \log(1-p)\left[(1-x_1) + (1-x_2) + \dots (1-x_n)\right]$$

$$\log L = (\Sigma x_i)(\log p) + (n - \Sigma x_i)\log(1-p)$$

$$\frac{d \log L}{dp} = (\Sigma x_i)\left(\frac{1}{p}\right) + (n - \Sigma x_i)\left(\frac{1}{1-p}\right)(-1)$$

$$\frac{d \log L}{dp} = \frac{1}{p}\Sigma x_i - \left(\frac{1}{1-p}\right)(n - \Sigma x_i)$$

$$\frac{d \log L}{dp} = 0 \Rightarrow \frac{\Sigma x_i}{p} = \frac{n - \Sigma x_i}{1-p}$$

$$(1-p)\Sigma x_i = pn - p\Sigma x_i$$

$$\Sigma x_i - \cancel{p\Sigma x_i} = pn - \cancel{p\Sigma x_i}$$

$$\boxed{p = \frac{\Sigma x_i}{n}}$$

PAGE No.
DATE / /

MLE for binomial.

$$p(x_i) = {}^nC_{x_i} \cdot p^{x_i} (1-p)^{n-x_i} \qquad x_i = 0, 1, 2, \dots n.$$

$$L = \prod_{i=1}^{n} p(x_i|\theta) = \prod_{i=1}^{n} ({}^nC_{x_i})\, p^{x_i} (1-p)^{n-x_i}$$

$$\log L = \sum_{i=1}^{n} \left[\log {}^nC_{x_i} + \log p^{x_i} + \log (1-p)^{n-x_i}\right]$$

$$\log L = \sum_{i=1}^{n} \log {}^nC_{x_i} + \sum_{i=1}^{n} x_i \cdot \log p + \sum_{i=1}^{n} (n-x_i) \log(1-p)$$

$$\log L = \sum_{i=1}^{n} \log {}^nC_{x_i} + \log p \sum_{i=1}^{n} x_i + \log(1-p) \cdot \sum_{i=1}^{n} (n-x_i)$$

$$\frac{d\log L}{dp} = 0 + \frac{1}{p}\sum_{i=1}^{n} x_i + \frac{-1}{(1-p)} \sum_{i=1}^{n} n-x_i = 0$$

$$\Rightarrow \frac{1}{p}\sum_{i=1}^{n} x_i = \frac{1}{1-p} \sum_{i=1}^{n} n-x_i$$

$$\Rightarrow \frac{1}{p}\sum_{i=1}^{n} x_i = \frac{1}{1-p}(n^2 - \Sigma x_i)$$

$$(1-p)\sum_{i=1}^{n} x_i = p(n^2 - \Sigma x_i)$$

$$\sum_{i=1}^{n} x_i - \cancel{p\sum_{i=1}^{n} x_i} = pn^2 - \cancel{p\Sigma x_i}$$

$$\Sigma n - x_i$$
$$= \Sigma n - \Sigma x_i$$
$$= n\Sigma 1 - \Sigma x_i$$
$$= n(n) - \Sigma x_i$$
$$= n^2 - \Sigma x_i$$

$$\boxed{p = \frac{\Sigma x_i}{n^2}} \rightarrow \text{mle of } p.$$

k = no. of success (size)
n = sample size
p = prob. of success in one try

```
from scipy.stats import binom
```

→ binom.pmf (k, n, p) gives probability of exactly 'k' occurances.
→ binom.cdf (k, n, p) gives prob. of ≤ k occurances.
→ binom.sf (k, n, p) gives prob. of > k

MLE for Poisson.

$$f(x_i) = \frac{e^{-\lambda} \cdot (\lambda)^{x_i}}{x_i!} \qquad i = 0, 1, 2, \ldots \qquad \text{MLE of } \lambda.$$

$$L = \prod_{i=1}^{n} f(x_i \mid \theta)$$

$$L = \prod_{i=1}^{n} \frac{e^{-\lambda} \cdot \lambda^{x_i}}{(x_i)!}$$

$$L = \frac{e^{-\lambda} \cdot \lambda^{x_1}}{x_1!} * \frac{e^{-\lambda} \cdot \lambda^{x_2}}{x_2!} * \cdots \frac{e^{-\lambda} \cdot \lambda^{x_n}}{x_n!}$$

$$\cancel{L = e^{-\lambda}\left[\frac{\lambda^{x_1}}{x_1!}\right]} \times$$

$$\log L = \log\left(\frac{e^{-\lambda} \cdot \lambda^{x_1}}{x_1!}\right) + \log\left(\frac{e^{-\lambda} \cdot \lambda^{x_2}}{x_2!}\right) + \cdots \log\left(\frac{e^{-\lambda} \cdot}{x_n}\right.$$

$$\log L = \left(\log e^{-\lambda} + \log \lambda^{x_1} - \log x_i!\right) + \cdots \left(\log e^{-\lambda} + \log \lambda^{x_n} - \ldots\right.$$

$$\log L = -n\lambda + (\log \lambda)(x_1 + \ldots x_n) - \Sigma \log x_i!$$

$$\log L = -n\lambda + \log \lambda \; \Sigma x_i - \Sigma \log (x_i)!$$

$$\frac{d \log L}{d\lambda} = -n + \frac{1}{\lambda} \Sigma x_i - 0$$

$$\frac{d \log L}{d\lambda} = 0 \Rightarrow \frac{1}{\lambda} \Sigma x_i = n \Rightarrow \boxed{\lambda = \frac{\Sigma x_i}{n}}$$

from scipy.stats import poisson — lambda = exp... per ti...

→ poisson.pmf(k, lam) exactly k occurance

→ poisson.cdf(k, lam) prob. of $\leq k$ — k = no of occu...

→ poisson.sf(k, lam) prob. of $> k$

norm.cdf → calculates the prob. for a given normal dist. value
norm.ppf → calculates norm dist value for which a given prob. is the required val.

norm.pdf → gives prob. density value.
norm.ppf → z crit. value is returned.
if $\alpha = 5\%$ for 2 tailed;
norm.ppf(0.025) → 1.96

~~N = 10~~
~~mu = 2~~
~~sigma = 2~~
~~x = np.random.normal.~~

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm
from scipy.optimize import minimize
```

normal distribution

```
mu = 2
sigma = 3
x1 = 7
x2 = -1
```
} assume.

```
l1 = norm.pdf(x1, loc=mu, Scale=Sigma)
l2 = norm.pdf(x2, loc=mu, Scale=sigma)
print(l1) → likelihood of x1
print(l2) → likelihood of x2
```
} i.e. likelihood of 2 individual events

```
print(l1 * l2) → combined likelihood.
```

now, we are going to know the points $x_1, \ldots x_n$ but we do not know mu or sigma.

```
np.random.seed(123)
N = 100
x = np.random.binomial(n, p, size)
```
n → no. of trials, p → prob., size → shape returned array

```
x = np.random.normal(loc=mu, scale=sigma, size=(N,))
print(np.mean(x))
print(np.std(x))
```
} → we find mean & sigma for this distribution by using code instead of the functions.

to find max, we do negative of minimum.

parameters mu & sigma
random nos we generate

```
def log-likelihood (p, x):
    p = [0, 1] → assume mean=0, stdev=1 → can be changed later
    mu = p[0]
    sigma = p[1]
    L = np.sum(np.log(norm.pdf(x, loc=mu, scale=sigma)))
    return -L
```

log likelihood i.e sum of the logs

p = 0.1
L = np.sum(np.log(binom.pmf(x, n, p)))

constraint that sigma has to be > 0.

```
def constraint (p):
    sigma = p[1]
    return sigma.
```

x = p
return x

kind of dictionary that tells solver we are doing an inequality and name of constraint is constraint.

```
cons = {'type': 'ineq', 'fun': constraint)

minimize (log_likelihood, p, args=(x, ), constraints=cons)
```

```
from scipy.stats import norm

fx = norm.pdf (x, loc = mean, scale = stdev)
Fx = norm.cdf (x, loc = mean, Scale = stdev)
```

→ likelihood
→ probability of $X \leq x$

stdev constant	mean const.
x = 32	x = 32
mean = [25, 26, 30, 32, 34]	mean = 32
sd = 4	sd = [2, 4, 6, 8, 10]
L = [ ]	L = [ ]
for i in range (0, len(mean)):	for i in range (0, len(sd)):
l = np.log(norm.pdf(x, mean[i], sd))	l = np.log(norm.pdf (x, mean, sd[i]
L.append (l)	L.append (l)
print (L)	print (L)
→ gives log likelihood for diff values of mean	plt.pl

MLE for Normal Dist.

$$pr(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

finding optimum parameter for μ, σ

$$L(\mu, \sigma \mid X) = L(\mu, \sigma \mid x_1) \times \cdots L(\mu, \sigma \mid x_n)$$

$$L = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/2\sigma^2} \times \cdots \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_n-\mu)^2/2\sigma^2}$$

$$\log L = \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/2\sigma^2}\right) + \cdots \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_n-\mu)^2/2\sigma^2}\right)$$

$$\log L = \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \log\left(e^{-(x_1-\mu)^2/2\sigma^2}\right)$$

$$\cancel{= \log\left(\sqrt{2\pi\sigma^2}\right)^{-1} - \frac{(x_1-\mu)^2}{\sqrt{2\pi}}}$$

$$\log L = \log\left(2\pi\sigma^2\right)^{-1/2} - \frac{(x_1-\mu)^2}{2\sigma^2} \log_e e \rightarrow 1$$

$$\log L = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(x_1-\mu)^2}{2\sigma^2}$$

$$\log L = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(x_1-\mu)^2$$

$$\log L = -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{1}{2\sigma^2}(x_1-\mu)^2 \leftarrow \text{1 term.}$$

log likelihood of 'n' terms :-

$$\log L = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}(x_1-\mu)^2 \cdots \frac{1}{2\sigma^2}(x_n-\mu)^2$$

ative wrt $\mu = 0$ & wrt $\sigma = 0$.

$$\frac{d\log L}{d\mu} = 0 - 0 - \frac{1}{\cancel{2}\sigma^2}(\cancel{2})(x_1-\mu)(-1) \quad \text{for 1 term.}$$

$$\frac{d\log L}{d\mu} = \frac{x_1-\mu}{\sigma^2} + \frac{x_2-\mu}{\sigma^2} + \cdots \frac{x_n-\mu}{\sigma^2} \quad \text{for n terms}$$

$$\frac{d \log L}{d\mu} = \frac{1}{\sigma^2}(x_1 + \cdots x_n - n\mu)$$

$$\frac{d \log L}{d\mu} = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} x_i - n\mu\right] \quad - \quad (i)$$

now derivative wrt σ

$$\frac{d \log L}{d\sigma} = 0 - n\left(\frac{1}{\sigma}\right) - \frac{(x_1 - \mu)^2}{2}(-2\sigma^{-3}) \quad \text{for 1 term}$$

$$= \frac{-n}{\sigma} + \frac{(x_1-\mu)^2}{\sigma^3} \quad \text{for 1 term.}$$

for 'n' terms

$$\frac{d \log L}{d\sigma} = -\frac{n}{\sigma} + \frac{(x_1-\mu)^2}{\sigma^3} + \frac{(x_2-\mu)^2}{\sigma^3} + \cdots \frac{(x_n-\mu)^2}{\sigma^3}$$

$$\frac{d \log L}{d\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\left[(x_1-\mu)^2 + \cdots (x_n-\mu)^2\right] \quad - \quad (ii)$$

equating both log likelihoods $= 0$

$d\log L/d\mu = 0$

$$\frac{1}{\sigma^2}\left[\sum x_i - n\mu\right] = 0$$

$$\rightarrow \sum x_i = n\mu$$

$$\boxed{\mu = \frac{\sum_{i=1}^{n} x_i}{n}}$$

$d\log L/d\sigma = 0.$

$$\frac{n}{\sigma} = \frac{1}{\sigma^3}\left[(x_1-\mu)^2 + \cdots (x_n-\mu)^2\right]$$

$$\sigma^2 = \frac{1}{n}\left[(x_1-\mu)^2 + \cdots (x_n-\mu)^2\right]$$

$$\boxed{\sigma = \sqrt{\frac{(x_1-\mu)^2 + \cdots (x_n-\mu)^2}{n}}}$$