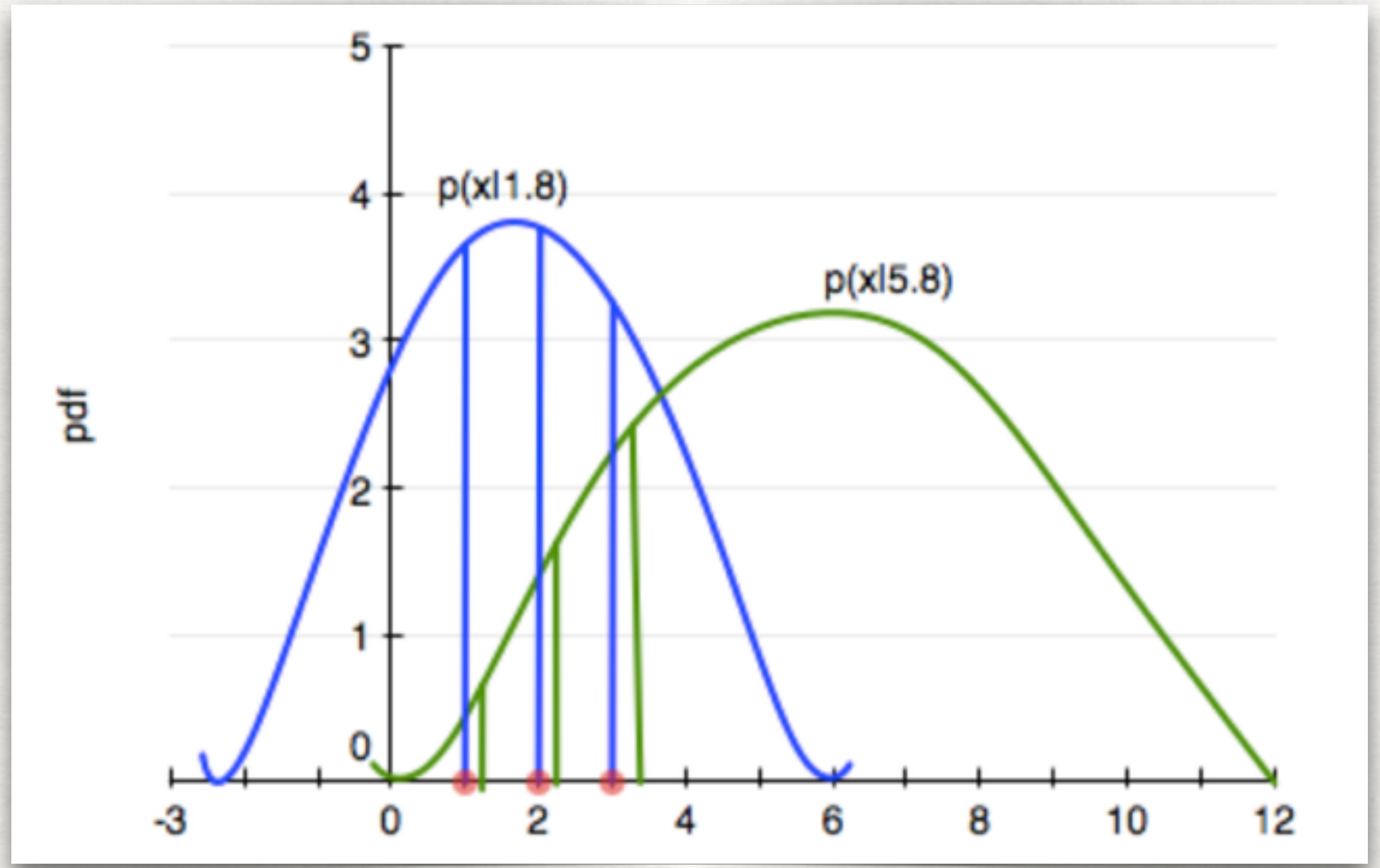
MLE AND LINEAR REGRESSION

MAXIMUM LIKELIHOOD ESTIMATION

MLE

- MLE asks is: how can we move and scale the distribution, that is, change θ, until the product of the 3 bars is maximized!



MLE

- Conditional on the fixed value of **λ**, which distribution is the data more likely to have come from?
- change **λ**, until the product of the 3 bars is maximized!
- That is, the product

$$L(\lambda) = \prod_{i=1}^{n} P(x_i \mid \lambda)$$

MLE

$$L(\lambda) = \prod_{i=1}^{n} P(x_i \mid \lambda)$$

- Measure of how likely it is to observe values x1,…,xn given the parameters **λ**.

- Maximum likelihood fitting consists of choosing the appropriate "likelihood" function L=P(X|**λ**) to maximize for a given set of observations. How likely are the observations if the model is true?

- Often it is easier and numerically more stable to maximize the log likelihood

$$\ell(\lambda) = \sum_{i=1}^{n} ln(P(x_i \mid \lambda))$$

MLE

- The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process.

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

- Given samples X = (x1,…,xn) follows exp(**λ**), what **λ** would give the best probability?

- **λ** = n/(x1+x2+ … + xn)

MLE

- {Xi} are IID , follow Uniform[0,λ]
- Given these samples X = (x1,…,xn) , what λ would give the best probability?

MLE

- $\{X_i\}$ are IID , follow Uniform$[0,\lambda]$
- Given these samples $X = (x_1,\ldots,x_n)$, what λ would give the best probability?
- $\lambda = \max\{x_1,x_2,\ldots,x_n\}$

MLE OF GAUSSIAN RANDOM VARIABLE

- Recall: Hessian method for two variables.
  How to find/determine max/min of f(x,y)?

- 1) we have two first partial derivative equations vanish

- 2) Rules for two variable Maximums and Minimums

1. Maximum

$$
\begin{aligned}
f_{xx} &< 0 \\
f_{yy} &< 0 \\
f_{yy}f_{xx} - f_{xy}f_{yx} &> 0
\end{aligned}
$$

2. Minimum

$$
\begin{aligned}
f_{xx} &> 0 \\
f_{yy} &> 0 \\
f_{yy}f_{xx} - f_{xy}f_{yx} &> 0
\end{aligned}
$$

3. Otherwise, we have a *Saddle Point*

MLE OF GAUSSIAN RANDOM VARIABLE

- How to find/determine max/min of f(x,y)?

- Recall: Hessian method for two variables.
  1) we have two first partial derivative equations vanish
  2) Rules for two variable Maximums and Minimums

1. Maximum

$$\begin{aligned} f_{xx} &< 0 \\ f_{yy} &< 0 \\ f_{yy}f_{xx} - f_{xy}f_{yx} &> 0 \end{aligned}$$

2. Minimum

$$\begin{aligned} f_{xx} &> 0 \\ f_{yy} &> 0 \\ f_{yy}f_{xx} - f_{xy}f_{yx} &> 0 \end{aligned}$$

3. Otherwise, we have a *Saddle Point*

- eg: Determine max/min of f(x,y)

$$f(x,y) = 10x + 10y + xy - x^2 - y^2$$

MLE OF GAUSSIAN RANDOM VARIABLE

- The parameters of a Gaussian distribution are the mean (μ) and standard deviation (**σ**).
- Given observations x1, . . . , xN , the likelihood of those observations for a certain μ and **σ**
- what **λ** = (μ, **σ**) would give the best probability?

$$p(x_1, \ldots, x_N | \mu, \sigma^2) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ \frac{-(x_n - \mu)^2}{2\sigma^2} \right\}$$

- the log likelihood is

$$-\frac{1}{2} N \log(2\pi\sigma^2) - \sum_{n=1}^{N} \frac{(x_n - \mu)^2}{2\sigma^2}$$

MLE OF GAUSSIAN RANDOM VARIABLE

- Find (μ, **σ**) such that the following function is maximized

$$-\frac{1}{2}N\log(2\pi\sigma^2) - \sum_{n=1}^{N}\frac{(x_n-\mu)^2}{2\sigma^2}$$

MLE OF GAUSSIAN RANDOM VARIABLE

- Find (μ, σ) such that the following function is maximized

$$p(x_1, \ldots, x_N|\mu, \sigma^2) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ \frac{-(x_n - \mu)^2}{2\sigma^2} \right\}$$

- Note: : To maximize (minimize) a function of many variables you use the technique of partial differentiation, and Hessian