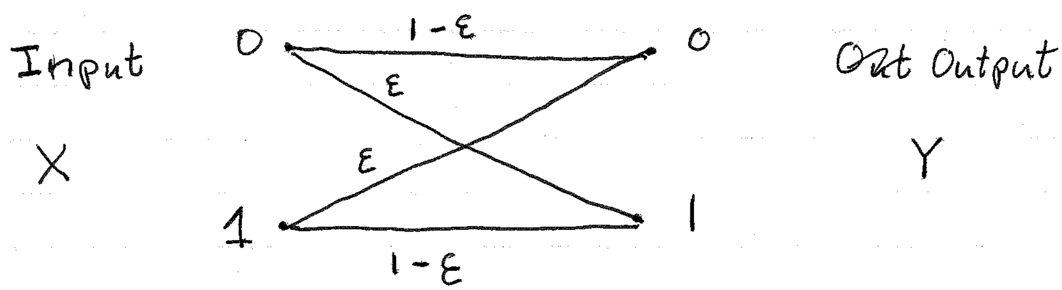
Estimation of R.V.

Engineering applications typically require the estimation of numerical quantities or parameter values given limited prior information. Probability models are useful for deriving "good" estimates of observed quantities in random experiments.

Ex Binary Communication Channel



In a communication system, we would like to estimate the value of X given the output observation $Y = y$.

Two methods:

1. What is the most likely value of X given $Y = y$?

Can find $Pr(X = x \mid Y = y)$ and let the estimate of X be the value that maximizes $Pr(X = x \mid Y = y)$.

Can express this as

$$\cancel{x}\ \hat{X} = \arg\max_{x} \Pr(X = x \mid Y = y)$$

The estimator above is referred to as the maximum a posteriori (MAP) estimator.

We can write $\Pr(X=x \mid Y=y)$ as

$$\Pr(X = x \mid Y = y) = \frac{\Pr(Y = y \mid X = x)\Pr(X=x)}{\Pr(Y = y)}$$

$$= \frac{\Pr(Y = y \mid X = x)\Pr(X = x)}{\sum_{x_i} \Pr(Y = y \mid X = x_i)\Pr(X = x_i)}$$

We are given $\Pr(Y = y \mid X = x)$, but need $\Pr(X = x)$ to find $\Pr(X = x \mid Y = y)$.

What if we do not know $\Pr(X = x)$?

2. What is the value of X that maximizes the likehood of the observation $Y = y$?

In this case, we let the estimate of X be the value $\cancel{X}$ x that maximizes

$\Pr(Y = y \mid X = x)$.

Can express this as

$$\hat{X} = \arg\max_{x} \Pr(Y = y \mid X = x)$$

The estimator above is referred to as the maximum likelihood (ML) estimator.

The MAP estimator for X given $Y = y$ is

$$\hat{X}_{MAP}(y) = \arg\max_{x} \Pr(X = x \mid Y = y)$$

if X, Y are discrete.

$$= \arg\max_{x} f_{X|Y}(x|y)$$

if X, Y are continuous.

The ML estimator for X given $Y = y$ is

$$\hat{X}_{ML}(y) = \arg\max_{x} \Pr(Y = y \mid X = x)$$

if X, Y are discrete.

$$= \arg\max_{x} f_{Y|X}(y|x)$$

if X, Y are continuous.

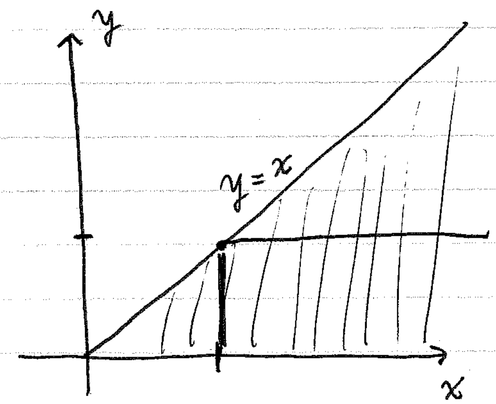
Ex: Let X, Y be r.v. with jpdf

$$f_{X,Y}(x,y) = 2e^{-(x+y)}, \quad 0 \le y \le x < \infty$$

Find the MAP and ML estimators of X given Y.

Want to find $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \quad f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$



$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$$= \int_0^x 2e^{-(x+y)}\,dy$$

$$= 2e^{-x}(1-e^{-x}), \quad x \ge 0$$

$$= 0, \quad \text{else}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

$$= \int_y^{\infty} 2e^{-(x+y)}\,dx$$

$$= 2e^{-2y}, \quad y \ge 0$$

$$= 0, \quad \text{else}$$

$$f_{X|Y}(x|y) = e^{-(x-y)} \quad , \quad 0 \le y \le x < \infty$$
$$= 0 \quad , \quad \text{else}$$

$$f_{Y|X}(y|x) = \frac{e^{-y}}{1-e^{-x}} \quad , \quad 0 \le y \le x < \infty$$
$$= 0 \quad , \quad \text{else}$$

$$\hat{X}_{MAP}(y) = \underset{x}{\arg\max}\; f_{X|Y}(x|y)$$
$$= \underset{x}{\arg\max}\; e^{-(x-y)} \quad 0 \le y \le x < \infty$$

$e^{-(x-y)}$ is maximized when $x-y=0 \Rightarrow x=y$

$$\Rightarrow \boxed{\hat{X}_{MAP}(y) = y}$$

$$\hat{X}_{ML}(y) = \underset{x}{\arg\max}\; f_{Y|X}(y|x)$$
$$= \underset{x}{\arg\max}\; \frac{e^{-y}}{1-e^{-x}} \quad , \quad 0 \le y \le x < \infty$$

Want $1-e^{-x}$ to be small so that $\frac{1}{1-e^{-x}}$ is large.

$1-e^{-x}$ is smallest when $x = y$

$\Rightarrow \hat{X}_{ML}(y) = y$

In this case $\hat{X}_{MAP}(y) = \hat{X}_{ML}(y)$, but this is not true in general.

How do we quantify how good our estimate is?

Let $\hat{X}$ be the estimate of X.

The estimation error between X and $\hat{X}$ is $X - \hat{X}$.

Often define some function $L(X-\hat{X})$ that we want to minimize on average.

Thus we want to minimize $E[L(X-\hat{X})]$.

Choices for $L(X-\hat{X})$ vary depending on the application. Often want to minimize mean square error (MSE) of the estimate, i.e.

$E[(X-\hat{X})^2]$.

Can be shown that the estimator $\hat{X}$ which minimizes $E[(X-\hat{X})^2]$ is $E[X|Y=y]$.

The minimum mean square error (MMSE) estimator of X given $Y=y$ is

$$\hat{X}_{MSE}(y) = E[X|Y=y]$$

Can be hard to find $E[X \mid Y=y]$

so sometimes "pretend" X, Y are jointly Gaussian because in this case

$$E[X \mid Y=y] = \rho_{XY} \frac{\sigma_X}{\sigma_Y} (y - \mu_Y) + \mu_X$$

Note that this estimator is a linear function of y.

The linear MMSE (LMMSE) estimator of X given $Y=y$ is

$$\hat{X}_{LMMSE}(y) = \rho_{XY} \frac{\sigma_X}{\sigma_Y} (y - \mu_Y) + \mu_X$$