ECE 459: Communications I

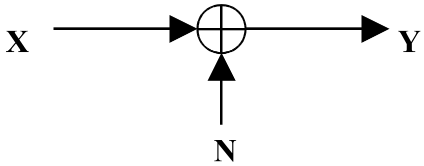
Supplemental Notes on Estimation

Prepared by: Suneil Hosmane
Based on Lectures by Prof. Hadjicostis
University of Illinois at Urbana-Champaign
Department of Electrical and Computer Engineering
© Copyright 2005 Christoforos Hadjicostis. All rights reserved.

Introduction to Estimation

Just like (binary) hypothesis testing is important in digital communication systems, **estimation** is important in analog communications.

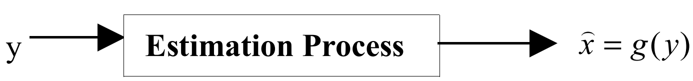
Consider the following simplified "communication channel":



X denotes the signal that we would like to transmit over a channel and is modeled as a random variable with some known pdf $f_x(x)$.

N represents noise in the channel and is modeled by as a random variable with normal distribution (Gaussian random variable) with mean = μ_n (typically 0) and variance = σ_n^2. (Noise is added to the signal **X** when it is transmitted through the channel.)

The receiver receives **Y** and needs to make an intelligent guess as to the actual signal **X** that was sent. How do we do this? We need a suitable estimator that can map our observation **Y** = y to some estimate about what was transmitted. The following sections discuss how we can choose our estimators.



y is probabilistically related to x (e.g., via a known joint pdf $f_{x,y}(x,y)$).
$\hat{x}$ is an estimate of x and is a function of y.

Consider the following example which is not a communication example but still a very relevant application of estimation.

Example: Class Grade Estimation

An exam was given at UIUC. The exam was out of a total of 80 points. We need to guess what a particular student (student **X**) scored in this exam. We only need to guess, 0-5, 6-10, 11-15, etc…. (interval guessing). What should our guess be? This is a difficult question because we have no prior information about how hard or easy the exam was.

Suppose that I now reveal to you the histogram for that particular exam. What would be your guess for student's X score?

The mean? The mode? The median?

It turns out that each of these (Mean, Mode and Median) are "optimal" estimators for different criteria.

Suppose we play the following game:

i) I randomly choose a student in the class and I write down the name on a piece of paper.
ii) Then you "guess" the student score.
iii) We repeat this many times (say N times).

What should you pick as your guess of a student's score? It depends on the rules of the game (optimization criterion).

If you want to minimize $\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2$ you should pick $\hat{x}_i = E[x \mid Info]$.

If you want to minimize $\frac{1}{N}\sum_{i=1}^{N}\left|\hat{x}_i - x_i\right|$ you should pick $\hat{x}_i = Median(x \mid Info)$.

If you want to minimize $\frac{1}{N}\sum_{i=1}^{N}\cos t(\hat{x}_i - x_i)$, where

$$\begin{aligned} \cos t(\hat{x}_i - x_i) &= 1 \text{ if } \hat{x}_i \neq x_i \\ &= 0 \text{ if } \hat{x}_i = x_i, \end{aligned}$$

you should pick $\hat{x}_i = Mode(x \mid Info)$.

(Note that the last calculation is very similar to the MAP rule (the mode(x | Info) is what has the maximum a posteriori probability.) Also note that in all of the above we are conditioning on the available information, i.e., the histogram that I revealed to you. Next we will focus on Minimum Mean Square Error (MMSE) estimation.

Minimum Mean Square Error (MMSE) Estimation

Suppose we want to choose $\hat{x}$ so that we minimize the following quantity:

$$E[(\hat{x}-x)^2] = E[(\hat{x}(y)-x)^2], \text{ or}$$

$$E[(\hat{x}(y)-x)^2] = \iint (\hat{x}(y)-x)^2 \cdot f_{x,y}(x,y)dxdy$$

If we make the substitution $f_{x,y}(x,y)dxdy = f_{x|y}(x|Y=y)$, we get

$$\min(\hat{x}(y)) \iint (\hat{x}(y)-x)^2 \cdot f_{x|y}(x|Y=y) \cdot f_y(y)dxdy$$

$$\min(\hat{x}(y)) \int \left[\int (\hat{x}(y)-x)^2 \cdot f_{x|y}(x|Y=y)dx\right] \cdot f_y(y)dy$$

Since $\hat{x}(y)$ is fixed for a given **Y** = y and since all quantities are positive, what we need to do is:

$$\min\alpha \int (\alpha - x)^2 \cdot f_{x|y}(x|Y=y)dx \text{ where we substituted } \alpha = \hat{x}(y)$$

To minimize the expression we take the derivative with respect to α .

$$\frac{\partial}{\partial\alpha} \int (\alpha - x)^2 \cdot f_{x|y}(x|Y=y)dx = 0$$

$$\int 2 \cdot (\alpha - x) \cdot f_{x|y}(x|Y=y)dx = 0$$

$$\alpha \int f_{x|y}(x|Y=y)dx = \int x \cdot f_{x|y}(x|Y=y)dx$$

$$\alpha = \int x \cdot f_{x|y}(x|Y=y)dx$$

$$\alpha = E[X|Y=y]$$

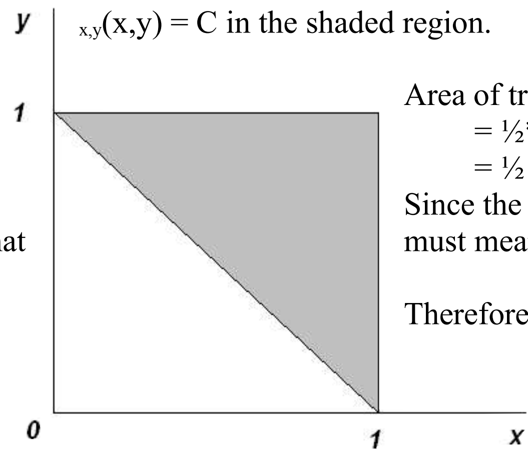
$$\therefore \hat{x}(y) = E[X|Y=y]$$

This is the MMSE estimator for X given the observation that **Y** = y.

Example: Find the MMSE estimator for X given Y=y.

Setup: **X**, **Y** are two Random Variables with joint density $f_{x,y}(x,y)$. Given **Y** = y, find $\widehat{x}_{MMSE}(y) = E[x \mid Y = y]$

$f_{x,y}(x,y)$ = C in the shaded region.



Area of triangle = ½ bh
= ½*1*1
= ½

Since the total area of $f_{x,y}(x,y) = 1$,
that must mean that ½ *C = 1.

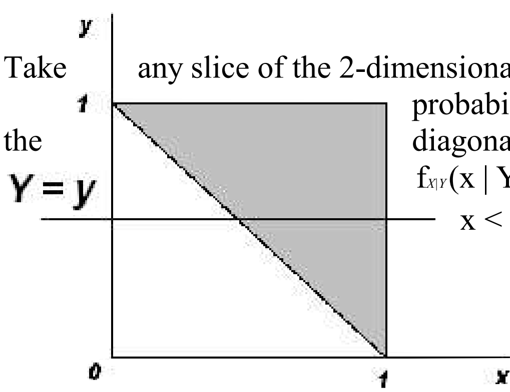
Therefore, C = 2.

$$f_{X|Y}(x \mid Y = y) = \frac{f_{x,y}(x, y)}{f_y(y)}$$

$f_{x,y}(x,y) = 2 \qquad 0 < x < 1 \qquad 1-x < y < 1$

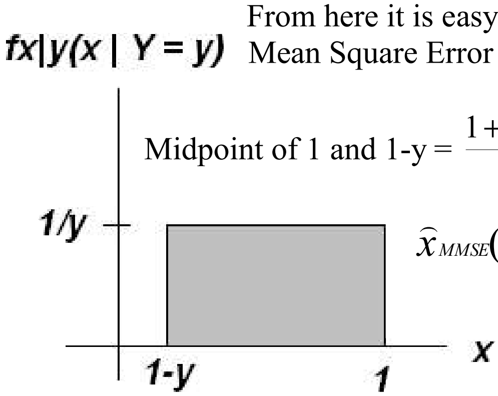
Since there is a constant distribution in the shaded region, $f_{X|Y}(x \mid Y = y)$ is a uniform probability density from 1-y < x < 1.

To illustrate, look at the following picture:



Take any slice of the 2-dimensional space along **Y** = y. Since there is an even probability distribution from the point of intersection of the diagonal line (x = 1 – y) and the vertical line at x = 1, $f_{X|Y}(x \mid Y = y)$ will have a uniform distribution from 1-y < x < y.

From here it is easy to find the mean, and hence, the Minimum Mean Square Error (MMSE) estimator.



Midpoint of 1 and 1-y = $\frac{1+(1-y)}{2} = \frac{y}{2}$

$$\hat{x}_{MMSE}(y) = 1 - \frac{y}{2}$$

Linear Minimum Mean Square Error (LMMSE) Estimators

LMMSE focuses on estimators that minimize the MSE but restricts itself to linear estimators. We do this because the distribution $f_{X|Y}(x \mid Y = y)$ is often hard to calculate or unavailable (whereas the qualities needed for LMMSE estimation are available). The LMMSE has the following form:

y ⟶ **Estimation Process** ⟶ $\hat{x}(y) = ay + b$

Our job is still to minimize the mean square error $E[(\hat{x}(y) - x)^2]$ given by

$$E[(\hat{x}(y) - x)^2] = \iint (ay + b - x)^2 f_{x,y}(x, y)dxdy\,.$$

To minimize this expression we must find appropriate coefficients ***a*** and ***b***. Therefore, we take two separate derivates (one with respect to ***a***, the other with respect to ***b***) and set them to 0.

$$E[(\hat{x}(y) - x)^2] = E[(ay + b - x)^2] = E[a^2y^2 + 2ayb + b^2 - 2ayx - 2bx + x^2]$$

$$\frac{\partial}{\partial a} E[(\hat{x}(y) - x)^2] = E[2ay^2] + E[2yb] - E[2yx] = 0$$

$$aE[y^2] + bE[y] = E[yx]$$

$$\frac{\partial}{\partial b} E[(\hat{x}(y) - x)^2] = E[2ay] + E[2b] - E[2x] = 0$$

$$aE[y] + b = E[x]$$

By solving this system of equations, we obtain the following closed form solution for the LMMSE estimator:

$$\hat{x}_{LMMSE}(y) = \mu_x + \rho(x,y)\frac{\sigma x}{\sigma y}(y - \mu_y)$$

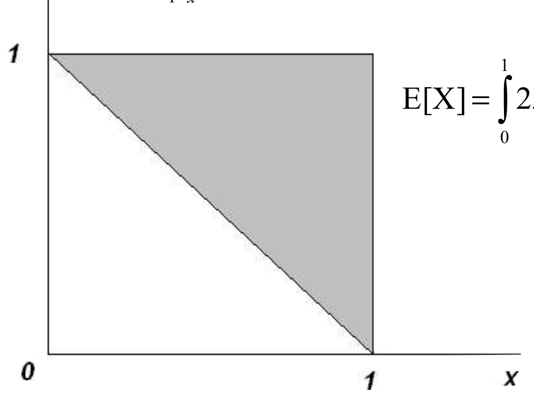
Here μ_x, σ_x (μ_y, σ_y) are the mean and standard deviation of **X** (**Y**) and $\rho(x, y)$ is the correlation coefficient between **X** and **Y.**

<u>Example: Revisiting the same problem but finding the LMMSE estimator for X given Y = y</u>

Setup: **X**, **Y** are two Random Variables with known $f_{x,y}(x,y)$. Find

$$\hat{x}_{LMMSE}(y) = \mu_x + \rho(x,y)\frac{\sigma x}{\sigma y}(y - \mu_y)$$

$$f_X(x) = \int_{1-x}^{1} 2dy = 2y\Big|_{1-x}^{1} = 2 - 2 + 2x = 2x$$



$$E[X] = \int_0^1 2x^2 dx = \frac{2x^3}{3}\Bigg|_0^1 = \frac{2}{3}$$

$$\text{var(X)} = E[X^2] - (E[X])^2 = \left(\int_0^1 2x^3 dx = \frac{x^4}{2}\Bigg|_0^1 = \frac{1}{2}\right) - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

Due to symmetry, E[X] = E[Y] and var(X) = var(Y).

$$\rho(X,Y) = \frac{\text{cov}(X,Y)}{\sigma x \cdot \sigma y} = \frac{E[XY] - E[X]E[Y]}{\sigma x \cdot \sigma y}$$

$$E[XY] = \int_0^1\left(\int_{1-x}^{1} 2xy\,dy\right)dx = xy^2\Big|_{1-x}^{1} = \int_0^1 \left(2x^2 - x^3\right)dx = \frac{2x^3}{3} - \frac{x^4}{4}\Bigg|_0^1 = \frac{5}{12}$$

$$\rho(X,Y) = \frac{E[XY] - E[X]E[Y]}{\sigma x \cdot \sigma y} = \frac{\frac{5}{12} - \left(\frac{2}{3}\right)^2}{\left(\frac{1}{18}\right)} = -\frac{1}{2}$$

$$\boxed{\hat{x}_{LMMSE}(y) = \frac{2}{3} - \frac{1}{2}(y - \frac{2}{3}) = 1 - \frac{y}{2}}$$

Notice we get the same result as in the unrestricted MMSE case. In general, the MMSE and LMMSE estimators will not be the same; However, if the MMSE estimator is linear, then MMSE = LMMSE. The following property always holds:

$$E[(\hat{x}_{MMSE}(y) - x)^2] \leq E[(\hat{x}_{LMMSE}(y) - x)^2]$$