7 Rec 7: Parameters $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$

Directions: Your instructor will spend the the first 40 minutes of the recitation period working some review problems and going over one or more Matlab experiments in the following. During the last 10 minutes of recitation, your proctor will give you a "Lab Form" that your recitation team completes, signs, and turns in. See the last page for an indication of what you will be asked to do on the Lab Form.

Due to time limitations, only a part of the following can be covered during the recitation period. However, you might want in the future to try some of the uncovered experiments on your own. They could give skills useful on some future homework problems and could lend insight into your understanding of the course from an experimental point of view.

This Week's Topics.

- Computation of $r_{X,Y}, \sigma_{X,Y}$ for discrete X, Y
- Estimation of $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$
- Correlation Properties of a Discrete Channel
- Correlation Receiver Design
- Correlation Matrix and Covariance Matrix

7.1 Exp 1: Computation of $r_{X,Y}, \sigma_{X,Y}$ for discrete X, Y

Let X, Y be RV's taking finitely many values. Then it is easy to compute $r_{X,Y}$ and $\sigma_{X,Y}$ via Matlab, using the joint PMF matrix. The following example illustrates the technique.

Example 1. Let X, Y have the following joint PMF array:

$$\begin{array}{c} \\ X=0 \\ X=1 \\ X=2 \end{array} \begin{array}{c} \begin{array}{ccc} Y=1 & Y=2 & Y=3 \end{array} \\ \begin{pmatrix} 1/6 & 1/6 & 0 \\ 0 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 \end{pmatrix} \end{array} \qquad (1)$$

Run the following Matlab script:

```
x=[0 1 2]; %enter in values of X
y=[1 2 3]; %enter in values of Y
M = [1/6 1/6 0
      0  1/6 1/6
     1/6  0  1/6];  %enter in joint PMF matrix
[X,Y]=ndgrid(x,y);
CORRELATION = sum(sum(X.*Y.*M))
```

```
MEAN_X = sum(sum(X.*M))
MEAN_Y = sum(sum(Y.*M))
COVARIANCE = CORRELATION - MEAN_X*MEAN_Y
```

It is not hard to evaluate $r_{X,Y} = E[XY]$ by hand in this case. Do that and compare with the figure `CORRELATION` given by the above script. Later, when you have time, you can investigate the wonderful Matlab command "`ndgrid`" in order to understand what it did in the script above. (The command "`ndgrid`" is similar to the command "`meshgrid`".)

7.2 Exp 2: Estimation of $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$

Suppose we have a random pair (X, Y). Suppose for some large n we observe or simulate the values of (X, Y) over n independent trials. We put the X observations in a vector `x` of length n and we put the Y observations in a vector `y` of length n. Here is how you use Matlab to estimate each of the parameters $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$ based on `x,y`:

- The quantity

  ```
  mean(x.*y)
  ```

  is a good estimate of the correlation $r_{X,Y} = E[XY]$.

- The quantity

  ```
  mean(x.*y) - mean(x)*mean(y)
  ```

  is a good estimate of the covariance $Cov(X, Y) = \sigma_{X,Y}$.

- The quantity

  ```
  (mean(x.*y) - mean(x)*mean(y))/(std(x)*std(y))
  ```

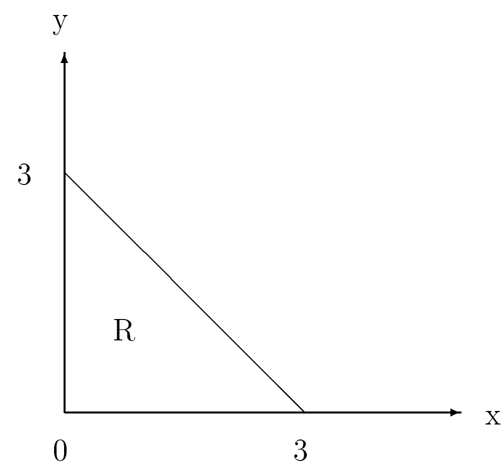
  is a good estimate of the correlation coefficient $\rho_{X,Y}$.

Example 2. Recall the ice cream experiment of Experiment 2 of Recitation 6. (Bill eats X ice cream cones and then runs Y miles.) Run the following script, which simulates 10000 observations of (X, Y):

```
clear
for i=1:10000
N=-1;
T=0;
while T<1
T=T-log(rand(1,1));
N=N+1;
end
x(i)=N;
y(i)=sum(rand(1,N+1)>1/2);
end
```

Now obtain Matlab estimates of $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$ using the vectors `x,y`. Re-run the preceding script and then re-compute the estimates to see if each of three estimates stays about the same as before. Now look at your covariance estimate (estimate of $\sigma_{X,Y}$). Based on this value, can you conclude whether X, Y are statistically independent or statistically dependent? Explain. (If you don't know the answer, ask your proctor.) Now look at your $\rho_{X,Y}$ estimate. Is it between 0 and 1 (meaning X, Y are positively correlated) or is it between -1 and 0 (meaning X, Y are negatively correlated). If X, Y seem to be positively correlated, explain why this makes sense. If you do not know why this makes sense, ask your proctor or look at page 176 of your textbook.

Example 3. As in Experiment 3 of Recitation 6, we select random pair (X, Y) uniformly from the region R, where R is the triangular region



Run the following script, which simulates 10000 observations of (X, Y):

```
clear;
i=0;
while i<10000
x_temp=3*rand(1,1);
y_temp=3*rand(1,1);
if x_temp+y_temp<3
i=i+1;
x(i)=x_temp; y(i)=y_temp;
else
end
end
```

Now obtain Matlab estimates of $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$ using the vectors `x,y`. Re-run the preceding script and then re-compute the estimates to see if each of three estimates stays about the same as before. Now look at your covariance estimate (estimate of $\sigma_{X,Y}$). Based on this value, can you conclude whether X, Y are statistically independent or statistically dependent? Explain. (If you don't know the answer, ask your proctor.) Now look at your $\rho_{X,Y}$ estimate. Is it between 0 and 1 (meaning X, Y are positively correlated) or is it between -1 and 0 (meaning X, Y are negatively correlated). If X, Y seem to be negatively correlated, explain why this makes sense. If you do not know why this makes sense, ask your proctor or look at page 176 of your textbook.

Example 4. If X, Y are independent then

$$r_{X,Y} = E[XY] = \mu_X \mu_Y$$

and

$$\sigma_{X,Y} = 0.$$

Suppose that X, Y are independent Uniform(0,1) RV's. What should $r_{X,Y}$ be? Now run the following script and see if the results conform to your expectations:

```
x=rand(1,50000);
y=rand(1,50000);
CORR_ESTIMATE = mean(x.*y)
COV_ESTIMATE = mean(x.*y)-mean(x)*mean(y)
```

Example 5. Let random pair (X, Y) be chosen uniformly from the circular region

$$\{(x, y) : x^2 + y^2 < 1\}$$

Are X, Y independent? Why or why not? What do you think the three parameters $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$ will be in this case? Run the following script to obtain the estimates of these parameters:

```
clear;
i=0;
while i<10000
x_temp=2*rand(1,1)-1;
y_temp=2*rand(1,1)-1;
if x_temp^2+y_temp^2<1
i=i+1;
x(i)=x_temp; y(i)=y_temp;
else
end
end
CORR_ESTIMATE = mean(x.*y)
COV_ESTIMATE = mean(x.*y)-mean(x)*mean(y)
RHO_ESTIMATE = ( mean(x.*y)-mean(x)*mean(y))/(std(x)*std(y))
```

What did you learn from this Example? If you have dependent RV's, can you say anything in advance about what the values of $r_{X,Y}, \sigma_{X,Y}, \rho_{X,Y}$ might be?

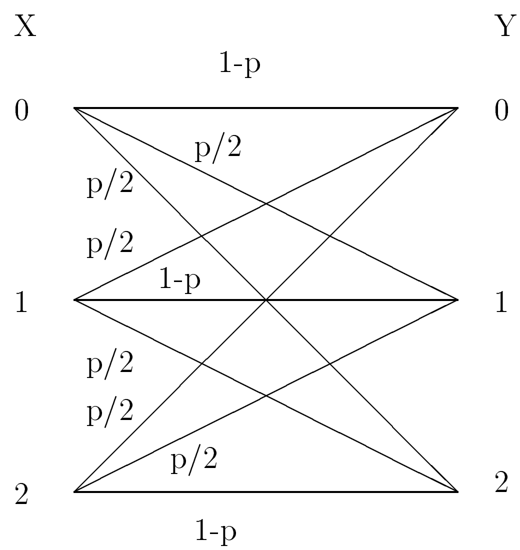
7.3 Exp 3: Correlation Properties of a Discrete Channel

In class, I examined correlation properties of a channel called the *binary symmetric channel* (BSC). The example which follows will show you that some of the correlation properties we exhibited for the BSC will also hold for other discrete channels.

Example 6. We examine a discrete channel with input and output alphabet $\{0, 1, 2\}$ and channel matrix

$$\begin{bmatrix} 1-p & p/2 & p/2 \\ p/2 & 1-p & p/2 \\ p/2 & p/2 & 1-p \end{bmatrix}$$

where p is a parameter between 0 and 1 (the probability the channel makes a transmission error). The "line diagram" of this channel is as follows:



Let X be a random input RV to this channel (assumed equiprobable), and let Y be the corresponding output RV. The correlation coefficient $\rho = \rho_{X,Y}$ of X and Y is given by:

$$\rho = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

The correlation coefficient ρ will be a function of the crossover probability p. In this Example, we will:

- Use Matlab to to obtain an estimated plot of ρ versus p.
- Examine the plot for certain ρ values to see what this tells us about the relationship between X and Y.

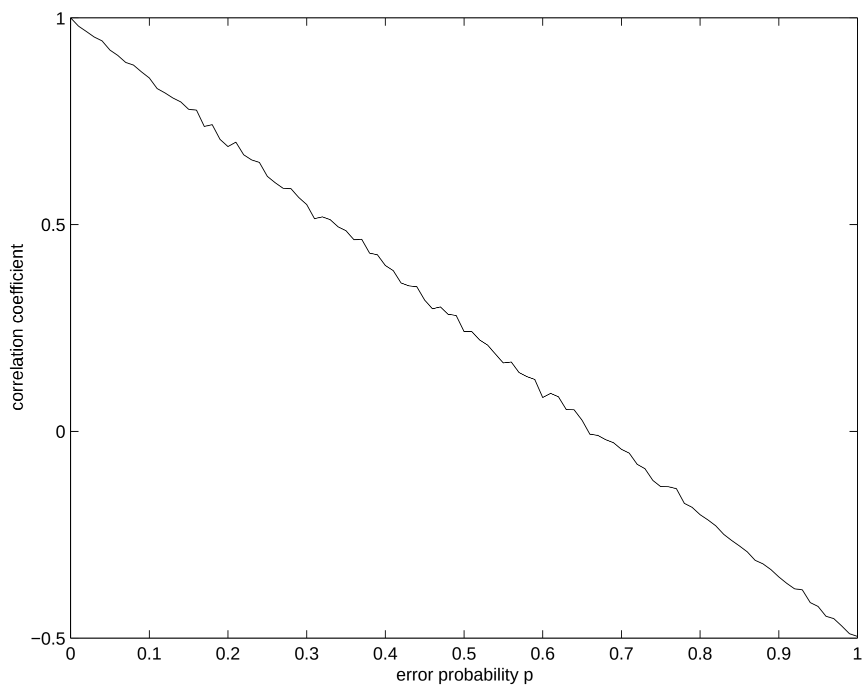
We will be using the following Matlab code to simulate a sequence `x` of equiprobable inputs and the corresponding sequence `y` of outputs from the channel in response to these inputs:

```
x=floor(3*rand(1,10000));
u=(rand(1,10000)<p);
y=rem(x+ceil(2*rand(1,10000)).*u,3);
```

Run the following program, which estimates the correlation coefficient ρ for channel input and output as a function of p:

```
p=0:.01:1;
for i=1:length(p)
q=p(i);
x=floor(3*rand(1,10000));
u=(rand(1,10000)<q);
y=rem(x+ceil(2*rand(1,10000)).*u,3);
rho(i)=mean((x-mean(x)).*(y-mean(y)))/(std(x)*std(y));
end
plot(p,rho)
xlabel('error probability p')
ylabel('correlation coefficient')
```

You should see a plot on your computer screen that looks something like this:



Things to Notice.

- One can show that ρ is the following straight line function of p:

$$\rho = (-3/2)p + 1. \tag{2}$$

  Does the "wiggly" straight line plot you see on your computer screen seem to be a good approximation to the straight line (2)? (Compare where the two straight lines start and end and compare the two slopes.)

- Notice that all the ρ values you see on your screen are between -1 and 1. This illustrates the following property of correlation coefficient:

$$-1 \leq \rho_{X,Y} \leq 1.$$

- Look at the plot in order to see what value of p yields a ρ value of 0. From the plot, this appears to occur at about $p = 2/3$. (Plugging, $\rho = 0$ into equation (2) and solving for p, you get exactly $p = 2/3$.) Plugging $p = 2/3$ into the channel matrix, you obtain

  ```
  1/3 1/3 1/3
  1/3 1/3 1/3
  1/3 1/3 1/3
  ```

  Notice that all 3 rows of the channel matrix are the same. Whenever you have a channel for which all of the rows of the channel matrix are identical, then channel input X and channel output Y will always be statistically independent (do you understand why?). A property of $\rho_{X,Y}$ states that if X, Y are independent, then $\rho_{X,Y}$ will automatically be equal to zero. We have just seen this property to be true in this special case.

- Notice from the plot that when $\rho = 1$, we have $p = 0$ and therefore the channel matrix is

  ```
  1 0 0
  0 1 0
  0 0 1
  ```

  In other words, the random pair (X, Y) always satisfies $X = Y$. That is, when we perform our experiment of running an input through the channel, the (input,output) pair will always fall on the straight line $y = x$ in the xy-plane. A property of $\rho_{X,Y}$ says that if $\rho_{X,Y} = 1$, then there must be a straight line $y = ax + b$ with positive slope a such that the observed value of (X, Y) will fall on the straight line $y = ax + b$ with probability equal to 1. We have just seen this property to be true in this special case.

To summarize, we have illustrated the following three properties of $\rho_{X,Y}$:

Property 1: $-1 \leq \rho_{X,Y} \leq 1$

Property 2: If X, Y are statistically independent, then $\rho_{X,Y} = 0$.

Property 3: If $\rho_{X,Y} = 1$, then there is a straight line relationship between X and Y in which the straight line has positive slope.

There is also the following property which did not show up in our experiment:

Property 4: If $\rho_{X,Y} = -1$, then there is a straight line relationship between X and Y in which the straight line has negative slope.

(This property did not show up in our experiment because we did not obtain any ρ value smaller than $-1/2$.)

7.4 Exp 4: Correlation Receiver Design

Consider the block diagram

$$X \rightarrow \boxed{\text{channel}} \rightarrow Y \rightarrow \boxed{\begin{array}{c}\text{correlation}\\ \text{receiver}\end{array}} \rightarrow \hat{X} = CY$$

As indicated, the output of the so-called "correlation receiver" is an estimate of channel input X of the form

$$\hat{X} = CY,$$

a constant C times the channel output Y. The constant C is chosen so that the mean-square estimation error

$$E[(X - \hat{X})^2] = E[(X - CY)^2] \tag{3}$$

is minimized. We will prove in class that the solution for C is

$$C = \frac{E[XY]}{E[Y^2]}. \tag{4}$$

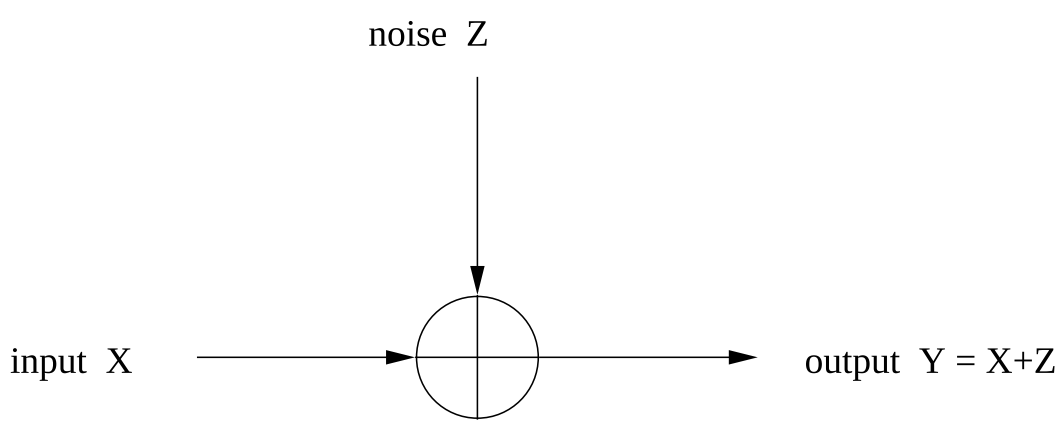
Because of the presence of the correlation $E[XY]$ in the preceding expression for C, you now see why the receiver is called the "correlation" receiver.

In this experiment, we are going to simulate a large number of inputs and outputs to a channel. Letting `x` be the vector of simulated channel inputs and letting `y` be the vector of simulated channel outputs, we are then going to plot

```
mean((x-Cy).^2)
```

as a function of C, which is an estimate of the mean-square estimation error (3). In this way, we will be able to verify that the formula (4) is approximately correct by seeing for what C value our plot reaches a minimum.

The channel we will be using is a "Gaussian additive noise channel":



The "channel noise" Z is a Gaussian RV which is independent of the input X.

Here is how the simulation of this channel will take place:

- The channel inputs will be simulated as

  ```
  x=randn(1,100000);
  ```

  In other words, we are taking the channel input RV X to be a standard Gaussian RV.

- The Gaussian channel noise samples will be simulated as

  ```
  z=2*randn(1,100000);
  ```

  In other words, we are taking the channel noise RV Z to be Gaussian with mean 0 and variance 4.

- Clearly, the channel outputs will be simulated as

  ```
  y=x+z;
  ```

Example 7. Run the following Matlab script, which estimates the constant C to be used in the correlation receiver:

```
x=randn(1,100000);
z=2*randn(1,100000);
y=x+z;
C_estimate = mean(x.*y)/mean(y.^2)
```

Now compute the exact value of C according to formula (4):

$$C = \frac{E[X(X+Z)]}{E[(X+Z)^2]} = \frac{E[X^2] + E[X]E[Z]}{E[X^2] + 2E[X]E[Z] + E[Z^2]}.$$

Plug in

$$\begin{aligned} E[X^2] &= 1 \\ E[X] &= 0 \\ E[Z^2] &= 4 \end{aligned}$$

Is your estimate for C pretty good? Store the exact value of C you just found for use in the next example.

Example 8. Run the following Matlab script, which will give you a plot of the estimated mean-square estimation error (3) as a function of C:

```
clear
x=randn(1,100000);
z=2*randn(1,100000);
y=x+z;
C=0:.01:.5;
for i=1:length(C);
esterror(i)=mean((x-C(i)*y).^2);
end
plot(C,esterror)
```

Eyeball the plot. Does its minimum point seem to coincide with the value of C you found in Example 7?

7.5 Exp 5: Correlation Matrix and Covariance Matrix

The purpose of this experiment is to introduce you to the concepts of correlation matrix and covariance matrix. I will use these concepts starting in next week's recitation to do some interesting things, some of which having to do with design.

Let X, Y be random variables. The *correlation matrix* of these RV's is defined to be the 2×2 matrix

$$\begin{bmatrix} E[X^2] & E[XY] \\ E[XY] & E[Y^2] \end{bmatrix}$$

Notice that the two diagonal elements are the second moments of the individual RV's, whereas the two off diagonal elements are both equal to the correlation $r_{X,Y} = E[XY]$.

On the other hand, the 2×2 matrix

$$\begin{bmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{bmatrix},$$

where $\sigma_{X,Y} = Cov[X, Y]$, is called the *covariance matrix* of the two RV's. Depending upon the application, it might be more convenient to deal with the correlation matrix than the covariance matrix, or vice-versa. You can go from either matrix to the other one by exploiting the equation:

$$\begin{bmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} E[X^2] & E[XY] \\ E[XY] & E[Y^2] \end{bmatrix} - \begin{bmatrix} \mu_X^2 & \mu_X\mu_Y \\ \mu_X\mu_Y & \mu_Y^2 \end{bmatrix}.$$

The last term can be written more compactly as

$$\begin{bmatrix} \mu_X^2 & \mu_X\mu_Y \\ \mu_X\mu_Y & \mu_Y^2 \end{bmatrix} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} \begin{bmatrix} \mu_X & \mu_Y \end{bmatrix}.$$

If RV's X, Y are statistically independent, then the covariance matrix is a diagonal matrix:

$$\begin{bmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}, \tag{5}$$

and the correlation matrix reduces to

$$\begin{bmatrix} E[X^2] & E[XY] \\ E[XY] & E[Y^2] \end{bmatrix} = \begin{bmatrix} E[X^2] & \mu_X\mu_Y \\ \mu_X\mu_Y & E[Y^2] \end{bmatrix}. \tag{6}$$

In this experiment, you will see how to estimate the correlation matrix and the covariance matrix from data points (x_i, y_i), and you will also verify the special forms of these matrices in the independent case.

Example 9: You estimate the correlation matrix and the covariance matrix from 50000 data points (x_i, y_i).

Step 1: Run the script:

```
clear
u=rand(1,50000); v=rand(1,50000);
x=3*u+v; y=-u+2*v;
```

You have generated 50000 observations of a random pair (X, Y), defined by

$$\begin{aligned} X &= 3U + V \qquad (7) \\ Y &= -U + 2V \qquad (8) \end{aligned}$$

where U, V are independent Uniform[0,1] RV's. That is, the i-th entry x_i of vector `x` and the i-th entry y_i of vector `y` yield the point (x_i, y_i), which is the i-th observation of (X, Y).

Step 2: Here is a Matlab one-liner estimating the correlation matrix of X, Y from the 50000 data points:

```
CORRMATRIX = [x;y]*[x;y]'/50000
```

Step 3: Here is a one-liner estimating the covariance matrix of X, Y from the 50000 data points:

```
COVMATRIX = [x-mean(x);y-mean(y)]*[x-mean(x);y-mean(y)]'/50000
```

Step 4: Compute the precise values of μ_X and μ_Y from the equations (7),(8). Then run the following script:

```
mX =0 %enter in here the mean of X
mY =0 %enter in here the mean of Y
[mX mY]'*[mX mY]
CORRMATRIX - COVMATRIX
```

You will see two 2×2 matrices on your screen. Do you understand why they are about the same? In next week's recitation, you will see how to compute the actual correlation matrix and the actual covariance matrix. Then, you will be able to return to this example to see if the estimated matrices CORRMATRIX and COVMATRIX make sense.

Step 5: Using the matrix COVMATRIX, generate an estimate for $\rho_{X,Y}$.

Example 10: In this example, you simulate observations (x_i, y_i) of a random pair (X, Y) in which X, Y are independent. You then examine the special form of the covariance matrix estimate and the correlation matrix estimate.

Step 1: Run the script

```
x=-log(rand(1,50000));
y=-log(rand(1,50000));
```

You are simulating values of (X, Y), where X, Y are independent exponentially distributed RV's each having mean 1.

Step 2: Generate estimated covariance and correlation matrices by running the script:

```
COVMATRIX = [x-mean(x);y-mean(y)]*[x-mean(x);y-mean(y)]'/50000
CORRMATRIX = [x;y]*[x;y]'/50000
```

See if your estimated covariance matrix is approximately equal to expression (5) and see if your estimated correlation matrix is approximately equal to expression (6). (Since you know X and Y each have mean and variance 1, you will be able to compute the second moments of these two RV's.) To make your results even more convincing, run the lines

```
round(COVMATRIX)
round(CORRMATRIX)
```

EE 3025 S2005 Recitation 7 Lab Form

Name and Student Number of Team Member 1:

Name and Student Number of Team Member 2:

Name and Student Number of Team Member 3:

****

Study Experiment 4 carefully, at least up through Example 7. I will ask a question concerning correlation receiver design.