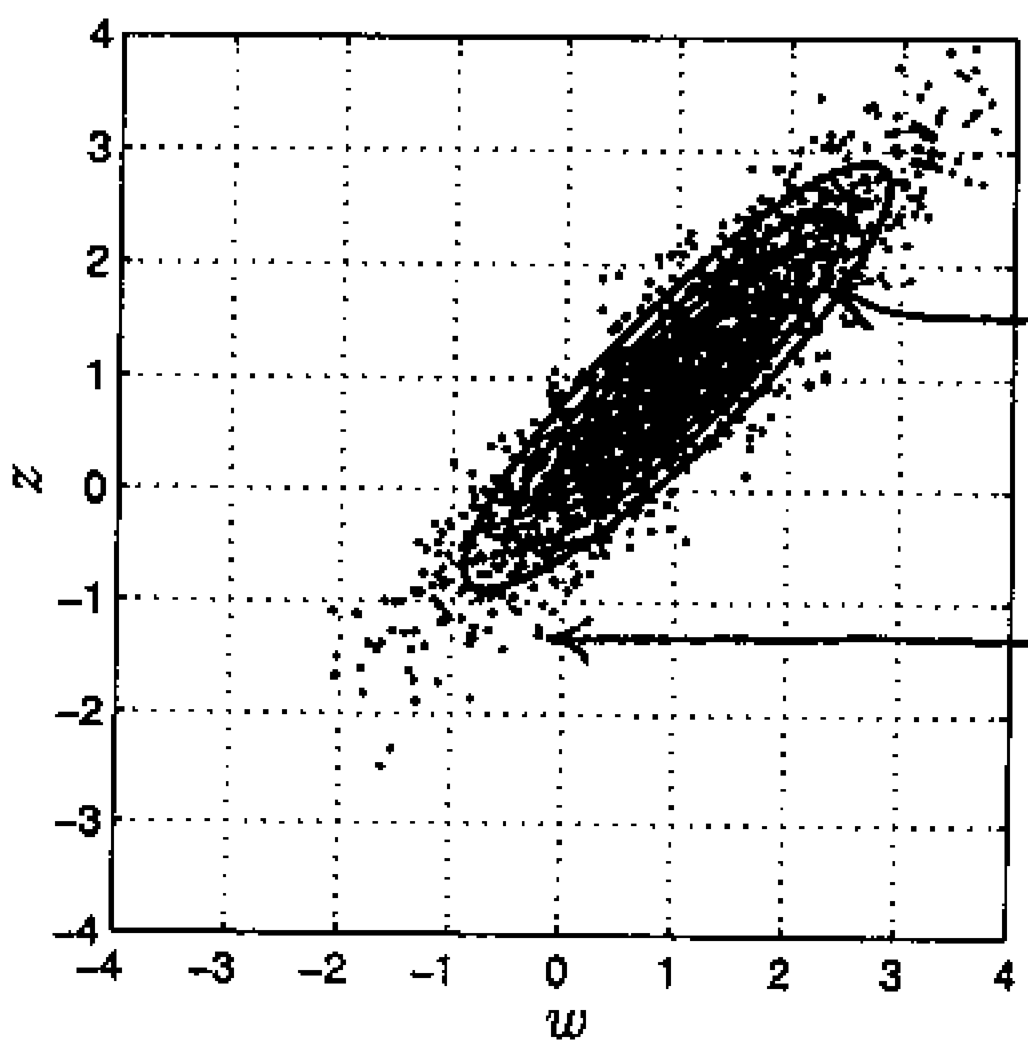
$$\Rightarrow \quad \sigma_W^2 = 1 \quad \Rightarrow \quad \sigma_W = 1 \qquad \sigma_Z^2 = 1 \Rightarrow \sigma_Z = 1$$

$$\text{cov}(W,Z) = \rho \sigma_W \sigma_Z = 0.9 \quad \Rightarrow \quad \rho = 0.9$$

$$\begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.9 & \underbrace{\sqrt{1-0.9^2}}_{\approx 0.436} \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

SEE PG. 418 FOR MATLAB CODE.



THEORETICAL PDF CONTOURS

OUTCOMES OF (W,Z)

Figure 12.21: 500 outcomes of bivariate Gaussian random vector with mean $[1\,1]^T$ and covariance matrix given by (12.40).

ESTIMATING MEANS AND COVARIANCES

ASSUME OUTCOMES (w_m, z_m) $m = 1, 2, \ldots, M$ ARE AVAILABLE

$$\hat{E}_{W,Z}\left[\begin{bmatrix} W \\ Z \end{bmatrix}\right] = \begin{bmatrix} \frac{1}{M}\sum_{m=1}^{M} w_m \\ \frac{1}{M}\sum_{m=1}^{M} z_m \end{bmatrix}$$

SAME AS BEFORE (SAMPLE MEAN)

$$= \frac{1}{M}\sum_{m=1}^{M}\begin{bmatrix} w_m \\ z_m \end{bmatrix}$$

MATLAB "FRIENDLY"

$$\hat{VAR}(W) = \frac{1}{M}\sum_{m=1}^{M}\left(w_m - \frac{1}{M}\sum_{i=1}^{M} w_i\right)^2$$

$$\hat{VAR}(Z) = \frac{1}{M}\sum_{m=1}^{M}\left(z_m - \frac{1}{M}\sum_{i=1}^{M} z_i\right)^2$$

$$\hat{COV}(W,Z) = \frac{1}{M}\sum_{m=1}^{M}\left(w_m - \frac{1}{M}\sum_{i=1}^{M} w_i\right)\left(z_m - \frac{1}{M}\sum_{i=1}^{M} z_i\right)$$

$$= \hat{COV}(Z,W)$$

TO SIMPLIFY NOTATION AND COMPUTATION (AND USEFUL FOR THEORETICAL WORK) NOTE:

$$\underset{2\times 1}{\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}} \underset{1\times 2}{\begin{bmatrix} b_1 & b_2 \end{bmatrix}} = \underset{2\times 2}{\begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{bmatrix}}$$

CALLED AN <u>OUTER PRODUCT</u> OR <u>DYAD</u>

LET $a_1 = W - E_W[W] = b_1$

$a_2 = Z - E_Z[Z] = b_2$

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} (W-E_W[W])^2 & (W-E_W[W])(Z-E_Z(Z)) \\ (Z-E_Z[Z])(W-E_W[W]) & (Z-E_Z[Z])^2 \end{bmatrix}$$

DEFINE $E\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} E[a_{11}] & E[a_{12}] \\ E[a_{21}] & E[a_{22}] \end{bmatrix}$

$$\Rightarrow E\left[\underbrace{\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}}_{\underline{a}} \underbrace{\begin{bmatrix} a_1 & a_2 \end{bmatrix}}_{\underline{a}^T}\right] = \begin{bmatrix} VAR(W) & COV(W,Z) \\ COV(Z,W) & \hat{VAR}(Z) \end{bmatrix}$$

THIS GIVES US ANOTHER DEFINITION (IN VECTOR/MATRIX NOTATION) OF $\underline{C}_{W,Z}$

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{bmatrix} W - E_W[W] \\ Z - E_Z[Z] \end{bmatrix} = \begin{bmatrix} W \\ Z \end{bmatrix} - \begin{bmatrix} E_W[W] \\ E_Z(Z) \end{bmatrix}$$

$$= \begin{bmatrix} W \\ Z \end{bmatrix} - E_{W,Z}\left[\begin{bmatrix} W \\ Z \end{bmatrix}\right]$$

$$\Rightarrow E_{W,Z}\left[\left(\begin{bmatrix} W \\ Z \end{bmatrix} - E_{W,Z}\left[\begin{bmatrix} W \\ Z \end{bmatrix}\right]\right)\left(\begin{bmatrix} W \\ Z \end{bmatrix} - E_{W,Z}\left[\begin{bmatrix} W \\ Z \end{bmatrix}\right]\right)^T\right] = \underline{C}_{W,Z}$$

SOMETIMES WRITTEN WITH $\underline{X} = \begin{pmatrix} W \\ Z \end{pmatrix}$

$$\underline{C}_X = E_{\underline{X}}\left[(\underline{X} - E_X[\underline{X}])(\underline{X} - E_X[\underline{X}])^T\right]$$

FINALLY, TO ESTIMATE $\underline{C}_{W,Z}$ USE

$$\hat{\underline{C}}_{W,Z} = \frac{1}{M}\sum_{m=1}^{M}\left(\begin{bmatrix} w_m \\ z_m \end{bmatrix} - \hat{E}_{W,Z}\left[\begin{bmatrix} W \\ Z \end{bmatrix}\right]\right)\left(\begin{bmatrix} w_m \\ z_m \end{bmatrix} - \hat{E}_{W,Z}\left[\begin{bmatrix} w_m \\ z_m \end{bmatrix}\right]\right)^T$$

SEE MATLAB CODE ON PG 418

CHAPTER 14 - CONT. N-DIMENSIONAL RVS

NOW VECTOR/MATRIX NOTATION IS ESSENTIAL!

$$\underline{X}(s) = \begin{bmatrix} X_1(s) \\ X_2(s) \\ \vdots \\ X_N(s) \end{bmatrix} \quad N \times 1$$

BEFORE $N = 2$ AND $X_1 = X$, $X_2 = Y$

$\underline{X}(S)$ TAKES ON VALUES

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

JOINT PDF NOW $p_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N)$
OR $p_{\underline{X}}(\underline{x})$. AS USUAL

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N)\, dx_1\, dx_2 \cdots dx_N = 1$$

AND $p_{X_1, X_2, \ldots, X_N}(x_1, x_2, \ldots, x_N) \geq 0$

$$P[A] = \int\int\cdots\int_A p_{\underline{X}}(\underline{x})\, d\underline{x}$$

MOST IMPORTANT N-DIMENSIONAL OR MULTIVARIATE PDF IS GAUSSIAN. DEFINED AS

$$p_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{N/2} \, \mathrm{DET}^{1/2}(\underline{C})} e^{-\frac{1}{2}\underbrace{(\underline{x}-\underline{\mu})^T \underline{C}^{-1}(\underline{x}-\underline{\mu})}_{\substack{\text{QUADRATIC FORM} \\ \text{(SCALAR)}}}}$$

WHERE
$$\underline{\mu} = E_{\underline{X}}[\underline{X}] = \begin{bmatrix} E_{X_1}[X_1] \\ \vdots \\ E_{X_N}[X_N] \end{bmatrix} \quad N \times 1$$

$$\underline{C} = \begin{bmatrix} \mathrm{VAR}(X_1) & \mathrm{COV}(X_1, X_2) & \ldots & \mathrm{COV}(X_1, X_N) \\ \mathrm{COV}(X_2, X_1) & \mathrm{VAR}(X_2) & \ldots & \mathrm{COV}(X_2, X_N) \\ & & \ddots & \\ \mathrm{COV}(X_N, X_1) & \mathrm{COV}(X_N, X_2) & \ldots & \mathrm{VAR}(X_N) \end{bmatrix}$$

$$N \times N$$

SHORTHAND $\underline{X} \sim N(\underline{\mu}, \underline{C})$

CAN ALSO BE WRITTEN AS

$$\underline{C} = E_{\underline{X}}\{(\underline{x}-\underline{\mu})(\underline{x}-\underline{\mu})^T\}$$

SEE EXAMPLE 14.1 FOR TYPICAL COMPUTATION USING $p_{\underline{X}}(\underline{x})$.

TO FIND MARGINAL PDF $p_{X_1}(x_1)$

$$p_{X_1}(x_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_{\underline{X}}(\underline{x})\, dx_2\, dx_3 \ldots dx_N$$

INTEGRATE OUT OTHER VARIABLES

OR $$p_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_{\underline{X}}(\underline{x})\, dx_3\, dx_4 \ldots dx_N$$

ETC.

$X_1, X_2, \ldots, X_N$ ARE IND. IF AND ONLY IF

$$p_{X_1,X_2,\ldots,X_N}(x_1, x_2, \ldots, x_N) = p_{X_1}(x_1)\, p_{X_2}(x_2) \cdots p_{X_N}(x_N)$$

<u>EXAMPLE</u>: INDEPENDENCE OF $N(\underline{\mu}, \underline{C})$ RVS

IF $\underline{C}$ IS DIAGONAL (ALL $\mathrm{COV}(X_i, X_j) = 0$ FOR $i \neq j$)

ALL X_i'S ARE IND.

LET $\underline{C} = \begin{bmatrix} \sigma_1^2 & & & \underline{0} \\ & \sigma_2^2 & & \\ & & \ddots & \\ \underline{0} & & & \sigma_N^2 \end{bmatrix} = \mathrm{DIAG}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$

$\mathrm{DET}(\underline{C}) = \prod_{i=1}^{N} \sigma_i^2 \qquad \underline{C}^{-1} = \begin{bmatrix} 1/\sigma_1^2 & & & \underline{0} \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ \underline{0} & & & 1/\sigma_N^2 \end{bmatrix}$

$$p_{\underline{x}}(\underline{x}) = \frac{1}{(2\pi)^{N/2} \left(\prod_{1}^{N} \sigma_i^2\right)^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \underline{C}^{-1} (\underline{x}-\underline{\mu})}$$

FOR A QUADRATIC FORM WITH DIAG. MATRIX

$$[u_1 \; u_2] \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = [u_1 \; u_2] \begin{bmatrix} a u_1 \\ b u_2 \end{bmatrix}$$

$$= a u_1^2 + b u_2^2$$

$$\Rightarrow p_{\underline{x}}(\underline{x}) = \frac{1}{(2\pi)^{N/2} \left(\prod_{1}^{N} \sigma_i^2\right)^{1/2}} e^{-\frac{1}{2} \sum_{i=1}^{N} \frac{(x_i-\mu_i)^2}{\sigma_i^2}}$$

$$= \prod_{i=1}^{N} \underbrace{\frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2\sigma_i^2}(x_i-\mu_i)^2}}_{N(\mu_i, \sigma_i^2)}$$

$$= \prod_{i=1}^{N} p_{x_i}(x_i)$$

$\Rightarrow$ X_i'S ARE IND !

TRANSFORMATIONS

STRAIGHTFORWARD EXTENSION OF (X,Y) RESULTS

$$Y_1 = g_1(x_1, x_2, \dots, x_N)$$
$$Y_2 = g_2(x_1, x_2, \dots, x_N)$$
$$\vdots$$
$$Y_N = g_N(x_1, x_2, \dots, x_N)$$

ONE-TO-ONE (CAN SOLVE FOR $x_1, x_2, \dots, x_N$ GIVEN $y_1, y_2, \dots, y_N$ UNIQUELY)

$$p_{Y_1,Y_2,\dots,Y_N}(y_1, y_2, \dots, y_N) = p_{X_1,X_2,\dots,X_N}(g_1^{-1}(\underline{y}), g_2^{-1}(\underline{y}), \dots, g_N^{-1}(\underline{y}))$$
$$\cdot \left| \mathrm{DET}\left(\frac{\partial(x_1, x_2, \dots, x_N)}{\partial(y_1, y_2, \dots, y_N)} \right) \right|$$

WHERE
$$\frac{\partial(x_1, x_2, \dots, x_N)}{\partial(y_1, y_2, \dots, y_N)} = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_N} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_N} \\ & & \ddots & \\ \frac{\partial x_N}{\partial y_1} & \frac{\partial x_N}{\partial y_2} & \cdots & \frac{\partial x_N}{\partial y_N} \end{bmatrix}$$

EXAMPLE : $\underline{X} \sim N(\underline{\mu}, \underline{C})$

$$\underline{Y} = \underline{G}\underline{X} \Rightarrow \underline{X} = \underline{G}^{-1}\underline{Y}$$

$$\frac{\partial(x_1, x_2, \dots, x_N)}{\partial(y_1, y_2, \dots, y_N)} = \underline{G}^{-1}$$ HOW TO VERIFY?

$$p_{\underline{Y}}(\underline{y}) = p_{\underline{X}}(\underline{G}^{-1}\underline{y})\,|\mathrm{DET}(\underline{G}^{-1})|$$

$$= \frac{1}{(2\pi)^{N/2}\,\mathrm{DET}^{1/2}(\underline{C})}\, e^{-\frac{1}{2}(\underline{G}^{-1}\underline{y}-\underline{\mu})^T \underline{C}^{-1}(\underline{G}^{-1}\underline{y}-\underline{\mu})}$$

$$\cdot\, |\mathrm{DET}(\underline{G}^{-1})|$$

BUT $\underbrace{\mathrm{DET}(\underline{I})}_{=1} = \mathrm{DET}(\underline{G}\,\underline{G}^{-1}) = \mathrm{DET}(\underline{G})\,\mathrm{DET}(\underline{G}^{-1})$

$$\Rightarrow \mathrm{DET}(\underline{G}^{-1}) = \frac{1}{\mathrm{DET}(\underline{G})}$$

$$p_Y(\underline{y}) = \frac{1}{(2\pi)^{N/2}\,\mathrm{DET}^{1/2}(\underline{C})\,|\mathrm{DET}(\underline{G})|}$$
$$\cdot\, e^{-\frac{1}{2}(\underline{G}^{-1}(\underline{y}-\underline{G}\underline{\mu}))^T \underline{C}^{-1}(\underline{G}^{-1}(\underline{y}-\underline{G}\underline{\mu}))}$$

AND $|\mathrm{DET}(\underline{G})| = \sqrt{\mathrm{DET}^2(\underline{G})}$
$= \sqrt{\mathrm{DET}(\underline{G})\,\mathrm{DET}(\underline{G}^T)}$

ALSO $(\underline{G}^{-1}(\underline{y}-\underline{G}\underline{\mu}))^T = (\underline{y}-\underline{G}\underline{\mu})^T\underline{G}^{-1^T}$

$$p_Y(\underline{y}) = \frac{1}{(2\pi)^{N/2}\,\mathrm{DET}^{1/2}(\underline{C}\,\underline{G}\,\underline{G}^T)}$$

SINCE $\mathrm{DET}(\underline{A}\underline{B}) = \mathrm{DET}(\underline{A})\,\mathrm{DET}(\underline{B})$

$$\cdot\, e^{-\frac{1}{2}(\underline{y}-\underline{G}\underline{\mu})^T \underbrace{\underline{G}^{-1^T}\underline{C}^{-1}\underline{G}^{-1}}_{(\underline{G}\underline{C}\underline{G}^T)^{-1}}(\underline{y}-\underline{G}\underline{\mu})}$$

AND SINCE $\mathrm{DET}(\underline{C}\,\underline{G}\,\underline{G}^T) = \mathrm{DET}(\underline{G}\,\underline{C}\,\underline{G}^T)$

$$\Rightarrow \quad \underline{Y} \sim N(\underline{G}\underline{\mu},\ \underline{G}\underline{C}\underline{G}^T)$$

CAN ALSO BE SHOWN TO HOLD IF $\underline{G}$ IS $M \times N$ WITH $M < N$.

EXPECTED VALUES

$$E_{\underline{X}}(\underline{X}) = E_{X_1, X_2, \ldots, X_N}\left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}\right) = \begin{bmatrix} E_{X_1}(X_1) \\ \vdots \\ E_{X_N}(X_N) \end{bmatrix}$$

$$E_{X_1, X_2, \ldots, X_N}[g(X_1, X_2, \ldots, X_N)] =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_N)\, p_{X_1 X_2 \ldots X_N}(x_1, x_2, \ldots, x_N)\, dx_1\, dx_2 \ldots dx_N$$

EXAMPLES : $g(X_1, X_2, \ldots, X_N) = \sum_{i=1}^{N} a_i X_i$

$$E_{X_1, X_2, \ldots, X_N}\left[\sum_{i=1}^{N} a_i X_i\right] \underset{\text{WHY?}}{=} \sum_{i=1}^{N} a_i E_{X_1, X_2, \ldots, X_N}[X_i]$$

$$= \sum_{i=1}^{N} a_i E_{X_i}[X_i]$$

$$g(X_1, X_2, \ldots, X_N) = \Big(\overbrace{\sum_{i=1}^{N} a_i X_i}^{Y} - \overbrace{E_{X_1, \ldots, X_N}\left[\sum a_i X_i\right]}^{E[Y]}\Big)^2$$

$$\Rightarrow E[g(X_1, X_2, \ldots, X_N)] = VAR\left(\sum_{i=1}^{N} a_i X_i\right)$$

TO FIND THIS WE USE SOME MATRIX ALGEBRA (RECALL $VAR(X_1 + X_2) = VAR(X_1) + VAR(X_2) + 2\,COV(X_1, X_2)$)

LET $\underline{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}$ $\quad \underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}$

$$\sum_{i=1}^{N} a_i X_i = \underline{a}^T \underline{X}$$ INNER PRODUCT

$$E_{X_1,\ldots,X_N}\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i E_{X_i}[X_i]$$

$$= \underline{a}^T E_{\underline{x}}[\underline{x}]$$ (RECALL DEFINITION)

$$\Rightarrow VAR\Big(\underbrace{\sum_{i=1}^{N} a_i X_i}_{Y}\Big) = E_Y\left[(Y - E[Y])^2\right]$$

$$= E_{\underline{x}}\left[\left(\underline{a}^T \underline{x} - \underline{a}^T E_{\underline{x}}[\underline{x}]\right)^2\right]$$

WHY?

$$= E_{\underline{x}}\left[\left(\underline{a}^T \underline{x} - \underline{a}^T E_{\underline{x}}[\underline{x}]\right)\left(\underline{a}^T \underline{x} - \underline{a}^T E_{\underline{x}}[\underline{x}]\right)^T\right]$$

$$= E_{\underline{x}}\left[\underline{a}^T\left(\underline{x} - E_{\underline{x}}[\underline{x}]\right)\left(\underline{a}^T\left(\underline{x} - E_{\underline{x}}[\underline{x}]\right)\right)^T\right]$$

$$= E_{\underline{x}}\left[\underline{a}^T\left(\underline{x} - E_{\underline{x}}[\underline{x}]\right)\left(\underline{x} - E_{\underline{x}}[\underline{x}]\right)^T \underline{a}\right]$$

$$= \underline{a}^T \underbrace{E_{\underline{x}}\left[\left(\underline{X} - E_{\underline{x}}[\underline{x}]\right)\left(\underline{X} - E_{\underline{x}}[\underline{x}]\right)^T\right]}_{\underline{C}_X} \underline{a}$$

$$= \underline{a}^T \underline{C}_x \underline{a}$$

IF $\underline{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad N = 2$, $VAR(X_1 + X_2) =$

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} VAR(X_1) & COV(X_1, X_2) \\ COV(X_2, X_1) & VAR(X_2) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= [1 \quad 1] \begin{bmatrix} VAR(X_1) + COV(X_1, X_2) \\ COV(X_2, X_1) + VAR(X_2) \end{bmatrix}$$

$$= VAR(X_1) + COV(X_1, X_2) + COV(X_2, X_1) + VAR(X_2)$$

AGREE WITH EARLIER RESULTS?

$(\underline{1}^T \underline{A} \underline{1} = \sum \text{ELEMENTS IN } \underline{A})$

IF $\underline{\Sigma}_X = DIAG(VAR(X_1), \ldots, VAR(X_N))$
(X_i'S UNCORRELATED),

$$VAR\left(\sum_{i=1}^{N} a_i X_i \right) = \sum_{i=1}^{N} a_i^2 VAR(X_i) \qquad \text{VERIFY!}$$

<u>EXAMPLE</u>: SAMPLE MEAN OF IID RVS

IID $\equiv$ INDEPENDENT AND <u>IDENTICALLY DISTRIBUTED</u>

MEANS THAT ALL RVS HAVE <u>SAME</u>
<u>MARGINAL</u> <u>PDF</u> OR

$$p_{X_1} = p_{X_2} = \ldots = p_{X_N} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x_i^2}$$

FOR EXAMPLE

AN ESTIMATE OF $E_{X_i}[X_i] = \mu$ SAME SINCE PDF SAME
IS SAMPLE MEAN

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

IS THIS A GOOD ESTIMATE? WHAT HAPPENS AS $N \to \infty$? RECALL DETECTION EXAMPLE ON PAGE 6 OF NOTES. THERE

$\mu = 0$ WHEN NO SIGNAL PRESENT

0.5 WHEN SIGNAL PRESENT

SINCE $X_i = 0.5 + W_i \leftarrow$ NOISE SAMPLE

W_i WITH $E[W_i] = 0$

$VAR(W_i) = \sigma^2$

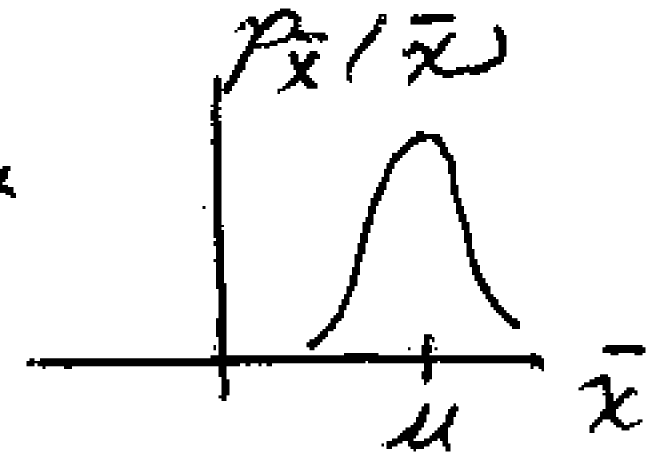
NOW CONSIDER $E[\bar{x}]$ AND $VAR[\bar{x}]$.

$$E[\bar{x}] = E_{\underline{x}}\left[\frac{1}{N}\sum_{i=1}^{N} x_i\right] = \frac{1}{N}\sum_{i=1}^{N} E_{\underline{x}}[x_i] \quad \text{WHY?}$$

$$= \frac{1}{N}\sum_{i=1}^{N} E_{x_i}[x_i] = \frac{1}{N}\sum_{i=1}^{N} \mu \quad \text{IDENTICALLY DIST.}$$

$\uparrow$ WHY? $= \mu$

SAYS MEAN OF PDF OF $\bar{x}$ IS μ

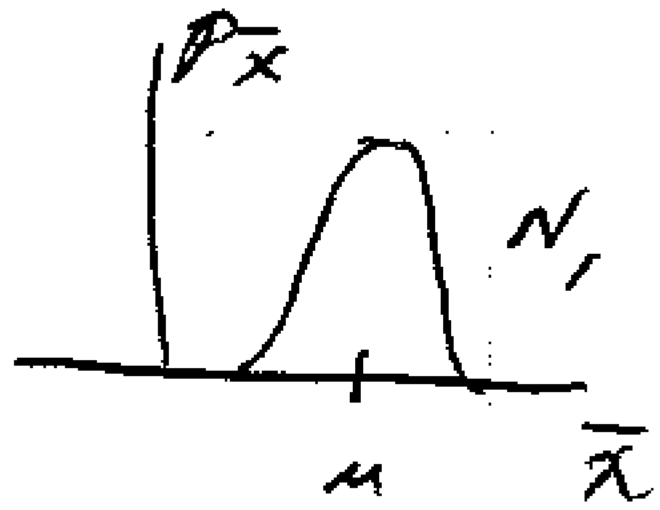


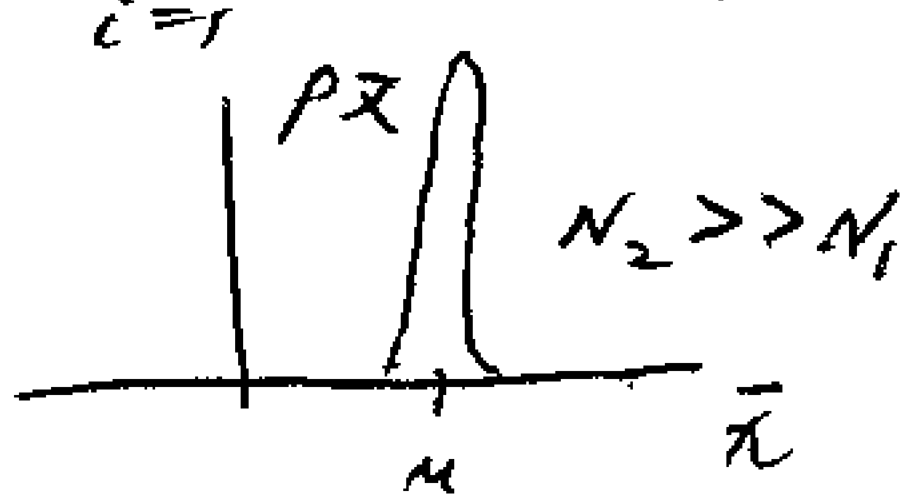
TO FIND VARIANCE USE

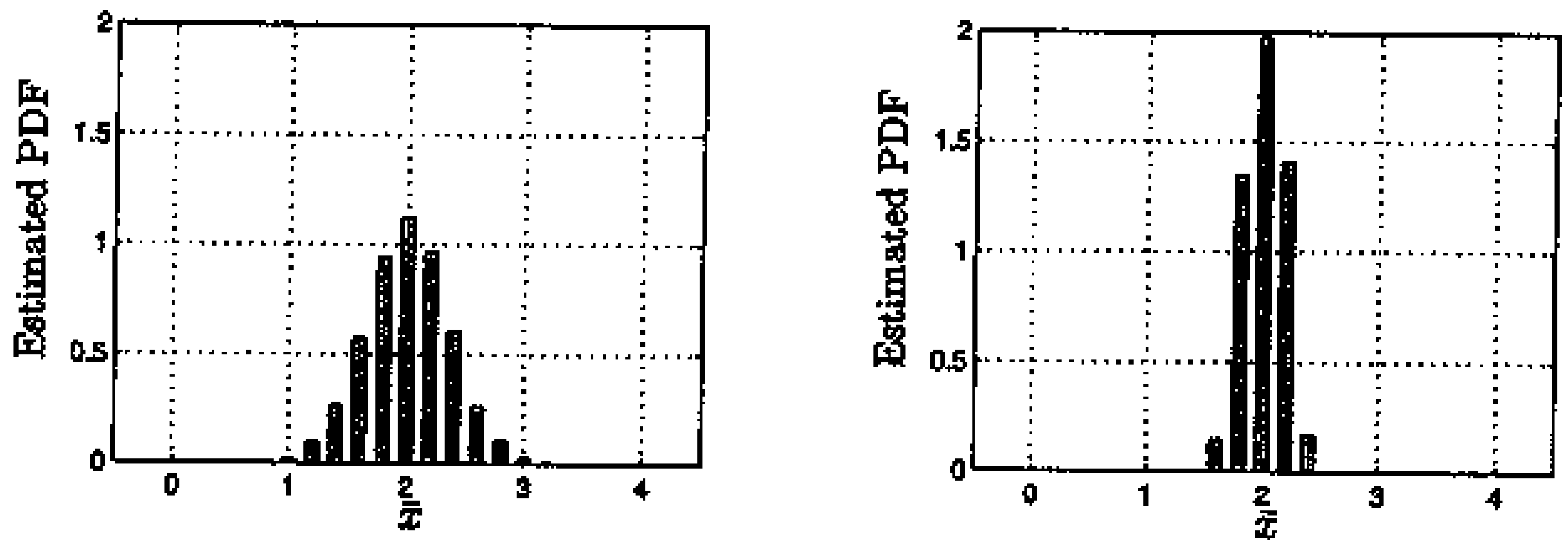
$$VAR\left(\sum_{i=1}^{N} a_i x_i\right) = \sum_{i=1}^{N} a_i^2 \, VAR(x_i) \quad x_i\text{'s UNCORRELATED?}$$

FOR $a_i = 1/N$, $VAR(\bar{x}) = \sum_{i=1}^{N} \frac{1}{N^2} \underbrace{VAR(x_i)}_{\sigma^2}$

$$= \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2 = \sigma^2/N \to 0 \quad \text{AS } N \to \infty$$







(a) $N = 10$ (b) $N = 100$

Figure 14.2: Estimated PDF for sample mean random variable, $\bar{X}$.

TRY SIMULATION
USE
x = 2+randn(N,1)
xbar = mean(x)
RUN THIS 1000
TIMES AND
USE hist

COMPUTER SIMULATION OF $\underline{X} \sim N(\underline{\mu}, \underline{C})$

1. PERFORM CHOLESKY DECOMPOSITION OF $\underline{C}$
   $\Rightarrow \quad \underline{C} = \underline{G}\,\underline{G}^T$
2. GENERATE REALIZATION $\underline{u}$ $(N \times 1)$
   WHERE $\underline{U} \sim N(\underline{0}, \underline{I})$ (u = randn(N,1))

3) FORM $\underline{x} = \underline{G}\,\underline{u} + \underline{\mu}$

SEE MATLAB CODE, PG. 476 FOR N=3

CHAPTER 15 - LIMIT THEOREMS

1) LAW OF LARGE NUMBERS - JUSTIFIES RELATIVE FREQUENCY INTERPRETATION OF PROBABILITY

EXAMPLE: TOSS FAIR COIN 1,000,000 TIMES AND COUNT NUMBER OF HEADS

FIND $\dfrac{\text{NUMBER OF HEADS}}{1{,}000{,}000} = 0.499$ THIS IS RELATIVE FREQUENCY

ASSIGN $P[\text{HEADS}] = \frac{1}{2}$ OR

$P(\text{HEADS}) \approx$ RELATIVE FREQUENCY
$=$ RELATIVE FREQUENCY AS NUMBER OF TOSSES $\to \infty$

2) <u>CENTRAL LIMIT THEOREM</u> – JUSTIFIES ASSUMPTION OF GAUSSIAN PDF OR "BELL SHAPED" DISTRIBUTION OF OUTCOMES

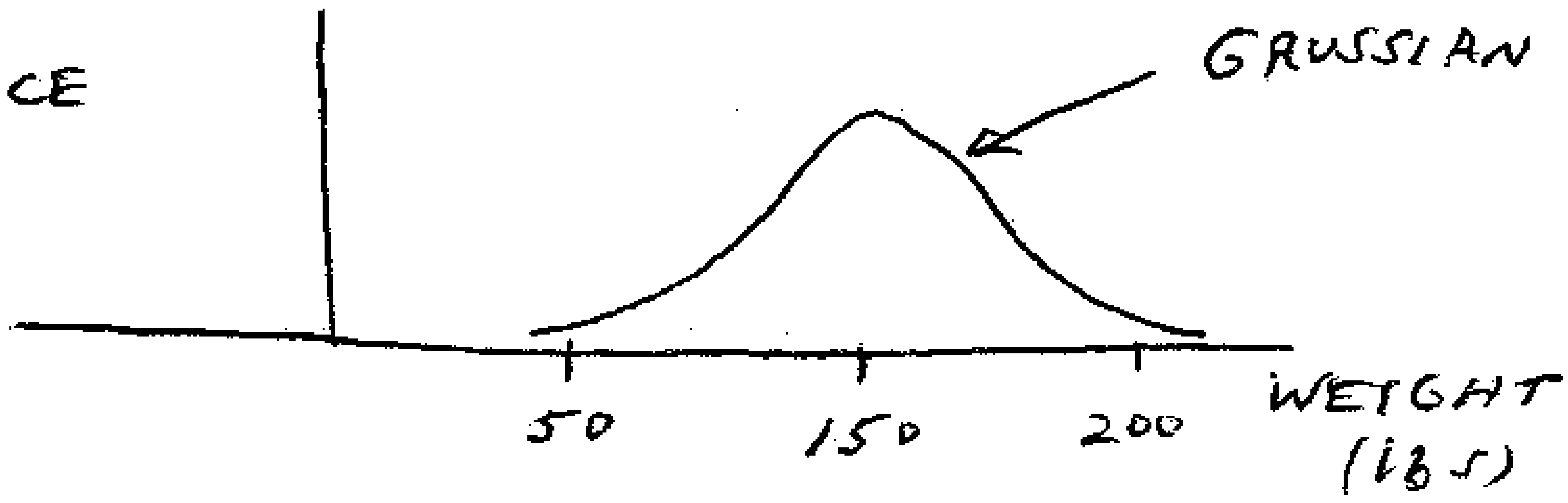
<u>EXAMPLE</u>: PERSON CHOSEN AT RANDOM
MEASURE WEIGHT

W = GENETIC lbs + JOB STRESS + DIET + EDUCATION + ...

LARGE NUMBER OF FACTORS – ALL MODELED AS RANDOM VARIABLES SINCE A PRIORI DON'T KNOW WHO IS CHOSEN

CENTRAL LIMIT THEOREM SAYS
$X_1 + X_2 + \ldots + X_N \sim$ GAUSSIAN AS $N \to \infty$
FOR X_i'S IID

IN PRACTICE



LAW OF LARGE NUMBERS

CONSIDER REPEATED TOSSING OF FAIR COIN.
OUTCOMES ARE $x_1, x_2, \ldots, x_N, \ldots$
WHERE

$$X_i = \begin{cases} 1 & \text{IF HEADS} \\ 0 & \text{IF TAILS} \end{cases}$$

ASSUME X_i's ARE IID -

ACTUALLY PROBS. SAYS OUTCOME OF ANY TOSS NOT DEPENDENT ON ANY OTHER TOSS ⇒ IND. AND SAME COIN USED AND TOSSED THE SAME WAY ⇒ IDENTICALLY DIST.

SETUP : OVERALL EXPERIMENT YIELDS A RANDOM OUTCOME MODELED BY

$$\underline{X} = [\underbrace{X_1 \, X_2 \, \ldots \, X_N}_{\text{IID}}]^T \quad (\text{WILL LET } N \to \infty)$$

$$p_{X_i}[k] = p_X[k] = \begin{cases} \frac{1}{2} & k=0 \\ \frac{1}{2} & k=1 \end{cases} \qquad \text{PMF}$$

LET $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$

NOTE THAT FOR ANY N, $\bar{x}$ IS A DISCRETE RV, FOR EXAMPLE,

$$N = 2 \qquad \bar{x} = \frac{1}{2}(x_1 + x_2)$$
$$= 0, \frac{1}{2}, 1$$

THUS, WE HAVE A PMF FOR $\bar{x}$. COMPUTING ITS MEAN AND VARIANCE

$$E_{\underline{x}}[\bar{x}] = E_{\underline{x}}\left[\frac{1}{N}\sum_{i=1}^{N} x_i\right] = \frac{1}{N}\sum_{i=1}^{N} E_{\underline{x}}[x_i]$$
$$= \frac{1}{N}\sum_{i=1}^{N} E_{x_i}[x_i] = \frac{1}{N}\sum_{i=1}^{N}\left(0\cdot\frac{1}{2} + 1\cdot\frac{1}{2}\right)$$
$$= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2} = \frac{1}{2}$$

$$\mathrm{VAR}(\bar{x}) = \mathrm{VAR}\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{VAR}(x_i) \quad \text{IND} \Rightarrow \text{UNCORR.}$$

$$\mathrm{VAR}(x_i) = E_{x_i}[x_i^2] - E^2_{x_i}[x_i]$$
$$= 0^2\cdot\frac{1}{2} + 1^2\cdot\frac{1}{2} - \left(\frac{1}{2}\right)^2$$
$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\mathrm{VAR}(\bar{x}) = \frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{4} = \frac{1/4}{N} \to 0 \text{ AS } N \to \infty$$

HENCE PMF "COLLAPSES" ABOUT $E_{x_i}[x_i] = \frac{1}{2}$.

SOMETIMES CALLED THE <u>BERNOULLI LAW OF LARGE NUMBERS</u> SINCE $x_i \sim$ BERNOULLI

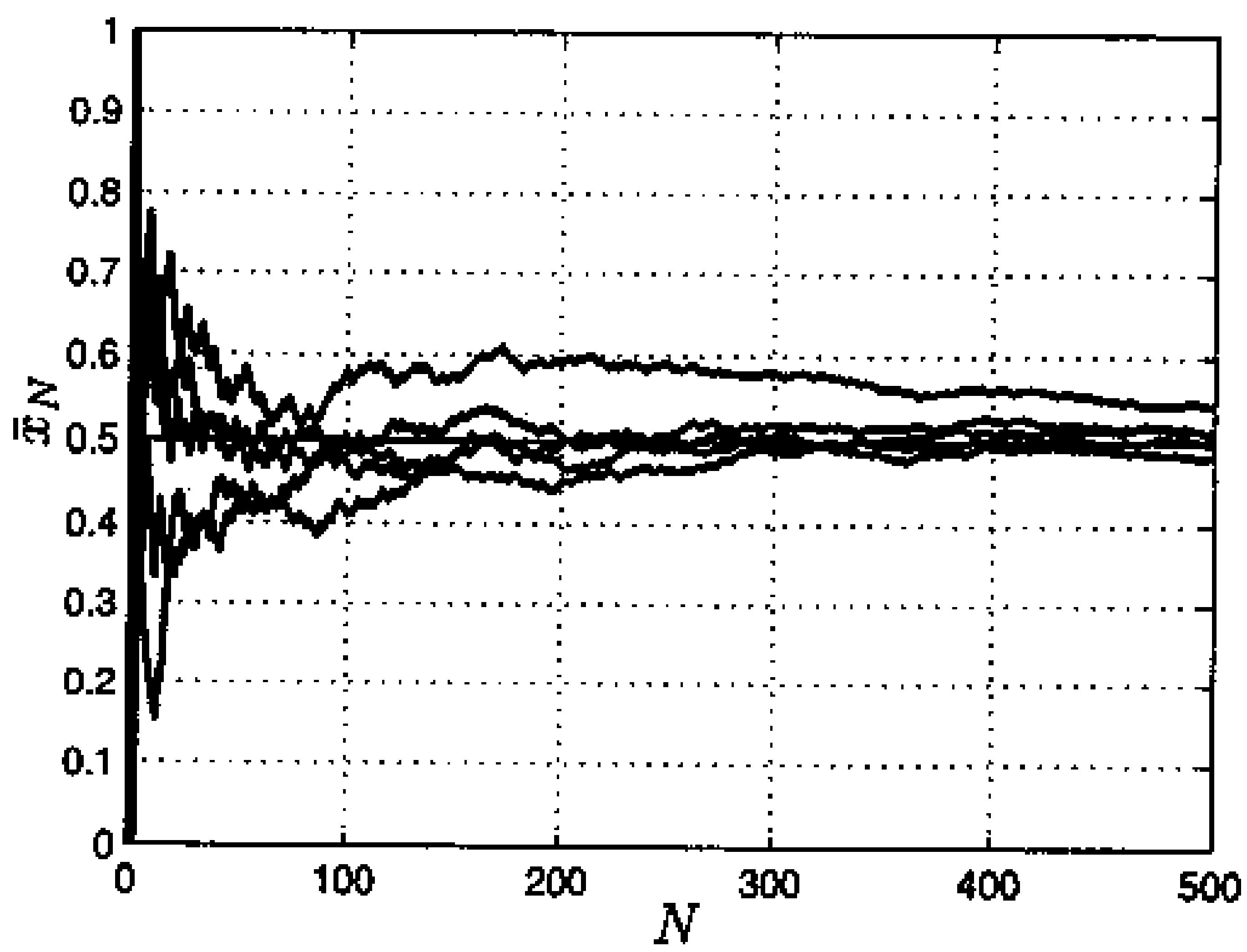


Figure 15.3: Realizations of sample mean random variable of N IID Bernoulli random variables with $p = 1/2$ as N increases.

LET $\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i$,

$\bar{X}_N \to \frac{1}{2} = E_X[X]$ AS $N \to \infty$

IN GENERAL THE LAW OF LARGE NUMBERS ASSERTS FOR IID X_i'S

$$\frac{1}{N} \sum_{i=1}^{N} X_i \to E_X[X] \quad \text{AS } N \to \infty$$

JUSTIFIES OUR USE OF SAMPLE MEAN AS AN ESTIMATOR OF $E_X[X]$ (PG. 58 OF NOTES) OR IN GENERAL

$$\hat{E[g(X)]} = \frac{1}{N} \sum_{i=1}^{N} g(x_i)$$

ACTUAL THEOREM IS:

IF $X_1, X_2, \ldots, X_N$ ARE IID WITH MEAN $E_X[X]$ AND VARIANCE $\sigma^2 < \infty$, THEN

$$\lim_{N \to \infty} \bar{X}_N = E_X[X].$$

MORE SPECIFICALLY WE MEAN

$$\lim_{N \to \infty} P[\underbrace{|\bar{X}_N - E_X[X]|}_{\text{DEVIATION OR ERROR}} > \epsilon] = 0$$

FOR ANY $\epsilon > 0$ (MAINLY ϵ CLOSE TO ZERO)

THIS TYPE OF CONVERGENCE CALLED

<u>CONVERGENCE IN PROBABILITY</u>

<u>PROOF</u>: $P[|\bar{X}_N - E_X[X]| > \epsilon] = P[|\underset{\uparrow\, Y}{\bar{X}_N} - \underset{\uparrow\, E[Y]}{E_X[\bar{X}_N]}| > \epsilon]$

RECALL CHEBYSHEV'S INEQUALITY

$$P[|Y - E[Y]| > \epsilon] \leq \frac{VAR(Y)}{\epsilon^2}$$

HERE $VAR(Y) = VAR(\bar{X}_N) = \sigma^2/N$

$$P[|\bar{X}_N - E_X[X]| > \epsilon] \leq \frac{\sigma^2/N}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2}$$

$$\lim_{N \to \infty} P[|\bar{X}_N - E_X[X]| > \epsilon] \leq \lim_{N \to \infty} \frac{\sigma^2}{N\epsilon^2} = 0$$

$$\Rightarrow \lim_{N \to \infty} P[|\bar{X}_N - E_X[X]| > \epsilon] = 0$$

EXAMPLE: RECALL DETECTION EXAMPLE WHEN SIGNAL $S = A$ IS PRESENT

$$X_i = S + W_i = A + W_i \quad i = 1, 2, \ldots$$

FOR W_i's IID WITH $E_W[W] = 0$ AND $VAR(W) = \sigma^2 < \infty$

$$\bar{X}_N \to E_X[X] = A \quad \text{AS } N \to \infty$$

WHEN NO SIGNAL IS PRESENT $X_i = W_i$ AND

$$\bar{X}_N \to E_X[X] = E_W[W] = 0$$

CENTRAL LIMIT THEOREM

LAW OF LARGE NUMBERS TELLS US ABOUT WIDTH AND LOCATION OF PDF/PMF OF $\bar{X}_N$

WIDTH $\to 0$ LOCATION $\to E_X[X]$.

HOW ABOUT PDF AS $N \to \infty$?

EXAMPLE: LET X_i's BE IID WITH

$$X_i \sim U(-\tfrac{1}{2}, \tfrac{1}{2})$$

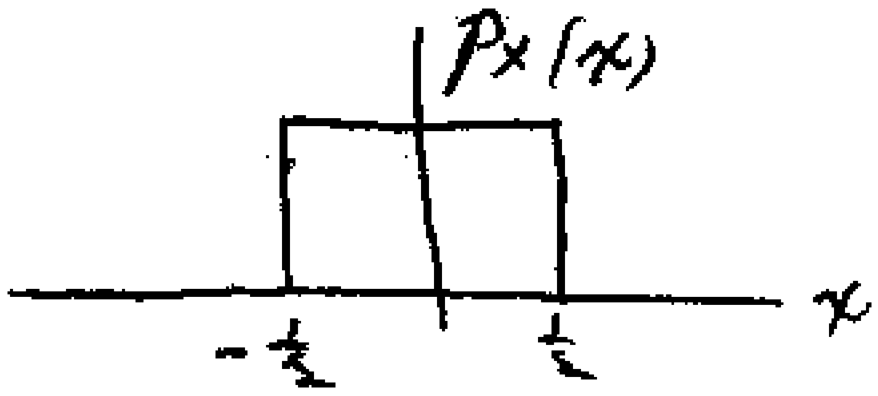
CONSIDER $S_N = \sum_{i=1}^{N} X_i$ NOW.

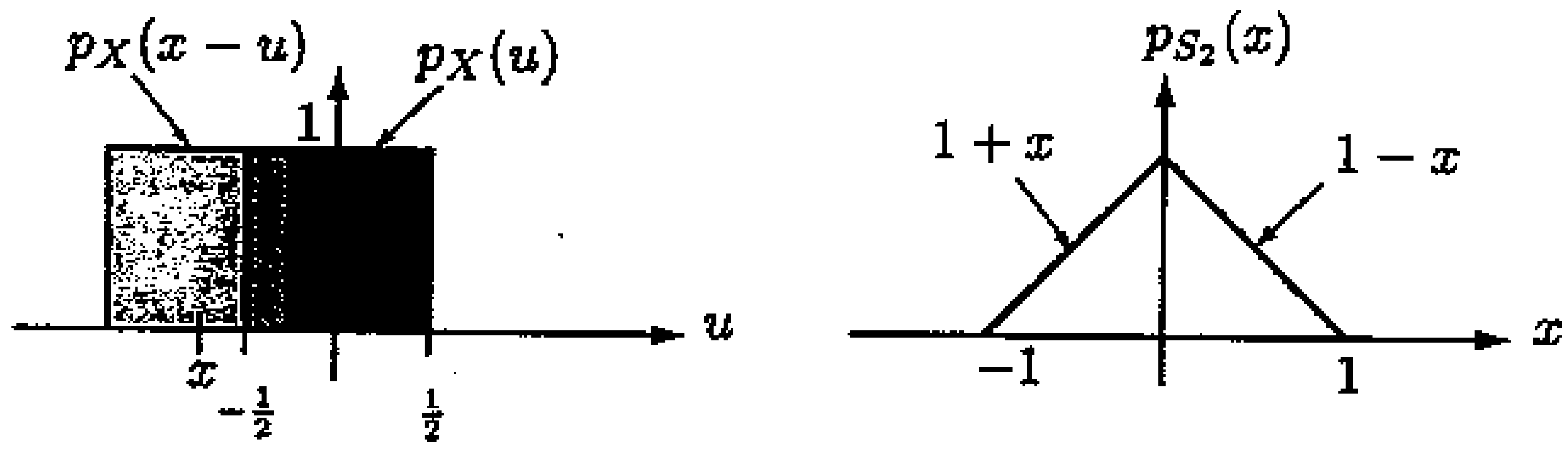
RECALL THAT PDF OF $X_1 + X_2$ IS OBTAINED BY CONVOLUTION OF PDFS

$$\Rightarrow \quad p_{S_2}(x) = p_X(x) \star p_X(x)$$

$$= \int_{-\infty}^{\infty} p_X(u)\, p_X(x-u)\, du$$

HERE





(a) Cross-hatched region contributes to integral

(b) Result of convolution

Figure 15.4: Determining the PDF for the sum of two independent uniform random variables using a convolution integral evaluation.

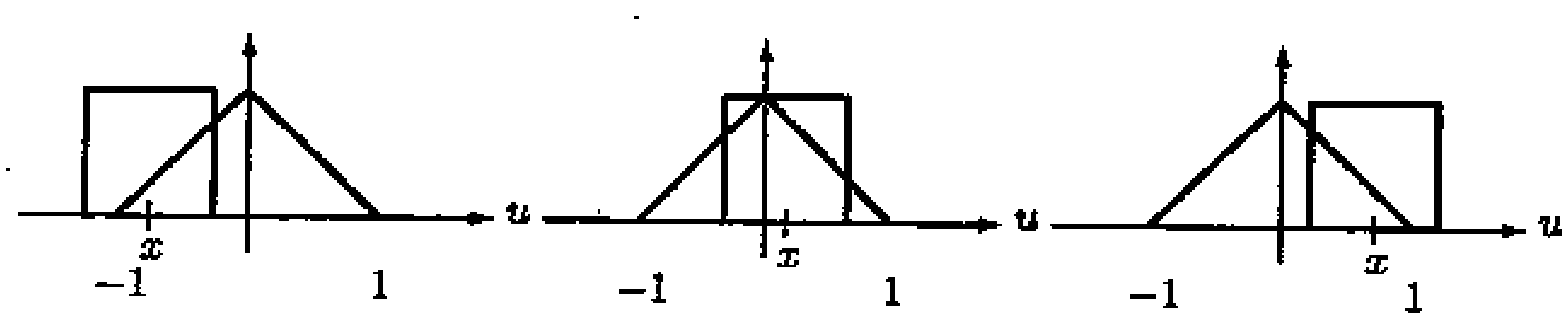
TO FIND PDF OF S_3 NOTE THAT

$$p_{S_3} = \underbrace{p_X \star p_X}_{p_{S_2}} \star p_X$$

$$\Rightarrow \quad p_{S_3}(x) = \int_{-\infty}^{\infty} p_{S_2}(u)\, p_X(x-u)\, du$$

KNOWN FROM ABOVE RESULT

$$= \int_{-\infty}^{\infty} p_{S_2}(u)\, p_X(u-x)\, du \qquad \text{WHY?}$$



(a) $-3/2 \le x \le -1/2$

(b) $-1/2 < x \le 1/2$

(c) $1/2 < x \le 3/2$

Figure 15.5: Determination of limits for convolution integral.