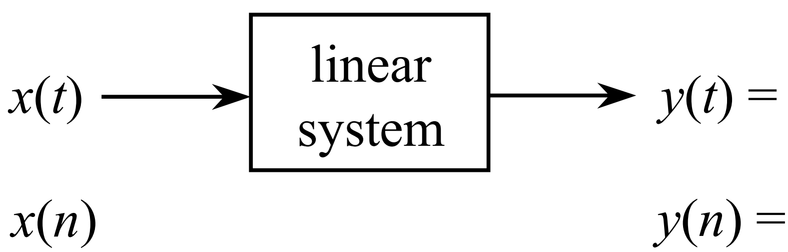
I. Signals and Systems Survival Kit

- Signals and systems (deterministic)
    - linear system
    - causal system
    - Z-transform and Fourier transform
- Signals and systems (random inputs)
  - basic probability concepts
  - correlation & covariance definition
  - statistical characterization of random signals
    - stationarity, wss, ergodicity, iid
    - white noise definition
  - statistical moments
  - correlation definition & properties for wss processes
  - how to compute correlation estimates
- Basic density functions and related properties
  - Gaussian, Q function, CLT
  - Rayleigh, Cauchy, Uniform, Chi-squared,
  - Non-central chi-squared
  - Monte Carlo performance

Signals and Systems
Survival Kit

I. Signals and Systems (deterministic review)

❖ Linear system



❖ Causal system

- Definition: a system is said to be causal if:

- Definition: a linear system is said to be causal if:

❖ Z-transform and Fourier transform

$$Z\left[x(n)\right]=$$

$$F_T\left[x(n)\right]=$$

- Applications to linear systems:

$$Z\left[y(n)\right] =$$

$$F_T\left[y(n)\right] =$$

- Why are Z-transform and Fourier transforms useful?

II. Signals and Systems (random input review)

❖ *Basic probability concepts*

- Distribution function (CDF)

- CDF Properties
  - –
  - –
  - –
  - –

- Probability density function (pdf) $f_x(x) =$
  - –
  - –

❖ *(Cross) Correlation and covariance function definitions*

$$R_{xy}\left(\tau_1, \tau_2\right) =$$

$$R_{xy}\left(k, l\right) =$$

$$C_{xy}\left(\tau_1, \tau_2\right) =$$

$$C_{xy}\left(k, l\right) =$$

❖ *Statistical Characterization of Random Signals*

- Random signals are characterized by joint distribution (or density) of samples
- $F_x(x_1, x_2, \ldots, x_k, n_1, \quad n_k)$

$$= \Pr\,[x(n_1) \leq x_1, \ldots\, x(n_k) \leq x_k]$$

- F(.) is highly complex to compute - difficult to obtain in practice

❖ <u>Stationarity</u>:

Definition: a RP is said to be stationary if any joint density or distribution function depends only on the spacing between samples, <u>not</u> where in the sequence the samples occur

<u>Example</u>: $f_x(x_1, x_2, \ldots, x_N; n_1, \ldots, n_N)$

$$= f_x(x_1, x_2, \ldots, x_N; n_{1+k} \ldots, n_{N+k})$$

for any value of k

If $x(n)$ is stationary for all orders $N = 1, 2, \ldots$

$x(n)$ is said to be strict-sense stationary.

Example: Stationary up to order 2 $\rightarrow$ wide-sense stationary.

❖ Wide-Sense Stationarity:

- *Definition*: a RS $x(n)$ is called wide-sense stationary (WSS) if

(1) the mean is a constant independent of "n"

(2) the autocorrelation depends only on the distance $\ell = n_1 - n_2$ (i.e., x(n) is a seq. of uncorrelated RVs)

Consequence: the variance is a constant independent of "n"

❖ Wide-Sense Stationarity (con't):

- *Definition:* $x[n]$ and $y[n]$ are said to be w.s. jointly stationary if:

  1) $x[n]$ and $y[n]$ are wss stationary

  2) $R_{xy}[n_1, n_0] = R_{xy}[n_1 - n_0]$

- When $x[n]$ and $y[n]$ are w.s.j stationary:

  $R_{xy}[n_1, n_0] = R_{xy}[n_1 - n_0] =$

  $C_{xy}[n_1, n_0] = C_{xy}[n_1 - n_0] =$

- Properties:

  $R_{xy}(k)=$

  $C_{xy}(k)=$

❖ Example

Let $x(n)$ be a real valued process of independent variables each with mean m and variance σ^2_x.

1) Compute: $R_x(k,n)$ & $C_x(k,n)$
2) Let $y(n)$ be defined as:

$$y(n)=x(n)+x(n-1)$$

Compute: $R_y(k,n)$ & $C_y(k,n)$

❖ Signal (time) Average:

$<x[n]>=$

❖ Ergodicity:

- in many applications only one realization of a RP is available
- in general, one single member doesn't provide information about the statistics of the process
- except when process is stationary +ergodic: statistical information can be derived from one realization of RP
- ***Def:*** a RP is called ergodic if:

  all ensemble averages = all corresponding time averages

- Ergodicity in the mean:

  Def: a RP is said to be ergodic in the mean if:

- Ergodicity in correlation:

  Def: a RP is said to be ergodic in correlation if:

❖ Example: Independent, Identically Distributed (I.I.D.) Random Process (RP):

A Random Process is said to be:

. an *independent process* if:
$f_x(x_1, x_2, \ldots, x_k; n_1, \ldots, n_k) = f_1(x_1; n_1) \ldots f_k(x_2; n_k)$

. if all RVs have the same pdf $f(x)$ => $x(n)$ is called I.I.D.

Note: I.I.D. processes have no memory, where a future value would depend on past values

- Mean of I.I.D. Process:

$$m_x(n) =$$

Autocovariance:

$$C_x(n_1, n_2) =$$

Autocorrelation:

$$R_x(n_1, n_2) =$$

❖ RP Example: White noise

Definition: A random sequence $w(n)$ is called a white noise process with mean m_w and variance σ^2_{w} iff

$$E\{w(n)\}=m_w$$
$$R_w(k)= \sigma^2_w(k)$$

Notes:

1) all frequencies contribute the same amount (as in the case of white light, therefore the name of "white noise")

2) if the pdf of $w(n)$ is Gaussian: it is called "white Gaussian noise"

❖ *Statistical moments*

- pdf information summarized by key aspects called statistical averages or moments

(1) <u>mean/average</u>

- $E\{x\} = m_x =$ if x is discrete

  $=$ if x is continuous

- important property of the mean $\rightarrow$ linearity!

  $E\{\alpha x + \beta\} =$

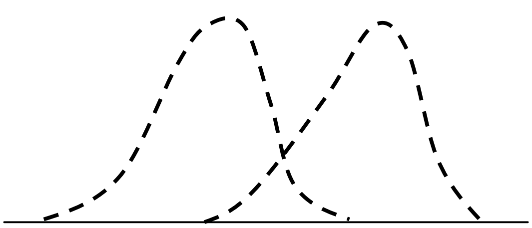
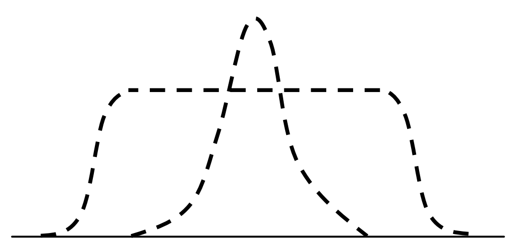
  $E\{g(x)\} =$

(3) <u>moments</u>

- $r_x^{(m)} = E\left\{x^m\right\} =$
- $\sigma_x^{(m)} = E\left\{\left|x - m_x\right|^m\right\} =$
- variance = σ_x^2
- variance property: $\sigma_x^2 = E\left[\left|x\right|^2\right] - \left(E\left[x\right]\right)^2$

  – proof:

❖ *Useful Moments:*

Skewness	Kurtosis
measures degree of asymmetry of distribution around the mean	measures relative flatness or peakedness of distribution about its mean
$k_x^{(3)} = E\left\{\left(\dfrac{x-m_x}{\sigma_x}\right)^3\right\} =$	$k_x^{(4)} = E\left\{\left(\dfrac{x-m_x}{\sigma_x}\right)^4\right\} - 3$
	

Note:

$\rightarrow\ k_x^{(4)} = 0$ for normal distribution

❖ *(Cross) Correlation and covariance function properties for wss processes*

$$R_{xy}(\tau) =$$

$$R_{xy}(k) =$$

$$C_{xy}(\tau) =$$

$$C_{xy}(k) =$$

❖ *Correlation and covariance functions main properties (for wss processes)*

❖ *How to compute correlation estimates*

Assume $\begin{cases} x\,(t) \text{ known } t = 0 \rightarrow t = T_0 \\ x\,(t) \text{ ergodic (why?)} \end{cases}$

- For discrete data: $\underline{x}$=[x (0), , x (N)]^T

  $\hat{R}_x(k\Delta t) =$

- Quality of estimate? → find mean and variance of $\hat{R}_x(k\Delta t) =$

  (1) $E\left[\hat{R}_x(k)\right] =$

$$(2)\ \ \mathrm{Var}\left[\hat{R}_x(k)\right] \cong \frac{N}{\left[N-k\right]^2}\sum_{i=-\infty}^{\infty}$$

$$\left[R_x^2(i) + R_x(i+k)\,R_x(i-k)\right]$$

$$\text{when } N >> k$$

Alternate Estimator: Biased Estimator

$$\widetilde{R}_x(k\Delta t) =$$

Quality of estimate:

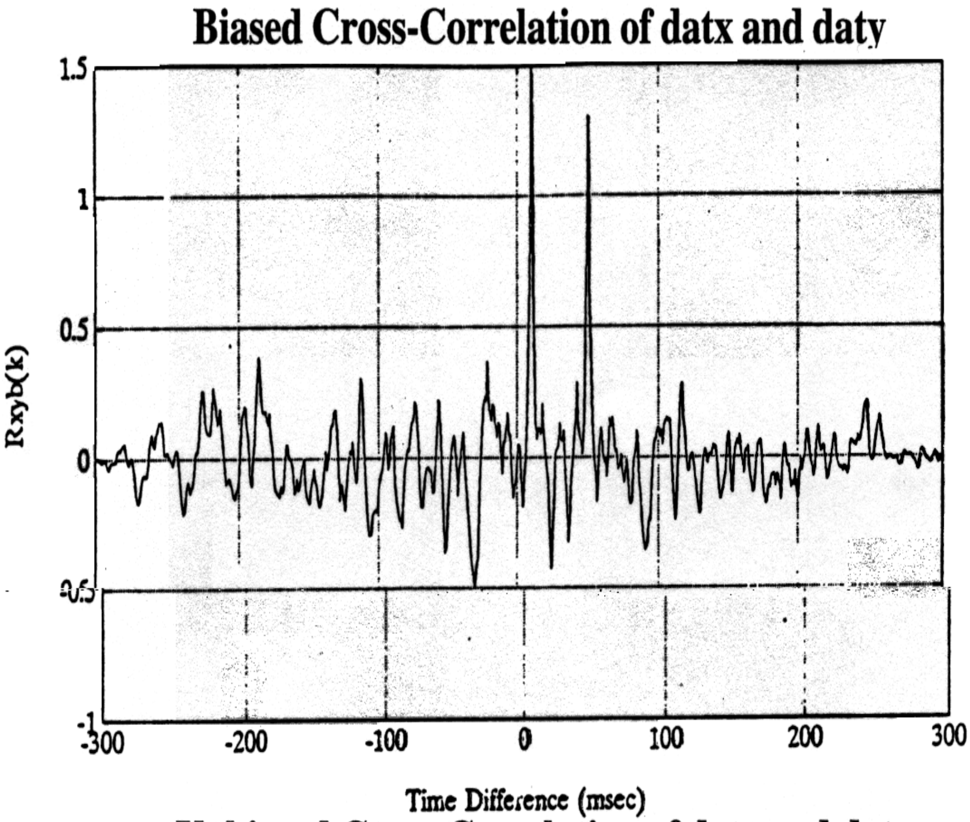
(1) $E\left[\tilde{R}_x\left(k\Delta T\right)\right] =$

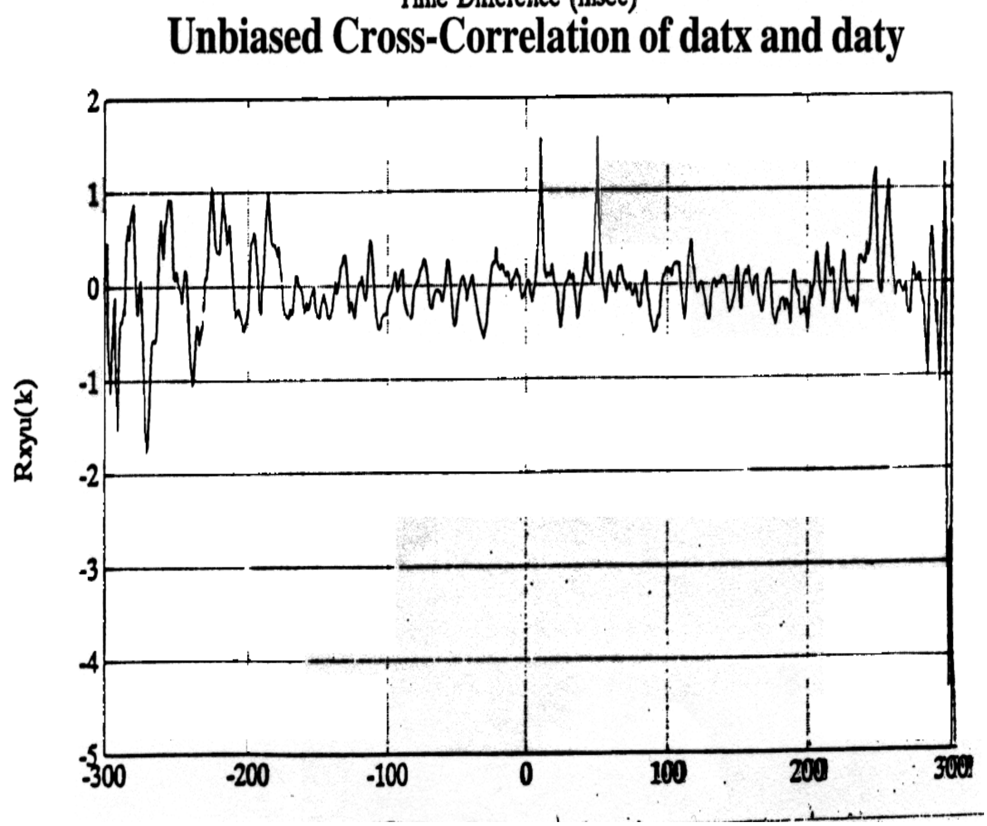
(2) $$\mathrm{Var}\left[\widetilde{R}_x(k\Delta T)\right] \cong \frac{1}{N+1}\sum_{-\infty}^{\infty} \left[R_x^2(i) + R_x(i+k)R_x(i-k)\right]$$

$$k > 0$$

Biased/Unbiased estimator Summary

Biased Estimator	Unbiased Estimator
$\tilde{R}_x(k\Delta T) = \dfrac{1}{N+1}\sum_{i=0}^{N-k} x(i)x(i+k)$	$\hat{R}_x(k\Delta T) = \dfrac{1}{N+1-k}\sum_{i=0}^{N-k} x(i)x(i+k)$
$E\left[\tilde{R}_x(k\Delta T)\right] = \dfrac{N-k+1}{N+1} R_x(k)$	$E\left[\hat{R}(k\Delta T)\right] = R_x(k)$
$\mathrm{Var}\left[\tilde{R}_x(k\Delta T)\right] = \dfrac{1}{N+1}\mathcal{L}(R_x)$	$\mathrm{Var}\left[\hat{R}(k\Delta T)\right] = \dfrac{N}{[N-k]^2}\mathcal{L}(R_x)$
when $k \to N$	
$\mathrm{Var}\left[\tilde{R}(k)\right] \to K < \infty$	$\mathrm{Var}\left[\hat{R}(k\Delta T)\right] \nearrow +\infty$
$E\left[\tilde{R}_x(k)\right] = \dfrac{1}{N+1} R_x(k)$	$E\left[\hat{R}(k)\right] \to R_x(k)$
when $N \to +\infty$	
$\mathrm{Var}\left[\tilde{R}_x(k)\right] \searrow$	
bias of $E\left[\tilde{R}_x(k)\right] \searrow$	


Biased Cross-Correlation of datx and daty
1.5
1
0.5
0
-0.5
-1
Rxyb(k)
-300
-200
-100
0
100
200
300
Time Difference (msec)

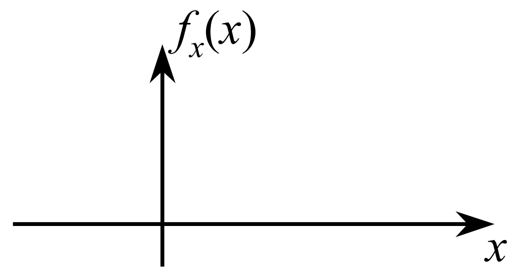

Unbiased Cross-Correlation of datx and daty
2
1
0
-1
-2
-3
-4
-5
Rxyu(k)
-300
-200
-100
0
100
200
300

III. Basic Density Functions and Related Properties

❖ *Gaussian density* $x \sim N\left(m_x, \sigma_x^2\right)$

- Real random variable

  $f_x(x) =$



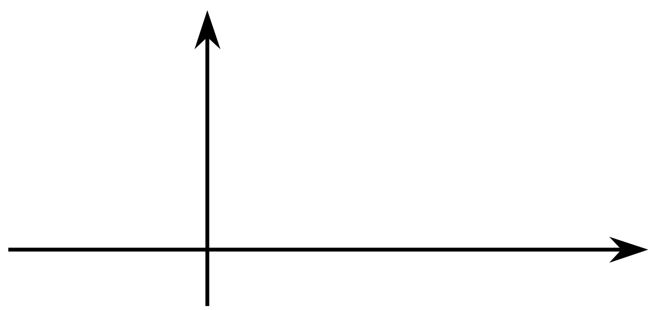
- Complex random variable

  $f_x(x) =$

- Gaussian density property

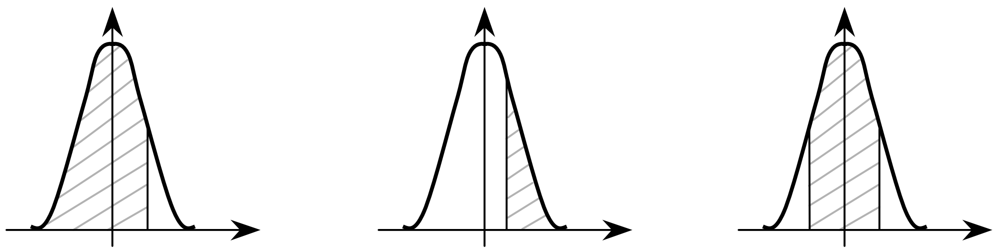
❖ *How to compute P(X > a) when X is Gaussian*

- $P(X > a) =$



- Q-function
  - $Q(y) =$

- How to apply the Q-function ?



Example:

The voltage X at the output of a noise generator is assumed to be $N(0, 1)$

Find $P(X > 2)$

$P(1 \le X \le 3)$

❖ *Central Limit Theorem (CLT):*

Describes the limiting behavior of the distribution function of a normalized sum of I.I.D. variables

Define:

$$z_n = \frac{s_n - nm}{\sigma\sqrt{n}}$$

$$where \quad s_n = \sum_{i=1}^{n} x_i; \; m = E[x_i]; \sigma^2 = \mathrm{var}[x_i]$$

As n gets large, $z_n \sim N(0,1)$

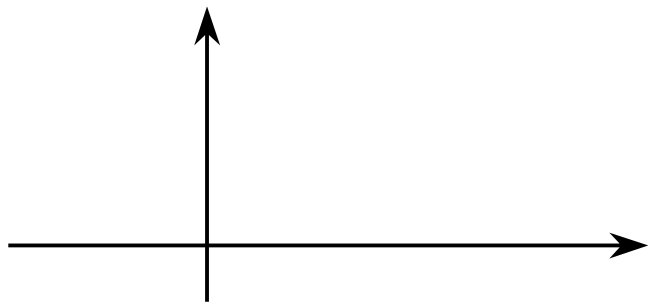
As n gets large, $s_n \sim N(nm, n\sigma^2)$

Example: Application of the CLT

Suppose orders at a restaurant are IID with a mean price m=\$8.00 and standard deviation σ=\$2.00.
Estimate the probability that the first 100 customers spend a total of more than \$840.00

❖ *Rayleigh density*

- $f_x(x) =$



- $E[x] =$

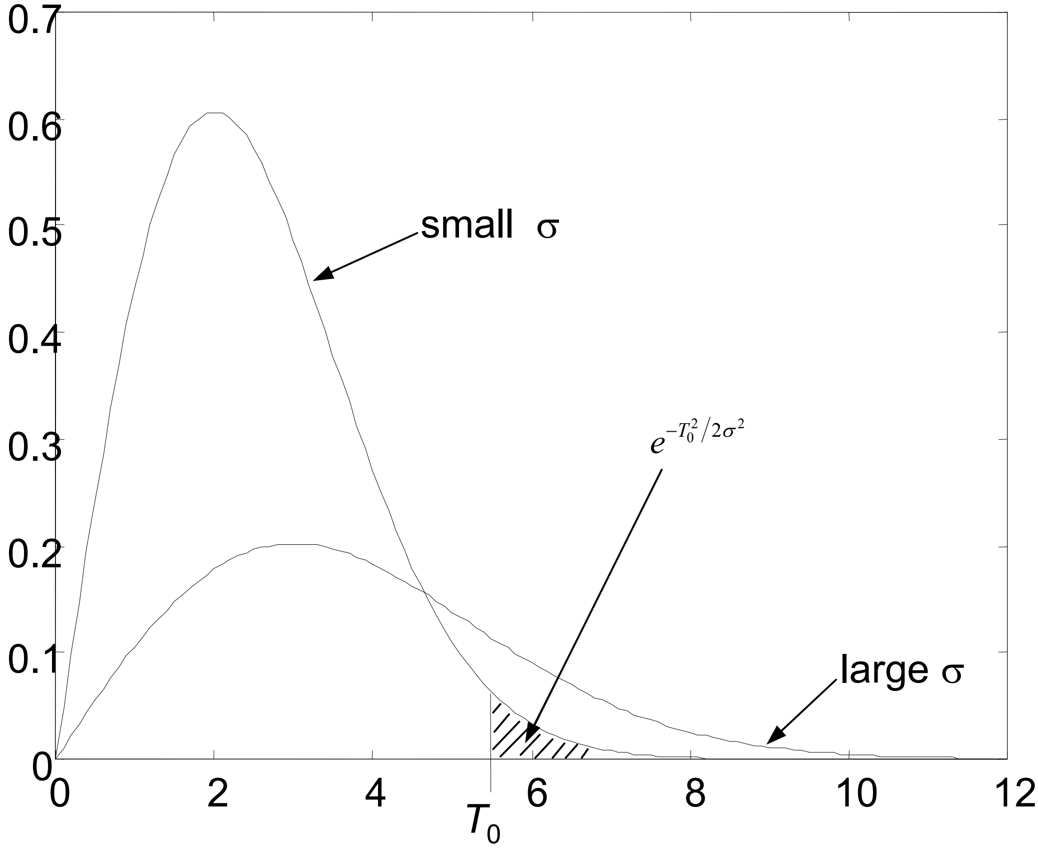
- $\sigma_x^2 =$

- Applications: Rayleigh densities are found in in communication applications when dealing with envelopes, etc...

Ex: $x \sim N\left(0, \sigma^2\right)$ x,y independent

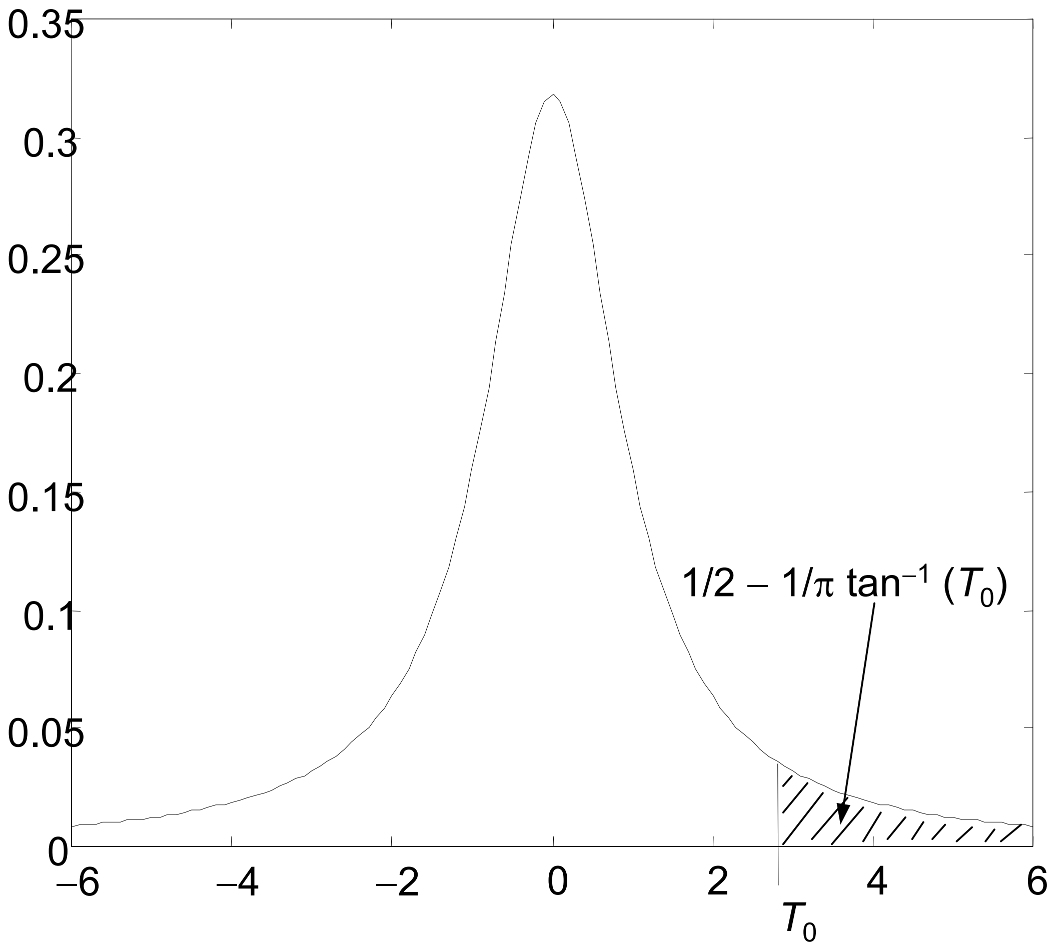
$y \sim N\left(0, \sigma^2\right)$

$f_{xy}\left(x, y\right) =$


small σ
$e^{-T_0^2/2\sigma^2}$
large σ
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
2
4
6
8
10
12
T_0

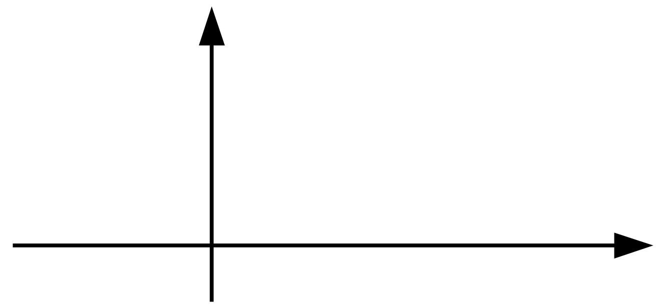
❖ *Cauchy density*

$$f_x(x) = \frac{1}{\pi\left(1+x^2\right)}$$



❖ *Uniform density*

- $f_x(x) =$



- $E\,[x] =$

- $\sigma_x^2 =$

❖ *Chi-squared density* $\left(\chi_N^2\right)$

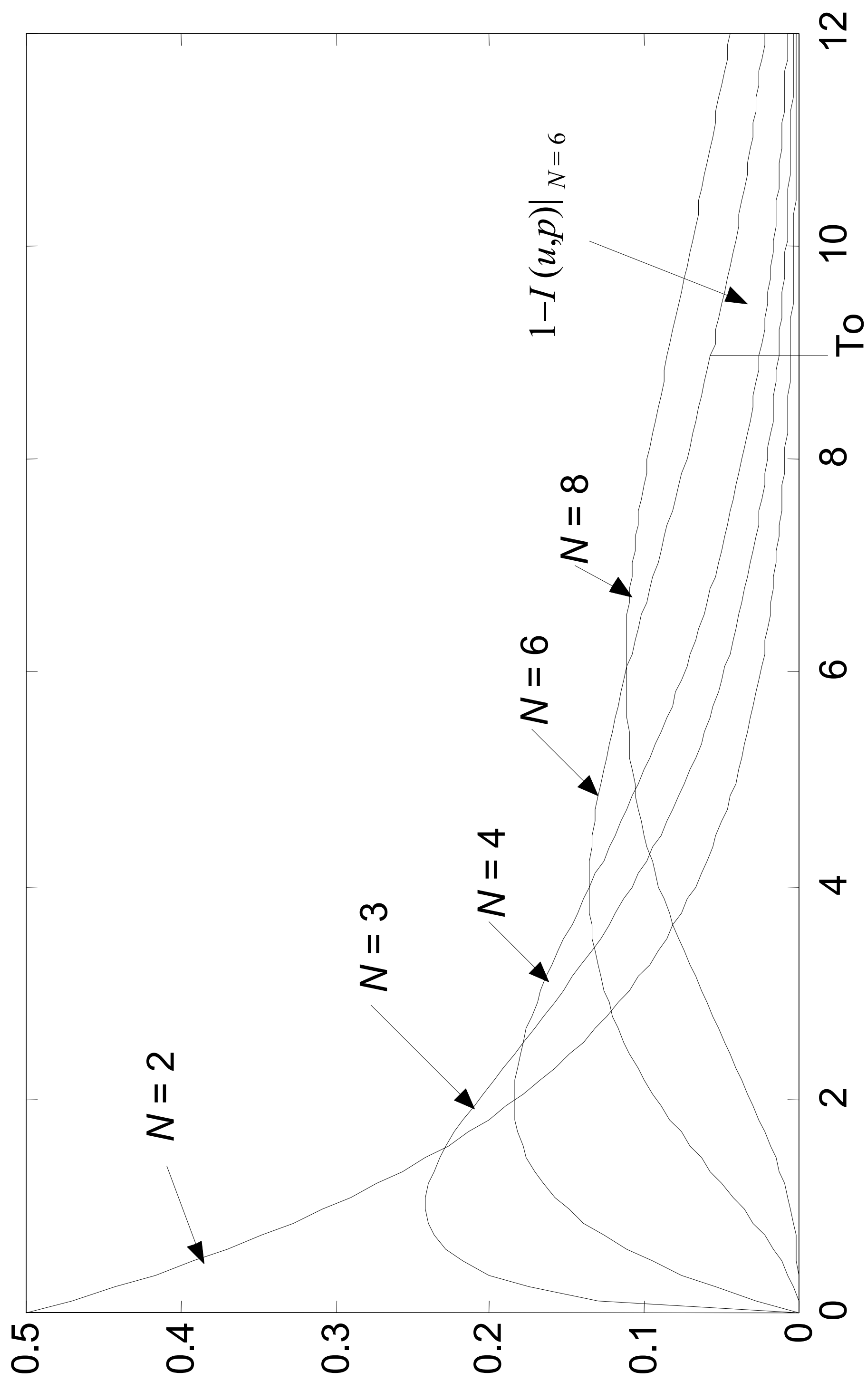
$$f_x\left(x\right)=\begin{cases}\dfrac{1}{2^{(N/2)}\Gamma\left(N/2\right)\sigma^N}x^{(N/2)-1}\exp\left(-\dfrac{1}{2\sigma^N}x\right); & x>0\\ 0; & x<0\end{cases}$$

N = number of degrees of freedom N ≥ 1)

$\Gamma(a)$ = Gamma function defined as:

$$\Gamma\left(a\right)=\int_0^\infty t^{a-1}\exp\left(-t\right)dt$$

- Pdf found when $x=\sum_{i=1}^{N}x_i^2 \quad x_i\sim N(0,\sigma^2)$


0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
N = 2
N = 3
N = 4
N = 6
N = 8
To
$1-I(u,p)|_{N=6}$

- Application:

  - Special case $N = 2$, $\sigma^2=1$

$$f_x(x) = \begin{cases} \frac{1}{2}\exp\left(-\frac{1}{2}x\right); & n > 0 \\ 0; & n < 0 \end{cases} \quad (\text{exponential pdf})$$

❖ *Non-central χ^2 density*

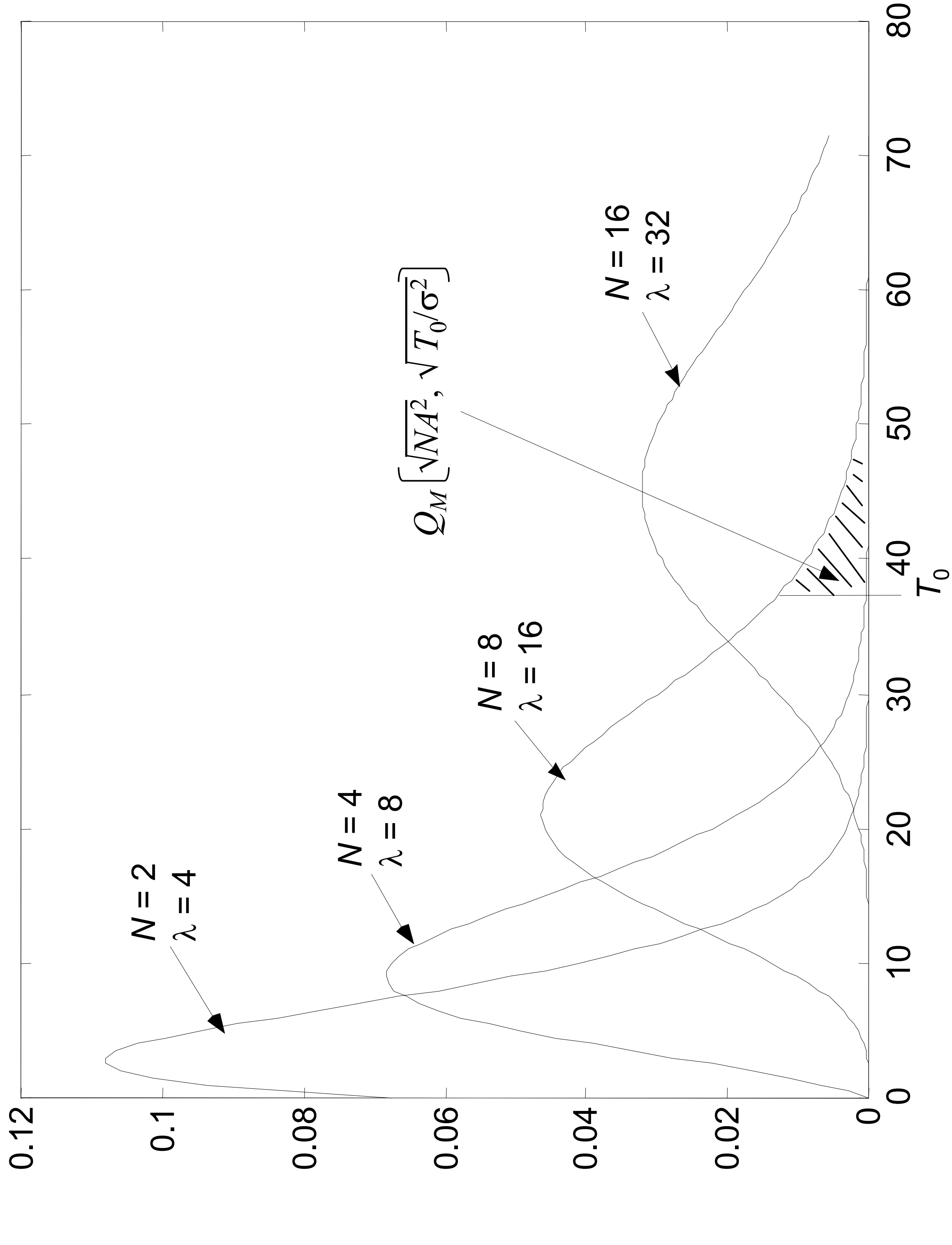
- Generalization of χ_N^2 density
- Found for

$$x = \sum_{i=1}^{N} x_i^2 ; \qquad x_i \sim N\left(A, \sigma^2\right) \quad \text{iid}$$

- Complete description:

$$f_x(x) = \frac{1}{2\sigma^2}\left(\frac{x}{\lambda}\right)^{(N-2)/4} \exp\left(-\frac{\lambda + x}{2\sigma^2}\right) I_{((N/2)-1)}\left[\frac{(x\lambda)^{1/2}}{\sigma^2}\right] U(x);$$

$\lambda = A^2 N$ <---- Non centrality parameter


$Q_M\left[\sqrt{NA^2},\ \sqrt{T_0/\sigma^2}\right]$
N = 2
λ = 4
N = 4
λ = 8
N = 8
λ = 16
N = 16
λ = 32
T_0
0
0.02
0.04
0.06
0.08
0.1
0.12
0
10
20
30
40
50
60
70
80

❖ *Monte Carlo performance evaluation*

- Computer evaluation of a probability
- Useful in cases where one cannot determine analytically or numerically expression of the form

$$P\ \{x > K\}$$

- Can be found in detection problems where we may wish to evaluate probability that a given statistic exceeds a threshold

Example:

– Assume we have a data set

$$\{x(0),\cdots,x(N-1)\}; \qquad x(n) \sim N\left(0,\sigma^2\right) \quad \text{iid}$$

– Assume we want to evaluate:

$$P\left\{\frac{1}{N}\sum_{n=0}^{N-1} x(n) > K\right\}$$

(1) Analytical derivation

(1) Monte Carlo derivation

(a) Generate data using

(b) Compute $T = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$

(c) Repeat (b) M times to yield M realizations of T

$$\{T_1, T_2, \cdots, T_M\}$$

(d) Probability evaluation:

1. Count the number of T_i's that exceed K: M_k
2. Estimate $P\{T > K\} = \frac{M_k}{M}$

- How to pick M ?

if a relative error

$$\varepsilon = \frac{|\text{true probability -- estimated probability}|}{\text{true probability}}$$

is desired 100 (1 – α) % of the time

we need: $$M \geq \frac{\left[Q^{-1}(\alpha/2)^2 (1-P)\right]}{\varepsilon^2 P}$$

where P = probability being estimated

Example: Determine $P\{T > 1\}$ with a relative error

$\varepsilon = 0.01\%$ for 95% of the time