ELEC4404 EXAMINATION FORMULA SHEET

Name	Pdf	Mean	Variance
Uniform	$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$	$\mu = \frac{a+b}{2}$	$\sigma^2 = \frac{(b-a)^2}{12}$
Gaussian, $N(\mu, \sigma^2)$	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Bernoulli	$p_X(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$	p	$p(1-p)$
Binomial, $B(n,p)$	$P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$P[X=k] = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ

A random process $X(t)$ is a **Gaussian random process** if the k-dimensional random variable vector $[X_1 \quad X_2 \quad \cdots \quad X_k]$, where $X_j = X(t_j)$ the joint pdf is given by:

$$f_{X_1 X_2 \cdots X_k}(x_1, x_2, \dots, x_k) = \frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{K}^{-1}(\mathbf{x}-\mathbf{m})}}{(2\pi)^{k/2}|\mathbf{K}|^{1/2}}$$

$$\mathbf{m} = \begin{bmatrix} m_X(t_1) \\ m_X(t_2) \\ \vdots \\ m_X(t_k) \end{bmatrix} \qquad \mathbf{K} = \begin{bmatrix} C_{XX}(t_1,t_1) & C_{XX}(t_1,t_2) & \cdots & C_{XX}(t_1,t_k) \\ C_{XX}(t_2,t_1) & C_{XX}(t_2,t_2) & \cdots & C_{XX}(t_2,t_k) \\ \vdots & \vdots & \ddots & \vdots \\ C_{XX}(t_k,t_1) & C_{XX}(t_k,t_2) & \cdots & C_{XX}(t_k,t_k) \end{bmatrix}$$

NOTE: Jointly *uncorrelated* Gaussian random variables are also *independent*.

Poisson process, $N(t)$, of rate λt: $P[N(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ for $k = 0,1,\dots$ where $E[N(t)] = \lambda t$, $\text{VAR}[N(t)] = \lambda t$ has: ***exponential interevent times***, ***uniformly distributed arrival times*** and possesses ***stationary and independent increments***: $P[N(t_1) = i, N(t_2) = j] = P[N(t_1) = i] P[N(t_2 - t_1) = j - i]$

Important Relations (for c constant)	Important Properties
1. $E[c] = c$ 2. $E[cX] = cE[X]$ 3. $E[c+X] = c + E[X]$ 4. $VAR[c] = 0$ 5. $VAR[cX] = c^2 VAR[X]$ 6. $VAR[c+X] = VAR[X]$	• If $E[XY] = 0$ then X and Y are **orthogonal.** • If $COV(X,Y) = 0$ then X and Y are **uncorrelated.** • If X and Y are **independent** then $E[XY] = E[X]E[Y]$, thus independent random variables are uncorrelated (since $COV(X,Y) = 0$).

pdf of a function of a random variable ($Y = g(X)$)

$$f_Y(y) = \sum_k \frac{f_X(x_k)}{|g'(x_k)|} \quad \text{where} \quad x_k = g^{-1}(y)$$

Expectation of $g(\mathbf{X})$

$$E[g(\mathbf{X})] = \iint_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} g(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x})$$

Important Identities and Formulae

$$2\cos A\cos B = \cos(A+B) + \cos(A-B) \quad 2\sin A\sin B = \cos(A-B) - \cos(A+B)$$

$$2\sin A\cos B = \sin(A+B) + \sin\ (A-B)$$

$$\cos(A \pm B) = \cos A\cos B \mp \sin A\sin B \qquad \sin(A \pm B) = \sin A\cos B \pm \cos A\sin B$$

$$\text{If } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ then } A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$VAR(X) = \sigma_X^2 = E[X^2] - E[X]E[X] \qquad COV(X,Y) = C_{XY} = E[XY] - E[X]E[Y] \qquad \rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}$$

$$S_{YX}(f) = H(f)S_X(f) \qquad S_Y(f) = |H(f)|^2 S_X(f) \qquad S_X(f) = \mathcal{F}\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau)e^{-j2\pi f\tau}d\tau$$

$$f_X(x) = \frac{dF_X(x)}{dx} \qquad F_X(x_1) = \int_{-\infty}^{x_1} f_X(x)dx$$

$$f_X(x|A) = \frac{P(A|X=x)f_X(x)}{P(A)} = \frac{P(A|X=x)f_X(x)}{\int_{-\infty}^{\infty} P(A|X=x)f_X(x)dx}$$

$$P[Y \text{ in } A|X=x] = \int_{y\in A} f_Y(y|x)dy \qquad P[Y \text{ in } A] = \int_{-\infty}^{\infty} P[Y \text{ in } A|X=x]f_X(x)dx$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy \qquad f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \qquad F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,v)dudv$$

$$\mathbf{R_X} = E[\mathbf{XX}^T] \qquad \mathbf{C_X} = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T]$$

Expectations of a Random Process

The mean of a random process:

$$m_X(t_0) = E[X(t_0)] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x f_{X(t_0)}(x)dx & \text{continuous-valued} \\ \displaystyle\sum_{k=m_L}^{k=m_H} x_k\, P[X_0 = x_k] & \text{discrete-valued} \end{cases}$$

The autocovariance of a random process:

$$C_{XX}(t_1,t_2) = E[\{X(t_1) - m_X(t_1)\}\{X(t_2) - m_X(t_2)\}] = R_{XX}(t_1,t_2) - E[X(t_1)]E[X(t_2)]$$

$$R_{XX}(t_1,t_2) = E[X(t_1)X(t_2)]$$

Wide-Sense Stationary (WSS)

1. $m_X(t) = E[X(t)] = m_X$ is independent of t
2. $R_{XX}(t_1,t_2) = E[X(t_1)X(t_2)] \equiv R_X(t_1 - t_2) = R_X(\tau)$ depends on the time difference $\tau = t_1 - t_2$.

Covariances

$$C_{XX}(t_1,t_2) = E[\{X(t_1) - m_X(t_1)\}\{X(t_2) - m_X(t_2)\}] = R_{XX}(t_1,t_2) - m_X(t_1)m_X(t_2)$$

$$C_{XY}(t_1,t_2) = E[\{X(t_1) - m_X(t_1)\}\{Y(t_2) - m_Y(t_2)\}] = R_{XY}(t_1,t_2) - m_X(t_1)m_Y(t_2)$$

<u>Power Spectral Density (PSD)</u>

$$S_X(f) = \mathcal{F}\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau)e^{-j2\pi f\tau}d\tau$$

$$R_X(\tau) = \mathcal{F}^{-1}\{S_X(f)\} = \int_{-\infty}^{\infty} S_X(f)e^{j2\pi f\tau}df$$

The PSD is a non-negative, real-valued and even function of f.

Response of Linear Systems to WSS Random Signals (in terms of power spectral densities and autocorrelation functions)

$$S_{YX}(f) = H(f)S_X(f) \qquad S_Y(f) = |H(f)|^2 S_X(f)$$

$$R_Y(\tau) = R_X(\tau) * \mathrm{h}(-\tau) * \mathrm{h}(\tau)$$

Spectral Factorisation

If we filter a white noise process $w(n)$ with variance σ^2 by a filter with transfer function $H(z)$ then we have the following spectral factorisation generating model for the output process $y(n)$:

$$S_y(z) = H(z^{-1})H(z)\sigma^2$$

Random binary wave process:

$$R_{XX}(t_i, t_k) = \begin{cases} A^2\left(1 - \frac{|t_k - t_i|}{T}\right) & |t_k - t_i| < T \\ 0 & |t_k - t_i| \geq T \end{cases}$$

Sinusoid with random phase: $R_{XX}(t_1, t_2) = \frac{A^2}{2}\cos(\omega_c(t_1 - t_2))$

Random Telegraph: $R_{XX}(t_1, t_2) = e^{-2\alpha|t_1 - t_2|}$

In **Welch's Method** the N-length realisation of the signal is partitioned into K sequences of length L which are offset by D samples, i.e. $x_i(n) = x(n + iD)$, such that $N = L + D(K - 1)$ and the K modified periodogram estimates are formed using the L-length subsequences and averaged, to yield the estimate:

$$\hat{P}_W(e^{j\omega}) = \frac{1}{KLU}\sum_{i=0}^{K-1}\left|\sum_{n=0}^{L-1} w(n)x(n + iD)e^{-jn\omega}\right|^2$$

where $U = \frac{1}{L}\sum_{n=0}^{L-1}|w(n)|^2$ and $w(n)$ is the window of length L. For $D = L/2$ (50% overlap) and a Bartlett window function it can be shown that:

$$\mathrm{var}\{\hat{P}_W(e^{j\omega})\} \approx \frac{9}{16}\frac{L}{N}P_x^2(e^{j\omega}) \approx \frac{P_x^2(e^{j\omega})}{K}$$

$$\mathrm{Res}\left[\hat{P}_W(e^{j\omega})\right] = 1.28\frac{2\pi}{L}$$

Important z-Transform Pairs

$$\delta(n) \leftrightarrow 1, \quad \text{All } z$$

$$a^n u(n) \leftrightarrow \frac{1}{1-az^{-1}}, \quad |z| > |a| \qquad a^n u(-n-1) \leftrightarrow \frac{a^{-1}z}{1-a^{-1}z}, \quad |z| < |a|$$

$$a^{|n|} \leftrightarrow \frac{1-a^2}{(1-az^{-1})(1-az)}, \quad |a| < |z| < |a^{-1}|$$

Infinite Series

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \quad , \quad \sum_{k=-\infty}^{\infty} r^{|k|} = \frac{1+r}{1-r} \qquad |r| < 1$$

$$\boldsymbol{\sum_{k=-1}^{-\infty} r^k = \frac{r^{-1}}{1-r^{-1}} \qquad |r| > 1}$$

2x2 Matrix Inversion

If $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $\mathbf{A}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ where $\det(\mathbf{A}) = ad - bc$

A. FOURIER TRANSFORM DEFINITION

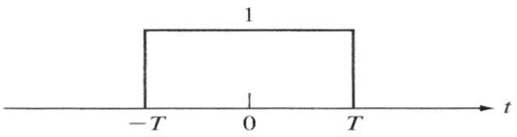
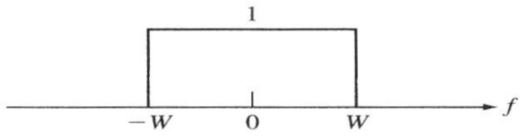
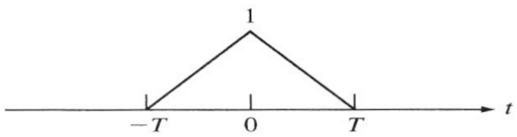
$$G(f) = \mathscr{F}\{g(t)\} = \int_{-\infty}^{\infty} g(t)e^{-j2\pi ft}\,dt$$

$$g(t) = \mathscr{F}^{-1}\{G(f)\} = \int_{-\infty}^{\infty} G(f)e^{j2\pi ft}\,df$$

B. PROPERTIES

Linearity:	$\mathscr{F}\{ag_1(t) + bg_2(t)\} = aG_1(f) + bG_2(f)$
Time scaling:	$\mathscr{F}\{g(at)\} = G(f/a)/\lvert a\rvert$
Duality:	If $\mathscr{F}\{g(t)\} = G(f)$, then $\mathscr{F}\{G(t)\} = g(-f)$
Time shifting:	$\mathscr{F}\{g(t - t_0)\} = G(f)e^{-j2\pi ft_0}$
Frequency shifting:	$\mathscr{F}\{g(t)e^{j2\pi f_0 t}\} = G(f - f_0)$
Differentiation:	$\mathscr{F}\{g'(t)\} = j2\pi fG(f)$
Integration:	$\mathscr{F}\left\{\int_{-\infty}^{t} g(s)ds\right\} = G(f)/(j2\pi f) + (G(0)/2)\delta(f)$
Multiplication in time:	$\mathscr{F}\{g_1(t)g_2(t)\} = G_1(f) * G_2(f)$
Convolution in time:	$\mathscr{F}\{g_1(t) * g_2(t)\} = G_1(f)G_2(f)$

C. TRANSFORM PAIRS

$g(t)$	$G(f)$
	$2T \sin 2\pi fT/(2\pi fT)$
$2W \sin(2\pi Wt)/2\pi Wt$	
	$T(\sin(\pi fT)/\pi fT)^2$
$e^{-at}u(t), \quad a > 0$	$1/(a + j2\pi f)$
$e^{-a\lvert t\rvert}, \quad a > 0$	$2a/(a^2 + (2\pi f)^2)$
$e^{-\pi t^2}$	$e^{-\pi f^2}$
$\delta(t)$	1
1	$\delta(f)$
$\delta(t - t_0)$	$e^{-j2\pi f t_0}$
$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$
$\sin(2\pi f_0 t)$	$(1/2j)\{\delta(f - f_0) - \delta(f + f_0)\}$

Maximum A Posteriori (MAP) estimate
Given observation $Y = y$ we can form an estimate of the desired or input $X = x$ by the MAP estimate:

$$\hat{x}_{MAP} = \underset{x}{\operatorname{argmax}} f_X(x|y) = \underset{x}{\operatorname{argmax}} P[X = x|Y = y]$$

Maximum Likelihood (ML) estimate
Given observation $Y = y$ we can form an estimate of the desired or input $X = x$ by the ML estimate:

$$\hat{x}_{ML} = \underset{x}{\operatorname{argmax}} f_Y(y|x) = \underset{x}{\operatorname{argmax}} P[Y = y|X = x]$$

Minimum MSE estimator is given by $g(.)$ as follows:

$$\hat{x}_{MMSE} = g^*(y) = \underset{g(.)}{\operatorname{argmin}} E[\left(X - g(Y)\right)^2] = E[X|Y = y]$$

Linear MMSE estimator is an unbiased estimate for X for observation Y:

$$\hat{X} = a^*Y + b^* = \rho_{XY}\left(\sigma_X \frac{Y - E[Y]}{\sigma_Y}\right) + E[X]$$

The **MSE** can be shown to be:

$$e^* = E[(X - (a^*Y + b^*))^2] = VAR[X] - a^*COV(X,Y) = VAR[X](1 - \rho_{XY}^2)$$

Discrete Wiener-Hopf Equations for WSS processes

$$\mathbf{R}\mathbf{h}_O = \mathbf{d}$$

or in matrix form (where $\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$ and $\mathbf{d} = E\{y(n)\mathbf{x}(n)\}$) and we define $r_x(k) = E\{x(n)x(n-k)\}$ and $r_{yx}(k) = E\{y(n)x(n-k)\}$):

$$= \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(M-1) \\ r_x(1) & r_x(0) & \cdots & r_x(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(M-1) & r_x(M-2) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} h_O(0) \\ h_O(1) \\ \vdots \\ h_O(M-1) \end{bmatrix} = \begin{bmatrix} r_{yx}(0) \\ r_{yx}(1) \\ \vdots \\ r_{yx}(M-1) \end{bmatrix}$$

That is, a time-invariant optimum FIR filter is implemented based upon the convolution:

$$\hat{y}_O(n) = \sum_{k=0}^{M-1} h_O(k)x(n-k) = \mathbf{h}_O^T\mathbf{x}(n)$$

where the filter co-efficients satisfy the *discrete-time Wiener-Hopf* equations:

$$\sum_{k=0}^{M-1} h_O(k)r_x(m-k) = r_{yx}(m) \qquad 0 \le m \le M-1$$

and the MMSE is given by:

$$P_O = P_y - \sum_{k=0}^{M-1} h_O(k)r_{yx}(k) = r_y(0) - \sum_{k=0}^{M-1} h_O(k)r_{yx}(k) = r_y(0) - \mathbf{h}_O^T\mathbf{d}$$

Mth order linear prediction of the ith sample

We want to predict or estimate $x(n-i)$ given known samples $x(n-k)$ by:

$$\hat{x}(n-i) = -\sum_{\substack{k=0 \\ k\neq i}}^{M} c_k x(n-k) = \mathbf{c}_i^T\mathbf{x}_i(n)$$

The predictor co-efficients, $\mathbf{c}_i(n)$ require solving:

$$\mathbf{R}_i\mathbf{c}_i = -\mathbf{r}_i, \quad \text{and the MMSE power is given by } P_O^{(i)}(n) = r_x(0) + \mathbf{r}_i^T\mathbf{c}_i$$

where:

$$\mathbf{R}_i = E\{\mathbf{x}_i(n)\mathbf{x}_i^T(n)\}, \;\; \mathbf{r}_i = E\{\mathbf{x}_i(n)x(n-i)\}$$

$$\mathbf{x}_i(n) = [x(n) \;\; x(n-1) \;\; \dots \;\; x\big(n-(i-1)\big) \;\; x(n-(i+1)) \;\; \dots \;\; x(n-M)]^T$$

$$\mathbf{c}_i = [c_0 \;\; c_1 \;\; \dots \;\; c_{i-1} \;\; c_{i+1} \;\; \dots \;\; c_M]^T$$

We have the <u>symmetric linear smoother (SLS)</u> if $i = L$ and $M = 2L$; the <u>forward linear predictor (FLP)</u> if $i = 0$, and the <u>backward linear predictor (BLP)</u> if $i = M$.

Optimum IIR Filter to estimate $y(n)$ given observations $x(n)$

	Non-causal	**Causal**
Design	$\hat{y}_O(n) = \sum_{k=-\infty}^{\infty} h_{nc}(k)x(n-k)$	$\hat{y}_O(n) = \sum_{k=0}^{\infty} h_c(k)x(n-k)$
$H(z)$	$H_{nc}(z) = \frac{R_{yx}(z)}{R_x(z)} = \frac{1}{\sigma_x^2 H_x(z)}\frac{R_{yx}(z)}{H_x(z^{-1})}$	$H_c(z) = \frac{1}{\sigma_x^2 H_x(z)}\left[\frac{R_{yx}(z)}{H_x(z^{-1})}\right]_+$
MMSE power	$P_{nc} = r_y(0) - \sum_{k=-\infty}^{\infty} h_{nc}(k)r_{yx}(k)$	$P_c = r_y(0) - \sum_{k=0}^{\infty} h_c(k)r_{yx}(k)$

Define $R_y(z) = \mathcal{Z}\{r_y(k)\}$ $R_x(z) = \mathcal{Z}\{r_x(k)\}$ $H(z) = \mathcal{Z}\{h(k)\}$
where $\mathcal{Z}\{.\} = z$ transform, then $R_{yx}(z) = H(z)R_x(z)$ $R_y(z) = H(z)H(z^{-1})R_x(z)$

LMS adaptive algorithm

$\hat{y}(n) = \mathbf{c}^T(n-1)\mathbf{x}(n)$ filtering

$e(n) = y(n) - \hat{y}(n)$ error formation

$\mathbf{c}(n) = \mathbf{c}(n-1) + 2\mu\mathbf{x}(n)e(n)$ coefficient updating

where:

$$P(n) = P_O + P_{tr}(n) + P_{ex}(\infty)$$
$$\tilde{\mathbf{c}}(n) = \mathbf{c}(n) - \mathbf{c}_O(n)$$

and

$$0 < \mu \ll \frac{1}{\sum_{k=1}^{M} E\{|x_k(n)|^2\}}$$
$$P_{tr}^{total} = \sum_{n=0}^{\infty} P_{tr}(n) \propto \frac{\tilde{\mathbf{c}}(0)}{\mu}$$
$$\mathrm{M} = \frac{P_{ex}(\infty)}{P_O} \cong \mu \sum_{k=1}^{M} E\{|x_k(n)|^2\} \equiv \mu M E\{|x(n)|^2\}$$

LSE FIR filter equations for order M

Measurement data:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^T(N_i) \\ \mathbf{x}^T(N_i+1) \\ \vdots \\ \mathbf{x}^T(N_f) \end{bmatrix}, \quad \text{where } \mathbf{x}(n) = [x(n) \quad x(n-1) \quad \cdots \quad x(n-M+1)]^T$$
$$\mathbf{y} = [y(N_i) \quad y(N_i+1) \quad \cdots \quad y(N_f)]^T$$

Normal equations and LSE error:

$$(\mathbf{X}^T\mathbf{X})\mathbf{c}_{ls} = \mathbf{X}^T\mathbf{y}$$
$$\hat{\mathbf{R}}\mathbf{c}_{ls} = \hat{\mathbf{d}}$$
$$E_{ls} = E_y - \hat{\mathbf{d}}^T\mathbf{c}_{ls}$$

No windowing: $N_i = M-1,\ N_f = N-1$

Full windowing: $N_i = 0,\ N_f = N+M-2$

Pre windowing: $N_i = 0,\ N_f = N-1$

END OF ATTACHMENT