Note: Vectors appear in bold-face, and random entities will be capitalized. Notation should be equivalent to what we’ve used in class; please ask about any notational conventions which aren’t clear to you. Unless otherwise specified, “linear” is to be interpreted as “affine.”

If your answer is a density function, be sure to include regions of support in the definition, if they are not the entire space. Simplify your answers as much as possible for full credit. If it is helpful to have the numerical value for an expression and you don’t have a calculator, you may make a reasonable estimate (for example, $\sqrt{3} \approx 1.73$) and continue. You will not be penalized, though, for not evaluating expressions in your final answer which require a calculator or tables which have not been supplied.

You are allowed only a single sheet, 8.5 x 11 inches, with both sides containing whatever notes you wish to bring into the exam.

Problem 1 (35) 

Problem 2 (30) 

Problem 3 (20) 

Problem 4 (20) 

Problem 5 (20) 

Total (125) 

Name______________________________
1. (5 pts. each) “True” means true in every case.
   (a) (5 pts.) A minimum variance, unbiased estimator attains the Cramér-Rao bound for variance.

   (b) If $A \subset B$, $P(A) \leq P(B)$.

   (c) The linear, least-squared error Bayesian estimate is orthogonal to all functions of the observations.

   (d) For covariance matrices of the random vectors $\mathbf{X}$ and $\mathbf{Y}$,
   
   $$\Lambda_{YY} \geq \Lambda_{YX}\Lambda_{XX}^{-1}\Lambda_{XY}$$

   (e) A set of jointly Gaussian RVs with non-singular covariance matrix can be linearly transformed into a set of i.i.d RVs.

   (f) Receiver Operating Characteristic curves lie on or above the line $P_D = P_F$.

   (g) For pdf $p(y; \mathbf{x})$, independent observations ($\mathbf{Y}$) lead to a diagonal Fisher Information matrix for parameters $\mathbf{x}$. 
2. Each of the elements of the vector $\mathbf{Y} = [Y_0 \ Y_1 \ ... \ Y_{N-1}]$ is conditionally Gaussian, with mean $M$. The conditional covariance matrix is diagonal, $\Lambda_{\mathbf{Y}|M} = \sigma^2 \mathbf{I}$.

(a) (10 pts.) Suppose $\sigma = 1$, $M = 2$ and $N = 4$, and define $\hat{M} = \sum_{i=0}^{N-1} Y_i$. Use the Chebyshev Inequality to bound the probability $P(\hat{M} > M + 2\sigma)$, and compare to the actual probability. Exploit anything you know to make the Chebyshev Bound as tight as possible.
(b) (10 pts.) Show how you can use these RVs to do the reverse of the usual random number transformation, using $Y$ to create a single uniformly distributed RV on the interval $(0,1)$. Demonstrate that the transformation truly yields the uniform RV.
(c) (10 pts.) If $N$ is even and $G = [1 \ -1 \ \ldots \ -1 \ -1]^T$ (dimension $N \times 1$), find the probability density function of $X = G^T Y$.
3. Some non-random parameter issues:
  (a) (10 pts.) Prove that when
  \[
  \frac{\partial}{\partial x} \ln p(y; x) = C(x)(g(Y) - x),
  \]
  and the pdf meets the regularity conditions for Cramér-Rao, \( g(Y) \) is an unbiased estimator of \( x \).

(b) (10 pts.) If \( \gamma = g(x) \) and we know \( p(y; \gamma) \), derive from the definition of Fisher information an expression for the information on \( x \) in terms of the Fisher information on \( \gamma \).
4. A corrupted signal is received as $Y[n] = S[n] + W[n]$, with $W[n]$ i.i.d. zero-mean Gaussian noise of variance $\sigma^2_W$, and

$$S[n] = \exp(-0.1n + \phi).$$

(a) (10 pts.) Find the maximum-likelihood estimate of $\phi$ from the observation of $Y$. 
(b) (10 pts.) Compute the Cramér-Rao Bound for unbiased estimates of $\phi$. 
5. The number of bees exiting a hive during any interval is a Poisson-distributed RV with parameter $\lambda T$ for interval of length $T$.
(a) (10 pts.) Find the distribution or density function for the RV which is the time the first bee exits after you start observing the hive at time $t = 0$. 
(b) (10 pts.) As $T \to \infty$, the number of bees which have left the hive, divided by $T$, may be expected to converge to $\lambda$, called the rate of this Poisson process. Prove anything you can concerning this convergence, and discuss what type of convergence you may have established.
Characteristic function: $\phi_X(\omega) = E[e^{j\omega X}]$
Moment generating function: $\theta_X(s) = E[e^{sX}]$
Joint Gaussian RVs:

$$p_X(x) = \frac{1}{(2\pi)^{N/2}|\Lambda_X|^{1/2}} \exp \left\{-1/2(x - \mu_X)^T \Lambda_X^{-1}(x - \mu_X) \right\}$$

$$\phi_X(\omega) = \exp \left\{j\omega^T \mu_X - 1/2\omega^T \Lambda_X \omega \right\}$$

Gamma function:

$$\int_0^\infty x^k e^{-x} dx = \Gamma(k+1), \Gamma(k+1) = k! \text{ for non-negative integers } k$$
Exponential distribution, with mean $\frac{1}{\theta}$, variance $\frac{1}{\theta^2}$: $p_X(x) = \theta e^{-\theta x} u(x)$
Poisson distributed counts with mean and variance $\theta$: $P(X = k) = \frac{e^{-\theta} \theta^k}{k!}$
Binomially-distributed RV:

$$P(X = k) = B(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Correlation coefficient of $X$ and $Y$: $\rho_{XY} = \frac{E[(X-\mu_X)(Y-\mu_Y)^*]}{\sigma_X \sigma_Y}$
Chernoff bound: $P(X \geq a) \leq e^{-at} \theta_X(t)$, with $\theta_X(t)$ the moment generating function.
Bienaymè-Chebyșhev: $P(X \geq a) \leq \frac{E[\varphi(X)]}{\varphi(a)}$ for $\varphi(x)$ non-negative, symmetric, increasing on $x \geq 0$.
Integration by parts: $\int u dv = uv - \int v du$

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

Bayesian binary hypothesis testing:

$$\frac{P(y|H_1)}{P(y|H_0)} \overset{H_1}{\underset{H_0}{\geq}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}$$
(Affine) Linear Least-Squared (LLS) Error Estimation (a.k.a. LMMSE)

$$\hat{X}_{LLS}(Y) = m_X + \Lambda_{XY} \Lambda_{YY}^{-1} (Y - m_Y)$$