Problem 1.1
For each item below, please list what (if any) courses that you have taken, books used, and grade received:
(a) probability  (b) statistics 
(c) linear algebra  (d) machine learning 
(e) optimization 

Problem 1.2
Let $A \in \mathbb{R}^{n \times n}$ be a matrix. Which of the following statements are equivalent to “$A$ is invertible”? Either give a proof of the equivalence, or a counterexample.
(i) The columns of $A$ span $\mathbb{R}^n$.
(ii) The rows of $A$ are linearly independent.
(iii) $\text{trace}(A) \neq 0$.
(iv) $\|Ax\|_2^2 > 0$ for all $x \neq 0$.
(v) $\det(A) \neq 0$

Problem 1.3
Consider a sequence of random variables $X_0, X_1, X_2, \ldots$ generated according to the following procedure. First we choose $X_0 \sim N(0, 1)$, and then for some number $|a| < 1$ we set $X_{t+1} = aX_t + W_t$ for $t = 0, 1, 2, \ldots$, where $W_t \sim N(0, \sigma^2)$, $W_t$ is uncorrelated with $X_t$, and $\{X_0, W_0, W_1, W_2, \ldots\}$ are mutually independent.
(a) Compute the joint distribution of $(X_0, X_1, X_2)$.
(b) Compute the distribution of $X_t$, as a function of $t$, $a$ and $\sigma^2$. 

1
(c) Are $W_t$ and $X_t$ independent? Are $W_t$ and $X_{t+10}$ uncorrelated?

(d) What happens to $\text{var}(X_t)$ as $t \to +\infty$?

Problem 1.4

Each cereal box contains an action figure, chosen uniformly from a set of four different action figures. The price of any given cereal box (in dollars) is an exponentially distributed random variable with parameter $\lambda$, and the prices of different cereal boxes are independent. For $i = 1, \ldots, 4$, let $T_i$ be a random variable corresponding to the number of boxes that you purchase in order to have $i$ different action figures. To be precise, after purchasing $T_2$ boxes (and not before), you have at least one copy of exactly 2 different action figures; and after purchasing $T_4$ boxes (and not before), you have at least one copy of all four action figures.

(a) What is the PMF, expected value, and variance of $T_1$?

(b) What is the PMF and expected value of $T_2$?

(c) Compute $E[T_4]$ and $\text{var}(T_4)$.

(d) Compute the moment generating function of $T_4$.

(e) You keep buying boxes until you have collected all four action figures. Letting $Z$ be a random variable representing the total amount of money (in dollars) that you spend, compute $E[Z]$ and $\text{var}[Z]$.

Problem 1.5

True or false: For each of the following statements, either give a counterexample to show that it is false, or provide an argument to justify that it is true. (Note: You will receive no points for just guessing the correct answer; full points will be awarded only when an answer is justified with an example or argument.)

(a) For any two events $A$ and $B$, if $P[A \mid B] > 1/2$, then $P[B \mid A] < 1/2$.

(b) If the moment generating function of $X_n$ is given by $M_{X_n}(t) = \exp(\frac{t^2}{2\sqrt{n}})$, then the sequence $\{X_n\}$ converges in probability to some real number.

(c) If $X \sim N(0, 1)$ and $Y \sim N(0, 1)$, then the Bayes' least squares estimator of $X$ given $Y$ is equal to the linear least squares estimator.
(d)) If the Bayes’ least squares estimator of $X$ given $Y$ is equal to $\mathbb{E}[X]$, then $X$ and $Y$ are independent.

**Problem 1.6**

Suppose that a pair of random variables $X$ and $Y$ has a joint PDF that is uniform over the shaded region shown in the figure below:

![Figure 1: Joint PDF of random variables $X$ and $Y$.](image)

(a) Compute the Bayes’ least squares estimator (BLSE) of $X$ based on $Y$. (*Note:* You should evaluate the required integrals; however, your answer can be left in terms of quantities like $1/e$ or $\sqrt{2}$).

(b) Compute the linear least squares estimator (LLSE) of $X$ based on $Y$, as well as the associated error variance of this estimator. Is the LLSE the same as the BLSE in this case? Why or why not? (*Note:* You should evaluate the required integrals; however, your answer can be left in terms of quantities like $1/e$ or $\sqrt{2}$).

(c) Now suppose that in addition to observing some value $Y = y$, we also know that $X \leq 1/e$. Compute the BLSE and LLSE estimators of $X$ based on both pieces of information. Are the estimators the same or different? Explain why in either case.